

24,25, 26 April, 2013, Azores, Portugal



CONFERENCE PROCEEDINGS

Vol.3

1st Annual International Interdisciplinary Conference, *AIIC 2013*

24-26 April 2013, Azores Islands, Portugal

(Conference place: **University of the Azores, Ponta Delgada**)

CONFERENCE PROCEEDINGS

Vol.3

1st Annual International Interdisciplinary Conference, *AIIC 2013*

24-26 April 2013, Azores Islands, Portugal

**EGALITE
(EUROPEAN SCIENTIFIC INSTITUTE Publishing)**

Impressum

Bibliographic information published by the National and University Library "St. Kliment Ohridski" in Skopje; detailed bibliographic data are available in the internet at <http://www.nubsk.edu.mk/>;

CIP - 616.43-097:616.44]-053.85
616.43-097:616.379-008.64]-053.85
COBISS.MK-ID 94077450

Any brand names and product names mentioned in this book are subject to trademark, brand or patent protection and trademarks or registered trademarks of their respective holders.

The use of brand names, product names, common names, trade names, product descriptions etc. even without a particular marking in this works is in no way to be construed to mean that such names may be regarded as unrestricted in respect of trademark and brand protection legislation and could thus be used by anyone.

Publisher: EGALITE (department : European Scientific Institute)
Street: "203", number "1", 2300 Kocani, Republic of Macedonia
Email: contact@eujournal.org
Printed in Republic of Macedonia

ISBN 978-608-4642-05-3

Copyright © 2013 by the authors, EGALITE and licensors
All rights reserved. Kocani, 2013

CONFERENCE PROCEEDINGS

Vol.3

1st Annual International Interdisciplinary Conference, *AIIC 2013*

24-26 April 2013, Azores Islands, Portugal

(Conference place: *University of the Azores, Ponta Delgada*)

June, 2013

Table Of Contents:

GLOBAL SOLUTIONS OF THE FUCHSIAN-CAUCHY PROBLEM IN GEVREY SPACES.....	1
<i>Faiza Derrab</i>	
ALGEBRAIC SCHUR COMPLEMENT APPROACH FOR A NON LINEAR 2D ADVECTION DIFFUSION EQUATION.....	7
<i>Hassan Belhadj</i> <i>Samir Khallouq</i>	
MAX-MIN EDGE MAGIC AND ANTIMAGIC LABELING.....	14
<i>J. Jayapriya</i> <i>K. Thirusangu</i>	
ON THE TOPOLOGICAL STRUCTURES ON Γ-SEMIRINGS.....	17
<i>Kostaq Hila</i> <i>Iilir Vardhami</i> <i>Kristaq Gjino</i>	
COMMON FIXED POINT THEOREMS FOR COMMUTING MAPPINGS ON A QUASIMETRIC SPACE.....	31
<i>Luljeta Kikina.</i> <i>Kristaq Gjino</i> <i>Iilir Vardhami</i>	
WEAK SEPARATION AXIOMS VIA D_ω, $D_{\alpha-\omega}$, $D_{pre-\omega}$, $D_{b-\omega}$, AND $D_{\beta-\omega}$ -SETS.....	37
<i>Mustafa Hasan Hadi</i>	
SYMBOLIC DATA ANALYSIS FOR THE ASSESSMENT OF USER SATISFACTION: AN APPLICATION TO READING ROOMS SERVICES.....	39
<i>Aurea Sousa</i> <i>Licínio Tomás</i> <i>Oswaldo Silva</i> <i>Helena Bacelar-Nicolau</i>	
ON USING SAGE TO SOLVE CONSTRAINED OPTIMIZATION PROBLEMS APPLYING THE LAGRANGE MULTIPLIERS METHOD.....	49
<i>Maria do Carmo Martins</i> <i>Francisco Martins</i>	
MATHEMATICS IN MOROCCAN UNIVERSITIES.....	55
<i>Zineb El Jaoussi</i> <i>Al Achhab Sakina</i> <i>Aziz Arbai</i> <i>Mohammed Khaldi</i>	
CONVERTING A DESERT BUGGY INTO SOLAR ELECTRIC VEHICLE.....	61
<i>Ahmed H. Osman</i> <i>Magdy. M. Massoud</i>	
THE PRINCIPLE OF RECORDING INFORMATION IN DISTRIBUTED ENVIRONMENTS VIA SULEIMENOV-MUN'S WAVES.....	68
<i>Dolayev M.</i> <i>Obukhova P.</i> <i>Panchenko S.</i> <i>Bakytbekov R.B.</i> <i>Ivlyev R.S.</i> <i>Chezhibayeva K.</i>	

POLARITY EFFECT ON CORONA DISCHARGE VOLTAGES IN SF₆-N₂ GAS MIXTURES.....	72
<i>M. Assia Guerroui</i> <i>Ahcene Lemzadmi</i>	
THE EFFECT OF CARBON BLACK ON D.C. CONDUCTIVITY OF EPOXY RESIN.....	76
<i>Ali S. Ali</i> <i>Kawakib J. Majeed</i>	
MATHEMATIC-PHYSICAL MODEL OF DIMENSIONING SYSTEM IN THE PROPAGATION OF MICROWAVE "WAVEGUIDE-SLUDGE" FROM WASTEWATER TREATMENT PLANTS.....	82
<i>Emilian-Narcis Riți-Mihoc</i> <i>Emil Riți-Mihoc</i> <i>Dan Porcar</i>	
INTERDISCIPLINARITY AROUND "CALÇADA PORTUGUESA" SIDEWALK.....	91
<i>Helena Melo</i>	
STUDY OF A THREE-AXIS PIEZORESISTIVE ACCELEROMETER WITH UNIFORM AXIAL SENSITIVITIES.....	95
<i>Abdelkader Benichou</i> <i>Nasreddine Benmoussa</i> <i>Kherreddine Ghaffour</i>	
SETTING UP A DATA ACQUISITION SYSTEM FOR SPARK ENGINES.....	100
<i>Ahmed H. Osman</i> <i>Magdy. M. Massoud</i>	
PLANING OF SMALL HYDROPOWER PLANTS AND KEZER RIVER.....	109
<i>Fevzi Onen</i> <i>Mahmut Eminoglu</i>	
THE URBAN ENVIRONMENTAL ENGINEERING CONDITION OF UDUPI TALUK: A GIS EVALUATION URBAN AMENITIES WITH A FOCAL POINT PREDICTION OF POTENTIAL IMPROVEMENT OF URBANIZATION.....	118
<i>H N Udayashankara</i> <i>B.S.Maddodi</i> <i>B.R.Raghavan</i> <i>Shema V.T</i>	
PLANE STATE PROBLEM ANALYSIS WITH FINITE-DIFFERENCE METHOD.....	123
<i>Harangus Katalin</i> <i>Kakucs Andras</i> <i>Galfi Botond Pal</i>	
SIMULATION OF AN AXIAL FLOW TURBINE RUNNER'S BLADES USING CFD.....	131
<i>Diaelhag Khalifa</i>	
A SYSTEM DYNAMICS MODEL TO ASSESS THE OPTIMAL NUMBER OF CREWS FOR CONDUCTING AIR CARGO OPERATIONS.....	138
<i>R. Pozzi</i> <i>C. Noe</i> <i>T. Rossi</i>	
COMPENSATION OF VOLUMETRIC EFFICIENCY BY TURBOCHARGING IN AN INSULATED DI DIESEL ENGINE WITH ALCOHOL AS FUEL.....	145
<i>S.Sunil Kumar Reddy</i> <i>V. Pandurangadu</i>	

FAULT DIAGNOSIS OF INDUCTION MOTOR.....	154
<i>Sanjana Malhotra</i>	
<i>M. K. Soni</i>	
INSTITUTIONAL EFFECT ANALYSIS COMPARING ENERGY EFFICIENCY RETROFITTING FOR EXISTING RESIDENTIAL BUILDINGS PATTERNS IN CHINA.....	160
<i>Junna Zhao</i>	
<i>Fanghong Lou</i>	
RECTANGULAR OPEN CHANNELS OF COMBINED ROUGHNESS.....	174
<i>E. Retsinis</i>	
<i>S. Bobotas</i>	
<i>J. Demetriou</i>	
CRITERIA FOR USING A ONE-DIMENSIONAL HYDRODYNAMIC MODEL FOR TWO BARRAGES SITUATED IN MEDITERRANEAN REGION.....	182
<i>Takkouk Saddek</i>	
<i>Xavier Casamitjana</i>	
SPURIOUS RESPONSE SUPPRESSION IN THREE POLES HAIRPIN BANDPASS FILTER USING DEFECTED GROUND STRUCTURES.....	191
<i>N. Amar Touhami</i>	
<i>A. Zakriti</i>	
<i>M. Lamsalli</i>	
<i>M. Boussouis</i>	
<i>A. Tribak</i>	
VISUALIZATION AND MEASUREMENTS OF BUBBLY TWO-PHASE FLOW STRUCTURE USING PARTICLE IMAGING VELOCIMETRY (PIV).....	197
<i>Hassan Abdulmouti</i>	
<i>Esam Jassim</i>	
DEVELOPMENT AND EVALUATION OF A DEFECT TRACKING MODEL FOR CLASSIFYING THE INSERTED DEFECT DATA.....	209
<i>Torky sultan</i>	
<i>Ayman E.Khedr</i>	
<i>Mostafa sayed</i>	
3GPP: LTE- AN INNOVATIVE TECHNOLOGY TOWARDS 4G WIRELESS NETWORKS.....	219
<i>J Vijay Franklin</i>	
<i>K.Paramasivam</i>	
IMPACT OF AN ACTION CYBER CRIME PREVENTION PROGRAMME ON IN-SCHOOL AGED CHILDREN'S ATTITUDE TO CRIME PREVENTION CONCEPTS IN CIVIC EDUCATION AND SOCIAL STUDIES.....	225
<i>Peter Adewale Amosun</i>	
<i>Olugbenga Adedayo Ige</i>	
THE ARCHITECTURE OF VIRTUAL SPACE MUSEUMS.....	238
<i>Anda-Ioana Sfinteş</i>	
A PROCESS OF SEAMLESSLY REPLACING CG ELEMENTS INTO LIVE-ACTION FOOTAGE.....	245
<i>Jin Zhi</i>	
SCADA LIVE FORENSICS: REAL TIME DATA ACQUISITION PROCESS TO DETECT, PREVENT OR EVALUATE CRITICAL SITUATIONS.....	253
<i>Pedro Taveras N.</i>	

FACTOR INFLUENCING MOBILE NUMBER PORTABILITY(MNP) IN KENYA: THE CASE OF AFRICA NAZARENE UNIVERSITY.....	263
<i>Githii Stephen Kagwathi</i>	
<i>John Njau Kamau</i>	
<i>Mary Muthoni Njau</i>	
<i>Elias Kiarie Kagiira</i>	
BUILDING VISION AND VOICE BASED ROBOTS USING ANDROID.....	276
<i>Pradeep N.</i>	
<i>Mohammed Sharief</i>	
<i>M. Siddappa</i>	
GROUND FLASH DENSITY OF THE BRASS COAST OF NIGERIA.....	284
<i>John Tarilanyo Afa</i>	
<i>Benjamin Ayebapreye Kelvin</i>	
CREATING PUBLIC ENVIRONMENTAL AWARENESS THROUGH MUSIC: “PLAY FOR NATURE” PROJECT.....	289
<i>A.Banu Bicakci</i>	
ANTHROPOGENIC IMPACT CONDITIONS ON WATER QUALITY IN THE ALLUVIAL AQUIFER OF THE REGION TADJENANET - CHELGHOU MOUNTAINS (EASTERN ALGERIA).....	294
<i>Abdelhamid Khedidja</i>	
<i>Abderrahmane Boudoukha</i>	
URBAN STORMWATER QUALITY AND QUANTITY IN THE CITY OF TALLINN.....	305
<i>Bharat Maharjan</i>	
<i>Karin Pachel</i>	
<i>Enn Loigu</i>	
FRENCH ENVIRONMENTAL SOCIOLOGY FACING INTERDISCIPLINARY RESEARCH ACROSS SOCIAL AND NATURAL SCIENCES.....	315
<i>Aspe Chantal</i>	
<i>Jacque Marie</i>	
VARIOUS ASPECTS OF ANTHROPOGENIC MATERIAL FLOWS IN THE CZECH REPUBLIC: PRESENTATION OF THE PROJECT.....	322
<i>Jan Kovanda</i>	
<i>Tomas Hak</i>	
LIFE CYCLE EVALUATION IN THE NATIONAL PARK MURÁŇ PLAIN (SLOVAKIA).....	332
<i>Judita Tomaskinova</i>	
<i>Jan Dubiel</i>	
GENERATION AND COLLECTION PRACTICES OF ORGANIC KITCHEN WASTE IN HOUSEHOLDS OF MANIPAL.....	339
<i>Lalitha Simon</i>	
<i>Shivali Patel</i>	
<i>Anysha Mattu</i>	
<i>Teresa Chirayil</i>	
ECOLOGICAL ROLE OF ANIMAL DIVERSITY IN SOIL SYSTEM (A CASE STUDY AT EL-RAWAKEEB DRY LAND RESEARCH STATION, SUDAN).....	345
<i>Maha Ali Abdel Latif</i>	
STATE OF THE ENVIRONMENT IN THE NIGER DELTA AREA OF ONDO STATE.....	351
<i>F.A. Olorunlana</i>	

PROPERTIES OF NEW ENVIRONMENTALLY FRIENDLY BIODEGRADABLE INSULATING FLUIDS FOR POWER TRANSFORMERS.....	358
<i>Pawel Rozga</i>	
COMPLIANCE OF WASTEWATER TREATMENT PLANTS IN JÄRVA COUNTY WITH THE EU URBAN WASTEWATER TREATMENT DIRECTIVE AND ESTONIAN NATIONAL REQUIREMENTS.....	365
<i>Raili Niine</i>	
<i>Enn Loigu</i>	
<i>Karin Pachel</i>	
RISK CAUSED BY THE SPATIAL FORMATION OF THE CITY OF BURSA (TURKEY) IN HISTORICAL PROCESS, CURRENT DEVELOPMENT PLANS AND LAND USAGES BASED ON DEVELOPMENT PLANS.....	376
<i>Ufuk Fatih Kucukali</i>	
IMPROVING MUNICIPAL MANAGEMENT ASSOCIATED WITH THE RESOURCES COLLECTION THROUGH IT: CAASIM CASE IN MEXICO.....	386
<i>Antonio Alejandro Arriaga-Martinez</i>	
<i>Sonia Ximena Diaz de Cossio-Priego</i>	
TREATED WASTEWATER USE AND ITS EFFECT ON WATER CONSERVATION, VEGETATIVE YIELD, YIELD COMPONENTS AND WATER USE EFFICIENCY OF SOME VEGETABLE CROPS GROWN UNDER TWO DIFFERENT IRRIGATION SYSTEMS IN WESTERN REGION, SAUDI ARABIA.....	395
<i>Khaled S. Balkhair</i>	
<i>Fathy S. El-Nakhlawi</i>	
<i>Saleh M. Ismail</i>	
<i>Samir G. Al-Solimani</i>	
PHYTODEPURATION PROCESS FOR THE RECYCLING OF WASTEWATER IN DAIRY.....	403
<i>Lydia Ferrara</i>	
MICROALGAE – BIODIESEL POTENTIAL PRODUCERS: A REVIEW.....	407
<i>Anna Krasowska</i>	
<i>Slawomir Jablonski</i>	
<i>Piotr Biniarz</i>	
<i>Malgorzata Plachetka</i>	
<i>Marcin Lukaszewicz</i>	
THE UNCERTAIN UNCERTAINTY OF RIVER BASINS: ACCOUNTING FOR UNCERTAINTY IN INTEGRATED WATER RESOURCES MANAGEMENT (IWRM).....	417
<i>Sarah Allen</i>	
PRELIMINARY TESTS OF THERMAL CONDUCTIVITY OF SELECTED SOIL TYPES.....	428
<i>Agata Ludynia</i>	
<i>Lukasz J. Orman</i>	
EFFECT OF COASTAL ENVIRONMENT IN CLAY FACING BRICKS AND ROOF TILES.....	432
<i>B. Sena da Fonseca</i>	
<i>J.A.R. Simao</i>	
<i>C. Galhano</i>	
CHARACTERISATION OF ALKALINE LIPASE FROM AN ARCTIC YEAST STRAIN <i>RHODOSPORIDIUM BABJEVAE</i> BD19.....	441
<i>Marcin lukaszewicz</i>	
<i>Slawomir Jablonski</i>	
<i>Anna Krasowska</i>	

DEPOPULATION OF TRADITIONAL MINING REGIONS IN CENTRAL AND EAST EUROPE: CASE STUDY OF THE UPPER-SILESIA BASIN (POLAND) AND THE DONETSK BASIN (UKRAINE).....	450
<i>Marcin Rechlowicz</i>	
<i>Maria Tkocz</i>	
THE SOUTHERN TELLIAN SERIES AT NUMMULITES (NE ALGERIA): STRATIGRAPHY AND STRUCTURAL CONSEQUENCES.....	460
<i>Mohamed Chadi</i>	
<i>El hadj Youcef Brahim</i>	
<i>Rami Djeflal</i>	
PEDOSTRATIGRAPHY, PEDOLOGICAL AND GEOCHEMISTRY OF KASHMIR LOESS: IMPLICATIONS FOR CHEMICAL WEATHERING HISTORY AND PALEOCLIMATIC RECONSTRUCTION.....	467
<i>Rakesh Chandra</i>	
<i>Ishtiaq Ahmad</i>	
LATE CRETACEOUS SYNSEDIMENTARY TECTONIC IN EASTERN ATLAS SAHARAN (NORTH EAST OF ALGERIA).....	479
<i>Benmansour Sana</i>	
GEOLOGICAL AND GEOTECHNICAL CHARACTERISTICS OF THE SOILS IN THE REGION OF SÉTIF.....	484
<i>Djenba Samir</i>	
TOWARD INDONESIA'S AGROINDUSTRIES COMPETITIVENESS: THE CASE OF BIOETHANOL DEVELOPMENT FROM SUGARCANE BASED INDUSTRIES.....	491
<i>Gita K. Indahsari</i>	
<i>Rudi Wibowo</i>	
RESPONSES OF RECOMBINANT INBRED LINES OF COWPEA [(<i>VIGNA UNGUICULATA</i> (L.) WALP] TO <i>STRIGA GESNERIOIDES</i> INFESTATION IN GHANA.....	503
<i>Aaron T. Asare</i>	
<i>Isaac K. A. Galyuon</i>	
<i>Francis K. Padi</i>	
<i>Emmanuel P. Otwe</i>	
<i>J.F. Takrama</i>	
DETERMINANTS OF HOUSEHOLD FOOD SECURITY IN THE SEKYERE-AFRAM PLAINS DISTRICT OF GHANA.....	514
<i>Robert Aidoo</i>	
<i>James Osei Mensah</i>	
<i>Thomas Tuffour</i>	
MOLECULAR IDENTIFICATION OF A DESTRUCTIVE PHYTOPATHOGENIC FUNGUS IN TOMATO FIELDS OF IRAN.....	522
<i>Parissa Taheri</i>	
<i>Atena Pourmahdi</i>	
MODEL POLICY DESIGN FOR THE BEEF CATTLE RANCH DEVELOPMENT IN SOUTH SULAWESI.....	525
<i>Machmud Achmad</i>	
<i>Sri Hartoyo</i>	
<i>Bustanul Arifin</i>	
<i>Muhammad Said Didu</i>	

DETERMINATION OF FARMERS' GOAL STATEMENTS PRIORITIES AND FACTORS EFFECTIVE ON DECISION (AN APPLICATION OF FUZZY PAIRWISE COMPARISON ANALYSIS).....	537
<i>Ahmet Sahin</i>	
<i>Cihat Gunden</i>	
<i>Murat Cankurt</i>	
<i>Bulent Miran</i>	
<i>Ibrahim Yildirim</i>	
BREAK EVEN ANALYSIS OF BROILER PRODUCTION IN THE ACCRA-TEMA AND KUMASI AREAS.....	545
<i>Mahama, E.A.</i>	
<i>Andah, E.K.</i>	
<i>Amegashie, D.P.K.</i>	
<i>Mensah-Bonsu, A.</i>	
THE ROLE OF SOCIO-CULTURAL IN RELATIONSHIP MARKETING: CASE FINDINGS ON FARMERS IN WEST JAVA.....	553
<i>Muchlis Ahmady</i>	
<i>Ujang Sumarwan</i>	
<i>Budi Suharjo</i>	
<i>Agus Maulana</i>	
TARAKAN STRUCTURAL MODEL OF IMPLEMENTATION OF THE PREFERENCE PROGRAM CITY GAS UNDER ANY THEORY OF PLANNED BEHAVIOUR.....	565
<i>Tubagus Haryono</i>	
<i>Ujang Sumarwan</i>	
<i>Asep Saefuddin</i>	
PREVALENCE OF POLYGLANDULAR AUTOIMMUNE SYNDROME TYPE III IN A GROUP OF ADULTS WITH THYROID DISEASES AND DIABETES MELLITUS.....	575
<i>Gherbon Adriana</i>	
INFLUENCING FACTORS IN MMR IMMUNISATION DECISION MAKING.....	582
<i>Marie C. Hill</i>	
<i>Carol L. Cox</i>	
PREDICTING NURSES' TURNOVER INTENTIONS BY DEMOGRAPHIC CHARACTERISTICS, PERCEPTION OF HEALTH, QUALITY OF WORK, AND WORK ATTITUDES.....	593
<i>Mahmoud Al-Hussami</i>	
<i>Muhammad Darawad</i>	
<i>Ali Saleh</i>	
<i>Ferial Ahmed Hayajneh</i>	
THE GENDERED KNOWLEDGE OF HIV/AIDS AMONG URBAN UNIVERSITY STUDENTS IN SOUTHERN ETHIOPIA.....	602
<i>Marianne Paul</i>	
<i>Basanti Majumdar</i>	
<i>Noori Akhtar-Danesh</i>	
<i>Sheryl Boblin</i>	
<i>Diana Sherifali</i>	
CHANGES LYMPH FLOW, HEART AND TRANSCAPILLARY EXCHANGE FLUID AND PROTEINS REFLEX-HUMORAL INFLUENCES.....	610
<i>Myrzakhanov N.</i>	
<i>Myrzakhanova M.N.</i>	
HRD MECHANISMS IN HEALTH CARE SECTOR IN J&K: A COMPARATIVE STUDY.....	614
<i>S.A. Mufeed</i>	
<i>Qurrat.A. Hamdani</i>	

ARE THERE CHANGES IN ORBICULARIS ORIS MUSCLE ACTIVITY DUE TO THE CONVENTIONAL PROSTHETIC TREATMENT OF COMPLETE EDENTATION ?.....640*S.Popsor**Claudia Soaita***HEPATOPROTECTIVE EFFECT OF BAY LEAVES CRUDE EXTRACT ON PRIMARY CULTURED RAT HEPATOCYTES.....647***Nahla A. Ayoub**Amani N. Hashim**Sahar A. Hussein**Nesrine M. Hegazi**Heba M. Hassanein**Mahmoud A. Nawwar***LARYNX PATHOLOGY AT PATIENTS WITH GASTRO-INTESTINAL TRACT DISORDERS.....656***Toguzbaeva Dinara Yerkenovna***ASSESSMENT OF THE REACTIVE HYPERAEMIA THROUGH QUANTIFICATION OF COLOUR CHANGES USING PIXEL ANALYSIS SOFTWARE.....659***T. Tanwi**S Manjunatha**B. Priya***EMERGENCY TROLLEYS: AVAILABLE AND MAINTAINED BUT ARE THEIR LOCATIONS KNOWN?.....667***Sawsan Ghonaimy**Yahya Khedr**Ahmed Boraei***DRUG SCENE CHANGES DURING TWO DECADES: SLOVAKIA 1993 – 2012.....671***Alojz Nociar***COMPARISON OF ELECTRICAL ACTIVITY OF PATELLA STABILIZER MUSCLES BETWEEN ATHLETES WITH AND WITHOUT PATELLOFEMORAL PAIN SYNDROME.....679***Nahid Khoshraftar Yazdi**Majid Mehdikhani**Stephan Starischka**Elke Zimmermann***EFFECT OF LUMBAR STABILIZATION EXERCISES VERSUS PRESSURE FEEDBACK TRAINING IN LOW BACKACHE PATIENTS.....687***Shadab Uddin**Fuzail Ahmed***TYPE 1/TYPE 2 CYTOKINE SERUM LEVELS AND ROLE OF INTERLEUKIN-18 IN CHILDREN WITH STEROID SENSITIVE NEPHROTIC SYNDROME.....696***Sherein A. Shalaby**Howaida M. Al-Edressi**Shereen A. El-Tarhoumy**Mohamed Fath EL-Bab**Mohamed A. Zolaly*

AN ASSESSMENT OF THE SPATIAL RELATIVITY OF HEALTH, LITERACY AND SOCIO-ECONOMIC GROWTH INDEX USING GIS TECHNIQUES: A CASE STUDY ON UDUPI SOUTH WEST COAST OF INDIA, KARNATAKA.....705*B S Maddodi**H N Udayashankara**B R Raghavan**Alina E Mathews***THE INFLUENCE OF REFERRAL SERVICE ON HOUSEHOLD ACCESS TO HEALTHCARE SERVICES IN ELDORET MUNICIPALITY, KENYA.....711***Nicholas Walter Otieno Ajwang'***FROM EPISTEMOLOGICAL TRENDS TO CLINICAL AND EDUCATIONAL PRACTICES IN SPEECH-LANGUAGE PATHOLOGY.....730***Nicole Keating**Chantal Mayer-Crittenden**Michele Minor-Corriveau**Manon Robillard**Roxanne Beanger***A DEA-SFA COMPARISON OF THE IMPACT OF ICT'S UTILIZATION.....736***Puiu Fatulescu***A MULTI-STATE MARKOV MODEL FOR PROJECTING HEALTH CARE SPENDING.....745***Francisco Javier Blanco-Encomienda***SEEKING MEDICAL ATTENTION AMONG JORDANIAN PATIENTS WITH ACUTE MYOCARDIAL INFARCTION.....753***Nidal F. Eshah***COMPARISON BETWEEN DIFFERENT CORD BLOOD STEM CELL POPULATIONS IN EFFICIENCY OF TRANSDIFFERENTIATION INTO HEPATIC LINEAGE.....762***Hala Gabr**Rania Zayed**Ahmed Rezk Elzayat***VERBAL AND NUMERICAL APTITUDE OF GRADUATE NURSING STUDENTS AT ENTRY LEVEL.....769***Maxie Andrade**Anice George**Shreemathi Mayya**Renee Furtado***PRODUCTION AND DEVELOPMENT OF NUTRACEUTICALS USING *BACILLUS SUBTILIS* NCIM 2708 UNDER SOLID STATE FERMENTATION BY RESPONSE SURFACE METHODOLOGY.....773***Md Faruque Ahmad**Syed Aamir Ashraf**ZR Azaz Ahmad Azad**Bibhu Prasad Panda***SOLID LIPID NANOPARTICLES FOR TOPICAL DELIVERY OF MELOXICAM: DEVELOPMENT AND IN VITRO CHARACTERIZATION.....779***Rawia M. Khalil**Ahmed Abd El-Bary**Mahfoz A. Kassem**Mamdouh M. Ghorab**Mona B. Ahmed*

AUTOMATED CLASSIFICATION OF BRAIN MRI USING COLOR-CONVERTED K-MEANS CLUSTERING SEGMENTATION AND APPLICATION OF DIFFERENT KERNEL FUNCTIONS WITH MULTI-CLASS SVM.....799*N.Rajalakshmi**V.Lakshmi Prabha***THE EFFECTS OF DIFFERENT DIETARY PROTEIN AND LIPID LEVELS AND OIL SOURCES ON THE GROWTH PERFORMANCE AND BODY COMPOSITION OF RAINBOW TROUT (*ONCORHYNCHUS MYKISS*, W.).....807***Murat Bilguven**Ibrahim AK***DETERMINATION OF CLENBUTEROL IN MEAT SAMPLES WITH ELISA AND GC-MS METHOD.....817***Zehra Hajrulai-Musliu**Risto Uzunov**Velimir Stojkovski**Elizabeta Dimitrieska-Stojkovik**Biljana Stojanovska-Dimzoska**Pavle Sekulovski***ANTIHYPERGLYCEMIC AND LIPID LOWERING ACTIVITIES OF ETHANOLIC EXTRACT OF *ERIOBOTRYA JAPONICA* SEEDS IN ALLOXAN INDUCED DIABETIC RATS.....825***Sabeeha Shafi**Nahida Tabassum***INFLUENCE OF E-BEAM POSTHARVEST IRRADIATION IN THE COLOUR OF FOUR EUROPEAN CHESTNUT FRUIT VARIETIES OF *CASTANEA SATIVA* MILL.....835***Amilcar L. Antonio**Marcio Caroch**Albino Bento**Andrezej Rafalski**Begona Quintana***SIMULTANEOUS SEPARATION/PRECONCENTRATION OF NICKEL AND COBALT BY DISPERSIVE LIQUID-LIQUID MICROEXTRACTION PRIOR TO DETERMINATION BY FAAS.....841***F. Sanchez Rojas**C. Bosch Ojeda**J.M. Cano Pavon***POLYCYCLIC AROMATIC HYDROCARBONS (PAHs) AND POLYCHLORINATED BIPHENYLS (PCBs) IN SOILS OF AGBABU, NIGERIA.....849***Olubunmi E. Fagbote**Edward O. Olanipekun***COMPARATIVE STUDY OF PEAT HUMIC ACIDS BY USING UV SPECTROSCOPY.....857***Oskars Purmalis**Maris Klavins***SYNTHESIS OF 5-BENZIMIDAZOLYL BENZOFURAN DERIVATIVES OF EXPECTED BIOLOGICAL ACTIVITY.....867***Fatma AbdAleem A.ElHag**Reem A.K. ElHarbi**Fatma M. Edress**Hoda I. El Diwani*

PHYSICOCHEMICAL SENSORS FOR THE DETERMINATION OF SILVER.....	873
<i>Wagiha , H. Mahmoud</i>	
<i>Fatma Abdul Aleem EL-hag</i>	
<i>Reem A.K. AL –Harbi</i>	
CHEMOINFORMATICS AND ITS APPLICATIONS.....	879
<i>Wenli XU</i>	
<i>Min LING</i>	
<i>Jing HU</i>	
<i>Ying HUANG</i>	
<i>Jia LI</i>	
<i>Jianhua YAO</i>	
PHYTODEPURATION PROCESS FOR THE RECYCLING OF WASTEWATER IN DAIRY.....	883
<i>Lydia Ferrara</i>	
ZINC DIPEPTIDE COMPLEX ([Zn(II)-GLY-TYR]⁺)-NINHYDRIN REACTION IN PRESENCE OF GEMINI SURFACTANTS: A KINETIC STUDY.....	887
<i>Mohd. Akram</i>	
<i>Dileep Kumar</i>	
<i>Kabir-ud-Din</i>	
PARASITIC INFECTIONS OF THE COMMON CARP <i>Cyprinus carpio</i> FROM LESSER ZAB RIVER IN KURDISTAN REGION, IRAQ.....	895
<i>Kamaram S. Mama</i>	
<i>Shamall M. A. Abdullah</i>	
PREVALENCE AND COMPARISON BETWEEN THE EFFICACY OF DIFFERENT TECHNIQUES FOR DIAGNOSIS OF <i>TOXOPLASMA GONDII</i> AMONG WOMEN IN ERBIL PROVINCE-IRAQI KURDISTAN.....	901
<i>Narmin R. Hamad</i>	
<i>Mohammed A. Kadir</i>	

GLOBAL SOLUTIONS OF THE FUCHSIAN-CAUCHY PROBLEM IN GEVREY SPACES

Faiza Derrab, MA

University Djillali Liabès/Sidi-Bel-Abbès, Mathematics Department, Exact Sciences Faculty, Algeria

Abstract:

We consider the Fuchsian Cauchy problem associated to linear partial differential equations with Fuchsian principal part of order m and weight μ in the sense of M. S. Baouendi and C. Goulaouic [2]. We obtain existence and uniqueness of a global solution to this problem in the space of holomorphic functions with respect to the fuchsian variable t and in Gevrey spaces with respect to the other variable x . The method of proof is based on the application of the fixed point theorem in some Banach spaces defined by majorant functions that are suitable to this kind of equations. We introduce new majorant functions as in [4] and [5] which allow us to simplify the proof given in [3].

Key Words: Fuchsian linear partial differential equations, Global solution, Gevrey spaces, Method of majorants, Fixed point theorem

Introduction

In [1], it is established by J.M. Bony and P. Schapira that the hyperbolicity of the operator is sufficient so that the Cauchy problem is well posed in the class of holomorphic functions (Bony-Schapira-Theorem). But, this result was established by C. Wagschal in [10] for operators not necessarily hyperbolic.

To demonstrate his local existence theorem for an holomorphic solution of a semilinear operator, A. Cauchy used his method known as the "Majorant series method" which consists to find the solution of the problem in the form of a serie and to prove the convergence of this serie by bounding above the modulus of the coefficients of this serie by those of another serie whose terms are positive and is convergent.

C. Wagschal [10] simplified local resolution of the nonlinear Goursat problem in spaces of holomorphic functions and in Gevrey spaces to the fixed point theorem. His technique consists to define Banach algebras, either through the formalism of majorant functions of A. Cauchy in the holomorphic case, either through the formalism of formal series in the Gevrey case, in which his problem is reduced to the search for fixed points of some map that he built from the original problem and he shows strictly contracting in balls of these algebras.

And since the method of fixed point developed by C. Wagschal [10] has been used successfully in other problems. It cites the work of P. Pongérard and C. Wagschal [7] its aim is to reduce the global resolution of the Cauchy-Kowalvski problem to the fixed point theorem in spaces of entire functions and entire functions of finite order.

Interest was directed later to study or to simplify the study of Fuchsian equations by the method of fixed point. The major difficulty is in the search for Banach spaces, so for majorant functions, suitable for the study of this type of equations.

P. Pongérard [9] extended the conclusions of [7] to Fuchsian operators in the class of entire functions and entire functions of finite order using this method of fixed point.

In [4], we have simplified the proof given by P. Pongérard in [9] by defining new majorant functions by introducing a new parameter ρ and we generalized his result to a differential operator with several Fuchsian variables.

In [3], the authors have extended the result of [9] to a global resolution in spaces of holomorphic functions with respect to the Fuchsian variable and in Gevrey classes with respect to the other variables. One of these classes of Gevrey is more general than the class introduced in [10] and the second is introduced by H. Komatsu in [6].

In this work, we simplify the proof given in [3] using the same technique introduced in [4] by defining new majorant functions by introducing a new parameter ρ . Thus, we establish a global solution for our problem which is holomorphic with respect to the Fuchsian variable and is in Gevrey spaces with respect to the other variable. Our study is limited to Gevrey class defined by H.Komatsu [6]. This same technique has enabled us in [5] to give global resolution for some nonlinear equations of Fuchs type in these same Gevrey classes.

The Problem Formulation And Result

We study Fuchsian linear partial differential equations in the space $\mathbb{C} \times \mathbb{R}^n$. We denote by t the generic point of \mathbb{C} and by $x = (x_1, \dots, x_n)$ the generic point of \mathbb{R}^n . Let Ω be an open set in \mathbb{R}^n . For a multiindice $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ we denote $D^\alpha = D_1^{\alpha_1} \dots D_n^{\alpha_n}$ where $D_j = D_{x_j}$ is the partial derivative with respect to x_j and by $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n$.

Let $m \geq 1$ be an integer, we denote by E a subset of $\{(l, \alpha) \in \mathbb{N} \times \mathbb{N}^n; l + |\alpha| \leq m, l < m, \alpha \neq 0\}$.

Let $0 \leq \mu \leq m$, we consider the Cauchy problem

$$\begin{cases} a(t, D_t)u(t, x) = \sum_{(l, \alpha) \in E} a_{(l, \alpha)}(t, x) t^{v+1+l-\mu} D_t^l D^\alpha u(t, x) + f(t, x); & (t, x) \in \mathbb{C} \times \Omega, \\ D_t^j u(t, x) = w_j(x), & 0 \leq j < \mu, \quad x \in \Omega, \end{cases} \quad (1)$$

where $a(t, D_t)$ is the linear differential operator defined by $a(t, D_t) = \sum_{l=\mu}^m a_l t^{l-\mu} D_t^l$, and a_l for $\mu \leq l \leq m$ are complex constants with $a_m \neq 0$.

$a(t, D_t)$ is then a Fuchsian principal part of order m and weight μ .

$v = v(l)$ is the integer number defined by $v = \max(\mu - l - 1, 0)$ and the coefficients $a_{(l, \alpha)}$ for $(l, \alpha) \in E$ are polynomial functions with respect to x of order strictly inferior to $|\alpha|$ with holomorphic coefficients in \mathbb{C}_t . It means that

$$\begin{aligned} & \text{for } (l, \alpha) \in E, \quad a_{(l, \alpha)}(t, x) \\ &= \sum_{|\beta| < |\alpha|} a_{l\alpha\beta}(t) x^\beta \text{ where } a_{l\alpha\beta} \text{ is an holomorphic function in } \mathbb{C}_t. \end{aligned} \quad (2)$$

We associate to the operator $a(t, D_t)$ the polynomial $P(\lambda) = \sum_{l=\mu}^m a_l \prod_{j=0}^{l-1} (\lambda - j)$ and we consider $\prod_{\phi} = 1$.

We obtain: $t^\mu a(t, D_t) = \sum_{l=\mu}^m a_l t^l D_t^l = P(tD_t)$. Then $P(tD_t)$ is a Fuchsian principal part of weight 0.

For the choice of f and w_j ($0 \leq j < \mu$), we introduce the following definitions. We recall the Gevrey class definition in the sense given by H.Komatsu in [6]

Definition 1: Let $d \geq 1$. A function $v \in C^\infty(\Omega)$ is said to be in the Gevrey class $G^{(d)}$ if for every compact set $K \subset \Omega$ and $h > 0$ there exists a constant $c = c_{K,h} \geq 0$ such that

$$\forall \alpha \in \mathbb{N}^n, \quad \sup_{x \in K} |D^\alpha v(x)| \leq c h^{|\alpha|} (|\alpha|!)^d.$$

Examples:

1. For $n \in \mathbb{N}$; $f_n(x) = \sin nx \in G^{(d)}(\mathbb{R})$. Then the vector space generated by $\{\sin nx, \cos nx, n \in \mathbb{N}\}$ is a subset of $G^{(d)}(\mathbb{R})$.

2. For $b \in \mathbb{N}$; $x \in \mathbb{R}^n$, $L(x) = x^\beta$ a polynomial of order $\beta \in \mathbb{N}^n$ such that $|\beta| \leq b$, we obtain $L(x) \in G^{(d)}(\mathbb{R}^n)$. Then the set $\mathbb{R}[x]$ of polynomials is a subset of $G^{(d)}(\mathbb{R}^n)$.

For U an open set in \mathbb{C} , we denote by $C^{\omega, \infty}(U \times \Omega)$ the algebra of functions $u: U \times \Omega \rightarrow \mathbb{C}$ which admit derivatives for every order with respect to x , are continuous in $U \times \Omega$ and holomorphic with respect to t .

Definition 2: We say that a function u is of $G^{(\omega, d)}(U \times \Omega)$ class if u belongs to $C^{\omega, \infty}(U \times \Omega)$ and if for every compact set $K \subset \Omega$ and $h > 0$ there exists a constant $c = c_{K,h} \geq 0$ such that

$$\forall \alpha \in \mathbb{N}^n, \quad \forall t \in U, \quad \forall x \in K; \quad |D^\alpha u(t, x)| \leq c h^{|\alpha|} (|\alpha|!)^d.$$

For $R > 0$, we denote by $D_R = \{t \in \mathbb{C}; |t| < R\}$.

$G^{(\omega,d)}(\mathbb{C} \times \Omega)$ denotes the set of functions $u \in C^{\omega,\infty}(\mathbb{C} \times \Omega)$ such that for all $R > 0$, $u \in G^{(\omega,d)}(D_R \times \Omega)$.

$G^{(\omega,d)}(\mathbb{C} \times \Omega)$ is a sub-algebra of $C^{\omega,\infty}(U \times \Omega)$.

The coefficients $a_{(l,\alpha)}$ assumed verifying (2), then we obtain the following theorem:

Theorem 1: If $P(\lambda) \neq 0$ for every integer $\lambda \geq \mu$, then for any functions $w_j \in G^{(d)}(\Omega)$, ($0 \leq j \leq \mu$) and $f \in G^{(\omega,d)}(\mathbb{C} \times \Omega)$; the Cauchy problem (1) admits a unique solution $u \in G^{(\omega,d)}(\mathbb{C} \times \Omega)$.

Remark 1:

1. The theorem1 establishes that the solution of the problem (1) inherits the $G^{(d)}$ Gevrey regularity of the data w_j , ($0 \leq j \leq \mu$).
- 1) If Ω is a bounded open set in \mathbb{R}^n , then we can extend the same study for Fuchsian operators where the coefficients $a_{(l,\alpha)}$ are polynomials of any order $N_{l,\alpha}$. In this case, the hypothesis (2) is written in the form:

$$\text{for } (l,\alpha) \in E, \quad a_{(l,\alpha)}(t,x) = \sum_{|\beta| < N_{l,\alpha}} a_{l\alpha\beta}(t) x^\beta \text{ where } a_{l\alpha\beta} \text{ is an holomorphic function in } \mathbb{C}_t.$$

Proof of theorem 1

A. Reduction of the Cauchy problem (1)

We set

$$P_1(u) = a(t, D_t)u - \sum_{(l,\alpha) \in E} a_{(l,\alpha)}(t,x) t^{v+1+l-\mu} D_t^l D^\alpha u.$$

Then by denoting $v(t,x) = u(t,x) - \sum_{j=0}^{\mu-1} \frac{t^j}{j!} w_j(x)$ and $g = f - P_1(\sum_{j=0}^{\mu-1} \frac{t^j}{j!} w_j)$;

the Cauchy problem (1) is equivalent to the problem $\begin{cases} P_1(v) = g \\ D_t^j v(t,x) = w_j(x), \quad 0 \leq j < \mu. \end{cases}$

By a second change of unknown: $v(t,x) = t^\mu z(t,x)$, the previous problem is reduced to solving the equation: $P_1(t^\mu z) = g$.

Using the relation: $t^\mu a(t, D_t)(t^\mu z) = P(tD_t)(t^\mu z) = t^\mu P(tD_t + \mu)z$; the Cauchy problem (1) is reduced to the resolution of the equation

$$P(tD_t)z = \sum_{(l,\alpha) \in E} a_{(l,\alpha)}(t,x) t^{v+1+l-\mu} D_t^l t^\mu D^\alpha z + g \quad (3)$$

and its solution is given by: $u = \sum_{j=0}^{\mu-1} \frac{t^j}{j!} w_j + t^\mu z$.

The coefficients $a_{(l,\alpha)}$ always verify the hypothesis (2). $P(tD_t)$ is a Fuchsian principal part of weight 0 which verifies $P(\lambda) \neq 0$ for every $\lambda \in \mathbb{N}$.

For proving that $g \in G^{(\omega,d)}(\mathbb{C} \times \Omega)$ we need the following lemma.

Lemma 1: Let $R > 0$ and $v \in G^{(\omega,d)}(D_R \times \Omega)$. Then for every $\alpha \in \mathbb{N}^n$, $D^\alpha v \in G^{(\omega,d)}(D_R \times \Omega)$.

For the reduction of the equation (3) we use the following lemma

Lemma 2:

1. $\exists c_0 > 0$ such that $P(\lambda) \geq c_0 \max(1, \lambda^m)$ for every $\lambda \in \mathbb{N}$.
2. For any $R > 0$, the operator $P(tD_t)$ is an automorphism of the vector space $G^{(\omega,d)}(D_{R'} \times \Omega)$ for every $0 < R' < R$. Its inverse is defined by:

$$P^{-1}(u)(t,x) = \sum_{k \in \mathbb{N}} \frac{t^k}{k!} \frac{D_t^k u(0,x)}{P(k)}.$$

In replacing z by $P^{-1}(z)$ in equation (3) and noting again z by u , then equation (3) is equivalent to:

$$u = \sum_{(l,\alpha) \in E} a_{(l,\alpha)}(t,x) t^{v+1+l-\mu} D_t^l t^\mu D^\alpha P^{-1}(u) + g \quad (4)$$

Thus, the resolution of the Cauchy problem (1) is reduced to looking for the fixed points of the map F defined by:

$$F(u) = H(u) + g \quad (5)$$

where the operator H is defined by

$$H(u) = \sum_{(l,\alpha) \in E} a_{(l,\alpha)}(t,x) t^{v+1+l-\mu} D_t^l t^\mu D^\alpha P^{-1}(u) \quad (6)$$

This takes us to introduce a Banach space associated to a majorant function where we establish that the map F is strictly contracting so we can apply the Banach fixed point theorem. We bring a new majorant function in this work compared to [3], we consider the majorant function of [9] and we apply the same technique introduced in [4], it concerns the introduction of the parameter ρ which allows us to simplify the proof presented in [3].

B. Banach space $C_{\Phi_{\rho,\rho R,\zeta}}^{\omega,\infty}(\Omega_R)$

Let D be an open neighborhood of the origine in \mathbb{C} and Φ be the formal serie $\Phi = \Phi(t,x) = \sum_{\alpha \in \mathbb{N}^n} \Phi_\alpha(t) \frac{x^\alpha}{\alpha!}$ where Φ_α is an entire serie $\gg 0$ which converges in D .

For any function $u \in C^{\omega,\infty}(U \times \Omega)$, the relation $u \ll \Phi$ is defined in [10] by

$$u \ll \Phi \Leftrightarrow (\forall \alpha \in \mathbb{N}^n), (\forall x \in \Omega), D^\alpha u(t,x) \ll \Phi_\alpha(t).$$

Let $R > 0$. For a parameter $\rho > 0$, we denote by $D_{\rho R} = \{t \in \mathbb{C}; \rho|t| < R\}$, $\Omega_{\rho,R} = D_{\rho,R} \times \Omega$ and we denote by convention $D_{\rho\rho,R} = D_R$; $\Omega_{\rho,\rho R} = \Omega_R$.

For a parameter $\zeta = (\zeta_1, \dots, \zeta_n) \in (\mathbb{R}_+^*)^n$ and a given integer $s \geq m$, we consider the Gevrey formal serie

$$\Phi_{\rho,\rho R,\zeta}^d = \Phi_{\rho,\rho R,\zeta}^d(t,x) = \sum_{p \in \mathbb{N}} (\rho t)^p (\rho R)^{s'p} \frac{(D^{sp} \phi_{\rho R,\zeta})^d(\zeta \cdot x)}{(sp)!} \quad (7)$$

where $\zeta \cdot x = (\zeta_1 \cdot x_1, \dots, \zeta_n \cdot x_n)$; $s' = s - 1 \geq 0$; ϕ is the majorant function defined by

$$\phi_{\rho R,\zeta}(\zeta \cdot x) = e^{\rho^{-1}(\zeta \cdot x)} \frac{1}{\rho R - \zeta \cdot x}. \quad (8)$$

We denote by $C_{\Phi_{\rho,\rho R,\zeta}}^{\omega,\infty}(D_R \times \Omega)$ the space of functions $u \in C^{\omega,\infty}(D_R \times \Omega)$ such that $\exists c \geq 0$; $u \ll \Phi_{\rho,\rho R,\zeta}^d$.

$C_{\Phi_{\rho,\rho R,\zeta}}^{\omega,\infty}(D_R \times \Omega)$ with the norm $\|u\|_{\Phi^d} = \min\{c \geq 0; u \ll c \Phi_{\rho,\rho R,\zeta}^d\}$ is a Banach space.

C. Proof of the contracting of F

For every compact set K in Ω of non-empty interior K° , we denote $K_R = D_R \times K^\circ$. If $\zeta = (\zeta_1, \dots, \zeta_n)$; $\zeta' = (\zeta'_1, \dots, \zeta'_n) \in (\mathbb{R}_+^*)^n$, we write $\zeta \leq \zeta'$ if $\zeta_j \leq \zeta'_j$ for every $1 \leq j \leq n$ and we write $\zeta < \zeta'$ if $(\zeta \leq \zeta' \text{ and } \zeta \neq \zeta')$.

Assume that the coefficients $a_{(l,\alpha)}$ verify the hypothesis (2) then for any function $g \in G^{(\omega,d)}(\mathbb{C} \times \Omega)$ and using some intermediate results not stated in this paper, we prove the following proposition.

Proposition 1: Let K be a fixed compact set in Ω of non-empty interior K° and let $R > 0$ be fixed. Then, for $\rho > 0$ and $\zeta \in (\mathbb{R}_+^*)^n$ verifying $\rho R > 1$ and $\zeta = (\zeta_0, \dots, \zeta_0) < (1, \dots, 1)$, there exists $\rho_0 > 0$ and a constant $c \in]0, 1[$ such that: for any $\rho \geq \rho_0$, there exists $b_\rho > 0$ for which

$$\forall b \geq b_\rho; F(B(0,b)) \subset B(0,b) \subset C_{\Phi_{\rho,\rho R,\zeta}}^{\omega,\infty}(K_R) \quad (9)$$

$$\forall u, u' \in C_{\Phi_{\rho, \rho R, \zeta}}^{\omega, \infty}(K_R); \quad \|F(u) - F(u')\|_{\Phi^d} \leq c \|u - u'\|_{\Phi^d} \quad (10)$$

where $B(0, b)$ is the closed ball of center 0 and radius b of $C_{\Phi_{\rho, \rho R, \zeta}}^{\omega, \infty}(K_R)$.

Under the hypotheses of this proposition, we deduce from the fixed point theorem that the map F admits a unique fixed point in $C_{\Phi_{\rho, \rho R, \zeta}}^{\omega, \infty}(K_R)$.

D. Construction of the fixed point of F in Ω_R

In the rest of the proof we fix $\zeta = (\zeta_0, \dots, \zeta_0) < (1, \dots, 1)$.

For the construction of the fixed point of F in Ω_R , we use the three following statements.

Lemma 3: Let $R > 0$ and let K be a compact set in Ω of non-empty interior K° . For any $\rho, \rho' > 0$ such that $\rho \geq \rho'$ we have

$$C_{\Phi_{\rho, \rho R, \zeta}}^{\omega, \infty}(K_R) \subset C_{\Phi_{\rho', \rho' R, \zeta}}^{\omega, \infty}(K_R)$$

and the canonical function of the inclusion is continuous of norm inferior to 1.

Proposition 2: Let $(K_j)_{j \in \mathbb{N}}$ be an exhaustive sequence of compact sets in Ω , then there exist an increasing sequence of positive numbers $\rho_j = \rho_{K_j}$ and a sequence $(u_j)_{j \in \mathbb{N}}$ of fixed points of F such that

- i) u_j is unique in $C_{\Phi_{\rho_j, \rho_j R, \zeta}}^{\omega, \infty}(D_R \times K_j^\circ)$,
- ii) $u_j \in \bigcap_{\rho \geq \rho_j} C_{\Phi_{\rho, \rho R, \zeta}}^{\omega, \infty}(D_R \times K_j^\circ)$.

Lemma 4: There exists a unique fixed point u of the strictly contracting F defined in $D_R \times \Omega$ satisfying

$$(\forall j \in \mathbb{N}), \quad u / K_j^\circ \in \bigcap_{\rho \geq \rho_j} C_{\Phi_{\rho, \rho R, \zeta}}^{\omega, \infty}(D_R \times K_j^\circ).$$

Next, we prove that this solution u defined in $D_R \times \Omega$ in lemma 4 is in the class $G^{(\omega, d)}(D_R \times \Omega)$. Then, using the principle of analytic continuation we show for every $R > 0$, the uniqueness of this fixed point of F in the class $G^{(\omega, d)}(D_R \times \Omega)$.

E. End of the proof of theorem 1

For every $R > 0$, the map F admits a unique fixed point u_R in $G^{(\omega, d)}(D_R \times \Omega)$. Then, using the reattachment of the solutions u_R , we define a unique fixed point of F in $G^{(\omega, d)}(\mathbb{C} \times \Omega)$.

To complete the proof of the theorem 1, we take $t^\mu u + \sum_{j=0}^{\mu-1} \frac{t^j}{j!} w_j$ as the unique solution in $G^{(\omega, d)}(\mathbb{C} \times \Omega)$ of the Cauchy problem (1).

Conclusion

This technique of introducing the parameter $\rho > 0$, allows us to define new majorant functions in this paper to simplify the proof given in [3]. We have also used this same technique in [5] with success to study some nonlinear equations of Fuchs type in this same class of Gevrey. Also we have used this technique in [4] to simplify the proof of [9] for linear Fuchsian operators with several variables.

References:

- [1] J. M. Bony and P. Schapira, "Existence et prolongement des solutions holomorphes des équations aux dérivées partielles," *Invent. Math.*, vol.17, 1971, pp. 95-105.
- [2] M. S. Baouendi and C. Goulaouic, "Cauchy problem with characteristic initial hypersurface," *Comm. on Pure and Appl. Math.*, vol. 26, 1973, pp. 455-475.

- [3] F. Derrab, A. Nabaji, P. Pongérard and C. Wagschal, “Problème de Cauchy Fuchsien dans les espaces de Gevrey,” *J. Math. Sci. Univ. Tokyo*, vol. 11, 2004, pp. 401–424.
- [4] F. Derrab and A. Nabaji, “Solutions holomorphes locale et globale pour un opérateur différentiel linéaire à plusieurs variables fuchsiennes,” *Osaka J. Math.*, vol. 42, 2005, pp. 653–675.
- [5] F. Derrab, “Solutions globales de certaines équations de Fuchs non linéaires dans les espaces de Gevrey,” *Pro. Am. Math. Soc.*, vol. 135, 2007, pp. 1803–1815.
- [6] H. Komatsu, “Linear hyperbolic equations with Gevrey coefficients,” *J. Math. Pures et Appl.*, vol. 59, 1980, pp. 145-185.
- [7] P. Pongérard and C. Wagschal, “Problème de Cauchy dans des espaces de fonctions entières,” *J. Math. Pures et Appl.*, vol. 75, 1996, pp. 409-418.
- [8] P. Pongérard, “Sur une classe d’équations de Fuchs non -linéaires,” *J. Math. Sci. Univ. Tokyo*, vol. 7, 2000, pp. 423-448.
- [9] P. Pongérard, “Problème de Cauchy caractéristique à solution entière,” *J. Math. Sci. Univ. Tokyo*, vol. 8, 2001, pp. 89-105.
- [10] C. Wagschal, “Le problème de Goursat non linéaire,” *J. Math. Pures et Appl.*, vol. 58, 1979, pp. 3096337.

ALGEBRAIC SCHUR COMPLEMENT APPROACH FOR A NON LINEAR 2D ADVECTION DIFFUSION EQUATION

Hassan Belhadj, Professor
Samir Khallouq, PhD Candidate

University of Abdelmalek Essaadi, FST, Department of Mathematics, Tangier, Morocco

Abstract:

This work deals with a domain decomposition approach for non stationary non linear advection diffusion equation. The domain of calculation is decomposed into $q \geq 2$ non-overlapping sub-domains. On each sub-domain the linear part of the equation is discretized using implicit finite volumes scheme and the non linear advection term is integrated explicitly into the scheme. As non-overlapping domain decomposition, we propose the Schur Complement (SC) Method. The proposed approach is applied for solving the local boundary sub-problems. The numerical experiments applied to Burgers equation show the interest of the method compared to the global calculation. The proposed algorithm has both the properties of stability and efficiency. It can be applied to more general non linear PDEs and can be adapted to different FV schemes.

Key Words: Non linear advection-diffusion problems, Structured mesh, Burgers equation, Finite volumes method (FVM), Schur Complement (SC)

The system of equations

Let us consider the following initial boundary value problem:

Find $c: \Omega \times (0, T) \rightarrow \mathbb{R}$ such that

$$(1.1) \quad \begin{cases} \frac{\partial c}{\partial t} - \nu \Delta c + \sum_{s=1}^2 \frac{\partial f_s(c)}{\partial x_s} = g & \text{in } \Omega \times (0, T) \\ c(x, t) = c_D(x, t) & \text{on } \partial\Omega \times (0, T) \\ c(x, 0) = c_0(x) & \text{in } \Omega \end{cases}$$

Where $\Omega \subseteq \mathbb{R}^2$ is a bounded polygonal domain and $(0; T)$, where $T > 0$, time interval. By $\bar{\Omega}$ and $\partial\Omega$ we denote the closure and boundary of Ω , respectively.

We assume that the data have the following properties [6, 7, 8]:

- $f_s \in C^1(\mathbb{R})$, $f_s(0) = 0$, $|f_s| \leq C_{f_s}$, $s = 1, 2$,
- $\nu > 0$,
- $g \in C([0, T]; L^2(\Omega))$
- C_D is the trace of some $C^* \in C([0, T]; H^1(\Omega)) \cap L^\infty(\Omega \times (0, T))$ on $\partial\Omega \times (0, T)$,
- $c_0 \in L^2(\Omega)$.

In virtue of assumption a), the functions f_s satisfy the Lipschitz condition with constant C_{f_s} , the functions f_s are fluxes of the quantity c in the direction x_s , its represent convective terms, the constant $\nu > 0$ is the diffusion coefficient.

We use the standard notation for function spaces (see, e.g. [9]): $L^p(\Omega)$, $L^p(\Omega \times (0, T))$ denote the Lebesgue spaces, $W^{k,p}(\Omega)$, $H^k(\Omega) = W^{k,2}(\Omega)$ are the Sobolev spaces, $L^p(0, T; X)$ is the Bochner space of functions p -integrable over the interval $(0, T)$ with values in a Banach space X ,

$C([0, T]; X)$ ($C^1([0, T]; X)$) is the space of continuous (continuously differentiable) mappings of the interval $[0, T]$ into X .

We shall assume that problem (1.1) has a weak solution (cf. [6,7]), satisfying the regularity conditions:

$$(1.2) \quad c, \frac{\partial c}{\partial t}, \frac{\partial^2 c}{\partial t^2} \in L^\infty(0, T; H^{p+1}(\Omega))$$

where an integer $p \geq 1$ will denote a given degree of polynomial approximations. Such a solution satisfies problem (1.1) pointwise. Under (1.2),

$$c \in C([0, T]; H^{p+1}(\Omega)), \quad \text{and} \quad \frac{\partial c}{\partial t} \in C([0, T]; L^2(\Omega))$$

Finite volume approach

The finite volumes approach consists in dividing the domain of calculation Ω into a finite number of control volumes (CVs) V_i ($i=1, \dots, N \times M$) with $\Omega = \bigcup_{i=1}^{N \times M} V_i$.

For a general CV we use the notation of the distinguished points (mid-point, midpoints of faces) and the unit normal vectors according to the notation as indicated in Figure 1 (right). The midpoints of neighboring CVs we denote with capital letters W, S, etc. (see Figure 1 left), these notations are given in [3].

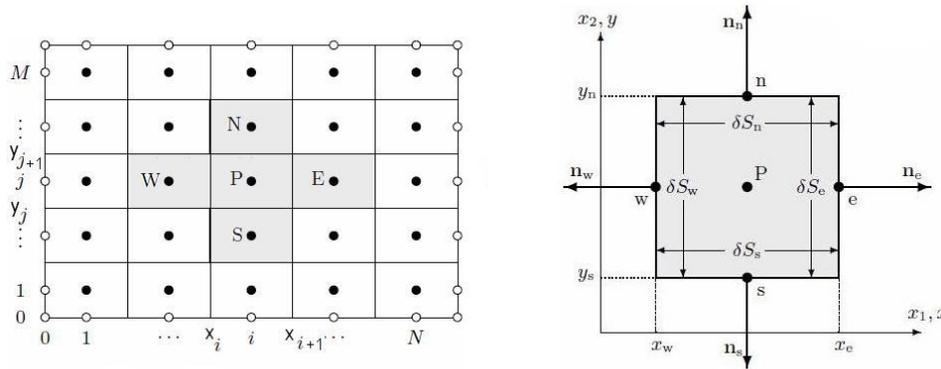


Figure 1. FV structured mesh of domain Ω

By integrating the equation (1.1) over an arbitrary CV V_P and applying the Green formula, we obtain:

$$(2.1) \quad \int_{V_P} \frac{\partial c}{\partial t}(t) dV_P + \sum_a \int_{S_{a,P}} \sum_{s=1}^2 f_s(c(t)) n_{a,s} dS_{a,P} - \nu \sum_a \int_{S_{a,P}} \nabla c(t) n_a dS_{a,P} = \int_{V_P} g(t) dV_P,$$

where $S_{a,P}$ ($a = e, n, w, s$) are the four faces of volume V_P (see Figure 1), $n_a = (n_{a,1}, n_{a,2})$ are the unit normal vectors to the face $S_{a,P}$ and $\mu(V_P)$ is the volume of cell V_P .

Approximating the linear operator $\partial_t - \nu \Delta$ by the implicit Euler method and the non-linear term by an explicit approximation, we get:

$$(2.2) \quad \mu(V_P) \frac{c_P^{n+1} - c_P^n}{\Delta t} + \sum_a \int_{S_{a,P}} \sum_{s=1}^2 f_s(c^n) n_{a,s} dS_{a,P} - \nu \sum_a \int_{S_{a,P}} \nabla c^{n+1} n_a dS_{a,P} = \mu(V_P) g_P^n,$$

where

$$g_P^n = \frac{1}{\mu(V_P)} \int_{V_P} g(x, t^n) dV_P,$$

and

$$c_P^0 = \frac{1}{\mu(V_P)} \int_{V_P} c_0(x) dV_P, \quad \text{or} \quad c_P^0 = c_0(x_P).$$

- For the discretization of diffusion term, we have considered a centred difference scheme.
- For the convective terms we use the numerical flux, for the CV V_P and $S_{a,P}$ ($a = e, n, w, s$):

$$(2.3) \quad \sum_{s=1}^2 f_s(c^n) n_{a,s} = \begin{cases} \sum_{s=1}^2 f_s(c_P^n) n_{a,s} & \text{if } K > 0, \\ \sum_{s=1}^2 f_s(c_I^n) n_{a,s} & (I = E, W, S, N) \text{ if } K \leq 0, \end{cases}$$

where

$$K = \sum_{s=1}^2 f_s(\bar{c}^n) n_{a,s}, \quad \bar{c}^n = \frac{1}{2}(c_P^n + c_I^n).$$

- For the approximation of the volume and surface integrals, we have employed the midpoint rule.

Let us denote that c_I^n is the concentration on the volume V_P ($I=P, E, W, N$ or S) at time t_n . The concentration variables c_I^{n+1} and c_I^n ($I=P, E, W, N$ or S) in equation (2.2) can be arranged as follows:

$$(2.4) \quad a_P c_P^{n+1} + a_E c_E^{n+1} + a_W c_W^{n+1} + a_N c_N^{n+1} + a_S c_S^{n+1} = b_P,$$

b_P is a constant depending on, the source term g_P^n , c_P^n , the discretized convection flux, the boundary and the initial conditions.

Finally, the numerical scheme is expressed as the linear system:

$$AC_P^{n+1} = b,$$

where A is a $(N \times M, N \times M)$ type matrix of coefficients a_I ($I=P, E, W, N$ or S), C_P^{n+1} and b are the vectors of c_P^{n+1} and b_P respectively.

Schur complement method

Domain decomposition

The domain Ω is decomposed into multi-domain nonoverlapping strip decomposition $\Omega_1, \dots, \Omega_q$ where $\bar{\Omega} = \bigcup_{i=1}^q \bar{\Omega}_i$ and $\Omega_i \cap \Omega_j = \emptyset$ when $i \neq j$ (figure 2).

Let Γ_{ij} denote the interface between Ω_i and Ω_j and $\Gamma = \bigcup \Gamma_{ij}$, and by n^i the normal direction (oriented outward) on Γ_{ij} for $i=1, \dots, q-1$ and $j=i+1$.

For simplicity of notation we also set $n = n^i$.

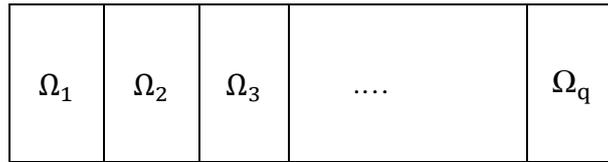


Figure 2. Non-overlapping strip decomposition

Considering a rectangular mesh of Ω , each subdomain Ω_i is partitioned into n_i ($i=1, \dots, q$) cells in X direction and m cells in Y direction (figure 3).

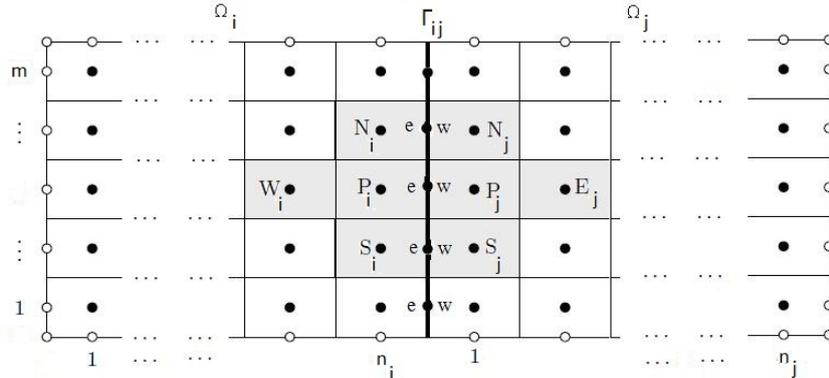


Figure 3. Domain decomposition and structured conforming mesh of domain Ω

The problem (1.1) can then be expressed as :

$$(3.1) \quad \begin{cases} \frac{\partial c_i}{\partial t} - \nu \Delta c_i + \sum_{s=1}^2 \frac{\partial f_s(c_i)}{\partial x_s} = g & \text{in } \Omega_i \times (0, T), i = 1, \dots, q \\ c_i(x, t) = c_D(x, t) & \text{on } (\partial\Omega_i - (\Gamma_{i+1} \cup \Gamma_{i-1})) \times (0, T) \\ c_i(x, 0) = c_0(x) & \text{in } \Omega_i \\ c_i = c_j & \text{on } \Gamma_{ij} \quad i, j = 1, \dots, q \\ \frac{\partial c_i}{\partial n} = \frac{\partial c_j}{\partial n} & \text{on } \Gamma_{ij} \end{cases}$$

The last two interface conditions are known as transmission conditions on Γ_{ij} .

The decomposed problem (3.1) is discretized on each sub-domain Ω_i , $i=1, \dots, q$ using the implicit finite volume scheme described in Section 2. For the interface conditions we have used the centred differences scheme. We obtain the following system for $i=1, \dots, q-1$ and $j=i+1$:

$$(3.2) \quad \begin{cases} a_{P_i} c_{P_i}^{n+1} + a_{W_i} c_{W_i}^{n+1} + a_{N_i} c_{N_i}^{n+1} + a_{S_i} c_{S_i}^{n+1} = b_{P_i} & \text{in } \Omega_i \quad (a) \\ a_{P_j} c_{P_j}^{n+1} + a_{E_j} c_{E_j}^{n+1} + a_{N_j} c_{N_j}^{n+1} + a_{S_j} c_{S_j}^{n+1} = b_{P_j} & \text{in } \Omega_j \quad (b) \\ c_{e_i}^{n+1} = c_{w_j}^{n+1} & \text{on } \Gamma_{ij} \quad (c) \\ c_{e_i}^{n+1} + c_{e_j}^{n+1} - c_{p_i}^{n+1} - c_{p_j}^{n+1} = 0 & \text{on } \Gamma_{ij} \quad (d) \end{cases}$$

where

$$\begin{cases} \sigma_i = e_i \text{ and } \sigma_j = e_j & \text{if } V_{P_i} \cap \Gamma_{ij} \neq \emptyset \quad (i = 1, \dots, q-1) \\ \sigma_i = E_i \text{ and } \sigma_j = W_j & \text{else} \end{cases}$$

b_{P_i} is a constant depending on, the source term $g_{P_i}^n$, $c_{P_i}^n$, the discretized convection flux, the boundary and the initial conditions in Ω_i , $i=1, \dots, q$.

Schur complement

The methods based on Schur Complement exists in two versions. The first one uses the Steklov Poincaré operator and the second one is an algebraic version.

For example in [1, 2, 4] and in [5], one finds presentations of these methods (for linear advection diffusion equation) used in the context of a finite elements method and finite volumes method, respectively.

In this work, we have used an algebraic version of Schur Complement technique.

Let C_i^{n+1} and C_Γ^{n+1} denote the vector of the unknowns of Ω_i ($i=1, \dots, q$) and Γ at time t_{n+1} (respectively), and b_i denote the vector of b_{P_i} .

The decomposed problem (3.2) can be written in the following matrix form:

$$(3.3) \quad \begin{bmatrix} A_1 & 0 & \dots & 0 & A_{1\Gamma} \\ 0 & A_2 & \dots & 0 & A_{2\Gamma} \\ \cdot & \cdot & \dots & \cdot & \cdot \\ 0 & \dots & 0 & A_q & A_{q\Gamma} \\ A_{\Gamma 1} & A_{\Gamma 2} & \dots & A_{\Gamma q} & A_{\Gamma\Gamma} \end{bmatrix} \begin{bmatrix} C_1^{n+1} \\ C_2^{n+1} \\ \cdot \\ C_q^{n+1} \\ C_\Gamma^{n+1} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ b_q \\ 0 \end{bmatrix}$$

with

A_i , $A_{i\Gamma}$ describe respectively (a) and (b) of system (3.2), and $A_{\Gamma i}$, $A_{\Gamma\Gamma}$ ($i=1, \dots, q$) describe respectively (c) and (d) of system (3.2).

The matrix A_i present the coupling of the unknowns in Ω_i , $A_{\Gamma\Gamma}$ it is related to the unknowns on the interface, $A_{\Gamma i}$ and $A_{i\Gamma}$ representing the coupling of the unknowns of each sub-domain Ω_i with those of the interface $\Gamma_{i,i+1}$ for ($i=1, \dots, q-1$).

The system (3.3) can be sought formally by block Gaussian elimination.

Eliminating C_i^{n+1} ($i=1,\dots,q$) in the system (3.3), yields the following reduced linear system for C_Γ^{n+1} :

$$(3.4) \quad SC_\Gamma^{n+1} = \chi_\Gamma,$$

where

$$\chi_\Gamma = - \sum_{i=1,\dots,q} A_{\Gamma i} A_i^{-1} b_i,$$

and

$$S = A_{\Gamma\Gamma} - \sum_{i=1,\dots,q} A_{\Gamma i} A_i^{-1} A_{i\Gamma},$$

S is the Schur Complement matrix.

After calculating, C_Γ^{n+1} , C_i^{n+1} can be obtained immediately and independently (in parallel) by solving

$$(3.5) \quad A_i C_i^{n+1} = b_i - A_{i\Gamma} C_\Gamma^{n+1} \quad (i=1,\dots,q)$$

Numerical Simulations

In this section, we shall verify the proposed approach by numerical experiments.

Let us apply FV mono-domain (FV-MonoD) and the combined FV method Schur Complement (FV-SC) to the 2D viscous Burgers equation [6, 7, 8]:

$$(4.1) \quad \frac{\partial c}{\partial t} - \nu \Delta c + c \frac{\partial c}{\partial x_1} + c \frac{\partial c}{\partial x_2} = g,$$

The spatial domain is the square $\Omega_i = (-1,1)^2$, the time interval $T = (0,1)$, $\nu = 0.01$, the initial data $c_0 = 0$ and the Dirichlet conditions $c_D = 0$. The right-hand side g is chosen so that it conforms to the exact solution [8]:

$$c(x, t) = (1 - e^{-2t})(1 - x_1^2)^2(1 - x_2^2)^2$$

As we want to examine the error of the space discretization, we overkill the time step so that the time discretization error is negligible.

Figure 4 (a,b,c,d) show respectively the analytical, the numerical mono-domain, the multi-domain ($q=2$) and the multi-domain ($q=9$) solutions.

Figure 5 shows the convergence of the proposed algorithm when varying the mesh of calculation.

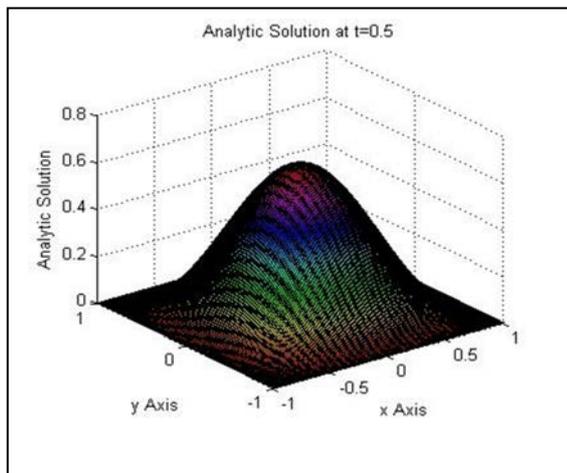


Figure 4a. Analytical solution

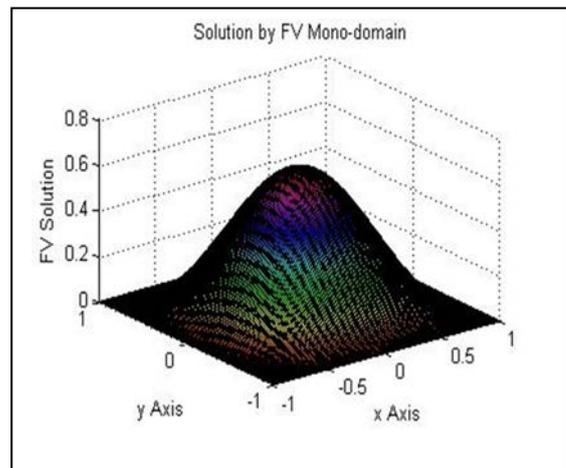


Figure 4b. Numerical mono-domain solution

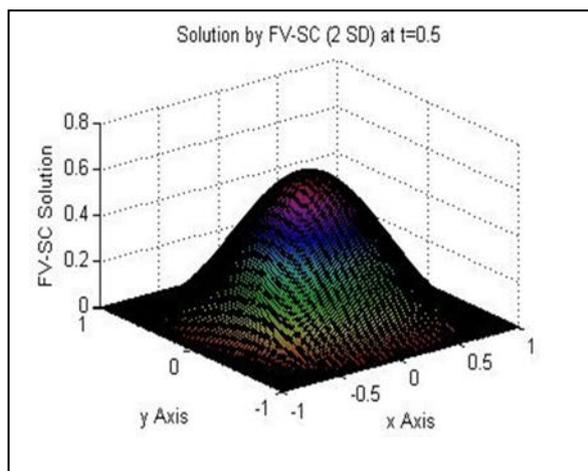


Figure 4c. Numerical multi-domain (q=2) solution

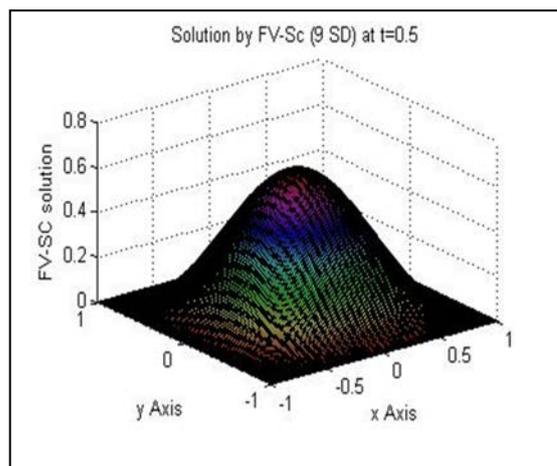


Figure 4d. Numerical multi-domain (q=9) solution

Figure 4. Numerical and analytical solution

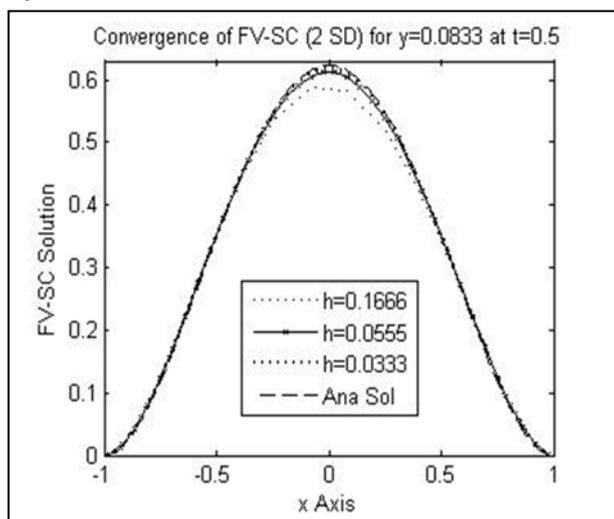
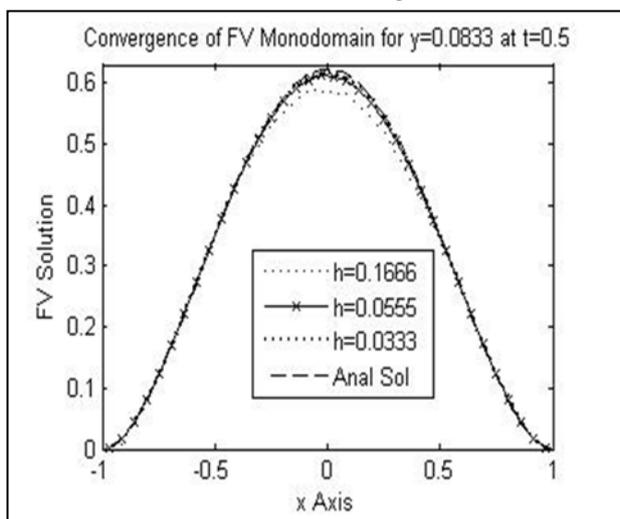


Figure 5. Convergence of numerical scheme

Conclusion

A new approach coupling implicit FV and Algebraic Schur Complement methods applied to a semi linear advection-diffusion equation, on 2D structured and conforming mesh, is presented.

The numerical experiments show that the proposed algorithm applied to a non-overlapping multi-subdomain decomposition has both the properties of stability and accuracy.

On the other hand, it reduces the calculation cost compared to global FV calculation.

As perspective of this work we project to develop a new algorithm integrating the non linear advection part implicitly. This algorithm will include for example Newton method to compute the advection term after each time step of the numerical scheme.

References:

[1] A. Quarteroni, A.Valli, Domain decomposition methods for partial differential equations, Clarendon Press, Oxford, 1999.
 [2] A. Quarteroni, A.Valli, Theory and application of stecklov-poincarre operators for boundary value problems, 179-203. Kluwer Academic Publishers, 1991.
 [3] Michael Schafer, Computational Engineering - Introduction to Numerical Methods. Springer-Verlag, Berlin, Heidelberg, 2006.

- [4] P. Tarek, A. Mathew, *Lecture Notes in Computational Science and Engineering 61- Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations*. Springer-Verlag, Berlin, Heidelberg, 2008.
- [5] S. Khallouq and H. Belhadj, Schur Complement Technique for Advection-Diffusion Equation Using Matching Structured Finite Volumes, *Advances in Dynamical Systems and Applications* ISSN 0973-5321, Volume 8, Number 1, pp. 51-62 (2013).
- [6] V. Dolejší, M. Feistauer, J. Hozman, Analysis of semi-implicit DGFEM for nonlinear convection-diffusion problems on nonconforming meshes, *Comput. Methods Appl. Mech. Engrg.* 196 (2007) 2813-2827
- [7] V. Dolejší, M. Feistauer, V. Sobotíková, A discontinuous Galerkin method for nonlinear convection-diffusion problems, *Comput. Methods Appl. Mech. Engrg.* 194 (2005) 2709.2733.
- [8] M. Bejček, M. Feistauer, T. Gallouët, J. Hájek, and R. Herbin, Combined triangular FV-triangular FE method for nonlinear convection-diffusion problems, *ZAMM Z. Angew. Math. Mech.* 87, No. 7, 499-517 (2007) / DOI 10.1002/zamm.200610332
- [9] A. Kufner, O. John, S. Fučík, *Function Spaces*, Academia, Prague, 1977.

MAX-MIN EDGE MAGIC AND ANTIMAGIC LABELING

J. Jayapriya, Research Scholar

Department of Mathematics, Sathyabama University, India

K. Thirusangu

Department of Mathematics, S.I.V.E.T College, Gowrivakkam, India

Abstract:

In this paper we introduce a new type of edge magic and antimagic labeling and study the same for various classes of graphs.

Key Words: Graph, labeling, function, magic, antimagic

Introduction

A graph labeling is an assignment of integers to the vertices or edges or both subject to certain conditions. Labeled graphs are becoming an increasingly useful family of Mathematical Models from a broad range of applications. The graph labeling problem that appears in graph theory has a fast development recently. This problem was first introduced by Alex Rosa in 1967. Since Rosa's article, many different types of graph labeling problems have been defined around this. This is not only due to its mathematical importance but also because of the wide range of the applications arising from this area, for instance, x-rays, crystallography, coding theory, radar, astronomy, circuit design, and design of good Radar Type Codes, Missile Guidance Codes and Convolution Codes with optimal autocorrelation properties and communication design. An enormous body of literature has grown around the subject in about 1500 papers. They gave birth to families of graphs with attractive names such as graceful, Harmonious, felicitous, elegant, cordial, magic antimagic, bimagic and prime labeling etc. A useful survey to know about the numerous graph labeling methods is the one by J.A. Gallian recently [2012][1]. All graphs considered here are finite simple and undirected.

The vertex-weight of a vertex v in G under an edge labeling is the sum of edge labels corresponding to all edges incident with v . Under a total labeling, vertex-weight of v is defined as the sum of the label of v and the edge labels corresponding to the entire edges incident with v . If all vertices in G have the same weight k , we call the labeling vertex-magic edge labeling or Vertex-magic total labeling respectively and we call k a magic constant. If all vertices in G have different weights, then the labeling is called vertex-antimagic edge labeling or vertex-antimagic total labeling. The edge-weight of an edge e under a vertex labeling is defined as the sum of the vertex labels corresponding to every vertex incident with e under a total labeling, we also add the label of e . Using edge-weight, we derive edge-magic vertex or edge-magic total labeling and edge-antimagic vertex or edge-antimagic total labeling.

Definition 1.1: A (p, q) graph G is said to be $(1, 0)$ edge-magic with the common edge count k if there exists a bijection $f: V(G) \rightarrow \{1, \dots, p\}$ such that for all $e = (u, v) \in E(G)$, $f(u) + f(v) = k$. It is said to be $(1, 0)$ edge-antimagic if for all $e = (u, v) \in E(G)$, $f(u) + f(v)$ are distinct.

Definition 1.2: A (p, q) graph G is said to be $(0, 1)$ vertex-magic with the common vertex count k if there exists a bijection $f: E(G) \rightarrow \{1, \dots, q\}$ such that for each $u \in V(G)$, $e \in \Sigma f(e) = k$ for all $e = (u, v) \in E(G)$ with $v \in V(G)$. It is said to be $(0, 1)$ vertex-antimagic if for each $u \in V(G)$, $e \in \Sigma f(e)$ are distinct for all $e = (u, v) \in E(G)$ with $v \in V(G)$.

Definition 1.3: A (p, q) graph G is said to be $(1, 1)$ edge-magic with the common edge count k if there exists a bijection $f: V(G) \cup E(G) \rightarrow \{1, \dots, p+q\}$ such that $f(u) + f(v) + f(e) = k$ for all $e = (u, v) \in E(G)$. It is said to be $(1, 1)$ edge-antimagic if $f(u) + f(v) + f(e)$ are distinct for all $e = (u, v) \in E(G)$.

In this paper we introduce a new type of edge magic and antimagic labeling and study the same for various classes of graphs.

Main Results

Definition 2.1: Let $G(V,E)$ be a simple graph with p vertices and q edges. An injective function $f : V(G) \rightarrow N$, is said to be max-min edge magic labeling if for every edge uv in E , $\max\{f(u),f(v)\}/\min\{f(u),f(v)\}$ is a constant where N is the set of natural numbers. f is said to be an max-min edge antimagic labeling if $\max\{f(u),f(v)\}/\min\{f(u),f(v)\}$ is distinct.

Theorem 2. 2: P_n admits max-min edge magic labeling.

Proof: Let the vertices $V=\{v_1,v_2,\dots,v_n\}$ and Let $f: V \rightarrow \{a^i\}$, $a \geq 2$, $1 \leq i \leq n$ $f(v_i) = a^i$, such that for all $xy \in E, f(y) = af(x)$ for constant a .

We calculate the edge labels $\lambda(uv) = \max\{f(u),f(v)\}/\min\{f(u),f(v)\}$.

For $1 \leq i \leq n$, $\lambda(v_i v_{i+1}) = \max\{f(v_i),f(v_{i+1})\}/\min\{f(v_i),f(v_{i+1})\} = a^{i+1}/a^i = a$.

Hence P_n admits max-min edge magic labeling.

Theorem 2.3: P_n admits max-min edge antimagic labeling for all n .

Proof: Let the vertices $V=\{v_1,v_2,\dots,v_n\}$ Let $f: V \rightarrow N$ such that $f(v_i) = i$; $1 \leq i \leq n$, and the edge set $E = \{v_i v_{i+1}; 1 \leq i \leq n-1\}$.

For $1 \leq i \leq n-1$, let $\lambda(v_i v_{i+1}) = \max\{f(v_i),f(v_{i+1})\}/\min\{f(v_i),f(v_{i+1})\} = (i+1)/i$,

For $1 \leq i, j \leq n-1$ clearly $\lambda(v_i v_{i+1}) \neq \lambda(v_j v_{j+1})$.

If $\lambda(v_i v_{i+1}) = \lambda(v_j v_{j+1})$

then $(i+1)/i = (j+1)/j$

$\Rightarrow ij + j = ij + i$

$\Rightarrow i = j$ which is a contradiction. Hence all edges are distinct, proving the theorem.

Theorem 2.4: C_n admits max-min edge antimagic labeling for all n .

Proof: The cycle graph has n vertices and n edges. Let the vertices $V=\{v_1,v_2,\dots,v_n\}$ Let $f: V \rightarrow N$ such that $f(v_i) = i$; $1 \leq i \leq n$, and

the edge set $E = \{v_i v_{i+1}; 1 \leq i \leq n-1\} \cup \{v_1 v_n\}$

for $1 \leq i \leq n-1$, $\lambda(v_i v_{i+1}) = \max\{f(v_i),f(v_{i+1})\}/\min\{f(v_i),f(v_{i+1})\} = (i+1)/i$,

and $\lambda(v_1 v_n) = \max\{f(v_1),f(v_n)\}/\min\{f(v_1),f(v_n)\} = n$.

Since $\lambda(v_i v_{i+1}) \neq \lambda(v_1 v_n)$,

If $\lambda(v_i v_{i+1}) = \lambda(v_1 v_n)$,

then $\Rightarrow (i+1)/i = n$

$\Rightarrow i+1 = ni$ which is a contradiction.

clearly all edges are distinct proving the theorem

Theorem 2.5: C_n^+ admits max-min edge antimagic labeling.

Proof: Let the vertices $V = \{v_1, v_2, \dots, v_{2n}\}$ Let $f: V \rightarrow N$ such that

$f(v_i) = 2i - 1$; $1 \leq i \leq n$, $f(v_i) = 2i$; $n+1 \leq i \leq 2n$.

and the edge set $E = \{v_i v_{i+1}; 1 \leq i \leq n-1\} \cup \{v_1 v_n\} \cup \{v_i v_{i+1}; n+1 \leq i \leq 2n\}$

for $1 \leq i \leq n-1$, $\lambda(v_i v_{i+1}) = \max\{f(v_i),f(v_{i+1})\}/\min\{f(v_i),f(v_{i+1})\} = (2i+1)/(2i-1)$.

also $\lambda(v_1 v_n) = \max\{f(v_1),f(v_n)\}/\min\{f(v_1),f(v_n)\} = n$.

and for $n+1 \leq i \leq 2n$, $\lambda(v_i v_{i+1}) = \max\{f(v_i),f(v_{i+1})\}/\min\{f(v_i),f(v_{i+1})\} = 2i+2/2i = i+1/i$.

Clearly all edges are distinct proving C_n^+ admits max-min edge antimagic labeling.

Theorem 2.6: The wheel graph admits max-min edge antimagic labeling.

Proof: The wheel graph has n vertices and $2(n-1)$ edges. Let the vertices $V = \{v_1, v_2, \dots, v_n\}$ Let $f: V \rightarrow N$ such that $f(v_1) = 1$, $f(v_i) = 2i-1$; $2 \leq i \leq n-1$, and the edge set $E = \{v_1 v_i; 2 \leq i \leq n-1\} \cup \{v_i v_{i+1}; 2 \leq i \leq n-1\} \cup \{v_2 v_n\}$.

For $2 \leq i \leq n-1$, $\lambda(v_1 v_i) = \max\{f(v_1),f(v_i)\}/\min\{f(v_1),f(v_i)\} = i/1 = i$,

For $2 \leq i \leq n-1$, $\lambda(v_i v_{i+1}) = \max\{f(v_i),f(v_{i+1})\}/\min\{f(v_i),f(v_{i+1})\} = (i+1)/i$,

and $\lambda(v_2 v_n) = \max\{f(v_2),f(v_n)\}/\min\{f(v_2),f(v_n)\} = n/2$.

Clearly all edges are distinct, proving the theorem.

Theorem 2. 7: $K_{1,n}$ admits max-min edge antimagic labeling .

Proof: Let the vertices $V = \{v_1, v_2, \dots, v_{n+1}\}$, and the edge set be $E = \{v_1 v_i; 2 \leq i \leq n+1\}$ now for $2 \leq i \leq n+1$, $\max\{f(v_1),f(v_i)\}/\min\{f(v_1),f(v_i)\} = i$. Clearly all edges are distinct and also are in Arithmetic progression, hence : $K_{1,n}$ admits max-min edge antimagic labeling .

Theorem 2. 8: The comb graph admits max-min edge antimagic labeling.

Proof: Let the vertices $V = \{v_1, v_2, \dots, v_n, u_1, u_2, \dots, u_n\}$. Let $f: V \rightarrow \mathbb{N}$ such that $f(v_i) = 2i-1; 1 \leq i \leq n$ and $f(u_i) = 2i; 1 \leq i \leq n$. Let $E = \{v_i u_i; 1 \leq i \leq n\} \cup \{v_i v_{i+1}; 1 \leq i \leq n-1\}$,

and also for $1 \leq i \leq n$ $\lambda(u_i v_i) = \max\{f(u_i), f(v_i)\} / \min\{f(u_i), f(v_i)\} = 2i / (2i-1)$ and for $1 \leq i \leq n-1$ $\lambda(v_i v_{i+1}) = \max\{f(v_i), f(v_{i+1})\} / \min\{f(v_i), f(v_{i+1})\} = 2i+1 / (2i-1)$ are distinct.

Hence the comb graph admits max-min edge antimagic labeling.

Corollary: binary tree, C_t^n (C_n), gear graph, C_n (C_n), corona, Sunflower graph, one vertex union of 't' isomorphic copies of any cycle is max-min edge antimagic labeling.

Conclusion

Max-Min edge antimagic labeling is studied for paths, cycle graph wheel graph, star graphs, binary tree, C_t^n (C_n), gear graph, C_n (C_n), corona, Sunflower graph, one vertex union of 't' isomorphic copies of any cycle. Also max-min edge magic labeling is observed for path graphs. Investigating max-min edge magic labeling for other classes is our future work.

References:

- [1] J.A Gallian, "A Dynamic Survey of Graph Labeling" Electronic Journal of Combinatorics, 19, 2012, # DS6.
- [2] F. Harary, Graphy Theory, Addison Wesley, Reading, Massachusetts, 1969.
- [3] A. Rosa, on certain valuation of the vertices of a graph, *Theory of Graphs* (Internat. Symposium, Rome, July 1966), Gordon and Breach, N.Y. and Dunond Paris (1967) 349-355.
- [4] W.D. Wallis, E.T. Baskoro, M. Miller and Slamin, Edge magic total labelings, *Australasian Journal of Combin.* 22 (2000) 177-190.

ON THE TOPOLOGICAL STRUCTURES ON Γ -SEMIRINGS

Kostaq Hila, Asoc. Prof.

Department of Mathematics & Computer Science, Faculty of Natural Sciences,
University of Gjirokastra, Albania

Ilir Vardhami, PhD

Kristaq Gjino, PhD

Department of Mathematics, Faculty of Natural Sciences, University of Tirana, Tirana, Albania.

Abstract:

In this paper, we introduce some special classes of ideals in Γ -semirings called prime k -ideal, prime full k -ideal, prime ideals, maximal and strongly irreducible ideals. Considering and investigating properties of the collection \mathcal{A} , \mathcal{T} , \mathcal{M} , \mathcal{B} and \mathcal{S} of all proper prime k -ideals, proper prime full k -ideals, maximal ideals, prime ideals and strongly irreducible ideals, respectively, of a Γ -semiring R , we construct the respective topologies on them by means of closure operator defined in terms of intersection and inclusion relation among these ideals of Γ -semiring R . The respective obtained topological spaces are called the structure spaces of the Γ -semiring R . We study a several principal topological axioms and properties in those structure spaces of Γ -semiring such as separation axioms, compactness and connectedness etc.

Key Words: Γ -Semiring; Prime k -ideal (ideal); (strongly) irreducible ideal; Hull-Kernel topology; Structure space

Introduction and preliminaries

Algebraic structures play a prominent role in mathematics with wide ranging applications in many disciplines such as theoretical physics, computer sciences, control engineering, information sciences, coding theory etc.

The theory of semiring was first developed by H. S. Vandiver [33] and he has obtained important results of the objects. Semiring constitute a fairly natural generalization of rings, with board applications in the mathematical foundation of computer science. Also, semiring theory has many applications to other branches. For example, automata theory, optimization theory, algebra of formal process, combinatorial optimization, Baysian networks and belief propagation (cf. [12, 13, 14]).

It is well known that the concept of Γ -rings was first introduced and investigated by Nobusawa in 1964 [27], which is a generalization of the concept of rings. The class of Γ -rings contains not only all rings but also all Hestenes ternary rings. Later Barnes [2] weakened slightly the conditions in the definition of Γ -ring in the sense of Nobusawa. After these two papers were published, many mathematicians obtained interesting results on Γ -rings in the sense of Barnes and Nobusawa extending and generalizing many classical notions and results of the theory of rings. Γ -semirings were first studied by M. K. Rao [28] as a generalization of Γ -ring as well as of semiring. The concepts of Γ -semirings and its sub- Γ -semirings with a left(right) unity was studied by J. Luh [26] and M. K. Rao in [28]. The ideals, prime ideals, semiprime ideals, k -ideals and h -ideals of a Γ -semiring, regular Γ -semiring, respectively, were extensively studied by S. Kyuno [21, 22, 24] (cf. [23]) and M. K. Rao [28, 29].

In Γ -semirings, the properties of their ideals, prime ideals, semiprime ideals and their generalizations play an important role in their structure theory, however the properties of an ideal in semirings and Γ -semirings are somewhat different from the properties of the usual ring ideals. In order to amend these differs, the concepts of k -ideals and h -ideals in a semiring were introduced and considered by D. R. LaTorre [25] in 1965. For the properties of some h -ideals in Γ -semirings, the reader is referred to the recent papers of T. K. Dutta and S. K. Sardar, K. P. Shum in [7,8,9, 10, 31].

The notion of Γ -semiring not only generalizes the notions of semiring and Γ -ring but also the notion of ternary semiring. We point out here that this notion provides an algebraic background to the non-positive cones of the totally ordered rings. We recall here that the non-negative cones of the totally ordered rings form semirings but the non-positive cones do not form semirings because the induced multiplication is no longer closed. For further study of semirings, Γ -semirings and their generalization and examples, the reader is referred to [7, 8, 9, 10, 16, 28, 29, 31].

In this paper, we introduce some special classes of ideals in Γ -semirings called prime k -ideal, prime full k -ideal, prime ideals, maximal and strongly irreducible ideals. Considering and investigating properties of the collection \mathbf{A} , \mathbf{T} , \mathbf{M} , \mathbf{B} and \mathbf{S} of all proper prime k -ideals, proper prime full k -ideals, maximal ideals, prime ideals and strongly irreducible ideals, respectively, of a Γ -semiring R , we construct the respective topologies on them by means of closure operator defined in terms of intersection and inclusion relation among these ideals of Γ -semiring R . The respective obtained topological spaces are called the structure spaces of the Γ -semiring R . In fact we define this topology on \mathbf{A} and topology on \mathbf{T} will be the subspace topology from \mathbf{A} since \mathbf{T} is a subset of \mathbf{A} . This topological space has been studied in different algebraic structures [1, 3, 4, 5, 6, 11, 18, 19, 20, 32]. Recently, in [17] we have studied the topological structure on semihypergroups. We study several principal topological axioms and properties in those structure spaces of Γ -semiring such as separation axioms, compactness and connectedness etc.

Recall first the basic terms and definitions from the Γ -semiring theory.

Let R and Γ be two additive commutative semigroups. Then R is called a Γ -semiring if there exists a mapping $R \times \Gamma \times R \rightarrow R$ (the image to be denoted by $a\alpha b$, for $a, b \in R$ and $\alpha \in \Gamma$) satisfying the following conditions:

1. $a\alpha(b+c) = a\alpha b + a\alpha c$;
2. $(a+b)\alpha c = a\alpha c + b\alpha c$;
3. $a(\alpha+\beta)c = a\alpha c + a\beta c$;
4. $a\alpha(b\beta c) = (a\alpha b)\beta c$ for all $a, b, c \in R$ and $\alpha, \beta \in \Gamma$.

Obviously, every semiring R is a Γ -semiring with $\Gamma = R$ where $a\alpha b$ denotes the product of elements $a, \alpha, b \in R$, but not conversely.

If R contains an element 0 such that $0+x = x = x+0$ and $0\alpha x = x\alpha 0 = 0$ for all $x \in R$, for all $\alpha \in \Gamma$, then 0 is called the zero element (absorbing zero) or simply the zero of the Γ -semiring R . A non-empty subset T of R is said to be a *sub- Γ -semiring* of R if $(T, +)$ is a subsemigroup of $(R, +)$ and $a\alpha b \in T$; for all $a, b \in T$ and for all $\alpha \in \Gamma$. A non-empty subset I of a Γ -semiring R is called an *ideal* of R if $I+I \subseteq I, \Gamma R \subseteq I, R\Gamma I \subseteq I$, where for subsets U, V of R and Θ of Γ ,

$$U\Theta V = \left\{ \sum_{i=1}^n u_i \gamma_i v_i : u_i \in U, v_i \in V, \gamma_i \in \Theta \text{ and } n \text{ is a positive integer} \right\}.$$

An ideal I of a Γ -semiring R is called a *k-ideal* if for $x, y \in R; x+y \in I$ and $y \in I$ implies that $x \in I$. For a Γ -semiring R , let $E^+(R) = \{x \in R \mid x = x+x\}$. A k -ideal I of R is said to be full if $E^+(R) \subseteq I$. A proper ideal P of a Γ -semiring R is called a *prime ideal* of R if $a\Gamma b \subseteq P$ implies $a \in P$ or $b \in P$ for all a, b of R . An ideal I of a Γ -semiring R is called *proper* iff $I \subset R$ holds, where \subset denotes proper inclusion, and a proper ideal I is called *maximal* iff there is no ideal A of R satisfying $I \subset A \subset R$. An element e of a Γ -semiring R is called *identity element* of R if $e\alpha x = x = x\alpha e$, for all $x \in R; \alpha \in \Gamma$.

Throughout this paper, R will always denote a Γ -semiring with zero and unless otherwise stated a Γ -semiring means a Γ -semiring with zero.

On topological space of prime k -hyperideals of Γ -semiring

Let us denote with \mathbf{A} the collection of all prime k -ideals and \mathbf{T} the collection of all prime full k -ideals of a Γ -semiring R . For any subset A of \mathbf{A} , we define

$$\bar{A} = \{I \in \mathbf{A} : \bigcap_{I_i \in A} I_i \subseteq I\}.$$

It can be easily seen that $\bar{\emptyset} = \emptyset$.

Theorem 2.1 Let A, B be any two subsets of \mathbf{A} . Then

1. $A \subseteq \bar{A}$.
2. $\overline{\bar{A}} = \bar{A}$.
3. $A \subseteq B \Rightarrow \bar{A} \subseteq \bar{B}$.
4. $\overline{A \cup B} = \bar{A} \cup \bar{B}$.

Proof. (1). It is clear that $\bigcap_{I_i \in A} I_i \subseteq I_i$ for each i and hence $A \subseteq \bar{A}$.

(2). By (1), we have $\bar{A} \subseteq \overline{\bar{A}}$. Conversely, let $I_j \in \overline{\bar{A}}$. Then $\bigcap_{I_i \in \bar{A}} I_i \subseteq I_j$. Now $I_i \in \bar{A}$ implies that $\bigcap_{I_t \in A} I_t \subseteq I_i$ for all $i \in \bar{A}$. Thus

$$\bigcap_{I_t \in A} I_t \subseteq \bigcap_{I_i \in \bar{A}} I_i \subseteq I_j$$

So $I_j \in \bar{A}$ and hence $\overline{\bar{A}} \subseteq \bar{A}$. Consequently, $\overline{\bar{A}} = \bar{A}$.

(3). Let us suppose that $A \subseteq B$. Let $I_i \in \bar{A}$. Then $\bigcap_{I_j \in A} I_j \subseteq I_i$. Since $A \subseteq B$, it follows that

$$\bigcap_{I_j \in A} I_j \subseteq \bigcap_{I_j \in B} I_j \subseteq I_i.$$

This implies that $I_i \in \bar{B}$ and hence $\bar{A} \subseteq \bar{B}$.

(4). It is clear that $\overline{A \cup B} \subseteq \bar{A} \cup \bar{B}$.

Conversely, let $I_i \in \overline{A \cup B}$. Then $\bigcap_{I_j \in A \cup B} I_j \subseteq I_i$. It can be easily seen that

$$\bigcap_{I_j \in A \cup B} I_j = \left(\bigcap_{I_j \in A} I_j \right) \cap \left(\bigcap_{I_j \in B} I_j \right).$$

Since $\bigcap_{I_j \in A} I_j$ and $\bigcap_{I_j \in B} I_j$ are ideals of R , we have

$$\left(\bigcap_{I_j \in A} I_j \right) \Gamma \left(\bigcap_{I_j \in B} I_j \right) \subseteq \left(\bigcap_{I_j \in A} I_j \right) \cap \left(\bigcap_{I_j \in B} I_j \right) = \bigcap_{I_j \in A \cup B} I_j \subseteq I_i.$$

We have I_i is a prime ideal of R and hence either $\bigcap_{I_j \in A} I_j \subseteq I_i$ or $\bigcap_{I_j \in B} I_j \subseteq I_i$ i.e. either $I_i \in \bar{A}$ or $I_i \in \bar{B}$ i.e. $I_i \in \bar{A} \cup \bar{B}$. Consequently, $\overline{A \cup B} \subseteq \bar{A} \cup \bar{B}$ and hence $\overline{A \cup B} = \bar{A} \cup \bar{B}$.

Definition 2.2 The closure operator $A \rightarrow \bar{A}$ gives a topology $\tau_{\mathbf{A}}$ on \mathbf{A} . This topology $\tau_{\mathbf{A}}$ is called the hull-kernel topology and the topological space $(\mathbf{A}, \tau_{\mathbf{A}})$ is called the structure space of the Γ -semiring R .

Let I be a k -ideal of a Γ -semiring R . We define

$$\Delta(I) = \{I' \in \mathbf{A} : I \subseteq I'\} \text{ and } C\Delta(I) = \mathbf{A} \setminus \Delta(I) = \{I' \in \mathbf{A} : I \not\subseteq I'\}.$$

Proposition 2.3 Let R be a Γ -semiring and I a k -ideal of R . Then any closed set in \mathbf{A} is of the form $\Delta(I)$.

Proof. Let \bar{A} be any closed set in \mathbf{A} , where $A \subseteq \mathbf{A}$. Let $A = \{I_i : i \in \Lambda\}$ and $I = \bigcap_{I_i \in A} I_i$. Then I is a k -ideal of R . Let $I' \in \bar{A}$. Then $\bigcap_{I_i \in A} I_i \subseteq I'$. This implies that $I \subseteq I'$. Consequently, $I' \subseteq \Delta(I)$. So $\bar{A} \subseteq \Delta(I)$.

Conversely, let $I' \in \Delta(I)$. Then $I \subseteq I'$ i.e. $\bigcap_{I_i \in A} I_i \subseteq I'$. Consequently, $I' \in \bar{A}$ and hence $\Delta(I) \subseteq \bar{A}$. Thus $\bar{A} = \Delta(I)$.

Corollary 2.4 Let R be a Γ -semiring and I a k -ideal of R . Then any open set in \mathbf{A} is of the form $C\Delta(I)$.

Let R be a Γ -semiring and $a \in R$. We define

$$\Delta(a) = \{I \in \mathbf{A} : a \in I\} \text{ and } C\Delta(a) = \mathbf{A} \setminus \Delta(a) = \{I \in \mathbf{A} : a \notin I\}.$$

Proposition 2.5 Let R be a Γ -semiring and $a \in R$. Then $\{C\Delta(a) : a \in R\}$ forms an open base for the hull-kernel topology $\tau_{\mathbf{A}}$ on \mathbf{A} .

Proof. Let $U \in \tau_{\mathbf{A}}$. Then $U = C\Delta(I)$, where I is a k -ideal of R . Let $J \in U = C\Delta(I)$. Then $I \not\subseteq J$. This implies that there exists $a \in I$ such that $a \notin J$. Thus $J \in C\Delta(a)$. It remains to show that $C\Delta(a) \subset U$. Let $K \in C\Delta(a)$. Then $a \notin K$. This implies that $I \not\subseteq K$. Consequently, $K \in U$ and hence $C\Delta(a) \subset U$. So we find that $J \in C\Delta(a) \subset U$. Thus $\{C\Delta(a) : a \in R\}$ is an open base for the hull-kernel topology $\tau_{\mathbf{A}}$ on \mathbf{A} .

Theorem 2.6 Let R be a Γ -semiring. The structure space $(\mathbf{A}, \tau_{\mathbf{A}})$ is a T_0 -space.

Proof. Let I_1 and I_2 be two distinct elements of \mathbf{A} . Then there is an element a either in $I_1 \setminus I_2$ or in $I_2 \setminus I_1$. Let us suppose that $a \in I_1 \setminus I_2$. Then $C\Delta(a)$ is a neighbourhood of I_2 not containing I_1 . Hence $(\mathbf{A}, \tau_{\mathbf{A}})$ is a T_0 -space.

Theorem 2.7 Let R be a Γ -semiring. $(\mathbf{A}, \tau_{\mathbf{A}})$ is a T_1 -space if and only if no element of \mathbf{A} is contained in any other element of \mathbf{A} .

Proof. Let $(\mathbf{A}, \tau_{\mathbf{A}})$ be a T_1 -space. Let us suppose that I_1 and I_2 be any two distinct elements of \mathbf{A} . Then each of I_1 and I_2 has a neighbourhood not containing the other. Since I_1 and I_2 are arbitrary elements of \mathbf{A} , it follows that no element of \mathbf{A} is contained in any other element of \mathbf{A} .

Conversely, let us suppose that no element of \mathbf{A} is contained in any other element of \mathbf{A} . Let I_1 and I_2 be any two distinct elements of \mathbf{A} . Then by hypothesis, $I_1 \not\subseteq I_2$ and $I_2 \not\subseteq I_1$. This implies that there exist $a, b \in R$ such that $a \in I_1$, but $a \notin I_2$ and $b \in I_2$, but $b \notin I_1$. Consequently, we have $I_1 \in C\Delta(b)$, but $I_1 \notin C\Delta(a)$ and $I_2 \in C\Delta(a)$, but $I_2 \notin C\Delta(b)$ i.e. each of I_1 and I_2 has a neighbourhood not containing the other. Hence $(\mathbf{A}, \tau_{\mathbf{A}})$ is a T_1 -space.

Corollary 2.8 Let \mathbf{M} be the set of all proper maximal k -ideals of a Γ -semiring R with identity. Then $(\mathbf{M}, \tau_{\mathbf{M}})$ is a T_1 -space, where $\tau_{\mathbf{M}}$ is the induced topology on \mathbf{M} from $(\mathbf{A}, \tau_{\mathbf{A}})$.

Theorem 2.9 Let R be a Γ -semiring. $(\mathbf{A}, \tau_{\mathbf{A}})$ is a Hausdorff space if and only if for any two distinct pair of elements I, J of \mathbf{A} , there exist $a, b \in R$ such that $a \notin I, b \notin J$ and there does not exist any element K of \mathbf{A} such that $a \notin K$ and $b \notin K$.

Proof. Let $(\mathbf{A}, \tau_{\mathbf{A}})$ be a Hausdorff space. Then for any two distinct elements I, J of \mathbf{A} , there exist basic open sets $C\Delta(a)$ and $C\Delta(b)$ such that $I \in C\Delta(a)$, $J \in C\Delta(b)$ and $C\Delta(a) \cap C\Delta(b) = \emptyset$. Now $I \in C\Delta(a)$ and $J \in C\Delta(b)$ imply that $a \notin I$ and $b \notin J$. If possible, let $K \in \mathbf{A}$ such that $a \notin K$ and $b \notin K$. Then $K \in C\Delta(a)$ and hence $K \in C\Delta(a) \cap C\Delta(b)$. It is

impossible, since $C\Delta(a) \cap C\Delta(b) = \emptyset$. Thus there does not exist any element $K \in \mathbf{A}$ such that $a \notin K$ and $b \notin K$.

Conversely, let us suppose that the given condition holds and $I, J \in \mathbf{A}$ such that $I \neq J$. Let $a, b \in R$ be such that $a \notin I, b \notin J$ and there does not exist any $K \in \mathbf{A}$ such that $a \notin K$ and $b \notin K$. Then $I \in C\Delta(a), J \in C\Delta(b)$ and $C\Delta(a) \cap C\Delta(b) = \emptyset$. This implies that $(\mathbf{A}, \tau_{\mathbf{A}})$ is a Hausdorff space.

Corollary 2.10 *Let R be a Γ -semiring. If $(\mathbf{A}, \tau_{\mathbf{A}})$ is a Hausdorff space, then no proper prime k -ideal contains any other proper prime k -ideal. If $(\mathbf{A}, \tau_{\mathbf{A}})$ contains more than one element, then there exist $a, b \in R$ such that $\mathbf{A} = C\Delta(a) \cup C\Delta(b) \cup \Delta(I)$, where I is the k -ideal generated by a, b .*

Proof. Let us suppose that $(\mathbf{A}, \tau_{\mathbf{A}})$ is a Hausdorff space. Since every Hausdorff space is a T_1 -space, $(\mathbf{A}, \tau_{\mathbf{A}})$ is a T_1 -space. Hence by Theorem 2.7, it follows that no proper prime k -ideal contains any other proper prime k -ideal. Now let $J, K \in \mathbf{A}$ be such that $J \neq K$. Since $(\mathbf{A}, \tau_{\mathbf{A}})$ is a Hausdorff space, there exist basic open sets $C\Delta(a)$ and $C\Delta(b)$ such that $J \in C\Delta(a), K \in C\Delta(b)$ and $C\Delta(a) \cap C\Delta(b) = \emptyset$. Let I be the k -ideal generated by a, b . Then I is the smallest k -ideal containing a and b . Let $K \in \mathbf{A}$. Then either $a \in K, b \notin K$ or $a \notin K, b \in K$ or $a, b \in K$. The case $a \notin K, b \notin K$ is not possible, since $C\Delta(a) \cap C\Delta(b) = \emptyset$. Now in the first case, $K \in C\Delta(a)$ and hence $\mathbf{A} \subseteq C\Delta(a) \cup C\Delta(b) \cup \Delta(I)$. In the second case, $K \in C\Delta(b)$ and hence $\mathbf{A} \subseteq C\Delta(a) \cup C\Delta(b) \cup \Delta(I)$. In the third case, $K \in \Delta(I)$ and hence $\mathbf{A} \subseteq C\Delta(a) \cup C\Delta(b) \cup \Delta(I)$. So we find that $\mathbf{A} \subseteq C\Delta(a) \cup C\Delta(b) \cup \Delta(I)$. Again, clearly $C\Delta(a) \cup C\Delta(b) \cup \Delta(I) \subseteq \mathbf{A}$. Hence $\mathbf{A} = C\Delta(a) \cup C\Delta(b) \cup \Delta(I)$.

Theorem 2.11 *Let R be a Γ -semiring. $(\mathbf{A}, \tau_{\mathbf{A}})$ is a regular space if and only if for any $I \in \mathbf{A}$ and $a \notin I, a \in R$, there exists a k -ideal J of R and $b \in R$ such that $I \in C\Delta(b) \subseteq \Delta(J) \subseteq C\Delta(a)$.*

Proof. Let $(\mathbf{A}, \tau_{\mathbf{A}})$ be a regular space. Let $I \in \mathbf{A}$ and $a \notin I$. Then $I \in C\Delta(a)$ and $\mathbf{A} \setminus C\Delta(a)$ is a closed set not containing I . Since $(\mathbf{A}, \tau_{\mathbf{A}})$ is a regular space, there exist disjoint open sets U and V such that $I \in U$ and $\mathbf{A} \setminus C\Delta(a) \subseteq V$. This implies that $\mathbf{A} \setminus V \subseteq C\Delta(a)$. Since V is open, $\mathbf{A} \setminus V$ is closed and hence there exists a k -ideal J of R such that $\mathbf{A} \setminus V = \Delta(J)$, by Proposition 2.3. So we find that $\Delta(J) \subseteq C\Delta(a)$. Again, since $U \cap V = \emptyset$, we have $V \subseteq \mathbf{A} \setminus U$. Since U is open, $\mathbf{A} \setminus U$ is closed and hence there exists a k -ideal K of R such that $\mathbf{A} \setminus U = \Delta(K)$ i.e. $V \subseteq \Delta(K)$. Since $I \in U, I \notin \mathbf{A} \setminus U = \Delta(K)$. This implies that $K \not\subseteq I$. Thus there exists $b \in K (\subset R)$ such that $b \notin I$. So $I \in C\Delta(b)$. Now we show that $V \subseteq \Delta(b)$. Let $M \in V \subseteq \Delta(K)$. Then $K \subseteq M$. Since $b \in K$, it follows that $b \in M$ and hence $M \in \Delta(b)$. Consequently, $V \subseteq \Delta(b)$. This implies that $\mathbf{A} \setminus \Delta(b) \subseteq \mathbf{A} \setminus V = \Delta(J) \Rightarrow C\Delta(b) \subseteq \Delta(J)$. Thus we find that $I \in C\Delta(b) \subseteq \Delta(J) \subseteq C\Delta(a)$.

Conversely, let us suppose that the given condition holds. Let $I \in \mathbf{A}$ and $\Delta(K)$ be any closed set not containing I . Since $I \notin \Delta(K)$, we have $K \not\subseteq I$. This implies that there exists $a \in K$ such that $a \notin I$. Now by the given condition, there exists a k -ideal J of R and $b \in R$ such that $I \in C\Delta(b) \subseteq \Delta(J) \subseteq C\Delta(a)$. Since $a \in K, C\Delta(a) \cap \Delta(K) = \emptyset$. This implies that $\Delta(K) \subseteq \mathbf{A} \setminus C\Delta(a) \subseteq \mathbf{A} \setminus \Delta(J)$. Since $\Delta(J)$ is a closed set, $\mathbf{A} \setminus \Delta(J)$ is an open set containing the closed set $\Delta(K)$. Clearly, $C\Delta(b) \cap (\mathbf{A} \setminus \Delta(J)) = \emptyset$. So we find that $C\Delta(b)$ and $\mathbf{A} \setminus \Delta(J)$ are two disjoint open sets containing I and $\Delta(K)$ respectively. Consequently, $(\mathbf{A}, \tau_{\mathbf{A}})$ is a regular space.

Theorem 2.12 Let R be a Γ -semiring. (A, τ_A) is a compact space if and only if for any collection $\{a_\alpha\}_{\alpha \in \Lambda} \subset R$, there exists a finite subcollection $\{a_i : i = 1, 2, \dots, n\}$ in R such that for any $I \in A$, there exists a_i such that $a_i \notin I$.

Proof. Let (A, τ_A) be a compact space. Then the open cover $\{C\Delta(a_i) : a_i \in R\}$ of (A, τ_A) has a finite subcover $\{C\Delta(a_i) : i = 1, 2, \dots, n\}$. Let $I \in A$. Then $I \in C\Delta(a_i)$ for some $a_i \in R$. This implies that $a_i \notin I$. Hence $\{a_i : i = 1, 2, \dots, n\}$ is the required finite subcollection of elements of R such that for any $I \in A$, there exists a_i such that $a_i \notin I$.

Conversely, let us suppose that the given condition holds. Let $\{C\Delta(a_i) : a_i \in R\}$ be an open cover of A . Suppose to the contrary that no finite subcollection of $\{C\Delta(a_i) : a_i \in R\}$ covers A . This means that for any finite set $\{a_1, a_2, \dots, a_n\}$ of elements of R ,

$$\begin{aligned} C\Delta(a_1) \cup C\Delta(a_2) \cup \dots \cup C\Delta(a_n) &\neq A \Rightarrow \\ \Rightarrow \Delta(a_1) \cap \Delta(a_2) \cap \dots \cap \Delta(a_n) &\neq \emptyset \Rightarrow \\ \Rightarrow \text{there exists } I \in A \text{ such that } I &\in \Delta(a_1) \cap \Delta(a_2) \cap \dots \cap \Delta(a_n) \Rightarrow \\ \Rightarrow a_1, a_2, \dots, a_n &\in I, \text{ which contradicts our hypothesis.} \end{aligned}$$

So the open cover $\{C\Delta(a_i) : a_i \in R\}$ has a finite subcover and hence (A, τ_A) is compact.

Corollary 2.13 If the Γ -semiring R is finitely generated, then (A, τ_A) is a compact space.

Proof. Let $\{a_i : i = 1, 2, \dots, n\}$ be a finite set of generators of R . Then for any $I \in A$, there exists a_i such that $a_i \notin I$, since I is a proper prime k -ideal of R . Hence by Theorem 2.12, (A, τ_A) is a compact space.

Proposition 2.14 Let R be a Γ -semiring. (T, τ_T) is compact space if $E^+(R) \neq \{0\}$.

Proof. Let $\{\Delta(I_i) | i \in \Lambda\}$ be any collection of closed sets in T with finite intersection property. Let I be the proper prime k -ideal which is also full k -ideal generated by $E^+(R)$. Since any prime, full k -ideal J of R contains $E^+(R)$, then J contains I . Hence $I \in \bigcap_{i \in \Lambda} \Delta(I_i) \neq \emptyset$.

Consequently, (T, τ_T) is compact.

Definition 2.15 A Γ -semiring R is called a k -Noetherian Γ -semiring if it satisfies the ascending chain condition on k -ideals of R i.e. if $I_1 \subseteq I_2 \subseteq \dots \subseteq I_n \subseteq \dots$ is an ascending chain of k -ideals of R , then there exists a positive integer m such that $I_n = I_m$ for all $n \geq m$.

Theorem 2.16 If R is a k -Noetherian Γ -semiring, then (A, τ_A) is countably compact.

Proof. Let $\{\Delta(I_n)\}_{n=1}^\infty$ be a countable collection of closed sets in A with finite intersection property (FIP). Let us consider the following ascending chain of prime k -ideals of R : $\langle I_1 \rangle \subseteq \langle I_1 \cup I_2 \rangle \subseteq \langle I_1 \cup I_2 \cup I_3 \rangle \subseteq \dots$.

Since R is a k -Noetherian Γ -semiring, there exists a positive integer m such that $\langle I_1 \cup I_2 \cup \dots \cup I_m \rangle = \langle I_1 \cup I_2 \cup \dots \cup I_{m+1} \rangle = \dots$. Thus it follows that $\langle I_1 \cup I_2 \cup \dots \cup I_m \rangle \in \bigcap_{n=1}^\infty \Delta(I_n)$. Consequently, $\bigcap_{n=1}^\infty \Delta(I_n) \neq \emptyset$ and hence (A, τ_A) is countably compact.

Corollary 2.17 If R is a k -Noetherian Γ -semiring and (A, τ_A) is second countable, then (A, τ_A) is compact.

Proof. The proof follows by Theorem 2.16 and the fact that a second countable space is compact if it is countably compact.

Remark 2.18 Let $\{I_i\}$ be a collection of prime k -ideals of a Γ -semiring R . Then $\bigcap I_i$ is a k -ideal of R but it may not be a prime k -ideal of R , in general.

For this we have the following proposition:

Proposition 2.19 Let R be a Γ -semiring and $\{I_i\}$ be a collection of prime k -ideals of R such that $\{I_i\}$ forms a chain. Then $\bigcap I_i$ is a prime k -ideal of R .

Proof. It is clear that $\bigcap I_i$ is a k -ideal of R . Let $A\Gamma B \subseteq \bigcap I_i$ for any two k -ideals A, B of R . If possible, let $A, B \not\subseteq \bigcap I_i$. Then there exist i, j such that $A \not\subseteq I_i, B \not\subseteq I_j$. Since I_i is a chain, let $I_i \subseteq I_j$. This implies that $B \not\subseteq I_i$. Since $A\Gamma B \subseteq I_i$ and I_i is prime, we must have either $A \subseteq I_i$ or $B \subseteq I_i$. It is impossible. Therefore, either $A \subseteq \bigcap I_i$ or $B \subseteq \bigcap I_i$. Consequently, $\bigcap I_i$ is a prime k -ideal of R .

Theorem 2.20 Let R be a Γ -semiring. (A, τ_A) is disconnected if and if there exist a k -ideal I of R and a collection of points $\{a_i\}_{i \in \Lambda}$ of R not belonging to I such that if $I' \in A$ and $a_i \in I', \forall i \in \Lambda$, then $I \setminus I' \neq \emptyset$.

Proof. Let (A, τ_A) be not connected. Then there exists a non-trivial open and closed subset of A . Let I be the k -ideal of R for which $\Delta(I)$ is closed as well as open. Then $\Delta(I) = \bigcup_{i \in \Lambda} C\Delta(a_i)$ where $\{a_i\}_{i \in \Lambda}$ is a collection of points of R . Now since $C\Delta(a_i) \subseteq \Delta(I), \forall i \in \Lambda$ for any $I_i \in C\Delta(a_i)$ we have $I \subseteq I_i$, therefore $a_i \notin I$ as $a_i \notin I_i, \forall i \in \Lambda$. For any $I' \in A$ and $a_i \in I', \forall i \in \Lambda$ we have $I' \notin \Delta(I)$, consequently $I \setminus I' \neq \emptyset$.

Conversely, let us suppose the the given condition holds. Then $\Delta(I) = \bigcup_{i \in \Lambda} C\Delta(a_i)$ is an open and closed non-trivial subset of A and hence (A, τ_A) is disconnected.

Definition 2.21 Let R be a Γ -semiring. The structure space (A, τ_A) of R is called irreducible if for any decomposition $A = A_1 \cup A_2$, where A_1, A_2 are closed subsets of A , we have either $A = A_1$ or $A = A_2$.

Theorem 2.22 Let R be a Γ -semiring and A be a closed subset of A . Then A is irreducible if and only if $\bigcap_{I_i \in A} I_i$ is a prime k -ideal of R .

Proof. Let A be irreducible. Let P, Q be two k -ideals of R such that $P\Gamma Q \subseteq \bigcap_{I_i \in A} I_i$. Then $P\Gamma Q \subseteq I_i$ for all i . Since I_i is prime, we have $P \subseteq I_i$ or $Q \subseteq I_i$ which implies for $I_i \in A$, $I_i \in \{\overline{P}\}$ or $I_i \in \{\overline{Q}\}$. Hence $A = (A \cap \overline{P}) \cup (A \cap \overline{Q})$. Since A is irreducible and $(A \cap \overline{P}), (A \cap \overline{Q})$, are closed, it follows that $A = A \cap \overline{P}$ or $A = A \cap \overline{Q}$ and hence $A \subseteq \overline{P}$ or $A \subseteq \overline{Q}$. This implies that $P \subseteq \bigcap_{I_i \in A} I_i$ or $Q \subseteq \bigcap_{I_i \in A} I_i$. Consequently, $\bigcap_{I_i \in A} I_i$ is a prime k -ideal of R .

Conversely, let us suppose that $\bigcap_{I_i \in A} I_i$ is a prime k -ideal of R . Let $A = A_1 \cup A_2$, where A_1, A_2 are closed subsets of A . Then $\bigcap_{I_i \in A} I_i \subseteq \bigcap_{I_i \in A_1} I_i, \bigcap_{I_i \in A} I_i \subseteq \bigcap_{I_i \in A_2} I_i$. We have

$$\bigcap_{I_i \in A} I_i = \bigcap_{I_i \in A_1 \cup A_2} I_i = \left(\bigcap_{I_i \in A_1} I_i \right) \cap \left(\bigcap_{I_i \in A_2} I_i \right).$$

Also

$$\left(\bigcap_{I_i \in A_1} I_i \right) \Gamma \left(\bigcap_{I_i \in A_2} I_i \right) \subseteq \left(\bigcap_{I_i \in A_1} I_i \right), \left(\bigcap_{I_i \in A_1} I_i \right) \Gamma \left(\bigcap_{I_i \in A_2} I_i \right) \subseteq \left(\bigcap_{I_i \in A_2} I_i \right).$$

Thus we have

$$\left(\bigcap_{I_i \in A_1} I_i \right) \Gamma \left(\bigcap_{I_i \in A_2} I_i \right) \subseteq \left(\bigcap_{I_i \in A_1} I_i \right) \cap \left(\bigcap_{I_i \in A_2} I_i \right).$$

Since $\bigcap_{I_i \in A} I_i$ is prime, it follows that

$$\bigcap_{I_i \in A_1} I_i \subseteq \bigcap_{I_i \in A} I_i \text{ or } \bigcap_{I_i \in A_2} I_i \subseteq \bigcap_{I_i \in A} I_i.$$

So we find that

$$\bigcap_{I_i \in A} I_i = \bigcap_{I_i \in A_1} I_i \text{ or } \bigcap_{I_i \in A} I_i = \bigcap_{I_i \in A_2} I_i.$$

Let $I_\beta \in A$. We have

$$\bigcap_{I_i \in A_1} I_i \subseteq I_\beta \text{ or } \bigcap_{I_i \in A_2} I_i \subseteq I_\beta.$$

Since $A_1, A_2 \subseteq A$, so $I_i \subseteq I_\beta$ for all $I_i \in A_1$ or $I_i \subseteq I_\beta$ for all $I_i \in A_2$. Thus $I_\beta \in \overline{A_1} = A_1$ or $I_\beta \in \overline{A_2} = A_2$, since A_1 or A_2 are closed, i.e. $A = A_1$ or A_2 .

Theorem 2.23 Let R be a Γ -semiring. $(\mathbf{A}, \tau_{\mathbf{A}})$ is disconnected if and if there exist a k -ideal I of R and a collection of points $\{a_i\}_{i \in \Lambda}$ of R not belonging to I such that if $I' \in \mathbf{A}$ and $a_i \in I', \forall i \in \Lambda$, then $I \setminus I' \neq \emptyset$.

Proof. Let $(\mathbf{A}, \tau_{\mathbf{A}})$ be not connected. Then there exists a non-trivial open and closed subset of \mathbf{A} . Let I be the k -ideal of R for which $\Delta(I)$ is closed as well as open. Then $\Delta(I) = \bigcup_{i \in \Lambda} C\Delta(a_i)$ where $\{a_i\}_{i \in \Lambda}$ is a collection of points of R . Now since $C\Delta(a_i) \subseteq \Delta(I), \forall i \in \Lambda$ for any $I_i \in C\Delta(a_i)$ we have $I \subseteq I_i$, therefore $a_i \notin I$ as $a_i \notin I_i, \forall i \in \Lambda$. For any $I' \in \mathbf{A}$ and $a_i \in I', \forall i \in \Lambda$ we have $I' \not\subseteq \Delta(I)$, consequently $I \setminus I' \neq \emptyset$.

Conversely, let us suppose the the given condition holds. Then $\Delta(I) = \bigcup_{i \in \Lambda} C\Delta(a_i)$ is an open and closed non-trivial subset of \mathbf{A} and hence $(\mathbf{A}, \tau_{\mathbf{A}})$ is disconnected.

Proposition 2.24 Let R be a Γ -semiring. $(\mathbf{T}, \tau_{\mathbf{A}})$ is connected space if $E^+(R) \neq \{0\}$.

Proof. Let I be the proper prime k -ideal generated by $E^+(R)$. Since any full k -ideal of R contains $E^+(R)$, then I belongs to any closed set $\Delta(I')$ of \mathbf{A} . Consequently, any two closed sets of \mathbf{A} are not disjoint. Hence $(\mathbf{T}, \tau_{\mathbf{T}})$ is connected

On topological space of maximal ideals of Γ -semiring

In this section, the structure space of all maximal ideals of a Γ -semiring R with identity e is considered and studied.

An ideal is maximal if there is no ideal containing properly it. Let \mathbf{M} be the set of all maximal ideals in a Γ -semiring R . We shall define two topologies on \mathbf{M} . For every $x \in R$, we denote by Δ_x the set of all maximal ideals containing x , by Ω_x the set $\mathbf{M} - \Delta_x$, i.e. the set of all maximal ideals not containing x . Let I be an ideal of R , we denote by Δ_I the set of all maximal ideals containing I .

We choose the family $\{\Delta_x \mid x \in R\}$ as a subbase for open sets of \mathbf{M} . We shall refer to the resulting topology on \mathbf{M} as Δ -topology (in symbol, \mathbf{M}_Δ). Similarly, we shall take the family $\{\Omega_x \mid x \in R\}$ as a subbase for open sets of \mathbf{M} (in symbol, \mathbf{M}_Ω).

Let M_1, M_2 be two distinct elements of \mathbf{M}_Δ . Then we have $M_1 + M_2 = R$. Therefore there are a, b such that $e = a + b$ and $a \in M_1, b \in M_2$, so we have $\Delta_a \ni M_1, \Delta_b \ni M_2$ and $\Delta_a \cap \Delta_b = \emptyset$. Hence we have

Theorem 3.1 *The topological space \mathbf{M}_Δ is a T_2 -space.*

Let now M be an element of \mathbf{M}_Γ , and $M \neq M_1 \in \mathbf{M}_\Omega$, then there is an element a such that $a \in M_1$ and $a \notin M$. Therefore $M_1 \notin \Omega_a$ and $M_1 \notin \bigcap_{x \in M} \Omega_x$. This implies $M = \bigcap_{x \notin M} \Omega_x$. Hence we obtain the following

Theorem 3.2 *The topological space \mathbf{M}_Ω is a T_1 -space.*

Let I be an ideal of R and $\{a_\lambda\}$ a generator of I , then we have

$$\Delta_I = \bigcap_{\lambda} \Delta_{a_\lambda}.$$

Therefore, the closed sets for the topological space \mathbf{M}_Ω have the form $\Delta_{I_1} \cup \Delta_{I_2} \cup \dots \cup \Delta_{I_n}$, where I_i are ideals of R . Let $I = \bigcap_{i=1}^n I_i$, if $M \in \Delta_{I_i}$ for some i , then $M \supset I_i$ and $M \supset I$. This implies $\Delta_I \ni M$ and we have $\bigcup_{i=1}^n \Delta_{I_i} \subset \Delta_I$. Let us suppose that there is a maximal ideal M such that $M \in \Delta_I - \bigcup_{i=1}^n \Delta_{I_i}$, then $M \in \Delta_I$ and $M \notin \bigcup_{i=1}^n \Delta_{I_i}$. Hence $M \supset I$ and M does not contain every $I_i (i = 1, 2, \dots, n)$. Therefore, since M is a maximal ideal, there are elements $a_i \in I_i$ and $m_i \in M$ such that

$$e = a_i + m_i (i = 1, 2, \dots, n).$$

Thus, we have

$$e = a_1 + a_2 + \dots + a_n + m, m \in M$$

and $a_1 + a_2 + \dots + a_n \in I$. This implies $I + M = R$. Hence, by $I \subset M$, we have $M = R$, which is a contradiction. This shows the following relation:

$$\bigcup_{i=1}^n \Delta_{I_i} = \Delta_I$$

and we have the following:

Theorem 3.3 *The closed sets for \mathbf{M}_Ω are expressed by sets Δ_I , where I is an ideal of R .*

By Theorem 3.3, we prove the following

Theorem 3.4 *The space \mathbf{M}_Ω is a compact T_1 -space.*

Proof. Let $\{\Delta_{I_\lambda}\}$ be a family of closed sets in \mathbf{M}_Ω with the finite intersection property, where I_λ are ideals in R . Therefore, any finite family of I_λ does not contain the Γ -semiring R . Hence the ideal I generated by $\{I_\lambda\}$ does not contain the identity e of R . This shows that I is contained in a maximal ideal M . Hence $M \in \bigcap_{\lambda} \Delta_{I_\lambda}$. Therefore, since $\bigcap_{\lambda} \Delta_{I_\lambda}$ is non-empty, \mathbf{M}_Ω is a compact space.

On topological space of prime ideals of Γ -semiring

In this section, the structure space \mathbf{B} of all prime ideals of a Γ -semiring R with identity e is considered and the relation of \mathbf{B} and the structure space \mathbf{M} of all maximal ideals of R is investigated. Throughout the section, we shall treat a commutative Γ -semiring R with identity e . An ideal P of R is prime if and only if $a\Gamma b \subseteq P$ implies $a \in P$ or $b \in P$. Since R has an identity e , any maximal ideal is prime, therefore $\mathbf{B} \supseteq \mathbf{M}$.

To introduce a topology τ on \mathbf{B} , we shall take $\tau_x = \{P \mid x \notin P, P \in \mathbf{B}\}$ for every $x \in R$ as an open base of \mathbf{B} . We have the following

Theorem 4.1 Let \mathbf{U} be a subset of \mathbf{B} , then

$$\overline{\mathbf{U}} = \{P' \in \mathbf{B} \mid \bigcap_{P \in \mathbf{U}} P \subset P'\},$$

where $\overline{\mathbf{U}}$ is the closure of \mathbf{U} by the topology τ .

Proof. Let $P' \in \{P' \in \mathbf{B} \mid \bigcap_{P \in \mathbf{U}} P \subset P'\}$ and let τ_x be a neighbourhood of P' , then $x \notin P'$, and

we have $x \notin \bigcap_{P \in \mathbf{U}} P$. Therefore, there is a prime ideal $P \in \mathbf{U}$ such that P does not contain x and

$\tau_x \ni P$. This shows that $P \in \overline{\mathbf{U}}$. Thus we have proved that the $\overline{\mathbf{U}}$ contains $\{P' \in \mathbf{B} \mid \bigcap_{P \in \mathbf{U}} P \subset P'\}$.

If a prime ideal P' is not in $\{P' \in \mathbf{B} \mid \bigcap_{P \in \mathbf{U}} P \subset P'\}$, then $\bigcap_{P \in \mathbf{U}} P - P' \neq \emptyset$. Hence, for $x \in \bigcap_{P \in \mathbf{U}} P - P'$, we have $x \in P, P \in \mathbf{U}$ and $x \notin P'$. This shows $P \notin \tau_x, P \in \mathbf{U}$ and $P' \notin \tau_x$.

Therefore $\tau_x \cap \mathbf{U} = \emptyset$ and hence $P' \notin \overline{\mathbf{U}}$. The proof is complete.

A similar argument for \mathbf{M} relative to Ω -topology implies the following

Proposition 4.2 Let \mathbf{U} be a subset of \mathbf{M} , then

$$\overline{\mathbf{U}} = \{M' \in \mathbf{M} \mid \bigcap_{M \in \mathbf{U}} M \subset M'\},$$

where $\overline{\mathbf{U}}$ is the closure of \mathbf{U} by the topology Ω .

In a similar way to the proof of the Theorem 2.1, we can prove the following

Theorem 4.3 The closure operation $\mathbf{U} \rightarrow \overline{\mathbf{U}}$ of \mathbf{B} satisfies the following relations:

1. $\mathbf{U} \subseteq \overline{\mathbf{U}}$.
2. $\overline{\overline{\mathbf{U}}} = \overline{\mathbf{U}}$.
3. $\overline{\mathbf{U} \cup \mathbf{B}} = \overline{\mathbf{U}} \cup \overline{\mathbf{B}}$.

Proof. We shall prove only the last relation (3). By Theorem 4.1, $\mathbf{U} \subset \mathbf{B}$ implies $\overline{\mathbf{U}} \subset \overline{\mathbf{B}}$ and hence $\overline{\mathbf{U}} \cup \overline{\mathbf{B}} \subset \overline{\mathbf{U} \cup \mathbf{B}}$. Let $P \notin \overline{\mathbf{U}} \cup \overline{\mathbf{B}}$, then $P \notin \overline{\mathbf{U}}$ and $P \notin \overline{\mathbf{B}}$. Hence $P \not\supseteq \bigcap_{P' \in \mathbf{U}} P' = P_{\mathbf{U}}$ and

$P \not\supseteq \bigcap_{P' \in \mathbf{B}} P' = P_{\mathbf{B}}$. The sets $\mathbf{B}_{\mathbf{U}}$ and $\mathbf{B}_{\mathbf{B}}$ are ideals. If $P_{\mathbf{U}} \Gamma P_{\mathbf{B}} \subset P$, for any elements a, b such that

$a \in P_{\mathbf{U}} - P, b \in P_{\mathbf{B}} - P$, we have $a\Gamma b \subseteq P$ and since P is a prime ideal, $a \in P$ or $b \in P$, which is a contradiction. Therefore, $P \not\supseteq P_{\mathbf{U}} \Gamma P_{\mathbf{B}} \supseteq P_{\mathbf{U}} \cap P_{\mathbf{B}} = P_{\mathbf{U} \cup \mathbf{B}}$. Hence $P \notin \overline{\mathbf{U} \cup \mathbf{B}}$.

Theorem 4.4 The topological space \mathbf{B} is a T_0 -space.

Proof. It is sufficient to prove that $(\overline{P_1}) = (\overline{P_2})$ implies $P_1 = P_2$. By $P_2 \in (\overline{P_1})$, then $P_2 \supset P_1$. Similarly $P_1 \supset P_2$ and we have $P_1 = P_2$.

Theorem 4.5 *The topological space \mathbf{B} is a compact T_1 -space.*

Proof. Let \mathbf{U}_λ be a family of closed sets such that $\bigcap_\lambda \mathbf{U}_\lambda = \emptyset$, then we have $\sum P_{\mathbf{U}_\lambda} = R$, where $P_{\mathbf{U}_\lambda} = \bigcap_{P \in \mathbf{U}_\lambda} P$. Let us suppose that $\sum P_{\mathbf{U}_\lambda} \neq R$. Then there is a maximal ideal M containing $\sum P_{\mathbf{U}_\lambda}$. Therefore $\sum P_{\mathbf{U}_\lambda} \subset M$ for every λ . Hence $\mathbf{U}_\lambda \ni M$ for every λ , and we have $\bigcap_\lambda \mathbf{U}_\lambda \ni M$, which is a contradiction. By $\sum P_{\mathbf{U}_\lambda} = R$, we have there exist $\gamma_1, \gamma_2, \dots, \gamma_n \in \Gamma$, such that $e = a_1 \gamma_1 a_2 \gamma_2 \dots \gamma_n a_n$, $a_i \in P_{\mathbf{U}_{\lambda_i}}$ ($i = 1, 2, \dots, n$). Hence $\sum_{i=1}^n P_{\mathbf{U}_{\lambda_i}} = R$. If $\bigcap_{i=1}^n \mathbf{U}_{\lambda_i} \neq \emptyset$, then for a prime ideal P of $\bigcap_{i=1}^n \mathbf{U}_{\lambda_i}$, we have $P \supset P_{\mathbf{U}_{\lambda_i}}$ ($i = 1, 2, \dots, n$) and hence $P \supset \sum_{i=1}^n P_{\mathbf{U}_{\lambda_i}}$. Therefore we have $\bigcap_{i=1}^n \mathbf{U}_{\lambda_i} = \emptyset$.

By the \mathbf{B} -radical $r(\mathbf{B})$ of the Γ -semiring R , we mean the intersection of all prime ideals of R , that is, $\bigcap_{P \in \mathbf{B}} P$. By the \mathbf{M} -radical $r(\mathbf{M})$ of R , we mean the intersection of all maximal ideals of R , that is, $\bigcap_{M \in \mathbf{M}} M$.

From $\mathbf{M} \subseteq \mathbf{B}$, we have $r(\mathbf{B}) \subseteq r(\mathbf{M})$. In the following proposition we give a condition to be $r(\mathbf{B}) = r(\mathbf{M})$.

Theorem 4.6 *The subset \mathbf{M} of \mathbf{B} is dense in \mathbf{B} , if and only if, $r(\mathbf{B}) = r(\mathbf{M})$.*

Proof. Let $\overline{\mathbf{M}} = \mathbf{B}$ for the topology τ . Then we have

$$\{P \mid \bigcap_{M \in \mathbf{M}} M \subset P\} = \mathbf{B}.$$

Hence

$$r(\mathbf{M}) = \bigcap_{M \in \mathbf{M}} M \subseteq \bigcap_{P \in \mathbf{B}} P = r(\mathbf{B}).$$

Since $r(\mathbf{B}) \subseteq r(\mathbf{M})$, therefore we have $r(\mathbf{B}) = r(\mathbf{M})$.

Conversely, if $P \in \mathbf{B} - \overline{\mathbf{M}}$, then $P \in \mathbf{B}$ and $P \notin \overline{\mathbf{M}}$. Therefore, there is a neighbourhood τ_x of P such that $\tau_x \cap \mathbf{M} = \emptyset$. Hence $r(\mathbf{B}) = \bigcap_{P \in \mathbf{B}} P$ is a proper subset of $\bigcap_{M \in \mathbf{M}} M$. Therefore $r(\mathbf{B}) \neq r(\mathbf{M})$, which completes the proof.

Definition 4.7 *If $r(\mathbf{M})$ is the zero ideal (0), then A is said to be \mathbf{M} -semisimple.*

From the Theorem 4.6, we have the following

Theorem 4.8 *If R is \mathbf{M} -semisimple, \mathbf{M} is dense in \mathbf{B} .*

On topological space of strongly irreducible ideals of Γ -semiring

In this section, the structure space \mathbf{S} of all strongly irreducible ideals of a commutative Γ -semiring R with identity e is investigated.

An ideal I of a Γ -semiring R is called *irreducible*, if and only if $A \cap B = I$ for two ideals A, B implies $A = I$ or $B = I$. An ideal I of a Γ -semiring R is called *strongly irreducible*, if and only if $A \cap B \subset I$ for any two ideals A, B implies $A \subset I$ or $B \subset I$. From $A \Gamma B \subset A \cap B$ for any two ideals A, B , it follows that any prime ideals are strongly irreducible and any strongly irreducible ideals are irreducible.

Let \mathbf{S} be the set of all strongly irreducible ideals of R . From the above, it is clear that $\mathbf{M} \subset \mathbf{B} \subset \mathbf{S}$. To give a topology σ on \mathbf{S} , we shall take $\sigma_x = \{S \in \mathbf{S} \mid x \notin S\}$ for every $x \in R$ as an open base of \mathbf{S} .

Theorem 5.1 *Let \mathbf{U} be a subset of \mathbf{S} , then we have*

$$\overline{\mathbf{U}} = \{S' \in \mathbf{S} \mid \bigcap_{S \in \mathbf{U}} S \subset S'\}$$

where $\overline{\mathbf{U}}$ is the closure of \mathbf{U} by σ .

Proof. Let $\mathbf{F} = \{S' \in \mathbf{S} \mid \bigcap_{S \in \mathbf{U}} S \subset S'\}$ and let $S' \in \mathbf{F}$. Let σ_x be an open base of S' , then, by

the definition of the topology σ , $x \notin S'$. Hence, we have $x \notin \bigcap_{S \in \mathbf{U}} S$. It follows from this that there is

a strongly irreducible ideal S of \mathbf{U} such that x is not contained in S . Hence $\sigma_x \ni S$. Therefore $S' \in \overline{\mathbf{U}}$ and $\mathbf{F} \subset \overline{\mathbf{U}}$.

To prove that $\mathbf{F} \supset \overline{\mathbf{U}}$, take a strongly irreducible ideal S' such that $S' \notin \mathbf{F}$. Then $\bigcap_{S \in \mathbf{U}} S - S' \neq \emptyset$. For an element $x \in \bigcap_{S \in \mathbf{U}} S - S'$, we have $x \in S (S \in \mathbf{U})$ and $x \in S'$. Hence $S' \in \sigma_x$ and $S \notin \sigma_x$ for all S of \mathbf{U} . Therefore $\mathbf{U} \cap \sigma_x = \emptyset$ and then we have $S' \notin \overline{\mathbf{U}}$. Hence $\mathbf{F} \supset \overline{\mathbf{U}}$. The proof of the theorem is complete.

We shall prove that the topological space \mathbf{S} for the topology σ is a compact T_0 -space. To prove that \mathbf{S} is a T_0 -space, it is sufficient to verify the following conditions:

1. $\mathbf{U} \subseteq \overline{\mathbf{U}}$.
2. $\overline{\overline{\mathbf{U}}} = \overline{\mathbf{U}}$.
3. $\overline{\mathbf{U} \cup \mathbf{B}} = \overline{\mathbf{U}} \cup \overline{\mathbf{B}}$
4. $\overline{S_1} = \overline{S_2}$ implies $S_1 = S_2$.

The conditions (1) and (2) are clear, and $\mathbf{U} \cup \mathbf{F}$ implies $\overline{\mathbf{U}} \subset \overline{\mathbf{F}}$. From this relation, we have $\overline{\overline{\mathbf{U}} \cup \overline{\mathbf{F}}} \subset \overline{\mathbf{U} \cup \mathbf{F}}$. For some element of S of $\overline{\mathbf{U} \cup \mathbf{F}}$, suppose that $S \notin \overline{\mathbf{U}}$ and $S \notin \overline{\mathbf{F}}$. From Theorem 5.1, we have

$$S \not\supset \bigcap_{S' \in \mathbf{U}} S' = S_{\mathbf{U}} \text{ and } S \not\supset \bigcap_{S' \in \mathbf{F}} S' = S_{\mathbf{F}}.$$

$S_{\mathbf{U}}$ and $S_{\mathbf{F}}$ are ideals. If $S_{\mathbf{U}} \cap S_{\mathbf{F}} \subset S$, by the definition of S , $S_{\mathbf{U}} \subset S$ or $S_{\mathbf{F}} \subset S$. Hence $S \not\supset S_{\mathbf{U}} \cap S_{\mathbf{F}} = S_{\mathbf{U} \cup \mathbf{F}}$. This shows $S \notin \overline{\mathbf{U} \cup \mathbf{F}}$.

To prove that $\overline{S_1} = \overline{S_2}$ implies $S_1 = S_2$, we shall use the condition (1). Then $\overline{S_1} \ni S_2$ and by the definition of closure operation, we have $S_1 \subset S_2$. Similarly we have $S_1 \supset S_2$ and $S_1 = S_2$. Therefore we complete the proof that \mathbf{S} is a T_0 -space.

We shall prove that \mathbf{S} is a compact space. Let \mathbf{U}_t be a family of closed sets with empty intersection. Let $S_{\mathbf{U}_t} = \bigcap_{S \in \mathbf{U}_t} S$, suppose that $\sum_t S_{\mathbf{U}_t} \neq S$, then there is a maximal ideal M containing the ideal $\sum_t S_{\mathbf{U}_t}$. Therefore we have $S_{\mathbf{U}_t} \subset M$ for every t . By the definition of $S_{\mathbf{U}_t}$, $\mathbf{U}_t \ni M$ for every t . Hence $\bigcap_t \mathbf{U}_t \ni M$, which contradicts our hypothesis of \mathbf{U}_t . Therefore $\sum_t S_{\mathbf{U}_t} = R$. Hence we have there exist $\gamma_1, \gamma_2, \dots, \gamma_n \in \Gamma$, such that $e = a_1 \gamma_1 a_2 \gamma_2 \dots \gamma_n a_n (a_i \in S_{\mathbf{U}_{t_i}} (i = 1, 2, \dots, n))$. Hence

we have $R = \sum_{i=1}^n S_{U_{t_i}}$. If $\bigcap_{i=1}^n U_{t_i} \neq \emptyset$, for every strongly irreducible ideal S of $\bigcap_{i=1}^n U_{t_i}$, $S \supset S_{U_{t_i}}$ ($i=1,2,\dots,n$) and $S \supset \sum_{i=1}^n S_{U_{t_i}}$. If $\bigcap_{i=1}^n U_{t_i} = R$, we can prove easily that \mathbf{S} is a compact space. If $\bigcap_{i=1}^n U_{t_i}$ contains a proper strongly irreducible ideal S , we have $S \supset \sum_{i=1}^n S_{U_{t_i}}$, which is a contradiction to $R = \sum_{i=1}^n S_{U_{t_i}}$. Therefore $\bigcap_{i=1}^n U_{t_i} = \emptyset$. Hence \mathbf{S} is a compact space. Thus we have proved the following

Theorem 5.2 *The topological space (\mathbf{S}, σ) is compact T_0 -space.*

By the \mathbf{S} -radical $r(\mathbf{S})$ of a Γ -semiring, we mean the intersection of all strongly irreducible ideals of it, i.e., $\bigcap_{S \in \mathbf{S}} S$. From $\mathbf{M} \subset \mathbf{B} \subset \mathbf{S}$, we have $r(\mathbf{M}) \supset r(\mathbf{B}) \supset r(\mathbf{S})$.

Theorem 5.3 *The subset \mathbf{B} of \mathbf{S} is dense in \mathbf{S} , if and only if $r(\mathbf{B}) = r(\mathbf{S})$.*

Proof. Let $\overline{\mathbf{B}} = \mathbf{S}$ for the topology σ , then we have

$$\{S \mid \bigcap_{P \in \mathbf{B}} P \subset S\} = \mathbf{S}.$$

Hence, we have

$$r(\mathbf{B}) = \bigcap_{P \in \mathbf{B}} P \subset \bigcap_{S \in \mathbf{S}} S = r(\mathbf{S}).$$

On the other hand, $r(\mathbf{B}) \supset r(\mathbf{S})$. This shows $r(\mathbf{S}) = r(\mathbf{B})$.

Conversely, suppose that $\mathbf{S} - \overline{\mathbf{B}} \neq \emptyset$, then there is a strongly irreducible ideal S such that $S \not\subset \overline{\mathbf{B}}$ and $S \in \mathbf{S}$. Therefore there is a neighbourhood σ_x of S which does not meet \mathbf{B} . Hence $r(\mathbf{S}) = \bigcap_{S \in \mathbf{S}} S$ is a proper subset of $\bigcap_{P \in \mathbf{B}} P$, and we have $r(\mathbf{S}) \neq r(\mathbf{B})$.

Corollary 5.4 *The subset \mathbf{M} of \mathbf{S} is dense in \mathbf{S} , if and only if $r(\mathbf{M}) = r(\mathbf{S})$.*

Corollary 5.5 *Let R be a Γ -semiring with 0. If R is \mathbf{M} -semisimple, then \mathbf{M} and \mathbf{B} are dense in \mathbf{S} .*

References:

- [1] Adhikari, M. R. and M. K. Das, *Structure spaces of semirings*, Bull. Cal. Math. Soc. **86** (1994), 313-317.
- [2] Barnes, W. E., *On the Γ -rings of Nobusawa*, Pacific J. Math. **18** (1966), 411-422.
- [3] Booth, G. L. and N. J. Groendwald, *On uniformly strongly prime Γ -rings*, Bull. Austral. Math. Soc. Vol **37** (1988), 437-445.
- [4] Chattopadhyay, S. and S. Kar, *On structure space of Γ -semigroups*, Acta Univ. Palacki. Olomuc., Fac. rer. nat., Mathematica **47** (2008), 37-46.
- [5] Dutta, T. K. and S. Chattopadhyay, *On uniformly strongly prime Γ -semigroup*, Analele Stiintifice Ale Universitatii "Al. I. Cuza" Iasi Tomul LII, s.I, Math., 2006, f.2.
- [6] Dutta, T. K. and S. Chattopadhyay, *On Uniformly strongly prime Γ -semigroup (2)*, Int. J. Algebra, Number Theory and Applications **1(1)** (2009), 35-42.
- [7] Dutta, T. K. and S. Dhara, *On uniformly strongly prime Γ -semirings*, Southeast Asian Bull. Math. **30** (2006), no. 1, 39-48.
- [8] Dutta, T. K. and S. Dhara, *On uniformly strongly prime Γ -semirings II*, Discuss. Math. Gen. Algebra Appl. **26** (2006), no. 2, 219-231.
- [9] Dutta, T. K. and S. K. Sardar, *On prime ideals and prime radicals of a Γ -semirings*, An. Stiint. Univ. "Al. I. Cuza" Iasi Mat. (N.S.) **46(2)** (2001), 319-329.

- [10] Dutta, T. K. and S. K. Sardar, *Semiprime ideals and irreducible ideals of Γ -semirings*, Novi Sad J. Math. **30(1)** (2000), 97-108.
- [11] Gillman, L., *Rings with Hausdorff structure space*, Fund. Math. **45** (1957), 1-16.
- [12] Golan, J. S., *The Theory of Semirings with Applications in Mathematics and Theoretical Computer Science*. Pitman Monographs and Surveys in Pure and Appl. Maths., No. **54** (1992), Longman, New York.
- [13] Golan, J. S., *Semirings and their applications*, Kluwer Academic Publishers (1999).
- [14] Hebisch, U. and H. J. Weinert, *Semiring Theory, Algebraic Theory and Applications in Computer Science*. World Scientific, 1998.
- [15] Buys, A. and N. J. Groenewald, *On strongly prime gamma rings*, Math. Japon. **30**(1985), no. 6, 839-848.
- [16] Hedayati, H. and K. P. Shum, *An introduction to Γ -semirings*, Int. J. Algebra **5** (2011), no. 13-16, 709-726.
- [17] Hila, K. and S. Abdullah, *On topological structure on semihypergroups*, Proceedings ICTA2011 Islamabad, Pakistan, July 4–10, 2011. Cambridge Scientific Publishers (2012), Edit. A. V. Arhangel'skii, M. ud D. Khan, L. D.R. Kocinac, pp. 217-225, ISBN 978-1-908106-17-9.
- [18] Holme, A., *A general theory of structure spaces*, Fund. Math. **58** (1966) 335-347.
- [19] Iseki, K., *Notes on topological spaces, V. On structure spaces of semiring*, Proc. Japan Acad. Vol. **32** (1956), no. 1956, 426-429.
- [20] Kohls, C. W., *The space prime ideals of a ring*, Fund. Math. **45** (1957), 17-27.
- [21] Kyuno, S., *On prime gamma rings*, Pacific J. Math. **75** (1978), no. 1, 185-190.
- [22] Kyuno, S., *Prime ideals in gamma rings*, Pacific J. Math. **98** (1982), no. 2, 375-379.
- [23] Kyuno, S., *Gamma rings*, Hadronic Press Monographs in Mathematics. Hadronic Press, Inc., Palm Harbor, FL, 1991. viii+113 pp. ISBN: 0-911767-49-5
- [24] Kyuno, S., *A gamma ring with minimum conditions*, Tsukuba J. Math. **5(1)**(1981) 47-65.
- [25] LaTorre, D. R., *On h-ideals and k-ideals in hemirings*, Publ. Math. Debrecen **12** (1965), 219-226.
- [26] Luh, J., *The structure of primitive gamma rings*, Osaka J. Math. **7**(1970), 267-274.
- [27] Nobusawa, N., *On a generalization of the ring theory*, Osaka J. Math. **1** (1964), 81-89.
- [28] Rao, M. M. K., *Γ -semirings I*, Southeast Asian Bull. of Math. **19**(1995), 49-54.
- [29] Rao, M. M. K., *Γ -semirings II*, Southeast Asian Bull. Math. **21** (1997), no. 3, 281-287.
- [30] Ravisankar, T. S. and U. S. Shukla, *Structure of Γ -rings*, Pacific J. Math. **80(2)** (1979).
- [31] Sardar, S. K., B. C. Saha and K. P. Shum, *h-prime and h-semiprime ideals in semirings and Γ -semirings*, Int. J. Algebra, **4(5)** (2010), 209-220.
- [32] Sen, M. K. and S. Bandyopadhyay, *Structure space of a semi algebra over a hemiring*, Kyungpook Math. J. **33(1)**(1993), 25-36.
- [33] Vandiver, H. S., *Note on a simple type of algebra in which cancellation law of addition does not hold*, Bull. Amer. Math. Soc. **40** (1934), 914-920.

COMMON FIXED POINT THEOREMS FOR COMMUTING MAPPINGS ON A QUASIMETRIC SPACE

Luljeta Kikina, Assoc. Prof.

Department of Mathematics & Computer Science, Faculty of Natural Sciences,
University of Gjirokastra, Albania

Kristaq Gjino, PhD

Illir Vardhami, PhD

Department of Mathematics, Faculty of Natural Sciences, University of Tirana, Tirana, Albania

Abstract:

A general fixed point theorem for commuting mappings on quasimetric spaces is proved. Let (X, d) be a complete quasimetric space with a constant $\beta \geq 1$ and let f, g be commuting mappings from X into itself. If f is continuous and satisfies the following condition:

$$g(X) \subset f(X).$$

and further, there exist $\alpha \in (0, \frac{1}{\beta})$ such that for all $x, y \in X$,

$$d(gx, gy) \leq \alpha \max \{d(fx, fy), d(fx, gx), d(fx, gy), d(fy, gx), d(fy, gy)\}$$

then f and g have a unique common fixed point in X .

This result generalizes and extends the main theorem from [1] and [2]. Some new results concerning fixed points for commuting mappings on quasimetric spaces are obtained too which extend the results obtained in [3].

Key Words: Fixed point, quasimetric space, complete quasimetric space, common fixed point

Introduction and preliminaries

In [1], Jungck proved the following fixed point theorem:

Theorem 1.1 *Let (X, d) be a complete metric space. Let f and g be commuting continuous self-mappings on X such that*

$$g(X) \subset f(X).$$

Further, let there exist a constant $a \in (0, 1)$ such that for all $x, y \in X$:

$$d(gx, gy) \leq ad(fx, fy).$$

Then f and g have a unique common fixed point in X .

In [2], C'iri c' proved the following fixed point theorem:

Theorem 1.2(C'iri c') *Let (X, d) be a complete metric space. Let f be a self-mapping on X such that for some constant $a \in (0, 1)$ and for every $x, y \in X$*

$$d(fx, fy) \leq a \max \{d(x, y), d(x, fx), d(y, fy), d(x, fy), d(y, fx)\}.$$

Then f possesses a unique fixed point in X .

In [3] some results concerning fixed points on a metric space are obtained, which generalize and unify fixed point theorems in [1] and [2].

In this paper we obtain some new results concerning fixed points for mappings on quasimetric spaces. We generalize and extend the results of the Theorem 1.1 and 1.2 for commuting mappings on quasimetric spaces. We also extend the results established in [3] for commuting mappings on quasimetric spaces.

First, we give a standard definition and notation which will be used in the sequel.

Definition 1.3 [4] Let X be an arbitrary set and R^+ the set of nonnegative real numbers. A function $d : X \times X \rightarrow R^+$ is called a quasidistance on X if and only if there exists a constant $k \geq 1$, such that for all x, y and $z \in X$ the following conditions hold:

- (1) $d(x, y) \geq 0$ and $d(x, y) = 0 \Leftrightarrow x = y$;
- (2) $d(x, y) = d(y, x)$;
- (3) $d(x, y) \leq k[d(x, z) + d(z, y)]$.

Inequality (3) is often called *quasitriangular inequality* and k is often called the *quasitriangular constant* of d . Of course, d is called a metric when $k = 1$.

A pair (X, d) is called *quasimetric space* if X is a set and d is a quasidistance on X . It is clear that for $k = 1$ we obtain the metric space.

The following example illustrate the existence of the quasidistance.

Example 1.4 Let $X = R \times R$ and $x = (x_1, x_2) \in X, y = (y_1, y_2) \in X$. The function $d : X \times X \rightarrow R^+$ such that

$$d(x, y) = \begin{cases} k|x_1 - y_1| + |x_2 - y_2|, & \text{for } |x_1 - y_1| \leq |x_2 - y_2| \\ |x_1 - y_1| + k|x_2 - y_2|, & \text{for } |x_1 - y_1| > |x_2 - y_2| \end{cases}$$

is a quasidistance with $k \geq 1$.

Let us verify the satisfying of the three conditions of the definition 1.3.

1.

$$d(x, y) = 0 \Leftrightarrow \begin{cases} x_1 - y_1 = 0 \\ x_2 - y_2 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = y_1 \\ x_2 = y_2 \end{cases} \Leftrightarrow x = (x_1, x_2) = y = (y_1, y_2)$$

2. $d(x, y) = d(y, x), \forall x, y \in X$ since $|x_1 - y_1| = |y_1 - x_1|$ and $|x_2 - y_2| = |y_2 - x_2|$.

3.

$$\begin{aligned} d(x, y) &\leq k[|x_1 - y_1| + |x_2 - y_2|] = k[|x_1 - z_1 + z_1 - y_1| + |x_2 - z_2 + z_2 - y_2|] \leq \\ &\leq k[|x_1 - z_1| + |z_1 - y_1| + |x_2 - z_2| + |z_2 - y_2|] = \\ &= k[|x_1 - z_1| + |x_2 - z_2|] + k[|z_1 - y_1| + |z_2 - y_2|] \leq \\ &\leq k[d(x, z) + d(z, y)], \forall x, y, z \in X. \end{aligned}$$

Main results

Let (X, d) be a complete quasimetric space with a constant $\beta \geq 1$ and let f, g be commuting mappings from X into itself. Let f be continuous and satisfies the following condition:

$$g(X) \subset f(X) \quad (1)$$

and further, there exist $\alpha \in (0, \frac{1}{\beta})$ such that for all $x, y \in X$,

$$d(gx, gy) \leq \alpha \max \{d(fx, fy), d(fx, gx), d(fx, gy), d(fy, gx), d(fy, gy)\} \quad (2).$$

Let x_0 be an arbitrary point in X and define the sequence (x_n) in X as follows: By the condition $g(X) \subset f(X)$ it follows that there exists $x_1 \in X$ such that $gx_0 = fx_1$. In the same way we define successively x_2, \dots, x_n . Let $x_{n+1} \in X$ such that

$$gx_n = fx_{n+1} = y_n; n = 1, 2, \dots$$

We denote

$$G(y_k; n) = \{y_k, y_{k+1}, \dots, y_{k+n}\}.$$

We denote by $\delta(G(y_k; n))$ the diameter of the set $G(y_k; n)$.

We shall prove the following lemma which is necessary to prove our main theorem.

Lemma 2.1 *If $\delta(G(y_k; n)) > 0, k \geq 0, n \in N$, then*

$$(1) \delta(G(y_k; n)) = d(y_k, y_j) \text{ with } k < j \leq k + n.$$

$$(2) \delta(G(y_k; n)) \leq \alpha \delta(G(y_{k-1}; n+1)).$$

$$(3) \delta(G(y_0; n+k)) \leq \frac{\beta}{1-\beta\alpha} d(y_0, y_1).$$

Proof. For i and j such that $1 \leq i < j$, we have

$$\begin{aligned} d(y_i, y_j) &= d(gx_i, gx_j) \leq \\ &\leq \alpha \max\{d(fx_i, fx_j), d(fx_i, gx_j), d(x_j, gx_j), d(fx_i, gx_j), d(fx_j, gx_i)\} \\ &= \alpha \max\{d(y_{i-1}, y_{j-1}), d(y_{i-1}, y_i), d(y_{i-1}, y_j), d(y_{i-1}, y_j), d(y_{j-1}, y_i)\} \end{aligned}$$

Thus

$$d(y_i, y_j) \leq \alpha \delta(G(y_{i-1}, j-i+1)) \quad (3)$$

By using the notion of the superior of a finite number distances we have

$$\delta(G(y_k; n)) = d(y_i, y_j)$$

for some pair i, j such that $k \leq i < j \leq k + n$.

If $i > k$, by (3) we have

$$\delta(G(y_k, n)) \leq \alpha \delta(G(y_{i-1}, j-i+1))$$

with $i-1 \geq k$ and $j \leq k + n$. Then

$$\delta(G(y_k, n)) \leq \alpha \delta(G(y_k, n))$$

a contradiction, since $0 \leq \alpha < \frac{1}{\beta} \leq 1$.

Thus we have $i = k$ and this completes the proof of (a). Further, during the proof of lemma is proved that

$$\delta(G(y_k, n)) = d(y_k, y_j) \leq \alpha \delta(G(y_{k-1}, j-k+1)) \leq \alpha \delta(G(y_{k-1}, n+1)).$$

Thus, (b) holds too. Now, let we prove (c).

By (a) and using quasitriangular inequality on quasimetric spaces we have

$$\begin{aligned} \delta(G(y_l; m)) &= d(y_l, y_j) \leq \beta [d(y_l, y_{l+1}) + d(y_{l+1}, y_j)] = \\ &= \beta d(y_l, y_{l+1}) + \beta d(y_{l+1}, y_j) \leq \beta d(y_l, y_{l+1}) + \beta \delta(G(y_{l+1}, m-1)) \end{aligned}$$

with $l < j \leq l + m$.

Using (b) we have

$$\delta(G(y_l; m)) \leq \beta d(y_l, y_{l+1}) + \beta \alpha \delta(G(y_l; m))$$

which implies that

$$\delta(G(y_l; m)) \leq \frac{\beta}{1-\beta\alpha} d(y_l, y_{l+1}) \quad (4)$$

since $1-\beta\alpha > 0$.

Using again (b) we have

$$\delta(G(y_k, n)) \leq \alpha \delta(G(y_{k-1}, n+1)) \leq \alpha^2 \delta(G(y_{k-2}, n+2)) \leq \dots \leq \alpha^k \delta(G(y_0; n+k))$$

For $l = 0$ and $m = n+k$, by (4) we have

$$\delta(G(y_0; n+k)) \leq \frac{\beta}{1-\beta\alpha} d(y_0, y_1).$$

Finally,

$$\delta(G(y_k; n)) \leq \frac{\alpha^k \beta}{1 - \beta\alpha} d(y_0, y_1).$$

This completes the proof of Lemma.

Now we give and prove our theorem as follows

Theorem 2.2 Let (X, d) be a complete quasimetric space with a constant $\beta \geq 1$, quasidistance d continuous and let f, g be commuting mappings from X into itself. If f is continuous and satisfies the following condition:

$$g(X) \subset f(X).$$

and further, there exist $\alpha \in (0, \frac{1}{\beta})$ such that for all $x, y \in X$, $d(gx, gy) \leq \alpha \max\{d(fx, fy), d(fx, gx), d(fx, gy), d(fy, gx), d(fy, gy)\}$. Then f and g have a unique common fixed point in X .

Proof. It is enough to find a point y such that

$$fy = gy.$$

Applying (2) for $x = gy$ we have

$$d(ggy, gy) \leq \alpha \max\{d(fgy, fy), d(fgy, ggy), d(fy, gy), d(fgy, gy), d(fy, ggy)\} = \alpha d(ggy, gy).$$

By the inequality $d(ggy, gy) \leq \alpha d(ggy, gy)$, since $0 < \alpha < \frac{1}{\beta} \leq 1$ it follows

$$d(ggy, gy) = 0.$$

This shows that $ggy = gy$. Thus, gy is a fixed point of g . On the other hand, since f and g commute we have

$$fgy = gfy = ggy = gy$$

which implies that gy is also a fixed point of f .

Case 1. Let $\delta(G(y_k; n)) = 0$ for some n and k . Then we have $y_k = y_{k+1}$ or

$$fx_{k+1} = gx_{k+1}.$$

It is clear that y_{k+1} is a common fixed point of f and g .

Case 2. Let $\delta(G(y_k; n)) > 0$.

For $n < m$ we consider the distance

$$d(y_m, y_n) \leq \delta(G(y_n, m-n)) \leq \alpha^n \frac{\beta}{1 - \beta\alpha} d(y_0, y_1)$$

and $\lim_{n \rightarrow \infty} d(y_m, y_n) = 0$. Thus, (y_n) is a Cauchy sequence with a limit y in X , since (X, d) is a complete quasimetric space.

Since f is continuous it follows that

$$\lim_{n \rightarrow \infty} fy_n = fy$$

and using the equality $gy_n = fy_{n+1}$ we have

$$\lim_{n \rightarrow \infty} gy_n = \lim_{n \rightarrow \infty} gy_n = \lim_{n \rightarrow \infty} fy_{n+1} = fy.$$

On the other hand we have

$$\begin{aligned} d(fy_{n+1}, gy) &= d(gy_n, gy) \leq \\ &\leq \alpha \max\{d(fy_n, fy), d(fy_n, gy_n), d(fy, gy), d(fy_n, gy), d(fy_n, gy_n)\}. \end{aligned}$$

Letting n tend to infinity, we get

$$g(fy, gy) \leq \alpha d(fy, gy)$$

which holds only for $fy = gy$, since $\alpha < 1$. By (2) it follows that y is a unique common fixed point

of f and g .

Corollary 2.3 *The Theorem 1.2 (C' iri c') is a corollary of the Theorem 2.2. If $f = I_x$, where I_x is the identity mapping in X , then the condition (2) just as in Lemma 2.1 can be written in the following form:*

$$d(gx, gy) \leq \alpha \max \{d(x, y), d(x, gx), d(y, gy), d(x, gy), d(y, gx)\}$$

If we replace g by f we obtain the condition of the Theorem 1.2(C' iri c')

Corollary 2.4 *The Theorem 1.1 is a corollary of the Theorem 2.2.*

It is obvious that whenever the condition $d(gx, gy) \leq \alpha d(fx, fy)$ is satisfied, the condition (2) is satisfied and so the corollary follows.

Corollary 2.5 *The Theorem 2.2 is a generalization of the Banach Theorem.*

If we take $k = 1$ and $f = I_x$ then by Corollary 2.4 the corollary follows.

Corollary 2.6 *The Theorems 1.1 and 1.2 can be proved on the quasimetric spaces with the constant β , where $0 < \alpha < \frac{1}{\beta}$.*

In the following theorem, as the domain of g is considered $f(X)$. We also replace the continuously condition of f by the continuously of $f^2 = ff$ as an weakly condition.

Theorem 2.7 *Let (X, d) be a complete quasimetric space with the constant $\beta \geq 1$, quasidistance d continuous, let f be a mapping from X into itself such that f^2 is continuous. Let*

$$g : f(X) \rightarrow X$$

be a mapping such that

$$gf(X) \subset f^2(X) \quad (5)$$

and for all x from the domain of fg and gf , $fgx = gfx$. Further, there exist $\alpha \in (0, \frac{1}{\beta})$ such that for all $x, y \in f(X)$ the condition (2) is satisfied.

Then, f and g have a unique common fixed point in X .

Proof. Let x_0 be an arbitrary point in $f(X)$. Let x_1 be a point in $f(X)$ such that

$$gx_0 = fx_1 = y_0.$$

This follows by (5). In the same way we define succesively x_2, \dots, x_n . Let $x_{n+1} \in f(X)$ such that

$$gx_n = fx_{n+1} = y_n; n = 1, 2, \dots$$

We denote:

$$fy_n = fgx_n = ffx_{n+1} = gy_{n-1} = z_n.$$

For $k > 0$ and $n \in N$, one can prove, just as in Lemma 2.1, that

$$\delta(G(z_n; n)) \leq \frac{\alpha^k \beta}{1 - \alpha\beta} d(z_0, z_1).$$

Therefore, (z_n) is a Cauchy sequence with a limit z in X .

Since f^2 is continuous it follows that

$$\lim_{n \rightarrow \infty} f^2 z_n = f^2 z$$

Further, we have

$$gfz_n = gff^2 x_{n+1} = f^2 fgx_{n+1} = f^2 z_{n+1}$$

which implies that

$$\lim_{n \rightarrow \infty} gfz_n = f^2 z.$$

Now by the condition (2) of the Theorem 2.2 we have

$$\begin{aligned} d(f^2 z_{n+1}, gfz) &= d(gfz_n, gfz) \leq \\ &\leq \alpha \max\{d(f^2 z_n, f^2 z), d(f^2 z_n, gfz_n), d(f^2 z, gfz), d(f^2 z_n, gfz), d(f^2 z, gfz_n)\}. \end{aligned}$$

Letting n tend to infinity, we get

$$d(f^2 z, gfz) \leq \alpha d(f^2 z, gfz)$$

which holds only in the case when $f^2 z = gfz$.

Finally, by

$$\begin{aligned} d(ggfz, gfz) &\leq \alpha \max\{d(fgfz, f^2 z), d(fgfz, ggfz), d(f^2 z, gfz), d(fgfz, gfz), d(f^2 z, ggfz)\} \\ &= \alpha d(gfz, gfz) \end{aligned}$$

it follows that $ggfz = gfz$. Therefore, gfz is a fixed point of g .

On the other hand, since we have

$$fgfz = f^2 gz = fgfz = ggfz = gfz$$

then gfz is also a fixed point of f . Therefore f and g have a common fixed point. By the condition (2) it follows that it is the unique common fixed point of f and g .

Theorem 2.8 Let (X, d) be a complete quasimetric space with the constant $\beta \geq 1$, quasidistance d continuous. If f, g are commuting mappings from X into itself such that f^2 is continuous satisfying the following condition

$$g(X) \subset f(X)$$

and further, there exist $\alpha \in (0, \frac{1}{\beta})$ such that for all $x, y \in X$

$$d(gx, gy) \leq \alpha \max\{d(fx, fy), d(fx, gx), d(fx, gy), d(fy, gx), d(fy, gy)\},$$

then f and g have a unique common fixed point in X .

Proof. By the condition $g(X) \subset f(X)$ follows

$$gf(X) = fg(X) \subset f^2(X)$$

which is the condition (5) of the Theorem 2.7. Since the condition (2) is satisfied, then by Theorem 2.2 and Theorem 2.7 it follows that f and g have a unique common fixed point.

References:

- [1] G. Jungck, *Commuting mappings and fixed points*, Amer. Math. Monthly, **83**(1976), 261-263.
- [2] Lj. B. Ćirić, *A generalization of Banach's contraction principle*, Proc. Amer. Math. Soc, **45** (1974), 267-273.
- [3] K. M. Das and K. Viswanatha naik, *Common fixed point theorems for commuting maps on a metric space*, Proc. Amer. Math. Soc., **77**(3), (1979), 369-373.
- [4] R. Toledano, *A note on the Lebesgue differentiation theorem in spaces of homogeneous type*, Real Analysis Exchange, **29**(1), (2003/2004), 335-340.
- [5] Shunrong Zhu, *Common fixed point theorems for commuting maps on a complete metric space*, J. Southeast Univ. (English Ed.) **17**(1), (2001), 82-84.
- [6] Shunrong Zhu, *Fixed points on complete metric spaces*, J. Southeast Univ. (English Ed.) **20**(2), (2004), 256-260
- [7] B. E. Rhoades, *A comparison of various definitions of contractive mappings*, Trans. Amer. Soc. **226**(1977), 257-290.
- [8] S. Sessa, *On a weak commutativity condition of mappings in fixed points considerations*, Publ. Inst. Math. **32**(46) (1982), 149-153.
- [9] M. I. Vulpe, D. Ostraih, F. Hoiman, *The topological structure of a quasimetric space. (Russian) Investigations in functional analysis and differential equations*, pp. 14-19, 137, "Shtiintsa", Kishinev, 1981.

WEAK SEPARATION AXIOMS VIA D_ω , $D_{\alpha-\omega}$, $D_{pre-\omega}$, $D_{b-\omega}$, AND $D_{\beta-\omega}$ -SETS

Mustafa Hasan Hadi

University of Babylon, Colledge Of Education For Applied Sciences, Mathematics Department, Iraq

Abstract:

In this paper we define new types of sets we call them D_ω , $D_{\alpha-\omega}$, $D_{pre-\omega}$, $D_{pre-\omega}$, $D_{b-\omega}$, and $D_{\beta-\omega}$ -sets and use them to define some associative separation axioms. Some theorems about the relation between them and the weak separation axioms introduced by M. H. Hadi in [1] are proved, with some other simple theorems. Throughout this paper, (X, T) stands for topological space. Let (X, T) be a topological space and A a subset of X . A point x in X is called **condensation point** of A if for each U in T with x in U , the set $U \cap A$ is uncountable [10]. In 1982 the ω -closed set was first introduced by H. Z. Hdeib in [10], and he defined it as: A is **ω -closed** if it contains all its condensation points and the **ω -open** set is the complement of the ω -closed set. Equivalently. A sub set W of a space (X, T) , is ω -open if and only if for each $x \in W$, there exists $U \in T$ such that $x \in U$ and $U \setminus W$ is countable. The collection of all ω -open sets of (X, T) denoted T_ω form topology on X and it is finer than T . Several characterizations of ω -closed sets were provided in [3, 4, 5, 6].

Key Words: Axioms, weak separation

In [7,8,9] some authors introduced α -open, pre -open, b -open, and β -open sets. On the other hand in [2] T. Noiri, A. Al-Omari, M. S. M. Noorani introduced the notions $\alpha - \omega$ -open, $pre - \omega$ -open, $\beta - \omega$ -open, and $b - \omega$ -open sets in topological spaces. They defined them as follows: A subset A of a space X is called: **$\alpha - \omega$ -open** [2] if $A \subseteq int_\omega (cl(int_\omega(A)))$ and the complement of the $\alpha - \omega$ -open set is called **$\alpha - \omega$ -closed** set, **$pre - \omega$ -open** [2] if $A \subseteq int_\omega (cl(A))$ and the complement of the $pre - \omega$ -open set is called **$pre - \omega$ -closed** set, **$b - \omega$ -open** [2] if $A \subseteq int_\omega (cl(A)) \cup cl(int_\omega(A))$ and the complement of the $b - \omega$ -open set is called **$b - \omega$ -closed** set, **$\beta - \omega$ -open** [2] if $A \subseteq cl (int_\omega (cl(A)))$ and the complement of the $\beta - \omega$ -open set is called **$\beta - \omega$ -closed** set.

Now let recall some condition introduced by M. H. Hadi in [1]: Let (X, T) be topological space. It said to be satisfy: The **ω -condition** if every ω -open set is $\omega - t$ -set. **2.** The **$\omega - B_\alpha$ -condition** if every $\alpha - \omega$ -open set is $\omega - B_\alpha$ -set. The **$\omega - B$ -condition** if every $pre - \omega$ -open is $\omega - B$ -set.

In Our paper we **firstly** introduce our dominion and some results related to it.

Definition 1. A subset A of a topological space (X, T) is called **D -set** [11] (resp. **D_ω -set**, **$D_{\alpha-\omega}$ -set**, **$D_{pre-\omega}$ -set**, **$D_{b-\omega}$ -set**, **$D_{\beta-\omega}$ -set**). If there are two open (resp. ω -open, $\alpha - \omega$ -open, $pre - \omega$ -open, $\beta - \omega$ -open, and $b - \omega$ -open) sets U and V with $U \neq X$ and $A = U \setminus V$.

Proposition 2. In any topological space satisfies ω -condition. Any D_ω -set is D -set.

Proposition 3. In any topological space satisfies $\omega - B_\alpha$ -condition. Any $D_{\alpha-\omega}$ -set is D -set.

Proposition 4. In any topological space satisfies $\omega - B$ -condition. Any $D_{pre-\omega}$ -set is D -set.

Proposition 5. In any topological space. Any $D_{b-\omega}$ -set with empty ω -interior is $D_{pre-\omega}$ -set.

Secondly now utilizing the weak D_ω sets we can define our separation axioms and a rather simple theorem related to it as follows:

Definition 6. Let X be a topological space. If $x \neq y \in X$, either there exists a set U , such that $x \in U, y \notin U$, or there exists a set U such that $x \notin U, y \in U$. Then X called

1. $\omega - D_0$ space, whenever U is D_ω -set in X .
2. $\alpha - \omega - D_0$ space, whenever U is $D_{\alpha-\omega}$ -set in X .
3. $pre-\omega - D_0$ space, whenever U is $D_{pre-\omega}$ -set in X .
4. $b - \omega - D_0$ space, whenever U is $D_{b-\omega}$ -set in X .
5. $\beta - \omega - D_0$ space, whenever U is $D_{\beta-\omega}$ -set in X .

Definition 7. We can define the spaces $\omega - D_i, \alpha - \omega - D_i, pre - \omega - D_i, b - \omega - D_i, \beta - \omega - D_i$, for $i = 0,1,2$. And $\omega^* - D_i, \alpha - \omega^* - D_i, \alpha - \omega^{**} - D_i, pre - \omega^* - D_i, \alpha - pre - \omega - D_i, pre - \omega^{**} - D_i, b - \omega^* - D_i, pre - b - \omega - D_i, \alpha - b - \omega - D_i, b - \omega^{**} - D_i, \beta - \omega^* - D_i, \alpha - \beta - \omega - D_i, pre - \beta - \omega - D_i, \beta - \omega^{**} - D_i$, and $b - \beta - \omega - D_i$, for $i = 1,2$, by replacing the sets: open, ω -open, $\alpha - \omega$ -open, $pre - \omega$ -open, $b - \omega$ -open, $\beta - \omega$ -open, by the D -set, D_ω -set, $D_{\alpha-\omega}$ -set, $D_{pre-\omega}$ -set, $D_{b-\omega}$ -set, and $D_{\beta-\omega}$ -set.

Theorem 8. Let (X, T) be a topological space. Then X is $\omega - D_1$, (resp. $\alpha - \omega - D_1, \omega^* - D_1, \alpha - \omega^* - D_1, \alpha - \omega^{**} - D_1, pre - \omega - D_1, pre - \omega^* - D_1, \alpha - pre - \omega - D_1, b - \omega - D_1, pre - \omega^{**} - D_1, b - \omega - D_1, b - \omega^* - D_1, pre - b - \omega - D_1, \alpha - b - \omega - D_1, pre - b - \omega - D_1, b - \omega^{**} - D_1, \beta - \omega - D_1, \beta - \omega^* - D_1, \alpha - \beta - \omega - D_1, pre - \beta - \omega - D_1, \beta - \omega^{**} - D_1, b - \beta - \omega - D_1$) if and only if it is $\omega - D_2$, (resp. $\alpha - \omega - D_2, \omega^* - D_2, \alpha - \omega^* - D_2, \alpha - \omega^{**} - D_2, pre - \omega - D_2, pre - \omega^* - D_2, \alpha - pre - \omega - D_2, b - \omega - D_2, pre - \omega^{**} - D_2, b - \omega - D_2, b - \omega^* - D_2, pre - b - \omega - D_2, \alpha - b - \omega - D_2, pre - b - \omega - D_2, b - \omega^{**} - D_2, \beta - \omega - D_2, \beta - \omega^* - D_2, \alpha - \beta - \omega - D_2, pre - \beta - \omega - D_2, \beta - \omega^{**} - D_2, b - \beta - \omega - D_2$).

Thirdly we introduce the so called ω -net point and a rather theorems related to it.

Definition 9. A point $x \in X$ which has only X as ω -neighbourhood is called an ω -net point.

Proposition 10. Let (X, T) be a topological space If X is $\omega - D_1$ space, then it has no ω -net point.

Theorem 11. If $f: (X, \tau) \rightarrow (Y, \sigma)$ is ω -continuous (resp. $\alpha - \omega$ -continuous, $pre - \omega$ -continuous, $\beta - \omega$ -continuous, $b - \omega$ -continuous) onto function and A is D_ω -set (resp. $D_{\alpha-\omega}$ -set, $D_{pre-\omega}$ -set, $D_{b-\omega}$ -set, $D_{\beta-\omega}$ -set) in Y , then $f^{-1}(A)$ is also D_ω -set (resp. $D_{\alpha-\omega}$ -set, $D_{pre-\omega}$ -set, $D_{b-\omega}$ -set, $D_{\beta-\omega}$ -set) in X .

Theorem 12. For any two topological spaces (X, τ) and (Y, σ) .

1. If (Y, σ) be an $\omega^* - D_1$ and $f: (X, \tau) \rightarrow (Y, \sigma)$ is an ω -continuous bijection, then (X, τ) is $\omega^* - D_1$.
2. If (Y, σ) be an $\alpha - \omega^{**} - D_1$ and $f: (X, \tau) \rightarrow (Y, \sigma)$ is an $\alpha - \omega$ -continuous bijection, then (X, τ) is, $\alpha - \omega^{**} - D_1$.
3. If (Y, σ) be a $pre - \omega^{**} - D_1$ and $f: (X, \tau) \rightarrow (Y, \sigma)$ is a $pre - \omega$ -continuous bijection, then (X, τ) is $pre - \omega^{**} - D_1$.
4. If (Y, σ) be a $b - \omega^{**} - D_1$ and $f: (X, \tau) \rightarrow (Y, \sigma)$ is a $b - \omega$ -continuous bijection, then (X, τ) is $b - \omega^{**} - D_1$.
5. If (Y, σ) be a $\beta - \omega^{**} - D_1$ and $f: (X, \tau) \rightarrow (Y, \sigma)$ is a $\beta - \omega$ -continuous bijection, then (X, τ) is $\beta - \omega^{**} - D_1$.

Theorem 13. A topological space (X, T) is $\omega^* - D_1$ (resp. $\alpha - \omega^{**} - D_1, pre - \omega^{**} - D_1, b - \omega^{**} - D_1, \beta - \omega^{**} - D_1$) if and only if for each pair of distinct points $x, y \in X$, there exists an ω -continuous (resp. $\alpha - \omega$ -continuous, $pre - \omega$ -continuous, $b - \omega$ -continuous, $\beta - \omega$ -continuous) onto function $f: (X, \tau) \rightarrow (Y, \sigma)$ such that $f(x)$ and $f(y)$ are distinct, where (Y, σ) is $\omega^* - D_1$ (resp. $\alpha - \omega^{**} - D_1, pre - \omega^{**} - D_1, b - \omega^{**} - D_1, \beta - \omega^{**} - D_1$) space.

References:

- [1]. M. H. Hadi, " Weak forms of ω -open sets and decomposition of separation axioms" , M. Sc. Thesis, Babylon University (2011).
- [2]. T. Noiri, A. Al-Omari and M. S. M. Noorani, " Weak forms of ω -open sets and decomposition of continuity", E.J.P.A.M. 2(1): 73-84 (2009).

SYMBOLIC DATA ANALYSIS FOR THE ASSESSMENT OF USER SATISFACTION: AN APPLICATION TO READING ROOMS SERVICES

Aurea Sousa, PhD

University of the Azores, Mathematics Department; CEEAplA, Portugal

Licínio Tomás, PhD

University of the Azores, History, Philosophy and Social Sciences Department, CES, Portugal

Oswaldo Silva, PhD

University of the Azores, Mathematics Department; CMATI and CES, Portugal.

Helena Bacelar-Nicolau, PhD

University of Lisbon, Faculty of Psychology, Laboratory of Statistics and Data Analysis, Portugal.

Abstract:

This paper re-examines and deepens the study of a portion of the data collected within the context of a wider 2007 research project conducted in the Autonomous Region of Azores. The 2007 study aimed to understand users' habits, attitudes and cultural practices, concerning reading and utilization of different library services, archives and museums. Based upon knowledge that only data analysis of a representative sample can supply, the study aimed to identify the aspects that should be prioritized in a process of restructuring the cultural services of leisure and reading to be implemented. This paper, utilizing data from the 2007 study, presents some results from the Ascendant Hierarchical Cluster Analysis (AHCA) of symbolic objects, according to the treatment to which they were submitted. These objects are described by different symbolic attributes pertaining to the latent variable '*Degree of Satisfaction*'. This variable was evaluated according to different dimensions of on-the-spot reading and consultation services. The aggregation criteria used in this study belong to a parametric family of methods and the similarity measure used is the weighted generalized affinity coefficient, for symbolic data. The validation of the clustering results is based on some validation measures.

Key Words: Ascendant Hierarchical Cluster Analysis, Symbolic Data, Affinity Coefficient, VL Methodology.

Introduction

The increasing use of databases, often large ones, in diverse areas of study, including Sociology, makes it pertinent to summarise data in terms of their most relevant concepts. These concepts may be described by types of **complex data**, also known as **symbolic data** [1].

In a symbolic data table, lines correspond to **symbolic objects** (SO's) and columns to **symbolic variables**, which may contain not just one value, as usual, but values, such as subsets of categories, intervals in real axes, or frequency distributions.

After a process of grouping, symbolic data can be used to describe the properties of the clusters obtained in terms of the initial variables, thus making it possible to summarize large sets of data. Some symbolic data sources can be found, for instance, in [1] and in [2]. It is important to stress that relational databases are an important source of symbolic objects, when we wish to study the properties of a set of units whose description needs the merging of several relations ([3]; [1]).

Stéphan [3] defines operators that can be used to retrieve groups of individuals from a relational database. The descriptions of groups are represented by assertions, in the formalism of symbolic data, which can be automatically obtained from a population or from a selected sample, and can be analysed using methods of Symbolic Data Analysis (SDA), which take under consideration a previous knowledge of the data (e.g., *metadata*) [4].

Based on the affinity coefficient between two discrete probability distributions as defined by Matusita [5], Bacelar-Nicolau [6] suggests the use of the affinity coefficient as a basic similarity coefficient between the columns or the lines of a data matrix. Later on she extended that coefficient to different types of data, including complex data (symbolic data) and variables of mixed types (heterogeneous data), possibly with different weights ([7]-[10]).

In the complex nature of cultural consumption, preferences show that in living practices there are different social denominators involved in the establishment of more long-lasting habits, which become the basis for daily reading routines. Despite several sorts of incentives, the interest in reading, in visiting museums or in attending libraries – as well as the regularity with which these occur – are still unequally distributed in the population. In most cases, these interests and activities are confined to groups with particular interests, such as those concerning to ludic or relaxing and recreational activities, cultural or professional enrichment.

The factors and the contexts related to the formation and development of such habits remain still somehow ambiguous, considering that this is still a relatively unknown process in the modern ways of life. There is, nonetheless, a certain number of indicators and information already contextualized and ready to be gathered and systematized through a set of categories, built and used specifically to serve the general purposes of this study.

The first part of the questionnaire used in the 2007 study, on the habits, attitudes and practices of readers and consumers of library services, included a filtering question: "*Have you ever visited or been to any public library in your region or island*", to which each of the 1684 respondents Answered (Yes or No).

The data matrix obtained from the question of filtering was used to provide the basis for the analysis which this article seeks to outline, with some results concerning the application of some methods of Ascendant Hierarchical Cluster Analysis of symbolic objects. The applied approach using symbolic objects allows us to exhibit differences on the users' degree of satisfaction according to a greater level of granularity, such as the island or the employment status. In this process was necessary to transform the standard data matrix (674 X 14) into two symbolic data matrices by the generation of symbolic objects.

We present, in Section 3, the main results of the Ascendant Hierarchical Cluster Analysis (of symbolic data) based on data from the wider 2007 study, conducted in the entire Autonomous Region of Azores. In the present work we consider different items/indicators used to measure the latent variable '*Degree of Satisfaction*', which was evaluated according to several aspects of on-the-spot consultation and reading services.

Used Methods of Cluster Analysis

The purpose of cluster analysis (of classical data as well as of symbolic data) is to build, from a data matrix ($N \times p$), a classification which is appropriate for a set E of data units or a set Y of variables, with the purpose of obtaining "homogenous" clusters of objects in a population Ω or E , so as to allow objects of the same cluster to present great similarity, whereas objects of different clusters will be much more different.

Given a matrix of proximity between symbolic objects, these can be classified through classical agglomerative hierarchical algorithm ([11]-[13]) or probabilistic ones [2]. Once we have obtained the matrix of the proximity between the elements of E , the classification can be obtained regardless of the fact that the data are symbolic [14]. This is the case of the algorithms of AHCA used in Section 3 of this paper.

The algorithms of AHCA for the case where the elements of a set $E = \{1, \dots, N\}$ of symbolic objects are described by p symbolic variables, Y_1, \dots, Y_p , take as their starting point, like they do in classical ones, the finest partition P_0^s , with N clusters of a single symbolic object. If, at the j -th stage, clusters A and B are put together, a new partition P_j^s , resulting from that stage, is defined as follows:

$$P_j^s = (P_{j-1}^s - \{A, B\}) \cup \{A \cup B\}$$
 The process includes $N-1$ stages and ends when the less fine partition $\{\{E\}\}$ has been obtained, which is composed of a single cluster containing all of the N symbolic objects. Thus, if in the first stage of the process there are N clusters, in the second one there will be $N-1$, and so forth, since in each new step a new cluster is formed, by joining two previously formed clusters [2].

Comparison measures

The data matrices analysed in Section 3.2 generically correspond to symbolic data tables $\underline{X} = (\xi_{kj})_{N \times p}$ referring to N data units (typically groups of individuals) and to p symbolic variables, Y_j , $j=1, \dots, p$, where Y_j is a modal variable with m_j modalities $\{1, \dots, m_j\}$, and each entry $\xi_{ij} = (n_{kj1}, \dots, n_{kjm_j})$ of \underline{X} contains either the absolute frequencies or the relative frequencies, x_{kiv} , of individuals (in unit k) which share category v of variable Y_j . In Table 1 each row $x_k = (\xi_{k1}, \dots, \xi_{kp})$ describes an individual modal symbolic object on the basis of p variables.

Table 1. Matrix of symbolic Data \underline{X}

Variables	...	Y_j	...
Data units
\vdots
k	...	$(x_{kj1}, \dots, x_{kjm_j})$...
\vdots
k'	...	$(x_{k'j1}, \dots, x_{k'jm_j})$...
\vdots

Taking as their starting point a table like Table 1, Nicolau and Bacelar-Nicolau ([15]; [8], [9]) suggest the use of the weighted generalized affinity coefficient in cases where symbolic objects are described by vectors of p probabilities or frequency distribution, or in case of some other data support to which this sort of description could be extended, such as histograms and variables whose values are intervals of real axes.

Let $A_{kk'} = (a_{jj'}^{kk'})$, $1 \leq j, j' \leq p$, be a table associated with the pair of symbolic objects $(O_k, O_{k'})$, where $a_{jj'}^{kk'}$ is the (partial, local or variable-specific) **affinity coefficient between the frequency distributions** ξ_{kj} , $\xi_{k'j'}$, corresponding to the variables j and j' ([8]; [15]):

$$a_{jj'}^{kk'} = \text{aff}(\xi_{kj}, \xi_{k'j'}) = \sum_{\ell=1}^{m_j} \sqrt{\frac{n_{kj\ell}}{n_{kj}} \cdot \frac{n_{k'j'\ell}}{n_{k'j'}}}, \quad 1 \leq j, j' \leq p, m_j = m_{j'}$$

where $n_{kj} = \sum_{\ell=1}^{m_j} n_{kj\ell}$ $\left(n_{k'j'} = \sum_{\ell=1}^{m_{j'}} n_{k'j'\ell} \right)$ is the number of individuals belonging to the data unit

(group k) for which the variable Y_j ($Y_{j'}$) has been observed and $0 \leq \text{aff}(\xi_{kj}, \xi_{k'j'}) \leq 1$. It is important to note that the value 1 is obtained if the frequency distributions ξ_{kj} and $\xi_{k'j'}$ are identical or proportional and that the result is value 0 if they are orthogonal. For a binary variable the variable-specific affinity similarity coefficient turns out to be the so called Ochiai coefficient ([8]).

Let $P_{kk'} = (\pi_{jj'})$, $1 \leq j, j' \leq p$, be a table of weights to be used in the convex linear combination, with $\pi_{jj'} \geq 0$ and $\sum_{1 \leq j, j' \leq p} \pi_{jj'} = 1$. The weighted generalised affinity coefficient between O_k and $O_{k'}$ is given by:

$$a(k, k') = \sum_{1 \leq j, j' \leq p} a_{jj'}^{kk'} * \pi_{jj'} \quad (I)$$

and takes values in $[0, 1]$. In fact, this coefficient is a general similarity coefficient, which uses all the information included in the matrix $(p \times p)$, $A_{kk'}$, of the similarity coefficients crossing the p laws which describe the symbolic objects O_k and $O_{k'}$, and is defined by a convex linear combination of the values of this matrix ([15]).

The used probabilistic aggregation criteria in the present work (AVL, AV1 and AVB) in the scope of the probabilistic approach of AHCA, named VL methodology (V for *Validity*, L for *Linkage*) resorts, essentially, to probabilistic notions for the definition of the comparative functions. In fact, the

VL-family is a set of agglomerative hierarchical clustering methods, based on the cumulative distribution function of basic similarity coefficients ([7]; [16]).

The analysis of the symbolic data matrices was done by using the weighted generalised affinity coefficient ([15]; [8]; [2]), with the weight option: $\pi_{jj}=1/p$ if $j=j'$ and $\pi_{jj}=0$ if $j \neq j'$ [15]. The measure of comparison between the elements was combined with classical aggregation criteria, *Single Linkage* (SL) and *Complete Linkage* (CL), and probabilistic ones, AVL, AV1 and AVB ([7]; [16]).

Validation measures

Let E be a set of elements to classify, of cardinal N , and $\{C_1, C_2, \dots, C_c\}$ a partition in c clusters of E . When the cardinality of E is high, it is useful to achieve a reduced or condensed hierarchy to the more representative levels. For this reason, the values of validation indexes ([2]; [17]) obtained from the values of the proximity matrix between elements, such as, for instance, the global statistics of levels, STAT, [6] can also be calculated, even in the case of symbolic data matrices.

On what validation of results is concerned, in this paper we chose to use some of the validation indexes mentioned in [2] and in [17], namely the STAT, DIF [6], $P(I2mod, \Sigma)$ and γ from Goodman and Kruskal indexes. To determine the appropriate number of clusters we used measures defined in the present section adapted for the case of similarity measures.

Considering $F = \{(i, j) : i, j \in E\}$, the global statistics of levels (STAT) ([6]) is defined by:

$$STAT = \frac{C_V}{\sqrt{\#(R) * \#(S) * \frac{\#(F) + 1}{12}}}, \text{ where} \quad (2)$$

$$C_V = C_B - \frac{\#(R \times S)}{2}, \quad C_B = \#(G(\text{PI}) \cap (R \times S)), \quad G(\text{PI}) = \{(ij, k\ell) \in F \times F : s_{ij} \geq s_{k\ell}\}$$

represents the graph of the initial preordination defined in $F \times F$, R is the set of the pairs of elements joined in the same cluster and S is the set of separated pairs.

Bacelar Nicolau ([6]) proposed the calculation of the values of STAT for each level ($K=1, \dots, m$, where m represents the total number of levels) of the hierarchy of classification and the values of the:

$$DIF(k) = STAT(k) - STAT(k-1). \quad (3)$$

The $P(I2mod, \Sigma)$ measure ([2]; [17]) is a normalization of the $P(I2, \Sigma)$ measure, reported in [14], and is defined as:

$$P(I2mod, \Sigma) \equiv \frac{1}{c} \sum_{r=1}^c \frac{\sum_{i \in C_r} \sum_{j \in E/C_r} s_{ij}}{n_r * (N - n_r)}, \quad (4)$$

where s_{ij} is a similarity measure between the elements i and j , n_r is the number of elements of the cluster C_r , N is the number of elements to classify and c is the number of clusters of the partition.

The γ index, proposed by Goodman and Kruskal, has been widely used ([14]). Comparisons are made between all within-cluster similarities s_{ij} and all between-cluster similarities $s_{k\ell}$. A comparison is deemed to be concordant (resp., discordant) if s_{ij} is strictly greater (resp., less) than $s_{k\ell}$. This index is defined by:

$$\gamma \equiv (S_+ - S_-) / (S_+ + S_-), \quad (5)$$

where S_+ (resp., S_-) denotes the number of concordant (resp., discordant) comparisons.

Result Analysis

Data presentation

The enquiry which was elaborated included, among others, a set of questions about the evaluation of the degree of satisfaction with the different indicators or dimensions referring to the quality of on-the-spot reading services, which can be consulted in Table 2.

Table 2. Variables/items/ indicators taken under consideration

V1-Accessibility and Easiness of Access to what you are looking for
V2-Opening Hours
V3-Assistance/ Waiting time
V4-Conditions for consulting/locating what you are looking for
V5-Reading conditions of spaces and facilities
V6-Conditions under which publications are loaned or handed out
V7- Thematic quality and interest
V8- Technical counselling and orientation
V9-Courtesy and availability of library staff
V10-Quality and thematic diversity of publications
V11-Quality of the catalogues and of the organization of publications for consultation
V12-Efficacy and pertinence of the consulted material
V13-Loan Services
V14-Overall service satisfaction

The questionnaire used in the 2007 study contained a filter-question: “Have you ever visited or been to any public library in your island or region?”, to which each of the 1684 respondents answered (Yes or no). Thus, the final classical data matrix is a matrix (674x14) constituted by 674 individuals, who answered “yes” to the above mentioned filter question, and by 14 variables/items which aimed at assessing the level of satisfaction regarding the quality of on-the-spot reading services.

The 674 respondents, who had already visited a public library in their island or region, indicated verbally their degree of satisfaction with statements containing five Lickert-type possible positions (1-Very unsatisfied, 2-Unsatisfied, 3-Neutral, 4-Satisfied, 5-Very satisfied). The answering modalities are distributed around a neutral or indifference axis, permitting, therefore, a precise determination of positional attitude regarding the evaluated service quality. This study also took into consideration the following nominal variables: “**Island**” (1- Santa Maria; 2- São Miguel; 3- Terceira; 4- Graciosa; 5- Faial; 6-Pico; 7-São Jorge; 8-Flores; 9-Corvo) and “**Employment Status**” (S1-“Employed on a permanent contract of employment”; S2 -“Self-employed”; S3-“Non-paid family work”; S4-“Looking for the first job”; S5-“Retired”; S6 -“Student/Working student”; S7- “Employed on a fixed-term contract”; S8-“Freelancer or similar”; S9-“Unemployed”; S10-“Homemaker”; S11-“Disabled/Handicapped/Unable to work”; S12-“Others”). Two matrices of symbolic data were also obtained from the above mentioned matrix (674 x 14). For this purpose some groups of individuals were created, according to the variables “Island in which the interview took place” and “Employment status”. Each matrix contains fourteen columns and a number of lines determined by the number of categories of the variables which were taken into consideration to form the groups. In Table 3, the notation F_i indicates the number of individuals included in category i of variable j , respectively with $1 \leq i \leq 9$ and $1 \leq i \leq 12$, for the variables “Island” and “Employment status”. The totals in this table are different because the two variables contain a different number of missing values.

Table 3. Nominal Variables – Application to real data

Island		Employment status	
Categories	F_i	Categories	F_i
Sta Maria	72	S1- Employed PC	259
S. Miguel	94	S2- Self-employed	49
Terceira	148	S3- Non-paid family work	6
Graciosa	82	S4- Looking for first job	5
Faial	48	S5- Retired	52
Pico	79	S6- Student/Working St.	133
S. Jorge	78	S7- Employed F-TC	103
Flores	27	S8- Freelancer	8
Corvo	46	S9- Unemployed	18
----	----	S10- Homemaker	31
----	----	S11- Disab./Hand./Unable	4
----	----	S12- Other	0
TOTAL	674		668

AHCA of the data units (Symbolic Data)

As mentioned in Section 2, two symbolic data matrices were created, according respectively to the island in which the interview took place and the respondents' employment status.

Table 4. A part of the symbolic data matrix –Islands

	V1	V2	...
<i>Sta Maria</i>	<i>S (1.00)</i>	<i>VS (0.04), N (0.13), S (0.74), U (0.08), VU (0.05)</i>	...
<i>S. Miguel</i>	<i>VS (0.74), S (0.26)</i>	<i>VS (0.54), N (0.06), S (0.35), U (0.04)</i>	...
<i>Terceira</i>	<i>S (0.57), N (0.22), U (0.17), VU (0.04)</i>	<i>VS (0.01), N (0.23), S (0.60), U (0.13), VU (0.03)</i>	...
<i>Graciosa</i>	<i>S (0.79), N (0.18), U (0.02)</i>	<i>VS (0.02), N (0.21), S (0.72), U (0.04), VU (0.01)</i>	...
<i>Faial</i>	<i>S (1.00)</i>	<i>VS (0.08, N (0.04), S(0.83), U(0.02), VU (0.02)</i>	...
<i>Pico</i>	<i>N (0.71), U (0.27), VU (0.03)</i>	<i>VS (0.03), N (0.51), S (0.34), U (0.08), VU (0.05)</i>	...
<i>S. jorge</i>	<i>S (0.88), N (0.03), U (0.01), VU (0.08)</i>	<i>N (0.14), S (0.74), U (0.06), VU (0.05)</i>	...
<i>Flores</i>	<i>S (1.00)</i>	<i>VS (0.07), N (0.04), S (0.74), U (0.15)</i>	...
<i>Corvo</i>	<i>S (1.00)</i>	<i>VS (0.04), N (0.11), S(0.76), U(0.04), VU (0.04)</i>	...

On what the Cluster analysis of the symbolic data units of Table 4 (islands) is concerned, and in accordance with the validation indexes used ([2]; [17]), the AVI tree shows at its most significant level (level 6), a partition in three clusters. Thus, we obtained three differentiated satisfaction profiles according to the different indicators under analysis, as can be seen in Table 5 and in Figures 1 to 4.

Figure 1. Dendrogram obtained from the AVI method

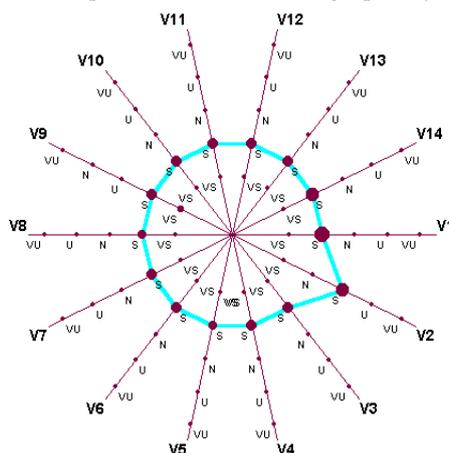
The first cluster is constituted by the islands of *Santa Maria*, *São Jorge*, *Flores*, *Corvo*, *Terceira*, *Graciosa* and *Faial*, to which the individuals who are generally pleased with the on-the-spot reading services belong. The Zoom Stars for all these islands were identical to those shown in Figure 2, that is, the most frequent answer provided by respondents in to these islands was "satisfied" for all fourteen variables. So, we present only the obtained one for the island of Faial.

Table 5. Values of some validation measures for the partitions – Islands - AVI

Nível	STAT	DIF	P(I2mod, Σ)	γ
1	1.6846	-----	0.8953	1
2	2.3481	0.6635	0.8897	1
3	2.8332	0.4851	0.8832	1
4	2.9504	0.1172	0.8755	0.8323
5	3.9638	1.0134	0.8615	0.893
6	5.0217	1.0578	0.8257	0.9937
7	2.8919	-2.1298	0.8385	0.6786

In the 2D Zoom Star, axes are linked by a line that connects most frequent values of each variable, so allows us to identify the main characteristics of the symbolic objects. Although the star shapes of all 2D Zoom Stars of the islands of cluster 1 are the same, there are some differences at the level of frequency distributions. Indeed, the little differences between the frequency distributions of the first and second sub-cluster of cluster 1 could be observed from a 3D Zoom Stars featuring the distributions corresponding to each variable with weighted values. Much of what is considered satisfying depends on the structural conditions of resource and equipment distribution, which attributes to each of the nine islands a specific position concerning reading and other services. Thus, this set of islands, with similar satisfaction profiles, corresponds to situations occurring in island libraries (Central or Municipal) whose traditional functioning captivates the regular, loyal users.

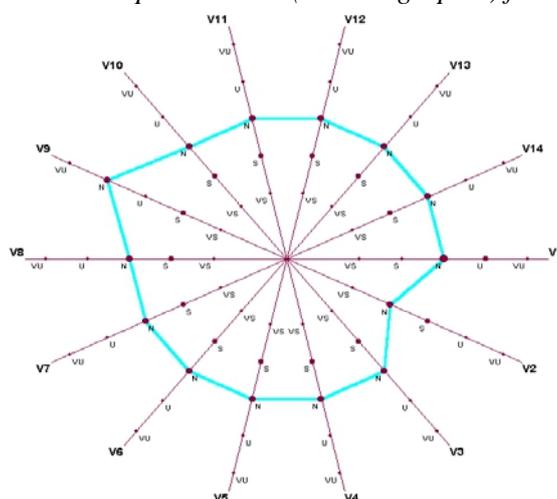
Figure 2. 2D Zoom Star representation (in a 2D graphic) for the island of Faial



In the case of the islands Terceira and Faial, although they are equipped with central libraries and archives (RLAAH- Regional Library and Archive of Angra do Heroísmo and RLAFA- Regional Library and Archive of Faial), at the time the study was conducted these were still functioning the same way, due to the unfit conditions of their facilities and equipments. There was, nevertheless, a project already underway for constructing new facilities for the Angra do Heroísmo Library, in the island of Terceira, and new facilities were ready to be inaugurated for the Library of Faial.

The second cluster is constituted by the island of Pico alone, where the most frequent answer given by the individuals of this island to the set of 14 variables was the neutral one (see Figure 3). This fact is probably due to the atypical or exceptional situation of the libraries in this island, at the time the data were collected on the field (2007). In fact, the three Municipal libraries of Pico were being remodelled at the time, in the context of a funding plan, the Funding Plan for Reconstructing and Remodelling Municipal Libraries, applicable to all of the national territory. The situation of indifference can perhaps be explained by the awaited expectations of service improvement after the transfer to new facilities.

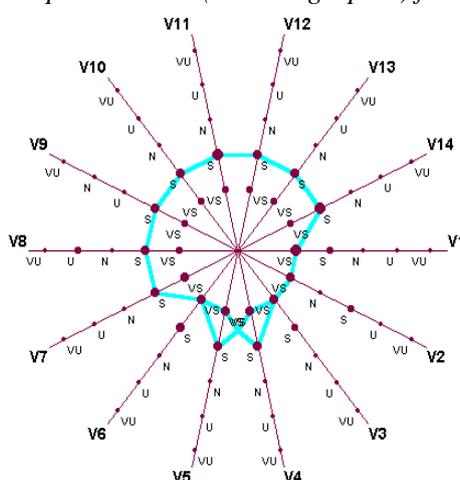
Figure 3. 2D Zoom Star representation (in a 2D graphic) for the island of Pico



The third and last cluster contains the island of São Miguel, the island where the degree of satisfaction was the highest (see Figure 4) on the variables regarding the *quality of the available materials/publications and the easiness of access to these materials (V1 to V6)*. At the time, the Regional Library and Archive of Ponta Delgada (RLAPD) was, in the archipelago, a paradigmatic case of transition to a new concept of the social and cultural function which is nowadays attributed to libraries. In fact, in the light of the public reading services this process which included remodelling the central services and moving to other facilities in more accessible buildings made a way for a new concept of promoting this sort of community services and cultural activities, intended for users differentiated by their needs (children, teenagers, elderly, etc.). Nowadays, more than in the recent past, this sort of space is meant to support consultation and study for university students, but mostly for the students from the different schools on the island which holds 56% of the region's population. We should also stress the fact that the main university campus is on this island and thus the public and the university libraries complement each other. The largest university campus, with the highest number of courses in the archipelago, is located in Ponta Delgada, in the island of São Miguel, and this has obvious implications for the materials and the subject areas encompassed by these libraries.

Figure 4. 2D Zoom Star representation (in a 2D graphic) for the island of São Miguel

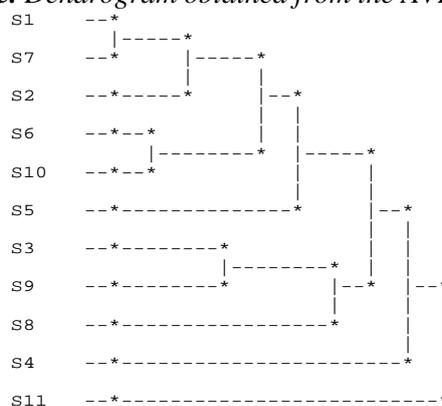
SMG



In relation to AHCA of the 11 symbolic objects obtained from the variable "Employment Status", the values of the STAT and γ of Goodman and Kruskal indexes, which can be seen in Table 6, allow us to conclude that the AV1 and AVB trees originated, at their most significant level (level6), a partition in five clusters (see Figure 5).

The first cluster contains the employment status of S1- "Employed on a permanent contract of employment"; S7- "Employed on a fixed-term contract"; S2 - "Self-employed"; S6- "Student/Working student"; S10- "Homemaker" and S5- "Retired", which encompasses the individuals who are in a paid working situation, the students and those who enjoy a greater amount of free time. In fact, the individuals from this group correspond to a similar profile of satisfaction with on-the-spot reading services.

Figure 5. Dendrogram obtained from the AVB method



The second cluster contains the working status S3- "*Non-paid family work*" and S9- "*Unemployed*", therefore relating to individuals who have no income. Each of the remaining three clusters corresponds to relatively atypical situations. It is therefore not strange that they can be found in the most significant partition, in clusters which contain a single symbolic data unit (symbolic object) and that the individuals included are but a small number of the studied sample, as can be seen in Table 3.

Table 6. Validation indexes – Employment status - AVB

Nível	STAT	DIF	P(I2mod, Σ)	γ
1	1.7008	-----	0.9005	1
2	2.1582	0.4573	0.8962	0.9057
3	2.9493	0.7912	0.8906	0.8922
4	3.0448	0.0955	0.8868	0.832
5	4.9658	1.921	0.8765	0.9752
6	5.7632	0.7975	0.8663	0.9968
7	5.4171	-0.3461	0.862	0.9069
8	5.54	5.8862	0.8339	0.9152
9	4.6262	4.6262	0.8046	0.9422

The third cluster includes the group of individuals who are *freelancers or similar* (symbolic object S8) and their profile is grounded, for the most part, on a situation of greater satisfaction, due to the adequacy of the materials to their regular and very specific consultation needs, done for working purposes, such as consulting of governmental decrees, official diplomas and manuals made available by the library to the general public. The fourth cluster includes the group of individuals who are S4- "*Looking for the first job*" (symbolic object S4). Finally, the fifth cluster, the one which most stands out from the remaining ones, includes the individuals who are S11- "*Disabled/Handicapped/Unable to work*" (symbolic object S11).

Worthy of notice is also the fact that, as can be seen in Table 6, the levels to which the maximum on-the-spot DIF values correspond are respectively levels 3, 5 and 8, which, according to Bacelar-Nicolau [6], indicates that it was on those levels that the formations of the most important, most internally coherent and best differentiated sub-clusters occurred. For this reason and for sociological interpretation purposes, we must mention, on level 3, the formation of cluster {S1, S7, S2}, which corresponds to workers who belong to a paid employment status, as can be seen in the dendrogram of Figure 5.

Conclusion

At a time when the use of increasingly complex databases has become the norm it is crucial to develop statistical methods which allow us to extract useful knowledge from these data while, when justifiable, making it possible to maintain their confidentiality. An advantage of the weighted generalised affinity coefficient in the field of Symbolic Data Analysis is the fact that it can be applied to different types of data. In this case, different proximity matrices are obtained for each sort of data, with a convex linear combination of these initial matrices being afterwards determined (Bacelar-Nicolau, 2000). We then obtain a single global proximity matrix based on a single measure of comparison between elements, to which different aggregation criteria, including the ones from VL methodology, can be applied.

In this paper we tried to show how the affinity coefficient can be applied in Hierarchical Cluster Analysis of units of symbolic data (symbolic objects), described by symbolic modal variables, keeping in mind the growing importance of this sort of variables. We also tried to show how some classical validation indexes can be extended to the case of symbolic data, since the validation of results is an essential step in Cluster Analysis. Finally, clusters referring to the selected partitions were constantly submitted to careful interpretation, as a way of obtaining useful knowledge on the context the data belong to, for, no matter how developed computers become, the correct analysis of data will always depend on human intervention.

References:

- [1] H.-H. Bock and E. Diday, (Eds.), “Analysis of Symbolic Data: Exploratory Methods for Extracting Statistical Information from Complex Data”, Series: Studies in Classification, Data Analysis, and Knowledge Organization. Berlin: Springer-Verlag, 2000.
- [2] A. Sousa, *Contribuições à Metodologia VL e Índices de Validação para Dados de Natureza Complexa*. Tese de Doutoramento, Universidade dos Açores, Ponta Delgada, 2005.
- [3] V. Stéphan (1996). Extracting Symbolic Objects from Relational Databases. *IEEE*, pp. 514-519, Available: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=00558606&tag=1>.
- [4] E. Diday, “From Data to Knowledge: Probabilistic Objects for a Symbolic Data Analysis”, in *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 19, 1995.
- [5] K. Matusita, “On the Theory of Statistical Decision Functions”, *Ann. Instit. Stat. Math.*, vol.III, pp.1-30, 1951.
- [6] H. Bacelar-Nicolau, “*Contribuições ao Estudo dos Coeficientes de Comparação em Análise Classificatória*”. Tese de Doutoramento, FCL, Universidade de Lisboa, 1980.
- [7] H. Bacelar-Nicolau, “*Two Probabilistic Models for Classification of Variables in Frequency Tables*”, in: H. H. Bock (ed.), IFCS-87, 1988, pp. 181-186.
- [8] H. Bacelar-Nicolau, “*The Affinity Coefficient for Complex Data*”, in: Analysis of Symbolic Data: Exploratory Methods for Extracting Statistical Information from Complex Data, H. H. Bock and E. Diday, Ed. Berlin: Springer-Verlag, 2000, pp.160-165.
- [9] H. Bacelar-Nicolau, “On the Generalised Affinity Coefficient for Complex Data”, *Biocybernetics and Biomedical Engineering*, vol. 22 (1), pp. 31-42, 2002.
- [10] H. Bacelar-Nicolau, H.; F. Nicolau; A. Sousa.; L. Bacelar-Nicolau, “Measuring similarity of complex and heterogeneous data in clustering of large data sets”, *Biocybernetics and Biomedical Engineering*, vol. 29 (2), pp. 9-18, 2009.
- [11] E. Diday, “*The Symbolic Approach in Clustering and Related Methods of Data Analysis: The Basic Choices*”, in: H. H. Bock (ed.) IFCS-87, 1988, pp. 673-684.
- [12] K. C. Gowda and E. Diday, “Symbolic Clustering Using a New Dissimilarity Measure”, *Pattern Recognition*, vol. 24 (6), pp. 567-578, 1991.
- [13] K. C. Gowda and E. Diday, “Symbolic Clustering Using a New Similarity Measure”, *IEEE Transactions on Systems, Man and Cybernetics*, vol. 22 (2), pp. 368-378, 1992.
- [14] A. D. Gordon, *Classification*, 2nd ed. London: Chapman & Hall, 1999.
- [15] F. Nicolau and H. Bacelar-Nicolau, “*Clustering Symbolic Objects Associated to Frequency or Probability Laws by the Weighted Affinity Coefficient*”, in H. Bacelar-Nicolau, F. C. Nicolau and Jacques Janssen, Ed., ASMDA-99, 1999, pp. 155-158.
- [16] F. Nicolau and H. Bacelar-Nicolau, “*Some Trends in the Classification of Variables*”, in C. Hayashi, N. Ohsumi, K. Yajima, Y. Tanaka, H. H. Bock and Y. Baba, Ed., IFCS-96, 1998, pp. 89-98.
- [17] A. Sousa; H. Bacelar-Nicolau; F. Nicolau, “A metodologia VL e Validação em Análise Classificatória de Dados de Natureza Complexa: Uma Aplicação a Dados Reais”, in Ferrão, M. et al., Ed., SPE-2006, 2007, pp. 775-788.

ON USING SAGE TO SOLVE CONSTRAINED OPTIMIZATION PROBLEMS APPLYING THE LAGRANGE MULTIPLIERS METHOD

Maria do Carmo Martins, PhD

CMATI & Department of Mathematics, University of Azores, Portugal

Francisco Martins, PhD

LaSIGE & Department of Informatics, Faculty of Sciences, University of Lisbon, Portugal

Abstract:

This paper combines Calculus and Programming to solve constrained optimization problems common in many areas, notably in Economics. It uses Lagrange multipliers, a well-known technique for maximizing (or minimizing) functions, and the free open-source mathematics software system Sage to compute the maximum (minimum) automatically. Moreover, Sage can be used interactively to work out the solution and to graphically interpret the results, which we find a valuable and practical approach in teaching such techniques to the undergraduate level. In this paper we carry out an exercise describing how these three interdisciplinary areas can work together.

Key Words: Constrained optimization problems; Octave programming language; Lagrange multipliers method

Introduction

Finding the extrema of a function subjected to some restriction is a well-studied problem in Mathematics with multiple, direct applications in Economics. For instance, it is plausible for a company to determine the amount of units of each product it must produce in order to maximizing its profits. Mathematical Analysis and Operational Research come into the rescue by providing tools for solving this kind of problems. Techniques like computing function derivatives or the Dantzig's Simplex algorithm [1] are the recurrent candidates, depending on the problem at hand (i.e., linear vs. non-linear problem). In this paper we stick to the methods provided by Mathematical Analysis, which are usually part of the undergraduate Calculus curricula. The novelty we propose is the using of mathematical software, in particular Sage, to interactively solve the problem and to help in visualizing the solution. We believe that using this approach is more appealing and enlightening for undergraduate students to grasp both the theoretical and the practical aspects of these methods.

Sage [2] is a high-level, open-source programmable system targeted to symbolic computation (amongst other things). Unlike the programming languages Fortran, C, or Java, Sage positions itself at a much higher level of abstraction, providing an environment for finding the exact roots of systems of non-linear equations, compute matrix inverses, integrate and differentiate functions, or plotting solutions' graph of a differential equation, to name a few of its features. We find that mastering such a tool (free of costs) can be of invaluable help, not just for the application we propose in this paper, but for many other subjects (e.g., Algebra, Statistics).

We use a very simple example throughout the paper:

A company wants to find its maximum production level that is modeled by the Cobb-Douglas function [3]

$$f(x, y) = 50x^{3/4}y^{1/4},$$

where x represents the units of labor and y represents the units of capital. Each labor unit costs €150 and each capital unit costs €250. The total expenses for labor and capital cannot exceed €40 000.

In this paper we recall how to solve this problem using the Lagrange multipliers [4, 5] and then illustrate how Sage becomes very handy in computing and illustrating how the method works.

The main text of the paper is organized as follows. The first section sets the mathematical grounds for the problem and recalls the application of the Lagrange multipliers method; the second section gives a brief introduction to Sage presenting the features needed for solving the running problem; finally, the third section illustrates the usage of Sage to solve, to visualize, and to understand the solution to the problem.

Main Text

Optimizing using Lagrange multipliers method

In general, a problem of finding the local maximum (resp. minimum) of a function $f(x, y)$ subjected to some restriction equation $g(x, y) = c$ can be directly solved by using two systems of equations, one for f and g , another for its first derivatives, provided that both functions have continuous first derivatives. Then, inspecting the second derivative of both functions we can decide whether the critical points correspond to a local maximum or minimum. However, this approach is rather cumbersome.

The Lagrange multipliers method presents a much more elegant solution and can easily be applied in this situation. The Lagrange function Λ (the *Lagrangian*) assumes the general form

$$\Lambda(x, y, \lambda) = f(x, y) + \lambda(g(x, y) - c), \quad (1)$$

where λ is a new variable, denoted the Lagrange multiplier. Briefly, if $f(x_0, y_0)$ is a maximum of function f , there exists some λ_0 such that the triple (x_0, y_0, λ_0) is a stationary point of the Lagrangian, i.e., a point where the first partial derivatives of Λ are zero. The multiplier λ_0 represents the rate of change in (x_0, y_0) with respect to c . The interested reader may refer to [4, 5] for further details on the Lagrange multipliers method.

For applying the Lagrange multipliers method to our running example, we start by determining the Lagrangian. The constraint in this problem comes from the fact that “*The total expenses for labor and capital cannot exceed €40 000*” that can be captured by the equation $g(x, y)$ defined as follows

$$150x + 250y = 40000$$

Then combining functions f and g as defined in (1), we get the Lagrangian

$$\Lambda(x, y, \lambda) = 50x^{3/4}y^{1/4} - \lambda(150x + 250y - 40000)$$

To solve the problem we need to compute the partial derivatives of Λ , provided below

$$\begin{aligned} \frac{\partial \Lambda}{\partial x} &= -150\lambda + \frac{75y^{1/4}}{2x^{1/4}} \\ \frac{\partial \Lambda}{\partial y} &= -250\lambda + \frac{25x^{3/4}}{2x^{3/4}} \\ \frac{\partial \Lambda}{\partial \lambda} &= -150x - 250y + 40000 \end{aligned}$$

and determine the stationary point of the Lagrangian by solving the system of non-linear equation formed by the partial derivatives (i.e., $\partial\Lambda/\partial x = \partial\Lambda/\partial y = \partial\Lambda/\partial\lambda = 0$) in order to x , y , and λ . Solution to the system is when $x = 200$, $y = 40$, and $\lambda = 0.167185$. The economical interpretation of the Lagrange multiplier is that it represents the percentage of each additional Euro spent on capital that will turn into production. For example, if an additional €10 000 were spent on capital, then it would be translated into $0.167185 \times 10\,000 = 1671.85$ additional units of production. If the problem has more than one constraint, then equation (1) needs to account for the additional restriction function and its multiplier. For a new constraint $h(x, y) = d$ equation (1) would be rewritten as

$$\Lambda(x, y, \lambda, \mu) = f(x, y) + \lambda(g(x, y) - c) + \mu(h(x, y) - d)$$

Sage in a nutshell

Sage covers many areas of Mathematics, namely Algebra, Calculus, Combinatorics, Numerical mathematics, and Number Theory. It is available on Linux, Windows, and Mac OS X operating systems; it is simple to install and present a clean and intuitive environment, making it ideal

for interactive sessions. Its usage can range from a sophisticated calculator to an advanced computer algebra system. Sage is an open source alternative to, and rivals with, Magma [6], Maple [7], Mathematica [8], and MATLAB [9], commercial systems widely used in engineering, science, and economics.

We cover variable and function definitions in order to be able to illustrate some symbolic manipulations, in particular for computing partial derivatives, solving systems of equations, and plotting two and three dimension graphics. Symbolic variables are introduced using the *var* function. For instance, $xVar = var('x')$, introduces a symbolic name x and binds it to variable $xVar$, possibly creating $xVar$. Then, whenever $xVar$ is used, it refers to the symbolic name x . We usual bind the symbolic name to a variable with the same name. The assignment, $yVar = xVar$, introduces a new variable $yVar$ that is bound to the same symbolic name x . So, expression $yVar + xVar$ evaluates to $2x$.

A function is a binding of a name to an expression. (More advanced functions can be created using the Python language.) The function *tripleSquare* that computes the triple of the square of x can be defined as $tripleSquare = 3 * xVar ^ 2$. Expression $tripleSquare(2)$ applies function *tripleSquare* to argument 2, yielding number 12.

A more interesting feature is plotting the graph of *tripleSquare* between 0 and 5. The task can be easily achieved using $plot(tripleSquare, (x, 0, 5))$. Many additional arguments may be supplied to function *plot*, like defining colors, graph title, various captions, but we omit such complexity in this paper.

Function differentiation and system solving finalize out brief tour on Sage. As for the former, we simply ask Sage $tripleSquare.diff(xVar)$. This computes the derivative of *tripleSquare* in order to $xVar$ (symbolic name x), yielding the expected result $6x$. The latter is a bit more involved; Function *solve* accepts as its first argument an equation to be solved. For instance, let us find when *tripleSquare* is zero; Expression $solve(tripleSquare == 0, xVar)$ will do the trick and solve yields $x = 0$. There are two things to notice: the usage of $==$ (double equal sign) to define an equation; the second argument of *solve* ($xVar$) indicates the variable in order to which the equation must be solved. In order to solve systems of equations, the first argument is a list of functions. Lists are comma-separated sequences of items (in this case of equations) enclosed in square brackets. The following example is self-explanatory.

$solve([x + y == 6, x - y == 4], x, y)$, yielding $x = 5$ and $y = 1$

Applying Sage to solve the problem using the Lagrange multipliers method

Figure 1 shows an excerpt of a Sage interactive session for addressing the running example. First we introduce three symbolic names x , y , l and bind them to three variables of the same name. Thereafter, we functions f and g . Notice that we define a function for the constraint, not an equation. This way is easier for dealing with g later in the session. However, we can rewrite equation $g(x,y) == 40\ 000$ in order to y easily using the *solve* function. Now we are ready to define the Lagrangian L , which is straightforwardly written using f and g . This line ends with $; L$. It is just a form of asking Sage to print the expression in a more convenient and standard notation. The three partial derivatives are obtained from expression L , asking Sage to derive according to some expression variable. In the present case, for instance, $L.diff(y)$, computes $\partial L / \partial y$. Determining the stationary point results from the direct application of the *solve* function. We provide the three equations involving the partial derivatives and ask for the system to be solved in order to x , y , and l . The expecting result appears clean and easy on the screen. The computation could be further condensed and written in a function that just receives L and returns the result. To additionally explore Sage we could plot graphs for the functions in order to better understand and visualize even the mathematical concepts underneath the, sometimes, not well understood notions of partial derivative, system solving, etc.

Lagrange multipliers method

```
x, y, l = var('x,y,l')
```

```
f = 50*x^(3/4)*y^(1/4); f
```

$$50x^{\left(\frac{3}{4}\right)}y^{\left(\frac{1}{4}\right)}$$

```
g = 150*x+250*y; g
```

$$150x + 250y$$

```
solve (g==40000, y)
```

$$\left[y = -\frac{3}{5}x + 160 \right]$$

```
L = f - l * (g - 40000); L
```

$$-50(3x + 5y - 800)l + 50x^{\left(\frac{3}{4}\right)}y^{\left(\frac{1}{4}\right)}$$

```
dfdx = L.diff(x); dfdx
```

$$-150l + \frac{75y^{\left(\frac{1}{4}\right)}}{2x^{\left(\frac{1}{4}\right)}}$$

```
dfdy = L.diff(y); dfdy
```

$$-250l + \frac{25x^{\left(\frac{3}{4}\right)}}{2y^{\left(\frac{3}{4}\right)}}$$

```
dfd1 = L.diff(l); dfd1
```

$$-150x - 250y + 40000$$

```
solve([dfdx == 0, dfdy == 0, dfd1 == 0], x, y, l)
```

$$\left[\left[x = 200, y = 40, l = \frac{1}{20} 5^{\left(\frac{3}{4}\right)} \right] \right]$$

Figure 1: Sage interactive session to solve the running problem

We present in Figure 2 a picture that, we believe, illustrates the solution to the problem. The figure is composed of four graphs: in green there is function f ; function g is depicted in red; in blue there is a plane marking the €40 000 restricting; and, finally, there is a black circle spotting the maximum. Notice that the mark is in the surface of f , and in the plane that is perpendicular to the blue plane and that passes in the intersection between the red and the blue planes. More graphs could be plotted to investigate the solution, but we leave it to the interested reader. For that reason we decided to include (top of Figure 2) the source code for producing and depicted the graph we present, so the reader can further explore the tool.

```
p1=plot3d (f, (x, 0, 400), (y, 0, 50), rgbcolor=(0,1,0))
```

```
p2=plot3d(g, (x,0,400), (y, 0, 50), rgbcolor=(1,0,0))
```

```
z=var('z')
```

```
p3=implicit_plot3d(z==40000, (x,0,400), (y, 0, 50), (z, 0, 40001))
```

```
p4=circle((200,40,f(200, 40)), 15, rgbcolor=(0,0,0), fill=true)
```

```
show(p1+p2+p3+p4)
```

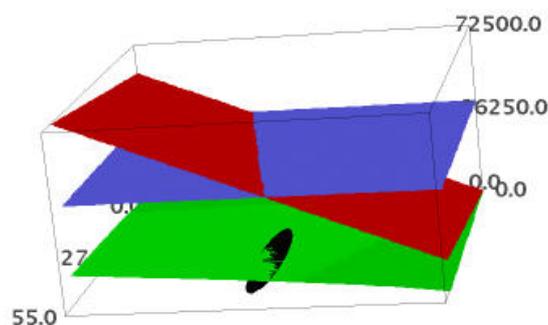


Figure 2: Graph interpreting the solution to the running example

Conclusion

We conducted a small experiment to illustrate the use of a tool, notably Sage, to aid in solving a maximization problem using Lagrange multipliers method. It happens that the problem comes from yet a third area of knowledge and that the mathematical solution can be interpreted using terms from that application domain: Economics. These interdisciplinary quests happen to be important for all the involved areas, since Economics benefits from the tools, both theoretical and practical, that have been researched over the years in Mathematics and Computer Science. On the other hand, Mathematics also benefits from this joint venture, since it gets inspiration in real world problems and then can try to come up with problem driven solutions. Computer Science benefits as well: first there is the opportunity for researching these computer aided systems that tackle all sorts of things automatically; moreover, there is the opportunity to disseminate software by different areas and get actual users performing complex tasks with the tools. Often, and unfortunately, the end users discover most of the bugs, performance issues, etc.

It is our conviction that this kind of tools should be used when teaching undergraduate and graduate courses, since mastering them constitutes a real benefit for expert end users.

References:

- [1] Murty, Katta G. *Linear Programming*. John Wiley & Sons, 2002.
- [2] Sage web site. <http://www.sagemath.org>.
- [3] Douglas, Paul, The Cobb-Douglas Production Function Once Again: Its History, Its Testing, and Some New Empirical Values. *Journal of Political Economy*, **84**(5):903-916, 1976.
- [4] Bertsekas, Dimitri P. *Nonlinear Programming (Second ed.)*. Athena Scientific, 1999.
- [5] Vapnyarskii, I. B. Lagrange multipliers, in Hazewinkel, Michiel, *Encyclopedia of Mathematics*, Springer, 2001.
- [6] Magma web site. <http://magma.maths.usyd.edu.au/magma/>.
- [7] Maple web site. <http://www.maplesoft.com>.
- [8] Mathematica web site. <http://www.wolfram.com/mathematica/>.
- [9] MatLab web site. <http://www.mathworks.com/products/matlab/>.

MATHEMATICS IN MOROCCAN UNIVERSITIES

Zineb El Jaoussi

Al Achhab Sakina

Aziz Arbai

Mohammed Khaldi

University Abdelmalek Essaadi, Morocco

Abstract:

This work deals with the teaching and learning of mathematics in Moroccan universities. Our goal is to better understand difficulties encountered by students, with math, and to study in which measure the integration of new technologies of information and communication may help them overcome these difficulties, by favoring interactions between representation by ICT and classical representation of the concerned course. On the basis of the instrumental approach of computers in teaching, we start from the hypothesis that succeeding to this type of mathematical tasks requires, beyond knowing adequate commands and their syntax, learning specific instrumented techniques, in a context taken in charge by the institution. In our thesis, a technique to represent called 3D technology occupies a central part, as a means to reduce the distance between the knowledge teachable and the knowledge to be taught.

The work includes an institutional analysis (ecological analysis of university curriculum) about the teaching of mathematics and the use of 3D technology, as well as two experiments, without and with the other.

Key Words: Mathematics, Information Technology and Communication ICT, 3D technology

Introduction

Our interest in this work is part of our experience as a student at the “University Abdelmalek Essaadi, Faculty of Tetouan” Morocco. Indeed, in our discussions with both teachers and students, we have made the empirical observation of many difficulties of teaching and learning mathematics. These difficulties affect first the students, who are struggling to capture mathematical concepts given to them and especially to visualize it in space. In addition, teachers are often helpless faced with these difficulties and obstacles of their students and do not see too how to act. The intention to solve these problems is the idea of integrating the new technologies of information and communication in education.

Here, for example, an excerpt from an interview we had with Mr. Benslimane, professor of mathematics at the Faculty of Tetouan, which pretty much sums up, in our view, the difficulties encountered:

“Even if the student knows the theory, how can he manipulate it to solve more complex problems. He needs to know how the base pass the abstract to the concrete. Much will depend on the viewing space of mathematical objects. This passage is very complicated, either the student has a visualization of the problem, or is very difficult to do. In general, the student should have this ability. But unfortunately it did not, because currently teaching mathematics does not follow a certain pedagogy. So there is this difficulty”

It seems that many points mentioned in this quotation require clarification or comments. For example, when the teacher talks about the transition from the abstract to the concrete, this phase of work is based primarily on a “view space”, he stresses that this is the point that raises the most difficulties for students. What is “viewing space”? Does it appeal only theoretical tools? If yes, which ones?

In fact, the teacher speaks of “visualization of the problem” and “vision in space.” The two expressions do they cover all the same (or more) didactic reality (s)? Which one? Putting themselves in the mathematical corpus commonly called Analytic Geometry of Space (AGS) is it the ability to

relate reality (the theory) and imagination? And more generally it is constantly able to make interpretations in the space of mathematical objects involved? This "ability" is it just a gift to "see" in space? Or correspond to knowledge and know-how to identify it is better to control it?

Answering these questions requires of course looking further into the study of mathematics degree program to better identify mathematics gaps involved and the roles of theory and interpretation in space.

At another point in the interview we had with Mr. Benslimane, he attributed that using the new technologies of information and communication can facilitated the transition from the abstract to the concrete in mathematics. This point deserves our attention. What ICT can bring to the teaching of mathematics, how can we integrate them effectively in this teaching, teaching in what way?

According to Y. Chevallard:

"An object (eg, a mathematical object) is an emerging system of practices which are handled material objects that are cut in different semiotic registers: register oral, spoken words or phrases, register sign language field script of, what is written or drawn (graphics, formalisms, calculation, etc..), that is to say, register writes. (1991, page 110)."

Our own teaching experience has led us to hypothesize that technological tools, including ICT, may be denied assistance in mathematics education, reducing the distance between the knowledge teachable and the knowledge to be taught.

In addition, several education institutions today introduced software to try to better manage mathematical objects that pose difficulties for students. For example, the Maple software, which helps calculate multiple integrals.

In the light of the first questions we already tightened our object of study by focusing on:

The transition from the abstract to the concrete in mathematics using new information technologies and communication. More specifically, the 3D technology.

This study motivated by the acknowledged difficulties of students and teachers is also justified by the need to assess the various initiatives implemented in recent years in Morocco and supported by the Ministry of National Education of Morocco.

Theoretical framework

We will build on approaches to the representation of mathematical objects. We will thus refer to the concept of registers of semiotic representation introduced by Duval in teaching (1993). This approach seems interesting because it will allow us to decontextualized and contextualized mathematical object. However, this approach alone is not sufficient to effectively study the issues that arise in the context of our research, since the institutional dimension is essential. We find, in this regard, support in the anthropological approach to teaching (TAD) developed by Chevallard (1999). Based on the use of ICT, our theoretical foundations will finally allow us to take into account the instrumental dimension of learning media environments. In this way, we build on the work of researchers in cognitive ergonomics, learning about the use of technological tools and in particular, the theory proposed by Verillon instrumentation & Rabardel (1995). We conclude with the presentation of the research questions we ask in the context of these frameworks and methodological and theoretical choices that we made.

The theory of instrumentation

To better understand the role of ICT in the teaching of mathematics, we use the theory of instrumentation. This theory, following the work in cognitive ergonomics concerns the use of learning technologies. We refer in particular to the theoretical framework developed by Rabardel (1995).

In recent years, many educational researchers have focused their attention on this theoretical perspective: Guin & Trouche (1999), Artigue (1997), Artigue (2002), Lagrange (2000), Guin & Trouche & Ruthven (2005) Haspekian (2005), BuenoRavel & Gueudet (2009), Drijvers (2000), Trouche (2000) ... etc.

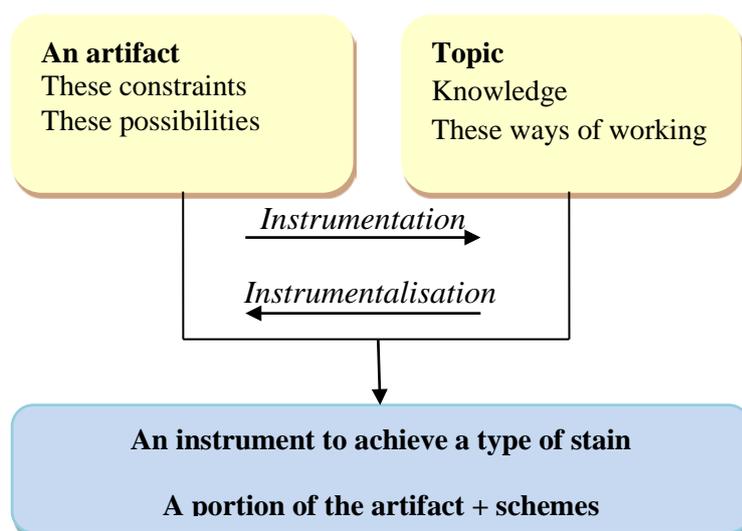
This emphasizes that:

"Recent works in cognitive ergonomics provide theoretical tools for understanding the process of ownership calculator's complex. Rabardel (1995), with regard to education in general and education in particular, offers a new approach, which essentially distinguishes technical tool (artifact), which is given to the subject, and instrument, which is constructed by the subject. The construction or instrumental genesis is a complex process due to the characteristics of the artifact (its

potential and its limitations), and the activity of the subject, knowledge and work habits earlier (Op. cit. 195).”

The instrument is not "given", it was built by the subject during a process of instrumental genesis. In this sense, Drijvers emphasizes that:

“The starting point of the theory is the idea that instrumentation tool is not automatically an effective and practical. A hammer, for example, an object is meaningless, except that when you have something to pound, it is transformed into a useful tool. This idea also applies to other objects, or computer software. The learning process in which an artifact is gradually becoming an instrument is called instrumental genesis. The user must develop skills to first recognize the tasks for which the instrument is suitable for and then perform in the environment of the instrument (Op. cit. 218).”



Instrumental genesis (Trouche, 2004)

The integration of technology in mathematical activity led to the construction of patterns of use, more responsive and less and less effective. According Rabardel schemata are multifunctional. Brought into play in specific situations, they help to:

- Understanding them (their epistemic function)
- Acting, transform, solve (this is their pragmatic function)
- Organize and control the action (this is their heuristic function)

It will be important for our work to characterize specific practices in mathematics education, and the use of 3D technology in the institution (University Abdelmalek Essaadi, Faculty of Sciences). We find support in the anthropological approach to teaching developed by Chevallard (1992) that we present below.

“The anthropological approach helps us think about the technical and instrumental dimension of mathematical work, which in didactic analyzes is often left in the background in favor of more conceptual analysis. (Ibid., p. 9).”

The anthropological didactic approach

The approach developed by Chevallard (1992) is an extension of the theory of didactic transposition. It considers mathematical objects, not as existing in itself, but as entities that emerge from practices in systems and institutions. These systems are praxeologies or described in terms of tasks in which the object is invested technologies can solve them, and through which discourse is used to explain and justify the technology. These can be viewed from the instrumental point of view (as explained Lagrange (2000, p. 169)).

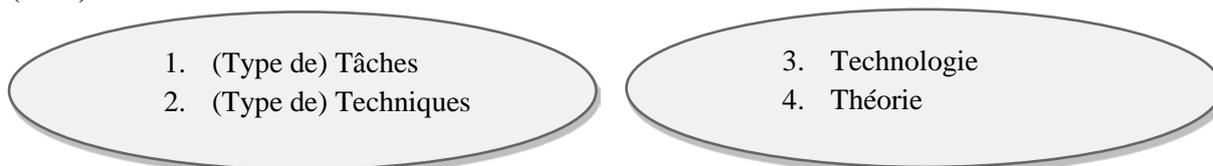
According Chevallard, science education, like all teaching, is in the field of social anthropology, that is to say, the study of man. Similarly there exists a religious anthropology or political anthropology, whose objects of study are respectively the religious or political Chevallard (1992) proposes to develop a didactic anthropology whose object of study is the learning to study, for example, the student with the problem mathematically.

As emphasized Chevallard mathematical knowledge as a special form of knowledge is the result of human action institution: it is something that occurs, is used, taught, or more generally, transposes institutions.

Chevallard proposes the notion of organization or praxeological praxeology (as the key concept) to examine institutional practices relating to an object of knowledge and in particular social practices in mathematics. He proposes to distinguish praxeologies that can be built in a class where we study this subject, analyzing how can build the study of this object, and allow the description and study conditions achievement. The praxeologies are described in terms of:

The notion of semiotic register of representation

In mathematics, the objects are only accessible through their representations. By Duval (1993):



“There is a paradox cognitive mathematical thinking: on the one hand, the understanding of mathematical objects can only be a conceptual understanding and, secondly, it is only by means of a semiotic representations activity of mathematical objects is possible. This paradox can be a real learning circle. (Op. cit. 38).”

Semiotic representation is a representation constructed from the mobilization of a system of signs. Its meaning is determined partly by the form in the semiotic system, on the other hand, the reference to the object represented. Geometric figures, a statement in language, an algebraic formula, a graph, are semiotic representations that are different semiotic systems. Treatments mathematical objects depend therefore opportunities representations themselves.

Duval (1995) explains the concept of semiotic register of representation as follows:

“Semiotic systems are capable of performing the three cognitive activities inherent in any representation. First, create a trace or sets perceptible traces that are identifiable as a representation of something in a given system. Then transform representations only by rules specific to the system in order to obtain other representations may constitute a relationship of knowledge in relation to the initial representations. Finally, convert the representations produced in a system of representation of another system, so that they allow explaining other meanings for what is represented. All semiotic systems do not allow these three fundamental cognitive activities ... But natural languages, symbolic languages, graphs, geometric figures, etc. The permit. We speak then register semiotic representation. (Op. cit. 20).”¹

Duval thus distinguishes three cognitive activities related registers of semiotic representation:

□ *The formation of a semiotic representation - using a (several) sign (s) to update or replace the target object. It is based on the application of compliance rules ¹ and the selection of a number of characters from the content preview. For example: composition of a text, drawing a geometric figure, drawing a diagram, writes a formula.*

□ *Treatment of representation - is the transformation of the representation in the registry even when it was formed. Treatment is an internal transformation to a register. For example, the calculation is a form of symbolic writing specific processing (numerical, algebraic calculus, integral calculus ...).*

□ *The conversion of representation - is the transformation of this representation to a representation of another register. For example: Translation is the conversion of a linguistic representation in a given language in a linguistic representation of another language.*

¹ Rules to be observed in the formation of a semiotic representation, such as grammar for natural language training rules in a formal system design constraints for the figures ... The formation of these rules is to ensure, first, the conditions of identification and recognition of the representation and, second, the possibility of their use for treatment. (Duval, 1993, p. 41).

The translation should not be confused with two activities, however, are nearby coding and interpretation. Interpretation requires a change of framework or context. This change does not involve a registry change but often mobilizes analogies.

"Coding" is the "transcription" of a representation in another semiotic system that it is given.

The conversion is therefore of particular importance. However, it is generally neglected in the teaching of mathematics, while, as noted Duval one of the essential conditions for the conceptual understanding of mathematical objects is available for the same purpose, several semiotic representations. The choice of an appropriate register of representation can facilitate processing (transformations of representations within a register).

Research Questions

As we already mentioned, our research focuses on the teaching and learning of mathematics in higher education and in particular to reduce the cognitive effort of mathematics. We will study the "object of knowledge", the tasks proposed to students, available technologies to solve their justifications and technological-theoretical depend on a set of data relating to the institution in which teaching takes place. In addition, analysis of the ecology of a mathematical object in an institution understands its meaning for this institution is to identify the organization mathematical object is in. Thus, we propose to study, in the context the TAD, the mathematical organization. This study should allow us to answer the first research question:

Q1. What is the mathematical organization in Moroccan universities?

Praxeological concepts of organization and reporting institutional offer, from an ecological study of the academic program, the tools to find answers to such questions. This question must be supplemented by other, more in relation to ICT (and more specifically 3D technology). Indeed, the academic program introduces the use of ICT in teaching and contemporary in particular 3D technology. As Lagrange points (2000, p.41)

"For techniques [instrumented] are meaningful, it should build praxeologies in which these techniques could be inserted and take a mathematical meaning."

Take into account the role of ICTs, including 3D technology in the university raises the following questions:

Q2. What is the academic report instrument to 3D technology at this university?

Q3. What constraints they impose the university, what conditions do they provide to the teaching of mathematics with 3D?

Analysis of 3D technology will allow us to better understand how this tool can reduce the cognitive effort of mathematics. We will try to press the three frames with the ostensive aspects and non-ostensive mathematical objects, their semiotics, and the instrumental aspects around these objects play important roles. This study will allow us also to identify potentialities, constraints on the actions and changes desired by the user 3D mathematical objects. To the analysis of these constraints, we will take into account the type considered by Trouche (2000). The following questions arise in the context of cognitive instruments:

Q4. What tools are available in the environment publicized 3D technology on: analytic geometry in space (GAE)²?

Q5. How through the use of these tools are changed techniques solving certain tasks?

Conclusion

Our work is the first investigation in a virgin field of education research. Level of education, very specific mathematical content and the use of 3D technology are all factors that have asked us to be innovative. It is a first draft that can be used especially for teachers Moroccans. We hope to have shown to accurately and varied work needed on the development of technology to reduce the distance between scholarly knowledge and expertise teachable. Our experiments show the potential of this technology. it will also show that you cannot just rely on this technology and process instrumentation must include the issue of transition from the abstract to the concrete in a more global manner by using certainly more explicit knowledge students.

² Both on the analytical representation and graphics or by functions of two variables, either by parametric equations.

References:

- Chevallard Y. (1999), Analysis of teaching practices in anthropological theory of teaching. *Research in Mathematics Education*, V. 19/2, p. 221-266.
- Chevallard, Y. (1992), Fundamentals of didactics: perspectives made by an anthropological approach. *Research in Mathematics Education*, V. 12, No. 1, p. 73-112.
- Chevallard, Y. (1994), The process of didactic transposition and theorizing In G. Arsac et al. *Didactic transposition to the test* (p. 135-180). The Wild Pansy. Grenoble.
- R. Duval (1993), registers of semiotic representation and cognitive functioning thought. *Annals of learning and cognitive science*. IREM Strasbourg v. 5, p. 35-65.
- R. Duval (1995), *Semiosis and human thought*, Bern: Peter Lang.
- Rabardel P. (1995), *Men and technologies - Cognitive Approach contemporary instruments*, Editions Armand Colin.
- Trouche L. (2005), Construction and driving tools in learning mathematical necessity orchestrations. *Research in Mathematics Education*, V. 25/1, p. 91-138.

CONVERTING A DESERT BUGGY INTO SOLAR ELECTRIC VEHICLE

Ahmed H. Osman, Assist. Prof., PhD

Zagazig University, Egypt

Magdy. M. Massoud, Assist. Prof., PhD

Prince Sultan University, KSA

Abstract:

The renewable energy became an essential demand for today's world life for all its importance in decreasing pollution and keeping our environment clean. The solar car is a step in saving our environment and energy resources. New project (supported from King Abdulaziz & his Companions Foundation for the Gifted and Technical & Vocational Training Corporation, KSA) is targeted to build a solar electric vehicle. The project was designed for talented students in the field of Vehicles & Engine Technology in Saudi Arabia. We initiated a program for one month duration, which includes different activities in designing, soldering, suspension, steering, electric motors, solar cells and batteries. The aim of this program is to prepare the students to cooperate as a team in achieving the production of the solar electric vehicle.

Starting from a very cheap desert buggy, the new solar vehicle is assembled in innovative manner with used material and component. The solar vehicle is assembled in the workshop of Riyadh College of Technology. The solar vehicle is successfully assembled, operated, tested and ran as fast as 50 km/hr for 3 hours with fully charged batteries. The solar panels will continue charging the batteries. On the other hand, the batteries can also be charged using regular 220 AC Volt. The solar cells extends the running distance by about 40%

Key Words: Renewable Energy, Solar Electric Vehicle

Introduction

As petrol price is rising day after day to a distinguished limit, acquiring a fuel vehicle become costly for everyone, meanwhile the vehicles using gasoline fuel is contributing by a high percentage in environment pollution and a bad effect on the global warming. Most of vehicles' manufacturers are doing their best for producing vehicles which can use different renewable energy sources. In this direction, we have done a contribution by achieving a solar electric vehicle in which we combined solar cells as auxiliary electric source with direct AC source as the main electric source. The AC electric source is used for charging the battery set using charger. The solar panels are covering the vehicle body to help in charging the batteries and to extend the running distance traveled by the vehicle.

Photovoltaic panels: efficiency and cost

The conversion process from light into DC electricity is based on the Bell Laboratories' researches in the 50's, where the principle discovered by Alexander-Edmond Becquerel (1820-1891) was applied for the first time (Gianfranco, 2010). The diffusion of using semiconductor technology has been growing exponentially in recent years (figure 1 and figure 2), due to the pressing need for renewable and carbon free energy (REN21, 2009).

The amount of solar energy is impressive: the 89 peta Watt of sunlight reaching the Earth's surface is almost 6,000 times more than the 15 tera Watt of average electrical power consumed by humans (Smil, 2006).

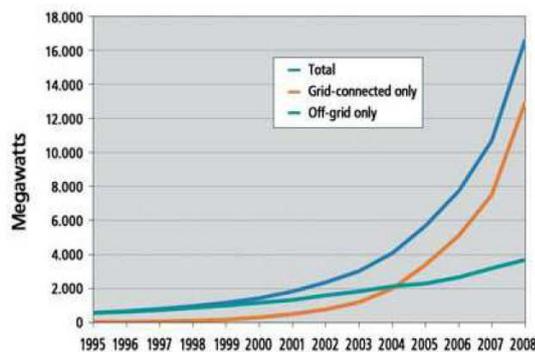
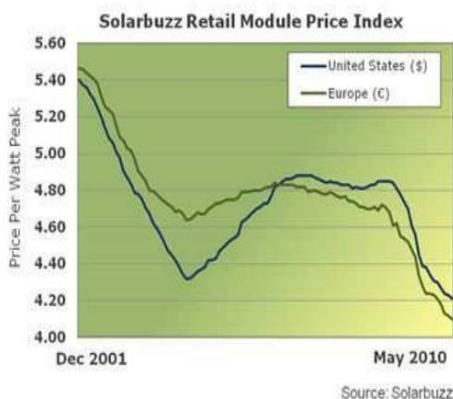


Fig. 1: Trends for cost of photovoltaic modules **Fig. 2:** Solar PV, world capacity 1995-2008

Solar Electric Vehicle Project

The program started with selection of 40 students out of 400 and then the program schedule was distributed among students. The executive team started to prepare for the vehicle whole scenario from start to end. The process started by buying a desert buggy. The problem that faced us was how to transform this very heavy machine to a car that can run by solar and direct electric power. A road map for work schedule is designed according to the availability of materials and according to the following key parts when designing a solar electric vehicle

1. Chassis, suspension and steering: how to build the frame of the car
2. Electrical System: how the solar panels will be assembled, batteries and motor work
3. Wheels, Axles and brakes: how to make wheels turn and how can you stop it
4. Transmission: how to transfer power from the motor to the wheels
5. Body Shell (profile): how the shell effects car performance

Design

The first step on the design was to get rid of the heavy wheels, and draw a complete design putting into consideration, the weight of the vehicle, the batteries location, the electric motor, the control unit, the charger, the electric wires, the solar panels and finally the final profile as illustrated in figure 3 and figure 4.

Chassis

The chassis constitutes a major problem for all its heavy metal construction, its narrow front and its small availability space to place the batteries, charger and controller. For this reason maximum cut of metal construction has been done keeping in mind that the new modification will not affect the safety of the final construction as shown in figure 5.

Suspension

Small modifications were done on the existing suspension as shown in figure 5.

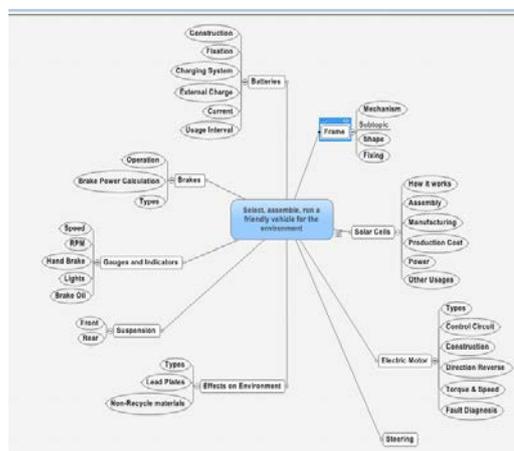
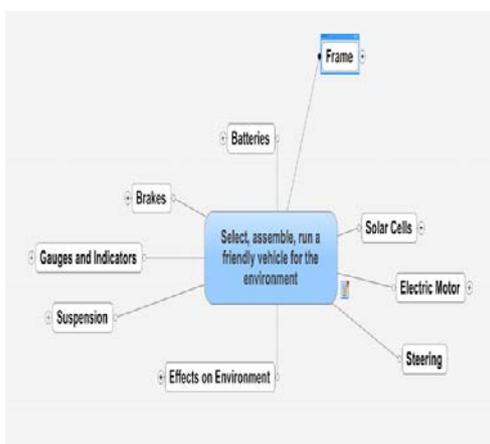


Fig. 3: Design plan**Fig. 4:** Detailed design plan**Fig. 5:** Chassis, front and rear suspension

Steering

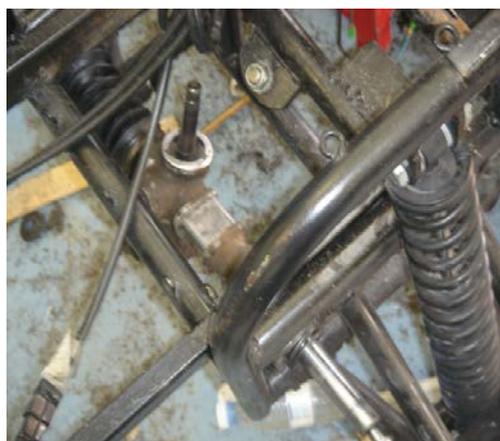
Rack and pinion were used instead of the existing steering to allow more steering control as shown in figures 6 and figure 7.

Electrical System

As shown in figure 8, the electrical system includes the following parts: Electric motor, batteries, Charger, Control unit, electric cables, and electric switches for start and stop, accelerator pedal connection and the solar panels.

Electric Motor

A DC electric motor of 3.6 kW and 48 volt is used. The power delivered to the load from the electric motor was sufficient to drive the vehicle with a reasonable traction and speed as shown in figure 9.

**Fig. 6:** Steering**Fig. 7:** Rack and pinion

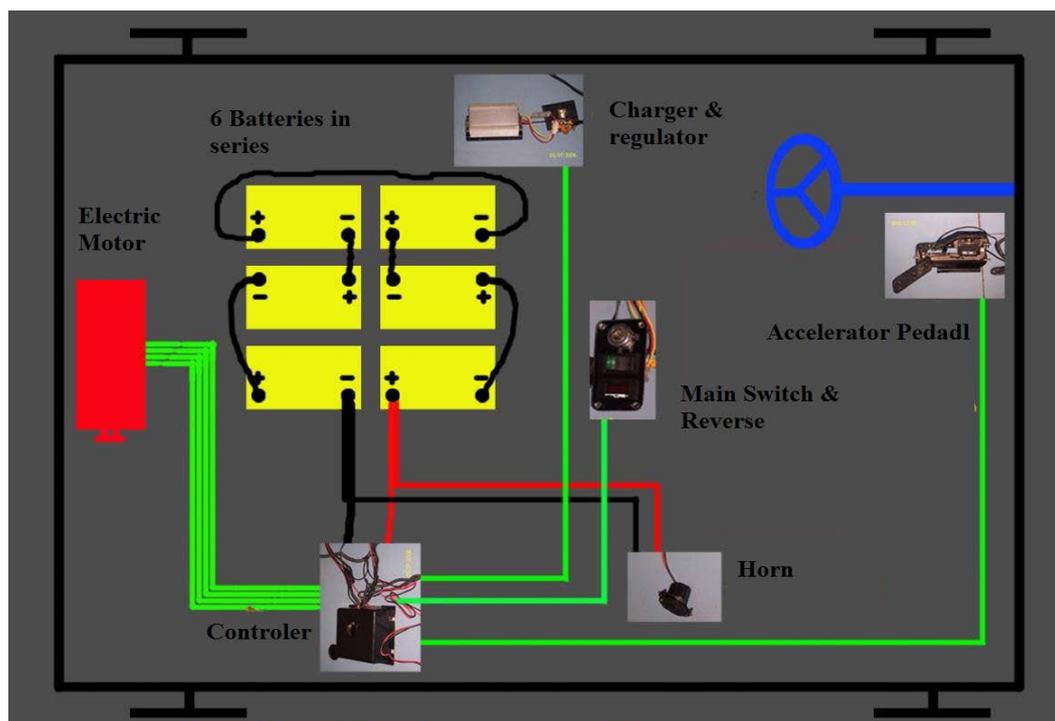


Fig. 8: Electric circuit connection system for the vehicle



Fig. 9: Electric motor installation and aligning Fig. 10: Batteries connection and installation

Batteries

To drive the motor of 48 Volt it needed 6 lead acid batteries with 8 volt each with capacity 180Ahr. The size of each battery and the weight was reasonable as shown in figure 10. When being completely empty, the batteries need 7 hours to be fully charged.

Control unit

A solid state control unit is used to control the electric motor speed & direction as shown in figure 11. Reversing and front driving will happen through the control unit as most DC gear motors are normally very easy to reverse; simply changing the polarity of the DC input will reverse the direction of the drive shaft. This changeover process can be achieved via a simple changeover switch.



Fig. 11: Control Unit



Fig. 12: Charger unit

Electric Charging

To recharge the batteries, a portable electric charger is used as shown in figure12. The charger specifications are as follows:

Input: AC230V 50/60 Hz 1.8KVA 8.2A; Output: DC 48V/25A; Characteristic: WSAVDE 0510 T1-3058T1-1; Protection mode: IP21.c1.1; Load: Battery working; Battery: 24 pb element/cells; Capacity: 150-210 Ah

The regulator is shown in figure 13



Fig. 13: Regulator unit

Electric switches, front, rear lights and horn

Figure 14 shows a set of electric switches that used to start, stop, reverse direction, lighting, honring, etc. Others were used to accelerate and slowing the speed. Figure 15 shows the real design of the electrical circuit.



Fig. 14: A set of cables and wires

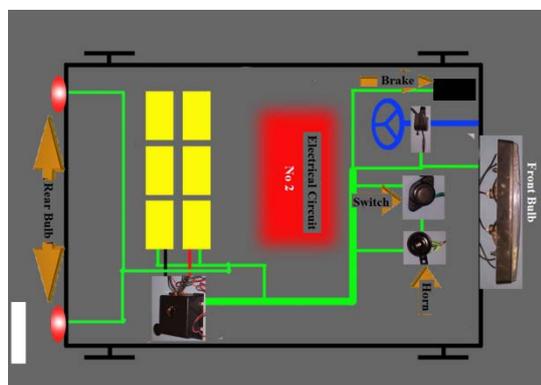


Fig. (15) Brake, front, rear bulbs, and horn wire connection

Brake system

A disk brake type is mounted on the rear axle as shown in figure 16. A hand brake beside the driver's chair is installed as shown in figure 17.



Fig. 16: Brake position in the vehicle



Fig. 17: Hand brake installation

Photovoltaic Panels

A Photovoltaic (PV) panel is an alternative name for solar panel. It consists of so many PV cells wired in parallel and in series. The parallel arrangement is utilized in increasing the current whereas the series arrangement is utilized in producing higher voltage. The front surface module is covered with tempered glass. The back surface is covered with a protective and waterproof material. An aluminum frame holds everything together in one unit. In the back of the module there are wire leads, to provide electrical connections.

Twelve solar panels each with 40 W were used as shown in figure 18 and 19. The solar panels covered the front, ceiling and the back of the vehicle. The solar panels alone were not sufficient to charge the batteries but it helped in decreasing the charging period, in addition it extended the running distance by about 40% (during sunlight period).

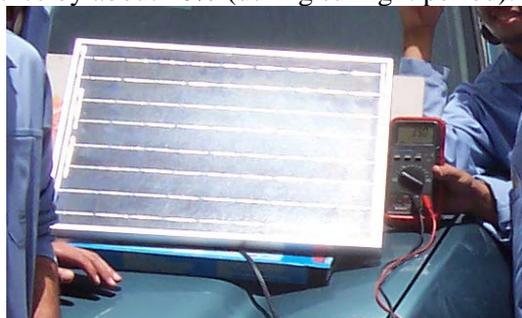


Fig. 18: Solar panel test



Fig. 19: Final Product

Profile and vehicle design specification

The profile of the vehicle is used as a support for the solar plates. After installation of the electric components, the batteries and the solar panels, aluminum sheets are cut and shaped according to the chassis and then it is assembled to the car as shown in figure 19.



Fig. 20 Team Work

Vehicle Specifications

- Length: 3090 mm
- Width: 182 mm
- Front Width: 1200 mm
- Rear Width: 1240 mm
- Base wheel Width: 970 mm
- Electrical motor: DC motor wound shunt, 3.6 KW, 48 V
- Control Unit: URTIC PMC MODEL 1244 MULTI MODE MOTOR CONTROLLER (48V)
- Batteries: 6x8V
- Maximum Speed 50 km/hr.
- Period: 3 hrs.

Conclusion

It's important to shift our energy usage into renewable and clean one. This conversion has significant positive economic and environmental effects. In this paper, the utilization of solar power in vehicle application was implemented. The selection of the appropriate components for the application was studied and assembled together to complete the system. The profile of the vehicle is designed such that the solar modules can be fitted on its surface.

According to the application (converting beach buggy into solar vehicle), the method of the power transmission to the rear axle must be changed. A rear axle with differential is used so that we can increase the speed. The assembled vehicle runs as fast as 50 km/hr for about 3 continuous hours. The solar panels extend this period by another 40%.

With further modification of the profile design and the method of fitting the solar module, we can increase our power generation from the sun.

References:

- [1] Gianfranco Rizzo, Ivan Arsie and Marco Sorrentino (2010): Hybrid Solar Vehicles, Solar Collectors and Panels, Theory and Applications. Dr. Reccab Manyala (Ed.), ISBN: 978-953-307-142-8,
- [2] REN21: Renewables - Global Status Report. 2009 Update,
- [3] Smil V.: Energy at the Crossroads, Global Science Forum Conference on Scientific Challenges for Energy Research. Paris, May 17-18, 2006,
- [4] James Larminie: Electric Vehicle Technology Explained. Oxford Brookes University, Oxford, UK and John Lowry Acenti Designs Ltd., UK John Wiley & Sons Inc, 2003.
- [5] Edward J. Simburger, Joseph T. Simburger: Electrical Engineer, Residential Photovoltaic, An Investment Vehicle for Retiree's. 4th World Conference on Photovoltaic Energy Conversion, May 8-12, 2006, Waikoloa, Hawaii.

THE PRINCIPLE OF RECORDING INFORMATION IN DISTRIBUTED ENVIRONMENTS VIA SULEIMENOV-MUN'S WAVES

Dolayev M., MS

Obukhova P., MS

Almaty University of Power Engineering and Telecommunications, Almaty, Kazakhstan

Panchenko S., MS

Al-Farabi Kazakh National University, Physico-Technical Faculty, Almaty, Kazakhstan

Bakytbekov R.B., MS

Al-Farabi Kazakh National University, Faculty of Chemistry and Chemical Technology, Almaty, Kazakhstan

Ivlyev R.S., MS

Kazakh-British Technical University, Almaty, Kazakhstan

Chezhimbayeva K., PhD

Almaty University of Power Engineering and Telecommunications, Almaty, Kazakhstan

Abstract:

A new approach to information recording in distributed media based on the forced wave excitation involved by temperature fluctuations was proposed. The scheme of the device providing the implementation of the proposed approach is considered.

Key Words: Waves, nonequilibrium systems, phase transition, stimuli-sensitive polymer

Introduction

The problem of recording of information in nanostructures remains one of the important issues for nanotechnology in general [1]. To solve this, a variety of approaches are offered, which are reviewed in [2]. Unfortunately, none of them has been able to find a practical implementation on an economically feasible level. This makes it relevant to the development of the simplified scheme for recording information in nanostructures.

In general theoretical positions, in order to work with nanowildscale elements, it is appropriate to use different kinds of self-organization processes. In particular, by focusing waves of appropriate scale to a certain point in space, you can realize local changes in environmental conditions. Such issue makes it relevant to the solution of recording information problem in distributed media of different nature, based on the use of waves, spontaneously arising in this media.

This paper describes the use of a recently discovered type of waves [3] (Suleimenov-Mun waves [4,5]) for the purpose of recording information in distributed environments. The Suleimenov-Mun waves can be defined as fluctuations, developing in a nanoequilibrium medium near the phase transition. As shown in [3], such waves appear due to the time delay between the moment of reaching a critical temperature (or another critical parameter) and the actual moment of the phase transition. Due to the fact that the phase transition in solutions of stimuli-responsive polymers can be initiated by various external influences, there may exist various types of the considered waves.

Experimental

The example of experimental results [6], confirming the theoretical predictions in[3], are shown in Figure 1. 2% solution of copolymer NIPAA:AA 90:10 at pH 6-7 in the temperature range from 26 to 40 °C was investigated. Selected polymer has a phase transition temperature of about 32 °C, at which there is an abrupt turbidity. This allows to measure the content of macromolecules, passed into the partially soluble form by turbidimetry (measuring the intensity of light passing through the solution). In particular, it is possible to investigate the variation in the number of these macromolecules by measuring the optical density fluctuations.

The relative intensity of the light transmitted through the solution vs. time was recorded, the time resolution of the equipment was 1 ms, the unit assumed the value corresponding to the intensity of the light transmitted through the cuvette at the initial time.

The solution was placed in a rectangular cuvette with the optically transparent walls cooled by external water jacket. The wire element located inside the cuvette provides heating of the solution to the desired temperature. Figure 1 shows that the state of the thermosensitive polymer solution near the phase transition is characterized by fluctuations in the optical density, which have a regular (periodic) character.

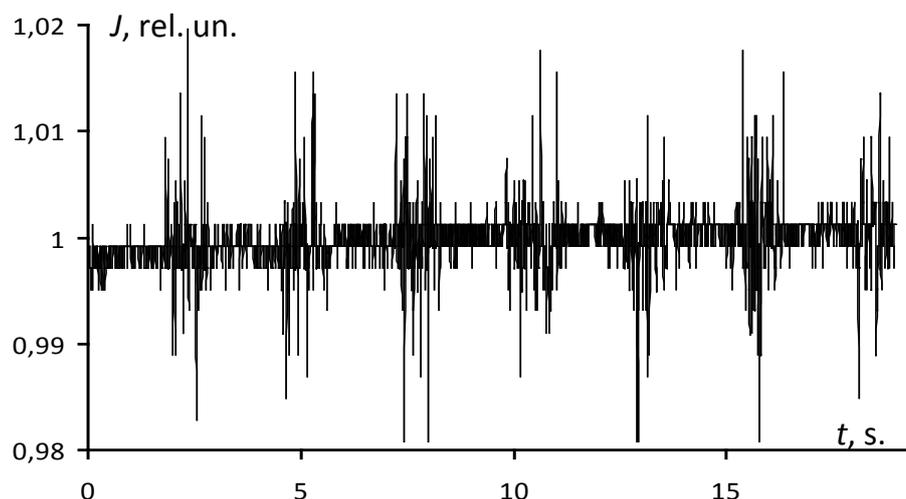


Figure 1 – Dependence of the relative radiation intensity passing through the cuvette vs. time; $T = 38\text{ }^{\circ}\text{C}$ [6].

Results and discussion

Qualitatively, the appearance of these fluctuations can be interpreted, starting from the assumption concerning the final rate of the phase transition. This assumption, in particular, means that there is a time delay between the moment of heating the local area within the polymer solution to the temperature T_{ph} and the time when molecules in this field changes its state.

This assumption seems more than justified, since the phase transition is accompanied by a change in size of the macromolecular coil. At temperatures below the T_{ph} the hydrophilic-hydrophobic balance is shifted towards hydrophilic interactions. It corresponds to the expressed interaction of macromolecule links with molecules of solvent. This, in particular, means that the macromolecular coils are in the swollen state in these conditions.

On the contrary, the hydrophilic-hydrophobic balance is shifted to the direction of strengthening the hydrophobic interactions by increasing the temperature T_{ph} . As a result, macromolecules partially lose their solubility and coils formed by them are compact. The transition from one state to another, at least requires sufficient time for diffusion (or directed movement) of macromolecule segments originally located on the periphery of the coil to its core. This shows a non-zero amount of time elapsed during the phase transition.

Similar conclusions can be obtained by examining the process of displacement of water from the coil when it is compressed. In this case, the parameter τ can be estimated as the time required for the outflow of the amount of water through the substance of the polymer.

We show that this assumption is sufficient to explain the occurrence of the predicted fluctuations in type observed in the work [3].

Let suppose that initially the temperature of solution was below T_{ph} . Assume that the solution is heated by some source of heat (in the simplest case, this may be a resistor immersed in solution). After turning on the heat source the solution will be heated and at a certain point in time the temperature rises to T_{ph} . However, due to the finite speed of such phase transition, as it will be demonstrated later, the solution is heated to a temperature above T_{ph} .

Further, the phase transition will cause some cooling of the solution, due to energy loss associated with overcoming the corresponding potential barrier. Provided that power level of the heat

source is close to the threshold value the cycle will be repeated again. The oscillation amplitude of a purely thermal character is small, but it can be significantly increased by directly heating the solution with electric current. It was theoretically described in [3].

Application

This type of waves can be used to record information in the media, where the phase transition occurs. The scheme realizing this approach is demonstrated in Figure 2.

The two-dimensional wave is used to record information, focusing through at some point. The amplitude exceeds the critical value in the area of focus, corresponding to stable phase transition while in the remaining sections of the medium amplitude has a value below the critical threshold.

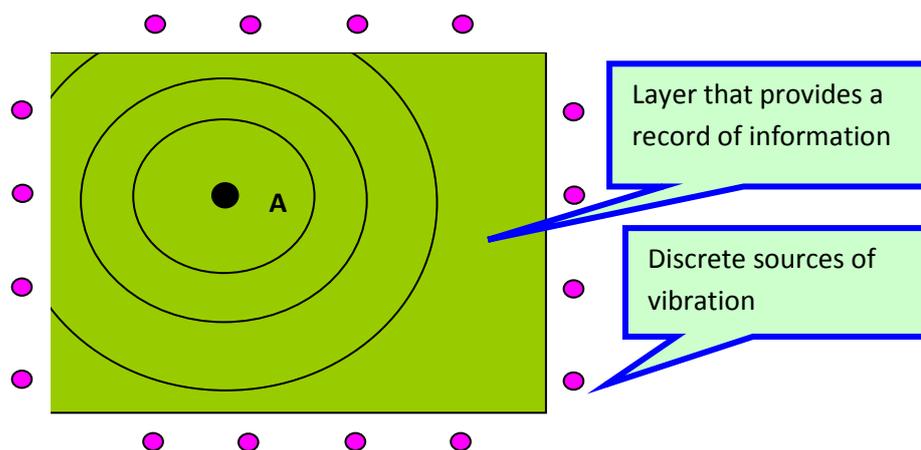


Figure 2. Scheme of a convergent cylindrical wave realization for recording information in a distributed environment.

Conclusion

The waves in Figure 2 are excited by the discrete sources of vibrations. For Suleimenov-Mun thermal waves conventional heating elements can be used. The synthesis of a required configuration of the wave is provided by the selection of phase delays between sinusoidal voltages applied to the individual elements. The record of information is implemented by irreversible changes in the environment, that are caused by the temperature (or other control parameter) exceeding a critical threshold.

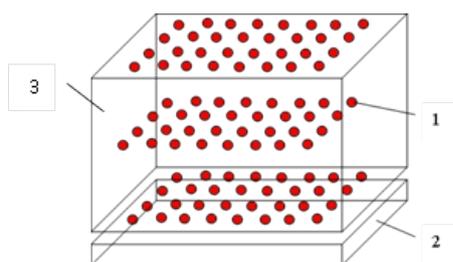


Figure 3. Using a grid made up of heating elements for the stabilization wave focusing, 1 – heating elements, 2 – medium, in which information is recorded, 3 – cuvette with a solution.

However, waves of this type have sufficiently complicated dispersion law [3,6]. Therefore, it is appropriate to use a three-dimension structure, shown schematically in Figure 3, for their excitation. The periodic arrangements of heating elements in this case allow for a stabilization of the waves used by a spatial period and, thus, to eliminate errors associated with instability of the dispersion of Suleimenov-Mun waves. The environment, in which information is recorded, is located directly below the grid formed from the heating elements. This medium can be made of, for instance, a temperature-sensitive hydrogel, which has the phase transition temperature of the solution used.

In this paper, a model structure was implemented, based on three layers each of which is a diode-resistive matrix (Figure 4).

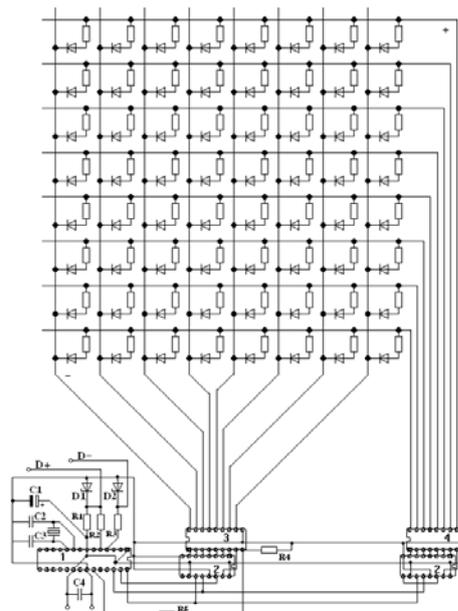


Figure 4. The scheme of the diode-resistor matrix for the stabilization system of a Suleimenov-Mun wave excitation

The use of SMD component provides a type of medium heating that can be described as “dotted”. The diodes are used to prevent the appearance of parasitic reverse currents. The control circuit is a controller on the chip ATMEGA8-16P.

Conclusion

The test experiments have shown that this type of a grid really provides a resonant excitation of waves of this type, and can stabilize the spatial period of the wave, whose amplitude reaches 20 – 30% based on a modulation of the optical density of medium. Thus, the initial experiments conducted in this study show that there is a possibility of focusing waves spontaneously arising in a liquid-phase systems.

References:

- Suleimenov I.E., Mun G.A. Osobennosti vozniknoveniya avtokolebatelnyh processov v zhidkofaznyh sistemah v oblasti fazavogo perehoda// *Izvestiya nauchno-tehnicheskogo obshestva “KAHAK”*// The features of self-oscillation processes in liquid-phase systems in the transition region// *Proceedings of the scientific and Technical Society “KAHAK”*, 2012, № 3 (37), p.109 [3].
- Suleimenov I., Mun G., Ivlev R., Panchenko S., Kaldybekov D. Autooscillations in Thermo-responsive Polymer Solutions as the Basis for a New Type of Sensor Panels, *Proc. 2012 AASRI Conference on Modeling, Identification and Control*, December, 10-20, Hong Kong, accepted [6].
- Yergozhin Ye.Ye., Aryn Ye.M., Suleimenov I.E., Mun G.A., Belenko N.M., Gabrielyan O.A., Park N.T., Negim El-S. M. El-Ash., Suleymenova K.I. *Nanotechnology versus the global crisis* / Seoul, Hollym Corporation Publishers, 2010, 300 p. [1].
- Yergozhin E.E., Zezin A.B., Suleimenov I.E., Mun G.A. *Gidrofilnye polimery v nanotehnologii i nanoelektronike (monografiya)* (In Russian language) // *Hydrophilic polymers in nanotechnology and nanoelectronics (monograph)*// *Library of nanotechnology*, Almaty-Moscow: LEM, 2008, p. 208 [2].
- Yermukhambetova B.B., Semenyakin N.V., Abdykalykova R.A., Timofeyev A.V., Dergunov M.A., Kaldybekov D.K., *Use of the Suleymenov-Mun Waves in PowerEngineering* // *8th International Conference on Nanosciences & Nanotechnologies* ", 7 - 15 July 2011 Thessaloniki, Greece. [4].
- Yeligbayeva G., Shaikhutdinova A., Irmukhambetova G., Bublik V., Shaikhutdinov E. *Excitation of Suleimenov – Mun’s waves at polymer composite surface* // *7th IUPAC International Conference on Novel materials and their Synthesis (NMS-VII) & 21st International Symposium on Fine Chemistry and Functional Polymers (FCFP-XXI)*, Shanghai, 16-21 October, 2011 [5].

POLARITY EFFECT ON CORONA DISCHARGE VOLTAGES IN SF₆-N₂ GAS MIXTURES

M. Assia Guerroui

Dr. Ahcene Lemzadmi

LGEG, Université 08 Mai 1945 Guelma, Algeria

Abstract:

This work is to investigate the behaviour of SF₆-N₂ mixtures that can be considered as a potential substitute of pure SF₆. The inception voltages have been determined from the measurements of the current-voltage characteristics for both polarities under high pressures and with highly inhomogeneous fields. The results show that the onset voltages increase with the increase of the gas pressure and the values of positive polarity are higher to those obtained with negative polarity.

Key Words: Inception voltages, corona discharge, sulphur-hexafluoride

Introduction

Sulphur-hexafluoride SF₆ has found a wide range of applications due to its superior insulating properties and chemical stability. As the size of the high voltage equipments increases the cost of insulating gas becomes appreciably high. The quantity of SF₆ released in atmosphere is therefore rising. Due to its ability to absorb and reemit IR make it a potent greenhouse gas [1-2]. With its high lifetime (more than 2000 years), it accumulates in the atmosphere and contribute to the global warming of the atmosphere, therefore SF₆ will be severely controlled in the next years. The by-products issued from decomposition of SF₆ exposed to electric discharges may be dangerous to the equipments (corrosion) and to the personnel (poisoning) [3]. These factors have stimulated the research of a replacement gas with little environmental impact. The most promising alternative is the use of gas mixtures of SF₆ with inexpensive common gases such as nitrogen (N₂). SF₆-N₂ gas mixtures have good dielectric strength, are non-toxic, non-flammable and they have a high arc quenching capacity with a good self healing ability. The purpose of this paper is to provide and discuss the measurements of the onset corona discharge voltages in SF₆-N₂ gas mixtures at higher pressure ranging from 3 to 15 bars and with different percentage of SF₆. The onset voltages were determined from the measurements of the current-voltage curves in both negative and positive polarities. A tip-plane configuration was used with the tip radius of few micro-meters and the gap between the electrodes is lower than 10 mm. Most of the data available deals with the breakdown voltages rather than the threshold voltage of corona discharge.

Experimental set-up

Experiments were made in a stainless-steel cell of 50 cc equipped with two quartz windows as shown in figure 1. Electrodes in a tip-to-plane configuration were mounted inside the cell. The tip electrode of few micrometers is made of tungsten and, is prepared by electrolyse technique, steel tips are also used. The gap between the electrodes varies from 5 to nearly 10 mm. The tip electrode is connected to the high D.C. voltage up to 60 kV. The stainless steel plane electrode with a radius of 12 mm is connected to a galvanometer which measures currents down to some microamperes. Before the cell was filled with the gas, pumping was undertaken pushing the vacuum down to nearly 3.10⁻² Pa. The gas was introduced in the cell without prior purification. The SF₆ used in these experiments is delivered with a purity of 99.97%. The measurements were made for pressures ranging from 3 bars to 15 bars. A partial pressure method is used to mix the gases in the cell. SF₆ was first introduced followed by nitrogen. It is supposed that SF₆ and N₂ do mix quite well. The current was measured when the voltage varies upwards and downwards and after each set of measurements a resting time is observed to allow the mixture to settle down. The tip electrode is regularly changed in order to limit

the radius variation due to the deposit of fluorine and sulphur. Analysis of used tips by electronic microscopy has been carried out.

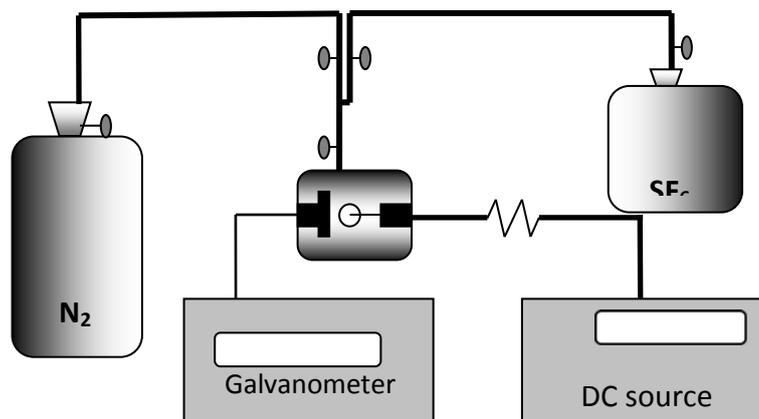


Figure 1, experimental set-up.

Results and discussion

The current-voltage characteristics $I = f(V_s)$ of corona discharge in SF_6-N_2 gas mixture show that the current follows an exponential law. The measured values of corona inception voltages (V_s) as a function of the ratio of SF_6 in the mixture are shown in figures 2 and 3 for negative and positive polarities respectively. The curves have the same tendency for both polarities and the V_s values increase with the increase of the amounts of SF_6 . This behaviour may be attributed to the decrease of the net ionisation coefficient $\bar{\alpha}$ [4-5].

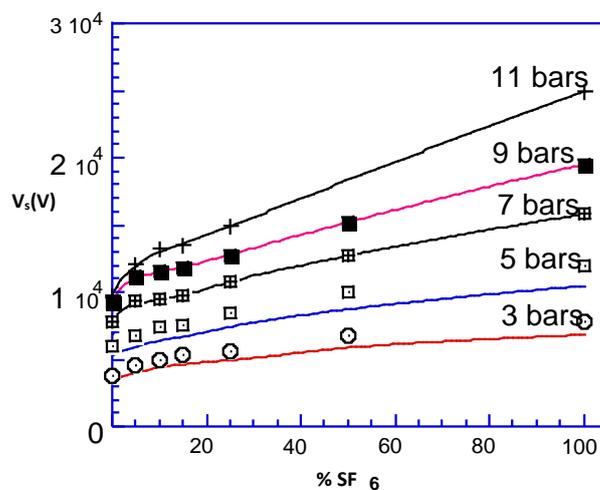


Figure 2. Curves of the onset voltages as a function of the ratio of SF_6 in the SF_6-N_2 mixture and gas pressure for negative polarity.

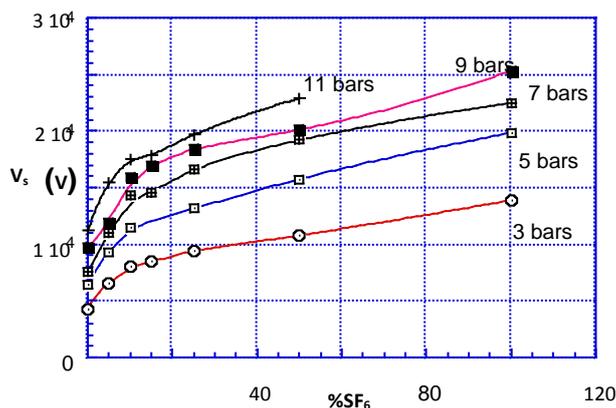


Figure 3. Curves of the onset voltages as a function of the ratio of SF₆ in the SF₆-N₂ mixture and gas pressure for positive polarity.

At higher amounts of SF₆, the onset voltages tend towards saturation. This tendency was observed in previous work [4]. The saturation may be attributed to the change of the surface of the tip electrodes. To check this assumption the surface was analysed using electronic microscopy. The analysis showed that there's some deposit on the tip after a set of electric discharge, this deposit is composed of fluorine and sulphur, but since tests are carried out with decreasing pressure such effect can't explain saturation for higher pressures, where the tip is quite free of deposit. The saturation may be the result of the concentration of space charges, which are more active at the vicinity of the tip electrode at higher pressure [5-6]. It is interesting to see that the onset voltages of the corona discharge are closer for low amounts of SF₆, which constitutes an argument for the replacement of SF₆ by SF₆-N₂ gas mixtures having very small percentage of SF₆. The onset voltages for positive and negative polarities are shown in figure 4, as a function of the gas pressure for 10% of SF₆ in the gas mixture. It can be seen that positive onset voltages are relatively higher [7]. This can be attributed to the mechanism of generation of initiatory electrons. Under positive polarity the main source for production of electrons is the detachment from negative ions and the difficulty for negative ions to reach the tip electrode. For negative polarity the field effect emission increases the probability of free electrons in the critical volume [8]. The effect of polarity is predominant for higher amounts of SF₆.

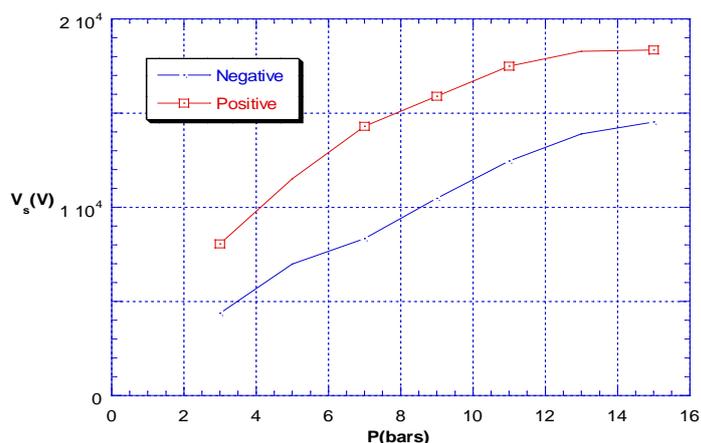


Figure 4, the onset voltages for positive and negative polarities, as a function of the gas pressure for 10% of SF₆ in the gas mixture.

Using the curves of figures 2 and 3, we were able to measure the working pressure for different SF₆-N₂ gas mixtures. In figure 8, are drawn the pressure values as a function of the percentage of SF₆ for the same onset voltage (13.8 kV) and for positive and negative polarities. In the present conditions, and for an onset voltage equals 13.8 kV a mixture of 100 % of SF₆ has a pressure

equals 3 bars for a positive polarity, whereas, for 10 % of SF₆ the pressure rises to 6.52, in order to withstand the same onset voltage. In negative polarity a mixture of 100 % of SF₆ has a pressure equals 7.3 bars; whereas, for 10 % of SF₆ the pressure rises to 13.12 bars, in order to withstand the same inception voltage (13.8 kV).

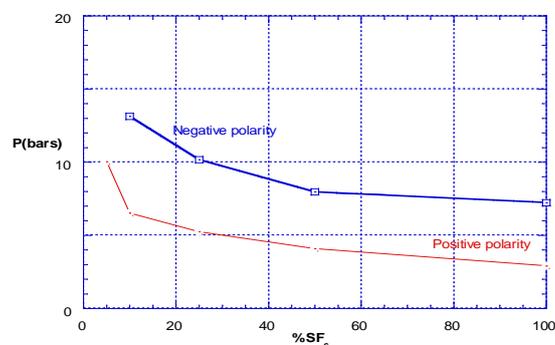


Figure 5, curves of the pressure as a function of the percentage of SF₆ for the same onset voltage (13.8 kV)

Conclusion

The corona onset voltages V_s increase with the increase of the amounts of SF₆ but at higher pressures the measured values tend to saturate. The positive onset voltages are relatively higher, especially for high percentage of SF₆. The polarity effect is predominant for higher amounts of SF₆. For the same onset voltage a mixture with 10 % of SF₆ must work approximately at twice the pressure of pure SF₆. This is true for both polarities.

References:

1. L. Niemeyer & F. Y. Chu «SF₆ and the atmosphere» IEEE Trans Elect Insul, vol 27, 1992.
2. L. G. Christophorou & R. J. Van Brunt «SF₆/N₂ Mixtures Basic and HV Insulation Properties» IEEE Trans on diel. and Elect. Insul., vol 2 N°5, 1995.
3. F. Y. Chu. " SF₆ Decomposition In Gas-Insulated Equipment" IEEE .Trans. on Elect. Insul, Vol. EI-21 N°5. vol 27, 1986.
4. A. H. Cookson and B. O. Pederson, "Analysis of the HV Breakdown Results for mixtures of SF₆ with CO₂, N₂ and Air", 3th Int. Symp. HV. Engineering, Milan, pp. 31-10. . 1979.
5. C. M. Cooke and R. Velazquez, "The insulation of ultra HV in coaxial Systems Using Compressed SF₆ gas", IEEE Trans. Power App. Syst, Vol.96, pp. 1491-1497, 1977.
6. V. K. Makdawala, D. R. James and L. G. Christophorou, "Effect of ionisation processes on the corona stabilisation breakdown in SF₆ and SF₆-mixtures", 4th Int. Symp. On High Voltage Engineering. Athens, Greece, pp.1-4, 1983.
7. Tomoyuki Hirata, Hideki Ueno, and Hiroshi Nakayama "Characteristics of N₂/SF₆ Mixture Gas in Creeping Discharge Developing in Narrow Gap with Backside Electrode" Electrical Engineering in Japan, Vol. 158, , 2007
8. F. A. M. Ritzk & M. B. Eteiba "Impulse breakdown voltage-time curves of SF₆ and SF₆-N₂ in coaxial cylinder gaps" IEEE, Transactions on Power Apparatus and Systems, vol. PAS-101. n°12, pp.4460-447, 1982.

THE EFFECT OF CARBON BLACK ON D.C. CONDUCTIVITY OF EPOXY RESIN

Ali S. Ali, Lecturer

Almuthanna University, College of Science, Physics department

Kawakib J. Majeed, Assistant Prof.

Baghdad University, College of Science, Physics department

Abstract:

D.C conductivity of epoxy composites was determined as a function of carbon concentration and temperature in the range (301-401) K. The samples were prepared by hand lay up method with different black carbon weight percentage ratios (2.5%, 5%, 7.5% and 10%) as sheets of dimensions (30x20x2)mm³. The electrical conductivity has been measured as a function of temperature by using DC power supply PE-1540, digital electrometer Keithely 616 and electrical oven. It has been found that the increase in carbon concentration caused increase in DC conductivity at room temperatures (RT) to maximum value of $9.42 \times 10^{-10} (\Omega \cdot \text{cm})^{-1}$ for carbon weight percentage of (10%). The Hall mobility and carrier concentration at laboratory temperature has been calculated, the Hall mobility values are increased from $1.41 \text{ cm}^2/\text{V} \cdot \text{sec}$ for the pure epoxy to $9.42 \text{ cm}^2/\text{V} \cdot \text{sec}$ for the composite with (10 wt %) C, and the carrier concentration values are increased from $4.2 \times 10^4 \text{ cm}^{-3}$ for the pure epoxy to $6.4 \times 10^6 \text{ cm}^{-3}$ for the composite with (10 wt %) C.

Key Words: Carbon black, DC conductivity, Epoxy Resin

Introduction

Much of the industrial progress is based on the use of organic materials such as polymers as insulators of heat and electricity. However, our modern lifestyle increasingly demands more from the polymers than their traditional role of insulators. With the rapid development of electronics industry, the demand for electrically conductive materials, such as electromagnetic wave interference shielding materials for personal computers and home electronic devices, flooring and ceiling materials, de-electrifying cloths and radar cross-section reducing protective fabrics for stealth technology has increased. This has led to the investigation and the subsequent commercial exploitation of organic polymers as conductors of electricity [1].

Conducting polymers have been the subjects of study for many decades as possible synthetic metals. However, the practical uses of conducting polymers are not very likely because of their poor mechanical properties and processability that rarely meet the industrial requirements. Thus, a unique combination of electronic and mechanical properties of composites of conducting polymers with conventional polymers seems to have great applications. The application of conducting polymers such as light emitting diode (LED)[2], Sensors[3], Photovoltaic cell[4], Field effect transistor (FET) [5], etc. The simplest example of the class of conducting polymers is polyacetylene (CH)_x, which was the first conducting polymers its can be doped by electron donors or acceptors lead to n-type or p-type polymers respectively [6]. It consists of weakly coupled chains of CH units forming two type of lattice; the first is called trans-(CH)_x which was the stable isomer and the other was called cis-(CH)_x is isomer[7].

Most of the research works on the conductive polymers composites focus on modification of the electrical properties by subjecting their structures to various physical and chemical conditions to produce new systems with modified conductivity. In this work black carbon was added to epoxy polymer to improve its conductivity [8-11].

Experimental

Material used Preparation

The material used to prepare the test samples were Epoxy resin (EP) type (Leyco-Pox 635) with one type hardener formulated amine was supplied by leyco chem. Leyde industry ,and black carbon supplied by industrial chemical and resin CO.LTD,kingdom of Saudi Arabia.

The test samples were prepared by hand lay-up method with diferent weight percentage ratios and dimensions (30x20x2)mm³, highly conductive silver paste is used to produce electrodes for DC contacts . The details of the test samples presented in table (1).

Table (1) details of the test samples

Sample No.	Composites Ratios
1	0% Black carbon +100 % epoxy
2	2.5% Black carbon + 97.5 % epoxy
3	5% Black carbon + 95 % epoxy
4	7.5% Black carbon + 92.51 % epoxy
5	10% Black carbon + 90 % epoxy

Testing Procedure

For the dc- conductivity measurements, samples with plane geometry have been used .The electrical contacts are made of fine silver paste. Kethley (616) digital meter was used for resistance measurement in the temperature range of (301-401K) temperature was monitored during the measurement using a chromel-alumel thermocouple.

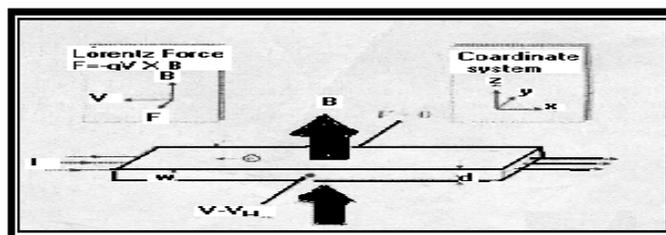
Dc conductivity of the samples as a function of temperature is studied. The rate of heating of the samples was controlled using an oven type (memert 854).The activation energy of the test samples is calculated, from the plot of Ln σ versus (10³/T) and by using Stuke's equation [10].

$$\sigma = \sigma_0 \exp(-E_a / k_B T) \dots\dots\dots(1)$$

This equation is basically derived to give the change of the electrical conductivity with temperature .E_a denotes the thermal activation energy of electrical conduction (σ) which is a parameter that depends on the material nature, (σ₀) is defined as the proportionality factor which represents the conductivity when temperature approaches infinity, T is the absolute temperature and k_B is Boltzman's constant

As a current –carrying conductor is placed in a transverse magnetic field, the Lorentz force on moving charge pushes them toward one side of the conductor producing a charge separation and as a result, a voltage arises in the direction perpendicular to both the magnetic field and the current and called Hall voltage as shown in Fig. (1)[12].

Fig (1) Schematic diagram of Hall effect [6]



Hall effect has been measured by using the electrical circuit which contain (D.C power supply (0-40)V, when the samples carrying a current expose a constant magnetic field (B=0.254)Tesla perpendicular to the electric field then an e.m.f which is called Hall voltage(V_H) is set up across the sample, the current I and V_H were recording by using Keithly Digital Electrometer 616. The

concentration and mobility of the carrier for the tested samples can be determined by the following equations [13].

The Hall coefficient (R_H) is found by:

$$R_H = \pm 1 / n q \dots \dots \dots (2)$$

If V_H is the Hall voltage across the sample, I is the current and B is the applied magnetic field, then the Hall coefficient is:

$$R_H = t V_H / I B \dots \dots \dots (3)$$

Where t is the thickness of the sample. If the conduction is due to the one carrier type (e.g: electrons) then we can find the mobility from the equation [14].

$$\mu = \sigma / n q \dots \dots \dots (4)$$

and

$$\mu = \sigma (R_H) \dots \dots \dots (5)$$

Results and Conclusion

Variation of d.c conductivity with temperature is the main tool in investigating the electrical properties of the composites by plots ($\ln\sigma$) vs. ($10^3/T$), the activation energy can be measured from the slope of straight lines of ($-\Delta E/K$).

The d.c conductivity σ for the composites has been determined as function of ($10^3/T$), it was noticed that at room temperature the d.c conductivity ($\sigma_{r,t}$) increased with the increasing of carbon weight percentage as shown in table (2) .

Table (2): Variation of D.C conductivity $\sigma_{r,t}$ with black carbon content at RT.

Sample	$\sigma_{r,t} (\Omega.cm)^{-1} \times 10^{-10}$
0% Black carbon +100 % epoxy	1.41
2.50% Black carbon +97.5 % epoxy	6.23
5% Black carbon +95 % epoxy	6.86
7.5% Black carbon +92.5 % epoxy	8.11
10% Black carbon +90 % epoxy	9.42

Fig. (2) shows the variation of ($\ln\sigma$) vs. ($1000/T$) for epoxy resin and its composites, it illustrates that the conductivity increased with the increasing of carbon weight percentage , The influence of temperature on the d.c conductivity has been explained by considering the mobility of charge carriers responsible for hopping. As temperature increases the mobility of hopping ions also increases thereby increasing conductivity. The electrons which are involved in hopping are responsible for electronic polarization in the composites [9,10].

The activation energy of epoxy resin and its composites were determined and listed in table (3).

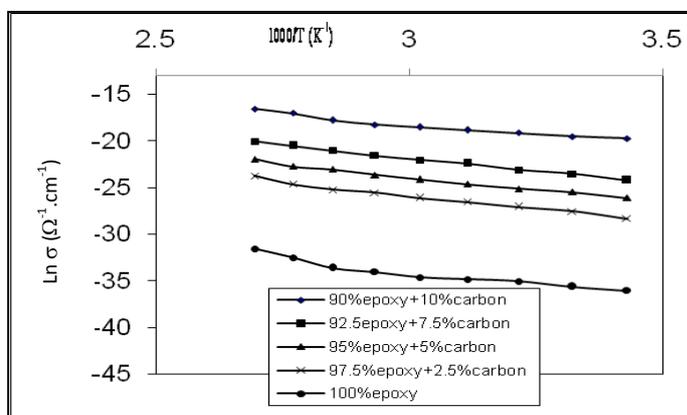


Fig. (2): Variation of $\ln \sigma (\Omega^{-1} cm^{-1})$ with $1000/T (K^{-1})$ for epoxy resin and its composites

Table (2) The activation energy values E_a as a function of C content for epoxy resin and its composites

Sample	E_a (eV)
0% Black carbon +100 % epoxy	4.3
2.50% Black carbon +97.5 % epoxy	4.06
5% Black carbon +95 % epoxy	3.8
7.5% Black carbon +92.5 % epoxy	3.4
10% Black carbon +90 % epoxy	3.1

Figs.(3-7) show the variation of Hall voltage as a function of currents for pure epoxy and its composites .from which the Hall coefficients R_H were determined and listed in Table (3) .

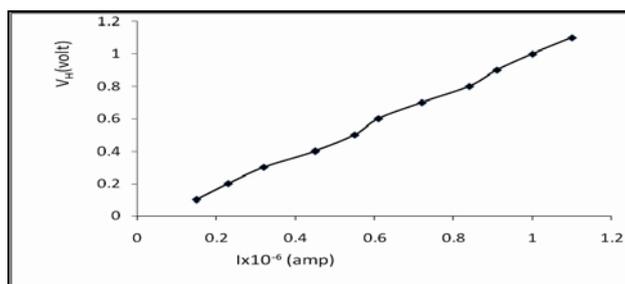


Fig.(3): Variation of Hall voltage as a function of current for pure epoxy

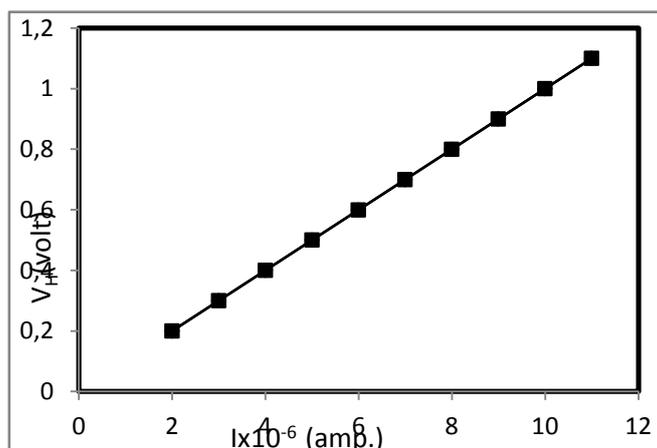


Fig.(4) Variation of Hall voltage as a function of current for composite of 2.5% carbon weight percentage

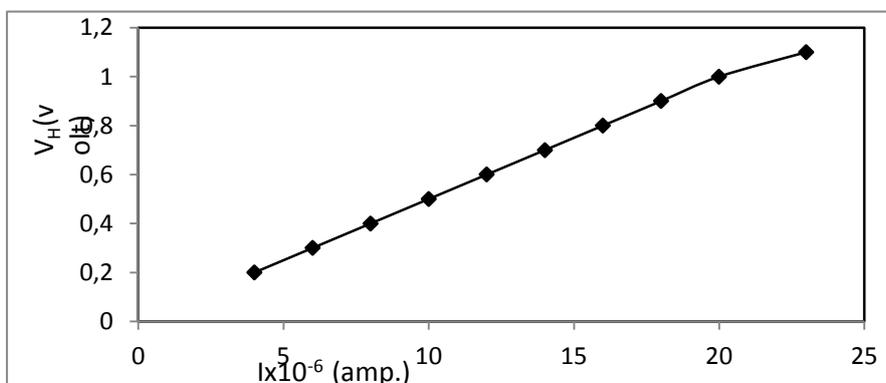


Fig.(5) Variation of Hall voltage as a function of current for composite of 5% carbon weight percentage

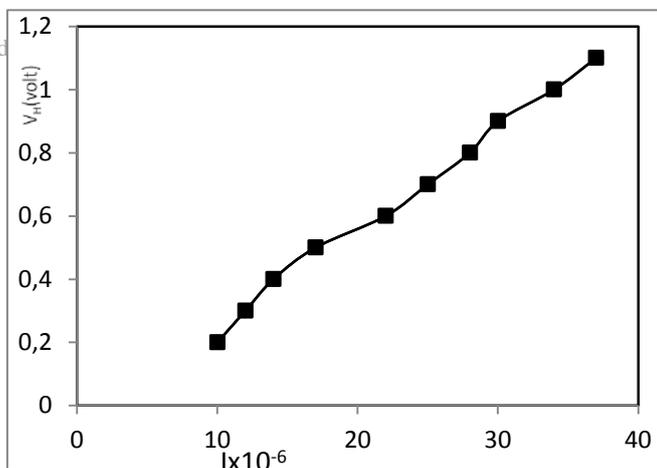


Fig.(6) Variation of Hall voltage as a function of current for composite of 7.5% carbonweight percentage

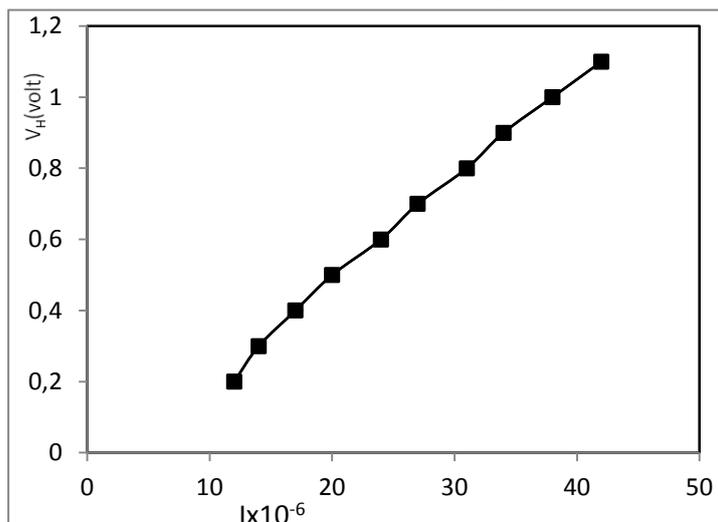


Fig.(7) Variation of Hall voltage as a function of current for composite of 10% carbonweight percentage

Table (3): Variation of Hall coefficients R_H and carriers concentration n_H and mobility μ_H with carbon content

1) Carbon(W _t %)	R _H (cm ³ /C)	2) n _H (cm ⁻³)	μ _H (cm ² /V.sec)x10 ⁻⁸
0	84	3) 4.2 x10 ⁴	0.018
2.5	33.7	4) 2.1 x10 ⁶	5.20
5	30.2	5) 3.7 x10 ⁶	8.17
7.5	27.6	6) 5.2 x10 ⁶	14.5
10	7.51	7) 6.4 x10 ⁶	19

Also the carrier's concentration and mobility were determined and listed in Table (3). It is found that the carrier's concentration of the composites much higher than that of pure epoxy due to the presence of excess carbon .It is found also that the mobility increases exponentially with increasing carbon concentration.

Conclusions

The conductivity at room temperatures increases with increasing C concentration .also the charge carrier concentration and Hall mobility increases with increasing C concentration.

References:

- (1) Ahmed H. and Marwa A., The D.C Electrical Properties of Polyvinylchloride-tin Composites, 2010, International Conference on Mathematical Applications in Engineering (ICMAE'10), 3-5 August, Kuala Lumpur, Malaysia.
- (2) Jen-Wei Huang and Shih Jung Bai, INSTITUTE OF PHYSICS PUBLISHING, Nanotechnology 16 ,1406–1410, (2005).
- (3) Hua Bai and Gaoquan shi, Sensors, 7, 267-307, (2007).
- (4) A. J. Breeze, Z. Schlesinger, and S. A. Carter, physical review B, Volume 64, 125205,(2001).
- (5) Wael A.S. Abdul Ghafor, Georgius A.AL-Adam, Basrah Journal of Sciece(C), Vol.23(1),1-9,(2005).
- (6) K. M. Zaidan, S.J. Abbas, Basrah Journal science, A, Vol.18,No. 1, 97-104, (2000)
- (7) Zeev Valy Vardeny and X. Wei, "HAND BOOK OF CONDUCTING POLYMERS", 2nd ed., New york, Basel. Hong Kong.(1998).
- (8) Ahmed M. S. And Zihlif A. M. , The electrical conductivity of polypropylene and Nickel- Coated carbon Fiber composite, 1992, J. Mater. Sc. Vol. 25, No. 706, Universit of Jordan, Amman, Jordan.
- (9) Bhattacharya S., Sachdev V.K., Tandon, Electrical properties of Graphite filled polymer composites, 2008, 2nd National Con R.P. ference Mathematical Techniques: Emerging Paradigms for Electronics and IT Industries
- (10) Comelio P., lazzeri V., and Waggle P., polymer preprints, 1996 Vol.37, No.2 (1996), PP. (17-21).
- (11) Hamzah M. , Saion E., Kassim A. and Yousuf M. , Temperature dependence of AC Electrical conductivity of PVA-PPy-FeCl₃ composites polymer Films, 2008, MPJ. Vol. 3, No. 2, p(24-31) , Malaysia.
- (12) D. Halliday & A. Resnick, "Physics", 4th ed, New York, Wiley and Sons. Inc, V.2, (1992) 745-46.
- (13) X. J. Du ., J. H. and Ying Z. , positive temperature coefficient effect in multwalled carbon nanotube high –density polyethylene composite, 2005, Vol. 86, J. Appl. Phys. Letters, China.
- (14) Kontos et al, electrical relaxation dynamics in TiO₂-polymer matrix composites, 2007, express polymer letters, Vol.1, No. 12, p (781-789).
- (15) N. CHAND and D. JAIN "Evaluation of a.c. conductivity behaviour of graphite filled polysulphide modified epoxy composites" Bull. Mater. Sci., Vol. 27, No. 3, pp. 227–233. June 2004.
- (16) D. Halliday & A. Resnick, "Physics", 4th ed, New York, Wiley and Sons. Inc, V.2, (1992) 745-46.

MATHEMATIC-PHYSICAL MODEL OF DIMENSIONING SYSTEM IN THE PROPAGATION OF MICROWAVE "WAVEGUIDE-SLUDGE" FROM WASTEWATER TREATMENT PLANTS

Emilian-Narcis Riți-Mihoc, Phd. Eng.

Emil Riți-Mihoc, Phd. Eng.

Dan Porcar, Phd. Eng.

Technical University of Cluj-Napoca, Roumanie

Abstract:

This paper presents a physico-mathematical model sizing system microwave field propagation through a waveguide of the waveguide composed of a rectangular section and a pyramidal funnel section rectangular variable distribution. This model was adapted to the conditions of "waveguide-pond sludge bed" treaty. The system is to define parameters and propagation equations for adaptation.

Key Words: Physical and mathematical model, microwave, sludge settling, waveguide

Introduction

Making the physical and mathematical model was developed in order to create a mathematical tool for calculating the size of the wave guide stand for laboratory research on the field of microwave thermal processing of sludge from wastewater treatment pond.

The model is based on the general theory of microwave propagation presented in the literature [1], [2], [3], [4], [5], adapting to the specific mathematical relations of "waveguide-bed mud pond". It takes into account phenomena occurring in electromagnetic wave propagation through a rectangular guide and mathematical relationships that define these phenomena.

Determination of model parameters

For the specific case of the plant designed to establish physical and mathematical model of microwave propagation system we took a structural analysis with the following composition: all the magnetron, wave guide, funnel distribution and resonant cavity, shown schematically in Figure 1.

Microwave transmission system, rectangular section; work on the fundamental propagation mode TE_{01} .

Determination of physical-mathematical model of microwave propagation in the system involves defining the parameters of the rectangular guide propagation.

Critical frequency is defined as the frequency below which no electromagnetic wave propagation occurs.

This frequency is calculated with:

$$f_c = \frac{c_0}{2} \sqrt{\left(\frac{m}{b}\right)^2 + \left(\frac{n}{a}\right)^2} \quad (1)$$

Critical wavelength is the wavelength corresponding critical frequency is determined with:

$$\lambda_c = \frac{c_0}{f_c} \quad (2)$$

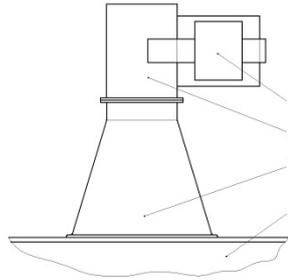


Figure 1. - The propagation of microwaves
 1 - magnetron assembly, 2 - waveguide, 3 - Funnels distribution, 4 - cavity resonant

Guide wavelength is calculated with:

$$\lambda_g = \frac{\lambda_0}{\sqrt{1 - (\lambda_0/\lambda_c)^2}} \tag{3}$$

Vacuum wavelength λ_0 , appropriate working frequency is calculated with:

$$\lambda_0 = \frac{c_0}{f} \tag{4}$$

Dephasing constant is calculated with:

$$\beta_{0,1} = \frac{2\pi \cdot f}{c_0} \sqrt{1 - (f_c/f)^2} \tag{5}$$

Dephasing constant change relative to 2π (i.e. $\beta_0/2\pi$), the frequency relative to the speed of propagation in free space c_0 (i.e. f/c_0), is represented in Figure 2.

In this case the dephasing constant does not vary in proportion to frequency, so the user is a dispersive wave.

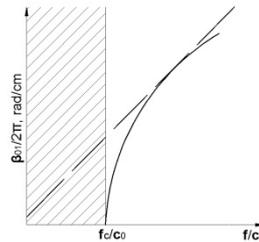


Figure 2. - Variation of constant lag in the reported frequency c_0 [1]

Group velocity is the actual speed of propagation of microwave energy is calculated with:

$$v_g = c_0 \cdot \sqrt{1 - (f_c/f)^2} \cdot 10^{-2} \tag{6}$$

Group velocity is less than the speed of light and by the asymptotic (figure 3.).

Wave impedance

If the wave magnetic field H_{01} (for mode propagation TE_{01}), the wave impedance is calculated with:

$$Z = Z_0 \cdot \frac{1}{\sqrt{1 - (f_c/f)^2}} \tag{7}$$

The:

$Z_0 = \sqrt{\mu_0/\epsilon_0} = 120 \cdot \pi = 377$ is the vacuum wave impedance.

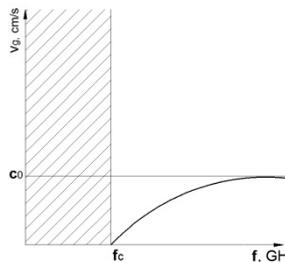


Figure 3.- Group-velocity variation with frequency [1]

In figure 4 is presented the wave impedance variation with frequency.

The expression for calculating the wave impedance (7) is valid for uniform waveguide (lossless) and air inside the propagation environment. Thus the frequency band $f > f_c$ wave impedance is purely resistive.

Distribution of electromagnetic field components in the guide is given by relations (8)...(10):

$$H_z = H_0 \cdot \cos \frac{\pi}{a} x \cdot \cos(\omega t - \beta_{01} \cdot z) \quad (8)$$

$$H_x = -\frac{\lambda_c}{\lambda_0} \cdot H_0 \cdot \sin \frac{\pi}{a} x \cdot \sin(\omega t - \beta_{01} \cdot z) \quad (9)$$

$$E_y = -\frac{\lambda_c}{\lambda_0} Z_0 \cdot H_0 \cdot \sin \frac{\pi}{a} x \cdot \sin(\omega t - \beta_{01} \cdot z) \quad (10)$$

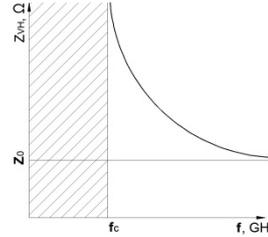


Figure 4. - Variation of impedance with frequency [1]

Curves of variation of the amplitude components x are presented according to figure 5, where it is shown as lines of electric and magnetic field in the waveguide where the wave H_{01}

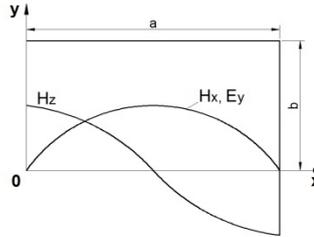


Figure 5. - Variation of intensity components H_z , H_x , and E_y depending on x [1]

From this figure two conclusions can be drawn: the electric field tangential to the wall is zero and shows many signs m and n of $H_{m,n}$ the semi-sine wave is present on the base a and the height b . If H_{01} for $x = a$ is a semi-sinusoid; and b the field is constant, i.e. not depending on y .

Attenuation constant determines the power losses in the waveguide. This constant is calculated from the dimensions of the waveguide section. For the wave H_{01} , the expression for calculating the attenuation constant is:

$$\alpha_{H_{01}} = \frac{\frac{2}{a} + \frac{2}{b} \left[1 + \left(\frac{\lambda_c}{\lambda_g} \right)^2 \right]}{\delta \cdot \psi \cdot Z_0 \cdot \lambda_c^2 / \lambda_0 \cdot \lambda_g} \quad (11)$$

Penetration depth δ is determined by the:

$$d = \frac{5}{\sqrt{f \cdot \psi \cdot 10^7}} \cdot 10^{-6} \quad (12)$$

From this relation results the penetration depth [m], if the frequency is measured in [GHz] and constant wall material forming the guide, in [s/m];

The dissipated power in the walls of the guidance system is determined by the total attenuation of the microwave route guidance system, and is given by:

$$\alpha = \alpha_{H_{01}} \cdot l \quad (13)$$

Where:

l - is the length of the trail and microwave to the material is composed of: magnetron waveguide length, and the length of the distribution funnel. For the distribution funnel was taken a coverage factor of 2, so that the total equivalent length is:

$$l = l_1 + 2 \cdot l_2 \quad (14)$$

Since,

$$\alpha = \log X \quad \text{and} \quad X = \frac{P_i}{P_e}$$

Where:

X - is the ratio of powers of entry into the route P_i and exit P_e .

results:

$$X = 10^{\frac{\alpha}{10}} \quad (15)$$

The dissipated power (which is converted into heat in the metal walls of the guidance system by the Joule-Lentz effect) is given by:

$$P_d = \frac{X-1}{X} \cdot P_i \quad (16)$$

The maximum allowable power is limited, depending on the guide size and accepted default losses. The maximum permissible power, transmitted through the waveguide for the wave $H_{0,1}$, in conditions of adaptation:

$$P_1 = 2500 \cdot \sqrt{1-\eta^2} \cdot a \cdot b \quad (17)$$

In case of inadapation, power in load is the difference between the corresponding power of the incident wave and power of the reflected wave, i.e.:

$$P_s = \frac{a \cdot b}{4 \cdot Z} \cdot (E_i + E_r) \cdot (E_i - E_r) \quad (18)$$

Because

$E_i + E_r = E_{\max}$; $E_i - E_r = E_{\min}$ and $\frac{E_{\max}}{E_{\min}} = \sigma$ (Defined as standing wave factor), we obtain:

$$P_s = \frac{a \cdot b \cdot E_{\max}^2}{4 \cdot Z \cdot \sigma} \quad (19)$$

That transmitted power is σ times less, than, for the same amount of electric field strength E_{\max} .

The power variation (or the squared electric field strength, E^2) in the case of adaptation is shown in figure 6.

Basically, for the modeling of microwave sizing guide must always consider the condition of adaptation, the condition is satisfied if the standing wave factor is $\sigma = 1,1 \div 1,2$.



Figure 6. - Variation of the square intensity of the electric field, in case of adaptation [1]

For the maximum value of the standing wave factor $\sigma = 1,2$, reflection coefficient becomes $|\Gamma| = \frac{\sigma-1}{\sigma+1} = \frac{0,2}{2,2} = 0,09$ and so the reflected power (reported to the incident power) represent 1% , which is practically acceptable.

Determination of the specific equation for the primary physical model of adaptation

In the case of pyramidal scheme distribution funnel (figure 1), adopted when designing mobile applicator laboratory stand, we can define the physical model of primary adaptation, whose principle scheme is shown in figure 7.

This model have of three areas of propagation: propagation in space guided, which is composed of the constant section guide and the variable cross-section guide, the pyramidal funnel; the propagation in free space

from the output of funnel and spread to the surface of sludge AA' and in the interior of the sludge bed, which is characterized as a mixture of two components, water and solid part.

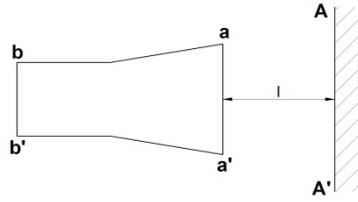


Figure 7. - Pyramidal funnel propagation system [1]

This propagation system can be equivalent with an electric circuit similar to that in figure 8. Sludge bed is a loss medium characterized in terms of dielectric constant and conductivity electromagnetic times exactly tangent of loss angle:

$$tg\delta_p = \frac{\sigma}{\omega \cdot \epsilon_p} \tag{20}$$

As the angle is small, it is considered to be equal to the value of angle tangent of loss angle in radians: $tg\delta_p \cong \delta_p$.

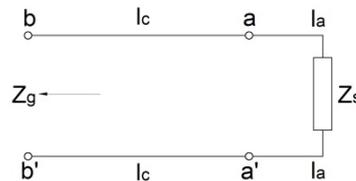


Figure 8. - Equivalent circuit of the adaptation system

Adaptation is the elimination of reflected wave (reflected power that is proportional to the square reflected wave shown). Elimination of reflected wave occurs when the transition from one environment to another is eliminated discontinuities, impedance i.e. at the left and right of the surfaces AA', aa' and bb' is conveniently turns on the same as shown.

It is necessary therefore to calculate the input impedance at the surface AA'.

Establishing the impedance calculation relationships of the at the entrance in the sludge bed:

Field distribution in the sludge bed, at the right of the surface AA' is exponential and there is no reflected wave, due to the high attenuation in the treated sludge. In this case the surface impedance is the impedance of semi-infinite space with the dielectric constant ϵ_p and the loss angle δ_p . Impedance of free space is $Z_0 = \sqrt{\mu_0/\epsilon_0}$ [Ω] and in the case of dielectric environment with loss, the impedance becomes:

$$Z = \sqrt{\frac{\mu_0}{\epsilon_0(\epsilon_p - j\delta_p)}} \tag{21}$$

where the angle of loss is given by:

$$\delta_p = \frac{\sigma}{j \cdot \omega \cdot \epsilon_0} \tag{22}$$

In this case – at 2,45 GHz frequency for the water $\epsilon_p = 77$ and $\delta_p = 0,15$, and the solid part $\epsilon_p = 4$ $\delta_p = 0,01$.

Equations for calculating the average electromagnetic constants and input impedance of sludge bed:

Sludge can be treated as a mixture of two components: water and the solid, so the average dielectric constant (ϵ_{pm}) is calculated with:

$$\epsilon_{pm} = \epsilon_{p1} \times P_1 + \epsilon_{p2} \times P_2 \tag{23}$$

Impedance at the entrance in the sludge bed is determined by the relationship:

$$Z_{AA'} = \frac{Z_0}{\sqrt{\epsilon_{pm}}} \quad (24)$$

The equations of variation of microwave reflection:

Input impedance in the sludge bed has different values depending on the water content, during drying process the water content is variable and the average dielectric constant is variable so that reflection of microwaves will occur.

Generally the reflection coefficient is given by:

$$\Gamma = \frac{Z_s - Z_{01}}{Z_s + Z_{01}} \quad (25)$$

Where:

Z_s - is the load impedance (for this case, the input impedance in the sludge bed);

Z_{01} - is the reference impedance, where the reflection coefficient is zero (in this case, the impedance of the water content for which is the adaptation).

The adjustment is made for the reference water content, calculating the reflectivity for the minimum and maximum water content in sludge and with (26) is determined the reflected power:

$$P_r = |\Gamma|^2 \cdot P_i \quad (26)$$

With relation (27) is determined the relationship between the maximum electric field and minimum electric field in the wave guide for the two water contents.

$$k = \frac{1 + |\Gamma|}{1 - |\Gamma|} \quad (27)$$

Permissible values of the water content in the sludge are given by the allowable amount of reflected power (P_r). In the general case of drying materials with high humidity, it can accept power reflected a permissible value of 10% of rated output (P_i) of the magnetron, in this case, reflectivity coefficient is:

$$|\Gamma|^2 = \frac{P_r}{P_i} = 0,1 \text{ i.e. } \Gamma = 0,316 \text{ and } k = \frac{1,316}{0,684} = 1,923 \cong 2$$

Establishing the line equivalent relations sludge-funnel:

The space between the sludge and funnel distribution mouth (figure 7) the distance l , between the surfaces AA' and aa' , has the equivalent circuit two-wire line with the length l_a , between the load Z_s (input impedance of the sludge bed) and terminals aa' (figure. 8).

Putting on that, the portion provided that propagation is lossless, the relationship for calculating the input impedance of a lossless line is:

$$Z_1 = Z_{01} \cdot \frac{Z_s + j \cdot Z_{01} \cdot \operatorname{tg} \beta_0 \cdot l}{Z_{01} + j \cdot Z_s \cdot \operatorname{tg} \beta_0 \cdot l} \quad (28)$$

Where:

Z_{01} - is the characteristic impedance of the line;

Z_s - is the load impedance;

β_0 - is the dephasing constant, with: $\beta_0 = \frac{2\pi}{\lambda_0}$;

l - is the length of the line.

Should be provided $\operatorname{tg} \beta_0 \cdot l = 0$ and when $Z_s = Z_i$, i.e. the load impedance is equal to the distribution at the mouth of the funnel. The distance between the shape and surface area at the mouth of the distribution funnel is derived from the relation:

$$\beta_0 \cdot l = N\pi, \quad N = 0,1,2,3... \quad (29)$$

And results:

$$l = N \cdot \frac{\lambda_0}{2} \quad (30)$$

So the distance between the distribution funnel mouth area and surface of the sludge is an integer value equal to half the size of the wavelength.

Funnel size optimization equations:

To establish the equations for optimization of the funnel waveguide dimensions must consider the variation of the module of the reflectance coefficient (Γ), according to the ratio l/λ_g , shown in figure 9.

Thus, the analysis of figure 9, results that (Γ) is small when the ratio l/λ_g is around

$$l = N \cdot \frac{\lambda_g}{2} .$$

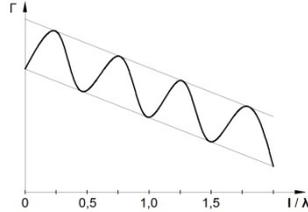


Figure 9. - The variation of the module (Γ) according to the ratio l/λ_g [1]

If the funnel section is quoted $a(x)$ in any point of its axis x (Figure 10), that can express this value by the relation:

$$a(x) = \frac{a_1 - a_0}{l_1} \cdot x + a_0 \tag{31}$$

Generally:

$$\int_0^{l_1} \frac{2\pi}{\lambda_g(x)} dx = (N - 2)\pi \tag{32}$$

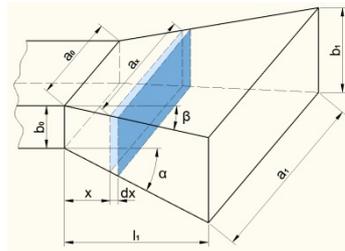


Figure 10. - The distribution funnel dimensions

And the guide wavelength at a distance x , we can write the relationship:

$$\lambda_g(x) = \lambda_0 / \sqrt{1 - \left(\frac{\lambda_0}{2a(x)}\right)^2} \tag{32}$$

So results:

$$\int_0^{l_1} \frac{2\pi}{\lambda_g(x)} dx = \int_0^{l_1} \frac{2\pi}{\lambda_0} \sqrt{1 - \left(\frac{\lambda_0}{2a(x)}\right)^2} dx = \frac{2\pi}{\lambda_0} \int_0^{l_1} \sqrt{1 - \left(\frac{\lambda_0}{2a(x)}\right)^2} dx \tag{33}$$

Noting:

$$T = \int_0^{l_1} \sqrt{1 - \left(\frac{\lambda_0}{2a(x)}\right)^2} dx \tag{34}$$

Results:

$$T = \frac{\lambda_0 \cdot l_1}{2(a_1 - a_0)} \left[\sqrt{\left(\frac{2a_1}{\lambda_0}\right)^2 - 1} - \sqrt{\left(\frac{2a_0}{\lambda_0}\right)^2 - 1} - \arccos \frac{\lambda_0}{2a_1} + \arccos \frac{\lambda_0}{2a_0} \right] \tag{35}$$

but:

$$\frac{2\pi \cdot T}{\lambda_0} = (N - 2) \cdot \pi \tag{36}$$

It follows the relationship for calculating the length of the distribution funnel:

$$l_1 = \frac{(N - 2) \cdot \pi \cdot \lambda_0}{T \cdot 2\pi} \text{ [cm]} \tag{37}$$

Calculation of tilt angles of the distribution funnel:

- the long side of the section:

$$tg\alpha = \frac{b_1 - b_0}{l_1} \quad (38)$$

- the short side of the section:

$$tg\beta = \frac{a_1 - a_0}{l_1} \quad (39)$$

The calculation relationships for the reflectance coefficient:

The reflection coefficient in the distribution funnel is:

$$|\Gamma| = \frac{\lambda_0}{l} \left[\frac{K_0^2 + K_1^2}{64\pi^2} - \frac{K_0 \cdot K_1}{32\pi^2} \right]^{1/2} \quad (40)$$

Where:

$$K_0 = \frac{\frac{b_1 - b_0}{b_0} - \frac{a_1 - a_0}{a_0} \left[\frac{1}{1 - (\lambda_0/2a_0)^2} \right]}{\left[1 - (\lambda_0/2a_0)^2 \right]^{1/2}} \quad (41)$$

$$K_1 = \frac{\frac{b_0 - b_1}{b_1} - \frac{a_0 - a_1}{a_1} \left[\frac{1}{1 - (\lambda_0/2a_1)^2} \right]}{\left[1 - (\lambda_0/2a_1)^2 \right]^{1/2}} \quad (42)$$

Notations:

TE_{01} - fundamental transverse mode of propagation of microwave power;

f_c - critical frequency, in [Hz];

f - operating frequency, in [Hz].

c_0 - speed of light, $c_0 = 3 \times 10^{10}$ in [m/s];

a, b - the base dimensions of the rectangular guide, in [m];

λ_c - critical wavelength, in [m];

λ_g - guide wavelength, in [m];

λ_0 - corresponding vacuum wavelength for the work frequency, in [m];

$\beta_{0,1}$ - dephasing constant, in [rad/m];

v_g - group velocity, in [m/s];

Z - impedance microwave propagation in various materials, in [Ω];

Z_0 - impedance to the propagation of microwaves in a vacuum space, in [Ω];

Z_m - average wave impedance, in [Ω];

$E_{m,n}, H_{m,n}$ - electric and magnetic intensity of the propagation mode indices m and n of electromagnetic waves;

H_{01} - magnetic intensity for the fundamental wave, $m = 0$ and $n = 1$;

$\alpha_{H_{01}}$ - attenuation constant, in [dB/m];

ψ - constant guide wall forming material, in [s/m];

d - depth of penetration of waves into the material, in [μm];

l - length of the route guidance system of the microwave, in [m];

α - total attenuation of the microwave, in [dB];

X - ratio between the power at the entrance and the power at the exit of the guide route;

P_d - power dissipated in the walls of the guidance system, in [W];

P_1 - limit power (maximum allowable by the guide), in [W];

$P_{r,70\%}$ - reflected power, in [W];

m, n - values of electromagnetic waves for example in the case of the transverse electric propagation mode TE_{01} , we have $m = 0$ and $n = 1$;

η - Field efficiency of transmission defined as: $\eta = \frac{f_c}{f}$;

σ - stationary wave factor;

Z_s - load impedance, in [Ω];

$Z_{0,1}$ - reference impedance, in [Ω];

$|\Gamma|$ - reflection coefficient module;

ε_{pm} - average dielectric constant of a mixture;

$\varepsilon_{p1}, \varepsilon_{p2}$ - dielectric constant of mixture components;

δ_p - angle of loss, in [rad];

p_1, p_2 - proportions of components in mixture, in [%];

N - wave number of the rank;

a_0, b_0 - funnel entrance dimensions, in [m];

a_1, b_1 - funnel exit dimensions, in [m];

l_1 - funnel length, in [m];

$K_0; K_1$ - reflection parameters calculated according to the funnel section size at the entrance

and exit;

Conclusion

This physic-mathematical model can be used to guide system sizing microwave applicator for mobile installations in the field of microwave processing of sludge from the wastewater treatment plants.

References:

- [1]. Riti-Mihoc. E.N., - *Studies and researches on the thermal processing of microwave field to sewage sludge of treatment plants*, Phd Thesis, Department of Engineering and Sustainable Development Entrepreneurship, Faculty of Materials and Environmental Engineering, Technical University of Cluj-Napoca, Cluj-Napoca, 2011.
- [2]. Dan, V., Riți-Mihoc, E., Riți-Mihoc, E. N. – *Thermal treatment of clearing sludge in microwave field*, 4th Research/Expert Conference with International Participations “QUALITY” 2005, Fojnica, B&H, 2005.
- [3]. Hill, J. M., Marchant, T. R. – *Modelling microwave heating*. Applied Mathematical Modelling; 20(1):3–15, 1996.
- [4]. Rulea, G. – *Tehnica Microundelor*, E.D.P., București, 1981.
- [5]. Metaxas, A. C., Meredith R. J. - *Industrial Microwave Heating*. Power Engineering Series 4. Peter Peregrinus Ltd. (on behalf of the IEE), 1993.
- [6]. Paresi, A. – *Etude de l'amélioration du rendement des techniques de séchage des matériaux épais par apport d'énergie électromagnétique*. Conception d'applicateurs, Thèse de Doctorat, L'Institut National Polytechnique de Toulouse, 1978.

INTERDISCIPLINARITY AROUND “CALÇADA PORTUGUESA” SIDEWALK

Helena Melo, PhD

CMATI & Department of Mathematics, University of Azores, Portugal

Abstract:

We find the "Calçada Portuguesa" sidewalk paving an interesting example of the collaboration of various interdisciplinary fields, such as Mathematics, Arts, Architecture, History, Tourism, Technology, Entrepreneurship, and many others. This work will focus on one possible use, the paving of sidewalks, highlighting the interest that arises from each mentioned areas, its concept, and its applications.

Key Words: Symmetry, frieze group, interdisciplinary

Introdução

Procurando o envolvimento de diversas áreas, tais como a Matemática, a Arte, a Arquitetura, a História, a Geologia, as Tecnologias, o Turismo e o Empreendedorismo, e possivelmente outras tantas mais, encontramos a temática dos passeios de Calçada Portuguesa.

Este trabalho incidirá sobre uma das suas utilizações, a pavimentação de passeios, destacando o interesse que desperta em cada uma das áreas mencionadas, o seu conceito e a sua aplicabilidade.

Desenvolvimento

Envolvendo as áreas de Matemática, Arte, Arquitetura, História, Geologia, Tecnologias, Turismo e Empreendedorismo, encontramos a temática dos passeios de Calçada Portuguesa. A pavimentação de passeios estimula a atenção de diversas destas áreas, entre outras, pelo seu conceito, pela sua história e grande aplicabilidade.

A Calçada Portuguesa é um legado da cultura romana. A arte de fazer pavimentações vem desde os tempos mais remotos da história do mundo. Desde as civilizações mais antigas eram utilizados materiais rochosos para pavimentar o interior e exterior dos edifícios. Acreditamos que os povos da Mesopotâmia foram os primeiros a fazer uso desta arte, seguidos pelo povo da Grécia Antiga e pela civilização romana. Devido a sua durabilidade, resistência, facilidade de manutenção de higiene, flexibilidade de montagem, composição plástica e beleza estética, este tipo de pavimentação, com estes materiais sólidos e resistentes, foi preferido para diversas construções arquitetónicas, em diferentes tempos históricos, sendo utilizado como uma expressão humana, religiosa e estética.

Com o terramoto de 1755, Lisboa iniciou o projeto de sua reconstrução, no qual um dos seus objetivos era a recuperação de ruas antigas e a abertura de novas, no espaço urbano. Na sua pavimentação utilizaram o que conhecemos hoje como Calçada Portuguesa, ou Mosaico Português. Neste seguimento, em meados do século XIX, com origem em Portugal, surge a Calçada Portuguesa que teve como um dos seus grandes impulsionadores o Tenente-General Eusébio Cândido Cordeiro Pinheiro Furtado (1777 – 1861), governador de armas do Castelo de São Jorge, em Lisboa, entre 1840 e 1846. A Calçada Portuguesa depressa se estendeu por todo o país e colónias, ultrapassando fronteiras, pois eram solicitados mestres calceteiros portugueses, trabalhadores especializados na colocação deste tipo de calçada, para executar e ensinar estes trabalhos no estrangeiro.

A Calçada Portuguesa é uma atividade repleta de história e tradição, mas a sua continuidade apresenta alguns problemas devido ao custo de manutenção das pedreiras e equipamentos, bem como pelas dificuldades ambientais que as pedreiras enfrentam nos dias de hoje.

Em termos técnicos a Calçada Portuguesa envolve duas fases: a extração da pedra e o calcetamento que abrange também o gosto artístico e a harmonia. As pedras normalmente utilizadas são o calcário, por apresentar uma superfície lisa e brilhante quando fragmentado, e o basalto. As

caraterísticas destas pedras, recorrendo a algum conhecimento de geologia, limitam muito as áreas da extração desta matéria-prima. As cores tradicionalmente usadas são o branco e o preto, embora seja possível a utilização de outras cores. Habitualmente a escala de tonalidades das pedras para a calçada é a branca, preta, cinzenta-escura; cinzenta-clara e rosa. O modo de extrair e de produção é idêntico para todas as variedades de calcário. As unidades de pedra para a pavimentação são partidas com um martelo, para a produção de calçada, com as dimensões 4/5 cm, 5/7 cm, 9/11 cm ou 12/13 cm. Para pavimentar passeios são recomendadas as de dimensões mais pequenas, já para vias de circulação de veículos são utilizadas as de maior dimensão.

A típica Calçada Portuguesa é originada pelo calcetamento com pedras de forma normalmente irregular de calcário e basalto, ou de calcário claro e calcário negro, formando padrões decorativos pelo contraste entre as pedras de cor diferente.

Os mestres calceteiros utilizam moldes para marcar as zonas de diferentes cores de modo a poderem repetir os motivos numa continuidade linear, os chamados frisos, ou nas duas dimensões do plano, denominados de padrões.

O conhecimento recente de geometria e álgebra, do século XX, demonstrou que há um número limitado de simetrias possíveis no plano. Para a pavimentação de passeios só há sete tipos distintos de frisos, mas para a pavimentação de uma grande área, podemos utilizar dezassete tipos diferentes de padrões.

A partir deste momento vamos focar mais pormenorizadamente os passeios com Calçada Portuguesa que por si promove muito interesse histórico, artístico, matemático, entre outros, para além de ser um gerador de trabalho.

Todos sabemos que um passeio é um caminho para peões que contorna uma rua e que são conhecidos desde a Antiguidade, mas que eram muito raros durante a Idade Média. Os passeios constituem uma parte importante do espaço público urbano, pois podem conter atividades comerciais em parte da sua extensão. Pela utilização, no geral, do contraste de duas cores, branco e preto, que corresponde, de certa forma, ao contorno, ou molde, do desenho considerado, equiparamos o seu estudo matemático ao grupo de frisos em que não ponderamos a cor utilizada na sua contituição, assim, os passeios revestidos com Calçada Portuguesa apresentam apenas sete tipos distintos. Se considerássemos mais cores na sua composição, estes apresentariam um número maior de tipos distintos. Por exemplo, examinando um passeio com efetivamente a combinação de duas cores ao longo da sua extensão, então teríamos cerca de trinta e um tipos diferentes de pavimentos. Com estamos tratando de Calçada Portuguesa, vamos então considerar apenas os sete tipos de frisos. Os frisos são muito utilizados na Arquitetura.

O conhecimento destes sete tipos remonta aos tempos do período Neolítico, compreendido entre 12000 a.C. e 4000 a.C. Há várias informações e alguns registos em objetos que corroboram este conhecimento. O povo do Antigo Egito também utilizava os frisos, do tipo geométrico, para ornamentar, por exemplo, as câmaras das pirâmides. Outros povos e outras civilizações, em todos os tempos, empregaram este tipo de ornamentação, e continua sendo aplicado, em larga escala, nos dias de hoje.

Vamos explorar um pouco matematicamente a pavimentação dos passeios de Calçada Portuguesa. Por ser uma construção numa só direção é evidente que há uma configuração básica, um motivo, que se repete, neste friso. Esta repetição corresponde a uma transformação geométrica denominada translação. Por sua vez, o próprio motivo pode admitir algumas simetrias, resultantes de outras isometrias, que o deixam invariante. Uma isometria é uma transformação geométrica que preserva a distância, o comprimento entre pontos. As simetrias que podem existir num motivo do friso são a simetria axial e a simetria pontual.

As simetrias axiais são provenientes das reflexões em reta. No caso dos frisos essas retas acompanham a direção do friso, ou são perpendiculares a esta direção. Podemos também ter reflexões em ponto, também conhecidas como meia-voltas, onde os pontos estão situados exatamente ao meio das extremidades que limitam o friso, e com uma distância igual a metade do comprimento do motivo.

Observamos que são considerados apenas as simetrias que para além de deixarem o motivo invariante também deixam invariante o friso, não mudando-o de direção.

Existem algumas notações de classificação para o friso. Uma das mais usuais é a dos “F”s, com índice inferior e índice superior facultativo, onde o índice inferior indica dois tipos de

transformações geométrica, a translação ou a reflexão em ponto, e a superior está relacionada com a reta de reflexão, ou a sua ausência.

Há todo um estudo geométrico e algébrico minucioso envolvendo os frisos que deleita qualquer matemático interessado no assunto. O modo como todos os elementos são cuidadosamente relacionados até ao mais pequeno detalhe faz como, mesmo o não matemático, queira descobrir mais sobre a sua origem, criação, aplicação, indo de encontro aos tópicos comuns a todos as épocas, relacionados com a expressão humana, religiosa e estética.

Como exemplo, apresentamos o passeio de Calçada Portuguesa da Rua Margarida de Chaves, em Ponta Delgada, na ilha de São Miguel, no Arquipélago dos Açores (figura 1).



Figura 1

Se observarmos bem este passeio, há um motivo básico que se repete (figura 2).



Figura 2

Por sua vez, este motivo pode ou não ter simetrias, das mencionadas anteriormente. Neste caso, não existem simetrias no motivo. Assim, neste friso notamos apenas uma translação do motivo em toda a sua extensão. Logo a sua classificação referirá apenas esta transformação geométrica, a translação, e será denotado por F_1 .

Para ilustrar melhor, consideremos agora o passeio do Largo da Matriz na mesma cidade de Ponta Delgada. (figura 3)



Figura 3

Olhando para este passeio, há novamente um motivo básico que se repete. (figura 4)



Figura 4

Já neste motivo podemos ver uma simetria pontual, como indicada na figura 5.

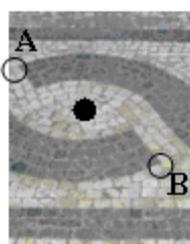


Figura 5

O ponto assinalado, com um círculo a preto, representa um centro de meia-volta, onde verificamos facilmente a reflexão por este ponto. Por exemplo, o ponto A, da figura 5, é refletido no ponto B, e vice-versa. Neste caso, o friso possui para além da translação, a reflexão em ponto, e por não possuir mais nenhuma outra simetria, nomeadamente, a de reflexão em uma reta, é classificado como F_2 . Todos os frisos com esta classificação, independentemente do motivo que apresentam, possuem apenas estas duas isometrias, a translação e a reflexão em ponto, ou meia-volta.

Para finalizar os exemplos, mostramos dois passeios que possuem simetria axial e a mesma classificação em relação aos sete tipos de frisos. Apresentamos passeios do tipo F_2^2 . Encontramos na Rua Dr. Bruno Tavares Carreiro e na Rua João Moreira, respetivamente, também em Ponta Delgada. (figura 6)



Figura 6

Não apresentamos todos os sete tipos de frisos, apenas demos alguns exemplos para dar uma ideia da vasta gama de passeios de Calçada Portuguesa que podemos obter considerando apenas os sete tipos de frisos possíveis e fazendo uso de diferentes motivos.

Como tópicos de pesquisa em diversas áreas, podemos questionar quais são os tipos, dos sete possíveis, que efetivamente são utilizados nos passeios de Calçada Portuguesa. Também podemos ir a busca da sua história; estudar quais as pedras utilizadas – calcário, basalto, ou outras; analisar o aspeto financeiro de um determinado projeto que envolva esta atividade; observar a sua aplicação ao nível da Arquitetura. Estes são alguns assuntos que podem ser desenvolvidos tendo como tema central os passeios de Calçada Portuguesa.

Não deixemos de mencionar o impacto que a Calçada Portuguesa provoca nos visitantes de outras partes do mundo. É evidente que este tipo de pavimento desperta o interesse turístico, pela sua beleza, pela sua história, pela sua execução, composição e construção. O Turismo pode tirar vantagens com roteiros que destaquem o aspeto artístico e histórico dos passeios de Calçadas Portuguesas.

Na verdade é muito mais fascinante andar sobre caminhos decorados em que a cada passo há uma descoberta de um padrão, de cores, ou até mesmo de simbolismo, dando ao nosso pensamento motivos para dar asas a nossa imaginação, do que andar por caminhos lisos, frios, sem emoção, que não nos transmitem sensações.

Todos ficamos maravilhados com o encanto de um lugar, quando este encanto pode ser absorvido por todos os nossos sentidos. Uma visão do belo, do alegre, nos torna alegres e de espírito aberto. Esta é um pouco das sensações transmitidas por quem passeia sobre um passeio de Calçada Portuguesa que, apesar de serem contruídos com pedras brancas e pedras pretas, nos transmitem um conforto secular.

Conclusão

Ao olhar para os trabalhos de Calçada Portuguesa absorvemos toda uma cultura, toda uma história, toda uma arte, numa simples e modesta coligação de pequenas pedras toscas e irregulares, com apenas dois tons, que unidas dão corpo a imaginação do ser humano. Sejam geométricos ou florais, figurativos ou invocativos, os desenhos colocados no chão são relatos silenciosos de episódios locais. Os passeios de Calçada Portuguesa têm esta característica, de levarmos a um tempo sem tempo, envolvendo-nos em muitas áreas do saber.

Referências:

- Ledergerber-Ruoff, Erika Brigitta. *Isometrias e Ornamentos no Plano Euclidiano*. Atual Editora, 1982.
- Matos, Ernesto. *Calçada portuguesa nos Açores: entre o céu e o horizonte / The portuguese stone pavement at the Azores: between the sky and the horizon*. Terceira, BLU edições, 2006.
- Alsina, C e Trillas E. *Lecciones de Algebra y Geometria*, curso para estudantes de Arquitectura. Editorial Gustavo Gili, S.A., 1984.

STUDY OF A THREE-AXIS PIEZORESISTIVE ACCELEROMETER WITH UNIFORM AXIAL SENSITIVITIES

Abdelkader Benichou, PhD Candidate

Nasreddine Benmoussa, PhD

Kherreddine Ghaffour, PhD

University of Tlemcen/Unit of Materials and Renewable Energies URMER, Algeria

Abstract:

Micro-electro-mechanical systems (MEMS) for automotive industry and biomedical applications (BioMEMS) have the fastest growth rate within the MEMS market. The Microsystems job market imposes to research laboratories and universities to respond by increasing the quality of MEMS engineering and informatics interdisciplinary training programs. In this fact, our work consists to study and develop a three-axis piezoresistive accelerometer having uniform sensitivities along to three axes. This sensor which is made of a heavy proof mass and four long beams, allow us to obtain high sensitivities, by reducing the resonant frequencies. Uniform axial sensitivities, with a transverse sensitivity, could be obtained using a three-axis sensor. The stress analysis of this sensor was performed in order to determine the positions of the piezoresistances, in the four flexure beams.

Key Words: Accelerometer, MEMS, Piezoresistance, Simulation

Introduction

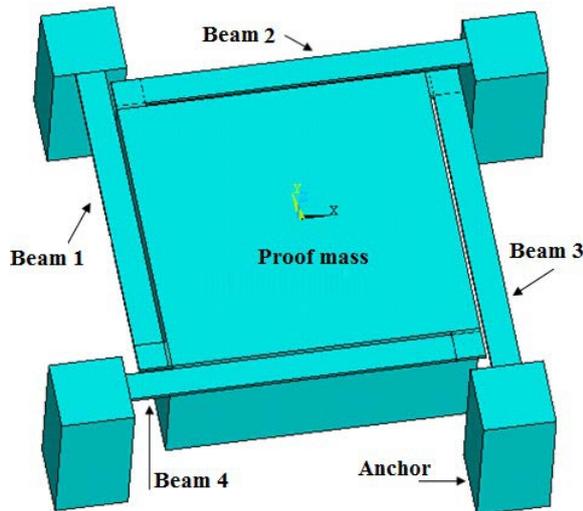
The micro-machined inertial sensors which are composed of accelerometers and gyroscopes have a significant percentage of sensors containing silicon. They can be found mainly in the automotive industry, the biomedical applications, the household electronics, robotics, vibration analysis systems, navigation systems. There are various techniques to transform the action of acceleration on the sensor into electric signal. These techniques are based on principles capacitive, piezoresistive, piezoelectric, and ect... The concept of accelerometer is not new, but its fabrication offers new market opportunities to Microsystems manufacturers, so the MEMS market has motivated continuous researches in this kind of sensors in order to minimize the size and to improve the performance.

As we know, the realistic applications create an enormous motivation for research on sensors MEMS, especially accelerometer. In this modern world, the applications require new sensors with a smaller size and high performances. In practice, rare are research which can provide an effective and complete methodology for the design of accelerometers.

The proposed of three-axis accelerometer

The three-axis accelerometer always requires small cross-axial acceleration, high and linear sensitivity. We proposed a flexure configuration that is show in figure (1) in order to meet these critical characteristics.

The parameters of the structure brought to the FEM process is shown in table (1).



Parameters	Size (Length, Wide, Thickness)
Proof mass	845x845x400 μm^3
Beams	975x80x10 μm^3
Anchors	200x200x200 μm^3
Global structure	1.5x1.5x0.5 mm^3

Figure 1: Structure of three-axis piezoresistive accelerometer.

Table 1: Parameters of structure.

When an external acceleration is applied to the sensor, the proof mass is deviated. The vertical component (A_z) of acceleration causes a vertical displacement of the mass. The second type of movement is caused by transverse accelerations (A_x and A_y). The deviation of the proof mass causes a variation of the stress on four surfaces of the beams. This can be measured by twelve p-type and n-type piezoresistances diffused and assembled by three Wheatstone bridge circuits.

These piezoresistances were aligned with the crystal direction $\langle 110 \rangle$ and $\langle \bar{1}\bar{1}0 \rangle$ of silicon (100). In silicon material, there are only three independent piezoresistive coefficients π_{11} , π_{12} and π_{44} . The longitudinal piezoresistive coefficient π_l is defined in the case the stress is parallel to the electric field. Similarly, the transverse piezoresistances coefficient π_t is defined in the case the stress is perpendicular to the electric field.

For the orientations $\langle 110 \rangle$ and $\langle \bar{1}\bar{1}0 \rangle$ of silicon (100), these coefficients can be expressed as follows:

$$\pi_l = \frac{1}{2}(\pi_{11} + \pi_{12} + \pi_{44}) \quad \pi_t = \frac{1}{2}(\pi_{11} + \pi_{12} - \pi_{44}) \quad (1) \quad (2)$$

Design and simulation using ANSYS

The finite element method (FEM) is applied to perform analyses of the stress distribution in the flexure beams. Considering the stress distribution, the piezoresistances are placed in order to eliminate the transverse sensitivities and to obtain maximum sensitivities for the three components of acceleration. The finite elements model of the sensing structure was analyzed by using software ANSYS. The boundary condition, by considering the fixed anchor, and the free proof mass charged in the medium by acceleration in the form of a force was applied.

Figure (2) shows the generation of mesh for the analysis by the finite element method.

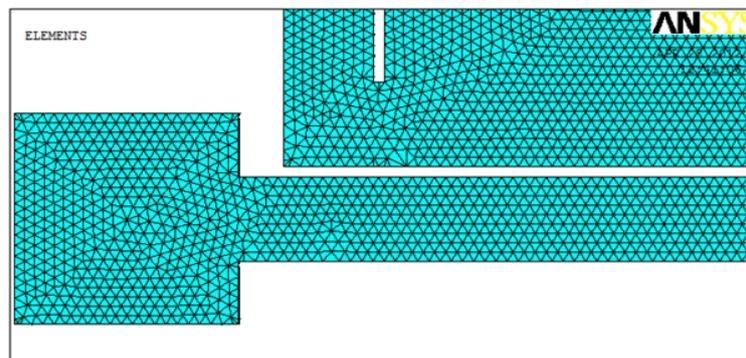


Figure 2: The mesh generation of the FEM model.

The stress distribution on the surface of the beams, caused by the A_z component of acceleration, is shown in figure (3). The principle of detection of the sensor is based on the characteristic of the p-type and n-type piezoresistances. The n-type piezoresistances decreases when the sensor is exerted by a tensile stress and contrary in the case of a p-type piezoresistances.

The figures (4.a) and (4.b) shows the stress analysis results along 1st and the 3rd beams when the sensor is exerted to a force caused by acceleration along axis z. From these figures, we can find the optimal locations for the piezoresistances Wheatstone bridge of component A_z .

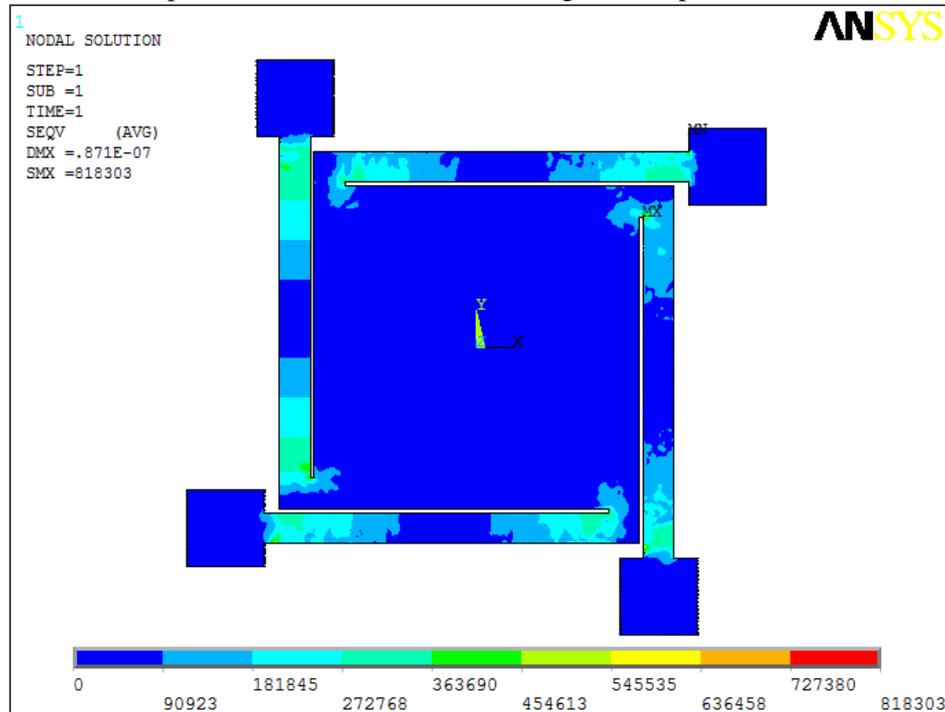


Figure 3: The stress distribution on the beams caused by the 1g acceleration A_z .

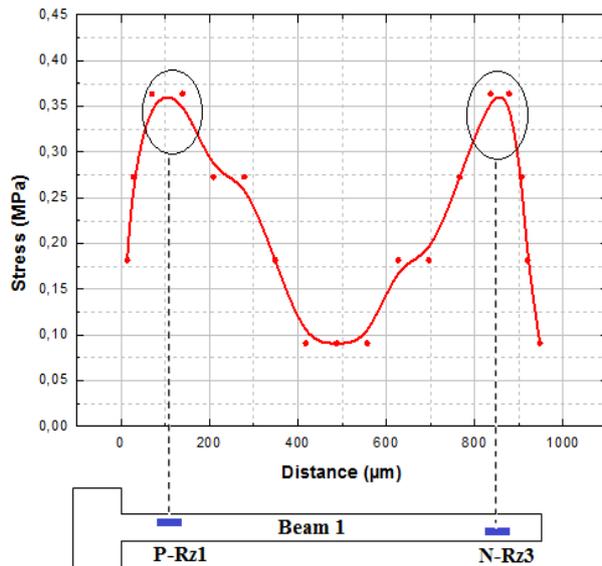


Figure 4.a

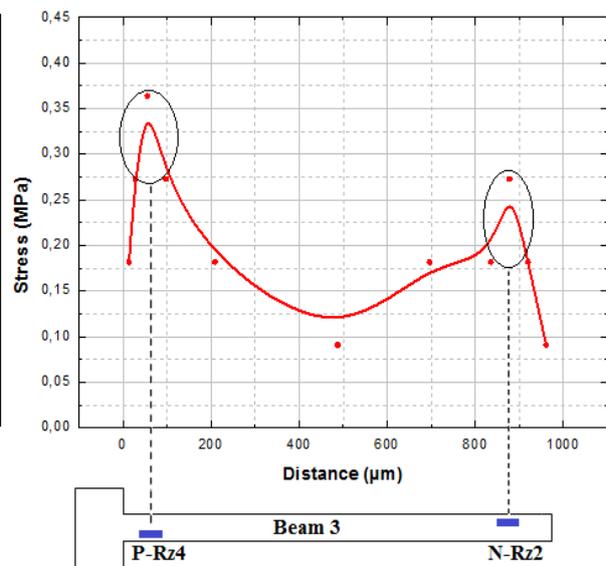


Figure 4.b

Figure 4: Longitudinal stresses on the surface of the 1st and 3rd beams due to 1g acceleration A_z .

By the same method, the acceleration components A_x and A_y can be detected by using four piezoresistances on 2nd and 4th beams of the Wheatstone bridge of component A_x , and four piezoresistances on 1st and 3rd beams of the Wheatstone bridge of component A_y . The positions of the piezoresistances are also indicated on the figures (5.a) and (5.b).

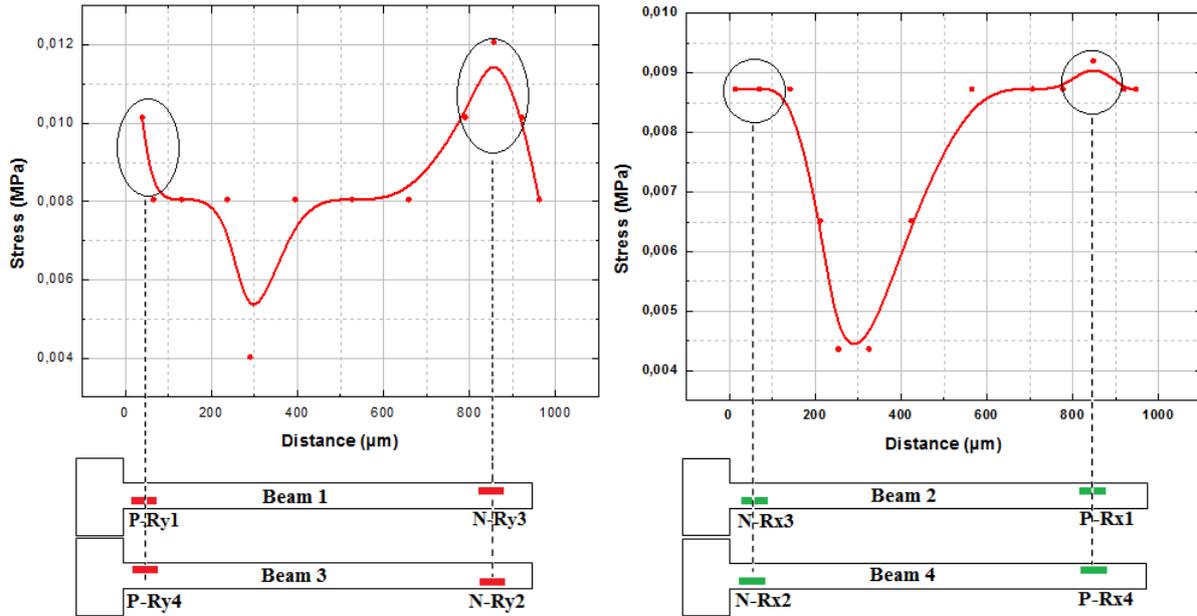


Figure 5.a

Figure 5.b

Figure 5: Longitudinal stresses on the surface of the 2nd and 4th beams due to 1g acceleration A_x and A_y .

From simulation results, we would find that two normal stresses are rather smaller when comparing to the longitudinal stress σ_l . The total resistance change is given by the following equation:

$$\frac{\Delta R_i}{R_i} = G_l \cdot \varepsilon = G_l \cdot \frac{\sigma_l^i}{E} \quad i = x, y, z \quad (3)$$

Where G_l is a longitudinal gauge factor: $G_l = \pi_l \cdot E + 1 + 2 \cdot \nu$ (4)

E is a Young's modulus; ν is a Poisson ratio and ε is the tensile strain.

In the equation (4), we have: $\pi_l \cdot E \gg 1 + 2 \cdot \nu$

$$\frac{\Delta R_i}{R_i} \approx \pi_l \cdot \sigma_l^i$$

Thus the equation (3) becomes:

$$\text{The electronics sensitivity can be given by: } S_i = \frac{V_{out}}{a_i} = \frac{\Delta R_i}{R_i} \cdot V_{in} = \pi_l \cdot \sigma_l^i \cdot V_{in} \quad (6)$$

Where S_i is the sensitivity to the i^{th} acceleration component, V_{in} and V_{out} are the input and output voltage, respectively.

The following figure gives the Wheatstone bridge circuits for the three acceleration components.

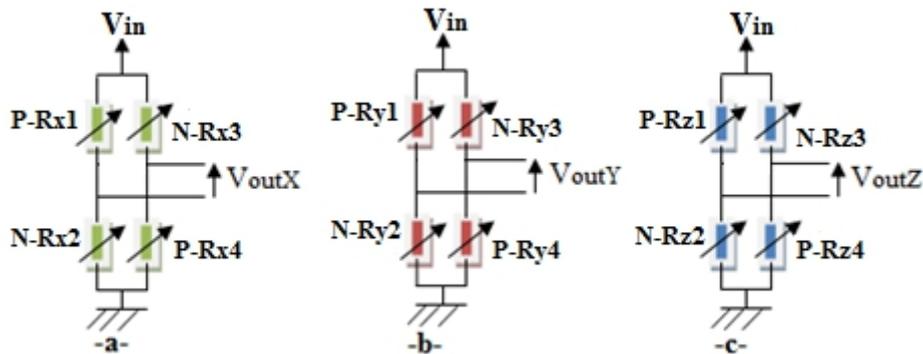


Figure 6: Wheatstone bridge circuits of three acceleration components.

Table (2) gives the increase (+), the decrease (-), or the invariance (0) of the piezoresistances by the application of the A_x , A_y and A_z components of acceleration.

	P-R _{Z1}	N-R _{Z2}	N-R _{Z3}	P-R _{Z4}	P-R _{Y1}	N-R _{Y2}	N-R _{Y3}	P-R _{Y4}	P-R _{X1}	N-R _{X2}	N-R _{X3}	P-R _{X4}
A_z	+	-	-	+	+	-	-	+	+	-	-	+
A_y	0	0	0	0	+	-	-	+	0	0	0	0
A_x	0	0	0	0	0	0	0	0	+	-	-	+

Table 2: Piezoresistance values changes of three acceleration components.

Conclusion

This work presents a design and simulation of three-axis piezoresistive accelerometer using MEMS technology. The sensing principle of the sensor is the piezoresistive effect. The most important aspect of Finite Element Analysis (FEA) in our design process is the analysis of the stress distribution in the four flexure beams. The stress analysis was performed in order to determine the positions of the piezoresistances on these beams, and consequently to eliminate the transverse sensitivities for obtaining optimal three acceleration components. This model of three-axis accelerometer is used in the field of biomedical applications (BioMEMS).

References:

- [1] EDITORS-IN-CHIEF; YOGESH B. GIANCHANDANI, O. TABATA, H. ZAPPE, "Comprehensive Microsystems" volume one, Copyright © 2008 Elsevier.
- [2] N. MALUF, K. WILLIAMS, "An Introduction to Microelectromechanical Systems Engineering" Second Edition, Copyright © 2004 ARTECH HOUSE, INC.
- [3] S. BEEBY, G. ENSELL, M. KRAFT, N. WHITE, "MEMS Mechanical Sensors", Copyright © 2004 ARTECH HOUSE, INC.
- [4] N. TIEN ANH, T. DUC TAN, "A Tree-Axis Piezoresistive Accelerometer with Uniform Axial Sensitivities" Cau Giay, Ha Noi, Viet Nam, Copyright © 2011 IEEE.
- [5] Y. KANDA, "Piezoresistance Effect of Silicon", Sensors and Actuators, 1991, pp.83.
- [6] Y. KANDA, "A Graphical Representation of the Piezoresistance Coefficients in Silicon", IEEE Trans. On electron Devices, Vol.29, n°1, 1982.

SETTING UP A DATA ACQUISITION SYSTEM FOR SPARK ENGINES

Ahmed H. Osman, Assist. Prof., PhD

Zagazig University, Egypt

Magdy. M. Massoud, Assist. Prof., PhD

Prince Sultan University, KSA

Abstract:

Measuring dynamic parameters on an engine is very important to diagnose and analysis the faulty problems and the quality of it. To achieve such measurements, it requires expensive data acquisition system (DAS) capable of measuring these parameters. The problem arises when there is a need to build such a system from scratch with a very limited budget and depending on your experience. This paper is dedicated to transfer experience and technology to mechanical engineers who have a little experience in setting up their dynamic test rigs and they do not have enough budgets to buy an expensive DAS [1].

The proposed system is intended to set up a low cost and configurable DAS consisting of sensors, analog to digital card A/D and software package for acquisition and processing.

The sensors are used to measure cylinder pressure, vibration of the cylinder block on 3D, different types of temperature (oil, inlet and outlet water temperature), crank angle and engine's RPM. The processing includes cycle separation to prepare the acquired data for heat release analysis. The Fast Fourier Transformation (FFT) is also used to analyze the data for fault diagnosis using vibration analysis. The developed system saves about 75% of the price of a similar on the shelf system.

Key Words: Data Acquisition System, Frequency Domain, Spark Ignition, Time Domain

Introduction

Data acquisition technology has taken giant leaps forward over the last 30 to 40 years [2]. Vehicle data acquisition covers a wide range of applications. This may be a simple data acquisition system in the form of a test process, testing the engine temperature of a new automobile design. However, vehicle data acquisition does not mean automobiles alone, but can include, race cars, buses, trucks, motorcycles, aircraft, boats, and other transportation. Vehicle data acquisition is not only measurement of gas or diesel engine parameters but, with using proper sensors and signal conditioning, it could be extend to display, report and control any other vehicle parameters.

The vibratory and acoustic behavior of the internal combustion engine is a highly complex one, consisting of many components that are subject to loads that vary greatly in magnitude and which operate at a wide range of speeds.

The acquisition of the cylinder pressure and its evaluation is one of the most powerful means for experimental optimization of internal combustion engines [3]. Behavior of engine working cycle may be investigated using this technique. Conditions influencing mechanical and thermal load of engine part may be determined as well. The quantification of engine running roughness is also a very useful result of evaluation of in-cylinder pressure records. The use of in-cylinder pressure record is extremely powerful tool as far as experimental data are confronted with results of mathematical model of engine working cycle. For production, testing or even prototype evaluation of vehicles (cars, trucks, etc.) the data acquisition requirements can be very demanding. This is made obvious from the fact that the types of parameters that would need to be measured include: Temperature (water, oil, transmission fluid, brakes, interior cooling system, and exhaust), Pressure (engine compression, tire, cooling system, and oil), Vibration (chassis, engine, doors, hood, trunk, and interior), Voltage (battery, lamps, relays, electric motors, fans, pumps, and instrument panel), Flow (water, oil, transmission fluid, coolant, and air). A vehicle data acquisition system might include hundreds of data

points

Basic Operation

The four steps of data acquisition are:

- Measure
- Record
- Access
- Analysis

Different types of Sensors represent the main key to achieve the measuring step, in which it measures the different aspect of parameters such as electronic signal, temperature, pressure, acceleration, linear movement, frequency and ON/OFF switch position which recorded by the data acquisition unit [4].

Recording is achieved throughout converting the sensor measurements into digital signal by A/D converter prior to logging. The sampled rates refer to the number of times per second the data is recorded i.e. 100Hz = 100 samples per second.

Access of the number of monitored channels and the sampling rate determine how long it takes to full or overwrite the data log memory, the download time and the amount of time to analyze the logged data. A higher hertz (Hz) rate creates more data. Once the data has been recorded it then needs to be accessed for analyzing and archiving.

Analysis, a software analysis program is used to process the data and prepare it in the required form. Finally one has to keep in mind that not all signals need to be recorded at a fast rate;

Data acquisition construction

A typical data acquisition system consists of these components, as shown in figure 1.

The data acquisition includes 3 phases as follows:

Input Phase

a) Physical Phenomenon (Pressure, Vibration etc..)

b) Sensors: A transducer is a device that converts input energy of one form into output energy of another form. For example, a microphone is a sensor that converts sound energy (in the form of pressure) into electrical energy, while a loudspeaker is an actuator that converts electrical energy into sound energy.

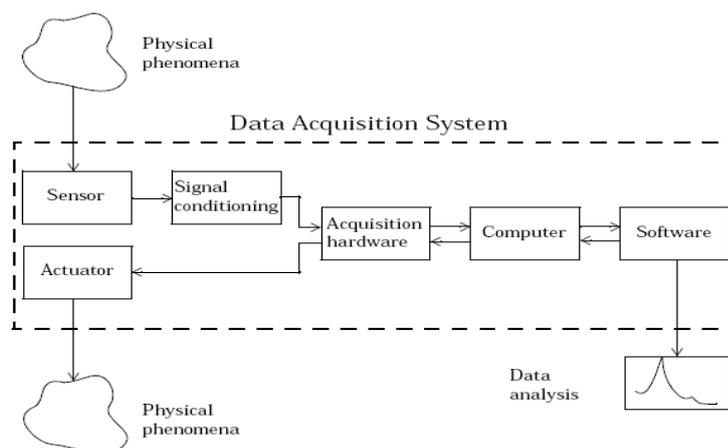


Figure 1: The data acquisition components

c) Signal Conditioning : Sensor signals are often incompatible with data acquisition hardware. To overcome this incompatibility, the signal must be conditioned. For example, you may need to condition an input signal by amplifying it or by removing unwanted frequency components. Output signals may need conditioning as well [5].

d) A/D Card: At the heart of any data acquisition system lays the data acquisition hardware. The main function of this hardware is to convert analog signals to digital signals, and to convert digital signals to analog signals.

e) Screw terminal (to connect wires)

Processing Phase

- a) Computer
- b) Software: Data acquisition software allows you to exchange information between the computer and the hardware. For example, typical software allows you to configure the sampling rate of your board, and acquire a predefined amount of data.

Output Phase

- a) Actuators (if there is a control system)
- b) Monitor
- c) Printed report

Hardware**Channel Types**

- Analog - Any input that continuously changes within a specified range. This type of input would traditionally have been displayed with a needle gauge.
- Digital - Any input or output which can only be in two states. e.g. ON or OFF, true or false, above or below.
- Frequency - Any input that is pulsed or oscillating. The rate at which the input is pulsing or oscillating is measured.
- Serial - Any device or sensor which is semi-intelligent and is capable of communicating via RS-232 serial communication.

Sensor Types**Table 1** Show various types of sensors' output

Analog	Digital	Frequency	RS232
Temperature	Switches	RPM	Gas analyzer
Pressure	in/out	Speed	
Flow	Status	Rotary encoder	
Current	Counter		
Volts	on/off		

What do you need to know?

- Resolution and range (4.1 & 4.2).
- How fast to sample (4.3).
- How many times to sample.
- Device and configuration (what is the maximum number).
- Connecting the signals the right way.
- What channels to sample.
- How to deal with the data.
- Data bus (type of slot inside computer e.g. PCI or external such as USB [6]).

Analog-to-Digital Conversion

- The A/D converter (ADC) converts an analog voltage into a binary number through the process of quantization.
- The ADC will have a full-scale voltage range over which it can operate.
- The number of bits will dictate how many discrete levels will be used to represent measured voltages. For example, an 8-bit converter with a full-scale voltage of 10 V will give you a resolution of 10V/256 which is 39.1 mV. Figure 2 shows a comparison between 16 bit card resolution and 3 bit. [7]

Signals' Resolution

$$\Delta = \frac{V_{FS}}{2^n} \left\{ \begin{array}{l} \frac{10V}{2^3} = 1.25V \\ \frac{10V}{2^{16}} = 0.15mV \end{array} \right.$$

Where V_{FS} is the full scale of the input signal

N is the card's resolution (12, 14 or 16 etc.)

Δ is the signal's resolution

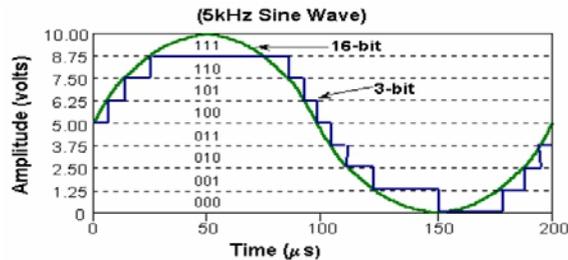


Figure 2 shows a comparison between 16 bit card resolution and 3 bit

Choosing a sampling or scan rate (scans/sec, or Hz)

The ADC samples according to a scan rate. How fast you sample should satisfy the Nyquist sampling theorem. The sampling frequency should be at least two times the highest frequency present in the signal. If Nyquist criterion is not satisfying there will be implications in how the signal is 'reconstructed'. Depending on your objective, you might choose **scan rate** to satisfy Nyquist criterion but, you might also want to have accuracy in **time** measurements. Can you see how you have to balance how fast you sample, how many samples you get, etc.?

Types of Measurement Systems:

1. Differential measurement system
2. Referenced single-ended (**RSE**)
3. Non-referenced single-ended (**NRSE**)

You may see these connection options on A/D hardware.

What Hidden things you have to take care about? [8],[9]

Sensors

When choosing the best analog sensor to use, you must match the characteristics of the physical variable you are measuring with the characteristics of the sensor. The two most important sensor characteristics are sensor's output & bandwidth

Sensor Output

The output from a sensor can be an analog signal or a digital signal, and the output variable is usually a voltage although some sensors output current.

Current Signals

Current is often used to transmit signals in noisy environments because it is much less affected by environmental noise. The full scale range of the current signal is often either 4-20mA or 0-20mA. A 4-20 mA signal has the advantage that even at minimum signal value, there should be a detectable current flowing. The absence of this indicates a wiring problem. You need to use this type if your measuring system is far from your control room.

Voltage Signals

The most commonly interfaced signal is a voltage signal. For example, thermocouples, strain gauges, and accelerometers all produce voltage signals. There are three major aspects of a voltage signal that you need to consider:

Amplitude

If the signal is smaller than a few millivolts, you may need to amplify it

Frequency

Whenever you acquire data, you should decide the highest frequency you want to measure. The highest frequency component of the signal determines how often you should sample the input. If you have more than one input, but only one analog input subsystem, then the overall sampling rate goes up in proportion to the number of inputs. Higher frequencies may be present as noise, which you can remove by filtering the signal before it is digitized. If you sample the input signal at least twice as fast as the highest frequency component, then that signal will be uniquely characterized. However, this rate may not mimic the waveform very closely.

Duration

How long do you want to sample the signal for? Data stored in ASCII format takes more space than data stored in binary format.

Sensor Bandwidth

In a real-world data acquisition experiment, the physical phenomena you are measuring have some limitation. The bandwidth is given by the range of frequencies present in the signal being measured. You can also think of bandwidth as being related to the rate of change of the signal. A slowly varying signal has a low bandwidth, while a rapidly varying signal has a high bandwidth. To properly measure the physical phenomena of interest, the sensor bandwidth must be compatible with the measurement bandwidth. You may want to use sensors with the widest possible bandwidth when making any physical measurement. This is the one way to ensure that the basic measurement system is capable of responding linearly over the full range of interest. However, the wider the bandwidth of the sensor, the more you must be concerned with eliminating sensor response to unwanted frequency components [4].

Signal Conditioning

Sensor signals are often incompatible with data acquisition hardware. To overcome this incompatibility, the sensor signal must be conditioned. The type of signal conditioning required depends on the sensor you are using. For example, a signal may have small amplitude and require amplification, or it may contain unwanted frequency components and require filtering. Common ways to condition signals include: Amplification, Filtering, Electrical Isolation, Multiplexing, and source Excitation [7].

Current DAS Application

The measuring system consists of many sensors and actuator which are shown in the figures (3-13). The sensors are RPM sensor, Oil pressure sensor, Cylinder head temperature sensor, Lambda (Oxygen) sensor, Inlet radiator temperature, Outlet radiator temperature, Dynamometer manual load, Cylinder pressure, Rotary Encoder (crank angle), and Vibration. The whole system is shown in figure 14.



Figure 3: RPM Sensor



Figure 4: Oil Pressure Sensor

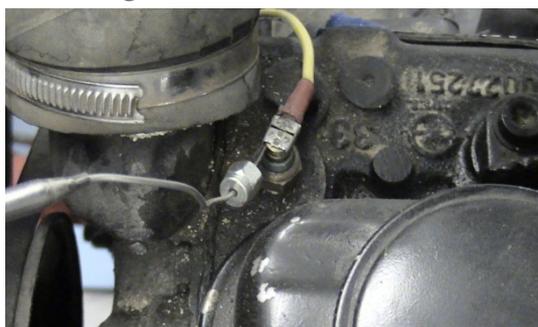


Figure 5: Cylinder Head Temperature Sensor



Figure 6: Oxygen Sensor

Software

The data acquisition software is developed using "TestPoint" [10].

TestPoint: It's a software package that supports acquisition from A/D boards and RS232 devices, and combines this with analysis and display capability. It provides object-oriented graphic style of creating custom test, measurement, and data acquisition programs.

By dragging and dropping *objects*, which are icons representing related tasks, to an *action list* in the desired order of execution, it is easy to create applications quickly without any complicated programming figure 15.



Figure 7: Inlet Radiator Temperature sensor



Figure 8: Outlet Radiator Temperature Sensor



Figure 9: Manual load control actuator



Figure 10: Cylinder pressure sensor



Figure 11: Rotary encoder (Crank Angle) Sensor and Dynamometer



Figure 12: 3D vibration Sensor

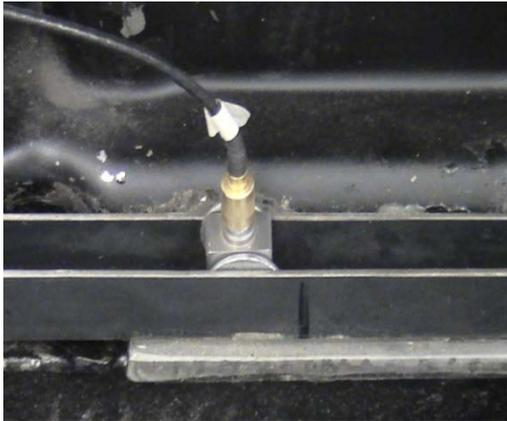


Figure 13: Installation of the Vibration Sensor



Figure 14: The whole system

Professional looking applications are created by simply placing the input, output, display, and control-button objects where you want them on the display window. Libraries for many popular instruments including Hewlett Packard oscilloscopes, spectrum analyzers etc. are supported. Data acquisition boards and GPIB cards are accessed through the same drag and drop action. During application development or test, a handy demo mode is available, allowing application development without the actual presence of hardware. Software package is necessary to process the acquired data [11]. A/D card without software is useless. A/D card with poor software is almost useless. A/D uses driver software which is the layer between the operating system and the hardware. It directly controls the registers of the A/D card, managing its operation and its integration with the computer resources, such as processor interrupts, Direct Memory Access DMA, and random memory [12].

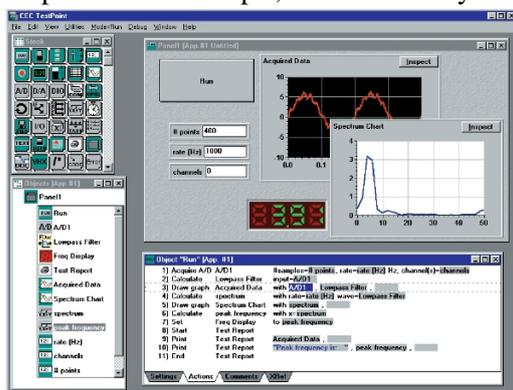


Figure 15: Interface architecture of the Software package

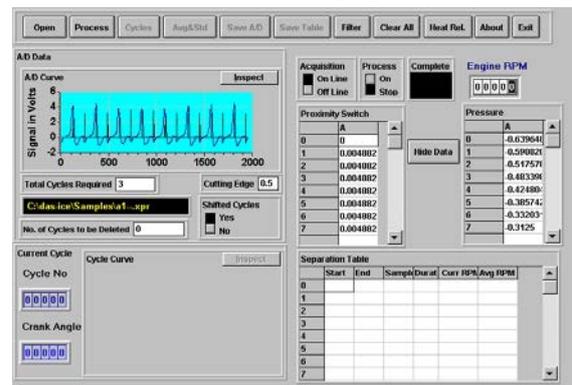


Figure 16: A typical DAS-ICE display showing the main menu system

Features of the developed package

Data Acquisition System for Internal Combustion Engine (DAS-ICE) provides an extensive set of data analysis tools, graphics, control and application interface development capabilities as well as the ability to acquire data from A/D. It runs in windows environment to provide the ability to tightly control timing and perform multi-tasking functions. Multitasking allows different parts of an application to run in parallel without complicated managers or disk-swapping methods. The synchronization of the acquired data (trigger) is very important (Top dead center of the acquired cylinder) [13]. The main interface of the application is shown in figure 16.

Measurement

Figures (17-18) illustrate a sample of the measurement results for the ignition cycles and the processed data to get average cycle. Multiply the magnitude of pressure by the sensor factor to get the physical real value in Pascal. Figures (19-20) illustrate a sample of processing the vibration signals in the frequency domain.

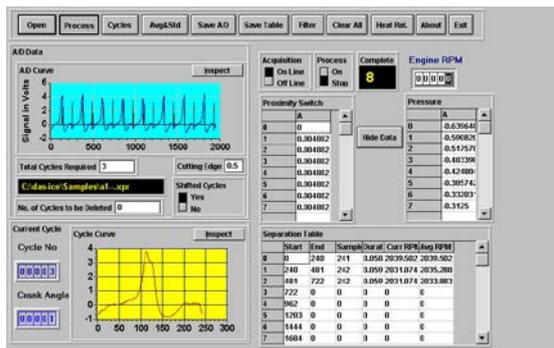


Figure 17: Acquiring 8 complete ignition cycles according to user request

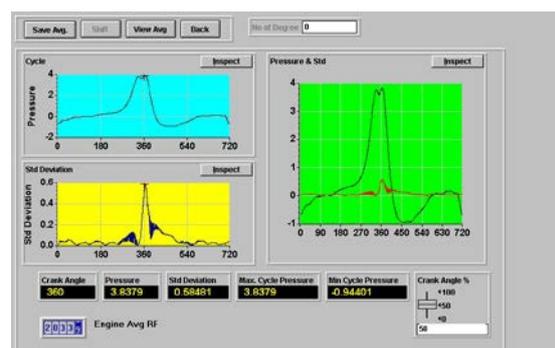


Figure 18: Processing the acquired data to get the average cycle

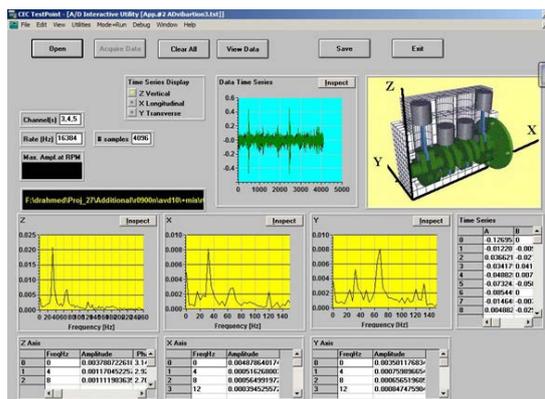


Figure 19: Acquiring vibration signals On the engine & Performing FFT

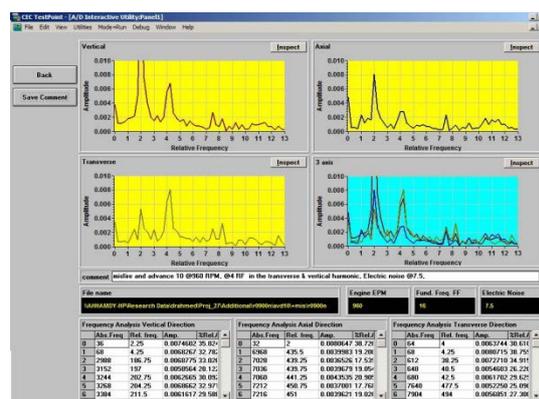


Figure 20: Processing the acquired data of the vibration signals

Conclusion

A configurable data acquisition system is built to acquire data from internal combustion engines. The system describes the basic steps for mechanical engineers who need to design their dynamic measuring system. The system is operational and capable of acting as a test bed for a wide variety of experimental investigation on Petrol engine.

The system takes into account different types of signals such as cylinder pressure, vibration in 3D, Oil pressure, crank angle, and engine RPM. Inlet water temperature to radiator, outlet temperature from radiator, oil temperature, and cylinder head temperature are also captured.

The system processes the acquired data to prepare the average cycle during the 4 strokes as a function of a synchronized crank angle. The vibration signals in time domain are converted into frequency domain using FFT. Beside online mode, the system can also process the data using offline mode. A digital filter is included in the system to filter the signals from noise. The developed system saves about 75% of the cost of a similar on the shelf system.

References:

- [1] P. Strachan, A. Oldroyd and M. Stickland: Introducing Instrumentation and Data Acquisition to Mechanical Engineers using LabVIEW, Int. J. Engng Ed. Vol. 16, No. 4, pp. 315-326, 2000.
- [2] "Selecting the right data acquisition system Application Note 1412", Agilent Technology.
- [3] Michal Takáts: In-Cylinder Pressure Recording and Data Acquisition System, Josef Bozek Research Centre of Engine and Automotive Technology, Czech Technical University in Prague, 2002.
- [4] Bentley, John P.: Principles of Measurement Systems, Second Edition; Longman Scientific and Technical, Harlow, Essex, UK, 1988.
- [5] "NI Signal Conditioning (SCXI) Hardware 2005", www.ni.com/signalconditioning.
- [6] John Park, Steve Mackay: Practical Data Acquisition for Instrumentation and Control Systems, Elsevier, 2003.
- [7] "NI Data Acquisition (A/D) Hardware 2005", www.ni.com/dataacquisition.

- [8] Ahmed H. Osman: **Data Acquisition System for Internal Combustion Engine**, Technical Manual V.2.0, Cairo, Egypt, 1998.
- [9] Ahmed H. Osman: **Multi-Function General Purpose Data Acquisition System**, Technical Manual V.2.1, Cairo, Egypt, 2002.
- [10] Testpoint Reference Manual.
- [11] “Engine Combustion Pressure Analysis Package for DL700 and WE7000”, Yokogawa.
- [12] Howard Austerlitz: **Data Acquisition Techniques Using PCs**, Second edition, Academic press, 2003.
- [13] LUBOMÍR MIKLÁNEK: **TDC Determination of IC Engine cranked by starter**, Josef Božek Research Centre, Czech Technical University in Prague, Czech Republic.

PLANNING OF SMALL HYDROPOWER PLANTS AND KEZER RIVER

Fevzi Onen

Mahmut Eminoglu

Dicle University, Eng. Faculty, Civil Eng. Department, Diyarbakir, Turkey

Abstract:

Investments of Hydroelectric Power Plants have been grown up and our country has become like a construction site by entering into force of the law concerning the transfer of the right of water use in Turkey since 2001. Turkey has a gross annual hydropotential of 433.000 GWh, which is almost 1% of world total potential. Its share is about 16% of the total hydropower capacity in Europe. This situation has increased the importance of planning and arousing interest of academic levels of hydroelectric power plant (HEP) and small hydropower plant (SHEP). Small-scale micro hydropower is both an efficient and reliable form of energy. The first small hydroelectric plant (SHEP) with a capacity of 88 kW was installed in Turkey in 1902. In this study, a concrete example of a not accumulated hydroelectric facility planning is discussed by aiming of consideration of Kezer River in Turkey hydroelectric energy potential. Finally Installed power and energy values of the total annual flow have been calculated using discharge, net head and turbine efficiency.

Key Words: Renewable energy, Small hydropower, Kezer river

Introduction

Energy plays a vital role in the socio-economic development and in raising the standard of living. Turkey is a rapidly growing country where both its population and economy are expanding each year, resulting in a corresponding increase in its energy demand. This increasing demand has to be met to keep a sustainable development in the economy and to raise the living conditions of the people. Although Turkey has many energy sources, it is a big energy importer. Turkey has a lot of potential to supply its own energy, which could be harnessed in order to avoid this energy dependency. Additionally, Turkey is a country with an abundance of renewable energy sources and can essentially provide all energy requirements from its own indigenous energy sources [1]. Table 1 shows Turkey's present and future energy consumption (taken from Ministry of Energy and Natural Resources of Turkey) with respect to various resources [2].

Clean, domestic and renewable energy is commonly accepted as the key for future life, not only for Turkey but also for the entire world. This is primarily because renewable energy resources have some advantages when compared to fossil fuels. Turkey has to adopt new longterm energy strategies to reduce the share of fossil fuels in the energy consumption [3].

Turkey has a total hydropower potential of 433 TWh that accounts for almost 1.1% of the total hydropower potential of the world and for 13.75% of European hydropower potential. Only 130 TWh of the total hydroelectric potential of Turkey can be used economically. By the commissioning of new hydropower plants, which are under construction, 43% of the economically usable potential of the country would be exploited. At present the hydropower energy is an important energy source for Turkey due to its useful characteristics such as being renewable, clean, and less impactful on the environment, as well as being a cheap and domestic energy source [4].

Turkey is divided into 25 drainage basins in regard to the concept of hydrology, as given in Table 2 [5]. Annual average flows of these basins are approximately 186 km³. The Euphrates and Tigris, among all basins, is the largest with about 185,000 km² of land area. Most of country's water potential lies in the southeast region (Euphrates and Tigris basin) with 28.5%, followed by Black Sea region with 13.3%, Mediterranean region with 10.8%, Marmara region with 4.5% and others.

Table 1. Turkey's final energy consumption by resources.

Type of Resources	2006	2010	2015	2020
Hard coal	14,721	17,282	26,884	48,156
Lignite	11,188	18,001	24,190	32,044
Asphaltite	259	301	301	301
Oil	32,551	41,184	50,420	60,918
Natural Gas	28,867	37,192	44,747	51,536
Nuclear	0	0	8229	8229
Wind	11	421	571	721
Solar	403	495	605	862
Fuelwood	4023	3383	3075	3075
Animal and Vegetable residues	1146	1034	926	850
Geothermal (Heat)	1081	1750	2836	4584
Geothermal (Electricity)	330	330	330	330
Hydraulic	3556	4903	7060	9419
Growth rates (%)		29	35	31

Table 2. Turkey's drainage basins [6].

No	Name of basin	Rainfall area (km ²)	Annual average flow (km ³)	Share in total (%)	Annual average yield (l/s/km ²)
1	Meric-Ergene	14,560	1.33	0.7	2.9
2	Marmara	24,100	8.33	4.5	11
3	Susurluk	22,399	5.43	2.9	7.2
4	K. Ege	10,003	2.09	1.1	7.4
5	Gediz	18,000	1.95	1.1	3.6
6	K. Menderes	6907	1.19	0.6	5.3
7	B. Menderes	24,976	3.03	1.6	3.9
8	W.Mediterranean	20,953	8.93	4.8	12.4
9	Antalya	19,577	11.06	5.9	24.2
10	Burdur	6374	0.5	0.3	1.8
11	Akarcay	7605	0.49	0.3	1.9
12	Sakarya	58,160	6.40	3.4	3.6
13	W.Black Sea	29,598	9.93	5.3	10.6
14	Yesilirmak	36,114	5.8	3.1	5.1
15	Kizilirmak	78,180	6.48	3.5	2.6
16	Konya(closed)	53,850	4.52	2.4	2.5
17	E.Mediterranean	22,048	11.07	6	15.6
18	Seyhan	20,450	8.01	4.3	12.3
19	Asi	7796	1.17	0.6	3.4
20	Ceyhan	21,982	7.18	3.9	10.7
21	Euphrates-Tigris	184,918	52.94	28.5	21.4
22	E.Black Sea	24,077	14.90	8	19.5
23	Çoruh	19,872	6.30	3.4	10.1
24	Aras	27,548	4.63	2.5	5.3
25	Van lake	19,405	2.39	1.3	5
	Total	779,452	186.05	100	

The obvious benefits of hydropower projects in Turkey or in any other country where hydropower potential exists abundantly, is associated with the generation of electrical power, which has the ability to both assist the sustainable economical development and increase the quality of life. Furthermore, they are labor-intensive during construction, as well as providing long term employment opportunities [7]. Another benefit of exploiting water resources is about environmental concern. Because it is a renewable, clean and green energy source, it is less harmful than fossil fuel sources causing dangerous gas emissions.

Small, mini and micro-hydro plants (usually defined as plants less than 10 MW, 2 MW and 100 kW, respectively) play a key role in many countries for rural electrification. Small scale hydro is mainly 'run of river,' so does not involve the construction of large dams and reservoirs. Therefore there have been minimal and ignorable environmental problems with small scaled hydro plants in comparison with those of HEPs with large dams. In medium head (5m < head < 15 m) or high head (head > 15 m) installations, water is carried to the forebay by a small canal. Low head installations (head < 5m) generally involve water entering the turbine almost directly from the weir. Small hydro technology is extremely robust (systems can last for 50 years or more with little maintenance) and also has the capacity to make a more immediate impact on the replacement of fossil fuels because unlike other sources of renewable energy. It can generally produce some electricity on demand with no need for storage or backup systems [8]. Turkey has a mountainous landscape with an average elevation of 1132 m that is about three times higher than the European average. This topography favors the formation of high gradient mountain streams which are suitable locations for SHEP development [9].

It is highly sensitive to environment like other renewable energy sources. Table 3 compares potential environmental impacts of renewable energy development. These negative effects, however, can be ignored compared to those of primary energy sources such as oil, coal and gas. In addition, SHEP plants also have effects on natural environment. They cause disadvantages such as blockage of fish passages and protection, and interruption of sediment transport. For this reason, effective fish passages for all fish species in the region where SHEP was set up must be achieved. Various local SHEP projects are under investigation and the preliminary reports of some of them are either under preparation or ready. The adoption of SHEP will lead to an improved rural economy through increased employment in their construction and provision of cheaper electricity for domestic use. In addition deforestation caused by using wood for heating will decrease.

Table 3. Comparison of potential environmental impacts of renewable energy sources.

Type of renewable	Potential environmental impact
SHEP	Blockage of fish passage and interruption of sediment transport
Wind	Noise, visual impacts, avian and bat mortality
Geothermal	Thermal pollution, damage to natural geothermal features, subsidence
Solar PV	PV panel disposal

There is a clear growth trend for SHEP plants, as can be seen in Table 4. Since 1990, number of SHEPS and their capacity has more than doubled. Installed capacity and power generation of SHEP plants are expected to be 260 MW and 968GWh/year by 2010 and 335 MW and 1250GWh/year by 2015. Table 4 gives the growth trend of SHEPs during the period 1996-2002 and short term and long term forecast of SHEP development [10].

Table 4. Small hydropower development in Turkey during 1996-2002 [11].

	1996	1997	1998	1999	2000	2001	2002	Forecast	
								2010	2015
Total number of SHEP	55	56	59	61	67	70	71	100	130
Capacity (MWh)	124.9	137.7	138.6	144.1	146.3	170.2	175.5	260	335
Generation (GWh)	499	500	524	533	636	664	673	968	1250

Material And Method

Introduction of the project area

The province of Siirt is located in the northeast of Southeastern Anatolia Region, 41°-57i-north latitude and 37°-55i east longitude and located on the Sirnak and Van, east, north and Bitlis, Batman from the west, south, surrounded by the provinces of Mardin and Sirnak. From north to south Mus Mountains, east of the province of Siirt area surrounded by the East of the Mountains, the Tigris River, is one of the major water catchment areas. All of the territory of the province fall into the Tigris river basin. Siirt's largest river, the Botan, is a major tributary of the Tigris. One of the deepest and narrowest valleys in the country, the Botan river valley opens onto the Bitlis river valley and the Tigris valley. Siirt's other major rivers are the Bitlis, Kezer and Zorava. Boasting over 350 km of

rivers and beautiful countryside. Siirt abounds in plateaus and high mountains. The city is backed by steep, high mountains to the north, which form part of the Southeastern Taurus range. Large plateaus such as the Cemiraki, Ceman, Herekol and Bacavan are used as pasturage. Siirt ongoing terrestrial climate and the most prominent features of the four seasons are experienced. Eastern and northern parts of the more hard and rainy winters, warmer summers in southern and southwestern regions. Summers are hot and dry. Surface water source of project is Kezer river and its side branches. Drainage area of kezer river is 1077 km² at 577 m elevation. Drainage area and water branches map is given in Fig 1.

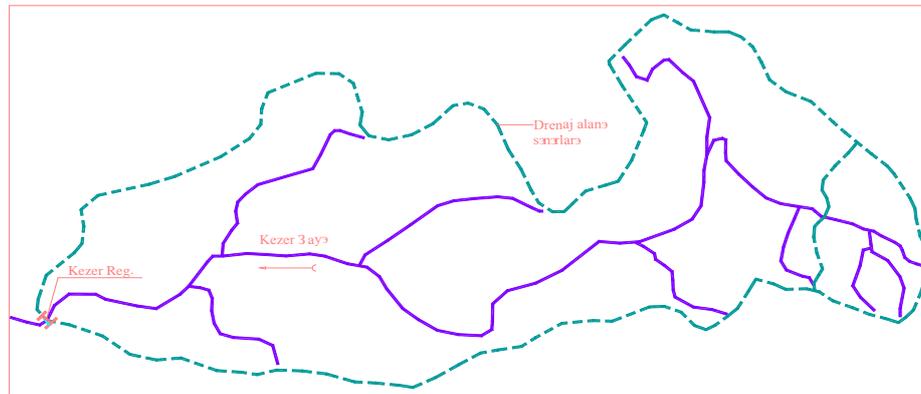


Fig 1. Kezer river drainage area and water branches map

Determining the project formulations

In order to determine project formulation we have used topographic map and google earth program and we have also done the investigations about land's geomorphological state, topographic state, impact of agricultural area, historical and cultural entities. Satellite images about project area are shown in Fig 2 and land pictures are shown in Fig 3. It is thought as not accumulated HEP after investigation mentioned above. It is much more suitable to put regulator location to 569 m. Free surface trapeze is used in order to transmit translated by regulator. Total length of transmission channel is 8200 m. Installation pool is made for providing connection with penstock made by transmission channel open to the atmosphere. Installing pool is made in order to perform pipe with pressure to store the water requirement at least for one minute for the turbine and to meet rebound in case of water entrance to turbine or water cut.

Water entrance from installation pool to tribunes in central is done with penstock. Penstock working with pressure is made with reinforced concrete and polietilen pipe or mainly steel pipe. Penstock's diameter can be changeable according to the discharge rate. Average speed in penstock is required as 3-6 m/s but maximum speed is required as 10 m/s [12] . Penstock is determined as 100 m in Kezer regulator and HEP project.



Fig 2. Kezer regulator, HEP regulator location and route of the transmission channel



Fig 3. Kezer regulator and HEP route of the transmission channel

Power house is the plant transforming water to energy and water comes from penstock. Water passed through turbines transform into mechanic energy. Energy made up of with water passed through turbines decreases with its productivity and energy transferred to generator is transformed to electrical energy. But, there can be seen decreases in the quantities of energy according to the productivity of generator. Voltage of the electricity is increased from medium to high and it is transferred to the switcyard. Operation of the water leaving the turbine again into the river bed is done by tail water channel. Elevation of the tail water channel of this project is 532 m.

Determining water discharge

There are records of observation station (AGİ) numbered as EİE 2624 from 1972 to 2010 and this AGİ is located on Pınarca village site 530 m on Kezer stream in the downstream of project working border around 577 m and 532 m. There is no need to investigate another stations or statistical analysis among stations and this AGİ has got enough time of period (39 years) and perfect information for representation of the project. Discharge and flow rates should be determined for regulator place in order to move this discharge and flow rates to regulator place. For this reason, field calculation should be made for the regulator place on 569 meters and records should be moved to regulator place with areas ratio method. Discharge results should be used to discharge continuity curve and then water discharge (in the 95 % of time) and quantity of firm (reliable) energy would be determined.

Installed capacity calculation

To calculate the installed power required to account for the net head. For this, first head losses to be calculated and calculation will be based on the difference between the gross head.

Flood calculation

Regulator to determine the capacity of the spillway flood accounts is required. For this purpose, regional flood frequency analysis method will be used.

Findings And Discussion

Climate and water resource of the region

Kezer rainfall area completely have features of climate of east anatolia. Regions parameters of rain, humidity, heat, evaporation, relative humid, wind and snow are determined by stations used by DSI (state water affairs) and DMI (state meteorological service). kezer regulator at the thalweg elevation of around 569.00 m is planned to turn water comes from rainfall area which is on this elevation. In Fig 4, rainfall area of Kezer regulator is determined on the map with the scale of 1/ 25000, rainfall area for regulator place is 1077.60 km² and rainfall area for power house is 1169.60 km². Meteorology observations stations near the project rainfall area were drawn using the thiessen poligons. Project rainfall area's observation stations are Baykan, Siirt, Sirvan, Bitlis, Hizan, Resadiye, Sarikonak, Tatvan. These stations representation rates are respectively 2.11%, 4.43%, 36,08% 19.67%, 13.79%,4.23%, 15.86%, and 3.82%.

Surface water source of project is Kezer river and its side branches. In the northeast of project rainfall area, Sapur river sprang from about 2700.00 m elevation, then named as Guzeldere with joining of large and small side branches from right or left shore. Guzeldere is then renamed as Baykan

river and its name is changed as Cobansuyu with joining of Ceviz river around 1050.00 m elevation. It is named as Koca river with joining of Harmanyeri river and then it is named as Kezer river with joining of side branches at 569.00 m elevation where it is reached to regulator place. Branches from right and left shores joined to Kezer river which flows in the South direction and it is reached to central place as Pınarca river at 532.00 m elevation.

Kezer river numbered as EİE 2624- Pınarca flow observation station (AGİ) is on the same stream with the project. Rainfall area of this station is 1169.60 km² and its approximate elevation is 530 m. Flow records of this station are decided to be used in water supply of Kezer regulator and HEP Project. The flow records of EİE 2624 AGİ is used in the calculations of average daily flow of Kezer regulator place planned to set up at 569.00 m elevation. 1972 - 2010 years of recorded daily average flows for the period covering the 39-year, monthly and annual average flow rates (Fig 5), monthly and annual total flow is used. This AGİ's average daily records of discharge for 39 years, average monthly and yearly discharge (Fig 5) and total monthly and yearly total flow rates are used. Accordingly, EİE 2624 AGİ's yearly average discharge is 19.78 m³/s and yearly average flow is 622.73 hm³.

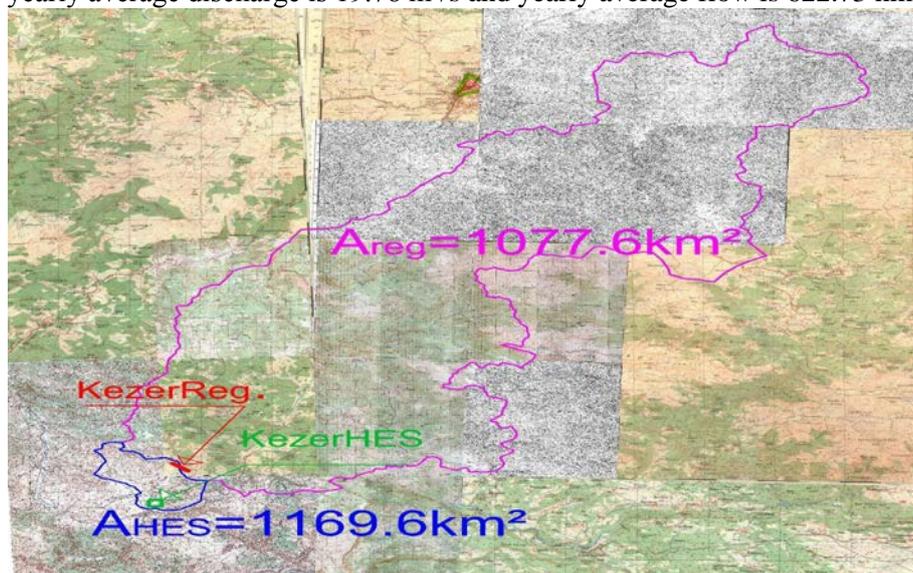


Fig 4. Regulator place and drainage area of the central place

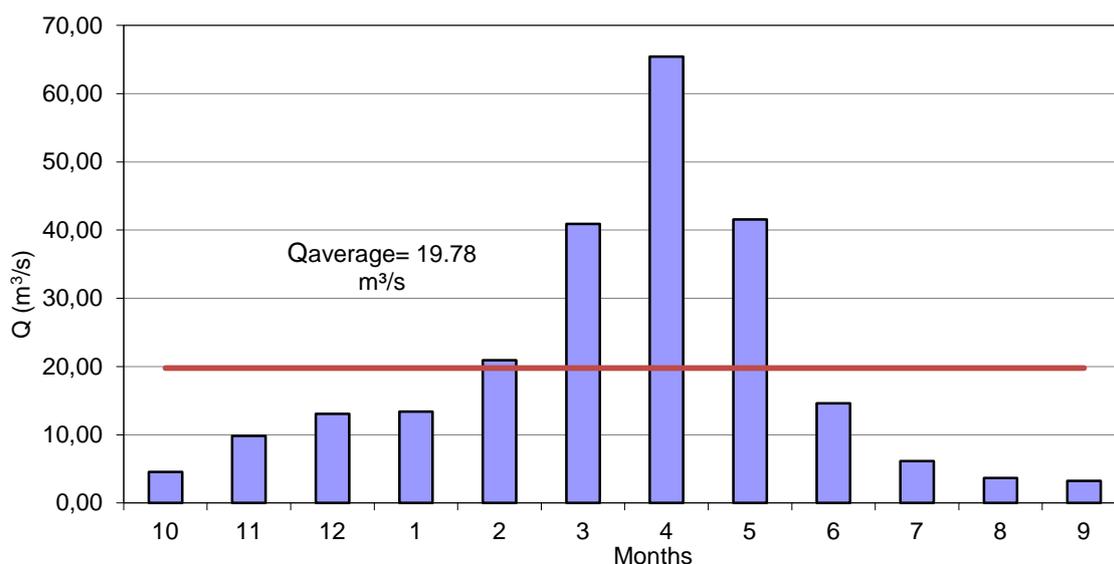


Fig 5. EİE 2624 AGİ, for many years monthly average flow rates

Discharge continuity curve drawn by using flow rates between the years of 1972 and 2010 is shown in Fig 6. Accordingly, yearly average discharge of Kezer regulator area is 17.16 m³/s and its yearly total flow is 540.08 hm³. Firm discharge available at least 95% of not accumulated HEP project is described as discharge on the source of river [13]. For this reason, discharge continuity curve is

used. Firm power is described as power obtained by this discharge and firm energy. Energy which is not firm is named as secondary energy.

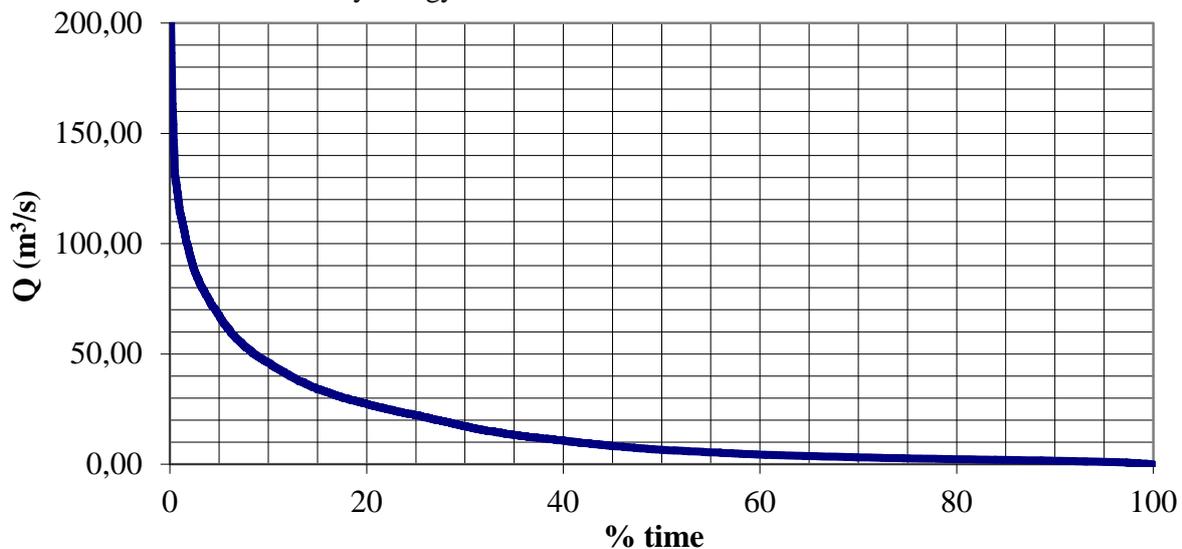


Fig 6. Regulator place discharge continuity curve

Discharge continuity curve drawn by discharge and temporal percentages is shown in Fig 6 [14]. According to discharge results obtained by discharge flow curve in the time of 95 % is 1.64 m³. The turbine power plant, set up with this discharge, is accepted as firm power and energy can be generated is accepted as firm energy.

Calculations of installed power

The installed power of project is calculated by the discharge and head, the equation is as follows:

$$P = 9.81 \times Q \times H \times \eta \quad (1)$$

where, P: installed power (Kwh), Q: discharge (m³/s), H: Net head, η : productivity of the turbine.

In this equation, we must find firstly head loss in order to find net head. Total head loss is calculated as 4.22 m. Gross head is 45 m (577-532) and net head is obtained as 40.78 m (45-4.22).

Design discharge is accepted as 36 m³/s by choosing discharge in 15% of time in the discharge continuity curve. In the literature calculation of firm power, installed power is Q_{50} and maximum design discharge is Q_{15} . Initial design discharge is described as the value of the time in the 15% and 30 % . Installed power P is calculated as 1325 kwh or mechanically 13.25MW (9.81x36x40.78x0.92). 0.98 which is the coefficient generator is multiplied by this energy and we find the electrical power. In this case installed power is 13.00 MW electric.

Energy production calculations

The amount of energy that can be produced in a year is calculated based on annual water volume and head. The annual volume of water taken as the average of all years removed tailrace,

$$E = 2.32 \cdot 10^3 \times V \times H_{\text{net}} \times \eta \quad (2)$$

where, E: annual amount of net energy generated (kwh), V: annual amount of water (m³), H: amount of net head (m), η : turbine efficiency. Total energy is calculated as 42,57 GWh/year (2.32.10⁻³x489.12.10⁶x40.78x0.92). Firm energy is calculated as 4.44 GWh/year (2.32.10⁻³ x 50.96.10⁶x40.78x0.92). Secondary energy is calculated as 38.13 GWh/year (42,57-4,44).

Choosing of turbines

According to the calculated discharge and head, It is chosen as Francis turbine. The productivity of turbine's exit power is determined as 0.92. The number of turbine is two because project design discharge is above the average discharge. Discharge of water is divided in to two named as branch in the central entrance after penstock and it will be passed through each turbine as 18.00 m³/s. Each turbine power is 6635 kwh. Horizontantly Francis turbine is used.

Flood Accounts

Regional flood frequency analysis is chosen in this project because the numbers of records are enough and flow data exactly represents the area. In these calculations annual recurrence flood discharge found by AGI's frequency analysis and this discharges moved to regulator and central by using formulations given below;

$$Q_{Reg} = (A_{Reg}/A_{AGI})^{2/3} \times Q_{AGI} \quad (3)$$

$$Q_{Santral} = (A_{Santral}/A_{AGI})^{2/3} \times Q_{AGI} \quad (4)$$

Given the recurrent times for AGI flood values calculated for regulator and central place is given in Table 5.

Table 5. Flood values

Recurrence year	Station place (m ³ /s)	Regulator place(m ³ /s)
2	328.98	307.84
2.33	345.68	323.47
5	480.81	449.91
10	574.64	537.71
25	687.98	643.77
50	769.24	719.8
100	848.13	793.62
500	1022.31	956.61

Conclusion

In this paper the evaluation of hidroelectrical potential of Kezer river (the branch of Botan river on Dicle basin) between the elevation of 532 m and 577 m is done. For this reason firstly, regulator which is thought as convert structure for plant is seen suitable to place it at the 569 m elevation. Water taken from the regulator is thought to transmit to central by transmission channel with a free surface flow which is 8200 m long. Installing pool is designed for controlling water entrance to central with penstock at the end of the transmission channel. Water discharge in the time of 95 % is found as 1.64 m³/s with drawing discharge continuity curve. Total of continuous and local head loss is calculated as 4.22 m (In the sedimentation pool, transmission channel and penstock) and it is subtracted from gross head (45 m) and real head is found as 40.78 m. Discharge, real head and the productivity of turbine are used to find the value of installed power as 13.25 Mwh and the value of total energy as 42.57 Gwh.

References:

- [1] Gokcol C, Dursun B, Alboyaci B, Sunan E. Importance of biomass energy as alternative to other sources in Turkey. *Energy Policy*;37(2):424-31, 2009.
- [2] MENR. Turkey's final energy consumption by resources. Ministry of Energy and Natural Resources; 2008 [Energy Statistics].
- [3] Hepbasli, A., Ozdamar, A., Ozalp, N., Present status and potential of renewable energy sources in Turkey. *Energy Sources* 23, 631–648, 2001.
- [4] Yuksek O, Komurcu MI, Yuksel I, Kaygusuz K. The role of hydropower meeting the electric energy demand in Turkey. *Energy Policy*; 34:3093-103, 2006.
- [5] DIE. State Institute of Statistics. Statistic yearbook of Turkey in 2003. Ankara: Prime Ministry, Republic of Turkey; 2004.
- [6] Ozturk M, Bezir NC, Ozek N. Hydropowerwater and renewable energy in Turkey: sources and policy. *Renewable and Sustainable Energy Reviews*;13(3):605-15, 2009.
- [7] Yuksel I. Hydropower in Turkey for a clean and sustainable energy future. *Renewable and Sustainable Energy Reviews*;12:1622-40, 2008.
- [8] Paish O. Small hydro power: technology and current status. *Renewable and Sustainable Energy Reviews*;6(6):537-56, 2002.
- [9] Kucukali S, Baris K. Assessment of small hydropower (SHP) development inTurkey:laws, Regulations and EU policy perspective. *Energy Policy*;37:3872-9, 2009.
- [10] Punys P, Pelikan B. Review of small hydropower in the new Member States and Candidate countries in the context of the enlarged European Union. *Renewable and Sustainable Energy Reviews*;11:1321-60, 2007.

- [11] Altınbilek D, Cakmak C. The role of dams in development, DSI Third international symposium. Ossiach, Austria; 2001.
- [12] Erdem M. Küçük Hidroelektrik Santrallerin Tasarım ölçütleri, Turkish (Master thesis), Pamukkale University, Denizli, 2006.
- [13] Buttanrı B. Türkiye’de Küçük Hidroelektrik Santrallerin Tarihsel Gelişimi ve Bugünkü Durum, Turkish (Master thesis.), ITU, İstanbul, 2006.
- [14] Bayazıt M, Avcı İ, Şen Z. Hidroloji Uygulamaları. Birsen yayınevi, No:29, Page: 1-235, İstanbul, 2001.
- [15] Emiroğlu M. Kezer Çayı’nın Hidroelektrik Enerji Potansiyelinin Değerlendirilmesi, Turkish (Master thesis) , Dicle University, Diyarbakır, 2012.

THE URBAN ENVIRONMENTAL ENGINEERING CONDITION OF UDUPI TALUK: A GIS EVALUATION URBAN AMENITIES WITH A FOCAL POINT PREDICTION OF POTENTIAL IMPROVEMENT OF URBANIZATION

H N Udayashankara

B.S.Maddodi

Department of Civil Engineering, Manipal Institute of Technology, Manipal

B.R.Raghavan

Department of Marine Geology, Geoinformatics Division, Mangalore University

Shema V.T

Department of Civil Engineering, Manipal Institute of Technology, Manipal

Abstract:

Urban Environmental Engineering is the application of science and engineering techniques to improve the natural environment to provide healthy water, air, and land for human habitation including infrastructure. The focal point is on the sustainable development of an urban entity where development of industrial and natural resources elevates and sustains energy needs of a civic population. This dissertation is aimed at developing a centralized system to carry out the analysis of urbanization of the Udupi Taluk with inputs from various interdisciplinary fields of demography, commerce, education, industries, health and tourism. To brawl with the problems of environmental degradation and to meet the challenges of sustainable development, it is suggested that the use of remote sensing and GIS in conjunction with geospatial data is of vital importance. There is need for the use of an urban information database that can be formulated using remote sensing data and GIS techniques. This is possible as GIS technology is unique in embracing spatial data of various parameters that could be modeled to bring out a multi-parametric relation in nature. The end result would provide spatial variations in the form of digital maps. These maps and their polygons would be imported in a GIS environment and the results would be used in query-models and thereby establishing the urban development of Manipal and the Udupi taluk. The output would shift the interpretations from single parametric to multi-parametric relation and this would enable the end users such as governmental agencies to pool data and adopt new strategies in dealing with urbanization in a comprehensive manner.

Key Words: Urban Environmental Engineering, Geographic Information System, Spatial Data, Digital maps, Query-models

Introduction

There is an asymmetrical urban growth, which is taking place all over the world, but the rate of urbanization is very fast in the developing countries, especially in Asia. In 1800 A.D., only 3% of the world's population lived in urban centers, but this figure ranged to 14% in 1900 and in 2000, about 47% (2.8 billion) people were living in urban areas. Statistics show that India's urban population is the second largest in the world after China, and is higher than the entire urban population of all countries put together barring China, USA and Russia.

Every major city of India faces the same proliferating problems of urban expansion, inadequate housing, poor transportation, flawed sewerage, erratic electric supply, and insufficient water supplies. The level of air, water, and land pollution has increased because of abysmal environmental management. This has a direct impact on the quality of the urban environment, affecting labor productivity and the overall socio-economic development. The circumstance of the

urban environment all over India is deteriorating so fast that the sustainability of the cities is threatened.

The increasing demands for information in urban planning and management sectors necessitate the application of remote sensing for sustainable development of urban areas. Thus in this context integrated geo-spatial technologies such as remote sensing (RS), geographic information system (GIS) and global positioning system (GPS) can tally up to interactive operations that would be an asset for assessing, understanding, and mapping utility and service facilities, as well as solving complex urban environmental problems. By utilizing remote sensing data and implementing GIS mapping techniques, changes in urban extent can be monitored and mapped for explicit developmental projects. The situation is grave in India due to unplanned growth of the cities in all directions. The current study will prove to be useful to assess some of the urban environmental issues which Manipal and Udipi taluk is currently facing with the help of geo-spatial tools.

Study Area

Udipi district is a fairly new one, carved out from the erstwhile Dakshina Kannada district of Karnataka in 1997. It lies between $74^{\circ}34'45''\text{E}$ to $75^{\circ}12'20''\text{E}$ longitudes and $12^{\circ}59'40''\text{N}$ to $13^{\circ}48'50''\text{N}$ latitudes. The district is covered in the Survey of India toposheet references of 48K and 48O series on 1:250,000 scale. **Udipi taluk** is in the Udipi District of the Indian state of Karnataka.

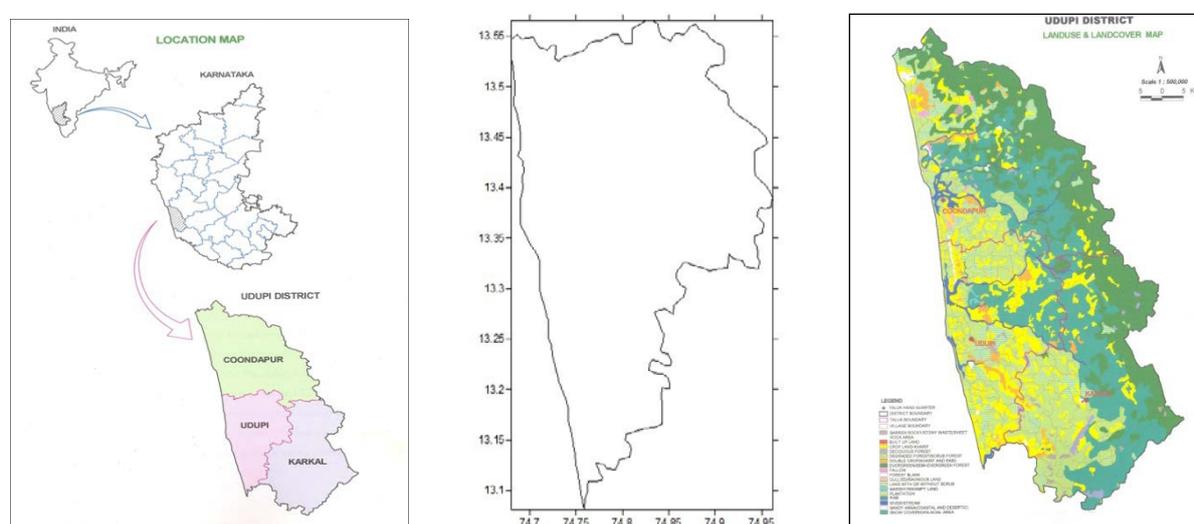


Figure 1: Study Area- Udipi Taluk with Geographical Co-ordinates

Role of Geo-information systems in sustainable urban development

The role of GIS in its decision making power in the field of sustainable urban development seems to be ideal. This would requisite the availability of 'sustainable indicators' to provide a solid basis for the study, which include health, education, employment, air quality, traffic, housing facilities and economic prosperity. Creating a powerful centralized information system describing the data and their locations would be necessary for the grass-root level planning of a city or district.

Remote sensing and GIS could serve as tools for monitoring and updating the changes. Therefore, 'geoinformation technology' is a means to integrate the sustainable indicators for suitable modeling, thereby helping in the evaluation of results and implementing an effective planning of an urban area.

Objective of the Work

The project aims to appreciate the spatial relations of population, schools, banks, urban facilities, worship places, tourist centers, and how it can be related to effectively managing the environmental resources. The objective of this work is the need to infuse GIS technology in conventional environmental engineering. The end result would provide spatial variations in the form of digital maps. These maps and their polygons would be imported in a GIS environment and the results would be used in query-models focusing on the urban development of Manipal and the Udipi taluk.

Methodology

The Udupi Taluk map is first digitized and Arc coverage of taluk map, a vector layer of stations and the attribute table are made. Following blanking process in Golden Surfer, the Contour Maps are drawn after gridding the dataset (point, line, area). Also, in this GIS platform, the incomplete polygons of the contour map are completed. In the ArcGIS environment, the selected polygon layer shows the corresponding polygon and its attribute table. All the polygons created are thus made to intersect, depicting all attribute values at once.

Query modeling of the created spatial layers is accomplished by assigning suitable conditions. Multiple queries can be established for the same spatial layer.

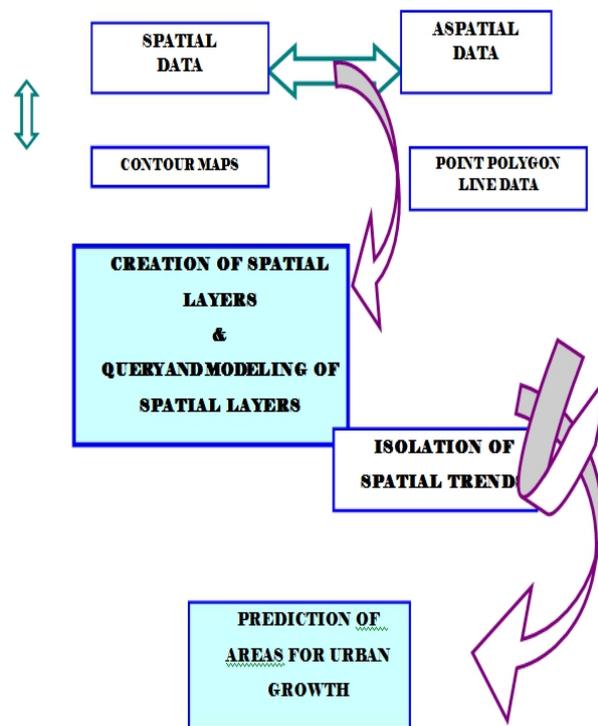


Figure 2: A Flowchart of the intended project program

Tools Required

The following is a short-list of details required:

- i. Primary and Secondary data on various spatial parameters.
- ii. Geospatial processing tools are used for fusing and integrating geospatial source content into software applications for the creation and update of geospatial data and information products. GIS software ESRI products such as the following are used:
 - i. ArcView10
 - ii. ArcGIS9.31
 - iii. ERDAS 9.1
 - iv. Golden Surfer 9.8

Data Collected

For determining the Urban Environmental Engineering status of Udupi Taluk, the data available for adjoining places in the taluk was analyzed. It included information on

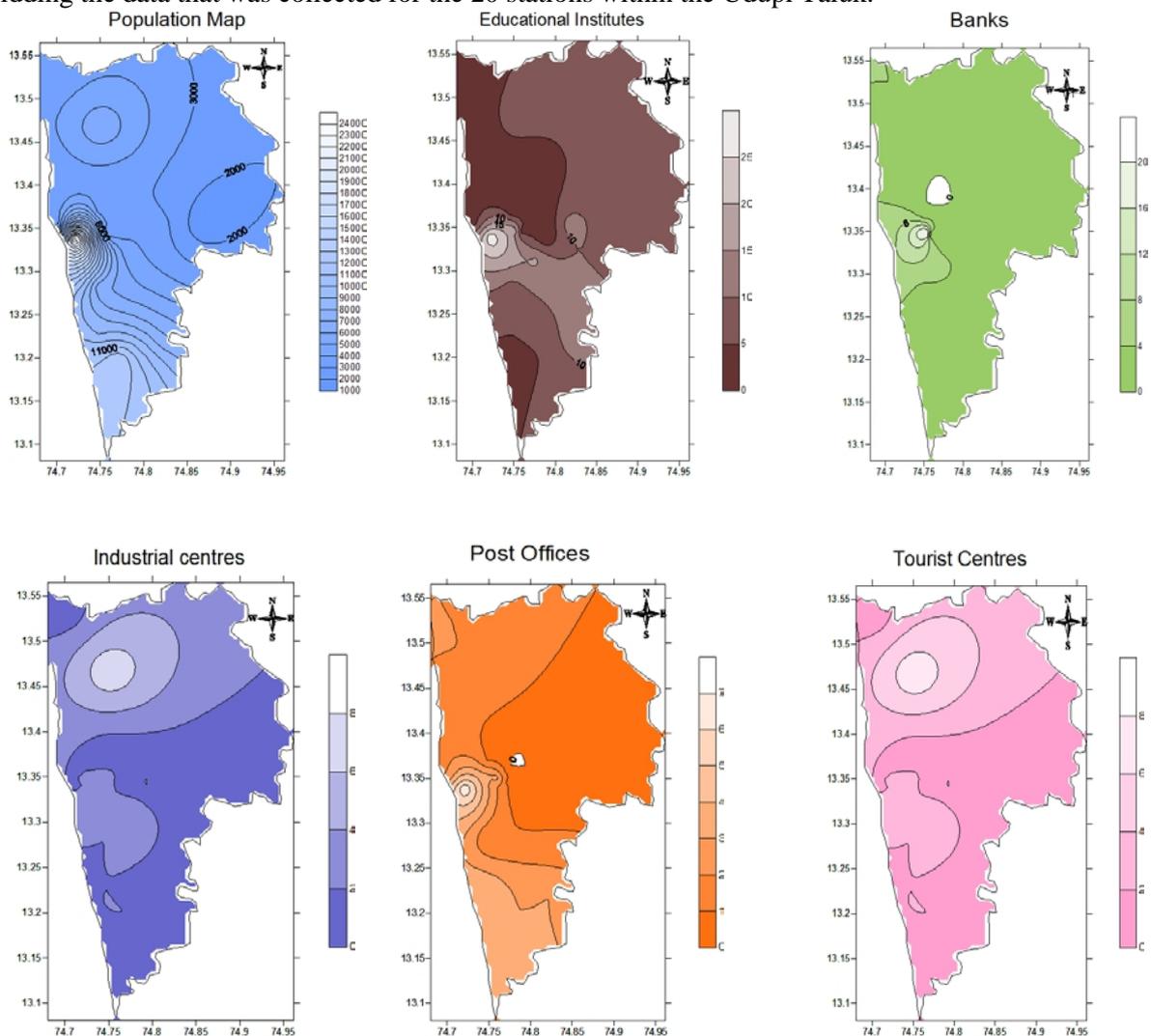
- i. Population statistics
- ii. Industrial networks
- iii. Educational Institutes
- iv. Worship places
- v. Banks
- vi. Post offices
- vii. Tourism centers

The following 20 stations were identified for data sampling

Manipal, Rajeev Nagar, Indralli, Kalsanka, Saagri, Parkala, Hiriyadka, Badagabettu; Udupi city; Barkur; Katpadi; Shirva; Padubidiri; Malpe; Ambagilu; Kaup; Saralabettu; Alevoor; Kota and Chitpady.

Results and Discussions

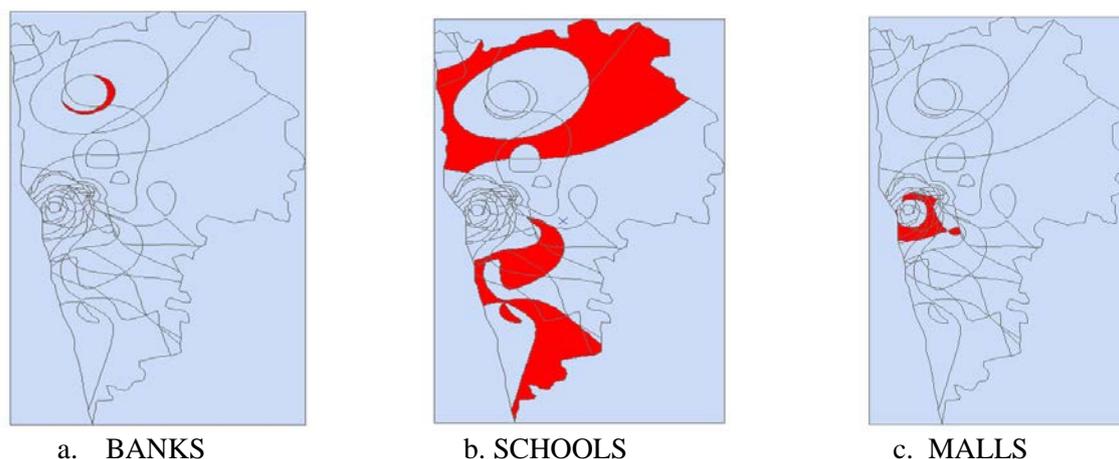
The contour maps for different data were obtained using Erdas Imagine and Golden Surfer by gridding the data that was collected for the 20 stations within the Udupi Taluk.



Query modeling is the most critical step of the study that leads to the prediction of future pockets of growth and development in a town, city, taluk, district, state or country. Querying of the data for different conditions yielded the following results.

- Query 1:* "Population" = '<5000' AND "Industry" = '6-8' AND "Banks" = '<4'
- Query 2:* "Banks" = '<4' AND "Industry" = '2-4' OR "Population" = '9000-13000' AND "Institute" = '5-10'
- Query 3:* "Industry" = '2-4' AND "Institute" = '15-20' OR "Banks" = '16-20' AND "tourist_pl" = '6-8'

The query will output results that incorporate the required conditions and the highlighted areas shows the corresponding region most suitable to set up the below mentioned urban facility.



Conclusions

The major problem associated with urban centers in India is that of unplanned expansion. From this study it is evident that using GIS techniques, an assessment of urban environmental issues can be done as they are important tools to map, assess, and monitor the changes in the urban environment. The deterioration of urban environmental quality is due to a combination of a growing population and rising urbanization levels, the consequences of which can be effectively monitored and assessed by using geo-spatial tools. GIS would help city planners to assess and set up urban facilities as per the requirements of the growth of the place. The officials of various government departments should be given thorough exposure and training of GIS for its application and implementation in the urban environmental management plans.

References:

- Boyd, D.S., G.M., Foody, *Ecological Informatics*, 6(1):25-36, (2011)
- Christine L. Rioux, David M. Gute, Doug Brugge, Scott Peterson, Barbara Parameter, Characterizing Urban Traffic Exposures Using Transportation Planning Tools: An Illustrated Methodology for Health Researchers. *Journal of Urban Health* 87:2, pages 167-188, (2010)
- David Briggs, *Journal of Toxicology and Environmental Health, Part A* 68:13-14, pages 1243-1261, (2005)
- H. Merbitz, M. Buttstädt, S. Michael, W. Dott, C. Schneider, GIS-based identification of spatial variables enhancing heat and poor air quality in urban areas. *Applied Geography* 33, pages 94-106, (2012)
- Javed, A, and Pandey, S “Land Use/Land Cover Analysis for Waste Disposal”, *GIS Development*, 8(6), Delhi, (2004)
- John Gulliver, Kees de Hoogh, Daniela Fecht, Danielle Vienneau, David Briggs, Comparative assessment of GIS-based methods and metrics for estimating long-term exposures to air pollution. *Atmospheric Environment* 45:39, pages 7072-7080, (2011)
- Jones, C. E., Vaughan, L., Haklay, M., Griffiths, S. GIS enhances collaboration: using the line to draw disciplines together. *Proceedings of the GIS Research UK 18th Annual Conference*. (pp.29-34). London: UCL, (2010).
- Keita Ebisu, Theodore R. Holford, Kathleen D. Belanger, Brian P. Leaderer, Michelle L. Bell, Urban land-use and respiratory symptoms in infants. *Environmental Research* 111:5, pages 677-684, (2011)
- Lata, K M, Sankar Rao, C H *et al.* “Measuring Urban Sprawl: A Case Study of Hyderabad”, *GIS development*, 5(12), (2001)
- Liu, Y., Ellul, C., Haklay, M. Planning Alerts for Community Maps. *Proceedings of the GIS Research UK 18th Annual Conference*. (pp.363-368). London: UCL, (2010).
- Paskins, J., Bell, S. J., Croxford, B., Haklay, M., Julier, S. Crossing Disciplines to Address Urban Sustainability. *European Association for the Study of Science and Technology (EASST)*. (2010).
- Small K & Song S, Population and employment densities: structure and change. *J. Urban Econ.* 36: 292-313, (1994)

PLANE STATE PROBLEM ANALYSIS WITH FINITE-DIFFERENCE METHOD

Harangus Katalin, PhD Candidate

Transilvania University of Braşov, Faculty of Mechanical Engineering, Romania

Kakucs Andras, Assoc. prof. Dr.

Sapientia-Hungarian University of Transylvania, Faculty of Technical and Human Sciences, Romania

Galfi Botond Pal, Dr. Eng.

Autoliv Inc., Braşov, Romania

Abstract:

This paper presents a finite-difference analysis of stresses and displacements of the plane elastic problems of orthotropic materials. Starting from the Airy stress function, we assume that in the case of orthotropic materials there is a function $\Psi(x, y)$ the partial derivatives of which determine the specific deformations with the material equations. We also use a potential function of the displacements, the partial derivatives of which lead to the stress fields with the help of the material equations. This will make the prescription of the mixed boundary conditions possible. Therefore, the description of the boundary conditions under the form of prescribed stresses (of the load distribution on the boundary) becomes possible, because there is a direct relation (differential equations) between the displacements and stresses.

Key Words: Displacement potential function, finite-difference method, orthotropic material

Introduction

In the linear elasticity theory it is assumed that the relations between stress and deformations are linear. This relation can be described by the formula

$$\{\sigma\} = [E] \cdot \{\varepsilon\}, \quad (1)$$

where $[E]$ is the elasticity matrix, a 6-by-6 matrix, which contains 36 material constants (Szalai, 1994).

For orthotropic materials in plane stress state, and in plain strain state, the two elasticity matrices $[E]$ are valid only if the directions of orthotropy coincide with the directions of the coordinate axes. Otherwise, the two arrays must be rotated (Curtu, 1984). Thus, the transformation leads to a full matrix (Kakucs, 2007)

$$[E] = \begin{bmatrix} \bar{E}_{11} & \bar{E}_{12} & \bar{E}_{13} \\ \bar{E}_{21} & \bar{E}_{22} & \bar{E}_{23} \\ \bar{E}_{31} & \bar{E}_{32} & \bar{E}_{33} \end{bmatrix}, \quad (2)$$

of which components are:

$$\begin{aligned}
\bar{E}_{11} &= E_{11} \cdot \cos^4 \theta + E_{22} \cdot \sin^4 \theta + (E_{12} + E_{21} + 4 \cdot E_{33}) \cdot \sin^2 \theta \cdot \cos^2 \theta, \\
\bar{E}_{12} &= E_{12} \cdot \cos^4 \theta + E_{21} \cdot \sin^4 \theta + (E_{11} + E_{22} - 4 \cdot E_{33}) \cdot \sin^2 \theta \cdot \cos^2 \theta, \\
\bar{E}_{13} &= (E_{11} - E_{12} - 2 \cdot E_{33}) \cdot \sin \theta \cdot \cos^3 \theta + (E_{21} - E_{22} + 2 \cdot E_{33}) \cdot \sin^3 \theta \cdot \cos \theta, \\
\bar{E}_{21} &= E_{12} \cdot \sin^4 \theta + E_{21} \cdot \cos^4 \theta + (E_{11} + E_{22} - 4 \cdot E_{33}) \cdot \sin^2 \theta \cdot \cos^2 \theta, \\
\bar{E}_{22} &= E_{11} \cdot \sin^4 \theta + E_{22} \cdot \cos^4 \theta + (E_{21} + E_{12} + 4 \cdot E_{33}) \cdot \sin^2 \theta \cdot \cos^2 \theta, \\
\bar{E}_{23} &= (E_{11} - E_{12} - 2 \cdot E_{33}) \cdot \sin^3 \theta \cdot \cos \theta + (E_{21} - E_{22} + 2 \cdot E_{33}) \cdot \sin \theta \cdot \cos^3 \theta, \\
\bar{E}_{31} &= (E_{11} - E_{21} - 2 \cdot E_{33}) \cdot \sin \theta \cdot \cos^3 \theta + (E_{12} - E_{22} + 2 \cdot E_{33}) \cdot \sin^3 \theta \cdot \cos \theta, \\
\bar{E}_{32} &= (E_{11} - E_{21} - 2 \cdot E_{33}) \cdot \sin^3 \theta \cdot \cos \theta + (E_{12} - E_{22} + 2 \cdot E_{33}) \cdot \sin \theta \cdot \cos^3 \theta, \\
\bar{E}_{33} &= (E_{11} - E_{12} - E_{21} + E_{22} - 2 \cdot E_{33}) \cdot \sin^2 \theta \cdot \cos^2 \theta + E_{33} \cdot (\sin^4 \theta + \cos^4 \theta),
\end{aligned} \tag{3}$$

where θ is the angle measured from the first direction of the orthotropy 1 to the x axis (figure 1.)

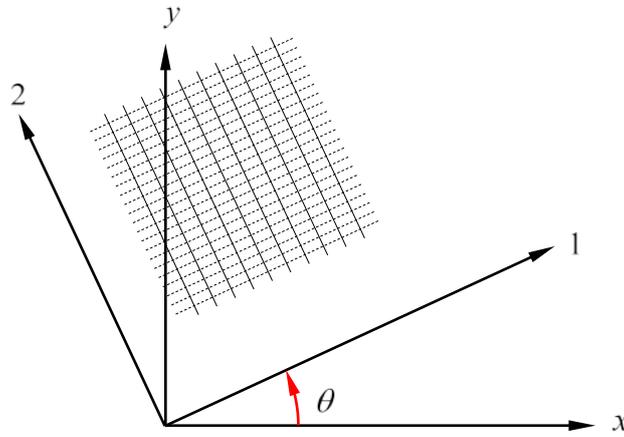


Fig. 1. Orthotropic direction

and

$$[\mathbf{E}_{12}] = \begin{bmatrix} E_{11} & E_{12} & 0 \\ E_{21} & E_{22} & 0 \\ 0 & 0 & E_{33} \end{bmatrix}, \tag{4}$$

is the matrix of elasticity in the directions of orthotropy (it has only three independent elements). Therefore, where the directions of orthotropy do not coincide with the axes, the elasticity matrix contains nine nonzero elements and is symmetric. In the general case of plane anisotropy, the elasticity matrix is also a full and symmetric, but it contains six independent elements.

Formulation for finite-difference solution

Starting from the idea of Airy stress function, we suppose that in the case of orthotropic materials there is a function $\Psi(x, y)$ of which partial derivatives give the projections of the displacement. We transcribe the equilibrium equations using Hooke's law in specific strains (considering $f_x = 0$), then with the geometrical equations, we obtain the followings (Harangus, 2012):

$$\begin{aligned}
& (\alpha_1 \cdot E_{11} + \alpha_4 \cdot E_{13}) \cdot \frac{\partial^4 \Psi}{\partial x^4} + \\
& + (\alpha_1 \cdot E_{13} + \alpha_1 \cdot E_{31} + \alpha_2 \cdot E_{11} + \alpha_4 \cdot E_{12} + \alpha_4 \cdot E_{33} + \alpha_5 \cdot E_{13}) \cdot \frac{\partial^4 \Psi}{\partial x^3 \cdot \partial y} + \\
& + (\alpha_1 \cdot E_{33} + \alpha_2 \cdot E_{13} + \alpha_2 \cdot E_{31} + \alpha_3 \cdot E_{11} + \alpha_4 \cdot E_{32} + \alpha_5 \cdot E_{12} + \alpha_5 \cdot E_{33} + \alpha_6 \cdot E_{13}) \cdot \frac{\partial^4 \Psi}{\partial x^2 \cdot \partial y^2} + \quad (5) \\
& + (\alpha_2 \cdot E_{33} + \alpha_3 \cdot E_{13} + \alpha_3 \cdot E_{31} + \alpha_5 \cdot E_{32} + \alpha_6 \cdot E_{12} + \alpha_6 \cdot E_{33}) \cdot \frac{\partial^4 \Psi}{\partial x \cdot \partial y^3} + \\
& + (\alpha_3 \cdot E_{33} + \alpha_6 \cdot E_{32}) \cdot \frac{\partial^4 \Psi}{\partial y^4} = 0,
\end{aligned}$$

$$\begin{aligned}
& (\alpha_1 \cdot E_{31} + \alpha_4 \cdot E_{33}) \cdot \frac{\partial^4 \Psi}{\partial x^4} + \\
& + (\alpha_1 \cdot E_{21} + \alpha_1 \cdot E_{33} + \alpha_2 \cdot E_{31} + \alpha_4 \cdot E_{23} + \alpha_4 \cdot E_{32} + \alpha_5 \cdot E_{33}) \cdot \frac{\partial^4 \Psi}{\partial x^3 \cdot \partial y} + \\
& + (\alpha_1 \cdot E_{23} + \alpha_2 \cdot E_{21} + \alpha_2 \cdot E_{33} + \alpha_3 \cdot E_{31} + \alpha_4 \cdot E_{22} + \alpha_5 \cdot E_{23} + \alpha_5 \cdot E_{32} + \alpha_6 \cdot E_{33}) \cdot \frac{\partial^4 \Psi}{\partial x^2 \cdot \partial y^2} + \quad (6) \\
& + (\alpha_2 \cdot E_{23} + \alpha_3 \cdot E_{21} + \alpha_3 \cdot E_{33} + \alpha_5 \cdot E_{22} + \alpha_6 \cdot E_{23} + \alpha_6 \cdot E_{32}) \cdot \frac{\partial^4 \Psi}{\partial x \cdot \partial y^3} + \\
& + (\alpha_3 \cdot E_{23} + \alpha_6 \cdot E_{22}) \cdot \frac{\partial^4 \Psi}{\partial y^4} + f_y = 0.
\end{aligned}$$

We determine the α coefficients in such manner to get the multipliers of the partial derivatives of the first equation equal to zero (in this case any function Ψ is a solution of the first equation):

$$\begin{aligned}
& \alpha_1 \cdot E_{11} + \alpha_4 \cdot E_{13} = 0 \\
& \alpha_1 \cdot E_{13} + \alpha_1 \cdot E_{31} + \alpha_2 \cdot E_{11} + \alpha_4 \cdot E_{12} + \alpha_4 \cdot E_{33} + \alpha_5 \cdot E_{13} = 0 \\
& \alpha_1 \cdot E_{33} + \alpha_2 \cdot E_{13} + \alpha_2 \cdot E_{31} + \alpha_3 \cdot E_{11} + \alpha_4 \cdot E_{32} + \alpha_5 \cdot E_{12} + \alpha_5 \cdot E_{33} + \alpha_6 \cdot E_{13} = 0 \quad (7) \\
& \alpha_2 \cdot E_{33} + \alpha_3 \cdot E_{13} + \alpha_3 \cdot E_{31} + \alpha_5 \cdot E_{32} + \alpha_6 \cdot E_{12} + \alpha_6 \cdot E_{33} = 0 \\
& \alpha_3 \cdot E_{33} + \alpha_6 \cdot E_{32} = 0.
\end{aligned}$$

Since six coefficients cannot be determined from these five equations, we must prescribe one of the values (Reaz, 2005). Therefore, we assign $\alpha_2 = 1$, and the five remaining coefficients are found by solving the system of equations (7), which can be solved by numerical methods. With the obtained α coefficients the second equilibrium equation will be the following:

$$\beta_1 \cdot \frac{\partial^4 \Psi}{\partial x^4} + \beta_2 \cdot \frac{\partial^4 \Psi}{\partial x^3 \cdot \partial y} + \beta_3 \cdot \frac{\partial^4 \Psi}{\partial x^2 \cdot \partial y^2} + \beta_4 \cdot \frac{\partial^4 \Psi}{\partial x \cdot \partial y^3} + \beta_5 \cdot \frac{\partial^4 \Psi}{\partial y^4} = \beta_0 \cdot f_y, \quad (8)$$

of which solution is the potential function sought by us. The coefficients of this equation are:

$$\begin{aligned}
& \beta_1 = \alpha_1 \cdot E_{31} + \alpha_4 \cdot E_{33}, \\
& \beta_2 = \alpha_1 \cdot (E_{21} + E_{33}) + \alpha_2 \cdot E_{31} + \alpha_4 \cdot (E_{23} + E_{32}) + \alpha_5 \cdot E_{33}, \\
& \beta_3 = \alpha_1 \cdot E_{23} + \alpha_2 \cdot (E_{21} + E_{33}) + \alpha_3 \cdot E_{31} + \alpha_4 \cdot E_{22} + \alpha_5 \cdot (E_{23} + E_{32}) + \alpha_6 \cdot E_{33}, \\
& \beta_4 = \alpha_2 \cdot E_{23} + \alpha_3 \cdot (E_{21} + E_{33}) + \alpha_5 \cdot E_{22} + \alpha_6 \cdot (E_{23} + E_{32}), \\
& \beta_5 = \alpha_3 \cdot E_{23} + \alpha_6 \cdot E_{22}, \\
& \beta_0 = -1.
\end{aligned} \quad (9)$$

In case that the orthotropy directions coincide with axes x and y , the expression of coefficients β_i simplifies and solving the equation becomes easier. Thus, if the angle θ is an integer multiple of the right angle, the coefficients β_2 and β_4 are equal to zero.

The problem is ultimately reduced to solving the equation (8); we propose the solving with finite differences. Therefore, in the point of coordinates (x, y) , which is the point (i, j) of the grid for calculating finite differences, we can write the following equation:

$$\begin{aligned}
& \frac{\beta_4}{4 \cdot h \cdot k^3} \cdot \Psi(i-1, j-2) + \frac{\beta_5}{k^4} \cdot \Psi(i, j-2) - \frac{\beta_4}{4 \cdot h \cdot k^3} \cdot \Psi(i+1, j-2) + \\
& + \frac{\beta_2}{4 \cdot h^3 \cdot k} \cdot \Psi(i-2, j-1) - \left(\frac{\beta_2}{2 \cdot h^3 \cdot k} - \frac{\beta_3}{h^2 \cdot k^2} + \frac{\beta_4}{2 \cdot h \cdot k^3} \right) \cdot \Psi(i-1, j-1) - \\
& - \left(2 \cdot \frac{\beta_3}{h^2 \cdot k^2} + 4 \cdot \frac{\beta_5}{k^4} \right) \cdot \Psi(i, j-1) + \left(\frac{\beta_2}{2 \cdot h^3 \cdot k} + \frac{\beta_3}{h^2 \cdot k^2} + \frac{\beta_4}{2 \cdot h \cdot k^3} \right) \cdot \Psi(i+1, j-1) - \\
& - \frac{\beta_2}{4 \cdot h^3 \cdot k} \cdot \Psi(i+2, j-1) + \frac{\beta_1}{h^4} \cdot \Psi(i-2, j) - \left(4 \cdot \frac{\beta_1}{h^4} + 2 \cdot \frac{\beta_3}{h^2 \cdot k^2} \right) \cdot \Psi(i-1, j) + \\
& + \left(6 \cdot \frac{\beta_1}{h^4} + 4 \cdot \frac{\beta_3}{h^2 \cdot k^2} + 6 \cdot \frac{\beta_5}{k^4} \right) \cdot \Psi(i, j) - \left(4 \cdot \frac{\beta_1}{h^4} + 2 \cdot \frac{\beta_3}{h^2 \cdot k^2} \right) \cdot \Psi(i+1, j) + \\
& + \frac{\beta_1}{h^4} \cdot \Psi(i+2, j) - \frac{\beta_2}{4 \cdot h^3 \cdot k} \cdot \Psi(i-2, j+1) + \\
& + \left(\frac{\beta_2}{2 \cdot h^3 \cdot k} + \frac{\beta_3}{h^2 \cdot k^2} + \frac{\beta_4}{2 \cdot h \cdot k^3} \right) \cdot \Psi(i-1, j+1) - \left(2 \cdot \frac{\beta_3}{h^2 \cdot k^2} + 4 \cdot \frac{\beta_5}{k^4} \right) \cdot \Psi(i, j+1) - \\
& - \left(\frac{\beta_2}{2 \cdot h^3 \cdot k} - \frac{\beta_3}{h^2 \cdot k^2} + \frac{\beta_4}{2 \cdot h \cdot k^3} \right) \cdot \Psi(i+1, j+1) + \frac{\beta_2}{4 \cdot h^3 \cdot k} \cdot \Psi(i+2, j+1) - \\
& - \frac{\beta_4}{4 \cdot h \cdot k^3} \cdot f(i-1, j+2) + \frac{\beta_5}{k^4} \cdot \Psi(i, j+2) + \frac{\beta_4}{4 \cdot h \cdot k^3} \cdot \Psi(i, j+2) = -f_y(i, j).
\end{aligned} \tag{10}$$

At each point of the finite-difference grid we write an equation like such. In these equations appear the values of Ψ function taken in the neighboring points, resulting in a system of equations to be solved in the $\Psi(i, j)$ nodal values.

For the boundary points, in (10) appear values of Ψ in some non-existing external nodes. These values also appear when we apply (10) for the nodes next to the boundary ones: these external nodes define a new virtual boundary beyond the physical one, increasing the number of the unknowns to be determined.

The system of equations can be solved only by writing boundary conditions: we will give these conditions in all boundary nodes, as prescribed displacements and/or loading forces.

The boundary conditions in the form of prescribed displacement

For easier applicability of this method, let's approximate the physical boundary with one which is made from horizontal and vertical lines adapted to the grid. In this case we define the boundary conditions as the projections of the displacement, as prescribed values of u and/or v . These projections are obtained by deriving of function Ψ , according to the relations:

$$\begin{aligned}
u &= \alpha_1 \cdot \frac{\partial^2 \Psi}{\partial x^2} + \alpha_2 \cdot \frac{\partial^2 \Psi}{\partial x \cdot \partial y} + \alpha_3 \cdot \frac{\partial^2 \Psi}{\partial y^2}, \\
v &= \alpha_4 \cdot \frac{\partial^2 \Psi}{\partial x^2} + \alpha_5 \cdot \frac{\partial^2 \Psi}{\partial x \cdot \partial y} + \alpha_6 \cdot \frac{\partial^2 \Psi}{\partial y^2},
\end{aligned} \tag{11}$$

If we express the value of u from the relation (11) with centered differences, we obtain the following equation:

$$\begin{aligned} & \frac{\alpha_2}{4 \cdot h \cdot k} \cdot \Psi(i-1, j-1) + \frac{\alpha_3}{k^2} \cdot \Psi(i, j-1) - \frac{\alpha_2}{4 \cdot h \cdot k} \cdot \Psi(i+1, j-1) + \\ & + \frac{\alpha_1}{h^2} \cdot \Psi(i-1, j) - \left(2 \cdot \frac{\alpha_1}{h^2} + 2 \cdot \frac{\alpha_3}{k^2} \right) \cdot \Psi(i, j) + \frac{\alpha_1}{h^2} \cdot \Psi(i+1, j) - \\ & - \frac{\alpha_2}{4 \cdot h \cdot k} \cdot \Psi(i-1, j+1) + \frac{\alpha_3}{k^2} \cdot \Psi(i, j+1) + \frac{\alpha_2}{4 \cdot h \cdot k} \cdot \Psi(i+1, j+1) = u(i, j). \end{aligned} \tag{12}$$

For v we obtain the same formula, the index of the α -s has to be increased by 3.

We can observe that applying the calculus for a grid node positioned on the boundary, it will be based on three points that are on the imaginary boundary. The concave corners will not raise issues or difficulties, however in the convex corners this scheme would include a point that does not belong to the imaginary boundary (figure 2).

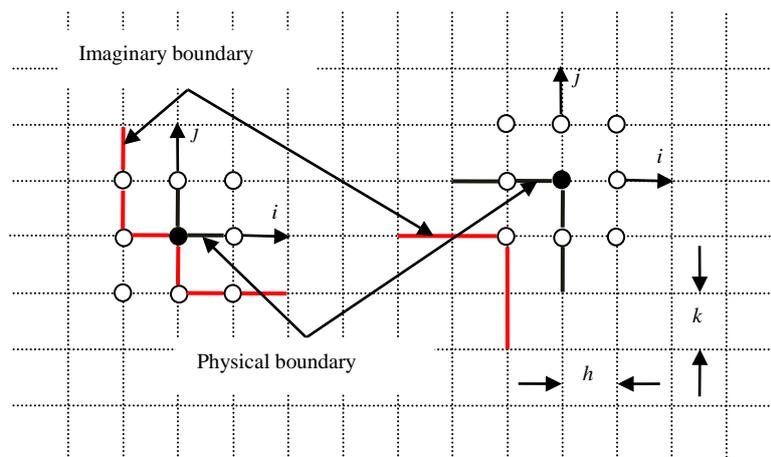


Fig. 2. The imaginary boundary that appears due to finite-difference approximation

In this case instead of centered difference approximation, we apply the derivatives' approximations with the help of forward or backward differences, depending on the corner position (Harangus, 2012).

The boundary conditions in the form of loading

Distributed stress, that loads the boundary is defined by its projections according to x and y directions, noted as p_x and p_y . This stress usually is described by an arbitrary function. During the mesh, this function is replaced by a step function (figure 3).

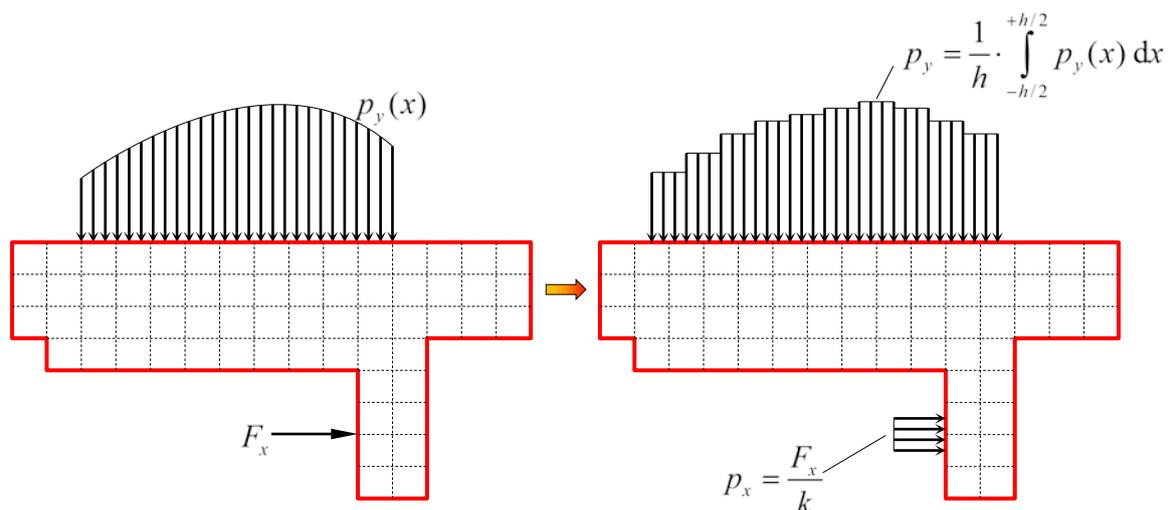


Fig. 3. Task discretion

If we cut an element from the area around a boundary point, the stresses along the boundary must be in balance with the external loads. Therefore we can write the relations which equal the projections of the exterior load with the stresses along the boundary (figure 4).

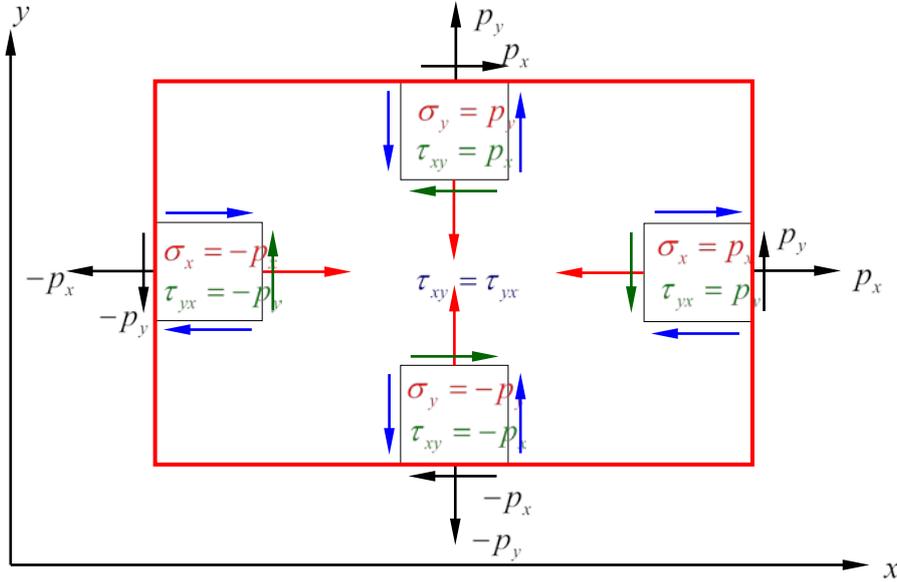


Fig. 4. Stresses along the boundary

Relations between specific stresses and strains are given by the generalized Hooke's law, the specific deformations and displacements of geometrical equations and the displacements and Ψ function by (11) formula. Thus we obtain the relation:

$$\begin{aligned}
 \sigma_x &= E_{11} \cdot \varepsilon_x + E_{12} \cdot \varepsilon_y + E_{13} \cdot \gamma_{xy} = \\
 &= E_{11} \cdot \frac{\partial u}{\partial x} + E_{12} \cdot \frac{\partial v}{\partial y} + E_{13} \cdot \frac{\partial u}{\partial y} + E_{13} \cdot \frac{\partial v}{\partial x} = \\
 &= E_{11} \cdot \left(\alpha_1 \cdot \frac{\partial^3 \Psi}{\partial x^3} + \alpha_2 \cdot \frac{\partial^3 \Psi}{\partial x^2 \cdot \partial y} + \alpha_3 \cdot \frac{\partial^3 \Psi}{\partial x \cdot \partial y^2} \right) + \\
 &+ E_{12} \cdot \left(\alpha_4 \cdot \frac{\partial^3 \Psi}{\partial x^2 \cdot \partial y} + \alpha_5 \cdot \frac{\partial^3 \Psi}{\partial x \cdot \partial y^2} + \alpha_6 \cdot \frac{\partial^3 \Psi}{\partial y^3} \right) + \\
 &+ E_{13} \cdot \left(\alpha_1 \cdot \frac{\partial^3 \Psi}{\partial x^2 \cdot \partial y} + \alpha_2 \cdot \frac{\partial^3 \Psi}{\partial x \cdot \partial y^2} + \alpha_3 \cdot \frac{\partial^3 \Psi}{\partial y^3} \right) + \\
 &+ E_{13} \cdot \left(\alpha_4 \cdot \frac{\partial^3 \Psi}{\partial x^3} + \alpha_5 \cdot \frac{\partial^3 \Psi}{\partial x^2 \cdot \partial y} + \alpha_6 \cdot \frac{\partial^3 \Psi}{\partial x \cdot \partial y^3} \right) = \\
 &= (E_{11} \cdot \alpha_1 + E_{13} \cdot \alpha_4) \cdot \frac{\partial^3 \Psi}{\partial x^3} + (E_{11} \cdot \alpha_2 + E_{12} \cdot \alpha_4 + E_{13} \cdot \alpha_1 + E_{13} \cdot \alpha_5) \cdot \frac{\partial^3 \Psi}{\partial x^2 \cdot \partial y} + \\
 &+ (E_{11} \cdot \alpha_3 + E_{12} \cdot \alpha_5 + E_{13} \cdot \alpha_2 + E_{13} \cdot \alpha_6) \cdot \frac{\partial^3 \Psi}{\partial x \cdot \partial y^2} + (E_{12} \cdot \alpha_6 + E_{13} \cdot \alpha_3) \cdot \frac{\partial^3 \Psi}{\partial y^3}. \tag{13}
 \end{aligned}$$

The boundary conditions are written by stresses with the help of Ψ function derivatives, rewriting these derivatives with finite-differences.

We exemplify determining conditions of the contour for the vertically side on the left, as follows (Harangus et al, 2012):

$$\begin{aligned}
& -\frac{c_4}{2 \cdot k^3} \cdot \Psi(j-2, i) - \left(\frac{c_2}{2 \cdot h^2 \cdot k} + \frac{c_3}{2 \cdot h \cdot k^2} \right) \cdot \Psi(j-1, i-1) + \\
& + \left(\frac{c_2}{h^2 \cdot k} + \frac{c_4}{k^3} \right) \cdot \Psi(j-1, i) - \left(\frac{c_2}{2 \cdot h^2 \cdot k} - \frac{c_3}{2 \cdot h \cdot k^2} \right) \cdot \Psi(j-1, i+1) + \\
& + \frac{c_3}{h \cdot k^2} \cdot \Psi(j, i-1) - \frac{c_1}{h^3} \cdot \Psi(j, i) + \left(\frac{3 \cdot c_1}{h^3} - \frac{c_3}{h \cdot k^2} \right) \cdot \Psi(j, i+1) - \\
& - \frac{3 \cdot c_1}{h^3} \cdot \Psi(j, i+2) + \frac{c_1}{h^3} \cdot \Psi(j, i+3) + \left(\frac{c_2}{2 \cdot h^2 \cdot k} - \frac{c_3}{2 \cdot h \cdot k^2} \right) \cdot \Psi(j+1, i-1) - \\
& - \left(\frac{c_2}{h^2 \cdot k} + \frac{c_4}{k^3} \right) \cdot \Psi(j+1, i) + \left(\frac{c_2}{2 \cdot h^2 \cdot k} + \frac{c_3}{2 \cdot h \cdot k^2} \right) \cdot \Psi(j+1, i+1) + \frac{c_4}{2 \cdot k^3} \cdot \Psi(j+2, i) = -p_x,
\end{aligned} \tag{14}$$

where

$$\begin{aligned}
c_1 &= E_{11} \cdot \alpha_1 + E_{13} \cdot \alpha_4, \\
c_2 &= E_{11} \cdot \alpha_2 + E_{12} \cdot \alpha_4 + E_{13} \cdot \alpha_1 + E_{13} \cdot \alpha_5, \\
c_3 &= E_{11} \cdot \alpha_3 + E_{12} \cdot \alpha_5 + E_{13} \cdot \alpha_2 + E_{13} \cdot \alpha_6, \\
c_4 &= E_{12} \cdot \alpha_6 + E_{13} \cdot \alpha_3.
\end{aligned} \tag{15}$$

Conclusions

This paper presents a finite-difference computational method for the integration of differential equations with partial derivatives describing the plane state of displacement or stress of the anisotropic materials. As shown in the paper, the problem can be expressed in stress leading to Airy function, which describes the second order partial differential stress field. With stress and material equations we can determine the specific strains. This method has the disadvantage of the impossibility to express directly the displacements.

By analogy with the Airy function, we used a "potential function" of the displacement, which made it possible to write mixed boundary conditions. The partial derivatives of this function are equal with the displacements in the directions of coordinate axes. Displacement derivatives, as in the derivatives of superior order displacement function give specific strains and by using material equations these superior order derivatives will lead to the stress field. Therefore becomes possible to write outline conditions as distributed load shape, there is a direct relationship (differential equations) between displacements and stresses. These relationships are approximated by finite differences.

In approximation with finite differences the real boundary was replaced by a boundary consisting of horizontal and vertical straight lines and the boundary conditions as prescribed loading led to some equivalence relations between loads and stresses. The denser the grid is, the more accurate the modeling of the load will be and the negative effects of the approximations made in the corner points will be more reduced. The disadvantage of this method is the fact that we can have body forces only in one direction.

References:

- Akanda M.A.S., Ahmed, S.R., Khan, M.R., and Uddin, M.W. A finite-difference scheme for mixed boundary-value problems of arbitrary-shaped elastic bodies. *Adv Eng Softw*, 31(3):173-84, 2000.
- Curtu, Ioan, and Nicolae Ghelmeziu. *Mecanica lemnului și materialelor pe bază de lemn. (Mechanics of wood and wood-based materials)*. București: Editura Tehnică, 1984.
- Harangus, Katalin. Model of orthotropic materials with the finite-difference method. *Research and Science Today, Academica Brâncuși Publishing*, 2(4):143-154, 2012.
- Harangus, Katalin, and Kakucs, András. Finite difference modeling of orthotropic materials. *Acta Universitatis Sapientiae, Electrical and Mechanical Engineering* 4, 2012.
- Kakucs, András. *A véges-elem módszer a szerkezetek számításában. (A finite-element method for the calculation of structures)*. Braşov: Editura Universităţii Transilvania, 2007.

- Kakucs, András. A véges-elem módszer alapjai. (Basics of the finite element method). Cluj-Napoca: Editura Scientia, 2007.
- Reaz, Ahmed S., Zubaer, Hossain M., and Wahhaj, Eddin M., A general mathematical formulation for finite-difference solution of mixed-boundary-value problems of anisotropic materials. *Computer and Structures*, 83:35-51, 2005.
- Szalai, József. A faanyag és faalapú anyagok anizotróp rugalmasság- és szilárdságtana. (Anisotropic elasticity and statics of the timber and wood-based materials).
- Szeidl, György. Boundary element methods in a dual formulation for plane problems of elasticity. The 3rd International Conference on Advanced Composite Materials Engineering, Braşov, 2010.
- Timoshenko, D., and Goodier, JN. *Theory of elasticity*. 3rd ed. New York: McGraw-Hill Book Company, 1979.

SIMULATION OF AN AXIAL FLOW TURBINE RUNNER'S BLADES USING CFD

Diaelhag Khalifa, M.Sc

Lecturer of Mechanical Engineering, King Abdul-Aziz University, Jeddah, KSA

Abstract:

This study aims to use computational fluid dynamics software package (CFX) to study and analyze flow's behavior in an axial flow turbine. This turbine used in low head and high flow rate hydropower plant. This study indicates that, performance of the designed blades is acceptable. These blades can be used in Kaplan turbine to produce a power with some addition design modification. These modifications can be made to reduce the tailing edge vortices (in suction side) of the blade and boundary layer forming in the leading edge of the blade. Also these modifications need furtherer analysis and more testing using another commercial CFD codes to investigate the results. Also this study focused to found the variations of velocity components and the pressure by average circumferential area (ACA) from inlet to outlet of the blades and used as factors to analyzed the flow inside the blades, the results of this analysis shows a good prediction of the flow behavior inside the blades and this lead to acceptable blade design, which can be used in Kaplan turbine.

Key Words: Designing, Simulation, CFD, Numerical, ANSYS, CFX10.0, Axial turbines, Verification, Visualization

Introduction

Hydro power plants generate one fifth of the total electrical power produced in world ^[1]. Even a small improvement of the hydrodynamic design and efficiency can contribute a great deal to the supply of the electric power. The efficiency of a hydropower plant depends on a number of parameters, such as: Turbine efficiency, Draft tube efficiency and Generator efficiency ^[2]. Most of the past studies have focused on the draft tube for increasing the efficiency of the plant, but a good draft tube design is not enough. Recent studies have shown that the efficiency improvement can also be realized by minor modification on the older design in the rest of the waterway i.e., in the draft tube, runner blades and spiral casing. Previous studies have shown that there is potential for increasing unit performance by a moderate modification of such runner blades. Runner blades have been found to have an efficiency loss due to the runner losses. A small increase in performance in these power stations represents a considerable economic value. This work will analyzed the flow field in the runner blades. This analysis is based on CFD simulations.

In axial flow turbine, water passes through the series of blade rows and changes its direction from radial to axial. Runner it the most important component of the turbine and its blade profile is designed at different sections from hub to casing to get the best performance and efficiency. The rotation of the runner and operation of the turbine either below or above the rated conditions cause variation of flow parameters from hub to tip. Hence, actual flow pattern in turbine space deviates from the simplifying assumptions made in design thus affecting the turbine performance. The experimental testing of turbine models at different operating regimes on specially designed test rigs is the conventional approach to assess the performance but its expensive and time-consuming tests.

The flow in axial flow turbine (Kaplan) is very complex including several flow phenomena, such as turbulence, separation, swirling flow and unsteadiness flow. Advanced fluid flows are described by the continuity and momentum equations, which can generally not be solved analytically. Therefore the numerical procedure in computational fluid dynamics (CFD) is of highest importance. CFD can be used to to check efficacy of alternate designs ^[3, 4] of turbines for optimization before final experimental testing of selected designs in resorted. However, in order to prove reliability of these tools for application to turbines, validations ^[5, 6] with known experimental results is required. In

present work, 3D viscous flow simulation with SST $k-\omega$ turbulence model is carried out in an experimental tested model of an axial flow hydraulic turbine using ANSYS CFX10.0 software. The variation of flow parameters from hub to tip of runner are presented in graphical form and average value of cascade parameters are computed at different operating regimes.

Definition of Geometry

The axial flow turbine consists of casing, stay rings, distributor, runner and draft tube. The energy transfer takes place in runner hence, present work is focused on runner blades only and therefore, analysis is carried from inlet to outlet where proper boundary condition can be applied. There are 12 stay vanes, 28 guide vanes and 6 runner blades in the model being analyzed. The blade rows of stay ring, distributor and runner are axi-symmetric and therefore, only single runner blade assembly is modeled for simulation using periodicity to minimize the total size of mesh. The flow parameters at inlet and exit of runner blade with velocity triangles are defined in Fig.1. The computational domain of runner blade is shown in Fig. 2, and the unstructured hexahedral mesh is generated in ANSYS workbench for all domains are shown in Fig. 3. The y^+ varies between 24 to 186, which is the acceptable range for automatic wall function treatment in boundary layer in SST $k-\omega$ turbulence model [7]. The maximum values other mesh quality parameters like face angle, edge length ratio, connectivity number are within acceptable limits of Ansys CFX10.0.

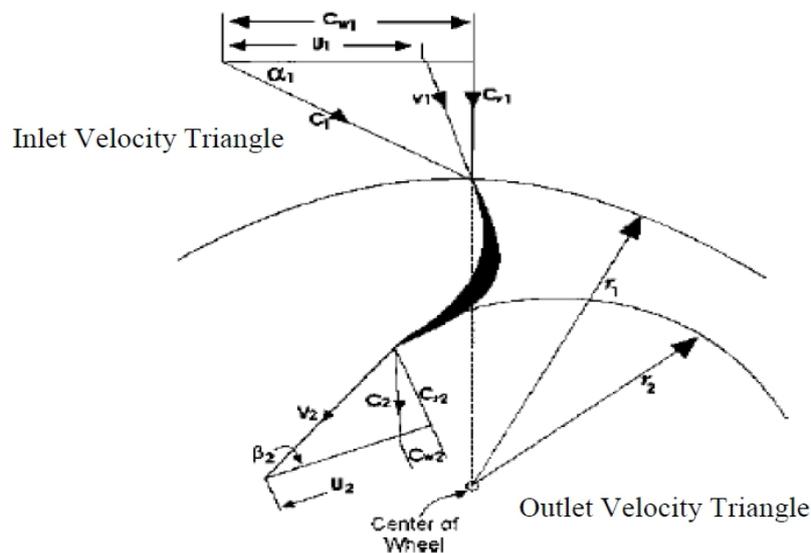


Fig. 1 – Velocity triangle at inlet and outlet of the runner blade

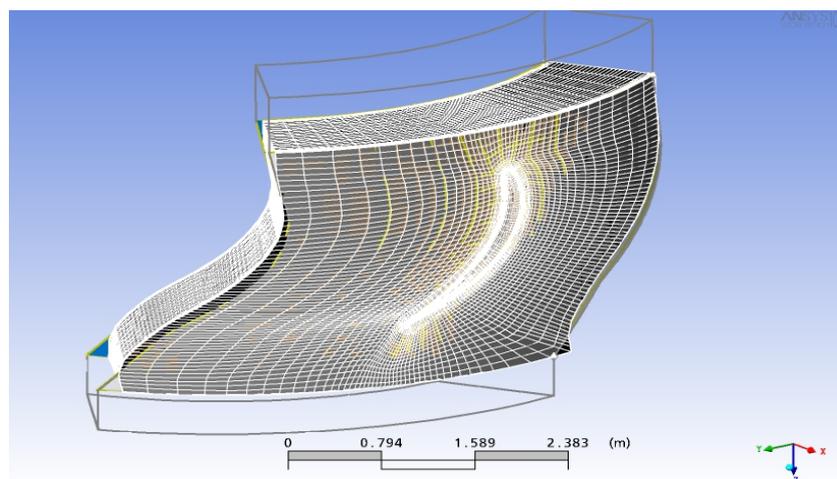


Fig. 2- Computational domain of runner blade

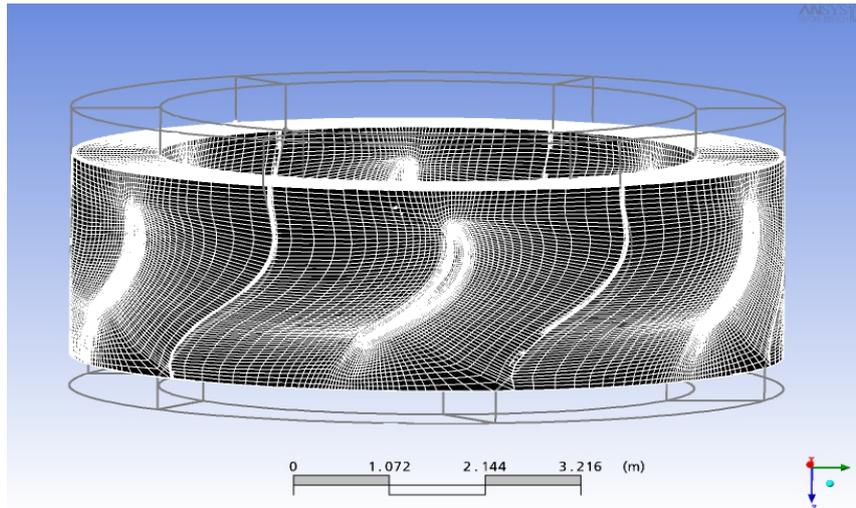


Fig. 3-Unstructured hexahedral mesh for all domains

Boundary Condition

The results obtained from the simulation in any flow domain depend on the specific boundary condition. In this work, the total flow rate and its direction are specified at stay vanes inlet as inlet boundary condition and static pressure specified at outlet of draft tube as outlet boundary condition. The rotational speed of runner is specified and other two blade rows are set stationary. All boundary walls are assumed smooth with no slip.

Governing Equations

The flow in the runner is assumed to be turbulent, and incompressible, the Reynolds Averaged Navier Stokes (RANS) equations consisting of continuity, momentum and energy equations are used. In order to solve these governing equations in Ansys workbench, the CFX10.0 solver has been utilized in this study.

$$\text{Continuity} \quad \frac{\partial u_i}{\partial x_i} = 0 \quad (1)$$

$$\text{Momentum} \quad \rho \frac{\partial u_i}{\partial t} + \rho \frac{\partial u_i u_j}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j} \left[\mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \right] \quad (2)$$

$$\text{Energy} \quad \rho \frac{\partial T}{\partial t} + \rho \frac{\partial T u_j}{\partial x_j} = \frac{\partial}{\partial x_j} \left(K \frac{\partial T}{\partial x_j} \right) \quad (3)$$

These equations are non-linear and their analytical solution is not possible and hence approximate solution is obtained by numerical method^[8].

Computation of Flow Parameters

The numerical analysis gives pressure and velocity distribution and the non-dimensional parameters are computed for presentation of results. The velocity components are divided by spouting velocity ($\sqrt{2gH}$) to get specific (non-dimensional) values of corresponding velocity. The following formulae are used for computation of different parameters^[9]:

$$\text{Pressure coefficient} \quad C_P = \frac{P-P_2}{\frac{1}{2}\rho W_2^2} \quad (4)$$

$$\text{Velocity coefficient} \quad C_V = \frac{W}{W_2} \quad (5)$$

$$\text{Flow deflection} \quad \varepsilon = \beta_1 - \beta_2 \quad (6)$$

$$\text{Degree of reaction} \quad \varphi = \frac{W_2^2 - W_1^2}{2gH} \quad (7)$$

$$\text{Circulation coefficient} \quad \tau = \frac{t(C_{U1} - C_{U2})}{D\sqrt{2gH}} \quad (8)$$

$$\text{Lift coefficient} \quad C_L = 2 \frac{t}{l} \sin \beta_m (\cot \beta_2 - \cot \beta_1) \quad (9)$$

$$\text{Runner energy coefficient} \quad \phi = \frac{gH_R D^4}{Q^2} \quad (10)$$

$$\text{Total Energy coefficient} \quad \psi = \frac{gHD^4}{Q^2} \quad (11)$$

$$\text{Total head} \quad H = \frac{TP_{SV1} - TP_{DET}}{\gamma} \quad (12)$$

$$\text{Head utilized by runner} \quad H_R = \frac{TP_1 - TP_2}{\gamma} - H_{FR} \quad (13)$$

$$\text{Hydraulic efficiency} \quad \eta_H = \frac{H_R}{H} * 100 \quad (14)$$

Results and Discussion

The CFD simulations are assumed converged when all the residuals are less than 10^{-7} , which is sufficient for most engineering problems. The velocity at points at the inlet, the centre and at the outlet is monitored and when there is no change in the results are considered converged. The distinct rise in the residual plot is due to the change in the differencing schemes.

The convergence of the SST turbulence model in this study is assumed converged when the residuals plots drop to 10^{-7} and the difference of mass flow in and mass flow out is very small compared to mass flow in. Fig (5.1) shows the residual plot of momentum and mass for three cases.

From the residual of mass and momentum for three cases, (A) the case of $225 * 10^3$ nodes is accepted because the residual plots is constant for long iteration and no change is the values of velocities components, (B) the case of $255 * 10^3$ nodes the residual plots is less than the residual target, (C) the case of $352 * 10^3$ nodes there is fluctuation in the value of v-component due to computer capabilities and grid quality, all the following results are taken for the case (B) of $255 * 10^3$ nodes.

Fig(5.2) show clearly the static pressure distribution between two successive blades of the turbine. Here, (it is very clear the high pressure exerted on the pressure side of blade and low pressure in blade suction side, this valid for all blades and according to the theorem of turbomachinery the pressure side is higher than suction side of the blade.

Results confirm that the pressure drops gradually from the inlet to outlet due to the extraction of fluid energy by the turbine runner.

Fig(5.3) show the pressure variation from inlet to outlet this means that the pressure decreased from inlet to leading edge and then increased at the leading edge after that decreased gradually along the blade until reach the trailing edge then increased to leaves with atmospheric pressure.

Fig(5.4) show the variation velocity component from inlet to outlet along the streamwise, the velocity component U in x-direction starting from small value near to zero and increased gradually along the streamwise (radial direction of blade), the velocity component V in Y- direction is very small starting from negative value and increased gradually along the streamwise until reach the value less than zero this in the direction of blade span, the velocity component W in Z-direction is semi constant along the streamwise this in the direction of rotational but the total velocity decreased from inlet to outlet due to rotation, swirling and vortex occur.

Fig(5.5) show the performance curves of the of blade under different operation condition, this indicate that the efficiency of the turbine blade increased with increasing of N rpm until reach maximum point and then decreased gradually (parabolic shape) this valid only for N from 120 to 150 rpm out of this range the efficiency is varies. The power also increased with the increase of N rpm to maximum also this is valid for N from 120 to 160 rpm.

Fig (5.6) show the streamlines at hub, shroud and blade, (A) at the hub this indicate that, there is vortex occur at the suction side of the blade near the trailing edge where the low pressure region. (B) at the shroud this means that the water flow path in the tip clearance. At (C) the streamlines in pressure surface of blade indicate the path of water flow over pressure side, (D) the streamlines at suction surface of the blade where the vortices occur and it's very clear at velocity vector.

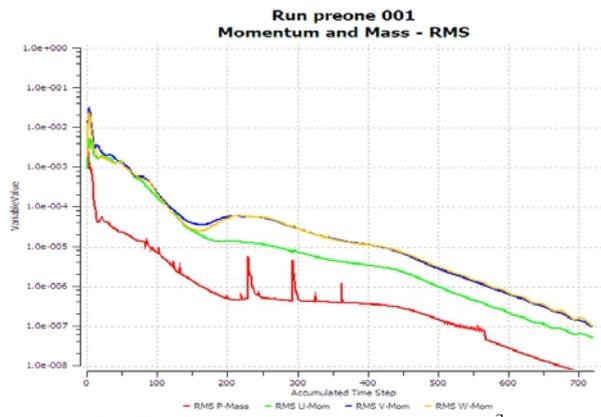


Fig. 5.1 Residual plots for (A) $225 \cdot 10^3$ nodes

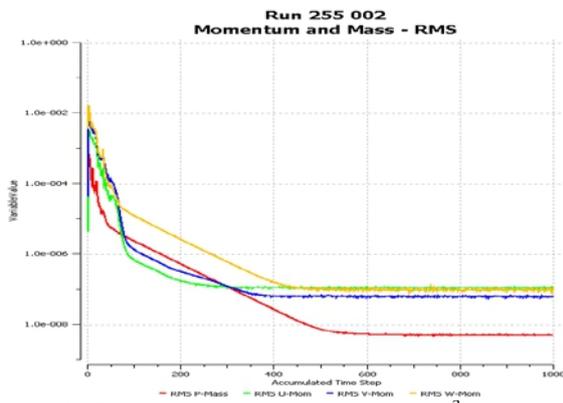


Fig.5.1 Residual plots for (B) $255 \cdot 10^3$ nodes

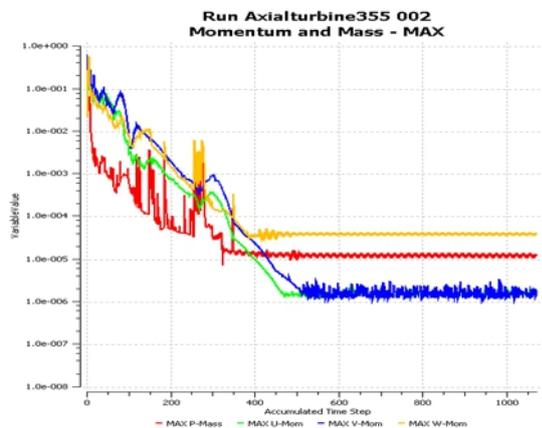


Fig: 5.1 Residual plots for (C) $352 \cdot 10^3$ nodes

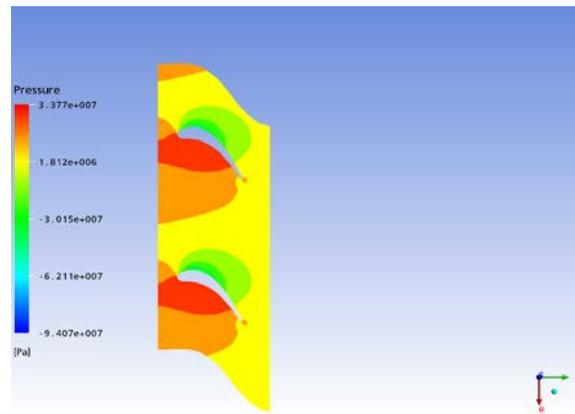


Fig: 5.2 Static pressure distribution on plane at mean radius

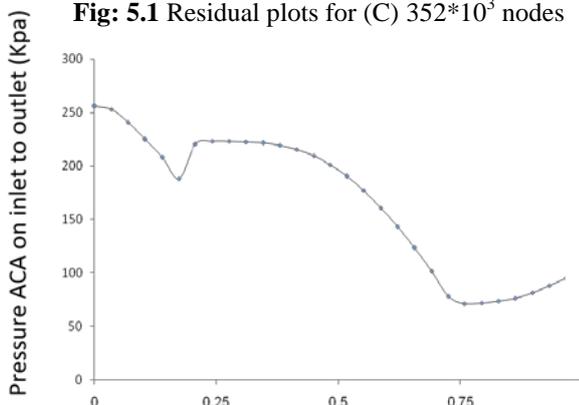


Fig: 5.3 Variation of Pressure ACA form inlet to outlet

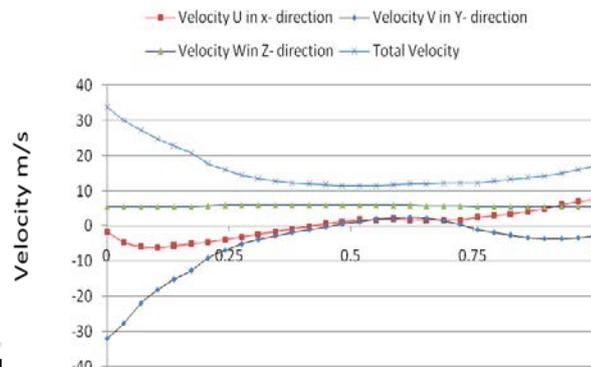


Fig: 5.4 Variation of velocities from inlet to outlet

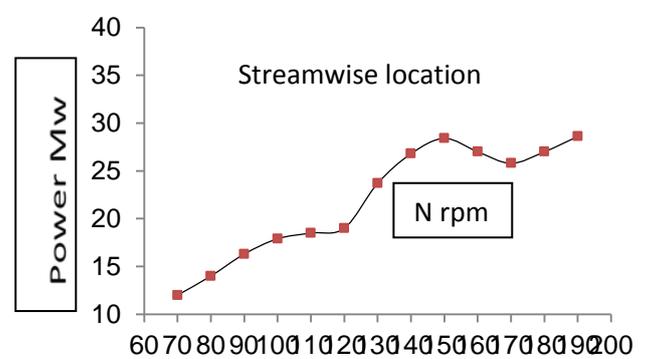
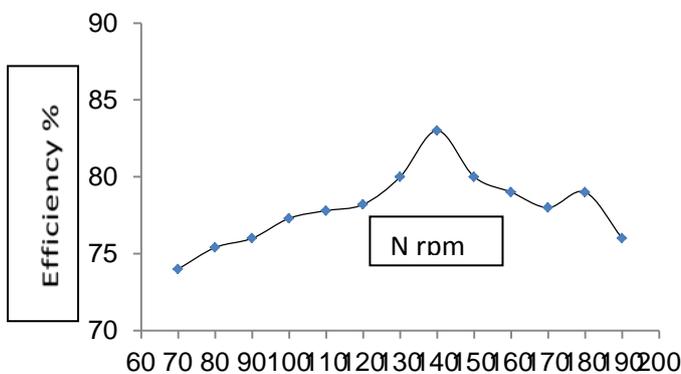


Fig (5.5) Performance Curves Efficiency and Power vs Rotational Speed

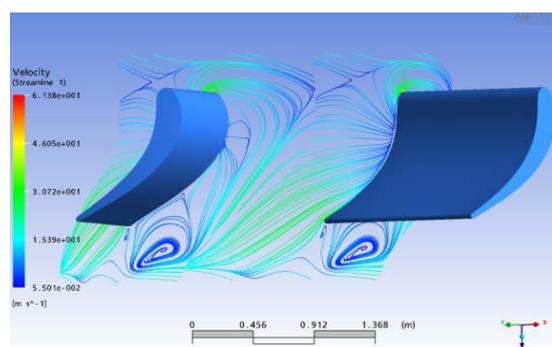


Fig: 5.6 (A)

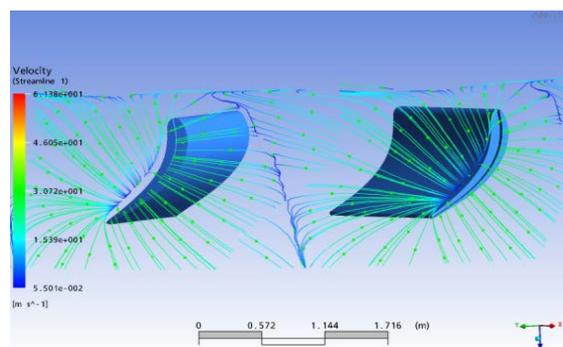


Fig: 5.6 (B)

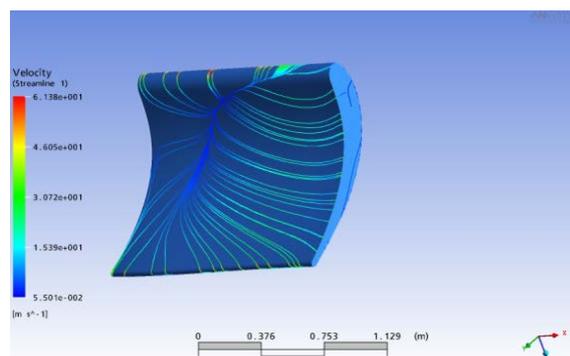


Fig: 5.6 (C)

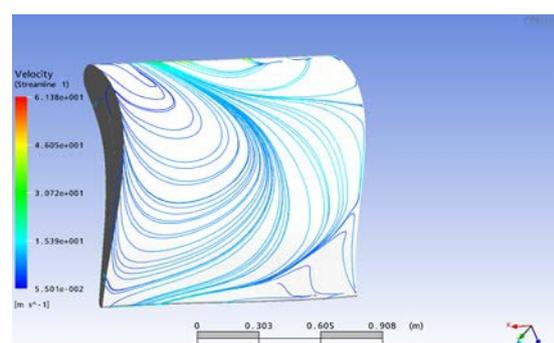


Fig: 5.6 (D)

Fig (5.6) streamlines (A) on Hub surface (B) on Shroud plane (C) On Blade-pressure surface (D) on Blade-suction surface.

Conclusion

The numerical simulation results shows similar pattern for velocity and pressure variation by average circumferential area (ACA) and the distribution between hub and shroud, efficiency and power output affected by the rotational speed of the runner. The maximum efficiency and power out but occurs at the same rotational speed. The total computed loss is minimum at the point of maximum efficiency. The streamline and pressure contour plots in different component confirm with actual flow behavior in axial flow turbine. The best operating regime can be easily identified from computed flow parameters, losses and flow pattern from numerical simulation. Hence, it is concluded that CFD approach can be used to study the flow pattern inside the turbine space and to optimize the design by different combinations of the design parameters and geometry at low cost in lesser time. Finally, the performances of optimized design need to be verified through model testing. This procedure will minimize time and the amount spent in development and optimization of hydraulic turbines.

Nomenclature

- C = absolute flow velocity (m/s)
- C_U = whirl velocity (m/s)
- D = diameter of turbine runner (m)
- G = gravitational acceleration (m/s^2)
- H = net head (m)
- H_{DTR} = head recovery in draft tube (m)
- H_{FR} = head loss in runner (m)
- L = blade chord (m)
- P = static pressure at any point on blade surface profile (Pa)
- Q = discharge through turbine (m^3/s)
- TP = total pressure at runner (N/m^2)
- TP_{SVI} = total pressure at stay vane inlet (N/m^2)
- TP_{DTE} = total pressure at draft tube outlet (N/m^2)
- T = Pitch of runner blades(m)

W = relative velocity at any point of blade surface (m/s)

α = guide vane angle from tangential direction ($^{\circ}$)

β = relative flow angle ($^{\circ}$)

β_m = mean relative flow angle ($^{\circ}$)

γ = specific weight of water (N/m³)

ρ = mass density of water (kg/m³)

subscript 1 and 2 denote values of parameters at inlet and outlet of runner respectively

References :

- 1- World Energy Council, “*Survey of Energy Resources*”, Technical Report, Sep.2006.
- 2- Krivchenko G., “*Hydraulic Machines: Turbines and Pumps*”, 2^{ed.}, Lewis, Boca Raton. 1994.
- 3- Guoyi P, J “*Fluid Engineering*”, 27 (2005)1183-1190
- 4- Wu J, Shimmel K, Tani K, Niikura K & Sato J, “*Fluid Engineering*”, 127 (2007) 159- 168.
- 5- Shukla, M, *CFD analysis of 3-D Flow and its validation for francis turbine*, M.Tch. Thesis, Maulana Azad National Institute of Technology, Bhopal, 2007
- 6- Rao V S & Tripathi S K , “*Pro Nat Seminar on CFD-The 3rd Dimension in Flow analysis & Thermal Design*”, Bhopal, India, (2007) 196-201.
- 7- ANSYS, “*Ansys CFX10.0 Help Manual*”, Bangalore: ANSYS Software Pvt. Ltd, 2005.
- 8- Lewis, R.I., “*Turbo Machinery Performance Analysis*”(Arnold London), 1996.
- 9- Raabe I J, “*Hydro Power- The Design, use and Function of hydro mechanical hydraulic and electrical equipment*”(VDI-Verlag, GmbH, Dusseldorf), 1985.

A SYSTEM DYNAMICS MODEL TO ASSESS THE OPTIMAL NUMBER OF CREWS FOR CONDUCTING AIR CARGO OPERATIONS

R. Pozzi

C. Noe

T. Rossi

School of Engineering, LIUC - Università Cattaneo, Castellanza, Italy

Abstract:

Notwithstanding the extensive studies involving the cargo terminal operations issues, a knowledge gap emerges. Literature has focused its research on the manpower assignment and the crew scheduling problems, whereas the determination of the cycle time of the activities to be performed, and, as a consequence, the job lead time, has been often neglected. Thus, filling such a gap is the goal of the present paper, which is aimed at developing and testing a model for defining the cycle time involved in air cargo terminal activities and the corresponding job lead time, in relation to both job characteristics and terminal resources. The causal diagram that depicts such relations is built and applied to a real process. Then, the assumptions based on the model are validated by means of data gathering and statistical analyses (ANOVA).

Key Words: Air cargo terminal, Activity cycle time, Job lead time, ANOVA

Introduction

In recent years, air transport has become one of the main modes for global freights transportation. Such a phenomenon is due to several reasons. From an economical point of view, according to the explanations proposed by (Yuan, Low, & Ching Tang, 2010), the change in consumption habits, entailing the shortening of product life cycles, can be pointed as enhancing the importance of just-in-time logistics and the need of fast transportation mode. From a management point of view, as supply chains have become global, the need of sourcing products and parts from allover the world, while not increasing the delivery time, has involved the need of a fast transport mode, once more. Of course, many other reasons, as the development of e-commerce may have affected and are going to influence the use of such a transport mode (Zhang, 2003). Due to these reasons, air cargo industry has become the fastest growing sector in the dynamic freight market (Coyle et al., 2003), and the forecasted growth rate of 5.8% per year over the next 20 years recalled by (Azadian, Murat, & Chinnam, 2012) hints at the increase of its importance.

One of the direct consequences of such a phenomenon is the increase in number of the air cargo handling companies (Coyle et al., 2003), and, as a result, of the competition within the air cargo sector. Such a high pressure on handling companies is leveraged, on the one hand, by the fact that customers request shorter and shorter lead time and, on the other hand, by more and more unpredictable demand. In order to survive in such a context, air cargo terminals address the challenge to handle large volumes of items efficiently and responding quickly to customer demands (i.e. ensuring both efficiency and effectiveness at the same time is the main issue) (Taylor, Choy, Chow, Poon, & Ho, 2012).

How can air cargo terminals decrease operating costs while maintaining customer service levels (i.e. restrained lead time)? In last years Literature has addressed the problem. In particular, researches on air cargo operations have pointed manpower costs as the largest part of all operating costs, and, consequently, the efficient management of manpower resources has been extensively studied (Rong & Grunow, 2009). Among the others, the crew scheduling planning problem has been broadly reviewed. Several approaches for solving the problem starting from the terminal manpower requirements (i.e. the number of crews needed within a certain time horizon) have been suggested (Yan, Chen, & Chen, 2006b), (Yan, Chen, & Chen, 2006a), (Rong & Grunow, 2009).

How to determine optimal terminal manpower requirements? Notwithstanding the extensive study of the cargo terminal operations problem, Literature lacks in the requirements determination. Nevertheless, properly assessing optimal manpower requirements is paramount to ensure both efficiency and effectiveness. As a matter of fact, in air cargo terminals the crews are usually composed of temporary workers. Every week the air cargo handling company requests to temping agencies a certain number of temporary workers, which depends on the number of crews needed in that week. So if such a number is overestimated or underestimated, despite the crew scheduling process, the air cargo handling company will incur higher costs or will not be able to face the customers demand respectively.

General personnel scheduling is frequently encountered in literature, and it can be defined as the problem of assigning a set of tasks to be performed in a given planning horizon to a group of workers, such that each crew does not exceed a limit on the total time it can spend working (Beasley & Cao, 1996). (Ernst et al., 2004) develop an exhaustive problem taxonomy and problems definition, providing a general framework for classifying much of the effort that has been carried out in the area of personnel scheduling. According to the authors, several application areas can be identified and, among them, the air cargo terminals are pointed. According to them, demand has to be first modeled determining how many staffs are needed at different times over some planning period. With reference to crew demand, (Ernst et al., 2004) state the importance of including a demand modeling step, even though many researches assume demand as either given or easily obtained and others refer to either forecast or heuristic models. Caprara et al. (2003) propose an approach that finds the solution of the manpower management problem in two steps. The determination of the minimum number of employees and the working days for each employee has to be first performed, then, the assignment of the duties to each employee has to be carried out. In such a case, manpower requirements are partially determined. Again, the authors do not completely perform the determination of manpower requirements. (Suryadi & Papageorgiou, 2004) deal with the allocation of maintenance crews to maintenance activities, and they address the problem applying mixed-integer non-linear programming. The authors exploit the statistic nature of equipment failure, and the content of work required in a certain period of time is determined by the piecewise constant function of the units failure rate. As a matter of fact, the presented approach is limited to a maintenance context, and not applicable to an air cargo one.

According to the review performed by (Ernst et al., 2004), personnel scheduling is already well established in the transportation sector in general and in the aviation industry in particular, and air cargo terminals crew scheduling is a widely discussed topic. Personnel scheduling bases on both the determination of the total content of work and the design of efficient assignment. With reference to the first sub-topic, the one investigated by the present research, both (Nobert & Roy, 1998) and (Rong & Grunow, 2009) make use of “demand leveling” to move a certain amount of freight service away from the peak times avoiding idle capacity before determining manpower requirements. On the other hand, according to (Nobert & Roy, 1998), since shipments’ characteristics (e.g. weight, number of pieces per shipment, handling unit) vary often, work standards are meant to be difficult to obtain, resulting in incorrect total content of work. (Rong & Grunow, 2009) distinguish service demand in build-up and break-down in the presented model. In the presented case study although demand of the outbound and inbound cargo in each hour is specified, time consumption parameters are not defined, hiding how to find the total content of work out.

As system dynamics (hereinafter SD) models are proved to be appropriate for studying manufacturing systems, offering lens on operations management systems (Georgiadis & Michaloudis, 2012) (Größler, Thun, & Milling, 2008), and, to our knowledge, the problem of the total content of work determination has not been solved, the present paper is aimed at developing a SD model to fill the above mentioned knowledge gap.

According to the proposed model, the total content of work within a certain time horizon depend on the cycle times characterizing the activities of the air cargo terminal operating process that must be executed for handling the jobs (i.e. the customer orders), which fall on the considered time horizon. In turn, the content of work of a job, which is approximated by its lead time, depends on both job characteristics and manpower involvement. The proposed model has been applied to a real-world terminal cargo managed by an Italian handling company with the aim to study the export process.

The paper is organized as follows: Section 2 outlines the proposed model; Section 3 describes the empirical case study the model is applied to and its validation by means of analysis of variance (hereinafter ANOVA); Section 4 provides concluding remarks and future research steps.

The proposed model

The aim of the proposed model is to express the relationships between both job characteristics and terminal resources and manpower requirements, by developing a causal diagram.

Variables and relationships

The independent variables of the proposed model refer both to job characteristics and to cargo terminal manpower involvement. With reference to job characteristics, the considered variables are:

- number of packages: the number of packages that compose the job;
- type of vehicle: the mean of transport where packages are loaded on;
- packages position: how the packages are positioned on the mean of transport;
- pallet height;

With reference to cargo terminal decisions, just one variable is considered:

- number of operators: the number of operators in the crew (i.e. working on the same job at the same time);

As the dependent variable of the proposed model is the lead time of the job and it equals the sum of the cycle times of each activity, the system variables should define the cycle times.

A causal diagram outlines the relationships between the independent and the dependent variables. In the diagram, system variables are impacted by job characteristics and/or cargo terminal facilities, impacting themselves on the lead time of the job. Whether when a variable is increasing/decreasing the impacted one is increasing/decreasing accordingly the impact is 'positive' or, on the contrary 'negative'.

Hypothesis

The system under consideration is an air cargo export process. The latter is composed of five subsequent activities: truck unload, control, pallet preparation, positioning and covering. For each job, all the activities have to be carried out. For each activity, its cycle time determines the amount of time needed by a crew to carry it out.

This model describes the relations between the independent variables and the lead time of the job. Based on the above hypothesis, we develop the causal diagram in figure 1.

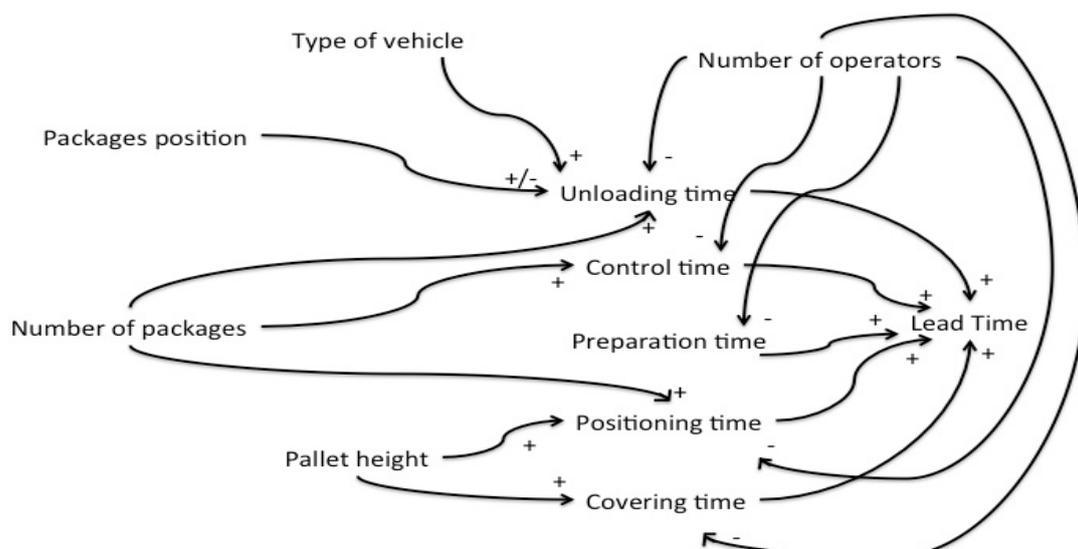


Figure 1 - Proposed model (causal diagram)

Model description

We will outline the effects of the decision variables, which are independent, on the dependent variable (i.e. lead time). In particular, as stated above, the independent variables are ‘number of packages’, ‘number of operators’, ‘type of vehicle’, ‘packages position’, and ‘pallet height’.

Starting from describing in detail the relations for the ‘number of packages’, we define ‘unloading time’ as the amount of time it takes to empty the vehicle. By definition, the higher number of packages the truck transports, the higher unloading time. Thus, in our opinion, the ‘number of packages’ positively influences ‘unloading time’. Similarly, the variable ‘number of packages’ positively influences ‘control time’, which deals with the amount of time needed to control the arrived packages. With reference to ‘positioning time’, by definition, the higher number of packages to be positioned, the higher positioning time. The latter is defined as the amount of time needed to place all the packages on the pallet, so that they can be loaded on the air cargo. Again, in our opinion, the variable ‘number of packages’ positively influences ‘positioning time’. Summarizing, the ‘number of packages’ positively influences ‘unloading time’, ‘control time’, and ‘positioning time’.

The ‘number of operators’ has an impact on ‘unloading time’: when the number of operators in the crew increases, the ‘unloading time’ decreases. Thus, the model represents a negative influence. Similarly, ‘number of operators’ negatively impacts on ‘control time’ and ‘positioning time’. With reference to ‘preparation time’, that is the amount of time needed to arrange the pallet where the packages are then placed, it decreases when the number of operators in the crew increases. The covering of the pallet, that requires a ‘covering time’, takes less time when the number of operators in the crew increases. Summarizing, negatively influences ‘unloading time’, ‘control time’, ‘positioning time’, ‘preparation time’, and ‘covering time’.

The ‘type of vehicle’ is involved just in the vehicle unloading activity. In our opinion, as different vehicles have different characteristics (e.g. dimensions), these influence the ‘unloading time’. Defining a vehicle ranking on their dimensions, and assuming that in a bigger vehicle a higher number of packages can be contained, it is reasonable to state that unloading a van would take less time than unloading an engine. Thus, the ‘type of vehicle’ influences ‘unloading time’ positively.

When packages to be exported are loaded on the vehicle, their optimal positioning is not always respected. Thus, whether ‘packages position’ is not correct, the ‘unloading time’ increases, while, whether they are correctly positioned the ‘unloading time’ decreases. Thus, the ‘packages positioning influences ‘unloading time’ positively or negatively.

With reference to ‘pallet height’, as it increases, the time to place an additional package to the others increases (i.e. ‘positioning time’ increases). It is reasonable to assume that placing packages at a low level is easier than placing them in a high position. Thus, ‘pallet height’ positively influences ‘positioning time’. Similarly, it is reasonable to assume that covering a high pallet is more difficult (i.e. involves more time) than covering a lower one. Summarizing, ‘pallet height’ positively influences both ‘positioning time’ and ‘covering time’.

Model validation

The causal diagram presented in figure 1 is validated in the next paragraphs by studying through data gathering and ANOVAs the effects of each independent variable on the state variables.

The validation procedure

The validation phase is performed by gathering as many activities cycle times as it is enough to perform statistical analyses. Moreover, the cycle times are gathered throughout the twenty-four hours of the day, so that all the situations are covered. The one-way ANOVAs on the data collected are performed by means of Minitab statistical software. The threshold p-value is set equal to 0,05 (5%). Thus, whether the p-value is lower than the threshold the variable is meant to influence statistically the amount of work of micro phase.

ANOVAs results

For each activity, the ANOVAs deal with contents of work and changes in supposed influencing variables. Thus, the results of the analyses include: the variables considered as influencing according to the model, the average contents of work, their standard deviation, the numbers of measurements and the calculated p-values.

Since all the activities are analyzed through the same procedure, an example of ANOVA is presented below (table 1, table 2, table 3, table 4, and table 5), while a results summary for all the activities is presented. The example deals with the “unload” activity, as it is the most representative

case. According to the proposed model, four variables influence ‘unloading time’: (i) type of vehicle; (ii) number of packages; (iii) number of operators; (iv) packages position.

Table 1 Example of ANOVA on Type of vehicle

Type of vehicle	Content of work mean [minutes]	Content of work standard deviation [minutes]	# Measurements	p-value	Statistical relevance
Van	4.35	3.41	26	0	YES
Engine	8.72	7.39	60		
Tractor-trailer	21.65	15.21	49		

The ANOVA confirms the relation modeled, as the content of work is higher when the type of vehicle to unload has bigger dimensions.

With reference to ‘number of packages’, as the types of vehicles differ in capacity, the variable is presented in relation to each type of vehicle.

Table 2 Example of ANOVA on Number of packages

Type of vehicle	Number of packages	Content of work mean [minutes]	Content of work standard deviation [minutes]	# Measurements	p-value	Statistical relevance
Van	1≤n<6	3.52	2.55	27	0	YES
	5<n<11	6.80	5.02	5		
Engine	1≤n<6	3.72	3.10	25		
	5<n<11	9.79	4.34	28		
	10≤n<16	13.67	5.92	9		
	15<n<21	23.00	9.99	2		
	>20	45.00	0	1		
Tractor-trailer	1≤n<6	3.89	1.76	9		
	5<n<11	11.2	4.73	10		
	10≤n<21	21.36	13.6	14		
	20<n<31	31.15	9.62	13		
	30<n<51	44.67	11.02	3		
	n>50	102.50	2.12	2		

Again, the ANOVA confirms the relation modeled (i.e. the content of work increases when the number of packages increases).

With reference to the variable ‘number of operators’, as the types of vehicles differ in dimensions, it is presented in relation to each type of vehicle.

Table 3 Example of ANOVA on Number of operators

Type of vehicle	Number of operators	Content of work mean [minutes]	Content of work standard deviation [minutes]	# Measurements	p-value	Statistical relevance
Van	1	3.83	3.21	24	0.038	YES
	2	9.00	2.82	2		
Engine	1	8.21	6.54	52	0.072	NO
	2	13.50	12.92	8		
Tractor-trailer	1	21.75	22.55	40	0.318	NO
	2	29.30	13.70	10		

In such a case, the ANOVA deny the relation that is proposed by the model (i.e. the negative influence of the variable ‘number of operators’ on ‘unloading time’). In particular, in case two operators unload a van, the unloading cycle time increases with statistical relevance. With reference to engine and tractor-trailer, the cycle time increases as well, but no statistical relevance is registered. The ANOVA results, at first sight, appear nonsense. On the other hand, it is reasonable that two operators unloading packages from a van get in the way each other, increasing significantly the cycle time. With reference to the engine and the tractor-trailer, as the throughput space is larger, the operators working together still increase the cycle time, but not significantly.

With reference to ‘packages position’, it is not strictly related to the type of vehicle.

Table 4 Example of ANOVA on Packages position

Packages position	Content of work mean [minutes]	Content of work standard deviation [minutes]	# Measurements	p-value	Statistical relevance
Correct	16.62	11.46	37	0	YES
Not correct	45.17	29.47	12		

The ANOVA confirms that positioning influences the ‘unloading time’.

Table 5 provides a summary of the relations between influencing variables and the system variable ‘unloading time’.

Table 5 Summary of Unloading time

Variable	Relation from the model	Validated?
Type of vehicle	Different vehicles have different characteristics (e.g. dimensions), these influence the ‘unloading time’	Yes
Number of packages	Positive influence	Yes
Number of operators	Negative influence	No
Packages position	Whether it is not correct, the ‘unloading time’ increases, while, whether packages are correctly positioned the ‘unloading time’ decreases	Yes

As previously stated, with reference to the other activities (i.e. control, preparation, positioning, and covering), their summary tables are presented (table 6, table 7, table 8, and table 9).

In relation to ‘control time’, the positive influence of the variable ‘number of packages’ in the system variable is confirmed. On the other hand, no statistical difference is registered between one or two operators working on the same job. This can be due to limited the control space.

Table 6 Summary of Control time

Variable	Relation from the model	Validated?
Number of packages	Positive influence	Yes
Number of operators	Negative influence	No

In relation to ‘preparation time’, the negative influence of the number of operators on the cycle time is validated.

Table 7 Summary of Preparation time

Variable	Relation from the model	Validated?
Number of operators	Negative influence	Yes

In relation to ‘positioning time’, the positive influence of the variables ‘number of packages’ and ‘pallet height’ is confirmed by the ANOVA. With reference to the ‘number of operators’, the negative influence is not significant. This can be due to the amount of time needed to set the positioning, which does not depend on the number of operators involved.

Table 8 Summary of Positioning time

Variable	Relation from the model	Validated?
Number of packages	Positive influence	Yes
Pallet height	Positive influence	Yes
Number of operators	Negative influence	No

With reference to the pallet covering activity, the positive relation identified by the model is confirmed by the ANOVA. With reference to the ‘number of operators’, the negative influence is not significant.

Table 9 Summary of Covering time

Variable	Relation from the model	Validated?
Pallet height	Positive influence	Yes
Number of operators	Negative influence	No

Conclusions and directions for future research

This paper focuses on topical and current managerial issues. As a matter of fact, air cargo terminal efficiency gained high significance in last years, due to economical and managerial points of view. With reference to the first, change in consumption habits, entailing the shortening of product life cycles is cited by literature. From a managerial point of view, the enlarging of the supply chains requires faster means of transportation. Thus, the paper is aimed at proposing a model for explaining the relations between both job characteristics and terminal resources and manpower requirements. By performing statistical analyses, i.e. one-way ANOVA, on the activities gathered cycle times, the relations between independent variables and system variables have been studied. On the basis of these, both general and particular conclusions are can be derived from the present research.

Referring to the model, it can be easily validated by means of data gathering and the performing of ANOVA by means of Minitab statistical software. On the other hand, since the amount of data to perform an effective ANOVA is quite high, the data gathering is a time consuming activity. Moreover, as the air cargo terminals are usually in service overnight, the cycle times has to be performed accordingly.

With reference to particular conclusions, the variable influence of the ‘number of operators’ on the cycle times deserves attention. As it is reasonable to infer that a higher number of operators reduces the cycle times, the ANOVA results deny this statement. In fact, some activities cycle time are disadvantaged from a higher number of operators, because of tight work spaces.

We reckon that the proposed model presents some limitations. First of all, not all the possible independent variables have been identified. Of course, many others influence the lead time of the job. It would be interesting to analyze whether there is a different effect of the lead time (i.e. on the system variables that compose the lead time). We believe these could be further steps of the research.

COMPENSATION OF VOLUMETRIC EFFICIENCY BY TURBOCHARGING IN AN INSULATED DI DIESEL ENGINE WITH ALCOHOL AS FUEL

S.Sunil Kumar Reddy

Associate Professor, Mechanical Department, N.B.K.R.I.S.T, Vidyanagar, Nellore, A.P

Dr. V. Pandurangadu

Professor, Mechanical Department, Jawaharlal Nehru Technological University, Anantapur. A.P

Abstract:

With the stringent emission legislation all over the world an intensive search is going on for diesel fuels which produce more environmental pollution. This simulated interest in alcohol-fueled diesel engines because it is renewable bio-based resource and it is oxygenated, thereby providing the potential to reduce particulate emissions in diesel engines and ethanol diffusion flames produce virtually no soot. But due to alcohols higher latent heat of vaporization, the higher temperatures available in the Insulated engine (IE) are used for vaporizing the alcohols and this decreases the ignition delay and aids combustion. With the higher temperatures in the insulated engines there will be drop in volumetric efficiency and this further increases the frictional horse power due to thinning of lubricant. So the volumetric efficiency drop is to be compensated by turbocharging to improve the thermal efficiency of insulated engine. Hence the present experimental work is planned accordingly. In the turbocharger the exhaust gases are expended in a turbine which is further coupled to engine. Further the turbine forces more air, proportionately more fuel into the combustion chamber, thereby enhancing the power output without increasing the cylinder capacity. Experiments are carried out in a single cylinder water cooled DI diesel engine with an air gap insulated piston, air-gap insulated liner and ceramic coated cylinder head and valves (Insulated engine) and the reduction in the volumetric efficiency is compensated with turbocharging. The insulated diesel engine with turbocharging gave the better performance and reduced smoke. Further the engine performance is also studied with the raise of intake boost pressure.

Key Words: Alcohols, Insulated engine, Super charging, Turbocharging, Volumetric efficiency

Introduction

In most of the diesel engines maximum amount of heat is lost to the cooling medium. With the lower temperatures in the combustion chambers the alcohols will not burn because of its higher latent heat of vaporization. So, insulated engine was developed with the insulation of the combustion chamber components. With this the heat rejection to the cooling medium is reduced and the same thing can be recovered in the form of useful work by expanding the exhaust gases in lower pressure turbines. Some important advantages of the insulated engines are improved fuel economy, reduced HC and CO emission, reduced noise due to lower rate of pressure rise and higher energy in the exhaust gases [2 & 3]. However, one of the main problems in the insulated engines is the drop in volumetric efficiency. This decrease in the volumetric efficiency is attributed to the decrease in the density of air entering the cylinder because of high wall temperatures of the insulated engine. The degree of degradation of volumetric efficiency depends on the degree of insulation. J.Cheong et al (7) conducted experiments on high speed direct Injection (HSDI) diesel engine and concluded that air mass is increased by 10-20% low speed range. With that smoke was reduced and fuel consumption was improved with the same fuel delivery and injection pressure.

Rakopoulos et al (8) developed a computer analysis for the study of performance of turbocharge diesel engine, operating under transient load conditions and it is validated with experimental results. Naser et al (9) concluded that the reduction was with the reduction of cylinder volume. With the increase of intake pressure the performance gains will be reduced due t decrease in

density and this can be compensated with intercooler. He developed a MAT LAB program to find the effect of intercooler on a multi-cylinder engine for operation at constant speed of 1600 rpm. Eyub et al (10) conducted experiments with turbocharger with intercooler and found that the engine power output can be increased by 154% with ideal intercooler while single turbocharger without intercooler can only increase 65%. Also the size of the cylinder can also be reduced to half by means of turbocharging and inter cooling.

Yashvir et. al., (11) conducted experiment on LML 125 cc and analyzed parameters like torque, power and specific fuel consumption with turbocharger. It is observed that power and torque of the engine increases from 7 to 11 KW and 9 to 13 NM at 7500 and 9000 RPM respectively. Ghodke et al (12) conducted experiment on turbocharger with intercooler. In the intercooler he reduced the outlet temperature to ambient and also increased the oxygen content in to the engine cylinder which leads to faster burn rates and ability to control exhaust emissions.

Objective

The main objective of this investigation is to experimentally determine effect of turbocharger on volumetric efficiency in an Insulated engine with alcohol as fuel at a reduced fuel injection pressure of 165 bar. The total experiment consists of

- (i) Insulated engine components preparation
- (ii) Experimental details
- (iii) Results
 - (a) Investigations with different piston materials – selecting the best piston material
 - (b) Investigations with turbocharger on the above best piston material

Insulated Engine Components Preparation

For the development of the Insulated engine (IE) in the present investigations the coating material selected must withstand for higher temperature and should also have sufficient strength. Among all the coating materials searched the partially stabilized Zirconium (PSZ) is found to be quite useful for adiabatic engine application because of its excellent insulating characteristics, adequate strength and thermal expansion characteristics [1, 3]. The insulated engine developed is having an air gap piston and liner, PSZ coated cylinder head and valves. The coating thickness on the components was based on the theoretical analysis and recommendations made by Wong et al [13]. The insulation methodology is explained in detail as follows.

Insulated Air Gap Piston

In this a 2 mm air-gap (whose thermal conductivity is low) is provided between a metallic crown and the standard piston made of Aluminum alloy. This air gap is optimized based on the literature available [1]. The metallic crown and standard piston were separated by copper and steel gaskets. Figure.1 shows the air gap insulated piston with brass insert.



Figure1: An air-gap insulated piston with brass insert

Insulated Cylinder Head and valves

The combustion chamber area of the cylinder head and the bottom surfaces of the valves are machined to a depth of 0.5 mm and are coated with PSZ material for the same depth [3]. The details of cylinder head and valves are as shown in the Figure 2.



Figure 2: PSZ coated cylinder head and valves

Insulated Cylinder Liner

A thin mild steel sleeve is circumscribed over the cast iron liner maintaining a 2mm layer of air in the annular space between the liner and the sleeve [2]. The joints of the sleeve are sealed to prevent seepage of cooling water into the air-gap region.

Experimental Details

A stationary, four stroke, 3.68 Kw direct injection Kirloskar water cooled single cylinder diesel engine is used to conduct experiments. If the engine is operated at normal injection pressures the amount of alcohol injected (due to low viscosity) in to the engine will be more and further it may cool the engine due to its higher latent heat of vaporization. So the fuel injection pressure is reduced to 165 bar for the experiment (12). With the high self ignition temperature of alcohol it takes more time for the vaporization. So the injection timing is made advanced to 27° bTDC. All the tests are conducted at the rated speed of 1500 rpm. The concentrations of smoke are measured with Bosch smoke meter; Air suction rate and exhaust air flow rates are with an air box method. Temperatures at the inlet and exhaust valves are monitored using Nickel-Nickel Chromium thermocouples. Time taken to consume 20 cc of fuel was noted using a digital stop watch. Engine RPM is measured using an electro-magnetic pick up in conjunction with a digital Indicator of AQUATAH make. The experimental set up used is as shown the following Figure.3.



Figure3. Experimental set up of Insulated Engine Test Rig

From the past literature it is observed that among various losses of heat in the combustion chamber, maximum heat transfer occurs through the piston. So in order to reduce the heat transfer, a piston is designed (similar to that of original piston) which is capable of absorbing heat from the hot combustion gases during the peak cycle temperature conditions and gives out the same to the incoming fresh charge during the suction and compression strokes of the next cycle. This improves the combustion efficiency and further thermal efficiency. So in order to retain the heat in the combustion chamber, to reduce the heat losses from the piston crown to the piston skirt and further to the exhaust a thermal barrier piston crown is designed with brass (BP) due to its lower thermal conductivity.

Further the combustion efficiency can be increased with good turbulence in the combustion chamber. So an attempt is made in this work with nine number of grooves on the brass crown piston in an insulated engine. The size of the groove on the piston crown is selected in such a way that maximum number of grooves can be generated. This brass crown is further knurled to increase its

surface area thus facilitating better heat transfer rate from the hot gases to the crown. The photographic views of the following two pistons are as shown in the figures (Fig 4 & 5).



Figure: 4 Photo Graphic of Brass crown piston (BP)



Figure: 5 Photo Graphic View of Brass Crown Piston with 9 grooves (BP9)

These crowns are same in the size of the original piston and can be adapted without any major modifications to the original design. In the present work investigations are carried out on brass crown piston (BP) and brass crown piston with nine grooves (BP9) in order to find the best one at which the insulated engine gives maximum performance and the same is compared with aluminum piston.

Investigative results with Brass piston in an Insulated engine

Initially the tests are performed at a constant speed of 1500 rpm in a normal engine at a constant injection timing (29° bTDC), the fuel flow rate and the load is varied. For all the operating conditions, the cooling water and lubricating oil temperatures are maintained constant throughout the experiment. For the insulated engine, due to higher operating temperatures and further lower ignition delays with insulation in the combustion chamber, the injection timing of 27° bTDC is found to give the optimum performance. The above tests are repeated in the insulated engine with alcohol as fuel.

In the first stage the engine is run with aluminum piston. Then the same is run with brass piston and brass piston with nine grooves. This is for finding the best one at which the volumetric efficiency drop is more due to higher temperatures in the combustion chambers. The results obtained are presented in the following graph.

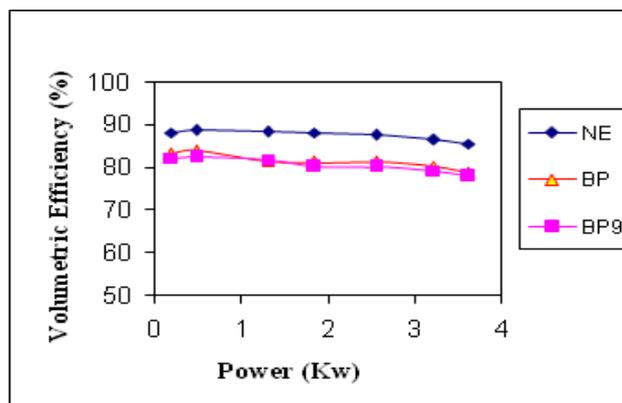


Exhibit 1: Comparison of Volumetric Efficiency with Power output

The effect of brass piston with different configurations on the volumetric efficiency in the insulated engine depicts in Exhibit 1. Due to higher operating temperatures, with insulated components, the intake air is heated to a higher temperature and consequently the mass of air drawn in each cycle is lower, resulting in a decrease in volumetric efficiency. For BP9 of the insulated engine the drop in volumetric efficiency is more and is about 1.2% as compared to BP and about 9.5% compared to normal engine at rated load.

So the drop in power output of an insulated engine can be compensated by turbocharging. Further the investigations are carried out with brass piston with nine grooves at which the volumetric efficiency is low due to higher operating temperatures with turbocharging equipment.

Investigative results of Brass piston with nine grooves (BP9) in an Insulated engine with turbocharging equipment

Generally the Turbochargers used are of forced induction type. They compress the air flowing into the engine so that the engine squeeze more air into the cylinder, and further more fuel can be added. Therefore, we get more power from the engine than the same engine without the charging. This improves the power to-weight ratio for the engine. In order to achieve this boost, the turbocharger uses the exhaust flow from the engine to spin a turbine, which in turn spins an air pump. Since it will be connected to the exhaust, the temperatures in the turbines are also very high. For the present investigations to pressurize the inlet air, internally powered turbocharging equipment with closed loop lubrication is fabricated. The schematic diagram of the turbocharging equipment is shown in Fig: 6. In the turbocharging the high temperature exhaust gases are expanded in a low pressure turbine for the power generation and this is further coupled to motor of the compressor [4, 5]. This compressor compresses the inlet air and supplies to the engine at slightly higher pressure. By controlling the inlet air, the engine is turbocharged at different inlet pressures.

In the present investigations the turbocharging equipment is connected to the insulated engine exhaust. The insulated engine is further run with brass piston with nine number of grooves (BP9) at various boost pressures. The results are summarized below.

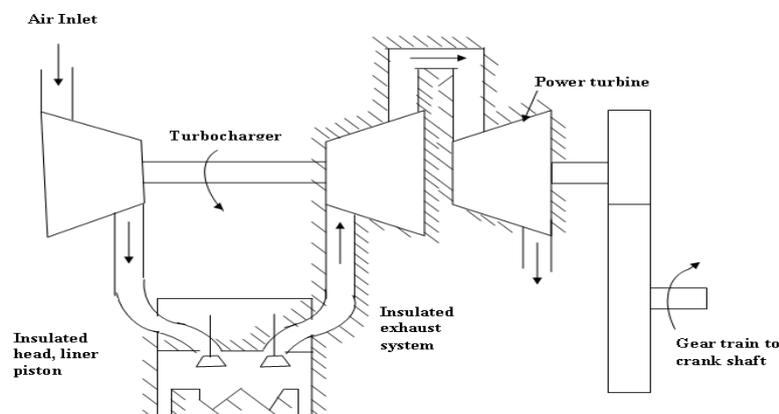


Figure 6. Turbocharged Insulated Diesel Engine

Effect of Turbocharging on the Volumetric Efficiency

The variation of volumetric efficiency with power output with intake boost pressure is shown in exhibit: 2. As the boost pressure is increased, more air is available for the combustion which further increases the combustion efficiency. At higher boost pressures excess air doesn't improve the combustion efficiency [12]. So the optimum boost pressure is 790 mm of Hg at which the drop in volumetric efficiency is compensated with turbocharger. The intersection points of the volumetric efficiency curves of turbocharger and the normal engine gives the inlet pressure of the turbocharger at which the volumetric efficiency drop is compensated. Because of the increase

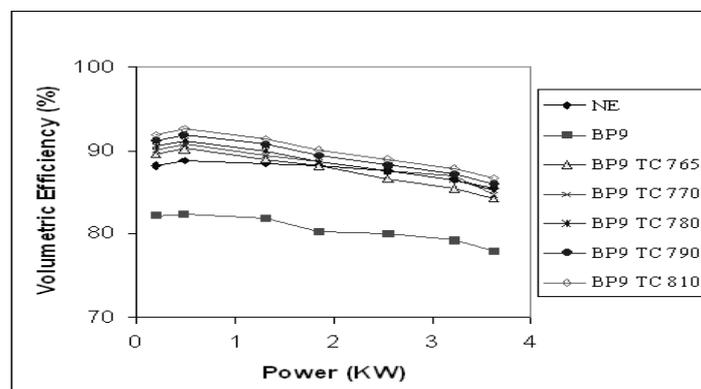


Exhibit 2 Variation of Volumetric Efficiency to power with Turbocharging

back pressure with turbocharging conditions, the inlet boost pressures are higher for compensating the volumetric efficiency drop in normal engine. It requires nearly 4% of intake boost pressure under turbocharging conditions for compensating the maximum efficiency drop of 10% in the normal engine. With the turbocharging the volumetric efficiency is compensated. The percentage of boost pressure required for the volumetric efficiency compensation is shown in the following exhibit.3.

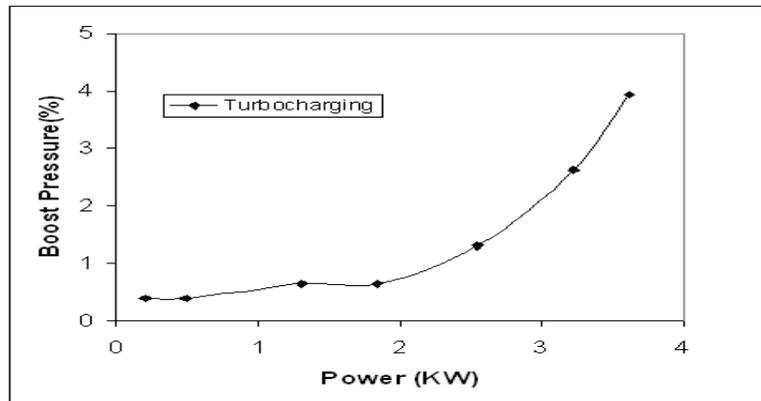


Exhibit 3. Comparison of Percentage of Boost Pressure Required for Volumetric Compensation with Power Output in Turbocharging

Brake Thermal Efficiency

The exhibit 3 shows the variation of brake thermal efficiency with power output for turbocharged condition. When the engine is turbocharged thermal efficiency is improved continuously with load. The maximum improvement is about 4.3% over insulated engine.

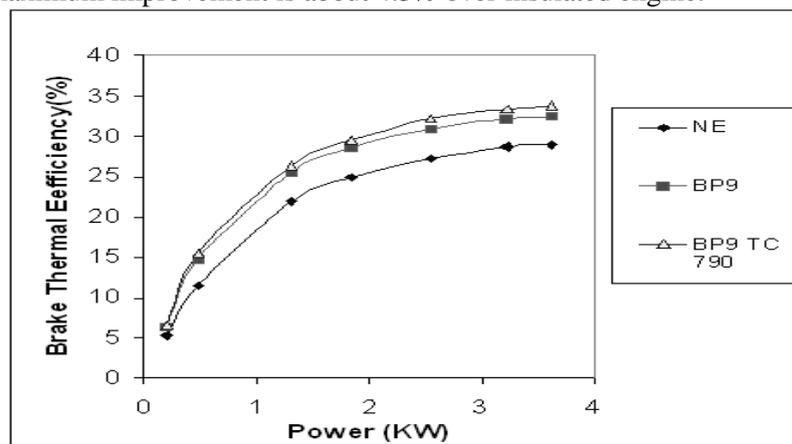


Exhibit 4. Comparison of Brake thermal Efficiency with Power Output for Volumetric Efficiency Compensation with Turbocharging

The improvement in thermal efficiency under turbocharging conditions is marginal due to following reasons:

- In the present work inlet boost pressures in turbocharging are moderate, because they were selected on the basis of volumetric efficiency compensation. At higher pressures still higher thermal efficiency could be obtained.
- Higher frictional losses due to increase in gas pressures.
- The engine had stability problem at higher intake pressures.

Combustion Parameters

In general due to less air availability in the combustion chamber the combustion will not be complete. But with the turbocharging more air will be available for the combustion and this will change the combustion parameters. The variations in the combustion parameters are shown below.

Peak Pressure

The peak pressure variation of turbocharging with power output is shown in exhibit 5. The peak

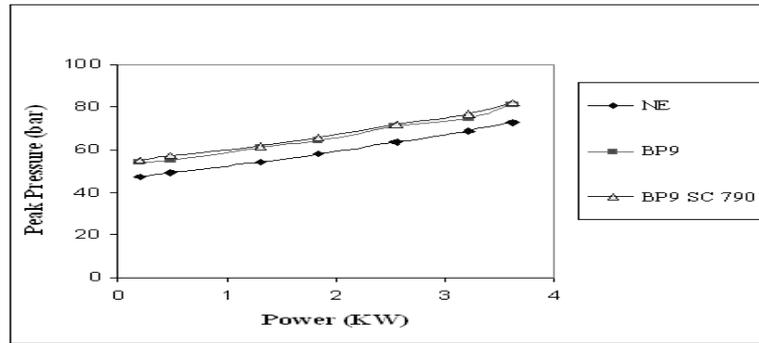


Exhibit 5 Comparison of Peak Pressure with Power Output for Volumetric Efficiency Compensation with Turbocharging

pressures of normal engine, Insulated engine and turbocharged Insulated engines are compared in the same figure. It is observed that the peak pressures are higher with turbocharged engine and is about 82 bar at the rated load. This is due to the complete combustion in the chamber.

Ignition Delay

The variation of ignition delay with power output for turbocharging conditions is shown in exhibit 6. With the turbocharging more amount of air enters into the combustion chamber which increases the combustion process and reduces the ignition delay. But at higher loads due to the high latent heat of alcohol, the ignition delay is slightly increased [6]. It is concluded that there is

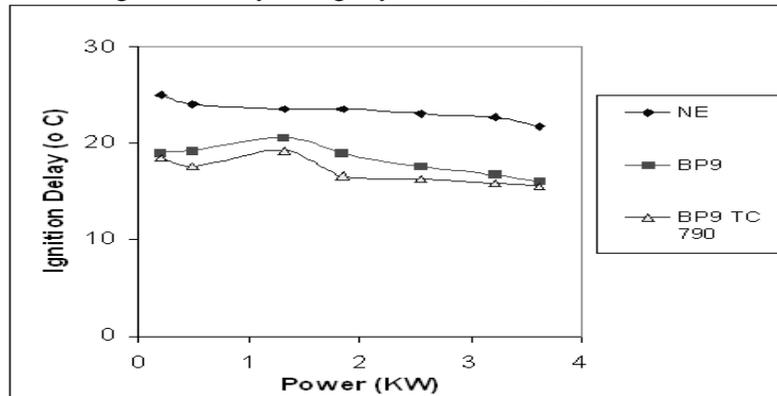


Exhibit 6. Comparison of Ignition Delay with Power Output for Volumetric Efficiency Compensation with Turbocharging

a reduction of 6.2^o CA for the turbocharged insulated engine compared to normal engine at rated load. So it will be beneficial to increase the turbocharging pressures in order to have a shorter ignition delays.

Exhaust Temperature

The increase in the exhaust temperatures are 20^oC to 50^oC with turbocharging compared to without charging. This is due to the increase of mass flow rate of air, reduction in the ignition

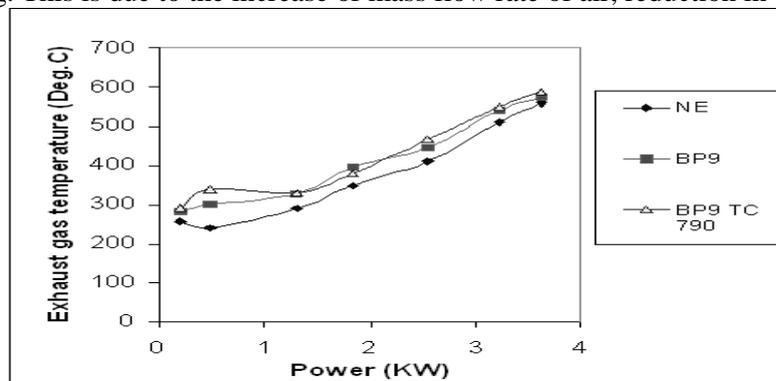


Exhibit 7 Comparison of Exhaust Gas Temperature with Power Output for Volumetric Efficiency Compensation with Turbocharging

delay and hotter combustion chamber which further increases the combustion process. Exhibit 7 shows the variation of exhaust temperatures with power output.

Smoke Number

Exhibit.8 shows the variation of exhaust smoke number with power output for turbocharging condition. It is concluded from the graph that there is a significant reduction in smoke level in turbocharged engine compared to normal engine at rated load condition due to complete combustion. As alcohol is bio based and is having rich in oxygen, literally it produced no soot.

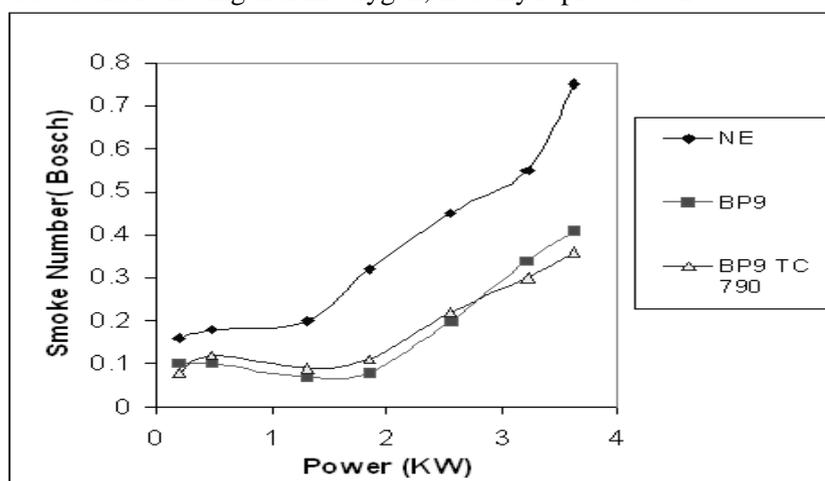


Exhibit 8 Comparison of Smoke Emissions with Power Output for Volumetric Efficiency Compensation with Turbocharging

Conclusions

The following conclusions are drawn based on the experimental investigations on an insulated diesel engine under turbocharging conditions:

1. The increase in the intake boost pressure improves the brake thermal efficiency of the engine.
2. For the compensation of drop in volumetric efficiency of the insulated engine 4% intake boost pressure is required for turbocharging.
3. Though the higher temperatures are available in the combustion chamber due to insulation, the increase in exhaust gas temperature is marginal. This is attributed to the higher latent heat of vaporization of alcohol.
4. As the alcohol contains oxygen and more air is available in the turbocharging for combustion, the ignition delay is reduced.
5. Due to the complete combustion of alcohol at higher temperatures the smoke emissions are also marginal.

References:

1. Miyari, Y., Matsnisa, T., Ozawa, T., Oikawa: "Selective heat insulation of combustion chamber wall, for a DI diesel engine with Monolithic Ceramics", SAE 2010141
2. Domingo, N. and Graves R.L. "A study of Adiabatic Engine Performance", on National Laboratory report under preparation (2003).
3. TKobori, Shigeharu "Combustion in LHR diesel engine", JSME international Journal, series II, Vol 35, No.1, 1992.4.
Michael C. Brands, "Vehicle Testing of Cummins Turbocompound Diesel Engine", SAE 1981, Paper 810073.
4. S.Sunil kumar reddy, Dr.V.Pandurangadu "effect of ceramic Insulation on DI diesel engine with various piston inserts", Vol.3, No.III/IV, IJMSEA, 2009
5. Mohd Muqem, Manoj Kumar, "Turbocharging of IC Engine: A Review", International Journal of Mechanical engineering and Technology, January, 2013, volume 4, pp 132-149

6. J. Cheong, Sunghwan Ch, Changho Kim, "Effect of Variable Geometry Turbocharger on HSDI Diesel Engine", Seoul 2000 FISITA World Automotive Congress June 12-15, 2000, Seoul, Korea.
7. C.D. Rakopoulos, E.G. Giakoumis, "Availability analysis of a turbocharged DI diesel engine operating under transient load conditions" *Energy* 29 (2004) 1085–1104, www.elsevier.com/locate/energy.
8. Naser B. Lajqi, Bashkim I. Baxhaku and Shpetim B. Lajqi, "Effect of Intercooler on Turbocharged Diesel Engine Performance" 13th International Research/Expert Conference, Trends in the Development of Machinery and Associated Technology" TMT 2009, Hammamet, Tunisia, October 2009.
9. Eyub Canli, Selcuk Darici and Muammer Ozgoren " Intercooler Effect on Conventional Supercharging Systems" International Scientific Conference, 19-20 November, 2010, Gabrovo
10. Yashvir Singh, Nishant Kr. Singh, Rakesh Prasad and Hemant Kr. Nayak, "Performance Analysis of Supercharging of Two Wheelers" *International Journal of Mechanical Engineering and Technology (IJMET)*, ISSN 0976 – 6340(Print) ISSN 0976 – 6359(Online) Volume 2 Issue 2, May – July (2011), pp. 63-69.
11. Pundlik R. GHODKE and J. G. SURYAWANSHI, "Advanced Turbocharger Technologies for High Performance Diesel Engine - Passenger Vehicle Application" *International Journal of Mechanical Engineering and Technology*, ISSN 0976 – 6340 (Print) ISSN 0976 – 6359 (Online), Volume 3, Issue 2, May-August (2012), pp. 620-632.
12. Ozer Can, smet Çelikten and Nazım Usta," Effects of ethanol addition on performance and emissions of a turbocharged indirect injection Diesel engine running at different injection pressures ", SAE Technical paper, 2004.
13. Wong, V., Kamo, R "Assessment of Thin Thermal Barrier Coatings for I.C. Engines", SAE 950980

FAULT DIAGNOSIS OF INDUCTION MOTOR

Sanjana Malhotra, Associate Professor

Manav Rachna College of Engineering, Faridabad, India

M. K. Soni, Executive Director

Manav Rachna International University, Faridabad, India

Abstract:

Although, Induction motors are highly reliable, they are susceptible to many types of faults that can become catastrophic and cause production shutdowns, personal injuries, and waste of raw material. Induction motor faults can be detected in an initial stage in order to prevent the complete failure of the system and unexpected production costs.

The purpose of this paper is the analysis of various faults of inverter fed induction machine. The laboratory tests thus conducted have been reported, and it is hoped that the research investigations reported would be very useful to the Power Electronics circuit industry.

Key Words: PWM Inverters, DC Motors, Drive Systems, Open & Short Circuit Test

Introduction

The study of induction motor behaviour during abnormal conditions due to the presence of faults; and the possibility to diagnose these abnormal conditions has been a challenging topic for many electrical machine researchers. Induction motor has been established as the workhorse of industry ever since the 20th century. Speed control of AC motors has been a continuously pressing requirement of industry, so as to ensure better production with a high degree of qualitative consistency. Although recent developments in Power Electronics and Controls have brought forth some very significant drive alternatives like the Switched Reluctance motor, Permanent magnet and Brushless DC Motor; these have not yet become very popular and cost effective for a wide range of applications, especially in the damp-proof, dust-proof and flame-proof environment. Therefore, the widespread use of induction motors is still economically viable as well as popular, and is likely to continue for the next few decades.

Variable speed drives are widely used in all application areas of industry. These include transport systems such as ships, railways, elevators, conveyors; material handling plants and utility companies for mechanical equipment e.g. machine tools, extruders, fans, pumps and compressors. The penetration of variable speed ac drives into these sectors has been further accelerated by the development of new power semiconductor devices and drive concepts, which further allow new functions and performance characteristics to be realized. The application of new Power Electronic components has also initiated a significant change in the market breakups between AC drives and DC drives. The rugged construction of AC drives has opened up a host of new application areas, thereby providing the user and also the manufacturer additional potential to increase their productivity and thus maintain their economic and technical competitive advantages.

Concept Of Drive Systems

While comparing the dynamic performance of a separately excited DC motor with that of an Induction motor, the latter presents a much more complex control plant. This is due to the fact that the main flux and armature current distribution of a DC motor is fixed in space and can be controlled independently; whereas in the case of AC motor, these quantities are strongly interacting. This design constraint makes the induction motor drive structure more complex and non-linear. The drive hardware complexity increases as more and more stringent performance specifications are demanded by the user. The complexity further increases because of the variable frequency power supply, AC signal processing and relatively complex dynamics of the AC machine.

PWM Inverters

One of the best possible methods to control the torque and speed of induction motor is to implement variable voltage and variable frequency inverters. Inverters used for variable speed drive applications should have the capability of varying both the voltage and frequency in accordance with speed and other control requirements. The simplest method to achieve this control is through a six step inverter. But this method suffers from the following limitations:

- (i) Presence of low order harmonics, because of which the motor losses are increased at all speeds causing derating of the motor.
- (ii) Torque pulsation is present at low speeds, owing to the presence of lower order harmonics.
- (iii) The harmonic content increases at low speeds, thus increasing motor losses. Also the increase in V/f ratio at low speed to compensate for the stator resistance drop may cause a higher motor current to flow at light loads due to saturation. These effects may overheat the machine at low speeds.

These limitations of a six step inverter drive are overcome in a pulse width modulated (PWM) inverter. The basic block diagram of a PWM inverter is shown in figure 1.

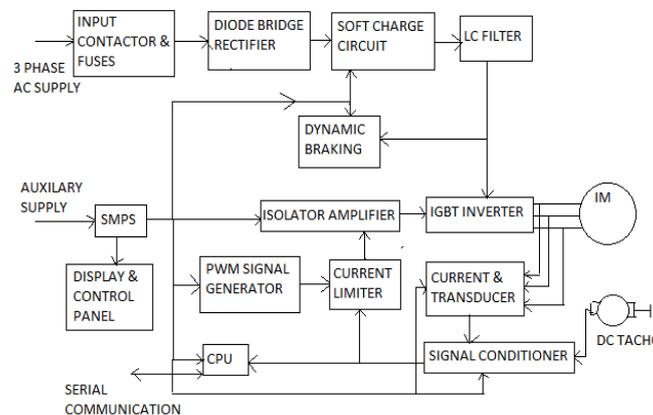


Figure 1. Block Diagram Of Inverter System

Because of a low harmonic content in the output voltage of diode bridge and also due to the presence of harmonics in the input current of a PWM inverter, the requirement of filter size in such systems is small. The drive system consequently delivers smooth low speed operation, free from torque pulsation, thus leading to lower derating of motor and higher overall efficiency. Also because of a constant DC bus voltage, a number of PWM inverters with their associated motors can be supplied from a common diode bridge. However, these advantages are obtained at the expense of a complex control system and higher switching loss due to high frequency operation.

Survey Of Various Faults

A wide range of motors are currently being used for industrial applications. They deliver a wide range of characteristics demanded for specific tasks. Motors for all types of duties and with various characteristics require adequate protection. Hence it is essential that the characteristics of motors be carefully examined and considered before applying protection systems. A three-phase voltage fed inverter can develop various types of faults as shown in figure 2.

- Input supply single line to ground fault F_1 .
- Rectifier diode short circuit fault F_2 .
- Earth fault on DC bus F_3 .
- DC link capacitor short circuit fault F_4 .
- Transistor base drive open fault F_5 .
- Transistor short circuit fault F_6 .
- Line to line short circuit at machine terminals F_7 .
- Single line to ground fault at machine terminals F_8 .
- Single phasing at machine terminals F_9 .

A three phase voltage fed inverter can develop any of the above stated faults, out of which the open base drive and shoot through are the most common.

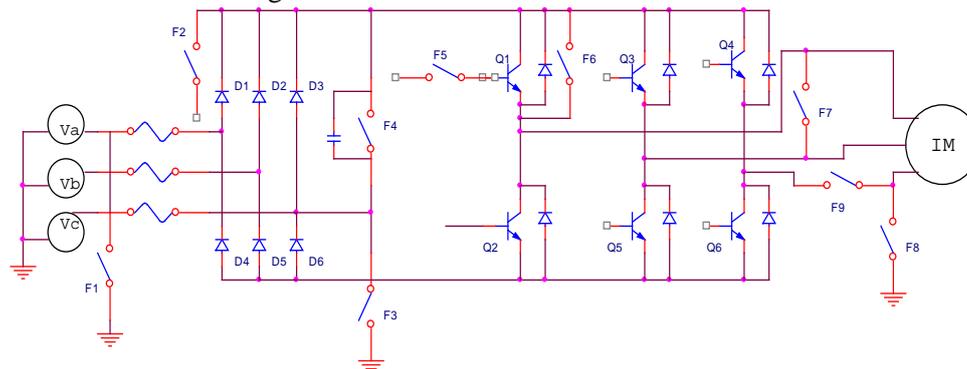


Figure 2. Failure Mode Of PWM Inverter

Analysis Of Various Faults

Different fault mode conditions on the machine side are applied and tested practically on the induction motor at a fixed load of 5Nm and fixed reference speed of 314.16 rad/sec (1500 RPM) with a current limit of 10A.

One-Phase Shorted To Ground

One of the motor terminals is shorted to ground when the motor is running under steady-state set speed conditions. The oscilloscope waveforms of speed-time, current-time and torque-time characteristics are shown in figure 3. From the waveforms, it is observed that although the torque pulsations are high, the motor can be run with reduced load depending on the application requirement. Also, the fault mode current is higher than the starting current. Hence, this aspect has to be taken into consideration in designing the drive system.

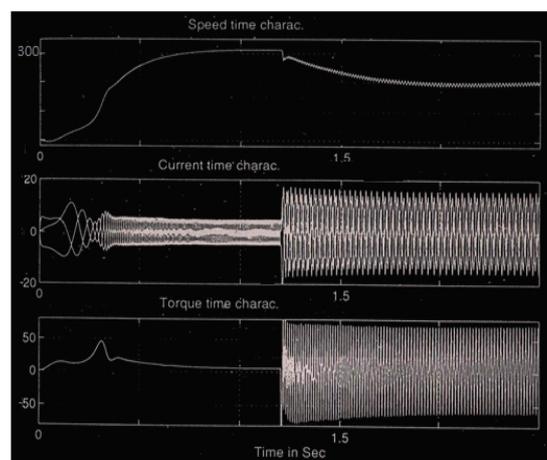


Figure 3. Single Phase Short Circuit to Ground

One-Phase Open Circuited

When the motor is running under steady-state conditions, it is possible that one phase may become open-circuit. The waveforms of figure 4 depict that fault is applied at time $t = 1.2$ sec. It is evident that although the torque pulsations are high, one can run the motor with reduced load. The torque pulsations in this case are relatively less as compared to the previous case where one of the phase has been short circuited to ground. Also, the fault mode current is almost equivalent to the starting current.

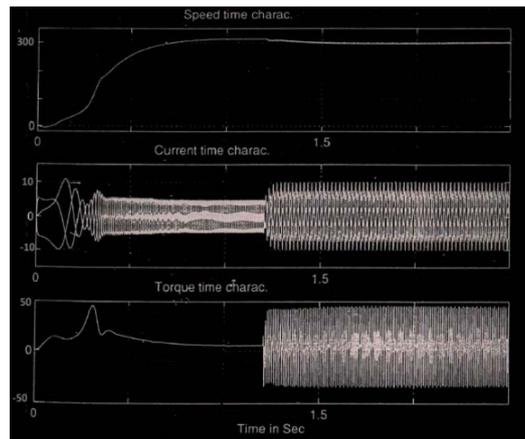


Figure 4. Single Phase Open Circuit

Two Lines Short Circuited

In this case, one of the motor terminal gets connected to the other, when the motor is running under set reference speed and set values of torque as well as current limit. From figure 5, it is observed that after the fault is applied at $t = 1.2$ sec, both current and torque slowly damp to zero. Thus the motor cannot continue to operate in this condition.

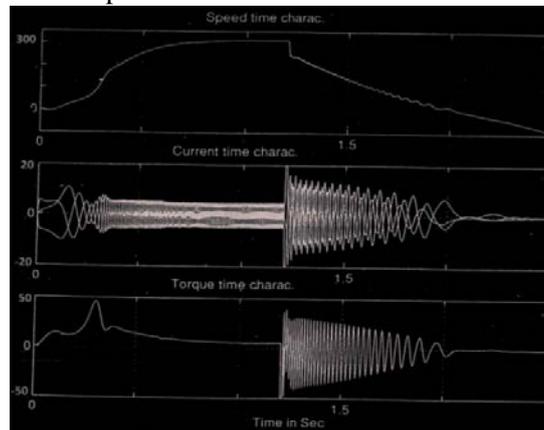


Figure 5. Line To Line Short Circuit

Two Lines Open Circuited

Two lines of the motor input terminals are open-circuited while the motor is running at a set value of torque, current limit and speed conditions. The oscillogram waveforms of figure 6 depict that after the fault is applied at $t = 1.2$ sec, both current and torque slowly damp to zero. The time taken by the current and torque characteristics to damp to zero is more as compared to the previous case where two lines were short circuited. These results indicate that the motor cannot continue to operate in this condition.

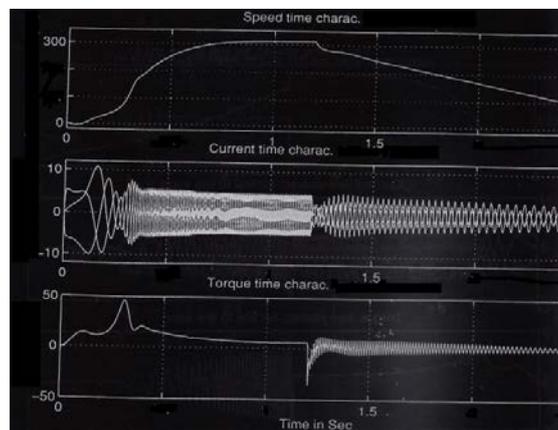


Figure 6. Two Lines Open Circuit

Three Phase Open Circuit

All the three lines of the motor terminals may get open circuited while the motor is running in steady state condition. As observed from figure 7, both current and torque immediately reach zero as soon as the fault occurs. Hence, it is not possible to continue running the motor under this condition.

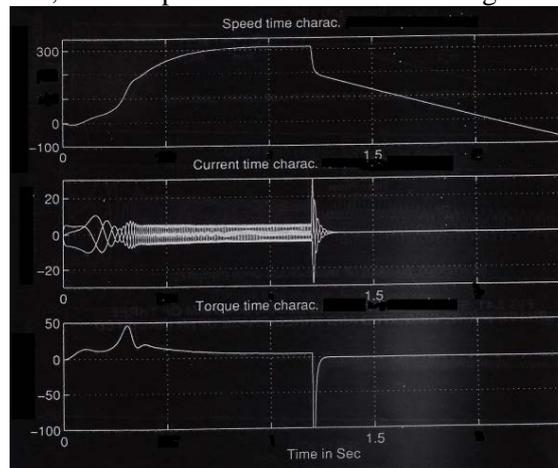


Figure 7. Three Phase Open Circuit

Three Phase Short Circuited

When the motor is running under set values of torque, current limit and speed conditions, it may happen that all the three lines of the motor terminals get short circuited. As shown in the waveforms of figure 8, as soon as the fault occurs, both current and torque immediately reach zero. This condition takes a little extra time as compared to the case of figure 7. As a result, the motor cannot continue to run in this condition.

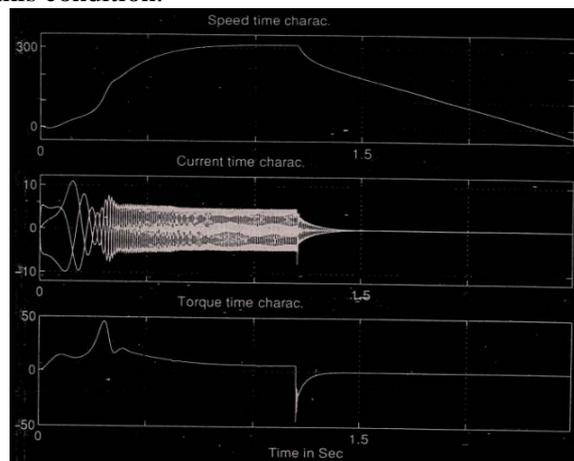


Figure 8. Three Phase Short Circuit

Conclusion

The comprehensive study carried out on the six cases of machine side faults revealed that the motor can run only under single line open or short circuit cases. The duration of time that the operator can allow the motor to operate under these two fault conditions is based upon the motor over-load carrying capacity and insulation capacity. The experimental investigation results should be kept in mind while designing a drive system since the power semiconductor devices are very sensitive to the fault conditions. It is hoped that the research investigations reported would be very useful to the industry in improving the Power Electronics circuit reliability.

References:

- General Electric SCR Manual; GE, Schenectady, N.Y. USA, 1979.
- T. L. Grant & T. H. Barton, "Control Strategies for PWM Drives", IEEE Transactions on Industry Applications, Vol. IA-16, March/April 1980, pp. 211-216.

- D. A. Grant, J. A. Houldsworth & K. N. Lower, "A New High Quality PWM AC Drive", IEEE Transactions on Industry Applications, Vol. IA-19, March/April 1983, pp. 211-216.
- W. Shepherd and L. N. Hulley; Power Electronics and Motor Control; Cambridge University Press UK, 1987.
- Hamid Nejjari & Mohamed El Hachemi Benbouzid, "Monitoring and Diagnosis of Induction Motors Electrical Faults Using a Current Park's Vector Pattern Learning Approach", IEEE Transactions on Industry Applications, Vol. 36, No. 3, May/June 2000.
- Thomas M. Jahns and Vladimir Blasko, "Recent Advances in Power Electronics Technology for Industrial and Traction Machine Drives", Proceedings of IEEE, Vol. 89, No. 6, June 2001.
- Klima, J., "Analytical investigation of an induction motor drive under inverter fault mode operations", IEE Proceedings Electrical Power Application, Vol. 150, No. 3, May 2003, pp. 255-262.
- B. Lu and S. Sharma, "A literature review of IGBT fault diagnostic and protection methods for power inverters," in Proceedings of the 43rd IEEE Industrial Applications Society Annual Meeting (IAS'08), October 2008.
- M. Abul Masrur, ZhiHang Chen and Yi Lu Murphey, "Intelligent Diagnosis of Open and Short Circuit Faults in Electric Drive Inverters for Real-Time Applications", IET-Power Electronics (Journal), March 2009.
- C. M. Bhatia, Sanjana Malhotra "Knowledge Based Protection Circuits for Converter and Inverter Applications", IEEE Conference, ICCPE 10, NSIT New Delhi, January 2011.

INSTITUTIONAL EFFECT ANALYSIS COMPARING ENERGY EFFICIENCY RETROFITTING FOR EXISTING RESIDENTIAL BUILDINGS PATTERNS IN CHINA

Junna Zhao

MES University of Waterloo, Canada

Fanghong Lou

MA Renmin University of China, China

Abstract:

Buildings in China are important contributors to the country's energy consumption and CO₂ emissions. Energy use in building in northern heating regions of China accounts for more than 40% of secondary energy consumption in urban areas. The total building areas has increased 75% over the past decade. Approximately 2.5 billion m² (nearly 1/3) of existing residential buildings in Northern China are worth retrofitting. Efficiency Retrofitting for Existing Residential Buildings (EERFERB) in China is undergoing a fast development support by both the Chinese Central Government and international governments and institutions. These international and domestic projects have not only played a significant role in enhancing China's building energy efficiency, but also significantly promoted the quality of housing conditions for low- and middle- income populations. It is recognized that projects with government cooperation among countries and central-to-local government projects have different implementation patterns. These differences provide valuable lessons for selecting efficient project delivering institutions. This research will develop criteria to compare international government-to-government projects and domestic central government projects. Two typical projects in Beijing and Tangshan will be evaluated from the perspectives of regulatory support, organizations, retrofitting effects, financing mode, management structure, incentive and technical support, and community sustainability improvement. Based on it, both the strengths and weaknesses of both retrofitting delivery patterns are presented. Finally, lessons and experiences are extracted and the implications for instruction and decision-making for the Chinese EERFERB policy design are identified.

Key Words: EERFERB, building energy efficiency, institutional effects, China

Introduction

Approximately half of worldwide new buildings in this century were built in China. It is predicted that building stock in China will reach 70 billion m² in 2020, which is three times of the number in 2000. With the dramatic increase of building stock and urbanization, building energy consumption has also nearly doubled in the past decade, from 0.4 billion tce (tons coal equivalent) to 0.7 billion tce. The total building energy consumption made up roughly 1/5 of Chinese total energy consumption. Among nationwide building energy consumption, space heating for Cold and Severe Cold zones accounts 40%. They are referred to together as the Northern Heating Zone, where installing space heating system is required by law. The energy consumption for heating per area in this zone is two to three times of the average in western countries. (Li 2009)

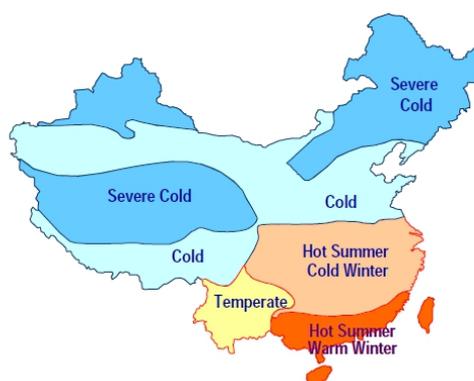


Figure 1 China Climate Regions Defined by the MOHURD
Source: (Huang and Deringer 2007)

Existing buildings are crucial for reducing the total building energy consumption. Ninety-five percent of existing buildings fail meeting the Chinese Building Energy Saving Standard and have potential energy for efficiency improvement. (Jin and Shakil 2010). The Ministry of Housing and Urban-Rural Development of the PRC (MOHURD) claimed that most buildings built before 2000 are non-energy saving buildings. Approximately 3.5 billion m² residential buildings in the Northern heating zone need to be retrofitted. These buildings were built over 20 to 30 years ago with high energy consumption and low living comfort. Some of them cannot be fully heated to 10°C in winter.

China has already retrofitted 0.182 billion m², half of the total residential buildings in Northern heating zone at the end of 2010. As a result, new heating supply system and heating meters have been installed, so that heating is charged based on heat load. The total retrofit is estimated to contribute with to 2 million tce energy consumption reduction and 5.2 million tons CO₂ emission reduction. For the 12th Five Years, the Chinese Central Government set a new goal of retrofitting over 0.4 billion m² residential building in the Northern heating zone. 6 million tce of energy saving will be achieved as a result of retrofitting for residential buildings, public building, and public institution buildings, (State Council of the People Republic of China 2012).

However, even though the Energy Efficiency Retrofit for Existing Residential Buildings (EERFERB) program is moving quickly and has enormous potential, the actual work in local levels has confronted hindrances originating from the retrofitting patterns. These include institutional failures, such as absence of standard, inefficient diffusion mechanism, inflexible fund-raising channel, and incompetent organizing and coordinating (Shilei, Wu and Sun 2009). Therefore, it is significant to analyze the institutional effect of the current retrofitting pattern. Experiences drawn from western countries have been discussed in previous studies (Shilei, Wu and Sun 2009), but the fundamental differences between western and Chinese buildings structures, and political contexts were ignored. Thus the feasibility of imported retrofitting patterns cannot be guaranteed.

This research is conducted to explore effective retrofitting patterns by comparing two kinds of institutional frameworks in existing Chinese projects. These two projects represent central government management-based institutions and international government co-management-based institutions respectively. Because these have been implemented under the same political, social and economic context, it is more explicit in answering which retrofitting pattern is more feasible to China. Analyzing and summarizing the experiences will give specific guidance to upcoming residential building retrofitting, and accordingly contribute to the overall achievement of energy saving goals.

For addressing this research question, Part Two introduces institutional effect assessment framework for comparative study. Part Three presents the China EERFERB institutional framework, including the retrofit subjects, retrofit content, funds-raising, organization and management and related policies. Part Four compares two retrofitting patterns. The institutional effects are addressed in analyzing the differences of institutional framework and evaluating the corresponding specific projects. Conclusion is explained in the Part Five.

Institutional Effects Assessment

Institutional arrangements are formal government organizational structures and the integration of policies and governance (Akgün, Leeuwen and Nijkamp 2012). These are set in order to arrange

the policies and government programs. The formulation and implementation of policies at all levels--- central government, provincial government, and local government, are under the framework of institutional arrangements. Thus, a quality institutional framework is seen pivotal for sustainable development (Haanyika 2008). The effect of institutional frameworks refers to the credibility, transparency and predictability of the operations under the framework. The frameworks are crucial to government programs because these institutions have a direct impact on extent of policy goals achievement and overall effectiveness of the program. In terms of house retrofitting programs, the institutional framework is also a central component of the retrofitting patterns. Embedded in Chinese government hierarchy, the institution framework determines the actual implementation mechanisms and even the content of house retrofitting.

In the previous case study (Stefopoulou, et al. 2008), legislation framework, administrative framework, and economical and market framework are three main domains of institutional frameworks. Similar research on rural electrification highlighted the significance of policy and institutions in dealing with energy saving issues, especially when related to high poverty and low skills group. The research defined key considerations of the Rural Electricity institutional framework as a statutory body with a clear mandate and responsibility, liaison facilitating participation by public and private players, and technologies support. This research discussed policy measures and strategies, and institution frameworks separately. (Haanyika 2008) . However, as policy and institutions are always discussed as a package, the institutions are also interpreted as “rules in use” (Brown 2003), and include formal constraints (constitutions, laws, legal standards), and informal constraints (self-imposed codes of conduct), and their enforcement characteristics (North 1990).

The assessment of institutional effects on two EERFERB retrofitting patterns is divided into two phases: an analysis of the institutional frameworks and coordination mechanisms, and a comparative analysis on the institutional effects on different retrofitting patterns.

EERFERB institutional and policy framework in China

EERFERB is the fundamental institutional background for EERFERB implementation. The origin and development of EERFERB is introduced. Furthermore, administrative framework and coordination hierarchy explains the inner coordination within the government hierarchy, and the implementation mechanism illustrates the steps of developing a specific project and the main stakeholders. As a crucial part, the sources and allocation of incentives for EERFERB are also addressed.

EERFERB Program in Northern Heating Zone

EERFERB program in China can be traced back to several small scale pilot programs in the 1990s. Lacking a corresponding institution, these attempts were not popularized (Li, 2009). In 2000, MOHURD issued *Technical specification for energy efficiency retrofit for existing buildings (JGJ129-2000)* , which took into effect on 2001 as an industry standard. It illustrates the principles and contents of EERFERB, including wall insulation, air tightness of doors and windows, and heating systems. However, as a technical standard, it is seen as a “dead letter” without either mandates requirements or incentives matched attached. This advancement also failed to bring about actual building retrofit. To implement EERFERB, pilot projects and isolated technical guidance are insufficient. To completely implementing a retrofitting program calls for a whole institutional mechanism with specific goals, regulatory support, technique support, and financial support.

During the 11th Five Years, to achieve the goal of 20% reduction in energy consumption per GDP, State Council issued *Information related to Comprehensive Energy Saving and Emission Reduction* in 2007, which set a goal of achieving 0.15 billion m² EERFERB in Northern Heating Zone by the end of 2011. This document eventually activated a large scale EERFERB in 15 Northern provinces.

Based on the achievement over the 11th Five Years, China National Development and Reform Commission (NDRC) and MOHURD issued a *Green Building Action Plan* on 1st January 2013. It pointed out that more than 0.4 billion m² out of 3.5 billion m² houses should be retrofitted by the end of 12th Five-Year (2015), and basically all eligible buildings in Northern Heating Zones should be retrofitted by the end of 2020. This ambitious goal and its challenges are shared by all related local governments. For example, Beijing new EERFERB goal for 12th Five Years is 0.15 billion m² (Beijing Municipal Commission of City Administration and Environment 2011), which is the total

number for the national EERFERB during 11th Five Years. (National Development and Reform Commission and MOHURD, 2013)

Administrative Framework and Coordination Hierarchy

MOHURD and Ministry of Finance (MOF) are responsible departments of EERFERB program. MOHURD is responsible for making EERERB plan, setting and allocating goals to the 15 provinces. Based on the received provincial goal, each provincial government signs a letter of guaranty, namely *Energy Efficiency Retrofitting for Existing Residential Buildings Protocol* with MOHURD to ensure the accomplishment. During the 11th Five Years, Beijing signed the highest goals among all provinces, namely 25 million m². Provincial governments further divide its goal to local governments, where the goals are further broken down to district/ county level. In this terminal level, specific EERFERB projects are developed. Legal guiding documents issued for EERFERB include (MOHURD 2008):

- A general guidance
 - *Implementation Guidance for The Promotion of Energy Efficiency and Heat Metering Retrofitting for Existing Residential Buildings in Northern Heating Region* (hereinafter to be referred as “*Implementation Guidance for EERFERB*”)
- Implementation and evaluation guidance for specific EERFERB project:
 - *Technical Guidance for Energy Efficiency and Heat Metering Retrofitting for Existing Residential Buildings in Northern Heating Region* (hereinafter to be referred as “*Technical Guidance for EERFERB*”)
 - *Project Evaluation Principles for Energy Efficiency and Heat Metering Retrofitting for Existing Residential Buildings in Northern Heating Region* (hereinafter to be referred as “*Project Evaluation Principles for EERFERB*”)

To further support EERFERB, MOF provides incentives based on rules in *Administrative Measures of Incentive Funds for Energy Efficiency and Heat Metering Retrofitting for Existing Residential Buildings in Northern Heating Region* (hereinafter to be referred as “*Measures of Incentive Funds for EERFERB*”). As the only guideline for incentive funds, it covers the scope, principle, and approaches for practical incentive funds. The incentive mechanism will be discussed in the part 3.4.

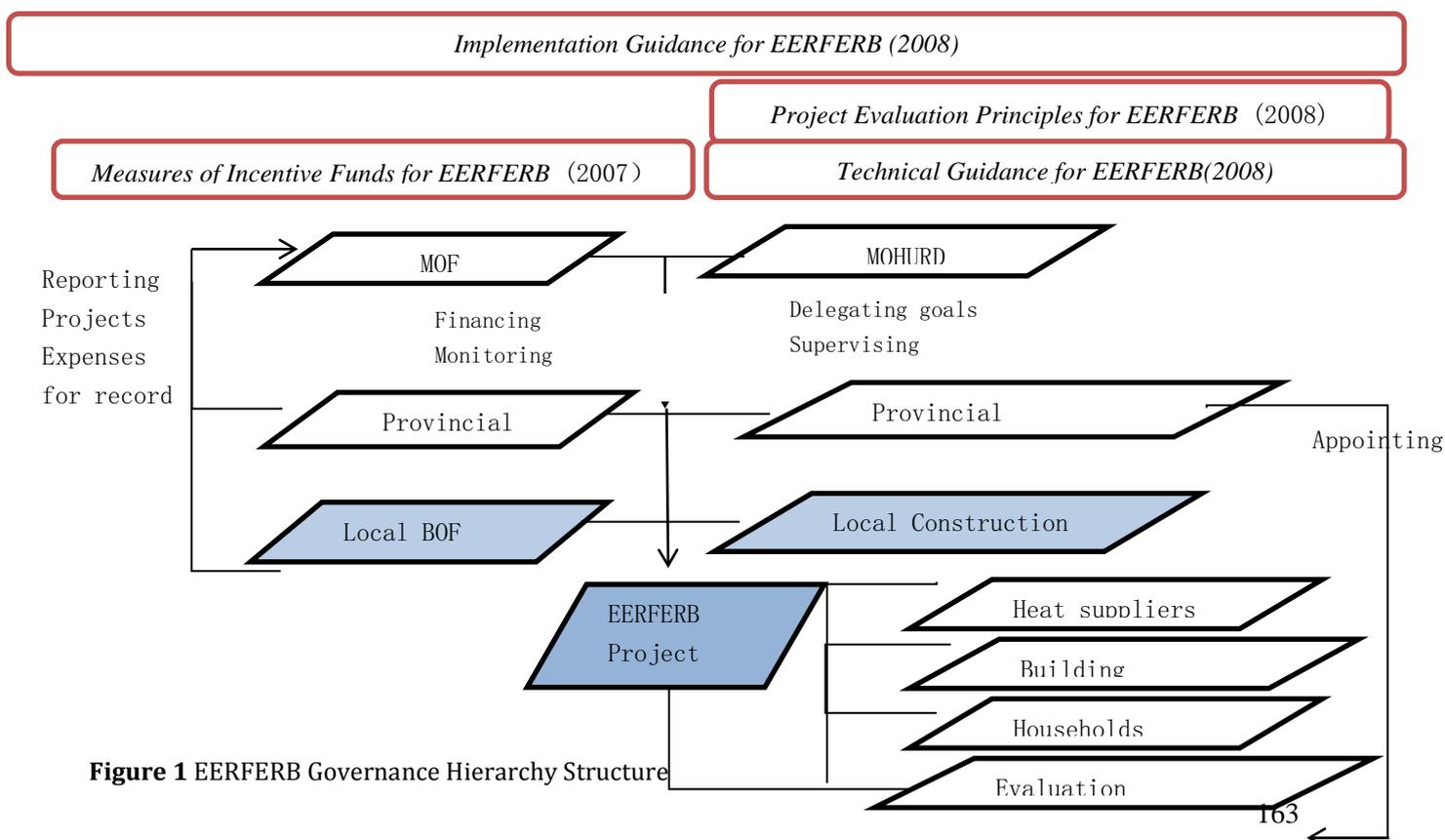


Figure 1 EERFERB Governance Hierarchy Structure

As shown in Figure 2, MOF is in charge of the disbursement of funds and clearing the incentive after retrofitting. MOHURD provides technical guidance, manages and evaluates provincial EERFERB programs. According to the *Implementation Guidance for EERFERB* (MOHURD; MOF 2008), the provincial Bureau of Construction (BOC) makes overall EERFERB plan and sets goals for local governments, while Bureau of Finance (BOF) makes retrofitting budgets for all allocated programs. Specific EERFERB projects are arranged and planned in the construction department of the district or county level. Corresponding budgets for projects are reported from local government to provincial government (MOF 2007). In the process of retrofitting construction, construction departments at all levels are administrators of their sub-level EERFERB work. Additionally, independent evaluation institutions are assigned by the provincial BOC to evaluate the specific projects in the territory (MOHURD 2008).

Apart from departments of construction and departments of finance, some other government departments and institutions are also involved. For instance, retrofitting for energy-saving residential buildings (ESB) in Beijing is undertaken by Municipal Administration Commission, instead of the Beijing Construction Committee on Housing and Urban (BCCHU), while BCCHU is in authority for non-energy-saving residential buildings (NESB)' retrofitting under a co-management system with Beijing BOF, the Beijing Municipal Development and Reform Commission (BJMDRC), the Beijing Municipal People's Government State-owned Assets Supervision and Administration Commission.

Implementation Mechanism

Implementation Process

Central Heating systems are the main heating supply systems in Northern China. The average single heating system can heat approximately 10 million m² (Li 2009). Thus to improve the overall energy efficiency of space heating, retrofitting must include the whole the heating chain, from heating sources to end users. The stakeholders are not only the local government and households, but also heat suppliers. Heat suppliers are the main implementers and involved throughout the eight step process: general building and demographic survey, households mobilization, building retrofitting plan and design, project budget making and incentive application, retrofitting construction, project quality examination and acceptance, energy saving effects evaluation, and sampling re-examination (Figure 3).

In the specific retrofitting construction process, either the heat supplier or households invest in upgrades one or more the following: heat source, thermal station, outdoor heating supply network, indoor heating systems and heat metering and doors/ windows. Without government incentive, the investments require high initial capital with long payback terms. Even though both heat supplier and households can benefit from long term energy saving, they still lack strong motivation to pay for the non-monetary social benefits. This market challenge can be eliminated to some extent through involving government efforts. Therefore, governments at all levels are still leading stakeholder of driving EERFERB in China.

Retrofitting Contents

Technical Guidance for EERFERB explicates three main entities for residential building retrofitting, namely building envelope, heating metering, and heating balance of the pipe network. As shown in Figure 4, to be eligible for retrofitting, the building should be safely used for 20 more years. This is determined through an initial safety examination, including seismic capability, building structure, and fire safety. Eligible buildings are further examined on energy consumption and energy saving potential before retrofitting. This pro-retrofitting examination is crucial because it provides the basic information for retrofitting design, cost and benefit estimation, retrofitting materials and updated products selection, and investment payback terms.

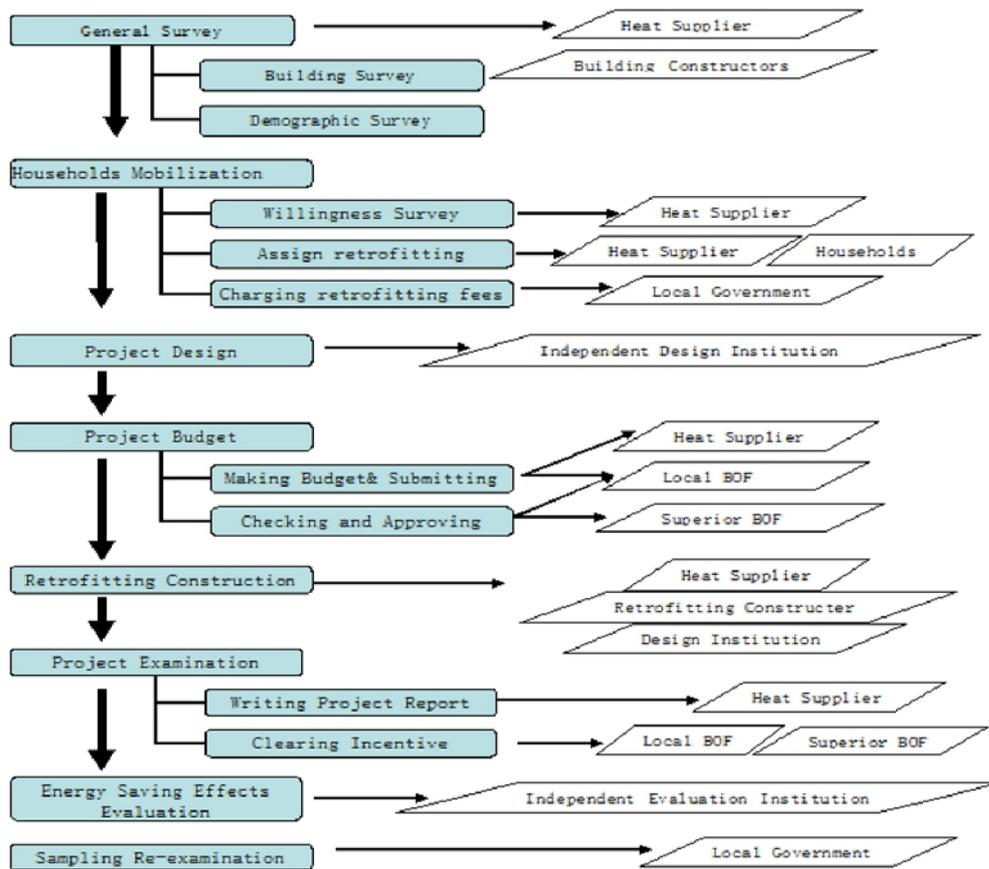


Figure 3 Implementation Process of EERFERB (MOHURD 2011)

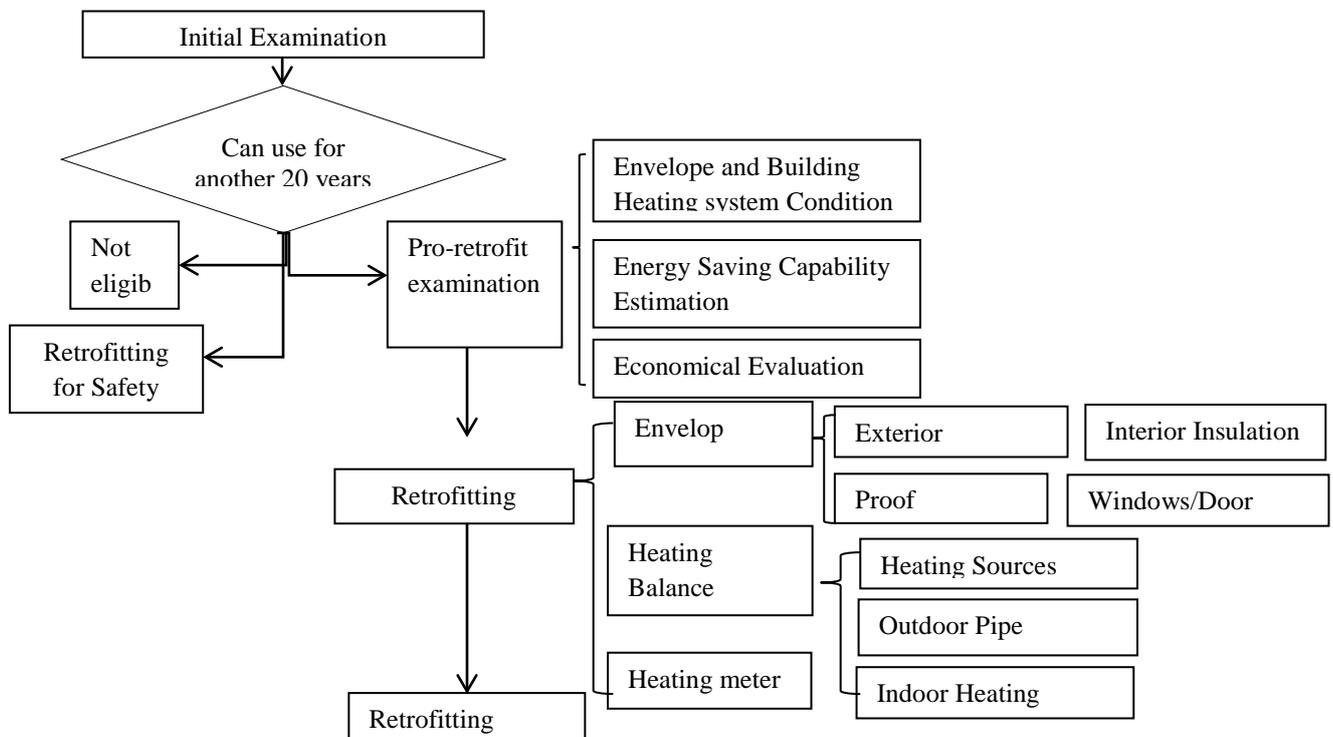


Figure 4 EERFERB Project Content and Retrofitting Process (MOHURD 2008)

Funding and Incentive Framework

The funding for EEREFBR comes from multiple sources (Figure 5). Central Government incentive and specific local government incentives are two main sources. Apart from governments, heating suppliers are required to self-finance through multiple channels, and households are also encouraged to invest in retrofitting, such as upgrading windows. The local funding sources vary by provinces.

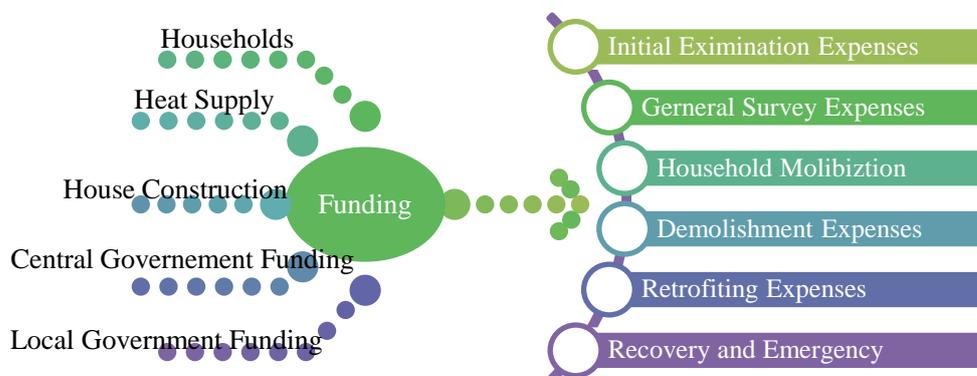


Figure 5 EEREFBR Funding Flow

Funding Sources Analysis

(1) Central Government Incentive

Central Government Incentive provides subsidies for envelop retrofitting, indoors heating meter, and heating sources and heating pipe network. The amount of incentive for each province varies based on the specific climate zone, retrofitting working load, energy saving effect and retrofitting progressing. To be specific, the incentive for particular area is calculated in the following form (MOF 2007).

Central Government Incentive for Particular Area = Incentive Base line for specific climate zone \times $[\sum (\text{Particular Entity Retrofitted Area} \times \text{Weight of the Particular Entity}) \times 70\% + \text{total retrofitting area} \times \text{Energy Saving Effect Coefficient} \times 30\%] \times \text{Progressing Coefficient}$

Where:

- Incentive Base line for specific climate zone is 55 RMB/ m² for Severe Cold zones, and 45 RMB/ m² for Cold Zones;
- Particular Entity refers to envelop retrofitting, indoor heat meter retrofitting, and heating sources and pipeline network retrofitting. Their perspective weights of the particular entity are 60%, 30%, and 10%;
- Energy Saving Effect Coefficient is determined by the actual energy consumption reduction;
- Energy Saving Effect Coefficient is divided into three levels (1.2, 1, 0.8) based on the retrofitting goal accomplishment time (before heating season of 2009, 2010, 2011 respectively)

(2) Local Government Incentive

Some provincial governments set budget funding to match Central Government incentives. Some areas also provide further subsidy for using heating meters after retrofitting. For instance, Tianjin has issued more subsidies for low-income groups and reduced their space heating bills.

Incentive Disbursement and Usage of Funding

Incentive disbursement in all levels is divided into two periods: prepay before retrofitting and clear incentive after retrofitting. In terms of Central Government Incentive, MOF prepays 6 RMB/ m² (equivalent to 0.98 CAN dollars) for installing heat meter. After retrofitting, MOF and MOHURD settle the accounts and offer all the left incentive based on the actual accomplished retrofitting projects. In 2007, the average incentive from central government is 9 RMB/m² (equivalent to 1.47 CAN dollars) (Zhong, 以及其他他人 2009). During the 12th Five Years, totally 13.5 RMB/m² will be

provided for heating meter system retrofitting, and 4.5 RMB/ m² (equivalent to 0.74 CAN dollars) for heating sources and pipe line networks (BJMCCAE; BJBOF 2011).

Institutional Effects on Comparative Retrofitting Patterns

A pilot project in Tangshan was established before the EERFERB program was widely applied. The following policies and institutions are synthesis of experiences from Tangshan project and existing government institutional framework. Beijing retrofitting pattern is a typical case. From Tangshan projects to Beijing projects, retrofitting pattern changed along with changes to institutional frameworks. The changes are mostly reflected in institutional frameworks, including the polices, contents, stakeholders, funding, and coordination mechanisms. These changes eventually result in significant impacts on retrofitting patterns and outcomes. Institutional effects are discussed from the perspective of policy roles, coordination, integration of private sectors, community acceptance, and financial and technique supports.

Institutional Frameworks of EERFERB projects in Beijing and Tangshan

Tangshan project is a model project for following up provincial projects, such as Beijing project. Lessons from Tangshan project was summarized and applied in the policy designed. What left different between Tangshan project and Beijing project is the institution frameworks, which resulted in various project achievements.

Central Government retrofitting pattern---Beijing

The EERFERB projects in Beijing are delivered under the central government-to- local government institution. It is a typical widely-used retrofit pattern shared by all 15 participating provinces. Existing residential buildings were divided into two categories in Beijing: energy-saving buildings (ESB) and non-energy-saving building (NESB). The energy-saving buildings were built after July 1998, when all new buildings were required to meet new energy efficiency standard. Retrofitting for ESB and NESB are undertaken by different departments, but the retrofitting mechanism and process are essentially the same.

International governments co-management retrofitting pattern---Tangshan

The Energy Efficiency in Existing Buildings (EEEB) project No.1 Quarter, Tangshan, is a demonstration project commissioned by the German Federal Ministry for Economic Cooperation Development (BMZ) and executed by MOHURD. This project lasted for 5 years (2005 to 2010). It reported success with an average 4-6°C increase of indoor temperature, and 30 % reduction of energy consumption within the first year of retrofitting (Shilei, Wu and Sun 2009). As a pilot project, the Tangshan EEEB project was an integrated with retrofitting for energy efficiency, and community light, façade, and building automation. To implement this project, a temporary co-management office, the Tangshan BEE office, was established with representatives from local community, local government, German government, Germany and Chinese experts, heat suppliers. A detailed comparison of retrofitting patterns is presented in Table 1.

Table 1 EERFERB Retrofitting Patterns Comparison

		Beijing Central Government EERFERB Projects		Tangshan Co-management EEER Project¹
		EERFERB for NESBs²	EERFERB for ESBs³	
Policy Role	Polies	Administrative Measures of Beijing energy efficiency retrofitting for existing NESBs Financing Measures of Beijing energy efficiency retrofitting for existing NESBs	Administrative Measures of Beijing energy efficiency retrofitting for existing ESBs Financing Measures of Beijing energy efficiency retrofitting for existing ESBs	Research Institute Standards and Norms (of MOHURD) Investment and financing modes of EEEB (not accessible)
	Authorities	BJMCHURD, BJMCCAE, Beijing Municipal Commission of Urban Planning (BJMCUP), Beijing Municipal Commission of Development and Reform (BJMCDR), BJ BOF	BJMCCAE, BJ BOF	MOHURD
Eligible Buildings Stakeholders	Governments	All Private NESBs BJMCHURD, BJMCCAE, BJBOF, BJMCDR, Municipal Industry Association, State-Owned Assets Supervision and Administration Commission of the People's Government of Beijing Municipality	ESBs built after 1998 BJMCCAE BJ BOF	three selected residential buildings MOHURD, Tangshan BEE office, BMZ, Tangshan Municipal Government(TMG)
	Companies and Institutions	District/ County Department of Construction District/ County BOF Local Heat Supplier Independent Project design Institution Construction companies Independent Project Evaluation Institution assigned by Beijing government	District/County MCCA District/ County BOF Same as NESBs	Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) Local Heat Supplier CABR – China Academy of Building Research, Information Centre of MOHURD, China Building Material Academy (CBMA)
	Households	Building owners	Building owners	Building owners

¹ China-Germany Existing Energy Efficiency Retrofitting Program Report (MOHURD, 2011)

² Administrative Measures of Beijing energy efficiency retrofitting for existing NESBs
Financing Measures of Beijing energy efficiency retrofitting for existing NESBs

³ Administrative Measures of Beijing energy efficiency retrofitting for existing ESBs
Financing Measures of Beijing energy efficiency retrofitting for existing ESBs

		Beijing Central Government EERFERB Projects		Tangshan Co-management EEER Project¹
		EERFERB for NESBs²	EERFERB for ESBs³	
Retrofitting Contents		<ul style="list-style-type: none"> • Building Envelope • Indoor Heating meter and temperature adjusting system • Heating Sources and Heating Pipeline Network 	<ul style="list-style-type: none"> • Indoor Heating meter and temperature adjusting system • Heating Sources and Heating Pipeline Network 	<ul style="list-style-type: none"> • Building Envelope • Indoor Heating meter and temperature adjusting system • Heating Sources and Heating Pipeline Network • Facade, community light, building automation
Responsible Implementation Actor		<ul style="list-style-type: none"> • Building Envelope Retrofitting: Original Building Construction companies • Heating Meter Retrofitting: Heat Supplier 	Heat supplier	Tangshan BEE office
Funding	Sources	Central Government Incentive, Municipal Government Incentive provides 100 RMB/ m ² , Heat supplier, Households	Apart from Central Government Incentive, Municipal Government Incentive and District/ County Government Incentive filled 80% of the rest expenses in 5:5 ⁴ . Heat suppliers take care of the rest.	5 million Euro from German government, 40 million RMB from China Central Government ⁵ , Households(20% of the total retrofitting expenses) (Shilei, Wu and Sun 2009)
	Principles of funding usage	Encouraging Multi-funding sources Government Incentives follow funding and incentive framework (3.4)	Government Incentives follow funding and incentive framework (3.4)	Information not accessible.
	Used for	<ul style="list-style-type: none"> • Indoor Heating meter and temperature adjusting system • Heating Sources and Heating Pipeline Network 	<ul style="list-style-type: none"> • Indoor Heating meter and temperature adjusting system • Heating Sources and Heating Pipeline Network 	<ul style="list-style-type: none"> • All retrofitting expenses
Technology Support and Project		Materials and Products are determined by heat supplier	Materials and Products are determined by heat supplier	German and Chinese techniques and products were selected by the Tangshan BEE

⁴ Municipal prepaid incentive is 6.3 RMB/ m² in 2011

⁵ As Tangshan EEER Project was part of overall EEER program, this is the total amount of German and Chinese governments funding. The specific funding for Tangshan project was not announced.

The required implementation process of both Beijing and Tangshan projects are the same. In launching the EERFERB program, lessons from the Tangshan project were drawn and applied to other areas. Thus, the required retrofitting steps are almost identical in both projects. However, the specific implementation mechanism and corresponding retrofitting effects vary due to the involved stakeholders, informal rules accompanied with different management and coordination frameworks, and funding.

Institutional effects on EERFERB projects

Policy Roles

Beijing's projects were conducted in line with national policies and municipal policies. The municipal policies are consistent with national policies and provide detailed guidance to specific projects. Unlike Beijing, however, Tangshan project did not have specific issued policy to follow when the project started. How to work on this project mainly depended on agreements among stakeholders.

Since the Tangshan project was a demonstration project for nation-wide promotion, experiences from its first implementation year (2005-2006) should be reflected on the national policies issued in 2007. For example, the retrofitting contents required in *Implementation Guidance for EERFERB* (MOHURD; MOF 2008) were selected among the retrofitting entities in Tangshan projects. Additionally, the step of household mobilization was a lesson drawn from Tangshan projects.

It is easy and infallible for local governments and retrofitting implementers to only do what the policies require. In national and municipal policies, the quality and energy efficiency coefficient of products and materials are vaguely demonstrated. Thus product and materials suppliers and construction companies are not regarded as key players in Beijing projects. Heat suppliers are in charge of choosing products mainly based on cost-benefits in Beijing, which is the responsibility of BEE office in Tangshan project. This results in a gap of energy efficiency improvement indemnification between Beijing projects and Tangshan projects.

Coordination inter- and inner- levels of government

Coordination in the Beijing projects is embedded in the top-down government hierarchy. As shown in Figure 2, governments in upper level supervise and finance the work in lower level. Achieving satisfactory EERFERB effects is also associated with the personal career of the local department head (MOHURD; NDRC; MOF; General Administration of Quality Supervision, Inspection and Quarantine of The people's Republic of China 2010). To gain the incentive from the central government, and confirm good personal working performance to upper level government, local governments are promoted to achieve EERFERB goals

In the Beijing experience, the responsibility of departments of finance and departments of construction at all levels seems clearly separated. The departments of finance care about incentive allocation and proper use, while the departments of construction concerns mainly about the retrofitting target achievement and outcome. A questionable blank remains on the effectiveness of investment on actual energy efficiency improvement.

Unlike Beijing, in Tangshan projects, BEE office guarantees communication across hierarchy boundaries. As energy consumption reduction was the shared goal within BEE office, it more focused on achieving actual energy saving.

Integration of private sectors and community acceptance

Local heating suppliers in China are government-owned corporations, which are created by the government undertaking heating supplying on the behalf of an owner government. In the Beijing pattern, private sectors such as retrofit design institutions, construction companies, and materials and products suppliers act as service or product suppliers, who are not responsible for energy saving outcomes. Project examination institutions are assigned by the provincial government. They act on behalf of provincial government to evaluate and check the retrofitting achievement. Therefore, there are no independent sectors from the market. Information is blocked from media, NGO and the public due to the lack of an information publicity platform.

Moreover, although households were required to pay for part of the retrofitting, they are not actually involved in the projects. Feedback about the energy efficiency effects is not provided. Without capability to examine the retrofitting effects, it is less possible for residences to express themselves and be heard.

In Tangshan projects, academia, as a third party, was included in decision making. It partly prevented project design from purely seeking economic benefits. Unfortunately, this is not passed down to wider national EERFERB projects.

Financial and technical support

Financial support is the core influential factor on the effectiveness of retrofitting. At central government level, more weights in incentives are set in envelop retrofitting and heating meter retrofitting. It results in strong preferences to these entities in Beijing municipal policies, and real EERFERB projects.

As both heat suppliers and households have low motivation toward retrofitting, the effectiveness highly depends on incentives from governments. The investigation results indicate that the average retrofit expense excluding the price of heat metering device was 200–230 Yuan/m² (equivalent to 32.77-37.69 CAN dollars) (Ding, et al. 2011), so the current 45-55 Yuan/m² (equivalent to 7.37-9.01 CAN dollars) central government incentive accounts for only 24% of the total expense. Many retrofitted buildings are located in poor urban areas, where local governments are not financially capable of providing enough incentives. The funding gap can only be filled by establishing more flexible financing institution. Investments from the market can be driven by obvious and visible energy efficiency improvement.

Conclusion

Based on a thorough analysis of institutional framework of EERFERB program in China, two retrofitting projects were selected to compare institutional effects on retrofitting effectiveness. The international government co-management based institution (i.e. the Tangshan case) shows higher capability achieving retrofitting effectiveness, with involvement of academia, co-management framework sharing goals, and more stable funding and technique supports. To carry forward EERFERB in China, a more flexible and transparent institution is needed. It is crucial to establish an efficient vertical and horizontal coordination systems to break constrains of hierarchy. The motivation towards retrofitting can be activated through clarifying energy efficiency goals and explicating benefits. Even though China has widely impelled energy efficiency retrofitting in Northern Heating Zones, there is still a long way to go in institution building. A self-motivated institution encompasses continuous investments from not only governments but also private sectors. A big potential market with 2.5 m² buildings for retrofitting has been recognized, so it is crucial to establish a sustained institution framework to make it happen.

References:

- Akgün, Aliye Ahu, Eveline van Leeuwen, and Peter Nijkamp. "A multi-actor multi-criteria scenario analysis of regional sustainable resource policy." *Ecological Economics*, 2012: 19-28.
- Beijing Municipal Commission of City Administration and Environment. *Drafting Instruction for Administration of Energy Efficiency Retrofitting for Beijing Existing Energy-Saving Residential Buildings*. January 21, 2011. http://www.bjmac.gov.cn/pub/guanwei/F/C/C6_1/201109/t20110914_22872.html (accessed March 10, 2013).
- BJMCCAE; BJBOF. *Financing Measures of Beijing energy efficiency retrofitting for existing NESBs*. September 14, 2011. <http://zfxgk.beijing.gov.cn/> (accessed March 1, 2013).
- Brown, Katrina. "Integrating conservation and development: a case of institutional misfit." *Front Ecol Environ*, 2003: 479-487.
- Ding, Yan, Zhe Tian, Yong Wu, and Neng Zhu. "Achievements and suggestions of heat metering and energy efficiency retrofit for existing residential buildings in northern heating regions of Chian." *Energy Policy*, 2011: 4675-4682.
- Haanyika, Charles M. "Rural electrification in Zambia: A policy and institutional analysis." *Energy Policy*, 2008: 1044-1058.
- Huang, Joe, and Joe Deringer. "Energy Efficiency Building Standards in China 2007." *Asia Business Council*. 2007. www.asiabusinesscouncil.org/docs/BEE/papers/BEE_Policy_China.pdf (accessed 3 16, 2013).
- Jin, Ruidong, and Alyas Shakil. "China Building Green Practice." *2010 ACEEE Summer Study on Energy Efficiency in Buildings*. CA, 2010.

- Li, Dongyan. "Fiscal and tax policy support for energy efficiency retrofit for existing." *Energy Policy*, 2009: 2113–2118.
- MOHURD. *China-Germany Existing Energy Efficiency Retrofitting Program Report*. October 12, 2011. http://www.mohurd.gov.cn/zxydt/201110/t20111012_206559.html (accessed March 10, 2013).
- Akgün, Aliye Ahu, Eveline van Leeuwen, and Peter Nijkamp. "A multi-actor multi-criteria scenario analysis of regional sustainable resource policy." *Ecological Economics*, 2012: 19-28.
- Beijing Municipal Commission of City Administration and Environment. *Drafting Instruction for Administration of Energy Efficiency Retrofitting for Beijing Existing Energy-Saving Residential Buildings*. January 21, 2011. http://www.bjmac.gov.cn/pub/guanwei/F/C/C6_1/201109/t20110914_22872.html (accessed March 10, 2013).
- BJMCCAE; BJBOF. *Financing Measures of Beijing energy efficiency retrofitting for existing NESBs*. September 14, 2011. <http://zfxgk.beijing.gov.cn/> (accessed March 1, 2013).
- Brown, Katrina. "Integrating conservation and development: a case of institutional misfit." *Front Ecol Environ*, 2003: 479-487.
- Ding, Yan, Zhe Tian, Yong Wu, and Neng Zhu. "Achievements and suggestions of heat metering and energy efficiency retrofit for existing residential buildings in northern heating regions of Chian." *Energy Policy*, 2011: 4675-4682.
- Haanyika, Charles M. "Rural electrification in Zambia: A policy and institutional analysis." *Energy Policy*, 2008: 1044-1058.
- Huang, Joe, and Joe Deringer. "Energy Efficiency Building Standards in China 2007." *Asia Business Council*. 2007. www.asiabusinesscouncil.org/docs/BEE/papers/BEE_Policy_China.pdf (accessed 3 16, 2013).
- Jin, Ruidong, and Alyas Shakil. "China Building Green Practice." *2010 ACEEE Summer Study on Energy Efficiency in Buildings*. CA, 2010.
- Li, Dongyan. "Fiscal and tax policy support for energy efficiency retrofit for existing." *Energy Policy*, 2009: 2113–2118.
- MOHURD. *China-Germany Existing Energy Efficiency Retrofitting Program Report*. October 12, 2011. http://www.mohurd.gov.cn/zxydt/201110/t20111012_206559.html (accessed March 10, 2013).
- MOF. *Information related to Administrative Measures of Incentive Funds for Energy Efficiency and Heat Metering Retrofitting for Existing Residential Buildings in Northern Heating Region*. 12 20, 2007. http://www.mohurd.gov.cn/zcfg/xgbwgz/200712/t20071228_159878.html (accessed 4 1, 2013).
- MOHURD. *China-Germany Existing Energy Efficiency Retrofitting Program Report*. October 12, 2011. http://www.mohurd.gov.cn/zxydt/201110/t20111012_206559.html (accessed March 10, 2013).
- MOHURD. *Information related to Technical Guidance for Energy Efficiency and Heat Metering Retrofitting for Existing Residential Buildings in Northern Heating Region*. 7 10, 2008. http://www.gov.cn/gongbao/content/2009/content_1205389.htm (accessed 4 1, 2013).
- MOHURD; MOF. *Implementation Guidance for the Promotion of Energy Efficiency and Heat Meter Retrofitting for Existing Residential Buildings in Northern Heating Region*. 5 21, 2008. http://www.mohurd.gov.cn/zcfg/jsbwj_0/jsbwjjskj/200806/t20080613_171707.html (accessed 4 1, 2013).
- MOHURD; NDRC; MOF; General Administration of Quality Supervision, Inspection and Quarantine of The people's Republic of China. *Suggestions on Further Implementing Heating Meter Reform*. 2 2, 2010. http://www.gov.cn/gzdt/2010-03/17/content_1557963.htm (accessed 4 1, 2010).
- North, D. *Institutions, institutional change and economic performance*. Cambridge, UK: Cambridge University Press, 1990.
- Shilei, Lv, Yong Wu, and Jinying Sun. "Pattern analysis and suggestion of energy efficiency retrofit for existing residential buildings in China's northern heating region." *Energy Policy*, 2009: 2102-2105.
- State Council of the People Republic of China. "State Council Inform of Twelveth Five Year Energy Saving Emission Reduction Plan." *The Central People's Government of the People's Republic of China*. 8 21, 2012. http://www.gov.cn/zwgk/2012-08/21/content_2207867.htm (accessed 3 10, 2013).
- Stefopoulou, Aggeliki, Konstantinos Soulis, Michael Papapetrou, Spyros Kyritsis, and Christian Epp. "Institutional and policy framework analysis in relation to the application of autonomous desalination systems-Greece." *Desalination*, 2008: 455-567.

Zhong, Y., W.G. Cai, Y Wu, and H. Ren. "Incentive mechanism design for the residential building energy efficiency improvement of heating zones in North China." *Energy Policy*, 2009: 2119-2123.

List of Acronyms

BJMCCAE Beijing Municipal Commission of City Administration and Environment
BJMCDR Beijing Municipal Commission of Development and Reform
BJMCHURD Beijing Municipal Commission of Housing and Urban-Rural Development
BJMCUP Beijing Municipal Commission of Urban Planning
BJ BOF Beijing Bureau of Finance
BJMPGSASAC Beijing Municipal People's Government State-owned Assets Supervision and Administration Commission
BMZ German Federal Ministry for Economic Cooperation and Development
CABR China Academy of Building Research
CBMA China Building Material Academy
CO₂ Carbon dioxide
EEEB Energy Efficiency in Existing Buildings
ESB Energy-Saving Building
EERFERB Energy Efficiency Retrofitting for Existing Residential Building
FYP Five Year Plan
GDP Gross Domestic Product
GTZ Deutsche Gesellschaft für Technische Zusammenarbeit
HSCW Hot Summer and Cold Winter
HSWW Hot Summer and Warm Winter
MOF Ministry of Finance
MOHURD Ministry of Housing and Urban-Rural Development
Mt Million ton
Mtoe Million ton of oil equivalent
NDRC National Development and Reform Commission
NESB Non-Energy-Saving Building
RISN Research Institute Standards and Norms (of MOHURD)
RMB Chinese Yuan
Tce tons coal equivalent
TMG Tangshan Municipal Government
11th Five-Year 2006-2010
12th Five-Year 2011-2014

RECTANGULAR OPEN CHANNELS OF COMBINED ROUGHNESS

E. Retsinis

S. Bobotas

Civil Engineer National Technical University of Athens, Greece

J. Demetriou

Associate Professor (RD) of Hydraulics, School of Civil Engineering,
National Technical University of Athens, Greece

Abstract:

In this experimental study the Manning roughness coefficient n is examined, through a considerable number of laboratory measurements with artificially roughened walls of a rectangular open channel-which always has a smooth bed. The roughness elements are vertical rubber strips, which are systematically varying as far as their projections into the flowing water and their distances among them are concerned. The case of the completely smooth rectangular channel is also examined. Since n has the disadvantage of being in the metric system of units, some suitable dimensionless groups of parameters are used. n appears not to be constant at any present roughness condition and this probably is due to the nature of any flow, where even at the same Run its behavior is not constant. Apart from this, n is also varying in the completely smooth channel case. The results are strictly holding for uniform-subcritical-turbulent-steady water flows in rectilinear rectangular open channels of small longitudinal slope.

Key Words: Rectangular Open Channels, Combined Roughness

Introduction

Fig. 1 shows the characteristics of the uniform-subcritical-turbulent-steady water flow within a rectangular open channel (various water depths, z_i , reference depths $z_{ri} < z_i$, width $b=24$ cm and longitudinal slope $J_o=0.0012$), where for depth z_i the Manning roughness coefficients is n_i (or n) while for z_{ri} depths corresponding roughness coefficients are n_{ri} , both determined from the Manning formula, that is $n_i=(1/V_i) \cdot R_i^{2/3} \cdot J_o^{1/2}$, with n_i in the metric system ($s \cdot m^{-1/3}$), and R_i =hydraulic radius, V_i =cross sectional velocity ($=Q_i/(z_i \cdot b)$), all in the same units' system.

In this investigation both channel walls were systematically supplied with adhered rubber strips ($\kappa \times h$, where always $\kappa=4$ mm and $h=8-4-0$ mm) at various distances $\lambda=50-100-200-400$ mm, forming a number of λ/h ratios, $\lambda/h=12.5-25-50-100$ or $\lambda/h=\infty$ (for $h=0$ -smooth walls). The channel bed was always smooth ($h=0$, $\lambda/h=\infty$), and rectilinear, while the various λ/h were alternating on the walls, creating some artificial combined roughnesses with considerable differences among them and to the channel bed (especially, for $h=8$ mm). The roughness is not only h dependent but also λ/h ratio dependent: Any small λ/h ratio (e.g. 12.5) corresponds to rougher wall (e.g. from $\lambda/h=25$).

Apart from the first part of this study, further on an effort is made to correlate various dimensionless quantities (including the Manning's n) in order to receive systematic results for all present laboratory measurements, i.e. overcoming the metric units dependence.

In this paper, as most interesting previous books or papers the following texts are considered, by Chow, (1959), Ramesh et al, (2000), Pyle, (1981), Robertson et al, (1973), Rosso et al, (1950), Ghosh, (1978), Demetriou, (2003), Demetriou et al, (1999), Demetriou, (2000), and finally, Demetriou, (2001).

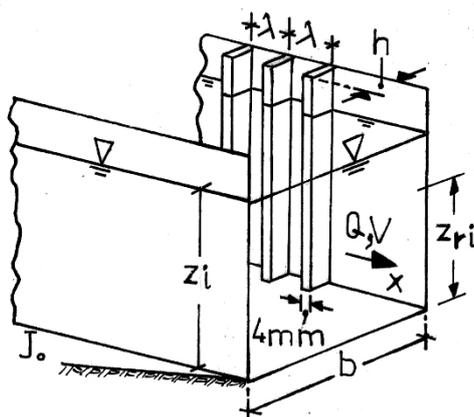


Figure 1. Flow and roughness characteristics.

The Experiments

All measurements were performed in the hydraulics Lab. of the School of Civil Engineering of the Nat. Technical Univ. of Athens (Greece), in a channel of 12 m long mainly with glass boundaries and steel bed.

8 Series with a number of $i=15 \div 17$ Runs (for each Series) were organized. z_i , R_i , Q_i , V_i , Re_i , Fr_i , were measured or calculated, while corresponding roughness coefficients (for each Run) were determined through the Manning formula. Table 1 shows a summary of all laboratory measurements and all aspect ratios (A.R.= b/z_i), where max n_i or min n_i do not always correspond to max or min A.R.

Table 1. Laboratory measurements.

Series N ^o	Runs (i)	z_i (cm)	Q_i (l/sec)	V_i (cm/sec)	h (mm)	λ (mm)	λ/h	n_i ($s \cdot m^{-1/3}$)	A.R.
1	i=1 to i=17	4.97 to 35.28	4.71 to 41.83	39.50 to 49.40	4	50	12.5	0.0106 to 0.0163	4.83 to 0.68
2	i=1 to i=17	4.70 to 35.53	3.84 to 41.83	34.07 to 49.05	4	100	25	0.01062 to 0.0160	5.11 to 0.68
3	i=1 to i=17	3.13 to 35.45	3.84 to 39.56	34.07 to 49.05	4	200	50	0.00577 to 0.0159	7.67 to 0.68
4	i=1 to i=17	4.92 to 34.30	5.43 to 45.79	46.05 to 55.68	4	400	100	0.00803 to 0.01524	4.88 to 0.70
5	i=1 to i=16	4.10 to 36.53	3.84 to 32.61	39.05 to 37.19	8	100	12.5	0.00867 to 0.01890	5.85 to 0.67
6	i=1 to i=15	5.25 to 36.93	5.43 to 38.71	43.13 to 43.67	8	200	25	0.00884 to 0.01870	4.57 to 0.65
7	i=1 to i=15	5.68 to 36.17	5.08 to 42.00	37.27 to 48.39	8	400	50	0.01061 to 0.0152	4.23 to 0.66
8	i=1 to i=16	4.62 to 36.25	3.84 to 52.08	34.68 to 59.87	0	-	∞	0.0100 to 0.0133	5.19 to 0.66

In any Run the smallest n_i value corresponds to the biggest A.R., but the biggest n_i value may be found at intermediate A.R. ratios. For the smooth channel $0.010 \leq n_i \leq 0.013$, as is referred in various books, e.g. by Chow, (1959). Finally, some of the initial results were rejected, while some other were smoothed out.

Results And Discussion

The method of interrelations

In order to elaborate the experimental results some reference depths, z_{ri} , with corresponding roughness coefficients, n_{ri} , were selected in any one Series, and so the ratios z_i/z_{ri} and n_i/n_{ri} were calculated. The typical selected z_{ri} depths were the closest to a scale of $z_{ri} \approx 10-15-20-25-30-35$ cm, where the finally chosen z_{ri} values were the actual depths in each Run. This method, i.e. of self-compared n_i/n_{ri} to z_i/z_{ri} in any Series, is named here as the method of interrelations.

Table 2 presents all results based on the above method, while Figs. 2 graphically show corresponding results, where n_i and n_{ri} are indirectly measured through the Manning formula.

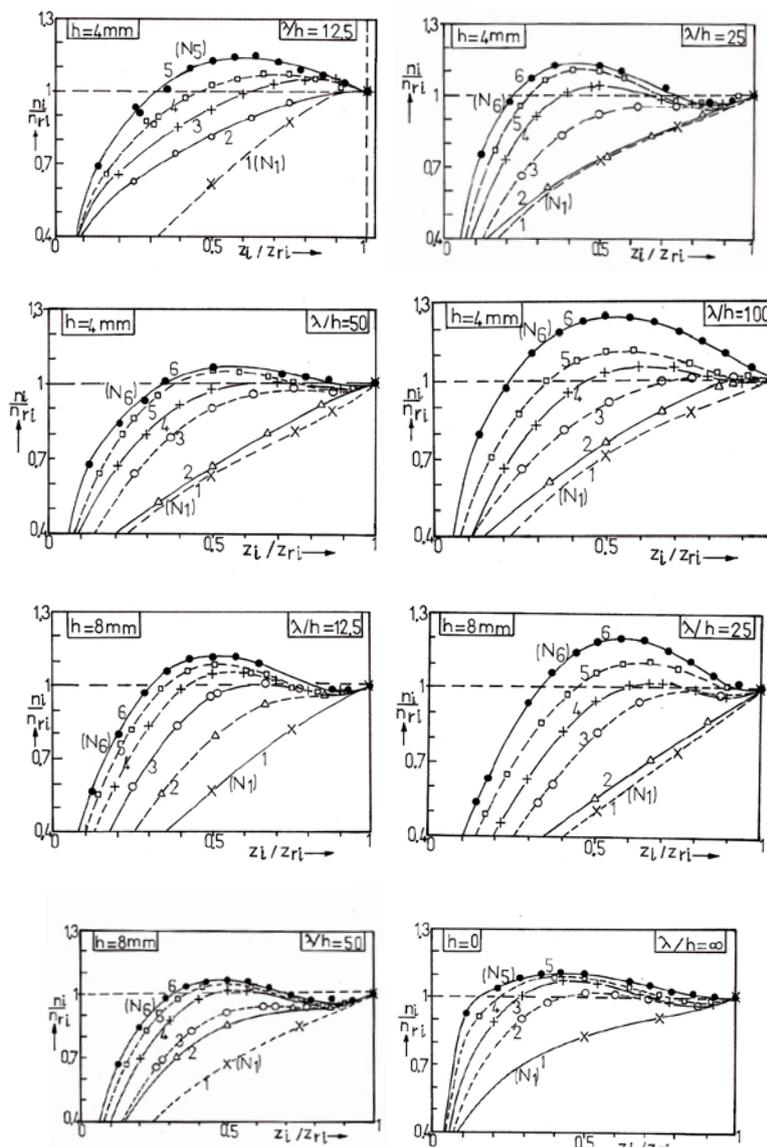
Table 2. Experimental measurements' elaboration.

Eqs. $N_i = n_i/n_{ri}$ vs $Z_i = z_i/z_{ri}$, in the form of $N_i = a \cdot Z_i^b \cdot [1 - Z_i]^c + Z_i$							
h=4 mm, $z_{ri} \approx 10-(15)-20-25-30-35$ cm				h=8 mm, $z_{ri} \approx 10-(15)-20-25-30-35$ cm			
$\lambda/h=12.5$	a	b	c	$\lambda/h=12.5$	a	b	C
N_1	0.465	1.383	0.663	N_1	0.652	2.017	1.182
N_2	0.712	0.274	0.877	N_2	3.464	1.933	1.676
N_3	0.889	0.310	0.740	N_3	4.564	1.620	1.812
N_4	1.489	0.510	1.909	N_4	4.962	1.296	1.892
N_5	1.930	0.550	1.090	N_5	4.583	1.105	1.862
$\lambda/h=25$	a	b	c	N_6	3.789	0.920	1.682
$\lambda/h=25$	a	b	c	$\lambda/h=25$	a	b	C
N_1	0.934	0.674	1.361	N_1	1.193	0.470	12.055
N_2	0.680	0.412	1.133	N_2	0.160	0.621	0.855
N_3	2.520	0.957	1.617	N_3	3.893	2.059	1.562
N_4	3.770	0.949	1.856	N_4	4.978	1.720	1.754
N_5	4.234	0.881	1.929	N_5	3.676	1.237	1.470
N_6	3.227	0.665	1.702	N_6	4.022	1.087	1.457
$\lambda/h=50$	a	b	c	$\lambda/h=50$	a	b	C
N_1	0.894	0.886	1.806	N_1	0.846	0.926	1.369
N_2	0.290	0.159	0.667	N_2	2.156	1.003	1.613
N_3	1.991	0.946	1.364	N_3	3.217	1.143	1.821
N_4	2.034	0.758	1.352	N_4	3.685	1.011	1.835
N_5	2.920	0.829	1.580	N_5	3.576	0.883	1.826
N_6	2.015	0.550	1.285	N_6	3.000	0.725	1.689
$\lambda/h=100$	a	b	c	$\lambda/h=\infty$	a	b	C
N_1	0.921	0.869	1.243	N_1	1.059	0.430	1.274
N_2	0.482	0.296	0.615	N_2	2.574	0.724	1.592
N_3	1.277	0.632	1.012	N_3	2.751	0.643	1.631
N_4	2.829	0.945	1.468	N_4	2.110	0.428	1.425
N_5	2.749	0.772	1.407	N_5	1.678	0.273	1.210
N_6	-	-	-				

The above curves are very smooth and for any one of them the ratio $N_i = n_i/n_{ri}$ may be compared to $Z_i = z_i/z_{ri}$, in the form of a typical equation

$$N_i = a \cdot Z_i^b \cdot (1 - Z_i)^c + Z_i \quad (1)$$

where the coefficients a, b, c, are also shown in Table 2. In this way to any Series a number of N_i equations (in total 45 equations) are produced, where each one of them corresponds to a respective curve of Figs. 2.



Figures 2. Numbered curves of n_i/n_{ri} vs z_i/z_{ri} for various h , λ/h , and corresponding equations N_i .

Other methods

To simplify the results, the roughness coefficients' symbol ($n=n_i$) corresponding to any depth z_i of each experimental Run is used again, while the index (i) is used only for z_i .

a) Apart from the previous description (Figs. 2 and Table 2), some more essential results are further presented, based on the following dimensionless groups among the Manning roughness coefficient n and g , R (=hydraulic radius), J_o , or λ/h ,

$$A=n \cdot g^{1/2} \cdot R^{-1/6}, A'=n \cdot g^{1/2} \cdot h^{-1/6},$$

$$B=(g \cdot R \cdot J_o)^{1/2} \cdot V^{-1}, A'=(g \cdot R \cdot J_o)^{1/2} \cdot (\lambda/h)^{1/2} \cdot V^{-1}.$$

A and A' mainly include Manning's n , while B and B' include $(g, R, J_o)^{1/2}$ and V , where for $\lambda/h \rightarrow 1$ it is $B' \rightarrow B$. n is in the metric system, but the above groups are dimensionless. A (or A') may combined to B (or B'), in order that some suitable descriptive curves are properly determined. In some cases the term $C=n \cdot (g \cdot z_n^{-1/3})^{0.5}$ is also used.

b) Smooth Channels

For the completely smooth ($\lambda/h=\infty$ on bed/walls), Fig. 3 shows all present experimental results (open circles) in terms of B vs C , where the corresponding (single) curve (for all present n, R, V, J_o) has the approximate descriptive equation

$$n \cdot (g \cdot z_n^{-1/3})^{0.5} = -0.0661 + 0.4855 \cdot [(g \cdot R \cdot J_o)^{0.25} \cdot V^{-0.5}], \tag{2}$$

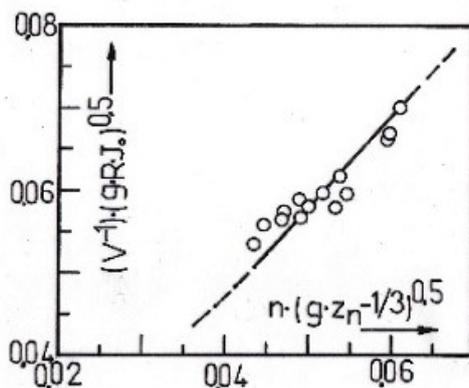


Figure 3. B vs C for completely smooth channel.

holding for $0.054 \leq B \leq 0.070$, $0.043 \leq C \leq 0.061$, and with a correlation coefficient $r^2 \approx 0.998$.

Other smooth channel cases are for $\lambda \rightarrow \infty$ (with $\kappa \neq 0$), or $\lambda \rightarrow \kappa$ (full contact of all roughness elements).

c) Rough walls and smooth bed.

Figure 4 compares A to B' (based on the present measurements) for $\lambda/h=12.5-25-50-100$

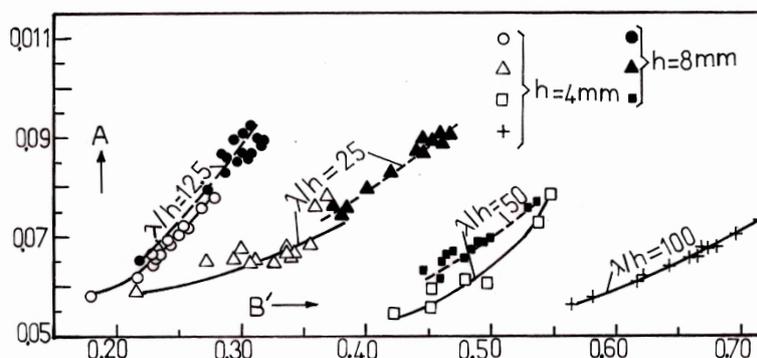


Figure 4. A vs B' for $\lambda/h=12.5-25-50-100$ ($h=4$ mm) and $\lambda/h=12.5-25-50$ ($h=8$ mm).

($h=4$ mm, open experimental symbols) and $\lambda/h=12.5-25-50$ ($h=8$ mm, black experimental symbols). All corresponding curves, through the experimental points, show that the $h=8$ mm results (dashed curves) are over the $h=4$ mm results (continuous curves), a fact which is considered as reasonable because of the increased wall roughness. Other facts are that all results produce separate curves, i.e. not a single equation can be given for all experimental results, and that when B' are increasing A are also increasing along each curve.

Among the (continuous) curves of $h=4$ mm a large space is offered for interpolations, and the same holds for the (dashed) curves ($h=8$ mm).

Contrary to the above results in Figs. 5, 6, 7, the same (as in Fig. 4) experimental results, are treated in terms of A' vs B. In Fig. 5 all results (for $h=4$ mm and all $\lambda/h=12.5-25-50-100$) are well concentrated around a single curve. This curve has a simple equation (through the experimental points-open circles), and since it holds for $\lambda/h=12.5$ to $\lambda/h=100$, it also holds in the entire field $12.5 \leq \lambda/h \leq 100$ (for $\kappa=4$ mm) and is described by the following equation,

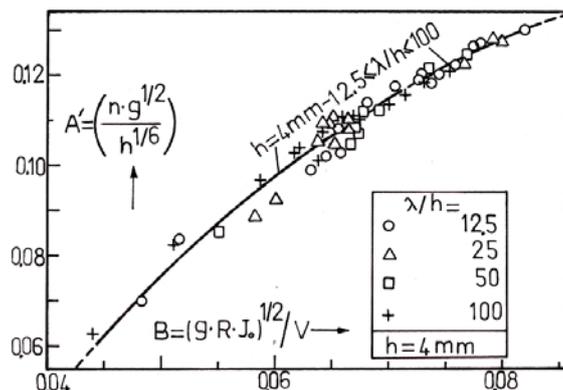


Figure 5. A' vs B , for all $h=4$ mm and $12.5 \leq \lambda/h \leq 100$ results.
 $A' \approx 0.3224 - 0.055 \cdot B^{-0.5}$, (3)

holding for $0.045 \leq B \leq 0.08$, $7 \leq A' \leq 12$ and $4.83 \leq A.R. \leq 7.67$. n is a function of h , R , J_0 and V , while when B are increasing n are also increasing, i.e. n is not constant throughout.

Fig. 6 also presents A' vs B for all $h=8$ mm and $\lambda/h=12.5-25-50$, experimental results (black symbols) which are also well concentrated around another single curve, holding for $h=8$ mm, the entire $12.5 \leq \lambda/h \leq 50$ range and $0.05 \leq B \leq 0.093$, $0.06 \leq A' \leq 0.113$, $0.65 \leq A.R. \leq 5.85$.

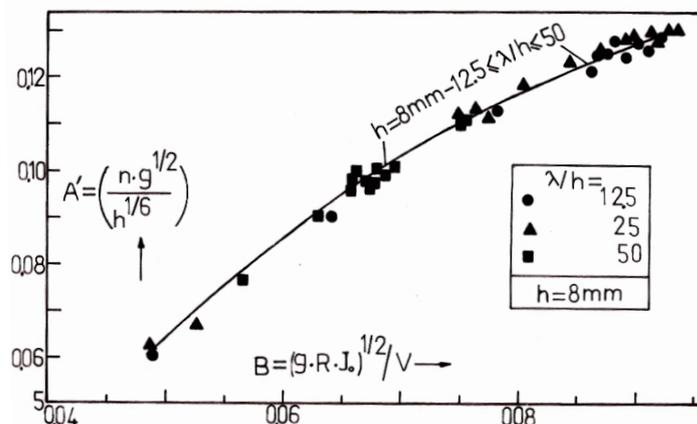


Figure 6. A' vs B , for all $h=8$ mm and $12.5 \leq \lambda/h \leq 50$ results.

Finally, Fig. 7 shows both curves (of Figs. 5, 6), in order to compare them. The curve for all $h=8$ mm experimental results is quite similar to the curve of $h=4$ mm results, and lies under the last curve at a constant vertical distance of $1.35 \cdot A'$. This fact may give a similar equation to the $h=8$ mm curve, since at same B value the lower curve lies at the above constant distance under the upper curve.

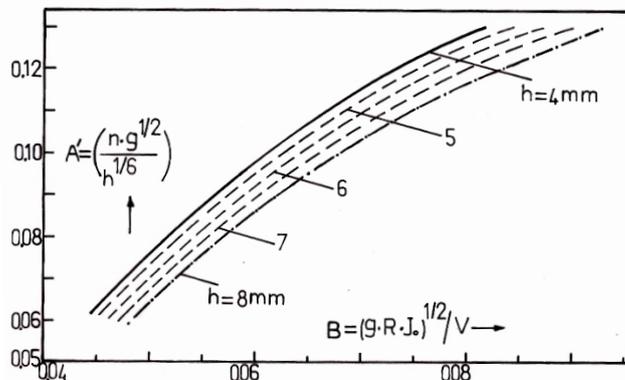


Figure 7. A' vs B , for all present rough walls' results and interpolations for $h=5-6$ and 7 mm.

If one wishes to explain why the $h=8$ mm curve is under the $h=4$ mm curve, this may be due to the combination of all quantities participating in the comparison, since n is proportional to $R^{2/3}$, $J_0^{1/2}$, and proportional to V^{-1} : When $h=8$ mm strips are used the flow (at $Q=\text{const.}$) has in general, larger velocities (in comparison to $h=4$ mm), since the channel's width which is used is constant for both cases and the flow is more intermixed.

As an important note on Fig. 7, since both curves, for $h=4$ and 8 mm are very similar among them, a family of intermediate (dashed) curves can be approximately traced by suitable interpolations, corresponding to $h=5-6$ and 7 mm, or even to more dense curves. More general, the entire group of curves of Fig. 7 holds for $4\text{mm} \leq h \leq 8\text{mm}$ and at least for $12.5 \leq \lambda/h \leq 50$.

It also may be noticed that for $h > 8$ mm corresponding curves are expected to be placed under the 8 mm curve, while for $h < 4$ mm (and even for $h=0$, smooth curve at $\lambda/h=\infty$) corresponding curves are expected to be placed over the 4 mm curve, i.e. apart from interpolations there is also a wide space for extrapolations.

The presentation of the results of Fig. 7 gives a rather broad possibility for a wide generalization of the results over the 4 mm curve, between $4 \leq h \leq 8$ mm, and below the $h=8$ mm curve, equally corresponding to a large λ/h range.

Finally, Chow, (1959), p. 98-99, referring to the Manning n establishment and its dimensions, is noting that "this was not a problem of concern to the forefathers of hydraulics" and "it is unreasonable to suppose that the roughness coefficient would contain time". Apart from this, the Manning formula has a factor of 1.49 for the English measuring system (in ft.), while this factor is 1 for the usual metric system. These problems could be solved if a dimensionless roughness coefficient is used. Although this has been already suggested in the past, it is not very clear who has made this suggestion, which actually, consists of the present terms A and B and their equation. If to A is given a new symbol $n^*(=A)$ then the dimensionless roughness coefficient is $n^*=A=(n \cdot g^{1/2} \cdot R^{-1/6})$, where-of course-the initial Manning equation (with n) is modified when n^* is used, $n^*=(g \cdot R \cdot J_0)^{1/2} \cdot (V^{-1})=B$.

Conclusions

In this experimental study the Manning n roughness coefficient is investigated for water uniform-subcritical-turbulent-steady flows, within rectilinear rectangular open channels of combined roughness, i.e. rough walls and smooth bed. The walls are artificially roughened through vertical strips (projecting into the flow at $h=4$ mm or $h=8$ mm) at regular distances λ , with $\lambda/h=12.5-25-50-100$ or $\lambda/h=\infty$ ($h=0$, smooth walls), while the bed constantly remains smooth ($h=0$, $\lambda/h=\infty$). The main conclusions are: 1) The application of the so-called method of interrelations, gives a large number of equations among the dimensionless roughness coefficients' ratios vs corresponding dimensionless flow depths. 2) These equations are giving some good relative results. 3) Since a big problem is coming from the fact that Manning roughness coefficients are in the metric system ($s \cdot m^{-1/3}$), four groups (A, A', B, B') of parameters are proposed in various combinations among n, g, R, h, J_0, V and λ/h , in order that some more general results may arise. 4) For the complete smooth channel ($h=0$, $\lambda/h=\infty$, everywhere) eq. (2) is determined based on the experiments, after the elaboration of them in terms of A vs B , this equation and Fig. 3 are concluded. 5) For all present experimental results with all various roughness conditions, Fig. 4, in terms of A vs B' , gives a number of systematic curves, for $h=4$ and 8 mm and all present λ/h combinations, where for $\lambda/h=\text{const.}$ all $h=8$ mm curves lie over $h=4$ mm (less rough) results. 6) When all present n results are elaborated in terms of A' vs B , then all experiments for $h=4$ mm (at all λ/h ratios) give (Fig. 5) a single curve, while a similar curve is produced for all $h=8$ mm results (Fig. 6 for all λ/h ratios). 7) Eq. 3 is proposed for the previous curve, concerning $h=4$ mm and all relevant λ/h ratios. 8) A method is also suggested for the corresponding curve for $h=8$ mm and all relevant λ/h ratios. 9) An extrapolation group of curves are also proposed, for $h=5-6-7$ mm and all associated λ/h ratios. The present study mainly presents an effort to rationalize the disadvantage of the Manning's roughness coefficient which is given in the metric system, by proposing combinations of all parameters in such a way that some dimensionless groups of expressions may be used. The application of such dimensionless groups of expressions may give in some cases, more general results, e.g. in the form of suitable curves described by proper equations. This application may also give, by systematic extrapolations, results for intermediate h values, while an overall result appears to be that Manning's n is not a constant throughout, i.e. it depends on various parameters and not only on roughness factors, especially in channels of combined roughnesses.

References:

- Chow V. T. *Open Channel Hydraulics*, Mc Graw-Hill, 1959.
- Demetriou J., Manassakis E. and Dimitriou D. Resistance Coefficient Change in Open Channels, 4th Nat. Congress of the Greek Committee for the Water Management, Proceedings, Volos, Greece, June Proceedings, Vol. B, pp. 91-97, 1999.
- Demetriou J. Irrigation Channels With Differential Roughness, 2nd Nat. Congress of Agricultural Mechanics, Volos, Greece, Sept., Proceedings, pp. 42-49, 2000.
- Demetriou J., Pourliotis K. and Sarantos P. Flow in Open Channels Lined With Different Materials, 6th Greek Congress on Mechanics, Thessaloniki, Greece, July, pp. 108-113, 2001.
- Demetriou J. Differential Roughness in Open Channels With Large Roughness Elements, EYE Congress, April, Thessaloniki, Greece, 7 pages, 2003.
- Ghosh S. Boundary Shear Distribution in Channels With Varying Wall Roughness, Institution of Civil Engineers, Proceedings, London, UK, Vol. 53, Part 2, Dec., pp. 529-543, 1972.
- Pyle R. and Novak P. Coefficient of Friction in Conduits With Large Roughness, *Journal of Hydraulic Research*, 19, N^o 2, pp. 119-140, 1981.
- Ramesh R., Data B., Bhallamudi M. and Narayana A. Optimal Estimation of Roughness in Open-Channel Flows, *Journal of Hydraulic Engineering*, ASCE, Vol. 126, N^o 4, April, pp. 299-303, 2000.
- Robertson J., Bajwa M and Wright S. A General Theory for Flow in Rough Conduits, *Journal of Hydraulic Research* 12, N^o 2, pp. 223-240, 1973.
- Rosso M., Shiara M. and Berlamont J. Flow Stability and Friction Factor in Rough Channels, ASCE, *Journal of Hydraulic Engineering*, Vol. 116, N^o 9, Sept., pp. 1109-1119, 1990.

CRITERIA FOR USING A ONE-DIMENSIONAL HYDRODYNAMIC MODEL FOR TWO BARRAGES SITUATED IN MEDITERRANEAN REGION

Takkouk Saddek

Maître Assistant, Université de Batna, Algérie

Xavier Casamitjana

Maître de conférence, Université de Girona, catalogne, Espagne

Abstract:

In order to characterise the barrage regime is usually very important to check the validity of assumptions on which its application is based. Accordingly, in this study we have estimated a set of numbers comprising the Lake number, the Wedderburn and Burger numbers, and inflow and outflow Froude numbers. Therefore we have used two years worth of time series data (2000-2001) consisting of temperature profiles, morphometry data, meteorological data, and inflow and withdrawal data of two reservoir located in the Mediterranean region.

Key Words: Hydrodynamic., Reservoir regime

Introduction

The understanding of lake hydrodynamics has made much progress in the last twenty years. However, it is still difficult for the general limnological practitioner to gain a quantitative description of the hydrodynamical regimes in a particular lake at a particular time (Imberger, J. (1994))

In general hydraulics characterisation of flow is based essentially on the Reynolds and Froude numbers. So flows are subdivided into laminar or turbulent and super or sub-critical. Therefore, an analogy methodology may be established to permit the limnologist to classify hydrodynamic regime in lakes. Such hydrodynamical regime is important because the mixing and transport process operating in a lake determine, to a large degree, the ecological response of the lake to meteorological forcing, inflows and outflows. Thus, one-dimensional model like the DYRESM is applied to a Mediterranean reservoir such as the Sau or Boadella, it is important to check the validity of assumptions on which its application is based. Accordingly, in this study we have estimated a set of numbers comprising the Lake number, the Wedderburn and Burger numbers, and inflow and outflow Froude numbers. For both reservoirs, we have used two years worth of time series data (2000-2001) consisting of temperature profiles, morphometry data, meteorological data, and inflow and withdrawal data. The temperature profiles have been converted to density profiles using the UNESCO (1981) state formula equation. Later, we have used these density profiles to calculate the non-dimensional numbers.

Lake Number

The Lake number L_N is a dimensionless parameter defined as the ratio of the moments of the stabilizing force of gravity. This. L_N describes the water upwelling from the hypolimnion to the surface layer, expressed by the equation:

$$L_N = \frac{(Z_g - Z_o)Mg \left(1 - \frac{Z_T}{Z_H}\right)}{A^{3/2} \left(1 - \frac{Z_g}{Z_H}\right) \rho_0 u_*^2} \quad (1.1)$$

where Z_o is the centre of gravity of the water mass and Z_g is the centre of volume for the entire lake body; Z_T is the height to the centre of the metalimnion; Z_H is the depth from the bottom

of the reservoir; M is the total mass of water [Kg]; g is acceleration due to gravity [$m.s^{-2}$]; ρ_0 is the average water density [$kg.m^{-3}$]; A is the surface area of the reservoir $A(Z)$; and u_* is the water friction velocity [$m.s^{-1}$].

Z_0 and Z_g are defined respectively as follows:

$$Z_0 = \frac{\sum_{i=0}^{Z_H} \rho_i(Z_i) Z_i \cdot A(Z_i)}{\sum_{i=0}^{Z_m} \rho_i(Z_i) \cdot A(Z_i)} \quad (1.2)$$

$$Z_g = \frac{\sum_{i=0}^{Z_m} Z_i \cdot A(Z_i)}{\sum_{i=0}^{Z_m} A(Z_i)} \quad (1.3)$$

$$u_*^2 = \frac{\rho_a}{\rho_o} C_D \cdot U_{10}^2 \quad (1.4)$$

where U_{10} is the wind velocity at 10 m above the water surface [$m.s^{-1}$]; C_D is the drag coefficient = $1.3 \cdot 10^{-3}$ [dimensionless]; and $\frac{\rho_a}{\rho_o}$ is the ratio between air and water densities = $1.2 \cdot 10^{-3}$ [dimensionless].

Under these circumstances, the isopycnals are expected to be primarily horizontal and little seiching of the seasonal thermocline or turbulent mixing in the hypolimnion are expected. For $L_N \ll 1$, stratification is weak with respect to wind stress. Under these circumstances, the seasonal thermocline is expected to experience strong seiching and the hypolimnion is expected to experience extensive turbulent mixing due to internal shear (Imberger 1989). Thus the hypolimnion water, very rich in nutrients, will reach the surface layer during the wind episode (Imberger 2001).

Wedderburn Number

The Wedderburn number represents the ratio of the baroclinic restoring force to the wind disturbance force, or the ratio of the restoring moment about the centre of the volume of the lake to the disturbance moment. W describes the upwelling of water from the metalimnion into the water surface, expressed by:

where g' is the modified acceleration due to gravity across the uppermost thermocline.

$$W = \frac{g' h^2}{u_*^2 \cdot L} \quad (1.5)$$

This is represented by $g' = \frac{\Delta\rho}{\rho_o} g$, where $\Delta\rho$ is the density difference between the surface layer and the mean water density, and h is the depth of the diurnal thermocline; u_* is the water shear velocity due to wind stress and approximated by the bulk aerodynamic formula as previously defined in the Lake number calculation; and L is the length of the lake.

e. For $W \gg 1$, tilting of the isotherms due to applied wind stress will be small and horizontal variations negligible. This coincides with strong stratification, light winds, and slow deepening of the mixed layer. For $W \ll 1$, deepening is dominated by internal shear production and will occur over a much shorter time scale than horizontal convection in the surface layer (Imberger & Patterson 1990). Where W is small and L_N large, only the upper region of the thermocline will respond to wind forcing. Where W and L_N are small, the lake as a whole should respond, and vertical mixing should occur throughout it (Imberger and Patterson 1990).

Burger Number

The Burger number S_i is an indicator of the influence of the earth's rotation on water motion in reservoirs; influence on the water internal waves. S_i is expressed by:

$$S_i = \frac{c_i}{L_w f} \quad (1.6)$$

where L_w is the width of the reservoir and $\frac{c_i}{f}$ is the Rossby radius; c_i is the wave velocity expressed by:

$$c_i = \sqrt{g' H} \quad (1.7)$$

$g' = g \cdot \frac{\Delta\rho}{\rho_0}$ is the reduced gravity; $\Delta\rho$ is the difference between surface water and mean H is the mean reservoir water depth, which depends on the inflow entering the reservoir and water withdrawal from the reservoir. f is the Coriolis parameter equal to the double rate of rotation of the earth at the latitude of the lake, $S_i = 1$ is the critical value indicating that the rotation is of the same magnitude as gravity. When $S_i \gg 1$, the internal oscillations increasingly take on the characteristics of simple gravitational seiches (Antenucci & Imberger 2001).

When $S_i \ll 1$, the waves have characteristics similar to those of an inertial oscillation, with the majority of the energy in the wave being in the form of kinetic energy.

Inflow Froude number

The regime behaviour of the river inflow entering the reservoir is described by the inflow Froude number, F_{ri}

$$\frac{h_i}{H} = \left[\frac{Q_i}{g_i^{1/2} H^{3/2} \cdot L_w} \right]^{2/3} = F_{ri}^{2/3} \quad (1.8)$$

where Q_i is the peak inflow discharge; H is the total depth of the reservoir; $g_i = \frac{\Delta\rho}{\rho_0}$ is the reduced gravity; $\Delta\rho$ is the difference between inflowing water density and the mean reservoir water density; ρ_0 is the mean water density; and $F_{ri} = 1$ is the critical value of the plunge or rise. Where $F_{ri} \gg 1$, the inflow is too large to separate as an underflow or an overflow. When $F_{ri} \ll 1$, the inflow separates as an underflow or an overflow. Once it has been established that the river water underflows ($\Delta\rho_i > 1$ and $h_i/H < 1$), it is necessary to carry out a more detailed analysis to estimate its entrainment into the downflow and thus the depth of the inflow intrusion.

Outflow Froude number

The outflow Froude number characterises the type of water withdrawal, expressed by the equation:

$$F_{r_0} = \left[\frac{Q_0}{g^{1/2} H^{5/2}} \right] \quad (1.9)$$

where Q_0 is the outflow discharge; $g' = \frac{\Delta\rho}{\rho_0}$ is the reduced gravity; $\Delta\rho$ is the difference between outflow water density and the mean reservoir water density; ρ_0 is the mean water density; H is the total depth of the reservoir. Where $F_{r_0} \ll 1$, then the outflow is selective from a depth corresponding to the outflow level.

Methods

Lake numbers were computed for each recorded temperature profile using equation 1.1. We used twenty measured over two years (2000-2001) for the Sau Reservoir and fifteen for the Boadella

Reservoir.. As water depth decreases and increases with water inflow and outflow, the water surface and the reservoir volume change. To calculate both, interpolation from the bathymetric data was necessary. To determine the daily velocity friction we used equation 2.4, with a drag coefficient of $1.3 \cdot 10^{-3}$, and a ratio between air and water densities of $1.2 \cdot 10^{-3}$. Wind velocity is given in meteorological data in Fig. 2.8 for the Sau Reservoir and in Fig. 2.21 Chapter 2 for the Boadella Reservoir. The mean water density was taken as $1000 \text{ kg} \cdot \text{m}^{-3}$. The depth of the reservoir Z_H and its area were estimated using morphometry data (Sections 2.4.1 and 2.5.1, Chapter 2). The estimated lake numbers are 712 in Sau reservoir and 636 in Boadella Reservoir

As described above for lake number, Wedderburn number values were estimated by linear interpolation between existing profiles. For each temperature profile using equation 1.5 daily friction velocity u_* had been previously defined in the calculation of the Lake number.

Modified acceleration due to gravity g' is the ratio of the difference in diurnal thermocline density from the mean water density, divided by the mean water density multiplied by gravity g . L is the average length of the reservoir: $L = 3000\text{m}$ for Sau and $L = 1500\text{m}$ for Boadella.

The Burger number was obtained using equation 1.6 in which the Rossby radius was calculated as the ratio between longwave phase velocity and inertial frequency.

Longwave phase speed is a function of reduced gravity g' , which depends on the difference between surface water density and mean water density divided by the mean reservoir water density. Inertial frequency depends on the latitude of a reservoir; the Sau and Boadella reservoirs are located in the same region and have approximately the same altitude.

The inflow Froude number was calculated using equation 1.8. The discharge or inflow rate Q_i is the peak inflow discharge entering the reservoir and was estimated from the inflow file: approximately $217 \text{ m}^3 \cdot \text{s}^{-1}$ for the Sau Reservoir and $111 \text{ m}^3 \cdot \text{s}^{-1}$ for Boadella. The reduced gravity g'_i is the difference between inflow density and mean reservoir density divided by the mean density. The average width of the reservoir is $B = 700\text{m}$ for Sau and 600m for Boadella. The total depth of the reservoir H is variable, depending on the volume entering and leaving it.

The outflow Froude number was computed using equation 1.9. The outflow Q_o was deduced from the withdrawal data. The reduced gravity for outflow is the ratio of the outflow density minus the mean water density to the mean water density. H is the reservoir water depth.

Results and discussions

Lake number

The L_N values obtained for the Sau Reservoir corresponding to lowest wind velocity 0.8m/s and the minimum is around 2 approximately in the middle of December 2001. Thus, $L_N > 1$ indicating that stratification is the dominant force when compared to wind stress. Minimum and maximum wind velocities are 0.8 and 2.9 (m/s) respectively. There is low turbulence and no mixing in deeper water, due to the isolation of the bottom waters from the surface. Consequently, However at the end of 2001 the Lake number was equal 2, indicating that the hypolimnion was relatively mixed.

The L_N of the Boadella Reservoir fluctuated from a minimum value of 5 in the beginning of June 2000 corresponding relatively to high wind velocity 5.6m/s (see section 2.5.2.4 Chapter 2) to a maximum of 2226 in the end of Mars 2000 corresponding to the lowest wind velocity 0.4 m/s . Compared to Sau the stratification is strong in the first year (2000) and during the second (2001). This is linked to low wind velocity, for which minimum value and maximum values are 0.8 and 2.9 for Sau and 0.7 and 2.1 for Boadella (see Figs. 2.8 and 2.21, Chapter 2.) It should also be noted that morphometry plays an important role in the determination of Lake number, and that determination of the thermocline is difficult. In both reservoirs, water upwelling from the hypolimnetic occurs at the winter end autumn in 2000 and 2001. Fig. 1.1 shows the differences in Sau and Boadella's Lake numbers. In this figure we see that lake number diminished corresponding to the peak inflow (see Fig. 2.10B, Chapter 2) in the end of the year 2000 means that may inflow also contribute in hypolimnion mixing.

However, in Boadella Reservoir the lack of profiles from the end of August until the end of April influence the estimation of lake number for comparison purposes versus inflow. Also, withdrawal being selective may enhance stratification. For both reservoirs through the period 2000 to 2001 Lake numbers are bigger than the critical value. This indicates that stratification is relatively strong and that the deepest part of the hypolimnion remains unmixed if not influenced by inflow/outflow and cooling forces. Unfortunately, there are fewer profiles for Boadella than there are for Sau. However, the variation in its Lake number is small, as is the case for Sau. Also, it should be noted that the wind velocity in Boadella, on the days when profiles were taken, was relatively low. We can therefore conclude that both reservoirs were not completely mixed, at least at the time the profiles were taken. It should also be noted, however, that the profiles were taken during the day. In winter, during the day, and the mean velocity of the wind is in general not high enough to overcome stratification. For example, on 09/02/2000:

$Z_g - Z_0 = 1.6 \cdot 10^{-3} m$, $u^* = 0.0017 m \cdot s^{-1}$ and $L_N = 60$. However, it is very likely that at night time $Z_g - Z_0$ was close to 0 and L_N went to 0. In these circumstances hypolimnetic mixing could be expected.

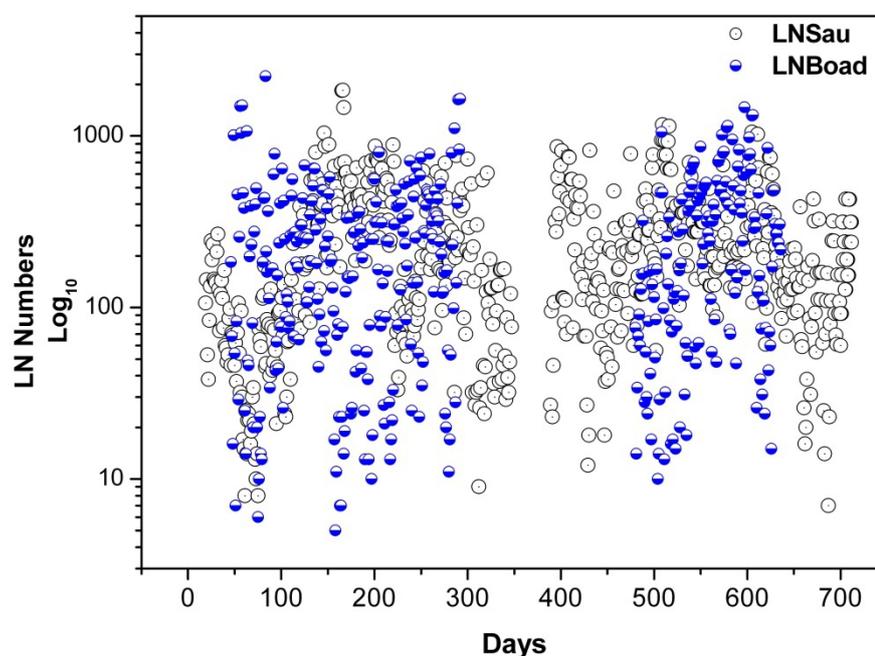


Fig 1.1 The differences in the Sau and Boadella Lake numbers

Wedderburn number

The Wedderburn numbers in the Sau Reservoir do change substantially, tending to oscillate over the two years between 0 and 750. Only one value is below 1 corresponding to the value in the beginning of November 2000 corresponding to the high wind velocity 5.6 m/s. 0, indicating that wind stress on the surface of the lake is able to overcome the stratification in the water column and metalimnetic water can be expected to be vented into the surface layer if we neglect other disturbing forces such as inflow/outflow and cooling.

Wedderburn numbers in the Boadella Reservoir range from approximately 0 to 454 at the end of May and are smaller than those of Sau, See Fig. 1.2 for Sau and Boadella Wedderburn number differences.

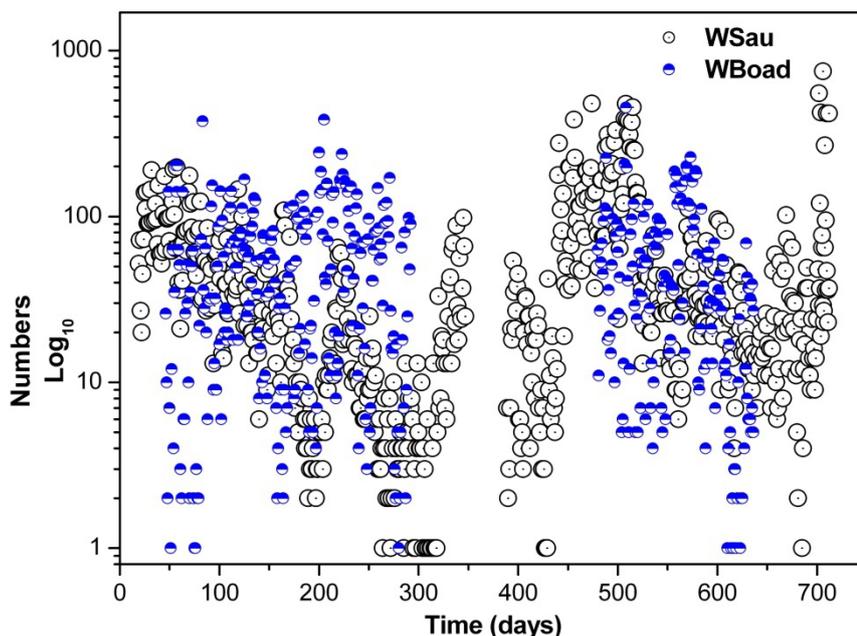


Figure 1.2 Sau and Boadella Wedderburn numbers

Burger number

The Sau Reservoir experiences substantial fluctuation in Si values throughout the two years, with all the values above 1. Burger numbers range from 1.20 to 2.3, with a mean value of approximately 1.75. both constant. The Burger number is high than the critical value 1.0 and therefore the rotational effects are insignificant. However, given that Si is close to 1, rotation would be discarded. The Sau and Boadella Burger numbers are shown in Fig.1.3

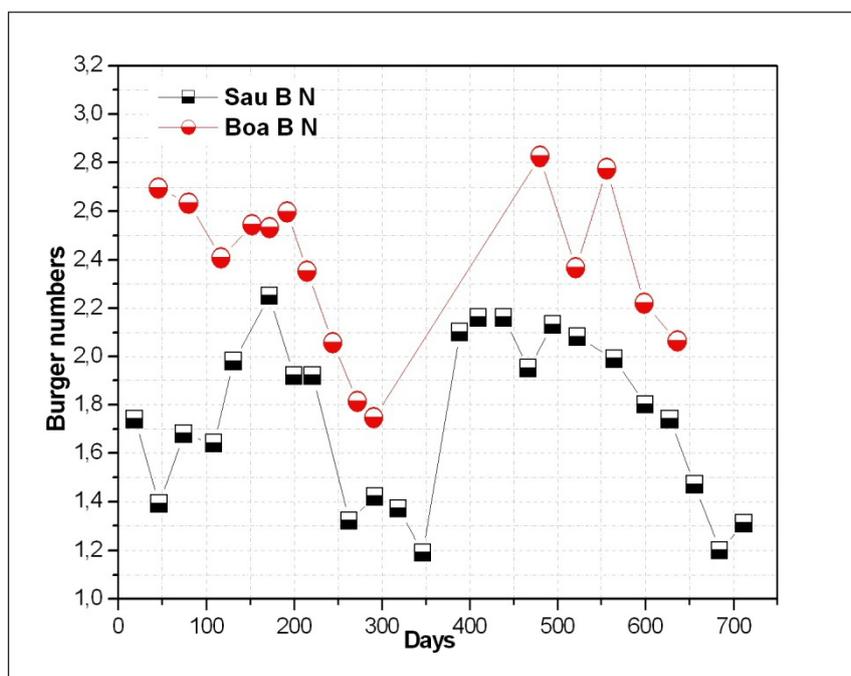


Figure 1.3 Comparison between the Sau and Boadella Burger numbers. Sau is represented by squares, and Boadella by circles.

Inflow Froude number

The average Sau inflow Froude number is 0.01. It is important to note that it would be necessary to carry out a more detailed analysis to estimate the entrainment into the inflow and the depth of the downflow insertion.

The inflow Froude number for Boadella is smaller for Sau. Arnera rivers which flow into Boadella. Fig.1.4 shows the inflow Froude numbers for Sau and Boadella.

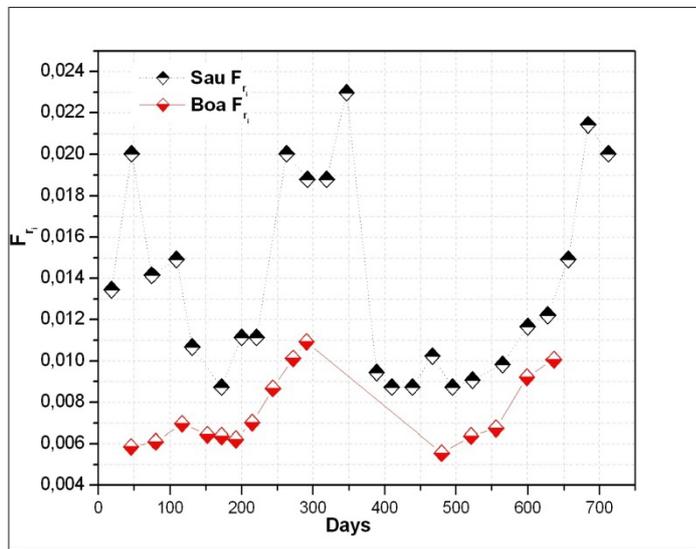


Figure 1.4 Sau and Boadella inflow Froude numbers

Outflow Froude number

Sau’s average outflow Froude number is about 0.01, which is lower than the critical value of 1.0. Boadella’s average outflow Froude number is 0.001, ten times smaller than Sau’s one. This indicates that in both reservoirs the outflow is selective. Fig. 1.5 shows the Sau and Boadella outflow Froude numbers.

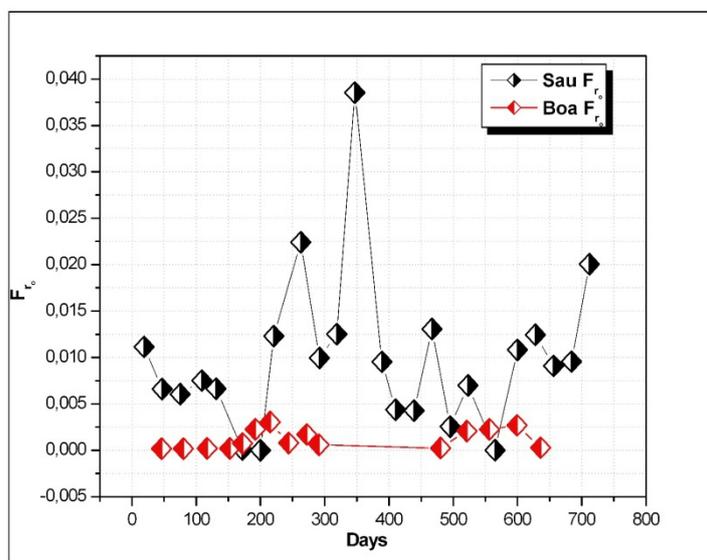


Figure 1.5 Sau and Boadella outflow Froude numbers

Conclusions

The Lake number L_N while the Wedderburn number W is an indicator of surface layer mixing. In both reservoirs, L_N is large and W is small for the profiles under study. This means that

the the one-dimensional DYRESM model can be used. (Imberger, J. 2001). Where $L_N < 1$, It should be pointed out, however, that to calculate the L_N we have used daily mean wind taken during the daytime. It is likely that there are times during the night and/or when there are very high winds that $L_N < 0$, Wedderburn number are not enough because it leads to an overestimation of these numbers especially when the days when profiles were taken are calm therefore daily temperature profiles and daily inflow are needed to judge perfectly whether the one dimensional hydrodynamic may be applied or not. For this reason, daily lake numbers, which were obtained by interpolation of daily temperature from monthly observed water temperature profiles, density structure is approximately horizontal and the one dimensional assumption is valid. Also we have estimate the daily wedderburn number as we did for lake numbers. Wedderburn numbers, are mostly greater than one indicating that tilting of the isotherms is small; this is due to the weakness of the wind stress.

The Burger number S_i is slightly larger than one, indicating that rotation might be discarded.

$F_{ri} \ll 1$ and $F_{ro} \ll 1$ indicate that in both reservoirs separation occurs

References:

- Afres, W and L. Andreas. 2003. Small-scale Hydrodynamic in lakes *Ann.Rev.Fluid Mech.* 35: 373-412.
- Ambrosetti, W and L. Barbanti. 2001. Temperature, heat content, mixing and stability in Lake Orta: a pluriannual investigation. *Limnol.* 60(1): 60-68.
- Ambrosetti, W., L.Barbanti and N. Sala. 2003. Residence time and physical processes in lakes. *Limnol.* 62(Suppl. 1): 1-15.
- Aminot, A. And R. K erouel. 2004. Dissolved organic carbon, nitrogen and phosphorus in the N-E Atlantic and the N-W Mediterranean with particular reference to non-refractory fractions and degradation. *Deep-Sea Research. I* 51: 1975-1999.
- Amrosetti, W and L. Barbanti. 2002. Physical limnology of Italian lakes. 1. Relationship between morphometry and heat content. *Limnol.* 61(2): 147-157.
- Andrad ttir, H.   and H. M. Nepf. 2000. Thermal mediation by littoral wetland and impact on lake intrusion depth. *Water Ressources Research.* 36(3): 725-735.
- Andrew, J. Tanentzap, D. P. Hamilton, and N. D. Yan. 2007. Calibrating the Dynamic Reservoir Simulation Model (DYRESM) and filling required data gaps for one-dimensional thermal profile predictions in a boreal lake. *Limnol. Oceanogr.Methods.* 5: 484-494.
- Antenucci, J. P and J. Imberger. 2003. The seasonal evolution of wind/internal wave resonance in Lake Kinneret *Limnol. Oceanogr.* 48(5): 2055-2061.
- Antenucci, J and A. Imerito. 2003 The CWR Dynamic Reservoir Simulation Model DYRESM. User Manual. Centre for Water Research. University of Western Australia.
- Antenucci, J., and A. Imerito. 2002 The CWR Dynamic Reservoir Simulation Model DYRESM. Science Manual. Centre for Water Research. University of Western Australia.
- Appt, J., J. Imberger, and H. Kobus. 2004. Basin-scale motion in stratified upper Lake Constance. *Limnol.Oceanogr.* 49(3): 919-933.
- Armengol, J., J. Toja, J., and A. Vidal. 1994. Seasonal rhythm and secular changes in Spanish reservoirs. In "Limnology Now: A Paradigm of Planetary Problems." (R. Margalef, ed.)
- Armengol, J., M. Comerma, J.C. Gracia, M. Romero, J.J. Rodriguez and A. Vidal 1999. Contribuci  al coneixement de l'ecologia acu tica de l'embassament de Sau evoluci  de l'embassament al 1998. *Aig es Ter de Llobregat.* pp 66
- Armengol, J., M. Comerma, J.C. Gracia, M. Romero, J.J. Rodriguez and A. Vidal 2000. Contribuci  al coneixement de l'ecologia acu tica de l'embassament de Sau evoluci  de l'embassament al 1999. *Aig es Ter de Llobregat.* pp 118.
- Armengol, J., J. Gracia, M. Comerma, M. Romero, J. Dolz, M. Roura, B. P. Han, A. Vidal and K. Simek. 1999. Longitudinal Processes in Canyon Type Reservoirs: The Case of Sau (N.E. Spain). *Theoretical Reservoir Ecology and its Applications* 313-345.

- Bruce, L. C., D Hamilton, J. Imberger, G. Gal, M. Gophen, T. Zohary and K. D. Hambright. 2006 A numerical simulation of the role of zooplankton in C, N and P cycling in Lake Kinneret, Israel. *Ecological Modelling*.193. 412-436.
- Carr, G. M., H.C. Duthie and W.D. Taylor. 1997. Models of aquatic plant productivity: a review of the factors that influence growth. *Aquatic Botany*. 59 195-215.
- Casamitjana, X., T. Serra, J. Colomer, C. Baserba and J. Pérez-Losada. 2003. Effects of the water withdrawal in the stratification patters of a reservoir. *Hydrobiologia*. 504: 21-28.
- Casamitjana, X., T. Serra, M. Soler, J. Colomer. 2002. A study of the evolution of the particle boundary layer in a reservoir, using laser particle sizing. *Water Research* 36: 4293-4300.
- Chan, T. U., D. P. Hamilton, B. J. Robson, B. R. Hodges and Chris. Dallimore. 2002. *Impacts of hydrological Changes on Phytoplankton Succession in the Swan River, Western Australia*. *Estuaries*. 25(68): 1406-1415.
- Colomer, J., T. Serra, M. Soler and X. Casamitjana. 2002. Sediment fluidization events in a lake caused by large monthly rainfalls *Geophysical research Letters*. 29, 0.1029/2001GL014299
- Colomer, J., T. Serra, M. Soler, X. Casamitjana. 2003. Hydrothermal plumes trapped by thermal stratification *Geophysical Research Letters* 30(21): 2092,
- Colomer, J., E. Roget and X. Casamitjana 1996. Daytime heat balance for estimating non-radiative fluxes of lake Banyoles, Spain. *Hydrological Processes*. 10: 721-726.
- Comerma, M., J. C. Garc, M. Romero, J. Armengol, K. Šimek.2003. Carbon flow dynamics in the pelagic community of the Sau Reservoir (Catalonia, NE Spain). *Hydrobiologia* 504: 87-98.
- Dallimore, C. J., J. Imberger and B. R. Hodges. 2004 Modeling a Plunging Underflow *Journal of Hydraulic Engineering*. **130**(11): 1068-1076.
- Dallimore, C. J., B. R. Hodges and J. Imberger. 2003 Coupling an Underflow Model to a Three-Dimensional Hydrodynamic Model. *Journal of Hydraulic Engineering*, ABCE. **748**
- Ewing, T., T.R. Romero, J. Imberger, J. Antenucci and A. Deen. 2004. A real-time reservoir decision support system. (6th International conference on Hydro informatics World) scientific publishing Company, ISBN 981-238-787-0
- Fernandes, R. L and J. Imberger. 2006. Bed roughness induced entrainment in a high Richardson number underflow. *Journal of Hydraulic Research*. **44**(6): 725-738.
- Fischer, J. 1979. Modelling of water quality processes in lakes and reservoirs. *Hydrological Sciences*. 24: 2-6.
- Fleenor, W. E. 2001. Effects and Control of Plunging Inflows on Reservoir Hydrodynamics and Downstream Releases PhD thesis pp 160
- Fleenor, W. E. 2001. Effects and Control of Plunging Inflows on Reservoir Hydrodynamics and Downstream Releases PhD thesis pp 160
- Fricker, P.D and H.M. Nepf. 2000. Bathymetry, stratification, and internal seiche structure *Journal of Geophysical research* **105** (6) 14237-14251.
- Fu, S. and B. R. Hodges. 2005. The Grid Scale Problem in Numerical Modeling of Flow Around Large Woody Debris in Rivers. University of Texas at Austin Graduate Student Research Conference. April 28, 2005
- Fu, S. and B. R. Hodges. 2005. The Grid Scale Problem in Numerical Modeling of Flow Around
- Gal, G., J. Imberger, T. Zohary, J. P. Antenucci, A. Anis, and T. Rosenberg. 2003. Simulating the thermal dynamics of Lake Kinneret. *Ecological Modelling*. **162**(1-2): 69-86.
- Hodges, B. R and S. K. Delavan. 2004. Numerical Diffusion in Hydrostatic Models of Internal Waves. 17th ASCE Engineering Mechanics Conference, June 13-16, 2204, University of Delaware, Newark, DE, EM.
- Hodges, B. R., J. Imberger, B. Laval and J. Appt. 2000. Modelling the Hydrodynamics of Stratified Lakes. *Hydroinformatics 2000 Conference*, Iowa Institute of Hydraulic Research, 23-27.
- Hurtado, J. V. Basin-Scale Hydrodynamics in a Mediterranean reservoir. Implications for the phytoplankton dynamics. PhD Theses. University of Girona. 133 pp.
- Imberger, J. 2001. Characterizing the dynamical regimes of a lake 6th Workshop on physical Process in Natural Waters University of Girona: 77-92
- Yeates. P.S., J. Imberger. 2003. Pseudo two-dimensional simulations of internal and boundary fluxes in stratified lakes and reservoirs. *Intl.J.River Basin Management*. IAHR and INBO. **1** (4). 297-319.

SPURIOUS RESPONSE SUPPRESSION IN THREE POLES HAIRPIN BANDPASS FILTER USING DEFECTED GROUND STRUCTURES

N. Amar Touhami

Abdel Malek Essaadi University/ Faculty of Sciences, Tetuan, Morocco

A. Zakriti

Abdel Malek Essaadi University/ National School of Applied Sciences, Tetuan, Morocco

M. Lamsalli

M. Boussouis

Abdel Malek Essaadi University/ Faculty of Sciences, Tetuan, Morocco

A. Tribak

National Institute of Posts and Telecommunications, Rabat, Morocco

Abstract:

The demand of high data rate in wireless communications requires higher system bandwidth. In this work, the design of wideband BPF with suppression of higher harmonics has brought about by considering the UWB BPF as a hairpin filter structure with DGS BSF. As a result, the structure is simple and compact. The expected result of a simple BPF together with DGS is to have a compact UWB bandpass filter with significantly higher harmonics suppression. The filter performances are exhibits $|S_{21}|$ more than -0.2 dB, $|S_{11}|$ less than -10dB, center frequency of proposed filter is around 4GHz with operating bandwidth 38% and can suppress unwanted passband to below -25 dB

Key Words: Defected Ground Structures, Bandpass filter, Stopband filter, Spurious response suppression

Introduction

In modern wireless and mobile communication systems, filters are always playing important and essential roles. Planar filters are particularly popular structures because they can be fabricated using printed circuit technology and are suitable for commercial applications due to their compact size and low-cost integration [1]. For a planar filter design, it is necessary to select proper resonator types since resonators are the basic components of a filter. To reduce the resonator size, several types of resonators such as the U-shaped hairpin resonators [2], [3], the open-loop resonators [4] and the folded open-line resonators [5], [6] have been proposed to design different kinds of bandpass filters.

The stringent requirements of modern microwave communication systems demand high performance and compact filtering structures. Recently, there has been an increasing interest in electromagnetic band gap (EBG) materials for applications in microwave and millimetre wave filters and other devices. Several compact and high performance filters have been reported using generic structures called the defected ground structures (DGS). Since DGS cells have inherently resonant properties, many of them have been used in filtering circuits to improve the stop and pass band characteristics. The DGS have been proposed for improving the spurious response of microstrip low pass filters [7-8] and coupled microstrip line bandpass filters [9-10]. In all these reports, DGS are not viewed as the central building blocks, they are rather used to enhance the response of already designed devices such as filters and couplers.

In this paper, a design of wideband BPF with suppression of higher harmonics has brought about by considering the UWB BPF as a three poles Hairpin BPF with DGS BSF. The expected result of a simple BPF together with DGS microstrip is to have a compact UWB bandpass filter with significantly higher harmonics suppression.

Three Poles HAIRPIN BPF Structure Design

In this paper, a design of wideband bandpass filter is based on a conventional Hairpin bandpass structure. Based on the dimensions listed in figure 1, simulation of three poles hairpin line BPF with CST software was performed as shown in figure 2. We consider only $|S_{11}|$ and $|S_{21}|$ because the filters designed are symmetrical so $|S_{11}|$ is very nearly $|S_{22}|$ and $|S_{21}|$ is very nearly $|S_{12}|$.

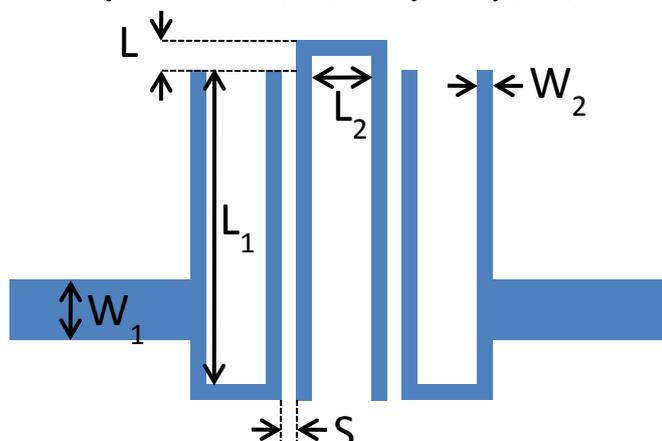


Figure 1: Layout of three poles hairpin line BPF on a 1.27mm thickness substrate with a relative dielectric constant of 6.15 ($W_1 = 1.85\text{mm}$, $W_2=0.5\text{mm}$, $S=0.3\text{mm}$, $L_1=10\text{mm}$, $L_2=L_3= 1\text{mm}$)

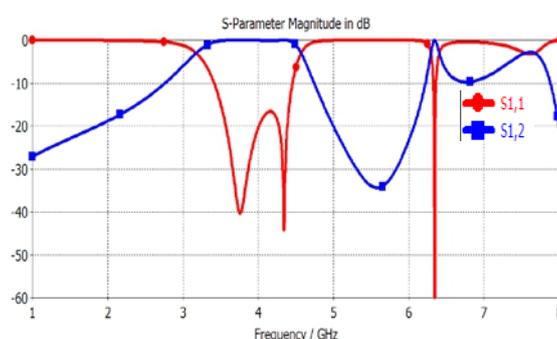


Figure 2: Simulation response of three poles hairpin line BPF

From the simulation response of three poles hairpin line BPF the simulated performance has 3 dB bandwidth from 3.2-4.7 GHz and shows second passband around 6.4 GHz. The response of hairpin line filter exhibits sufficient bandwidth that covers an ultra-wideband application but its undesired second passband must be minimized. The deployment of a bandstop filter as cascading section with the bandpass filter is a classical technique to reduce unwanted passband. By taking the advantages of an attenuation pole of the designed DGS having resonant frequency at the second passband, the unneeded passband is drastically reduced.

DGS BSF Design

Several Compact and high performance components have been reported by using the generic structure called the defected ground structure (DGS) for the microstrip line. DGS BSF has been attractive to obtain the function of unwanted frequency rejection and circuit size reduction.

The structure of DGS BSF can provide cutoff frequency and attenuation pole in some frequencies without any periodic array of DGS [11]. The defected areas can be realized by dumbbell slot, rectangular, circular slot, arrow slot shape and slot variation as shown in Figure 3.

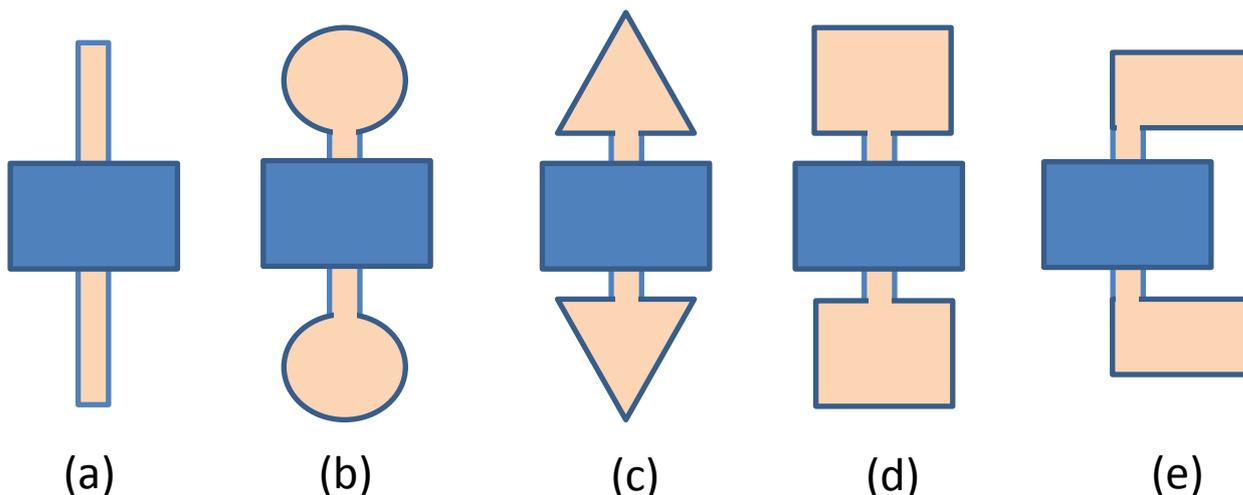


Figure 3: Example of DGS BSF shapes (a) rectangular (b) circular (c) arrow and (d) dumbbell (e) slot variation

The three dimensional view of microstrip line with DGS shape shown in Figure 4(a), the etched lattice shape of the proposed DGS section is located on the backside metallic ground plane.

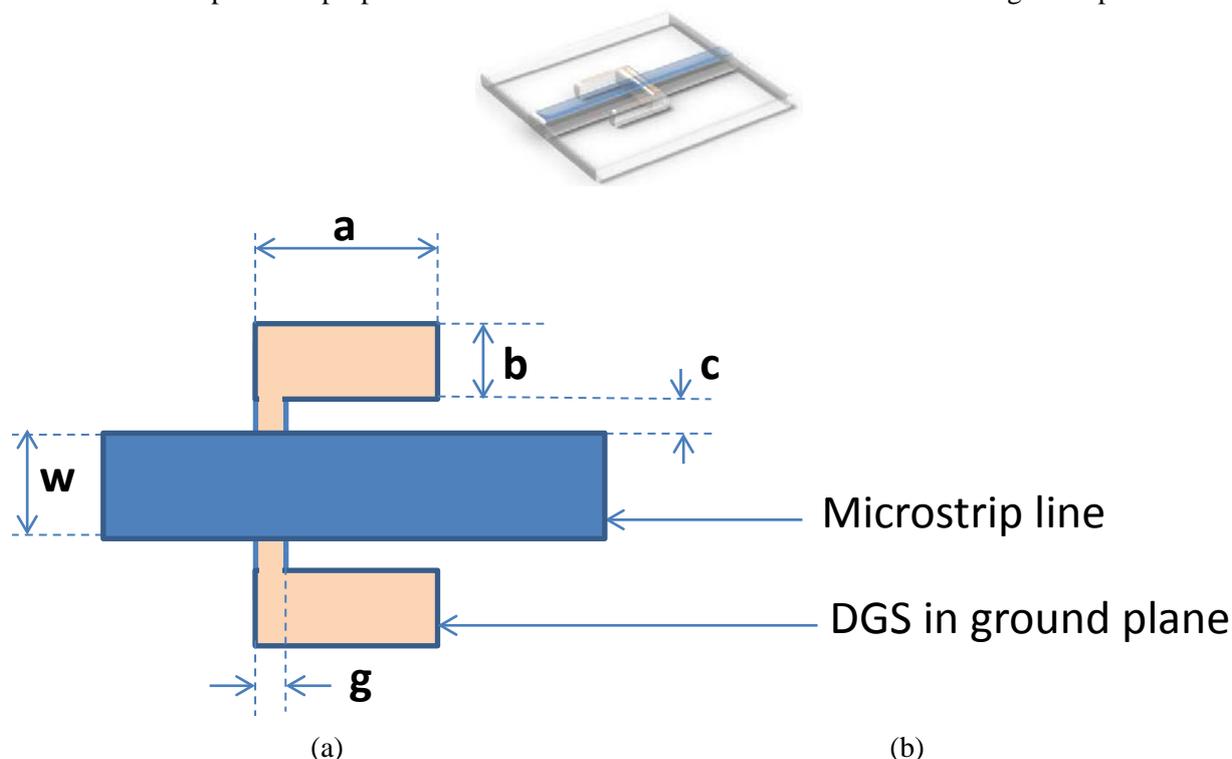


Figure 4: (a) Three dimensional view of microstrip line with DGS slot variation shape (b) Layout of DGS

A DGS BSF has four design parameters (a , b , c and g) (see figure 4b). Figure 2 shows the second passband of the three poles hairpin line BPF occurring around 6.4 GHz. Therefore, the DGS BSF must be designed to have attenuation pole at unwanted that frequency. In this paper, we use simulation oriented to design and optimize for DGS BSF layout. In order to investigate the frequency characteristics of the DGS section, we simulated the DGS unit section and considered the effect of parameter that affect to $|S_{21}|$. The simulation results show that they are bandstop filter characteristics, as expected. The cutoff frequency mainly depends on the square area in a ground plane. There is also attenuation pole location, which is due to the gap width.

Figure 5 shows the response of $|S_{21}|$, when increase a by 0.5 mm per one step the result is attenuation pole decreased.

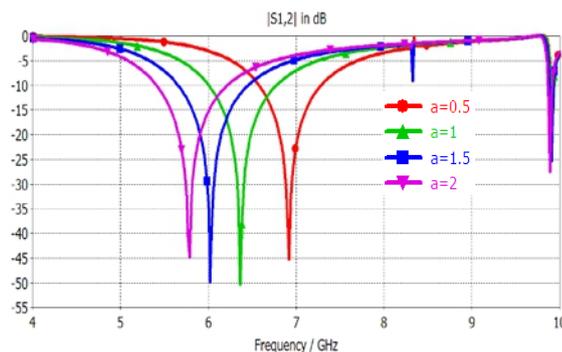


Figure 5: Investigate parameter a (varies a , $b = 4$ mm, $c = 1$ mm and $g = 0.2$ mm)

Figure 6 shows the response of $|S21|$, when increase b by 1 mm per one step the result is attenuation pole decreased as same as a .

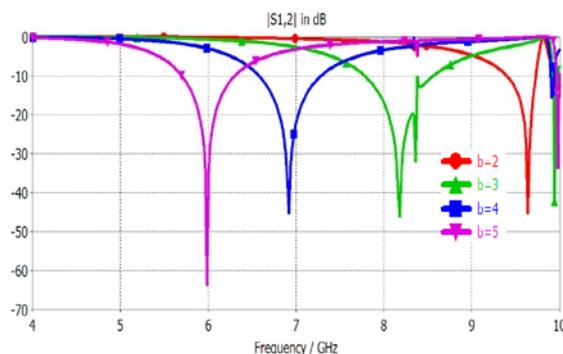


Figure 6: Investigate parameter b ($a = 0.5$ mm, varies b , $c = 1$ mm and $g = 0.2$ mm)

Figure 7 shows the response of $|S21|$, when increase c by 1 mm per one step the result is attenuation pole decreased as same as a and b .

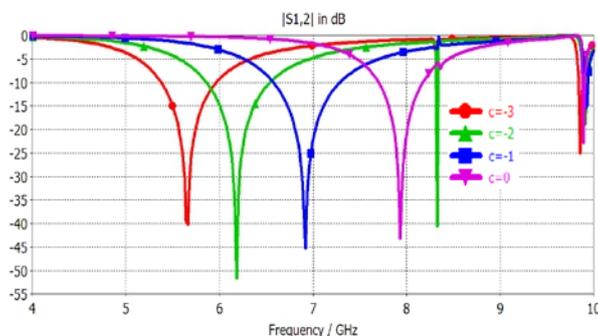


Figure 7: Investigate parameter c ($a = 0.5$ mm, $b = 4$ mm, varies c and $g = 0.2$ mm)

Figure 8 shows the response of $|S21|$, when increase g by 0.2 mm per one step the result is attenuation pole increased.

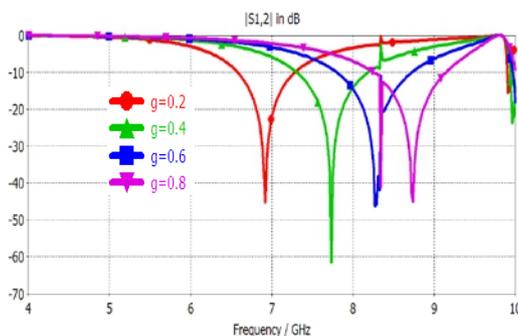


Figure 8: Investigate parameter g ($a = 0.5$ mm, $b = 4$ mm, $c = 1$ mm and varies g)

From investigating all parameters of the DGS BSF, we can summarize the effect as in Table1.

TABLE I. SUMMARIZED OF INVESTIGATION PARAMETERS

Increase	Effect to attenuation pole
a	decrease
b	decrease
c	decrease
g	Increase

So the greatest response used for suppression second passband which occurs from three poles hairpin line BPF and is able to cover the passband has dimension of the DGS BSF at $a = 0.5$ mm, $b = 4$ mm, $c = 1$ mm and $g = 0.2$ mm (see figure 9)

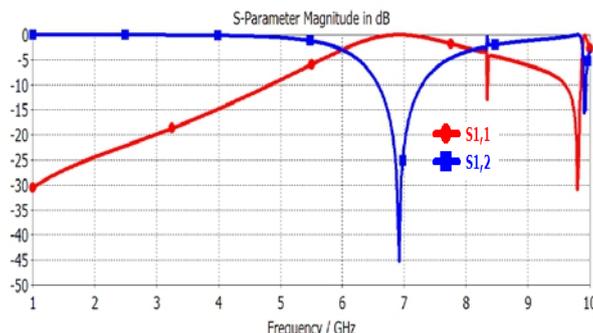


Figure 9: Simulation response of DGS BSF

HAIRPIN Filter with DGS BPF Design

From the structure of the three poles hairpin line BPF and DGS BSF, the structure of wideband BPF which significantly suppresses unwanted passband is shown in Figure 10.

In figure 11 we can see the simulation results of the hairpin BPF incorporating DGS structures. The Simulated $|S_{21}|$ achieve below -20 dB in the stopband and less than -0.2 dB in the pass band. Comparing with $|S_{11}|$ of the conventional BPF, $|S_{11}|$ of the DGS BPF is poorer, but the operating bandwidth increased.

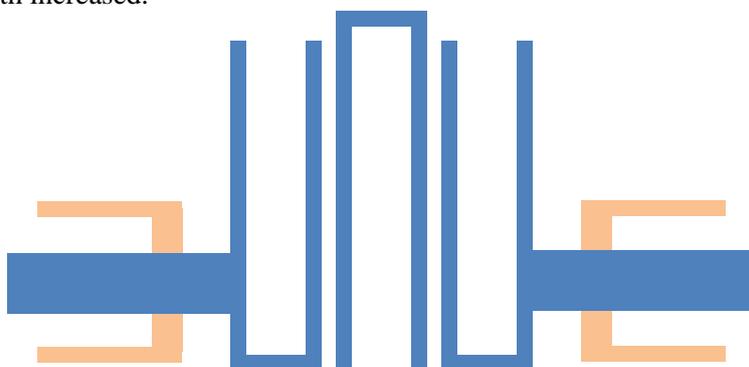


Figure 10: Layout of three poles hairpin line BPF with defected ground structure

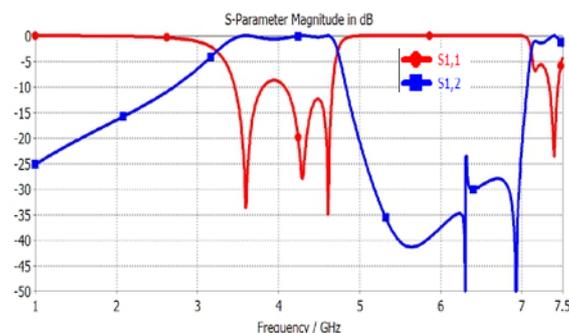


Figure 11: Simulation response of three poles hairpin line BPF with DGS BPF

Conclusion

A three poles hairpin line bandpass filter with defected ground structure of microstrip line is introduced and demonstrated in this paper. The conventional bandpass filter is firstly designed and its response proposed unwanted passband. Cascading DGS with that three poles hairpin line BPF introduces attenuation pole contributing the suppression of unwanted second passband. The filter performances are exhibits $|S_{21}|$ more than -0.2dB, $|S_{11}|$ less than -10 dB, center frequency of proposed filter is around 4GHz with operating bandwidth 38% and can suppress unwanted passband to below -20 dB.

References:

- [1]. D. M. Pozar, *Microwave Engineering*, 2nd ed. New York: Wiley, 1998, ch. 8.
- [2]. E. G. Cristal and S. Frankel, "Hairpin-line and hybrid hairpin-line/halfwave parallel-coupled-line filters," *IEEE Trans. Microw. Theory Tech.*, vol. MTT-20, no. 11, pp. 719–728, Nov. 1972.
- [3]. G. L. Matthaei, N. O. Fenzi, R. J. Forse, and S. M. Rohlfing, "Hairpincomb filters for HTS and other narrowband applications," *IEEE Trans. Microw. Theory Tech.*, vol. 45, no. 8, pp. 1226–1231, Aug. 1997.
- [4]. J. S. Hong and M. J. Lancaster, "Couplings of microstrip square openloop resonators for cross-coupled planar microwave filters," *IEEE Trans. Microw. Theory Tech.*, vol. 44, no. 11, pp. 2099–2109, Nov. 1996.
- [5]. J. T. Kuo, M. J. Maa, and P. H. Lu, "A microstrip elliptic function filter with compact miniaturized hairpin resonators," *IEEE Trans. Microw. Guided Wave Lett*, vol. 10, no. 3, pp. 94–95, Mar. 2000.
- [6]. M. Sagawa, K. Takahashi, and M. Makimoto, "Miniaturized hairpin resonator filters and their application to receiver front-end MIC's," *IEEE Trans. Microw. Theory Tech.*, vol. 37, no. 12, pp. 1991–1997, Dec. 1989.
- [7]. D. Ahn, J.S. Park, C.S. Kim, J. Kim, Y. Qian and T. Itoh, "A design of the low-pass filter using the novel microstrip defected ground structure", *IEEE Trans. Microwave Theory Tech.*, vol.49, n0.1, pp. 86- 93,Jan. 2001.
- [8]. F.R. Yang, Y. Qian and T. Itoh, " A novel uniplanar compact PBG structures for filter and mixer applications", *IEEE Microwave Theory Tech. Symposium, MTT (S)-1999*.
- [9]. J. S. Yun, G. Y. Kim, J.S. Park, D. Ahn, K. W. Kang and J. B. Lim, " A design of the novel coupled line bandpass filter using defected ground structure", *IEEE Microwave Theory Tech. Symposium, MTT (S)-2000*.
- [10]. F.R. Yang, R. Coccioli, Y. Qian and T. Itoh, "Analysis and application of coupled microstrips on periodically patterned ground plane", *IEEE Microwave Theory Tech. Symposium, MTT (S)-2000*.
- [11]. Dal Ahn, Jun-Seok Park, Chul-Soo Kim, Juno Kim, Yongxi Qian and Tatsuo Itoh. "A Design of the Low-Pass Filter Using the Novel Microstrip Defected Ground Structure." *IEEE Transactions On Microwave Theory and Techniques*. vol. 49, no. 1, Jan. 2001.

VISUALIZATION AND MEASUREMENTS OF BUBBLY TWO-PHASE FLOW STRUCTURE USING PARTICLE IMAGING VELOCIMETRY (PIV)

Hassan Abdulmouti, Associate Professor

Esam Jassim, Assistant Professor

Mechanical Engineering Dept. Prince Mohammad Bin Fahd University (PMU). Al Azezaya, Eastern Province, KSA

Abstract:

Gas injection through a bottom nozzle is very popular and has wide applications. Hence; many researchers have carried out extensive model experiments by focusing on flow field's structure using air bubbles. The bubble plume, which is a typical form of bubble flow, is known as one of the transport phenomena that have the capability to drive a large-scale convection due to the buoyancy of the bubbles. The technique of using a surface flow generated by the bubble plume is utilized as an effective way to control and collect surface floating substances in naval systems, lakes, seas, rivers, oceans especially the oil layer formed during large oil spill accidents. The surface flows generated by bubble plumes are considered to be key phenomena in various kinds of reactors, engineering processes and industrial processes handling a free surface. The motivation of this research is to broaden the understanding and demonstration of the following points: 1) The most important applications of bubbly flow and gas-liquid two-phase flow. 2) The differences of surface flow generation mechanisms among single-phase liquid jet, single phase buoyant plume, and bubble plume. 3) The relationship between the dominated parameters of bubbly flow that are demanded to describe the characteristics of the bubble generating surface flow. Such flow depends on the gas flow rate, the bubble size, void fraction, bubble velocity and the internal two-phase flow structure of the bubble plume. Laboratory experiments have been carried out in order to investigate the multi-dimensional motion of water and bubbles. The data are obtained by applying image processing and Particle Imaging Velocimetry (PIV) measurements to two kinds of visualized images: The first is visualization of the whole field around the bubble plume, and the second is that of the flow structure of bubble for the different sections of bubble regions. The surface flow is effectively generated in case of bubble plume compared to liquid jet flow since the distortion point appears in the vicinity of surface. The flow structure and bubble parameters are sensitively modulated by the gas flow rate and bubble size.

Key words: Multiphase Flow, Bubble Plume, Surface Flow, Flow Visualization, Particle Imaging Velocimetry (PIV), Bubble

Introduction

In many engineering fields such as materials, chemical, mechanical, and environmental engineering, the techniques of gas injection have been widely utilized for improving chemical reactions, waste treatment, gas mixing and resolution, heat and mass transfer, and other engineering process.

Bubbly two-phase flows have various flow structures, which are not observed in a single-phase flow, due to the complexity of the translational motion and the volumetric change of the bubbles. The variety of the flow structure often considerably impacts the performance of hydraulic machinery and chemical and bioreactor in which the bubbly media is used as a working fluid. In particular, the local flow behavior such as the mutual interaction between bubbles and vortices influences the statistical characteristics of the flow. For instance, both the generation and deformation of vortices in the liquid phase due to the inhomogeneous buoyancy caused by the local distribution of

bubbles, and the accumulation of bubbles in the vortex cores and turbulence generation due to the bubble migration, enhance the flow instability and the turbulence modification. The bubble plume is a suitable object to analyze the detailed flow structure in bubble flow because it involves various interactions between the bubble and the ambient liquid flow when the bubble rises from the bottom to the upper free surface (Hassan 2002, 2003, 2006, 2011, Hassan and Tamer 2006, Murai and Matsumoto 1998, Matsumoto and Prosperetti 1997, Matsumoto and Murai 1995, Hassan et. all. 2001).

On the other hand, the behavior of bubbles in a liquid is fundamentally different from that of solid particles in a gas. Bubbles have essentially no mass compared to the surrounding liquid, but particle inertia dominates the surrounding gas. Therefore, bubble motion leads the fluid while particles lag and the relative velocity is generally positive for bubbles but negative for particles. In fact, there is an interaction between the bubble plume and its environment. Hence, measurements of the environment can be used to derive the entrainment flux to the plume. However, it is required to develop an expression for the relative velocity in order to study the response of bubbles to the surrounding liquid (Hassan 2003, 2006, 2011, Hassan and Tamer 2006, Stewart and Crowe 1992).

Flows including bubbles called “bubbly flow” frequently show a very complicated behavior due to a strong interaction between bubble and vortex motions, and such phenomena significantly govern the performance of various industrial and environmental systems and processes in which bubbles are used. These processes require to be improved further since their performances are strongly swayed by the characteristics of the bubble motion. In this case, we have, not only to grasp the rough flow pattern of the two phases formed by bubble injection or boiling, but also to quantitatively recognize the flow field in order to maintain or improve the performance of the system. If we could predict or accurately measure the detailed flow structure, we would be able to significantly improve the efficiency, stability, and controllability of the devices by using the obtained data.

Bubble plumes are observed in various engineering disciplines, e.g. in industrial, material, chemical, mechanical, civil, and environmental applications such as chemical plants, nuclear power plants, naval engineering, the accumulation of surface slag in metal refining processes, the reduction of surfactants in chemical reactive processes, chemical reactions, waste treatment, gas mixing and resolution, heat and mass transfer, aeronautical and astronautical systems, biochemical reactors as well as distillation plants, etc (Hassan 2002, 2003, 2006, 2011, Hassan and Tamer 2006, Hassan, et. al. 2001, Murai et. al. 2001, Abdel Aal et. al. 1966, Goosens and Smith 1975, Al Tawell and Landau 1977, Chesters et. al. 1980, Bankovic et. al. 1984, Sun and Faeth 1986, Szekely et. al. 1988, Gross and Kuhlman 1992, Bulson 1968).

Flows induced by a bubble plume are utilized in many industrial processes. The main features of this kind of flow are:

(1) A large scale circulation of the liquid phase can be generated in natural circulation systems like lakes, agitation tanks, etc.

(2) Strong rising flows can be induced by the pumping effect as in air-lifting pumps.

(3) High speed surface flows may be developed at the free surface, by which the density and the transportation of the surface floating substances can be controlled.

(4) High turbulence energies can be produced in the two-phase region due to the strong local interaction between individual bubbles and the surrounding liquid flow (Hassan 2002, 2003, 2006, 2011, Hassan and Tamer 2006, Murai et. al. 2001).

Surface flows generated by bubble plumes are considered as a key phenomenon in many processes in bioreactors, chemical plants, modern industrial technologies, such as metal refinement, and future-type nuclear power plants, in addition to the many applications of bubble plumes mentioned above. These processes, which are expected to be improved by applying the bubble plume, require the control of both concentration and transportation of surface-floating substances, i.e. solidized materials or impurities, as well as the stabilization of the interface motion itself in order to guarantee their designed performances. Hence, the flow in the vicinity of a free surface induced by a bubble plume was utilized as an effective way to control surface floating substances on lakes, rivers, seas, oceans, as well as in various kinds of reactors and industrial processes handling a free surface (Hassan 2002, 2003, 2006, 2011, Hassan and Tamer 2006, Hassan, et. al. 2001, Sheng and Irons 1992, Iguchi et. al. 1991, Tomiyama et. al. 1994).

The main reason for surface flows induced by bubble plumes to be utilized in so many fields

mentioned above is the simplicity of installation. However, from the fluid mechanical point of view, it is important to focus on the advantages of the bubble plume comparing to various other types of jet flows. As an example, Fig. 1 shows schematic figures of three types of the jet flows:

- 1) Single-phase liquid jet flow (water jet flow).
- 2) Single-phase buoyant jet flow.
- 3) Bubble plume induced flow.

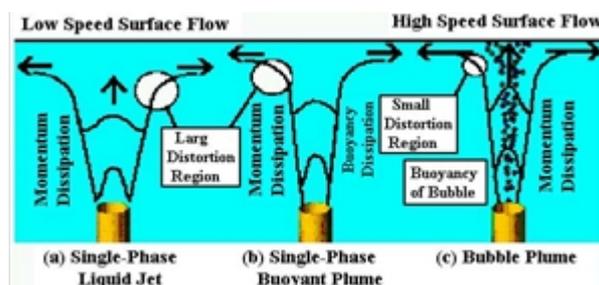


Fig. 1 Different surface flow generation mechanisms

1) In the case of the single-phase liquid jet flow (water jet flow), the maximum velocity in the jet decreases monotonically due to turbulent momentum diffusion, so that the width of the jet expands near the surface. The resultant thick layer surface flow will have a slow velocity.

2) In the case of a single-phase buoyant jet flow, the maximum velocity peak is flattened by the buoyancy effect. However, the buoyancy is dissipated due to turbulent mixing near the free surface so that no thin high speed surface flow occurs.

3) On the contrary, the bubble plume induced flow keeps its initial buoyancy even close to the surface since bubbles do not diffuse immediately (it is called immiscible buoyant jet). Furthermore, an individual bubble has motion characteristics different from surrounding liquid; therefore, the buoyancy distribution does not simply diffuse. The resultant surface flow is thinner and faster than in the first two cases. Hence, the surface flow is rather effective in case of the bubble plume compared to the first two flows because the distortion point occurs near the surface (Hassan 2002, 2003, 2006, 2011, Hassan and Tamer 2006, Hassan, et. al. 2001, Murai et. al. 2001).

Many researchers have carried out extensive model experiments by focusing on the flow field using air bubbles as gas injection through a bottom nozzle. This model is the most popular and has wide applications. Since bubble plumes have been used with varying degrees of success more information on the mentioned above subjects should be accumulated because there is still possible improvement to get higher efficiency for generating the surface flow (Hassan et. al. 2001, Gross, et. al. 1992, Sun, et. al. 1986, Hussain and Narang 1984, Hara, et. al. 1984, Chesters, et. al. 1980, McDougall 1978, Abdel-Aal, et. al. 1966, Hussain and Siegel 1976, Leitch and Baines 1989, Murai and Matsumoto 1998, Taylor 1955, Jones 1972, Murai, et. al. 2001).

This paper is concerned with the characteristics of bubble induce the surface flow (the detailed structure of surface flow generation mechanism), which depends on the gas flow rate, the bubble size, and the internal two-phase flow structure of the bubble plume. Particle Imaging Velocimetry (PIV) including Particle Tracking Velocimetry (PTV) and numerical simulation model using (E-L model) (Murai and Matsumoto 1996 and 1999) are applied after carrying out two kinds of flow visualization in order to evaluate their parameter relationships. The first one is visualization of the whole field around the bubble plume, and the second that of the different sections of bubble regions in order to clarify the relationship between bubble parameters (gas volume flow rate, mean bubble diameters and void fraction). It is confirmed by this paper that the experimental results resemble the numerical results. The relationship between the maximum surface flow velocity and the bubble generation condition in the surface flow generation process is explained.

Experimental Apparatus, Method and Conditions:

Experimental apparatus for carrying out the experiments of bubble parameters and the flow pattern around a bubble plume (whole field flow structure) is constructed as shown in figure 2. The inner tank size is 1300 mm in length, 1000 mm in height, and 110 mm wide, made of transparent acrylic resin. The experimental and simulation conditions are listed up in Table 1.

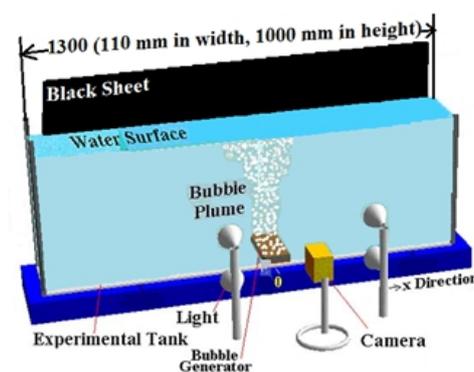


Fig. 2 Schematic diagram of experimental setup

Table 1: Experimental and simulation conditions

Parameter	Value
Density of water	$\rho=1000 \text{ kg/m}^3$
Kinematic viscosity of water	$\nu =10^{-6} \text{ m}^2/\text{s}$
Initial water height	$H=0.2 \sim 0.6 \text{ m}$
Atmospheric pressure	101 kPa
Temperature of environment	22-25 °C
Density of gas (air)	1.25 kg/m^3
Maximum gas flow rate	$8.0 \times 10^{-6} \text{ m}^3/\text{s}$

The bubble generator is installed at the center of the bottom part of the tank. Four kinds of bubble generators (with different injector nozzles) are applied for the experiments. Table 2 shows the experimental conditions for bubble generators. The gas flow rate is precisely controlled by a pressure regulator and a flowmeter. A lighting setup (direct lightning method) with a black back sheet background and two halogen lamps of 1000 W is used to clearly visualize and take the images of the experiments and for the PIV measurements. The visualized flows are recorded by a digital video camera (Panasonic DMC-GH2H) that captures 30 fps. The digital images are preprocessed through the video to JPEG converter image software and Adobe After Effects CS6 image processing software. The preprocessing entails sharpening, binarizing and smoothing of the images, and labeling bubbles and particles.

Table 2: Experimental conditions for bubble generators

	Bubble Generator-1	Bubble Generator-2	Bubble Generator-3	Bubble Generator-4
Type	Block	Straight	Round Shape	Single Nozzle
Number of Nozzles	14	16	39	1
Nozzle Diameter	0.75 mm	1 mm	1 mm	5 mm
A (mm ²) Injector	65×20	120×1	150×45	19.63
Air Tube/ Room	80×55×25	6.5 mm	6.5 mm	6.5 mm

Bubble Parameters Calculations:

Table 3 shows the calculations of mean (average) bubble diameter for the four types of bubble generators and for water height of 200, 400 and 600 mm respectively and for six ranges of gas volume flow rates [$Q_{g,1}$ to $Q_{g,6}$]. Therefore, seventy two cases (conditions) are handled in these experiments. The values in this table are calculated by using the time average of 180 consecutive frames in the image processing (6 seconds). The averaged bubble diameter, the void fraction, the width of bubble injection and the standard deviation are calculated by measuring more than 1200 bubbles in the local VTR images inside the bubble plume using image processing. These images are taken by recording local pictures of different regions; the injector region of the bubble generator, the middle reigns and near the free surface. Figure 3 shows samples of bubble images for different condition of bubble generators. The bubble diameter is defined by the equivalent bubble diameter using ellipsoidal approximations for the bubble shapes. The equivalent bubble diameter is estimated by the vertical and

the horizontal lengths of each bubble of the image object, which are obtained by using the JPEG converter image software and Adobe After Effects CS6 image processing software after binarizing the images. The measurement uncertainty (standard deviation) for the bubble diameter is estimated to be around (0.01~0.015 mm) of the averaged bubble diameter according to the pixel resolution. This reflects the accuracy of the experimental results.

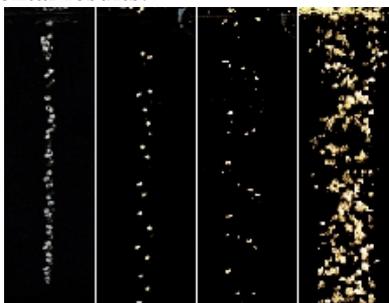


Fig. 3 Samples of Bubble Images for Different Condition of Bubble Generators

Table 3: Mean bubble diameter for water height H=200, H=400 and H=600 mm

Gas volume flow rate Q_g [m ³ /s]	Mean bubble diameter D (mm) for water height H=200, 400, 600 mm											
	Bubble Generator-1			Bubble Generator-2			Bubble Generator-3			Bubble Generator-4		
	H=200	H=400	H=600	H=200	H=400	H=600	H=200	H=400	H=600	H=200	H=400	H=600
$Q_{g1}=0.56 \times 10^{-6}$	2.8	6.01	6.87	3.0	6.89	7.80	3.80	7.00	8.90	6.20	8.10	9.54
$Q_{g2}=1.11 \times 10^{-6}$	3.0	6.30	7.54	3.4	7.17	8.10	4.02	7.55	9.20	6.70	8.70	9.91
$Q_{g3}=1.67 \times 10^{-6}$	3.3	6.70	8.06	3.6	7.84	8.90	4.95	8.25	9.80	6.95	9.60	10.3
$Q_{g4}=2.22 \times 10^{-6}$	3.6	7.15	8.41	3.9	8.09	9.20	5.90	8.55	10.1	7.40	10.01	10.7
$Q_{g5}=3.33 \times 10^{-6}$	4.0	7.60	9.01	4.1	8.72	9.50	6.80	8.95	10.6	7.90	10.80	11.2
$Q_{g6}=7.00 \times 10^{-6}$	6.5	10.80	12.30	7.1	11.80	13.5	9.12	12.70	15.6	10.80	15.89	17.8

Figure 4 shows the relationship between the mean (average) bubble diameter and the gas volume flow rates for the four types of bubble generators for water height of 200, 400 and 600 mm respectively. These figures confirm that as the gas volume flow rates increases the mean (average) bubble diameter increases.

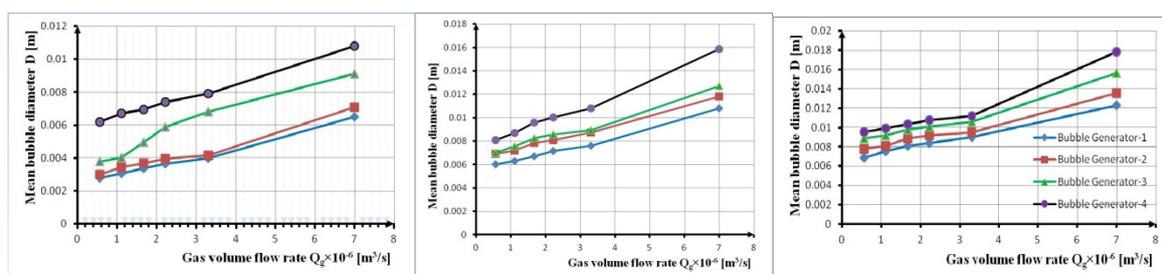


Fig. 4 The relationship between gas volume flow rate and mean bubble diameter for H=200, H=400 and H=600 mm respectively

On the other hand, the void fraction (α) is calculated by using the equation ($\alpha=Q_g/A \times V_b$) (Hassan et. al. 2001, Hassan 2002, 2003, 2006, 2012, Isao et. al 1993, Murai et. al. 1998, 2001 and Matsumoto et. al. 1995-B), where A is the area of calculation in the injector region, (injector surface of the bubble generator) as shown in table 4. The bubble rising velocity V_b is about 0.20 to 0.30 m/s. The bubble rise velocity is unsteady at the beginning (at the nozzle exit) and after a short period of time it reaches the terminal rise velocity. The measurement uncertainty for the rising bubble velocity is estimated to be about 2%. The relative velocity between the bubbles and the liquid flow corresponds well to the terminal rising velocity of the bubble in a quiescent liquid.

Table 4: Void fraction for bubble generator type

Gas volume flow rate Q_g [m ³ /s]	Void fraction α			
	Bubble Generator-1	Bubble Generator-2	Bubble Generator-3	Bubble Generator-4
$Q_{g1}=0.56 \times 10^{-6}$	0.0022	0.0009	0.0004	0.1425
$Q_{g2}=1.11 \times 10^{-6}$	0.0039	0.0017	0.0007	0.2567
$Q_{g3}=1.67 \times 10^{-6}$	0.0053	0.0023	0.0010	0.3539
$Q_{g4}=2.22 \times 10^{-6}$	0.0065	0.0028	0.0012	0.4350
$Q_{g5}=3.33 \times 10^{-6}$	0.0091	0.0039	0.0018	0.6050
$Q_{g6}=7.00 \times 10^{-6}$	0.0179	0.0078	0.0035	1.1881

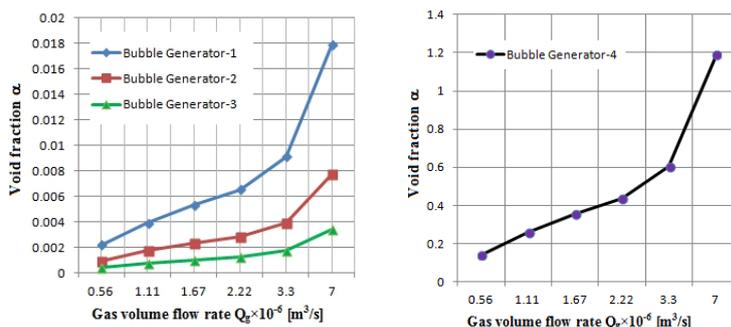


Fig. 5 The relationship between gas volume flow rate and void fraction α

Figure 5 shows the relationship between the void friction of the four types of bubble generators and the gas volume flow rates. These figures confirm that as the gas volume flow rates increases the void friction increases. Moreover, the void friction increases with the gas flow rate at a power index of around 0.8 to 1.0 [$\alpha \propto Q_g^{(0.8-1.0)}$]. The measurement uncertainty for the void fraction is estimated to be about 2 to 3%.

Figures 6 and 7 show the relationship between mean bubble diameter and water height in the tank for the four bubble generator types and for the six gas volume flow rate. It is clear from these figures that the bubble size increases as the water height in the tank increase.

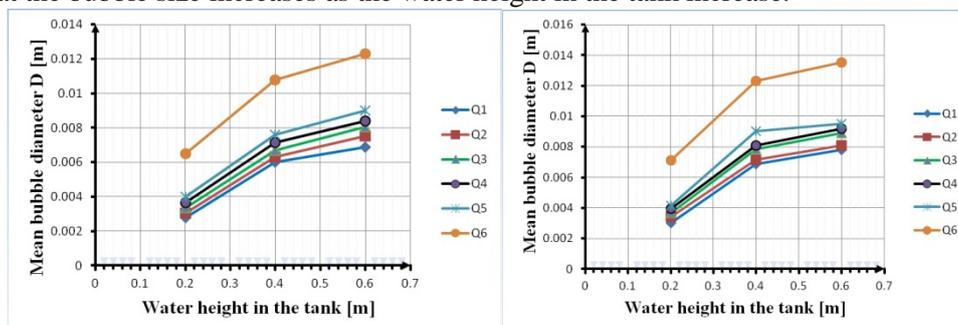


Fig. 6 The relationship between mean bubble diameter and water height in the tank for bubble generator -1 and bubble generator -2 and for the six gas volume flow rate

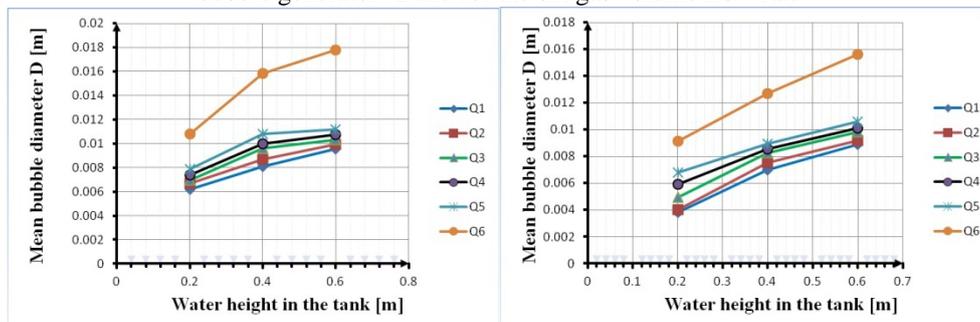


Fig. 7 The relationship between mean bubble diameter and water height in the tank for bubble generator -3 and bubble generator -4 and for the six gas volume flow rate

Flow Pattern around a Bubble Plume (Whole Field Flow Structure):

In order to clarify the flow pattern of the internal liquid flow (the whole field flow structure around the bubble plume) in water tank, spherical particles made of a high-porous polymer with diameters of 200 to 600 μm and a density of 1010 kg/m^3 , are seeded in the entire tank as tracer particles for the PIV measurement and then the flow is visualized. The recorded images are ported to a computer. The flows are measured in this section by using the PIV technique measurements for [case-1]: bubble generator-2 and gas flow rate value of $Q_{g5}=3.33\times 10^{-6}$, [case-2]: bubble generator -1 $Q_{g6}=7.00\times 10^{-6}$. The height of water (H) in the tank is 600 mm.

Fig. 8 shows a sample the recorded image of the seeded flow field around a bubble plume for the case-1 and case-2. In these images the bubble plume is located in the middle of the images, while the particles are distributed around the bubble plume. After these images are ported to a computer, velocity vector maps are obtained by using the BDCC (Brightness Distribution Cross-Correlation) method (Adrian et. al. 1991 and Kimura et. al. 1986) as shown in Fig. 9.

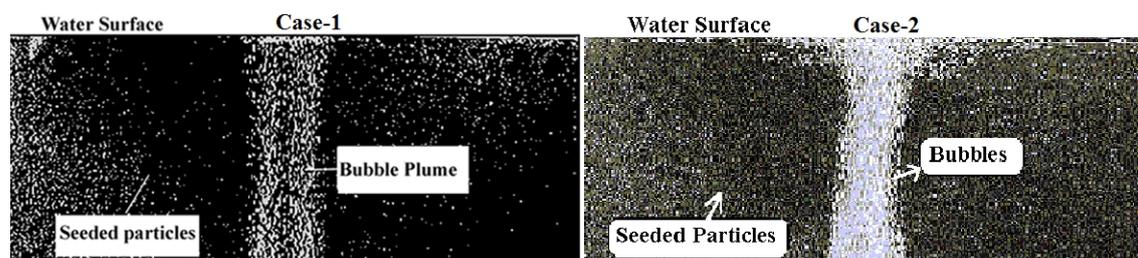


Fig. 8 Seeded image for PIV measurements (the image of the global motion) for [case-1] and [case-2]

It follows from these figures that the resulting flow is steady and symmetric relative to the bubble plume center (especially when a small gas flow rate is given). The flow pattern depends on the gas flow rate, and as the gas flow rate increases, the magnitude of velocity (the mean velocities of liquid phase and gas phase) increases and the effective area of the bubble plume (of the surface flow) expands in horizontal direction. The detailed flow mechanism can be explained as follows. The main upward liquid flow is driven along the bubble plume by the rising bubbles. Then the main flow reaches up to the free surface. The momentum of the upward flow becomes maximum near the liquid surface. Just under the free surface the upward flow rapidly changes its orientation into a horizontal flow. Then, a pair of liquid circulations is generated beside the bubble plume. By this symmetrical circulation, the main vertical flow is sharply converted to a strong surface flow. After same time, the pair of liquid circulations induces a whole scale circulation of the liquid over the entire tank. Also, since it induces an entrainment flow at the bottom region, total flow rate of main flow increases and more effective surface flow is generated.

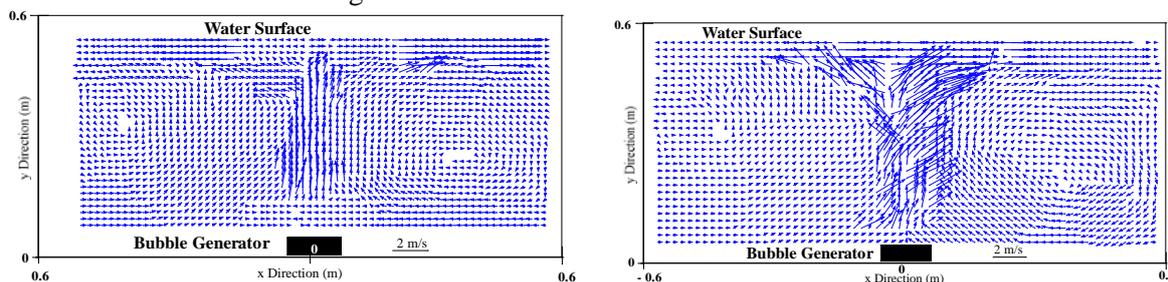


Fig. 9 Velocity vector map for [case-1] and [case-2]

The numerically simulation predicted results (which are done for comparing their results with PIV results) are obtained by solving seven governing equations (Matsumoto and Murai 1995, Murai and Matsumoto 1996, Hassan 2003, 2007). These calculations are for [case-1]. The simulation period for numerical predictions is about 0 to 120 seconds. The numerically predicted results shown in Fig. 10 indicates a time-development of whole-field flow structure (liquid velocity vector and bubble distributions) generated by the bubble plume. Although at the initial stage ($T < 2$) only a vertical liquid flow is induced along the bubble plume, the flow region expands widely and a strong surface flow is

stably maintained and well-developed stage. Also, the predicted local flow near the bubble plume is in good agreement with the PIV measured velocity map.

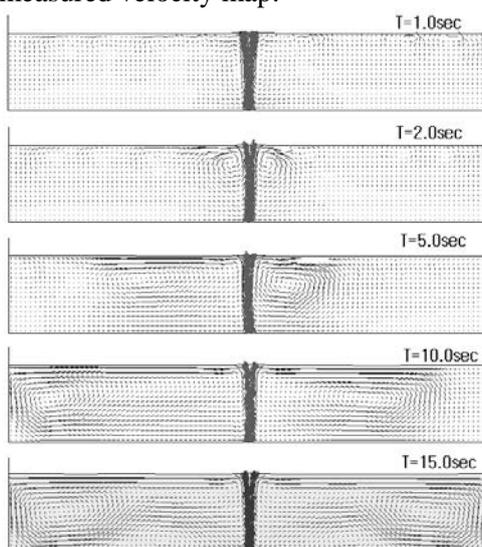


Fig. 10 Global field flow structure (liquid velocity vector and bubble distribution) for [case-1]

Fig. 11 shows samples of the whole field structure instantaneous streamlines during periodic swaying motion of the bubble plume (for the same cases “case-1 and case-2”), which are calculated by time-integrating of the measured velocity vector maps after calculating the stream function. The following two matters can clearly be recognized in these figures.

(1) The main upward flow converges and accelerates along the bubble plume. This phenomenon is completely different from a single-phase jet structure.

(2) The flow near the liquid surface is much larger than that in other regions.

Table 5 presents the relationship between gas flow rate and the width area of the bubble plume on the surface (the area which contains bubbles on the free surface “the width of the surface flow” in the horizontal direction), which is measured from video images. It is confirmed by this table that this area increases approximately proportional to the square root of the gas flow rate.

Table 5: The relationship between gas flow rate and area width of the bubble plume on the free surface

Gas flow rate (m ³ /s)	Area width (m)
$Q_g = 2.22 \times 10^{-6}$	0.225
$Q_g = 3.33 \times 10^{-6}$	0.415
$Q_g = 7.00 \times 10^{-6}$	0.530

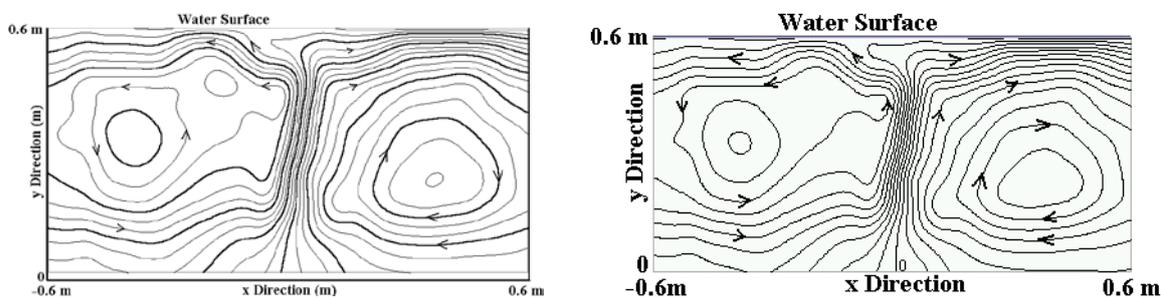


Fig. 11 Streamlines obtained by PIV for [case-1] and [case-2]

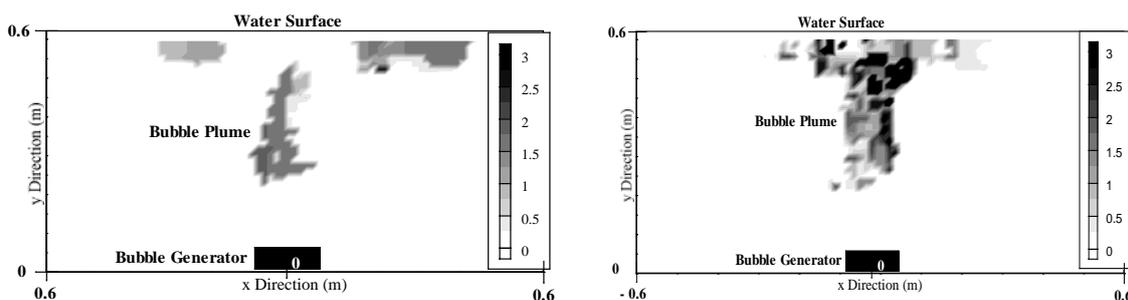


Fig. 12 Kinetic energy distribution (m/s^2) for [case-1] and [case-2]

Moreover, in order to clarify the detailed structure of the flow, the two-dimensional distribution components of the specific kinetic energy and the vorticity are calculated. These values are calculated from the measured velocity vector map, where the velocity inside the bubble plume contains two different phases, a liquid phase and a gas phase. Fig. 12 shows the kinetic energy distribution for the same cases “case-1 and case-2”, while Fig. 13 shows the vorticity distribution for the same cases.

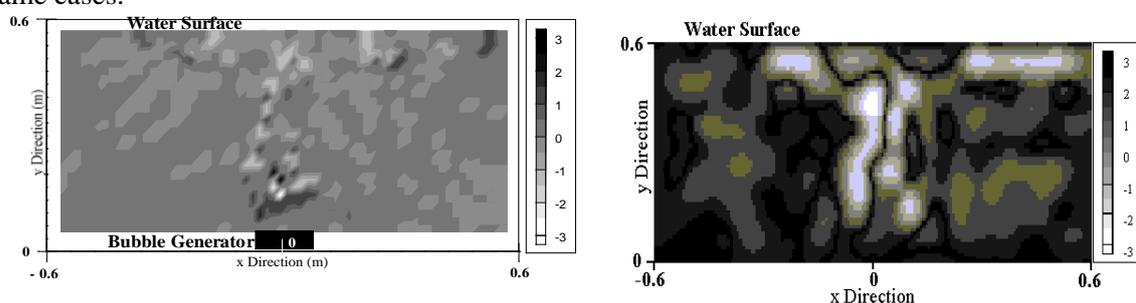


Fig. 13 Vorticity distribution (s^{-1}) for [case-1] and [case-2]

It is clear that the highest kinetic energy is generated far up (in the center of) the bubble plume and in the vicinity of the free surface. This observation confirms the idea that the bubble plume can generate a strong and wide surface flow over the bubble generation system. Beyond that, the figures of the vorticity distribution show that the rapid distortion or the rotation flow occurs in the area under the free surface where the upward flow changes its orientation from vertical into horizontal direction.

As our future works, the similarity between laboratory model and an actual application scale should be analyzed as well as three-dimensional interferences. Moreover, the relationship between mean bubble diameter and water height in the tank should be addressed for larger range of water height in the tank in order to find the ideal parameters for real application.

Conclusion

As a result, it is clear that the bubble plume “which is a typical bubble flow” is a key phenomenon to be studied and investigated, and it is an effective tool for many applications and can indeed contribute to various improvements. The motivation to study the bubble plume is the demands to improve its performance and its applications. Especially in many engineering fields as materials, chemical, mechanical, modern industrial technologies, and in the environment in order to protect the natural environment, the naval planets, navel systems, rivers, lakes, etc from pollution. Hence these engineering fields are expected to benefit from the improvement and development of the bubble plume.

Flow in the vicinity of a free surface, which is induced by a bubble plume is utilized as an effective way to control surface floating substances on lakes and oceans, as well as in various kinds of processing and industries handling free surfaces.

Surface flows are more effectively generated by bubble plumes compared to liquid jet flows because the distortion point appears in the vicinity of the surface.

Flow visualization, image analysis using PIV and numerical simulation of the bubble plume are carried out in order to improve the applicability of the bubble plume. The flow pattern for the whole field flow structure of the bubble plume is demonstrated. The parameters of bubbles are

calculated. The experimental results and the numerical results show good analogy. The flow structure is sensitively modulated by the gas flow rate and bubble size, and the main results can be summarized as follows:

- 1) There are two large circulation flow regions of liquid near the bubble plume (at the right and the left side of the bubble plume).
- 2) The local liquid flow pattern around the bubble plume depends on the gas flow rate. It is recognized that as the gas flow rate increases, the magnitude of velocity increases and the effective area of the bubble plume (the width of the surface flow) expands in the horizontal direction.
- 3) Inside the bubble plume and near the free surface, the velocity of the two-phase flow is higher while it is slower in other regions. Hence, high speed two-phase flow is maintained and further accelerated along the vertical axis, and it produces large entrainment flow in the lower region. This is due to the effect of the buoyancy of bubbles. Hence, the generation of this high speed flow is considered a main contribution to induce a strong surface flow. If a vertically rising liquid jet is applied to induce the surface flow instead of the bubble plume, the high speed upward flow is not maintained near the free surface due to the turbulent momentum dissipation and the lack of the buoyancy inside the jet, and in this case the power efficiency is considerably less due to the dissipation of momentum under the free surface.
- 4) The spacing of the streamlines becomes smallest at the free surface. Moreover, near the free surface, the liquid flow in the horizontal direction is maintained over long distances. Hence, the maximum velocity in the horizontal direction is observed to be near the free surface. This means that the horizontal velocity is fastest on the free surface since there is no shear stress acting on the free surface. This is also one of the reasons why such a wide and thin surface flow is generated by the bubble plume.
- 5) The highest kinetic energy is generated at a long distance (far up) inside the bubble plume and in the vicinity of the free surface. This observation confirms the fact that the bubble plume can indeed generate a strong and wide surface flow over the bubble generation system.
- 6) High vorticity distribution is generated by the surface flow, which induced by the bubble plume, and these phenomena appear in a layer under the free surface. In this layer the liquid flow rapidly changes its orientation from the vertical to the horizontal direction. Therefore, it can be said that the initial surface flow is rapidly generated in this layer. Therefore, the rapid distortion of the liquid phase results just under the free surface. This is qualitatively different from (not found so clearly in) the case of a single-phase liquid jet flow whose speed is equivalent to the bubble plume. These results indicate that a surface flow is more effectively generated by means of bubbles than by a liquid jet flow because the distortion point appears in the vicinity of surface.

References:

- Abdel-Aal H. K., Stiles G. B. and Holland C. D. 1966. Formation of Interfacial Area at High Rates Gas Flow Through Submerged Orifices. *AIChE J.* 12, pp. 174-180.
- Adrian, R. J. 1991. Particle-Imaging Techniques for Experimental Fluid Mechanics. *Annu. Rev. Fluid Mech.* Vol. 23. Pp. 261-304.
- Al Tawell A. M. and Landau J. 1977. Turbulence Modulation in Two-Phase Jets. *Int. J. Multiphase Flow* 3, pp. 341-353.
- Bankovic A., Currie, I. G. and Martin W. W. 1984. Laser-Doppler Measurements of Bubble Plumes. *Phys. Fluids* 27, pp. 348-355.
- Bulson P.S. 1968. The Theory and Design of Bubble Breakwaters. Proc. 11th Conf. Coastal Engng, London. 995.
- Chesters A. K., Van Doorn M. and Goossens L. H. J. 1980. A General Model of Unconfined Bubble Plumes from an Extended Source. *Int. J. Multiphase Flow* 6, pp. 499-521.
- C. W. Stewart and C. T. Crowe 1992. Bubble Dispersion in Free Shear Flows. *Int. J. Multiphase Flow* Vol. 19. No. 3, pp. 501-507. 1992. Great Britain.
- Goossens L. H. J. and Smith J. M. 1975. The Hydrodynamics of Unconfined Bubble Columns for Mixing Lakes and Reservoirs. *Chem. Eng. Tech.* 47, 951. PP. 249-261.
- Gross R. W. and kuhlman J. M. 1992. Three-Component Velocity Measurements in a Turbulent Recirculating Bubble-Driven Liquid Flow. *Int. J. Multiphase Flow* 18, 413-421.

- Hara, S., Ikai, M., Namie, S. 1984. Fundamental Study on an Air Bubble Type of Oil Boom. *Trans. Ship-Making Society of Kansai-Japan*. Vol. 194.
- Hassan Abdulmouti, Yuichi Murai, Yasushi Ohno and Fujio Yamamoto 2001. Measurement of Bubble Plume Generated Surface Flow Using PIV. *Journal of the Visualization Society of Japan*. Vol. 21. No. 2. Pp. 31-37
- Hassan Abdulmouti 2002. The Flow Patterns in Two Immiscible Stratified Liquids Induced by Bubble Plume. *The International Journal of Fluid Dynamic*. Vol. 6. Article 1.
- Hassan Abdulmouti 2003. Visualization and Image Measurement of Flow Structures Induced by a Bubbly Plume. Ph. D. thesis. Fukui University.
- Hassan Abdulmouti. Bubbling Convection Patterns in Immiscible Two-phase Stratified Liquids. *International Journal of Heat Exchangers (IJHEX)*. Vol. VII. No. 1. Pp. 123-143. ISSN 1524-5608. June 2006.
- Hassan Abdulmouti and Tamer Mohamed Mansour. Bubbly Two-Phase Flow and Its Application. 10th International Congress on Liquid Atomization and Spray Systems (ICLASS-2006). Aug. 27-Sept. 1. Kyoto, Japan. 2006.
- Hassan Abdulmouti. Surface Flow Generation Mechanism Induced by Bubble Plume. *Yanbu Journal of Engineering and Science (YJES)*. Second issue. PS-M02-28 (50-67). 2011.
- Hussain N. A. and Siegel R. 1976. Liquid Jet Pumped by Rising Gas Bubbles, *J. Fluids Eng.*, March 8. Pp. 49-62.
- Hussain N. A. and Narang B. S. 1984. Simplified Analysis of Air-Bubble Plumes in Moderately Stratified Environments. *ASME. Journal of Heat Transfer*. Vol. 106, pp. 543-551.
- Iguchi M., Takehuchi H. and Morita Z. The Flow Field in Air-Water Vertical Bubbling Jets in a Cylindrical Vessel, *ISIJ International J.*, 31, 3. 1991, 246-253.
- Isao Kataoka, Akimi Serizawa and D. C. Besnard. 1993. Prediction of Turbulence Suppression and Turbulence Modeling in Bubbly Two-Phase Flow. *Nuclear Engineering and Design* 141. 1993. 145-158. North-Holland.
- Jones W. T. 1972. Air Barriers as Oil-Spill Containment Devices. *J. Soc. Pet. Engng.*. Pp.126-142.
- Kimura, I., Kimura, H., Takamori, T. 1986. Image Processing of Flow around a Circular Cylinder by Using Correlation Techniques. *Proc. 4th Int. Int. Symp. Flow Visualization, Hemisphere*. Pp. 221-226.
- Leitch, A. M., and Baines, W. D. 1989. Liquid Volume Flux in a Weak Bubble Plume, *J. Fluid Mech.*, 205. Pp. 77-98.
- Matsumoto Y., Murai Y.. (1995). The Numerical Simulation of Bubble Plume in Container Have Free Surface. *Trans. JSME B*, 61-588. 2818-2825.
- Matsumoto Y. and Murai Y. 1995. *Trans. Jpn. Soc. Mech. Eng.*, Vol.61, No. 588, B 1995, pp.2818-2825.
- Matsumoto, Y., Murai, Y. 1995-B. Numerical Simulation of Bubble Plume in a Tank with Free Surface. *Trans. Jpn. Soc. Mech. Eng.*. Vol. 61. No. 588. B. Pp (54-61).
- Matsumoto Y. and Prosperetti A. 1997. Edit., *Proc. ISAC'97 High Performance Computing on Multi-phase Flows*.
- McDougall T. J. 1978. Bubble Plumes In Stratified Environments. *Journal of Fluid Mechanics*, Vol. 85, 1978, pp. 655-672.
- Murai Yuichi, Ido Takehiro, Ishikawa Masa-aki Yamamoto Fujio. (1998). Development of the Post-Processing Method for PIV Measurement Result Using Computational Fluid Dynamics Procedure. *Transaction of the Japan Society of Mechanical Engineers* 64B-626. Pp.109-116.
- Murai Y., Matsumoto Y., Sou K., Yamamoto F.. (1998). The Numerical Simulation of Disorder Structure Formed by Bubbly Buoyant Force. *Trans. JSME B*, 64-626. 3257-3262.
- Murai Yuichi, Ohno Yasushi, Abdulmouti Hassan, Yamamoto Fujio 2001. Flow in the Vicinity of Free Surface Induced By a Bubble Plume. *JSME*. 067, 657, B.
- Murai Y. and Matsumoto Y. 1996. Numerical Simulation of Turbulent Bubble Plumes Using Eulerian- Lagrangian Bubbly Flow Model Equation. *Proc. Int. Symp. on Mathematical Modeling of Turbulent Flows* pp. 233-238.
- Murai Y. and Matsumoto Y. 1998. Numerical Analysis of Detailed Flow Structures of a Bubble Plume. *JSME International Journal, Series B*, Vol. 41, No. 3, 1998. Pp.568-575.

- Murai, Y., Matsumoto, Y. 1999. Eulerian Analysis of Bubbly Two-Phase Flows Using CIP Scheme. *Computational Fluid Dynamics Journal*. Vol. 8 No. 1 April.
- Sheng, Y.Y., and Irons, G.A.: Measurements of the Internal Structure of Gas-Liquid Plumes, *Metallurgical Trans.*, Vol. 23-B. 1992, 779-788.
- Sun T. Y. and Faeth G. M. 1986 a. Structure of Turbulent Bubbly Jets-I. Methods and Centerline properties. *Int. J. Multiphase Flow* 12, pp. 99-114.
- Szekely, J., Carlson, G. and Helle L. 1988. *Ladell Metallurgy*. Springer. Berlin.
- Taylor Sir Geoffery. The action of a Surface Current Used as a Breakwater, *Proc. Royal Society, A.*, Vol. 231, 1955, p. 466-478.
- Tomiyama, A., Uegomori, S., Minagawa, H., Fukuda, T. and Sakaguchi, T. Numerical Analysis of Bubble-Induced Natural Circulation based on Multidimensional Two-Fluid Model, *Trans. JSME (Jpn. Soc. Mech. Eng.)*, 60-580, B, 1994, 9-14.

DEVELOPMENT AND EVALUATION OF A DEFECT TRACKING MODEL FOR CLASSIFYING THE INSERTED DEFECT DATA

Prof. Torky sultan

Ayman E. Khedr

Information Systems Department Faculty of computers and Information,
Helwan University Cairo, Egypt

Mostafa sayed

Information Systems Department Council state Cairo, Egypt

Abstract:

Defect tracking systems play an important role in the software development organizations as they can store historical information about defects. There are many research in defect tracking models and systems to enhance their capabilities to be more specifically tracking. Furthermore, there are different studies in classifying bugs in a step by step method to have clear perception and applicable method in detecting such bugs. This paper shows a new proposed defect tracking model for the purpose of classifying the inserted defects reports in a step by step method for more enhancement of the software quality. Besides, an evaluation of experiment made for measuring the proposed factors results for defects classification.

Key Words: Bugs, Defects, Bug Tracking Systems, Defect Tracking Models, Software Quality

Introduction

In many software development organizations, bug tracking systems play an important role as they allow different types of users communicating with each other (i.e. developers; testers and customers) to assure that they have the same perception about problems or requesting new features. In addition, bug tracking systems can keep track of more historical information stored of the bugs. Earlier attempts were made for enhancing the defects tracking models and defect classifications [16]; [17]; [6]; [7]; [8] and [15]. But there were no general farm work for concentrating on different tracking phases. Besides, there were no interests with the insertion factors for classifying defects through those attempts. These issues make difficulty for retrieving accurate information from defects tracking tools in the future.

To address these issues, we have developed a general defects tracking model for classifying the inserted defect data in a step for more enhancing defects tracking model quality. In this paper we developed the model based on the previous works of [7] and [15]. Also we did evaluate the proposed model using an experiment. The model was helpful in describing and explaining different phases of defect tracking model, as well as the insertion factors for classifying defects.

This paper provides new proposed defects tracking model concentrating on the factors for the insertion of defects reports through tracking tools. In addition, it provides a theoretical overview of the literature on defects tracking systems and overview of different aspects of it and their components (section 2). The existing attempts to improve the defects tracking systems are highlighted in our synthesizing framework (section 3). Also, the paper provides a conceptual framework design for the proposed defects tracking model (section 4). The paper ends with an experiment for evaluating the proposed model (section 5), followed by section summary of (section 6).

Overview

This section aims at providing a detailed discussion of the background overviews about defects tracking systems.

There are many software tools that play an important role in tracking defects of software and which are called "Defects Tracking Systems". Jalbert defined them as "Allow users to report, describe, track, classify and comment on bugs reports and feature requests" [14].

Defects Tracking Systems can be separated systems that can integrate, and contribute in software development process. They can keep, with details of defects reports and information associated with resolving it, in a database storage. Lethbridge, Singer and Forward indicated that developers view the defects tracking systems as important repositories of historical information [20]. Furthermore, software defects data is an important source to the organizations for the software process improvement decisions and that "ignoring defects data, can lead to serious consequences for an organizational business" [11]. In addition "they may be part of an integrated suite of configuration management tools, where the status of the defect may act as a trigger or key for other events within the system" [1].

There is no doubt that software quality which is used in detecting defects, is one of the important factors for evaluating the software process development. Weinberg (1983) documented that an error costing a company 1.6 billion dollars, and was the result of changing a single character in a line of code [31]. Also, Curhan mentioned that "some types of defects have a much higher costs to fix due to the customer impact and the time needed to fix them, or the wide distribution of the software in which they are embedded " [6].

Moreover, large number of software companies use Software Tracking Tools to achieve the goals of the Configuration Management. Janák defined configuration management as "the process of controlling and documenting changes to a developing system" [15]. Also, software tracking tools help quality control engineers to accomplish their jobs as good as possible to discover, and prevent the occurrence of bugs by tracking them.

The Software Tracking Tools are simply built based on defects tracking models. Edwards and Steinke (2006) simply discussed the defects tracking model, as they divided it into the following two stages: ((repair /resolution)-(verification)) and the following three changes of status: (discovery – resolved – closed) [7].

Microsoft Team Systems used a four-stage defects tracking model for Capability Maturity Model Integration (CMMI); the model expanded and evolved the "open" stage into the following two stages: "proposed" and "active" stage. Although the model enhanced the three- stage defects tracking model, it still works as a framework describing the status and phases of bugs that should followed. The three statuses (deferred – rejected – duplicate) duplicated through two positions, the proposed stage and the active stage [22]. There were no remarks about how to examine and register the bugs.

Edwards et al. (2006), proposed the Full Product Life Cycle Defect (FPLC) Model, which was an extension of IBM/Rational Model with changes to include the test and project management interfaces. The model discussed in details, the five statuses of the defects tracking model which are: Submitted, Open, Postponed, Resolved and Closed. Although the model mentioned perfectly the duplication problem of defects; it still has some remarkable scope for more enhancements.

The research dealt with the status "reject" as not a closed status. It coped with it as a circulating process where it should be a "Closed" status. Also another remarkable note about the postponed defect, Edwards et al. (2006), reported that "Placing any defect in a Postponed status is a tacit admission that it should be repaired, but at a later time." [7] which means that it has the priority to be repaired not to be a closed.

One of the famous defects tracking tools used by quality control engineers is Bugzilla defects tracking system. The work flow of the model showed that it classified the new bugs into the following two categories: the first one comes from a user with a confirmation right, and the second comes from any user but it will not be confirmed till it has enough votes. Also, it concentrated on quality control engineer roles in checking the appropriate solution which being satisfied, verified, closed or didn't conform with the solution [15].

Although the default IBM Rational Clear Quest Ticket mentioned the workflow path that defect process has taken, and which "Starts when the defect is discovered and ends when the defect is resolved, hopefully repaired, for the most immediate release of the software application" [15]. It still has some shortcomings as the "rejected" status could be at any state. It may be after investigation, the approved state or after the task opened and in all the cases, it should be closed. Also the approved

status should be one of the roles of quality control engineer; who should check it as the defect may not exist only in a new project process, but also may exist at the maintenance process.

Synthesis Model for the Classification of the Bugs

The last section discussed the different overviews of the defects tracking systems. Their workflow models, the status and paths of the defects through the process of discovering the defect. Also, it mentioned the literature reviews of different research at the same point that dealing with defects in their overall aspects. The proposed model concerned with the following two points: First the different classifications of the bugs; and the second is the different phases of tracking the defects. The two points will be discussed in more details in the next paragraphs.

This section covers discussing the proposed model and how it makes filtration and classification of the bugs. However, a bug in its default way, is discovered where an action or value is not achieved as it decided or going in an unexpected way. We divided the classification and track of the bugs into the following Phases: (Submission – Examination – Registration – Tracking) which interact with each other and will be discussed in more details later.

The Submission Phase

The first phase of the proposed model will help us in understanding the classification process of the bugs and the issues in the submission phase.

The role of this model in tracking bugs, starts after discovering an incorrect action or value to the system. Describing and stating the problem is a significant component for retrieving a suitable solution. Stimson (1998) mentioned that "A problem well-stated is half-solved" [28], this push us to define of who discovered the bug, and where it is discovered.

There are two different groups of users who can discover the presence bugs. The first Group is: "The Normal User" who deals with the system after released to him to achieve a specific function or goal. The second Group is: "The Authorized User" who participates at any phase of the development process. He may be one of triage team or development engineering team. The bug is usually discovered in two positions: the first position is through the development process. The second position is after releasing the product to users.

Section (2), discussed a number of different defects tracking models; where there were a number of these models which coped with "the submission phase" as the first step of filtration and classification defect reports. The adapted one with our model was "Bugzilla tickets workflow". We will modify it to be more compatible with the proposed model.

Bugzilla Workflow Model, classified the detectors of the bugs into two categories:-

- 1- Anyone who has enough votes.
- 2- A user with confirmation rights.

According to the classification of users in Bugzilla Workflow Model, we will classify the users into three categories:-

- 1- Authorized user with confirmation rights.
- 2- Trusted User.
- 3- Normal User.

The first Category: (Authorized User) who is discovering the bugs inside the location of the development process. The authorized user may be one of the quality control engineers i.e. developing engineer or may be anyone who has the ability to discover a bug. The second category: (Trusted User), who can be defined as the user who has the ability, and good experience for dealing with the product or system; also has a recorded history of detecting bugs. The third category: (Normal User) who has the ability but little experience for dealing with the product or system, and has short recorded history of detecting bugs. That is, the Normal User has the ability to inform the presence of a bug but hasn't the priority element without a confirmation of an "Authorized User".

The Examination Phase

The examination phase begins after the end of the submission phase. In this phase, the outside or inside user who participates in the development process decides that there is a vision for a defect.

This phase has its own priority as Hooimeijer and Weimer (2007) documented that "Bug report triage and evaluation are the significant part of modern software engineering for many large projects" [12]. It is the first phase of preventing the distortion of recording un-wanted data or

duplication announcement of bugs by checking the database for recorded bugs before and through each time of registering a newly discovered bug.

According to the work and efforts made by Mays, Jones, Holloway and Studinski, at IBM 1990 for defects prevention. They analyzed the faults that appeared in order to understand them using casual analysis. In addition, detecting the way of prevents defects from appearing in the future. They showed the role and importance of the action team whose responsibility was to detect and store the appeared faults in the database and make a checklist to be updated with the new faults. Also the important role of "triage team", mentioned by Black (1999), who assured that the triage team can review, evaluate the defects and assigning them to the development team [3]. For more information in details about triage team see [21].

When the Bug examination process is done, it is followed by rules and strategy of checking tests through quality control engineers. Furthermore, bug examination, is the last phase of deciding whether either the bug was recorded before with a suitable solution, or it will be a new classified bug.

The last statement lead us into the following three states after having the bug's examination recorded history such as:-

- 1- Bug not found and not registered before.
- 2- Bug found with the same condition and need to be in (reopened state).

The first point will be discussed in more details, in the next section as it will be the default path. This point was achieved by following different test case scenarios, and confirmation that there was a bug with the same conditions registered before. Therefore, a "reopen state" can be released by an authorized user in the examination phase in order to prevent duplication defect reports.

The Registration Phase

The registration phase follows the examination phase. It is an important phase for retrieving useful information in the future because there are different factors for classifying defects reports which were registered in this phase. The next paragraphs, will discuss different classification of defects schemes.

Although there are large number of research on classification of defects schemes, they faced a number of problems including ambiguous, and confusion of error causes [23].

One of the first works for classifying defects, was made by Endres in 1973 of IBM. He classified the defects into six general categories including: Machine Error, User Error, and Documentation Error. Also, he classified each defect by 'type'. But this type of classification scheme was very complex [8].

According to Basili, and Perricone's categories, the error classified as one of the following Categories: Requirements incorrect, Functional Specification misinterpreted, a design error which spans several modules, an implementation error in a single module, misunderstanding of the external environment, error in the use of compiler, clerical error, and error due to previous wrong correction of an error [2].

Another work was made by Sullivan and Chillarege (1992) to analyze the different error classifications. They made their work based on defects reported at customer sites in two large IBM database management products, DB2 and IMS. They compared the errors type; defects type and error triggers classifications [30].

Fredericks and Basili (1998) made analysis to find defects and how organizations dealt with it. They focused on achieving three goals that can be defined as significance factors of building a new defect tracking models. These goals are: Detecting the Nature of Defects; Detecting the Location of Defects, and when the Defects are Inserted.

In the early 1990's, IBM developed two new Technologies using defects data. The First Technology: "Defect Prevention", which involves development teams contributing to a knowledge database containing: common defects, how they can be prevented, and how to easily detect them. The Second Technology: "Orthogonal Defect Classification", which involves using statistical methods to predict the phase in which a defect originated, rather than relying on the subjective evaluation of an engineer [9].

The Defect Tracking Model had evolved at the end of the nineties; the Defects Classification Scheme mentioned three Elements of defects Categorization. The First Element was the Location where bug is discovered through the development process. The Second Element was the type of

defect where may be classified through each phase of the development process has its own kind of defects. The Third Element was the value of each defect which can be measured [25].

According to Rus (2002), there is a defect classifying schema that was developed and used by IBM called "Orthogonal Defect Classification". He defined it as "A measurement concept for software development that uses the defect stream as a source of information on the product and the development process "[26]. He divided it into two classes of defect attributes: the First Class associated with the defect discovery and contained the elements: (activity, trigger, and impact). The Second Class associated with the removal of the defect and contained the elements: (target, defect type, qualifier, source, and age).

With the last different views of the classification schemes of defects, it appeared that there were a number of factors that describe the defects, and these factors are so important. We will concentrate on the following two Factors that are seen on the degree of importance from quality control perspective.

1- The First Element is "Bug Location":

The locations of the bugs are determined by, in which stage the bug appeared or discovered; and in which place in the system or the application it appeared. Fry and Weimer (2010) defined fault localization as: "Is the task of determining if a program or code fragment contains a defect, and if so, locating exactly where that defect resides" [10].

Furthermore, another element for describing the location of bugs is to describe where the bug was discovered through the system. Also we have to describe the surrounding environment of the system as an element in classifying the location of the bugs such the version of the system, the kind of operating system that the system works under.

2- The Second Element is "Bug Type":

The 'Bug Type' varies from one system to another because the different tools which were used to create such systems, have their own limitations and shortcomings according to the study made by Ko et al., [17]. However, we have to put a dynamic framework for defining the bug type according to the various tools used to build the systems; with respect to the major general bug type.

Sullivan et al., 1992 Classified the Software Defects Type as: function defect, data structure/algorithm defect, assignment/checking defect, interface defect, timing/synchronization defect, and build/package/merge defect [29]. We proposed a general defect types which are Interface defect, calculation defect, loading defect, security defect, documentation defect, enhancement defect and business logic defect that may appear in any system.

The Tracking Phase

This phase is concerned with the traceability of the defects that were registered in the system before. There were a number of scenarios expected from the proposed model to achieve. One of these scenarios is the traditional scenario which begins with a new classified defect on the status "Initial".

At this stage, the development engineer who is responsible for fixing such defect starts to work under the status: "under development" with a new date recorded of the beginning of the development process. After the development process of the defect is finished, its status changes to "development completed" with respect to recording date of finishing this process as bugs fixing is a time-consuming process [12]. With small calculations of dates between "Under Development" and "Development Completed", we can measure how much time it took the development team to fix this defect. After the status "Development Completed" is finished, a Regression Test Phase is going to achieve. When finishing all test cases and scenarios for the defects, the quality control engineers release the status "Test Complete" then the status "Closed" for finishing the scenario.

The last scenario showed different statuses which defect moves through it, concerning the time element that recoded before and recording every change on defect status, as mentioned by Tammana and Faught (1998) "A defect tracking system that lets the user query the defect database is useful not only to generate summary reports, but also to track the status and the progress of a project that's underway" [30]. Therefore, through the power of DBMS (data base management system), we can achieve a powerful tracking of defects through this last scenario.

The Following Table, Abbreviated the different status of tracking statuses that the defects take through different phases of the product development phases and whose responsibility to check.

	Status	Meaning	Responsibility
1	Initial	Declaration status for Fixing the Defect.	Quality Team
2	Under Development	Declaration status for starting Development on Defect.	Development Team
3	Development Complete	Declaration status for finishing development on defect	Development team leader
4	Under Test	Declaration status for starting test	Quality Team
5	Test Complete	Declaration status for finishing test	Quality Team Leader
6	User Acceptance Test Complete	Declaration status for user acceptance that the defect fixed	Users and Quality team leader
7	Closed	Declaration status for finishing all process needed for fixing defect	Project Manager
8	Need more Details	Declaration status for misunderstanding the requirement or function needed	Development Team
9	Postponed	Declaration status for postponing task for a time or next versions	Project Manager Development Team Leader
10	Refused	Declaration status for Refusing the Task for unacceptable requirement or function needed.	Development Team leader
11	Reopen	Declaration status for reopening a defect with a same condition of a recorded defect before	Quality team leader

Table 1: Proposed Defect status Scheme

Conceptual Framework Design for the Proposed Defects Tracking Model

Based on the previous discussions in sections 1, 2 and that derived from the development of the research synthesis model for different ways of defect classifications that were presented in the preceding section, the ultimate conceptual model is given in Figure (1). The present research adopts the following model for classifying bugs which appear through different development phases especially in the maintenance and testing phases. As mentioned by Boehm and Basili, the maintenance phase consumes over 70% of the total life cycle cost of the software development projects [5]. The model developed was based on the previous work of [7], [15] and on our general synthesized model for classifying and tracking defects (section 3 & 4). It is quite suitable for a case study. This model will guide us through our exploration for classification of defects to enhance quality control for the software development and maintenance processes.

Our model will focus, in details, on the phases of preventing defects and classifying them. Moreover, it concentrates on the bug examination, location and type factors for the insertion process of defect reports. Due to the highly exploratory nature of this research, all isolated conceptual variables/factors only represent the initial ideas about the discussion of defects tracking phases and classification method to deal with in the future.

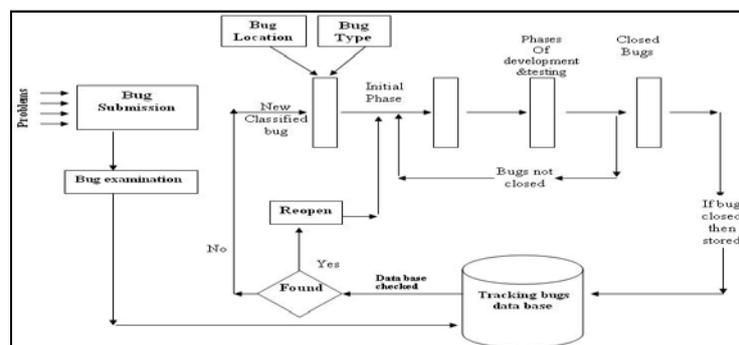


Figure 1: Proposed Defect Tracking Model

Evaluation

To evaluate the utility of the model, we performed a case study for locally made tracking tool used in an Egyptian software development company called (4s Systems). In addition, an experiment was made based on the data collected through the case study to achieve the purpose of evaluation. Also, we used in our research methodology observations, and formal interviews with a well prepared interviews protocol to face and analyze problems of defect tracking systems.

Experiment Descriptions

According to the conceptual framework of the proposed model, which built based on the discussed literature review, also the database gained and analyzed. We believe that experiments are suitable solutions for proving deductions. Different definitions of experiments and its advantages and disadvantages were cited in [4], [18] and [27]. However, according to the proposed model, we detected three factors (Bug Examination, Bug Location and Bug Type) which were different between this model and the others. An experiment was designed for measuring each factor separately and will be discussed in the next subsections.

Bug Examination Factor

We choose five random users from quality control team that have at least two years experience. Also, we choose a random sample consisting of 50 defect reports for the overall data of all applications which were registered in the database. Afterwards, well defined and classified 20 defect reports were well prepared for the experiment. We divided the sample into two parts: First Part contains new ten bug scenarios and the Second Part had ten bug scenarios duplicated from the 50 bugs in the database. We requested the users to select randomly, four bug scenarios from the 20 bugs to register them on the tracking tools.

With re-doing the last designed experiment with new factor of bug examination that represents in proposed screen; new trends appeared. The function of the new examination screen was to force the user to check the historical defect database at least ones per new record. Also, another function of the examination screen was to track the number for each user attempt to check the database before registering new record. This would help users in discovering the duplicated scenarios to be reopened (in reopen status). The results were Coordinated in the following Table 2.

Bug Type Factor

The experiment was designed using a random sample of defects reports that contain (50) bugs; then we classified these reports with the top managers of departments of the organization who have at least 20 years of experience.

The defect reports, which represents the data sample was classified according to the experience consulted top managers of department and in parallel with the researcher view to be a reference data after running the experiment.

All samples scenarios classified in a group with a serial number key to be easily detectable when participants were randomly taken. Also, we have discussed all bugs type and what does it mean to all participants before beginning classifying and registering them on the dummy database as a step on the designed experiment.

The Bug Location Factor

The objective of the bug location experiment, was to classify the location of each bug scenario in a numeric path to become more precise in detecting and describing the bugs where they occur. Also, another objective was to measure the ability of the participants to use a unified method in classifying the paths of defect instead of being written thought the text. Hence, achieving the last objectives will help in restricting the inserted data about classifying the defects; as the user has no choices to register a new defect report except from classifying the path of defects.

According to the problem which faced the research in detecting the place where the bugs took place, we proposed to describe the places into a numeric manner, and in a classification structure as screen, report etc. However, paths of all places through the application, were described in numeric path stored on relational database tables.

The experiment began with an empty database expected from the random sample consisting of (50) bugs that were previously classified. We designed a screen that would help the users who participate in the experiment, to classify the path of defects thought text of defects description scenario into controlled detectable choices. As the paths of bugs were described previously in text without a unified method or detectable choices (e.g. file /choose registration screen / new record).

Afterwards, we distributed the defect reports randomly and equally on the users after inserting the screen of application paths which was connected to the database. In addition, we requested them to detect the path of the bugs in the application after reading, and extracting them from the task description field. The path of the bugs usually described through the task description field in tracking tool used by the company.

Experiment Results

As mentioned in section 5.1 the model proposed three factors Bug Examination, Bug Location and Bug Type. The experiment results will be show according to each factor separately in the next subsections.

The Bug Examination Factor

According to the results showed in the Table 2, we focused only on two dependent values i.e.: Registered and Reopen State. Also, we compared the two results after and before the experimentation. Analyzing the results specially the registered state after running the experimentation compared with the tracking users attempts factor, showed that there was a change of the curves values in figure 2.

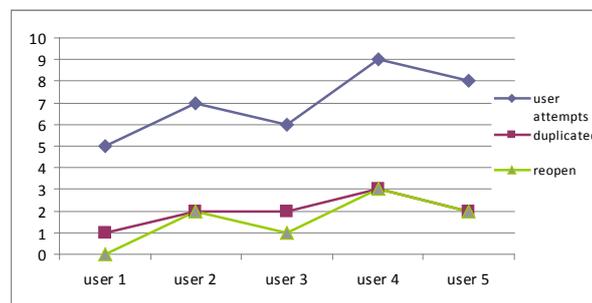


Figure 2: The Causal relationship between the Factors (User's attempts and reopen).

There was a positive relationship between the numbers of user's attempts to check for a registered defect with reopen factor corresponding to the duplicated factor of defects.

	Tracking Users Attempts	Duplicated Bugs	New Bugs	Registered	Reopen State
User (1)	5	1	3	4	0
User (2)	7	2	2	2	2
User (3)	6	2	2	3	1
User (4)	9	3	1	1	3
User (5)	8	2	2	2	2
Total		10	10	12	8

Table 2: Results of examination factor after the experiment

The Bug Type Factor

Bug type	Number of Standard Sample (x)	Total (T)	Means of Bug Type versus to Total Sample	After Registered (p)	Error Ratio (z)
Interface Error	5	50	10%	4	2%
Calculation Error	12	50	24%	10	4%
Loading Error	10	50	20%	10	0
Security Error	9	50	18%	9	0
Documentation Error	2	50	4%	2	0
Enhancement Error	4	50	8%	5	-2%
Business Logic Error	8	50	16%	10	-4%

Table 3: Results of Analysis of Bug Type Factor Data.

We used a well prepared standard sample data to each type of bug type factor which Referred to with factor (x). Also, we used simple statistical law $Z = ((x-p)/T) * 100$ for measuring the error ratio where the factor (Z) is the error ratio between x and p versus the total sample.

The negative error ratio percentage (-2% & -4%) were interpreted to the incorrect defect reports registered as wrong in the defect types enhancement and business logic error attributed to human error factor.

The Bug Location Factor

	Sample (x)	Total (T)	Mean	After Registered (p)	Error Ratio(z)
Screen	20	50	40%	20	0
Report	15	50	30%	15	0
Other	15	50	30%	15	0

Table 4: The Results of Bug Location Factor after the experiment

As shown in the last table, after using our proposed bug location factor of proposed model with the standard sample data, we achieved high percentage of success with zero error ratio percentage.

The results of error ratio with zero value confirmed the possibility of converting paths of errors that described in the reports of problems to be tracked digitally and easily recognizable by the human element used in the process of recording errors. Hence, we could uniformly unify the description process of different tracks within applications. This added a new element in the process of registering errors in addition to the previous factors the bug examination and bug type to facilitate future retrieval and tracking errors.

Conclusions

This Paper described terms and findings from significant earlier research, thereby forming a conceptual context and foundation for the exploratory observational study that was central to this research.

The following are the Four Main Findings drawn from the study:

1-The Bug Examination Factor is considered as the greatest effect on the process of tracking defects reports. As with increasing the efforts of examining the tracking history of database; it decreases the level of registered defect duplication.

2- The Bug type has significant effects on evaluating the tracking system as a method in structuring the inserted data to have structured correct information.

3- The Bug location has significant effects on detecting location of defects precisely through software development and enhancement.

4- The Quality Control field depending directly on human factor in triaging defects. This appears obvious through the Error Ratio in classifying defect type and bug location factor.

Considering the Limitations of the Conceptual Study Model, the following agenda for further research is proposed:

1-User's Authorization and Privileges through the submission phase of defect tracking models, have to be measured and enhanced.

2- Securing sensitive information for customers registered through the database tracking tools have to be researched.

References:

- [1](Avram, 2007) Gabriela Avram, Anne Sheehan and Daniel K. Sullivan, Defect Tracking Systems in Global Software Development – a work practice study, Limerick, Ireland, 2007.
- [2](Basili, 1984) Victor R. Basili, and Barry T. Perricone, Software Errors and Complexity: An Empirical Investigation, Communications of the ACM, P.42-52, 1984.
- [3](Black, 1999) Rex Black, Managing the Testing Process, Wiley, 1999.
- [4](Blaxter, 2006) Loraine Blaxter, Christina Hughes and Malcolm Tight , How to Research THIRD EDITION ,2006 , p135.

- [5] (Boehm, 2001) Barry Boehm and Victor R. Basili, Software Defect Reduction Top 10 List, p.135–137, 2001.
- [6] (Curhan, 2005) Lisa A. Curhan, Software Defect Tracking during New Product Development of Computer System, Massachusetts institute of technology, 2005.
- [7] (Edwards, 2006) Jim Nindel-Edwards and Gerhard Steinke, A Full Life Cycle Defect Process Model That Supports Defect Tracking, Software Product Cycles, And Test Iterations, Communications of the IIMA ,Volume 6 Issue 1, 2006.
- [8] (Endres, 1975) Albert Endres, an Analysis of Errors and Their Causes in System Programs, Proceedings: International Conference on Reliable Software, P327-336, 1975.
- [9] (Fredericks, 1998) Michael Fredericks and Victor Basili , A State-of-the-Art-Report for Using Defect Tracking and Analysis to Improve Software Quality , 1998.
- [10] (Fry, 2010) Zachary P. Fry and Westley Weimer, A Human Study of Fault Localization Accuracy, 2010.
- [11] (Grady, 1996) Robert B. Grady, Software Failure Analysis for High-Return Process Improvement Decisions, Hewlett-Packard Journal, 47(4), 1996.
- [12] (Hooimeijer, 2007) Pieter Hooimeijer and Westley Weimer, Modeling Bug Report Quality, In Automated Software Engineering, pages 34–43, 2007.
- [13] (IBM, 2005) IBM® Rational® ClearQuest® Deployment Kit, Usage Model Guidelines, Version 1.1, Jan., p. 4, 2005.
- [14] (Jalbert, 2008) Nicholas Jalbert and Westley Weimer, Automated Duplicate Detection for Bug Tracking Systems, IEEE, 2008.
- [15] (Janák, 2009) Jiří Janák, Issue Tracking Systems, Brno, spring, 2009.
- [16] (Just, 2008) Sascha Just, Rahul Premraj and Thomas Zimmermann, Towards the Next Generation of Bug Tracking Systems, 2008.
- [17] (KO, 2003) Andrew J. KO and Brad A. Myers, Development and Evaluation of a Model of Programming Errors, 2003.
- [18] (Kothari, 2004) C.R. Kothari, research methodology methods and techniques (second revised edition), 2004, p121.
- [19](Lenin, 2010) R.B. Lenin, R. B. Govindan and S. Ramaswamy, Predicting Bugs in Distributed Large Scale Software Systems Development, SCS M&S Magazine, 2010
- [20] (Lethbridge, 2003) Timothy C. Lethbridge , Janice Singer and Andrew Forward , How Software Engineers Use Documentation: The State of the Practice, IEEE Software Volume 20 (Issue 6), 2003.
- [21] (Mays, 1990) R.G. Mays, C. L. Jones, G. J. Holloway and D. P. Studinski, Experiences with Defect Prevention, IBM Systems Journal, 29 (1) (January):4-32,1990.
- [22] Process Guidance documentation provided in the documentation for Microsoft Visual Studio, 2012, <http://msdn.microsoft.com/en-us/library/ee332488.aspx>
- [23] (Ostrand, 1984) Thomas S. Ostrand, and Elaine Weyuker, Collecting and Categorizing Software Error Data in an Industrial Environment, the Journal of Systems and Software, 4:289-300, 1984.
- [24] (Ostrand, 2004) Thomas J. Ostrand and Elaine J. Weyuker, A Tool for Mining Defect-Tracking Systems to Predict Fault-Prone Files , 1st International Workshop on Mining Software Repositories, united kingdom, 2004 .
- [25] (Pfleeger, 1998) Pfleeger, and Shari Lawrence, software engineering theory and practice upper saddle river, NJ: prentice hall, 1998.
- [26] (Rus, 2002) Ioana Rus, Combining Process Simulation and Orthogonal Defect Classification for Improving Software Dependability, Chillarege Press, 2002.
- [27] (Sjøberg, 2005) D.I.K. Sjøberg, J.E. Hannay, O. Hansen, V.B. Kampenes, A. Karahasanovic, N.-K.Liborg, A.C. Rekdal, A survey of controlled experiments in software engineering, IEEE Transactions on Software Engineering 31 (9),2005, P 733–753.
- [28] (Stimson, 1998) Carl Stimson, the Quality Improvement Story, 1998.
- [29] (Sullivan, 1992) Mark Sullivan and Ram Chillarege , A comparison of Software Defects in Database Management Systems and Operating Systems ,IEEE,1992 .
- [30] (Tammana, 1998) Prathibha Tammana and Danny Faught, Software Defect Isolation, 1998.
- [31] (Weinberg, 1983) Gerald. M. Weinberg, Kill That Code, Info systems, 1983, P.49.

3GPP: LTE- AN INNOVATIVE TECHNOLOGY TOWARDS 4G WIRELESS NETWORKS

J Vijay Franklin

Lecturer, Department of Computer Science and Engineering,
Bannari Amman Institute of Technology, Sathyamangalam

Dr. K.Paramasivam

Professor, Department of Electronics and Communication Engineering,
Bannari Amman Institute of Technology, Sathyamanglam

Abstract:

3GPP: LTE –Long term evolution is baseline to the modern 4th generation communication technology, that is an enhanced version of Universal Mobile Telecommunication System (3G). LTE is introduced by 3GPP. LTE is standardized in the form of REL-8, this is operated in data rate 100 Mbps for downlink and 50 Mbps for uplink in the 20 MHz bandwidth. LTE can accommodate up to 200 users in cell (5 MHz). The objective of LTE is to reduce costs, providing high performance, enhances the features, improves the data transfer speed, and reduce resource utilization. The main advantages are it provides end –to-end Qos, Enhanced multimedia broadcast service, spectrum flexibility. It uses OFDM for modulation. In OFDM a large no of sub carriers are used to convey the data. Each and every sub carrier carries one data. Data will be split into no of data streams. And also it uses SC-FDMA, for single carrier modulation. For downlink LTE uses OFDMA with CP. MIMO (multiple-input Multiple-output) is an antenna technology used to improve the performance of communication. In this multiple antenna are used to send and receive multiple data streams. LTE uses a new resource scheduling algorithm called multi carrier channel dependent resource scheduling

Key Words: LTE, HSPA, OFDM, Qos, Handoff

Introduction

Modern wireless technologies have grown in tremendous way. LTE is a new technology that has been developed by 3GPP 2 LTE focus on enhancement in UMTS system. It optimizes the architecture of 3G system. LTE operates 15 to

100 times faster than current 3GPP system. LTE supports both frequency division duplex (FDD) and time division duplex (TDD) modes, allowing operators to address all available spectrum types. Nowadays people are using HSPA (high speed packet access) technology for send and receiving data through internet. LTE will replace HSPA by providing the following benefits,

- *Simplicity*-supports flexible carrier bandwidth
- *Capacity*-provides key features for uplink and downlink
- *Performance*-operates in very high speed eg:160 Mbps by Ericsson
- *Wide range of terminals*- operates in heterogeneous terminals

It allows the existing HSPA to create UMTS with high data rates, low latency and wider spectrum. LTE managing multimode devices that is compatible with existing 3G technologies. Cost efficiency plays an important role in the success of LTE in 4G technology evolution. This paper describes the architecture, Qos, Security, Handovers and advantages of LTE advanced system.

LTE Architecture

LTE architecture is designed for improving the performance, reducing the cost and delivering the services in more efficient way. Fig shows the Service Architecture Evolution (SAE) of LTE. It has two nodes, eNodeB (evolved Node B) as base station of LTE, and AGW (SAE Gateway). All base stations are interconnected with IP based network AGWs are connected with eNodeB by using IP-based interfaces. LTE system standardizes the existing interfaces of HSPA for the interconnection of networks.

The existing architecture of WCDMA, CDMA2000, HSPA, and SGSN are integrated for developing the design of LTE. This helps the system to achieve the handovers in various networks. The architecture has several modules, each and every module have specific functionalities. LTE uses retransmission technology for regulating the packets in eNodeB. For this base station have control plane and buffer for high speed data retransmission S1 interface is used to connect eNodeB and Evolved packet core, X2 is used to interconnect eNodeBs. eNodeBs is providing compression and encryption methodologies to the users. It is responsible for Routing.

Protocol Architecture

Fig.2 shows the protocol structure of LTE, Physical layer is used to offers the information to higher channels by using transport channels. Physical layer is also called as Layer 1 of the protocol stack.

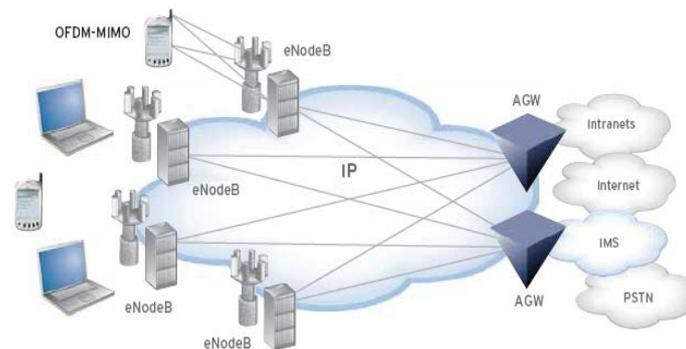


Fig1: Architecture of 3GPP: LTE

Transport channels describe characteristics of the data is to transmit. In layer 1 uses different channels for uplink and down link those are,

Down Link Transport channels

- *Down Link Shared Channel (DL-SCH)* - Shared among UEs with in cell.
- *Broadcast Channel (BCH)* - broadcasts the information to UE
- *Paging channel(PCH)* - Supports for Broad cast
- *Multicast Channel(MCH)* -Broad cast to multiple cells

Uplink transport channels

- *Uplink shared channel(USCH)*- shared between all UEs within a cell
- *Random access channel(RACH)* - access to UTRAN

Layer 2

Layer 2 of the protocol stack is divided into three sub layers *Medium Access layer(MAC)*, *Radio link control(RLC)* and packet data convergence protocol it describes what to be transferred by using the transport channels of Layer 1

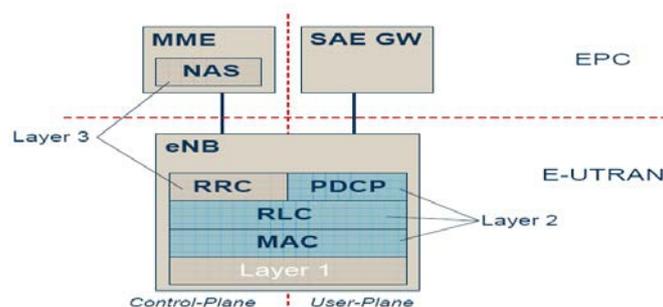


Fig2: Protocol Architecture

These logical channels are divided into control channels and traffic channels. Control channels transfers the control plane information. Traffic channels transfers the user plane information

Control channels

- Broad cast control channels(BCCH)-Broad casting control information
- *Common Control Channel(CCCH)*-transfers control information between UEs
- *Dedicated Control Channel(DCCH)*-Point to point channel transfers data between UEs and network
- *Multicast Control Channel (MCCH)*-point to multi point channel transfer MBNS information.

Traffic channels

- *Dedicated Traffic Channel (DTCH)*-dedicated to one UE for uplink and down link
- *Multicast Traffic channel(MTCH)*-one to many channel transfers traffic information to entire network

Radio Resource control

RRC is layer3 of LTE protocol stack, which is used to takes care of control signals and handovers between eNodeBs in the network, it is responsible for paging, intera-cell handovers, Qos, Broadcasting, security and reporting & control information. And also it takes care of the establishment of radio barriers.

Orthogonal frequency division multiple access (OFDM)/ Downlink

In LTE system uses OFDM for downlink transmission. OFDM splits the entire system is divided into multiple data streams. Each stream is modulated independently by low rate data stream. OFDM provides flexibility, spectrum efficiency and throughput for data streams. It allows the user to access the provided bandwidth. TDD is used by user to access the bandwidth .data is allocated in terms of resource blocks. Each block consists of 24 sub carriers in frequency domain. Total number of blocks available is depends upon the bandwidth. Size of the block is same for all blocks. it uses FFT for converting the signals into data streams. Allocation of blocks will be performed in scheduling process that is taken place in eNodeB. In down link operation the following are the major tasks.

- Scheduling
- Symbolization & Cell search
- Link adaptation
- Hybrid ARQ

The first task is scheduling the time and resources are allocated in a downlink format.

Adaptive modulation and coding technique is used for link adaptation. Hybrid ARQ is retransmission protocol; it retransmits the incorrectly received data packets.

Single carrier Frequency division Multiple Access (SC-FDMA)/ UPLINK

For uplink Data transmission. In OFDM Technology it uses multiple data streams, it does not require for uplink the data. So that it uses SC-FDMA scheme for uplink User equipment needs the flexible and power efficient methodology. SC-FDMA provides these constrains in effective manner.

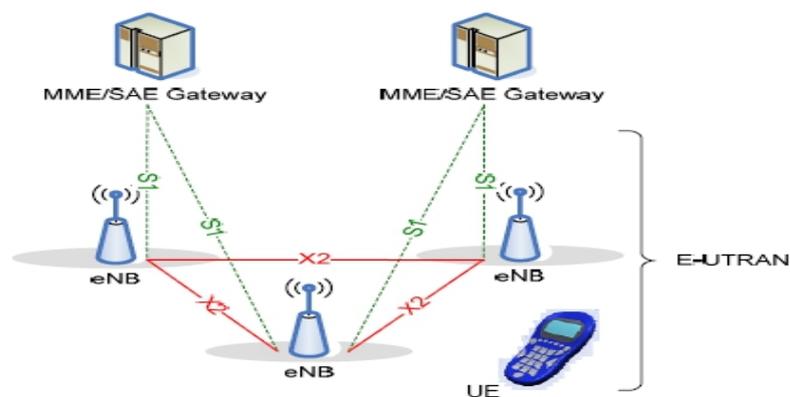


Fig3: Interfaces in LTE Source: [16]

It has the features of OFDM and PAPR (Peak –to-Average Power Ratio) .SC-FDMA uses DFT for converting the signals. It uses the data sub carriers of OFDM scheme it coverts the binary input to modulated sub carriers. In OFDM it uses FFT but here SC-FDMA uses DFT.

Multiple Input Multiple Output (MIMO)

The communication link between the sender and receiver is made-up of MIMO system. MIMO is innovative antenna technology that is used in LTE. In this each antenna uses same spectrum for transmission. At the time of transmission the data stream is divided among multiple antennas, these antennas improve the data rate of transmission. It allows multiple UEs can transmit at same time. Each eNodeB consist of two antennas for sending and two antennas for receiving the data. MIMO is used to achieve throughput and spectral efficiency. The capacity of wireless environment is improved by the combination of MIMO and OFDM. It enhances the spectral efficiency. it maximizes the spectrum of usage. The magnitude of data streams for uplink and downlink will be increased. The entire data stream is mapped with the multiple antennas of MIMO. MIMO performs the beam forming operation, and enhance the air interfaces for transmission.

Quality of service

LTE provides end to end Qos. Qos is defined as facility to provide the guarantee of the performance of data flow for heterogeneous users. Qos ensures the service quality for each user. This will be depends upon the requests made by users. In the LTE system all Mobile broadband services are performed by packet switched connections. LTE Qos has two parameters one is for delay sensitive packets from the user equipment and another is for monitoring the packet lose ratio.

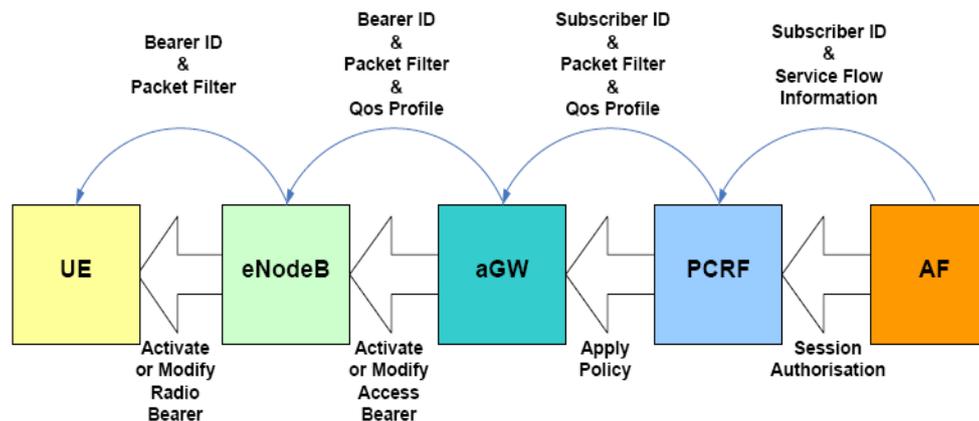


Fig 4: QoS Architecture Source: [2]

Fig 4 shows QOS architecture for LTE in this Application Function (AF) initiates the session authorization activity which has information about Subscriber ID (SID) and service flow. By using this information policy charging and rules function generates QoS profile and packet filtering function. Access Gate Way (aGW) is used to perform the activation, modification of access barrier operation and applies the end to end Qos policy. Enhanced Node B (eNodeB) is a base station (as per standards of 3GPP) transfers the profiles of quality policy and barrier ID for user Equipment (UE), in the first phase AF gives the details need to frame the Quality policy. PCRF frames the QP and applies the policy in the aGW. QP assures the following

- Guaranteed delivery
- End to end Qos
- Applied to all traffic

Handover in LTE

Hand over mechanisms in LTE are defined in 3GPP TS 23.401 standard. The user equipment is moving one place to another place the UE has change its eNodeB, aGw, and MME. The LTE interface architecture consist of s1 and X2 interfaces, based on these two the handovers or classified into two categories

- S1 based handovers: For change in eNodeB, aGw and MME
- X2 based handovers: For change in SGW

Hand over mechanism of LTE has several control transfer operations. The Source eNodeB will have the

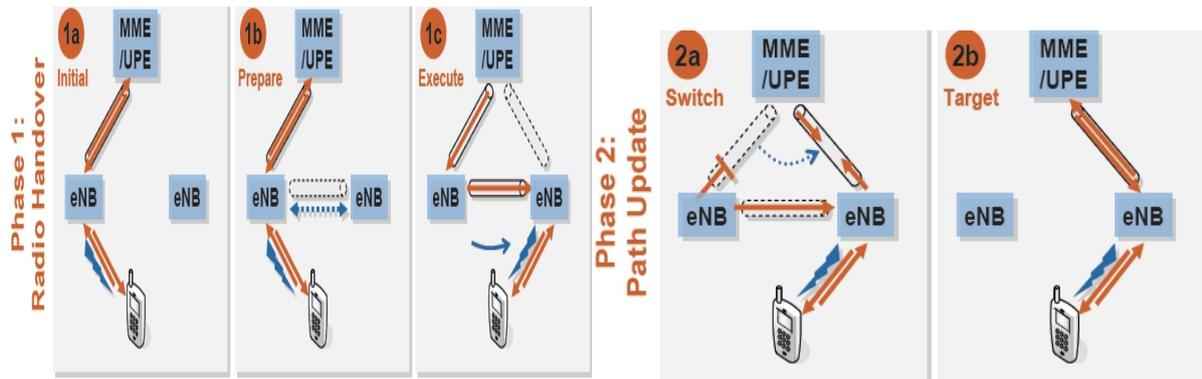


Fig 5: Hand over in LTE Source: [2]

Collection of measurement reports and its sends the handover request to the target eNodeB. The target eNodeB replies the ACK message to the source. Downlink allocation and handover command will be issued to UE by source eNodeB. The UE will get detach from the Source eNodeB and perform the delivery of SN status and data. This data will be stored in the buffer of target eNodeB the signal request. Synchronization and handover conformation will be performed as per the path switch and its response of user plane control of SGW.

Security in LTE

LTE uses specific security functions for data transmission; it concentrated on signaling protection, user plane protection, network domain security, authentication and key agreement. The same UMTS AKA is reused in LTE; it provides confidentiality and integrity for data transmission. LTEs security mainly deploys on the subscribers authentication and traffic protection. The enhanced version of 3G AKA protocol is used to provide subscriber authentication between UE and MME. Access management entity (ASME)^[9] is used to provide the protection for architecture of LTE over traffic.

The authentication mechanism consist of for major operations those are eNodeB authentication, Integrity verification, device authentication and device binding authentication can be provided by either operator of the network or mutual authentication by eNodeB or authentication by certificates. The third is performed by IKEv2 protocol ASME provides the keys for authentication and as well as encryption of data for MME and encryption schemes for the interfaces in LTE system.

In this system the keys (K_ASME) are derived from the ASME and to allocate to MME; the keys derived from K_ASME are send to UE. LTE provides end –to- end security to the interfaces. This system uses crypto algorithms for security, it uses 128-EEA1 and 128-EIA1 and also AES is used for encryption and decryption in UE. LTE has plenty of other security mechanisms like Device integrity check, location locking, access control mechanism, clock synchronization The following figure shows integrity protection optimum cost.

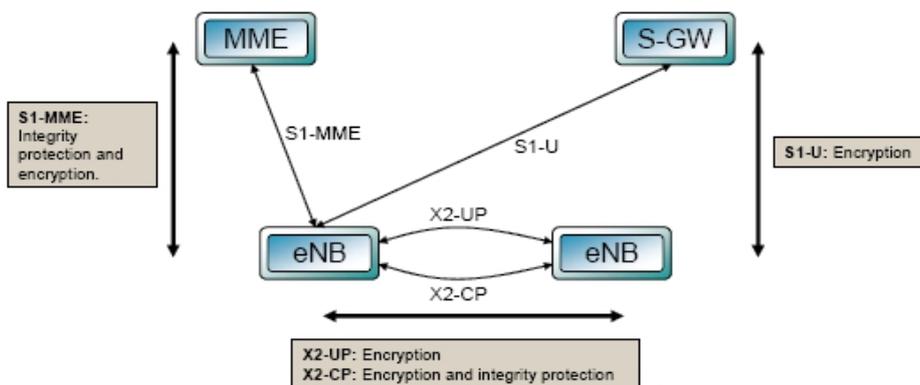


Fig 6: Encryption in LTE Architecture Source: [4]

Conclusion

LTE 3GPP: is designed to provide an incredible performance with high data rates, high security, low latency and high flexibility. And also it provides end to end QoS, Handovers with spectral efficiency and cost. This paper describes the introduction to the methodologies and the behavior of 3GPP: LTE system. The tremendous features of LTE will give higher end technology with the and UE.

The architecture of LTE has the compatibility to accommodate future enhancements like LTE – Advanced. IMS towards the Rel-9, Rel-10 of 4G. The access methodology of this LTE supports instantaneous radio interfaces; it uses new technologies like MIMO, it improves the power consumption, reduced interferences and increased quality of signal. LTE provides advanced security mechanisms for data transmission and hand over. The security functions of LTE will provide a secured architecture for business environments. LTE architecture has established a baseline to future 4G wireless networks.

Glossary:

HSPA	High Speed Packet access
LTE	Long Term Evolution
OFDM	Orthogonal FDMA
MIMO	Multiple input, multiple outputs
MME	Mobile Management Entity
PAPR	Peak to Average Ratio
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
SC-FDMA	Single Carrier FDMA
3GPP	3G partnership program
aGW	access Gateway
QoS	Quality of service

References:

- [1] 3GPP TS Series <http://www.3gpp.org/>
- [2] Dr.Ing.Michel schop, Siemens Networks, "Trends in Mobile Network Architectures Wimax 3GPP LTE", White paper Nov-2006.
- [3] Dr.Jayesh,Kotetecha,Jason Wang,"LTE MIMO Techniques in 3GPP-LTE,NOV-08
- [4].Dr.Erik Dahlman, Ericsson Research,"3G long term Evaluation", Nov-2005
- [5] Erik Dahlman, Hannes Ekström, Anders Furuskär, Ylva Jading, Jonas Karlsson, Magnus Lundevall, Stefan Parkvall, Ericsson Research" The 3G Long-Term Evolution – Radio Interface Concepts and Performance Evaluation"2006
- [6] Erik Dahlman, Anders Furuskär, Ylva Jading, Magnus Lindström and Stefan Parkvall," Key features of the LTE radio interface" Ericsson Review No. 2, 2008
- [7] Adrian Scrase," Examining the 3GPP Release 8 Standards" LTE World Summit, Berlin, May 18-20, 2009
- [8] Stephen Hayes," *Exploring 3GPP Rel9*", LTE World Summit, May 19, 2009
- [9] Charles Brookson," 3GPP Security:LTE/SAE and Home (e)NB" LTE Summit - Berlin, Germany - 18-20 May 2009.
- [10] Hyung G.Myung "Technical Overview of 3GPP LTE" May, 2008
- [11] Amit Mukhopadhyay, Ph. D," LTE/EPC – 4G Wireless", October 2008.
- [12] M I Barkway,"LTE: What, Where, and When", Jan 2009,
- [13] Stefan Parkvall," LTE – Long-Term Evolution" Jan 2008
- [14] Nortel" Long Term Evolution (LTE): The vision beyond, 3 G"White paper.
- [15] Ericsson"HSPA, LTE and Beyond"Feb 2009.
- [16] Valterri Niemi,"3GPP Security Hot topics:LTE/SAE Comman Topics.Jan 2008.

IMPACT OF AN ACTION CYBER CRIME PREVENTION PROGRAMME ON IN-SCHOOL AGED CHILDREN'S ATTITUDE TO CRIME PREVENTION CONCEPTS IN CIVIC EDUCATION AND SOCIAL STUDIES

Peter Adewale Amosun

University of Ibadan, Ibadan

Olugbenga Adedayo Ige

Adekunle Ajasin University, Akungba Akoko

Abstract:

This study evaluated the impact of a participatory cyber crime prevention programme on in-school aged children's attitude to cyber crime prevention in Nigeria, consequent on the global annual ranking of Nigeria as one of the most cyber lawless countries by The National White Collar Crime Center and Federal Bureau of Investigation in the United States of America. The study used the participatory action research paradigm to develop an informal cyber crime prevention programme, and further utilized pretest-posttest quasi experimental design to evaluate the impact of the action cyber crime prevention programme on secondary school students' attitude in Civic Education and Social Studies. The sample of the study comprised two hundred and eighteen students purposefully selected from intact classes in six secondary schools in Nigeria. Percentages, Mean, Standard Deviation and Analysis of Covariance (ANCOVA) were used to analyze the data. Multiple Classification Analysis aspect of ANCOVA was used to determine the magnitude of performances across the groups, while line graph was used to disentangle significant the interaction effects. Results showed that there is significant main effect of treatment on in-school aged children's attitude to cyber crime prevention in civic education and social studies. Consequent on these, the study recommended the use of Action Cyber Crime Prevention Programme for developing countries experiencing the menace of cyber crimes perpetrated by school-age children.

Key Words: Impact, Action Cyber Crime Prevention Programme, In-School Aged Children, Learning Outcomes, Social Studies, Civic Education

Introduction

The new Social Studies and Civic Education curriculum that was put together in year 2007 reflect depth, appropriateness and interrelatedness of the contents of the curriculum. Emerging issues which covered value reorientation, peace and dialogue including human rights education, family life HIV/AIDS education, entrepreneurial skills, crime prevention themes and Information and Communication Technology (ICT) were infused into the relevant contents of the new 9-year basic education curriculum (NERDC, 2007). Adesina and Adeyemi (2007) stated that the field of Social Studies has come to stay in Nigerian primary and secondary schools, as such the importance attached to its study cannot be exaggerated. It is believed to be the correct drug that can be used to heal the ailment of moral decadence (Cyber Crime) and instill in the youth a sense of decent behaviour (Cyber Crime Prevention). To this end, the objectives of Social Studies specifically are to enable the children develop an understanding of their immediate surroundings; develop certain skills which will enable the children to deal with and manage the forces of the world in which they live; and learn how to live harmoniously where many different groups co-exist (Obidoa, 1991). Adesina and Adeyemi (2007) further clarify that Social Studies can be seen as a programme of study in Nigerian secondary schools which is used to inculcate in learners the knowledge, skills, attitude and actions considered important in human relationship in society. These are reasons why Action Cyber Crime Prevention Programme will best be taught through Social Studies and Civic Education. Comparative Education Study and Adaptation Centre (1993) defines Social Studies as a subject concerned with the way man lives in and

interacts with his social and physical environments and how science and technology, for instance, computer, help him to live well in those environments.

In addition to these, the Internet Crime Reports of year 2001 to 2010 published by the National White Collar Crime Center and the Federal Bureau of Investigation in the United States of America show the extent to which Nigerians perpetrate cyber crimes. Table 1 show the top ten countries in the world whose citizens are enmeshed in cyber crimes from year 2001 to 2010.

Table 1: Global Cyber Crime Ratings for Year 2001 to 2010.

Countries	Year 2001	Year 2002	Year 2003	Year 2004	Year 2005	Year 2006	Year 2007	Year 2008	Year 2009	Year 2010
United States	3.7%	4.0%	3.7%	3.7%	3.7%	3.7%	3.7%	3.7%	3.7%	3.7%
Nigeria	2.7%	5.1%	2.9%	2.87%	7.9%	5.9%	5.7%	7.5%	8.0%	5.8%
Canada	1.2%	1.2%	1.2%	1.2%	1.2%	1.2%	1.2%	1.2%	1.2%	1.2%
Kenya	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%
United Kingdom	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%	0.8%
South Africa	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
Australia	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
Indonesia	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
France	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
Spain	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
Sweden	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
Belgium	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
Germany	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
Italy	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
Japan	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
South Korea	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
India	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
China	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
Malaysia	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
Philippines	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%

Source: ICS 2001 – 2009 Internet Crime Reports (January 1, 2001 to December 31, 2010) prepared by the National White Collar Crime Center and the Federal Bureau of Investigation.

* The annual ranking of Nigeria in percentage (%).

Table 1 show that Nigeria was second in year 2001 among the top ten countries noted for cyber crimes with 2.7%. In year 2002, Nigeria retained the second position with 5.1% while Nigeria was third with 2.9% in year 2003. In year 2004, Nigeria still maintained the third position with 2.87% even though this year recorded the lowest perpetration of internet crimes involving Nigerians. In year 2005 Nigeria moved to second position with 7.9%, and third repeatedly in years 2006, 2007, 2008 and 2009 with 5.9%, 5.7%, 7.5% and 8.0% respectively. In year 2010, Nigeria retains the third position with 5.8%, Table 1 shows a downward trend in the perpetration of scamming activities by Nigerians in 2010. This decline may not be unconnected with austere conditions caused by the global economic meltdown. Internet fraudsters could only thrive when the economic conditions of the target (mark) countries are in good condition. However, it should be noted that the Internet Crime Reports only shows the ten most cyber lawless countries and leaves out the remaining countries below the global top ten mark.

For educational researchers, the need to protect children from the dangers posed by cyber crimes has become imperative. It is most unfortunate that over ninety percent (90%) of 8 to 16years old in the category mentioned earlier viewed pornography online while doing home work (Chiemeke and Longe, 2007). Although empirical data about cyber crimes and its impact on the life of the teeming youth and children are not readily available, fears are already very rife in some sectors about the possible negative consequences obnoxious and uncensored Internet contents will have on the psycho-social wellbeing of varying categories of users, especially school-age children in Nigeria.

The Internet is both a source of prospect for school-age children in Nigeria and a source of concern. The prospect is that the Internet offers such an enormous range of positive educational experiences and materials. Yet, children online may not only be vulnerable to harm through exposure to sexually explicit materials, adult predators and peddlers of hate, but also perpetrate any of these crimes while on the Internet. The technical nature of the Internet has not evolved in a way that makes control over content easy to achieve (Dick and Herbert, 2003). It has become imperative for educational researchers to evolve protective schemes like Action Cyber Crime Prevention Programme for Internet savvy children and teenagers in a country where the struggle for survival keeps parents at work while having little or no time to monitor what their children do with the Internet at home, in their schools and other Internet access points.

Theoretical Framework

This study is based on the constructivist theory of learning which provides the theoretical base for the participatory action research paradigm adopted for this study, space transition theory, which is a criminological theory that take cognizance of the milieu in which cyber crime is committed, Lombroso theory of atavism, which has been described as the most notorious of the evolution approaches to deviant behaviour and Routine Activity Theory, formerly applied to aggregate level crimes has been considered appropriate for cybercrimes .

The concept of constructivism has roots in classical antiquity, going back to Socrates's dialogues with his followers, in which he asked directed questions that led his students to realize for themselves the weaknesses in their thinking. The Socratic dialogue is still an important tool in the way constructivist educators assess their students' learning and plan new learning experiences. Jean Piaget and John Dewey developed theories of childhood development and education, what is now called 'Progressive Education' that led to the evolution of constructivism. Piaget believed that humans learn through the construction of one logical structure after another. He also concluded that the logic of children and their modes of thinking are initially entirely different from those of adults. The implications of this theory and how he applied them have shaped the foundation for constructivist education. Dewey called for education to be grounded in real experience. He wrote, "If you have doubts about how learning happens, engage in sustained inquiry: study, ponder, consider alternative possibilities and arrive at your belief grounded in evidence." To Dewey, inquiry is a key part of constructivist learning. Among the educators, philosophers, psychologists, and sociologists that added new perspectives to constructivist learning theory and practice are Lev Vygotsky, Jerome Bruner and David Ausubel. This philosophy of constructivism underlies the methodology of action research and they are achievable under collaborative endeavours. The constructivist theory of learning is only relevant to the process of developing the Action Cyber Crime Prevention Programme.

Gagnon and Collay (2006) stated that there are four epistemological assumptions at the heart of constructivist learning. These are:

- I. Knowledge is physically constructed by learners who are involved in active learning;
- II. Knowledge is symbolically constructed by learners who are making their own representations of action;
- III. Knowledge is socially constructed by learners who convey their meaning and making to others;
- IV. Knowledge is theoretically constructed by learners who try to explain things they do not completely understand.

The scholars further maintained that these epistemological assumptions can be seen in the works of Dewey, Montessori, Piaget, Bruner and Vygotsky, and that constructivism represents a paradigm shift from education based on behaviourism to education based on cognitive theory. Fosnot (1996) maintains that behaviourist epistemology focuses on intelligence, domains of objectives, level of knowledge and reinforcement, while the constructivist epistemology assumes that learners construct their own knowledge on the basis of interaction with their environment. One of the major arguments of the constructivist is that effective learning involves action or active participation of the learners which involves participatory activities where groups of students interact and help one another to learn. Participatory action model is not having students talk to one another, either face to face or in a conference while they do individual assignments; rather it involves working together to achieve a common purpose. For instance, students can be assisted to work together in a participatory mode, form a club, and elect officers. By learning to vote, they are learning the concept of democracy and at the same time acquiring the values of fairplay, tolerance and teamwork. From the constructivist point of view, there are some shifts in learning today. These shifts are:

- from instruction to construction and discovery;
- from teacher-centered to learners centered education;
- from absorbing learning materials to learning how to navigate and how to learn;
- from school learning to life long learning;
- from learning as torture to learning as fun;
- from the teacher as a transmitter to the teacher as a facilitator,

This philosophy of constructivism underlies the methodology of action research and they are achievable under collaborative endeavours. The constructivist theory of learning is only relevant to

the process of developing the Action Cyber Crime Prevention Programme, consequent on the selected students active participation in the development of the Action Cyber Crime Prevention Programme.

However, it may be necessary to examine some criminological theories that are relevant to the milieu in which cyber crimes are committed. Jaishankar (2008) propounded that there is a need for a separate theory of cyber crime because the general theoretical explanations were found to be inadequate as an overall explanation of the phenomenon of cyber crime. Consequent on these inadequacies, Jaishankar (2008) developed space transition theory in order to explain the causation of crime in the cyberspace. Space transition theory is an explanation of the nature of the behaviour of the person who brings out his/her conforming space and non-conforming behaviour in the physical space to cyber space. Space transition involves the movement of person from one space to another (e. g from physical space to cyberspace and vice versa). The theory argues that, people behave differently when they move from one space to another. The positions of the theory are:

- I. Persons, with repressed criminal behaviour (in the physical space) have a propensity to commit crime in cyberspace, which, otherwise they would not commit in physical space, due to their status and position.
- II. Identity flexibility, dissociative anonymity and lack of deterrence factor in the cyberspace provide the offenders the choice to commit cyber crime.
- III. Criminal behaviour of offenders in cyberspace is likely to be imported to physical space which, in physical space may be exported to cyber space as well.
- IV. Intermittent ventures of offenders in to the cyberspace and the dynamic spatio-temporal nature of cyberspace provide the channel to escape.
- V. (a) strangers are likely to unite together in cyberspace to commit crime in the physical space.
(b) associates of physical space are likely to unite to commit crime in cyber space.
- VI. Persons from closed society are more likely to commit crime in cyberspace than person from open society.
- VII. The conflict of norms and values of cyber space may lead to cyber crimes.

The space transition theory views the emergence of cyber space as a new locus of criminal activity. This theory shows that secondary school students, if, not educated on the ills of cyber crimes can physically unite in their school premises to perpetrate acts that are criminal which they would not ordinarily commit on their school compounds.

Another criminological theory relevant to this study is the theory of atavism which Fulcher and Scott (2007) described as the most notorious of the evolutionary approaches to deviant behaviours. The theory was propounded by Cesare Lombroso, who holds that many criminals were born with atavistic features. Lombroso found out that criminals have definite biological failings that prevent them from developing to a fully human level. The criminals showed certain ape-like characteristics or sometimes merely saving features that gave them the distinct anatomical characteristics from which they could easily be identified: large jaws, long arms, thick skulls, and large monkey-like ears and so on. These atavistic features, Lombroso argues, also lead them to prefer forms of behaviour that are normal among apes and savages, but are criminal in human societies. He claims that about forty percent of all criminals are born criminals of this kind, and are driven into criminality by their biology. Other law breakers are simply occasional, circumstantial offenders and do not have the atavistic characteristics of the born criminal.

The excesses of Lombroso's theory and the racial assumptions that underpinned it have long been discarded. However, many people still see criminality as resulting from innate characteristics. Violence and aggression, for example, are often seen not only as specifically male characteristic, but in their extreme forms as being due to genetic peculiarities. It has been proposed, for example, that many violent criminals have an extra Y chromosome in their cells. Thornhill and Palmer (2000) suggested that rape can be explained as a consequence of normal, genetically determined male behaviour. In the 1990s, the success of the Human Genome project led to many strong claims about the genetic basis of crime. The idea of the born criminal was supported in a report that pimping and petty theft appear to be genetically conditioned, but a person's genes have little influence on their propensity for committing crime of violence (Independent, 15 February 1994), violence was reported to be due to a mild brain dysfunction in early life health care for pregnant women could reduce violent crime by over twenty percent (Independent, 8 March 1994). The link between biology basis to violent behaviour, the ways in which this is expressed and the consequence that flow from it depend upon the

meanings that are attached to it and the particular social situation in which it occurs (Fulcher & Scott, 2007). Though, the excesses of Lombroso's theory and the racial assumptions that underpinned it have long been discarded, his findings led other researchers to study specific traits that could make human beings to commit crimes. One of the biological factors discovered by scholars is the presence of extra Y chromosomes in 1-3% of violent criminals in prison cells. This may be the factor responsible for males involvement in cyber crimes than their female colleagues across the globe. Action Cyber Crime Prevention Programme is proposed consequent on educational researchers that education could be used to tame the excess 'Y' chromosome responsible for male students cyber lawless behaviour in Nigeria.

In addition to these, Routine Activity Theory (RAT) propounded by Lawrence Cohen and Marcus Felson in 1979 to fill the shortcomings in existing models that failed to adequately address crime rate trends since the end of World War II and formerly applied to aggregate level crimes has been considered appropriate for cybercrimes (Pease, 2001; Yar, 2005). Kigerl (2011) saw Routine Activity Theory (RAT) as an ecological approach to crime causation, the accessibility, location, and presence or absence of environmental characteristics, and certain types of people are what prove predictive of criminal behaviour. Cohen and Felson (1979) posited that Routine Activity Theory requires three situations to be true and meet in space for there to be a crime that is committed. These include: a motivated offender, a suitable target, the absence of a capable guardian.

Akers and Sellers (2004) went further to explain that a motivated offender must be someone willing to commit a crime should he or she have the opportunity. Felson and Clarke (1998) further stated that a suitable target is one of the motivated offender values (credit card information). The motivated offender would have observed that the suitable target is visible, accessible, and able to be misappropriated by the offender. In addition to these, anything that obstructs the offender's ability to acquire the target (for example, antivirus, encryption) i.e. a capable guardian must be absent. Kigerl (2011) stated that research has discovered some support for the three Routine Activity theory conditions whether a crime occurs in cyberspace. Actions such as time spent online, more use of internet banking, making more online purchases and risky online behaviours make people more suitable to offenders. Research have shown that people with these actions are likely to be recipients of cyber bullying (Holt & Bossler, 2009), receive more phishing attacks (Hutchings & Hayes, 2009), are more often targeted by internet fraud (Pratt, Holtfreter, & Reisig, 2010) as well as other forms of fraud (Holtfreter, Reisig, & Pratt, 2008) and lost time due to malware infection (Bossler & Holt, 2007). Choo (2008) stated that lack of antivirus, antispyware, and firewalls (capable guardians) are associated with more malware victimizations. Consequent on these, there appears to be some support for Routine Activity Theory in explaining the likelihood of being the target or victim of a cyber crime attack. However, it is not easy to predict whether a country will more likely be the source of a cyber attack or be the place of residence for a disproportionate number of cyber criminals. Some countries are known to be more problematic than others in terms of cyber crime activity (Kigerl, 2011). This theory shows that the failure of the capable guardians like encryption, anti-virus and internet security programme that computer scientists entrusted secondary school students to, is responsible for their exploitation of the anonymity afforded them by the Internet to defraud unsuspecting people in the cyber space. Consequent on this discovery, we feel 'Education' through the Action Cyber Crime Prevention Programme would be a more capable guardian to students than encryption, anti-virus and internet security programs in the cyber space.

Purpose of the Study

Consequent on the annual rankings of Nigeria by the National White Collar Crime Center and the Federal Bureau of Investigation as the most notorious country whose citizens perpetrate cyber crimes in Africa and third most notorious country noted for cyber crime acts in the world, it is expedient to find a means of changing this orientation. The median age of the cyber crime offenders is found to be less than thirteen years, which is an age range peculiar to secondary school students in Nigeria. The absence of enabling laws coupled with the fact that laws sometimes do not serve as a deterrent, education is the key factor that could be used to change long-established attitudes and habits to end the menace of cyber crimes. And in the absence of any formal specific programme targeting secondary school students' crime prevention knowledge and attitude in Civics and Social Studies, this study, therefore, determined the impact of an informal Cyber Crime Prevention Programme (CCPP)

through the action or participatory approach. The study further determined the moderating effects of gender and computer literacy on the dependent measures.

Hypothesis

There is no significant main effect of treatment on students' attitude to cyber crime prevention

Methodology

This study adopted a participatory action research paradigm to develop an informal Cyber Crime Prevention Programme as research have shown that participatory action research paradigm could foster value acquisition among students (Koshy, Koshy and Waterman, 2010). The study further adopted the pretest-posttest, control group quasi experimental design (Shadish, Cook and Campbell, 2002) to determine the impact of the Action Cyber Crime Prevention Programme on secondary school students' learning outcomes in Civics and Social Studies in Ondo State.

Selection of Participants

The target population of this study comprises secondary school students in Ondo State. Six secondary schools were purposefully selected and assigned three a piece to the experimental and control groups for the purpose of this study in Akure, the capital city of Ondo State. The state capital is purposefully selected for this study because it has all the modern social amenities especially electricity, Internet Service Providers and cafes. Intact JSS 2 classes sample were selected in each of the six secondary schools that permitted their students' to partake in the study. Four hundred and twenty-six students started the study, but only two hundred and eighteen students were regular during the period the study lasted. The students dropped off consequent on their non residence on school compound. JSS 2 students were used for this study because, Civic Education and Social Studies are being taught at JSS level in Nigerian secondary schools, the students were not preparing for the external Junior Secondary School Examination, and as such they were more available to participate in the development of ACCPP. In addition to these, the students will be able to provide the initial leadership for the Cyber Crime Prevention Club, as the students have four more years in the secondary level of education; this would enable them to continue the Cyber Crime Prevention Clubs in their respective schools.

Research Instruments

Five instruments were used for the study. These were:

1. Students' Attitude to Cyber Crime Prevention Scale (SACCPS)
2. Computer Literacy Test (CLT)
3. Conventional (Class) Lecture Method Guide (CLMG)
4. The Focus Group Discussion Guide (FGDG)
5. Action Cyber Crime Prevention Programme for Experimental Group

Students' Attitude to Cyber Crime Prevention Scale (SACCPS)

This is a 22 item scale developed by the researcher to measure the attitudes of selected secondary school students to Cyber Crime Prevention. The scale is made up of two sections; Section A elicits responses on personal details like name, sex, school and membership of any Cyber Crime Prevention Club. Section B contains 22 scaled questions which consist of modified likert 4-point scale of Strongly Agree (SA), Agree (A), Disagree (D) and Strongly Disagree (SD). The scale was designed to measure the following:

- a. Attainment of requisite crime prevention attitudes for identifying crime related problems and issues in the immediate environment.
- b. Inculcation of the right values and skills that will enable students to actively participate in the cyber crime prevention process.
- c. Development of a sense of moral responsibility in preventing incidences of cyber crime in the immediate environment.

The items will be scored by allotting 4 points to SA, 3 to A, 2 to D and 1 to SD for positively worded statements. This will be reversed for negatively worded statements.

An initial 22 items were drawn and subjected to peer and expert review to determine their appropriateness with reference to the targeted learners. This was also used to establish its content validity. Based on their inputs, modifications were effected and the 22 items were then tried out on a sample of JSS 2 students that will not be part of the main study. For reliability, copies of the draft

questionnaire administered were subjected to Cronbach Alpha method and a reliability coefficient of 0.76 was obtained.

Computer Literacy Test (CLT)

The CLT was designed by the Office of Education, North American Division of Seventh Day Adventists in 2005 and revised in 2009. The test has two sections; section A has 60 questions which evaluates literacy, while section B has 10 questions, which tests competency. The 60 questions in section A are in 6 parts with 10 questions each. There are 13 options lettered A to M, with a correct option matching each question. Section B has 10 objective test items with options A – D which include only one correct option. The initial 70 questions were subjected to peer and expert review to ensure the construct and content validity of the test as well as its appropriateness for the target learners. The test items were thereafter reduced from 70 to 38 based on too high or too low difficulty indices, the results of the analysis were used to pick items that have discriminating index of 0.4 to 0.7. For reliability and average item difficulty, the final draft of the test was subjected to KR-20 which yielded 0.78.

The Conventional (Class) Lecture Method Guide (CLMG)

The CLMG was designed by the researcher to guide teachers in the control group. This was to ensure uniformity. The CLMG was prepared on each of the concepts selected for the study. The Conventional (Class) Lecture Method Guide is made up of five steps; these are:

- The teacher introduces the concept
- The teacher discusses facts or ideas on the concepts in steps
- The teacher gives notes on the concept
- The teacher asks questions
- The teacher gives assignment to students.

The Focus Group Discussion Guide (FGDG)

The researcher constructed the Focus Group Discussion Guide (FGDG). The instrument was constructed to enable secondary school students make decisions and develop the Action Cyber Crime Prevention Programme. The FGDG has 14 items and is made up of four (4) sections. To ascertain the validity of this instrument, the researcher made use of twenty-five (25) students from a co-educational school that were not involved in the study. Each item was discussed and any item that the students found difficult was amended. To ascertain the validity of the FGDG, the researcher made use of 20 JSS 2 students from a co-educational school that will not be involved in the study. Each item was discussed and the two items that the students find difficult was amended.

Development of the Action Cyber Crime Prevention Programme

The Roberts (2007) curriculum model was adapted for the development of the Action Cyber Crime Prevention Programme. The model comprises 3 stages, namely:

- I. Planning the curriculum.
- II. Establishing curriculum content through action approach
- III. Validating the curriculum

The model is presented in the Figure I:

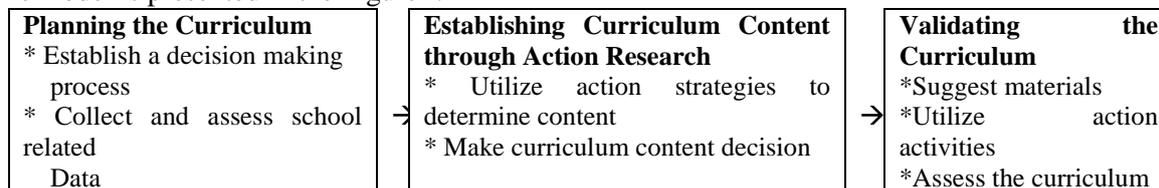


Fig. I: Curriculum model for the study (Adapted from Roberts, 2007)

Planning the Action Cyber Crime Prevention Programme

Focus Group Discussion Moderators (the selected teachers and research assistants) were selected and trained. The focus group discussion moderators moderated the group activities and led discussion towards making decisions in each discussion group. A training format was used to train the Focus Group Discussion Moderators. They were trained to achieve action means of reaching decision among group members and engage in participatory efforts toward solving problems.

The Junior Secondary School (JSS) Social Studies and Civic Education Syllabi was obtained, the crime prevention concepts were identified. The concepts that were identified are: citizenship, ICT,

problems of ICT, social issues and problems, negative behaviours, and values. JSS II Social Studies and Civic teachers were interviewed in each of the experimental schools to discover the method used to teach Crime Prevention concepts in the consulted syllabi. Some of the students in the experimental groups were selected randomly and interviewed.

Establishing the Objectives and Contents of the Cyber Crime Prevention Programme through Action Approach

The objectives and contents of the Action Cyber Crime Prevention Programme were developed collaboratively during the focus groups discussion. The selected JSS 2 students in the experimental schools engaged in the focus group discussion.

Focus Group Discussion

The focus group discussion was used to establish the objectives and contents of the Action Cyber Crime Prevention Programme. The focus group discussion took place in each of the three schools in the experimental group. Each focus group was made up of eight to ten participants. All the focus groups in each school had their discussion sessions simultaneously. The duration of the focus group discussion was two weeks. The outline for the focus group discussion was: introduction, setting the ground rules, overview of the topic, questions, discussion, and summary.

Validation of the Action Cyber Crime Prevention Programme

The developed Action Cyber Crime Prevention Programme was validated in each experimental school with the establishment of Cyber Crime Prevention Club. The validation involves the following steps:

Drafting the Cyber Crime Prevention Club's Constitution

The research assistants and the cooperating teachers educated the participants on the importance of drafting a constitution to guide the activities of the Cyber Crime Prevention Club. In each experimental school, copies of draft constitution prepared by the researcher were given to the participants, who grouped themselves and review every item in the draft constitution with the assistance of the participating teacher. The President of each review group later presented the positions of each group to all the participants. The participants harmonized the positions of each group on each item by consensus or voting. The participating teacher collected the harmonized constitution and read out the contents and called for motion to adopt the reviewed draft of the constitution, which the participants did. It could be observed that some marked differences exist in the contents of the constitution of each school, consequent on the peculiarities of each experimental school.

Election of Club Executives and Executive Meetings

The participants elected one to two students as electoral officer/officers in each school who conducted the election with the assistance of the researcher, participating teachers and research assistant students were asked to nominate candidates for each post, while the nominated candidates were given a day to campaign and sell their manifestos to their colleagues. The election took place a day after campaign in each experimental school, the electoral officers counted the votes, collated the results and announced immediately. The electoral officers in consultation with the participating teacher nullified elections into some positions in two of the three experimental schools where the number of votes casted, exceeded the accredited voters. The electoral officers re-conducted such elections, counted, re-collated and announced the results immediately.

The elected club executives were presented to the participants in each experimental school. In each school, the executives met, deliberated and decided on inauguration of the club, and inaugural meeting of the club.

Inauguration of the Action Cyber Crime Prevention Club

The Cyber Crime Prevention club was formally inaugurated in each school with the participating teacher who became the club staff Adviser, and Principal, who became the patron in attendance or represented by the Vice Principal. During the inauguration ceremony, the researcher mentioned the benefits that the participants and their schools could derive from selfless participation in the activities of the club such as acquisition of knowledge for effective crime prevention in the cyber space. The researcher presented a notice board, files, and banners to the club in each school, to enable the club take-up effectively.

Inaugural Meeting of the Cyber Crime Prevention Club

After the inauguration of the Cyber Crime Prevention Club in each experimental school, the clubs conveyed the inaugural meetings to discuss programmes the club would carry out and strategies to adopt in encouraging other students to join the club next academic session.

Lecture

Two lectures were organized for the Cyber Crime Prevention Club in each experimental school. The first lecture examined communication, ICT and Problems of ICT'. A lecturer from Adekunle Ajasin University, Akungba who was invited to 'IT Security Conferences for the Next Generation' in Malaysia and Germany to present his research on cyber crime prevention delivered the lecture in each experimental school. The students were grouped, each group had at least two laptops connected to the internet throughout the duration the lecture lasted. It was observed during the first lecture that participants with high computer literacy and competency helped their low and moderate computer literate colleagues to navigate through difficult computer tasks.

The second lecture centred on citizenship and negative behaviours. A High Court Magistrate and Legal practitioner delegated by Ondo State Judiciary spoke on 'What the Nigerian Judiciary, Police force, and Nigeria Prisons' does to Cyber Crime offenders. The legal practitioner taught for twelve years in the secondary school cadre before joining the bar, this enable him to sequence the lecture he gave to the students.

Drama

The Cyber Crime Prevention Club in each school staged a drama titled 'Food by Deceit' to educate the students who were non-club members and teachers on cyber crime prevention. The drama chronicled the life of Raymond, a young promising chap that got entangled in web of 'Yahoo Yahoo' (Cyber Crimes) syndrome which ultimately sent him to the abysmal dungeon of cul-de-sac. Students in one of the experimental schools composed a song titled 'E ma Yahoo' meaning 'do not involve in Yahoo Yahoo'.

Debate

The participants in each experimental school had debate on the topic 'ICT has done more harm than good to humanity'. In each school, the President, vice-President and Editor-In-Chief debated on the topic with three non-executive members of the Cyber Crime Prevention Club.

Post Test

The treatment lasted for ten weeks in the experimental schools. The researcher subsequently administered the instruments on the participants in the experimental and control schools.

Method of Data Analysis

The quantitative data was collated, coded and analyzed using Analysis of Covariance (ANCOVA). The Multiple Classification Analysis was employed to determine the magnitude of performances across the groups. A line graph was used to disentangle significant interaction effects.

Results

Hypothesis: There is no significant main effect of treatment on students' attitude to cyber crime prevention.

Table 2: Summary of ANCOVA of Posttest Attitude Scores by Treatment, Gender and Computer Literacy

Source of Variance		Hierarchical Method				
		Sum of Squares	Df	Mean Square	F	Sig
Covariates	PREATT	8029.586	1	8029.586	109.949	.000
Main Effects	(Combined)	12021.196	4	3005.299	41.151	.000
	TREATMENT	11618.533	1	11618.533	159.092	.000*
2-Way Interactions	GENDER	8.055	1	8.055	.110	.740
	COMPTLIT	394.609	2	197.304	2.702	.069
	(Combined)	711.911	5	142.382	1.950	.088
	TRTMENT *	4.283	1	4.283	.059	.809
	GENDER	151.950	2	75.975	1.040	.355
	TRTMENT *	465.427	2	232.713	3.187	.043
3-Way Interactions	COMPTLIT	81.894	2	40.947	.561	.572
	GENDER *					
	COMPTLIT	20844.587	12	1737.049	23.785	.000
		14971.193	205	73.030		

TRTMENT*GENDER* COMPTLIT	35815.780	217	165.050		
Model					
Residual					
Total					

* Significant at $p < 0.05$

Table 4.4 shows that there is significant main effect of treatment on students' attitude to cyber crime prevention ($F_{(1;205)} = 159.092$; $p < .05$). This means that the difference in the attitude of students exposed to Action Cyber Crime Prevention Programme and their control group counterparts is significant. Hypothesis 1b is, therefore, rejected.

Table 3: MCA of Posttest Attitude Scores by Treatment, Gender and Computer Literacy
Grand Mean = 64.89

Variable + Category	N	Predicted Mean		Deviation		Eta	Beta
		Unadjusted	Adjusted for factors and covariates	Unadjusted	Adjusted for factors and covariates		
TREATMENT Cyber Crime Prevention Prog	75	77.09	75.19	12.19	10.29	.689	.582
	43	58.50	59.49	-6.39	-5.40		
GENDER Male	1	65.03	64.45	.13	-.45	.010	.035
	08	64.77	65.34	-.13	.44		
COMPUTER LITERACY Low Moderate High	98	61.46	63.39	-3.44	-1.51	.299	.110
	86	65.81	65.82	.91	.92		
	34	72.50	66.93	7.60	2.03		
R = .748							
R Squared = .560							

Table 4.5 shows that students exposed to the Action Cyber Crime Prevention Programme had higher adjusted posttest attitude mean score ($\bar{x} = 75.19$; Dev. = 10.29) than their control group counterparts ($\bar{x} = 59.49$; Dev. = 5.40). To this end, the Action Cyber Crime Prevention Programme was more effective than the conventional teaching.

Discussion

Results obtained shows that the students' exposed to the Action Cyber Crime Prevention Programme obtained higher adjusted posttest attitude mean scores. The programme proved effective to the conventional lecture method, consequent on the relevance of the programme and retention of the participatory activities that allowed students to hear their views, learn from their colleagues and the support of the school and teachers for the programme. The Action Cyber Crime Prevention Programme was more effective because the learners were at the centre of the action activities; hence prevention attitudes were developed effectively.

Conclusion

This study has established that the action Cyber Crime Prevention Programme is effective at enhancing students' knowledge of crime prevention and attitude to cyber crime prevention. The programme can be used to not only foster the learning of cyber crime prevention concepts, but attitude to cyber crime prevention irrespective of students' gender and computer literacy level. Consequent on the findings of the study, it could be concluded that the Action Cyber Crime Prevention Programme have great potentials at enhancing students' attainment in crime prevention concepts and attitude to

cyber crime prevention. The participatory action research paradigm adopted to develop the Action Cyber Crime Prevention Programme changed the learning landscape of the students not only from torture to fun but from instruction to discovery which developed in the students civic values and skills that would enable them to shun on-line activities that may project the image of Nigeria in a bad light. The formation of the Cyber Crime Prevention Clubs to sensitive members of the school community and public at large on dangers inherent in cyber crimes developed in the students civic values and skills like cooperation, unity, love, sympathy, consideration of others welfare, selflessness, hardwork, positive use of knowledge, honesty, obedience, respect, fair play, patience and loyalty. In this case, Civic Education and social studies teachers should make conscious efforts to utilize the Action Cyber Crime Prevention Programme when teaching crime prevention concepts in Civic Education and Social Studies. Also, the government and relevant agencies and professional associations should organize regular workshops and in-service teacher education programmes for social studies and civil education through which emerging issues could be discussed. Schools should also put in place facilities that can encourage students to participate in non formal research activities.

Implications for Teacher Education

This study developed and validated an Action Cyber Crime Prevention Programme for secondary school students in Ondo-State, Nigeria, consequent on the growing concern at finding solutions to the incidences of cyber crime among school – age children in Nigeria. The students were encouraged to develop a Cyber Crime Prevention Programme through the bottom-up approach that could be used to educate students' against involvement in cyber crimes. The findings of the study show that the sustained implementation of the programme would enable the students to inculcate good values while roaming the cyber space through the internet.

Another major implication of this study is the acquiring of the values and traits of group roles like positive use of knowledge, patience, consideration of others, cooperation, love, unity and loyalty by the students. The action approach adopted helped the students to acquire these values, thus showing the importance of participatory approach in equipping learners with values and traits requisite for effective cyber crime prevention. This finding shows the need to adopt the participatory approach to complement the conventional method used to teach Civic Education and Social Studies in Nigerian secondary schools.

In addition to these, curriculum planners could adopt the participatory approach used in this study for curriculum development in order to involve students, who would benefit from such educational programme to participate in designing the programme. The 'Bottom-up' approach, rather than the top-down approach should not only be used to develop Civic Education and Social Studies curriculum, but to teach the subjects consequent on the trans-disciplinary nature of Civic Education and Social Studies.

The Action Cyber Crime Prevention Programme is recommended to teachers for educating secondary school students in Nigeria on Cyber Crime prevention. Students should be encouraged by the teachers to construct personal ideas, identify conceptions and misconceptions and allowed to correct their misconceptions when learning crime prevention concepts in Civic Education and Social Studies Education. Government should empower teachers with participatory learning facilities in secondary schools and adopt bottom up approach in developing curricula in Civic Education and Social Studies. Government could also use the Action Cyber Crime Prevention Programme to train commercial cyber café operators on the risks of indulging secondary school students who use Internet to defraud foreign nationals.

The views represented in this paper, and any errors that accompany them, are, of course, ours.

References:

- Adesina, A.D.O. and Adeyemi, B.A. 2007. Teaching to Achieve Social Studies Values: A case of re-education of Teachers. Retrieved Oct. 10, 2010 <http://www.deta.up.ac.za/archive2007/presentations/word/Teaching%20to%20Achieve%20Social%20studies%20values%20ADEYEMI.pdf>
- Akers, R.L. & Sellers, C.S. 2004. Criminological theories: Introduction, evaluation and application (4th ed., p.33). Los Angeles, CA: Roxbury. and Technology. Publication of the National Association of Academies and the University of Texas. Winter, 2003.

- Bossler, A.M & Holt, T.J. 2007, November 14. Examining the utility of routine activities theory for cyber crime. Paper presented at the annual meeting of the American Society of Criminology. Atlanta Marriot Marquis, Atlanta,GA.
- CESAC. 1993. Nigerian Secondary Schools Social Studies Project. Book 1. Lagos: NERC.
- Chiemekwe, C. S, and Longe, O. B. 2007. Information and Communication Penetration in Nigeria: Prospects, Challenges and Metrics. *Asian Journal of Information Technology*, 6(3) 280-287.
- Choo, K.K.R. 2008. *Organised Crime Groups in Cyberspace: A Typology* Springer Science + Business Media, LLC.
- Cohen, L.E. & Felson, M. 1979. Social change and Crime rate trends: A routine activity approach. *American Sociological Review*. 44.
- Dick, T and Herbert, L. 2003. Youth, Pornography and the Internet. *Issues in Science*
- Felson, M. & Clarke, R.V. 1998. Opportunity makes the thief: Practical theory for crime prevention. *Policing and Reducing Crime Unit: Research, Development and statistic Directorate*. 98: 1-36.
- Fulcher, J. and Scott, J. 2007. *Sociology 3rd Edition*. Oxford University Press. Retrieved June 28 from <http://docs.google.com/viewer?a=v&q=cache:loyxiY3axbE>
- Gagnon, G.W. and Golly, M. 2006. *Constructivist Learning Design*. Retrieved Apr.25, 2010 from <http://www.google.com>
- Holt freter, k., Reisig, M.D., & Pratt, T.C. 2008. Low Self-control, routine activities and fraud victimization. *Criminology*. 46: 189-220.
- Holt, T.J. & Bossler, A.M. 2009. Examining the applicability of lifestyle- routine activities theory for cyber crime victimization. *Deviant Behaviour*. 30: 1-25.
- Holtfreter, K., Beaver, K. M., Reisig, M. D., & Pratt, T. C. 2010. Low self-control and fraud offending. [DOI: 10.1108/13590791011056264]. *Journal of Financial Crime*, 17(3), 295-307.
- Hutchings, A. & Hayes, H.2009 Routine Activity theory and phishing victimization: who got caught in the 'net'? *Current issues in Criminal Justice*.20: 432-451.
- Independent. 1994. 'Genes linked to Violence and Crime Condemned'. February 15. Retrieved June 28, 2011 from <http://www.independent.co.uk/news/genes-link-to-violence-and-crime-condemned-1573089.html>
- Independent. 1994. 'The Human Genome Project'. March 8. Retrieved June 28, 2011 From <http://www.independent.co.uk/news/genes-link-to-violence-and-crime-condemned-1573089.html>
- Jaishankar, K. 2008. Space Transition Theory of Cyber Crimes. In Schnallagar, F. & Pittaro, M. (Eds.). *Crimes of the Internet* (Pp. 283-301). Upper Saddle River, NJ: Prentice Hall.
- Kigerl, A. 2011. Routine Activity Theory and the Determinants of High Cybercrime Countries. *Social Science Computer Review*. 000(00).1-17.
- Koshy E. Koshy V. Waterman H. 2010. *Action research for health care*. London: SAGE Nigeria: Prospects, Challenges and Metrics. *Asian Journal of Information Technology*, 6(3) 280-287.
- Nigerian Educational Research and Development Council (NERDC). 2007. *UBE Civic Education Curriculum*. Abuja. Federal Government Press.
- Obidoa, A. 1991. A Re-appraisal of Liberal Arts and Social Studies Education in Nigerian Schools in National Secondary School Curriculum Review Conference Proceedings, Ibadan: Macmillan Nigeria Publishers Limited.
- Pease, K. 2001. Crime futures and foresight: Challenging criminal behaviour in the information age. In D. Wall (Ed.), *Crime and the internet*. London, England: Routledge
- Roberts, M.R. 2005. *Curriculum Development Process*. Northern Arizona,CTE 592.
- Roberts, M.R. 2007. *Curriculum Development Process*. Northern Arizona,CTE 592.
- Shadish, W.R., Cook, T.D. & Campbell, D.T. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin Company.
- The National White Collar Crime Center .2010. *IC3 2009 Internet Crime Report* (January 1, 2009 to December 31, 2009). Retrieved May 3, 2008 from <http://www.IC3.state.gov>. Pg. 1-29.
- The National White Collar Crime Center .2011. *IC3 2010. Internet Crime Report* (January 1, 2010 to December 31, 2010). Retrieved May 3, 2008 from <http://www.IC3.state.gov>. Pg. 1-29.
- The National White Collar Crime Center. 2002. *IFCC 2001 Internet Fraud Report* (January 1, 2001 to December 31, 2001). Pg. 1-27. Retrieved Apr. 27, 2007 from <http://www.IC3.gov>

- The National White Collar Crime Center. 2003. IFCC 2002 Internet Fraud Report (January 1, 2002 to December 31, 2002). Pg. 1-27. Retrieved Apr. 27, 2007 from <http://www.IC3.gov>
- The National White Collar Crime Center. 2004. IFCC 2003 Internet Fraud Report (January 1, 2003 to December 31, 2003). Pg. 1-28. Retrieved Apr. 27, 2007 from <http://www.IC3.gov>
- The National White Collar Crime Center. 2005. Internet Fraud-Crime Report (January 1, 2004 to December 31, 2004). Pg. 1-29. Retrieved Apr. 27, 2007 from <http://www.IC3.state.gov>.
- The National White Collar Crime Center. 2006. Internet Crime Report (January 1, 2005 to December 31, 2005). Pg. 1-27. Retrieved Apr. 27, 2007 from <http://www.IC3.state.gov>.
- The National White Collar Crime Center. 2007. Internet Crime Report (January 1, 2006 to December 31, 2006). Pg. 1-27. Retrieved Apr. 27, 2007; from <http://www.IC3.state.gov>.
- The National White Collar Crime Center. 2008. IC3 2007 Internet Crime Report (January 1, 2007 to December 31, 2007). Retrieved May 3, 2008 from <http://www.IC3.state.gov>. Pg. 1-29.
- The National White Collar Crime Center. 2009. IC3 2008 Internet Crime Report (January 1, 2008 to December 31, 2008). Retrieved May 3, 2008 from <http://www.IC3.state.gov>. Pg. 1-29.
- Thornhill, R. and Palmer, C.T. 2000. *A Natural History of Rape: Biological Bases of Sexual Coercion*. Cambridge: MA. MIT Press.
- Yar, M. 2005. The Novelty of Cyber Crime: An assessment in Light of Routine Activity Theory. *European Journal of Criminology*. 2: 407-427.

THE ARCHITECTURE OF VIRTUAL SPACE MUSEUMS

Anda-Ioana Sfinteş, PhD c. arch.

University of Architecture and Urbanism „Ion Mincu” Bucharest, Romania

Abstract:

Using the internet is more than a trend if we are to acknowledge the fact that it does not only shape its users, but it also reflects the changes that take place at a social level through everything that the user itself gives back to the network, through users interactions in virtual space, through consumption and demands virtually stated etc. As we pass from searching to sharing, playing or shopping on-line, we realize that between these theoretically distinct actions the boundaries blur. We identify this movement towards unclarity all around us and more and more profound lately (for example, in *guerilla* theatre or multifunctional buildings). In this context, *the virtual museum*, *eMuseum* or *online museum* are exhibition forms that further expand, blur, interweave defining characteristics of the museum with other kinds of activities not rarely considered opposed to it. Many museums have already added up to their physical exhibitions virtual tours or on-line exhibitions (some even created especially for the internet). The experiential differences between on-line/on-site museums are many and important, answering to different needs. However, all the more so as we admit that, by visiting an exhibition, we do not only experience the display layout, but also the space itself, we can further ask ourselves what happens with the curatorial discourse and architecture in virtual space. The aim of this article is to find a response to this question by analyzing the place, role and impact of architecture in the case of virtual space museums.

Key Words: Museum architecture, virtual space, discourse

Introduction

Nowadays, we tend to search for everything firstly on the internet. It's easier, faster, more convenient and, depending on the website, it can also be quite reliable. The ceaseless technological upgrading facilitates the access to more and more complex information, but also to better communication and interaction activities which further became social demands and needs. Museums developed their websites in order to accommodate these facilities and even implemented their own programs which better responded to their new related missions (like, for example, those of addressing themselves to a broader audience, of allowing interpretation and negotiation of meaning and of becoming community involved). The ongoing *Taking Part* survey (“Taking Part 2012/13 Quarter 3: Statistical Release” 2013) shows that, in the case of museums, the digital engagement grew to 29.9% from 15.8% in 2005/6 and that the online visitors have been interested especially in finding out information about exhibitions, events and particular subjects, in ordering tickets and taking a virtual tour. These all new available features and their popularity further raise questions about expectations, goals and better forms of communication between museums and their visitors: *what do museums wish to offer and what do visitors wish to discover?* In this context, we wish to interrogate the role that architecture yet plays and the way in which is hence uses, presented and understood.

The blurring of the boundaries between leisure / education / consumption transforms the museum websites into places of (re)presentation / information / exhibition / marketing / kindling. As the compounding elements of the museum as an institution accumulate diverse meanings and purposes, the virtual space augments this trend through the various possibilities of utilization it offers. Museum websites can awake curiosity, encourage socialization and community involvement etc. Architecture can also be exploited differently (and partially it is), and uncovering its role in virtual space can lead to further developments which could help museums accomplish their newest established goals.

Museums on the web

Museums gradually went from being the institution dedicated to the acquisition, preservation, conservation, research and display of human life vestiges to getting involved in the life of its visitors, crossing over its boundaries in an attempt to face constantly changing social, economical and political needs and desires. As Susan Crane states, the museum and its contents are even being shaped by these changes (Crane 2006), as we can see, for example, in the case of museums dedicated to communities (like immigrants for whom creating their own representational space means creating a benchmark, helping them to adapt to a new society while keeping their old values alive). Anyway, blurring the boundaries also meant expanding the museum branch to include enjoyment and social work activities, as well as interpreting the idea of museum in order to incorporate more conceptual, objectless, interactive and virtual exhibitions. Sharon Macdonald edited *A Companion to Museum Studies* (MacDonald 2006a) issuing from the expansion, in range and variability, of museum events and programs and from “the development of museum «franchises», «blockbuster» shows, iconic «landmark» architecture [...], «superstar» museums [...] and «meta-museums»” (MacDonald 2006b, 5) which must not only be understood as identity statements but also as activities aiming to be spectacular, promoting culture and symbols in the context of globalization.

Virtual space museums take all these preoccupations even further while adding up their own aims and purposes. The evolution of museum concept and curatorial practices – the opening up towards a wider public, the shifts that took place from *old* to *new museology*, from simply displaying objects to introducing a narrative discourse and allowing the subjective interpretation of the exhibitions (see Fleming 2005, Psarra 2005, Hillier and Tzortzi 2006, Giebelhausen 2006), from considering the visitor as neutral and passive to his contemporary central and dynamic importance for the curatorial discourse (see Bollo and Dal Pozzolo 2005, MacDonald 2006b, Hooper-Greenhill 2006, Ravelli 2006) – all these led the way to go online. Museums started creating their own virtual pages and further to organize virtual tours and exhibitions, to see to their own blogs and to send and receive feeds in order to keep themselves and their public informed and up to date. Thus, the visitor can now participate not only physically and psychologically at the negotiation of meaning through “constructing, sharing and interpreting a range of content, attitudes and values” (Ravelli 2006, 3), but also virtually.

In this train of thought, we can also refer to the social work of museums as many of the online programs developed extend the support offered to the individuals and communities related. What prevails in this case in the virtual space isn't necessarily engaging with the exhibits, but with the others, let it be the museum personnel or other virtual visitors. Lois Silverman stated in *The Social Work of Museums* (2009) that museums serve the self, the close pair, the family and the group. Some of the ways of serving still remain valid online. For the self, for example, visiting a virtual museum might also facilitate relaxation and introspection (although less intense than in the case of an actual visit), but programs concerning health education, public health mobilization, enhancement of health care environments and public service, promotion of stability, support for change, social consciousness, religious and spiritual tolerance can easily be doubled online, promoted and even explained more thoroughly, encouraging participation and interaction (through comments) when possible. Communication, affiliation and membership, role enactment, personal storytelling can all even be enhanced online as former barriers fall (like those concerning physical distance or the reticence of opening one's mind to someone face to face) while others rise (coming, for example, from the uncertainty and mystery in which the others are being entangled, thus from the psychical distance). The same happens for the group: the virtual museum offers resources for public communication, promotion of civic engagement, consciousness-raising, representation, intergroup contact and service, multigroup collaboration. The virtual space emphasizes the idea of museums as “public spaces in which definitions of cultures and their values may be actively contested and debated” (Mason 2006, 18).

Anyway, not many of the online museums have also features implying human interaction. Some only concentrate upon presenting the institution, the collections and exhibition. Here we can also make a distinction between the interactive sites which make available virtual tours or objects and information self-handling and those which do not. The more complex a site is, the more we can talk about its role in a broader context. Michelle Henning (2006) advocates for the role of new media in organizing and structuring knowledge, in inverting the hierarchies between front and back regions of the museum. Although she refers mainly to the new media devices which accompany the actual

exhibit, frequently the same digital information (like database images, recordings, texts) can also be consulted online. Having access to such information may facilitate the self organization of museum collections according to preferences, desires or needs. The institutions thus renounce to their “traditional authoritative voice” and “break down disciplinary boundaries and hierarchical systems of cataloguing” (Witcomb 2003, 121). Henning also concludes that new media objects can model “ways of thinking and understanding which are non-hierarchical and decentralized, and privileging allegorical and arbitrary associations, correspondences, and resonances” (Henning 2006, 315).

The true challenge of virtual museums is still at its beginning. Although the sites keep evolving, the same do the demands addressed to them. Ross Parry concludes at the *Museums in the Information Age: Evolution or Extinction?* debate, held at The Science Museum in London in 2012, that the future of museums in the information age will be *social, situated, sensory* and *semantic*. By *social* he refers to ultra-socialization and personalization, by *situated* – to providing specific content to the visitor based on his location, while *sensory* raises concerns about the changing relationship between humans and digital world, making it necessary for the museums to recognize and use accordingly the blurring of the boundary between digital and not-digital. The *semantic* evolutionary direction alludes to the ability to make online connections between items.

Museum architecture developed as well, in line with all the fundamental transformations of the institution. The shift from *old to new museology* also meant changing architectural form, purpose and concepts “from public monument to spatial experience, from scientific centre to popular destination and landmark, and from forming a social event to shaping national and cultural aspirations” (Psarra 2005, 81). The subsequent shift towards virtual museums will clearly put architecture in a new light. In this context, we wish to investigate, first of all, the role that architecture currently plays online for museums, opposite to the interplay between space and curatorial discourse on-site. The aim is that of spotting current strengths and weaknesses and thus to discover opportunities for further development. Anyway, it is clear now that in the center of any approach concerning museums will continue to be the individual as *user, audience member, learner, customer* (Parry in *Museums in the Information Age: Evolution or Extinction?* 2012), visitor, consumer.

Museum architecture in virtual space

By such comparison between the real and the virtual form of the same space we have to actually ask ourselves what is being gained and what is being lost by passing from one to the other. We best acknowledge the impact and importance of the real feel when the physical presence or absence strikes us (see Lawson 2001). Daveen Koh, a *Cornell Daily Sun* blogger, describes his revelation as follows:

I remember being in a roomful of Monets at the Metropolitan Museum of Art this past summer, and all I could think of was that I was in a **roomful** of Monets. And perhaps, the hordes of tourists, saddled with large cameras, felt the same excitement. We were in the company of the famous. **We were standing on sacred ground.** (our emphasis, Koh 2012)

We thus see not only the role played by the physical presence of objects and people in inducing a certain morale, but also the reference to space: the feeling of a **roomful** of Monets and saying that you stay on **sacred ground**, they both come clearly from the strong relationship between the paintings and the room they occupy. Anyway, we aren't necessarily aware of the impact that the built environment has upon us. We do experience it, but our reaction to it is rather unconscious, affective, subjective, associative and contextual (see Moore 1979, Rapoport 1990 and Psarra 2009).

In order to identify the role of architecture in on-line museums we shall first of all review the on-site connections between exhibits and built environment. The *space syntax* theory seeks to uncover the interrelationship between spatial configuration and display and its impact upon the visitors' route, exhibit experimentation and interpretation (see Hillier and Tzortzi 2006). Kali Tzortzi (2007) questions the junction of curatorial decisions and space. She states that a) space can be exploited in order to amplify the display; b) the layout of objects can emphasize the spatial characteristics; c) space can be neutrally treated. Either way, space is being experimented as well, beside the exhibit, and it influences the subjective perception and interpretation of the museal space.

Some of the questions that Bill Hillier and Kali Tzortzi (2006) pose in regard to spatial design vs. museum layout change online, while others vanish. It doesn't make sense anymore to ask ourselves if the spatial design does “make a difference to how a gallery works as a social space” (Hillier and Tzortzi 2006, 282) as long as the halls of the museums are empty and frozen in the virtual tours. As for the experience of visiting (if we can still consider it an experience), can we say that the spatial

organization into sequences shape that experimentation and influence the further movement? At least, this last question isn't entirely out of place online, as long as the sequences of virtual tours are even easier to detect and the movement isn't bound to an itinerancy anymore. Anyway, the *space syntax* theory represents a good reference point in identifying the changing role of architecture in virtual museums. Starting from one of the two basic ideas of the theory – “space is not just the *background* to human activity and experience, but an *intrinsic* aspect of it” (authors' emphasis, Hillier and Tzortzi 2006, 283) – raises the first challenge in our inquiry, the answer to the question: *is the architecture of virtual space museums only a background?* To answer that, we shall first of all review the forms, aims and impact of museum architecture online.

David Fleming (2005) wonders where does the museum space begin and ascertains that “[t]he whole point of marketing and publicity, and image-building and branding, is to prepare people to make contact with a museum” (Fleming 2005, 55) Thus, he uncovers the idea of a psychological space of each museum in which you enter way before the visit, or even before deciding to take the visit. As an extreme example we can take the excessively marketized *Guggenheim Bilbao*. As a cultural facility, housed inside an iconic structure, the museum had the role of enhancing Bilbao's visibility and value at an international level (Plaza, Haarich, and Waldron 2012). The image of the spectacular building designed by Frank Gehry has been heavily used by mass-media, thus creating a very good name awareness not through artworks but through architecture (Caldwell 2000). Even on the <http://www.guggenheim.org/bilbao> main page, the cover photography depicts a general exterior view of the museum, accompanied by few lines praising the uniqueness of the structure. The purpose of using the image of the building has been, first of all, an economical one, aiming at attracting tourists and investors capable of renewing the city. Through these images, people got familiar with the museum even if, at that time, they had no intention of visiting Bilbao or the *Guggenheim*. Similar examples of museums presenting the history and architectural concepts of their iconic building are: *The Jewish Museum* in Berlin, designed by Daniel Libeskind (<http://www.jmberlin.de/main/EN/04-About-The-Museum/01-Architecture/00-architecture.php>), *Tate Modern* in London, designed by Herzog & de Meuron (<http://www.tate.org.uk/about/who-we-are/history-of-tate>), *MAXXI – The National Museum of the 21st Century Arts* in Rome, designed by Zaha Hadid (<http://www.fondazionemaxxi.it/museo/progetto-architettonico/?lang=en>) etc. Less extreme is the case of any other virtual space museum which presents, on its own website, images with the building and the exhibits, wishing rather to make the future visitors acquainted with the museum and not to sell itself through architecture.

Another way of presenting the physical image of the museum online is through the virtual tours. Much more elaborated and detailed than the simple images, the virtual tours offer the possibility of moving throughout the exhibition halls, of zooming in and out, of looking around, up and down, dependent on the available technology. This facility allows access to the museum halls and collections to anyone who possesses an internet connection, the adequate hardware and software and, last but not least, the skills to manipulate the programs employed. Thus, although the virtual tours are spatial boundary breakers they can also limit access for certain people. Lianne Mctavish (2006) analyses the experience of visiting an online museum and she notes as strengths the transgression of physical boundaries by jumping between halls and floors and the absence of coercion when it comes to following a path. On the other hand she acknowledges the limits of the dynamic movement offered by virtual reality museums, as you have to choose your place from preset positions. A very important aspect in online museums is that “[v]irtual viewers are offered a limited bodily experience, which stresses visual (and occasionally also aural) perception.” (Mctavish 2006, 233) Anyway, this compensation is rather contradictory. On the one hand, the empty halls, disembarassed of the others, their behavior or their remarks, stress the individual engagement with the objects displayed, but on the other hand they neglect the role the others play in understanding, learning, socializing etc. In the same time, the space recreated online is elusory and, by being emphasized, it actually draws the attention from the objects towards insignificant details. The result is that “[a]s details take precedence, the main subject of virtual galleries becomes less clear; distinctions between foreground and background are blurred. The boundaries of art works are implicitly questioned” (Mctavish 2006, 231).

In the physical space of museums, the interrelationship between architecture and curatorial discourse is being assessed depending on “fundamental spatial qualities – such as, hierarchy, axiality and perspective – and key configurational properties – as, for instance, *integration, connectivity* and

control” (author’s emphasis, Tzortzi 2007, 72.7). Hierarchy, axiomaticity and perspective are less opened to interpretation online. The lack of movement liberty, except between fixed points, blurs the visitors’ own searches and discoveries and it hinders the shaping of a full spatial understanding. It is also harder to get a sense of orientation through the mediation of mouse clicks. Sophia Psarra asserts that “*the relationship between geometric and spatial properties can be understood as based on the varying degree of geometrical control over the potential for variance in the structure of visual fields observed with bodily movement*” (author’s emphasis, Psarra 2009, 227). Online, the bodily movement, as well as the visual fields, are mediated and limited. If we also consider the lack of human scale and the lack of temporal dimensions usually mediated through the narrative (see Silverstone 2002), we see that the sense of space and time are both distorted. However, the curators have the chance of framing online the perspectives, installations and objects that stress the most the curatorial desires and discourse. Integration and connectivity are being limited as well and bound to the same fixed positions. The sudden pass from one point to the other cancels the fluency of the physical space. The liberty of not respecting a certain path and of jumping between halls gives much more control to the virtual visitor, but it also makes much more difficult the understanding of the general layout and the perception of spatial characteristics related to it. We can thus say that museum virtual tours rather emphasize the objects by themselves. For example, *The National Museum of the Romanian Peasant* in Bucharest has a curatorial discourse based on the relationship between the layout of the objects and the spatial configuration (see Sfinteş 2012). The visual and spatial integration complete the search driven narrative discourse (see Bernea, Nicolau, and Huluiţă 2001). It is obvious that the creators of the virtual tour (<http://www.tur.muzeultaranuluiroman.ro/>) tried to catch, in images, the interrelationship between halls, *strong/weak* objects, installations and space, but frequently the spatial characteristics (like the narrow spaces) led to unpleasant skews which actually harm the reading of the space. Consequently, the attention is driven towards objects, separately. *The Smithsonian National Museum of Natural History* in Washington D.C. (<http://www.mnh.si.edu/vtp/1-desktop/>) and *The Louvre* (<http://www.louvre.fr/en/visites-en-ligne>) in Paris both have great panoramic virtual tours, with high resolution images, but they still cannot be compared with the dramatic physical experimentation of space and exhibition. This only comes to emphasize the fact that virtual tours aren’t meant to substitute the actual visit, but rather to bestir curiosity and to better prepare the visitor for it, as “[t]he emotive experience of seeing the real requires the real and no surrogate will do” (Knell 2003, 140).

However, there are many examples of well established museums which do not offer virtual tours. Instead, they oriented themselves towards social networks. Particularly important in our analysis is the video streaming on YouTube. Short videos, eventually available also on the museum website, present different aspects of the exhibition, sometimes better capturing architectural details and the curatorial decisions regarding the relationship between space and layout. Adding sounds and speeches related to the images, the short films augment the degree of understanding as wished by curators, while diminishing the freedom of movement, albeit limited, available in the virtual tours. However, although you can get a better sense of spatial characteristics (because of the skew correction, the human scale, the various perspectives upon the same objects or installations), orienting yourself can still be difficult as you no longer have access to a larger context. A few museums using video streaming on YouTube are: *MAXXI – The National Museum of the 21st Century Arts* (<http://www.youtube.com/user/MuseoMAXXI>), *Musée de Quai Branly* in Paris (<http://www.youtube.com/user/quaibrantly>) and *The Jewish Museum* in Berlin (<http://www.youtube.com/user/jewishmuseumberlin>). They all have videos taking into consideration (on various degrees) the building as well. Although it might be perceived rather as background, it cannot be ignored, thus contributing to a more thorough understanding of the exhibition. The enactment can even highlight less obvious aspects and encourage new readings.

Conclusion

The museum evolution towards new forms and concepts is regarded by many with fear and seen as a decline and an alienation from the main purposes of acquiring, preserving, conserving, researching and displaying material and immaterial evidences of human life. Entertainment and consumption become key features, interfering “with the purity of the cultural experience” (Fleming 2005, 59) and transforming the museum into “a place in which culture is exploited in order to create turnover” (Lampugnani 2006, 252). Online museums don’t go far from this trend, promoting entertaining and commercial events, but they also extend their limits virtually thus making their

exhibitions and activities known to a broader audience. On the other hand, we saw that online the individual engagement with the objects can even grow and the exhibits can regain their central position. Clearly the on-site experience still has major advantages compared to the on-line visit, but depending on each one's needs that led to accessing virtual space museums, different features can be found helpful and even lead to the later (re)turn to the physical space. In other words, as Michelle Henning says, "New media's greatest promise is to be found [...] in the part it plays in a return to curiosity" (Henning 2006, 316). We cannot expect from virtual museums to be able to replace the actual experience, but we also cannot ignore the opportunities they offer, from reaching people unable to physically visit and keeping their public informed to putting their collections into a new light and expanding their goals and missions far beyond their physical reach. Going online is a "win some, lose some" situation. The question that rises then is: *how should the virtual space be exploited in order to offer a unique, diverse experience, related to the aims of the institution and to the needs and expectations of its online visitors?*

In this article we concentrated upon the architecture of virtual space museums. We saw that pictures depicting the building are used on the internet in order to familiarize the future visitors with the building or to create a better name awareness through the connection established with the iconic structures (when the case) rather than with the works of art. Virtual tours get closer to illustrating the whole experience of visiting (objects, display, space, layout etc.), but the difficulties encountered in manipulating the software, the limited quantity of information made available due to economical reasons (less objects presented) as well as the technological deficiencies are real drawbacks. Based on the above analysis, we consider that architecture and its relationship with the exhibition are best understood in videos. In this case, even if we do not have access to the whole museums but rather to small areas, the presentation is usually better detailed, which makes the reading more thorough. The producer/curator has the opportunity of choosing the best angles for an object or an installation, by contrast with the spots, capturing the most information in a spin, used in virtual tours. Anyway, the reasons for entering a museum website may vary greatly, as the reasons for searching museum architectural instances online. Different features answer different needs and sometimes we need to see the same aspects from different angles for a better understanding. As long as architecture is an intrinsic aspect of "human activity and experience" and not just a background (Hillier and Tzortzi 2006, 283), maybe we should try to capture it and its interaction (influence and transformations) with humans online also. Further online programs and upgraded virtual tours might incorporate, for example, sounds taken from the site (and we shall acknowledge the fact that sound is very important in reading space), human scale and interaction, panoramic views which better transform 2D images into 3D and 4D. A good example to follow might be the games industry which already developed online games with real time human interaction, but this idea itself could steer long debates as museums have to continually negotiate their "proper" (Mctavish 2006, 229) social role.

References:

- Bernea, Horia, Irina Nicolau, and Carmen Hulață. *Câteva gânduri despre muzeu, cantități, materialitate și încrucișare / Dosar sentimental*. n.p.: Ars Docendi, 2001.
- Bollo, Alessandro, and Luca Dal Pozzolo. "Analysis of Visitor Behaviour Inside the Museum: An Empirical Study." AIMAC Conference article. Montreal, 2005 http://neumann.hec.ca/aimac2005/PDF_Text/BolloA_DalPozzoloL.pdf, accessed February 03, 2012
- Caldwell, Niall G. "The Emergence of Museum Brands." *International Journal of Arts Management* 2 (3): 28–34, 2000.
- Crane, Susan A. "The Conundrum of Ephemerality: Time, Memory, and Museums." In *A Companion to Museum Studies*, edited by Sharon MacDonald, 98–109. Cornwall: Blackwell Pub., 2006
- Fleming, David. "Creative Space." In *Reshaping Museum Space: Architecture, Design, Exhibitions*, edited by Suzanne MacLeod, 53–61. London: Taylor & Francis, 2005.
- Giebelhausen, Michaela. "Museum Architecture: A Brief History." In *A Companion to Museum Studies*, edited by Sharon MacDonald, 223–244. Cornwall: Blackwell Pub., 2006
- Henning, Michelle. *Museums, Media And Cultural Theory. Issues in Cultural and Media Studies*. New York: McGraw-Hill Education, 2006.
- Hillier, Bill, and Kali Tzortzi. "Space Syntax: The Language of Museum Space." In *A Companion to Museum Studies*, edited by Sharon MacDonald, 282–301. Cornwall: Blackwell Pub., 2006

- Hooper-Greenhill, Eilean. "Studying Visitors." In *A Companion to Museum Studies*, edited by Sharon MacDonald, 362–376. Cornwall: Blackwell Pub., 2006
- Knell, Simon J. "The Shape of Things to Come: Museums in the Technological Landscape." *Museum and Society* 1 (3) (November): 132–146, 2003.
- Koh, Daveen. "Back to the Future." *The Cornell Daily Sun*. <http://cornellsun.com/node/53014>., 2012, accessed Mars 26, 2013
- Lampugnani, Vittorio Magnago. "Insight Versus Entertainment: Untimely Meditations on the Architecture of Twentieth-century Art Museums." In *A Companion to Museum Studies*, edited by Sharon MacDonald, 245–262. Cornwall: Blackwell Pub., 2006
- Lawson, Bryan. *The Language of Space*. Oxford: Architectural Press, 2001.
- MacDonald, Sharon, ed. *A Companion to Museum Studies*. Cornwall: Blackwell Pub., 2006a
- MacDonald, Sharon. "Expanding Museum Studies: An Introduction." In *A Companion to Museum Studies*, edited by Sharon MacDonald, 1–12. Cornwall: Blackwell Pub., 2006b
- Mason, Rhiannon. "Cultural Theory and Museum Studies." In *A Companion to Museum Studies*, edited by Sharon MacDonald, 17–32. Cornwall: Blackwell Pub., 2006
- Mctavish, Lianne. "Visiting the Virtual Museum: Art and Experience Online." In *New Museum Theory and Practice: An Introduction*, edited by Janet Marstine, 226–246. Malden: Blackwell Publishing, 2006.
- Moore, Gary T. "Environment Behavior Studies." In *Introduction to Architecture*, edited by James C. Snyder and Anthony J. Catanese, 46–71. New York: McGraw-Hill, 1979.
- Plaza, Beatriz, Silke N. Haarich, and Courtney Waldron. "Branding Bilbao: Assessing the Role of an Art Museum in Creating a Valuable Place Brand." *ACEI 2012 - 17th International Conference on Cultural Economics*, Doshisha University, Kyoto, 2012. http://www.jace.gr.jp/ACEI2012/usb_program/pdf/1.7.4.pdf., accessed February 03, 2013
- Psarra, Sophia. "Spatial Culture, Way-finding and the Educational Message. The Impact of Layout on the Spatial, Social and Educational Experiences of Visitors to Museums and Galleries." In *Reshaping Museum Space: Architecture, Design, Exhibitions*, edited by Suzanne MacLeod, 78–94. London: Taylor & Francis, 2005.
- Psarra, Sophia. *Architecture and Narrative: The Formation of Space and Cultural Meaning*. n.p.: Taylor & Francis e-Library, 2009.
- Rapoport, Amos. *The Meaning of the Built Environment: A Nonverbal Communication Approach*. Tucson: University of Arizona Press, 1990.
- Ravelli, Louise. *Museum Texts: Communication Frameworks*. n.p.: Taylor & Francis e-Library, 2006.
- Sfinteş, Anda-Ioana. "Muzeul Țăranului Român: Arhitectură, Discurs, Interpretare". Master thesis, Master Antropologie, Bucureşti: Şcoala Naţională de Studii Politice şi Administrative, 2012.
- Silverman, Lois H. *The Social Work of Museums*. n.p.: Taylor & Francis e-Library, 2009.
- Silverstone, Roger. "The Medium Is the Museum: On Objects and Logics in Times and Spaces." In *Towards the Museum of the Future: New European Perspectives*, edited by Roger S. Miles and Lauro Zavala, 161–176. n.p.: Taylor & Francis e-Library, 2002.
- Tzortzi, Kali. "Museum Building Design and Exhibition Layout: Patterns of Interaction." 6th International Space Syntax Symposium, Istanbul: 72.1–72.16. 2007. <http://www.spacesyntaxistanbul.itu.edu.tr/papers/longpapers/072 - Tzortzi.pdf>., accessed October 31, 2012
- Witcomb, Andrea. "Media Impacts." In *Re-Imagining the Museum: Beyond the Mausoleum*, 116–127. n.p.: Taylor & Francis e-Library, 2003.
- *** Museums in the Information Age: Evolution or Extinction? Leicester Exchanges Live Debate. Science Museum, London. <http://www.museumsassociation.org/video/07122012-museums-in-the-information-age>., 2012, accessed March 25, 2013
- *** "Taking Part 2012/13 Quarter 3: Statistical Release." Department for Culture, Media and Sport, 2013. <https://www.gov.uk/government/publications/taking-part-2012-13-quarter-3-statistical-release>., accessed April 04, 2013

A PROCESS OF SEAMLESSLY REPLACING CG ELEMENTS INTO LIVE-ACTION FOOTAGE

Jin Zhi

Duncan of Jordanstone College of Art & Design, United Kingdom

Abstract:

This research focused on inserting a computer-generated element into live-action footage and replacing unwanted existing objects in the footage. In addition, creating a realistic and seamless visual representation in the field of digital compositing. The purpose of this paper is to cover a detailed working process of digital compositing in order to clarify the production process and provide a clear idea for those entry-level artists to improve an overall understanding of the digital compositing and visual effects and hopefully inspire further collaboration and participants particularly between academia and industry.

Key Words: Digital compositing, CG elements, Realistic, Visual effects

Introduction

Along with the increase of digital compositing in feature films, the role of digital compositors has become of unprecedented importance. Compositors thus sit at the core of most image manipulation pipelines. In order to meet the needs of the visual effects industry, universities, art and film schools have established relevant professional courses in recent years. Nevertheless, current educational concepts still need to promote a sound scientific understanding in order to steer entry-level visual effects artists in the right direction.

This study has addressed several main factors from the perspective of theory and practice. As a general discussion, the research briefly outlines the historical perspectives of digital compositing, and continues to investigate the current situation of the digital compositing and visual effects industry. Subsequently as an experiment, a vintage film clip with complicated character movement was utilized. In the section concerning footage analysis, the research explores all possible working difficulties. In order to clarify this production process, a vintage film – *The Lavendar Hill mob* (1951) has been chosen for this experimental project, the researcher intends to integrate a CGI element into the film footage seamlessly. Through a series of practical aspects, the study covered analyses and discussions on analyzing footage, rotoscoping, producing clean plates, 3D camera tracking, creating CGI models, setting virtual lights, animating, rendering and final compositing, etc, and furthermore, the quality of the final project is intended not only for the purpose of demonstration, but also to reach a level of professionalism in line with current industry standards.

Definition Of Digital Compositing

Digital compositing is closely related to photomontage and the digital manipulation of multiple images (moving sequences or still images) into an integrated and seamless whole.

Photomontage is the process of cutting and joining photographic sources in order to recreate a series of photographic prints. The most famous photomontage *The Two Ways of Life* was created in 1857 [1]. Followed shortly by a photograph called “Fading Away” (H.P. Robinson) [2] that was produced by a photographer Henry Peach Robinson in 1858. These series of works actively set out to challenge the then-dominant medium of paint and other forms of photography. In today’s digital world, the traditional method of photomontage is realized through image editing software employed to reproduce a photo-realistic image. Instead of relying solely on a specific effect, a successful digital composite is able to make a particular shot look real. In other words, the goal of compositing is making all the elements (background plate, CGI elements, matte paintings, etc.) look like they were shot by the same camera under the same lighting conditions and existing in a same real world location.

Color images represent the full range of colors with red, green and blue components. A full range of the visible spectrum can be represented by using a combination of these three primary colors at different intensities, and moreover every single channel contains black and white data information [3]. Thus these channels can be thought of as monochrome images in their own channel. In digital images, every single pixel of each channel can be represented by an arbitrary number of bits. Typically each of these channels has 8 bits. The fourth channel, an additional channel which stores opacity of the color can be used as a matte (an alternative name for the matte is the ‘alpha channel’)[4]. The concept of the alpha channel was introduced by A.R.Smith in the late 1970s (Wikipedia, 2010) [5] and subsequently fully developed by Thomas Porter and Tom Duff in 1984 [6]. They used RGBA to indicate these four channels. Alpha has a value between 0 and 1. A value of 0 means that the pixel is transparent, but on the other hand, a value of 1 means that the pixel is 100 percent opaque. In mathematics, any number times 0 equals to zero and conversely, 1 multiplies any number will remain the result unchanged. In computer-generated images (CGI), if an image contains an alpha channel, it is usually referred to as a pre-multiplied image. For instance, if an image has value at 0.0,0.5,0.0,0.5 (R,G,B and Alpha), it implies a half green image with a 50% transparent alpha channel on it. If the color were fully green, its RGB channels would be (0.0,1,0.0) [7]. A pre-multiplied image is one where the RGB channels have been multiplied by the alpha channel. This means it is a four-channel image in which the red, green and blue channels have already been multiplied by the integrated matte (alpha) channel. The R.G.B channels of a pre-multiplied image will never have a higher value than its alpha channel.

Research Methods And Findings

The goal of this project is to explore a systematic process of inserting CGI elements into digital moving images, with the aim of producing a seamless and photo-realistic composite.

Project Concept

The project will explore and examine issues from erasing unwanted parts from the footage and making clean plates, to the entire complex process of 2d tracking, 3d camera match moving, building 3d models, multi-pass rendering and integrating CGI elements into the original footage in order to seamlessly replace one existing object in the footage. As an example, one object on the original footage will be completely removed and replaced with a CG coke glass bottle by the end of this project. Through a detailed discussion and demonstration, the study will clearly reveal a series of processes undertaken to replace objects and insert CG elements in feature films.

Footage Analysis

The overall length of this project is approximately in 16 seconds (00:00:16:00) and it was captured as a PAL TV broadcast format. The following image (Fig 1.) shows one frame of the entire clip. The object - toy Eiffel Tower in the man’s hand will be replaced by a CG coke bottle.



Figure 1. One snapped image at frame 150.

Working range of this project is 55 frames long (frame 130 – frame 185), however, the difficulties on this part are fairly challenging. Fig 2. Indicates a number of unwanted objects, which have been marked at frame 165. Creating a clean plate for the luggage case (inside part). This is because the character grabs one toy from the opened brief case. A clean empty space is needed after erasing the toy from its original position. In addition, several places also need clean plates, such as

shadow of the toy on the wall and the man's chest when he was passing the toys from one person to another. Fig 3. Shows these unwanted objects in three different frames.



Figure 2. Shows a series of key frames between frame 130 and 185.



Figure 3. Shows a number of unwanted objects, which need to be removed in order of create a clean plate.

Creating Mattes and Clean Plates

Creating mattes and clean plates is a tedious process, however, mattes and clean plates are essential elements in digital compositing, especially for producing a photo-realistic composite. Rotoscoping [8] is one of the main tasks in this project, which will directly determine the success of the final result. Rotoscoping started from frame 130 and finished at 185 frames for most objects in the scene. The compositor had to repeatedly roto every object, which overlapped with the target frame by frame within frame 130 to 185. Because the position of the target is in a medium shot, it therefore required highly accurate mattes, otherwise the final composite would reveal the edges of the original background around the mattes. Fig 4. Shows a pair of finished mattes.



Figure 4. Shows two finished mattes, which are used for bringing objects back to their original position after compositing.

The clean plates were created using a “Time Offset” method. The concept of this method is to take advantage of the time offset between each movement of the objects. For instance, if an object moves from position A to B from frame 100 to 150, it means the position of frame 100 is empty and clean when the object has moved to frame 150. Fig 5. Shows the result of creating clean patches (Left) and a completed clean plate (Right)



Figure 5. Shows the result of creating clean patches (Left) and a completed clean plate (Right) at frame 150.

3D Camera Tracking

3D camera tracking is an important section. CG elements need to have the same camera lens distortion when composite with live-action footage in order to look as they were shot by the exactly the same camera at the same time [9]. In 3D tracking section, the original plate needs to be undistorted before the 3D matching tasks. See Fig 6.

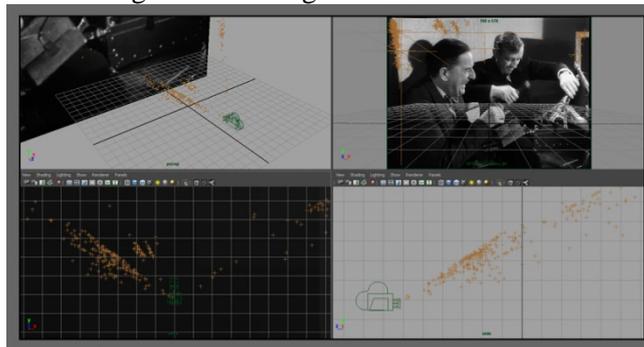


Figure 6. Shows a match moving scene with 3D tracking data.

Creating CG Elements

The footage - *The Lavender Hill Mob* was filmed in 1951. If a CG coke bottle is replaced into the scene, we need to ensure that the shape of CG glass bottle matches the style of the coke bottle in 1951. Such details are fairly important in compositing. In some cases a small detail can bring an entire scene back to life. The name of coke bottle from 1951 to 1958 was called 'Hobbleskirt Coke, 6 FL OZ' [10], it was embossed with the Coca-Cola logo and capped with a white and red cap. See Fig 7.



Figure 7. Shows the style of the coke bottle in 1951, and a CG coke bottle model (Right) based on this bottle.

Multi-pass Rendering

Caustics

Caustics occur on transparent objects. When light rays pass through a transparent object, due to the principle of physical optics, light rays will be enveloped by the curvature of the surface of this object and projected onto another surface [11]. In computer graphics, caustics effect is accomplished by ray-tracing attributes through Mental Ray rendering engine of 3D applications, such as Maya. See Fig 8.

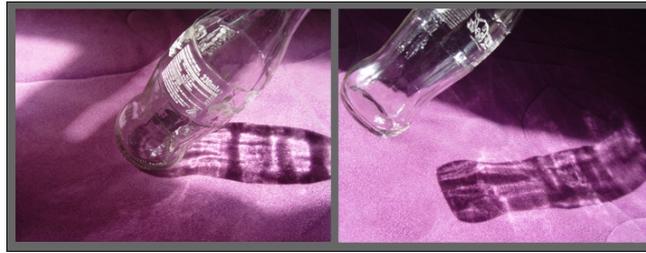


Figure 8. Shows a comparison of examining caustics from a real glass coke bottle.

In 3D software, caustics rendering is based on the luminance value (luminance value must be higher than 0) of the shader attributes of the dummy CG object, these receive caustics from transparent objects. The issue is that if the caustics are to be received by CG surfaces, the luminance value of the CG surface must be 0. In compositing, 0 means completely transparent, any value above 0 means semi-transparent and 1 means 100% opaque, so if the value is 0 the rendered caustics can be composited with live-action footage in digital compositing. However, there is a known issue with Maya the mental ray engine, which is not fully compatible with Maya nodes.

Test one: Test one was performed by applying a “Surface Shader” (the “Surface Shader” uses Maya default shading algorithms which black the object with a zero value) to the surface of a test CG object, then position this object behind a transparent character, the result of rendering was a fully black image without caustics cast on it.

Test two: In test two, a “Background Shader” was applied to the same surface of this testing CG objects (Background is a Maya shader, it leaves a transparent area around the outside of the test objects). It was assumed that this method would provide a result, which only rendered the caustics area and rest of the parts with a value of 0, so it could be used for compositing purposes. However, the result was not successful and good enough for digital compositing.

Solution: The solution for rendering an image with successful caustics on was to change the “Color” and “Weight” attributes of the shader, which is used for receiving caustic objects. In addition, the value of caustic Phones was increased to 3000000, photon intensity was increased to 800000 and the photon exponent value to was reduced to 0.01. The following image Fig 9. Shows a comparison result of creating caustics. Fig 10. Indicates an acceptable value of “Color” and “Weight” in Maya render attributes.

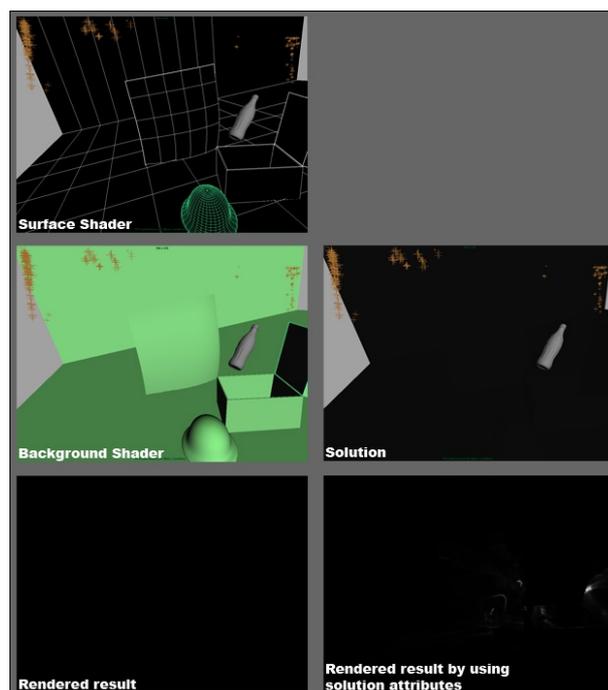


Figure 9. Shows a comparison result of creating caustics.

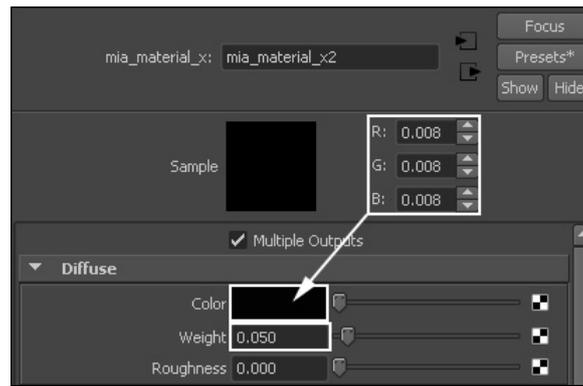


Figure 10.Multi-pass rendering

Multi-pass images give digital compositors a great deal of control over the look of the shot. Compositors can make an overall improvement to each pass as well as adjust the final composite more artistically. In addition, compositors can take advantage of a specific pass by interactively relighting the scene, adding motion blur or even retexturing the CG elements without spending extra time re-rendering them in the 3D software.

The most often used multi-passes are: Beauty pass, diffuse pass, specular pass, reflection, refraction, shadow pass, depth pass, normal pass, ambient occlusion pass, ID pass, motion vector, etc. These separated render passes can be seamlessly combined together later on through the use of mathematic formulas in digital compositing software. See Fig 11.

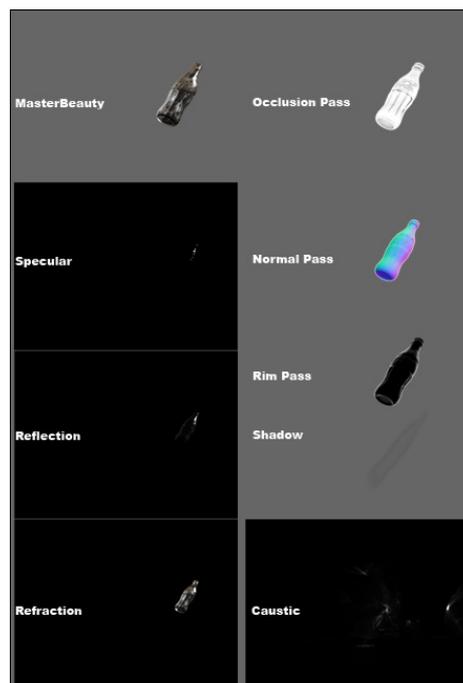


Figure 11. Shows a series of multipass images.

Final Compositing

Once all the multi-pass images have been rendered compositors can make an overall improvement to each pass as well as adjust the final composite more artistically. With a full understanding of mathematical formulas of digital compositing and a strong aesthetic ability, a seamless composite will be created. Advanced compositing effects can be achieved by using an alpha channel with a series of math operations. The most common compositing operation is to combine image A (foreground) and B (Background) through a number of math formulas. However, at this stage of combining multi-passes the main operation formulas are: Over, Plus, Multiply [12].

Over operator

$$\text{Over} = (A \times B) + [(1 - M) \times B] \quad (1)$$

Over operator is the most often used formula in digital compositing. It takes image A on top of image B, by using a matte of image A.

First of all, the foreground image A is multiplied by its matte image, this allows the matte of A to cut out image A. As mentioned in previous sections, in digital compositing, black implies 0 and white is represented as 1. This step will produce an image with black areas around it. Once this cutout has been created the result is then added on the top of image B, which is an inverted matte of A multiplies image B (Background image).

Plus operator

$$\text{Plus} = A(r,g,b) + B(r,g,b) \quad (2)$$

It simply added two images together. The order in which two images are asymmetrical, and can be switched over. This operator usually uses for adding reflection pass, refraction pass, specular pass, etc. elements together.

Multiply operator Multiply operation will mostly use for Occlusion pass, which multiplied by a combination of other multi-pass images. The result can enrich soft shadow of CG objects and enhance the look of realism.

$$\text{Multiply} = A \times B \quad (3)$$



Figure 12. A comparison of original footage and the final composition.

Conclusion

This experimental project focused on the process of replacing an object with a CG object for feature films. Instead of defining a series of complicated theories the compositor attempted to cover a number of key steps through a piece of challenging vintage film material. The difficulties involved in using this footage are comparable to real VFX industry scenarios.

By translating a series of complex theories into a practical study through one challenging real world example of digital compositing, the basic structure of the operating procedures of digital compositing have been laid out and a series of experimental results have been demonstrated. Although this study is finished, the technique of digital compositing and visual effects is still developing, and more realistic and breathtaking visual effects will be presented in the future in feature films. With the

aim of creating ever more lifelike visual effects the field of digital compositing will continually evolve in this way.

References:

- [1] B rinkmann, R. *The Art and Science of Digital Compositing*, SanDiego: Morgan Kaufmann, 2008, pp. 4-6.
- [2] Annette Labedzki. *His Most Famous Photograph (Fading Away)*.
- [3] Wallace, B.A. "Merging and Transformation of Raster Images for Cartoon Animation", Proc. ACM SIGGRAPH '81, Vol 15, No. 3, 1981, ACM, pp. 253-262.
- [4] Wikupedia Foundation, Inc., 2010. http://en.wikipedia.org/wiki/Alpha_compositing
- [5] Wikupedia Foundation, Inc., 2010. http://en.wikipedia.org/wiki/Alvy_Ray_Smith
- [6] Thomas Porter, Tom Duff. *Compositing Digital Images*. Computer Graphics. Vol. 18, Number3. 1984.
- [7] Wikupedia Foundation, Inc., 2010. http://en.wikipedia.org/wiki/Alpha_compositing
- [8] Wikupedia Foundation, Inc., 2010. <http://en.wikipedia.org/wiki/Rotoscoping>
- [9] Steve Wright. *Digital Compositing for Film and Video*, Focal Press, 2010, pp.197-199
- [10] The Cocacola Company. *Hobbleskirt Coke, 6 FL OZ" 1951-1958*
<http://www.flickr.com/photos/60541684@N00/268751126>
- [11] Wikupedia Foundation, Inc., 2011. http://en.wikipedia.org/wiki/Caustic_%28optics%29
- [12] K. Elissa, "Title of paper if known," unpublished.

SCADA LIVE FORENSICS: REAL TIME DATA ACQUISITION PROCESS TO DETECT, PREVENT OR EVALUATE CRITICAL SITUATIONS

Pedro Taveras N., MSC.

Pontificia Universidad Católica Madre y Maestra, Dominican Republic

Abstract:

SCADA (Supervisory Control and Data Acquisition System) systems were originally created to be deployed in non-networked environments. Therefore they lack of adequate security against Internet-based threats and cyber-related forensics. In recent years, SCADA systems have undergone a series of changes that might increase the risks to which they are exposed. Among these risks it can be observed that its increased connectivity may permit remote controls over the Internet, or the incorporation of general purpose tools, thus incorporating already known vulnerabilities of these. Any cyber-attack against SCADA systems demands forensic investigation to understand the cause and effects of the intrusion or disruption on such systems. However, a SCADA system has a critical requirement of being continuously operational and therefore a forensic investigator cannot turn off the SCADA system for data acquisition and analysis. This paper leads to the creation of a high level software application capable of detecting critical situations like abnormal changes of sensor reads, illegal penetrations, failures, physical memory content and abnormal traffic over the communication channel. One of the main challenges is to achieve the development of a tool that has minimal impact over the SCADA resources, during the data acquisition process.

Key Words: Cyber security, SCADA, Live Forensics, Intrusion Detection

Introduction

The security of SCADA systems is especially relevant in the field of Critical Infrastructure. A failure of critical infrastructure could have direct impact for society to the extent of affecting entire nations and their environment.

Any government network infrastructure or industrial based SCADA (Supervisory Control and Data Acquisition) or DCS (Distributed Control Systems), designed to automate, monitor and control critical physical processes, including manufacturing and testing, electric transmission, fuel and water transport, is subject to potential attacks.

Supervisory Control and Data Acquisition comprise all application solutions that collect measurements and operational data from locally and remotely controlled equipment. The data is processed to determine if the values are within tolerance levels and, if necessary, take corrective action to maintain stability and control. Its basic architecture comprises a centralized server or server farm, RTU (Remote Terminal Units) or PLC (Programmable Logic Controller) to manage devices; consoles from which operators monitor and control equipment and machinery.

SCADA systems were originally created to be deployed in non-networked environments. Therefore they lack of adequate security against Internet-based threats and cyber-related forensics.

Most industrial plants now employ networked process historian servers for storing process data and other possible business and process interfaces. The adoption of Ethernet and transmission control protocol/ Internet protocol TCP/IP for process control networks and wireless technologies such as IEEE 802.x and Bluetooth has further reduced the isolation of SCADA networks (Zhu, Anthony & Sastry, 2011).

In recent years, SCADA systems have undergone a series of changes that might increase the risks to which they are exposed. Among these risks it can be observed that its increased connectivity may permit remote controls over the Internet, or the incorporation of general purpose tools, thus incorporating already known vulnerabilities of these.

SCADA systems, in particular, perform vital functions in national critical infrastructures, such as electric power distribution, oil and natural gas distribution, water and waste-water treatment, and transportation systems. They are also at the core of health-care devices, weapons systems, and transportation management. The disruption of these control systems could have a significant impact on public health, safety and lead to large economic losses (Cardenas, Amin, Huang, Lin & Sastry 2011).

As a consequence, there is an increasing interest in the security/forensic research community on SCADA systems. This is mostly due to the heightened focus of governments worldwide on protecting their critical infrastructures, including SCADA systems (Ahmed, Obermeier & Naedele, David, Chaugule & Campbell, 2012).

Securing SCADA systems is a critical aspect of Smartgrid security. As sophisticated attacks continue to target industrial systems, the focus should be on planning and developing new security techniques that will adapt to the SCADA environment and protocols (Rodrigues, Best & Pendse, 2011).

Immediate needs identified in this area include the collection of evidence in the absence of persistent memory, hardware-based capture devices for control systems network audit trails, honeypots for control systems as part of the investigatory process, radio frequency forensics and intrusion detection systems for SCADA control systems (Nance, Hay & Bishop, 2009). However, post-mortem analysis tools require the investigator to shut down the system to inspect the contents of disks and identify artifacts of interest. This process breaks network connections and unmounts encrypted disks causing significant loss of potential evidence and possible disruption of critical systems (Chan & Venkataraman, 2010).

Computer forensics relies on log events for searching evidence of a security incident. However, the massive amounts of generated events along a lack of standardize logs complicate the analyst tasks (Herrerias & Gomez, 2007).

Digital forensics investigators are experiencing an increase in both the number and complexity of cases that require their attention. Most current digital forensic tools are designed to run on a single workstation, with the investigator issuing queries against copies of the acquired data evidence. With current generation tools, the single workstation models works reasonably well and allows tolerable case turnaround times for small forensic targets (for example < 40GB). For much larger targets, these tools are too slow to provide acceptable turnaround times (Richard & Roussey, 2006).

The challenge, however, is to mitigate the vulnerabilities that occur once a networked device becomes accessible from the internet. Attacks ranging from DDoS to backdoor intrusion are possible on industrial networks and power and SCADA systems. Although network firewalls can stop a significant amount of malicious traffic, there are several techniques hackers can use to bypass these security devices. The complexity of the infrastructure can make it difficult to detect malicious behavior (Rodrigues *et al.*, 2011).

Research Problem

Any cyber-attack against SCADA systems demands forensic investigation to understand the cause and effects of the intrusion or disruption on such systems. However, a SCADA system has a critical requirement of being continuously operational and therefore a forensic investigator cannot turn off the SCADA system for data acquisition and analysis. Current forensic tools are limited by their inability to preserve the hardware and software state of a system during investigation.

Research Goal and Target

Process control systems (SCADA Systems) generated much discussion as an area that the security community recognizes as a security threat, but not yet perceived by industry to be as much of a threat. As a result, this field lags behind most technical fields in the area of security (Nance, Hay & Bishop, 2009).

Study and research security vulnerabilities related to networked Supervisory Control and Data Acquisition (SCADA) systems, in order to *develop a forensic computing model to support incident response and digital evidence collection process*. Forensic investigation can play a vital role in a protection strategy for SCADA systems and may assist in the prosecution of attackers, but also in a deep analysis of the underlying SCADA IT system, for example, in the case of non-malicious events such as malfunctioning hard disks or other hardware. However the critical nature of SCADA systems

and the 24/7 availability requirement entails forensic investigators spending as little time on a live SCADA system as possible, necessarily performing live data acquisition and then subsequent offline analysis of the acquired data (Ahmed *et al.* 2010).

Relevance and Significance

In the last years there has been an increasing interest in the security of process control and SCADA systems. Furthermore, recent computer attacks such as the Stuxnet worm, have shown there are parties with the motivation and resources to effectively attack control systems (Cardenas *et al.*, 2011)

SCADA systems are deeply ingrained in the fabric of critical infrastructure sectors. These computerized real-time process control systems, over geographically dispersed continuous distribution operations, are increasingly subject to serious damage and disruption by cyber means due to their standardization and connectivity to other networks (Zhu & Anthony, 2011).

In recent times it has been noticed that hackers implement newer techniques to launch attacks that can evade traditional security devices. It is therefore important to secure the SCADA systems from process related threats (Rodrigues *et al.*, 2011).

Compromising such a system with intrusion attacks can lead not only to high financial losses but, more importantly, to the endangerment of public safety. The danger is even higher considering that critical infrastructures are not immune to these threats and that they may be potentially more vulnerable than common information technology systems. Hence intrusion protection for critical infrastructures is an obvious need (Linda, Vollmer & Manic, 2009).

Reliability of many SCADA systems is not only dependent on safety, but also on security. Recent attacks against SCADA systems, by sophisticated malware, demands forensic investigation to understand the cause and effects of the intrusion on such systems so that their cyber defense can be improved.

A SCADA system has a critical requirement of being continuously operational and therefore a forensic investigator cannot turn off the SCADA system for data acquisition and analysis. In this case, live forensics is a viable solution for digital investigation in SCADA systems (Ahmed *et al.*, 2012).

In real life, logs are rarely processed by stakeholders due to 1) the large number of entries generated daily by systems and 2) a general lack of security skills and resources (time) (Hadziosmanovic *et al.*, 2011). However, the use of the classical post-mortem analysis approach is becoming problematic especially for large-scale investigations involving a network of computers. In addition, the amount of time available for processing this data is often limited (Su & Wang, 2011).

Review of Literature

A substantial body of research exists in the area of forensics models for live acquisition over SCADA systems. Related research work is discussed on this section.

There is a growing need for systems that allow not only the detection of complex attacks, but after the fact understanding of what happened (Tang & Daniels, 2012).

Several researches address threats in SCADA systems. For the identification of threats, authors typically use questionnaires and interviews. To detect anomalous behavior, authors use approaches based on inspecting network traffic, validating protocol specifications and analyzing data readings. Process-related attacks typically cannot be detected by observing network traffic or protocol specifications in the system. To detect such attacks one needs to analyze data passing through the system, and include a semantic understanding of user actions (Hadziosmanovic *et al.*, 2011).

A group of researchers who met at the Colloquium for Information Systems Security Education (CISSE 2008) to brainstorm ideas for the development of a research for Digital Forensic, concluded that actual SCADA systems are potentially more vulnerable to attack and more likely to need associated digital forensics capabilities. Unfortunately, most process control systems were not built to track their processes, but merely to control them. As a result, significant research and development categories were identified under this area, including among the most important: mechanism form the collection of evidence in the absence of persistent Memory and hardware-based capture devices for control (Nance, 2009). Figure 1 shows the list of topics in need for further development. It can be noticed that areas for Live Acquisition and Control Systems.

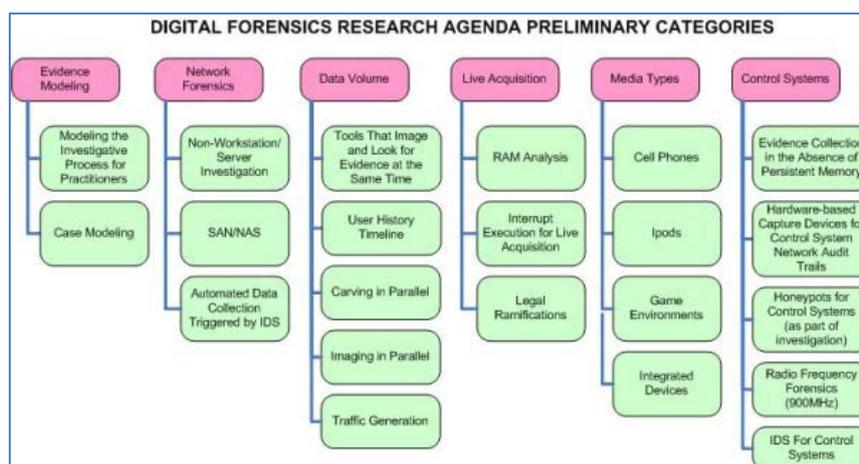


Figure 1: Research Topics for Digital Forensics

Chen & Abu-Nimeh (2011), developed a deep research over the case of Suxnet malware. According their report, this was the first malware written exclusively to attack SCADA platform.

Stuxnet experience has shown that isolation from the Internet isn't an effective defense, and an extremely motivated attacker might have an unexpected combination of inside knowledge, advanced skills, and vast resources. Existing technologies would have difficulty defending against this caliber of attack (Chen & Abu-Nimeh, 2011). Therefore the need of new forensics methods that goes beyond the traditional prevention mechanism.

Ahmed, Obermeier & Naedele, David, Chaugule & Campbell (2012), propose a forensic mechanism denominated Live Forensics as a viable solution for SCADA systems. Live data acquisition involves acquiring both volatile data (such as the contents of physical memory) and non-volatile data (such as data stored on a hard disk). It is different from traditional dead disk acquisition, which involves bringing the system offline before the acquisition, where all volatile data is lost.

However, despite the importance of live data acquisition, it is still unclear how contemporary live data acquisition tools should be run on a SCADA system so that they minimize risk to SCADA system services (Ahmed *et al.*, 2011).

Aldenstein (2006) established that the possibility of implementing live forensics over SCADA systems relies on the capability of the operating system to provide the list of running processes. Therefore, he recognized the need for tools capable of examining the raw memory of a machine. These tools are analogous to the static tools that open the raw disk device and impose the file system structure on it to extract files, directories, and metadata (Adelstein, 2006).

Sutherland *et al* (2008), performed exploratory studies for live forensics within Windows operating systems environment and also determined the need for more invasive tools that allows better access to information related to memory, network and system activity were assessed to determine the impact on the file system, system registry, memory and the usage of DLLs.

Hadziosmanovic *et al.* (2011) proposed a tool-assisted approach to address process related threats. They presented an experimental study where SCADA threats that unlikely to happen or that does not occur on a systematic manner are detected and logged for investigation. An example could be when an attacker manages to get valid user credentials and performs disruptive actions against the process. However this effort was limited to post-mortem log analysis containing data for single event operations and does cover anomalous command process sequences. Likewise, it was determined that an attacker might gain unauthenticated remote access to devices and change their data set points. This can cause devices to fail at a very low threshold value or an alarm not to go off when it should. Another possibility is that the attacker, after gaining unauthenticated access, could change the operator display values so that when an alarm actually goes off, the human operator is unaware of it. This could delay the human response to an emergency which might adversely affect the safety of people in the vicinity of the plant (Zhu & Anthony, 2011).

SCADA systems are increasingly commonly being attached to networks, and typically offer no persistent storage for logging of network activity. The challenge for the digital forensic research

community is to develop methods to allow an investigator to determine how these devices interacted with the network during a time period of interest (Nance, Hay & Bishop, 2009).

There is continuing interest in researching generic security architectures and strategies for managing SCADA and process control systems. Documentation from various countries on IT security does now begin to recommend recommendations for security controls for (federal) information systems which include connected process control systems. Little or no work exists in the public domain which takes a big picture approach to the issue of developing a generic or generalizable approach to SCADA and process control system forensics (Sly & Stinikova, 2009).

Collection of adequate records or logs of events that happened near incident time is crucial for successful investigation. Logging capabilities of SCADA systems are geared towards discovering and diagnosing process disturbances, not security incidents, and are thus often not adequate for forensic investigation (Fabro & Cornelius, 2008)

Kilpatrick *et al* (2008) developed an architecture based on the Modbus TCP (Transmission Control Protocol) using two control devices and one HMI (Human Machine Interface) station. This architecture comprised two agents and a central warehouse. Various agents were deployed over the SCADA network. These agents captured network traffic containing real time data and stored them into the warehouse. Relational databases query mechanisms were used in the event of a forensic investigation. However, Ahmed *et al* (2011) established that state of the art forensic analysis tools do not support the unique features of diverse SCADA environments, which include supporting SCADA protocols and numerous SCADA applications' proprietary log formats etc. Thus plugins or modules for contemporary forensic tools need to be developed to augment the forensic analysis in SCADA systems.

Nehimbe & Nehibe (2012) proposed a timed series methodology to analyze forensic logs. During their research they concluded that actual for forensic tools may not necessarily generate the needed results. Due to two basic limitations on these tools: Some of them only have recovery and imaging capabilities and some intrusion analysis tools are flawed in terms of how they analyze intrusion logs.

Hunt & Slay (2010) proposed an approach named security information event management (SIEM) with the purpose to provide a tool that allows any networked system to auto adapt itself based on forensic logging. Their works showed that a SIEM system is an ideal point at which to store log data emanating from security devices and the point at which forensic logging needs to occur. However, although they were able to achieve the implementation of forensically sound log files in some systems; their approach is by no means universal. They concluded that their works still falls short of addressing the core domain of real-time forensically sound adaptive security.

With the purpose of rebuilding an attack scenario Herreria & Gomez (2007), proposed a log correlation model to support the evidence search process in a forensic investigation. In this work, they proposed a system composed by a set of agents in order to collect, filter, and to normalize events coming from diverse log files. Events may come from systems logs, application logs, and security logs. Once events are joined together in the same place and under the same format, they are sent to a correlation engine. The engine compares and processes the events in a global fashion in order to follow all actions taken by the attacker (Herrerias & Gomez, 2007).

Su & Wang (2011) developed a formula using probability theory and mathematical statistics to quantitatively calculate the degree of memory change on a live system. Their conclusions states that since the live memory state frequently changes is natural limitation for the purpose of live forensics. In their experiments they tried to restore to the same system state each time, however, the real state has been changed after one or two seconds. Therefore, they were only able determine and approximate of the system memory in every repeated process.

Further work is required to assess tools over various operating systems. This would be of value to the forensic investigator, but the way memory is handled and its analysis varies greatly between Windows Service packs let alone other operating systems; as a result the area of memory forensics is deeply complex and requires a significant amount of time and effort invested by the forensic examiner to begin to comprehend how memory works in modern Operating Systems (Sutherland, Evans, Tryfonas & Blyth, 2008).

Other research approaches are related to autonomic attack detection and response. Cardenas *et al.* (2011) showed that by incorporating a physical model of the system they were able to identify the

most critical sensors and attacks. They also proposed the use of automatic response mechanisms based on estimates of the state of the system. However, they concluded that this methodology might be problematic, especially when the response to a false alarm is costly (Which could be the case for SCADA environment). As a result their model should be considered as a temporary solution before a human investigates the alarm.

Approach

While there have been a good number of research efforts investigating the suitability of forensics mechanism for SCADA, this work would be different in that it is intended to develop an investigation into *what is required to develop a forensic computing model to support incident response and digital evidence collection process, without interfering with the “always running” condition of SCADA platforms.* The intended approach is an extension of the works from Ahmed *et al.* (2011), who proposed a forensic mechanism denominated Live Forensics as a viable solution for SCADA systems. Live data acquisition involves acquiring both volatile data (such as the contents of physical memory) and non-volatile data (such as data stored on a hard disk). It is different from traditional dead disk acquisition, which involves bringing the system offline before the acquisition, where all volatile data is lost.

From a forensic perspective, a SCADA system can be viewed in different layers based on the connectivity of the various SCADA components and their network connectivity with other networks such as the Internet.

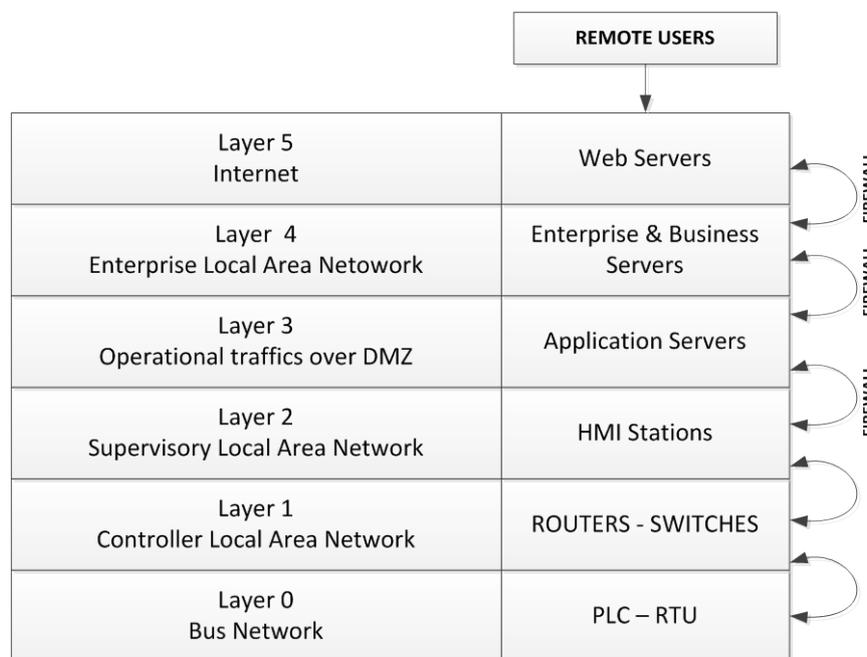


Figure 2: Layer Model for SCADA

The lowest layer represents the physical elements designed to interact directly with the industrial hardware or machinery. These devices are connected via bus network. Layer 1 receives electrical input signals which are decoded as a bit streams over a standard network protocols. The result is transferred to the upper layers for analysis and controlling response. Layer 3 and above, represents the enterprise network which is also interconnected to the Supervisory systems. At this stage all traffic containing database content and applications supporting the business logics for the operation is managed. As stated by Amehd *et al.*, (2010), live forensic analysis for the SCADA system must focus on the Layers 0, 1 and 2.

The initial approach has the intention of developing a forensic watch dog by means of a finite state automaton that would function as an agent that is constantly listening SCADA events. When any particular event is sensed, the input values are read and compared to a set of predefined rules in order to decide the change of state. Figure 2 represents the proposed automaton model.

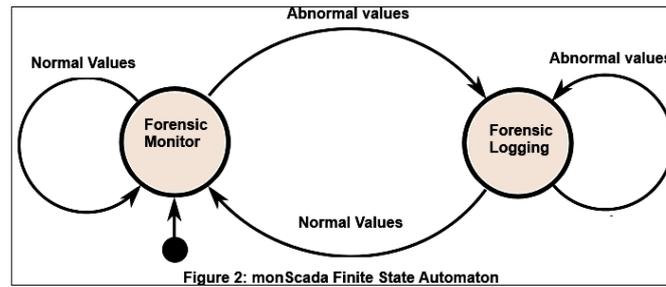


Figure 2: monScada Finite State Automaton

Figure 3: Finite State Automaton as SCADA Live Monitor

This two state automaton or agent constantly monitors the state of the SCADA system, including measures from: System Variables, Sensor Tags, Network Traffic and Command Executions. These values are checked against a set of behavioral rules. If a read is detected to be above the normal range, the agent automatically switches to Forensic mode and initiate the logging of forensic information. Ideally it would need a separate backup system to continuously dump the abnormal lectures from the SCADA tags and creates a record of this event appending all the information available about the system state. Including, but not limited to: CPU load, sensor names, sensor values, state of the physical memory, state of the virtual memory, state of the networking variables, state of mounted disk and network drives, list of active process in memory including name, executable name, working directory, command line, user name, user ids and group ids, threads, connections, file descriptors, etc. The information logged would be exclusively related to the period of time of the anomaly. Once the system start reading normal values, it switches back to Forensic Monitor and stop logging.

Normally a SCADA system reads every sensor or control registry on the system. These registries are known as tags and the logging frequency can vary from a read every second or even every 300 milliseconds. A typical SCADA system can have up to 40,000 tags. A system of such magnitude can generate approximately 400GB of data for every 24 hour period of operation. This calculation is based estimating an average size of 120 bytes per record. It can be seen, that the live acquisition is just part of the challenge. Dealing with vast amounts of data needs to be considered. These volumes requires manipulation by means of database query processors and moreover, require a fast capture and writing process that must be able to: (1) keep up with the logging process at the same time that new data comes into the system; and (2) all this must be done without incorporating additional workload over the monitored SCADA system. In other words, needs to be accomplished in a non-invasive manner.

Another challenge imposed to the intended solution is that SCADA system components can be found running on legacy hardware and operating systems. In such cases, a SCADA system provides limited system resources for data acquisition and therefore demands lightweight data acquisition tools and the gathering process might not represent a large resource consumer. However the data conversion process, if a relational database would be implemented for a better data analysis, would requires specialized hardware to reduce the processing time and speed up files conversion and querying processes.

Another important aspect from this experiment that would be suitable for further development is that current forensic analysis tools do not provide a standard support for the variety of SCADA hardware versions, protocols and log formats. Therefore, we have an interesting opportunity to expand this experiment with the development of plugins and applications and interface layers in order to increase the number of SCADA forensic tools as an expansion of the works of Hadziosmanovic et al., (2011), who stated that despite the fact that there are several vendors, system architectures in various SCADA systems are similar and the terminology is interchangeable.

In conclusion, future work leads to the creation of a high level software application capable of detecting critical situations like abnormal changes of sensor reads, illegal penetrations, failures, physical memory content and abnormal traffic over the communication channel. One of the main challenges is to achieve the development of a tool that has minimal impact over the SCADA resources, during the data acquisition process. In previous exercises it was observed that the processes for acquiring low level information, such as processes or memory information does not represents an

extensive load on the actual system that is processing the task. However, it is expected that the amount of demanded resources increase as the number of SCADA tags and the frequency of logging increases. Therefore on real live SCADA system, the acquisition process could be competing for resources that should be available for the normal operation of the SCADA systems.

Barriers and Resources

From the literature review it can be determined multiple barriers and issues that could be anticipated for the development of this work. The next section presents a summary of the know limitations and challenges determined by previous research efforts on this field.

Forensic data gathered from a live system can provide evidence that is not available in a static disk image. Live forensics also operates with different constraints—specifically, the evidence gathered represents a snapshot of a dynamic system that *cannot* be reproduced at a later date. Standards for acceptance are evolving, and legal precedents are still being established (Adelstein, 2006).

Given that volatile data in a running system changes continuously, Ahmed et al., (2011) established two main challenges involved during live data acquisition over operational SCADA systems: (1) Live data acquisition needs to be performed as quickly as possible after an incident in order to capture any traces of the incident on volatile data before the processes or services on the running system overwrite useful data; (2) Cryptographic hash of the actual evidence on the compromised system and its acquired copy, which is used for all the examination and analysis. If, however, the compromised system remains live, the state of the data may change between the copying and the hash calculation, rendering hashing ineffective as an integrity check (Ahmed *et al.*, 2011).

From a forensic standpoint, modifying the original system memory state is unavoidable, therefore, it is needed that changing as little as possible on the process of collecting live forensics. For a real live forensics case, it should content digital forensics that collected by the forensic tools, the analysis and evaluation of the uncertainty. However, it is difficult to measure how much of the volatile memory is modified by a forensics tool. Moreover, it is difficult (if not impossible) to calculate the extent of the memory's impact caused by a running process on the volatile memory. So, measuring the extent of the volatile memory changes caused by running a live forensic tool becomes more and more important (Su & Wan, 2011).

Because the architecture of production operating systems prevents applications from accessing kernel memory and storage devices without using the kernel, kernel-based rootkits will always be a threat to live analysis. Future directions in live analysis techniques involve the use of specialized hardware to collect the raw memory and storage data for a dead analysis (Carrier, 2006).

Research has shown that an attacker with control of the target system can manipulate memory mappings so that the CPU and devices on the PCI or Firewire buses don't necessarily get the same view of memory. In such cases, attempts to acquire the memory's contents could crash the target system or enable the attacker to mask sections of memory without that masking being apparent to the investigator (Hay, Bishop & Nance, 2009).

Furthermore, factors like the continuous availability demand, time-criticality, constrained computation resources on edge devices, large physical base, wide interface between digital and analog signals, social acceptance including cost effectiveness and user reluctance to change, legacy issues and so on make SCADA system a peculiar security engineering task (Zhu & Anthony, 2011).

Finally, no matter how well the simulations and models emulate a possible solution, any given conclusion needs to be tested over a real SCADA systems. Real SCADA systems are expensive to build and thus require significant research funding. Access to sample information or security failure scenarios could be difficult because the critical nature of SCADA systems demands the owners and operators not share any information about their system.

Conclusion and Expected Contributions

Applying traditional information security mechanism directly to SCADA systems is not possible. SCADA systems cannot afford non-deterministic delays in performance, security controls that require a lot of memory, block access for safety or relatively long intermediate processes. Security measures that can be applied to SCADA systems should consider this special operating paradigm.

Process-related attacks typically cannot be detected by observing network traffic or protocol specifications in the system. To detect such attacks one needs to analyze data passing through the system, and include a semantic understanding of user actions (Hadziosmanovic *et al.*, 2011).

In conclusion, there is no generic model for understanding the forensic computing processes necessary to gather digital evidence from Process Control and SCADA systems. Therefore, the need for developing a forensic computing model to support incident response and digital evidence collection process is justified.

As a consequence this work could help to improve critical infrastructure protection and provide appropriate tools that could be used for dealing with incident responses and forensics analysis over interconnected SCADA systems.

References:

- Adelstein, F. (2006). Live forensics: diagnosing your system without killing it first. *Commun. ACM*, 49(2), 63–66.
- Ahmed, I., Obermeier, S., Naedele, M., & III, G. R. (5555). SCADA systems: Challenges for forensic investigators. *Computer*, 99(1), 1.
- Andersson, G., Esfahani, P. M., Vrakopoulou, M., Margellos, K., Lygeros, J., Teixeira, A., ... Johansson, K. H. (2012). Cyber-security of SCADA systems. In *Innovative Smart Grid Technologies, IEEE PES* (Vol. 0, pp. 1–2). Los Alamitos, CA, USA: IEEE Computer Society.
- Boca, L., Croitoru, B., & Risteiu, M. (2010). Monitoring approach of Supervisory Control and Data Acquisition downloadable data files for mission critical situations detection. In *International Conference on Automation, Quality and Testing, Robotics* (Vol. 3, pp. 1–6). Los Alamitos, CA, USA: IEEE Computer Society.
- Cárdenas, A. A., Amin, S., Lin, Z.-S., Huang, Y.-L., Huang, C.-Y., & Sastry, S. (2011). Attacks against process control systems: risk assessment, detection, and response. In *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security* (pp. 355–366). New York, NY, USA: ACM.
- Carrier, B. D. (2006). Risks of live digital forensic analysis. *Commun. ACM*, 49(2), 56–61.
- Chan, E., Venkataraman, S., David, F., Chaugule, A., & Campbell, R. (2010). Forenscope: a framework for live forensics. In *Proceedings of the 26th Annual Computer Security Applications Conference* (pp. 307–316). New York, NY, USA: ACM.
- Chen, T. M., & Abu-Nimeh, S. (2011). Lessons from Stuxnet. *Computer*, 44(4), 91–93.
- Hay, B., Bishop, M., & Nance, K. (2009). Live Analysis: Progress and Challenges. *IEEE Security & Privacy*, 7(2), 30–37.
- Hay, B., & Nance, K. (2008). Forensics examination of volatile system data using virtual introspection. *SIGOPS Oper. Syst. Rev.*, 42(3), 74–82.
- Herrerias, J., & Gomez, R. (2007). A Log Correlation Model to Support the Evidence Search Process in a Forensic Investigation. In *Systematic Approaches to Digital Forensic Engineering, IEEE International Workshop On* (Vol. 0, pp. 31–42). Los Alamitos, CA, USA: IEEE Computer Society.
- Hunt, R., & Slay, J. (2010). The Design of Real-Time Adaptive Forensically Sound Secure Critical Infrastructure. In *Network and System Security, International Conference On* (Vol. 0, pp. 328–333). Los Alamitos, CA, USA: IEEE Computer Society.
- Kilpatrick, T., Gonzalez, J., Chandia, R., Papa, M., & Sheno, S. (2008). Forensic analysis of SCADA systems and networks. *Int. J. Secur. Netw.*, 3(2), 95–102.
- Linda, O., Vollmer, T., & Manic, M. (2009). Neural Network based Intrusion Detection System for critical infrastructures. In *Neural Networks, IEEE - INNS - ENNS International Joint Conference On* (Vol. 0, pp. 1827–1834). Los Alamitos, CA, USA: IEEE Computer Society.
- Morris, T., Vaughn, R., & Dandass, Y. S. (2011). A testbed for SCADA control system cybersecurity research and pedagogy. In *Proceedings of the Seventh Annual Workshop on Cyber Security and Information Intelligence Research* (pp. 27:1–27:1). New York, NY, USA: ACM.
- Nance, K., Hay, B., & Bishop, M. (2009). Digital Forensics: Defining a Research Agenda. In *Hawaii International Conference on System Sciences* (Vol. 0, pp. 1–6). Los Alamitos, CA, USA: IEEE Computer Society.

- Nehinbe, J. O., & Nehibe, J. I. (2012). A Forensic Model for Forecasting Alerts Workload and Patterns of Intrusions. In *Computer Modeling and Simulation, International Conference On* (Vol. 0, pp. 223–228). Los Alamitos, CA, USA: IEEE Computer Society.
- Peterson, D. (2009). Quickdraw: Generating Security Log Events for Legacy SCADA and Control System Devices. In *Conference For Homeland Security, Cybersecurity Applications & Technology* (Vol. 0, pp. 227–229). Los Alamitos, CA, USA: IEEE Computer Society.
- Richard, III, G. G., & Roussev, V. (2006). Next-generation digital forensics. *Commun. ACM*, 49(2), 76–80.
- Rodrigues, A., Best, T., & Pendse, R. (2011). SCADA security device: design and implementation. In *Proceedings of the Seventh Annual Workshop on Cyber Security and Information Intelligence Research* (pp. 25:1–25:1). New York, NY, USA: ACM.
- Su, Z., & Wang, L. H. (2011). Evaluating the Effect of Loading Forensic Tools on the Volatile Memory for Digital Evidences. In *2010 International Conference on Computational Intelligence and Security* (Vol. 0, pp. 798–802). Los Alamitos, CA, USA: IEEE Computer Society.
- Sutherland, I., Evans, J., Tryfonas, T., & Blyth, A. (2008). Acquiring volatile operating system data tools and techniques. *SIGOPS Oper. Syst. Rev.*, 42(3), 65–73.
- Tang, Y., & Daniels, T. E. (2005). A Simple Framework for Distributed Forensics. In *2012 32nd International Conference on Distributed Computing Systems Workshops* (Vol. 2, pp. 163–169). Los Alamitos, CA, USA: IEEE Computer Society.
- Yen, P.-H., Yang, C.-H., & Ahn, T.-N. (2009). Design and implementation of a live-analysis digital forensic system. In *Proceedings of the 2009 International Conference on Hybrid Information Technology* (pp. 239–243). New York, NY, USA: ACM.
- Zhu, B., Anthony, Joseph, & Sastry, S. (2011). A Taxonomy of Cyber Attacks on SCADA Systems. In *2011 IEEE International Conferences on Internet of Things, and Cyber, Physical and Social Computing*. San Diego, California: IEEE Computer Society.

FACTOR INFLUENCING MOBILE NUMBER PORTABILITY(MNP) IN KENYA: THE CASE OF AFRICA NAZARENE UNIVERSITY

Githii Stephen Kagwathi

John Njau Kamau

Department of Business administration, Africa Nazarene University

Mary Muthoni Njau, PhD Candidate

University of Nairobi

Elias Kiarie Kagiira

Department of Business Administration , Africa Nazarene University

Abstract:

The much awaited mobile number portability took effects in Kenya on April 1st 2011. The objective of porting was to accord consumers flexibility of shifting from one service provider to another without worrying of loosing their preferred numbers. The free movement of subscribers was expected to open a new competition front within the telephony market, with the small operators expected to benefit. However the porting process has not been so smooth. Over a year since the process started, only 47,206 subscribers have switched out of the possible twenty two million.

Methodology: The study used descriptive design whereby a sample of 142 subscribers was taken. Ten factors were identified as influencing mobile number portability in kenya. These were exit barriers, cost, process, customer service, satisfaction among others.

Key Words: Mobile Number Portability (MNP), Mobile Service Providers

Introduction

Kim and Shin (2007) defines mobile number portability as the ability for subscribers to retain their phone numbers when changing subscription from one mobile service provider to another. Odunaik (2010), described MNP as having a great opportunity which can increase acquisition and to a greater extent encourage healthy competition among telephone operators. The process however is a complex one and one that require goodwill from all parties involved.

Sutherland (2007) noted that MNP although ignored or overlooked in creation of the telcommunication markets, it is becoming a feature that regulatory commuissions or bodies are considering necessary to: reduce switching costs; facilitate subscriber choice; and ensure effective competition in the market.

Between 2004 and 2008 Communication Commission of Kenya (CCK) carried out a consultations that revealed that the market was ready for Mobile Number Portability (MNP). Based on this the commissions set the deadline to December 2010 but had to deferr it to April 2011 to give mobile service providers more time to acquire and test their equipment (www.cck.go.ke). The process is being facilitated by Porting Access Kenya, CCK and the mobile service providers.

The history of MNP started in 1990s whith Singapore implementing a limited version of this functionality, Hongkong implemented in 1999, Spain in 2000, and Australia in 2001 the list go on and on (www.scribd.com). Hence porting is not a new concept and according to CCK, Kenya joins 62 other countries around the world who have implemented it, South Africa and Egypt included. The main objective according to the commission for the service is to depeen the level of competition in the mobile telcommunications market and enhance consemmer choice. Shin and Kim (2007) seem to agree that MNP will result to mobile service providers actively competiting and providing innovative as well as improved customer service, in order to retain and expand their subscription base.

Prior to the introduction of the service, subscribers have had to use multiple Subscriber Identity Module (SIM) cards or buy two mobile handsets, thus incurring additional unnecessary costs

and inconvenience (www.cck.go.ke). The inconvenience, CCK go ahead and say that, it has inhibited subscribers from taking advantage of the growing competition in the industry. Despite the introduction of competition in the Kenyan telecommunication industry the lock-in effects of subscribers has helped the incumbents to retain market dominance, a scenario Shin and Kim (2007) seem to have identified in Korea. It is from this back drop that the CCK and many other regulators around the world sought to implement MNP. Availability of MNP is expected to benefits consumers as competition between the players intensify. Some of the specific benefits that are likely to go to customers as noted by Shin and Kim include: lower price, greater choice, higher quality, and greater range of services.

One of the key determinant of customer switching from one product to another, and in this case from one service provider to another, is the switching cost associated with the process (Polo, 2009). Porter as quoted by Ruyter, Wetzels, and Bloemer (1997, p439) defines switching cost as the costs involved in changing from one service provider to another. According to Polo (2009) switching costs are recognized as an important driver of customer retention that leads to stable and long-lasting relationships.

Information Permanent Secretary Dr. Bitange Ndemo (2011) noted that the portability may be undermined by the fact that many Kenyans already have more than one handset and SIM cards which enable the services of two or more service providers. Besides the issue of two handsets, there is also the fact that the process has been hit by a long delays before porting, and the porting fee which has been capped at Kshs 200 (2.5USD) has been said to be way above the cost of a new SIM card (www.capitalfm.com).

While there has been live discussions as to whether MNP increases competition in mobile markets as noted by Shin and Kim (2007), the more underpinning question may be whether subscribers are able to port freely without significant switching barriers. As Xavier and Ypsilanti (2008) argued switching patterns provide an important indicator that the demand-side of a market is well developed and that consumers are sufficiently empowered to participate actively. The two goes ahead to argue that the ability and willingness of consumer to switch is critically important and where switching is impeded or discouraged this could impact not only on the demand-side but also potentially raise supply side barriers-new entrants.

The motivation to switch (port) is generally a function of consumers' estimate of the performance of their existing provider; and whether or not they believe there are better alternatives available from other providers on the aspects of service that matter to them (Xavier and Ypsilanti, 2008). Where the market is perceived to be undifferentiated or where the current provider is perceived to be the best on the market on the criteria that are important, there may be no expected benefit from switching.

The aim of this paper was to assess why subscribers are or not switching. Issues covered by this research include:

- The extent of switching/porting by customers between the mobile providers;
- What is driving subscribers to port or what is holding them to their current mobile provider;

In order to fulfill the above objectives, the paper is divided into the following sections: Section 2- industry overview, literature review on number portability and switching costs; section 3- research methodology; section 4- the study findings and section 5- conclusion and recommendations.

Overview of Kenya Mobile Market

For a long period, Kenyan telecom industry has been dominated by Telkom which was a parastatal, but significant changes took place around 2004 when it lost its monopoly in the fixed-line and international bandwidth. With issuance of licences to other operators and carriers the competition landscape changed drastically (accessed on 5th May 2011).

Mobile services in Kenya started around 1993, but given high cost that was associated with both the hand set and the service there was low interest by Kenyans and by the end of 1999 there were 20,000 subscribers only (www.africantelecomsnews.com accessed on 5th May 2011).

There are currently four mobile service providers in Kenya namely Safaricom, Airtel, Essar Telecom, and Orange. The sector has witnessed profound changes in recent past just like in other countries, from technological advancement to increased regulations. The results has been new markets, new entrant, and new challenges (accessed on May 5th 2011). As at 31st March 2012 the total number of mobile subscribers had hit 29.2 million (Quartely Sector Statistic Report -3rd Quarter

January-March 2012 by CCK). This number represent a growth of 4.0 per cent. The report put Safaricom ahead of the other three providers with a market share of 65.3%, followed by Airtel (15.3%), Essar Telecom (8.7%), and Orange with 10.6%.

Safaricom was the first to enter the Kenyan market and to launch GSM-based mobile service in Kenya around 1999 (). From the same website majority of Safaricom shares are owned by Vodafone Kenya (40%), Kenyan government having 35% stake and the rest (25%) being owned by the public. It started as a department of Kenya Posts & Telecommunications Corporation, which was the monopoly operator (accessed on May 5th, 2011).

Airtel started its operation in Kenya in 2000 as Kencell, it was then rebranded to Zain in 2008 and finally to Airtel in 2010 (accessed on May 5th 2011). The company is the second largest in terms of market share and subscriber base after Safaricom (Quartely Sector Statistic Report 1st Quarter July-Sept 2010/2011 by CCK). The company, according to the report gained 4.4 percentage points of market share over the quarter under the review.

Essar Telecom Kenya Limited (Yu) got the licence to operate in Kenya in 2003, but its operation was delayed to 2008 due to court cases (www.africantelecomsnews.com). The company is a unit of India based group-Essar, which is a diversified business corporation with interest in areas such as manufacturing, shipping, energy, power among others. The group has its foot print in more than 15 countries around the globe. The company launched its mobile service network under brand "Yu" in November 2008 (www.yu.co.ke accessed on May 5th 2011). By September 2010 the company had a market share of 6.7% which was a drop from 7.4% in June same year (Quartely Sector Statistic Report 1st Quarter July-Sept 2010/2011 by CCK).

Telkom Kenya (Orange/France Telecom) ws established as a telecommunications operator in April 1999 under the Companies Act (www.telkom.co.ke accessed May 5th, 2011). The company partnered with France Telcom Group which saw the launch of Orange brand in Kenya in 2008. Besides offering mobile services, the company also offer fixed line services and Internet services.

Mobile penetration has been increasing steadily over the years and at the end of the quarter it stood at 74.0 per cent compared to 71.3 per cent last quarter (Quartely Sector Statistic Report -3rd Quarter January-March 2012 by CCK). Contrary to this increament, fixed line services witnessed a general decline in the same period, a fact attributed partly to increase in uptake of mobile telephony.

Future Outlook

Given that the number of subscriptions has trippled in the last five years from 7.3 in 2006 to the current figures of 22 million subscribers (Quartely Sector Statistic Report 1st Quarter July-Sept 2010/2011 by CCK), it is estimated that by the end of 2014 the number will hit 33.2 million achieving a penetration rate of 79 percent (www.africantelecomsnews.com). The CCK quarterly report note that the mobile market domenstrate increased subscriptions while a decline in fixed line will continue to be felt.

Competitive pressure is expected to remain intense among the four providers now than before given the introduction of MNP. The report contends that, will require providers to diversify their services.

Empirical Review

Mobile Number Portability

MNP allows mobile subscribers to retain their telephone numbers when they change mobile service providers (Sutherland, 2007). Implementation of MNP in Kenya aims to deepen the competitive environment within the mobile telecommunication industry and as a result enhance subscribers' choice as CCK Director-General Charles Njoroge noted. These sentiments were shared by Shin and Kim (2007) who noted that the implementation of MNP in Korea main objective was to benefit consumers though reduced prices as a result of competition between providers.

Murillo (2007) noted that number portability around the world started to gain greater relevance in the wake of market liberalization, but importantly with the advent of wireless telephone service. Lin et al., as quoted in Murillo (2007,p26) suggested three types of number portability; 1) Location portability where the user is able to retain their telephone number even after changing the address; 2) Service portability is where the subscriber retain the number when they change services within the same provider; and 3) Operator portability, which is being focused in this research, being a situation where subscribers maintain their number even after they change service providers.

In Kenya MNP was to kick off in December (2010) but had to be delayed up until 1st April 2011 in order to give services providers more time to put their houses in order and to test their equipments (www.cck.go.ke). However, the process has not been without teething problems as predicted by CCK Director General and confirmed by Information Permanent Secretary Dr. Bitange Ndumo. Some providers have developed cold feet to the process and have been accused of frustrating subscribers who want to port. There has also been technical hitches, by April 27th there were 10,000 successful requests, a whopping 15,000 requests were pending and 11,000 timed out (Daily Nation, April 28, 2011). There are accusations and counter-accusations which have resulted to one mobile provider threatening to sue the company overseeing number portability for what it terms as defamation and economic sabotage (Daily Nation, May 4, 2011).

The above mentioned “fightings” by mobile service providers in Kenya is captured by Murillo (2007) sentiments that, of the three types of number portability, MNP is the one that operators fear most given that it gives the subscribers the ability to move without losing their number. He goes on to note that MNP can force providers to improve their offerings be it in terms of quality or price, an objective the CCK sought to achieve. Implementation of MNP in Japan was also faced with the same opposition as Sutherland (2007) noted.

There is also cost element associated with porting that Murillo (2007) identified which could be the reason why some operators resist the process, this include: the cost of upgrading the network which was the cause for a delay in the Kenyan case and the cost of porting a subscriber’s number when they make that request. He noted that once a subscriber move to a competing provider the current provider loses that revenue stream and it will be costly for them to regain them in terms of marketing cost.

Murillo (2007) argued that implementation of MNP is as a result of market failure. The incumbent providers especially those who have a large subscriber base do not want to lose them, therefore they have great incentive to prevent the implementation of the process. Buehler and Haucap as quoted by Shin and Kim (2007,42) also hold the same perspective when they argued that, MNP is likely to benefit the new entrant and hurt the incumbent providers. This seem to be the case in Kenyan market.

The Porting Procedure

Subscribers wishing to port their numbers are expected to fill in the MNP Form that is available from the mobile service provider they intent to switch to. Together with this they are required to present original documents that would identify them such as identity card (I.D.), Passport or Armed forces I.D. card for authentication. When the number is owned by an organization an official letter duly signed by relevant authority must be presented (www.cck.og.ke).

According to CCK, customer wishing to port will be required to pay a fee of Kshs. 200 after which they will be issued with a new SIM card. They will continue to use their current operator until the process is complete. During this waiting period customers are required to ensure they save their contacts, clear any balance airtime or any balance in their mobile money transfer account. For those porting from providers who have postpaid services and airtime lending, bills must be cleared and any borrowed airtime be paid up (www.cck.og.ke).

To kick start the process subscribers will have to send word PORT or “HAMA” to 1501 using their current SIM card. Where the short message (SMS) fails the subscriber will be notified and advised to contact his/her new operator. When the porting process is complete, subscriber will receive an SMS from PORTING containing information on closer of the account, advice to use the new SIM card from the new provider, or Porting error (www.cck.og.ke).

The final stage is where the subscriber will replace the current SIM card with the new one from the new provider, thereby start enjoying the services of the new provider (www.cck.og.ke).

Pros and Cons of Porting

One of the advantage accorded to subscribers by MNP is the reduction of switching costs. Switching cost include both monetary costs as well as time and psychological effort of dealing with the new provider (Ruyter, wetzels and Bloemer, 1997). This was well captured by CCK Director General Mr. Charles Njoroge when he noted,

“One of the major factors that have been discouraging consumers and business firms from changing mobile service providers is the inconveniences of losing contacts with friends,

family and business associates. For business, change of telephone numbers could have cost implications in regard to advertising.”

Sutherland (2007) seem to agree with the above views, when he suggested that MNP present great advantages to small businesses and sole traders especially in regard to repeat business or personal recommendations. He gave examples of dress makers, painters plumbers, and taxi drivers, lose of their number would significantly affect their business.

In regard to financial element of switching cost, Polo (2011) argued that, where the current provider's price is too high, the higher the potential monetary savings from switching provider, which will lead to lower switching costs. He noted that competitors are likely to use price to stimulate customer behaviour, they do this by affecting the cost of switching providers for the customers of the focal firm. This is clear in the Kenyan case where airtel is making it free for those who want to port to their network in addition a reward of 1,000 free airtel reward points and one gets 25% bonus on any top up for 3 months (<http://africa.airtel.com/kenya> accessed on May 6th, 2011). Airtel behaviour seem to resonate within what Polo and Sese (2009) found out in their research, that marketing instruments, price and advertising can increase the size of the company customer base as it reduces the switching costs of its rival's customers.

MNP present and opportunity for mobile service providers to enlarge their market share (Sutherland, 2007). Where providers have competitive offerings they are likely to have more subscribers turning to their network. However, Polo and Sese (2011) warn that where subscribers are not satisfied with after-sales service quality, even where the switching cost is high, they are likely to consider changing their service providers.

There is likelihood that services offered in the telecommunication will improve in the advent of MNP as level of competition increases and this will directly be a plus to consumers. As Sutherland (2007) observed in his research most people who change providers do so due to poor quality of services such as lack of coverage, excessive prices or even poor customer care. These may be addressed with implementation of MNP, he reckons.

However, by switching to a new provider, a subscriber may end up losing special services and facilities that were offered by the previous provider (<http://factoidz.com> accessed on 27 April 2011). Porting is not an instant process, thus subscribers wishing to port must be willing to wait for a while before they can start enjoying services from their new provider. The waiting can be even longer or frustrating where the SMS to 1501 time out.

The process is costly too, as noted earlier implementation of this process requires investment on equipments as well as testing by the mobile service providers. Porting itself has cost element associated to it from mobile service providers perspective as was indicated by Sutherland (2007).

From the subscribers perspective the process can also be costly in instances where the providers are using two different technologies, they may be forced to buy a new handset that will be compatible with the new provider's technology (Sutherland, 2007). Finally from the ministry view, MNP is a strategy of enhancing competition within the telecommunication market and a ploy to reduce entry barriers for the new entrants (www.cck.org.ke).

The Experience of MNP in Other Countries

Kenya is not the first country to implement MNP, in fact it join other 62 countries (www.cck.org.ke). Sutherland (2007), in his work which sought to review the experience with the implementation of MNP found out the following;

1. By 2006 a third of consumers had ported in Hong Kong, Australia which implemented MNP years later had almost the same rate.
2. In United State of America (USA) the monthly switch average 1.5 to around 3.0 percent per month.
3. In European Union they introduced fixed number portability while member countries initiated MNP individually. By 2005 the total numbers that had ported in the whole of European Union (EU) was only 25 million which was about 5 percent of the total population.
4. Sutherland (2007) specifically noted that less than 1 percent of porting characterized Austria, Germany, France, Portugal and Greece

Factors influencing MNP

Srinagesh and Mitchell as quoted by Shin and Kim (2007, p42) identified that MNP has tremendously contributed positively to the competition environment of mobile market in US. Generally it was noted by Sutherland (2007) that very few countries have managed over 10 percent portability. Some of the reasons put across for the failure or low porting include; porting costs, obstacles created by mobile providers, and not being aware of the service. Ofcom research as quoted in Xavier and Ypsilanti (2008, p22) seem to agree that lack of information (not being fully aware) contribute to low rate of porting. In their research 48 percent of the respondents felt they didn't know enough to make the right choice.

Ofcom as quoted in Xavier and Ypsilanti (2008, p22) found out that more than 36 percent of mobile subscribers in United Kingdom (UK) had switched their mobile service provider in the past four years. However the research also pointed out that even those who had not ported during that period, had made some changes to their existing service within their service provider such as change of tariff or package. The research disclosed that the UK subscribers who were engaged in porting were likely to be young and males.

In Portugal the number of subscribers who had switched was however much less compared to that of UK. A survey conducted by Anacom as quoted in Xavier and Ypsilanti (2008, p23) found that only 19 percent had switched. Of those who switched about one-third gave their reason for switching as "most of my contacts are clients of the new operator". 66 percent of those who never ported, said they did see the reason to do so as they were satisfied with their current provider's services.

Ofcom's research also found out that the key drivers for switching was price and interest in technology. 85 percent of the participants agreed that low cost and overall value for money were very significant to their decision of whether to port or not. Regarding reasons given for not switching, Ofcom identified that 54 percent were satisfied with their current mobile service provider, 36 percent saying they were fairly satisfied.

The research also identified that most subscribers (64 percent) were unwilling to switch due to reluctance to leave a known and trusted provider for another who they were unfamiliar. This indicates the value of loyalty in telecommunication market.

In their conclusion Xavier and Ypsilanti (2008) noted that where subscribers are to switch, service providers are less likely to charge excessive price or supply poor quality services. They also noted that consumer awareness of alternative services is paramount.

On their part Shin and Kim (2007) concluded that in Korea MNP has not achieved the regulator's goal of freeing up switching barriers instead it enhanced them as providers sought to lock-in their customers. This notion was shared by Polo and Sese (2011) when they noted regardless of implementation of MNP by a number of regulators which were aimed at addressing switching cost or barriers, there still exist high switching costs in the mobile phone market inspite of the efforts. Shin and Kim also noted that since the implementation of MNP there has been an upsurge of advertisement wars and special offers in order to lock-in existing clients and steal those of rivals. This scenario is live in Kenya, one of the mobile service providers has been involved in intense marketing campaign encouraging people to *hama* (port).

Kangangi (2011) noted that porting is likely to be high where marketing campaigns are intense. He also pointed out that the porting speed in Kenya (2 days) is too long compared to other countries such as USA (2 hours), 20 minutes in the Republic of Ireland, 3 minutes in Australia and seconds in New Zealand (Kangangi, 2011). This clearly point that consumer would prefer a shorter time to port if the process is to be successful.

Customers who have more than one mobile lines are also unlikely to port (Kangangi, 2011)

Given the foregoing findings it is important to document Kenya experience in regard to MNP. This research will seek to identify the extent and factors influencing porting between mobile providers in Kenya.

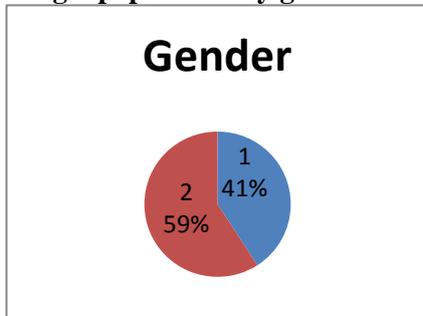
Research Methodology

The target population of the study were the 1500 students of Africa Nazarene University. As the study was descriptive, descriptive design was adopted with research strategy being a case study. The study was also cross sectional as it was done in one week in October 2012. A stratified sampling technique was used where a total of 142 participant were selected for the study. This represents approximately 10% of the total population which was considered a representative sample.

The participants were served with a questionnaire that had 40 variables considered important in decision making regarding porting. They were asked to rate the extent to which they agreed with statement made on different variables. Likert scale with points 1-5 was used where 1 represented least agreed and 5 strongly agreed. The survey lasted for seven days after which data collected was analyzed using SPSS version 17.

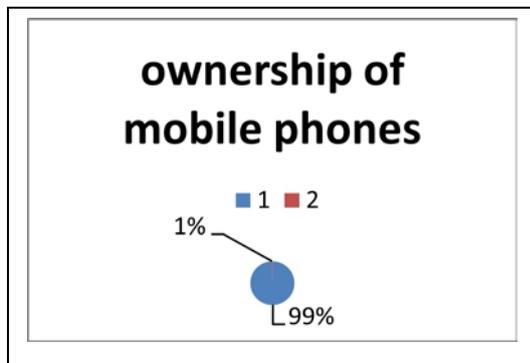
Results

Target population by gender



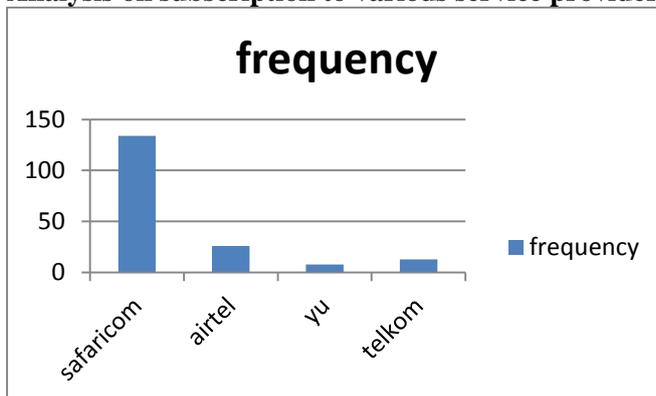
Out of all participants 59% were ladies and men constituted 41%. This reflected well the gender proportion in the college as is the case in the entire student population where there are slightly more women compared to men.

Analysis on Ownership of mobile phones



Analysis on ownership showed that 99% of participants owned a mobile phone and only 1% does not. The rate of mobile subscription in Kenya is 79% and given that these were university students, these findings are not a surprise. The researcher expected a hundred percent mobile phones ownership given that the participants are youth with higher technology orientation.

Analysis on subscription to various service providers

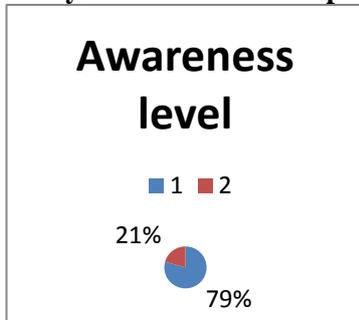


Analysis on subscription to the various service providers showed that Safaricom had 74% subscription of market share. Airtel had 14%, Yu mobile 5% while Orange had 7%. To assess whether there was a significant difference between these proportions and the records held by CCK a chi-square test was carried out.

	Observed	Expected (current market share)	O - E	(O - E) ²	(O - E) ² /E
Safaricom	74	65.3	8.7	75.69	1.159
Airtel	14	15.3	-1.3	1.69	0.110
Orange	07	10.6	-3.6	12.96	1.223
Essar	05	8.7	-3.7	13.69	1.574
					4.066

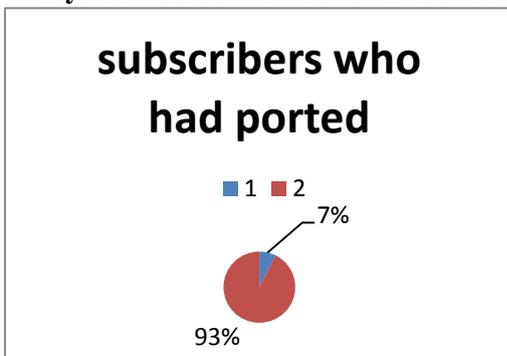
The chi-square test was used to test the null hypothesis that there was no significant difference between the record held by CCK and data obtained from the study at 0.05 significance level and 3 degrees of freedom where the critical chi square value is 7.815. The test chi square test statistic is indicated in the above table as 4.066 which lies within the acceptable region. We therefore fail to reject the null hypothesis and conclude that there is no significant difference between the CCK records on market share of all companies and the obtained data from the sample of the study.

Analysis on awareness of porting



Data collected on awareness level indicated that the concept of porting is not strange among the subscribers to various service providers. Majority of respondents, 79% indicated they were well aware that they could change from one service provider without losing their current subscription numbers while only 21% were uninformed on the same.

Analysis on actual number of subscribers who had ported



Data on the numbers of subscribers who had ported were rather scaring and inconsistent with the level of awareness. Whereas porting was supposed to provide subscribers with value for their money, majority of subscribers, 93% had not ported and only 7% had attempted to do so. This contradicts not only market theories that suggest that consumers seek value for their money but also commonsense.

Factor analysis

Factor analysis was conducted and the issues considered were appropriateness of the model, communalities, total variance explained and the factors extracted from the process.

Test of appropriateness of the model

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.699
Bartlett's Test of Sphericity	Approx. Chi-Square	897.974
	df	351
	Sig.	.000

To test the appropriateness of the factor model, Bartlett's test of sphericity was used to test the null hypothesis that the variables were uncorrelated. The results of PCA in the above table demonstrate that the null hypothesis, that the population correlation matrix is an identity matrix is rejected by the Bartlett's test of sphericity. The approximate chi-square statistics is 897.974 with 351 degrees of freedom which is significant at the 0.05 level. The values of KMO statistic (.699) is also larger (>0.5). Thus factor analysis was considered an appropriate technique for analyzing the correlation matrix.

Communalities

Using principal component analysis (PCA) to extract communalities among the variables, the result indicate that all factors accounted for a significant proportion of variance as they range from 0.454 to 0.823.

Total variance

We conducted a factor analysis on the 139 questionnaire sent to participants of the study. Using eigenvalues greater than 1.0, evaluation of scree plots, total explained variance, and factor loadings greater than 0.35 as criteria for identifying meaningful factors (see Nunnally and Bernstein, 1994), we identified 10 factors (or categories).

Total Variance Explained

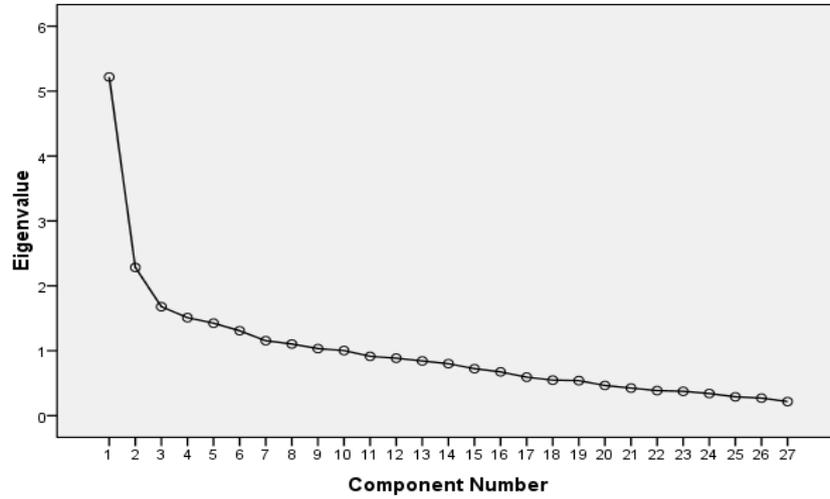
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.218	19.325	19.325	5.218	19.325	19.325	2.295	8.498	8.498
2	2.284	8.460	27.784	2.284	8.460	27.784	2.193	8.124	16.622
3	1.678	6.216	34.001	1.678	6.216	34.001	2.037	7.545	24.167
4	1.510	5.591	39.592	1.510	5.591	39.592	1.856	6.872	31.039
5	1.424	5.274	44.865	1.424	5.274	44.865	1.802	6.673	37.713
6	1.308	4.846	49.711	1.308	4.846	49.711	1.755	6.499	44.212
7	1.154	4.276	53.987	1.154	4.276	53.987	1.629	6.035	50.246
8	1.104	4.089	58.076	1.104	4.089	58.076	1.439	5.330	55.577
9	1.032	3.824	61.899	1.032	3.824	61.899	1.362	5.044	60.620
10	1.002	3.712	65.611	1.002	3.712	65.611	1.347	4.991	65.611
11	.914	3.385	68.996						
12	.885	3.279	72.275						
13	.842	3.119	75.394						
14	.801	2.965	78.359						
15	.723	2.679	81.037						
16	.674	2.496	83.533						
17	.593	2.195	85.729						
18	.546	2.024	87.752						
19	.540	1.999	89.751						
20	.465	1.722	91.473						
21	.425	1.574	93.047						
22	.385	1.427	94.474						
23	.374	1.386	95.860						
24	.341	1.263	97.123						
25	.290	1.074	98.198						
26	.270	.999	99.196						
27	.217	.804	100.000						

Extraction Method: Principal Component Analysis.

From the above table ten factors have been extracted with Eigen values greater than 1 accounting for 65.611% of the total variance. No factor explains more than ten percent of variance and therefore all factors are quite significant

The Scree plot below indicates total variance associated with each factor and shows a distinct break between steep slope of the large factors and the gradually trailing off of the rest of the factors. The scree plot further supports a 10-factor model.

Scree Plot



Rotated Component Matrix

	Component									
	1	2	3	4	5	6	7	8	9	10
VAR00007	.804	.061	.090	.003	.140	.144	-.065	.034	.187	-.085
VAR00008	.770	.027	-.042	.141	-.060	-.002	.199	.039	.088	.096
VAR00006	.564	.133	.237	.088	-.148	.367	.200	-.117	-.221	.072
VAR00025	.020	.709	.154	-.016	.266	.022	-.172	.088	-.016	-.030
VAR00026	.110	.687	.083	.007	-.023	.117	.259	-.077	.044	-.045
VAR00018	-.127	.597	-.026	.096	.078	.285	.139	.137	.297	.137
VAR00023	.364	.492	.198	.260	-.029	-.380	-.173	.066	-.027	.179
VAR00027	.215	.353	-.172	.292	.173	.049	.269	.250	.012	-.033
VAR00022	-.018	.189	.638	.259	.025	.103	-.054	.082	.082	.106
VAR00024	.192	.169	.636	-.212	.168	-.221	.118	-.009	.276	.072
VAR00012	.174	-.149	.619	-.132	.229	.323	.196	.094	.026	-.136
VAR00017	-.228	.096	.454	.394	.039	.111	.210	.395	-.150	-.012
VAR00011	.309	-.164	-.020	.691	.013	-.054	.053	.064	-.086	-.127
VAR00021	-.023	.209	.079	.676	.035	.006	.076	.056	.234	-.030
VAR00003	.048	.118	.076	-.008	.848	.008	.041	.102	.089	.045
VAR00019	-.198	.178	.316	.318	.566	-.085	.036	-.006	-.045	.298
VAR00002	.120	.394	.202	.057	.408	.141	.280	-.250	-.063	.174
VAR00005	.140	.164	.099	-.052	-.056	.751	-.012	.075	.023	.045
VAR00010	.187	.115	.009	.471	.193	.555	.075	-.279	.183	-.018
VAR00016	.035	.097	.218	.042	-.028	.026	.848	-.033	.173	.104
VAR00015	.377	.067	-.087	.216	.284	.039	.579	.147	-.062	.013
VAR00004	.078	.027	.095	.060	.037	-.034	-.019	.810	.012	.131
VAR00014	-.028	.172	.373	.050	.182	.272	.124	.418	.236	.123
VAR00009	.072	.052	.092	.034	.159	.179	.213	.193	.661	-.183
VAR00020	.210	.105	.217	.213	-.117	-.082	-.083	-.255	.636	.261
VAR00001	.028	-.006	.002	-.120	.095	-.003	.082	.121	-.041	.835
VAR00013	.060	.089	.217	-.035	.392	.405	.047	.130	.174	.503

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 18 iterations.

The rotated factor matrix sorted out 10 factors that were deemed significant for this study. Each of the factor had loadings greater than 0.3. The rotation method was Varimax with Kaiser Normalization where items were sorted from one with the highest loading to the one with lowest. The results of the rotated matrix helped to identify and interpret the following as key factors influencing porting process based on the factor loadings.

Factors influencing porting

Factor	Factor interpretation	% of variance Explained	Factor loadings	Variables included in the factor
F1	Customer Service	8.498%	.804	Porting results in service providers becoming more innovative
			.770	Porting makes service providers improve customer service
			.564	Porting enhances consumer choice of service providers
F2	Loss incurred	8.124%	.709	Fear by subscribers of losing their number affects porting
			.687	The inconveniences of losing contacts with friends, family and business associates affects porting
			.597	For businesses, change of telephone numbers could have cost implications in regard to advertising.
			.492	Before porting one is required to ensure they save their contacts.
			.353	Where subscribers are not satisfied with after-sales service quality, even where the switching cost is high, they are likely to consider changing their service providers
F3	Constraints	7.545%	.638	Before porting one must clear any balance airtime or any balance in their mobile money transfer account
			.636	The cost incurred by service provider of porting a subscriber's number when they make that request affects porting
			.619	Possession of more than one handset and SIM cards by Kenyans undermines porting
			.454	"Fightings" by mobile service providers in Kenya is hindrance to porting
F4	Competitors incentives	6.872%	.691	Higher quality of service by competitor leads to porting
			.676	Mobile service providers provide great incentive to customers to prevent the implementation of the porting process
F5	Preparedness	6.673%	.848	Kenyans are generally not ready to port lowering the pace`
			.566	The cost of upgrading the network affects porting speed by subscribers
			.408	Mobile number portability was initially ignored or overlooked in creation of the telecommunication markets in Kenya
F6	Competition	6.499%	.751	Porting deepens the level of competition in the mobile telecommunication's market
			.555	Porting may be stimulated by lower price of service from the competitor
F7	Satisfaction with service provider	6.035%	.848	Satisfaction by performance of current service provider is an impediment to porting
			.579	The ability and willingness of consumer to switch is critically important in porting process
F8	Porting Facilitating cost	5.330%	.810	mobile service providers require special equipment to facilitate porting
			.418	The porting fee which has been capped at Ksh 200 is a hindrance to porting process

F9	Exit barriers	5.044%	.661	The lock-in effects by service providers hinders porting
			.636	MNP is likely to benefit the new entrant and hurt the incumbent providers.
F10	Process	4.991%	.835	The process of porting is complex and challenging
			.503	The process of porting is hit by long delays before completion

Conclusion and recommendations

The study had set out to identify factors that influence mobile number portability in Kenya. This was aroused by the fact that it was expected that majority of subscribers would port to reap the benefits of reduced cost of service. However results indicated that only 7% of subscribers had attempted to do so. The factors therefore affecting the process must have been significant and require to be addressed. This study has revealed that customer service offered by current service providers was a great impediment to porting accounting for 8.498% of total variance. The service provider especially market leader (Safaricom) provided superior customer care and therefore the subscribers found no good reason for porting. We can conclude that good customer care as perceived by subscribers supercedes incentives such as cost reduction that would be enjoyed if customers switched. This is in agreement with finding of Shin and Kin (2007). Many subscribers also feared they would lose their contact and even business opportunities if they ported. This was accounted for by 8.124% of total variance. Fear of porting and its consequences is therefore as a significant factor. CCK ought to assure subscribers that their fear is unwarranted since they would retain their numbers in their current state.

The cost to be incurred by service providers as well as cost incurred by subscribers were also identified as significant factors. The fees attached to porting was considered unnecessary and contributed to delay in the process of porting. The level of preparedness to port was a major factor affecting porting process accounting for 6.673% of total variance implying that a large proportion of subscribers and service providers were not well prepared to engage in the process. Satisfaction with current service providers accounting for 6.035% of variance suggest that rewards associated with porting were not strong enough to attract customers to port. The service providers especially the market leader had put exit barriers in form of incentives that hindered the porting process accounting for 5.044% of total variance. Market instruments such as advertising, sales promotion and added services such as money transfer were affecting the porting process adversely. These benefit and marketing activities locked subscribers and accounted for 5.044% of variance. Coupled with this, is level of competition among the service providers which accounted for 6.499% of total variance. The market leader (Safaricom) sets the bar very high in terms of competition frustrating the efforts of other providers to offer substantial competition.

The process of porting is perceived to be very complex and challenging for most subscribers accounting 4.991% of variance. This is in agreement with findings of Odunake (2010) where the process was considered to be a major drawback to the process.

This study makes the following recommendations:

1. The level of awareness is relatively low (21%) and therefore there is need to increase awareness informing subscribers of the benefits that would accrue to them if they engaged in the porting process.
2. The process of porting should be simplified and the cost to be incurred waived off so as to facilitate the process.
3. Assurances should be given to subscribers that porting does not result in loss of their current numbers but only a change of service provider and this could alleviate fears of loss perceived by subscribers.

References:

Daily Nation April 28 2011

Daily Nation May 4 2011

Durukan. T, Bozaci. I, Dogan. T, "Mobile Number Portability in Turkey: An Empirical Analysis of Consumer Switching Behavior" European Journal of Social Sciences 2011, Vol. 20 No. 4.

- Maicas. P, Polo. Y, Sese. J, "Mobile Number Portability in Europe: International Review of Economics & Finance-INT REV ECON FINANC, Vol. 18, No. 4, 2009 pp. 611-623.
- Morillo. M, "Number Portability in Central America" Emerald Group Publishing Ltd Vol. 9 Iss 4, 2007, pp 25-37
- Odunaike Solomon, "The Impact of Mobile Number Portability on TUT students on-line Connectivity". ISECON Proceeding V27, 2010 n1348.
- Polo. Y, Sese. J, "How to Make Switching Costly; The role of Marketing and Relationship Characteristics" Journal of Service Research November 2009 vol. 12 no. 2 119-137
- Ruyter, de, K, J.M.M. Bloemer and M. Wetzels. "On the Relationship between perceived services quality, service loyalty and switching costs" International Journal of Service Industry Management 1998, pp. 436-454.
- Shin. D, Kim. W, "Mobile Number Portability on Customer Switching Behavior: in the case of the Korean Mobile Market" Emerald Group Publishing Ltd Vol. 9 No. 4 2007 pp 38-54
- Sutherland. E, "Mobile Number Portability" Emerald Group Publishing Ltd Vol. 9 No. 4 2007, pp 10-24 ISSN 1463-6697.
- Xavier. P, Ypsilanti. D, "Switching Costs & Consumer Behavior: Implication for Telecommunication regulation" Emerald Group Publishing Ltd Vol. 10 Iss 4, 2008 pp 13-29
- www.telkom.co.ke Accessed on 26th April 2011
- <http://factoidz.com> Accessed on 27th April 2011
- www.africatelecomsnews.com Accessed on 5th May 2011
- www.yu.co.ke Accessed on 4th May 2011
- www.safaricom.co.ke Accessed on May 6th 2011
- www.capitalfm.com Accessed on 5th May 2011
- <http://africa.airtel.com> Accessed on May 6th 2011

BUILDING VISION AND VOICE BASED ROBOTS USING ANDROID

Pradeep N. , Assistant Professor
Dr. Mohammed Sharief, Professor
Dr. M. Siddappa, Professor

Dept. of Computer Science and Engineering, SSIT, Tumkur, Karnataka, India

Abstract:

This project uses Android, along with cloud based APIs and the Arduino microcontroller, to create a software-hardware system that attempts to address several current problems in robotics, including, among other things, line-following, obstacle-avoidance, voice control, voice synthesis, remote surveillance, motion & obstacle detection, face detection, remote live image streaming and remote photography in panorama & time-lapse mode. The rover, built upon a portable five-layered architecture, will use several mature and a few experimental algorithms in AI, Computer Vision, Robot Kinetics and Dynamics behind-the-scenes to provide its prospective user in defense, entertainment or industrial sector a positive user experience. A minimal amount of security is achieved via. a handshake using MD5, SHA1 and AES. It even has a Twitter account for autonomous social networking.

Key Words: Voice Synthesis, Voice Control, Robot-User Interaction, Face Detection, Motion and Object Detection, Remote Surveillance, Panoramic and Time-lapse Photography, Line-following and Obstacle-Avoiding Robotics, AI, Computer Vision, Image Processing, Network Security

Introduction

Android, which is actually a software stack, running atop the Linux kernel, is one of the most successful operating systems available to mobile users today. This success can be attributed to the fact that Android makes development of application softwares ('apps') really easy by allowing developers to code in Java (among many other languages) and thus enabling them to take advantage of the innumerable packages and modules made available by the Java community. Also the *android.** package provides several APIs to access features that are traditionally unique to mobile platforms (if not only Android). Examples of such APIs are the ones that we can use to access the readings of the innumerable sensors that an Android smart-phone may have, like proximity meter, compass, GPS, accelerometer and, most importantly, the camera. The presence of these sensors often make Android a more lucrative computing platform compared to 'traditional' platforms like PCs. Like any other Linux distribution, Android supports the entire TCP/IP stack, right from working with raw sockets to making application-layer level HTTP requests. Android can connect to the Internet using any of Wi-Fi, 3G, GSM or EDGE. Moreover, it can usually make telephone calls and send and receive SMSes (which are attractive features, but almost always absent in PCs). It is because of Moore's Law that an Android phone of today is much more powerful than a desktop PC from last decade. It is normal for an Android phone to ship in with more than a 1 GHz processor and 512 MB RAM.

These facts have recently increased people's curiosity on using Android for uses, which till now were rarely associated with mobile phones. Devices like smart TVs and dashboards of automobiles are being run on Android. NASA is planning to launch three nano-satellites (named Alexander, Graham and Bell) running on Android smart-phones under a project named PhoneSat. We simply wanted to explore the possibility of using Android to control a robot.

Description Of The Robot

From the perspective of the Android phone, the robot is yet another application running on it. The application acts as a broker between the user and the robot hardware.

The communication between the user and the robot takes place through HTTP or SMS. The robot (comprising of the Android phone and other hardware) itself may connect to the Internet through Wi-Fi or 3G/GSM/EDGE. The user interacts with the robot using a RIA (Rich Internet Application) that relies heavily on HTML5, CSS3 and JQuery. Security, being one of the major quality attribute, is achieved using a basic technique. The robot authenticates the user by matching MD5 and SHA1 hash of passwords. This prevents the transmission of the actual password string across the network. Once authenticated, all further communication between the user and the robot are encrypted using AES with the password as the key.

The communication layer interacts with the user to receive new commands. When a new command arrives, it notifies the Controller Layer. The Controller Layer is responsible for stopping the current task and starting the new task as required by the user's commands.

The tasks that the rover can perform are basically Java classes. Some classes may contain methods to perform more than one tasks. These tasks enable some, but not all, features of the robot. The features, that are not implemented directly using Java in the Android application, are implemented using JavaScript in the RIA. Examples of such features are 'Voice Recognition' and 'Motion and Object Detection'.

Because the Communication Layer, the Controller Layer and the individual features are all separated from one another, we were able to pack in a whole lot of features into our robot.

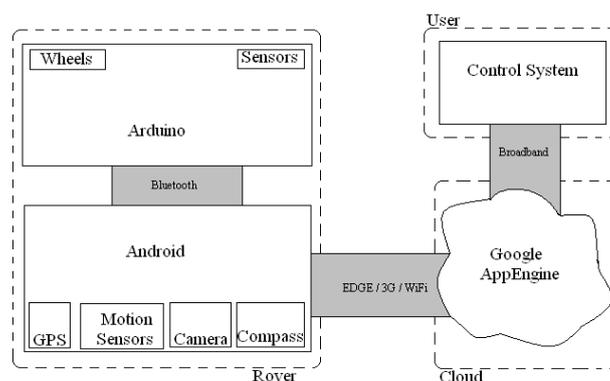
The features that the robot has are enumerated below:

- 1.Line Following
- 2.Obstacle Avoidance
- 3.Voice Synthesis
- 4.Voice Control
- 5.Robot-User Interaction
- 6.Face Detection
- 7.Motion and Object Detection
- 8.Remote Surveillance
- 9.Time Lapse Photography
- 10.Panoramic Photography

Features 1-8 address problems in robotics, while Features 9 and 10 address problems in photography. These features, working together, provide a robotic system that will enable its user to operate it with ease and efficiency.

Description Of The Architecture

It can be said that, the outcome of our project would be a robotic software platform that allows the user to control the robot over HTTP. Our initial plan was to create a cloud application that acts as a synchronizing server between the robot and the user. The interested party does a HTTP POST to the cloud whenever it needs to update any information. Availability of new information is checked by repeatedly polling the cloud application for new data.



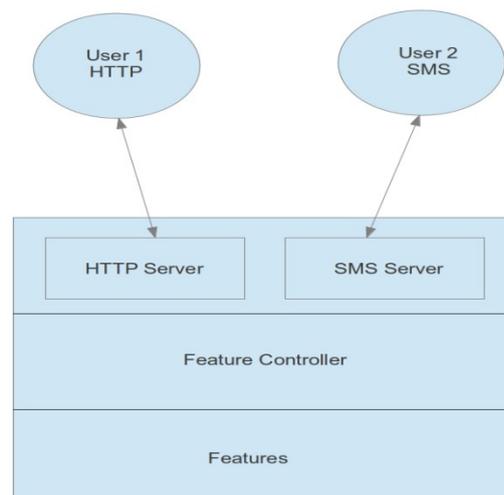
While this initially looked like a good scheme, the architecture has a few defects which surfaced only when we actually implemented and tested the architecture. They are enumerated below:

1. Redundancy : Because the cloud acts as a broker between the robot and the user, an unwanted redundancy is introduced. This is because one party, when it needs to inform something to the other party, sends the data to the server on the cloud and not actually to the other party.

2. Latency : The cloud can inform any party about any new data only when the said party polls for new data using HTTP GET. This introduces a latency between the time when the data is posted and when it is retrieved. This latency, ranging from less than a hundred milliseconds to more than a few seconds, although not astronomical, is undesirable. Although, W3C has introduced Web Sockets for HTML5 to enable a full-duplex connection over HTTP in late-2011, it is not yet widely available.

3. Resource Utilization : Repeated polling means that the rover will post data to the cloud and poll the cloud server for new user data at all times. These posting and polling would take place even when there is no need for such activity. Thus resource is misused to a great extent. And on a mobile platform like Android (a typical phone may have only 1 GHz processor with 512 MB RAM), this mis-utilization of resource affects the system adversely.

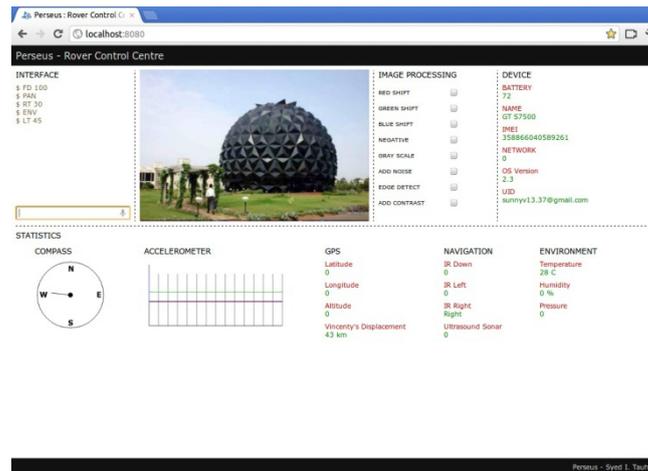
To remove these disadvantages, we removed the server application on the cloud and replace it by creating and embedding our own custom HTTP server in the robot. We created a Communication Layer that rests between the user and the core robotic features. The Communication Layer consists of two Java classes. One acts as as HTTP server and the other as an SMS server. These are implemented as Java threads using Android's Service. The HTTP server is created by extending and overwriting a few classes of NanoHTTPD. The SMS server is created simply by using Android's APIs. For obvious reasons, some features (for example, Video Surveillance) are inaccessible through the SMS server.



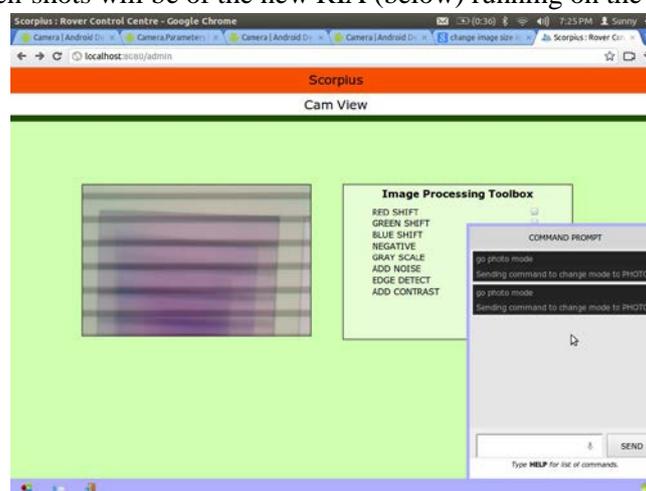
The cloud-based architecture is now deprecated, but an working implementation using Google appEngine on Python is available in the source code repository of the project. The repository is available at bit.ly/scorpius.

The removal of disadvantages of using a cloud based server made possible by integration of an embedded server within the application did enable us to create a new, more dynamic RIA which also has features like 'Voice Synthesis' and 'Motion and Object Detection' built into it.

Given below is an screen-shot of the RIA before the change in architecture. The entire web application is in a single page and the RIA couldn't boast of any feature. But on such a high-latency architecture, it was most apt.



All subsequent screen-shots will be of the new RIA (below) running on the new architecture.



Description Of The Features

Several features are available to the robot through the Android application. Only three of them rely on the physical robot (line-following, obstacle avoiding and panoramic photo), the rest can be used on any Android phone, even without the robot.

Most of the features are implemented using Java, directly on the Android phone, but a few, like 'Motion and Object Detection' and 'Voice Recognition' are implemented using JavaScript on the RIA, which is hosted, as aforementioned, using an embedded HTTP server on the Android phone.

We use a wide variety of APIs, both locally and over the cloud, to implement these features. This saves time and resource by virtue of reusable codes, thus letting us, proverbially speaking, 'stand on the shoulders of giants'. All libraries and packages are in open source and so is this application.

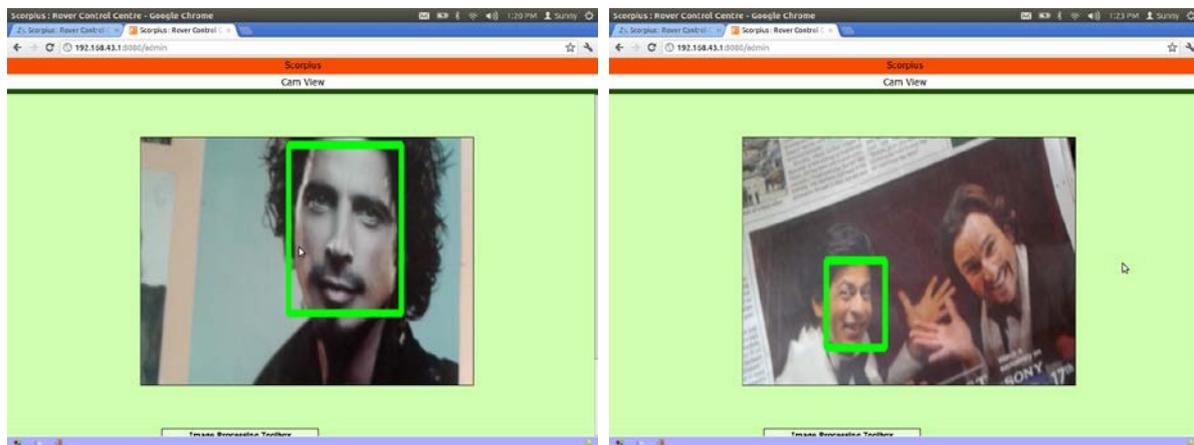
In the section below, we try to describe the various features, along with a summary explaining their implementation techniques.



Voice Recognition allows a user to control the robot using only his voice.

We implemented it using Google Chrome specific APIs. These APIs enable users to use Google Voice Recognition software over the cloud using just a few additional HTML5 tags and JavaScript-callbacks.

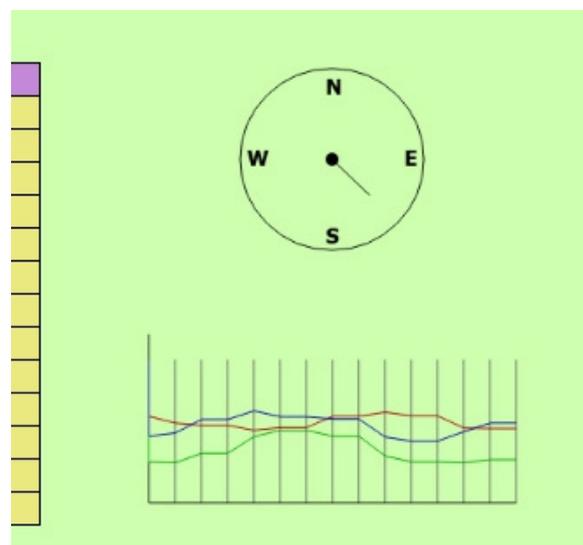
Voice Synthesis allows the user to type a string of text remotely and have the Android phone read out the string aloud. This is implemented using the Android's TTS (Text To Speech) APIs from the following packages : android.speech.tts.TextToSpeech and android.speech.tts.TextToSpeech.OnInitListener.



In the **Face Detection** mode, we process subsequent frames using the Android's camera. If and when a face is detected, a green box is drawn around the face on the output image. It is implemented using openCV and the LBP Cascade Classifier through XML in the Android application. The screenshots show that the Cascade Classifier does a good job of finding faces in an image, but as can be seen from the third image, it still is not perfect.

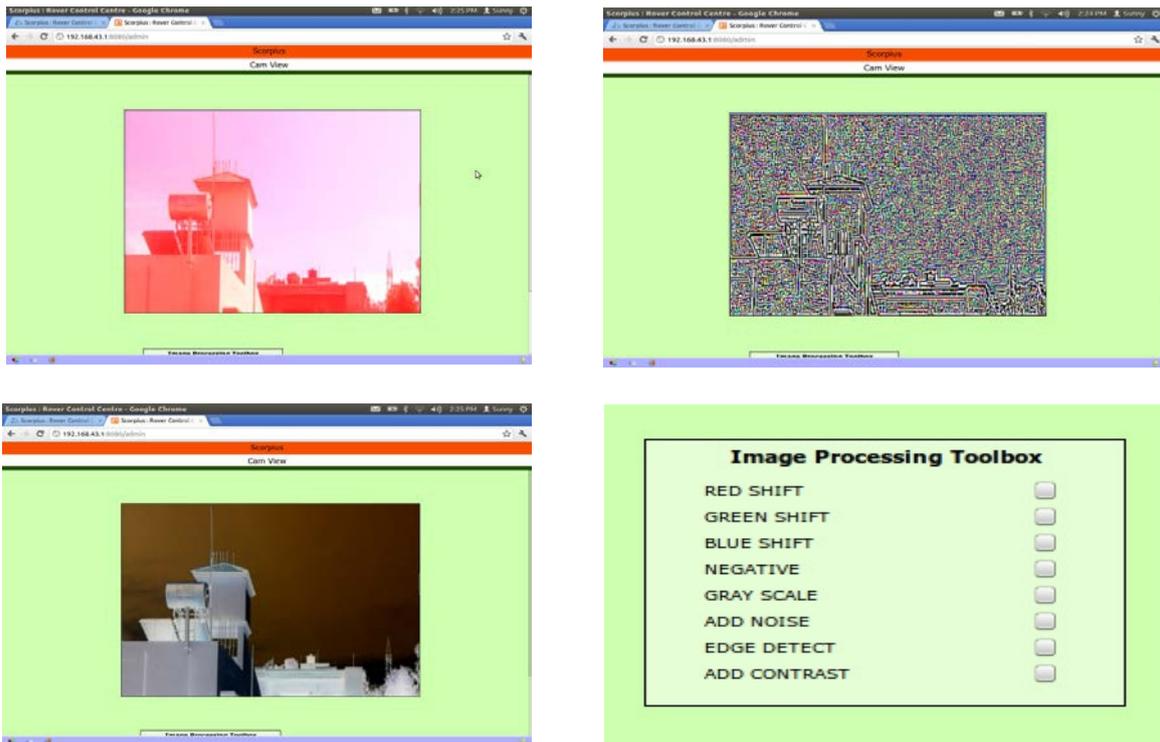
Motion and Object Detection is done by analyzing the sensor readings from Android phone. We primarily use three sensors – accelerometer for motion and vibration detection, proximity meter for object detection and compass to detect presence of electrical devices by looking for effects of electromagnetism.

Name	Status
Compass	134
Latitude	26.331844
Longitude	91.00237
Acceleration - X	3.064578
Acceleration - Y	-7.5082164
Acceleration - Z	4.903325
Proximity	5
IR Down	-
Sonar Front	-
Sonar Left	-
Sonar Right	-
Battery	68
Temperature	350



The Motion and Object Detection takes place in the RIA, which is an innovative, though still experimental, technique for **Robot-User Interaction**. The RIA is made using W3C standard tools, namely HTML5, CSS3 and JQuery-enhanced JavaScript. Some of the HTML5 specifications are very recent, and supported only on recent browsers. We expect the user to have the most recent version of Google Chrome.

The RIA uses pixastic.js for some real-time image processing whenever any video streaming is taking place.



The user can access the various image processing options on the RIA from a toolbox. The user can apply any single option or any combinations of the options by selection the required effect from the toolbox.

Remote Surveillance is the most basic form of live streaming, without the application of any computer vision algorithm on the input image.

Panoramic Photography enables us to capture wide-angled images using standard-lens camera. This is done by capturing several images and stitching them together using openCV.



Time-lapse Photography is the technique of capturing a video at a slow frame-rate and playing it back using normal frame-rate.

Line Follower is the name given to any robot that can move about in its environment by following a line on the ground. Traditionally people use IR sensors to detect lines on the ground, but we'll use only the Android's camera and openCV. There are two major techniques for finding and following a line using Computer Vision. One is to use Hough Line Transform, followed by detection of the vanishing point of the lines. We move the robot towards the angle which directs to the vanishing point. The other method is to set up a 8-bit Color Blob Detector and move the robot in such a way as to keep the blob at the center of the image.

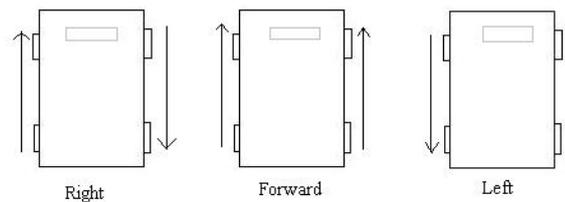
Obstacle Avoidance is another problem in robotics that is traditionally solved using hardware sensors, but which we attempt to solve using the Android's camera and openCV. We do this by looking at individual contours and attempting to navigate by keeping the robot from any colliding with any contour. As of this writing, the Obstacle Avoidance algorithm is yet to be implemented in the Android application.

The Robot-Android Interaction

The Android phone is connected to the robot using either microbridge or ADK using the USB. The Android provides two floating point variables in the range -1 to +1. One indicates angle of direction of motion, where -1 indicates extreme left, +1 indicates extreme right and 0 indicates straight motion. The other variable indicates speed of motion, where -1 indicates fast reverse, +1 indicates fast forward and 0 indicates no motion. These values can also be obtained from the embedded HTTP server over TCP/IP.

The Android application needs the robot to have just two motors with wheels. It doesn't take any reading from the robot and doesn't need the robot to have any sensor. But from the perspective of the robot, the Android phone is just another sensor. The robot may have its own sensors and use them in its own calculation (for example, to make hardware switch that turns off the robot in case of danger), but the Android phone need not know about those sensors.

As mentioned previously, the robot needs just two wheeled motors capable of bi-directional motion. These are used for all robot kinetics using differential drive, as illustrated below.



Description Of Security

Security, as any course in Software Architecture would tell us, is one of the most important 'quality attributes' in a software. Moreover, we need to be more careful about security in an application like ours that depends on wireless communication. Also, HTTP, unlike HTTPS, uses only plain-text, no cyphers. Hence, an application like ours would be vulnerable to sniffing and spooking attacks. Therefore, we have implemented some basic security within the application. The pseudocode for the Android application is given below:

1. Initialize auth = false, IP='127.0.0.1'
2. Wait for user to login from the RIA using MD5 and SHA1 hash of password. Check against user-saved password. If password matches, save user IP and turn auth = true. Else, exit.
3. If auth=true and IP = requestingUsersIP, then all communication between the RIA and the rover are to be encrypted using AES with the password as the key and then transmitted.

The following is the pseudocode for the RIA:

1. When user enters the password, the RIA authenticate the user by sending MD5 and SHA1 hash of entered string. If authentication succeeds, let user access control. Else, reload.
2. All communication between the RIA and the rover are to be encrypted using AES with the password as the key and then transmitted.

This security technique has the advantage of never transmitting password or plain-text over the network.

Observations

The cloud-based architecture, which looked promising initially, turned out to be a redundant system with unwanted latency. We removed these problems by using an embedded HTTP server, but that too comes with its own baggages.

Because the HTTP server has to run at all times, the battery drains rapidly. But this should not be a problem when running the phone in 'accessory mode' using ADK. The second problem is that most low-end mobile phones are not very effective in running multi-threaded applications like these because of the absence of multiple cores in the processor. But, even with these minor defects, the application would be suitable for usage in many areas.

Conclusion

Developing this project has enabled us to understand many aspects of creating a full-fledged real-time system. No other system of similar style and scale exists in the knowledge of the project members. We understood parts of AI and Computer Vision while making this project. The project will get refined as our understand AI and CV improves.

References:

1. *Learning openCV Computer Vision and the openCV library* - Gary Bradski and Adrian Kaehlar
2. *OpenCV 2 - Computer Vision Application Programming Cookbook* - Robert Laganière
3. *Android - A Programmer's Guide* - J. F. DiMarzio
4. *Beginning Android* - Mark L. Murphy
5. *Beginning Arduino* – Michael McRoberts
6. *The Robot Builder's Bonanza – 2nd Edition* – Gordon McComb
7. <http://developer.android.com>
8. <http://www.arduino.cc>
9. <http://www.opencv.org>
10. <http://www.talkingelectronics.com/>

GROUND FLASH DENSITY OF THE BRASS COAST OF NIGERIA

John Tarilanyo Afa

Benjamin Ayebapreye Kelvin

Electrical and Electronics Engineering, Niger Delta University, Bayelsa State, Nigeria

Abstract:

With the present industries and the expected industrial growth in the Brass Coast, it was necessary to study the ground flash density in the area. It is a lightning prone area because of the location and the environmental conditions.

Data for thunderstorm days in the year was collected from 2008 to 2012 from the meteorological department of the oil company and meteorological department of the College of Education Okpoma. Several observations were compared and the data were analysed. The ground flash density was found for each year in the two locations. The ground flash density (GFD) was about 10 flashes/km²/year to 15 flashes/km²/year.

This data is necessary because it will help Electrical and Telecommunication Engineers and operators to involve safer and better methods to protect and minimize danger to life and equipment.

Key Words: Ground flash density, thunder storm day, lightning activities, keraunic level, probability density

Introduction

Brass is becoming the most industrialized place in Bayelsa State due to the presence of the oil terminal of Agip Oil Company and the liquefied Natural Gas Company. It is expected that the oil refinery company will soon take off.

Brass is situated at the coast of the Sombrero River that emptied into the Atlantic Ocean. The Brass town is not more than 500m to the Atlantic Ocean and it is surrounded by rivers and tall mangrove swamps [1]. This is an area of high keraunic level with intermittent and unexpected stormy weather with the ocean breeze. During the month of November the stormy weather is persistent that it affects the economic life of the people.

A Survey on Lightning Distribution

Research on lightning distribution in Africa is limited. Several global detection satellites provide coverage of the continent [2], though the spatial resolution and temporal coverage is limited. The most comprehensive flash density results for the continent was produced by Christian et al (2003) using the Optical Transient Detector (OTD) and Lightning Imaging Sensor (LIS). The Congo Basin was identified as the lightning hotspot of the world. In this regard, South Africa has played leading role in the field of lightning research in Africa [2, 3]. Since then several modern techniques have been used for effective lightning detection.

Nigeria has been living with the global records with little effort to have local records except with the multi-national companies and aviation industries operating in some localities. These records can only be restricted to some areas and can not give a global record of the country. For the development of a country or industrialization, these records are vital for the power and telecommunication industries.

It has been established that the highest flash densities occur in coastal area, mountainous regions, regions frequented by migrating synoptic scale cycloves and convergence zones [2, 4]. That is why air movements and different meteorological data affect lightning activities in localities.

A ten year average ground flash density map on U.S.A. [4] indicated areas of different (GFD), elevation or convectional air movement. Among the areas of high GFD are

- The state of Florida as well as all the Southern States along the Gulf of Mexico. Other areas are the state of Georgia and Southern Carolina. To the West state, Arizona is the only state with GFD levels as high as 8 flashes/km²/year.
- Rakov *et al* [5] said that lightning discharges produced by winter storms in the coastal area of the sea of Japan exhibit a number of features that have not been observed during the summer months or in any season in other geographic locations.
- The Iberian Peninsula is also distinguished by its high incidence of lightning due to elevation.
- Sariano and De la Rose [6] indicated that the most frequent lightning activities tend to occur near the Mediterranean coast where warm, humid air increases convection.

The Brass Coast like any other coastal area in the globe is prone to high ground flash density, therefore the study was necessary to have a lightning record for this area.

Stochastic Characteristics of Lightning Strokes

The most important lightning return stroke parameters are:

- Peak current, I_p
- Current front time t_f
- Return stroke velocity V
- Ground flash density

These parameters are stochastic in nature. Analysis of field data shows that the statistical variation of the peak current, I_p and current front time t_f of the return stroke current fit lognormal distribution [7]. The probability density function of I_p can be expressed as:

$$\rho(I_p) = \frac{e^{-0.5 f_1}}{\sqrt{2\pi} \cdot I_p \cdot \sigma_{\ln I_p}} \quad (1)$$

$$\text{Where } f_1 = \left(\frac{\ln I_p - \ln I_m}{\sigma_{\ln I_p}} \right)^2$$

Similarly, the probability density function of t_f can be expressed as:

$$\rho(t_f) = \frac{e^{-0.5 f_2}}{t_f \cdot \sigma_{\ln t_f} \cdot \sqrt{2\pi}} \quad (2)$$

$$\text{Where } f_2 = \left(\frac{\ln t_f - \ln t_m}{\sigma_{\ln t_f}} \right)^2$$

The joint probability density function, $p(I_p t_f)$ is given by

$$\rho(I_p t_f) = \frac{e^{-\frac{0.5}{1-\rho}(f_1 - 2\rho\sqrt{f_1 \cdot f_2} + f_2)}}{2\pi (I_p \cdot t_f) (\sigma_{\ln I_p} \cdot \sigma_{\ln t_f}) \sqrt{1-\rho}} \quad (3)$$

Where ρ is coefficient of correlation

I_m is mean current

t_m is mean time

σ is standard deviation

The statistical parameters of return stroke current are as follows [8, 9]

Median time to crest, $t_{fm} = 3.83 \mu s$, Log (to base e) of standard deviation $\sigma(\ln t_f) = 0.553$

Median Peak Current $I_{pm1} = 61.1 kA$, Log (to base e) of standard deviation $\sigma = 1.33$

Median Peak Current I_{pm2} is 33.3 kA, Log (to base e) of standard deviation $\sigma = 0.605$

The coefficient of correlation $\rho = 0.47$.

Some authors [7, 8] proposed the following empirical relationship between the return – stroke peak current and its velocity as

$$V = \frac{c}{\sqrt{1 + \frac{500}{I_p}}} \quad (\text{m/s}) \quad (4)$$

Where c is velocity of light.

Besides the I_p and t_f , the ground flash density is the next significant parameter in estimating the lightning performance of power system. The ground flash density n_g varies regionally and seasonally and the geographic region varies by a large margin from year to year usually it is a long-term average value and ideally it should take into account the yearly variation that take place within a solar cycle-believed to the period within which all climate variation that produce different GFD level occur.

The ground flash density can generally be defined as the average number of lightning flashes per square kilometer per year in a geographic region.

De la Rosa [4] discussed how to determine the ground flash density as a function of thunder day (T_α) or Keraunic level or thunder thour (T_h). This is important where GFD is not available from the lightning location.

Basically any of these parameters can be used to et a rough a approximation of GFD. Using the expression described in Anderson et al [9] and MacGorman et al [10] respectively.

The ground falsh density n_g is

$$n_g = 0.04T_\alpha^{1.25} \text{ flashes } 1\text{km}^2/\text{year} \quad (5)$$

$$n_g = 0.05T_h^{1.1} \text{ flashes } 1\text{km}^2/\text{year} \quad (6)$$

Methodology

Meteorological data (pressure, daily temperature and thunder storm) were taken from 2008 to 2012. In order to have records for some other location, records were taken from the College of Education. The records available was from 2010 to 2012 and some of the records not detailed.

These data were compared to the data taken from the different centres. From the given records the five years lightning days were tabulated as shown in table 1.

Result and Analysis

A flash which is made up of one or more strokes may take place between two clouds known as cloud to cloud (CC), flashes within the same cloud is inter cloud (IC) and flashes from clouds to ground is known as cloud to ground (CG).

If the CC to CG is 3:1, then

$$\text{The number of lightning days (CG)} = \frac{\text{Total Number of Thunder Storm day}}{n}$$

The ground flash density was calculated from the equation 5.

From table 1, all results (the number of lightning days, GFD) are shown.

Table 1: Thunderstorm Days For The Brass Coast and Inland site.

MONTH	BRASS COAST					BRASS INLAND SITE				
	2008	2009	2010	2011	2012	2008	2009	2010	2011	2012
JAN	7	4	6	2	8	4	2	3	6	8
FEB	28	16	10	26	17	15	12	12	20	18
MAR	63	50	48	57	48	40	48	38	50	50
APR	70	61	60	62	63	42	42	37	64	48
MAY	65	57	55	60	59	50	47	49	52	55
JUN	54	44	39	50	51	35	40	36	45	45
JUL	40	30	36	42	37	32	29	25	37	30
AUG	28	21	27	30	25	21	30	22	18	16
SEP	42	36	29	40	36	20	20	18	36	19
OCT	62	49	54	52	40	41	50	42	50	35
NOV	20	23	18	24	28	20	18	16	21	20
DEC	6	8	66	4	7	8	6	4	5	6
TOTAL NO. OF FLASHES	485	399	388	449	419	328	344	302	404	350
FLASHES CG	122	100	97	112	105	82	86	76	101	88
GFD		13	12	15	13.5	10	11	9	13	11

In fig. 1, the diagram given the variation of lightning days for the period of five year in the Brass Coast.

From the given record, the average GFD is

- (i) The Brass Coast average GFD is 14
- (ii) Inland site average GFD is 11.

The average annual temperature and relative humidity (2008 – 2012) were recorded in table 2 and 3 respectively.

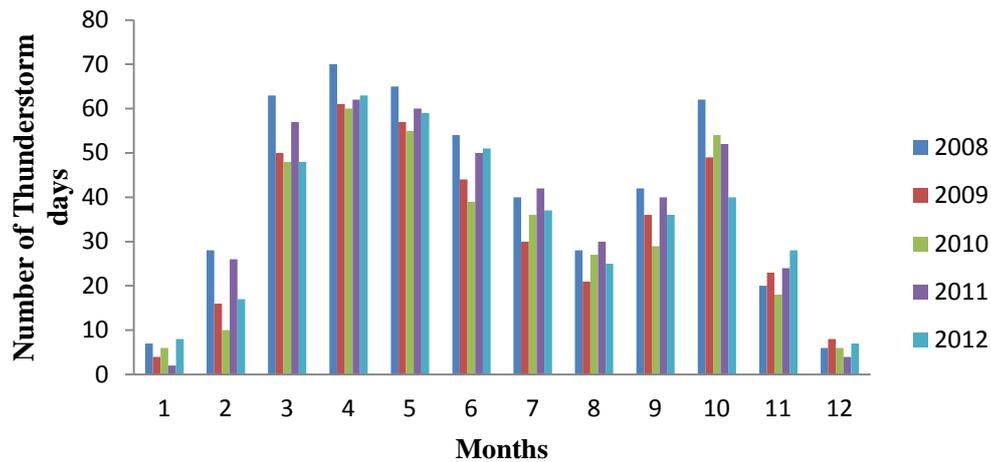
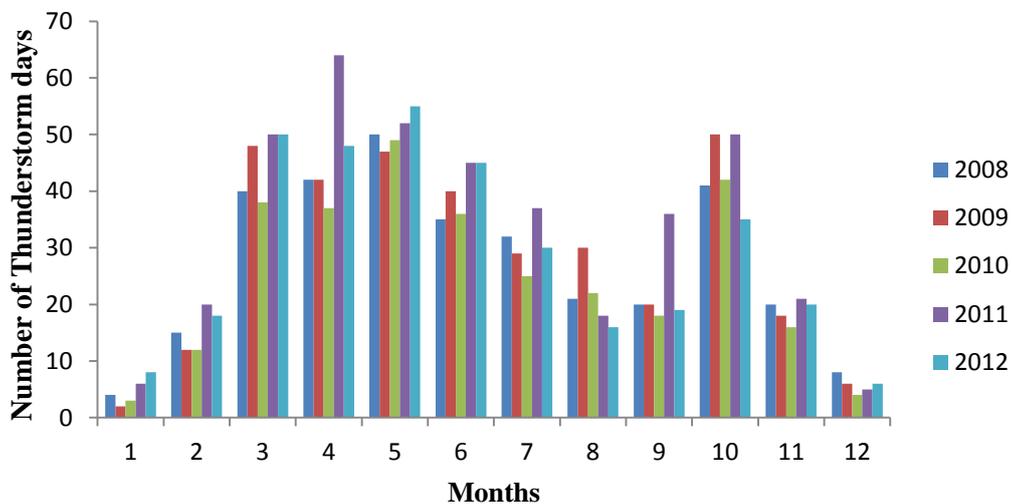
Table 2: Mean Annual Temperature (Average for 2008 – 2012)

MONTH	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEPT	OCT	NOV	DEC
MAX	33.8	33.8	33.9	32.8	32	30.4	29.4	29.2	29.8	30.8	32.4	33.5
MIN	21.5	23.8	24.0	23.5	23.1	22.8	22.1	22.8	22.9	22.9	23.3	23

Table 3: Annual Relative Humidity (Average for 2008 – 2012)

MONTH	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEPT	OCT	NOV	DEC
RELATIVE HUMIDITY	75	79	80	84	87	89	90	91	90	89	85	76

From the results, yearly thunderstorm days was presented in a histogram as shown in figure 1.

**(a) Brass Coast****(b) Brass Inland site****Figure 1: Histogram of Thunderstorm days for each month of the years 2008 – 2012.**

Discussion

Thunderstorm may be as a result of orographic uplift that occur in mountainous area. Mountain slopes are heated more rapidly than valley causing air to rise against the slope initiating vertical development of convective air mass. Thunderstorm is also associated with cold front. Warm air is displaced as cold air masses move underneath at frontal surface resulting in rigorous uplift. This explains the more frequent thunderstorm days in March, April and May and at October. These are periods of transition from dry sunny seasons to the rainy seasons and vice-versa [11]. These periods are marked with sever and prolonged thunderstorms with multiple flashes. The rainy periods (April to

September) are periods of heavy rain with high humidity (90 percent) and the thunderstorm are more devastating. It is obvious that more cloud to ground lightning are experienced.

Diendorfer et al [12] conclude that the percentage of single stroke flashes vary significantly from storm to storm even in the same region and depend probably on seasons and type of thunderstorm. In the Brass coast is more significant that lightning activity occurred mainly overland mass with an average land to ocean ratio of 10:1. This explains the marked difference of the ground flash density of the Brass coast to the inland site.

Conclusion

Lightning parameters are the bases for the design of lightning protection equipment and for the calculation of lightning radiated fields and their interactions with power and telecommunication lines. Lightning peak current is one of the most important lightning parameters and all national and international standards are based on lightning current measurements. Modern multiple station lightning location systems are necessary because their output besides lightning co-ordinates, includes estimates of lightning peak current and number of strokes per flash (multiplicity).

Also, some location systems masses the first stroke and misidentifies the second stroke as the first one. Statistical evaluations based on data for a few storm day may not be a true representation of the total lightning activities in a region therefore frequent measurements and for several years are necessary. Identifying the number of thunderstorm days may only serve some geographical purpose but for scientific data and analysis more modern lightning location systems are necessary in the country. More so, most systems do not record the lightning with low frequency.

Recommendation

It is necessary to recommend that the government, scientific organizations and institutions of learning need to encourage researchers in this area in the country to have reliable records.

References:

- Alagoa, E.J. *The Land and People of Bayelsa State, Central Niger Delta*, Onyoma Research Publisher, Port Harcourt, 1990.
- Bhavika, B. *The influence of Terrain Elevation on Lightning Density in South Africa*, Master Thesis, University of Johannesburg, South Africa, 2007.
- Christian, H.J. Global Lightning Activity, Proceedings of the 12th International Conference on Atmospheric Electricity Versailles June 9th – 13th 2003, France, 2003.
- De La Rose, F. *Characteristics of Lightning Stroke Electric Power System*, CRC Press – Taylor and Francis group, New York, 2006.
- Rakov, V.A., Uman, M.A. and Thottapillil, R. Review of lightning properties from electric field and TV observations. *Journal of Geophysical Research*, 99(D5): 10745 – 10750, 1994.
- Sariano, L.R. and De La Rose, F. Study of Lightning event duration and flash rate in the Iberian Peninsula using cloud to ground lightning data. *Journal of Atmospheric Research*, 61: 189 – 201, 2002.
- Lundholm, R. Induced overvoltage surge on transmission lines and their bearing on the lightning performance at medium voltage network, *Transaction on Chalmers University*, Gothenburg, Sweden, 188: 1 – 117, 1957.
- Rusck, S. Induced lightning overvoltage on power transmission line with special reference to the overvoltage protection of low voltage networks, *Transmission on Royal Institute of Technology*, Sweden, 120: 1 – 118, 1958.
- Anderson, R.B. and Eriksson, A.J. Lightning parameters for engineering applications, *Electra*, 69: 65 – 102, 1980.
- MacGorman, D.R., Maier, M.W. and Rust, W.D. Lightning strike density for the contiguous United States from thunderstorm duration records NUREG/CR – 3759, Office of Nuclear Regulation Research, U.S. Nuclear Regulatory Commission, 44 Washington DC, 1984.
- Afa, J.T. Lightning Activities and Ground Flash Density in Niger Delta Coast, *Journal of Engineering and Applied Science*, 7(4): 339 – 341, 2012.
- Diendorfer, G., Schulz, W. and Rakov, V.A. Lightning Characteristics based on data from the Austrian Lightning locating system. *IEEE Transactions on Electromagnetic Compatibility*, 40(4): 452 – 464, 1998.

CREATING PUBLIC ENVIRONMENTAL AWARENESS THROUGH MUSIC: “PLAY FOR NATURE” PROJECT

A. Banu Bıçakcı, Asst. Prof. Dr

Yeditepe University, Turkey

Abstract:

This paper examines a voluntary initiative named ‘Play for Nature’ that utilizes music as a mediator to create public environmental awareness. Strengths and weaknesses of the project are scrutinized from a public relations viewpoint. Data is collected via related media coverage followed by an online interview held by the initiator and organizer of the project in Turkey, Fırat Çavaş. It is concluded that the project that is able to reach millions of people, by using the powers of music and social media, should be managed with a strategic communication approach, as to fulfill the main objective. The study offers Play for Nature an applicable communication frame for the future chapter of the project.

Key Words: Environmental awareness, music, strategic communication, Play for Nature, corporate social responsibility

Introduction

The global climate change and environmental destruction have been raised to a level of major concern in recent years. However, public environmental awareness and education still need to be enhanced as to take significant precautions against all anti-environmentalist activities. Since creating awareness is an important first step toward building audience understanding, influencing opinion and motivating behavior; environmentalist movement should take creative activities, utilizing arts as a way to communicate with the audiences, as an opportunity to create environmental conscience.

This paper examines a voluntary initiative named ‘Play for Nature’ that utilizes music as a mediator to create public environmental awareness, with a rhetorical perspective. Main objective of this study is to figure out how a civil initiative may be successful in reaching its target audiences without support of any corporation or a governmental body, only by using musicians as opinion leaders and role models. Additionally its media selection is questioned, since the project’s utilization of traditional media is very limited and the overall communication is based upon social media. Strengths and weaknesses of the project are examined from a public relations viewpoint. Data is collected via related media coverage followed by an online interview held by the initiator and organizer of the project, Fırat Çavaş, in September 2012.

Environmentalism and Public Environmental Awareness

Environmentalism, as a social movement and political ideology, varies from the anthropocentric pole to the eco-centric pole. Resource conservation, human welfare ecology, preservation, animal liberation and eco-centrism are major environmentalist streams (Eckersley, 1992:34). Anthropocentrism considers humans as the center of the universe and the top of all creation; human being is the dominator of the nature and natural environment is created for human usage. Anthropocentric attitudes are deeply rooted in Western culture whereas eco-centric form of public awareness is opposed to human-centered system of values while it denotes a nature-centered worldview (Rowe, 1994). Regarding Turkey’s geo-political position and culture, Turkish people’s environmental attitudes are assumed to be in between anthropocentrism and eco-centrism.

Even though both developing and developed countries are facing serious environmental problems (Hoerisch, 2002), the tendency to give priority to environmental goals is much stronger in wealthy countries than in poorer ones (Diekman and Franzen, 1999:548). Countries with higher economic ratings and countries successful in dealing with the local environment are expected to place more emphasis on global environmental problems. According to a global research made by Dunlap, Gallup and Gallup in 1993, for 61% of the Turkish population, environment is a ‘very serious’ issue,

however only 18% of the people think that environment is the most important problem in Turkey (Diekman and Franzen, 1999). In 2003, a nationwide study on environmental attitudes of Turkish people has been accomplished (Tuna, 2004). The research includes environmental issues such as environmental politics, environmental knowledge, environmental degradation, pollution and energy usage. This study shows that highly educated respondents and the ones with prestigious occupations have more environmental commitment (Tuna, 2004). More recent studies on environmental attitudes of Turkish people are mostly focusing on the attitudes of young generations. As concluded in Tuncer et. al (2005), there is a widespread support for conservation of the environment among high school students in Ankara (capital city of Turkey). Another study shows that Turkish university students in Konya (one of the biggest cities in Turkey) are concerned about environmental issues, however only 8,53% of them regularly follow the related publications and just 7,73 of them hold a membership to an environmental organization (Önder, 2006).

Concerning the sensitivity of businesses to environment, it is possible to recall the environment oriented corporate social responsibility (CSR) activities. In Turkey, although the major CSR topics include education and societal issues such as healthcare, there are serious CSR projects on environmental issues, as well. Many corporations are sharing sustainability reports or index results with their stakeholders. Today, there are 226 participants of United Nations Global Compact¹, from Turkey

http://www.unglobalcompact.org/participants/search?commit=Search&keyword=&country%5B%5D=202&joined_after=&joined_before=&business_type=all§or_id=all&listing_status_id=all&cop_status=all&organization_type_id=&commit=Search, yet there is not enough data on these organizations' overall environmental attitudes and behavior. A research made particularly on hotel business in Turkey has found that the policies and practices of Ankara hotels generally lack attributes relevant to environmental protection and conservation. The authors concluded that the hotel managers mostly lack the necessary environmental knowledge and interest to meet the objectives of societal and environmental responsibility (Erdoğan and Barış, 2007).

Another research topic deals with the effectiveness of various communication channels utilized in creating public environmental awareness. A comparative study proves that the role of media in creating environmental awareness is definitely the dominant one compared by personal experiences in daily life, personal discussions and formal education at schools, both in developing and developed countries (Hoerisch, 2002). A recent study investigates the use of integrating mobile technologies, data services and multimedia messaging services (MMS) to increase university students' use of mobile technologies and to develop environmental awareness. The study concludes that these new technologies are useful in developing environmental awareness among young people (Uzunboylu, Çavuş and Erçağ, 2009).

According to United Nations (UN) education program there are several ways to increase public environmental awareness and education. Promoting responsible action in the community through the media by involving key public players, decision-makers and opinion-builders in such campaigns is one of these ways. <http://www.unep.org/DEC/OnLineManual/Enforcement/InstitutionalFrameworks/PublicAwarenessEducation/tabid/99/Default.aspx>. A project called 'Play for Nature' aims to increase the level of awareness towards environmental issues, by making celebrities involved in their project, in which music is a mediator between this initiative and the public.

Play for Nature Project: The Need for a Strategic Communication

'Play for Nature' is not an association or a foundation; it is a project realized by more than 200 volunteer, amateur and professional musicians. Inspired by an American originated project called 'Playing for Change' <http://playingforchange.com/>, 'Play for Nature' project was developed in 2008, in Turkey, with the objective of creating and enhancing public awareness of the environment and natural life. This is a civil initiative which tries to take attention to the environmentalist NGOs

The UN Global Compact is a strategic policy initiative for businesses that are committed to aligning their operations and strategies with ten universally accepted principles in the areas of human rights, labour, environment and anti-corruption <http://www.unglobalcompact.org/AboutTheGC/index.html#environment> and <http://www.unglobalcompact.org/AboutTheGC/index.html#anti-corruption>

activities and to help raise their funds <http://www.dogaicincal.com/index.asp?sayfa=anasayfa>. While music is a universal language that unites everybody and helps product/idea recall and recognition (Stewart and Punj, 1998), to reach the highly heterogeneous target audience via music seems to be a proper tactic. They explain their essence as follows:

Our starting point was the shared, universality of environmental problems and music. Nature as a provider of housing and resources has a strong connection with music, which is a common language of all human beings. [...] Now, it's our responsibility to find solutions and carry them into action; we should remember and remind people of this every chance we get. Despite knowing that change is a necessity, we still don't seem to listen or do what we're supposed to do. So, can a song then be useful and remind us of what we should do every time we listen to it? We hope that it may at least bring some awareness and consciousness to the people who listen to it. <http://www.dogaicincal.com/index.asp?sayfa=playfornature>

As Çavaş claims, this is a project to raise awareness; they do not have a concrete aim as collecting funds or planting trees (Diker, 2012:52). Çavaş calls Play for Nature a social responsibility project with the goal of directing people to get information from nature associations. So, a number of nature associations are their 'natural' stakeholders. However, he is complaining that other than TÜRÇEK (one of these associations) none of them has responded and none of them has sent an e-mail of appreciation. In addition to that, none of the existing associations has had any intention to meet and cooperate with them, even though this project is realized to help fulfill their objectives. Other stakeholders of the project are media, volunteers and sponsors. This volunteer project makes barter agreements with their sponsors; a software company provides the software they need, a technology company gives them the necessary equipment and all the rest, from recording to mixing, is handled by Çavaş himself (Diker, 2012:58).

Play for change is a social responsibility project. Any social responsibility campaign should start with research. This research includes the search for the organization itself, the current situation and the publics (Lattimore et al., 2009). Çavaş mentions that, Play for Nature, claiming to be an amateur social responsibility project, have neither defined its target publics nor made any research on the public's previous environmental attitudes. This is the major weakness of the project, which disables them to have clear objectives.

The project takes the advantages of social media to spread their videos to the public. Facebook and Twitter are the most utilized media. Via social media, the project can reach the public easily and quickly; however it has some disadvantages as Çavaş claims: Being easily criticized by people. There are interesting comments reaching them over the social media, for instance one of the audiences has criticized the musicians' instruments being made from wood, highlighting the paradox between protecting the Mother Nature and playing wooden instruments. Some traditional media, including the major national TV channels (TRT, CNN Turk, Show TV, ATV, Kanal D, TV 8, Cine 5, Fox TV, Dream TV, Kral TV) and radio channels (TRT, Radyo Klas) are supporting the project, as well <http://www.dogaicincal.com/index.asp?sayfa=biz&pid=3> . Conducting a communication audit for understanding the channel effectiveness would be beneficial, since there has not been such an attempt. Media planning should be based on research outcomes.

There are many celebrities that are voluntarily supporting the project, by singing the project's songs. Among these names, there are well known singers and actors of Turkey such as Levent Yüksel, Cahit Berkay, Gökçe and Levent Üzümcü. Celebrities are opinion leaders for the public and they have the power to take the public attention to a certain subject. Thus, their support and participations are significant to reach masses and affect their attitudes positively on environmental issues. Since the volunteer celebrities are only supporting the project with their voices as Çavaş mentions, it would be better to lead them to use their social media accounts as to spread the project and give environmentalist messages. It is noticeable that Play for Change project, except its name, which actually is not a message but the theme of the initiative, does not have clear messages for their public. Without messages, it is not possible for the target public to understand their real objective and the major peril is that the audience will listen to the songs just for amusement. Designing campaign messages consistent with the soul and main goal of the project is a part of strategic communication approach, and it would be useful in realizing the objectives.

Regarding the evaluation of the project, Çavaş emphasizes that 'Play for Nature 4' project's video has had 2 million clicks in 5 months time, from a single link on the YouTube. The project gets

lots of positive feedback from primary and high school teachers, as well. Play for Change is adored because of its musical quality, synthesizing a lot of music genres and for making folk songs popular among the younger audience (Diker, 2012:56). As visible from the comments, the positive feedbacks are irrelevant to environmental issues, which are primary aims of the project. While the project is lacking objectives depending on existing data, it is not possible to make a real evaluation to measure the success. The question ‘How much this project has contributed to public environmental awareness?’ is still without an answer.

Conclusion

Climate change, global warming or other environmental issues are not priorities for the Turkish government and major corporate social responsibility activities are focused on education and health issues in Turkey. As a result of minimal media coverage on environmental issues and lack of critical perspective on controversial topics such as nuclear power and ecological risks, Turkish public is not enough aware of the environmental threats. To find a common ground for those who are not aware of the environmental perils and the urgency of creating a positive behavior towards the environment is quite problematic. Play for Nature is an amateur initiative that aims to raise the public environmental awareness through music. However the study exposes that the project lacks a strategic communication approach to help meeting the objectives.

While this amateur and well-intentioned project seems to reach huge masses, it is not clear that whether they have helped people creating positive attitudes towards environment. It would be beneficial to conduct a strategic communication plan for the future versions of the project including the following points:

- Organizational analysis: What are the capabilities of the organization? What are the strengths and weaknesses? Who are the volunteers? Who are the major stakeholders? What is the major problem with the environmentalist NGO’s: defining the problem and providing an ongoing communication with them. An analysis of the feedbacks for the previous chapters of the project.
- Situational analysis: A scientific research on the present environmental attitudes of the public. It might be sponsored by one of the relevant volunteers.
- Audience research: Defining the audiences: Demographic and psychological factors included.
- Defining long term and short-term goals depending on the research findings: What will be the main strategy and the tactics? What are the sociologic, economic, political and cultural barriers? How can they be surpassed?
- Designing the messages for various publics: Relevance of the lyrics and the messages should be considered
- Media selection: Scientific media planning according to the research findings. Integrating social media with the traditional media would be useful since the target audience is too broad. Mobile technologies should be considered.
- Opinion-leaders: Making volunteer celebrities the project ambassadors. Provide a sustainable support of them for the initiative, not only once by singing a song. Persuading them to talk about the project both in traditional and digital media.
- Evaluation: Measuring the level of success periodically and if necessary, re-designing more realistic and measurable goals.

As a conclusion, Play for Nature is successful in reaching the people and attracting their attention, however, the project is lacking a strategic communication approach as to reach the basic objective of creating public environmental awareness. This study proposes above-mentioned outline to manage the communication of the future chapters of Play for Nature strategically.

References:

- Diekman, Andreas. and Franzen, Axel. The Wealth of Nations and Environmental Concern, *Environment and Behavior*, 31(4), 540-549. 1999.
- Diker, A.H. Hepimiz için Çalışıyorlar- They Play for All of Us, *Anadolu Jet Magazin*, 35, 52-58. 2012
- Eckersley Robin. *Environmentalism and Political Theory: Toward an Ecocentric Approach*. NY: State University of New York Press. 1992.

- Erdoğan, Nazmiye and Barış, Emin Environmental Protection Programs and Conservations Practices of Hotels in Ankara, Turkey, *Tourism Management*, 28(2), 604-614. 2007.
- Hoerisch, Hannah. A Comparative Study on Environmental Awareness and Environmentally Beneficial Behavior in India, Report for CMS ENVIS Centre. 2002. online access: <http://cmsenvis.cmsindia.org/researchstudy/beneficial.pdf>
- Lattimore, Dan, Baskin, Otis, Heineman, Suzette T. and Toth, Elizabeth.L. *Public Relations: The Profession and the Practice*, 3rd ed., New York: McGraw-Hill. 2009.
- Önder, Serpil. A Survey of Awareness and Behaviour in Regard to Environmental Issues among Selçuk University Students in Konya, Turkey, *Journal of Applied Sciences*, 6(2), 347-352. 2006.
- Rowe, Stan J. Ecocentrism: the Chord that Harmonizes Humans and Earth. *The Trumpeter*. 11(2), 106-107. 1994.
- Stewart, D.W. and Punj, G.N. Effects of Using a Nonverbal (Musical) Cue on Recall and Playback of Television Advertising: Implications for Advertising Tracking, *Journal of Business Research*, 42, 39-51. 1998.
- Tuncer, Gaye, Ertepinar, Hamide, Tekkaya, Ceren and Sungur, Semra. Environmental Attitudes of Young People in Turkey: Effects of school type and gender, *Environmental Education*, 11(2), 215-233. 2005.
- Uzunboylu, Hüseyin Çavuş, Nadire and Erçağ E. Using Mobile Learning to Increase Environmental Awareness, *Computers & Education*, 52(2), 381-389. 2009.
- Tuna, Muammer. Public Environmental Attitudes in Turkey, *Proceedings of 3rd Global Conference on Ecological Justice and Global Citizenship*, 2004. online access: www.interdisciplinary.net/ptb/ejgc/ejgc3/tuna%paper.pdf.

ANTHROPOGENIC IMPACT CONDITIONS ON WATER QUALITY IN THE ALLUVIAL AQUIFER OF THE REGION TADJENANET - CHELGHOU M LAID (EASTERN ALGERIA)

Abdelhamid Khedidja

Laboratory of natural hazards and land use. University of Batna, Algeria

Abderrahmane Boudoukha

Research Laboratory of Applied Hydraulics. University of Batna, Algeria

Abstract:

The region of Tadjenanet - Chelghoum Laid is located in the upper valley of Oued Rhumel. It has shown in recent years an important agricultural and industrial development, which resulted in increased occupancy of the natural environment and therefore a deterioration of water quality of surface and underground aquifer located in the alluvium of Mio Plio Quaternary.

This study aims to determine the critical impact of natural and anthropogenic pollution on physico-chemical water of the shallow aquifer zone of Tadjenanet - Chelghoum Laid.

Chemical analysis of this water showed a rather marked salinity, due to dissolution and leaching of surrounding formations, carbonate, and gypsiferous alluvial. The high concentration of nitrate in irrigated areas at the periphery of Oued Rhumel, reflects the agricultural activities, marked by a wide variety of crops, marked by an unmanaged employment of chemical fertilizers, especially nitrogen.

Key Words: Risk, pollution, groundwater, alluvium, aquifer, nitrate, Tadjenanet- Chelghoum Laid

Introduction

Diffuse pollution from agriculture is a major cause of the deterioration of groundwater quality in rural areas. Nitrates and pesticides are the main source of pollution of aquifers in the region Tadjenanet - Chelghoum Laid [7]. The pollution risk is accentuated further by the lack of protective cover causing pollutants directly into the reservoir formation [3].

Contamination of groundwater in the alluvial aquifer of Tadjenanet Chelghoum Laid by nutrients (nitrate and ammonium) is often related to misuse of fertilizers in agriculture, the decomposition of plant materials or animal manure, and domestic wastewater and industrial applications [8].

Through a piezometric and hydrochemical studies in the unconfined aquifer and from the results of physico-chemical analysis have provided the data necessary to define the risk caused by the phenomenon of pollution.

In this study we tried to assess the risk of groundwater contamination by nitrates. We should also mention that the concentration of nitrate in groundwater is changing at an alarming rate [7].

Agricultural development, often leads to rapid pollution of groundwater by the chemical forms of nitrogen, especially of nitrates, due to their high solubility and low affinity for ion exchange [8], [10]. Nitrates are the most oxygenated nitrogen and are a very soluble form; they are the result of the nitrogen cycle which is a nutrient essential to plant life. The spatial variation of nitrate concentrations is mainly related to agricultural activity that develops on the surface, the nature of the formations in the unsaturated zone and the conditions of oxidation-reduction [3].

Materials And Methods

Geological and hydrogeological settings

Region Tadjenanet -Chelghoum Laid covers an area of 1130 km². It is part of the western watershed of the great Kabir Rhumel located in eastern Algeria, which drains much of the southern slope of the tell Setif (Figure 1).

The plain is bounded by a ridge whose edges consist of mountains peaks [Dj.Tnoutit (1189m), Dj.Tafrent (1069), Dj. Ed Dess (1212m) and Dj Grouz (1188m)]. The topography of the plain is

nearly flat with a slope not exceeding 2%, and altitudes ranging from 920m to 720m North West and the South East. This morphology is monotonous sometimes interrupted by a few hills scattered across the plain. The area is mainly covered by formations of the Mio-Plio-Quaternary corresponding to clays, marls, silts, alluvium and calcrete. The Eocene limestone outcrop mainly on the boundaries but also in some places in the center of the plain (Figure.2). The region is drained primarily by Oued Rhumel and some secondary wadis as El Mehri, and Ouskourt in the south and Maamra, Boutouil, Boumrah, and Dekri in the north. The main river is controlled by the dam of Hammam Grouz in the North East. The average annual rainfall reaches the 372 mm estimated over a period of sixteen years (1988/89 - 2003/2004) at the station Grouz Hammam. Establish the water balance using Thornthwaite and relationships Tixeront-Berkaloff shows that the ETR is the order of 343 mm contrary infiltration does not exceed 6 mm, while the runoff can reach 24mm.

The various geological, geophysical and hydrogeological activities in the region [7], [2], [1] and [11] showed the existence of three permeable geological formations.

- An aquifer in the fractured limestone formations of the Cretaceous that characterize tablecloth constantinian neritic and all allochthonous South Setif. [7], [2], [6].

- A shallow aquifer in the Quaternary alluvium developed at Oued Rhumel which is in destocking phase.

- A continuous aquifer formations in fluvio-lacustrine Mio-Pliocene age.

The first part covers the entire North Eastern region in the Cretaceous carbonate massif of Djebel Grouz. The second aquifer is thin and closely linked to waterways. In times of low water, it feeds the rivers, on the contrary in times of high water the reverse occurs. Along the Oued Rhumel, this sheet is partly fueled by neritic limestone aquifer of Cretaceous age. However, it is interesting to note that some of the alluvial aquifer has high mineralization following an intense evaporation, [7]. Its groundwater level is relatively close to the surface, from 0 to 3m [5], it has little interest in terms of hydrogeology.

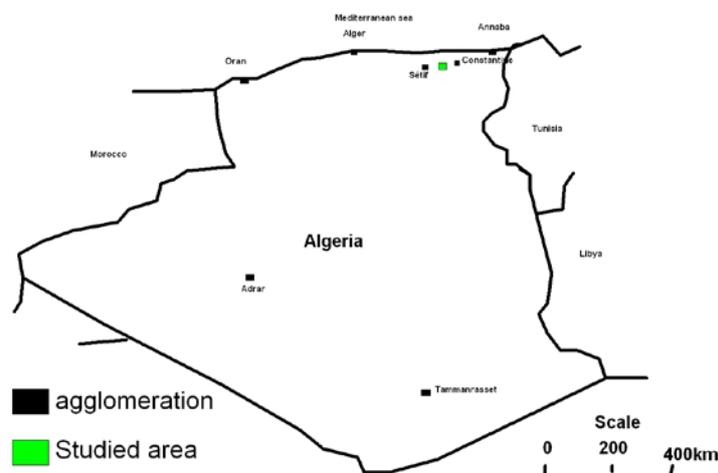


Figure.1 Location of the studied area

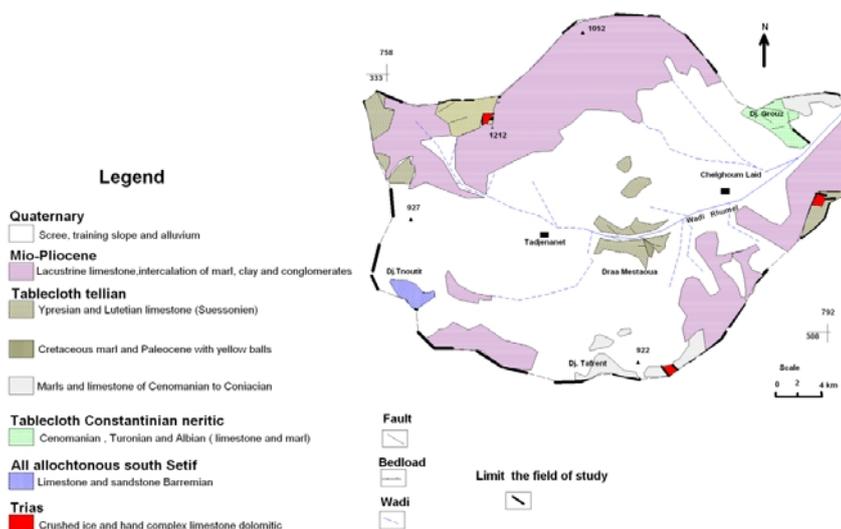


Figure 2 . Geological map of the sub - basin of Tadjenanet - Chelghoum Laid

According to J.M.Vila -1977

For the third aquifer located in the formations of the Mio-Pliocene has good potential. The thickness of the aquifer varies from 100 to 150m and consists of fluvio-lacustrine usually with reddish coloration quite pronounced. These formations crop out much more on the periphery of the watershed. These lacustrine limestone, red marl, silt and sand red. (Figure.3).

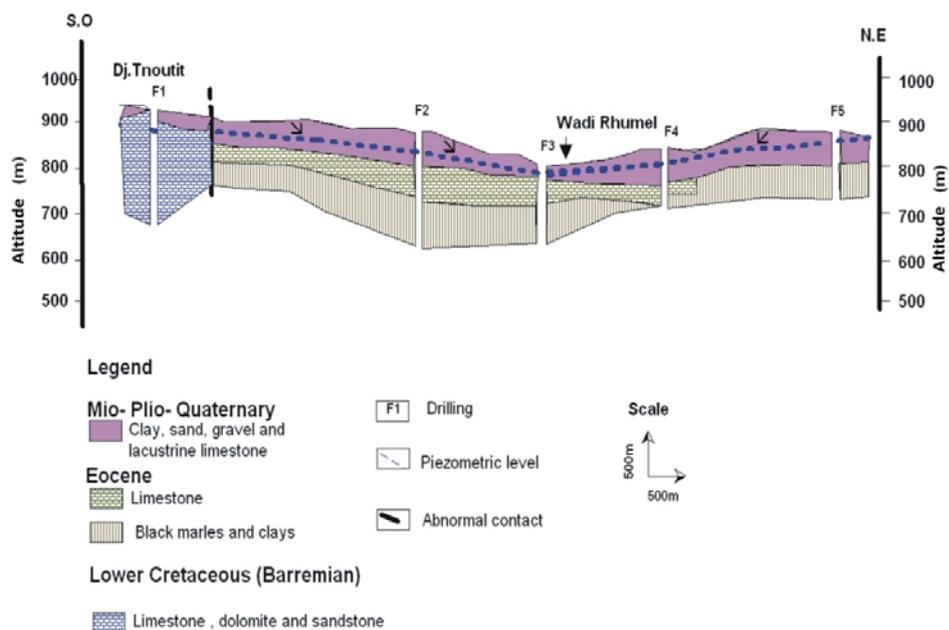


Figure 3. Hydrogeological section

The state the groundwater was analyzed using two piezometric campaigns, the first conducted at the end of the period of low water (September 2006) and the second at the end of the recharge period (May 2007). Analysis of piezometric maps of two periods (Figure.4) showed that the water has kept the same structure with:

- A limit to inflow from the massive carbonate that surround the plain to the north, south-west and north;
- A limit to outflow to the North East is the outlet of Oued Rhumel;

- A tight boundary east of the field characterized by the presence of impermeable formations and marl. The bedrock is formed by clays and marls;

- A low piezometric confused with the bowl of the dam

The change of transmissivity (Figure 5) was analyzed using nine pumping test conducted at constant flow in small diameter wells distributed over the majority of the during the month of May 2007. The interpretation using the Theis model showed that the values of transmissivity decrease with the flow direction from upstream to downstream. This decrease is due to lateral changes in lithology resulting variation of hydrodynamic parameters of $2.5 \cdot 10^{-4}$ to $1 \cdot 10^{-4} \text{ m}^2 / \text{s}$ [7].

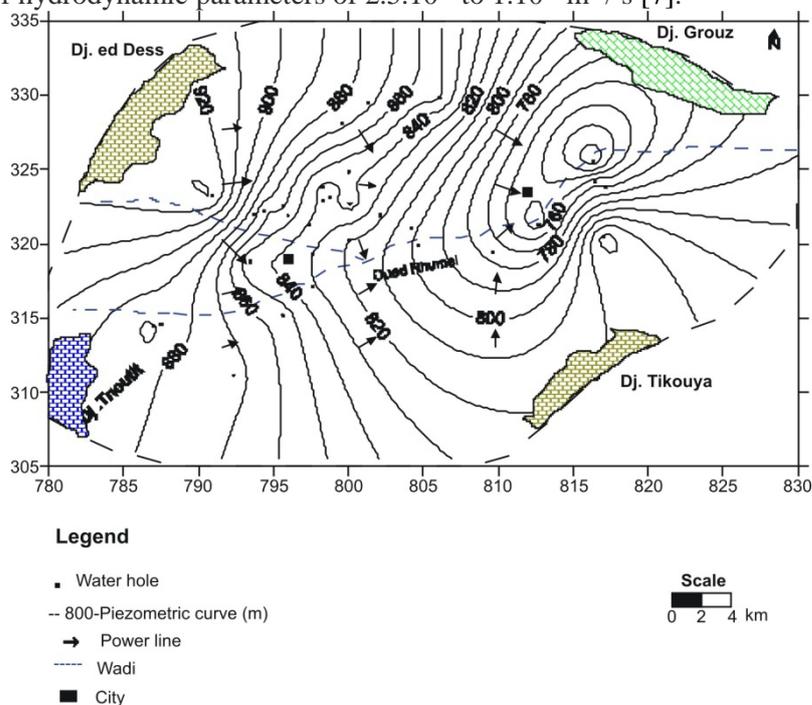


Figure 4. Piezometric map

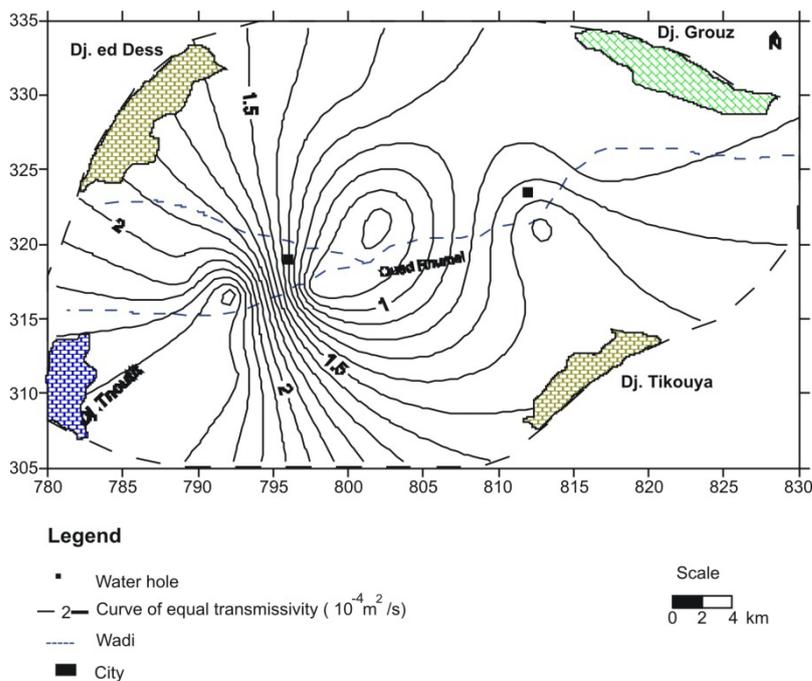


Figure 5. Map of transmissivity in $10^{-4} \text{ m}^2 / \text{s}$

Sample Collection

All hydrochemical work involves collecting water samples, while respecting the collection techniques prescribed. The analysis of water samples can be employed:

- To have an idea about the identity physico-chemical water and its mode of spatial distribution in the environment.

-To show the effects of groundwater quality on the risk of degradation of agricultural land.

-To approach the problem of potable water in the region.

To achieve this goal, a sampling campaign was conducted during the recharge period (May 2007). According to a mesh size of about 1 km², we were able to select 34 samples collected at water points (wells and boreholes) that we have previously selected (Figure.6).

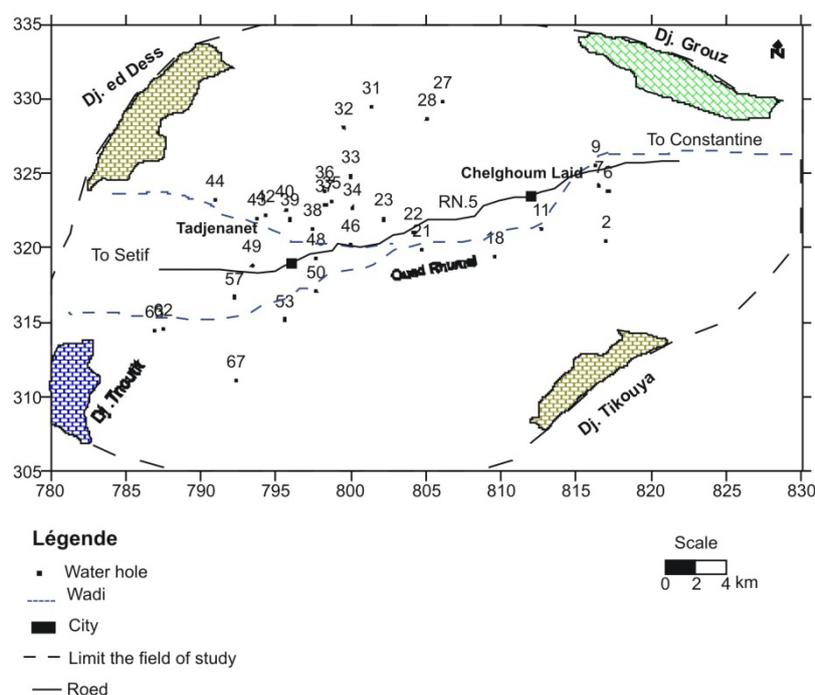


Figure.6 Inventory map of water points have been the subject of a physico-chemical analysis

Chemical analysis were performed at the laboratory of WWTP of Batna and in a private laboratory analysis of soil and irrigation water. The chemical elements were measured were: Ca^{++} , Mg^{++} , Na^+ , K^+ , HCO_3^- , SO_4^{--} , Cl^- , NO_3^- , NO_2^- , NH_4^+ , PO_4^{--} .

The physico-chemical parameters ($T^\circ\text{C}$, pH and conductivity) are measured in situ using a pH meter and a conductivity type WTW. Major ions (Na^+ , K^+ , Ca^{++} , Mg^{++} , HCO_3^- , Cl^- , and SO_4^{--}) were determined by atomic absorption spectrometer for cations, by titration for chloride and alkalinity. Nitrogen elements were determined by spectrophotometry, nitrate in the presence of silicate give paranitosalicylate sodium, colored in yellow (415nm) [9],

Results And Discussion

Water chemistry of the alluvial aquifer is characterized by wide variation in concentrations of chemical elements: Na^+ (51-143mg / l) and Cl^- (88-339mg / l), SO_4^{--} (30-1100mg / l) Ca^{++} (65-304mg / l), Mg^{++} (30-172mg / l), and HCO_3^- (122-512mg/l), NO_3^- (13-486mg/l), NO_2^- , NH_4^+ , PO_4^{--} (0.01-2.2mg/l). (Table I).

The origin of these chemicals is linked to the geological nature of the land to contact with water by dissolution of gypsum in the Triassic and Quaternary formations, dissolution of limestone and dolomite characteristic of the Cretaceous formations at the periphery of the study area. Only nitrogen compounds have a different origin, related to the use of chemical fertilizers and / or organic agriculture and the decomposition of organic matter.

Table I. Change in physicochemical parameters in the alluvial aquifer of Tadjenat Chelghoum-Laid

N° Water hole	pH	Cond µs/cm- à25°C	T°C	Ca ⁺⁺ (mg/l)	Mg ⁺⁺ (mg/l)	Na ⁺ (mg/l)	K ⁺ (mg/l)	HCO ₃ ⁻ (mg/l)	SO ₄ ⁻ (mg/l)	Cl ⁻ (mg/l)	NO ₃ ⁻ (mg/l)	NO ₂ ⁻ (mg/l)	NH ₄ ⁴ (mg/l)	PO ₄ ⁻⁻⁻ mg/l
2	7,52	1846	17,9	175,2	96,35	102,35	11,2	485,5	430	291,3	142,2	0,01	0,1	1,2
6	7,6	1071	18,6	220,3	123,2	109,2	7,8	245,6	375	310,1	34,11	0	0	0
7	7,56	1285	18,4	254,22	111,25	106,3	10,2	310,3	560	302,6	97,9	0	0	0
9	7,5	2530	19	259,71	122,64	115,6	8,9	230	1100	298,2	107,65	0,03	0,1	1,1
11	7,52	3170	16,9	256,54	114,96	143,8	10,2	512,4	640	319,5	134,23	0,02	0,56	0,8
18	7,12	3610	17,5	304,6	172,44	132,7	11,3	340,6	910	337,25	126,7	0,05	0,3	0,5
21	7,2	1529	19	291,61	165,43	65,8	4,5	430,2	530	250,3	58,48	0	0	0
22	7,48	1018	18,1	182,34	109,65	75,3	6,7	245,1	135	185,6	182,07	0,03	0,2	0,5
23	7,52	3020	17	272,54	111,12	121,6	9,9	165,5	910	339,15	486,41	0,05	0,56	2,2
27	7,5	863	17,6	135,4	85,91	75,6	5,8	265,3	280	105,6	126,7	0,01	0,1	0,4
28	7,6	1330	16,8	115,25	65,01	61,2	6,9	201,3	320	185,75	91,7	0	0	0
31	7,54	1698	17,3	125,04	61,32	95,9	9,2	322,4	390	223,56	135,11	0,01	0,4	0,9
32	7,41	1032	15,6	120,24	43,05	48,7	4,3	226,08	50	88,75	100,35	0	0	2
33	7,29	1445	18,5	104,2	50,28	61,6	6,8	256,2	380	142	91,7	0	0	0
34	7,42	1356	15,4	115,81	41,39	51,2	4,5	212,1	45	101,35	174,98	0,02	0,1	0,8
35	7,2	950	15,5	121,84	32,49	59,9	5,8	221,2	30	117,15	73,09	0	0	0
36	7,66	1170	19,2	110,53	41,2	62,3	5,9	231,5	48	109,16	86,38	0	0	0
37	7,5	745	19	95,42	35,07	53,1	4,2	220,6	42	96,58	87,27	0	0	0
38	7,35	1230	15,1	88,17	48	63,6	7	244	48	152,65	94,8	0	0	0
39	7,45	1121	16,9	87,35	39,08	53,7	6	236	46	143,4	93,03	0	0	0
40	7,48	1190	18,2	81,25	38,25	56,2	4,8	255,75	66	148,5	131,57	0	0	0
42	8,27	990	16	96,21	35,21	55,2	4,9	265,2	45	128,6	105,81	0,01	0,2	0,6
43	7,8	1047	16,4	128,25	30,6	56,7	6,1	180,6	29	131,35	98,78	0	0	0
44	8,63	910	17	145,6	45,94	65,2	7,3	220,8	52	158,6	47,84	0	0	0
46	8	1719	18,8	91,15	50,2	70,25	7,5	256,2	48	165,3	57,59	0	0	0
48	8,5	2080	16,6	171,54	87,24	106,3	8,8	475,8	680	294,65	52,27	0	0	0
49	7,9	1717	17	155,42	45,2	93,5	7,9	196,29	51,2	176,4	142,65	0,03	0,1	0,4
50	7,8	1109	17,5	128,25	33,48	81,7	7,8	175,68	34	156,2	116,51	0	0	0
53	7,6	1100	17,5	65,68	96,58	96,33	3,9	135,25	170,6	250,75	22,2	0	0	0
57	7,9	1079	17,4	105,81	46,92	60,8	6,1	170,8	39	142	27,5	0	0	0
62	7,8	983	16,9	96,35	44,53	72,39	7,1	271,2	153,25	126,2	13	0	0	0
63	7,4	1083	17	108,2	45,02	26,35	0,4	122	159,55	180,65	18,31	0	0	0
67	7,5	1330	17,5	86,92	63,42	39,2	1,9	134,75	200,3	163,18	29,35	0	0	0

(May2007)

Facies chemical and potable water

The postponement of the test results on physicochemical digraph logarithmic (Figure.7) shows the existence of two major water types: calcium bicarbonate and calcium sulfate. The calcium bicarbonate facies is the most dominant, with 51% of water samples analyzed. This group has a very high hardness (> 30 ° F), which provides water to the water quality average to poor irrigation with high salinity risk and low risk of sodicity.

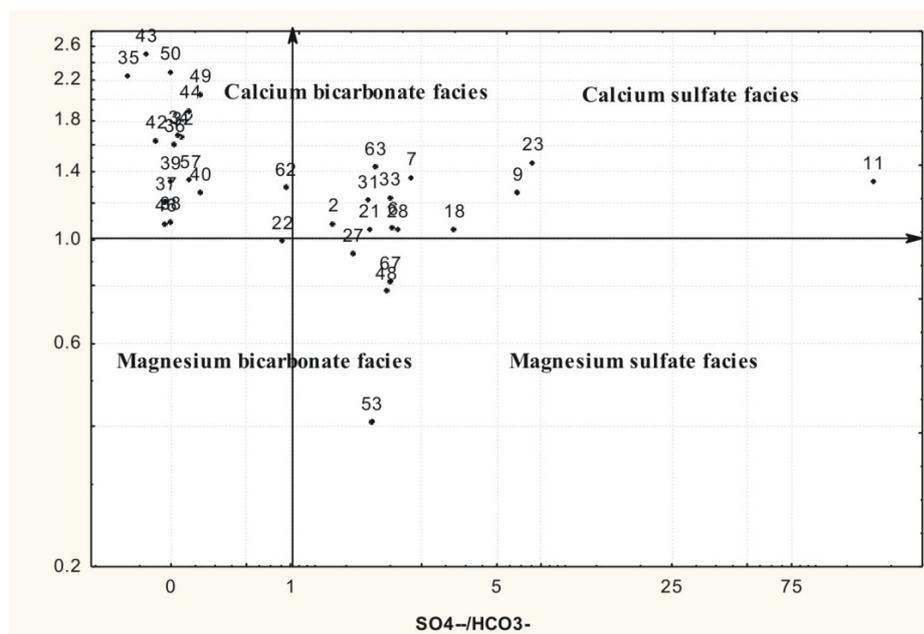


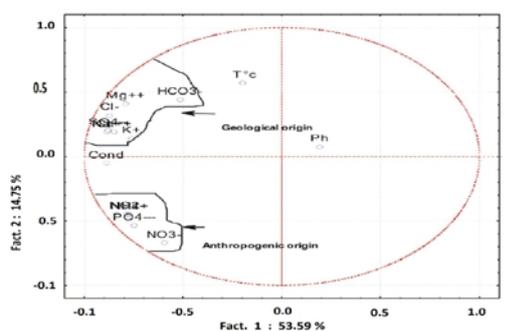
Figure 7. Distribution of facies chemical water points

The principal component analysis was performed on an array of 33 samples and 14 variables: Ca^{++} , Mg^{++} , Na^+ , K^+ , HCO_3^- , SO_4^{--} , Cl^- , NO_3^- , NO_2^- , NH_4^+ , PO_4^{---} , pH, $T^\circ\text{C}$, Cond. Several significant correlations between different chemical elements, with a strong correlation ($r > 0.6$) observed between Ca^{++} , Mg^{++} , Na^+ , K^+ , SO_4^{--} , Cl^- and conductivity. Other less significant correlations were found between conductivity and nitrate ($r = 0.475$) indicating the high use of fertilizers (chemical and / or organic) agriculture.

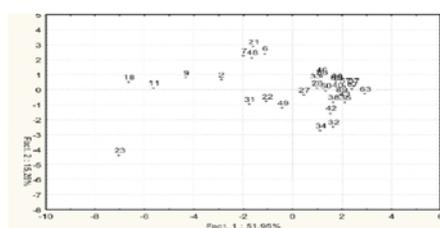
Negative correlations were observed between pH and Ca^{++} , Mg^{++} , SO_4^{--} , Cl^- , NO_3^- , NO_2^- , NH_4^+ , PO_4^{---} indicating the role of pH in the dissolution of evaporite formations above. The analysis was carried to two factors and only 68% of the total variance could be cast (Figure.8.a).

- The horizontal axis expresses F1 53% of the variance and is related to Ca^{++} , Mg^{++} , Na^+ , K^+ , HCO_3^- , SO_4^{--} , Cl^- and Cond.

- The vertical axis F2 which represents 15% of the variance is marked by opposition to elements of the temperature of anthropogenic NO_3^- , NO_2^- , NH_4^+ and PO_4^{---} . The projection of individuals has revealed that water points characterized by these elements are located in center field and near the wadi where the use of chemical fertilizer is intense (Figure.8.b)



a. Projection of variables on the factorial (1 x 2)



b. Projected individuals on the factorial (1 x 2)

Figure 8. ACP graphical representation of chemical data. May 2007

Study of nitrate

An examination of the map plotted from data sampling campaign in May 2007, (Figure.9) shows that the area's most vulnerable to nitrate pollution are located in the central part of the plain. This zone is characterized by the presence of a thick permeable layer that promotes the migration of nitrogen compounds to the saturated zone, following a continuous supply of nitrogen fertilizers. Strong content is recorded in the well N^o. 23 with a concentration of over 486 mg / l. low levels are recorded in the southwestern part of the land. This zone is characterized by a relatively deep groundwater level (75 m) and by the existence of a thick clay layer that protects the water against the infiltration of fertilizers. The lowest value was recorded in well N^o. 62 with a concentration of about 13mg / l. The chemical analysis showed that nearly 79% of water points have values higher than 50 mg / l.

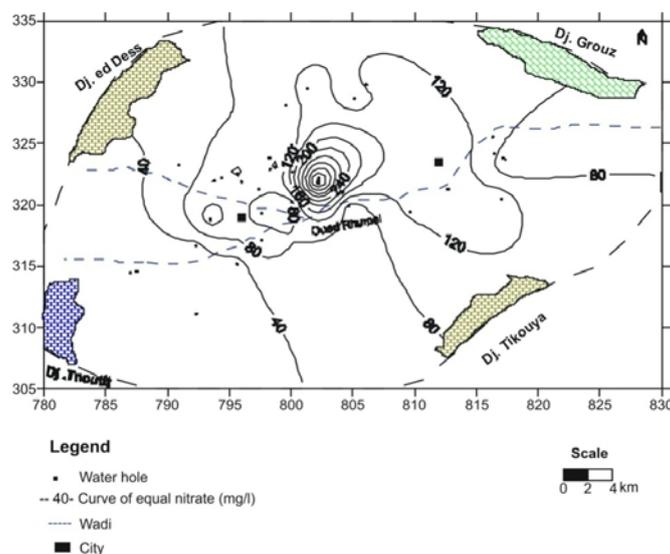


Figure 9. Variation of the nitrate concentrations in alluvial groundwater

Influence of water depth on the content of nitrogen

Examination of the graph on the evolution of the concentration of nitrate according to the depth of groundwater level (figure.10) shows that most water points are aligned around a positive slope which indicates an opposite trend of the nitrate content according to the depth of the water. This phenomenon is due to nitrate reduction following a decrease in oxygen content. This state is clearly visible especially in the western part of the land where the permeability is important. Points situated at great depths have been observed by high nitrate levels, this is due to the effect of the intensive use of chemical fertilizers and high soil permeability, which allows nitrate to seep deep

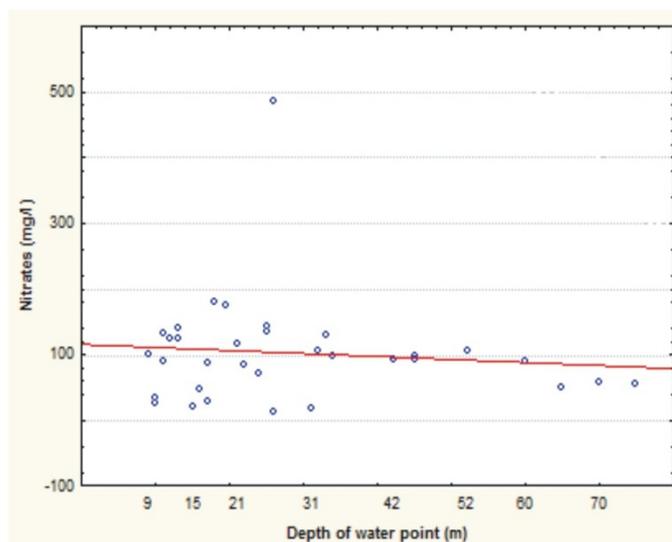


Figure 10 . Relation nitrate – depth of the water point

Relationship nitrate-nitrite and ammonium nitrate

The other two nitrogen forms (nitrite and ammonium) appear from a certain concentration limit for nitrate (100mg / l). They are due mainly to the reduction of the nitrate form. Pairing these two forms only in the central part of the land (Figures.11 and 12).

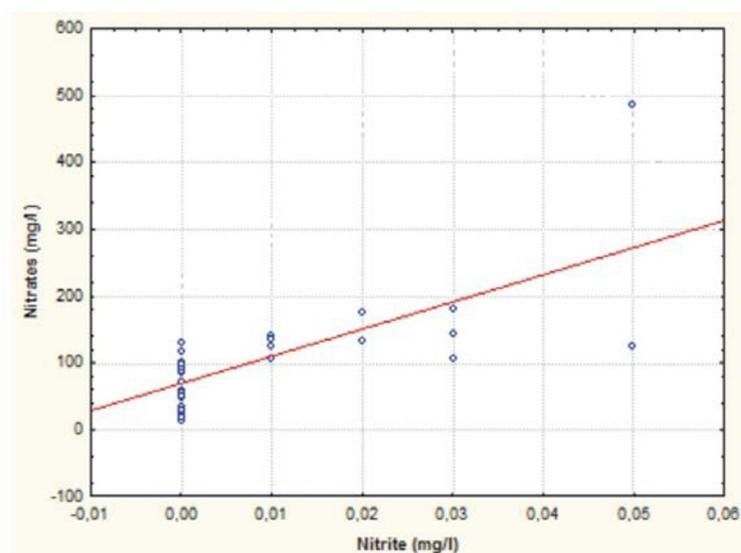


Figure11. Relation nitrate-nitrite

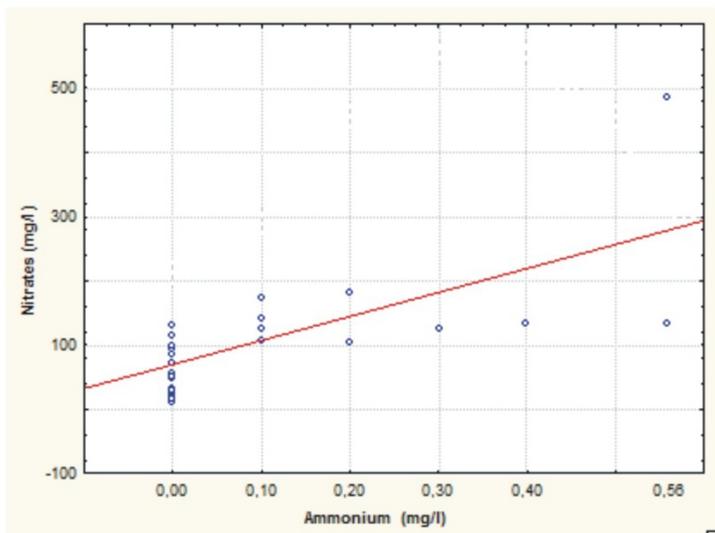


Figure 12. Relation ammonium nitrate

Effect of hydraulic head on the content of nitrates

The infiltration of water from rainfall and irrigation in the water will play a major role in the variations of nitrate content, as a result of charging.

The representation of nitrates according to the static level of all water points (Figure.13) shows that the points of water with low static level have a high content of nitrates due to the ease of ion nitrate to be transported by water through the unsaturated zone of small thickness. With a particular case that characterize water points located in central and eastern plains results in the presence of a large hydraulic gradient and the effect of intense pumping causes a leaching of agricultural land whose use of chemical fertilizers (N, P and K) in these locations is in an abusive manner.

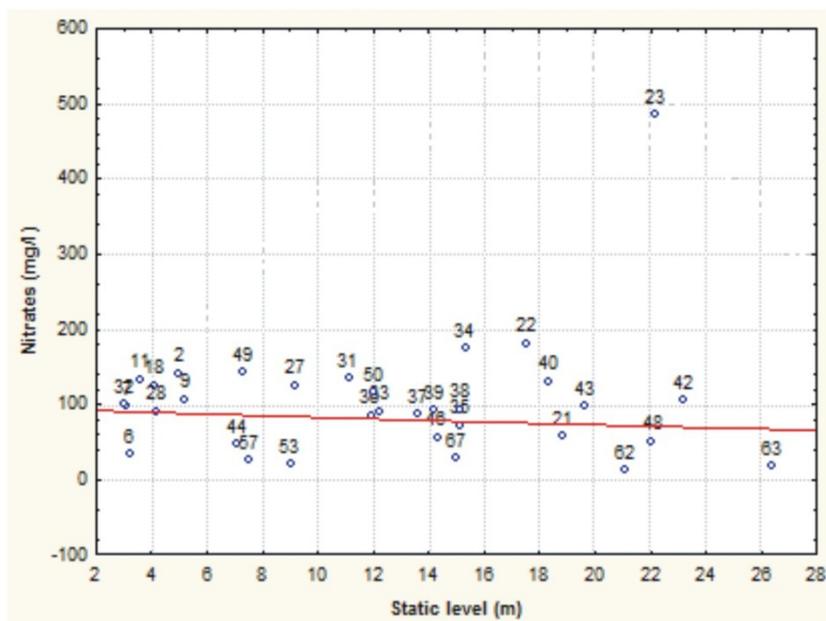


Figure 13. Relation nitrate - static level (May 2007)

Conclusion

The study area is characterized by the presence of a sedimentary formation of the Mio-Plio-Quaternary. We also note the presence of carbonate formations of the Cretaceous on the east and west edge of the field. These courses play an important role in feeding the alluvial aquifer. The water table of the web showed the presence of a groundwater flow direction West-East supported by a supply from carbonates formations surrounding plain.

The water chemistry was used to assess the physical and chemical quality of these waters. The dominant facies of the waters of this aquifer is of type calcium bicarbonate to calcium sulfate in association with gypsiferous clays Mio-Plio-Quaternary.

The practice of intensive agriculture in the region allows spreading irrational fertilizer rich in nitrates, phosphates and potassium-threatening deterioration of the physico-chemical quality of water from the alluvial aquifer. The areas most affected in the region are the central area and the area east mainly because of the shallow groundwater level.

References:

- [1]. Bensouilah, S,1995. Contribution to the hydrogeological study of Sétif high plains through the upper valley of the river upstream of Wadi Rhumel Athménia.Th.Magi.Univ.Constantine.184p..
- [2]. CGG, 1973. Geophysical Survey Area Chelghoum Laid. DHW Mila-Algeria.
- [3].Debieche, T.H,2002. Changes in water quality (salinity, nitrogen and heavy metals) due to salt pollution, agricultural and industrial. Application to the lowlands of Seybouse Nort-East of Algeria. PhD thesis, U. F. R. Science and Technology, Univ. Franche-Comte, 235p.
- [4].Durozoy, G, 1960. Water resources of massive Cretaceous limestones of the Constantine region, study of applied hydrogeology, hydraulics and management of rural infrastructure, service and scientific studies, Algiers.
- [5].Farah, A.S, 1991. Study of the behavior of a wadi in the Mediterranean region semi-arid North Africa and its natural and anthropogenic causes, Rhumel Oued, Algeria. PhD thesis, Univ. Orleans, 271P .
- [6].Issaadi, A, 1981. Hydrogeological study of massive and Guerion Fortars. Thesis, 3rd cycle, Univ. Haouari Boumedienne, Algiers .
- [7].Khedidja, A, 2001. Study of the vulnerability of the lake limestone aquifer in the region of Bir El Arch, W. Setif. Magister thesis, Univ. Tebessa, Algeria | 102p .
- [8]. Macko, S.A. ; Ostrom, N.E, 1994. Pollution studies using stable isotopes. In :Lajtha K. And Michener R., Stable isotopes in ecology.Blackwell,Oxford, pp.45-62.
- [9]. Rodier, J, 1978. Analyse de l'eau, edi. Dunod, 430p.
- [10]. Stumm, W.; Morgan, J.J,1996. Aquatic chemistry ,3rd edn.Wiley- Interscience, New York,1022p.
- [11]. Vila, J.M, 1977. Geological Map of the East of Algeria, Sheet No. N-0-Setif 3-4 in the 200-000th, Department of Geological Map of Algeria, 1977 ..

Graphical Abstract

Figure 1. Location of the studied area ., INC Alger2000.

Figure 2. Geological map of the sub- basin of Tadjenanet- Chelghoum Laid -J.M.Villa 1977.

Figure 3.Hydrogeological section. Khedidja Abdelhamid.2007.

Figure 4 Piezometric map. Khedidja Abdelhamid.2007.

Figure 5.Map of trasmissivity in $10^{-4}m^2/s$. Khedidja Abdelhamid.2007.

Figure6. Inventory map of the water points have been the subject of a physico-chimical analysis. Khedidja Abdelhamid.2007

Khedidja Abdelhamid.2007

Figure 7.Distribution of facies chemical water points .Debieche. T.2002.

Figure 8. ACP graphical representation of chemical data May 2007. Khedidja Abdelhamid.2007

Figure 9.Variation of the nitrate concentrations in alluvial groundwater. Khedidja Abdelhamid.2007

Figure 10.Relation nitrate – depth of the water point Khedidja abdelhamid.2007

Figure 11.Relation nitrate – nitrite. Khedidja Abdelhamid.2007

Figure 12.Relation ammonium – nitrate. Khedidja Abdelhamid.2007

Figure 13. Relation nitrate – static level (May2007). Khedidja Abdelhamid.2007

URBAN STORMWATER QUALITY AND QUANTITY IN THE CITY OF TALLINN

Bharat Maharjan, PhD

Karin Pachel, PhD

Enn Loigu, PhD

Department of Environmental Engineering, Tallinn University of Technology, Tallinn, Estonia

Abstract:

This study aimed to provide an overview of stormwater quality and quantity, the impact on waterbodies and the likelihood of its usage in stormwater management. The potential impacts were assessed using statistical analysis following HELCOM, EU Directive and Estonian requirements. Further, Seasonal behaviour, variability, tentative sample size and frequency were examined using strongly dependent parameters obtained from correlation studies. Results show that the average concentrations of pollutants are not extremely high except microbiological parameters. Most basins have positive correlation of 0.4 - 0.6 between flow and suspended solids (SS), as well as of 0.4 - 0.95 between total phosphorus (TP) and SS. As for seasonal variation, large amount of SS is transported in spring whereas in summer, runoff and SS are consistent against winter and autumn. However, at a 70% confidence interval, there is considerable uncertainty in the mean flow and concentrations. Flow and SS have higher uncertainty than conductivity, BOD7, total nitrogen (TN) and TP. It was discovered that most of the samples belong to a small range of daily rainfall (<5 mm) and there is no measurement for first flush. Variability and inadequate representation of rainfall range calls for comprehensive sampling and validation of the data intended to use in stormwater management programs.

Key Words: Urban stormwater, pollution load, monitoring program, seasonal variation

Introduction

Urbanisation with its uncontrolled impervious surfaces increases stormwater runoff and transports pollutants to the receiving waterbodies. These pollutants not only have an adverse effect on human health but also to indigenous plants and animals (Jacobson, 2011; Christensen et al., 2006; Leecaster et al., 2002). Sediment from stormwater runoff is a potential problem source (Lau & Stenstrom, 2005; German & Svensson, 2002). In order to prevent and minimise stormwater runoff volumes and the pollution load, the Baltic Sea member states jointly pooled their efforts through the Helsinki Commission towards the ecological restoration of the Baltic Sea (HELCOM, 2002). Furthermore, the EU Water Framework Directive (WFD) as well as the Estonian Water Act (EWA) (RTI, 2011) have set a target to protect all waters against pollution and to achieve the good status of all waters by promoting sustainable water and wastewater management (EC, 2000).

The eutrophication of inland waters and the sea is one of the major environmental problems in the Baltic Sea Region, including Estonia (Kotta et al., 2009; Iital et al., 2010; Elofsson, 2010). The urban runoff load has made a substantial contribution towards raising nutrient levels in waterbodies (Taylor et al., 2005; King et al., 2007). HELCOM has adopted an action plan to considerably reduce the anthropogenic nutrient load by 2021 (HELCOM, 2007).

The revised Environmental Charges Act (RTI, 2005) did not elicit the expected reduction in pollutant discharge into waterbodies because the stormwater pollution load is not easily measurable. The stormwater load measurement expenses are significantly higher than the collected tax returns. The specialists in the Ministry of the Environment had not yet defined exactly what kind of mean concentration should be measured (Lääne & Reisner, 2011). There is real need to study urban stormwater pollution in order to develop methods for the reduction of stormwater pollution exports to the sea (Hood et al., 2007), including both flood control and pollution control.

To address these problems and to select appropriate water protection measures, the first objective that needs to be set is to activate the assessment of the status of water, including a comprehensive interpretation of the monitoring data that form the basis for the planning and implementation of protection measures. In addition, low variable and representative data are standard requirements for stormwater management approaches, as they are susceptible to the actual total pollution load and the mean concentration of pollutants. This study will provide a status update on stormwater quality and quantity in the city of Tallinn through analysis of the monitoring data. The main objectives of the study are to assess stormwater quality and quantity; the spatial and seasonal variation of stormwater discharges and pollution load; and to identify, the likelihood of data to be meaningful, representative and verifiable quality on the basis of existing routine monitoring programme, so that they can effectively aid in managing stormwater runoff.

Material and Method

Site Description

Tallinn, the capital of Estonia, is situated on the southern coast of the Gulf of Finland, in north-western Estonia. It has total area of 156 sq km with a population of 417,741, and population density of 2,614 per sq km. The average precipitation in Estonia is 550–750 mm and the mean runoff 280–290 mm per year. The climate in Tallinn is fairly cold in winter with an average temperature of 1.93 °C and a maximum low of -32 °C, a cool spring with little precipitation, a moderately warm summer with an average temperature of 8.64 °C and a high of 32.3 °C, and a rainy autumn.

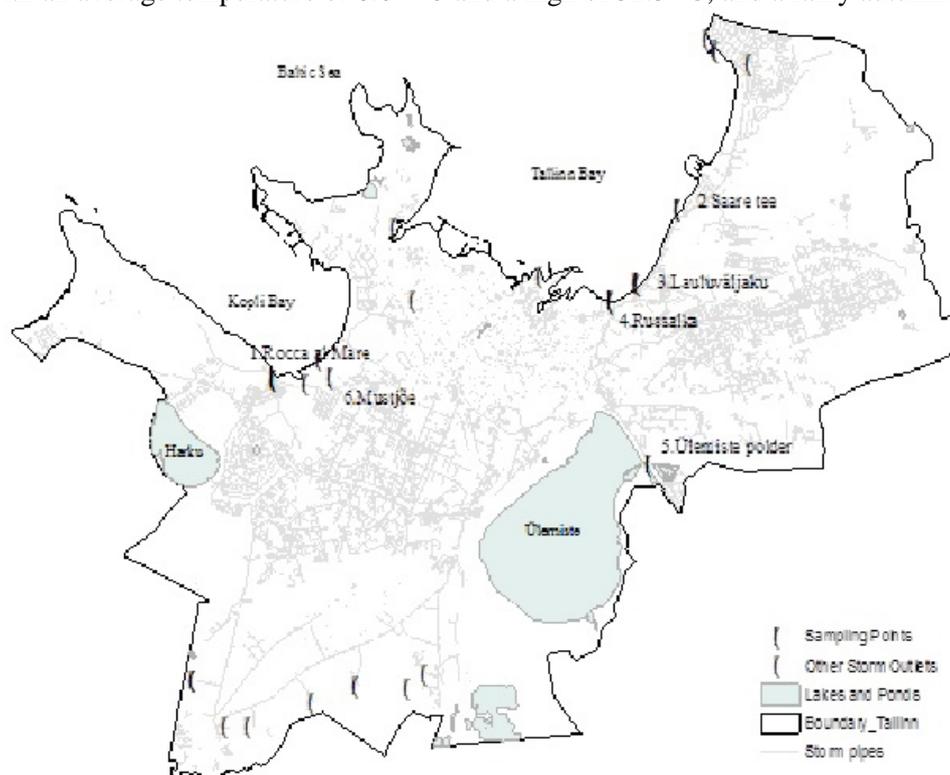


Figure 1: Sampling sites and their location in Tallinn

The area of the stormwater system of Tallinn is about 6,500 hectares and the length of stormwater conduits was 414 km in 2011 (Tallinna Vesi, 2012). Stormwater from residential and industrial areas is either diverted to municipal Waste Water Treatment Plants (WWTPs) and treated with sewage or is collected in a separate stormwater system and mainly disposed to waterbodies without any treatment. The city centre has a combined sewerage system while the other parts have mostly separate systems. There are 23 stormwater outlets that mostly discharge water to the coastal sea (*Figure 1*) in the Tallinn catchment area. Six major storm outlets: Rocca al Mare, Saare Tee, Lauluväljak, Russalka, Ülemiste and Mustoja are included in the monitoring program organised by the Tallinn City Environment Department. For this study, these six outlets are examined, between

them covering a catchment area of almost 4,000 ha, as they are supposed to form a separate stormwater system in Tallinn city. Among them, the Mustoja outlet is the largest and serves almost 30% of the total area. The second biggest is Lauluväljak and the smallest is Saare tee (*Table 1*).

Table 1: Area of Drainage basins and their characteristics

No.	Basin	Area, ha	% Coverage	Receiving water body	Characteristics
1	Rocca al Mare	816	21.5	Kopli Bay	mostly storm and surplus water from the apartment house areas, from pools in the zoo during water exchange, an increase in impervious areas is noticed in the catchment.
2	Saare tee	156	4.1	Tallinn Bay	storm and drainage waters from private house area, collected via open ditches and Varsaallika spring (basin 1.6 km ²), sewage discharges from the private house area can occur
3	Lauluväljak	961	25.3	Tallinn Bay	mostly high density area with impervious area one third of total area. Runoff collected from residential and industrial, sewage discharges can occur from this area
4	Russalka	734	19.3	Tallinn Bay	mostly consists of the Ülemiste polder storm and drainage waters and Lake Ülemiste surplus water after heavy and continuous rains, and during the meltwater period
5	Ülemiste polder			Tallinn Bay	storm and drainage waters from the industrial district and private house areas, airport treated stormwater and runway stormwaters, Ülemiste polder drainage water
6	Mustoja Paldiski Road	1,128	29.7	Kopli Bay	mostly storm and drainage waters from private house areas, apartment houses and industrial district collected via ditches, open channels and pipes into the Mustoja River, increase in impervious areas is noticed in the catchment, sewage discharges from one of the private house areas can occur
Total		3,795			

Data source, sampling procedure and chemical analysis

Stormwater monitoring has been carried out since the late 1980s, but it only became regular in the 1990s. For this study, stormwater monitoring data for the years 2005 and 2008-2011 have been obtained from monitoring reports (Pauklin et al., 2005-2011). In this monitoring system, grab samples were collected 4–6 times a year from the stormwater outlets (see *Figure 1*). The data was measured only once in 2010. The sampling procedure adhered to the sampling requirements in Council Regulation no. 30, 5 May 2002, of the Estonian Ministry of Environment. For 2012, samples were measured by both Tallinn University of Technology and AS Tallinna Vesi. Other samples were taken by the Estonian Environmental Research Centre, all of which are competent bodies according to EN ISO/IEC 17025:2005 for conducting tests in the field of water analysis (accreditation scope on the Estonian Accreditation Centre). Data for 24 hour precipitation from the Tallinn-Harku Meteorological Station located approximately 20 km from the study area was obtained from the Estonian Meteorological and Hydrological Institute (EMHI) for 2005–2012. Samples were tested for parameters such as conductivity, pH, temperature, suspended solids (SS), total nitrogen (TN), total phosphorus (TP), biological oxygen demand (BOD₇), hydrocarbons (HC), Escherichia Coli and Enterococci using analytical methods based on ISO 10523, ISO 5667-10, EVS-EN 25814, EVS-EN 27888, EVS-EN 872, ISO 5815-1, SFS 5505, EVS 9377-2, EVS 9308-1, EVS-EN ISO 7899-2 and EVS-ISO 6340, respectively.

Data Analysis

Normal statistical analysis was carried out to estimate arithmetic means, median, quartiles, correlation coefficient, coefficient of variance (CV) and confidence intervals (CI). The relationship between parameters were analysed through correlation coefficient to obtain prior parameters according to which seasonal variation were observed. Further, with CV and CIs, it was attempted to assess variability of data. Finally, variability and representativeness of data to rainfall intensity were scrutinized through sample size and frequency.

The grab samples for a year did not amount to more than six, except for Mustoja (consists 12 in 2005). It was known from rainfall data of the available data source period that the main parameter of hydrology (average annual rainfall) did not vary significantly. The highest deviation from the mean was 18% only in 2005. Thus, these samples from six years for each basin were combined to attain a higher number of samples, assuming that there was no excessive change in the urban environment.

In terms of the estimation of average total mass emission, it is viable to measure grab sampling with continuous flow measurements over a specific time period (day, week, month), instead of instantaneous flow measurement (Fogle et al., 2003; HELCOM, 2006). For instantaneous flow and concentration measurement, the load calculation was carried out by multiplying the average load by 365 days. Therefore, the mean flow and load over six years in each basin were deemed the average annual flows and loads that are discharged into the waterbodies.

To analyse seasonal variation, the twelve months were categorised with regard to the hydrological year as spring (February, March and April), summer (May, June and July), autumn (August, September and October) and winter (November, December and January). In this way, the data were separated according to the sampling date and grouped into 4 seasons irrespective of yearly variation.

Sampling time during a storm event affects runoff. With correct sampling frequency, it avoids the bias of the first flush and better characterises the mass emission of the event (Lee et al., 2007). To evaluate the sampling programme in terms of sampling number and frequency, it was assumed that there was a constant area of impervious surfaces throughout the study period so that flow can be mainly related to rainfall intensity, though the correlation between daily rainfall intensity (DRI) and runoff was 0.64. Snow cover period was separated and excluded from the analysis because snow melt affects hydrology in a different way. The rainfall data was stratified into three sizes according to rainfall range (small: <5 mm, medium: 5–20 mm, and large: >20 mm). The number of samples that can address rainfall range is hard to determine in regard to grab samples because a grab sample is taken at a particular flow and time, and finding the rainfall intensity that generates that particular flow is almost impossible. Therefore, approximate DRI according to minimum and maximum flow was sought from 24 hour precipitation data. Then the rainfall for other discharges was interpolated to put into the range, and the amount of rainfall within that range was calculated. This rainfall number is actually the number of samples within that particular range. After comparing with the required number of samples, the percentage deficit and surplus was calculated.

Result and Discussion

Stormwater general statistics

The flow and pollution parameters from sampling for six years at six stormwater outlets are summarised in *Table 2*. The total number of samples for most of the parameters is 156. However, some have a lesser number than that to calculate mean flow and mean concentration. HELCOM and Regulation No 269 of the Government of Estonia, 31 July 2001, on the procedure for discharging wastewater into waterbodies or soil, provided limiting values for SS as 40 mg/l and HC as 5 mg/l in stormwater runoff (RTI, 2001). The European Union, as well as Estonia, has restricted microbiological parameters exceeding 1000 cfu/100ml *Escherichia Coli* and 400 cfu/100ml Enterococci for good bathing water quality (EU, 2006; RTI, 2008). There are three public beaches on the Tallinn coastline that are not far from the stormwater outlets. The ecological status of the Tallinn coastal sea was estimated as moderate (The Estonian Environment, 2012). The trophic level in the coastal sea is still quite high despite the fact that the pollution load of Tallinn WWTPs has decreased remarkably since 1990 and discharges via deep outlets do not extend to the coast; therefore, stormwater is still affecting the coastal sea.

It is noticeable that there is large variation in flow, conductivity, SS, TN, TP and pathogens. There is a higher consistency of pH that falls near the neutral range, implying that there is negligible impact from any kinds of industries. Even extreme pH values vary between 6 and 9, and lower or higher values that exceed the limits can be toxic to aquatic organisms. Saare Tee (sampling point 2) has the lowest but Mustoja (sampling point 6) has the largest flow. It reflects the fact that outflows at Saare Tee are from a small drainage basin and at Mustoja from a large drainage basin. It is also true the Russalka (sampling point 4) sometimes exceeds the runoff of the Mustoja basin. In such a case, the runoff is most likely due to the captured overflows of Ülemiste Lake during storm events.

The observed pollutant concentrations are not substantially high, excluding microbiological parameters. The mean concentration for SS is below the permissible level of 40 mg/l. Comparing flow with the transport of this pollutant, the results are found to be opposite in the case of Saare Tee. The discharge from Saare Tee is more concentrated than Mustoja. But in the case of Mustoja basin, there is a high variation in the measurement of SS. Higher readings are recorded occasionally; therefore, the maximum discharge is more than twice the mean value. Rocca al Mare (sampling point 1) is the most polluted basin in terms of mean SS. The basin has water exchange activities inside. This is probably the major contributing factor for such a large value. Ülemiste polder (sampling point 5) has natural stormwater treatment systems – polder areas – that treat stormwater and decrease the harmful effects on the receiving waterbodies. It is found that a few SS samples are above the limit of HELCOM and the Estonian stormwater requirement at 12.3%. The result shows that there are no significant effects from the SS discharged at the outlets of Lauulväljaku, Russalka and Ülemiste sites. However, this is

hard to conclude for other sites because the maximum amount of these parameters is very high and it is essential to look at what factors affected those basins to cause such high values.

Table 2: Pollution parameters concentrations

S. Pt.	Q, l/s	Temp., °C	Diss.O, mgO/l	pH	Conduct. µs/cm	SS, mg/l	BOD ₇ , mgO/l	Ntot, mgN/l	Ptot, mgP/l	HC, mg/l	E. coli, CFU/100ml	Enterococcid, CFU/100ml
Limit		40								5		
1	Samples 23	24	24	24	24	24	24	24	24	16	10	10
	mean 95.8	8.6	9.7	7.5	818.3	38.2	10.6	4.1	0.4	0.2	571,900	27,340
	range 22.9 - 244.2	3.1 - 15	4.2 - 16	7.11 - 8.09	39.5 - 1,556	3 - 178	1.9 - 41	1.94 - 7.21	0.18 - 1.4	0.02 - 1.31	14,000 - 5,100,000	4,800 - 56,000
2	Samples 24	24	24	24	24	24	21	24	24	0	10	10
	mean 38.4	8.9	9.3	7.7	1,420.5	22.8	6.9	4.6	0.4	NA	174,743	16,018
	range 2 - 188.7	3.5 - 14.7	2.5 - 15	7.2 - 8.01	82.9 - 7,600	2 - 220	1.9 - 23	2.72 - 7	0.17 - 1.37	NA	3,600 - 1,200,000	500 - 100,000
3	Samples 19	24	24	24	24	23	13	24	24	0	4	4
	mean 80.5	9.4	10.2	7.8	1,008.9	8.4	8.1	5.0	0.2	NA	92,775	11,150
	range 13.8 - 432.9	6 - 13.9	6.5 - 15	7.2 - 8.36	78.7 - 2,220	2 - 56	3 - 35	3.1 - 9.87	0.08 - 0.8	NA	3,800 - 240,000	1,700 - 21,000
4	Samples 26	26	26	26	26	23	25	26	26	0	10	10
	mean 150.0	10.3	9.8	7.9	760.5	18.3	9.1	6.8	0.1	NA	50,975	6,218
	range 23.4 - 724.5	2.5 - 16.5	0.5 - 16.7	7.44 - 8.18	58.5 - 4,100	2 - 80	3.3 - 45	1.81 - 18	0.02 - 0.3	NA	1,350 - 320,000	160 - 46,000
5	Samples 4	26	26	26	26	21	19	26	26	0	10	10
	mean 115.9	5.8	7.4	7.4	596.0	6.2	7.9	7.5	0.1	NA	432	89
	range 37.1 - 334	0.5 - 19.5	0.2 - 18	7.09 - 7.82	55.5 - 1,015	2 - 17	2.3 - 37	1.07 - 45	0.02 - 0.41	NA	0 - 1,200	0 - 350
6	Samples 32	32	31	32	32	32	30	32	32	16	4	4
	mean 184.3	9.7	9.9	7.7	558.4	32.0	5.7	4.2	0.3	0.2	29,850	3,818
	range 108 - 450.2	4.9 - 14.7	4.7 - 16	7.16 - 8.08	41 - 1,279	2 - 416	1.4 - 21	2.6 - 9.64	0.08 - 2.2	0.03 - 0.69	3,400 - 5,1000	470 - 10,000

The mean concentration of dissolved oxygen in stormwater varies typically from 7.4 to 10.2 mg/l. However, the impact on the oxygen balance is important if secondary pollutants such as oxygen demanding sediments exist. All the measured values for HC are below the permit level and it implies there are no effects due to hydrocarbons in the waterbodies.

Nutrients are a major problem for eutrophication in the Baltic Sea and urban runoff and stormwater from Tallinn city have also added a considerable amount to the sea. The mean concentration of TN and phosphorus exceed the second class – good status limit values of natural surface water in all basins. The limit values are 3 mgN/l and 0.08 mgP/l respectively. However, as shown in *Table 2*, the total N and P concentration in stormwater are substantially less than those of treated wastewater. The limit values by special water permit for Tallinn WWTP outlet are 10 mgN/l and 1 mgP/l, respectively. Microbiology varies quite a lot, with the highest values occurring in the Rocca al Mare outlet, which consists of water from the pools of the zoo. It is possible that some sanitary waste in those basins mixes with runoffs.

Total mass emission

In many studies, the average mass emission from the catchment is estimated using EMC for which composite samples or numbers of grab samples over number of storm events are required. Selecting a single grab sample from many events provides a snapshot of water characteristics for each event, but it will not tell the entire story of the whole pollutograph for any one event. The value of single grab samples is sensitive to the point in time where the grab sample is made (Davis & McCuen, 2005). There is high uncertainty in the estimation of actual mass emission from the grab samples, but it is planned to provide a general overview of mass emission to determine amount of pollutants that are discharged from the specific outlet.

Table 3: Calculated average total mass and specific mass emission for the study period (2005–2012)

Sam. Pt.	Flow		SS		BOD ₇		Ntot		Ptot	
	Total th. m ³ /yr	Specific l/s*ha	Total t/yr	Specific kg/ha	Total t/yr	Specific kg/ha	Total t/yr	Specific kg/ha	Total t/yr	Specific kg/ha
1	3,022.3	117.45	116.2	142.37	30.6	37.47	12.2	14.98	1.2	1.49
2	1,211.6	246.28	48.4	310.07	8.4	53.53	5.4	34.55	0.7	4.22
3	2,537.8	83.74	42.7	44.46	55.5	57.76	13.5	14.07	0.6	0.67
4	4,731.8	204.42	77.8	105.93	45.7	62.27	34.5	47.05	0.4	0.58
5	3,655.8		27.2		24.9		37.7		0.3	
6	5,813.0	163.41	303.8	269.30	38.5	34.16	24.7	21.86	2.1	1.82
	20,972.3		616.0		203.6		128.1		5.3	

Table 3 shows the mass emissions for each basin in terms of the total for and specific of the catchment area. Due to the unavailability of an actual area of the Ülemiste polder, the specific weights were not calculated. It is evident that, on average, Mustoja (sampling point 6) emits the highest SS

with the largest volume of runoff. The specific load of this basin is also comparatively large. In contrast, Saare Tee (sampling point 2) is a small basin and also emits small discharges. However, it has a large specific load for SS, BOD₇, nitrogen and phosphorus. Rocca al Mare (sampling point 1) is also a significant basin for SS, BOD₇, nitrogen and phosphorus, though it emits less pollution per hectare of area than Saare Tee. The highest specific load of BOD₇ and nitrogen is released from Russalka basin (sampling point 4). Lauluväljak and Ülemiste are mild in terms of their discharging pollutant load. The average amount of mass through these six basins is 616 t/yr of SS, 203.6 t/yr of BOD₇, 128.1 t/yr of TN and 5.3 t/yr of TP through 20,972 thousand cubic metres of runoff.

As conducted by AS Tallinna Vesi, the stormwater amounts and pollution loads are not measured but are calculated using a formula based on the drainage area and annual rainfall for annual reporting to the environmental authorities. These values are smaller than in *Table 3*. The possible reasons for this could be that the meteorological station is too far and does not adequately describe the actual situation in basins and the different methodological bases.

Correlation with flow and suspended solids

It is attempted to correlate runoff pollutants with flow at every discharge point, as shown in *Figure 2*. In all the sampling sites, SS and BOD₇ have positive correlation with discharge at that particular time of sampling, while other parameters show positive correlation at certain sampling sites and negative at other sites. The Lauluväljak and Mustoja basins have a good correlation of SS at nearly 0.6. Ülemiste and Saare tee have nearly 0.4, while the remaining basins have a relation of less than 0.4.

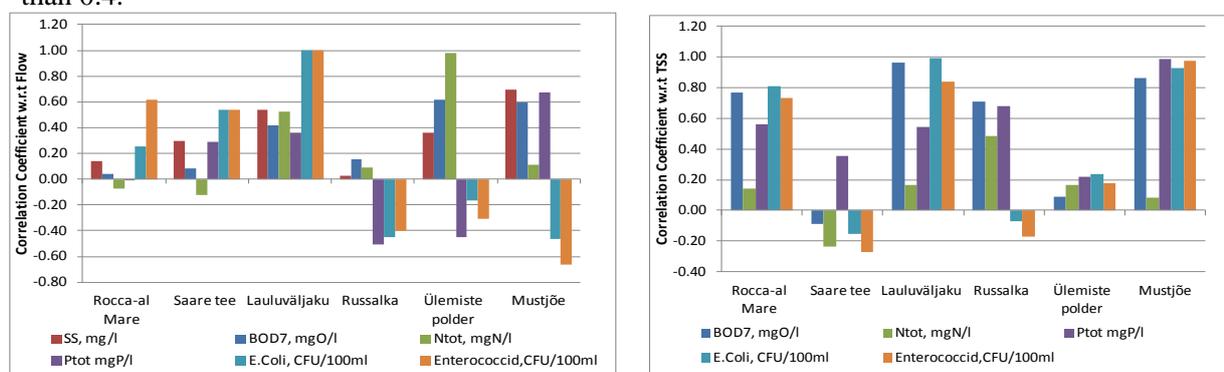


Figure 2: Correlation coefficients with respect to flow (left) and with respect to SS (right)

In relation to flow, parameters aside from SS do not have a strong one-sided correlation. As in the figure on the right, they are again correlated with SS and it is found that in three of the six sites (1, 3 and 6), parameters such as BOD₇, phosphorus and microbiological parameters have a strong positive relationship (range 0.5 - 0.95) with SS. Nutrients, especially TP, always show a positive increment with SS at correlation 0.4 - 0.95, though one site indicates a low figure at nearly 0.2. In the case of microbiological parameters, the Rocca al Mare, Lauluväljak and Mustoja basins are more sensitive to the amount of SS.

Seasonal variation

Normally, the rainy season results in a high amount of runoff from urban areas, while the dry season induces considerably low. Also, the ice melting period is very sensitive to a rise in water levels in drains and channels. During spring (see *Figure 3*), there is usually a high water depth in the conduits and channels. The spring runoff is mainly due to meltwater rather than rainfall, while the autumn and winter runoff is entirely due to precipitation. The mean runoffs at outlets are higher in the winter season than in autumn and summer. Viewing the range of runoff, it is also possible that a greater runoff can occur during autumn but the variability is high. Nevertheless, Russalka showed quite a high flow in winter. This is due to the fact that surplus water in Ülemiste Lake discharged into the overflow channel during heavy rainfall. Generally, summer is the low rain season.

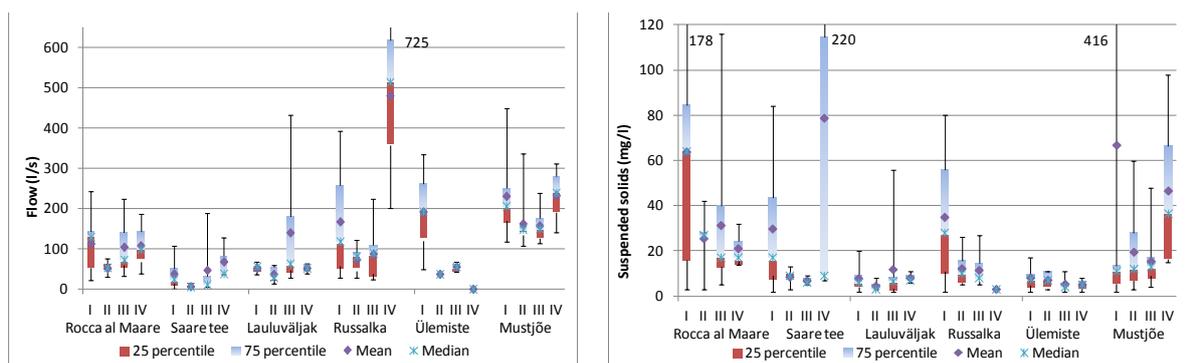


Figure 3: Variation of flow (left) and SS (right) for spring (I), summer (II), autumn (III) and winter (IV)

The emission of SS depends on storm event duration and intensity, antecedent dry days and impervious surfaces. Large storm events do not necessarily develop into a large amount of SS, but the first flush is the main concern in this regard (Davis & McCuen, 2005). *Figure 3* (right) shows that all basins discharge high SS during the spring. This is likely due to the process of ice melting after a long accumulation of contaminants and washing off activities entering the nearest drains. In autumn, besides Mustoja and Saare Tee, all other basins discharge higher SS than in winter. The summer has less variability while autumn and winter have a high variability in emissions.

Therefore, it is valid for all basins that the spring season is crucial for transporting SS but the same is hard to conclude for other seasons. Summer is more consistence, while autumn and winter are variable for discharging SS. From *Figure 3*, it is clear that half of flows greater than the median values are distributed over a large range. In other words, there is a huge bias towards the upper part. The same result can be noticed in the suspended solid concentration.

Accuracy of Means

The mean of flow and concentration is of great value in estimating total volume and mass emission from the drainage area. *Figure 4* shows a coefficient of variance (CV) for various monitored stormwater parameters for different drainage basins. A CV greater than 1 has a higher variation than mean. Flow, SS, TP and microbiological parameters have greater variation in data than mean, while other parameters have less variability.

Table 4 seeks to determine how much deviation of mean could occur in the analysis of the existing data. Positive and negative CI for mean of flow, conductivity, SS, BOD₇, TN and TP are calculated according to a range from 99% (p-value 0.01) to 70% (p-value 0.3) confidence levels. The mean parameters at 99% confidence interval could vary up or down 18.8 - 161.6% (flow), 26.3 - 56.7% (conductivity), 36.2 - 106.9% (SS), 28.7 - 72.6% (BOD₇), 12.5 - 57.3% (TN) and 27.0 - 64.8% (TP) depending upon the drainage area. Although confidence widths at 70% confidence interval are comparatively narrow, they still vary by plus minus 7.6 - 30.6% (flow), 10.6 - 22.8% (conductivity), and 14.6 - 43.03% (SS) 11.6 - 29.2% (BOD₇), 6.3 - 23.1% (TN) and 10.9 - 26.1% (TP).

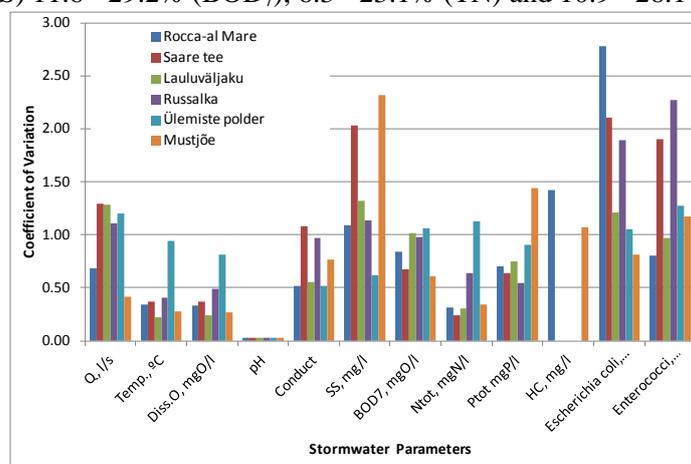


Figure 4: Coefficient of variance of parameters

Table 4: Deviation of mean between two confidence intervals

S. Sit	p-value	Q, l/s		Conductivity, mS/cm		SS, mg/l		BOD ₇ , mgO/l		Ntot, mgN/l		Ptot mgP/l	
		Mean	CI	Mean	CI	Mean	CI	Mean	CI	Mean	CI	Mean	CI
1	(0.01-0.3)	95.84	[(+)-36.78% to (+)-14.8%]	818.26	[(+)-27.12% to (+)-10.91%]	38.21	[(+)-57.57% to (+)-23.17%]	10.56	[(+)-44.21% to (+)-17.79%]	4.09	[(+)-16.44% to (+)-6.61%]	0.44	[(+)-37.14% to (+)-14.94%]
2	(0.01-0.3)	38.42	[(+)-67.83% to (+)-27.29%]	1420.52	[(+)-56.74% to (+)-22.83%]	22.79	[(+)-106.94% to (+)-43.03%]	6.89	[(+)-37.71% to (+)-15.17%]	4.61	[(+)-12.5% to (+)-5.03%]	0.44	[(+)-33.6% to (+)-13.52%]
3	(0.01-0.3)	80.47	[(+)-75.95% to (+)-30.56%]	1008.93	[(+)-28.95% to (+)-11.65%]	8.43	[(+)-70.88% to (+)-28.52%]	8.10	[(+)-72.55% to (+)-29.19%]	5.00	[(+)-16.06% to (+)-6.46%]	0.23	[(+)-39.41% to (+)-15.86%]
4	(0.01-0.3)	150.04	[(+)-56.25% to (+)-22.63%]	760.48	[(+)-50.67% to (+)-20.39%]	18.26	[(+)-61.37% to (+)-24.69%]	9.14	[(+)-50.69% to (+)-20.4%]	6.81	[(+)-32.68% to (+)-13.15%]	0.13	[(+)-26.96% to (+)-10.85%]
5	(0.01-0.3)	115.93	[(+)-161.63% to (+)-65.03%]	596.03	[(+)-26.39% to (+)-10.62%]	6.19	[(+)-36.16% to (+)-14.55%]	7.87	[(+)-63.16% to (+)-25.41%]	7.53	[(+)-57.34% to (+)-23.07%]	0.10	[(+)-48.19% to (+)-19.39%]
6	(0.01-0.3)	184.33	[(+)-18.81% to (+)-7.57%]	558.39	[(+)-33.97% to (+)-13.67%]	31.97	[(+)-103.99% to (+)-41.84%]	5.75	[(+)-28.71% to (+)-11.55%]	4.18	[(+)-15.66% to (+)-6.3%]	0.26	[(+)-64.78% to (+)-26.06%]

There is large uncertainty of mean even when the confidence level is reduced to 70% (Table 4). Among them, sampling site 6 (Mustoja) has a relatively narrow deviation of means except for SS and TP. The confidence interval width for concentration narrows as the sample size increased and does not decrease proportionately for more than seven samples (Leecaster et al., 2002). In this case, the sample size is comparatively high (32 samples), but the main influencing factor is the range of data. As in Figure 4, it has a large range of measurement, which also illustrates why means of TN have relatively low deviations. In summary, flow and SS have higher uncertainty than conductivity, BOD₇, TN and TP in both confidence levels. There is a significant decrease in confidence width from 90-70% but at 70% confidence level, there is still considerable uncertainty in the mean flow and concentrations.

The above results show the variability in the stormwater data according to the mean value. With such high variability, statistical inferences will be highly uncertain. Therefore, further scrutinisation of sampling method in terms of sampling size and frequency is performed.

Scrutinizing sample size and frequency

Rainfall is categorised based on the size of daily rainfall intensity (DRI). The percentage distribution of rainfall is deemed as small amount of 69%, medium amount of 27% and large amount of 4%. At least 20 samples out of 30 are required for monitoring five years during the snow-free period. To sufficiently address the small, medium and large amount of rainfall, 14, 5 and 1 samples are required, according to percentage distribution of rainfall. During the study period, sites 1 - 5 deficits required number of samples or sample size as shown in Table 5. In sampling site 1, nearly 50% small rainfalls are not addressed but it lacks totally the runoff measurements of large DRI. In sampling site 2 and 3, samples are mostly collected when small storms are occurring, but most of the samples in the medium and large daily rainfall are missing. Sampling site 5 has the worst sampling frequency because only some of the medium DRI samples are covered. Finally, sampling sites 4 and 6 are good in terms of sampling for medium and large DRI and also attained relatively better confidence interval. Also, they have relatively good measurements for small DRI. Thus it is noticeable that there is no sufficient sample size and most of flows are captured for small range of DRI (<5mm).

Understanding and quantifying first flush is necessary for predicting environmental impacts on receiving waters and for the efficient design of treatment practices. The first flush wash off usually has the highest concentrations of pollutants, so it is this flush that can prove detrimental to healthy waterbodies. The pollutant loads in runoff after this first flush (over 12 mm of runoff) are assumed to be much smaller and should not have a significant impact on downstream ecology (Davis & McCuen, 2005). As in Table 5, the antecedent dry days (at least 7 days) before the runoff starts are counted. The numbers of those days are 12, 6 and 1 with corresponding small, medium and large rainfall during the snow-free period. There is one such sample for each site in snow cover period, which has 0.7 mm of 24 hrs precipitation and has a higher amount of SS, but it is difficult to suggest on the basis of this data how much antecedent dry days and rainfall can affect SS in total. No sample was measured during the snow-free period that can address such antecedent dry days, so it is hard to estimate the contribution of SS due to first flush on total mass emission, and it is difficult to obtain the sample size required to address those SS. It could probably increase the mean concentration and ultimately increase not only the mass emission of suspended solids but also positively related nutrients and pollutants like phosphorus, BOD₇ and microbiological parameters.

Table 5: Categorized rainfall size and approximate number of flow samples corresponding to the rainfall range (negative denotes deficit and positive denotes surplus)

Range, mm	size	Actual DRI		Reqd no. of sample	Approx. samples						Deficit and Surplus						7 days Antecedent dry		
		no of Rain days	% size of DRI		1	2	3	4	5	6	1	2	3	4	5	6	no of Rain of days	% size of DRI	Sample addressing Antecedent dry days
<5	Small	417	69%	14	6	12	11	10	0	10	-57%	-13%	-20%	-28%	-100%	-28%	12	3%	NA
5-20	Medium	163	27%	5	8	3	2	6	1	11	48%	-45%	-63%	11%	-82%	103%	6	4%	NA
>20	Large	23	4%	1	0	0	0	1	0	2	-100%	-100%	-100%	31%	-100%	162%	1	4%	NA

Conclusion

In this study, the monitoring data was analysed to obtain stormwater quality and quantity status in the city of Tallinn. The pollutant concentrations are not very high, compared to surface water quality classes, the stormwater status could be classified as moderate, aside from microbiological parameters. However, it cannot be suggested that the impact from stormwaters are negligible because the maximum concentrations observed were quite high for those basins and the status of the coastal sea is estimated as moderate. The high values of the microbiological parameters refer to possible occurrence of sewage discharges in the stormwater system, except the Ülemiste polder. It is observed that the polder basins function well in minimising the stormwater pollutants, especially in relation to sedimentation.

In more than half basins, positive correlation is found between flow and SS (0.4 - 0.6) as well between SS and TP (0.4 - 0.95). There is significant decrease in confidence width from 99–70% but there is still considerable uncertainty in the mean flow and concentrations at the 70% confidence interval. Flow and SS have higher uncertainty than conductivity, BOD₇, TN and TP at both confidence intervals. The variability in the stormwater is significantly larger than the mean value. Samples to inadequately address the entire rainfall, absence of information for first flush and high variability of data are particular shortcomings of this monitoring programme. Therefore, the stormwater monitoring programme as well the data should be revised in order to use for the further management approaches.

The sampling time during a storm event is quite important in order to prevent variability and improve sample representativeness of grab samples. Meanwhile, flush concentration can also influence substantially in the calculation of mass emission. This sampling time varies with hydrology, impervious surface as well as topology and the basin soil characteristics.

Single storms can be efficiently characterised with small bias and standard error by taking 12 samples with flow proportioned composite samples. The uncertainty of the overall average concentrations becomes reasonably steady as more samples are collected. In all methods, composite samples are taken either to measure total flow or mean concentration for a storm event or both. These composite samples are minimally required to measure flow since it is totally dependent on storm events, thereafter to provide platforms for validation of data.

The rainfall data are fundamental inputs for the analysis of stormwater runoff. Accuracy is achieved when the rainfall station is near to a sampling site. In our study, it is 20 km from Tallinn city centre. It is recommended that installation of a recording rain gauge on site or as close to the sampling site as possible is essential.

Heavy metals are important pollutants from stormwater runoff. These pollutants are detrimental to the waterbodies. Also, salting activity in highway for snow melting provides chlorides ions in the discharges. Proper monitoring of these metals and ions should also be included in the stormwater monitoring programme.

Acknowledgements

European Social Foundation is financing task 1.2.4. Cooperation of Universities and Innovation Development, Doctoral School Project “ Civil Engineering and Environmental Engineering” code 1.2.0401.09- 0080 has made publishing of this article possible.

References:

Christensen, A.M., Nakajima, F. and Baun, A. (2006) Toxicity of water and sediment in a small urban river (Store Vejleå, Denmark). *Environmental pollution* (Barking, Essex : 1987) 144(2), 621-625.

- Davis, A.P. and McCuen, R.H. (2005) *Stormwater Quality* Springer US.
- EC (2000) Directive of the European Parliament and of the Council 2000/60/EC, Establishing a Framework for Community Action in the Field of Water Policy Official Journal of the European Communities L327 1, 73.
- Elofsson, K. (2010) Cost-effectiveness of the Baltic Sea Action Plan. *Marine Policy* 34(5), 1043-1050.
- EU (2006) Directive 2006/7/EC of the European Parliament and the Council of 15 February 2006 concerning the management of bathing water quality and repealing
- Fogle, A.W., Taraba, J.L. and Dinger, J.S. (2003) Mass load estimation errors utilizing grab sampling strategies in a karst watershed *JAWRA Journal of the American Water Resources Association* 39(6), 1361-1372.
- German, J. and Svensson, G. (2002) Metal content and particle size distribution of street sediments and street sweeping waste. *Water Sci Technol* 46(6-7), 191-198.
- HELCOM (2002) HELCOM Recommendation 23/5, Reduction of Discharges from Urban areas by the Proper Management of Storm Water Systems.
- HELCOM (2006) Guidelines for the compilation of waterborne pollution load to the Baltic sea (PLC-water), p. 30.
- HELCOM (2007) HELCOM Baltic Sea Action Plan, p. 101, Krakow, Poland.
- Hood, M., Reihan, A. and Loigu, E. (2007) Modeling urban stormwater runoff pollution in Tallin, Estonia. *International Symposium on New Directions in Urban Water Management*.
- Iital, A., Pachel, K., Loigu, E., Pihlak, M. and Leisk, U. (2010) Recent trends in nutrient concentrations in Estonian rivers as a response to large-scale changes in land-use intensity and lifestyles. *Journal of Environmental Monitoring* 12(1), 178-188.
- Jacobson, C.R. (2011) Identification and quantification of the hydrological impacts of imperviousness in urban catchments: a review. *Journal of environmental management* 92(6), 1438-1448.
- King, K.W., Balogh, J.C. and Harmel, R.D. (2007) Nutrient flux in storm water runoff and baseflow from managed turf. *Environmental Pollution* 150(3), 321-328.
- Kotta, J., Kotta, I., Simm, M. and Pöllupüü, M. (2009) Separate and interactive effects of eutrophication and climate variables on the ecosystem elements of the Gulf of Riga. *Estuarine, Coastal and Shelf Science* 84(4), 509-518.
- Lääne, A. and Reisner, R. (2011) Water sector reforms in Estonia, A Background paper for the OECD Global Forum on Environment: Making Water Reform Happen, Paris.
- Lau, S.-L. and Stenstrom, M.K. (2005) Metals and PAHs adsorbed to street particles. *Water Research* 39(17), 4083-4092.
- Lee, H., Swamikannu, X., Radulescu, D., Kim, S.-j. and Stenstrom, M.K. (2007) Design of stormwater monitoring programs. *Water Research* 41(18), 4186-4196.
- Leecaster, M.K., Schiff, K. and Tiefenthaler, L.L. (2002) Assessment of efficient sampling designs for urban stormwater monitoring. *Water Research* 36(6), 1556-1564.
- Pauklin, T., Otsa, E., Alumaa, P. and Kakum, T. (2005-2011) Tallinna sademevee väljalaskude seire, Estonian Environmental Research Centre.
- RTI (2001) Heitvee veekogusse või pinnasesse juhtimise kord. Retrieved from <https://www.riigiteataja.ee/akt/13290813>.
- RTI (2005) Keskkonnatasude seadus (Environmental Charges Act), Parliament. Retrieved from <https://www.riigiteataja.ee/akt/114032011039>.
- RTI (2008) Regulation of the Government of Estonia nr 74 „Nõuded suplusveele ja supelrannale“ RTI 2008,16,117, Government of Estonia. Retrieved from <https://www.riigiteataja.ee/akt/12950336>.
- RTI (2011) Veeseadus. Retrieved 12 Oct, 2012, from <https://www.riigiteataja.ee/akt/110032011010>.
- Tallinna Vesi (2012) Environmental Report 2011, p. 3, AS Tallinna Vesi.
- Taylor, G.D., Fletcher, T.D., Wong, T.H.F., Breen, P.F. and Duncan, H.P. (2005) Nitrogen composition in urban runoff—implications for stormwater management. *Water Research* 39(10), 1982-1989.
- The Estonian Environment (2012) *Estonian Environmental Indicators 2012*. Antso, K. and Kaukver, K. (eds), Estonian Environment Information Centre.

FRENCH ENVIRONMENTAL SOCIOLOGY FACING INTERDISCIPLINARY RESEARCH ACROSS SOCIAL AND NATURAL SCIENCES

Aspe Chantal, Maître de Conférences – HDR - Sociology

Mediterranean Institute of Marine and terrestrial Biodiversity and Ecology,
Aix-Marseilles University / CNRS

Jacque Marie, Maître de Conférences - Sociology

Population Environment Development Laboratory, Aix-Marseilles University / IRD

Abstract:

Interdisciplinary research between social and natural sciences appear both as constitutive of environmental sociology and the main problem for its disciplinary and institutional recognition. Interdisciplinary research programs are widely acknowledged in the environmental field, particularly due to their linkage with social demand. Conversely, researchers embarked in this practice are often perceived as marginal from their discipline's standpoint. Based on the French experience, this paper shows how interdisciplinary research plays a reflexive role for environmental sociology by theorizing the "co-construction" of ecological and social systems. Interdisciplinarity is not a purpose for environmental sociology but a necessary approach to confront its theoretical breakthroughs and the social and political understanding of the environmental issue to the ecological reality.

Key Words: Interdisciplinarity, natural and social sciences, environmental research

Introduction

The interdisciplinary approach makes it possible to challenge findings and theoretical breakthroughs in environmental sociology. Firstly, this occurs through the acknowledgment of the material and ecological dimensions of societies' production. Then, this approach provides a reflexive return on sociology's role and position in both the political and social construction of the environmental issue. The aim of this paper is not to support interdisciplinary work as a purpose for environmental sociology but to show how it contributes to raising questions and advances in the discipline. Interdisciplinarity is a necessary approach to achieve these ends.

The gradual construction of environmental sociology has been supported by sociologists' commitment within interdisciplinary programs and research, even though the disciplinary and institutional recognition of this approach is debated, denied (Mol, 2006) or more often disregarded. However, interdisciplinary research practices have developed to respond to the limitations of disciplinary paradigms that are insufficient in analyzing the complex environmental study object, and have partly redefined the conceptualization of the nature-society relationships. These practices—although still marginal—are inevitable in today's understanding of environmental sociology. To feed this debate, our argumentation will be three-fold. First, we will develop the historical link between the development of an interdisciplinary scientific research field and public orders, which question both the institutional legitimacy of research and the sociologist's impact. Sociologists' commitment to interdisciplinary research has enabled them to reconsider the epistemological dimensions of their scientific practices—particularly by theorizing the "co-construction" of ecological and social systems—and participate in model building. Lastly, we will examine the impact of today's interdisciplinary practices since these lie at a cross-roads between two fields and may generate a critical debate on the environmental issue.

Interdisciplinary research across social and natural sciences: between public orders and scientific relevance

Many French sociologists first embarked on this approach for two main reasons. First, political and public orders begun in the late 1960s under the impetus of the DATAR (land planning' local government) and research programs launched by the DGRST (Directorate general for scientific research) at the national level and UNESCO at the international level have been focusing on the future of marginalized rural areas. Secondly, the sociologists participating in these studies predominantly came from rural sociology; they will respond to the call of natural sciences that started to reintroduce man as an integrated ecosystem component.

Various researchers from both social and natural sciences have met in the 1960s to deal with a spatial-related study object. This period is the time of "reconstruction" and "land management" in France. Consequently, long-term planning becomes the key of the DATAR policy, which will have to "imagine the future" to ensure balanced land management and contribute to economic modernization. Researchers were then approached to analyze the declining farming area's future.

The DATAR's concerns will be supported by various multidisciplinary programs, particularly through many Committees across time named "Biological control" (1968); "Biological balance and control" (1971); "Renewable natural resources management" (1975); "Ecology and rural planning" (1978) of the DGRST; and the "Diversification of rural development models" (1983) of the Ministry of Research. Lastly, the concepts of environment and interdisciplinarity will be first and foremost officially recognized with the creation of the Interdisciplinary Program of environmental research (PIREN) of the CNRS in 1978.

Most interdisciplinary research will focus on "marginal spaces." The major studies are probably those carried out within the research program "Causses-Cevennes" of Interdisciplinary Program of environmental research. This program is aimed at "analyzing the interface between environmental and social factors that characterize marginalized rural areas. This analysis precedes the exploration on how these areas may be provided with ecological dynamics for development (concept of ecological management)" (Jollivet, 1983).

We can wonder why these first national interdisciplinary programs have focused on these marginalized areas. These territories—places of interest of "community utopias" (Léger, 1979) — were also used as observatories by researchers, who consider them just like the social movement as areas free of nearly any human impact. These lands could become study areas for sociologists since the analysis of a society's rejections, waste and margins makes it possible to underscore its marks, values and norms. A part of the land empties while another one fills up. But what occurs in the empty space may also reveal what happens in the full one. Assessing lands that temporarily have no economic value and that are in a way "economically in reserve" reveals the spatial ethical representation favored by a society. Since there was not even "one" evaluation criterion, these "reserves" had to be explored in all their dimensions: ecological, geological, geographical, cultural, historical, linguistic, etc.

This approach for a shared study object—the space—that was first delineated at the local level, and enabled researchers from various disciplines but also local actors and scientists to meet. From the outset, interdisciplinarity between social and natural sciences faced two major barriers. To meet the challenges of interdisciplinarity, researchers first need to be open minded to scientific cultures other than their own. Secondly, there is the problem of direct confrontation with local actors. Exchanging on a shared issue and focusing on a common territory and a scale understood by various disciplines, akin to a form of "*bricolage*", has made it possible to raise some questions that would not have arisen otherwise. Interdisciplinarity became a scientific challenge as soon as the interdisciplinary research program on environment (PIREN) was developed by the CNRS. The 1960s "*bricolage*" was out of the question then, and the terminology changed while complexity became the key concept. Interdisciplinarity must be theorized, be given a scientific basis and a better-known set of knowledge.

This program was the figurehead of social and natural sciences interdisciplinarity in France due to its legitimacy gained by the CNRS name and to CNRS funds promoting the integration of social sciences into environmental research programs. Pioneer researchers from the first stage or working as "free riders" (with regard to their affiliated institution or laboratories) will henceforth speak in a more visible manner and with increased recognition from the scientific community.

Nevertheless, these researchers—mainly sociologists—remain in the minority. This obstacle will turn out to be a relative asset, such as an investment similar to those made by “Crusaders”. A core set of shared knowledge will gradually be built, enabling ideas to circulate, researchers to be demarginalized and interdisciplinarity to stand officially within the scientific circle. All these factors will lead to a gradual shift from secret confinement to an exposed openness.

Research programs’ prompting to develop a new approach will indirectly generate contradictory breakthroughs for sociology. On a theoretical level, it enabled the reintroduction of the “natural environment” variable into the social construction and on the pragmatic level, to tackle the challenge of responding to “social demand”.

Theorization and research practices

Although interdisciplinary practices between the social and natural sciences appear marginal or even marginalized in sociology, they provide a framework for reflection and reflexivity for environmental sociology and more broadly for various disciplines using this approach. In this line of thought, M. Jollivet and M. Legay have distinguished an interdisciplinarity “of proximity” from an “extended” one that reintroduces the link with social actors. They set interdisciplinarity back into a “normal” process of scientific production that characterizes the link between science and society. The issue is to clarify the interdisciplinary approach whose difficulty is mostly based on “*the structural duality of problems that it has to deal with*” (Jollivet, Legay, 2005).

The historical analysis has taught us that most sciences are based on the redefinition of gained knowledge, methods and conceptualization of other disciplines. Therefore, this is also the case for the two disciplines concerned in this paper: sociology and ecology. Nothing will come from “ex nihilo”, and even less for sciences. Thus, new disciplines will emerge from the direct encounter of other distinct disciplines, such as molecular biology (biology and chemistry) or by a more complex pathway, borrowing some reflection and analysis components here and there, such as geography, botany, cybernetics, thermodynamics, etc. for ecology. North American sociology seems to follow this pathway today to tackle environment-society interactions when it uses the following labels “ecosociology” (Vaillancourt, 1992), “environmental sociology” or “ecological sociology” (Dunlap and Catton, 1979, 1996). The epistemological broadening necessary for the production of integrating concepts refers to the transformation of the thinking in the social sciences, particularly sociology. Shouldn’t there be a connection between the development of innovative paradigms in sociology on the nature-society relationship and the increasing number of studies on the environmental issue? “*Social process and nature’s dynamics continually cross the great divide, and hence it is counterproductive for the social science to remain solely on one side. (...) The analytical starting point of this expansion is an reconceptualization of what is meant by “action” in the social sciences so that it can capture the interaction between social constructions and nature’s constructions*” (Murphy, 2009).

Above all, interdisciplinarity between social sciences and natural sciences in France was an approach or a practice that must take into account specific questions, methods and theoretical stances raised by different disciplines, and which actually needs open-minded researchers. Although openness has unanimous support among the “interdisciplinarity circle”, opinions are divided regarding the form of the interdisciplinary approach. Would it be better that the common research object be constructed at the start, or should researchers first progress in its understanding before reformulating it as each one advances in his knowledge? Does each discipline need to follow its own rationale and share results towards an “interdisciplinary takeover”—with shared knowledge to be distinguished from disciplinary knowledge? May a researcher practice “interdisciplinarity” alone by collecting other scholars’ findings and building an integrated analysis? Does a discipline have to play the conductor’s role to make everyone play the same part? All these questions underscore the richness of the debate and we believe that there is no need to take sides since each stance is useful and research-specific—institutional demand, comprehensibility of the prior research object, existing data—but also depends on the charisma of some researchers.

The interdisciplinary debate between social and natural sciences is renewed through model-building. There are high expectations that modeling will force researchers to agree on the key issues or make them give up some others if necessary. “*Its formal requirements urge to precisely state the terms of the questions, work on well-defined data, and follow the rules for analysis that are adapted to the variables, and take them completely into account. It is likely to be a rigorous guide in the*

confrontation between disciplines. [This is true] to the point of observing that it must be used ultimately." (Jollivet, 2009). Several types of models are developed, particularly those that aid the decision-making process: Adaptive management (Holling, 1978), patrimonial approach (developed by Ollagon, 1989), multi-agent systems (Barreteau, 1998; Collectif ComMod, 2005), strategic analysis and role-playing games (Mermet, 2005). All these approaches seek to take into account not only the researchers' data and questions but also all the stakeholders in an interactive manner. In this perspective, scientific data has to be modeled to make it possible to proceed to decision-making while including the decision-makers' issues.

These various endeavors that are worthy of being engaged in a civic-minded action of science may challenge social sciences on their actual epistemological effectiveness. The stance of social sciences, particularly sociology, remains a problem. First, social sciences have a small tradition in modeling their data and even less in forecasting. They prefer using past and present observations and analyses and are careful not to predict social dynamics that are somewhat unpredictable.

The contemporary use of the resilience concept—from physics, used in psychoanalysis and then in ecology—reaches the sphere of social sciences today. It makes it possible to rethink the ecosystem dynamics, while considering its capacities to resist disturbances and those for finding an unstable equilibrium again, which has "learned the lessons" of previous disorganizations. This concept has begun to be used in social sciences to adapt to social or environmental changes through relevant social organizations (Adger, 2000). Therefore, social and natural sciences are recognized in the dynamic approach of ecosystems and social systems. However, can there be other perspectives than following these dynamics, if not in real-time but at least according to scales adapted to these two systems?

The practice of interdisciplinarity does not seem confined to the margins of the scientific approach but its challenge is still present. Three issues currently appear as fundamental for the practice and theorization of interdisciplinarity. This approach allows the incorporation of considerations on each discipline's specific concepts and terminology in the heart of a debate regarding the forms of diffusion and appropriation within scientific discourse. The complex social and natural reality may emerge from a conceptual clarification observed through the analysis of reciprocal dynamics. This last element encounters the problem of overcoming different research positions between distant disciplines for the benefit of a shared methodology. The systemic approach of reality as well as resorting to models may form a common pathway (Jollivet, Legay 2005). The rigorous methodology probably insures an equal representativity of disciplines involved in the research projects, since interdisciplinary programs are often controlled by a discipline, particularly by natural sciences (Leroy, 1995, 2004). The challenge to build an interdisciplinary approach is not only scientific but also social and political. This complex approach not only aims at responding to a social demand, but sheds light on new issues and social and political questions.

Opening up to other analyses and discourses and the integration of the environment-society interface have helped to rethink reductionism principles, but have also set the limits of the efficiency of the system concept and the boundaries of the analysis object. Following L. Von Bertalanffy (1973), many authors have tried to suggest a standard language to provide a universal approach to the evolution and functioning of systems. However, just like a catalyst, the use of the system concept turns out to be both essential and relatively ineffective. It was used as an abstract model facilitating the collaboration of disciplines to think about reality in its complexity, but seldom as an operational methodology. Nevertheless, the systemic approach will enable natural sciences to question sociologists on the functioning of social systems to understand the social conditions of ecosystem reproduction; it will also allow sociologists to relativize the social determinism and reconsider the interaction between the environment and social behaviors in a different way.

These questions focus not only on operationalisation but even more on understanding interrelationships between environmental changes and social practices. The integration of complexity, which has partly questioned the causal explanation or resulted in the emergence of the multiplicity of effects and causes, has led to a debate between researchers involved in interdisciplinary programs. As a result of studies on environment, disciplines tackled discipline-specific objects with a new stand, enhanced by the knowledge from other disciplines, methods and know-how, and thus question their object with this fresh point of view.

The research field first influenced the disciplines, which in return have yielded new types of questions on the reality. The systemic approach plays a key role in the construction of the representations of reality and the frame of reference of social realities (relationships between individuals), but also in the construction of the relationships between nature and society.

If collaboration has increased between disciplines since the 1990s, it is also due to a changing relationship to time. Changes that were thought about over very long-time periods for many disciplines now appear as though they may be impending and mostly relatively unpredictable due to the complexity of the causes and effects. While “small” was “beautiful” (Schumacher, 1978) in the 1970s, it is viewed today as having considerable consequences on global phenomena and participating in the emergence of brutal transitions. The “small”, the “singular” or the “local” are not only epiphenomenons compared with the “global” and facing global uncertainties. Thus, they have resumed their full place within research, particularly in the contemporary approach to development. Henceforth, problems are addressed on a changing scale and this constitutes a new challenge for sociological research.

Reflexivity of interdisciplinarity: towards a co-construction of ecological and social systems

The changing vision of reality provided by the interdisciplinary approaches has also had an impact on the development of the disciplines in question, particularly ecology and sociology (Aspe, Jacqué, 2012).

The breakthroughs leading to another sociological stance must not conceal the epistemological problems still raised for the sociologist, despite a tendency to open up to other disciplinary approaches. The current possible multidisciplinary mixing has caused effects that were unforeseen by research programs: scholars and their divergent discourses have become a study object for sociologists. The following question is then raised: how can a partnership be organized when the analysis definitely needs distance? In other words, how can researcher’s knowledge be enhanced by other disciplines to understand the social realm, while setting it as a social product underlying actors’ strategies?

In addition, the sociology’s analytical approach and the necessary hindsight to its study object make it hard to meet the increasing demand of political decision-makers, to provide short- and long-term solutions to immediate environmental issues, and to develop different management for the future. Whereas most French sociologists remain skeptical about possibilities for collaboration with other disciplines, some scholars such as physicists, chemists, geneticists, and naturalists raise questions on the limits of their fields, and underscore the need of the natural sciences’ viewpoint on their object, methods and results.

As far as natural sciences are concerned, they have also relativized their disciplinary approaches by dropping some concepts that are unable to show the dynamic interactions between nature and society. For instance, the gradual renouncing of the concepts of “climax” or ‘natural equilibrium” for a dynamic and contextualized approach of ecosystems attests to the new stance of ecologists who prefer using the paradigm termed “co-adaptation”, “which requires emphasis on the adaptation processes through change” (Blandin, 2009). Incorporating human and social dimensions to comprehend the environmental issues makes it possible to go beyond the Manichean opposition between the destructive man and a nature that is unavoidably threatened, and thus complying with the constructed-based dimension of nature (Lévêque, 2008).

These considerations echo our interdisciplinary research results about water management in Provence. The social dynamic occurring while uses and conflict on farm water are changing has ecological consequences. The sociologist seeks to integrate this new ecological dimension to understand strategies in terms of natural resources management and appropriation while replacing social conflicts and relationships into territorial logics (Aspe, 2012).

Looking back on Durkheim’s work, many sociologists have often suspected natural sciences to deflect social sciences from their true objective leading them to the dangerous “social biologism”, where mechanisms and laws observed on natural objects would be tacked onto human behavior in society. This debate is doubtlessly open but must not conceal what the scientific community and perhaps more specifically sociology, is asked to do today.

This request is perhaps the “price of fame. “ The scientific explanation of the world, followed up by “experts” through media, has increasingly prevailed over the pragmatic and religious explanation. Scientists—some of them without knowing—were faced with their involvement as a

social actor (debate between fundamental and finalized research) but also as decision-maker and “advisor to the Prince.” International conferences bring together side-by-side politicians and leaders of non-governmental and scientific institutions. Called upon to arbitrate the debate and decisions, the scientific community plays a double role, while asserting on the one hand that it can only provide assumptions and on the other hand that its opinion cannot be ignored.

Conclusion

Two major questions lie at the root of these debates. First, don't humans cause irreversible actions? Then, what prospective sustainable model can be made to perpetuate humanity in the future? Sociologists are not left indifferent to these questions since these refer to ethics and social organization facing human societies' future.

Likewise, sociologists are concerned with life sciences and their sudden growth, particularly in genetics. The impacts of these discoveries have a direct social implication whether they give rise to fears (“mad cow disease crisis,” transgenesis, etc.) or hopes (positive medical outcomes, increased of life expectancy, and reduced hunger). Indirectly, life sciences question the limits of living beings, in this case humans who are also characterized as a social being. Will the increase in the human being's life expectancy, the resistance to diseases, and genetic changes affect humans' life within society? To what extent will the changes in biological beings have impacts on the social beings? It is difficult today to answer to these questions. However, they seem to lie at the core of scientific and societal debate. The sociologist should realize the scope of changes caused by the transformation of knowledge in the other disciplines, tackle the complex interference between biological and social construction of human beings, and maintain the rightful place in an overall exploration of the future of societies.

References:

- Adger, W. N. Social and ecological resilience: are they related? *Progress in human geography*, 24: 347-364, 2000.
- Aspe, C. ed. De l'eau agricole à l'eau environnementale. Résistance et adaptation aux nouveaux enjeux de partage de l'eau en Méditerranée. Paris : Quae Editions, 2012.
- Aspe, C. and M. Jacqué. Environnement et Société. Une analyse sociologique de la question environnementale. Paris : Quae Editions / Maisons des Sciences de l'Homme, 2012.
- von Bertalanffy, L. Théorie générale des systèmes. Paris: Dunod, 1973.
- Barreteau, O. Un système multi-agents pour explorer la viabilité des systèmes irrigués : dynamique des interactions et modes d'organisation. PhD thesis. Montpellier: Ecole nationale du génie rural, des eaux et forêts, 260p., 1998.
- Blandin, P. De la protection de la nature au pilotage de la biodiversité. Paris: Quae, Col. Sciences en question, 2009.
- Collectif ComMod. La modélisation comme outil d'accompagnement, *Natures, Sciences, Sociétés*, 13 : 165-168, 2005.
- Dunlap, R.E., Catton W. R. Environmental Sociology, *Annual Review of Sociology* 5: 243-273, 1979.
- Dunlap, R.E. Toward an Ecological Sociology: the development, current status and probable future of environmental sociology, *Annals of the international institute of sociology*, 3 : 263-284, 1996.
- Holling, C.S. Adaptive environmental assessment and management. Londres : London John Wiley, 1978.
- Jollivet, M. Présentation du programme Causses - Cévennes du Piren, *Etudes rurales*, 89-91 : 331-333, 1983.
- Jollivet, M. La prise en compte de la société dans les recherches sur l'environnement, *Lettres des Programmes Interdisciplinaires de Recherche du CNRS*, 4 : 13-16? 1991;
- Jollivet, M. ed. Sciences de la nature, sciences de la société. Les passeurs de frontières. Paris : CNRS Editions, 1992.
- Jollivet, M. and J.M. Legay. Canevas pour une réflexion sur une interdisciplinarité entre sciences de la nature et sciences sociales, *Natures Sciences Sociétés* 13 : 184-188, 2005.
- Jollivet, M. Eléments de théorie pour une recherche interdisciplinaire sur les interfaces natures/sociétés, in D. Hervé and Laloë F. Modélisation de l'environnement : entre natures et sociétés. Paris: Quae, Col. Indisciplines, 9-20, 2009.

- Léger, D. Les utopies du retour, *Actes de la Recherche en sciences sociales* 29 : 45-74, 1979.
- Leroy, P. Environmental sciences has a vocation. Nijmegen University Press, 1995.
- Leroy, P. Sciences environnementales et interdisciplinarité : une réflexion sur les Pays-Bas, *Natures Sciences Sociétés* 12 : 274-284, 2004.
- Lévêque, Ch. La biodiversité au quotidien. Le durable à l'épreuve des faits. Paris: Quae Editions, 2008.
- Mermet, L., R. Billé, M. Leroy, J.B. Narcy and X. Poux. L'analyse stratégique de la gestion environnementale: un cadre théorique pour penser l'efficacité en matière d'environnement, *Natures, Sciences, Sociétés*, 13 : 127-137, 2005.
- Mol, A. From environmental sociologies to environmental sociology ? A comparison of U.S. and European environmental sociology, *Organization and environment* 19: 5-27, 2006.
- Murphy, R. Leadership in disaster: learning for a future with global climate change. McGill-Queen's University Press, 2009.
- Schumacher, E.F. Small Is Beautiful - une société à la mesure de l'homme. Paris : Contretemps/ Le Seuil (ed franc.), 1978.
- Vaillancourt, J.G. Sociologie de l'environnement: de l'écologie humaine à l'écologie sociale, in R. Teissier, J.G.Vaillancourt eds. La recherche sociale en environnement, nouveaux paradigmes. Montréal: Presses Universitaires de Montréal, 1992.

VARIOUS ASPECTS OF ANTHROPOGENIC MATERIAL FLOWS IN THE CZECH REPUBLIC: PRESENTATION OF THE PROJECT

Jan Kovanda, PhD

Tomas Hak, PhD

Charles University in Prague, Environment Center, Czech Republic

Abstract:

This article presents a project supported by Grant Agency of the Czech Republic and carried out at the Charles University in Prague, Environment Center. The project is being carried out in 2012-2014 and its overall goal is to evaluate various aspects of anthropogenic material flows in the Czech Republic such as the trends of overall environmental pressures related to material flows, the shifts of these pressures among countries due to international trade, efficiency of material use, recycling and product specificity of material-based pressures. The project applies approaches of ecological economics including economy-wide material flow accounting and hybrid input-output life cycle assessment method. These approaches are considered well established and suitable tools for monitoring and evaluation of human-induced material flows and related environmental pressures. The project has four objectives: 1) Extension of time series of economy-wide material flow indicators for the Czech Republic and assessment of their trends; 2) Compiling a material flow indicator related to recycling; 3) Extension of material flow indicators by water accounts; 4) Identification of product groups within the Czech economy, which are related to largest environmental pressures.

The article further shows some first results of the above project, which include the decrease in environmental pressure related to use of materials in 2009 and 2010, increasing rate in recycling and increasing pressure exerted on the environment due to surface water abstraction.

Key Words: Anthropogenic material flows, hybrid input-output life cycle assessment method, EW-MFA indicators, recycling rate, water abstraction

Introduction

This article presents a project supported by Grant Agency of the Czech Republic and carried out at the Charles University in Prague, Environment Center. The project is of highly interdisciplinary character and deals with various aspects of anthropogenic material flows in the Czech Republic.

The overall environmental pressure and impact caused by human societies is to a large extent induced by the consumption of energy and materials. Arguments supporting this claim have been put forward by a number of authors (e.g. Schmidt-Bleek, 1993; Ayres and Simonis, 1994; Weizsäcker and Lovins, 1997). Materials have to be first extracted from the environment. Then they are used to produce goods or deliver services required to satisfy human needs. Some are recycled, but they are all eventually released back into the environment in the forms of emissions and waste. All these stages of material processing have some environmental pressures and impacts. These include structural landscape changes, loss of biodiversity, acidification, eutrophication, global climate change and others (Giljum et al., 2005). In order to measure material and energy flows and to mitigate the related problems, material flow analysis has been conceived. The aim of this approach is to monitor material and energy flows at various levels of detail, and to provide indicators which contribute to management of resource use and output emission flows from both economic and environmental points of view (OECD, 2008). As convenient measures of sustainability, material flow indicators focusing mostly on an economy-wide level have been compiled for a range of both developed and developing countries (for instance, Adriaanse et al., 1997; Matthews et al., 2000; Giljum, 2004; Weisz et al., 2006; Weisz et al., 2007a).

Conceptual and methodological approaches

The project is based on two major methods: material flow analysis and the hybrid input-output life cycle assessment method.

Material flow analysis (MFA) is considered an approach of ecological economic, which study how ecosystems and economic activity interrelate (Proops, 1989). MFA uses some economic accounting principles and derives indicators, which relates both to economic variables such as gross domestic product (GDP) and to various environmental problems.

As a follow-up to pilot studies such as Steurer (1992), Schütz and Bringezu (1993), the Ministry of the Environment Japan (1992), Adriaanse and colleagues (1997) and Matthews and colleagues (2000), a first attempt to standardize economy-wide material flow analysis (EW-MFA) was undertaken by the statistical office of the European Union, which published a method guide for economy-wide material flow accounts and derived indicators (Eurostat, 2001). The standardization process was continued with the publication of a “compilation guide” for EW- MFA (Weisz et al., 2007b) and with the OECD work program on material flows (OECD, 2008). The aim of the EW-MFA is to quantify the physical exchange between a national economy, the environment and foreign economies on the basis of the total material mass flowing across the boundaries of the national economy. The ultimate goal of the analysis is to get a material balance, i.e., the state when material inputs into the economy equal material outputs summed with additions to the physical stock of the economy (for instance, traffic infrastructure, buildings and durable goods).

Material inputs into the economy consist primarily of extracted raw materials and produced biomass that have entered the economic system. Material outputs consist primarily of emissions to air and water, landfilled wastes and dissipative uses of materials (e.g. fertilizers, pesticides and solvents). The methodology also includes a concept of unused extraction (also called hidden flows). Unused extractions are material flows that have taken place as a result of resource extraction, but which do not directly enter the economic system. Examples include biomass left back in forests after logging, overburden from extraction of raw materials (such as in open-cast coal mining), earth movements resulting from the building of infrastructure, dredged deposits from rivers, etc. Foreign trade also plays an important role in the analysis because it represents an important material flow across the boundaries of the economic system. Imports of commodities are placed on the input side, while exports are placed on the output side of the material balance. Used and unused extraction are associated with foreign trade in the same way that domestic economic activities are (e.g. movement of overburden associated with imported coal), and are identified as indirect or up-stream material flows associated with imports and exports.

Based on the input and output flows, a large array of EW-MFA indicators can be compiled. These indicators include:

- Direct material input (DMI), which equals used domestic extraction (excavated raw materials, harvested biomass) plus imports;
- Raw material input (RMI), which includes used domestic extraction and up-stream material flows of imports related to used extraction (called raw material equivalents of imports - RMEIM);
- Total material requirement (TMR), which includes DMI plus unused domestic extraction and up-stream material flows of imports related both to used and unused extraction;
- Domestic processed output (DPO), which comprises emissions to air, landfilled wastes, the material load in wastewater and dissipative flows;
- Total domestic outputs (TDO), which includes DPO and unused domestic extractions;
- Domestic material consumption (DMC), which is calculated as DMI minus exports;
- Raw material consumption (RMC), which is calculated as RMI minus up-stream material flows of exports related to used extraction (called raw material equivalents of exports – RMEEX);
- Total material consumption (TMC), which equals TMR minus exports and their up-stream material flows related both to used and unused extraction; and
- Net additions to stock (NAS), which measures the physical growth rate of an economy.

EW-MFA indicators can serve various purposes, which include (OECD, 2008):

- Assessment of overall physical scale of the economy and total environmental pressure related to use of materials;

To study overall physical scale of the economy over time, it is advisable to refer to material flow indicators in absolute terms. It can be assumed that growing volume/mass of any material flow indicator will result in growing environmental pressure and impacts (Weizsäcker and Lovins, 1997; Bringezu et al., 2003).

- Assessment of shifts in environmental pressures among countries and regions related to international trade;

Many industrialized countries have decreased their amounts of domestically extracted and processed materials by importing them from other countries. The shift of pressure related to extraction and processing of these materials has taken place between states and world regions mainly to the detriment of developing countries (Schütz et al., 2004). To capture these shifts, it is necessary to study physical imports and exports and their up-stream flows.

- Measuring the efficiency of use of materials;

Relating input and consumption material flow indicators to national account aggregates, such as GDP, allows for measuring the efficiency by which an economic system transforms used materials into economic output. Such indicators reflect material productivity, i.e. the ratio of GDP over the material flow indicator, or material intensity, i.e. the ratio of the material flow indicator over GDP. Assessment of material intensity and productivity is complementary to analysis of decoupling, which study whether there is a decoupling between economic goods such as GDP growth and environmental bads such environmental pressures from material use.

- Assessment of foreign material dependency of countries and their material security;

Material flow indicators can be further used for monitoring of foreign material dependency. Economies fulfil their material demands partly from their own territory and partly by importing materials from other countries. The higher the share of imports in domestic material input and domestic material consumption is, the more the economy is susceptible to incidental shortage of particular commodities abroad, increase in their price or to upheaval of other barriers to foreign trade.

In order to assess environmental pressures related to whole life cycles of materials and products and evaluate properly shifts in environmental pressures between countries and regions it is necessary to include up-stream material flows (Commission of the European Communities, 2005a). This is done for imports and exports, which are then expressed in terms of all materials needed worldwide to produce imported/exported commodities (RMEIM, RMEEX). Two methods can be applied to quantify up-stream flows: a life cycle inventory (LCI) and hybrid input-output life cycle assessment method (hybrid IOA-LCA). The first method is quite straightforward. The volume of imports/exports is multiplied with coefficients, which express the volume of all materials that entered the particular product system throughout its life cycle. Hybrid IOA-LCA, on the other hand, is a combination of LCI and input-output analysis (IOA). IOA was first introduced by Leontief (1936, 1970) in order to monitor economic transactions among sectors. It can be, however, extended by some environmental variable such as extraction of resources or emissions. Hybrid LCA starts with the input-output model (IOM), which is represented by an input-output table (IOT) in monetary units. This model can be described by following equation:

$$1) \quad x = (I - A)^{-1} \cdot y,$$

Where x is a vector of total product output of the system, I is identity matrix, A is a matrix of technological coefficients and y is a vector of total final demand.

A matrix can be compiled for a product-by-product model or a sector-by-sector model, each of them can be further broken down into two types according to the assumed technology of production/sales of products. A sector-by-sector model under the assumption of fixed product sales structure of products will be used in this project. For this model, A matrix is calculated as follows:

$$2) \quad A = ST \cdot (\text{diag}(t - IM))^{-1} \cdot U \cdot (\text{diag}(g))^{-1}$$

Where S is a supply matrix as appears in the supply table, T denotes matrix transposition, t is the vector of total product use, IM is a vector of imports, U is a use matrix as shown in the use table and g is a vector of total sector output.

In the next step, an environmental extension represented by matrix F will be added to the model.

$$3) \quad e = F \cdot (I - A)^{-1} \cdot y$$

$$4) \quad F = F_r \cdot (\text{diag}(g))^{-1}$$

Where e is a vector of total induced material flows in the form of raw material equivalents and F_r is a matrix of used domestic extraction (DE) in physical units broken down by economic sectors.

The above calculation would assume that the imported commodities are produced abroad using the same production technology as the corresponding commodities in the domestic economy. Since this assumption need not hold for a range of products (especially when taking into account imports from developing), the results can be significantly distorted. To overcome this shortcoming, there are in general two possibilities: to build a multi-regional input output model, which uses country specific input output tables for the exporting countries, or to integrate life cycle inventory (LCI) data into the model for commodities, for which the domestic technology assumption does not hold. In this project, the second approach will be applied and LCI data will be used for natural gas, crude oil, metal ores and basic metals, which are not produced in the Czech Republic at all or only in minor quantities. These data will be integrated into the F_r matrix, use tables, supply tables and the vector of imports. This integration is denoted as ' in the following equations.

RMEIM and RMEEEX are calculated for product groups, but there is a sector-by-sector model so far. It is therefore necessary to transform the total final demand according to the sectors into the total final demand according to the products. This will be done by following equation:

$$5) \quad y = S^T \cdot (\text{diag}(t - IM^c))^{-1} \cdot y_p$$

Where y_p is a vector of total final demand according to the product groups.

To calculate RMEIM and RMEEEX the vector y_p will be sequentially substituted by the vectors of imports (the original vector of imports) and exports.

$$6) \quad RMEIM = F^c \cdot (I - A^c)^{-1} \cdot S^T \cdot (\text{diag}(t - IM^c))^{-1} \cdot IM$$

$$7) \quad RMEEEX = F^c \cdot (I - A^c)^{-1} \cdot S^T \cdot (\text{diag}(t - IM^c))^{-1} \cdot EX$$

It is also possible to calculate directly RMI and RMC indicators, which comprise RMEIM and RMEEEX, by using vectors of total final demand and domestic final demand.

$$8) \quad RMI = F^c \cdot (I - A^c)^{-1} \cdot S^T \cdot (\text{diag}(t - IM^c))^{-1} \cdot y_p = DE + RMEIM$$

$$9) \quad RMC = F^c \cdot (I - A^c)^{-1} \cdot S^T \cdot (\text{diag}(t - IM^c))^{-1} \cdot y_{pd} = DE + RMEIM - RMEEEX$$

Where y_{pd} is a vector of domestic final demand according to the product groups.

For more detailed description of the calculation procedure including discussion on various types of the model, which can be employed for calculation, see Weinzettel and Kovanda (2009).

Project objectives

The project has several objectives which contribute to the solution of selected research questions relevant for material flow analysis. Selection of these objectives takes into account current state of the art and previous Czech and international projects related to this topic. The objectives are principally fourfold:

1) Extension of time series of EW-MFA indicators available for the Czech Republic and assessment of their trends

EW-MFA indicators for the Czech Republic have been compiled within previous research and development projects of Ministry of the Environment (10/2/00, 320/2/03, 1C/7/14/04, SP/4i2/141/08) and Grant Agency of the Czech Republic (205/04/0582, 205/08/1475), which were worked on by applicant of this project proposal. Some of these indicators, namely DMI and DMC, are now calculated by the Czech Statistical Office (e.g. Czech Statistical Office, 2010a). This is not however true for more complex indicators such as RMC, TMR, TDO and TMC, which will be in the core of this objective. Time series of these indicators are now available up to 2008. Their extension for 2009-2012 within this project will help, for instance, to answer a question how world economic crisis of 2009 and 2010 has influenced material flows and related pressures in the Czech Republic. The indicators including up-stream flows will be calculated using both simple LCI and hybrid LCA, the results will be compared and it will be suggested which method delivers the more precise results. The calculated indicators will be analysed to answer following questions:

- Was there a change in the overall physical scale of the economy and total environmental pressure related to use of materials?
- Was there an increase in shifts in environmental pressures between the Czech Republic and other countries?
- Was there an increase in the efficiency of use of materials?

- Was there an increase in foreign material dependency of the Czech Republic?

2) Compiling a material flow indicator related to recycling

EW-MFA indicators as defined by Eurostat (2001) focuses on use of primary materials and on the flows of material back into the nature. The cyclical use of materials is not taken into account, as it is considered a flow, which does not cross the system boundary between economy and the nature. On the other hand it is widely acknowledged that the higher the recycling, the lower the need for primary materials (Commission of the European Communities, 2005b). The EW-MFA indicators and cyclical use of materials are therefore mutually interdependent. This was faced by Japanese “Fundamental plan for establishing a sound material-cycle society“ (Ministry of the Environment Japan, 2003), which designed a new indicator bringing together EW-MFA and recycling. It is named cyclical use rate and is defined as amount of cyclical use of materials divided by the sum of cyclical use of materials and DMI indicator. Since the “Fundamental plan for establishing a sound material-cycle society“ has been approved in 2003, the Japanese Ministry of the Environment monitors this indicator with the goal to reach 15 percent cyclical use rate by 2015.

The aim of this task is to compile cyclical use rate indicator for the Czech Republic for 2001-2012 (data on recycling are not consistent before and after 2001 due to changes in waste statistics) and to answer following questions:

- Is there an increasing trend in cyclical use rate in the Czech Republic?
- Is the cyclical use rate in the Czech Republic comparable to countries like Japan? Providing that not, what are the major barriers to reach similar cyclical use rate?

3) Extension of material flow indicators by water accounts

Material flow accounts usually exclude water and air flows (Eurostat, 2001). This is because water flows are typically ten times more voluminous than flows of other materials, which would be hardly visible in the total material balance. This is the reason why it is suggested to establish and keep water flow accounts separately from material flow accounts (Schandl et al., 1999). Within this project objective we establish the water accounts parallel to material flow accounts, which will be based on the same accounting principles as EW-MFA. The length of the time series will depend on data availability, but will go back to the year of 2000 at least. The water accounts will open the opportunity for further research under the “water footprint” methodology (Chapagain and Hoekstra, 2004). Besides the analysis of trend of water use, the major question to answer by establishing the water flow accounts will be as follows:

- What is the relation between consumption of water and the overall availability of water resources in the Czech Republic?

4) Identification of product groups within the Czech economy, which are related to largest environmental pressures

As acknowledged by various current research projects, the management of material flows and related pressures and impacts could be much more effective when focusing on products with highest material flows mobilized during their whole lifecycles. This issue has been addressed on the European level (Moll and Watson, 2009), but no similar study has been available for the Czech Republic yet. The lifecycle-wide material flows and emissions related to product groups will be calculated using a hybrid LCA approach by placing import, export and final demand vectors in equations 6)-9) on a diagonal. The results will be analysed from the following viewpoints:

- What are the products groups responsible for highest material flows? Are these the same in the Czech Republic as on the European level?
- Based on the review of the best available techniques for these product groups, what are the suggestions for reduction of related material flows and emissions?

Due the data intensity of this work, this analysis will be carried out for one pilot year for which the most recent data will be available.

First project results

The project is being carried out in the period 2012-2014. There are, however, some first results available already by now.

- 1) Extension of time series of EW-MFA indicators available for the Czech Republic and assessment of their trends

The analysis of trends of EW-MFA indicators revealed that many indicators went down in newly compiled years – the decrease was obvious above all in 2009 and could be attributed to global economic crisis. The overall physical size of economy and related environmental pressure therefore shrank in 2009 and 2010 (Figure 1).

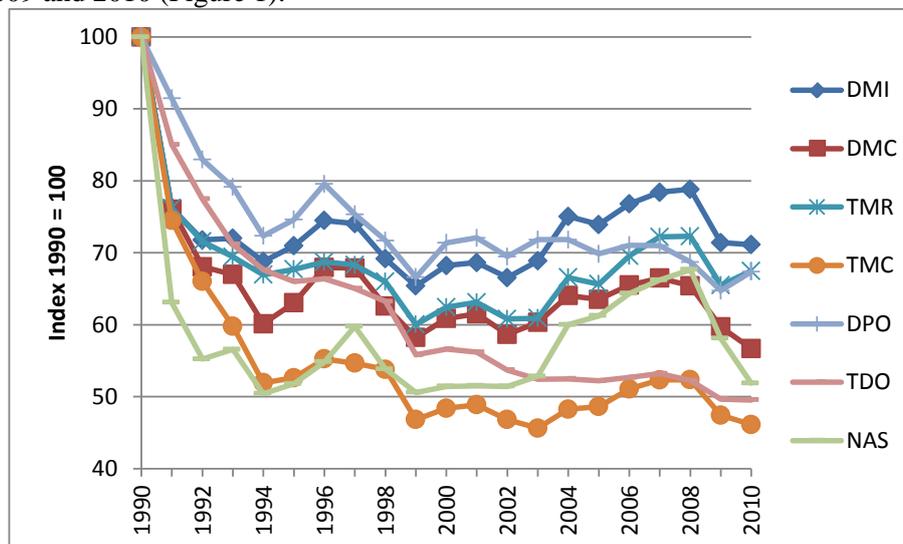


Figure 1: Overall trends of selected EW-MFA indicators, Czech Republic, 1990-2010

The indirect flows of imports and exports dropped in 2009, but grew again in 2010 on the levels of 2008. It means that there were not more profound changes in terms of shifts in environmental pressure between the Czech Republic and other countries. Conversely, the material productivity continued to grow in 2009 and 2010. This trend was, however, connected to unfavorable increase in foreign trade dependency, which grew by 3 percentage points between 2008 and 2010.

2) Compiling a material flow indicator related to recycling

The indicator of cyclical use rate was calculated in a standard and in a modified form, where cyclical use of materials included also imports of wastes, secondary materials and scarp while the same imports were subtracted from DMI. This modification can be substantiated by the fact that one of the goals of the indicator is to express the ratio of consumption of secondary (recycled) materials and primary raw materials.

The compilation and analysis of the indicator showed that it grew in the Czech Republic - it was 1.32 % in 2002, but 2.86 % in 2010. The modified form of the indicator showed values higher by 32 % on average and a very similar trend - it went up from 1.89 % in 2002 to 3.56 % in 2010 (Figure 2). The modification was therefore reasonable: the indicator attained significantly different values and stressed the importance of imports of waste and scrap for cyclical use of materials.

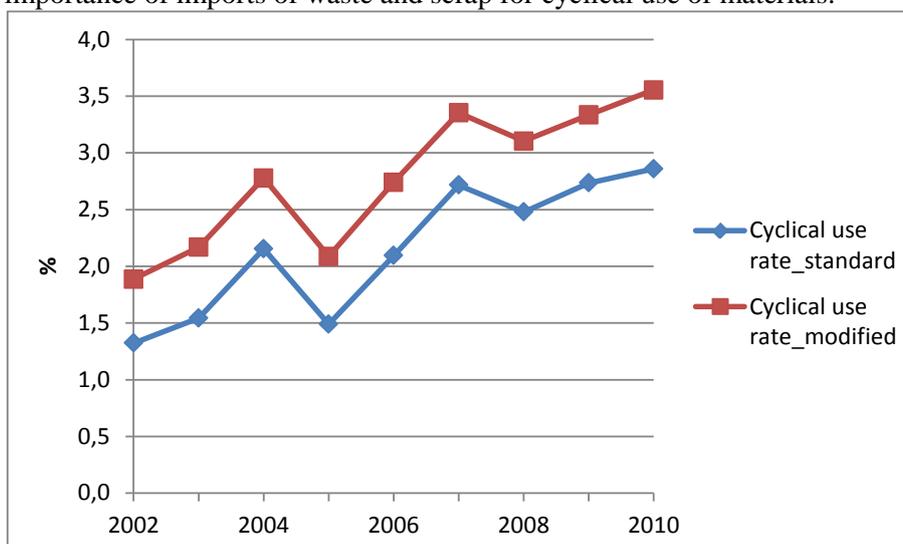


Figure 2: Cyclical use rate indicator (standard and modified form), Czech Republic, 2002-2010

Compared to Japan, for instance, the indicator recorded significantly lower values in the Czech Republic – only 2.72 % as compared to 13.5 % in 2007. The breakdown of the indicator by particular material groups revealed that the Czech Republic lagged especially behind in the cyclical use of biomass and non-metallic minerals (construction minerals). On the other hand the difference was not so large in the case of the cyclical use rate of metals (Ministry of the Environment Japan, 2010).

3) Extension of material flow indicators by water accounts

We have compiled water accounts for the Czech Republic for 2000-2011 so far. The water accounts comprised water abstraction in million cubic meters (m³) broken down by particular river basins (Labe, Vltava, Ohře, Odra, Morava), by surface water and ground water and by particular purchasers (water supply systems for public use, agriculture, energetics, industry, other including construction). Moreover, the renewable water resources were determined, which can be defined as the available volume of water for consumption in every year. They were determined as flow rate in major water courses with 95% security for surface water and as an educated guess of Ministry of Agriculture for ground water.

With the exception of Labe and Morava river basins, the volume of water abstraction from surface sources went slightly down in 2000-2011, but the increase in abstraction from Labe river basin was so profound, that the total water abstraction from surface water sources went up in this period (Figure 3). On the other hand, the total water abstraction from ground water sources went down in 2000-2011. The only river basin which showed an increase in ground water abstraction was Vltava (Figure 4).

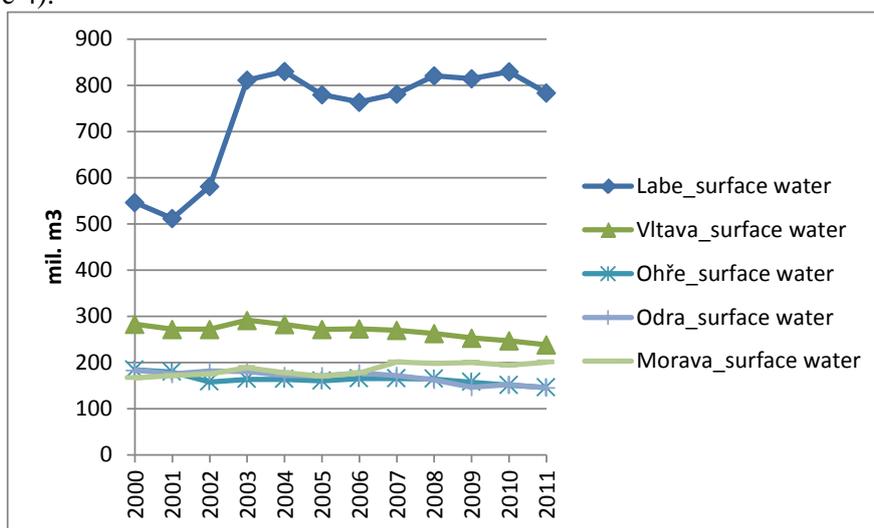


Figure 3: Water abstraction from surface sources, Czech Republic, 2000-2011

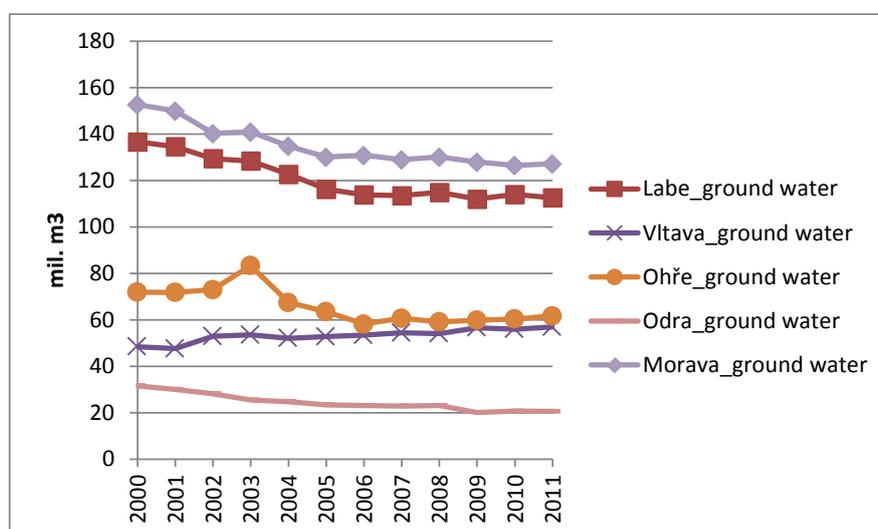


Figure 4: Water abstraction from ground water sources, Czech Republic, 2000-2011

4) Identification of product groups within the Czech economy, which are related to largest environmental pressures

From the viewpoint of raw material extraction, the most demanding product groups included motor vehicles, basic metals, buildings, machinery and equipment, hard and brown coal, electricity and gas, food, chemical substances and products, civil engineering works, agricultural products, non-metallic mineral products, computers and optical appliances, fabricated metal products, electrical appliances, non-metallic raw materials, real estate services, coke and refinery products. Taking into account the fact that environmental pressure related to non-metallic minerals is lower compared to fossil fuels, biomass and most significantly to metal ores, lower priority from above product groups could be given to buildings, civil engineering works, non-metallic raw materials and non-metallic products, as their production is linked to non-metallic minerals. The most emission intensive product groups are quite similar to the most raw material intensive product groups, but their order is different and for emissions to the air also land transport showed up as very important: it reached 9th position in case of greenhouse gases and PM total, 14th position in case of SO₂, 7th position for NO_x and 8th position for CO.

Conclusion

Although the project described in this article has not been finalized yet, there are some first project results available right now. It can be concluded that many EW-MFA indicators in the Czech Republic went down in newly compiled years and due to this fact the overall physical size of the economy and related environmental pressure shrank in 2009 and 2010. It further showed that the Czech Republic suffered from increasing foreign trade dependency, but the efficiency of transformation of material inputs into economic output grew. The work on the cyclical use rate indicator proved that the recycling is on the upswing in the Czech Republic, but compared e.g. to Japan, the overall recycling rate is still much lower. The calculation of modified version of this indicator also emphasized the significance of imported wastes, secondary materials and scrap for overall recycling. We have further compiled the water accounts, which showed an increase in total surface water abstraction in 2000-2011 mostly due to growth in water abstraction from Labe river basin. Unlike pressure on the environment related to material flows of solid materials covered by EW-MFA, the environmental pressure related to surface water abstraction thus went up quite significantly in the first decade of the 21st century. Regarding the product groups most demanding on extraction of raw materials, they include basic metals, fabricated metal products, motor vehicles and machinery. These commodities require large amounts of metal ores to be mined and processed. This is not a favorable discovery, as mining and processing of metal ores is usually related to significant impacts on the environment and human health.

Acknowledgments

The research on the subject presented in this article was undertaken as a part of the research and development project No. P402/12/2116 Various aspects of anthropogenic material flows in the Czech Republic, funded by the Grant Agency of the Czech Republic. This support is gratefully appreciated.

References:

- Adriaanse, A., S. Bringezu, A. Hammond, Y. Moriguchi, E. Rodenburg, D. Rogich, and H. Schütz. Resource flows: The material basis of industrial economies. Washington, DC: World Resource Institute, 1997.
- Ayres, R. U. and U. E. Simonis (Editors). Industrial metabolism: Restructuring for sustainable development. Tokio: UNU Press, 1994.
- Bringezu, S., H. Schütz, and S., Moll. Rationale for and interpretation of economy-wide material flow analysis and derived indicators. *Journal of Industrial Ecology* 2 (7), 2003: 43–64.
- CENIA. Report on the environment in the Czech Republic in 2009. Prague: Ministry of the Environment, 2010.
- Chapagain, A.K., Hoekstra, A.Y. Water footprints of nations. Research report Series No. 16. Paris: Institute for Water Education, 2004.
- Commission of the European Communities. Review of the EU sustainable development strategy (EU SDS) - Renewed strategy. Brussels: Commission of the European Communities, 2006.

Commission of the European Communities. The sixth environment action programme of the European Community. Decision No 1600/2002/EC of the European Parliament and of the Council. Brussels: Commission of the European Communities, 2002.

Commission of the European Communities. Thematic strategy on the sustainable use of natural resources. Communication from the commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions. Brussels: Commission of the European Communities, 2005a.

Commission of the European Communities. Thematic strategy on the prevention and recycling of wastes. Communication from the commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions. Brussels: Commission of the European Communities, 2005b.

Czech Hydrometeorological Institute. Emission balances of the Czech Republic. http://www.chmi.cz/files/portal/docs/uoco/oez/emisnibilance_CZ.html.

Czech Statistical Office. Annual national accounts database. <http://apl.czso.cz/pll/rocenka/rocenka.indexnu>.

Czech Statistical Office. Material flows accounts in the CR 2004 - 2009 (selected indicators). Prague: Czech Statistical Office, 2010a.

Czech Statistical Office. Waste generation and treatment in CR in 2009. Prague: Czech Statistical Office, 2010b.

Ecoinvent. National life cycle inventory database Ecoinvent, data v1.3. Dübendorf: Swiss Centre for Life Cycle Inventories, 2006.

Eurostat. Economy-wide material flow accounts and derived indicators. A methodological guide. Luxembourg: Eurostat, 2001.

Giljum, S. Trade, material flows and economic development in the South: The example of Chile. *Journal of Industrial Ecology* 8, 2004: 241–261.

Giljum, S., T. Hak, F. Hinterberger, and J. Kovanda. Environmental governance in the European Union: Strategies and instruments for absolute decoupling. *International Journal of Sustainable Development* 8(1/2), 2005: 31–46.

Government Council for Sustainable Development. Progress report on the Czech Republic sustainable development strategy. Prague: Ministry of the Environment, 2009.

Government Council for Sustainable Development. Strategic framework for sustainable development in the CR. Prague: Ministry of the Environment, 2010.

Leontief, W. Environmental repercussions and the economic structure: An input-output approach. *Review of Economics and Statistics* 52(3), 1970: 262–271.

Leontief, W. Quantitative input-output relations in the economic systems of the United States. *Review of Economics and Statistics* 18(3), 1936: 105–125.

Matthews, E., C. Amann, S. Bringezu, M. Fischer-Kowalski, W. Hüttler, R. Kleijn, Y. Moriguchi, C. Ottke, E. Rodenburg, D. Rogich, H. Schandl, H. Schütz, E. van der Voet, and H. Weisz. The weight of nations: Material outflows from industrial economies. Washington, DC.: World Resources Institute, 2000.

Ministry of Agriculture. Water Management Information Portal. <http://www.voda.gov.cz/portal/cz/>.

Ministry of the Environment Japan. Establishing a sound material-cycle society: Milestone toward a sound material-cycle society through changes in business and life styles. http://www.env.go.jp/en/recycle/smcs/a-rep/2010gs_full.pdf, 2010.

Ministry of the Environment Japan. Fundamental plan for establishing a sound material-cycle society. www.gdrc.org/uem/waste/japan-3r/4-basicplan.pdf, 2003.

Ministry of the Environment Japan. Quality of the environment in Japan 1992. <http://www.env.go.jp/en/wpaper/1992/index.html>, 1992.

Moll, S. and D. Watson. Environmental pressures from European consumption and production: A study in integrated environmental and economic analysis. Copenhagen: European Environment Agency, 2009.

OECD. Measuring material flows and resource productivity. Volume I: The OECD guide. Paris: OECD, 2008.

Proops, J. Ecological economics: rationale and problem areas. *Ecological Economics* 1, 1989: 59–76.

- Schandl, H., W. Hüttler, and H. Payer. De-linking of economic growth and materials turnover. *The European Journal of Social Sciences* 12 (1), 1999: 31–45.
- Schmidt-Bleek, F. *Wieviel Umwelt Brecht der Mensch? MIPS—das Maß für ökologisches Wirtschaften* [How much environment Leeds humankind? MIPS—The measure for ecological economic activities]. Berlin: Birkhauser, 1994.
- Schütz, H. and S. Bringezu. Major material flows in Germany. *Fresenius Environmental Bulletin* 2: 443–448, 1993.
- Schütz, H., S. Moll, and S. Bringezu. Globalisation and the shifting environmental burdens. Material trade flows of the European Union. Wuppertal: Wuppertal Papers 134, 2004.
- Steurer, A. *Stoffstrombilanz Österreich 1988* [Material flow accounting Austria 1988]. Social ecology working paper. Vienna: IFF Social Ecology, 1992.
- Weinzettel, J. and J. Kovanda. Assessing socioeconomic metabolism through hybrid life cycle assessment: The case of the Czech Republic. *Journal of Industrial Ecology* 13(4), 2009: 607–621.
- Weisz, H. Accounting for raw material equivalents of traded goods: a comparison of input–output approaches in physical, monetary, and mixed units. Social ecology working paper. Vienna: IFF Social Ecology, 2006.
- Weisz, H., F. Krausman, N. Eisenmenger, H. Schütz, W. Haas, and A. Schaffartzik. Economy-wide material flow accounting "A compilation guide". Luxembourg: Eurostat, 2007b.
- Weisz, H., F. Krausmann, and S. Sangarkan. Resource use in a transition economy: Material and energy flow analysis for Thailand. Social ecology working paper. Vienna: IFF Social Ecology, 2007a.
- Weisz, H., F. Krausmann, C. Amann, N. Eisenmenger, K. Erb, K. Hubacek, and M. Fischer-Kowalski. The physical economy of the European Union: Cross-country comparison and determinants of material consumption. *Ecological Economics* 58(4), 2006: 676–698.
- Weizsäcker, E. U. and A. Lovins. *Factor four: Doubling wealth – halving resource use*. London: Earthscan, 1997.
- Wiedmann, T. A review of recent multi-region input–output models used for consumption-based emission and resource accounting. *Ecological Economics* 69(2), 2009: 211–222.
- Wiedmann, T., M. Lenzen, K. Turner, and J. Barrett. Examining the global environmental impact of regional consumption activities, Part 2: Review of input–output models for the assessment of environmental impacts embodied in trade. *Ecological Economics* 61(1), 2007: 15–26.

LIFE CYCLE EVALUATION IN THE NATIONAL PARK MURÁŇ PLAIN (SLOVAKIA)

Judita Tomaskinova, RNDR., PhD

Department of Environment, Faculty of Natural Sciences,
Matej Bel University in Banská Bystrica, Slovakia

Jan Dubiel, Ing.

Department of Biology and Ecology, Faculty of Natural Sciences,
Matej Bel University in Banská Bystrica, Slovakia

Abstract:

The article presents the evaluation results of „life cycle“ of the National Park Muráň Plain (Slovakia) based on self-assessment of stakeholders by IPAM methodology (Integrative Protected Area Management). Based on this methodology, we identified areas of management that should be targeted for improvement also recommendations are proposed to streamline management. Based on the results of evaluations, we can state that Basic Planning Phase in life cycle of evaluated area is least effective developed. For improvement of other phases is first of all necessary to improve communication and to ensure transparency in the process of providing information. In terms of relevant documents, which will contribute to more effective integrated management of nature and landscape protection in the NP Muráň Plain, we suggest a high attention to the most expeditious approval of the proposed zoning of the national park.

Key Words: National Park Muran Plain, life cycle protected area, integrated management, protected area, IPAM

Introduction

The role of protected areas (PAs) is the protection of biodiversity and the most valuable natural and landscaping parts of nature and landscape. Concurrently, the mission of protected areas has expanded from biodiversity conservation to improving local socioeconomic benefits in the context regional sustainable development. Many best practice examples show that nature protection can be a good prerequisite for local and regional economic development such as in the Bavarian Forest national park. With effective and efficient management, sustainable development – economic, ecological and social – can be promoted for the *advantage* of the region and the whole economy (Getzner, Jungmeier, Lange, 2010, Vološčuk, 2008). Naughton-Treves, Buck Holland & Brandon (2005) also state that by global mandates, protected areas are now supposed to do far more than conserve biological diversity. These areas are charged with improving social welfare, guarding local security, and providing economic benefits across multiple scales, objectives traditionally relegated to the development sector. Managers of protected areas, although often accused of being unconcerned with social issues, have significantly altered their approach in an attempt to meet the new role for protected areas. In many cases, conservation organizations formed new partnerships with development agencies and institutions, as well as citizens' groups. Together they have pursued an array of strategies linking conservation with development that generally fall into three broad groups: community-based natural resource management, community-based conservation, and integrated conservation and development. Švajda (2008a) states that by the evaluation of protected areas management in our conditions (Eastern Europe) every important role plays the fact that our protected areas mostly came into existence in the period of communist regime which was characterized by absence of or by very low level of discussion with all relevant stakeholders. At the same time he claims that to the key reasons for generally low support for local communities and local inhabitants to conservation of nature belongs also complicated and different land ownership.

Based on international experience, guidance, expert assessment and our evaluation is the best way Integrated Protected Area Management on the base IPAM, which inclusive the design and development of a protected area during the “life cycle”.

Material and Methods

The National Park Muráň Plain with its wild mountain karstic landscape and minimum human interventions is located between central and eastern parts of Slovakia. It lies in the Spiš-Gemer Karst, in the Slovenské Rudohorie Mountains (part of the Western Carpathian Mountains). National park Muráň Plain is one of the youngest national parks in Slovakia declared in October 1997 and opened on 27 May 1998.

We have been evaluating the „life cycle“ in the National Park Muráň Plain (Slovakia) in years 2011 – 2012 by methodology integrative protected area management (IPAM). IPAM can divide the “life cycle” of each PA into four principal phases (Planning Phase is split in two periods – Basic Planning Phase and Detailed Planning Phase) (Fig. 1) and 29 related “Fields of Activity” (FoA) (Tab. 1) (Jungmeier et al., 2005, Getzner et al., 2012).

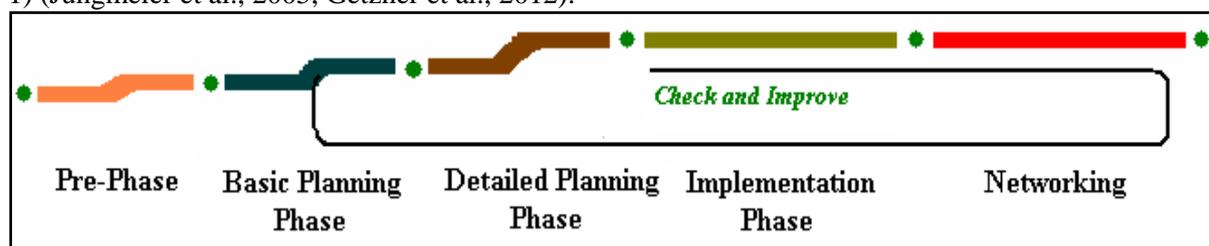


Fig. 1 The principal phases of protected area management life cycle (Modified by: Wagner et al., 2005)

This methodology secures interactive communication focused on identification of problems together with finding of the optimal solutions for the protected areas management in the middle and east of Europe. Expert system IPAM consists of three components: **self-assessment, standardized measures and knowledge basis**, which are designed to provide any information needed for the development of particular protected area. For the evaluation of individual activities in every phase of protected area management by stakeholders (groups involved) was used the method of questionnaire survey. Each question of the questionnaire enables three-scale evaluation (not yet started – started - finished), that forms input data for computer software processing. By the detailed evaluation of protected area management the methodology concentrates on the detailed evaluation of individual phases of protected area life cycle, whereas in the planning phase the basic and detailed planning is evaluated separately. Expert system warns about placement of control and improvement before the actual phase of implementation of the integrated management (Jungmeier et al., 2005). The overview of the evaluated areas of activity in individual phases of protected area management life cycle is described in Tab. 1.

Tab. 1 The evaluated areas of activity in individual phases of protected area management life cycle (Source: Jungmeier et al., 2005)

Phase	Fields of the activity	
<i>Pre-Phase</i>	Development of ideas and visions	
	Realisability controlling	
	Communication and participation I.	
	Integration into the system of protected areas	
<i>Planning Phase</i>	<i>Basic Planning Phase</i>	Planning guide
		Communication and participation II.
		Basic research
		Planning
	<i>Detailed Planning Phase</i>	Determination and introduction
		Keynote address and essential papers
		Management plans based on ecosystems
		Proposal of (regional) economical plans
		Specific planning (subsidiary plans)
		Personal and organizational development
Region development of protected area		

<i>Implementation Phase</i>	Evaluation of management efficiency
	Framework of research and monitoring
	Management of data and information
	Optimization of financial situation
	Information, interpretation, education
	Visitors management, services and infrastructure
	Marketing and communication with public
	Cooperation scheme
	Communication and participation III.
	Evaluation of influence and limits
<i>Networking</i>	Universal integration into the net
	Integration into economical net
	Integration into social net
	Integration into ecological net

We used the Three-aspectual system of classification for the identification of stakeholders (Fig. 2) By the term *stakeholder* we label all persons, institutions or organizations, that influence management running or they are somehow influenced by its running (Zelený, 2008).

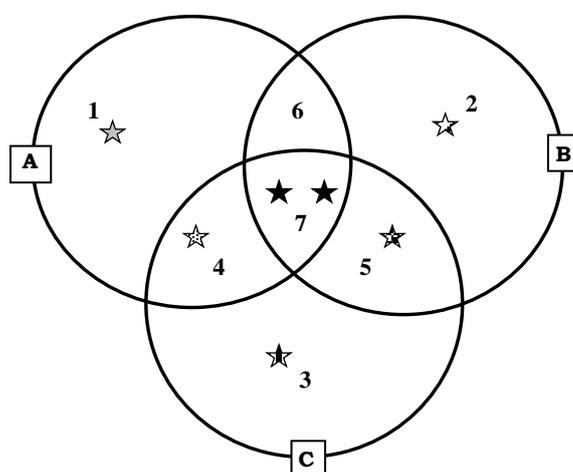


Fig. 2 Three-aspectual system of stakeholder classification (Source: Zelený et al., 2010)

Combination of the aspects: **A** – power of the influence, **B** – legitimacy (of requirements), **C** – evaluation of awaited stakeholder's approach to potential co-operation.

Groups of stakeholders: **1.** Stakeholders with the possible influence on management, but with relatively low level of legitimacy and insistency on fulfilling own requirement, **2.** Stakeholders with requirements legitimacy in some way, but with relatively low level of influence and insistency on their fulfilling, **3.** Stakeholders relatively insisting on fulfilling of their own requirements, but with relatively low level of their legitimacy and influence, **4.** Relatively “dangerous” stakeholders thanks to the combination of their influence and insistency, with which they are demanding the fulfilling of their requirements and rights, but which have a relatively low level of legitimacy, **5.** Stakeholders with the combination of legitimate requirements and rights, which they are demanding to be fulfilled, but with a relatively low level of influence, **6.** Stakeholders with the combination of legitimate requirements and rights, with a relatively high rate of influence, but with a lower demanding of their fulfilling, **7.** Key stakeholders.

Results

Evaluation of management effectiveness is one of really important part, even essential part in the life cycle of a protected area, because it makes a space for continuous improvement. Group of environmental stakeholders was the most suited to our set criteria (according to sphere of interests) for evaluation of the protected area using expert system IPAM. Along the life-cycle of Muráň Plain were evaluated 29 areas of management at intervals of three scalable evaluation of life-cycle phases

evaluated PA made by key stakeholders. We present complex assessment of whole life – cycle after assessment by expert system IPAM in the following graph (Fig. 3). From the management analysis results follows that is important to improve Base Planning Phase, Detailed Planning Phase and also the Networking. Efficiency improvement of communication and participation in every phase of life cycle of National Park Muráň Plain has a priority. For its improvement is necessary to work out a stakeholders mind map and also a project on communication, which should come from the specification of target groups, determination of order and the way of communication with help of existing nets, structures and media and it shouldn't forget about the difference between core information for all of them and specific detailed information.

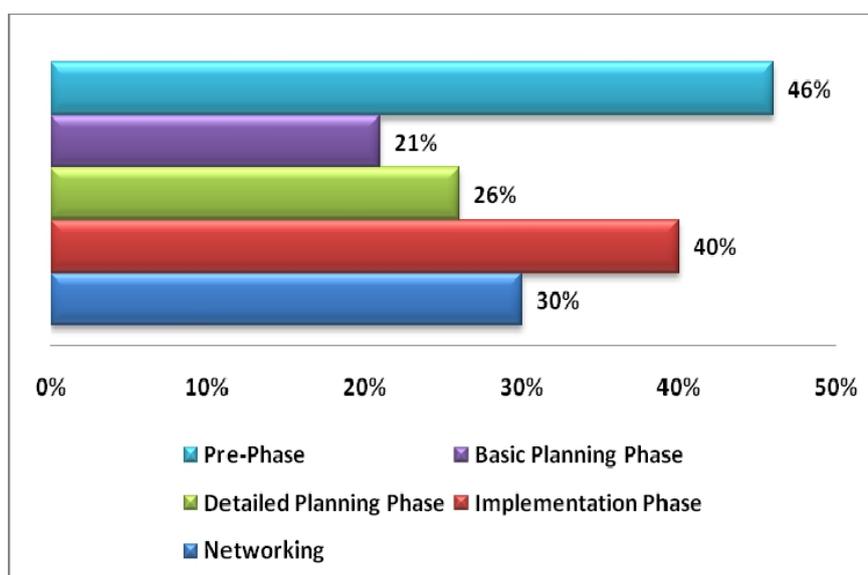


Fig. 3 Complex evaluation of basic phases of life – cycle by expert system IPAM

The weakest part in the management of evaluated National Park is Basic Planning Phase (Fig. 4). Very important recommended step of this phase to take action for immediate kick-off of the fields of activities: Planning Handbook and Implementation Planning.

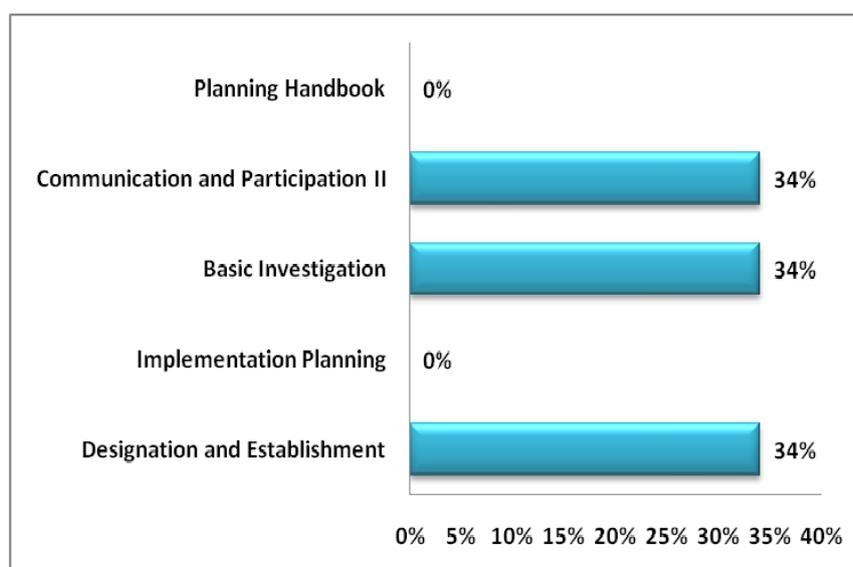


Fig. 4 Evaluation of Basic Planning Phase

We propose to elaborate „Plan puzzle“ to get a picture of all questions, disciplines and planning issues and identify core aspects for streamlining of this phase. Define and set up the planning process using professional project management tools. Very important is specify the content and

dimension of the planning, the procedure, the methodologies applied and technical details (like specifying planning documents). To streamline the management of FoA Implementation Planning is necessary to pay attention to the development of Zonation Plan. In this phase, top management of national park must be prepared for a difficult procedure and intensive discussions with all stakeholders. Depending on the management category of the Protected Area, a set of minimum requirements needs to be considered when fixing the outer boundaries and the inner zones (e.g. size of Protected Area, types of zones, management objectives, type and extent of land use etc.). This planning task is extremely complex as it nowadays basically means reaching an agreement on a sovereign act by applying participatory methods.

The boundaries and zonation planning is embedded in the contradictory context of:

- Public / sovereign versus private / participatory act
- Individual versus collective decision
- Self-determination versus heteronomy
- Conservation requirement versus land use

To overcome these contradictory aspects set up a multiple-stage process with the landowners providing repeated exit opportunities. This process should contain the following steps:

- Outline the process (define the steps, the timing, the responsibilities etc.)
- Find out name and contact details of all landowners involved
- Define conservation requirements
- Present the PA idea / planning process to all the landowners (e.g. an information evening assuring that everyone involved gets the same information at the same time)
- Hold conversations with each landowner (individual or on a group level)
- Go ahead with in this iterative procedure up until you achieve an agreement on an individual level
- Aggregate and fix the boundaries and zonation

Top management must bear in mind that the minimum criteria as requested by the management categorisation system need to be met and reflected in the final version of the boundaries and zonation planning.

On the basis of this complex analysis we state the recommended steps that are necessary to be taken in every FoAs of NP Muráň Plain life cycle for a purpose of efficiency improvement of its management (Tab. 2).

Tab. 2 High priority recommendations of Fields of Activity for National park Muráň Plain

Field of Activity	Recommended Action
Feasibility Check	Chance-risk-analysis
	Acceptance zoning
Communication and Participation I	Communication design
Planning Handbook	Plan puzzle
Communication and Participation II	Checklist transparency
Implementation Planning	Boundaries and Zonation plan
Designation and Establishment	(Intern)National application
Ecosystem-based Management Plans	Priorities and measures
	Calculation of costs and finances
Design of (Regional) Economic Programmes	Product / Service - Portfolio
	Product / Service - Platform
Personnel and Organisational Development	Personal development plan
Evaluating Management Effectiveness	Management cycle
	Monitoring and benchmarking
Financing (Business Plan)	List of benefits
	New incomes
Communication and Participation III	Permanent communication
Development of Protected Area's Region	Regional Economic Program
	Info-Platform
Networking General	Action plan on network optimization
Networking Ecological	Ecological profile

Discussion and Conclusion

It is indisputable that protected areas have many positive benefits. Their benefits extend from local people who live near this beautiful and healthy landscapes, to nations that depend on their environmental services (for example clear water, fresh air, regulate the climate, regulate the water, erosion control, biofiltration, recreation, etc.). They are not only ecological but also socio-economic benefits. Stockmann (2007), Diller, (2008), Švajda (2008b) state that evaluations of PAs management in general may fulfil several aims and objectives in modern societies. It is not only necessary to evaluate products (outputs) but to assess also outcomes in terms of legitimacy of public activities, and to provide steering mechanisms for public decision-makers. We agree with Pflieger (2008) who state that during the life-cycle of a protected area (PA), the evaluation of management effectiveness of the PA administrating and managing authorities including the whole PA context becomes increasingly important, both for securing and improving the conservation of biodiversity, but also for the acceptance of stakeholders and funding bodies. Currently, there are numerous approaches to evaluating management effectiveness of parks; many international institutions have drafted and implemented such evaluation instruments. At the beginning of any evaluation exercise, PA management and policy makers have to be convinced of the usefulness of the evaluation to increase the probability that recommendations will indeed be implemented. The implementation of the evaluation results is crucial since the evaluation might lead to higher costs than benefits since the frustration of those involved in the evaluation process might be significant (Hockings et al., 2006). However, the lack of implementing evaluation results is, at the moment, one of the biggest problems in assessing protected area management effectiveness. Evaluations are carried out, but recommendations are not implemented in daily management (Steindlegger, 2007).

Widely publishes the results of life-cycle evaluation of PAs, brings far more benefits than just management recommendations. The debate on appropriate indicators and evaluation tools, has intensified in the last decade. However, any learning and progress in and about protected areas management is clearly connected to the intensity of the public debate provoked by the evaluation process.

Effectively implemented integrated protected area management in the whole life cycle brings the following expected effects:

Expected Effects

- Integration of protected area management into regional economy and rural development
- Improving the quality of protected area management
- Raising public awareness of the complex tasks of protected area management
- Providing an effort to link protected area management with the tasks, instruments and tools of spatial planning
- Support to implement European standards, policies, procedures and technologies

There will be a need to reconsider the organisational structure in National park Muráň Plain, not necessarily with a view to changing for its own sake, but to determine whether it is fit for delivering the strategic re-orientation and the engagement of staff with a range of skills to cope with partnership working, stakeholder engagement and delivering societal benefit. The ideal structural model is shown in Fig. 5.

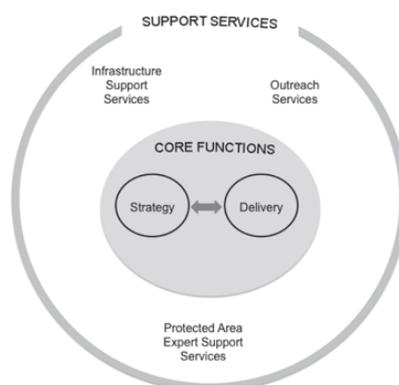


Fig. 5 Ideal structural model of a PA (Source: Authors' draft based on Croft , 2010)

It is not a structural diagram as there is no ideal structure. Managers should ensure that the key functions performed by the organisation distinguish between the 'core functions' and 'support services'. The former are those that directly achieve the mission by developing the strategy and ensuring the ability to review performance, and delivering the strategy on the ground within the PA and in cooperation with the 'communities of interest'. Support services are those that ensure the organisation is operated as efficiently as possible:

- PA Expert Support Services are natural and social scientific expertise on all aspects of knowledge gathering, analysis, information provision for management of the PA and for dissemination to the public and community of interests;

- Infrastructure Support Services include information technology, information management, human resources, and financial management; and - Outreach Services include partnership working, education and communication, commerce and business development, and fund raising (Getzner, Jungmeier, Lange, 2010).

References:

- Diller, Christian. Evaluierungen und Regional Governance: Funktionen der Evaluierung von und in regionalen Steuerungsstrukturen. *Zeitschrift für Evaluierung* 2008, 8:273-301.
- Getzner Michael, Jungmeier Michael, Lange Sigrun. *People, Parks and Money: Stakeholder involvement and regional development*. Klagenfurt: Verlag Johannes Heyn, 2010.
- Getzner, Michael et al. *Evaluating Management Effectiveness of National Parks as a Contribution to Good Governance and Social Learning*. Klagenfurt: InTech, 2012.
- Hockings Marc, Stolton Sue., Leverington Fiona, Dudley Nigel, Courrau José. *Evaluating effectiveness: a framework for assessing management effectiveness of protected areas*. Gland (Switzerland)/Cambridge (UK): IUCN, 2006.
- Jungmeier, Michael et al. *IPAM-Toolbox. Transnational Results (Expert System, Toolbox and Best Practice)*. Klagenfurt: E.C.O. Institute for Ecology Ltd., 2005.
- Naughton-Treves Lisa, Buck Holland Margaret, Brandon Katrina. The role of protected areas in conserving biodiversity and sustaining local livelihoods. *Annu. Rev. Environ. Resour.* 2005. 30:219–252.
- Pfleger, Bernd. Management effectiveness assessment in Austria – site consolidation scorecard. In: Stolton, Sue ed.: *Assessment of management effectiveness in European protected areas*. Germany: BfN and EUROPARC Federation on the Island of Vilm, 2008.
- Steindlegger, Gerald. WWF International – Member of the WWF Global Protected Area Team. Personal communication, 2007.
- Stockmann, Reinhard. Evaluation in der Gesellschaft: Entwicklung, Stand und Perspektiven. *Zeitschrift für Evaluation*, 2007, 6:195-222.
- Švajda, Juraj. Participatory conservation in a post-communist context: The Tatra National Park and Biosphere Reserve, Slovakia. *International Journal of Biodiversity Science and Management*, 2008a. 4:200-208.
- Švajda, Juraj. *Monographic study of national parks*. Zvolen: Vydavateľstvo Technickej univerzity vo Zvolene, 2008b.
- Vološčuk, Ivan. The threat of natural systems and the development of concepts of nature conservation. Zvolen: Asociácia karpatských národných parkov, Slovenská ekologická spoločnosť pri SAV, Slovenský národný komitét pre program UNESCO Človek a biosféra, 2008.
- Zelený, Ján. *Environmental management and corporate social responsibility*. Banská Bystrica: Faculty of Natural Sciences, University Matej Bel in Banskej Bystrici, 2008.
- Zelený, Ján et al. *Environmental policy, management and stakeholder management : Environmental and integrated management*. Banská Bystrica: Fakulta prírodných vied UMB, 2010.
- Wagner, Johann, Jungmeier, Michael, Kirchmeier Hanns, Velik Iris, Zoller Daniel. *IPAM toolbox – integrative protected area management. An expert system for integrative planning and management of protected areas*. Klagenfurt: Office of the Carinthian Government, 2005.

Acknowledgements

The study was written as part of the grant tasks APVV-0663-10 and VEGA 2-0065-11.

GENERATION AND COLLECTION PRACTICES OF ORGANIC KITCHEN WASTE IN HOUSEHOLDS OF MANIPAL

Lalitha Simon

Shivali Patel

Anysha Mattu

Teresa Chirayil

KMC International Center, Manipal University, Manipal, India

Abstract:

The present study was conducted to understand the organic kitchen waste management practices in households of Manipal. The general steps in waste management are 1. Generation 2. Collection 3. Sorting 4. Separation 5. Transfer 6. Transport and 7. Disposal. This study was aimed at understanding the management practices at the first two steps only i.e. Generation and Collection. Data was collected by interviewing the household members, with the help of a structured questionnaire. The study brought out the various types of wastes generated, the constituents and their quantum in organic kitchen wastes as well as the activities which happen at the generation and at the two waste collection points.

Key Words: Organic kitchen waste, generation, collection

Introduction

Disposal of household waste is one of the biggest challenges that towns and cities all over the world is facing. Multifold are the problems associated with the disposal of waste from all three streams of waste i.e. municipal waste, commercial and industrial waste and construction or demolition waste. Kitchen waste forms a significant constituent of municipal waste. The composition of municipal waste varies greatly from country to country and changes significantly with time. Municipal waste is generally categorized as (i) Biodegradable waste: food and kitchen waste, green waste, paper (ii) Recyclable material: paper, glass, bottles, cans, metals, certain plastics, fabrics, clothes, batteries (iii) Inert waste: construction and demolition waste, dirt, rocks, debris (iv) Electrical and electronic waste (WEEE) - electrical appliances, TVs, computers (v) Composite wastes: clothing, tetra packs, waste plastic such as toys. (vi) Hazardous waste including paints, chemicals, light bulbs, fluorescent tubes, spray cans, fertilizer and containers (vii) Toxic waste: pesticides, herbicides, fungicides and (viii) Medical waste [1]. These wastes have different characteristics and can be divided into those which are bio-degradable and not, as well those which can be re-cycled or not. The ability to manage them further, comes from the understanding of wastes.

Municipal waste management is a challenging task. It consists of various stages such as recycling, composting, land filling and waste-to-energy via incineration. The primary steps are generation, collection, sorting and separation, transfer and transport, and disposal [2]. Collection includes the gathering of the generated waste which will either be disposed in a landfill disposal site or taken to a further processing facility. There are two collection points; one at the individual house and the other at common points of waste collection in a locality. The third step is the sorting and separation. Separation and sorting of waste at both the household and common collection points are usually done. One usual categorization of household waste is by segregating the waste into biodegradable and non-biodegradable as well as recyclable and non-recyclable. The fourth step is the transfer and transport. This element involves two main steps. First, the waste is transferred from a smaller collection point to a larger transport point. This waste is then transported, usually over longer distances, to a processing or disposal site. The final step is the disposal. The further processing depends on the philosophy of the municipality, concern for environment, and the sensitivity of the

residents and possibility of commercial viabilities. This study focuses on the nature of organic kitchen waste and the generation and collection of it at different households.

This study focused on the organic waste from households [5]. Kitchen waste contains lots of rich nutrients which include carbohydrates, lipids and protein compounds that are not harmful. A more detailed analysis of kitchen waste shows that it can be further divided into two categories i.e. waste of vegetarian food items and waste of non-vegetarian food items. A study done in an urban community in a state of North India, is shown in Table-2[6].

Table 1: Constituents of organic kitchen waste.

Constituents	% composition*
1. Cooked rice	23.00-59.50
2. Cereals	8.76-18.93
3. Cooked vegetables	5.68-16.27
4. Chappatti	18.67- 41.90
5. Vegetable waste residue	1.37-2.59
6. Salad	17.40-29.50
7. Misc/ non vegetarian items	0.79-2.16

*the average, based on a study conducted on a few households observed over a few days.

Kitchen waste has chemical characteristics which are beneficial as well as non- beneficial for the natural environment. Kitchen waste is usually acidic. Since there is a wide variation in food content, there can be a wide variation in Chemical Oxygen Demand (COD) content. COD indicates its decomposing nature and thus its bio degradable characteristics. Some of the chemical characteristics of kitchen waste are shown in (Table-2) [6].

Table 2: Chemical characteristics of organic kitchen waste

Parameter	Median value
Density, kg/m ³	797 ± 51.33
pH	5.9 ± 0.49
Alkalinity, mg/l	44 ± 20.11
COD, gm/l	396.82 ± 265.95
TS, %	23.16 ± 4.01
VS, % of TS.	97.64 ± 3.19
Moisture content %	76.85 ± 4.01
Carbon %	38.68 ± 1.34
Nitrogen %	1.5 ± 0.26
C/N	24.7 ± 4.83
Phosphorus %	0.3255 ± 0.006
Potassium %	0.225 ± 0.035

Objectives

To understand the generation and collection practices, of organic kitchen waste, in households of Manipal.

Methodology

This study was aimed to find out the kitchen waste generation and collection practices of households. The study was limited to the first two steps in the waste management cycle i.e. kitchen waste generation and collection. The study is limited to knowing the nature and type of generation and collection of only organic kitchen waste.

The study was limited to Manipal in the Udupi district of Karnataka. According to the Census of India 2011 Udupi district has a population of 2,35,000 households. Manipal is a semi urban township in Udupi district and is about 6 km in diameter and has a population of about 26000 households. Manipal has both municipality and panchayat areas. Most of the households are in the middle and upper class category and it is on this household sample, that the study was carried out.

Sampling

The sample of households selected was limited to 100 in number, as there was a lot of homogeneity in characteristics of most of the households in Manipal. The population was divided into 3 strata 1. Households with an income, Rs. 1.0 lakh and above per month 2. income of households between Rs. 50,000 and Rs. 1.0 lakh and 3. income of households below Rs. 50,000 per month.

Stratified proportionate random sampling was used for sampling. The sample size consisted of 20 households from the first, 60 from the second and 20 from the third strata. But care was taken to choose households in the municipality and panchayat areas and also from vegetarian and non-vegetarian households, graduates and non- graduates, as well as those who own a house or living in rented house.

Table 3: Characteristics of the sample studied

Strata of households	Vegetarian %	Non Vegetarian %	Graduate and above %	Below graduation %	Own house %	Rented house %
Rs. 1.0 lakh and above	45	55	75	25	62	38
between Rs. 50,000 and Rs. 1.0 lakh	40	60	65	35	52	48
Below Rs. 50,000 per month.	55	45	50	50	44	56

Data collection and analysis

Data was collected through the interview method, using a structured questionnaire. A 7-point Likert scale was used to measure satisfaction of households towards waste collection.

Data was quantitatively as well as qualitatively analyzed. Mean, percentages and rating scores were calculated.

Findings

Every household was concerned about the proper management of kitchen waste. At an average, about 820 gms of waste per day is generated from each house.

Composition of kitchen waste

House hold waste consisted of kitchen cooking waste, plastic, paper, glass, metallic items, clothes, house cleaning wastes etc (Table-4).

Table 4: Composition of Household waste.

Constituents	% of composition
Kitchen cooking waste	36
Plastic	20
Paper	15
Glass,	6
Metallic items	8
Clothes	6
House cleaning wastes	3
Misc.	6

The most frequently generated wastes are kitchen wastes, plastic and paper. A major component of plastic is milk cover and the plastic carry bags from shops. About 20% of the households separately keep milk plastic covers to be sold to scrap merchants' over a period of time. Paper wastes are the result of writing paper and other paper stationery items used by students and elders in the house. Newspapers are generally separately kept and sold as scrap to scrap merchants by most houses. Kitchen waste which is the major constituent of house hold waste is generated and collected almost daily.

Composition of Organic kitchen waste

The major constituent of Kitchen waste are those associated with cooking. Most households prepare food two times a day; once in the morning and once in the evening. The cooking in the morning is usually more elaborate. Breakfast cooking includes preparing dosa, idly, poori, chappathi, upma, sandwich etc. Lunch is both vegetarian and non- vegetarian. Since the study was done in a coastal region, fish is a more popular non vegetarian item cooked and consumed.

Kitchen waste consisted of vegetable peels, spoiled fruits, food remains after consumption, spoiled food and other eatable items, meat waste, fish waste etc.(Table- 5). Vegetable peels, cereal remains, cooked food remains and spoilt food are the most regularly generated kitchen waste.

Table -5: Constituents of organic kitchen waste

Constituents	% of composition
Vegetable peels & cereals	25
Spoiled fruits	7
Cooked food remains after consumption	28
Spoiled food and other eatable items	14
Meat waste	8
Fish waste	12
Others	6

Waste collection at first point

Most of the households keep all the wastes in a single dustbin. Only 10% do any kind of separation of the household wastes. A large % of the households keep the waste of the day, tied in plastic bag in the dust bin. 60% of the household keep these tied bags or dust bins in front of their house, to be taken by the waste collecting agency. The remaining put it in the public dustbin kept in the public place. In most of the cases, this is done in the mornings. Only a very few percent casually throw the waste in the nearby open area.

Usually the collection at the first point is done by the maid or the lady of the house. In most of the cases, transferring to the second collection point is done by the lady of the house. While a large percentage of households transfer waste to the second collection point only once in a day, about 20% of the households do it two times a day.

Waste collection at second point

Most of the household complained of the irregularity in picking waste from the second collection point. One of the biggest fall out of this, is pollution and other side effects in the local area. These solid wastes, when improperly and irregularly collected can be an environmental hazard. It can be a major cause of water pollution. If these wastes get into drinking water, it can cause diseases like cholera, dysentery etc. These wastes when eaten by domestic animals can also lead to their death due to diseases as well as choking. These wastes are good breeding ground for mosquitoes and flies. This improper dumping can pollute nearby water bodies and cause water pollutions leading to death of fish. When solid wastes are not removed properly, it can clog drainage channels and gutters and block the flow of the sewage. This may cause flooding and other difficulties. Irregular lifting of waste from the second point can produce bad odor in the locality, which pollutes the air quality of that place. It spoils the natural beauty of the place too.

Awareness of biodegradation and composting of kitchen waste

Bio degradation or composting is possible after the first collection point in the case of organic kitchen waste. About 80% of the households studied, know that the organic kitchen wastes are bio degradable. Nevertheless only 40% of the households showed interest in making compost out of the organic kitchen waste. It has not been done because of lack of specific knowledge of how to do it or lack of time or shortage of space. However, most of them said that, they would like to manure their plants using this compost manure.

Biodegradation is the chemical dissolution of materials by bacteria or other biological means[7]. Biodegradable matter is generally organic material such as plant and animal matter and other substances originating from living organisms, or artificial materials that are similar to plant and animal matter to be put to use by microorganisms. In nature, all the materials have the capability to be broken down into their raw material. The process of biodegradation may be different for different substances but in general, biodegradable substances will be decomposed into carbon dioxide, methane, and water as the final products.

Composting can be more amateurally done by either using a bin or using a pit in the garden[8]. If one, uses a bin, one can add a shovel or two of garden soil because this will contain the microorganisms that will break down the organic matter. If one uses a pit in the ground, those organisms are already present in the soil, to 'act' on the waste. The organic matter is continually added in layers. Weeds can also be a good component of the compost pile, along with food scraps and even fireplace ash and coffee filters added sparingly. Water can be added if the waste mixture is dry. It may take a couple of months to turn the waste into compost manure. Of the 60% of household who

were not keen on composting, some (20%) are willing to attempt, if sufficient knowledge on the process is imparted by an expert and if the process is not cumbersome.

Satisfaction towards waste collection

Households expressed overall dissatisfaction (Table-6) at the waste collection at the second point, by the waste collecting agency. But, most of them would like the waste collecting agency to regularly lift the waste from the first collection point as well as second collection point. Majority of the households (70%) are paying about Rs.30 per month for the waste disposal and most of them would not mind if they have to pay double the amount, if the collection is regularly done.

Table-6: *Satisfaction of households at the second waste collection point

Factors	Satisfaction score
Regularity of collection.	4.5
Maintenance of the dust bin	3.0
Cleanliness around the dust bin	3.5
Method of collection	5.5
Sensitivity of people who collect	5.0
Quality of mode of transport	6.0
Fee charged	7.0

*Score on a scale of 1-10. 1- being low satisfaction and 10- being high satisfaction.

Usage of technology in waste collection

Udupi does not use any technology in waste collection or disposal. But the adjacent district of Mangalore, decided to adopt a technology to monitor on-line whether its waste collecting contractors have cleared solid waste from bins and containers in the city daily. The contractors would be given away cell phones from the corporation which they should use for taking the photographs of the bins and containers daily. Immediately after clearing a bin the contractor should click its photograph and press an option in the pre-loaded software in the cell phone. The picture would get uploaded in a website (which could be linked to the city corporation's website) within 10 seconds with the date, time and location of the bin or container. On the other hand, the website would have a map of the city with the locations (which are called geo reference stations) of bins and containers and black spots (places minus bins and containers where people daily dump solid waste). A green icon would appear on the location where a bin or the container had been cleared. If not there would a red icon indicating that it had not been cleared. It would be mandatory for the contractor to upload the photograph of bins which had not been cleared daily. Photographs of such bins would be recorded in red icons. The households of Manipal are willing to pay extra for waste collection, if such technology is implemented in Udupi district too, and if it can increase the effectiveness of waste collection.

Conclusion

The primary steps in waste management are: generation, collection, sorting and separation, transfer and transport, and disposal. The study was intended to understand the organic kitchen waste generation and collections practices and was limited to kitchen waste collection at the first and second collection points. The study was done on middle and upper class households in the Manipal area of the Udupi district of Karnataka.

Each household generated about 820 gms of waste per day. Out of that, the biggest constituent was organic kitchen waste followed by plastic and paper. These three are the most regularly generated wastes too. In kitchen waste, the biggest contributor is cooked food remains after consumption, followed by waste of vegetable peels and cereals followed by spoiled food. Fish waste came fourth. Cooked food waste, vegetable waste and spoiled food are the three most regularly disposed kitchen waste items. Surprisingly, very few made efforts for separation of waste items at the source into either biodegradable or non-biodegradable or even those which can be recycled.

The second collection point, for many households. Since it is not done regularly and scientifically, it has led to dissatisfaction of the local area residents. Another negative impact which the households perceive of this is the problems due to pollution. Most of the households felt that the public dustbins (2nd collection point) as well as the area around it, is maintained poorly making the place highly unhygienic. A positive aspect was that a sizable number of households showed interest in trying out further processing of the organic waste .

References:

1. Schultz, George P. Facility Planning for a Public Service System: Domestic Solid Waste Collection. *Journal of Regional Science*. 9 (1969) 1-18.
2. Buck, Karen, and Ruth B. McKay. Removing Maximum Kitchen Organics from the Waste Stream: A Case Study. *J. Environmental Systems* 30 (2003-2004)1-18.
3. Waste Data Overview, June 2011: Department of Environment, Food and Rural Affairs, UK. Household Waste Per Capita. Pages. 38-39.
4. Sarkhel, Prasenjit. Economics of Household Waste Mangement in Kolkata: Proposed Steps Towards Improved Efficiency. *Contemporary Issues and Ideas in Social Sciences*. (March 2006) 1-20.
5. SID5 Research project Final Report, Department of Environment, Food and Rural Affairs, UK.
6. Mhaisalkar V. A. and A. R. Tembhurkar. Studies on Hydrolysis and Acidogenesis of Kitchen Waste in Two Phase Anaerobic Digestion. *Journal of the IPHE*. (2007-2008) 10-18.
7. Premakumara, D.G.J. Survey of Household Solid Waste Generation and Public Awareness on Waste Separation and Compositing Practices in Cebu City.(2011) 1-27.
8. Al-Maliky, Salam J. Bash, and Zainab Qahtan ElKhayat. Kitchen Food Waste Inventory for Residential Areas in Baghdad City. *Modern Applied Science*. 6 (2012)1-8.

ECOLOGICAL ROLE OF ANIMAL DIVERSITY IN SOIL SYSTEM (A CASE STUDY AT EL- RAWAKEEB DRY LAND RESEARCH STATION, SUDAN)

Maha Ali Abdel Latif, Assistant Prof. Researcher

Desert Research Institute, National Centre for Research, Khartoum Sudan

Abstract:

Interest in soil biodiversity characterization and its function in soil ecosystem have come from the need to develop this ecosystem and manipulate soil biota to encourage a greater reliance on ecosystem self-regulation. Observations on the impacts of agricultural managements on populations and communities of soil fauna and their interactions confirm that high input, intensively managed systems tend to promote low diversity while lower input systems conserve diversity. Animal diversity contributes to soil ecology and functioning such as breakdown and cycling of nutrients in the soil. These animals are essential to the proper function of the soil ecosystem in different environments. In arid environment, the role of soil Collembola, mites and nematodes in decomposing plant litter was assessed and compared in El-Rawakeeb Dry land Research Station, (latitudes 15° - 2° and 15° - 36° N and longitudes 32° - 0° and 32° - 10° E) as affected by application of organic manure (bovine manure) or pesticide (neem leaf powder) over one year. Thirty litter bags each one contain 10 gm oven dried litter of *Cajanus, cajan* were buried into the treated and control plots. Six bags per plot along with their respective soil were retrieved four times the year. Mean mass loss in litter was correlated to population density of Collembola, mites and nematodes. Some soil properties eg. temperature, moisture, particle size distribution and organic matter contents were correlated to litter mass loss. Results showed that manuring enhanced role of decomposers in decomposing plant litter, while neem application mostly retarded it.

Key Words: Collembola, mites, nematodes organic manure, neem leaves powder, soil ecosystem

Introduction

Litter decomposition is a biological process by which plant litter is broken down and nutrients held in its combination are released (Deyer *et al*, 1990). It begins usually with large soil animals such as earthworms, arthropods and gastropode. These organisms breakdown plant litter into smaller fragments, which are further decomposed by bacteria and fungi (Pidwiry, 2001). Such decomposition takes several months to several years to complete depending on climatic and edaphic factors. Additionally, litter inherent decomposability are factors that can affect decomposition (Couleausae *et al*, 1995).

Animal decomposers are various groups of animals that feed on waste products produced within a food web. They contribute to the process of litter decomposition and to the maintenance of soil fertility. They also impose regulatory effect on decomposition through their trophic interactions with microorganisms and thus can enhance the primary production, (Setala and Huhta, 1991). Coleman and Crossaly (1996) illustrated that decomposers animals include protozoans, earthworms, micro arthropods such as mites and Collembolan and macro arthropods such as insects, arachnids, millipedes and centipedes. Seastedt (1984), illustrated that the presence of free living mites increases litter mass loss and primary productivity. These observations agreed to the recent findings of Khalil *et al* (1999) who added that soil mites could be used as bio indicators for ecotoxicological responses and environmental monitoring.

Nematodes are the most abundant soil decomposers animals. They are found in surprising numbers and usually classified according to their feeding habits, (Brady, 1990). Moreover, they represent major reserves of nutrients that upon their death returned to the cycle and used as food for

other invertebrates. Nematodes affect decomposition directly as microbiovore feeding on primary decomposers. They can also affect decomposition as phytophages since they feed on the living protoplasm and plant which may facilitate their decay, (Yastes, 1979).

Soil decomposers animals are very important functional groups. They gain this importance through their contribution to different biological processes especially to organic matter decomposition and nutrients cycling. There is, therefore, a reason concerning the study of the potential role of decomposers in litter decomposition. The present study involves evaluation of the role of soil Collembola, mites and nematodes in litter decomposition in relation to manuring and application of neem leaf powder. It is also meant to clarify the impact of these animals in enhancing plant growth and productivity.

Materials and Methods

Study Site

Field experiments were conducted at El-Rawakeeb Research Station which lies in a El-Rawakeeb dry land which occupies the area southwestern Omdurman Governorate. It is located between latitudes 15° - 2° and 15°- 36° N and longitudes 32° -0° and 32° - 10° E. It thus lies within the tropical semi – arid region of the Sudan and its climate is characterized by a short rainy season that extends from July to October with peak in August as shown in Fig. (1).

According to El-Hag *et, al.* (1994), the average rainfall was 100 – 180 mm, and the evaporation potential was 1800 mm thus the relative humidity is low. The summer season usually, extends for a long period with a maximum temperature of 43 C° during May (Fig.2). The soil temperature was lower than the ambient temperature throughout the year. The geological formation of the area is mainly basement complex overlain by superficial deposits of the Nubian sand stone.

The soil is generally characterized by sandy texture, poor organic nitrogen and carbon, moderate bicarbonate and potassium and high sodium, calcium and chloride contents, El-Hag *et, al.* (1994).

The natural vegetation of El-Rawakeeb area is composed mainly of *Acacia sp.* (e.g. *Acacia tortilis*) and different grasses (e.g. *Aristida sp.*). This natural vegetation is replaced in the cultivated sites by some cereals (e.g. *sorghum bicolor*) and few legumes (e.g. *Cajanus cajan*). The system of land use is mainly pastoral except in low land where traditional agriculture is practiced. Within El-Rawakeeb Research Station, irrigated fodder, vegetables and cereal crops are grown.

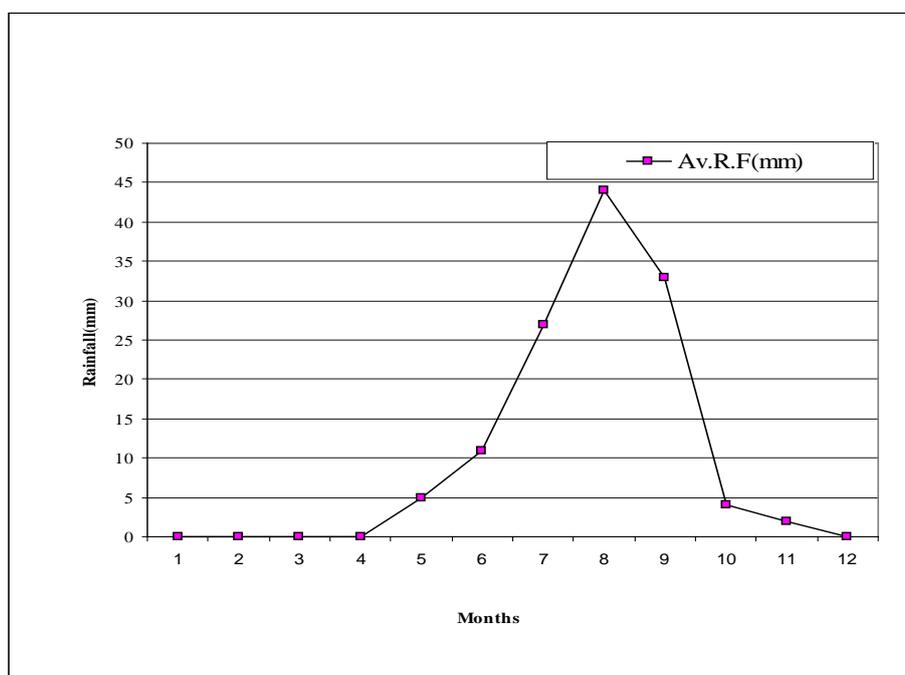


Fig.(1): Mean monthly rainfall in Khartoum state during the study period

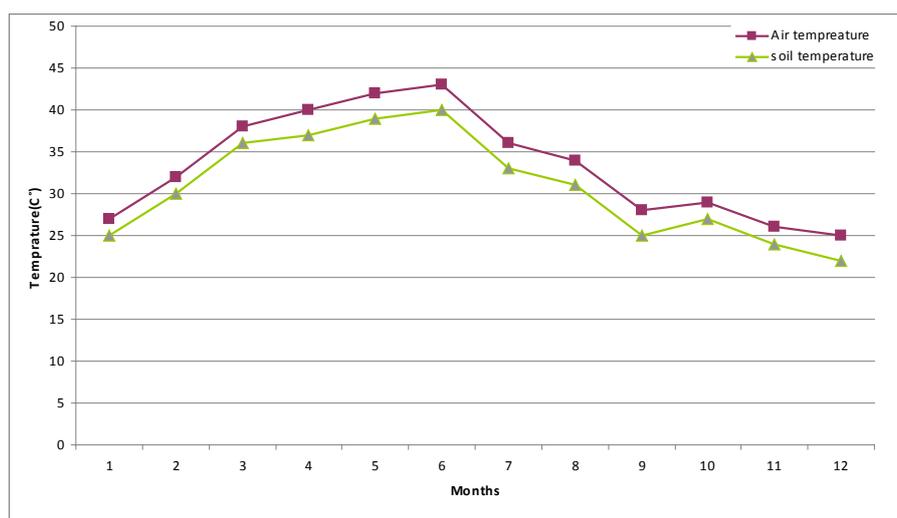


Fig.(2): Mean monthly air and soil temperature as recorded - El -Rawakeeb dry soil during the study period

Preparation and application of treatments

Fertilizers

Organic manure in the form of cattle manure was used as fertilizer. Fresh manure was collected from Shambat Dairy Production Unit, University of Sudan. It was then subjected to sunrays for 6 hrs to eliminate any possible faunal content and used as organic manure.

Pesticides

Leaves of neem tree (*Azadirachta indica* A. juss) were manually collected, air-dried and ground using a pestle and a mortar. Neem leaves powder was then used as pesticide.

Preparation of litterbags

Litter bags were prepared following Cepeda-Pizarro *et al.* (1992). Fresh leaves materials were collected from individuals of *Cajanus cajan* grown in the study area. Immediately after collection, the leaves were dried in sunlight for a week and then at 50 °C in an oven for 24 hrs. When dry, leaves were fragmented into large fragments using a mortar. Ten gm leaves' fragments were weighed and placed in bags made of cloth mosquito netting, of (10 X 5 cm) size.

Experimental layout

Litter bag experiment was conducted at El-Rawakeeb Dryland Research Station Prior to experimentation the study area was divided into 45 plots each plot is of 3 X 2 m in area in a complete randomized block design. The 45 plots were then grouped into three groups each group is of 15 plots, which were then kept as permanent sampling plots. The whole plots were cultivated with *Cajanus cajan* legume which is commonly known as pigeon peas in English or lubya Sudani in Arabic languages. According to the method described by Hafiz (1998), 12-kg of cattle manure were weighed and added to the first 15 plots. These were used as manure treatment plots.

Neem leaves powder was added to the soil as one gm/plant hole in the second 15 plots according to Siddig (Pers Comm). These plots were used as neem treatment plots. The rest 15 plots were kept as control.

A total of 90 litterbags were buried to a depth of 10 cm of each sampling plots. Sampling took place every 120-day. The sampling periods were winter, autumn and summer. Nine bags along with their respective soil were randomly retrieved from the field on each sampling event. The retrieved bags were reweighed and their mass loss was obtained. Soil samples were analysed for their some of physicochemical properties. Following methods described by the International Centre for Agricultural Research in Dry Areas (ICARDA).

Extraction of animal decomposers

Nine soil samples were weekly taken randomly from each sampling plot. They were then introduced into polyethylene bags and taken to laboratory for decomposers extraction as follows:

Extraction of nematodes

Nematodes were extracted from soils of the study using Baermann trays technique as described by Freckman and Baldwin (1986). A soil sample was spread evenly into a thin layer over a

sieve lined with tissue paper. The sieve containing the sample was placed in plastic tray. Sustainable amount of water was added to saturate the soil sample. The nematodes became active, swam out of the soil, and sank on the bottom of the tray. The nematodes were collected in water and preserved in 75 % ethanol for further identification.

Extraction of mites and Collembola

Mites and Collembola were extracted using Tullgren funnel apparatus as described by Griffith (1996). This method is based on taking a soil sample and applying a physical or a chemical stimulus to drive the animal into a collecting fluid. Tullgren funnel apparatus consists of a metal funnel measuring 25 cm in diameter. A fine mesh screen of 0.2 mm size was fitted to the neck of the funnel to retain the soil sample. Since dividing 1 kg soil sample into sub samples produced no difference in the encountered fauna, the funnel was loaded with the whole 1 kg soil. The funnel was then fitted to a beaker containing 70 % ethanol as a preservative medium. The soil sample was subjected to the direct illumination of a 60-watt light bulb placed 10 cm above soil surface for 24 hrs. The contents of the collecting beaker were transferred into a Petri dish and then examined under a binocular microscope. The fauna collected were further preserved in 75% ethanol for further identification.

Statistical analysis

Litter decomposition in EL-Rawakeeb treated soil were analyzed as affected by the application of cattle manure and Neem leaves powder using the Statistical package for social sciences (SPSS), (1997) under Windows.

Results and Discussion

Soil Properties

Effects of manuring and application of neem leaf powder were evaluated through the study period. Results shown in table (1) explained that manuring increased water content and decreased temperature value. It also decreased soil contents of silt and sand but increased clay content as compared to the control. Manuring decreased pH and increased organic C and N contents. These results could be ascribed to the fibrous content of cattle manure which may add to its cementing effect of manure to soil particles. These effects may improve soil texture and increase its ability to retain water content by improving its water holding capacity. Also, manuring may alter cations and anions contents and improve its organic matter content. These findings were in conformity with Hafiz (1994) who studied the impacts of cattle manure application on soil properties. He concluded that cattle manure has a significant positive effect on soil bulk density texture and water content:

Application of neem leaf powder to El-Rawakeeb soil induced varying impacts on its physicochemical properties as given in table (1). Physically, neem application increased water and clay contents but decreased temperature, sand and silt contents. Chemically, it increased soil organic contents and C and N while induced insignificant change on soil pH. These results might be due to the amendment effects and neem beside its biocidal effects Radwanski and Wickens (1981) evaluated the impacts of neem products on soil properties. They concluded that neem products added significant amounts of organic matter, improved soil texture and increased its cations contents.

Table (1): Physicochemical properties of El-Rawakeeb soil as affected by application of cattle manure or neem leaf powder

Soil properties	Treatments		
	Cattle manure	Neem leaf Powder	Control
Water content	0.251a	0.158b	0.107c
Temperature	30.40a	31.04a	34.68b
Sand	68.30a	70.20a	72.70b
Silt	17.60a	18.30a	21.20b
Clay	14.10a	11.50b	6.10c
pH	7.17a	7.05b	7.06b
Organic N	0.10a	0.08a	0.02b
Organic C	1.54a	1.45a	0.013b

Means with same letter do not significantly differ, otherwise they do according to Duncan's multiple range test.

Litter decomposition in treated soil

Litter decomposition was measured as mean litter molasses over four sampling periods in response to manuring or neem application. Results shown in table (2) indicated that mean litter mass

loss varied between treatments. Manuring increased litter mass loss during the first, second and fourth sampling periods but induced insignificant change in litter mass loss for the third sampling period as compared to the control. Increase in mean litter mass loss could be due to the possible effects of manuring on soil properties which in turn may enhance litter breakdown. Tester (1990) observed that manuring improved soil properties and altered litter decomposition consequently.

Application of neem leaf powder to El-Rawakeeb soil induced significant decreases in mean litter mass loss during the first, second and fourth sampling periods as given in table (2). These results could ascribe to neem toxicity which may reduce soil biological activity and reducing litter decomposition accordingly. Similar results were given by Parker *et al* (1984).

Maunring or neem application induced insignificant change in litter decomposition during the third sampling period as compared to the control (table 2). These results might be due to the effect of the rainy season which may mask the effects of treatments. Also, the rainy season flavoured decomposers to flourish and enhance the rate of litter mass loss. Christensen (1985) studied litter decomposition under field conditions. He observed that the percentage of litter mass loss was significantly higher during the rainy season than the other seasons of the experimentation year. Moreover, Reddy *et al* (1989) added that seasonal fluctuation of decomposers greatly affected litter decomposition.

Table (2): Litter decomposition in treated soil as recorded for four sampling periods.

Sampling periods	SS	MS	df	F-Values	
				F-calculated	F-tabulated
Period I	3.28	1.64	2	4.16*	3.17
Period II	21.59	10.79	2	3.59*	3.17
Period III	8.01	4.00	2	0.92Ns	3.17
Period IV	200.94	100.47	2	228.11**	3.17

* Significantly difference at P = 0.05 ** Significantly difference P = 0.01 Ns = insignificant difference.

Impact of decomposers animals on litter decomposition

Impact of Collembola, mites or nematodes on litter decomposition was assessed in El-Rawakeeb soil treated with cattle manure or neem leaf powder. Results shown in table (3) indicated that the three faunal groups were significantly positively correlated to mean litter mass loss upon manuring. These results could be ascribed to the substantial effect of manuring on decomposers animals which may in turn intensify their decomposing activity. Wenner and Dindal (1990) evaluated the impact of manuring on the functional role decomposers fauna. They found that manuring significantly increased population densities of fauna and intensified their ecological role in soil processes. Chiagnon *et al* (2000) noticed that decomposers animals responded to manuring interms of population density and functional role.

Table (3) showed that only Collembola correlated positively to litter mass loss upon neem application than the other two groups (mites or nematodes). This might be due to the ability of Collembola to resist neem toxicity by increasing its population density.

Moriarty (1978) explained that Collembola showed adaptive response to biocides application by increasing individual chances of survival and reproduction rate.

The insignificant correlation to litter of mites or nematodes to litter mean mass loss upon neem application could be attributed to neem toxicity which may retard ecological role of these animals. Wright and Coleman (1988) revealed that neem application reduced mites population density and reduced their biological activities as decomposers. Similarly, Brady (1990) observed negative effects of neem products on nematodes under normal conditions (control soil), mean population density of each of Collembola, mites and nematodes exhibited significantly positive correlation to litter mass loss as shown in table (3). This could be due to the direct effect of these animals as decomposers. Lussenhop (1980) indicated soil arthropods affected litter mass loss significantly due to their small sizes and great abundance in soil ecosystem. Paker *et al* (1986) observed positive correlation between arthropods or nematodes and litter mass loss under normal (untreated) soil.

Table (3): Correlation between decomposers animals and litter mean loss in El-Rawakeeb soil treated with cattle manure or neem leaf powder.

Taxon	Treatments		
	Cattle manure	Neem leaf powder	Control
Collembola	0.904**	0.831*	0.821
Mites	0.884**	0.693	0.810*
Nemastodes	0.814**	0.502	0.809*

* Significant at P = 0.05 ** Significant at P = 0.01

References:

- Brady, N.C (1990). *The Nature and Properties of Soil*. 8th edition. Published by Macmillan publishing company, New York.
- Cepeda Pizarro, Maldonado, M. B., Uitches, J. L., Rojas M. A. and Perira N. P. (1992): A litter bagstudy of mite density (Actinedida and Orlatida) in a triplex litter and soil of a desert site in North Chifle. *Journal of Arid Environment* (23): 177-188.
- Chagnon, M; Hebert C and Pare D. (2000): Community Structure of collembolan in Suger maple forests: relations to hummabs type and seasonal trends.
- Christensen, B.T. (1985). Wheat and barley straw decomposition under field conditions: Effect of soil type and plant cover on weight loss, nitrogen and Potassium contents. *Soil Biology and Biochemistry*. 17 (5): 691 – 697.
- Coleman, D.C and Crossley, D.A (1996) *Fundamentals of soil ecology*. Academic, San Diego.
- Couleausae, M. M., Bottner, P. and Berg, B. (1995). Litter decomposition, climate and litter quality. *Trends in Ecology and Evolution*, 10: 63 – 66.
- Dyer M., Meentemeyer, V. and Berg, B. (1990). Apparent controls of mass loss rates of leaf litter on a regional scale. Litter quality vs. climate. *Scandinavian Journal of Forest Research*. 5: 311-323.
- Hafiz T.F. (1998): Responses of vegetative growth, early flowering and yield of two lomato cultivars to form yard manure and mineral fertilizer application. *Jour of Arab Agri. Research Org*. V. 4 (1): 143-155.
- Hafiz, A. A. R. (1994): Comparative changes in soil physical properties by mixtures of manure from various domestic animals. *Soil Sci*. (118): 53-59.
- ICARDA (1996): A soil and plant analysis manual adopted for the West Asia and North Africa Region, Published by ICARDA, Aleppo, Syria.
- Khalil, M. A., Abdel-latif H. M. and AL-Assiuty, A.I. (1999): Changes in oribatid faunal structure associated with hand conversion from annual crop into orchard, In: *pedobiologia* 43, 85-96 (1999). Jrbn and Fischer berlag.
- Lussenhop J. (1980): Microbial and microarthropod – detrital processing in prairie soil. In: *Ecology*, (26) 4, pp. 964 – 972. Published by the ecological society of America.
- Moriarty, F. (1978): Terrestrial animals. In: G.C. Buller (ed) *Principles of Ecotoxicology*, SCOPE, 12. Chichester John. Wiley and Sons: 769-786.
- Parker, L., Elkins, N., Alden E. and Whitford, W. (1986). Decomposition and soil biota after reclamation of coal mine spoils in an arid region. *Biology and Fertility of Soils*. 4: 129 – 135.
- Pidwirny M. J. (2001): *Fundamental to physical geography*, Mpidwirny @ Okangan VE. Ca.
- Radwanski S.A. and Wickens G.E. (1981): Vegetative fallows and potential value of the neem tree (*Azadirachta indica* in the tropicas. *Econ. Bot.* 35(4): 398-414.
- Reddy MV, Reddy VR., Kumar, VPK, Yule, DF and Herman, M. (1992): Soil management and seasonal community structure of soil Microarthropods in semiarid tropical Alfisols.
- Seasted, TR (1984): The role of microarthropods in decomposition and mineralization processes. *Ann. Rev. Entamol.* 29: 25 – 46.
- Setala, H. and Huhta V. (1991) Soil fauna increase *Betula pendula* growth:. Laboratory experiments with coniferous forest floor. *Ecology*. 72: 665 – 671.
- SPSS (1997). *Data handling and biostatistics: use of SPSS for windows*. Published by the Danish Bilharziasis Laboratory.
- Tester, C.F. (1990). Organic amendments effects on physical and chemical properties of a sandy soil. *Soil Science Society of America Journal*. 54 (3): 827 – 831.
- Wenner, M.R. and Dindal, D.L. (1990): Effects of conversion to organic agricultural practices on soil biota. *American Journal of Alternative Agriculture* 5(1): 24-32.
- Yeastes, G.W (1979). Soil nematodes in terrestrial ecosystems. *Journal of Nematology*. 11: 113 – 229.

STATE OF THE ENVIRONMENT IN THE NIGER DELTA AREA OF ONDO STATE

F.A. Olorunlana

Department Of Geography And Planning Sciences, Adekunle Ajasin Univeristy, Akungba-Akoko

Abstract:

The geomorphological agents at work in the coastal zone are winds, the astronomically-generated tides, the wind-generated waves, and various forms of current flow. These generate specific geomorphic processes and interact to produce an energy input which shapes and modifies the coast. Through feedback mechanisms, the developing morphology influences geomorphological processes, thereby becoming a factor influencing subsequent coastal evolution. Man, through his activities, is also a variable in this geomorphological equation.

Within the last three decades, accelerated coastal erosion, sea incursion and flooding have surfaced and reached alarming dimensions in some other parts of the niger Delta, notably along the transgressive mud coast in Ondo state. This paper examines the state of the environment in the Western Niger Delta area of Ondo State. It identifies the environmental problems in the area and also at the environmental and socio-economic impacts of these problems. Possible and practicable recommendations were proffered in an attempt to reduce or check the trend and scenario of these problems in the area.

Key Words: Environment, Niger delta, Ondo State

Introduction

Studies on coastal landforms, processes and change are becoming more and more relevant as human intervention (at scales ranging from sandmining on a single beach to global warming) increases (Viles, 1990). It is very apparent that coastal zone is of great importance to every maritime nation of the world. Data derived from United Nations Population Division (1993) revealed that many of the leading megacities in the world are located on the coast. Consequently, the coastal zone of most countries of the world represents areas of many activities; for example communication, food production, settlement, recreation, mining, harbor making and so on. In addition to human activities or moderation on the coast, the coast itself, is a zone of intense energy input naturally, this energy transported by waves arrives at the coast and is available for work. Therefore, the coast represents a zone of complex and dynamic process naturally which shape and re-shape the configuration of the zone continuously. Geomorphologically, the zone can best be referred to as the most restless part of the continent.

Obvious advantages of the coastal zone to maritime nations make it expedient that deliberate, systematic studies of the areas be carried out for the purpose of proper coastal management and planning. Failure to study and manage each coastline properly with regard to ongoing natural and human activities in this area poses a great hazard and threat of lives and investment in the future. The entire coastal zone of Nigeria and elsewhere represent a challenging environment for applied geomorphological research where a continuum of relationship is observable.

Coastal Erosion

Coastal geomorphologists are becoming increasingly concerned with applied studies for a variety of reasons, including economic considerations (Viles, 1991). Dees and Davis (2001) described erosion as the means of wearing away of a shore of a body of water covered by sand, gravel or larger rock fragments by an instance or product of erosive action. The problem of coastal erosion and coastal stabilization is attracting greater attention from recent studies. Coastal erosion is not confined to Nigeria or West Africa coastline alone. It is a global phenomenon creating problems in virtually all the continents, such as United States of America, South America, France, Japan, Australia

and Africa. The coastal characteristics vary a lot from one coast to another on a regional basis, even along the same country like Nigeria, but the fundamental processes of wave's tides, and current actions at the coast remain the same everywhere.

Studies by Davies (1986) have shown that in the United States of America, the coastline is eroding at an average of more than one metre per year on the Pacific West coast and exceeding thirty metres per year in parts of Louisiana's coast on the Atlantic side. The magnitude of coastal erosion and coastal recession varies from place to place over time. Erosional rates along the Spain is said to vary from thirteen metres to more than thirty metres per year according to measurement. In Japan, problem of coastal erosion is so acute along the coastline that presently Japan has witnessed perhaps the world's most concentrated to arrest coastline erosion (Ibe, 1990; Nordstorm, 1994).

In Miami Beach, for example, despite the provision of a great number of steel groynes and vertical seawall, for millions of dollars, erosion still continues. The same story is at Florida beach, after the dredging of the inlet, the jetties that were built in 1918-1925 had blocked the southward littoral drift almost completely. The consequence was heavy erosion on the southern side of the inlet. Over many years now, attempts were made to combat erosion by construction of a great number of groynes, but this has not been very successful (Steers, 1971).

Coastal erosion is of course not limited to lowlying coast like in a stretch of depositional coast such as forms a larger part of Nigerian coastline. It also affects rocky coasts with cliffs and wavecut platforms (Faniran, 1986). For example, in the studies by Valentine (1971), along the Romans coastline of the Northsea between Flamborough head and Kilmsea, the low hills of Holderness reach the coast and form cliffs. As reported by Valentine (1971), along 61.5 kilometre of the Holderness coast, 307 points were selected at which to measure the retreat of the land between 1852 and 1952. on average, the cliffs receded by 120 metres in that period, approximately 720 hectares of land were lost and hence about 100 million cubic metres of materials were carried away. It was noticed that erosion was most severe at the south-east than north. Also, the most general cause of erosion was attributed to the energy of the sea. For future coastal protection, it was suggested that severest points of erosions should be considered first.

The Nigerian coast is richly endowed with a variety of natural resources, especially forests and mangroves, fisheries, touristic resources and mineral resources such as crude oil and gas which account for over 90% of Nigeria's foreign exchange earnings. Geomorphologically, the Nigerian coastline can be divided into four main physiographic zones: the Barrier-Lagoon complex, the Transgressive Mud beach or Mahin Mud coast, the Niger Delta and the Strand Coast. These different sections of the Nigerian coastline are associated with differing erosive activities; a result of combination of natural and anthropogenic factors.

The Niger delta geomorphic region in its entirety that is including the flanks is extensive, composite and multifaceted. The erosion rates in this section of the Nigerian coastline varies, ranging from between 18-24 m annually at Ugborodo/Esravos station, 20-22m annually at Forcados station, 16-19m annually at Brass station, 15-20m annually at Kulama station and 20-24m annually at Bonny station and 10-14m annually at Opobo (Ibe, 1988). The strand coast is about 85km long lying between Imo River and the Nigerian border with Cameroon in the east. The Strand coast is fronted by flat beaches which change into a beach ridge plain behind with a few small swamp systems. An average of between 10-13m erosion rates were recorded at Ibeno-Eket station (Ibe, 1988).

The perception of the deleterious effects of human activities on the coastal environment system in Nigeria dates back to 1910 when the construction of moles to halt the movement of sand into Lagos Harbour began to induce progressive accretion of Lighthouse Beach behind the west mole and the retrogradation of Victoria Beach east of the mole (Webb, 1960; Uzoroh, 1971). Similar changes have also taken place at Ogborodo (Esravos estuary) as a result of the construction of moles which interfere with the natural movement of sand along the coast and have altered, therefore, the age long pattern of erosion and deposition in the area. These man- induced accelerated processes are still very active in these coastal areas.

Within the last three decades, accelerated coastal erosion, sea incursion and flooding have surfaced and reached alarming dimensions in some other parts of the Niger Delta, notably along the transgressive mud coast in Ondo State; Patani, Bomadi, Burutu and Ogulaha along Forcados River in Delta State; Oron, James Town and Mkpong Utong along Cross River; Utaiwa along Jaja creek; and

Brass and Ibeno-Eket in Akwa Ibom State (CREMS, 2000). In fact, Asangwe (1993) identified more than fifty erosional sites along the national coastline.

Coastal erosion is a normal geomorphological process worldwide. However, the acceleration of coastal erosion is largely attributed to direct and indirect interference by man with the normal hydrodynamics along the coast. With the exception of Victoria Beach, Awoye area and Ogorodo (Escravos), the causes of accelerated erosion along the Nigerian coast are still unknown (CREMS, 2000). The alarming ecological and socio-economic impacts of these processes, however, have prompted individual environmental scientists, research institutions and governmental agencies to proffer and implement "lasting" solutions to these problems. Invariably, these solutions are not products of detailed long-term studies of the hydrodynamics of coastal processes, accelerated erosion and flooding in the affected areas. Also importantly, environmental issues were not taken into consideration in the decision and implementation of the coastal protection works. Consequently, many have failed woefully to solve the problems. In fact, in most cases, they have further undermined the stability of the coastal environment. All these underscore the need for research and database for coastline protection, coastal zone planning and management. This is particularly desirable in respect of the Ondo State coast owing to its uniqueness which makes it unwise to transfer to the area solution which has worked in other but dissimilar areas.

The Study Area

The area lies east of the West African lagoon system and on the western or Benin Flank of the Niger Delta Basin. The numerous lagoons in the coastal area have all disappeared due to silting and spread of aquatic weeds; the only survivor is the Mahin Lagoon. All the creeks and rivers in the area drain into Benin River. It is only here that the sandy beach materials of the Nigerian coastal plain are replaced by mud and lacustrine deposits. The general stratigraphy of Ondo State coast, therefore, differs from that of the Niger Delta by having about 60m mud overlying the Benin sands. The location of its eastern part in an unstable fault zone on the Benin Flank of the Niger Delta has important implications for geomorphological processes in this environmentally sensitive region. Unlike in other coastal states where wide wave beaten barrier beaches are backed inland by a sequence of equally extensive raised beaches and lagoons or mangrove swamps, the coastal plain in Ondo State is not only narrow (less than 700 metres wide) but it is backed by 30km-60km wide freshwater marshes and freshwater swamps with an intricate network of interconnected creeks swamp and lacustrine marshes. The narrow coastal plain constitutes, therefore, a fragile buffer zone between the tranquil waters of the swamps and the menacing waves of the Atlantic Ocean. Extensive breaching of this narrow coastal plain will result in massive incursion of the sea into the inland swamps with serious implications for national security and economy.

Soil and vegetation follow closely the geomorphological pattern. The main types of vegetation are:

- i. Mangrove forest and coastal
- ii. Vegetation.
- iii. Freshwater swamp forest.
- iv. Savanna and stunted rain forests of the sand ridges and mainland margin.

The mangrove forest on the coastal plain has been extensively cleared by the people for use as fuelwood, especially around the settlements where they have been replaced by coarse salt-resistant grass. Striking feature of vegetation in the area is the dessication of about 10,000 hectares of freshwater swamp forest in the Awoye area. It is attributed to salinization induced by salt water intrusion.

The soils are hydromorphic. The older sand ridge complexes develop brown and orange sand soils while the more recent one near the coast bear light grey sand soils. Where there are depressions on the ridge complexes, the soils can be described as poorly drained light grey sands. The swamps flats are characterized by swampy 'organic' soils which in the major part consist of decomposed and partly decomposed matter, while areas affected by tides bear saline soils.

Coastline Recession

Coastline recession in the Western Niger Delta area of Ondo State is principally by wave attack on the unconsolidated sediments of the clay ridge since the coast is not under the influence of longshore currents and the nearshore littoral Guinea current is very weak. In the absence of the protection offered by mangroves, the energy of waves and tides is concentrated on the bare surface of

the clay ridge exposed to the direct impacts of waves. As the waves overrun the plain during high tide, a network of rills develops on the plain by the backwash of the waves. These rills are progressively deepened and widened by subsequent uprush and backwash currents to form deep U-shaped gullies with vertical heads. The bigger gullies have terraced walls, the terrace being on the same level as the sub-tidal platform. These gullies become avenues for penetration of waves into the coastal plain as the gully heads advance at a very rapid rate into the plain. The measured rates of gully head retreat vary from 5.7m to 15.8m per annum (Ebisemiju, 1987). Most of the gully heads are now less than 50m from the town. Some gullies develop tributaries. Opposing tributaries of adjoining gullies cut deeply into their common interfluvies and eventually meet. In this way, a block of the plain is detached from the 'mainland'. As this block is attacked from all sides by waves, its total destruction is very rapid and a sub-tidal platform is produced. The platform grows as other detached blocks are eroded. The landward limit of effective wave action is marked by a mud scarp less than 1m high. The retreat of this scarp face is through the collapse of chunks of clay which first accumulate at the base of the scarp until they are broken into smaller fractions and transported seaward by backwash or carried into the coastal plain during high tide. Thus during low tide, the backshore area is highly indented by gullies and the sub-tidal platform is dissected into blocks of varying sizes. During high tide, the platform is submerged. The sea inlets dug by fishermen also serve as zones of rapid coastal erosion.

During high tides and the infrequent storm surges, the low coastal plain is overtopped by waves and tides. These initiate a sheetwash process and an imperceptible but widespread removal of surface sediments and their transportation into the inland creeks and swamps. In this way, the coastal plain is progressively lowered. Surface lowering by sheetwash also takes place during rainstorms.

The accumulation of ocean water in depressions on the coastal plain aids both coastline recession and ground surface lowering. These flood pools have the effect of making the water content of the soils higher and thereby weakening the physio-chemical interparticle bonds, reducing the coherence of the unconsolidated soil, and rendering the flood pool area more vulnerable to sheetwash erosion by the energy uprush and backwash currents of the waves and tides. The depressions are thereby deepened and widened until adjacent ones coalesce and become incorporated into the growing sub-tidal platform and plays a major role, therefore, in its expansion and subsequent destruction of the coastal plain.

Probable Causes Of Coastline Recession

The Niger Delta has been growing seaward by the accumulation of sediments brought down by the River Niger itself. At the same time, erosion phenomenon has been very aggressive and persistent in some parts of the Delta. The Niger Delta has also been experiencing regional subsidence arising from isostatic response to loading of the Delta and its margins with sediments from the Niger/Benue drainage system (Allen, 1964, 1965; Allen and Wells, 1962; Burke, 1972; Mascle, et al, 1976). Ebisemiju (1985, 1987) postulated that this regional subsidence process had probably been accelerated by the withdrawal of oil and gas from the Agbada Formations which underlie the unconsolidated deposits of silt and clay materials in the area.

It is also significant that the local people link the incidence of accelerated coastal erosion in the area with the commencement in the 1970s of crude oil and gas exploration in the area. A particularly striking feature of the mud coast of the desiccation off the swamp forest. The central and lateral bounds of the stressed forest area closely match those of the zone of accelerated coastline recession. The area appears to be expanding outward, especially westward, from a central point some 5km inland from the coast.

Seismic investigations by oil prospecting companies also could have induced local subsidence. Seismic investigations have involved the detonation of explosive charges at or below the ocean floor in the nearshore zone. Local people, especially at Aiyetoro, claimed in a memorandum to the NNPC in 1981 that seismic investigations conducted in the area generated shock waves which caused extensive damages to the King's concrete palace and other structures in the town; this claim was refuted by the Corporation. It is well known that, however, that the force of such explosion could cause liquefaction of surficial sediments and consolidation of the subsurface sediments, resulting in a potential increase in sediment erodibility and water depth. The potential disturbance of the nearshore bottom induced by such activities is consistent with the conditions necessary to promote coastal erosion. Increased water depth within the nearshore zone would enable larger waves to penetrate

further inland than otherwise experienced. This would result in coastal erosion and inundation of inland forests by sea water.

The association of space and time of accelerated marine erosion and dessication of forests would imply that the transgression of the sea which induced these major ecological disasters were indeed triggered by a single geomorphic process, viz, localised accelerated of land subsidence. The high concentration of oil wells within the affected area further supports this theory since it defines a possible mechanism for the consolidation of subsurface host and overburden materials causing land subsidence. Seismic shaking of the southern parts of Nigeria in 1985 indicate that the continental shelf here may not be so stable after all. It is, however, not unlikely that the seismicity was triggered by subsidence associated with hydrocarbon extraction in Awoye area, as has been reported from other parts of the world (Yerkes and Castle, 1976).

Overtopping of the coastal plain by waves and tidal floods, also could have been further accelerated by rise in sea level in response to global warming. It has been predicted that temperature will rise by between 2^oC and 3^oC and sea level will rise by 0.5m to 1m by the year 2100. the resulting deepening of the continental shelf will result in increase in effective wave height and other changes in wave and tidal dynamics. These will further increase the rates of coastal erosion, flooding and salt water intrusion along low-lying coasts like the mud coast of Ondo State.

The Environmental And Socio-Economic Impacts Of Coastal Erosion

The coastline in Awoye has receded by about 3.31km between 1974 and 1996 at annual rates which varied from 31m to 19m. by 1981, coastline recession in Awoye area had caused the incorporation of 487 hectares of the coastline plain into the Atlantic Ocean within a period of eighteen years: 1973 to 1991 (Ebisemiju, 2001). By 1996, the cumulative losses had risen astronomically to 3415 hectares. During this period, the coastal plain was reduced in width by about 62% and breached in four places at Jirinwo, Odofado, Awoye and Ago Nati. This has resulted in the loss of scarce land for settlement and animal husbandry on the narrow coastal plain. At Aiyetoro, the coastline has receded by about 0.57km between 1974 and 1999 resulting in the incorporation of about 300 hectares of the coastal plain into the Atlantic Ocean. The present rate of coastline recession is 35m a year. At this rate, the coastal plain which is only about 50m wide here will be breached. Other areas experiencing very rapid rates of coastline recession are Ori-Oke Iwamimo to Ogororo and Abereke.

The coastal processes have also adversely affected the fishing economy of the people. Fishing in both the freshwater creeks, lagoons and swamps and in the nearshore zone of the Atlantic Ocean has always been the main source of revenue for the local people. The turbidity of ocean waters in the nearshore zone is high. The waters of the nearshore zone, therefore, constitute an unfavourable ecological environment for marine fishes which have consequently deserted the zone forcing fishermen to travel far into the sea. Unfortunately, they are unable to finance the additional investments which deep sea fishing entails.

The major direct impacts of the erosion menace, obvious even to the least discerning observer, is the large scale destruction of buildings, canals and mangrove forest especially in Awoye area. Evidence abounds both around the present coastline, within the inter-tidal platform and in the nearshore zone to indicate previous settlement sites; there are in the form of ironwood stakes of abandoned houses.

Accelerated coastline recession and tidal floods have forced people to abandon their houses and move inland to relatively safe locations. On the average, settlements are moved from one location to other every four years. Today, the original site of Awoye is in the Atlantic Ocean at a distance of 3km from the present shoreline. This semi-nomadic movement of the people does not allow erection of permanent structures but those that can be easily dismantled whenever they are being threatened by tidal floods and coastal erosion. This periodic relocation of settlements inland, therefore, is a significant economic social and cultural disruption to the people. The massive breaching of the coastal plain in Aiyetoro area will undoubtedly produce a catastrophic event that will wipe out all the settlements in this heartbeat of Ilaje coast, with the impacts felt throughout the freshwater swamps to the north.

The deleterious effects of accelerated coastal erosion and sedimentation on the environment, economy and social life of the communities in the Awoye/Imoluma area are pervasive and most of the people have resigned themselves to a situation beyond their control. Indeed, the continuing threat to life and property by coastal erosion and tidal floods, the cost of rebuilding houses every two to four

years, the hazards and high costs of deep sea fishing, the loss of the rich nearshore and freshwater fishing grounds and the lack of alternative gainful employment opportunities have forced the young and middle-aged men and women to emigrate to other riverine states in Nigeria and also to neighbouring West Africa countries, especially the Cameroon Republic, Gabon, Congo and Ghana where they engage primarily in fishing and remit part of their earnings for the upkeep of the older folks who remain behind.

Erosion Control Measures

Two categories of coastal erosion are distinguishable: long-term (chronic) and short-term (storm-induced) coastal erosion. Generally, chronic erosion is more difficult to address than storm-induced erosion (PACE, 2003). Generally, because of the dynamic nature of coastal areas, various methods for protecting the coastline from the consequences of coastal erosion problems exist; some structural and others non-structural. The structural measures have been described as 'hard' measures, which has to do with the constructions of various engineering devices, which are physical and concrete, meant to protect or reduce subsequent occurrences of a given hazard events. Some of the hard coastline defense works that have been used in managing coastal erosion and retreat include groins, jetties, breakwaters, sea-walls, wave-return walls and tetrapods (Charlier, et al., 1989a, Charlier, et al., 1989b). The non-structural responses include the use of satellite forecasting or telecommunications for issuing warnings (Anderson, 1994).

Sand replenishment, revetment and other structures such as groins and jetties which are structural measure and designed to retard littoral drift of materials are not relevant to the resolution of the coastal erosion crisis in the area since littoral drift is virtually non-existent along the entire length of the mud coast. Rather, the searchlight should be focused on those structures which prevent incoming waves from breaking close to the coastline, and at the same time dampen their ferocity. The control works capable of achieving these are sea walls (dykes) and breakwaters.

A widely canvassed engineering solution to coastal erosion in the area is the construction of a dyke to follow the edge of the high waterline, and high enough to prevent over-topping by tides waves, with the seaward toe protected with some form of armour layer, and the construction of revetment on this side to reduce the action of waves and currents on the dyke.

Suggestions For Effective Management Of Coastal Erosion Along The Nigerian Coast

Coastal erosion, as a natural occurrence and process may not entirely be stopped, but can be properly managed in such a manner as to maximize the gains associated with its occupation and physical development while reducing the risk in such human occupation. Two categories of coastal erosion have been distinguished- long-term (chronic) and short-term (storm-induced) coastal erosion. Generally, chronic erosion is more difficult to address than storm-induced erosion (Dees, et al., 2001). A combination of methods has been recognized in the management of coastal erosion in the world. Charlier et al. (1989a, 1989b) has divided the methods into 'hard' and 'soft' methods. However, despite the use of the hard and soft or a combination of the hard and soft methods of coastal erosion management, it has been found that three quarter's of the world's beaches are in retreat, some at a rate of a year. Thus, in this paper, Integrated Coastal Zone Management (ICZM) is been recommended for the Nigerian coastal area.

ICZM is defined as a dynamic process in which a coordinated strategy is developed and implemented for the allocation of environmental, socio-cultural and institutional resources to achieve the conservation and sustainable use of the coastal area (Nwilo, 2004). ICZM has been described as a process for decision-making: it should be continuous, iterative and should recognize the contributions of stakeholders and the natural dynamism, both physical and ecological, of the coastal environment (Smith, et al., 2004). A primary goal of ICZM is to overcome the compartmentalized approach to managing coastal resources by harmonizing the decisions of diverse jurisdictions and levels of government. For instance, EU has recently recommended the promotion of participatory planning in coastal management and encouragement to develop systems that allow the monitoring and dissemination of coastal zone information (Jude, et al., 2004). In particular, ICZM aims to bridge the traditional divide between management of the land and the water. ICZM, therefore, is also about building institutions that facilitate this integration and is also founded on the principles of sustainable development, recognizing that the coastline is the fount of resources of great value to human communities and that these resources should be managed in ways that conserve their value for future generations (Jude, et al., 2004).

However, some of the challenges facing the successful implementation of an ICZM in Nigeria for the management of the Nigerian coastal erosion can be broadly grouped into political, socio-economic and manpower factors. These challenges must be properly addressed as part of the concerted efforts aimed at avoiding adverse effects of coastal erosion phenomenon.

References:

- Asangwe, C.K. (1993): Issues and Problems in controlling threats of erosion along the Nigerian coastline. In coastlines of West Africa, American Society of Civil Engineers, 309. Charlier, R.H., CHR. DE Meyer and D. Decroo, 1989:
- CREMS, (2000): Amelioration of Coastal Erosion and Sea Incursion in Ondo State. Report prepared for Ondo State Government.
- Davies, D.W. (1986): The Retreating Coast. In *Journal of soil and Water Conservation*, Vol. 41, Pp. 146-157.
- Dees, K. and Davis, G. (2001): Galveston Coastal Erosion: An overview of Causes, Problems and where we stand, Environmental Law Project, June 29, 2001.
- Ebisemiju, F.S. (1987): An evaluation of factors controlling present rates of shoreline retrogradation in the Western Niger Delta, Nigeria. *Catena* 14, 1-12.
- Ebisemiju, F.S. (2001): Environment and Development Issues in the Niger Delta Area of Onod State, Nigeria. Paper presented at the Stakeholders' Forum/Workshop, Niger Delta Development Commission, July 2, 2001
- Faniran, A. (1986): African Landforms: An Introductory Geomorphology. Heinemann Educational Books Nig. Ltd. Pp. 223-270.
- Ibe, A.C. (1988): Coastline Erosion in Nigeria, Ibadan University Press, Ibadan.
- Ibe, A.C. (1990): Global Climate Change and the Vulnerability of the Nigerian Coastal zone to accelerated sea-level rise: Impacts and Responses Measures. Nigeria Institute of Oceanography and Marine Research Technical Paper No. 32. February, 1990.
- Nordstorm, K.F. (1994): Beaches and Dune of Human Altered coasts in Progress in Physical Geography, Edward Arnold. Pp. 497-516
- Parliamentary Assembly/Council of Europe, (2003): Erosion of the Mediterranean Coastline: Implication for tourism. Report of Committee on Economic Affairs and Development Group of the European People's Party (EPP/CD).
- U.N.D.P., (1993): "World Urbanized Problems in United Nation Population Development 1993, New York.
- Uzoroh, E.J., (1971): Recent Rates of Shoreline Retreat at Victoria Beach, Lagos, Nigeria. *Nigeria Geographical Journal* XIV. Pp 49-58.
- Valentine, H. (1971): Land Loss at Holderness in Applied Coastal Geomorphology. J.A. Steers.
- Viles, H.A., (1990): Coastal Geomorphology into the 1990s. *Progress in Physical Geography*. Vol. 14, No. 2, 1990. Pp. 238-250.
- Viles, H.A. (1991): Coastal Geomorphology. *Progress in Geography*. Pp. 182-192.
- Webb, J.E. (1960): The Erosion of Victoria Beach: Cause and Cure. Ibadan University Press. Ibadan.

PROPERTIES OF NEW ENVIRONMENTALLY FRIENDLY BIODEGRADABLE INSULATING FLUIDS FOR POWER TRANSFORMERS

Pawel Rozga, Dr., PhD Eng.
Technical University of Lodz, Poland

Abstract:

The problem of using the new environmentally friendly insulating liquids like synthetic and natural esters for power transformers was presented in this article. The positive and negative properties of the esters were described on the basis of a comparison with the most popular insulating liquid like mineral oil. The results of the own experiences and experimental studies were taken into account as well as the results of works presented in the world literature. Conclusions from the analysis indicate that the wide knowledge about ester properties is necessary to correct design process and to proper exploitation of the power transformers filled in esters.

Key Words: Power transformers, insulating fluids, environmental protection, high voltages

Introduction

Power transformers are the one of the most important components of electrical power system serving the transmission and distribution of electricity in the high voltage three-phase electrical power grids and from these grids to the low voltage grids. From this reason, the transformers are the ubiquitous in the human environment. Generally, the transformers are not the devices that generate particularly troublesome environmental problems. However, taking into account the contemporary transformation of public awareness, forcing the pro-ecological look on the each technical device, especially on it, that appears in large number of units in the human environment, it should be searched also for power transformers some new solutions, which improve their ecological properties (Kycior K. 2007, Mosinski F. 2009). A significant majority of the total number of transformers, produced in all ranges of voltage and power, are the oil power transformers. The so-called dry transformers (with resin insulation) are the margin of exploited units. Therefore, in the assessment of environmental risks connected with the transformer in service, the mineral oil is the most important element to be taken into consideration. Of course, the aspects such as the risk of electric shock, noise and vibration as well as energy and power losses are also important, but taking into account the effects for the environment from these types of interactions, these, connected with mineral oil, seem to be the most dangerous (Kucharska A. 2007, Mosinski F. 2009).

Mineral oil is a mixture of different kinds of hydrocarbon compounds (naphthenic C_nH_{2n} , such as cyclohexane C_6H_{12} , paraffin C_nH_{2n+2} , such as hexane C_6H_{14} , and aromatic C_nH_n , such as benzene C_6H_6). The proportions of the contents of the individual components depend on the composition of the starting petroleum and, for example, most prized naphthenic transformer oils can be obtained only with the certain types of petroleum. Therefore, in the case of spillage of oil to the environment, the oil is not a neutral for it. Additionally its amount (exemplary in the oil power transformer having the nominal voltage 110 kV the mass of oil is about 7-8 thousand kilograms, while in the transformer 400 kV such mass is nearly 80 thousand kilograms) intensifies the hazard. Commonly used mineral transformer oil has an ecological threat index twice higher than water, but it concerns the fresh oil without any polychlorobiphenyls admixtures (PCBs). The aged oils (those derived from the long-operated transformers) may, however, contain significant quantities of polycyclic aromatic compounds (benzene, toluene, xylene), which have a carcinogenic properties. Ecological threat index can then be increased up to the three times in comparison with fresh oil. Generally, in the case of transformer failure, the mineral oil leaked to the soil may be the source of its contamination (degradation of oil-soaked layer), and in consequence also the contamination of water bodies and flora

and fauna, which is in the area of contaminated space. Due to the relatively high solubility in water, the hydrocarbon compounds can migrate with infiltrating rainwater by the aeration zone to the first aquifer. For example, 1 kg of the waste oil that has leaked from the transformer to the water reservoir makes it unfit to drink in the volume of 5 million liters (Kucharska A. 2007, Mosinski F. 2009). Numerically, the measure of a potential contamination of the given fluid is determined by the level of its biodegradability. For mineral oil, this factor is very low, and in accordance to the OECD 301 standard, is only 10%. It means that after 28 days from the entering the oil to the environment, only such small part of this fluid surrenders to self degradation. Thus, the failures concerning the oil leakage are undoubtedly harmful to the environment (air, water, soil), and the costs of removing their effects are usually very large. Prevent them or limitation their effects, although expensive, has an importance meaning for the environment. It is believed that this role is fulfilled by the special place with the pans for leaking oil, by the systems of oil and water separation, and by the monitoring of the transformer and its equipment (Mosinski F. 2009, Rozga P., Skowron A., 2012). The example of the oil pan installation is shown in Fig. 1.

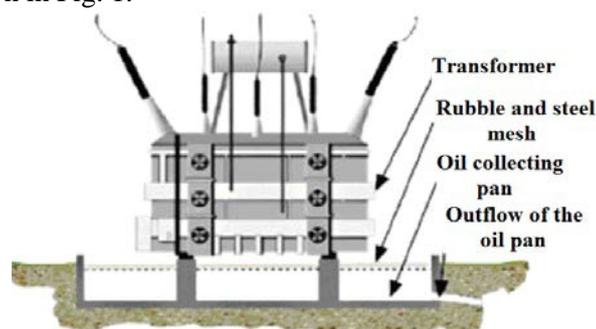


Fig. 1. Example of the oil pan installation.

The volume of designed and built oil pans should be sufficient to bring all of the oil contained in the transformer. In the cases where the transformers are equipped additionally with the water sprinkler systems, the volume of the pan is increased by an additional 10-20%.

Mentioned on the beginning of this chapter the problem of PCBs is now generally margin. Currently used transformer oils meet the criteria for the PCB content, which, in accordance to the specified requirements, should be less than 50 ppm (Kaminska A. 2007).

Another problem of environmental nature is the combustibility of mineral oil due to its relatively low flash point. Combustion products of mineral oil being the result of its ignition occurred because of transformer failure (e.g. short circuit), are considered to be dangerous and cause air pollution. 1000 kg of burnt, under unfavorable conditions, mineral oil emits circa 10 kg of harmful substances to the atmosphere. Additionally, the side product accompanying the mineral oil combustion is the dense and very black smoke (Fig. 2).



Fig. 2. Example of transformer failure with fire

Therefore, one of the pro-ecological solution, which is connected directly with the use of mineral oil in power transformers, is to replace it by a new environmentally friendly insulating liquids produced on the basis of natural and synthetic esters, which importantly decrease the negative influence of the transformer on the environment (Oommen T. V., Clairborne C. C., Mullen J. T. 1997, Borsi H., Gockenbach E. 2005, Martins M. A. G. 2010). Such alternative is necessary, especially in

the areas having highly restricted regulations concerning the environmental protection and fire prevention. The pro-ecological liquids, from the environmental protection point of view, are characterized by the high level of biodegradability reaching even 95%, and from the fire prevention point of view, have the significantly higher than mineral oil flash point (above 300 °C to circa 140 °C characterizing the mineral oil).

It is however important, that in spite of the good ecological properties of esters, they can bring many problems, especially in the phase of design and manufacture of transformers, what resulted from the short time of existing the esters on the transformer market, and incomplete knowledge in this field. Therefore, the properties of new, environmentally friendly electro-insulating liquids will be presented in comparison to the analogical properties of mineral oil. The comparison will be based on the fundamental experimental studies (also author's). The similarities (mainly in the field of electrical strength) and the differences (in the behavior of liquids in the heating process, in the field of lightning strength, in the field of chemical parameters) will be presented and critically discussed.

The problem of using biodegradable liquids in power transformers

Environmentally friendly insulating fluids, which are an alternative to mineral oils, are the synthetic esters and natural esters, for which the base are vegetable oils produced from canola, soybean, corn etc. In Fig. 3, the molecular structures of two commercial dielectric liquids - natural and synthetic ester are shown.

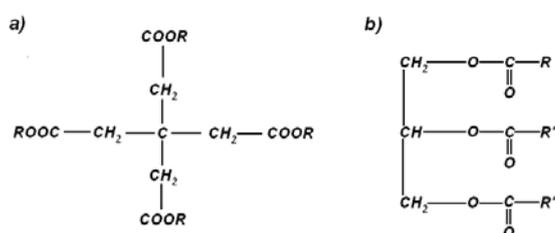


Fig. 3. The chemical structures of synthetic ester (a) and natural ester (b); R, R', R'' – organic groups.

Both liquids have been more than ten years in the test phase, and prototype units filled by these fluids are with nominal voltages up to 238 kV and above 100 MVA in capacity. Natural ester is an ester based on soybean oil. The major molecular component of this ester is mixed triglyceride built from the glycerol molecule and three molecules of long chain fatty acids linked together by ester bonds. In a one triglyceride molecule may be a variety saturated and unsaturated fatty acids consisting even from the 22 carbon atoms in the chain. The synthetic ester is an organic synthetic chemical compound - pentaerythritol ester, which consists of four ester groups -COOR which are at the end of the cross structure of the compound, and in which the organic groups R may be either the same or different.

In the Table 1, the basic physico-chemical and dielectric parameters for both esters and mineral oil are summarized (FR3 Data Sheet 2008, Midel 7131 Data Sheet 2010). These parameters concern the fresh fluids as received by the manufacturers, without the use of any treatment.

Table 1. Basic parameters of synthetic ester, natural ester and mineral oil.

	Units	Synthetic ester	Natural ester	Mineral oil
Physico-chemical properties				
Density at 20°C	kg / dm ³	0,97	0,92	0,88
Specific Heat at 20°C	J / kg K	1880	1848	1860
Thermal Conductivity at 20°C	W / m K	0.144	0.177	0.126
Kinematic Viscosity at 20°C	mm ² / s	70	85	22
Kinematic Viscosity at 100°C	mm ² / s	5.25	8.4	2.6
Pour Point	°C	-60	-21	-50
Fire Point	°C	316	360	170
Flash Point	°C	260	316	150
Fire Hazard Classification to IEC 61100 / IEC 61039	-	K3	K2	0
Biodegradability	%	89	97	10
Dielectric properties				

Breakdown Voltage	kV	> 75	> 75	70
Dielectric Dissipation Factor Tg δ at 90 ^o C	-	< 0.008	< 0.003	< 0.002
Permittivity at 20 ^o C	-	3.2	3.1	2.2

As was mentioned above, among the main physico-chemical properties of the esters, particular meaning, from the side of ecological aspect, has two of them: biodegradability and flash point. For both of esters, their biodegradability is significantly higher than mineral oil, what confirms the better ecological properties of these liquids. Taking into account that after 28 days both of esters decay in above 60%, according to the OECD 301 standard, they can be admitted as a biodegradable. Thus, using esters in the places having restricted environmental regulations is not a problem. It is important to notice, that in the case of esters it is not necessary to use the oil pans, what finally reduces the costs of installation of the transformers.

The second of the mentioned parameters – flash point – also specifies the esters as the environmentally friendly liquids. Performed for the all three insulating fluids the test of “open burner” indicates on the flash temperatures identical as determined by the manufacturers in the data sheets. For the mineral oil the temperature, after delivered the open fire from the acetyl burner, rose violently and after 4 minutes the ignition of mineral oil took place in the temperature of about 130 °C. Then the oil burned without the open fire and during the combustion process emitted the dense black smoke. Whereas in the case of ester fluids the temperature rose slower and after the 70 minutes the process was stopped because the ignition did not occur and the temperature reached nearly 300 °C. The better properties of esters are also visible looking at the intensity of the smoke emitted by the burned liquids. In the case of mineral oil smoke is black and very dense while for the both esters smoke is emitted in the smaller amounts and its dense is also much lower. In percentages, the volume of waste gases emitted by the burned esters in relation to the gases emitted by burned mineral oil is only about 15%.

From the insulating properties point of view the most important parameter is breakdown voltage. This voltage, according to the IEC 60156 standard, is determined by the introduction of the liquid sample to the special apparatus with profiled metal electrodes creating the uniform electrical field distribution and then subject to the rising electrical field resulting from the AC voltage increasing. The average value of six breakdowns occurring in the specified conditions is taken as an AC breakdown voltage. Of course, the value of breakdown voltage depends on the quality of tested liquid (water and impurities content), so in order to compare the values of breakdown voltages of different dielectric liquids, they should have the similar quality (Li J., Grzybowski S., Sun Y., Chen X., 2007, Dang Viet-Nung, Beroual A., Perrier C. 2011). In the Fig. 4, the results of the measurements of AC breakdown voltage are presented. These results confirm identical properties connected with the electrical strength of synthetic ester, natural ester and mineral oil.

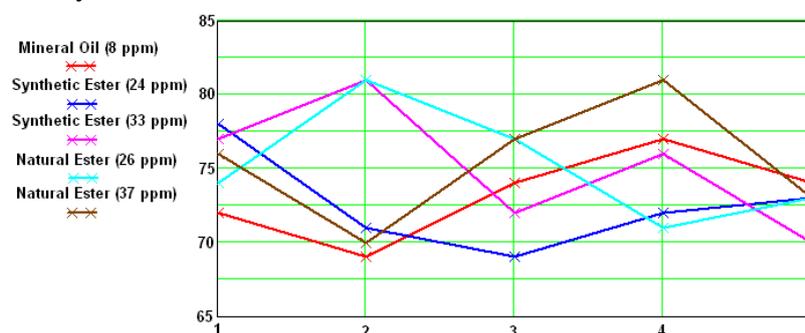


Fig. 4. AC breakdown voltage for different types of insulating liquids measured according to the IEC 60156 standard.

Looking at the influence of moisture content on the electrical strength, the esters behave more favorable. It is because the esters are characterized by the constant value of breakdown voltage apart from the moisture content up to 600 ppm (Martins 2010, Dang Viet-Nung, Beroual A., Perrier C. 2011). For the mineral oil, its strength goes down even at the small increasing of moisture (at 20 ppm the breakdown voltage is only 40 kV while for good conditioned oil this value is in the range of 65-75 kV).

Similarly other parameter characterizing the dielectric properties of the liquids, that is the electrical permittivity, also indicates on the possibility of better using the esters in the insulating systems of power transformers. The insulating systems of the transformer windings are usually a combination of electro-insulating fluid and solid insulating materials like paper or transformerboard. These systems are series, series-parallel or parallel. In the series systems, which are the most often systems in the transformers, the distribution of the stresses is inversely proportional to the electrical permittivity of the materials created it. The system has a higher strength if the more uniform stress distribution is. The higher value of electrical permittivity of the esters (3.1-3.2) than the mineral oil (2.2) thus influences profitable on the stress distribution in the system: electro-insulating fluid – cellulose (electrical permittivity of cellulose is 4).

Although there are many positive properties of the biodegradable esters concerning the electrical strength at AC voltage, their lightning strength, that is the strength for a short voltage pulses having steep front (1.2 μ s) and long back (50 μ s), is not as good as in the case of mineral oil. The studies on the mechanism of initiation and propagation of electrical discharges developing under above mentioned lightning pulses indicated that only at inception voltage or at the 50% breakdown voltage, the discharges developing in the same way independently on the type of dielectric liquids. When the voltage increases, thus, when the voltage multiplicity in relation to the 50% breakdown voltage rises, the discharges, both in the natural and synthetic esters stands more energetic and at the lower than in the case of mineral oil overvoltages, the developing of discharges in esters is determined by liquid phase ionization. More energetic discharges are more dangerous for the solid insulation of the transformer (paper and transformerboard) causing its damage in the form of local carbonization. Therefore, when the one of the esters is used as an insulating medium, it is indicated, for the sake of lightning strength, using the wider insulating gaps in the transformer. It is particularly important in the case of big transformer units having the rated voltages equal or higher than 110 kV (Tran Duy C., Lesaint O., Denat A., Bonifaci N. 2009, Liu Q., Wang Z. D. 2011).

One of the negative properties of the esters, but without any direct influence on the possibility of using esters in power transformers, is the susceptibility on the action of concentrated heat flux (Rozga P., Skowron A. 2012). This property was observed during the process of preparing of the esters for the filling the transformer, and then confirmed by the laboratory measurements. The experimental studies indicated that the influence of the heating elements, used in the heating process of the esters, determined by the unitary surface load expressed in W/cm^2 , causes the growth of the dielectric dissipation factor $\tan\delta$ of the esters. This growth, at the value of unitary surface load identically as used during the heating process of the mineral oil (2 W/cm^2), is such meaningful, that $\tan\delta$ of the esters exceeds the boundary values determined in the standards although the fluids were not exploited. The exemplary curve presenting the dependence of the $\tan\delta$ versus unitary surface load for synthetic ester, performed during experimental works, was shown in Fig. 4.

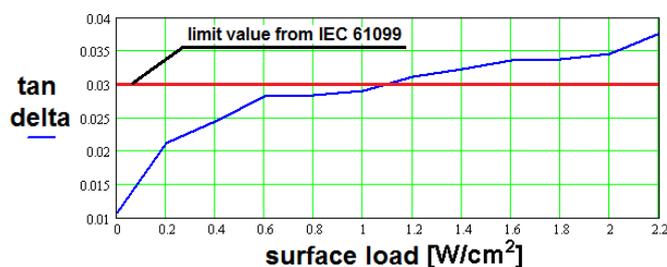


Fig. 5. Dependence of dielectric dissipation factor versus unitary surface load for a synthetic ester.

The described phenomena do not concern mineral oil. Its treatment does not change the dielectric properties of mineral oil. The differences in the different behavior of esters and mineral oil result from the different physico-chemical properties, which directly influence on the phenomena occurring during the heating. In the case of mineral oil, the laminar flow of the oil, because of its relative low density and first of all low viscosity, is relatively fast. Thanks to it, the oil reminds in the direct contact with the surface of the heater only by the short time. In the case of esters, which are characterized by the much higher density and viscosity, contact with the surface of the heater is much longer. It causes that the layers of the liquids “stick” to the surface of the heater and their overheating

taking effect of disintegration and decomposition of molecular bonds. Finally, in the liquid volume the polar particles appear and growth of above mentioned dielectric dissipation factor takes place. Taking into account the large experience in the field of mineral oil treatment and lack of precisely defined parameters of the heating process of the esters, the presented problem is quite significant from the production phase point of view. For the esters the production process becomes lengthen because similar heating conditions as for mineral oil is not possible to keep.

The higher density and first of all the higher kinematic viscosity of the esters is also important at the consideration of their influence on the transformer cooling conditions. Generally, the working transformer generates the energy losses. These losses are in the windings, in the magnetic core and in the metal elements of the construction. The losses transform in the heat, what causes the heating of the individual parts of transformer. The requirement for the good work of transformer is the effective heat abstraction outside it. The mechanism of the heat abstraction looks as following: the heat arising in the individual internal parts of transformer gets out of its (heat conduction); from where with the usage of the cooling medium (electro-insulating fluid) is directly or indirectly carried away to the surroundings. The higher kinematic viscosity of the insulating fluids is, the better the cooling function is fulfilled. Additionally, the kinematic viscosity, both for esters and for mineral oil, depends on the temperature, what was presented in Fig. 5. However, the mineral oil has a viscosity much lower and because of it, it seems to notice that mineral oil fulfils better the cooling functions in the power transformer than biodegradable esters.

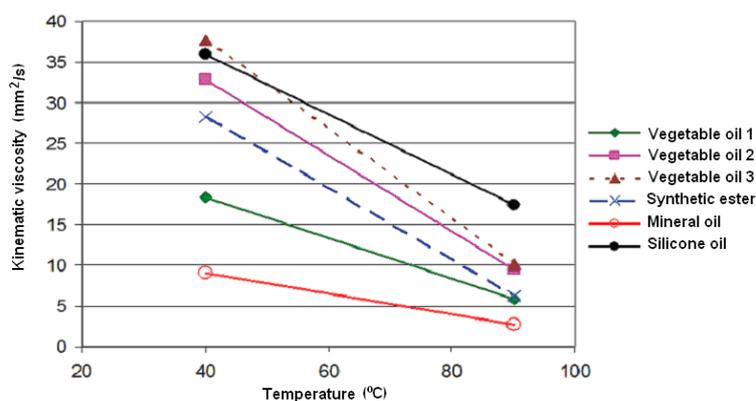


Fig. 5. Dependence of kinematic viscosity versus temperature for selected dielectric fluids.

From the other hand, the higher density of esters causes that these liquids do not circulate as freely and with the proper flow velocity in the narrow cooling channels of the windings and magnetic core as mineral oil. Flowing ester cannot reach all fragments of the insulating systems of transformer windings. Thus, the construction of insulating system of power transformer filled by ester should take into account its physico-chemical properties in order to provide the correct circulation of the cooling fluid both in the natural and forced circulation. Hence the systems forcing the liquid circulation in the transformer filled by ester should be characterized by higher power, and the oil channels should be wider than in the case of transformer with mineral oil (Hulshorst W.T.J., Groeman J.F. 2002, Pukel G. J. 2009, Martins M. A. G. 2010).

Conclusion

Biodegradable synthetic and natural esters are the good alternative for mineral oil, especially in the situations when the power transformer has to be installed in the places for which the restricted environmental regulations are specified. Next to the many positive aspects like the higher biodegradability, high flash point and good properties concerning the AC electrical strength, esters have also negative features, which should be taken into account in the design phase and during the exploitation of the transformer with esters. The parameters of the designed transformer must be based on the knowledge about these negative features. Designer should know that the esters are susceptible to the action of a concentrated heat flux, have the worse cooling properties than mineral oil, and lower lightning strength. The final product – power transformer filled by ester – will be in effect the more expensive device, but taking into consideration the pro-ecological tendencies reigned on the

transformer market and still increasing demand for new environmentally friendly products, in the future the cost will be certainly lower.

References:

- Borsi H., Gockenbach E., Properties of ester liquid Midel 7131 as an alternative liquid to mineral oil for transformers. IEEE Intern. Conf. on Diel. Liquids, 2005.
- Dang Viet-Nung, Beroual A., Perrier C., Comparative study of statistical breakdown in mineral, synthetic and natural ester oils under AC voltage. IEEE Intern. Conf. on Diel. Liquids, 2011.
- Hulshorst W.T.J., Groeman J.F., Energy saving in industrial distribution transformers. KEMA report reference 40130017-TDC 02-24226A, 2002.
- Kaminska A., Very dangerous waste – PCB in oil (in Polish). *Ekologia*, No 2, 2007.
- Kucharska A., Review of water-soil protection systems from electro insulating oil leaks from electrical power transformers in law aspects (in Polish), *Polityka Energetyczna*, T. 2, 2007.
- Kycior K., Environmental protection against the oil leakage (in Polish). *Ekologia*, No 1, 2007.
- Li J., Grzybowski S., Sun Y., Chen X., Dielectric properties of rapeseed oil paper insulation. Annual Report Conf. on Electr. Insul. and Diel. Phenom., 2007.
- Liu Q., Wang Z. D., Streamer characteristic and breakdown in synthetic and natural ester transformer liquids under standard lightning impulse. IEEE Trans. Diel. Electr. Insul., No 1, 2011.
- Martins M. A. G., Vegetable oils, an alternative to mineral oil for power transformers – Experimental study of paper aging in vegetable oil versus mineral oil. IEEE Electr. Insul. Mag., No 6, 2010.
- Mosinski F. Ecological aspects of power transformer exploitation (in Polish). *Wiadomosci Elektrotechniczne*, No 3, 2009.
- Oommen T. V., Clairborne C. C., Mullen J. T., Biodegradable electrical insulation fluids. *Proced. of Electr. Insul. Conf.*, 1997.
- Perrier C., Beroual A., Experimental investigations on mineral and ester oils for power transformers, Annual Report Conf. on Electr. Insul. and Diel. Phenom., 2007.
- Pukel G. J., Schwarz R., Schatzl F., Bauman F., Gertstl A., Environmental friendly insulating liquids - a challenge for power transformers. *Cigre Southern Africa Regional Conf.*, 2009.
- Raoul D., High and extra high voltage power lines, health and environment. *Electra*, No 256, 2011.
- Rozga P., Skowron A., Changing the dielectric parameters of the transformer fluid based on synthetic ester in the condition of a concentrated heat flux. *Proce. of Intern. Conf. on high voltage Eng. and Appl.*, 2012.
- Tran Duy C., Lesaint O., Denat A., Bonifaci N., Streamer propagation and breakdown in natural ester at high voltage. IEEE Trans. on Diel. and Electr. Insul., No 6, 2009.
- Envirotemp FR3 Fluid – Material Safety Data Sheet, 2008.
- Midel 7131 - Comparison to Alternative Technologies, 2010.
- Midel 7131 Transformer fluid - Technical Data Sheets. www.midel.com, 2010.
- IEC 60296 Fluids for electrotechnical applications – Unused mineral insulating oils for transformers and switchgears.
- IEC 60156 Insulating liquids – Determination of the breakdown voltage at power frequency – test method.
- IEC 61099 Insulating liquids – Specifications for unused synthetic organic esters for electrical purposes.
- ASTM D6871-03:2008 Standard specification for natural (Vegetable Oil) ester fluids used in electrical apparatus.
- IEEE C57.147-2008 Guide for acceptance and maintenance of natural ester fluids in transformers.
- IEEE C57.106-2006 Guide for acceptance and maintenance of insulating oils in equipment.

COMPLIANCE OF WASTEWATER TREATMENT PLANTS IN JÄRVA COUNTY WITH THE EU URBAN WASTEWATER TREATMENT DIRECTIVE AND ESTONIAN NATIONAL REQUIREMENTS

Raili Niine, MA
Enn Loigu, Prof
Karin Pachel, PhD

Department of Environmental Engineering, Tallinn University of Technology, Tallinn, Estonia

Abstract:

The aim of this study was to assess the treatment efficiency conformity of Wastewater Treatment Plants to the Estonian and European Union requirements in the area of Estonian nitrate vulnerable zone as well as evaluate the existing national wastewater monitoring programme and the efficiency of the self-monitoring programme. Monitoring the results and total pollution load of effluent discharged to the receiving water body from 2007 to 2008 are used to assess the treatment level of WWTPs and the need to establish higher treatment requirements. Estonian national standards, due to the fact that Estonian water bodies are small and vulnerable to pollution, are stricter than the Urban Wastewater Treatment Directive (UWWTD) requirements. If the requirements given in UWWTD are not sufficient to achieve a good status for water bodies, it is proven that wastewater discharge adversely affects the receiving water body and this discharge is one of the important point-pollution sources for water bodies, additional stricter wastewater treatment requirements are needed. WWTPs with a more than a 2,000 population equivalent (p.e.) should be regulated vigorously to guarantee their compliance with the requirements, because WWTPs with more than 2,000 p.e. count for 80% of the total pollution load. A pollution load from WWTPs of less than 2,000 p.e. on the receiving water bodies is marginal; however, the extent of the impact of each individual WWTP depends on, among other things, the characteristics of the wastewater, the turbulence of the receiving water body and the area's sensitivity as well as other pressure sources.

Key Words: Wastewater treatment, pollution load, effluent, receiving water body

Introduction

In Estonia, Järvamaa is one of the three counties that is designated a nitrate vulnerable zone according to European Nitrates Directive (European Community (EC), 1991/1). According to Regulation no. 17 of the Estonian Government, 2003, there is one nitrate vulnerable zone consisting of two subareas – Pandivere and Adavere-Põltsamaa. In Estonia, the rivers of almost all catchment areas (excl. the islands) spring from the slopes of the Pandivere Upland karst area (Environmental Report 7, 1993). In a nitrate vulnerable zone, all water bodies are highly affected by the non-point sources from agriculture and are also easily influenced by point sources from wastewater treatment plants (WWTPs) (Valdmaa et al., 2008). It is important to control the nutrient load, particularly phosphorus pollution, even that of small settlements within agricultural and rural catchments (Jarvie et al., 2006; Iital et al., 2010). Therefore, WWTPs have to comply with the discharge requirements established by both the Urban Wastewater Treatment Directive (UWWTD) and the Estonian national standards (ENS). The European Union integrated water policy main document, which establishes a framework for EU actions in the field of water policy – Water Framework Directive (WFD), sets a goal to protect all waters against pollution and to achieve good status for all waters, promoting sustainable water and wastewater management (European Community (EC), 2000; Ministry of Environment, 2011). One of the most important surface water pressures is point-source pollution (IMPRESS, 2002). Furthermore, the Baltic Sea countries adopted an action plan to achieve the good ecological status of the Baltic Sea by 2021 (Helcom, 2007). One of the main issues covered by the Baltic Sea Action Plan is the further reduction of nutrient inputs in order to limit the eutrophication of

water bodies. Estonian water bodies are sensitive to nutrients and have a high eutrophication risk. To limit the eutrophication process, it is important to decrease total phosphorous (TP) and total nitrogen (TN) loads. Several studies have researched the impact of WWTPs on receiving water bodies (Kontas et al., 2004; Millier et al., 2011; Dickenson et al., 2011). For instance, Kontas et al. (2004) studied the concentrations of inorganic nutrients, phytoplankton chlorophyll- α , and N/P ratios before and after the treatment plant. Dickenson et al. (2011) has reported on the presence of trace organic chemicals in municipal wastewater effluents. This paper uses a set of common trace organic chemicals as indicators to assess the degree of impact and attenuation of trace organic chemicals in receiving streams. Also, other Baltic countries are considering the problem of phosphorus and nitrogen removal from wastewater and, therefore, several studies have investigated the different methods of removing nutrients from wastewater. One such study, Vaboliené et al. (2007), evaluated the effect of biological nitrogen removal from wastewater on biological phosphorus removal, and Dauknys et al. (2009) analysed the influence of substrate on the biological removal of phosphorus. These studies use the results of only a few WWTPs and these analyses do not assess the cumulative impact of effluent from all regional WWTPs on receiving water bodies. This study attempts to check the compliance to the established WWTPs requirements for different sized WWTPs in Järva County and assesses the cumulative impact on receiving water bodies. Drawing from study results and taking into account the pollution load and impact on receiving water bodies, it is possible to assess the compatibility of the validated wastewater requirements in Estonia. The monitoring data of WWTPs in Järva County were compared with the requirements set by both UWWTD and ENS. ENS is based on the regulation of the Estonian Government (Government of Estonia, 2001) and water permit requirements (Ministry of the Environment, 2009/2). A further task was to estimate and compare existing national wastewater monitoring and the self-monitoring results and efficiency of WWTPs as well as make proposals to improve monitoring programmes and recommendations to change legislation. These programmes should help to evaluate the wastewater impact on the recipient and the operating efficiency of WWTPs.

Methodology for Compliance Check

The monitoring results were based on control sampling conducted by environmental authorities from 2007 to 2008. The total pollution load is calculated using monitoring results and the National Environmental Register (Ministry of the Environment, 2009/1) to get real flow rate of WWTP effluent. 154 samples from Järvamaa WWTP effluents were taken and analysed for biochemical oxygen demand (BOD₇), suspended solids (SS), TN, TP and chemical oxygen demand (COD) concentrations by the Environmental Research Centre of Estonia. The UWWTD only set limit values for WWTPs with a pollution load more than 2,000 p.e. because it aims to establish minimum treatment requirements (European Community (EC), 1991/2). The effluent limit values for WWTPs with a pollution load less than 2,000 p.e. were not regulated by these requirements. Separate criteria are used to assess the compliance of these WWTPs with the requirements. The common limit values given in the permits for the special use of water (Ministry of the Environment, 2009/2) for establishing the criteria for WWTPs with a pollution load less than 2,000 p.e. were used. Both UWWTD requirements and ENS are listed in Table 1 and 2, respectively.

Table 1. Criteria for the assessment of compliance with the UWWTD requirements

Pollution indicator	Limit values, mg/l			
	<2,000 p.e.*	2,000-9,999 p.e.	10,000-99,999 p.e.	≥100,000 p.e.
BOD ₇	25	25	25	25
COD	125	125	125	125
SS	35	35	35	35
TP	-	-	2	1
TN	-	-	15	10

* The Urban Wastewater Treatment Directive does not establish common criteria and only requires an appropriate treatment level. These criteria are developed by taking into account the aim of the directives and requirements given in the permits for the special use of water.

Table 2. Criteria for the assessment of compliance with the ENS

Pollution indicator	Limit values, mg/l				
	<500 p.e.*	500-1,999 p.e.*	2,000-9,999 p.e.	10,000-99,999 p.e.	≥100,000 p.e.
BOD ₇	25	25	15	15	15
COD	125	125	125	125	125
SS	25	25	25	15	15
TP	-	2	1.5	1	1
TN	-	-	-	15	10

* Estonia do not establish common standards. These criteria are developed by taking into account the aim of the directives and requirements given in the permits for the special use of water.

When comparing Table 2 with Table 1, two major differences can be found: 1) ENS regulates BOD₇, COD and SS for WWTPs with a pollution load from 0 p.e. to 500 p.e. and, additionally, TP from 500 to 1,999 p.e. 2) For WWTPs with a pollution load greater than 2,000, the ENS for BOD₇, SS, and TP, are 40%, 57%, and 100% higher than the UWWTD requirements, respectively.

The compliance check of WWTPs to the UWWTD requirements is only made for those WWTPs that fall within the scope of the directive. Because the UWWTD does not establish uniform requirements to be achieved for agglomerations with a pollution load of less than 2,000 p.e., requiring only an appropriate treatment level, the limit values for this study are set to achieve the good status of water bodies required by the WFD. The requirements for WWTPs with a pollution load less than 2,000 p.e. must be proportionate to the requirements for WWTPs with a pollution load more than 2,000 p.e.. Also, the requirements must comply with environmental protection objectives. The limit values for WWTPs with a pollution load less than 2,000 p.e. are given for BOD₇, COD and SS. A TN and TP removal requirement is given only for the WWTPs with a pollution load of more than 10,000 p.e. that discharge the effluent to the pollution sensitive water body according to the UWWTD. Since the whole territory of Estonia is designated as a pollution sensitive area, the criteria also demand TP and TN removal for WWTPs with a pollution load more than 10,000 p.e., as set by the UWWTD in Table 1.

To comply with the requirements, a minimum number of effluent samples are required each year, depending on the size of the WWTP. The compliance check for WWTPs with more than 2,000 p.e. is based on the principle that 3 monitoring results of 4 samples must be in conformity with the criteria set down in Table 1 or Table 2. During this study, national monitoring consists of 4 samples per year for all WWTPs with more than 2,000 p.e. and at least 2 samples per year for WWTPs with a pollution load of 500 – 2,000 p.e. All the samples on the basis of this analysis have been analysed in accredited laboratories and all water samplers have the certification of water sampling. The WWTPs of less than 2,000 p.e. are in conformity with the requirements, if all quality indicators (the criteria set down in Table 1 or Table 2) are met in more than 50% of all monitoring results. If the number of samples that confirm and the number of samples that does not confirm with the criteria set out in Table 1 or Table 2 is equal, the compliance check is based on the average monitoring results made by the water user.

Monitoring Results

Based on the 2008 monitoring results, the average effluent concentrations in mean (\bar{x}) and median (Me) of WWTPs in Järva County are shown in Figure 1. The mean concentration is influenced by a single extreme value of some WWTPs. The median characterises the average values of pollution indicators more adequately because the median value represents the centre of the data distribution.

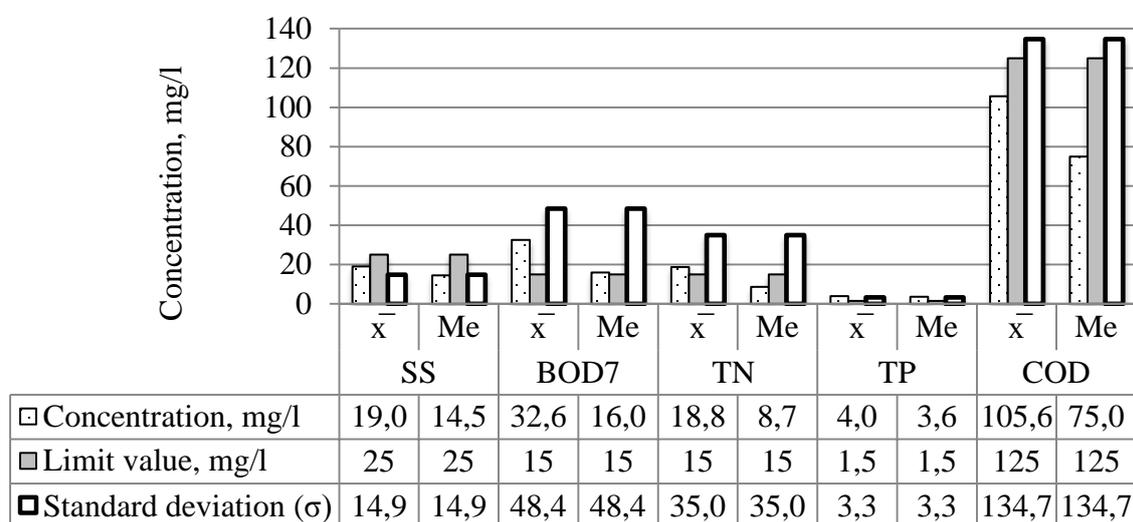


Figure 1. Average concentrations of effluent of urban WWTPs in Järva County in 2008.

Figure 1 reflects all of the 46 urban WWTPs, apart from Väätsa landfill. The limit values vary according to the size of the WWTP. The effluent average results are affected by the results of these WWTPs, which are excluded from the UWWTD. The average results of the WWTPs that are not within the scope of the directive are worse by 20% when compared to the WWTPs results given in Figure 1. If Figure 1 did not also reflect the results of the urban WWTPs, which do not belong within the scope of the UWWTD, the maximum difference would be the median of organic matter. The median of organic matter for urban WWTPs covered by the directive is 14 mg/l, which is 12.5% lower than shown in Figure 1. The median for other indicators would be lower for urban WWTPs covered by the directive between 3.5-6.7 %, compared to that given in Figure 1. There were 21 urban WWTPs that do not belong to the directive, which constitutes 42 % of all WWTPs covered in this study. Separately analysed monitoring results for the urban WWTPs that belong within the scope of the directive and the WWTPs outside the scope of the directive are given in Table 3.

Table 3. Average concentrations of the effluent of WWTPs in Järva County

	SS, mg/l		BOD ₇ , mgO ₂ /l		TN, mgN/l		TP, mgP/l		COD, mgO ₂ /l	
	\bar{x}	Me	\bar{x}	Me	\bar{x}	Me	\bar{x}	Me	\bar{x}	Me
Urban WWTPs belong to the scope of the UWWTD	17.7	14.0	24.7	14.0	15.3	8.4	3.7	3.4	85.1	70.0
Urban WWTPs outside the scope of the UWWTD	22.6	18.0	54.7	22.0	28.7	9.9	4.8	4.2	163.3	80.0
Results difference (%)	21.5	22.2	54.8	36.4	46.7	15.2	23.3	19.0	47.9	12.5

Figure 1 shows that effluent concentration is in compliance with the UWWTD requirements, except TP results, because the UWWTD does not provide TN and TP removal requirement for WWTPs with a pollution load less than 10,000 p.e. Figure 1 and Table 3 represent only two (Paide

and Järva-Jaani) WWTPs effluent results with a pollution load more than 10,000 p.e., because the remaining 44 treatment plants have pollution loads less than 10,000 p.e. According to the mean of the results, some WWTPs have problems in removing organic matter. Since the average concentration of organic matter is relatively small (16 mg/l), the removal of organic matter causes difficulties for a few WWTPs in Järva County. Comparing the average concentrations of the effluent of WWTPs that belong within the scope of the UWWTD and are added to the WWTPs results that do not serve the agglomeration, the average of the results differs up to 55%.

WWTPs Compliance Check to UWWTD and ENS Requirements

WWTPs compliance check is based on the monitoring results. In 2007, 53 samples were taken from 29 different WWTPs to check compliance with the UWWTD requirements. The compliance check results show that 11 urban WWTPs (37.9 % from all plants) were in conformity with the UWWTD requirements. Only 4 plants (Türi, Paide, Koeru and Järva-Jaani) of these 29 treatment plants were with pollution loads more than 2,000 p.e. One of the WWTPs (Koeru) with a pollution load more than 2,000 p.e was in conformity with the UWWTD requirements. The compliance check to ENS was based on the 71 sampling results, which were taken from the 45 different treatment plants. Some of these plants are not within the scope of UWWTD. 12 WWTPs (26.7%) were in compliance with the ENS requirements.

In 2008, 59 samples were taken from 28 different WWTPs to assess the compliance of Järva County WWTPs with the UWWTD requirements. 18 treatment plants (already 64.3%) were in compliance. Of the WWTPs with pollution loads more than 2,000 p.e., only Paide WWTP met the requirements. The compliance check to ENS was based on the 83 samples that were taken from 47 different WWTPs. 23 WWTPs (48.9%) were in compliance with the requirements.

Taking into account that many WWTPs in Järva County are not in compliance with the established requirements, the results of concentrations of TP, TN, and BOD₇ were also evaluated. Figures 2-4 show the concentrations variability, as well as 10- and 90-percentiles of the concentrations of BOD₇, TP, and TN.

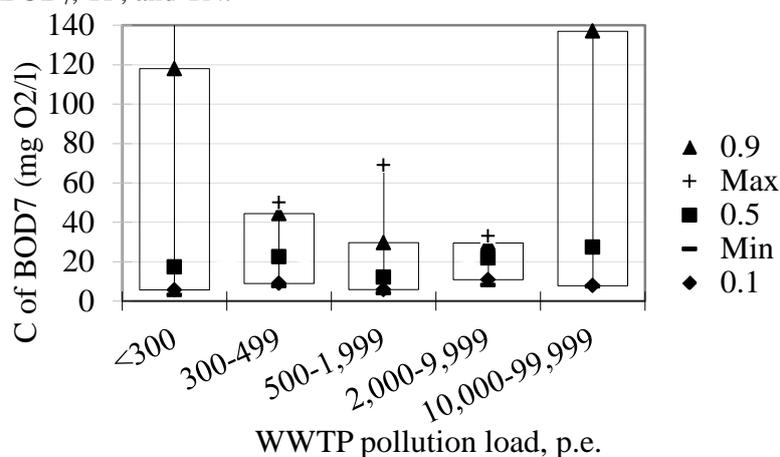


Figure 2. WWTP effluent variability of BOD₇, together with 10- and 90-percentiles in 2008

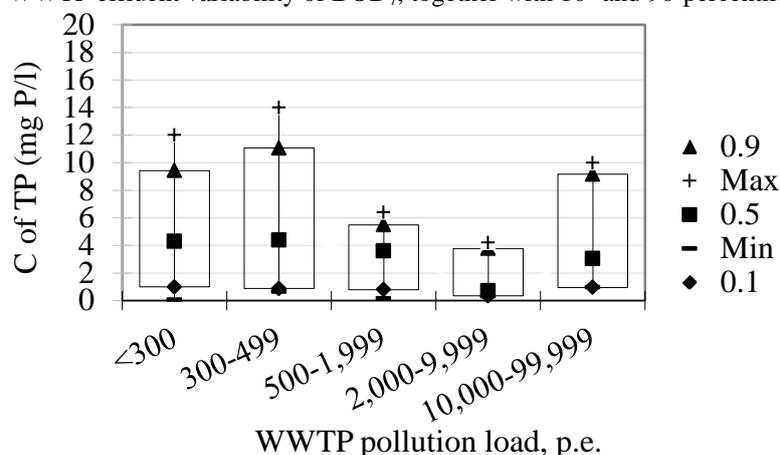


Figure 3. WWTP effluent variability of TP, together with 10- and 90-percentiles in 2008

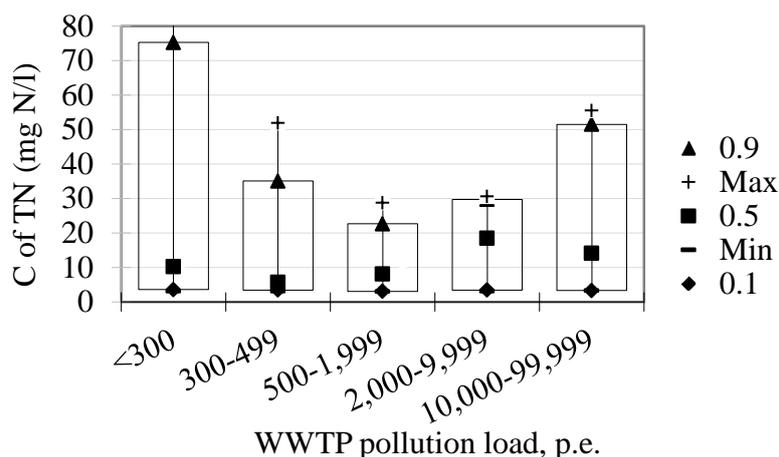


Figure 4. WWTP effluent variability of TN, together with 10- and 90-percentiles in 2008

As Figures 2-4 show, very high concentration variability exists in all WWTPs categories. We can conclude that even bigger WWTPs (more than 2,000 p.e.) have problems with the removal efficiency of all pollutants.

The results are based on the monitoring results made by Estonian Research Centre in accredited laboratories. The monitoring results obtained by WWTPs operators – the self-monitoring required according to the permits for the special use of water (Ministry of the Environment, 2009/2) – were not used in pollution load calculations, as Figure 5 indicates a significant difference in the self-monitoring and national monitoring results. The average monitoring results in 2008 made by WWTPs operators are even better; 65.2% of WWTPs are in compliance with the ENS requirements and 79.3% of WWTPs are in compliance with the UWWTD requirements using WWTPs operators self-monitoring results. Since annual average results instead of individual sample results were used, the difference between the methodologies of assessment of the compliance of WWTPs with the requirements is shown in Figure 5.

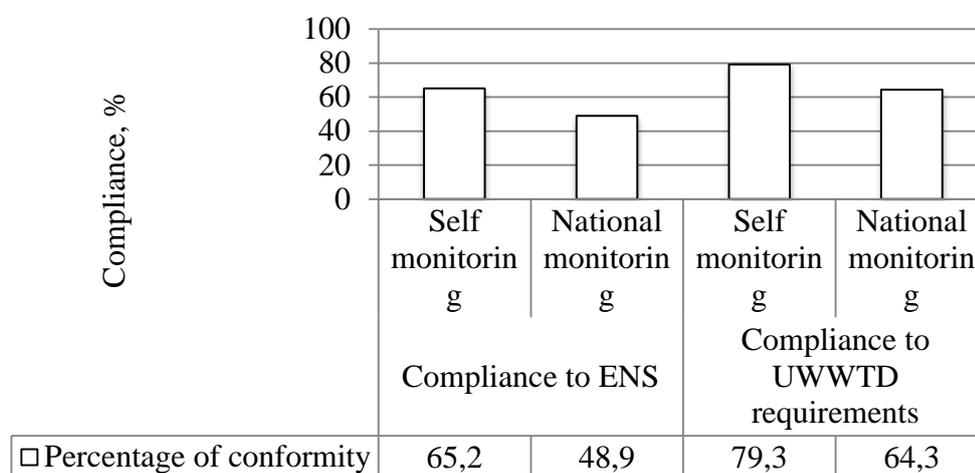


Figure 5. Percentage of compliance of WWTPs according to different methodologies, 2008

Problematic wastewater pollution indicators that do not conform to the established limit values

According to ENS, the marked part of pollution load must not discharge to the environment as shown in Figure 6. If all the WWTPs meet the requirements, the quantity of organic matter impact on water bodies would be reduce by at least 5.4 tons per year and TP 1.9 tons per year. The origin of the above normative pollution load is presented in Figure 6. An above normative pollution load discharged by WWTPs is shown with a pollution load more than 2,000 p.e. and less than 2,000 p.e., respectively.

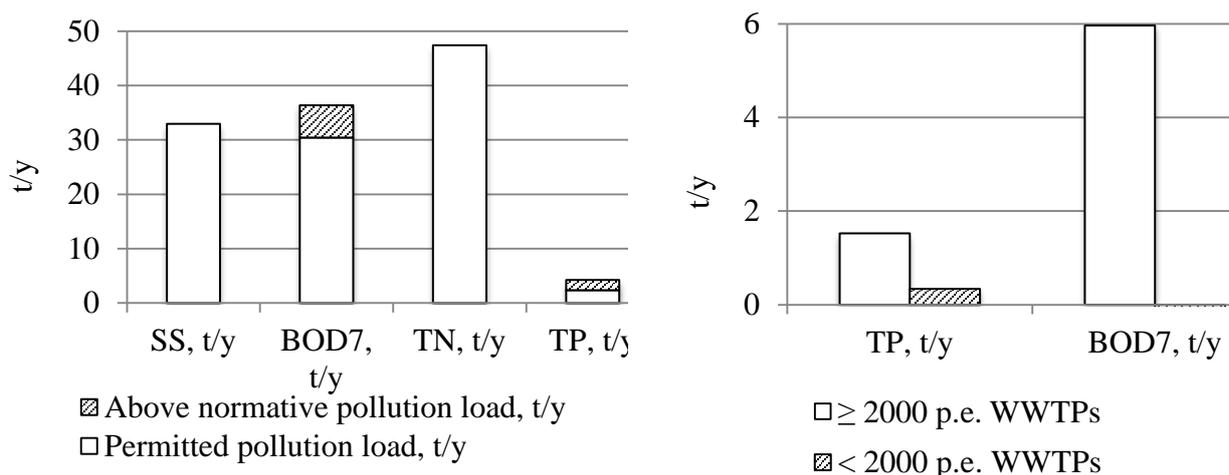


Figure 6. Total pollution load discharged to the receiving water bodies (left) and the origin of the above normative pollution load (right) in Järva County, 2008.

Figure 6 shows WWTPs with a pollution load of 2,000 p.e. or more overloading the receiving water bodies with an above normative TP pollution load over 4 times more than WWTPs with a pollution load less than 2,000 p.e. It should also be pointed out that only 4 WWTPs have a pollution load more than 2,000 p.e. in Järva County and 42 WWTPs (without Väätsa landfill) have a pollution load less than 2,000 p.e.

The excessive input of organic matter distribution between WWTPs with a pollution load more than 2,000 p.e. and less than 2,000 p.e. is worse than for an excessive TP load. An above normative organic matter pollution load comes entirely from WWTPs with a pollution load more than 2,000 p.e. The actual total quantity of organic matter from WWTPs with less than 2,000 p.e. is 0.59 tons lower per year than the amount permitted by the established standards. The discharge from WWTPs with a pollution load more than 2,000 p.e. has to be reduced to protect the sensitive water bodies in Estonia. There is a need to introduce third level high-grade treatment in WWTPs, which will improve treatment efficiency and result in an immediate impact (Pachel, 2010; Bryhn, 2009; Humborg et al, 2007).

Distribution and origin of WWTPs effluent pollution load discharged to the receiving water body

The distribution of pollution load is presented in terms of BOD₇, TP, and TN, which affect the receiving water body. Figure 7 shows that biodegradable organic matter derives 82% from WWTPs with a pollution load more than 2,000 p.e. and only 18% of the total pollution load comes from less than 2,000 p.e. WWTPs. TP and TN derive 76% from the WWTPs with a pollution load more than 2,000 p.e.

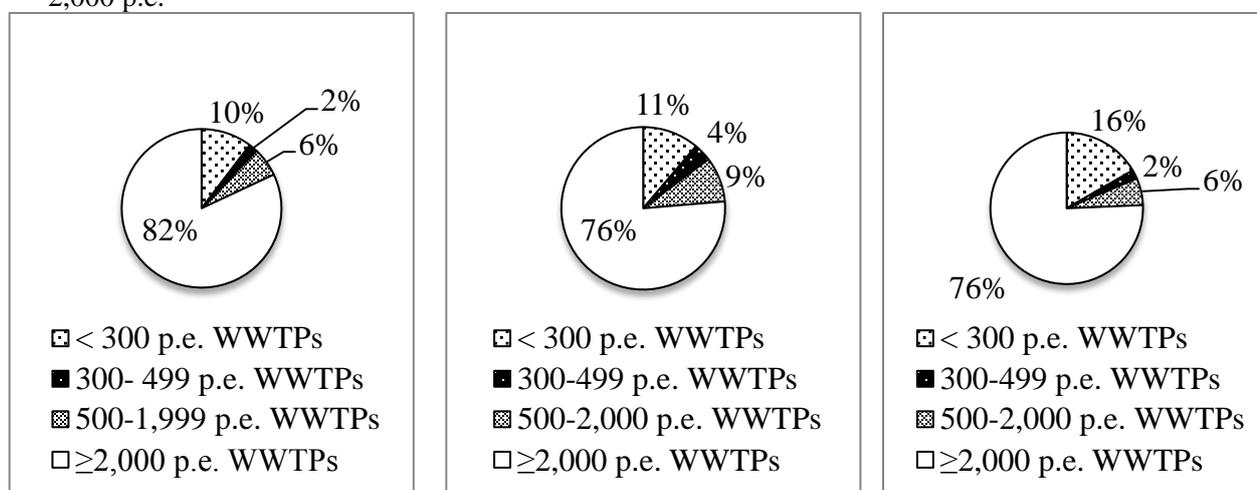


Figure 7. Distribution of BOD₇ (left), TP (centre) and TN pollution load in Järva County in 2008

Figure 7 suggests that to reduce the pollution load, which is discharged to the receiving water bodies, the discharge from WWTPs with more than 2,000 p.e. has to be reduced because WWTPs with a pollution load less than 2,000 p.e. contribute only ca 20% of the total biodegradable load discharged into the water bodies.

Discussion

The study analyses the impact of different sized WWTPs on the environment. Also, the study is unique as it was first time the cumulative impact of different sized WWTPs on the environment was analysed, using state-controlled monitoring results as well as self-monitoring results for comparison. The study reflects all 46 urban WWTPs results in Järva County and the results show that 42 WWTPs form only 18% of the entire BOD₇ pollution load and only 4 WWTPs with a pollution load of 2,000 p.e. or more form 82% of the entire BOD₇ pollution load. Similarly TP and TN pollution load distribution between WWTPs with a pollution load of less than 2,000 p.e. and more than 2,000 p.e. was 24% and 76%, respectively. The major sections of the WWTPs were constructed between the 1970s and the 1990s and have been in operation without any significant renovation; therefore, they are depreciated and outdated. Nutrient removal in these WWTPs is problematic. Eutrophication was not recognised as an important issue some decades ago. The main problem is relatively high phosphorus concentration in WWTP effluents. The high content of phosphorus causes problems in the recipient waterbodies, resulting in algae blooms and excessive plant growth. Figure 8 indicates phosphorus cumulative distribution; each point on the figure characterises a single WWTPs average result in 2007 based on 1-4 monitoring samples.

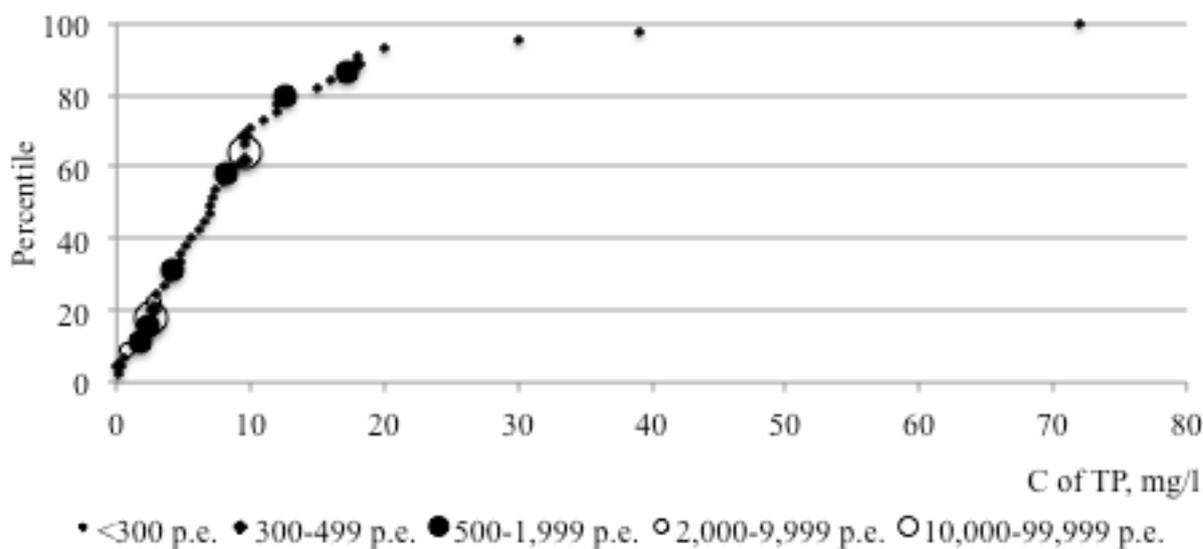


Figure 8. Percentile values of TP of different sized WWTPs in 2007

The 50-percentile value of TP for all types of WWTPs is 7.1 mgP/l; the outlet concentrations of only 5 WWTPs are below 2.0 mgP/l. To improve the ecological status of small rivers, it is necessary to improve TP values for WWTP effluents. Furthermore, existing national monitoring programmes, the data collection system, and data analyses need improvement in order to adequately assess treatment efficiency and the impact on the status of water bodies, especially in minimal runoff periods when water flow is low.

Järva County rivers have a small catchment area and, therefore, are relatively poor in water. This situation causes problems in using rivers as recipients for wastewater discharge because of insufficient dilution. Several studies point out that the majority of P tends to be retained within river systems during low-flow periods i.e. at times of greatest eutrophication risk (Jarvie et al., 2006, Pachel, 2010, Millier and Hooda, 2011, Reddy et al, 1999, Jarvie et al., 2006/2, Bukaveckas et al., 2005, Nemery and Garnier, 2007). According to the national wastewater monitoring programme, the measurements of water discharge in WWTP outlets and mixed river profile are not included at all. This is essential for calculating the actual wastewater load and its impact in low-flow rate periods. Therefore,

legislation ought to be improved so that the monitoring frequency could depend on the size of the WWTP, together with the condition of the WWTP and the status of the receiving water body.

Also during this study, the above normative pollution load to obtain comprehensive information regarding the status of WWTPs was analysed. This was done to get basic knowledge on whether the established wastewater treatment standards are in conformity with the required investments and usage cost of WWTPs while taking into account that the impact of WWTPs on the environment must be minimal. As the study results show, it is essential to bring into focus WWTPs with a pollution load of 2,000 p.e. and more, while these plants have the biggest impact on the environment. On the other hand, these WWTPs are not in conformity with the established wastewater treatment standards, which cause water bodies to be overloaded with TP and biodegradable organic matter pollution. Unfortunately, it is not possible to compare the study results with earlier results due to a lack of these kinds of studies. Furthermore, as the state-controlled WWTPs effluent monitoring results show, the monitoring results differ between national and self-monitoring results by as much as 25%; therefore, it is not suitable to mix these data. As a result, self-monitoring results do not reflect the actual WWTP situation due to the low reliability of the monitoring results. During this study, we found that it is crucial to change the WWTP effluent monitoring requirements and data collecting system. Otherwise, we do not have enough reliable information about the effluent impact of WWTPs on the environment.

Also, the study shows that it is not necessary to have strict treatment standards at present, as too many WWTPs are not in compliance with today's established standards and, therefore, cause water bodies to be overloaded with TP and biodegradable organic matter pollution as a result. In Järva County, bigger WWTPs are also unexpectedly not in compliance with the established standards.

There are now several activities being implemented to reduce the significant input of organic pollutants and nutrients into water courses in Estonia. In recent years, important efforts to reduce the phosphorus load have been put into the upgrading of existing WWTPs as well as the construction of new high-grade plants with phosphorus removal and also the renewal of existing sewers and the construction of new ones in order to connect more settlements to public WWTPs. Considering the fact that most of the Estonian WWTPs are in a renovation phase at the moment, it would be necessary to analyse the WWTP effluent contaminants and nutrient ratio following the renovation of these WWTPs.

Conclusions

In Estonia, much effort has been devoted to reducing the anthropogenic loads of TN and TP to water bodies, in terms of upgrading the existing WWTPs and building new high-grade plants with nitrogen and phosphorus removal. Nevertheless, the effluent impact of WWTPs on the water bodies is considerable. The UWWTD does not impose any specific effluent limit values for WWTPs with a pollution load less than 2,000 p.e., and therefore the criteria for a compliance check have been developed in this study. However, when assessing the compliance of WWTPs with the ENS, it must acknowledge that the national requirements are much stricter than the WWTPs can actually accomplish. Only half of the WWTPs are able to comply with the national requirements. Due to the total pollutant load and origin of above normative pollution load, it is imperative to address the priority of WWTPs with a pollution load of more than 2,000 p.e., since these plants represent more than 80% of the total WWTPs pollution load and these treatment plants also have the biggest impact on the above normative pollution load because of the non-conformity treatment level.

The study indicates that due to non-compliance, WWTPs will cause rivers to overload with nutrients. The reasons for non-compliance have been due to a lack of purification capacity and operational and maintenance problems. The effluent inlets of WWTPs have the biggest adverse effect in terms of TP content in receiving water bodies. Also, the study shows that to minimise the adverse effects of effluent from WWTPs, it is not necessary to establish stricter treatment requirements, because it was apparent that the treatment levels of many WWTPs were not in compliance with the validated requirements. At the same time, it is essential to make crucial changes in legislation to improve the national monitoring programmes and data collection system. As the study results show, the self-monitoring system of the operators and the national monitoring results are not reliable enough to assess the actual state of affairs of WWTPs and the WWTP effluent impact on the receiving water bodies. On the other hand, national monitoring frequency and scope are not representative enough to make comprehensive conclusions. Therefore, we suggest changing the existing monitoring system so that all monitoring goes under state-control as well as establishing more frequent national monitoring.

This will abolish the existing double-system, improve data reliability and give more adequate information about the status of water bodies.

In summary, the study shows that:

- 1) the biggest impact on water bodies are WWTPs that have a pollution load more than 2000 p.e.;
- 2) WWTPs less than 2000 p.e. have the highest impact in terms of TP and WWTPs with a pollution load of 2000 p.e. and more having the highest impact in terms of BOD₇ and TP. Therefore, WWTPs TP removal efficiency and established TP limit values need urgent re-evaluation;
- 3) WWTPs self-monitoring results differ from the national monitoring results; therefore, the monitoring system needs essential changes.

Furthermore, the authors found that it is essential to undertake more complex studies in Estonia in order to analyse the cumulative impact on environmental processes.

In summary, WWTPs with more than 2,000 p.e. should be regulated vigorously to guarantee their compliance with the requirements, because WWTPs with more than 2,000 p.e. count for 80% of the total pollution load in Järva County. The number of WWTPs with less than a 2,000 p.e. impact on the receiving water bodies is marginal. Measures also have to be adopted for WWTPs with less than 2,000 p.e. to comply with the ENS, since the extent of the impact of each individual WWTP depends on, among other things, the characteristics of the wastewater and water turbulence. Also, since a very small WWTP may affect the status of a water body depending on the characteristics of the wastewater and receiving water body, particularly careful handling should be applied to the up-streams of watercourses. Considering the fact that most of the Estonian WWTPs are in a renovation phase at the moment, it would be necessary to analyse the WWTPs effluent contaminants and nutrient ratio following the renovation of these WWTPs and also re-evaluate their conformity in 2014–2015. After that, it will be possible to analyse contiguous data that were discovered during this study conducted before the WWTPs renovation phase with data collected from 2014–2015 after the renovation of the WWTPs.

References:

- Bryhn AC. Sustainable Phosphorus Loadings from Effective and Cost-Effective Phosphorus Management Around the Baltic Sea. *PLoS ONE* 4(5): e5417. doi:10.1371/journal.pone.0005417, 2009.
- Bukaveckas, P.A., Guelda, D.L., Jack, J., Koch, R., Sellers, T. and Shostell, J. Effects of point source loadings, sub-basin inputs and longitudinal variation in material retention on C, N and P delivery from the Ohio River basin. *Ecosystems*, 8, 825–40, 2009.
- Dauknys, R., Vabolienė, G., Valentukevičienė, M., Rimeika, M. Influence of substrate on biological removal of phosphorus. *Ekologija*, 55, 220-225, 2009.
- Dickenson, E.R.V., Snyder, S.A., Sedlak, D.L., Drewes, J.E. Indicator compounds for assessment of wastewater effluent contributions to flow and water quality. *Water Research*, 45, 1199-1212, 2011.
- Environmental Report 7. Eesti jõgede ja järvede seisund. Water Pollution and Quality in Estonia. Environment Data Centre (EDC), National Board of Waters and Environment, Finland, Helsinki, 1993.
- European Community (EC). Directive of the European Council 91/676/EEC Concerning the protection of waters against pollution caused by nitrates from agricultural sources. *Official Journal of the European Communities* L 375, 1–8, 1991.
- European Community (EC). Directive of the European Council 91/271/EEC Concerning Urban Waste-Water Treatment. *Official Journal of the European Communities* L 135, 40–52, 1991.
- European Community (EC). Directive of the European Parliament and of the Council 2000/60/EC Establishing a Framework for Community Action in the Field of Water Policy. *Official Journal of the European Communities* L 327, 1-73, 2000.
- Government of Estonia. Regulation No 17, 21/01/2003 „Pandivere ja Adavere-Põltsamaa nitraaditundliku ala kaitse-eeskiri”. *RT I* 2003, 10, 49, 2003.
- Government of Estonia. Regulation No 269, 31 July 2001 “Heitvee veekogusse või pinnasesse juhtimise kord”. *RT I* 2001, 69, 424, 2001.
- Helcom. Helcom Baltic Sea Action Plan. HELCOM Ministerial Meeting. Krakow, Poland, 2007.

- Humborg, C., Mörth, C.-M., Sundbom, M., Wulf, F. Riverine transport of biogenic elements to the Baltic sea – past and possible future perspectives. *Hydrology and Earth System Sciences Discussions*, 4, 1095-1131, 2007.
- Iital, A., Brandt, N., Loigu, E., Kloga, M. Impact of changes in inputs to the water quality of the shallow Haapsalu Bay, the Baltic Sea. *Journal of Environmental Monitoring*, 8, 1531-1536, 2010.
- IMPRESS. Guidance for the Analysis of Pressures and Impacts in Accordance with the Water Framework Directive. Common Implementation Strategy Working Group 2.1, Office for Official Publications of the European Communities. Brussels, 2002.
- Jarvie, H.P., Neal, C., Paul J.A. Withers P.J.A. Sewage-effluent phosphorus: A greater risk to river eutrophication than agricultural phosphorus? *Science of the Total Environment*, 360, 246–253, 2006.
- Jarvie, H.P., Neal, C., Juergens, M.D., Sutton, E.J., Neal, M. and Wickham, H.D. Within-river nutrient processing in Chalk streams: The Pang and Lambourn, UK. *J. of Hydrol.*, 2, 330, 101–25, 2006.
- Kontas, A., Kucuksezgin, F., Altay, O., Uluturhan, E. Monitoring of eutrophication and nutrient limitation in the Izmir Bay (Turkey) before and after Wastewater Treatment Plant. *Environment International*, 29, 1057-1062, 2004.
- Millier, Helen K.G.R., Hooda, Peter S. Phosphorus species and fractionation –Why sewage derived phosphorus is a problem. *Journal of Environmental Management*, 1-5, 2011.
- Ministry of the Environment. Environmental Information Database EELIS. (visited 2009-05-01), 2009.
- Ministry of the Environment. Environmental permits information system at <http://klis.envir.ee/klis/>. (visited 2009-05-06), 2009.
- Ministry of Environment. Ida-Eesti ja Lääne-Eesti vesikonna veemajanduskavad. <http://www.envir.ee/vmk> (visited 2011-01-15), 2011.
- Némery, J. and Garnier, J. Origin and fate of phosphorus in the Seine watershed (France): agricultural and hydrographic P budgets. *J Geophys Res-Biogeosciences*, 112(G3) Article No. G03012, 2007.
- Pachel, K. Water Resources, Sustainable Use and Integrated Management in Estonia. 359 PhD Thesis, TUT Press, 20-33, 2010.
- Reddy, K.R., Kadlec, R.H., Flaig, E. and Gale, P.M. Phosphorus retention in streams and wetlands: a review. *Crit Rev Environ Sci Technol*, 29, 83–146, 1999.
- Vabolienė, G., Matuzevičius, A., Valentukevičienė, M. Effect of nitrogen on phosphate reduction in biological phosphorus removal from wastewater. *Ekologija*, 1. 80-88, 2007.
- Valdmaa, T., Rooma, L., Iital, A. An overview of the implementation of nitrate vulnerable zone action plan for the 2004-2008, the assessment of effectiveness of measures, the evaluation of monitoring programs. Tallinn, 2008.

RISK CAUSED BY THE SPATIAL FORMATION OF THE CITY OF BURSA (TURKEY) IN HISTORICAL PROCESS, CURRENT DEVELOPMENT PLANS AND LAND USAGES BASED ON DEVELOPMENT PLANS

Ufuk Fatih Kucukali, PhD
Istanbul Aydin University , Turkey

Abstract:

In Turkey, it is observed that there is an overgrowth of population, resolutions on land usage have deteriorating effects on natural ecological balance and consequently environmental problems reached threatening levels for all life forms. It is important to remove the impacts of urbanization, which consume irreplaceable natural resources, through planning methods and techniques. Examination of the current land usage of Bursa Nilüfer Creek Başköy – Kestel Section and its subbasins reveal that location organization was radically changed due to industrialization, and functional areas such as settlements and industrial areas, located on first class agricultural lands and on rock formation with geologically high permeability, create pollutions on the natural resources forming the basin, over acceptable parameters of international standards, as a result of governmental decisions.

In this context, resolutions, implementations and their reflections of land usage, population, residence facts and unplanned developments regulated according to development plans prepared in historical perspective and existing development plans in Bursa, the research area, have been discussed within a 20-year process; findings were evaluated and the type and dimension of ecological risks were assessed against international parameters.

Key Words: Basin, Ecological Planning, Ecological Risk, Land Use

Introduction

In order to conduct pollution and risk analysis of Bursa Nilüfer Stream Başköy – Kestel Section and sub-basins, decisions, practices brought by master plans drawn up in the historical perspective of Bursa and current zoning plans, and their reflection on the land use have to be examined.

Space Formation of the city of Bursa in historical process

Analyzing what kind of a development line Bursa city had in the historical process is important in terms of understanding formations in the city space of the 21st century Bursa. In the historical process, changes in the city of Bursa will be addressed in two different eras: Urban transformation in Bursa up until 1980s, and urban transformation in Bursa from 1980 to present day.

Urban Transformation in Bursa up until 1980s

When looking at Bursa in the era up until 1980s, it is observed that, the city had a prominent place in the Republic of Turkey in economical and social terms, as was the case during the Ottoman Empire era. Importance attached to the city during Republic era can be understood from the plans made for the city (Tekeli, 1999).

The first planning study for the city of Bursa was conducted by Karl Löcher in 1924. Such plan , influenced by ‘garden – city movement’ of the era, had an understanding that ignored the existing land uses, and did not play a role in the shaping of urban space of Bursa.

Henri Prost who conducted planning studies of Bursa between 1938 – 1944 adopted an understanding similar to axial planning understanding applied by Baron Haussmann in Paris. Foundations of current transportation separation in Bursa were laid in line with the plan by Prost. According to such understanding, it was planned to adapt the city to motor vehicle traffic, and the plan played an active role in determination of the 1st and 2nd stage road routes. Especially widening of Altıparmak Avenue and connection of Çekirge city center led an immense traffic jam and air pollution

risk attributable to transportation came into being along the road route. As a result of precipitation of such air pollution over the agricultural fields along the road routes, it is observed that risks of soil pollution and, thereby, water pollution came into being and pollution phenomenon are detected along these road routes as a result of the pollution analyses conducted today.

In 1960, under supervision of Luigi Piccinato, Bursa Master Plan with a scale of 1/5.000 was (Fig. 1) drawn up. Piccinato suggested that the city be developed along Ankara – Bursa – Mudanya road, that is, in a linear format on a east – west axis. In order to support such development, he planned to develop small scale industry along Ankara road which is the eastern edge of the city and suggested an Organized Industrial Zone be established along Mudanya road (Dostoglu and Vural, 2004). In line with suggestion by Piccinato, Turkey's first Organized Industrial Zone was opened in 1966 within Bursa.

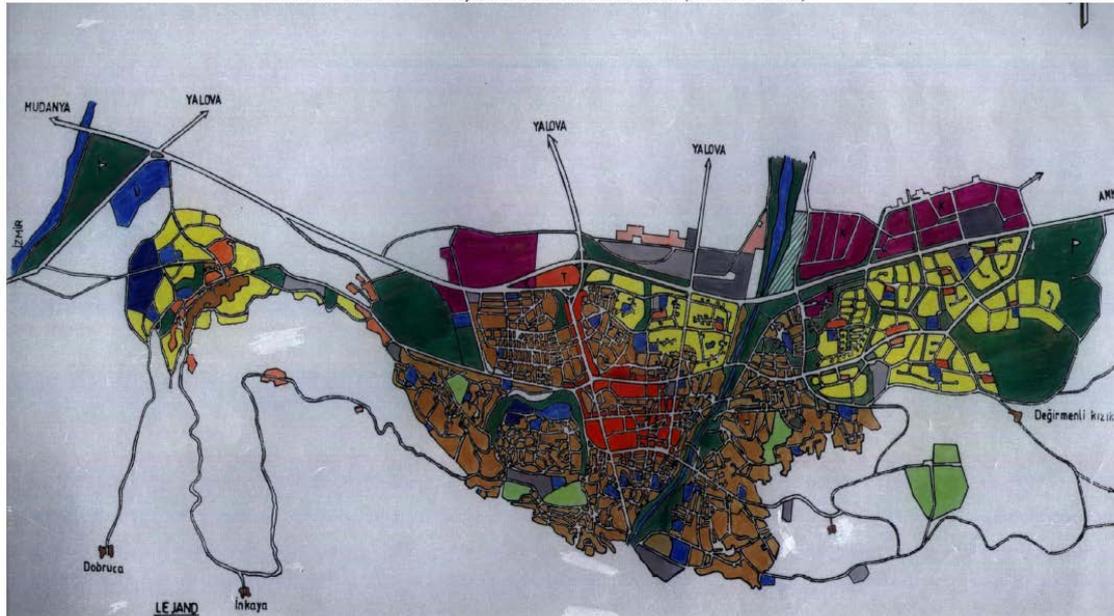


Fig. 1 Bursa Master Plan with a scale of 1/5.000 dated 1960 (Piccinato Plan)

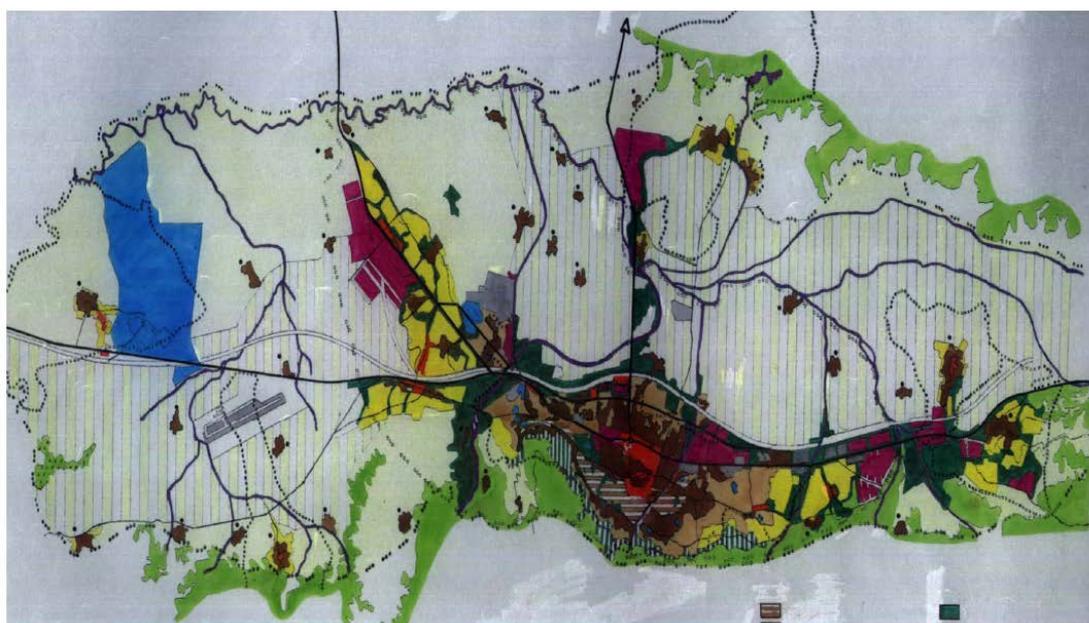
Owing to influence created by establishment of Organize Industrial Zone, immense pressure caused by migration flowing into the city neutralized the population and density projections of Piccinato plan. Population target of Piccinato Plan was 250.000 in 1980, yet population of city had already reached 350.000 in 1975 (Arslanoğlu, 2004).

This affected physical structure of the city. New factories have been established in Mudanya Avenue and city has developed in the western direction. Large scale industrial establishments have heralded the western direction of city's development. New housing settlements have been formed in the region in connection with Renault, Sönmez Filament and Nergis plants being located along Mudanya Avenue. Besides, Tofaş located along Yalova Highway, Karsan located along Izmir highway, and Yeşim located along Ankara highway have been built along main roads outside the city space. In parallel to diversification and increase of city's industry in the 1970s, shanty settlements started to appear in Bursa Plain to the north and the east. Back in that era, 'Plain Conservation Protocol' dated 1977 was drawn up, which determined those portions of Bursa Plain that will be conserved, yet this protocol could not stop shanty settlements and industrial establishments from spreading all over the plain (Kaplanoglu and Cengiz, 2005).

In Bursa, Master Plan Bureau prepared many master plans and revision plans after 1960 (Table 1). Master Plan dated 1976 was a plan that tried to decentralize the city (Fig. 2). Almost all of the plan suggestions remained outside the border of municipality This plan also addressed Demirtaş, Kestel, Gürsu, Görükle, Çalı, Kayapa, Hasanağa and Akçalar Municipalities which are within vicinity of Bursa. Macro form and development directions pertaining to Bursa city were determined by the plan dated 1976 (Albatan, 1999).

Table 1. Master plans and revision plans prepared in Bursa after 1960

Plan Name	Scale	Preparation Date	Institution
Master Plan (Piccinato Plan)	1/5.000, 1/10.000	1960	Ministry of Development and Housing
Master Plan	1/25.000	1976	Bursa Master Plan Office
Master Plan	1/5.000	1984	Bursa Master Plan Office
Master Plan (Revision)	1/25.000, 1/5.000	1990	Bursa Metropolitan Municipality
Coastal Zone Environmental Improvement Plan	1/25.000	1990	Bursa Metropolitan Municipality
Master Plan (Revision)	1/25.000, 1/1.000, 1/5.000	1995	Bursa Metropolitan Municipality
West Planning Region Master Plan	1/25.000	1997	Bursa Metropolitan Municipality
Kayapa Mass Housing Area Master Plan	1/5.000	1997	Bursa Metropolitan Municipality
Environmental Master Plan	1/100.000	1997-1998	Provincial Directorate of Public Works and Bursa Metropolitan Municipality
Center Planning Region Master Plan	1/25.000	1998	Bursa Metropolitan Municipality

**Fig. 2** Master Plan dated 1976

Urban Transformation in Bursa from 1980 to the present day

Previously prepared plans were no longer applicable and plans had to be renewed because of the settlement areas required due to increasing city population and shanty settlements as from 1980. 1984 Master Plan prepared as a result aimed to stop the city from extending in plain area to the north of the city. According to such plan, new plain conservation areas have been established and plain conservation protocols have been drawn up. Also, plan studies were intended to cover İnegöl, Mustafa Kemalpaşa, Karacabey, Mudanya, Gemlik, Yenişehir counties and a 'Bursa Metropolitan Area' was defined. A Bursa City Whole concept was created which included Görükle, Kayapa, Hasanağa, Çalı and Akçalar, to the west, Gürsu and Kestel, and Demirtaş municipalities to the south which entered into a process of integration with the city center. Within framework of this plan, Fethiye, Beşevler, İhsaniye Ataevler to the west and Kaplıkaya to the east were zoned for housing settlements aimed at low or medium income families (Akkılıç, 2002). Since these settlement areas have been planned without proper infrastructure and over the fertile agricultural fields, agricultural fields have been lost,

risks such as pollution of underground waters due to infiltration according to parent rock characteristics and discharging of domestic pollutions into Stream Nilüfer through surface flow have appeared. According to planning decisions made in 1984, risk of water pollution that has appeared has been confirmed by the evaluation of current pollution analyses.

As from 1976, master plans and plan revisions have been made almost every 5 years and urban extension has been tried to be kept under control by opening new areas. In 1990, Bursa Metropolitan Municipality revised Master Plan (1/25.000 and 1/5000) and an additional housing area that is 1970 hectare large was suggested in the west of the city. Owing to such area, in addition to population of 1976 plan, an area that would house additional 240.000 people was opened and housing areas that would be enough for a population of 1.2 million were suggested in the whole city. Having made another Master Plan Revision in 1995, an additional housing area of 1500 hectare and immense housing areas to house a population as large as 420.000 have been suggested. Likewise, agricultural and forest areas have been lost in these areas as well. This revision has paved the way for inaccurate land use over those areas with liquefaction problem in geological terms. Under the Zoning Law dated 1985 (Law on Zoning numbered 3194), neighboring municipalities were authorized to prepare and approve their own zoning plans. Thus, apart from Bursa Master Plans, zoning plans were prepared by municipalities such as Demirtaş, Görükle, Kayapa, Kestel and new areas have been opened (Tosun, 2007). These plans drawn up under the Law on Zoning within the borders of municipality in a way that does not conform to borders of catchment basin indicate that planning was done from a fragmental point of view. Nonconformities between those plans drawn up and current land use demonstrate that plans were totally null and void or remained in force for a very short period of time. Therefore, it has been deemed necessary to prepare Environmental Plan where strategies and planning decisions are made at the macro scale.

In 1997, in collaboration with Provincial Directorate of Public Works and Settlement, “Province of Bursa 2020 Environmental Strategy Plan” with a scale of 1/100.000 was prepared and approved by the ministry in 1998. Target of Bursa 2020 Environmental Strategy Plan with a scale of 1/100.000 is creating an environment that is sustainable and inhabitable until 2020, establishing industrial development targets, planning principles of Bursa within framework of development plan of Turkey and in line with retaining historical identity of Bursa, and ensuring that City extends and develops in a healthy way in line with these principles (Bursa Metropolitan Municipality, 1998). Corer planning principle of Bursa 2020 Environmental Strategy Plan with a scale of 1/100.000 in line with the planning target is not making Bursa merely an industrial or agricultural city as a result of conservation and development principles to surmount one another. Bursa Metropolitan Master Plan Bureau, in accordance with 2020 Bursa Environmental Strategy Plan with a scale of 1/100.000 as approved in 1998, generated and approved Central Region Master Plan with a scale of 1/25.000. In 2004, boundaries of Bursa metropolitan area were expanded by 30 kilometers in radius. New boundaries of Bursa metropolitan area include 3 central counties (Osmangazi, Yıldırım, Nilüfer), 4 counties (Mudanya, Gemlik, Gürsu and Kestel) and 17 first phase municipalities (Akçalar, Hasanağa, Kayapa, Çalı, Göknükbelen, Kirazlı, Barakfaki, Görükle, Emek, Demirtaş, Ovaakça, Güzelyalı, Kurşunlu and Umurbey). Neighboring area of Bursa Metropolitan Municipality is 300.000 hectare large and its population has increased to 1.528.720 (Bursa Metropolitan Municipality, 1997).

Current Zoning Plans

The last time when boundaries of Bursa Metropolitan Municipality were established was when it was done so under law numbered 5216 and boundaries reached their most extensive form. Boundaries of the area defined under the law as a circle with a radius of 30 km covers an area of around 300 square kilometers. In the area covering the larger province of Bursa, “Bursa 2020 Environmental Plan with a scale of 1/100 000” approved by Ministry of Public Works and Settlement applies. Some definitions have been put forward by taking into account development tendencies within framework of Environmental Plan. One of these definitions is “Bursa Metropolitan Area.” Within scope of Bursa Metropolitan Area, seven sub-planning areas have been defined. These include (Fig. 3) Central Area Plan, Western Area Plan, Mudanya Area Plan, North Area Plan, Gemlik Area Plan, East Area Plan, Alaçam (Uludağ) Area Plan.

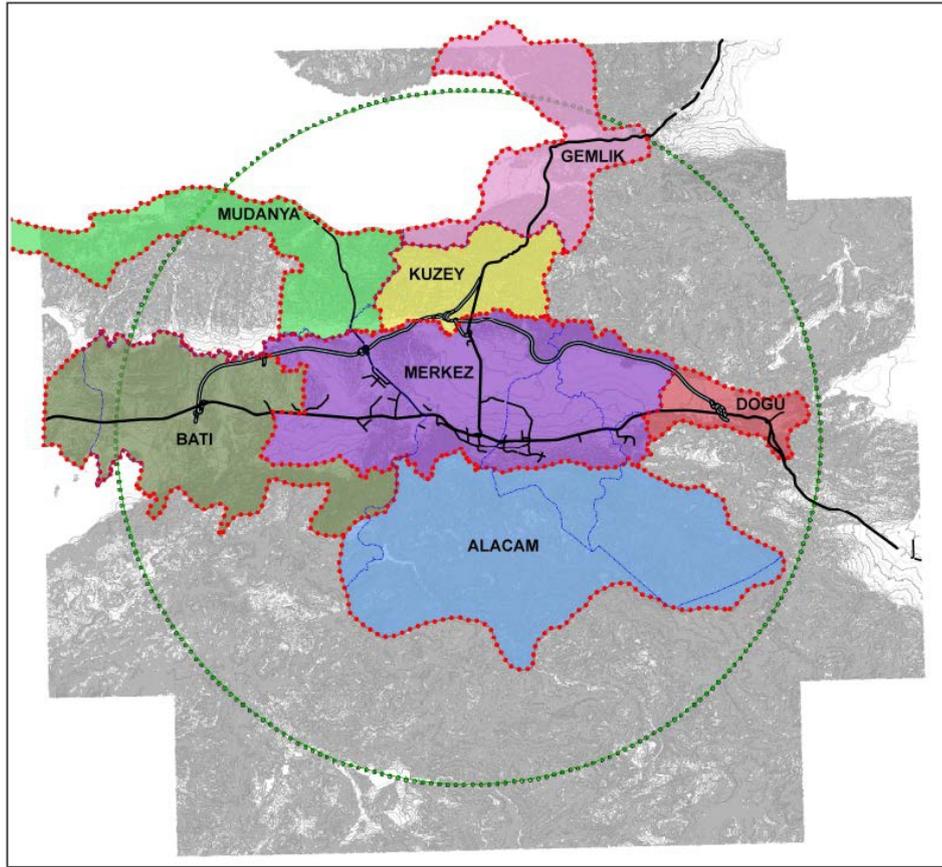


Fig. 3 Bursa Metropolitan Area planning areas

While establishing planning areas of Master Plan with a scale of 1/25.000, only managerial borders and development tendencies have been taken into consideration. Basin based planning where physical structure, natural resource components and ecological relations are relied upon has not been guiding / decisive in respect of the planning studies conducted in Bursa.

Decisions and practices brought by the existing zoning plans to the study area have to be evaluated together with positive and negative aspects thereof . Therefore, planning areas that are included in the study area have to be assessed. The planning areas include: Central Area Plan, Western Area Plan, Eastern Area Plan and a portion of Alaçam (Uludağ) Area Plan.

Bursa Metropolitan Municipality boundaries cover the planning areas defined in the upper scale plan. These planning areas correspond to a portion equal to 170 square kilometers of boundaries of Bursa Metropolitan Municipality that is as large as 300 square kilometers. Planning areas are city spaces established by certain transportation axes and areas which comprise the macro form of these city spaces. Areas beyond such 170 square kilometer large area are predominantly vast forest areas and areas used for agriculture.

Central Area Plan

Boundaries of Metropolitan Municipality and Gürsu, Kestel and Demirtaş Municipalities are defined as land registration boundaries of the villages Adaköy, Hasanköy, Cambazlar, Ağaköy, Kumlukalan, Doğanköy, and Yolçatı and planning area covering Çayırköy Plain.

Target of Central Area Master Plan with a scale of 1/25.000 is, by adopting Bursa 2020 Environmental Plan with a scale of 1/100.000 as target, within the seven planning areas established inside the boundaries of Metropolitan Municipality and neighboring areas, is ‘creating sustainable, inhabitable environment and retaining agricultural, touristic, historical identity, assuring conservation – utilization balance and realizing healthy urban development and extension targets in line with the industrial development targets within scope of development policies of Turkey (Fidan and others, 2005).

It is seen that main criteria guiding the planning in the plan report are composed of the following titles;

- Population structure – projections
- Existing land use criteria and existing formation
- Existing zoning plans (conservation plans, registered plans, plan amendments etc.)
- Structure of property
- Opinions and letters of official authorities
- General structure, morphology and geological structure of the land
- Climate and the other data

As can be derived from the criteria which guide the planning, existing land use criteria, social-economical data and development tendencies guide the planning as the main criteria, only the morphology, geology and climate status is assessed within natural structure characteristics and they are evaluated as secondary criteria which offer data input for the planning. In the planning understanding with basin approach, all natural structure components have to be weighted according to degree of inter-relation and guide the planning. It can be understood from this phenomenon and existing planning understanding in Bursa, risks faced by the natural resources in the study area will continue to grow.

Population of Bursa city center had an overall regular increases between 1935 and 1960. The fact that high rises in the population took place after 1960 can be attributed to migration received by the city due to industrial investments. Over the years, when looking into population movements in Bursa, it is observed that population of Bursa city center doubles every ten years as of 1960s (DİE, 1993; DİE, 2002). Changes seen in population structure after 1960 predominantly reflect changes in economic structure of Bursa.

By having a look at industrial development of the recent years in Bursa, it is observed that employment capacity of agriculture has stabilized, there is a great increase in manufacturing industry, yet employment capacity of the services industry grow a lot faster than the other industries. Despite rapid development and growth in industry, employment does not grow at the same pace due to advanced technology or the other factors. In terms of growth rate, industry with the fastest growth rate is wood, forest and furniture industry with 26,8%. The second fastest growth rate is enjoyed by electrics and electronics industry with 20,2%, and transportation and forwarding is at the third place with 17,7%.

Given growth rates of industries, it is obvious that there is a causality relation between wood, forestry products and furniture industry, the fastest growing industry, and losses in the forests of the study area. Likewise, if relation between transportation and forwarding industry, the third fastest growing industry, and pollution measurements conducted along the road routes of the study area is evaluated; it is observed that air, soil and water pollutions tend to increase in line with the transportation.

Conclusion

When zoning plans shaping the land use within study area and drawn up for planning areas within “Bursa Metropolitan Area” are evaluated, they do not seem to overlap the current land use. Also, land use tendencies that give rise to risks which are subject of thesis can be seen in the plans and plan notes.

Looking into distribution of land use shown on zoning plans according to sub-basins (Table 2); the highest value for areas to be forested seems to be present in Ayvalı Branch with 4,03 square kilometers. Accordingly, region park decision has been made for sub-basin of Ayvalı Branch which has a 1,39 square kilometer large area. Likewise, Ayvalı Branch Sub-Basin is the largest sub-basin with a park area that is 2,76 square kilometers large areas that are defined as areas to be protected as agricultural field with a 33,95 square kilometer large area according to plan decisions.

Table 2. Area distribution of land uses of Bursa Metropolitan Area Zoning Plans according to sub-basins of Stream Nilüfer catchment Basin

LAND USE (km2)	Sub-basins										TOTAL (km2)	
	Ayvalı	Cilimboz	Demirtaş	Gökdere	Hacivat	Hasanağa	Nilüfer	Üçpınar	Y.Karaağaç	Other Basins		
Military Airport											1,383	1,383
Afforestation Area	4,031	0,011	0,642	0,533	0,455	1,274	1,431	0,060	0,139	5,123	13,699	
Forest	30,328	0,565	11,361	0,226	1,836	7,374	8,988	0,143	1,915	45,567	108,303	
Purification	0,448							0,080		0,306	0,834	
Military Area	0,181			0,038				0,214		0,115	0,548	
Regional Park	1,390		0,144	0,114						2,594	4,242	
Framework										0,214	0,214	
University	4,998									6,333	11,331	
Health										0,025	0,025	
Social	1,788	0,185		0,410	0,723	0,267	0,372		0,016	2,429	6,190	
Sports	1,039			0,190	0,020	0,043	0,218			5,547	2,057	
Technical Infrastructure	0,270	0,101		0,026	0,036		0,297			1,433	2,163	
Stream	0,046					0,262		0,002	0,001	0,577	0,888	
Fair										0,210	0,210	
Storage	0,146								0,219	0,080	0,445	
Urban Study Areas	1,563		0,031	1,814	0,170					2,094	5,672	
Rural Study Areas					0,038	1,473	0,281		1,093	2,403	5,288	
Small Industrial Sites	1,648			1,105	0,158	0,971	0,271		0,527	1,345	6,025	
Organized Industrial Zone	4,306		0,032			0,036	0,281		0,737	9,392	14,784	
Industry	2,518					1,781	0,280		0,065	1,662	6,306	
2. and 3 Degree of Commerce	1,480				0,207	0,439	0,281			0,255	2,662	
Wholesale Trade Areas										0,194	0,194	
Trade	1,410	0,434	0,027	1,064	0,505		0,795			4,011	8,246	
Central Business Areas		0,716		1,793						0,039	2,548	
Very Rare Density Residential Development Areas	3,892		0,400	0,063	0,766	0,993	0,664		0,044	2,739	9,561	
Rare Density Residential Development Areas	0,145		0,272				1,157			8,651	10,225	
Medium Density Residential Development Areas	8,290	0,067	0,004	0,870	0,409	6,874	0,576			7,961	25,051	
Development of High Density Residential Areas	0,152		0,628	0,047	0,273	0,153				0,628	1,881	
Very Rare Density Residential Housing Areas	1,195	0,783	0,332	1,742	0,810		1,189			5,873	11,924	
Rare Density Residential Housing Areas										3,276	3,276	
Residential Medium Density Residential Areas	5,190	2,760	1,559	4,840	6,214	1,173	1,935			14,815	38,486	
Residential High Density Residential Areas	4,012	0,621	0,345	2,176	1,459		1,679			3,375	13,667	
Development Rural Settlement Areas	0,521		0,075		0,093	0,583	0,031		0,184	3,643	5,130	
Residential Rural Settlement Areas	1,701		0,002		0,213	0,370	0,183		0,102	3,958	6,529	
Cemetery	0,731			0,039						0,890	1,660	
Specific Project Areas							0,099			1,207	1,306	
Park	2,769	0,804	0,139	0,956	0,942	2,294	1,066		0,826	7,250	17,046	
Recreation	0,119	0,375		0,253		0,398	0,065			3,050	4,260	
Specialty Crop Areas			0,799			1,279	1,076			2,855	6,009	
Other Agricultural Areas	14,968		6,283			3,000	3,095	0,001	6,517	54,806	88,670	
Quality of Agricultural Areas	33,947		6,578	0,960	6,487	7,530	0,011	0,001	4,374	105,654	165,542	
Liquidation										0,963	0,963	
Technopark	0,043									0,540	0,583	
Terminal										0,208	0,208	
Daily Tourism	0,007						0,173			0,467	0,647	
Tourism		0,082					0,220			0,255	0,557	
Path Protection Strip	0,322		0,259			0,030			0,334	2,377	3,322	

Besides, Ayvalı Branch Sub-Basin has been planned in a manner that it will cover the largest area in the study area with a total surface area of 10,19 square kilometers for industrial activities according to zoning plans in terms of distribution of storage areas, urban work areas, non-urban work areas, small scale industrial sites, organized industrial sites and industrial areas. Likewise, Ayvalı Branch Sub-Basin has been planned in a manner that it will cover the largest area in the study area with a total area of 25.10 square kilometers in terms of very scattered, scattered, medium and high density housing development areas, very scattered, scattered, medium and high density residential housing areas as well as development and residential rural settlement areas. When plan decisions related to study area are evaluated in terms of natural structure features, it is seen that Ayvalı Branch Sub-Basin is composed of alluvia, lime stone, metamorphic rocks, travertine and hillside debris in terms of lithologic formation. In Ayvalı Branch Sub-Basin which is dominated by these rocks with large clearance, soil and underground water pollution risk is high in terms of lithologic structure. In terms of soil groups, the soil groups with the highest value in Ayvalı Branch Sub-Basin are brown forest soils and vertisol. These soils have a clay structure, and soil and underground water pollution risk is low. Besides, pollution discharge is likely to reach Ayvalı Creek through surface flow. From perspective of decisions made in the zoning plans for these sub-basins, although area to be forested, regional park and park plan decisions seem to be correct from lithology perspective, caution should be used for industrial activities and residential areas which may give rise to underground water pollution

and soil pollution risk. Since sub-basin is composed of clay soil groups in terms of soil groups, precautions which would prevent pollutions carried by surface water from reaching the surface water have to be taken.

Based on consideration of land uses derived from satellite footage, 5 land uses have been evaluated in details (residential, agriculture, forest, industry and trade) and modifications taking place in the years 1989-1999-2009 have been evaluated. Based on the land uses laid down in the zoning plans, these 5 land uses have to be evaluated according to use and compared to current land use. At this point, area distribution of residential, forest, agricultural, industrial and commercial land uses is as follows:

When comparing the land uses of zoning plans to current land use, since all of the forest and agriculture areas inside the study area have not been planned, it is not possible to make a comparison for forest and agricultural fields in case of zoning plan – current land use comparison due to lack of data. Besides, comparisons made for housing, industry and trade (Table 3) demonstrate that zoning plans support the increasing trend in the current land uses.

According to current land use, while there is a housing area of 118,08 square kilometers, zoning plan stipulates a housing area of 125,73 square kilometers. As a result, zoning planners support a housing area of 13,65 square kilometers and increase of current housing areas by 12,18%. Looking from perspective of sub – basins, at some sub-basins (Cilimboz, Gökdere, Hacivat, Üçpınar, Y. Karaağaç branches) housing areas are in excess of values stipulated in the zoning plans . Production of houses is encouraged at the other sub-basins by the zoning plans.

According to current land use, while there is an industrial area of 22,54 square kilometers in study area, zoning plan stipulates an industrial area of 37,53 square kilometers. As a result, zoning planners support an industrial area of 14,99 square kilometers and increase of current industrial areas by 66,50%. When looking from perspective of sub-basins, while it is observed that industrial activities intensify at Ayvalı branch, zoning planners believe that industrial investments have to be made along Hasanağa Branch and Demirtaş Branch.

Table 3. Area distribution of housing, agriculture, forest, industry and commercial land uses of Bursa Metropolitan Area Zoning Plans according to sub-basins of Stream Nilüfer catchment Basin

	LAND USE (km2)	Sub-basins										TOTAL (km2)
		Ayvalı	Cilimboz	Demirtaş	Gökdere	Hacivat	Hasanağa	Nilüfer	Üçpınar	Y.Karaağaç	Other Basins	
Housing	Very Rare Density Residential Development Areas	3,89		0,40	0,06	0,77	0,99	0,66		0,04	2,74	125,73
	Rare Density Residential Development Areas	0,15		0,27				1,16			8,65	
	Medium Density Residential Development Areas	8,29	0,07	0,00	0,87	0,41	6,87	0,58			7,96	
	Development of High Density Residential Areas	0,15		0,63	0,05	0,27	0,15				0,63	
	Very Rare Density Residential Housing Areas	1,20	0,78	0,33	1,74	0,81		1,19			5,87	
	Rare Density Residential Housing Areas										3,28	
	Residential Medium Density Residential Areas	5,19	2,76	1,56	4,84	6,21	1,17	1,93			14,81	
	Residential High Density Residential Areas	4,01	0,62	0,34	2,18	1,46		1,68			3,38	
	Development Rural Settlement Areas	0,52		0,07		0,09	0,58	0,03		0,18	3,64	
	Residential Rural Settlement Areas	1,70		0,00		0,21	0,37	0,18		0,10	3,96	
Agriculture	Special Crop Areas			0,80			1,28	1,08			2,85	260,22
	Other Agricultural Areas	14,97		6,28			3,00	3,10	0,00	6,52	54,81	
	Quality of Agricultural Areas	33,95		6,58	0,96	6,49	7,53	0,01	0,00	4,37	105,65	
Forest	Afforestation Area	4,03	0,01	0,64	0,53	0,46	1,27	1,43	0,06	0,14	5,12	122,01
	Forest	30,33	0,57	11,36	0,23	1,84	7,37	8,99	0,14	1,91	45,57	
Industry	Storage	0,15								0,22	0,08	37,53
	Urban Areas	1,56		0,03	1,81	0,17					2,09	
	Urban Study Areas					0,04	1,47	0,28		1,09	2,40	
	Small Industrial Sites	1,65			0,11	0,16	0,97	0,28		0,53	1,35	
	Organized Industrial Zone	4,31		0,03			0,04	0,28		0,74	9,39	
	Industry	2,52					1,78	0,28		0,07	1,66	
Trade	2. and 3 Degree of Commerce	1,48				0,21	0,44	0,28			0,26	13,66
	Wholesale Trade Areas										0,19	
	Trade	1,41	0,43	0,03	1,06	0,51		0,79			4,01	
	Central Business Areas		0,72		1,79						0,04	

According to current land use, while there is a commercial area of 3,59 square kilometers in study area, zoning plan stipulates a commercial area of 13,66 square kilometers. As a result, zoning planners support a commercial area of 10,07 square kilometers and increase of current commercial areas by 280,50%.

When comparing Bursa Metropolitan Alanı zoning plans and current land use in terms of sub-basins (Table 4); for Ayvalı Branch Sub-Basin, an increase at housing areas by 7,51 square kilometers (42,69%), at industrial areas by 2,07 square kilometers (25,52%) and at commercial areas by 2,89 square kilometers. In Cilimboz Branch Sub-Branch, it is observed that housing areas of the current land use have exceeded housing areas shown on the zoning plans by 2,24 square kilometers (34,62%) and an increase by 0,58 square kilometers is planned for commercial areas. At Demirtaş Branch Sub-Basin, for housing areas, an increase by 1,63 square kilometers (81,91%), for industrial areas an increase by 0,01 square kilometers and for commercial areas an increase by 0,03 square kilometers is foreseen.

Table 4. Comparison of Bursa Metropolitan Area Zoning Plans and current land use

	LAND USE (km2)	Sub-basins										TOTAL (km2)
		Ayvalı	Cilimboz	Demirtaş	Gökdere	Hacivat	Hasanağa	Nilüfer	Üçpınar	Y.Karaağaç	Other Basins	
Housing	According to the Development Plan	25,10	4,23	3,62	9,74	10,24	10,15	7,41	0	0,33	54,92	125,73
	Current Status	17,59	6,47	1,99	14,47	14,32	2,75	7,31	0,09	0,77	46,32	112,08
	Diffirence	7,51	-2,24	1,63	-4,73	-4,08	7,40	0,10	-0,09	-0,44	8,60	13,65
	Exchange Rate	42,69%	-34,62%	81,91%	-32,69%	-28,49%	269,09%	1,37%	-100,00%	-57,14%	18,57%	12,18%
Industry	According to the Development Plan	10,18	0	0,06	1,92	0,37	4,26	1,12	0	2,64	16,98	37,53
	Current Status	8,11	0	0,01	0,45	0,16	0,49	0	0	1,53	11,79	22,54
	Diffirence	2,07	0	0,05	1,47	0,21	3,77	1,12	0	1,11	5,19	14,99
	Exchange Rate	25,52%	0,00%	500,00%	326,67%	131,25%	769,39%	100,00%	0,00%	72,55%	44,02%	66,50%
Trade	According to the Development Plan	2,89	1,15	0,03	2,86	0,71	0,44	1,08	0	0	4,50	13,66
	Current Status	0	0,57	0	0,98	0,68	0	0,43	0	0	0,93	3,59
	Diffirence	2,89	0,58	0,03	1,88	0,03	0,44	0,65	0	0	3,57	10,07
	Exchange Rate	100,00%	101,75%	100,00%	191,84%	4,41%	100,00%	151,16%	0,00%	0,00%	383,87%	280,50%

(based on Bursa Metropolitan Area zoning plans and Ikonos 2009 satellite footage)

In Gökdere Branch Sub-Branch, it is observed that housing areas of the current land use have exceeded housing areas shown on the zoning plans by 4,73 square kilometers (32,69%) and an increase by 1,88 square kilometers is planned for commercial areas. In Hacivat Branch Sub-Branch, it is observed that housing areas of the current land use have exceeded housing areas shown on the zoning plans by 4,08 square kilometers (28,49%) and an increase by 0,03 square kilometers is planned for commercial areas. At Hasanağa Branch Sub-Basin, for housing areas, an increase by 7,40 square kilometers, for industrial areas an increase by 3,77 square kilometers and for commercial areas an increase by 0,44 square kilometers is foreseen. At Nilüfer Branch Sub-Basin, for housing areas, an increase by 0,10 square kilometers (1,37%), for industrial areas an increase by 1,12 square kilometers and for commercial areas an increase by 0,65 square kilometers is foreseen. At Üçpınar Branch Sub-Basin, it is observed that housing areas in the current land use have exceeded housing areas shown on the zoning plans by 0,09 square kilometers.

In Y. Karaağaç Branch Sub-Branch, it is observed that housing areas of the current land use have exceeded housing areas shown on the zoning plans by 0,44 (57,14%) square kilometers and an increase by 1,11 square kilometers is planned for industrial areas (Küçükali, 2012).

When the relationship between these change and planning decisions and the current pollution due to land-use change in the 20-year period are evaluated; we see that there are discrepancies between the zoning plans made for the planning areas in “Bursa Metropolitan Area” shaped for land use in the field of research, and the current land use in the areas of, observed. This fact reveals the importance of pollutions caused by changes in land use led by legal planning decisions, as well as the pollutions caused by unplanned and illegal uses of land. Ecological risks arise on land-use development plans, especially for the Research area, due to land use plans and plan notes that support the increase in residential and industrial land uses. Especially due to these processes that cause loss of agricultural lands and forest areas, natural thresholds are exceeded and natural resources are disposed.

References:

- Akkılıç, Y. Encyclopedia of Bursa, BURDEF Publications, issue: 3, pp. 102-108, Bursa, Turkey, 2002.
- Altaban, Ö. A Balance Sheet of Urbanization in Greater Bursa, Establishment of the Ottoman Empire 700 11th anniversary of the Bursa Region Building Life Congress, Bursa Branch of the Chamber of Architects Press, pp. 169 – 192, 1999.

- Arslanoğlu, E. Regulation of the General Master Plan, Project: Architect Luigi Piccinato, Event Journal, pp. 36 – 54, 2004.
- Report Strategic Plan 2020 in Bursa, Bursa Metropolitan Municipality, the Ministry of Public Works and Settlement, Bursa Governor's Office, Turkey, pp. 122-138, 1998.
- Bursa Metropolitan Municipality, 1/100.000-Scale Environmental Master Strategy Plan Report 2020, Publication of the Bursa Metropolitan Municipality Master Plan Bureau, Bursa, Turkey, pp. 192-193, 1997.
- DİE, State Institute of Statistics, 1990 Population Census, Population, Social and Economic Structure of the State Institute of Statistics, Ankara, Turkey, Issue: 1577, 1993.
- DİE, State Institute of Statistics, 2000 Population Census, Population, Social and Economic Structure of the State Institute of Statistics, Ankara, Turkey, Issue: 2718, 2002.
- Dostoğlu, N. and Vural, T. Löcher, Prost and Urban Development Contributions Piccinato'nun in Bursa, Event Journal, pp. 122 – 126, 2004.
- Fidan, N., Alan, J. and Güniçen, E. Bursa: S Bursa Healthy Vision and Plans for the World Health Organization Healthy Cities Project IV. Phase (2003-2007), Bursa Metropolitan Municipality Healthy Cities Project Coordinator, Bursa, Turkey, 2005.
- Kaplanoğlu, R. and Cengiz, İ. From Prusa to Bursa: Urbanizing the Way to Europe, Istanbul Publication, ISBN: 9759549783, İstanbul, Turkey, 2005.
- Küçükali, U. F. Pollution and Risk Analysis for Başköy-Kestel Section and Sub-basins of Bursa Nilüfer Creek, PhD Thesis, Yıldız Technical University, Istanbul, Turkey. 2012.
- Tekeli, İ. Date of Bursa Three Different Conversion Period, the Founding of the Ottoman Empire 700 11th anniversary of the Bursa Region Building Life Congress, Bursa Branch Chamber of Architects Press, Turkey, pp. 7-30, 1999.
- Tosun, E. K. Spatial, Social and Cultural Change in Cities in the Globalization Process: Bursa Example, PhD Thesis, Uludağ University, Bursa, Turkey, 2007.

IMPROVING MUNICIPAL MANAGEMENT ASSOCIATED WITH THE RESOURCES COLLECTION THROUGH IT: CAASIM CASE IN MEXICO

Antonio Alejandro Arriaga-Martinez, PhD

Universidad Anahuac México Sur

Sonia Ximena Diaz de Cossio-Priego, Master

Universidad del Pedregal

Abstract:

The study presented is research based on the impact of the pay stub in the management of the Water Commission (CAASIM), an institution that works in some municipalities that belong to the state of Hidalgo in Mexico. Largely, the findings show significant changes such as reduced operating costs, increase in the amount of revenue, decrease in the volume of water consumed per household, reduced staff dedicated to bill collecting and the use of a convenience store as a principle payment point.

The findings in this research, considering data over several years, completely refute the belief that people in Latin America are seeking to evade the payment of their obligations to the government, but that there is a need for more efficient administrative mechanisms in order to provide adequate service for citizens.

Key Words: E-government, payment slip, efficient municipal management, Mexico

Introduction

The analysis of the impact of technology can be measured through different aspects, however one should be careful to create a framework so that the findings obtained are consistently expressed.

This paper seeks to frame the answer to the research question: What impact has the implementation of the pay stub payment system had for CAASIM service? In a deductive manner, improving public administration electronically arises from how the expansion of e-government initially serves and as an ulterior goal. We also refer to the factors affecting the adoption of technologies, depending on the level of economic development of the country. We show some recommendations made in previous studies on models that can effectively promote the adoption of information and communications technologies in developing countries, as in the case of Mexico.

The idea of e-government has been preceded by the structure of electronic commerce and for this, without a doubt, the main motivator is to reduce transaction costs and increase the speed with which these are made possible by the digitization of information. The literature supports the implementation of technology through the Transaction Cost Theory by Ronald Coase². This study seeks to confirm said theory with information from CAASIM. We've added to this document some cases and effects of e-government in the Caribbean and finally, we describe the technology used in the CAASIM payment system as well as their processes and those persons involved.

CAASIM uses barcode technology, which can be scanned and read directly into a computer. Thus the data is instantaneously interpreted, virtually and without errors, by a bar code reader. A payment system has been developed which involves a pay stub and is a standard that can be used in retail stores, banks and privately-owned billing establishments.

The companies that provide utilities (electricity, gas, water) and other providers (insurance,

² Ronald H. Coase es ganador del Premio Nobel de Economía en 1991 por su aportación del significado de los costos de transacción en 1961.

government agencies) have sought to improve customer service through an automated payment system. This requires that the information that all parties involved in the payment process must have be such that it can be captured electronically. One of the most effective methods to do so, from the standpoint of cost, is the method implemented by CAASIM: the barcode.

The flow of information accompanying the physical payment requirement is communicated between business partners in various ways. EDI (Electronic Data Interchange) is the optimal way to convey that information.

This case is considered to be an example of how information and communication technologies, in a simple way, can help improve the quality of people's lives. This is also an example of how a municipal e-government begins to change the way the government administration or its agencies deliver some services to citizens. It is also determined in the investigation, how other government agencies tend to offer similar options for the government's own services in a growing stage of e-government at the local or municipal level. Finally, thanks to these findings, we can conclude that people in Latin America (specifically in Mexico) are not necessarily looking to evade their responsibilities to the government, but there is an urgent need to improve services so that the payment of utilities is more accessible to society in general.

Literature Review

Because the application of IT by CAASIM can be catalogued as e-government, we must look further into the literature on the subject. E-government, is an agent of change, Prins (2001) defines e-government as the delivery of online government services, reducing government bureaucracy, increasing access to democracy and reinforcing agencies' accountability towards citizens' needs. For the European Union, e-government is "the use of information and communication technologies in public administration, combined with organizational change and new skills to improve public services and democratic processes, strengthening support to public policies." (European Information Society, 2004) The implementation of e-government has different applications either to automate and integrate different information spaces in order to simplify and maximize the benefits of technology, or to modernize public services and provide services to citizens. (Navarra and Cornford, 2003) Aside from specific objectives, the transfer of public administration processes from a manual format to an electronically automated format in real time engages, in some countries, a radical redesign of processes at local and national levels.

According to the company Accenture (2004), e-government allows for high performance as it provides better results at lower costs and maximum value for each resource used.

In the case of a municipality, certain resistance has been identified in the adoption of e-government in various countries; these factors recognize some limitations, such as lack of awareness (Reffat, 2003), lack of confidence (Navarra and Cornford, 2003), lack of strategy and infrastructure. (Reffat, 2003)

Chen, et. al. (2007) generated recommendations after a comparison of the state of e-government among developed and developing countries. According to the authors, financial problems are prevalent in developing countries and can be resolved by a strict tax structure. To improve infrastructure, outsourcing is an option that needs to be carefully considered due to lack of internal, technical team. Technicians are a major issue in the implementation of e-government in developing countries.

Another item to consider is collaboration, where the removal of borders between different government agencies is the biggest problem in e-government. Government departments should first define their structure, then the transactions need to be simplified and the decision-making process must be known. Simplifying transactions, citizens are more likely to access the site. Before any underdeveloped nation can intend to change its public administration structures to a technological infrastructure, they must first determine to what degree the population has access to internet.

Until now, the implementation of e-commerce confirms Coase's theory, which states that the more people and entities transacting (applied to this case, electronically), the less the individual cost will be, and the transactions will be faster and more frequent. Coase believed that the circumstances of some transactions could make the government a more efficient participant than those private firms affected by market forces. The primary goal of e-government is "the gradual transformation of government business at many levels." An important example of the results of e-government in developing countries is the case of the Caribbean, where he stressed that the e-government project can be a

tangible driver in the reduction of the digital divide in the region.

A reduced digital divide can result in a wider range of services available via e-government because a higher proportion of the public would have access.

Some more developed nations in the region, such as Brazil and Mexico, have reported success in infrastructure development projects in e-government. In the case study presented, we analyze the impact generated when a public service organization includes technology in the payment made by customers. In order to understand this, we must consider the definition of the payment system implemented by CAASIM.

An electronic payment is defined (Raja & Velmurgan 2008) as a paid service that uses information and communications technology including an integrated circuit card, cryptography and telecommunications networks. The activities needed to design the infrastructure for payment systems become more complex, since competition and innovation constantly promote the search for better combinations of efficiency, reliability, security and system stability, in order to provide payment services to a number greater of individual users and institutions. (Raja & Velmurgan 2008)

On a pay stub you can find the following information, which can be read by barcode (Rosario, 2008): customer information, supplier information, a detailed invoice of the services provided, a reference number, the amount payable and payment terms (deadline or place of payment).

Furthermore, it is important to mention the key steps for using coded pay stubs:

Table 1: Adapted from Rosario, 2008

Step 1: The billing entity and payment receiving agency must enter into an agreement to work within the appropriate, legal framework regarding the payment receiving system, considering the agency.

Step 2: The payment receiving agency informs and empowers all receiving points regarding the processing of pay stubs.

Step 3: The billing entity emits coded payment stubs for their customers.

Step 4: These customers (invoicees) take the coupon to the agency receiving payments. They scan the pay stub, payment is made and a receipt is issued.

NOTE: the working agreement between the billing entity and payment receiving agency should consider what would be an appropriate action if the deadline has passed.

Step 5: The payment receiving agency files all pay stubs received. The information is collected, organized, and then transmitted to the party issuing the relevant invoice or through the banking system, after an agreed upon period of time. The funds are then transferred to the bank account assigned.

Step 6: The billing entity updates its database.

This table (Rosario, 2008) clearly shows the flow of information and transactions involving the use of the pay stub for services. The technological infrastructure and administrative processes should adjust effectively, in view of the benefits that customers require and that the competition demands.

Methodology

In this study we have agreed to use the case method (Yin, 1994) to address the situation, because the strategy of approaching the phenomenon is very important in analyzing the different characteristics. According to scientific methodology, research is descriptive, suggesting a depth characterized by collecting data that allow the researcher to identify what is being observed, as a first approximation to the phenomenon. This is because there is no other similar study to the case of Mexico. For Hernandez (2010), descriptive studies are used to analyze how a phenomenon and its components manifest. The purpose of the research is to describe situations and events. He seeks to specify important properties of individuals, groups, communities or any other phenomenon that is under analysis. Various aspects are measured or evaluated, as well as dimensions or components of the phenomenon to be researched. For its interrogative purpose a nonexperimental investigative line must remain for the study, since none of the variables are manipulated by the research team (Hernandez, 2010). The approach from which the study will be addressed is quantitative, therefore deductive (Kerlinger, 1988), because it uses the collection, analysis and interpretation of data to answer the research question.

It should also be considered that a clear quantitative approach does not limit the case study technique (Eisenhardt, 1989), but strengthens it by extracting specific information (questionnaires, document analysis, observation) which can be reviewed and assessed in the analysis unit chosen. In this case, because the information is not public, it was necessary to carry out personal visits in order to obtain the information presented here.

It is important to mention that the case study method does not compromise its value because it exhibits a specific study, rather because of the depth with which said case is researched. (Pettigrew, 1997) The case study method associates the broad collection of evidence, which supports science to develop theories (Eisenhardt, 1989) bringing together the knowledge in areas in which there is no clear consensus. The researcher assumes that there is an external, complex reality, knowable only through observation and triangulation of information (and that observation is fallible). In addition, realism supposes the possibility of assessing the reliability and validity of the knowledge, something that is not so evident in constructivism and critical theory (Craig, 1990). The choice of unit analysis corresponds to the characteristics found in e-government at the municipal level, where the importance of the study lies. An example of the implementation of the pay stub, for the collection of a public service and their respective revenue in concentrated areas and accessible measurement.

Mexican Water Works: CAASIM (Comisión de Agua y Alcantarillado de Sistemas Intermunicipales)

CAASIM is a public agency of the State Public Administration (Hidalgo, Mexico) called the "Commission of Water and Inter-municipal Sewer Systems." The Commission's Mission Statement is "To provide well-being to society through quality water services, with efficiency and optimization of agency resources through advanced technology, implementing a water culture in accordance with global needs." (CAASIM, 2011)

The legal nature of the decentralized municipal management in the state of Hidalgo comes from the publication of the National Water Act (1992), which grants jurisdiction to the municipalities to operate and run for public urban use. In the state of Hidalgo, the commission was established by Government Decree published in the Official Journal of the State on March 19, 1992, but changes on May 17, 2010, established the objectives to be fulfilled by said institution. Among these is the obligation of "Providing public services: water supply, drainage, sewer treatment, disposal and reuse of wastewater in the population where solicited."

Also attributions are granted to allow for optimum performance, these include:

- I. Plan, schedule, budget, study, project, build, rehabilitate, expand, operate, manage, conserve and improve systems for drinking or treated water, sewage, water treatment, disposal and reuse of wastewater: in terms of the law;
- II. Charge for the service under the terms of the State Water and Sewage Law, its Regulations and other applicable norms, contributions as rights, exploitation and their accessories, or other to be perceived through providing public services;
- III. Apply dues and fees that correspond to each of the activities or actions to be made for the provision of public services.

Similarly, the official document illustrates the organic statute which details the administrative structure and includes General Management, Technical Secretariat, Directorate of Legal Affairs and Communications Technology Directorate, Directorate of Administration and Finance, Supply Management, Commercial Management, User Support, Comptroller Division, Directorate of Construction and Hydraulic Operations Directorate. The situation which precedes access to drinking water in urban and rural areas of the state of Hidalgo suggests shortcomings in distribution and fluid loss, but not in the supply of the resource (Amaya, 2010). Similarly, the water problem in this area has been characterized as a state of exhaustion and over-exploitation of aquifers exacerbated by financial problems and limited administrative capacity, as is stated by Amaya (2010), in his study of inter-municipal water management models in Mexico.

To approach the fulfillment of their duties, CAASIM has deployed the instrument that is the subject of this study, a pay stub with barcode technology, implemented as a means of payment that is accepted in various retailers, unrelated to the Water Commission. It is worthwhile to mention that certain convenience stores have an agreement with the Water Commission and there are at least 73 locations within the state of Hidalgo for the collection of these payments.

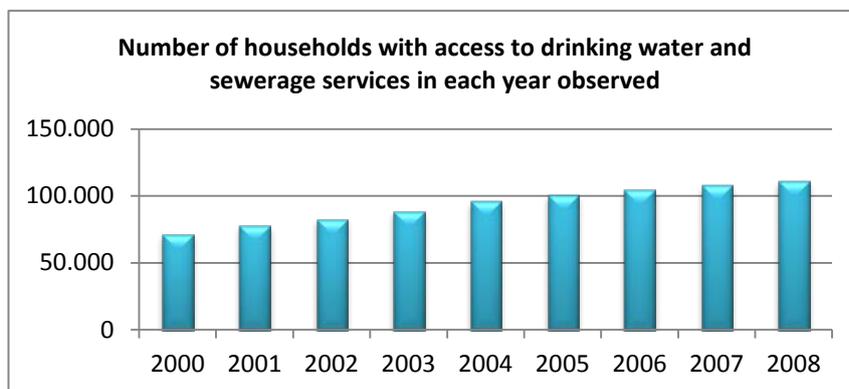
The collection of data in the Water Commission went on for two months with documents provided by the institution and the review of external sources that allowed comparison and objectivity. The formats used include the measurement year and the spaces where you put the information residing in the information systems of the Commission.

Findings

In the section below, the compendiums of information are displayed, after the data collection originating in CAASIM where we obtained the following comparative, in which the period from 2000 to 2009 is the most complete.

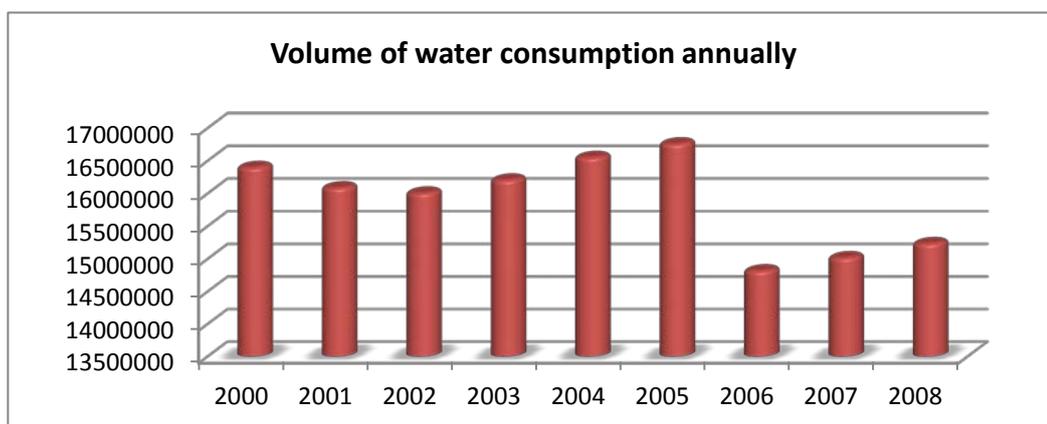
Initially statistics are established on the number of households with access to drinking water and sewerage services in each year observed (INEGI, 2009). The trend in the growth of the number of connections per year is shown in the graph below, suggesting a parallel impact on the productivity of CAASIM:

Graph 1: Developed by the authors



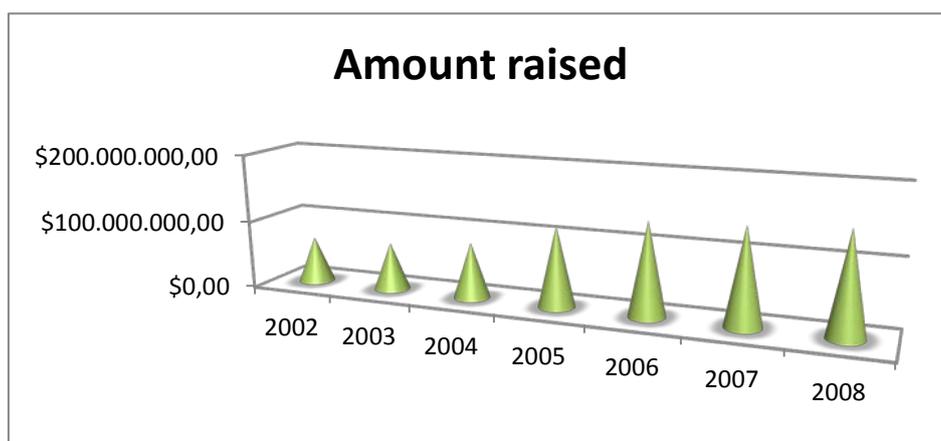
Similarly, the concentration of information to determine the volume of water consumption annually suggests a steady decline, as shown in the following table and graph:

Graph 2: Developed by the authors



We overwhelmingly assert a substantial increase in the amount raised by the Water Commission during the periods observed:

Graph 3: Developed by the authors



In the final review of each of the documents obtained and summarized in the tables and charts here presented, we can see that despite the increase in the number of connections, some interesting indicators clearly manifest, which are as follows:

- **The annual water consumption decreases:** The consumer has more access to information use and makes monthly payment. It can be noticed that this impacts water consumption.
- **The amount collected increases:** Because the options of places to pay have increased, the amount collected has as well.
- **There is a decrease of 62.9% in costs associated with the collection payment.**
- **It has been proven that the collection staff decreased from 35 to 13 individuals:** The change in the payment system allowed for eliminating costs associated with the place and the people who did this work

The primary place of payment is the convenience store. It has been determined that the use of the pay stub along with the convenience store chain has been a success, and the consumer has established it as the preferred means of payment.

This is evident in the following tables:

Table 1 Developed by autors	Associated Collection Costs
Year	Amount
2006	\$6,650,000.00
2007	\$2,470,000.00
Savings	62.90%

Table 2 Developed by authors	Associated Collection Personnel
Year	Personnel
2006	35
2007	13
Personnel reduction	22

Table 3 Developed by authors	Mean of payments, Commissions and Number of locations	
Means of Payments	Commission Value	Locations
Bancomer	\$5.00	8
Banorte	\$5.00	3
Banco del Bajío	\$5.00	1
OXXO	\$2.00	40
CAASIM office	N/A	6

Analysis and Discussion

After observing the results of the investigation, we can confirm that the view from the Transaction Cost Theory is found in the case of CAASIM, not only by reducing operating costs, which Coase emphasizes in his theory, applied also to government transactions-(Fenwick, John & Stimac, 2009) and personal transactions, but because the consumer has been able to adapt to this innovation, for several reasons, among which the most important is the fact that the convenience store is a known and accepted as a safe, trustworthy context for the customer.

The customers' benefits are represented by the amount of payments made by users now have a greater number of points of payment and the convenience this generates. For invoicing, payment is made in a timely manner, resulting in increased income for CAASIM. The results together promote the service of the Water Commission, as a service with added value. Another impact discovered in the study is the presence of movements in the relationship consumption-payment, since the greater the use, the more the household must pay, which allows to consider the user's awareness of use when he periodically registers the payments and receives specific information about his consumption on the pay stub.

It is clear, from another perspective, that the implementation of the pay stub suggests an economic recovery for the Water Commission, which has been among other things, presumably used in the development of municipal infrastructure in the regions for which it is responsible, so that the number of households with access is greater. Within the multidimensional assessment of the information gathered, there is a stronger sense of poverty rates in those who manifest shortcomings and limitations of access to basic services in local populations, with respect to the supply of drinking water. (CONEVAL, 2010) From the perspective of the impact of the use of information technology some

authors (Srivastava, & Teo, 2007) have mentioned that it is possible to analyze how e-government can help nations achieve their social and economic goals. Specifically, they explain the significant relationships between the degree of development of e-government and administrative efficiency and location of resources, which affects the social equity and business competitiveness of a country. Equally, this paper shows that those benefits which locally promote a decrease in the gap among the social groups most deprived of public services.

Upon completion of this research and reflecting upon the initial question that asked what would the impact of the implementation of the payment system through pay stubs be for CAASIM, it is feasible to argue that at the municipal level has been a successful transformation payment system to choose outlets such as convenience stores that have a highly valued and privileged presence (i.e., in large quantities and with good location), in addition to charging a fee which is not significant for the consumer. We have also identified that the citizen chooses the means of payment that are most attractive (easy access, minimal extra charges) and this supports the fulfillment of the obligation to pay monthly, an issue which previously, without information systems and electronic collection, for CAASIM, was unthinkable. The decrease in costs for the Water Commission has been instrumental in strengthening and expanding this system in the municipalities of the state of Hidalgo, as the benefits are flattering. The justifications for the dissemination of the issue revised here are based on several studies that address some doubts about the importance of e-government. For example, Lee, et. al. (2008) suggest two primary aspects to understand the benefits of e-government. In the first the transformation of government operations results in benefits for citizens, companies and the government itself. This means that the needs of citizens can be better recognized and can help companies become service providers of both individuals and the government. This will also influence the reduction of operating costs by increasing government efficiency in internal operations. The second aspect is the positive transformation of the relationship between government-subordinates through improved interactivity, making this interaction smoother, faster and more responsive. (Lee, et al 2008)

It is important to recognize that there is no documentation on the effects of pay stub as a system in the public sector in Mexico, and this paper proposes a line of research to put into practice, so that other members of government administration find the application of technology accessible in the context of Mexico. It is worthwhile to remind the reader of the methodological importance implemented, because when you fail to distinguish several sources of evidence and the boundaries between the observable facts and context are not clearly stated, the case study becomes important once again. (Yin, 1994)

This method also allows researchers to use a "controlled opportunism" to respond flexibly to new discoveries while collecting new data. (Eisenhardt, 1989) The methodological approach specifically within the discipline of information technology, analyzing e-government, has expressed greater inclination by observation methods and review of documents, allowing researchers to verify and validate the empirical findings by triangulation. (Sarikas, 2007)

Data analyzes were performed comparing the different discoveries including identifying issues, which are categorized and sorted into various lines. (Sarikas, 2007) Verification and validation of the findings, by triangulating data from primary sources with official documents published and used as secondary sources, were taken up to ensure to avoid any bias that might arise. (Choudri, 2005). For Dhillon, et. al. (2008), the case studies offer the potential to generate alternative explanations of diverse perspectives, making way for the researcher to detect contradictions or misunderstandings. They are also appropriate for review where the goal is to study current events, and where it is necessary to control behavior or behavioral variables, according to Yin (2003).

CAASIM is an example of how the government approaches the citizen and provides more efficient service, changing the business model for the benefit of all involved. With absolute certainty and precision in its analysis, Fenwick, et al (2009) recognized at the conclusion of the study on the economic necessity of e-government that the race for wealth (among countries) will be won by the governments with the best design which will allow millions of transactions, combining efficiency with trust (minimizing transaction costs with information security).

An important part of the findings is the irrefutable proof that the user has not sought to evade payment, but rather the opposite. An increase in payment frequency has been achieved and also encourages the user to be more responsible in water consumption.

Conclusion

The integration of advanced technology tools, with some fundamental processes in the management of government institutions, denote the ability of corporate leaders to adapt to the benefit of citizens, instruments like the barcode, which engaged in a pay stub promotes the respective collection of a public service.

The recognition of various factors, including the degree of acceptance of a new means of payment for services, the expansion of the number of payment points for the water and sewer system in the municipalities of the state of Hidalgo, the certainty of citizens regarding the reliability of this innovation, the relevance in billing information per household, the accuracy in the amount of water used and the appropriate fee, among others, refer to the virtue that local authorities who are in charge of drinking water in many areas Hidalgo have held.

The contribution herein obtained is relevant through the analysis and evaluation of the facts described, not only for the areas who directly benefit, nor for citizens experiencing the changes, but for the country at large, because this line of research, that is: the technology applied to the quality of life of the population, a critical input in the review that is done to the payment of municipal management of water and sewerage services in organisms such as CAASIM. Enthusiastically the Water Commission has engaged in a process of change that, beyond the technological aspect, moves into the cultural range, even where environmental consciousness and narrowing of the government-citizen relationship is affected, where improvement in fulfilling the obligations of these institutions has been expressed in order to increase efficiency, giving people a new way of perceiving the public function. It is also a contribution which should encourage the researcher's approach to exercise similar cases for further understanding of this phenomenon (the service payment pay stub), which is increasing in commercial nature, but still is in consolidation in terms of federal, state or municipal public administration.

Finally, the assessment of the consummation of certain economic principles that have value in today's world satisfies researchers, allowing them to accept the successful implementation of theoretical and practical models in real benefit of individuals, with immediate and plausible translation for various socio-economic sectors, demonstrating with greater knowledge the long road yet to be traveled to achieve a more just society for all its members, with precise awareness of the scope and limitations of administrative processes. The will to continue to evolve towards e-government with high response and without excess, much of the people do not conceive to be part of their reality, but it has been shown, at least in this case, that when citizens receive good service, they are willing to pay for it.

References:

- Accenture. (2005). E-Government leadership: High performance, maximum value. Government Executive Series, <http://www.egov.gov.sg>
- Amaya Ventura, Ma. de Lourdes. (2010). La Gestión Intermunicipal del agua en México. La revue en sciences de l'environnement, Hors-série 7
- CAASIM (2011). Retrieved August 20, <http://caasim.hidalgo.gob.mx/>
- Chen, Y., Chen, H.M., Ching, Russell K.H., Huang, Wayne W. (2007, April-June). Electronic Government Implementation: A Comparison between Developed and Developing Countries. *International Journal of Electronic Government Research*, 3(2), 45-61. Retrieved from <http://www.igi-pub.com>
- Choudrie, Weerakkody, & Jones. (2005) Realising e-government in the UK: rural and urban challenges. *School of Information Systems, Computing and Mathematics. The Journal of Enterprise Information Management*, Vol. 18 No. 5. Retrieved de www.emeraldinsight.com/1741-0398.htm
- CONEVAL (2009). Metodologías de medición multidimensional de la pobreza. Consejo Nacional de Evaluación de la Política de Desarrollo Social. Retrieved de www.coneval.gob.mx
- Coase, R. H. (1937), The Nature of the Firm. *Economica*, 4: 386-405. doi: 10.1111/j.1468-0335.1937.tb00002.x
- Craig, Edward (1990). Artículo "Realism and antirealism", *Rutledge Encyclopedia of Philosophy*.
- Dhillon, G. & Weerakkody, V. (2008). Moving from E-Government to T-Government: A Study of Process Reengineering Challenges in a UK Local Authority Context. *International Journal of Electronic Government Research*, Volume 4, Issue 4. IGI Global

- Eisenhardt, K. M. (1989), "Building Theories from Case Study Research", *Academy of Management Review*, Vol.14, No. 4, pp. 532-550.
- Fenwick, W., Erin, J., & Stimac, J. (2009, Marzo). The necessity of e-government. *Santa Clara Computer and High - Technology Law Journal*; 25, 3; Retrived from ABI/INFORM Global
- Hernández S, Roberto. (2010) *Metodología De La Investigación*. Mexico, Mc Graw-hill.
- INEGI (2010) Retrived on January 10, 2010 <http://www.inegi.org.mx/>
- Kerlinger, F. N. (1988). *Investigación del comportamiento*, México: McGraw-Hill.
- Lee, H., Irani, Z., Osman, I., Balci, A., Ozkan, S. & Medeni, T. (2008) Toward a reference process model for citizen-oriented evaluation of e-Government services *Transforming Government: People, Process and Policy*, Vol. 2 No. 4.
- Ley de Aguas Nacionales. *Diario Oficial de la Federación*, México, 1º Diciembre de 1992
- Navarra, D.D. and Cornford, T. (2003), "A policy making view of e-government innovations in public governance". *Proceedings of the Ninth Americas Conference on Information Systems*. Tampa, Florida.
- Pettigrew, A. M. (1997), "What is a processual analysis?" *Scandinavian Journal of Management*, Vol. 13, No. 4, pp. 337-348.
- Prins, C., (2001). *Electronic government: variations on a concept* in J, J.E. (Ed.), *Designing E-Government*, Kluwer Law International, Dordrecht.
- Raja, J. & Velmurgan, M. Senthil. (2008, April). E-payments: Problems and Prospects. *Journal of Internet Banking and Commerce*, Vol. 13, no.
- Reffat, R. (2003), *Developing a successful e-government*. School of Architecture, Design Science and Planning, University of Sydney, Sydney.
- Rosario V., Giselle. *Modelo Integral Genérico Payment Slip. Payment Slip para Pagos de Servicios*. Retrived Junio 28, 2009. http://www.amece.org.mx/amece/Documentos/estandares/estandares_identificacion/Modelo_Integral_Generico_Payment_Slip.pdf
- Sarikas O. & Weerakkody, V. (2007). Realising integrated e-government services: a UK local government perspective. *Transforming Government: People, Process and Policy*, Vol. 1 No. 2. Retrived de www.emeraldinsight.com/1750-6166.htm
- eGovernment for Europe's public services of the future (2004) <http://www.uoc.edu/inaugural03/eng/article/index.html>
- Srivastava, S. & Teo, T. (2007) E-Government Payoffs: Evidence from Cross-Country Data. *Journal of Global Information Management*, Volume 15, Issue 4, IGI Global
- Yin, R. (1994), *Case Study Research, Design and Methods*. Sage Publications, Thousands Oaks, CA.
- Yin, R. (2003). *Case study research: Design and methods*, 3rd Edition. London: Sage Publications.

TREATED WASTEWATER USE AND ITS EFFECT ON WATER CONSERVATION, VEGETATIVE YIELD, YIELD COMPONENTS AND WATER USE EFFICIENCY OF SOME VEGETABLE CROPS GROWN UNDER TWO DIFFERENT IRRIGATION SYSTEMS IN WESTERN REGION, SAUDI ARABIA

Khaled S. Balkhair
Fathy S. El-Nakhlawi
Saleh M. Ismail
Samir G. Al-Solimani

Faculty of Meteorology, Environment and Arid Land Agriculture,
King Abdulaziz University, Jeddah, Saudi Arabia

Abstract:

Irrigation with wastewater is a widespread observable fact in developing countries, especially those exhibiting arid climate conditions. Utilization of large volumes of domestic wastewater effluent is part of an overall integrated water resources management. The objective of this study was to investigate the effect of irrigation by six different treated domestic wastewater qualities on the yield of Cabbage and Lettuce crops under surface and Sub-surface drip irrigation systems. Field experiments were carried out during 2011 and 2012 seasons at the Agriculture Research station of King Abdulaziz University located at Hada Alsham area north east of Jeddah city. The six water qualities were prepared by diluting treated wastewater with local groundwater of the region. The dilution percentages were 0%, 20%, 40%, 60%, 80%, and 100% treated wastewater. Crop water requirement were calculated based of Penman Monteith equation for dry land condition and supplied daily to the crops. The obtained results showed significant impact of wastewater quality and irrigation system and their interactions on vegetative yield and yield components of both crops. Irrigation with 60% and 80% wastewater under sub-surface irrigation system produced the highest yield and yield components in both seasons. Irrigation Water Use Efficiency (IWUE) of cabbage crop increased under the use of subsurface drip irrigation while IWUE of lettuce crop increased under surface drip irrigation system. Among dilution percentages, the 60% gave the highest IWUE in both crops. In conclusion, this study implies an advantage of using treated wastewater in growing vegetable crops. This in turn saves considerable quantities of local groundwater resource.

Key Words: Wastewater, Irrigation system, Vegetabel crops, Yield, WUE, Water conservation

Introduction

Shortage of fresh water in arid and semi-arid regions like in Saudi Arabia increased the demand for using wastewater in agriculture sector. Using treated wastewater in agriculture is gaining tremendous popularity because of the wide range of benefits that accompany it. These benefits include conservation of fresh water, high level of organic matter and recycling of nutrients, thereby reducing the need to invest in chemical fertilizers (Al-Sha'lan, 2001). In Saudi Arabia, there are more that 30 desalination water treatment plants producing around 3 million m³/day. Part of it is reused for agricultural purposes as an alternative resource for irrigation water. In an economical study of the use of recycled sewage water in irrigation of field crops, Al-Abdulqader and Al-Jaloud (2003) stated that it saved up to 45% in fertilizers cost for wheat crop and 94% for alfalfa crop compared to irrigation with well water due to the fact that sewage water contains the essential elements needed by such crops. They also indicated that the usage of treated sewage water in irrigating of wheat and alfalfa crops increased their yield by 11 and 23%, respectively, and consequently increasing the profit by 14 and 28%, respectively as compared to irrigation with well water.

In Dakar, Senegal, more than 60% of the vegetables especially lettuce and onion are grown in urban areas using a mixture of groundwater and untreated wastewater. About 20 million hectares in developing countries are irrigated with raw sewage or partially treated wastewater (Dreschel et al., 2002). Using wastewater in crop production sometimes decrease the quantity and quality of the yield however, it is possible to achieve high yields of crops without deterioration of their quality by using treated wastewater for the irrigation of crops under controlled conditions. This is evident was from large number of the previous researches as in (Najafi et al., 2003, Jimenez, 2005, Munir and Ayadi 2005, Esmailiyan et al., 2008 and Zavadil, 2009).

In spite of the benefits of using wastewater in crop production, the production is faced by some risks from heavy metal accumulation and microbial pollution. The main risk associated with wastewater irrigation is infection with intestinal helminthes (Mara and CainCross 1989). Depending on the source of the wastewater it might contain chemical pollutants and heavy metals that can accumulate in the soil and crops thereby posing a threat to human health. These risks can be greatly reduced by treating the wastewater before using it or by applying some precautions while using it. Such precaution include: partially treated, diluted with fresh water and using subsurface drip irrigation. The contamination of crops irrigated with municipal wastewater was not detected if the wastewater had been treated in stabilization ponds or disinfected, and if subsurface drip irrigation had been used (Vaz da Costa et al. 1996; Najafi et al. 2003; Panoras et al. 2003; Al-Lahham et al. 2003; Aiello et al. 2007). Excessive contents of heavy metals in crops irrigated with wastewater have not been reported (Zavadil 2009). Sometimes increasing the level of microbiological contamination than that reported by the WHO guidelines can be used for irrigation (Blumenthal et al., 2000 WHO, 2006). Many literatures indicate that the use of treated wastewater is partially safe for irrigation of some crops. The objective of this study was to investigate the effect of irrigation by six different qualities of treated domestic wastewater on the yield of Cabbage and Lettuce crops under surface and Sub-surface drip irrigation systems. This study and its results is part of a comprehensive project aims to investigate water resources conservation and the impact of irrigation by treated wastewater on the pollution and yield of some vegetable crops as well as soil contamination.

Materials and Methods

Experimental design

Cabbage and Lettuce cultivation experiments were carried out at the Agricultural Research Station of King Abdulaziz University (KAU), located at Hada AlSham village; 110 km north east of Jeddah city, Saudi Arabia. The soil at the experimental site is classified as sandy loam. The two vegetable crops were cultivated in strip plot design (Split block) with 4 replications; each of 2x3 m in size. The main plot treatments consist of two irrigation systems, surface and subsurface, while the sub plot treatments designed and arranged in strips containing six wastewater qualities.

Irrigation water source

Bani Malik wastewater treatment plant which is located in Jeddah city was the main source of irrigation water. Water was conveyed to the field site by trucks and stored into two large size reservoirs, each for an irrigation system. Reservoirs were also connected to six different storage tanks corresponding to six different wastewater qualities. The dilution process occurs within the storage tanks based on the desirable ratio of wastewater to local groundwater. The six wastewater qualities are: 0%, 20%, 40%, 60%, 80%, and 100%. The percentages indicate the wastewater portion. For example, 20% means one fifth as volume of wastewater mixed with four-fifths of local groundwater. Accordingly, the 0% and 100% corresponds to local groundwater and undiluted wastewater respectively.

Irrigation systems

Each water quality treatment of its corresponding irrigation system consists of storage tank with a capacity of 5000 L, disk filter, pump, controller, drip lines and solenoid to control flow time and irrigation interval. In sub-surface drip irrigation systems the field was leveled and the dripper lines were installed at 10 cm deep on 40 cm between two adjacent dripper lines. The distance between drippers was 45 cm with a discharge of 0.9 G/h (*RAIN BIRD LD- 06- 12-1000 Landscape drip 0.9 G/h @18"*). The downstream end of each dripper line was connected to a manifold for convenient flushing. Inlet pressure on each tape was about 1.5 bars. The system uses 125 micron disk filter to prevent blockage. The lay-out of the surface drip irrigation was exactly the same as in subsurface drip except for the positions of dripper lines, where they installed on soil surface.

Irrigation water requirements and supply

The required irrigation water was calculated based on crop water requirement (Evapotranspiration) and total available soil moisture. Evapotranspiration for each plant was calculated from reference evapotranspiration and crop coefficient as follows:

$$ET_c = K_c \times ET_o$$

Where:

ET_c : crop evapotranspiration (mm/day).

ET_o : Reference evapotranspiration (mm/day).

K_c : Crop Coefficient.

Reference evapotranspiration were calculated using Penman-Monteith equation as described by Allen et al., (1998). Also, crop coefficient values listed by Allen et al., (1998) for vegetable crops were used.

Data collection

Before harvesting, 10 random guarded plants per plot were labeled and three different traits were measured for each label. In Lettuce crop (*Paris Island CV.*), Plant height (cm), fresh foliage weight (g)/plant, and fresh foliage yield (t)/ha, while in Cabbage crop (*Brunswick CV.*), head diameter (cm), fresh head weight (g) /plant and fresh head yield (t)/ha were collected and measured for two seasons 2011-2012.

Statistical Analysis

The collected data of each crop in each experiment was statistically analyzed using the analysis of variance procedures and mean separation under the criteria of Least

Significant Difference (LSD) test. The analysis was carried out based to the used experimental design, and after applying the assumptions of the statistical analysis according to El-Nakhlawy (2010).

Results and Discussion

Cabbage crop

Means of the studied traits of cabbage crop with their statistically comparisons using L.S.D. are presented in Table 1. Results indicated that sub-surface irrigation system significantly dominated over the surface irrigation in all studied traits in the first season (2011) while no significant differences were observed in the second season (2012). However, head yield/ha were the highest under the sub-surface irrigation in both seasons with values of 18.167 and 17.667 t/ha in 2011 and 2012 seasons, respectively. As for wastewater treatments effects, the obtained measurements revealed that the 0.0 % wastewater irrigation which is the local groundwater (LGW) had the highest mean values in all agronomic traits in the first season while no significant differences were detected in the second season. Apparently, yield and yield components are maximums when LGW is used for irrigation. This result is very clear in the first season, however, there are no distinct differences found among wastewater qualities in the second season.

Table 1. Means of the studied agronomic traits and irrigation water use efficiency of Cabbage crop under the effects of irrigation systems and wastewater qualities during 2011 and 2012 seasons.

Treatme nt	Agronomic traits								IWUE (kg/mm/ha)	
	Plant Height (cm)		Head diameter (cm)		Head weight/Plant (g)		Head yield/ha (t)		2011	2012
	2011	2012	2011	2012	2011	2012	2011	2012		
Irrigation system										
Surface	11.29 b*	15.16	9.43 b	11.2 8	154.15 b	229.44	15.679 b	16.178	25.7 b	76.8 a
Sub- surface	13.88 a	14.84	15.0 a	11.2 3	180.04 a	236.08	18.167 a	17.667	30.1 a	76.9 a
Waste water percentage in Irrigation water (%)										
0.0	16.63 a	15.12	20.15 a	11.1 8	225.61 a	232.22	20.650 a	26.286	37.7 a	81.7 ab
20	11.65 c	14.54	11.3 b	11.9 4	176.11 b	232.3	18.075 b	25.481	25.1 cd	69.1 e
40	12.66 b	14.66	10.85 b	11.0 8	170.06 b	230.92	17.188 b	27.954	26.9 bc	74.6 cd
60	13.43 b	14.89	11.06	11.1	173.29	231.94	15.825 c	26.661	28.9 b	84.1 a

			b	0	b					
80	10.70 d	15.06	9.98 b	11.5 0	136.62 c	233.94	15.137 cd	26.367	23.4 d	70.4 de
100	10.44 d	15.18	9.13 b	11.4 1	125.79 d	235.38	13.662 d	26.286	25.4 cd	77.6 bc

*, Means followed by the same letter(s) are not significantly different according to LSD test at $p \leq 0.05$.

Irrigation water use efficiency (IWUE) under the current study is defined as crop yield per unit of applied water. The last column of Table 1 shows the obtained values of IWUE for both seasons under surface and subsurface irrigation systems as well as wastewater qualities. The value of IWUE increased under sub-surface drip irrigation in the first growing season compared to surface irrigation. Conversely, no change was obtained in the second growing season. As of wastewater quality effect, LGW resulted in the highest IWUE value in the first season followed by 60%, and 40% wastewater qualities respectively. The least IWUE was obtained under the application of 80% wastewater quality. On the other hand, the second season results demonstrated slight change in the order of wastewater quality influence. The highest IWUE (84.1 kg/mm/ha) was obtained under the 60% wastewater quality while the lowest (69.1 kg/mm/ha) was under 20%.

Table 2 shows the interaction between irrigation system and wastewater quality of cabbage traits. Clear discrepancies were found in the results between the two seasons. Significant differences were depicted in the first season on plant height, head weight/plant, and head weight/ha, while no significant effects were noticed between trait means in the second season. An interesting negative regression between head yield/ha and applied water qualities was observed in both surface and sub-surface irrigation systems, especially in the first season. The highest head yield (22.525 t/ha) was produced under LGW and the lowest (11.13 t/ha) was under the application of 100% wastewater. On the other hand, no significant differences were shown in the values of yield under sub-surface irrigation system in all applied water qualities in the 2nd season.

Means of head weight/plant results were almost similar in trend to that of head yield/ha. The negative regression between wastewater and head weight/plant also applies here not only in the first season as in the case of head yield/ha but also in the 2nd season. Though, the second season showed no significant differences. Similarly, the two other yield components, head diameter and plant height followed the same trend but less affected. On the other hand, the highest IWUE was obtained under LGW in both irrigation systems in the first season. In the second season, however, the highest IWUE were obtained with the use of 60% (90.62 kg/mm/ha) and LGW (85.50 kg/mm/ha) water qualities under sub-surface and surface irrigation system respectively.

Lettuce crop

Comparison of agronomic traits of lettuce crop under surface and sub-surface drip irrigation systems are shown in Table 3. Results confirmed that vegetative yield/ha and weight/plant were significantly higher in surface irrigation than that of sub-surface irrigation system. Vegetative yield/ha were 46.83 and 48.98 t/ha under the surface irrigation and 40.59 and 42.76 t/ha under the sub-surface irrigation in 2011 and 2012 seasons, respectively.

Table 2. Means of agronomic traits and irrigation water use efficiency of Cabbage crop under the effects of the interaction between irrigation systems and wastewater qualities during 2011 and 2012 seasons.

Irrigati on System	Waste water (%)	Agronomic traits								IWUE (kg/mm/ha)	
		Plant Height (cm)		Head diameter (cm)		Head weight/Plant (g)		Head yield/ha (t)		2011	2012
		2011	2012	2011	2012	2011	2012	2011	2012		
Surface	0.0	15.58	15.57	11.00	11.13	222.53	232.57	22.525	26.761	37.55	85.50
	20	11.00	14.79	10.13	10.99	165.35	229.95	18.200	24.650	25.43	58.75
	40	11.58	14.57	9.86	11.03	156.90	225.13	15.200	28.912	26.51	71.87
	60	12.05	14.99	9.80	11.05	154.25	226.87	14.275	26.431	27.28	78.27
	80	9.33	15.32	11.21	11.83	107.83	230.52	12.750	25.756	18.05	80.35
Sub- surface	100	8.22	15.70	6.64	11.64	101.76	231.60	11.125	25.900	19.30	79.72
	0.0	17.43	14.67	11.03	11.24	224.70	231.87	18.775	25.811	37.88	77.85
	20	12.25	14.30	11.02	11.49	186.88	234.65	19.950	26.311	24.73	79.45

	40	13.57	14.76	10.14	11.14	183.23	236.71	19.175	26.996	27.35	77.40
	60	14.88	14.80	10.16	11.16	172.23	237.01	17.375	26.892	30.42	90.62
	80	12.18	14.08	10.19	11.17	165.40	237.36	17.525	26.979	28.75	60.50
	100	12.60	14.66	11.21	11.19	149.83	238.89	16.200	26.800	31.53	75.40
	LSD(0.05)	1.27	NS	1.64	NS	22.31	NS	5.169	NS	3.85	6.31

The effect of wastewater quality application on yield/ha and weight/plant was almost the same in both seasons. Obtained values of these two yield components can be classified in two categories. The first category which is the irrigation with LGW, 60%, and 100% wastewater qualities produced higher yield/ha and weight/plant than the second category which is 20%, 40, and 80%. The lowest values of yield/ha and weight/plant were obtained under the cultivation of the crop under 20% wastewater treatment in both seasons.

Surface drip irrigation significantly increased IWUE during the first growing season compared to sub-surface drip irrigation. However, in the second season the case was reverse, where IWUE was higher under subsurface irrigation (Table 3). The high values of IWUE recorded during the first growing season were under the irrigation with 60%, 100% and LGW qualities, where they were almost similar; followed by 40% and 80%. The least IWUE was obtained under 20% treatment. It can be concluded that the trend of IWUE is similar to that of yield/ha and weight/plant. In the second growing season the highest IWUE was obtained by 100% and 60% with no significant differences between them; followed by LGW, 80%, and 40%. The least IWUE was obtained under the 20% wastewater quality.

Table 3. Means of the studied agronomic traits and irrigation water use efficiency of Lettuce crop under the effects of irrigation systems and wastewater qualities during 2011 and 2012 seasons.

Treatment	Agronomic traits						IWUE (kg/mm/ha)	
	Plant Height (cm)		Vegetative Weight//Plant (g)		Vegetative yield/ha (t)		2011	2012
	2011	2012	2011	2012	2011	2012		
Irrigation system								
Surface	24.13	27.13	520.25 a	564.33 a	46.83 a	48.98 a	202.3 a	188.0 a
Sub-surface	23.92	25.92	461.46 b	529.29 b	40.59 b	42.76 b	178.9 b	196.5 a
Waste water percentage in Irrigation water (%)								
0.0	23.88 b	24.85 a	582.63 a	645.88 a	52.49 a	52.98 a	226.7 a	204.1 ab
20	20.38 c	22.21 b	338.50 c	351.75 e	31.70 b	30.4 d	131.5 c	128.6 c
40	24.25 b	24.76 a	450.88 b	506.50 d	40.23 b	48.3 c	173.8 b	181.5 b
60	25.38 ab	26.46 a	599.38 a	592.13 bc	53.94 a	50.73 b	233.0 a	215.0 a
80	24.13 b	24.08 b	428.85 b	570.38 c	33.80 b	48.21 c	146.0 bc	204.3ab
100	23.88 b	26.32 a	598.13 a	614.25 ab	53.85 a	51.84 ab	232.6 a	219.7 a

*, Means followed by the same letter(s) are not significantly different according to LSD test at $p \leq 0.05$.

Table 4 shows results of the interaction between irrigation systems and wastewater qualities. A significance interaction was obtained in the first season only. The 60% wastewater quality and LGW produced the highest vegetative yield/ha under surface irrigation system with values of 57.95 and 56.63 respectively. On the other hand the 80% and 100% wastewater quality responded identically with a value of 57.2 under sub-surface irrigation.

Highest weight values of vegetative weight/plant were 644.23, 628.25, 586.77, and 561.01 under the surface irrigation corresponding to 60%, 0.0%, 40%, and 100% wastewater respectively. The order of water quality however is different in the sub-surface irrigation system. The highest values of weight/plant were obtained with 100% and 80% wastewater quality corresponding to values of 635.25 and 554.50 respectively (Table 4).

Table 4. Means of agronomic traits and irrigation water use efficiency of Lettuce crop under the effects of the interaction between irrigation systems and wastewater qualities during 2011 and 2012 seasons.

Irrigation System	Waste water (%)	Agronomic traits						IWUE (kg/mm/ha)	
		Plant Height (cm)		Vegetative Weight/Plant (g)		Vegetative Yield (t/ha)		2011	2012
		2011	2012	2011	2012	2011	2012		
Surface	0.0	24.50	25.00	628.25	625.70	56.63	45.90	244.60	194.50
	20	21.24	24.50	398.11	346.00	38.30	29.63	154.65	125.57
	40	25.76	26.75	586.77	480.00	52.80	52.80	228.08	181.40
	60	25.48	25.50	644.23	583.25	57.95	48.42	250.32	205.20
	80	22.00	25.50	303.21	569.00	57.30	46.43	117.90	196.72
	100	25.75	26.00	561.01	571.75	50.50	52.99	218.17	224.55
Sub-surface	0.0	23.23	26.00	537.12	666.00	48.34	50.60	208.78	214.40
	20	19.49	24.75	279.00	357.50	25.10	31.07	108.43	131.67
	40	22.76	25.50	315.20	533.00	49.92	42.87	119.42	181.67
	60	25.23	25.00	448.11	601.00	40.30	53.02	215.65	224.70
	80	26.25	25.00	554.50	571.75	57.20	45.00	174.08	211.87
	100	26.51	26.50	635.25	625.70	57.20	50.70	246.97	214.82
LSD(0.05)		2.21	NS	103.25	NS	9.33	NS	40.16	NS

The highest IWUE during the first growing season obtained under 60% wastewater quality followed by LGW under surface drip irrigation. In subsurface drip irrigation, the highest IWUE obtained by 100% followed by 60% and LGW. In the second growing season however there was a slight change in the order of the high and low values, where the highest IWUE was found with 100% followed by 60% and LGW under surface irrigation system. Likewise, under subsurface drip irrigation system, the highest IWUE was obtained under 60% followed by 100% and LGW.

The descending yield of Cabbage crop with ascending concentration of wastewater especially in the first season could be attributed to adverse impact of wastewater on the crop due to the accumulation of micro and toxic elements in stem and leaves of the plant. Consequently, the metabolic processes and total yield of cabbage tend to decrease. This response depends on the interaction between crop genetic makeup and the wastewater concentration in irrigation water (G x E interaction) as reported by Ba-Sahri et al., (2007) and Khan et al., (2009). A study by Ahmed (2011) has also supported our results where he reported that the application of increasing doses of municipal wastewater reduces the overall yield of some vegetable crops.

On the other hand, the positive responses of using wastewater in irrigation of lettuce crop, especially under the sub-surface irrigation system, may be due to the increase in the absorption of macro and micro elements in the wastewater which was reflected in the increase of leaf area and yield components and finally the increase in yield/ha.

In general, the obtained results illustrated the significant interaction between the economic part of the crop and water quality, especially the (%) of wastewater in the irrigation water of the agricultural crops. The positive effects of irrigation with 40% and 60% of wastewater might be due to the increase in the nutrients of the soil under wastewater irrigation. These nutrients may improved the physical and nutrient contents of the soil, hence significantly increased the total chlorophyll and carotene and established good growth and increased biomass and yield of the crop. This justification is supported by many independent studies (Singh and Agrawal, 2009, Khan et al., 2011 and Thapliyal et al., 2011). Moreover, Zavadil (2009) reported that the primary-treated wastewater increased the yield of all vegetables crops, the increase having been statistically significant in most cases.

As indicated by the results of IWUE especially in cabbage crop, sub-surface drip irrigation generally increase IWUE compared to surface drip irrigation. Since the amount of irrigation water supply was the same for both irrigation systems, hence increasing IWUE is an indicator for higher crop production. The increase in IWUE under sub-surface drip irrigation might be due to the minimal losses under subsurface drip compared to surface drip. In addition, no surface evaporation in sub-surface drip occurs, compared to drip irrigation (Phene et al. 1992, Ismail and Almarshadi, 2011). Soil moisture content under subsurface drip irrigation was more uniform as compared with surface drip. Soil water data of sub-surface irrigation suggest little or no potential for deep percolation losses. As a result, crop production is increased with the same amount of water supply, consequently IWUE is

increased. Support to this point can be found in Hutmacher et al. (1992) and Senyigit et al., (2011) studies. The scope of losses due to deep percolation and surface evaporation is greatly reduced by sub-surface drip irrigation as found by Alam et al. (2002). Sometimes surface drip irrigation increase yield production compared to sub-surface irrigation as found in Lettuce. The results could be due to the root distribution behavior of lettuce where they highly presented in surface layer. The Increase of IWUE in the second season compared to the first season might be due to the change in planting date, metrological conditions, length of the growing period and photo period.

Conclusion

The response of the two growing seasons to yield and yield components was different. In the first season there was almost an inverse proportionality between the agronomic traits and the irrigated wastewater quality in Cabbage crop. As wastewater quality increase, yield and yield components decrease. This proportionality does not exist in the second season where no significant differences were found in the values of all agronomic traits. The response of Lettuce crop on the other hand is quite different from that of Cabbage. There was no clear relationship between the agronomic traits and the applied water qualities. However, the Local groundwater quality, 60% and 100% wastewater qualities produced the highest yields in both seasons. Thus, there was a similarity in the response of the two seasons. On the other hand, contrasting proportionality was found under the interaction between irrigation system and wastewater quality not only in the two growing seasons but also between the two crops.

There was found high discrepancies in the IWUE values between the two seasons and less discrepancy between the two irrigation systems. In addition, the IWUE response to Cabbage and Lettuce crops was unlike. The sub-surface irrigation system produced higher value of IWUE compared to surface irrigation in Cabbage crop. However, the case is reverse in Lettuce crop. There was also a significant increase in IWUE values in the second season of Cabbage crop. This increase reaches up to two to three times. IWUE is an indicator of not only of crop yield but also water savings. In this study, a considerable saving of water resources was attained. Among all investigated wastewater treatments, 60% found the best because it saved 60% of natural clean water resources. In another word, 60% of treated wastewater which could have been disposed unsafely in the environment has been utilized in the production of vegetable crops.

Acknowledgment

This work was funded by King Abdulaziz city for Science and Technology (KACST), Saudi Arabia, under the program of NSTIP program Grant No. (10-WAT1080-3). The authors therefore acknowledge with thanks KACST and the Deanship of Scientific Research (DSR) at King Abdulaziz University (KAU) for their technical and administrative support.

References:

- Aiello R., Cirelli L.G., Consoli S. (2007). Effect of reclaimed wastewater irrigation on soil and tomato fruits: A case study in Sicily (Italy). *Agricultural Water Management*, 93: 65–72.
- Al-Abdulqader, M.A. and A.A. Al-Jaloud (2003). Economics of using treated sewage water in agriculture in Saudi Arabia, paper presented at The Sixth Gulf Water Conference held by the Water Science and Technology Association in Riyadh, Kingdom of Saudi Arabia, in 8 – 12 March, 2003.
- Alam, M., Trooien, T.P., Rogers, D.H. and Dumler, T.J. 2002. An efficient irrigation technology for alfalfa growers. *Journal of Extension*, 40, (3).
- Al-Sha'lan, S.I. (2001). Effect of treated sewage water on water movement in soil using undersurface irrigation system. A project funded by KACST, #AI 10-20, Riyadh, Saudi Arabia.
- Al-Lahham O., El Assi N.M., Fayyad M. (2003). Impact of treated wastewater irrigation on quality attributes and contamination of tomato fruits. *Agricultural Water Management*, 61: 51–62.
- Blumenthal, UJ, A Peasey, G Ruiz-Palacio and DD Mara (2000). Guidelines for wastewater reuse in agriculture and aquaculture: recommended revisions based on new research evidence. WELL Study No. 68 part 1, June 2000. <http://www.lboro.ac.uk/well/resources/well-studies/full-reports-pdf/task0068i.pdf>
- Dreschsel, P. Blumenthal, U.J. and Keraita B. (2002). Balancing health and livelihoods. Adjusting wastewater irrigation guidelines for resource-poor countries. *Urban Agriculture Magazine*. 8: 7-9

- Esmailiyan Y., Heidari M., Ghanbari A. (2008). Effect of Municipal Wastewater with Manure and Chemical Fertilizer on Grain Yield and Yield Components in Corn KoSc 704. *Journal of Agronomy*, 7: 227–280.
- Ismail, S.M. and M.H. Almarshadi, 2011. Effects of water stress applied with sub-surface drip irrigation on forage productivity and water use efficiency of alfalfa under precise irrigation practice in arid climate. *American-Eurasian Journal of Sustainable Agriculture*, 5: 97-106
- Jimenez B. (2005). Treatment technology and standards for agricultural wastewater reuse: a case study in Mexico. *Journal of Irrigation and Drainage Engineering – ASCE*, 54: 23–33. Management, 93: 65–72.
- Mara D.D. and Cain-Cross S. (1989): *Guidelines for the Safe Use of Wastewater and Excreta in Agriculture and Aquaculture - Measure for Public Health Protection*. World Health Organization, Geneva.
- Munir J.M., Ayadi M. (2005.) Forage yield and nutrient uptake as influenced by secondary treated wastewater. *Journal of Plant Nutrition*, 27: 351–356.
- Najafi P., Mousavi S.F., Feizy F. (2003). Effects of using treated municipal wastewater in irrigation of tomato. In: *Proc. ICID 20th European Regional Conf. Montpellier, September 14–19, 2003*, 8.
- Panoras A., Evgenidis G., Bladenopoulou S., Melidis V., Doitsinis A., Samaras I., Zdragkas A., Mathi Th. (2003). Corn irrigation with reclaimed municipal wastewater. *Global Nest: the International Journal*, 5: 39–45.
- Phene, C.J., Hutmacher, R.B., Ayars, J.E., Davis, K.R., Mead, R.M., Schoneman, R.A. 1992. Maximizing water use efficiency with subsurface drip irrigation. *International Summer Meeting of the American Society of Agricultural Engineers*, Paper No. 922090. St. Joseph Michigan
- Senyigit, U. A. Kadayifci, F. O. Ozdemir, H. Oz and A. Atilgan, (2011). Effects of different irrigation programs on yield and quality parameters of eggplant (*Solanum melongena* L.) under greenhouse conditions. *African Journal of Biotechnology* Vol. 10(34), pp. 6497-6503
- Vaz da Costa S.M., Bastos R.K.X., Mara D.D. (1996). Bacteriological aspects of wastewater irrigation. *Tropical Public Health Engineering (TPHE) Research Monograph no. 8*, University Leeds (Department of Civil Engineering), Leeds, UK.
- WHO (2006). *Guidelines for the safe use of wastewater, excreta and gray water. Volume 2: Wastewater use in agriculture*. Available at http://www.who.int/water_sanitation_health/wastewater/gsuweg2/en/index.html (accessed October 31, 2006).
- Zavadil, J. 2009. The Effect of Municipal Wastewater Irrigation on the Yield and Quality of Vegetables and Crops. *Soil & Water Res.*, 4: 91–103

PHYTODEPURATION PROCESS FOR THE RECYCLING OF WASTEWATER IN DAIRY

Lydia Ferrara

Chair of Hydrology Chemistry, Department of Pharmacy, University of Naples "Federico II", Naples

Abstract:

The problem of saving water is very important for the need to rationalize the use of the water resource. We have tested a method founded on phytodepuration for waste water of cheese factory, for to recover and recycle the water used in food industry, a sector characterized by high consumption of water.

The phytodepuration represents an innovative method for the disposal of wastewater. Some plants have a natural capacity to absorb and / or degrade toxic substances and contaminants to the environment, for the presence of a rhizosphere microflora, which is able to metabolize and make them available for their growth.

In the laboratory we developed a little system for phytodepuration using a plastic basin in which were placed some plants of *Cyperus papyrus* immersed in a quantity of wastewater remained constant over time. We performed the analysis before the process and then a distance a few months to verify the variation of the parameters most important pollutants. At the end of our experience, approximately one year, it was possible to draw the following conclusions: the *Cyperus papyrus* proved effective to lessen the concentration of organic substances: the value of COD was lower than the initial wastewater; the concentration of phosphorus, of zinc was decreased and the concentration of chlorides also, which remained constant until the end of the experience. The plants after the stage of adjustment is not showing signs of suffering and therefore can be considered suitable for such use.

Key Words: Wastewaters, phytodepuration, *Cyperus papyrus*

Introduction

The problem of disposing of industrial waste is a subject of constant disputes because of the power polluting the environment. In fact very often you can see streams, rivers and lakes are polluted so as to make it almost impossible to aquatic life and seriously affect the health of the surrounding cultivated areas. The blame for this degradation is commonly attributed to industrial discharging their effluents into waterways regardless of the ecological damage that can provide. It is created in such a way that environmental degradation is depleting the water resources, which until a few years ago seemed unlimited, whether because the surface waters contain higher amounts of pollutants, or because their pollution through the soil, has achieved in many cases even the underground water reserves. Whether the discharges were repaid a little doses in surface water bodies, they could be slowly degraded by aerobic microorganisms present, supported by the presence of dissolved oxygen. Otherwise the pollution load is always very high and the oxygen is insufficient for the oxidation of organic matter for the survival of aquatic species. In this case are involved the anaerobic microorganisms that establish a process of decay with the development of unpleasant odors while simultaneously appears the phenomenon of eutrophication.

The Italian legislation on waste water was low and confused until 1976, when he was adopted and published the law 319/76, known as "Legge Merli". that was subsequently supplemented by the "Technical Standards" and amended by Law 650/79, which provided subsidies for the construction of sewage treatment plants enlarging the powers of municipalities as responsible for most of the tasks performed until that point by the provinces. In 1980 and 1981 were published some ministerial directives for the regions to which were committed to regulate discharges into sewers and installations for sewage treatment.

The law Merli and all its amendments have been replaced by the D. Lgs. 152/99, currently in force, which sets maximum limits for a number of parameters that characterize a water discharge such as color, the concentration of lead, sulphides, organic solvents and other substances. These values vary depending on the water body to which it is intended the treated water, in particular, are more stringent for discharges to surface water in relation to the wastes and can be modified by the regions according to local need.

The dairy industry in Campania is a highly productive in terms of turnover and places of employment, but the pollution caused by wastewater from the product is superior to twice that of domestic waste. The milk during the process of dairies production suffers a slow pasteurization followed by the addition of rennet, which determines the coagulation of casein. It is observed as a clear separation between the solid mass, used for the subsequent production of mozzarella cheese, and whey, which is further treated to produce ricotta. After removal of the ricotta the whey waste still contains substances of very high nutritional value, and could be an additional source of income considering other possible uses. The use of the whey as food for animals, particularly pigs, is hampered by the limited pig farms in the Campania territory, unable to absorb the large amount of whey produced each day. With regard to human nutrition the whey has very good properties so it is recommended not only to athletes but also to children, women in menopause and all those people who need a further dose of mineral salts. However till now not yet developed a practical and economic method for its use as an integrator, and various applications are proposed for disposal of this waste to particularly high costs

An innovative method for the disposal of this particular wastewater rich in organic matter is the phytodepuration

Materials and methods

AOAC methods for determination of COD, organic matter, chlorides,

Conductivity meter Crison micron C.M 2200 (Crison instruments Srl Lainate-MI)

pHmeter Φ 50 at 20°C BeckmanCoulter (Cassina De' Pecchi – Milano Italia)

In laboratory was set up a simple system for the purifying using *Cyperus Papyrus*, plant typical of Mediterranean regions, which grows naturally along rivers, whose purifying capacity have been highlighted recently by several authors. The benefits that make this species interesting for the purpose of purification are easy availability, cultivation, breeding, excellent placement and environmental landscape, good adaptation to different climatic conditions

We analyzed three samples of wastewater coming respectively from the manufacture of mozzarella (I, II, III) whose characteristics are reported in Table1. The variability of the values can be attributed to the composition of the milk to start., The sample to purify was obtained by mixing equal parts of the three wastewater.

In a rectangular basin of 30 x70 cm well exposed to sunlight have been placed on a layer of 10 cm of soil 20 young shrubs and 10 liters of drinking water. The water level was maintained constant throughout the experimental phase, while stirring the water itself was performed every day to avoid the formation on the surface of a layer of fatty substances. The wastewater was analyzed before being subjected to treatment taking into account parameters such as COD, organic matter, chlorides, pH, conductivity as indices of assessment of purifying plant, as well as total phosphorus and mineral elements.

After a week were initially added 10 ml of serum every 20 days after which we taked a water sample to evaluate the parameters of pollution. The values of these parameters are given in the following graphs (Figs1,2,3)

These initial tests have allowed to evaluate the resistance of the plant in the presence of wastewater. afterwards we increased the rate of serum added, 100 ml every 20 days, and taked a sample of which we explored the various parameters. The values are given in the following graphs (Figs 4,5,6)

Samples	I	II	III
Ph	3.90	3.98	4.01
Conductivity mS/cm	16.70	17.78	17.77
Dry weight g/L	42.68	36.6	68.0
Organic g/L	13.52	12.2	11.2
COD mg/L	8000	7218	6627
Na ⁺ mg/L	370.0	395.0	550.0
K ⁺ mg/L	875	600	962.5
Cl ⁻ mg/L	2481	2127	2481
Total phosphorus mg/L	870	750	800
%Total protein mg/L	12.37	10.42	10.93
Lactose mg/L	4.6	4.3	4.4

TAB 1 Analysis of wastes

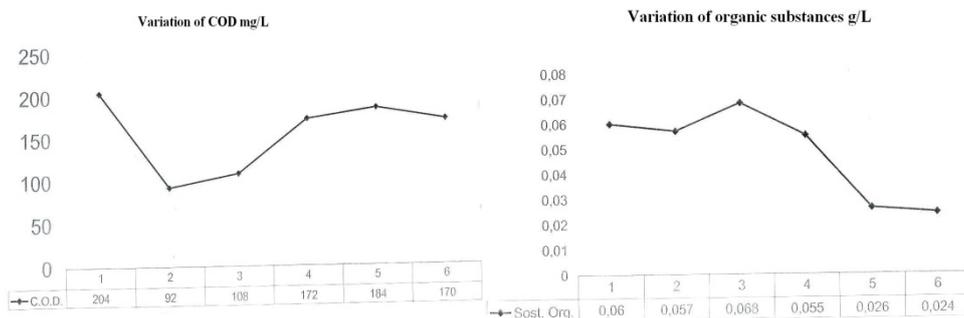


Fig. 1 COD variations

Fig. 2 Organic substances variations

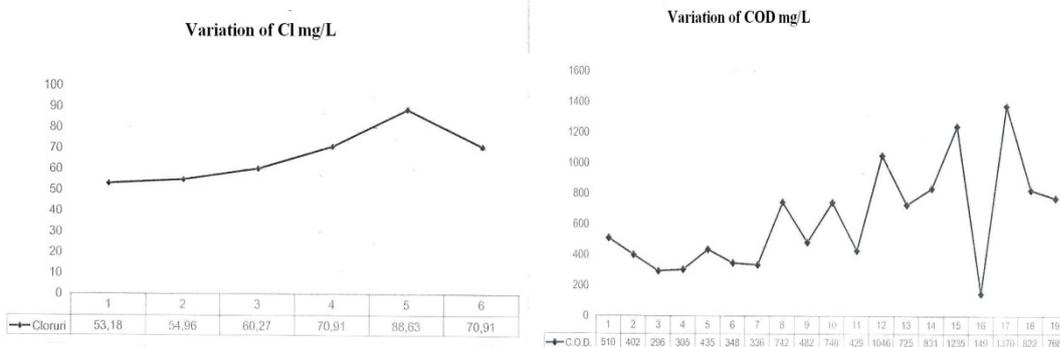


Fig. 3 Cl variations

Fig. 4 COD variations

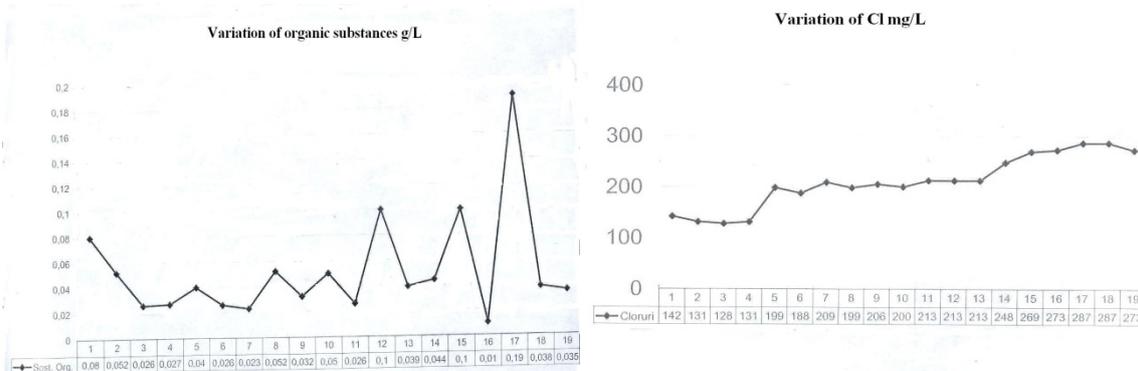


Fig. 5 Organic substances variations

Fig. 6 Cl variations

Conclusion

The techniques of phytodepuration can be successfully used for various applications, as installations for purification of waste water, or industrial wastewater depuration, or the rehabilitation of contaminated sites: their use in any case requires a preliminary assessment to check the adequacy of the chosen system

For each class of contaminants is necessary to discriminate which of the phytodepuration techniques currently known is best suited..

We can choose between phytoextraction, based on absorption of contaminants radical; phytotransformation, in which contaminants are degraded by the metabolism of the plant; phytostimulation or plant-assisted bioremediation based on the stimulation of biodegradation by microbial activity in the plant root zone; phytostabilization, using plants to reduce the mobility and migration of contaminants in soil.

The choice of plants to decontaminate soils and waters, based on the natural capacity of plants to absorb, accumulate and /or degrade, due largely to stimulation of the rhizosphere microflora, molecules in the environment in which they live. The results of research and field testing have demonstrated the applicability of such systems to a broad group of contaminants, including many metals, radionuclides and organic solvents such as chlorinated, pesticides, insecticides and explosives.

The plant species used are usually aquatic plants or highly hydrophilic since those systems are used to purify water. The process occurs through the cooperative growth of macrophytes and micro-organisms associated with them. The plants absorb the nutrients (mainly inorganic salts) present in water to be purified through the roots ; the radical development of the plant species used acts as a coupling for micro-organisms, whose activity is favored by the release of atmospheric oxygen that, absorbed by equipment of the plants is transferred to the roots and released into the surrounding environment. In practice it has a wet ecosystem in which various components plants, microorganisms, soil, solar radiation, contribute to the removal of pollutants

On the process of plant prepared in our laboratory we can make some considerations. After an initial phase during which the plant selected has shown the difficulty of adaptation, as shown in the graphs by a descent of the initial values of the parameters considered, we could record a good performance with the advance of time.

In particular, the activities of the plant compared to the COD and organic substances is very high, thanks to the constant oxygenation provided by the continuous movement in the basin

For the chlorides, however the activity of plant was initially reduced, until to reach an almost constant trend.

Note that there are no such phenomena of eutrophication due to the ability by this plant to absorb phosphorus, which is responsible with the nitrogen of this phenomenon.

References:

Official Methods of Analysis of A.O.A.C International 1990

Forgione P., Ferrara L., Schettino O. The use of *Eichornia crassipes* for the improvement of some characteristics of water *Bollettino della Società italiana di biologia sperimentale*. 1982; 58(5):241-5.

Ferrara L., Forgiane P, Schettino O... The use of aquatic plants in the correction water pollution from chemical agents, absorption of zirconio and cadmium . *Bollettino della Società Italiana di Biologia Sperimentale* . LXI (9) 1985 ,1343-8

Ferrara L., Forgione P., Schettino O., Rullo V. Purification of waste water from olive-oil pressing plants using hydrophytes. Preliminary study on the use of *Eichornia crassipes* *Bollettino della Società italiana di biologia sperimentale*. 05/1989; 65(4):303-9.

Ribeyre F., Boudou A... "Experimental study of inorganic and methylmercury bioaccumulation by four species of freshwater and sediment contamination sources". *Ecotoxicology and environmental safety*, 28, 270-286, 1994

AYAZ-AKCA. "Treatment of wastewater by natural systems". *Environ Int.*, 26 (3): 189-195, 2001

MICROALGAE – BIODIESEL POTENTIAL PRODUCERS: A REVIEW

Anna Krasowska, Prof.
Slawomir Jablonski, MA
Piotr Biniarz, MA
Malgorzata Plachetka, MA
Marcin Lukaszewicz, PhD

Faculty of Biotechnology, University of Wrocław, Wrocław, Poland

Abstract:

Declining quantity of fossil fuels force scientific community to think about alternative energy sources. Thus, it is critical to focus on renewable resources and development of new technologies. Solar energy seems to be sufficient for actual energy demand. The question is what is the most efficient method for its capture, storage and distribution. As diesel stands for the third part of fossil fuels used in transport, cultivation of microalgae and extraction of lipids out of cells is one of the possibilities. Biodiesel produced from microalgae-derived lipids offers notable environmental benefits e.g. reducing the greenhouse effect by utilization of CO₂ emissions or sewage treatment.

The review includes descriptions of species selection for biodiesel production, genetic modifications leading for to higher efficiency and production process improvement, biomass and metabolites recovery from cultures, transesterification process leading to biodiesel, waste utilization and business value.

Microalgae can potentially offer substantially higher yields than other oil-producing crops and they can grow beside fresh water also in saline water or even sewage. Additionally microalgae do not compete with productive farmland thus there is no competition with food chain.

In spite of many advantages of biodiesel production from algae, there are a lot of limitations blocking its real competition with petrodiesel. Production costs seem the most substantial problem. Out of all recognized methods of microalgae production, the culture of microalgae in photobioreactors seems the most favorable for biodiesel production, however the costs are discouraging. Probably small modular systems may be an attractive solution with positive economical rationale.

Key Words: Microalgae, biodiesel, transesterification, renewable energy

Introduction

Petroleum originating fuels are currently regarded as largely environment-hostile due to their influence on carbon dioxide accumulation in the atmosphere which contributes towards greenhouse effect formation. Moreover, petroleum resources undergo progressive stock-out for their seam renew is very slow. Deficiency of fossil fuels may result in global famine as technological development, including modern agriculture advance, is predominantly petrol dependent (Hall 2009).

Renewable fuel sources are permanently searched for, especially ones suitable for transportation. One of the resources are products originating from living organisms (animals, plants and micro-organisms). Fuels from these sources are called biofuels whereas biodiesel is a combustible material used in compression-ignition engines (by Diesel).

At present, algae seem the only recognized source which, quite potentially, is able to replace totally fossil fuels (Fig. 1) in transportation (Chisti 2007), despite the fact that existing cultivation and harvesting technologies have to be improved to meet the economical requirements for fuel production. The term “algae” comprise both prokaryotic and eukaryotic, mono- and multicellular organisms of sizes from several micrometres to over of dozen metres (Sheehan, Dunahay et al. 1998). They can occur in all fresh and salt, cold and warm waters of all geographical zones. Their production reached in 2011 up to 10000 tonnes of dry mass (Griffiths, Dicks et al. 2011). The main products recovered from the algae were high value compounds such: carotenoids, polyunsaturated fatty acids, cosmetics and nutrient supplements.

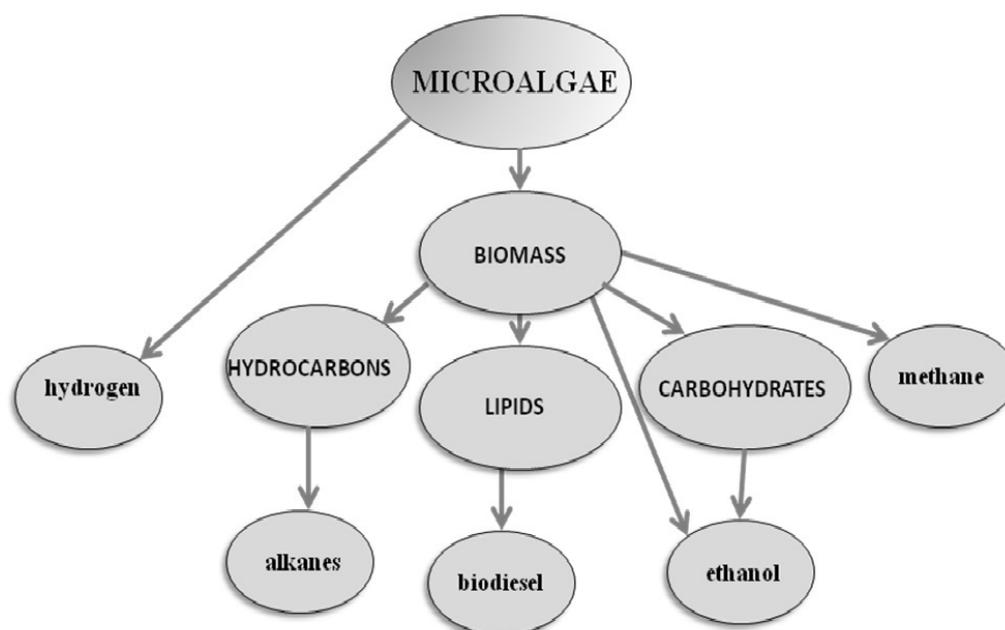


Fig. 1: Potential fuels from microalgae.

Microalgae as biodiesel sources

Microalgae is the term assigned to microscopic organisms which are several to a few hundred micrometres long. They are predominantly monocellular organisms often living in colonies. Microalgae can be the source of several types of biofuels: methane produced during anaerobic digestion of algae biomass (Spolaore, Joannis-Cassan et al. 2006), hydrogen produced photobiologically in anaerobic conditions (Ghirardi, Zhang et al. 2000) and biodiesel derived from lipids accumulated as reserve material in algae cells (Xu, Miao et al. 2006; Demirbas 2008). The most convenient fuel for transport is biodiesel, but it needs further processing of microalgae biomass for the better recovery of energy. The complete utilisation of algal biomass may involve the combination of technologies mentioned (Wiley, Campbell et al. 2011).

The algae are potentially more efficient in oil production in comparison to common oil seed crops due to the higher productivity per area. Higher oil yield arises from high biomass production rate and high lipid content (Wiley, Campbell et al. 2011).

Many among classified species of algae have growth rate below 1 d^{-1} . While algae do not have organs characteristic to the higher plants whole cell surface can be involved in photosynthesis process. Moreover higher oil yield arises from the fact that lipids are accumulated in whole cell, while in oil crops only seeds contain significant amount of oil and are collected and processed (Griffiths, Dicks et al. 2011).

Algae species capable of accumulating large amount of lipids are found in many taxonomic groups (Fig. 2). However, chlorophytae represent the biggest group within which the species with the average content of 25.5% of lipids in dry biomass have been identified. Examinations of cyanobacteria did not reveal highly oil-bearing species. In this group lipids average contents in dry mass reached only 9.8% and did not show neutral lipids accumulation (Basova 2005; Hu, Sommerfeld et al. 2008).

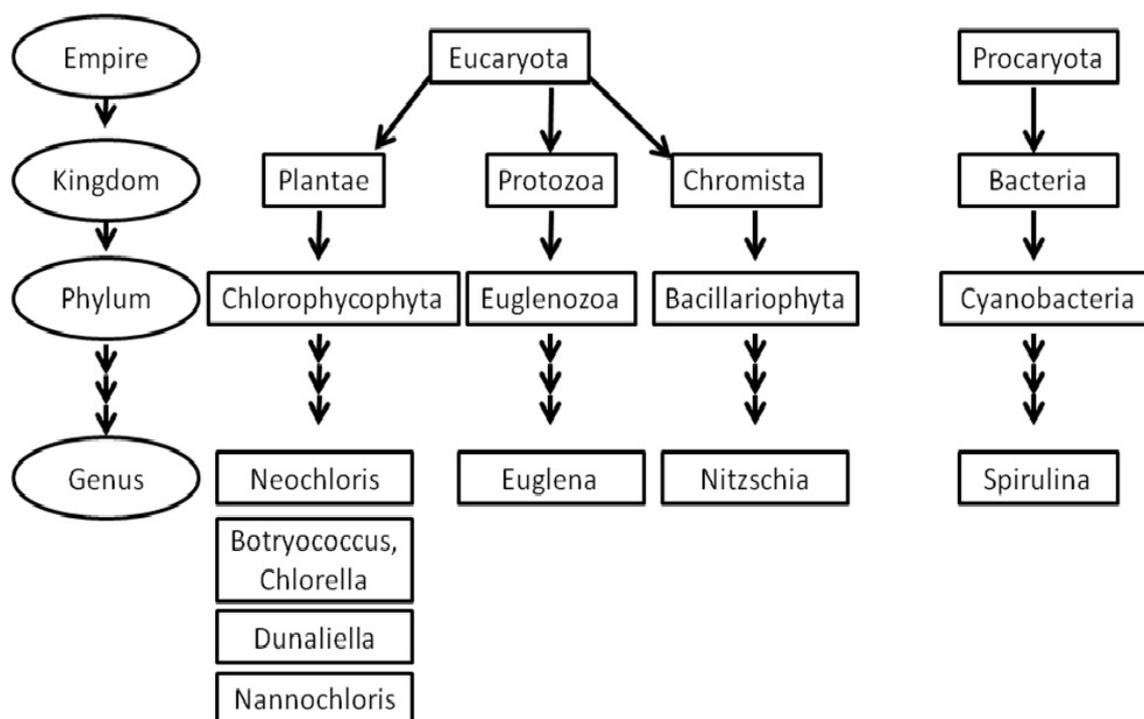


Fig. 2: Taxonomy of some algae species potentially useful for biodiesel production.

Algae is a large group of polyphyletic organisms. Most of the species potentially useful for biodiesel production belongs to green algae (*Chlorophycophyta*) and diatoms (*Bacillariophyta*). Although Cyanobacteria species accumulate lower levels of lipids (e.g. *Spirulina* 9%) but within this group there are species which are able to fix atmospheric nitrogen which could be useful for culture nutrition. Taxonomic classification based on (Guiry and Guiry 2010)

Moreover the amount of lipids in chlorophytae may be raised up to 45% of dry weight by stress or nutrient starvation (Hu, Sommerfeld et al. 2008). The increase of lipid concentration in stress exposed and ageing cells refers mainly to neutral lipids and triacylglycerols in particular. The observed phenomenon is a result of lipids metabolism shift from membrane lipids synthesis to neutral lipids storage. This can cause the increase of triacylglycerols (TCG) synthesis *de novo*, as well as the conversion of present membrane lipids to TCG. As a result, triacylglycerols may constitute up to 80% of lipids total contents in the cell (Hu, Sommerfeld et al. 2008). It is very important to notify that high lipid content does not necessarily reflects the overall lipid production (Griffiths, Dicks et al. 2011). During the starvation period the growth rate is reduced and the total lipid production may be lower in comparison to well nourished culture.

On the other hand the lipid composition has considerable influence on the technology of biodiesel production and product quality (Li, Du et al. 2013). Lipids derived from algae cultured without stress contain significant amounts of polar lipids (phospholipids and glycolipids) and limited content of TCG (up to 40 % of total lipids) (Harwood and Guschina 2009; Wang and Wang 2012). The best material for biodiesel production are TCG while polar lipids are unfavourable since they are a cause of emulsification and catalyst depletion (Mendow, Monella et al. 2011). Lipids other than TCG may also reduce the fuel quality by increasing the content of sulphur and phosphorus (Mendow, Monella et al. 2011). Despite reduced growth rate and total lipid production rate, starvation of algae may be beneficial due to the increased content of TCG.

Fatty acids composition in algal cells

Algae produce both saturated and unsaturated fatty acids with different number and position of unsaturated bonds and various length of carboxylic chain. However, both saturated and unsaturated fatty acids with even number of carbon atoms prevail there (Cobelas and Lechado 1989; Makri, Bellou et al. 2011).

Fatty acids in algae are more diverse in comparison to higher plants. Acids with three or more double bonds (up to six) are present. Another special feature of algae is high content of polyunsaturated fatty acids (PUFA) with very long chains (longer than C22) (Hu, Sommerfeld et al.

2008). Algae are the source of DHA rich oil used in vegetarian diet supplementation instead of cod liver oil (Masuda, Tanaka et al. 2003; Pyle, Garcia et al. 2008).

While algae could be excellent source of PUFA in dietary supplementation, in biodiesel production the amounts of fatty acids with four or more double bonds should be as small as possible. Such acids as well as their esters are definitely more susceptible to oxidation during fuel storage decreasing their quality (Fukuda, Kondo et al. 2001; Chisti 2007; Hu, Sommerfeld et al. 2008). It seems especially important in the case of biodiesel which is to be used in vehicles. European standards (Standard EN 14214) allow only 1% mol contents of methylic or ethylic esters of 4 or more double bonds fatty acids (Knothe 2006). Many of algae derived oils are not up to this standards. However, they can be used in biodiesel production due to partial catalytic lipids hydrogenation – the technology used during margarine production but increasing the process costs (Jang, Jung et al. 2005; Dijkstra 2006)

Algae species selection for biodiesel production

The amount of lipids in algae biomass varies largely between species. For the last decades, thousands of algae have been analysed in respect of lipids content and culture facility both in a laboratory and on a industrial scale (Hu, Sommerfeld et al. 2008). The collection of 3000 species of high lipids productivity have been made up which, after tests, isolation and thorough characteristics were limited to 300 most valuable species including mainly green algae and diatoms (Sheehan, Dunahay et al. 1998).

Rodolfi's team (Rodolfi, Chini Zittelli et al. 2009) examined 30 algae strains isolated from both fresh and salt waters. They were initially cultured in a laboratory in order to define the strains characteristic for relatively big efficacy in biomass production as well as high lipids content. Elicited results allowed to select two fresh-water species (*Chlorella* sp. F&M-M48 and *Scenedesmus* sp. DM) and two salt-water species (*Nannochloropsis* sp. F&M-M24 and *Tetraselmis suecica* F&M-M33). These microorganisms were subsequently moved to culture tubes under artificial light to examine the influence of limited access to nitrogen and phosphorus. The highest lipids accumulation simultaneously with the smallest biomass production drop was revealed for *Nannochloropsis* sp. F&M-M24. Observation of light intensity and starvation reaction in a mount rack photobioreactor was another stage followed by the culture movement to 110 litre photobioreactor GWP type (Green Wall Panel) to carry on an experiment in conditions close to large-scale culture.

Described methodology is suitable for the screening of a large number of algae strains. For strains with high growth rate and lipid accumulation the economical simulation can be prepared. In order to do this, NER (Net Energy Ratio) is calculated. NER is the ratio between the produced energy (the energy embedded in lipids used in biodiesel production and potential energy from biomass remains) and the energy introduced to the system during the cultivation and processing of biomass. If the value of NER for particulate system is higher than 1 it may be economically profitable (Jorquera, Kiperstok et al. 2010).

Genetic modifications

The genetic modifications of algae are not necessary to form achieve relatively high lipids contents and productivity. However, they can be used for the significant increase of processes business value (by e.g. elimination of photoinhibition) (Rodolfi, Chini Zittelli et al. 2009).

The use of genetically modified algae may be helpful in improving culture stability. In large scale, the contamination of selected strains often occurs. Competition between organisms reduces production efficiency. In cultures of algae resistant to herbicides the growth of undesirable microorganisms could be hindered with this chemical compounds (Gressel 2008).

Algae are very sensitive to temperature shifts. Maintaining stable temperature in cultivation vessels (heating and cooling during sunny days) is very expensive, thus obtaining strains resistant to this changes would enable the reduction of costs (Shlyk-Kerner, Samish et al. 2006; Gressel 2008).

The ability of light utilisation by algae is characterized by a light saturation constant, that is the intensity of light at which the specific biomass growth rate is half of its maximum. Light saturation constants for microalgae tend to be much lower than the maximum sunlight level that occurs at midday. Moreover intense light can lead to photoinhibition resulting from damage of the photosynthetic apparatus to the reduction of growth rate. Uneven light intensity (the highest at the surface) is also the problem of photobioreactors. Adequate genetic modifications can unsensitize photosynthetic apparatus to intense light and suppress photosaturation and photoinhibition (Chisti

2007). On the other hand, in low light intensities growth of modified algae will be reduced in comparison to wild type organisms (Polle, Kanakagiri et al. 2003; Gressel 2008). Green algae have a tendency to assemble large arrays of light-absorbing chlorophyll antenna molecules in their photosystems, which in higher light intensity absorb more photons than photosynthesis can utilize, resulting in dissipation of light energy by the first layers of cells. To advance light access to the cells located in “dark zone”, antenna size could be reduced. Decreasing antennas size (e.g. by chlorophyll particles number half-restriction), light access to deeper culture layers is considerably improved. Besides, such a modification would definitely enlarge photosaturation constant value which would allow cells further growth at high light intensity (Polle, Kanakagiri et al. 2003).

Lipids accumulation can be stimulated by redirecting metabolic pathways to lipids from starch. In *Chlamydomonas reinhardtii* when starch biosynthesis was blocked lipids bodies content increased 30-fold (Wang, Ullrich et al. 2009). In addition, *C. reinhardtii* used in this experiment was cell wall-less mutant (Davies and Plaskitt 1971) enabling much easier and cheaper lipids extraction.

Lipids composition and structure can be improved by mutation within desaturases. The result of such a procedure is flax cultivar called Linola with modified fatty acids proportions (Łukaszewicz, Szopa et al. 2004). In the case of algae for biodiesel production, modification (repression) should aim at FAD2 desaturase equivalent. In higher plants species this enzyme is responsible for the synthesis of PUFA (Krasowska, Dziadkowiec et al. 2007). The elimination of this protein may result in the reduction of PUFA and increase of monounsaturated fatty acids with chain average length like oleic acid (18:1).

The modification of metabolite synthesis gives great possibilities being intensively developing domain on both basic and applied researches levels. Quite theoretically, there is possibility to obtain fatty acids ethylic esters directly in the cell which would significantly decrease biofuel production costs and solve the problem of glycerol disposal being waste product on biodiesel production.

Biodiesel synthesis out of triacylglycerols

Biodiesel, being methylic or ethylic esters of fatty acids, is synthesized in the process of transesterification of TCG with methanol or ethanol. This reaction can also be based on other alcohols like propanol, butanol or amyl alcohol, however, methanol and ethanol are predominantly used due to lower costs (Fukuda, Kondo et al. 2001). Methanol is most often produced out of natural gas or coal so, in contrast to ethyl ones, methyl esters cannot be fully derived from renewable energy sources. Transesterification needs 3 alcohol molecules for every TCG to produce 3 molecules of methyl esters. To achieve 95% efficacy of esters generation, the reaction is performed in alcohol significant excess (Fukuda, Kondo et al. 2001). Thus, there are still many possibilities of the process optimization.

Transesterification catalysis methods

TCG transesterification reaction may undergo catalysis in three ways: by the use of acids, bases or enzymes (Sharma, Chisti et al. 2001; Meher, Vidya Sagar et al. 2006). Only recently, transesterification process optimization has been found to be possible with microwaves application. Acidic catalysis is slow but it is mostly suitable for transesterification of oils with water and high content of free fatty acids (Fukuda, Kondo et al. 2001).

Basic catalysis is the highest-speed transesterification method (it arrives 4000 times faster than acidic catalysis with the use of the same amount of catalysts), however, it can be applied only in the case of water-free oils, otherwise saponification occurs and newly formed soaps decrease catalysis effectiveness and disturb glycerol separation from post reaction mixture. Recently basic transesterification method was optimised enabling single transesterification stage, lack of industrial wastes, the bleaching stage without absorbers and possibility of using raw material “first-pressing” (Kołodziej, Vogt et al. 2008).

Enzymatic catalysis with the use of lipases is also possible, however, its costs are very high (Fukuda, Kondo et al. 2001; Li, Du et al. 2013). Only biocatalysis with whole BSP (Biomass Support Particles) immobilised cells use provides lipases with high stability and long activity thus the whole process seems simple and easy to apply in industry (Ban, Kaieda et al. 2001; Ban, Hama et al. 2002).

Biomass and metabolites recovery from algae cultures

To collect TCG out of algae cultures, the following steps should be executed: 1) recovery of biomass out of the culture, 2) extraction of the compound, 3) purification of the compounds from crude extract (Molina Grima, Belarbi et al. 2003). There are three basic methods of biomass recovery:

sedimentation, filtration and centrifugation, all of them being burdened with substantial defects. Each of the methods can be preceded by flocculation (Uduman, Qi et al. 2010). Other techniques, like electrolytic methods are still being investigated and are used to a lesser extent (Show, Lee et al. 2012). Choosing the method, the features such as biomass density, volume and microalgae cells size (most often of 3-30 μm diameter) should be considered (Molina Grima, Belarbi et al. 2003). It is important to point out that biomass recovery may amount to 20-40% of the production total costs (Gudin and Therpenier 1986), so the optimization of this step might be essential for increasing the production of biodiesel from microalgae.

The process of filtration brings positive effects in the case of small amounts of biomass, especially when microorganisms used for biofuel production are bigger than bacteria. Unfortunately, the species which seem potentially promising in biodiesel production (e.g. *Chlorella* sp. or *Dunaliella* sp.) have the sizes similar to those of bacteria's, so filters regain requires relatively high energy and equipment input, and the process is relatively slow (Molina Grima, Belarbi et al. 2003). Although, as the membrane filtration is an intensively developing and more and more widely used in industry, the adaptation of this technology for the larger amounts of algal biomass seems of great promise (He, Bagley et al. 2012).

Flocculation reinforced sedimentation is a widely used method in sewage-works which potentially requires the smallest input (Shelef 1978). After such a waste treatment process, biomass contain large amount of water which increases dehydration costs (Mohn 1978). Exact and precise dehydration is especially important at base catalysis due to saponification reactions (Fukuda, Kondo et al. 2001).

Centrifugation is the most effective but also most expensive method for the recovery of biomass from algae cultures. This process is short and recovered biomass includes small amounts of water (Molina Grima, Belarbi et al. 2003). For the recovery of more than 95% of cells the centrifugal force as high as 13 000g is needed. Therefore, harvesting of microorganisms is energy-consuming. The centrifugation of 1 m^3 of algae culture consumes 1 – 1.3 kW/h (depending on the system used) whereas the filtration consumes from 2 to 3 times less energy (Sim, Goh et al. 1988). With reducing centrifuge acceleration, the amount of harvested biomass is quickly decreasing. At 6000g only 60% of cells are recovered, which is why it is of no use in the reduction of biomass recovery costs.

Collected biomass should be immediately subjected to further processes, otherwise it can get spoiled within several hours. The first stage of biomass transformation is dehydration, which is another, high energy-consuming process (Show, Lee et al. 2012). There are several methods of microalgae biomass dehydration: sun rays seasoning, spray drying, sublimation drying or drying drums application. Each of these techniques has its own merits and drawbacks. Selection of the method for biomass drying should be therefore considered carefully taking into account the scale of operation (Show, Lee et al. 2012).

Lipid containing cells should be broken in order to extract lipids which, in massive processes, is based on homogenizers application. It allows far more efficient extraction of lipids out of biomass (Chisti and Moo-Young 1986; Middelberg 1994). Oil extraction from disrupted cells can be performed with organic solvents like hexane (Chisti 2008). If the remaining biomass is considered as a feedstock source, organic solvents should be replaced with non-toxic ones e.g. supercritical CO_2 can be used as an extracting agent (Pereira and Meireles 2010; Soh and Zimmerman 2011).

Wastes recycling

Contemporary technologies should be complete and consider disposal of all substrates and their products (Fig. 3). Potentially, the biomass remaining after the extraction of lipids contains large amounts of proteins. In order to make full use of them for animal feed purpose, they cannot contain insanitary residues e.g. organic solvents used in lipids extraction. Additionally, seaweeds exploited species should be approved safe and applicable for animal feeding and alimentary purposes to procure added value from the biomass remaining after the extraction of lipids.

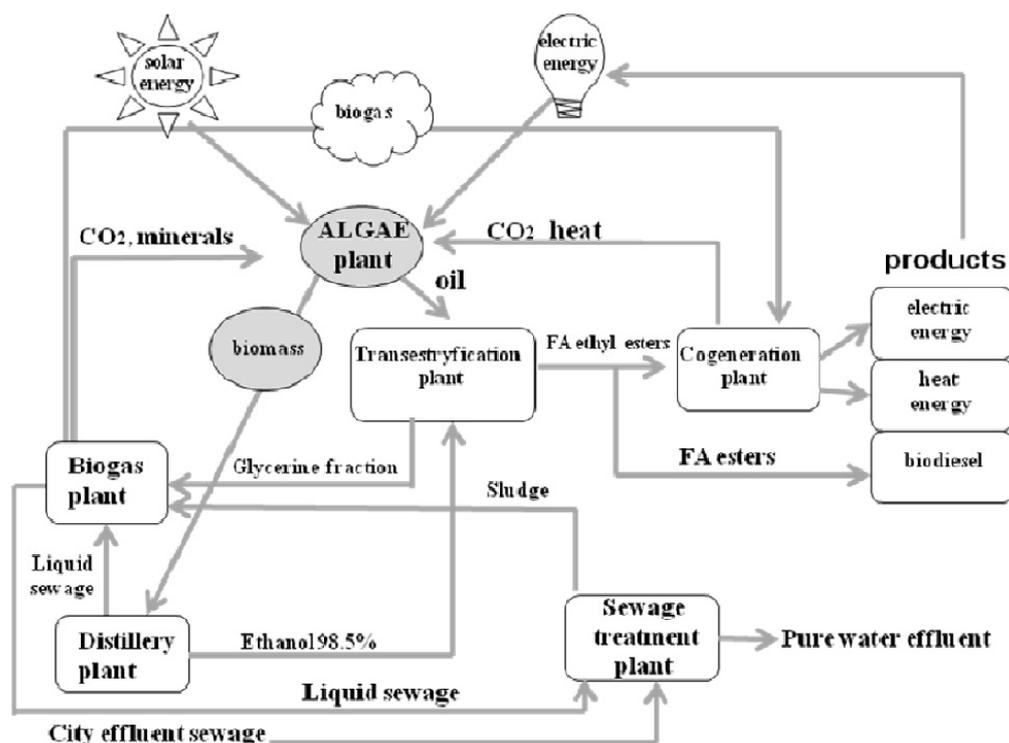


Fig. 3: Global-scale concept of the waste treatment and the biofuel-energy cogeneration centre

Another by-product in algae lipid processing is glycerol from transesterification. Recent technology of glycerol application in fodder-yeasts has been introduced by Skotan (<http://www.skotansa.pl>). The method has been invented and elaborated at Wrocław University of Environmental and Life Sciences with the use of *Yarrowia lipolytica* species (Rywińska and Rymowicz 2010). The surveys carried out by FDA have revealed *Y. lipolytica* as GRAS (generally recognized as safe) and confirmed applicability in e.g. animal feeding (registered by FDA 21 CFR 170.36). These types of tests should be also carried on algae which could be used in biofuels production.

Glycerol can be also used as a substrate for production of biogas. Addition of glycerol efficiently improved production of biogas from substrates such as: sewage sludge (Fountoulakis and Manios 2009) and cattle slurry (Robra, Serpa da Cruz et al. 2010). It is possible that anaerobic digestion of biomass obtained from algae could be improved with glycerol supplementation. Another interesting possibility of glycerol recycling is microbiological conversion of this by-product to more valuable compounds such as: citric acid, erythritol (Rywińska and Rymowicz 2010), ethanol and butanol (Yazdani and Gonzalez 2007).

Association of biodiesel production from algae with anaerobic digestion process could arise other benefits. The utilization of process water from the production of biogas is a serious problem. High concentration of biogenic elements such as nitrogen and phosphorus makes this water dangerous for the environment. On the other hand these elements are essential for growth of microorganisms and could be used to improve the cultivation of algae. Also, sewage from other industrial processes and household wastewater can be potentially used as a source of nutrients for microalgae.

The second method for increasing the production of algae biomass is recirculation of carbon dioxide produced in process of anaerobic digestion. Photosynthesis efficiency is often limited due to low concentration of CO₂, and photooxidation process which is the effect of low rubisco specificity. Previously, algae strains that can tolerate up to 12% CO₂ were identified (Pulz 2001). Increased availability of CO₂ should reduce this energy-consuming process and improve biomass production.

Business value of algae derived biodiesel

The production of biodiesel from algae is relatively expensive. The production costs of 1 kg of algae biomass amounts from 2.95 to 3.80\$ per kg (Chisti 2007), but may be decreased in the future to 0.34\$. Seabiotic LTD from Israel is one of enterprise which have work intensely to achieve comparatively low costs of dry algae production (<http://www.seabiotic.com/>). The costs of growing

and processing of algal biomass for the providing of 1 L of oil were therefore estimated between 1.40\$ and 1.81\$. To get the final price of algae-derived biodiesel these costs have to be doubled (for processing, distribution, etc.), which makes from 2.80\$ to 3.62\$ per L (Chisti 2007). This numbers should be compared with the price of traditional biodiesel (B99-B100) as high as 1.12\$ per L in January 2013 (1) or petroleum price of approximately 105\$ per barrel (0.88\$ per L) in March 2013 (2). In the situation of economic stagnation and problems in public finances of majority of European countries and United States, the production biofuels derived from algae seems rather doubtful due to difficulties with capital raising.

On the other hand, scale-down rather than scale-up has been observed in industry lately. In industrial output, a series of parallel operating microbioreactors can be applied which would prove extremely effective if carried from laboratory to industrial scale. It gives basis for the presumption that small-scale modular technologies adapted to particular conditions will prove very useful in future and will decrease costs of the production of biodiesel for local use. Additionally, use of different wastes as a source of nutrients and carbon for microorganisms can increase the profitability of algae biodiesel production. Also local governments show interest in the idea of surface waters treatment. The combination of these factors can make small investment easier for financing.

Conclusions

In spite of many advantages of biodiesel production from algae, there are a lot of limitations blocking its real competition with petroleum derived diesel. Production costs seem the most substantial problem. Although many companies culture algae on a large commercial scale (omega-3 and omega-6 acids production – Eau Plus or cosmetic component – Fitoplancton Marino), microalgae culture for biodiesel production still remains within small-scale laboratory interests. There are many companies working upon R&D but no project concerning biofuels derivation from algae has started so far on industrial scale (Wagner 2007).

Out of all recognized methods of microalgae production, the culture in photobioreactors seems the most favourable for biodiesel production, however the costs are discouraging. The investments cost of whole plant is one of the major limiting factors. It seems that small modular systems may be an attractive solution with positive economical rationale. Algae production plant should also be integrated with other existing installation like sewage treatment, biogas and power cogeneration plant (Fig. 3). Co-localisation these production processes results in lowered operating and investments costs. Also, such a waste treatment and the biofuel-energy centre generate a number of additional by-products that can be sold on or used for self-consumed.

Algae derived biodiesel is environmentally friendly which in the case of petroleum shortage may replace petrodiesel. However, to make this product cost-effective, petroleum price should rise or the costs of biodiesel derivation from algae should drop significantly. The presence of alternative technologies of fuels production are optimistic in the face of discussion about petroleum stock out.

References:

- (1). Clean Cities Alternative Fuel Price Report; http://www.afdc.energy.gov/pdfs/afpr_jan_12.pdf
 - (2). "OPEC Basket Price." from http://www.opec.org/opec_web/en/data_graphs/40.htm.
- Ban, K., S. Hama, et al. (2002). "Repeated use of whole-cell biocatalysts immobilized within biomass support particles for biodiesel fuel production." *Journal of Molecular Catalysis B: Enzymatic* 17: 157-165.
- Ban, K., M. Kaieda, et al. (2001). "Whole cell biocatalyst for biodiesel fuel production utilizing *Rhizopus oryzae* cells immobilized within biomass support particles." *Biochemical Engineering Journal* 8(1): 39-43.
- Basova, M. M. (2005). "Fatty acid composition of lipids in microalgae." *International Journal on Algae* 7: 33-57.
- Chisti, Y. (2007). "Biodiesel from microalgae." *Biotechnology Advances* 25(3): 294-306.
- Chisti, Y. (2008). "Biodiesel from microalgae beats bioethanol." *Trends in Biotechnology* 26(3): 126-131.
- Chisti, Y. and M. Moo-Young (1986). "Disruption of microbial cells for intracellular products." *Enzyme and Microbial Technology* 8(4): 194-204.
- Cobelas, M. A. and J. Z. Lechado (1989). "Lipids in microalgae. A review. I Biochemistry." *Grasas y Aceites* 40: 118-145.

- Davies, D. R. and A. Plaskitt (1971). "Genetical and structural analyses of cell-wall formation in *Chlamydomonas reinhardtii*." *Genetics Research* 17(01): 33-43.
- Demirbas, A. (2008). "Production of Biodiesel from Algae Oils." *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 31(2): 163-168.
- Dijkstra, A. J. (2006). "Revisiting the formation of trans isomers during partial hydrogenation of triacylglycerol oils." *European Journal of Lipid Science and Technology* 108(3): 249-264.
- Fountoulakis, M. S. and T. Manios (2009). "Enhanced methane and hydrogen production from municipal solid waste and agro-industrial by-products co-digested with crude glycerol." *Bioresource Technology* 100(12): 3043-3047.
- Fukuda, H., A. Kondo, et al. (2001). "Biodiesel fuel production by transesterification of oils." *Journal of Bioscience and Bioengineering* 92(5): 405-416.
- Ghirardi, M. L., L. Zhang, et al. (2000). "Microalgae: a green source of renewable H₂." *Trends in Biotechnology* 18(12): 506-511.
- Gressel, J. (2008). "Transgenics are imperative for biofuel crops." *Plant Science* 174(3): 246-263.
- Griffiths, M. J., R. G. Dicks, et al. (2011). *Advantages and Challenges of Microalgae as a Source of Oil for Biodiesel. Biodiesel - Feedstocks and Processing Technologies*. M. Stoytcheva and G. Montero. CC BY 3.0 license.
- Gudin, C. and C. Therpenier (1986). "Bioconversion of solar energy into organic chemicals by microalgae." *Adv Biotech Proc* 6: 73-110.
- Guiry, M. D. and G. M. Guiry. (2010). "AlgaeBase." World-wide electronic publication, from <http://www.algaebase.org>.
- Hall, C. A. S., Day J.W. (2009). "Revisiting the Limits to Growth After Peak Oil." *American Scientist* 97(3): 230-237.
- Harwood, J. L. and I. A. Guschina (2009). "The versatility of algae and their lipid metabolism." *Biochimie* 91(6): 679-684.
- He, Y., D. M. Bagley, et al. (2012). "Recent advances in membrane technologies for biorefining and bioenergy production." *Biotechnology Advances* 30(4): 817-858.
- Hu, Q., M. Sommerfeld, et al. (2008). "Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances." *The Plant Journal* 54(4): 621-639.
- Jang, E. S., M. Y. Jung, et al. (2005). "Hydrogenation for Low Trans and High Conjugated Fatty Acids." *Comprehensive Reviews in Food Science and Food Safety* 4(1): 22-30.
- Jorquera, O., A. Kiperstok, et al. (2010). "Comparative energy life-cycle analyses of microalgal biomass production in open ponds and photobioreactors." *Bioresource Technology* 101(4): 1406-1413.
- Knothe, G. (2006). "Analyzing biodiesel: standards and other methods." *Journal of the American Oil Chemists' Society* 83(10): 823-833.
- Kołodziej, H. A., A. Vogt, et al. (2008). "Sposób wytwarzania estrów etylowych lub metylowych wyższych kwasów tłuszczowych oraz instalacja do realizacji tego sposobu." P-386610.
- Krasowska, A., D. Dziadkowiec, et al. (2007). "Cloning of Flax Oleic Fatty Acid Desaturase and Its Expression in Yeast." *Journal of the American Oil Chemists' Society* 84(9): 809-816.
- Li, Y., W. Du, et al. (2013). "Effect of phospholipids on free lipase-mediated methanolysis for biodiesel production." *Journal of Molecular Catalysis B: Enzymatic* 91(0): 67-71.
- Łukaszewicz, M., J. Szopa, et al. (2004). "Susceptibility of lipids from different flax cultivars to peroxidation and its lowering by added antioxidants." *Food Chemistry* 88(2): 225-231.
- Makri, A., S. Bellou, et al. (2011). "Lipid synthesized by micro-algae grown in laboratory- and industrial-scale bioreactors." *Engineering in Life Sciences* 11(1): 52-58.
- Masuda, T., A. Tanaka, et al. (2003). "Chlorophyll antenna size adjustments by irradiance in *Dunaliella salina*; involve coordinate regulation of chlorophyll *a* and oxygenase (CAO) and *Lhcb* gene expression." *Plant Molecular Biology* 51(5): 757-771.
- Meher, L. C., D. Vidya Sagar, et al. (2006). "Technical aspects of biodiesel production by transesterification—a review." *Renewable and Sustainable Energy Reviews* 10(3): 248-268.
- Mendow, G., F. C. Monella, et al. (2011). "Biodiesel production from non-degummed vegetable oils: Phosphorus balance throughout the process." *Fuel Processing Technology* 92(5): 864-870.

- Middelberg, A. P. J. (1994). The release of intracellular bioproducts. *Bioseparation and bioprocessing: a handbook*, 2. G. Subramanian. Weinheim, Wiley-VCH: 131-164.
- Mohn, F.H. (1978). "Improved technologies for the harvesting and processing of microalgae and their impact on production costs." *Arch Hydrobiol Beih Ergeb Limnol* 1:228-53.
- Molina Grima, E., E. H. Belarbi, et al. (2003). "Recovery of microalgal biomass and metabolites: process options and economics." *Biotechnology Advances* 20(8): 491-515.
- Pereira, C. and M. Meireles (2010). "Supercritical Fluid Extraction of Bioactive Compounds: Fundamentals, Applications and Economic Perspectives." *Food and Bioprocess Technology* 3(3): 340-372.
- Polle, J. E. W., S.-D. Kanakagiri, et al. (2003). "tla1, a DNA insertional transformant of the green alga *Chlamydomonas reinhardtii* with a truncated light-harvesting chlorophyll antenna size." *Planta* 217(1): 49-59.
- Pulz, O. (2001). "Photobioreactors: production systems for phototrophic microorganisms." *Applied Microbiology and Biotechnology* 57(3): 287-293.
- Pyle, D. J., R. A. Garcia, et al. (2008). "Producing Docosahexaenoic Acid (DHA)-Rich Algae from Biodiesel-Derived Crude Glycerol: Effects of Impurities on DHA Production and Algal Biomass Composition." *Journal of Agricultural and Food Chemistry* 56(11): 3933-3939.
- Robra, S., R. Serpa da Cruz, et al. (2010). "Generation of biogas using crude glycerin from biodiesel production as a supplement to cattle slurry." *Biomass and Bioenergy* 34(9): 1330-1335.
- Rodolfi, L., G. Chini Zittelli, et al. (2009). "Microalgae for oil: Strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor." *Biotechnology and Bioengineering* 102(1): 100-112.
- Rywińska, A. and W. Rymowicz (2010). "High-yield production of citric acid by *Yarrowia lipolytica* on glycerol in repeated-batch bioreactors." *Journal of Industrial Microbiology & Biotechnology* 37(5): 431-435.
- Sharma, R., Y. Chisti, et al. (2001). "Production, purification, characterization, and applications of lipases." *Biotechnology Advances* 19(8): 627-662.
- Sheehan, J., T. Dunahay, et al. (1998). A look back at the U.S. Department of Energy's Aquatic Species Program: biodiesel from algae.
- Shelef, G. (1978). "Photosynthetic biomass production from sewage." *Arch Hydrobiol Beih* 11(3-14).
- Shlyk-Kerner, O., I. Samish, et al. (2006). "Protein flexibility acclimatizes photosynthetic energy conversion to the ambient temperature." *Nature* 442(7104): 827-830.
- Show, K.-Y., D.-J. Lee, et al. (2012). "Algal biomass dehydration." *Bioresource Technology*.
- Sim, T. S., A. Goh, et al. (1988). "Comparison of centrifugation, dissolved air flotation and drum filtration techniques for harvesting sewage-grown algae." *Biomass* 16(1): 51-62.
- Soh, L. and J. Zimmerman (2011). "Biodiesel production: the potential of algal lipids extracted with supercritical carbon dioxide." *Green Chemistry* 13(6): 1422-1429.
- Spolaore, P., C. Joannis-Cassan, et al. (2006). "Commercial applications of microalgae." *Journal of Bioscience and Bioengineering* 101(2): 87-96.
- Uduman, N., Y. Qi, et al. (2010). "Dewatering of microalgal cultures: A major bottleneck to algae-based fuels." *Journal of Renewable and Sustainable Energy* 2(1): 012701-15.
- Wagner, L. (2007). Biodiesel from Algae oil. Mora Associates Research Report
- Wang, G. and T. Wang (2012). "Characterization of Lipid Components in Two Microalgae for Biofuel Application." *Journal of the American Oil Chemists' Society* 89(1): 135-143.
- Wang, Z. T., N. Ullrich, et al. (2009). "Algal lipid bodies: stress induction, purification, and biochemical characterization in wild-type and starchless *Chlamydomonas reinhardtii*." *Eukaryot Cell* 8(12): 1856-68.
- Wiley, P. E., J. E. Campbell, et al. (2011). "Production of biodiesel and biogas from algae: a review of process train options." *Water Environ Res* 83(4): 326-38.
- Xu, H., X. Miao, et al. (2006). "High quality biodiesel production from a microalga *Chlorella protothecoides* by heterotrophic growth in fermenters." *Journal of Biotechnology* 126(4): 499-507.
- Yazdani, S. S. and R. Gonzalez (2007). "Anaerobic fermentation of glycerol: a path to economic viability for the biofuels industry." *Current Opinion in Biotechnology* 18(3): 213-219.

THE UNCERTAIN UNCERTAINTY OF RIVER BASINS: ACCOUNTING FOR UNCERTAINTY IN INTEGRATED WATER RESOURCES MANAGEMENT (IWRM)

Sarah Allen, Candidate for MA of Global Governance

Balsillie School of International Affairs, University of Waterloo, Waterloo, Ontario, Canada

Abstract:

Integrated Water Resources Management was officially established under the 1992 Dublin Principles. A set of standard principles was thought to be the ‘state-of-the-art’ solution to problems of ensuring good management, governance, distribution, and usage across the world. Therefore, this ‘perfect’ solution has become a leading discourse in water governance and management. However, current water related problems leave the Principles unrealistic and too idealistic for real success. Without responding to the lack of transparency and accountability, IWRM will not have the capacity to manage uncertainty in the future as global environmental changes continue. Following a case study from the Mekong River Basin, this paper will discuss how IWRM does not accurately account for issues of uncertainty in the water system and provides a plausible solution to this problem.

Key Words: River basins, uncertainty, water resources management

Introduction

Water governance and management practices have been changing drastically over the past few decades to accommodate the world’s growing population with a finite water supply. Partnerships, networks, frameworks, and dialogues, are implemented at each level of water governance, attempting to create optimal management strategies. The water sector is far from reaching these goals and “though the water sector has lagged behind in explicitly addressing water challenges in a governance framework, ‘fixing’ various water-related challenges...is now increasingly seen in terms of getting the ‘right’ governance system in place.”³ Implementing Integrated Water Resource Management (IWRM) principles into water governance and management strategies is seen as the ‘state-of-the-art’ solution to water related problems. However, this method to ‘fix’ water-related challenges has been met with varying criticisms.

IWRM of today is based on the Dublin Principles of 1992 and were developed at the World Summit in Rio. Throughout this paper I will attempt to address the perception that IWRM is too idealistic and does not cope nor account for uncertainty within the water system.

Developing IWRM in 1992 was an important step towards improving water governance. Environmental movements of the 1960s and 1970s opened the eyes of many within the water sector to the future of the world’s vulnerable resource. Standard principles that could ensure good management, governance, distribution, and usage across the world seemed the practical solution to a complicated problem. Water is important and we are facing a management crisis and needs to be addressed sooner rather than later.

Using a case study of the Mekong River Basin (MRB) will solidify my argument that IWRM does not consider the consequences or account for issues of uncertainty. Poor implementation of the principles in regions without adequate transparency and accountability makes them almost irrelevant fixtures in policy documents and signed agreements. But, at the same time, best practices through IWRM are strived for among water experts meaning that improvement to the framework is welcome. Through implementing interactive on-line mapping systems, major sources of uncertainty can be addressed. Different levels of uncertainty exist and will be elaborated on below.

³ Tropp, H. (2007). *Water governance: trends and needs for new capacity development*. Stockholm, Sweden: Stockholm International Water Institute.

In this paper I will first provide background information on the MRB, which I will use throughout my paper to give context to my argument. Second, I am going to provide an overview of IWRM in terms of water governance and management, the Dublin Principles of 1992, and the IWRM framework. Third, as a core section to my paper, I will examine issues of uncertainty and the consequences for IWRM of failing to recognize it. I will then conclude the paper with one practical recommendation, which has the capacity to comprehensively address the issues of uncertainty at hand by improving transparency and accountability at all levels.

Introduction to the Mekong River Basin – A Case Study

The Mekong River, located in Southeast Asia, as seen in Figure 1⁴, is one of the world's largest rivers and starts at the Tibetan Plateau in China and flows through Laos, Myanmar, Thailand, Cambodia, and Vietnam before emptying into the South China Sea. The river supports over 70 million people. This heavy reliance revolves around livelihoods of: "fishing, agriculture, hydroelectric power, transportation, biodiversity, and so on".^{5&6} The large populations and diverse uses of the ecosystems leave the river vulnerable to complexities from increasing numbers of development projects and other transboundary governance issues in the region, which is not in the scope of the paper. Figure 2 also highlights future uncertainty from development projects by showing the quantity of existing dams, those under construction and the plans as of 2009. Development projects are a major source of conflict and the reason why shortcuts are taken by development teams to hide the multitude of uncertainties that accompany hydropower projects.

I am using the MRB for a case study because of the river's transboundary nature and its socio-political issues: "widespread poverty, high population growth, a history of conflict and which caused by the weak governance structure".⁷ It is an interesting case because despite these socio-political issues the lower four riparian states came together to become the first transboundary river to attempt reducing regional conflict through the establishment of The Mekong Committee in 1957.⁸

Integrated Water Resources Management in Depth

Water Resources Management Through Water Governance

Among development mentalities and frameworks, governance has always played a central role. Surprisingly, water governance has only recently become significant, moving from "something that was close to a political taboo...to being more widely accepted as a critical issue that needs to be addressed in order to come to grips with unsustainable development and poverty".⁹ This has resulted in the recognition of water governance challenges and has allowed for a more open look at the roles of politics, corruption and power in management practices.

⁴ Institute for Environmental Security. (Producer). (2009). *Dams in the mekong river basin*. [Print Map]. Retrieved from http://www.envirosecurity.org/espa/MekongRiverBasin/maps_images.php

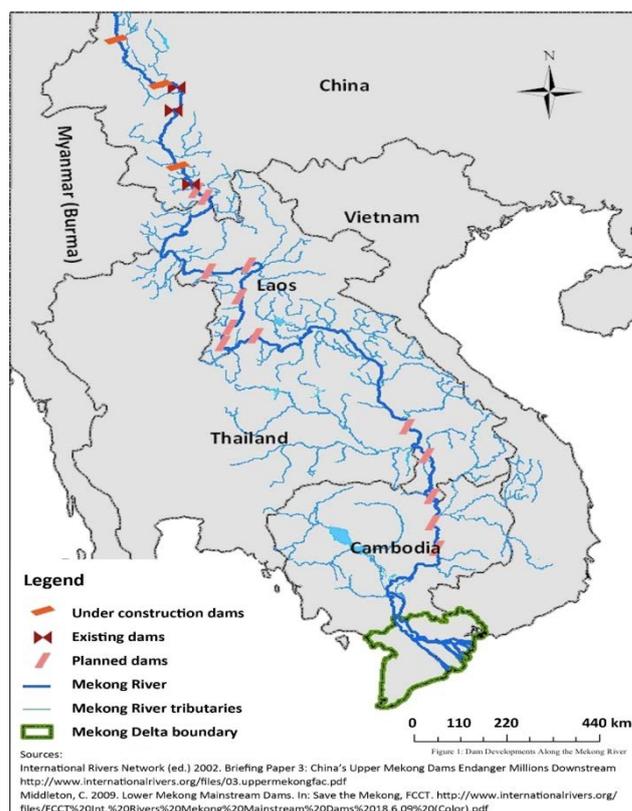
⁵ Belay, A. A., Haq, S. A., Chien, V. Q., & Arafat, B. (2010). The challenges of integrated management of mekong river basin in terms of people's livelihood. *J. Water Resources and Protection*, (2), 61-68.

⁶ Petropoulos, S., & Valvis, A. (2011). International relations and environmental security: Conflict or cooperation? contrasting the cases of the maritza-evros-meric and mekong transboundary rivers. In J. Ganoulis, A. Aureli & J. Fried (Eds.), *Transboundary Water Resources Management: A Multidisciplinary Approach* Germany: Wiley-VCH.

⁷ Belay, et. al (2010)

⁸ Varis, O., Ranhaman, M.M., & Stucki, V. (2008). Integrated Water Resources Management Plans: The Key to Sustainability? *Water & Development Publications*. Helsinki University of Technology, 173-183.

⁹ Tropp (2007)



Water management has traditionally focused on infrastructure development and technology investments to improve access to water. However, it is common for these not to reach their fullest potential. For this reason, governance and management must co-exist to where governance is more than just government actions. Governance needs to involve the horizontal (across different sectors and between urban to rural regions) and vertical approaches (moving between the local to international scale) with the importance of civil society and private sector.¹⁰

Moving to the present, water management still see the importance of developing infrastructure but now encompasses resource allocation, resource protection, and incorporating incentives for efficient water use while ensuring financial sustainability throughout.¹¹ As I will discuss below, water management is a complex process. Financing water management techniques needs constant monitoring and adaptability as “new information and technology gradually become available under changing and uncertain external impacts, such as climate change.”¹²

Dublin Principles 1992

Prior to the 1992 Dublin Principles, water resources management was not a new concept. It dates back to the 1977 United Nations Water Conference in Mar del Plata, Argentina when dialogue and networking emerged between water policy actors.¹³ Current IWRM strategies are modelled after the Dublin Principles and were developed at the World Summit in Rio de Janeiro in 1992 in response to criticism of previous top-down approaches.

Principle 1, as described by the Global Water Partnership, is: “Fresh water is a finite and vulnerable resource, essential to sustain life, development and the environment”.¹⁴ This finite resource recycles itself continuously through the global water cycle. Water can be tainted by large amounts of pollution making it unusable for centuries but, as a finite resource, it cannot be created; making sustainable usage and management practices especially important. Due the interdisciplinary nature of water, the integration of management needs to account for both supply and demand of the resource. A

¹⁰ Tropp (2007)

¹¹ Lenton, R., & Muller, M. (2009). *Integrated water resources management in practice: Better water management for development*. Sterling, Va, USA: Global Water Partnership, Earthscan.

¹² Van der Keur, P., Henriksen, H.J., & Refsgaard, J.C. (2008). *Identification of Major Sources of Uncertainty in Current IWRM Practice. Illustrated for the Rhine Basin*. Water Resource Manage. 1677-1708.

¹³ Conca, K. (2005). Expert networks: The elusive quest for integrated water resources management. In P. Dauvergne (Ed.), *Handbook of Global Environmental Politics* Edward Elgar Pub.

¹⁴ Global Water Partnership. (2012). *What is iwrn?* Retrieved from <http://www.gwp.org/en/The-Challenge/What-is-IWRM>

hydrographical management approach is used, meaning it recognizes the interconnectedness of aquatic ecosystems and environmental health. Under this approach, a river basin is the assigned unit for water management.¹⁵

Principle 2 is used to explain how water as a resource affects everyone: “Water development and management should be based on a participatory approach, involving users, planners and policy-makers at all levels.”¹⁶ Principle 2 revolves around true participation where all actors, stakeholders and affected people are involved to reach long-term success. Participation is interpreted differently at different levels; meaning agreements are not always met, especially between governments and vulnerable groups or communities. This principle attempts a bottom-up approach to decision making – starting with consensus from those with the most at stake.¹⁷

Principle 3 plays off of principle 2 and centres around the importance of women: “Women play a central part in the provision, management and safeguarding of water.”¹⁸ Women play an important role in the transportation and usage of domestic and agricultural water. However, the role of women is viewed as being lesser than men during decision-making processes. Gender research shows that there are strong linkages between gender equality, improved local water management and improved water access.

Principle 4 is, to me, the most compelling of the principles: “Water is a public good and has a social and economic value in all its competing uses”.¹⁹ Water must be recognized as a human right and should be accessible at an affordable price to everyone because it is vital to support life. At the same time, there are costs associated with management, distribution, infrastructure, etc. – for this reason water is also an economic good. This recognition is important for effective distribution across different sectors and users.

IWRM Framework

The IWRM focus strives to reach equilibrium of the ‘three Es’ – efficiency, equity and environment. The three Es’ seek to promote and ensure sustainable development while addressing water challenges of the present and future.²⁰ Figure 2²¹ shows the general framework for IWRM that promotes economic efficiency of management instruments, ecological stability through enabling the environment, and social equity through institutional roles. This framework was designed to avoid and mitigate for fragmentation of possible water management solutions. Through creating an enabling environment for policies, legislation and international cooperation, the institutional roles are more appropriately and equitably filled. This allows for management tools to be the most effective and produce desirable results under the Dublin Principles of 1992.²²

¹⁵ Wenger, R., Rogger, C., & Wymann von Dach, (2003). S. Swiss Agency for Development and Cooperation (SDC). *Integrated Water Resources Management (IWRM): A way to sustainability*. InfoResources Focus.

¹⁶ Global Water Partnership (2012)

¹⁷ Global Water Partnership (2012)

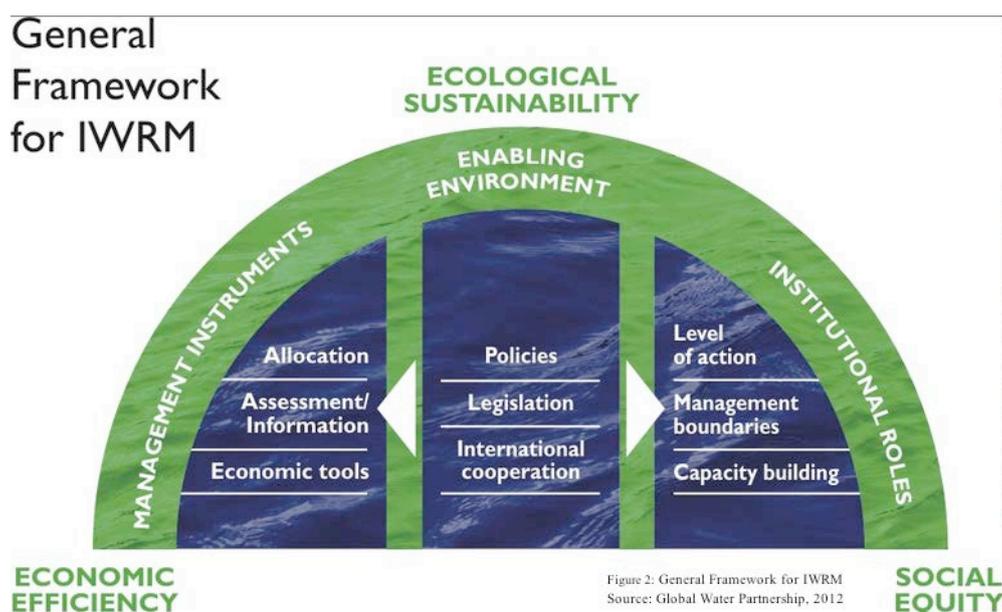
¹⁸ Global Water Partnership (2012)

¹⁹ Global Water Partnership (2012)

²⁰ Lenton & Muller (2009)

²¹ Global Water Partnership (2012)

²² Global Water Partnership (2012)



Issues of Uncertainty

Complexity Leading to Uncertainty

Uncertainty is seen throughout every management decision; therefore, it should be an integral part of the IWRM framework. Management frameworks are generally well received as theoretically practical solutions. However, outside of the written application, complexity sets in as implementation of IWRM produces different cases of path-dependency resulting from seemingly uniform decisions made at the beginning. Nothing is linear in our current environmental system and changes “can trigger sudden and surprising changes”²³ to systems. Over the past decade it has become clearer that in order to handle uncertainty, flexible, adaptable and interdisciplinary frameworks are the best option²⁴. Uncertainty sets in at the very first levels of management “because the first interventions in water resources management are driven by individual users who abstract and store water for their particular purpose. But the interconnected nature of the water cycle means that individual actions often have [unforeseen] impacts”.²⁵

Transboundary water systems, are complex due to high levels of uncertainty. Here, complexity is understood to be how the ecosystem adapts to changes brought on by the unpredictability of human impacts, global environmental change, and changing human values.²⁶

Types of Uncertainty

Understanding the discourse of uncertainty in IWRM is useful for recognizing different types within data, models, frameworks, and in specific cases. When used successfully, it translates to better understanding of IWRM strategies and management scenarios. The first step is to identify the nature of the uncertainty – ontological or epistemic. Ontological refers to uncertainty that cannot be reduced and is “due to inherent variability in the system”.²⁷ Epistemic uncertainty, on the other hand, can be reduced depending on the type and source and is “due to imperfect knowledge of the system”.²⁸ However, with more information uncertainty can, at times, increase with the idea that ‘the more we know, the more we do not know’.

Figure 3: Uncertainty Scale, shown below, looks at the transition of uncertainty from determinism to total ignorance. The state of determinism is when, based on the information gathered,

²³ Galaz, V. (2007). Water governance, resilience and global environmental change – a reassessment of integrated water resources management (IWRM). *Water Science & Technology*, 56, 4, 1-9.

²⁴ Brugnach, M., Dewulf, A., Henriksen, H. J., & Van der Keur, P. (2011). More is not always better: Coping with ambiguity in natural resources management. *Journal of Environmental Management*, (92), 78-84.

²⁵ Lenton & Muller (2009)

²⁶ Van der Keur, et. al (2008)

²⁷ Van der Keur, et al (2008)

²⁸ Van der Keur, et al (2008)

everything is known – complete certainty in a situation. This is the ideal situation but is never achieved due to complexity of the water system.²⁹

Statistical uncertainty is generally referred to as potential measurement errors or any form of inaccuracy in data collection or interpretation. Moving left on the scale; *scenario* uncertainty is not based on statistics or data. This type is in the analysis of policy and is used to explain potential system changes based “as a function of known controls like changes in management, technology and price structure”.³⁰ Scenario uncertainty is largely based on past experiences and considers potentially known outcomes when the probabilities of those outcomes are unknown. *Qualitative* uncertainty appears when statistical explanations cannot be provided and scenario outcomes are also relatively unknown. *Recognized* ignorance is true when those involved are aware that there is a lack of knowledge and information on a given topic but do not know how to move forward. In this case, predictions are often impossible regardless of continued research. This is also the point of indeterminacy. Finally, *total ignorance* is when there is a complete lack of awareness that the situation contains gaps in knowledge and information.³¹

For the sake of the recommendations section below it is necessary for me to outline uncertainty in a natural, technical and social context based in the Mekong River Basin. Technical uncertainty it is an incomplete understanding or a lack of knowledge on the technology and efficiency of an existing hydropower dam. Natural uncertainty is relatively broad where there is overall limited understanding of processes involved in collecting further data and knowledge at a project site. In this case, it can be reduced and understanding gained from surrounding ecosystems. Referring back to Figure 1: Dam Developments Along the Mekong River, the proposed dams are presented to stakeholders, and potentially affected community members, in very simplistic terms. These proposals are often misrepresenting the reality of future harms a dam could have on the local environment. Finally, social uncertainty is with regards to potentially limited knowledge of the social and economic system aspects. An example “could be uncertainty on statistical economic data from a river basin, because all existing data bases are disaggregated into administrative units and are not coinciding with river basin boundaries”.³² With proper distribution and communication of the knowledge at hand this type of uncertainty can be reduced.

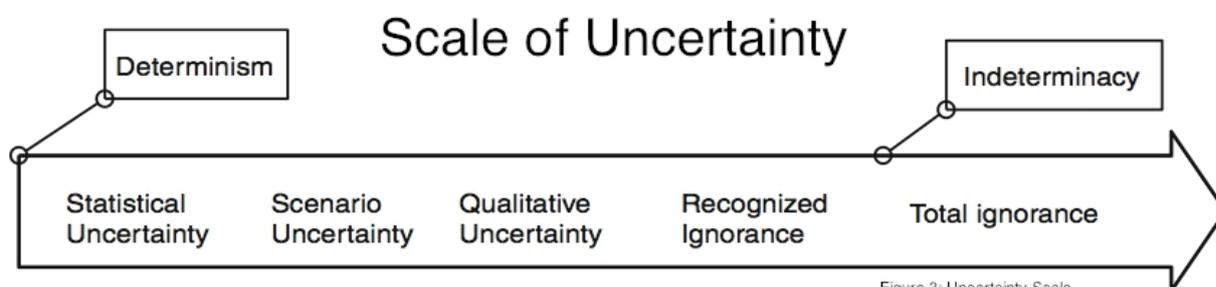


Figure 3: Uncertainty Scale

Brugnach, et al explains the importance of coping with ambiguity in a river basin ecosystem. Ambiguity is distinct from the types of uncertainty listed on the scale above and results come from the “simultaneous presence of multiple valid, and sometimes conflicting, ways of framing a problem.”³³ Ambiguity shows the inconsistencies between how different actors translate and understand problems. Unfortunately, as more actors are added into a decision-making process ambiguity is sometimes unavoidable. As a result, a clear solution, distinguishing who should be involved and the next course of action is not always available. It suggests that more than one method might be available for a given system, meaning ambiguity is not a direct knowledge gap. The question is then posed for how to cope with ambiguity in a river basin ecosystem under IWRM strategies? To cope, different strategies must be recognized at all stages of management. This means full participation is key; listening to all stakeholders to decide where and what values should be incorporated during decision-making.³⁴

²⁹ Van der Keur, et al (2008)

³⁰ Van der Keur, et al (2008)

³¹ Van der Keur, et al (2008)

³² Van der Keur, et al (2010)

³³ Brugnach, et al (2011)

³⁴ Brugnach, et al (2011)

Sources of Uncertainty

As visualized in Figure 3: Scale of Uncertainty, uncertainties vary. However, identifying the types can help identify where the sources of uncertainty are. IWRM has guidelines for determining these sources that generally focus on certain processes within the framework, which are “notably on uncertainty associated with the modelling process and monitoring data”.³⁵ The importance of having guidelines is to localize uncertainty to ensure informed actions within the decision-making process. Many of the guidelines I found in the literature range from generic and ambiguous to very specific. These varying guidelines can be useful for context specific cases and the range of users within the management process.

The following flow chart, Figure 4: Finding Uncertainty in Decision-Making³⁶, is one of the guidelines I found to be very descriptive in the literature and will outline the steps that are taken in a best case scenario of the IWRM decision-making process. Each of the five steps is a representation of either a single or group of possible uncertainties. The literature by Sigel, et al describes the flaws of the guidelines: “As these steps aim to describe the entire selection process, [technically] a description of sources of uncertainty based on this structure can claim to be complete. However...a different structure may lead to other descriptions of the sources of uncertainty”.³⁷ This represents statistical uncertainty and highlights the biased thought process that these are part of a linear system. The below guidelines disregard the IWRM’s first principle of following the hydrological approach where a river ecosystem exhibits non-linear behaviour and is within a complex system.

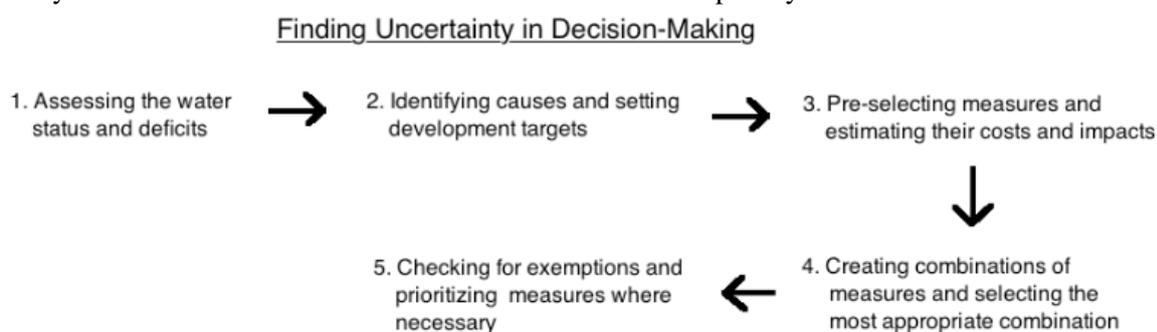


Figure 4: Finding Uncertainty in Decision-Making

One of the biggest problems contributing to IWRM’s inability to address uncertainty in the MRB is due to a lack of political commitment or an uneven level of commitment between states sharing the same basin. This idea will be addressed below when looking at China’s relationship with the lower riparian states. Without political commitment there is minimal state representation, power and influence and none of the necessary components for effective action plans. In order for IWRM to be effective, state commitment needs to be backed “by a solid and accurate understanding of the dimensions of the problem”³⁸. Being informed to the point of statistical uncertainty will ensure for the best possible management outcomes.

The competitive nature of states and the desire for state sovereignty can lead to even lower levels of cooperation or an inability to see the potential of cooperation in the future. Disengagement of information sharing and technology transfers between riparian states can also be from a lack of mutual trust – especially when involved in competition for shared resources and have similar interests.³⁹

Analysing Uncertainty in the Mekong River Basin

Another somewhat simpler form of guidelines under the IWRM framework are a set of three questions posed in preparation of a failure due to uncertainty in the system. I found this set of guidelines useful because it incorporates types of uncertainty into the equation and thus, attempts to cover a broader scope with fewer steps. These questions are posed sequentially:

³⁵ Van der Keur, et al (2010)

³⁶ Sigel, et al (2010)

³⁷ Sigel, et al (2010)

³⁸ Ganoulis, J. & Salame, L. (2011). *A Risk-Based Integrated Framework for Conflict Resolution in Transboundary Water Resources Management*. In J. Ganoulis, A. Aureli & J. Fried (Eds.), *Transboundary Water Resources Management: A Multidisciplinary Approach*. Germany: Wiley-VCH

³⁹ Ganoulis & Salame (2011)

- “1) When could the system fail?
- 2) How often is failure expected?
- 3) What are the likely consequences?”⁴⁰

Steps one and two are addressed to determine the uncertainty of the system. Question one is found by putting together a critical scenario – scenario uncertainty – and determining the potential outcomes based on past experiences. Question two examines the potential frequency of failure through probability calculations and statistical analysis – statistical uncertainty. Finally, question three puts the results of the above questions together to determine any potential losses or gains – qualitative uncertainty.⁴¹

Using the Mekong case study I will use these guidelines to address the uncertainties of China’s absence in signing regional agreements with the other riparian Mekong states in the context of building dams for energy production. By looking at the process of answering the questions it will become clear that based on the IWRM’s current structure and idealistic values, it does not accurately account for uncertainty in the system.

As China’s population grows, so do domestic energy demands, thus increasing the amount of dam proposals along the upper Mekong River – refer back to Figure 1: Dam Developments Along the Mekong. This figure shows that, in China alone, there are three existing dams, three under construction, and plans to construct two more as of 2009.

1) *When could the system fail?*

As this question examines scenario uncertainty, the first step would be to examine the results of past dams built in China. Consider the expected timelines of the current infrastructure and when those are expected to fail or need repairs. China abstained from signing the 1997 Convention on the Law of Non-navigational Uses of International Watercourses or onto the Mekong River Commission. Nor will they sign, knowing that it is the only treaty governing the use of shared freshwater and, that without 35 ratifications, it will not become international law. In 1996, Manwan dam was constructed in Yunnan Province, China and sits as a founder for the planned construction of eight more in the province. The Manwan dam will be used below to provide more context for questions two and three.

2) *How often is failure expected?*

For the purposes of this paper I will not be doing my own calculations for probability, I will instead outline a formula to develop an assessment of the risk of failure. The formula is laid out as follows to reach a level of statistical uncertainty:

(L) = Behaviour based on external stresses on the system

(r) = Resistance of the system

* These terms occur at random under probability assumptions*

Based on when L exceeds its limit, either a *failure* or an *incident* will present itself.

“FAILURE or INCIDENT: $L \geq r$, where SAFETY or RELIABILITY: $L < r$ ”⁴²

Therefore, assuming randomness under probability, the probability of failure is a risk.

“RISK = probability of failure = $P(L \geq r)$ ”

In this instance RISK is defined as: “the possibility of losses”

Used in an equation:

RISK = (Hazard) x (Vulnerability), or RISK = (Probability) x (Consequences) = (Expected Consequences)⁴³

In order to reach statistical uncertainty under these guidelines, regulations for an acceptable data analysis are at minimum 20 years. The data series used for the Manwan Dam falls short with only 12 years – 1992-2003. Also, to make matters worse, literature shows that there are major sources of recognized uncertainty in the rating curve, which is the relationship between stage and water discharge. It was established in 1975 and was used without updates for close to 20 years until 1994.⁴⁴

3) *What are the likely consequences?*

Based on results from the above questions, as well as from findings in *Probe International* and the *World Wildlife Fund Global*, I am distinguishing three consequences as the most severe for

⁴⁰ Ganoulis & Salame (2011)

⁴¹ Ganoulis & Salame (2011)

⁴² Ganoulis & Salame (2011)

⁴³ Ganoulis & Salame (2011)

⁴⁴ Xi Xi, et al (2008)

downstream ecosystems. The first is a loss of river sediment, or delta instability. Sediment becomes caught in the dams and as more are constructed along the river, the amount reaching the Lower Mekong Basin is significantly reduced. Reduced sediment flow leaves the basin “vulnerable to sea level rise and saline intrusion brought on by climate change.”⁴⁵ The second is a decline in fish species diversity. Inefficient and lack of accountability during construction will potentially hinder fish migration necessary for fish productivity by 60%. Populations inhabiting the river basin are more than 75% dependent on the river livelihood stability. Any combination of the above consequences can result in damaged livelihoods as well as general ecosystem health.^{46,47}

Overall, great uncertainty resides with China because it holds the power within the MRB. Geographically and politically it has the advantage for future negotiations on river management. Uncertainty increases as China’s energy needs grow and will likely affect flow levels for the Lower River Basin.⁴⁸ Based on the results of question 2, it is clear that without proper management and transparency of appropriate information, IWRM is not going to improve and will not be able to account for uncertainty in the system.

Recommendation

Reading the literature of uncertainty within IWRM, I became aware that only a small portion of the literature dealt with the notion that uncertainty can largely be improved by implementing new forms of regulation and ensuring transparency and accountability. Man-made problems are everywhere in the water sector and produce many unknowns. However, a majority of these uncertainties can be reduced to scenario or statistical uncertainty by the simple act of collecting more information and improving how this information is presented.⁴⁹ Galaz’s article describing water governance and resilience lays out the need for improved information collection very clearly: “The increased potential for regime shifts, surprises and conflict...from uncertain hydrological changes, and the desires of competing water users...calls for a much closer examination of the ways in which economic tools...can be used to facilitate or hinder adjustment to the effects of global environmental change.”⁵⁰

Uncertainty needs to be harnessed. Dynamics of the river basin ecosystem need to be under as much control as statistical uncertainty can allow. Water quality is generally more complicated and difficult to monitor and regulate than water quantity because technical knowledge is needed for both monitoring and enforcement.⁵¹ To harness the uncertainty and lessen the complexity of monitoring and enforcement a new method and strategy is necessary. I will be giving a close look to using mapping technologies as a management tool to increase transparency and accountability to account for issues of uncertainty in IWRM. This is based on the assumptions that large amounts of information and knowledge are necessary for solving environmental problems and that access to all available knowledge is the foundation of good decision-making.⁵² Caution and sensitivity is necessary in many parts of the world, especially for environmental issues in the MRB. This mapping technology has the potential to reduce knowledge gaps resulting from sensitivity as well as political and power dynamics in the Basin.

States and territorial boundaries are a result of the globe’s modern system of political organization. These boundaries are the outcome of a major and complex European transformation. Modern cartography was a large factor to this transformation: “new mapmaking technologies changed how actors thought about political space, political organization, and political authority.”⁵³ I think there

⁴⁵ WWF Global (panda.org). (2012). *Mekong River Basin: Damming the Mekong*. Some Rights Reserved. http://wwf.panda.org/what_we_do/footprint/water/dams_initiative/examples/mekong/

⁴⁶ WWF Global (panda.org) (2012)

⁴⁷ Probe International (2008). *Neighbours 'face harm from Chinese dam project'*. Energy Probe Research Foundation. Probe International. <http://eprf.probeinternational.org/node/5593>

⁴⁸ Petropoulos & Valvis (2011)

⁴⁹ Ganoulis & Salame (2011)

⁵⁰ Galaz (2007)

⁵¹ Lenton & Muller (2009)

⁵² Sigel, et al (2010)

⁵³ Branch, J. (2011). *Mapping the Sovereign State: Technology, Authority, and Systemic Change*. International Organization Foundation. 1-36.

is potential for current and future digital cartographic technologies to make drastic changes again, especially for global environmental governance and IWRM.

Reliable information and up to date information is essential to every aspect of IWRM and the same is true for any form of governance. Necessary and essential information includes, at the working level, daily data levels of river basins – flows, discharge, and quality. Also, at a more complex level there is a need for “engineering structures, ecological needs of flow, quality at ecologically significant parts of the area, and calendars of local cultural events which all support the management needs at this level...patterns of demand, economic development indicators...”⁵⁴ and the list goes on. Attempting to incorporate all types of information into one database that is accessible to all relevant stakeholders is a complicated process. Therefore, of course there are limits to this form of mapping technology but it is outside the scope of this paper.

Knowledge is still limited on how mapping technologies would realistically function as a form of increasing transparency and accountability within IWRM to account for uncertainty. However, there are cases of major success in the forestry sector of global environmental governance. I believe that using results and starting points of the forestry sector there is great potential for successful implementation in the water governance sector as well.

Forestry presents many major environmental concerns that are different than those of the water sector. However, both have the same end goals that revolve around improving transparency and accountability – why should the means to those ends be different? Ecosystem mapping uses interactive maps of Geographic Information Systems (GIS) mapping techniques and satellite imagery to monitor physical forest areas. It shows and explains where, when, and how deforestation is occurring in Central West Africa. One project funded by the World Resources Institute (WRI) is the *Moabi* system used in the Democratic Republic of Congo (DRC). They take a bottom-up approach by going to the public and requesting the submission of information in an offline setting that is transferred online every 16 days. *Moabi* then approached the public and private sector to lobby for provided information. Results were very surprising as information poured into the *Moabi* database. The website and database worked as both advocacy and an information sharing space, as well as a means for private sector actors to get name recognition and build a reputation.⁵⁵

Conclusion

I briefly conclude this paper on accounting for uncertainty within IWRM frameworks by outlining the importance of improved transparency and accountability. The answers to IWRM problems and criticisms will not be found in past experiences. A restructuring of IWRM to include new approaches into the framework is necessary using present day technology. Satellite imagery and GIS mapping systems are, what I think to be, the way forward. They have potential to account for transparency and ensure accountability at all levels of the IWRM process, which, as described above, is extremely complex and full of varying uncertainties at every level and at the end of each decision.

References:

- Belay, A. A., Haq, S. A., Chien, V. Q., & Arafat, B. (2010). The challenges of integrated management of mekong river basin in terms of people's livelihood. *J. Water Resources and Protection*, (2), 61-68.
- Branch, J. (2011). *Mapping the Sovereign State: Technology, Authority, and Systemic Change*. International Organization Foundation. 1-36.
- Brugnach, M., Dewulf, A., Henriksen, H. J., & Van der Keur, P. (2011). More is not always better: Coping with ambiguity in natural resources management. *Journal of Environmental Management*, (92), 78-84.
- Conca, K. (2005). Expert networks: The elusive quest for integrated water resources management. In P. Dauvergne (Ed.), *Handbook of Global Environmental Politics* Edward Elgar Pub.
- Galaz, V. (2007). Water governance, resilience and global environmental change – a reassessment of integrated water resources management (IWRM). *Water Science & Technology*. 56, 4, 1-9.

⁵⁴ McDonnell, R.A. (2008). *Challenges for Integrated Water Resources Management: How Do We Provide the Knowledge to Support Truly Integrated Thinking?* International Journal of Water Resources Development. 24:1, 131-143.

⁵⁵ Harvard Kennedy School of Government. (2012). *Ecosystem Report: Natural Resource Governance*. Transparency Policy Project. Harvard Kennedy School of Government.

- Ganoulis, J. & Salame, L. (2011). *A Risk-Based Integrated Framework for Conflict Resolution in Transboundary Water Resources Management*. In J. Ganoulis, A. Aureli & J. Fried (Eds.), *Transboundary Water Resources Management: A Multidisciplinary Approach*. Germany: Wiley-VCH
- Global Water Partnership. (2012). *What is iwrn?* Retrieved from <http://www.gwp.org/en/The-Challenge/What-is-IWRM>
- Harvard Kennedy School of Government. (2012). *Ecosystem Report: Natural Resource Governance*. Transparency Policy Project. Harvard Kennedy School of Government.
- Institute for Environmental Security. (Producer). (2009). *Dams in the mekong river basin*. [Print Map]. Retrieved from http://www.envirosecurity.org/espa/MekongRiverBasin/maps_images.php
- Institute for Environmental Security. (Producer). (2009). *Dams in the mekong river basin*. [Print Map]. Retrieved from http://www.envirosecurity.org/espa/MekongRiverBasin/maps_images.php
- Lenton, R., & Muller, M. (2009). *Integrated water resources management in practice: Better water management for development*. Sterling, Va, USA: Global Water Partnership, Earthscan.
- McDonnell, R.A. (2008). *Challenges for Integrated Water Resources Management: How Do We Provide the Knowledge to Support Truly Integrated Thinking?* *International Journal of Water Resources Development*. 24:1, 131-143.
- Petropoulos, S., & Valvis, A. (2011). International relations and environmental security: Conflict or cooperation? contrasting the cases of the maritza-evros-meric and mekong transboundary rivers. In J. Ganoulis, A. Aureli & J. Fried (Eds.), *Transboundary Water Resources Management: A Multidisciplinary Approach*. Germany: Wiley-VCH.
- Probe International (2008). *Neighbours 'face harm from Chinese dam project'*. Energy Probe Research Foundation. Probe International. <http://eprf.probeinternational.org/node/5593>
- Sigel, K., Klauer, B., & Pahl-Wostl, C. (2010). Conceptualising uncertainty in environmental decision-making: The example of the EU water framework directive. *Ecological Economics*. 69, 502-510
- Tropp, H. (2007). *Water governance: trends and needs for new capacity development*. Stockholm, Sweden: Stockholm International Water Institute.
- Van der Keur, P., Henriksen, H.J., & Refsgaard, J.C. (2008). *Identification of Major Sources of Uncertainty in Current IWRM Practice. Illustrated for the Rhine Basin*. *Water Resource Manage*. 1677-1708.
- Van der Keur, P., Brugnach, M., Dewulf, A., et al. (2010). *Identifying Uncertainty Guidelines for Supporting Policy Making in Water Management Illustrated for Upper Guadiana and Rhine Basins*. *Water Resource Manag*. 3901-3938.
- Varis, O., Ranhaman, M.M., & Stucki, V. (2008). *Integrated Water Resources Management Plans: The Key to Sustainability?* *Water & Development Publications*. Helsinki University of Technology, 173-183.
- Wenger, R., Rogger, C., & Wymann von Dach, S. Swiss Agency for Development and Cooperation (SDC), (2003). *Integrated Water Resources Management (IWRM): A way to sustainability*. InfoResources Focus.
- WWF Global (panda.org). (2012). *Mekong River Basin: Damming the Mekong*. Some Rights Reserved. http://wwf.panda.org/what_we_do/footprint/water/dams_initiative/examples/mekong/
- Xi Xi, L., Jian-Jun, W., & Grundy-Warr, C. (2008). *Are the Chinese dams to be blamed for the lower water levels in the Lower Mekong?* *Water & Development Publications*. Helsinki University of Technology.

PRELIMINARY TESTS OF THERMAL CONDUCTIVITY OF SELECTED SOIL TYPES

Agata Ludynia, MSc

Lukasz J.Orman, PhD

Kielce University of Technology, Poland

Abstract:

The paper discusses the topic of thermal conductivity investigation of soil. This problem is especially crucial for the design and exploitation of ground heat exchangers. Such systems located in the soil can either extract heat from the ground or transport it there for season storage. Nowadays, these and similar installations are used in connection with renewable energy systems – for example heat pumps. The paper presents the preliminary test results and characteristic properties of selected kinds of soil. Further tests could help to determine the usefulness of different soil types for heat exchanger applications.

Key Words: Conductivity, soil, thermal properties

Introduction

Thermal conductivity is a property of engineering materials and its value is often determined in order to assess the insulation characteristics of building materials. However, in view of the current trend to efficiently use energy, soil thermal properties become interesting and scientific effort is made in this field. The precise knowledge of the thermal conductivity of soil can help to properly design ground heat exchangers. The energy from the ground might be used, among other applications, to heat homes in winter or provide domestic hot water throughout the whole year.

Abu-Hamdeh and Reeder (2000) focused their research on the influences of bulk density, moisture and salt as well as organic matter of soil samples - repacked and sieved. They wrote that conductivity rose with soil density and moisture concentration. Abu-Hamdeh (2003) in a later paper also analysed the influence of the concentration of water and bulk density of soils. Tang et al. (2008) investigated compacted bentonite and the influence of dry density, water content, saturation on conductivity. Abuel-Naga et al. (2009) presented experimental tests showing that the values of thermal conductivity were higher with the rise in soil density. O'Donnell et al. (2009) considered a relation between thermal conductivity and moisture of soil. In terms of theoretical studies, Singh et al. (2011) presented a possibility of using a method of artificial neural networks for the prediction of effective thermal conductivity of moist porous systems.

Samples and measurements

Sample preparation and preliminary tests of basic properties is quite significant from the scientific point of view. The paper by Grobelska and Ludynia (2010) discusses a sample preparation method.

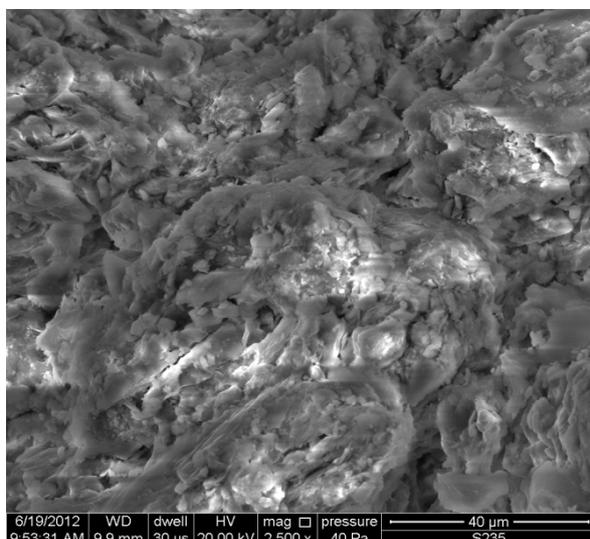
In the present research two kinds of soil (sasiCl and Cl) have been selected for the experiments and Table 1 shows their basic properties determined directly through measurements or calculated basing on the results of the laboratory tests.

Table 1. Basic properties of the investigated soil types.

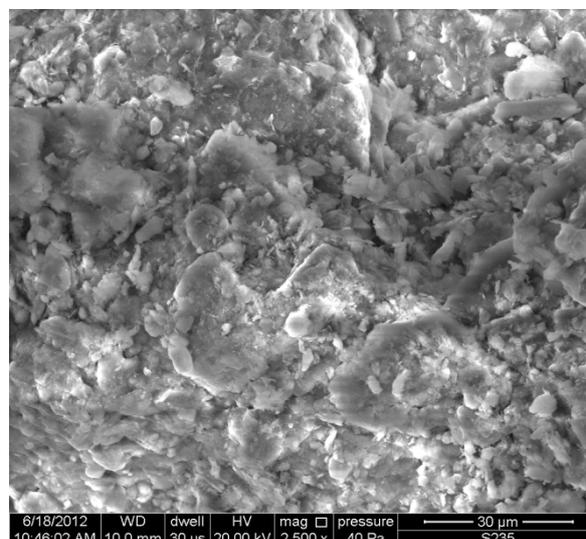
Property	sasiCl	Cl	Testing method
bulk density of soil [ρ], g/cm ³	2,32	2,05	ring method
density of solid particles [ρ_s], g/cm ³	2,69	2,72	pycnometer method
dry density of solid particles [ρ_d], g/cm ³	2,13	1,65	calculated
porosity [n], -	0,21	0,39	calculated
void ratio [e], -	0,26	0,65	calculated
fine fraction (f_{si+cl}), %	51,35	91,5	laser method
water content w , %	9,04	24	dry method

Additionally, the SEM images have been made in order to have an insight into the surface structure of the analysed kinds of soil. The results of the scanning microscopy tests have been presented below in Figures 1a and 1b for the magnification of 2500 times.

a)

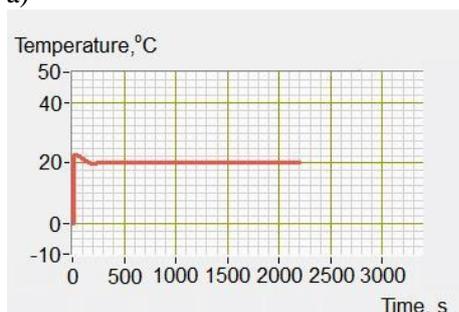


b)

**Fig. 1.** SEM images of sasiCl (a) and Cl (b).

The measurements of thermal conductivity have been conducted on samples of ca. 12x12cm in the a plate apparatus. This equipment comprises two parallel plates located in an insulated chamber to reduce heat losses. One plate acts as a heater (kept at the temperature of ca. 20°C) and the other a cooler (temperature ca. 0°C). The temperatures could be changed. The conductivity is determined by the equipment based on the measured temperature difference, thickness of the analysed sample as well as heat flux transferred through the sample and measured during the test. The correct steady – state measurement is carried out if the temperatures of the plates reach the given values. The heat flux should also be constant. Figures 2a and 2b show example graphs of changes in temperatures of the plates while Figure 3 – an example heat flux curve.

a)



b)

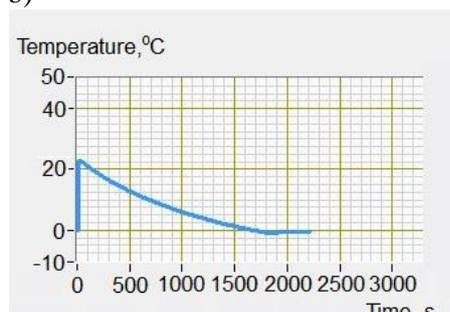
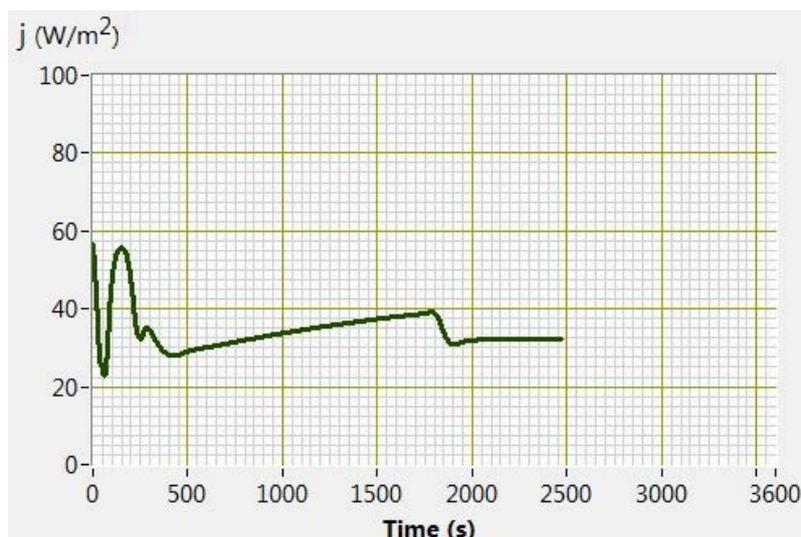


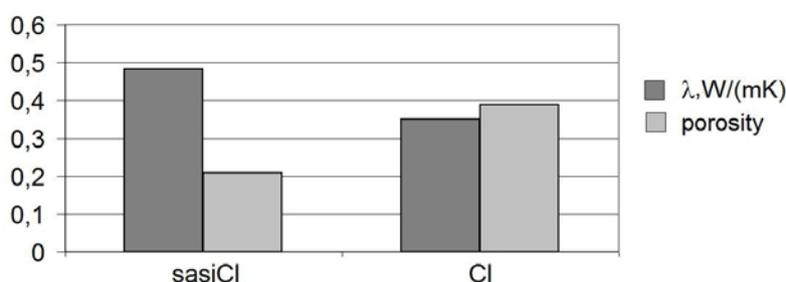
Fig. 2. Example temperature changes of the heater (a) and the cooler (b).**Fig. 3.** Example changes of the heat flux.

Test results

The presented research enabled to determine thermal conductivity of the analysed soil samples. The results have been presented in Table 2 and Figure 4.

Table 2. Thermal conductivity of the tested soil samples.

Sample:	sasiCl	Cl
Thermal conductivity, W/(mK)	0,483	0,351

**Fig. 4.** Thermal conductivity and porosity results.

The above test results are in general agreement with literature data on thermal conductivity of soil which indicates that the experimental method is applied correctly. The higher value was recorded for sample which has lower porosity while the sample of higher porosity had lower conductivity. This can be explained by the presence of voids that increase the insulating properties of the materials. Basing on the good understanding of the soil thermal conductivity, ground heat exchangers might be designed more accurately in given conditions.

Conclusion

Current trend in energy recovery is common in almost every country. Ground heat exchangers can be used such practical applications as heat pumps systems. Consequently, a detailed knowledge about the soil thermal parameters is crucial in a proper design of the whole installation. Nevertheless, more thorough analysis and calculations are necessary in this field. More experiments could also provide a basis for creating a model of thermal conductivity of different types of soil.

References:

- Abuel-Naga, Hossam M., Dennes T. Bergado, Abdelmalek Bouazza and Michael J. Pender. Thermal conductivity of soft Bangkok clay from laboratory and field measurements, *Engineering Geology*, Vol. 105, 3–4, pp. 211–219, 2009.
- Abu-Hamdeh, Nidal H. and Randall C. Reeder. Soil thermal conductivity effects of density, moisture, salt concentration, and organic matter. *Soil Science Society of America Journal*, Vol. 64 No. 4, pp. 1285-1290, 2000.
- Abu-Hamdeh, Nidal H. Thermal Properties of Soils as affected by Density and Water Content, *Biosystems Engineering*, 86 (1), pp. 97–102, 2003.
- Jonathan A. O'Donnell, Vladimir E. Romanovsky, Jennifer W. Harden and David A. McGuire. The Effect of Moisture Content on the Thermal Conductivity of Moss and Organic Soil Horizons From Black Spruce Ecosystems in Interior Alaska, *Soil Science*, Vol. 174, 12, pp. 646-651, 2009
- Grobelska, Edyta and A. Ludynia. The influence of sample preparation by the laser diffraction method on the particle size distribution of bentonite, *Structure and Environment*, 4, pp. 41- 47, 2010.
- Ramvir Singh, R.S. Bhoopal and Sajjan Kumar. Prediction of effective thermal conductivity of moist porous materials using artificial neural network approach, *Building and Environment*, Vol. 46, 12, pp. 2603–2608, 2011.
- Anh-Minh Tang, Yu-Jun Cui and Trung-Tinh LE. A study on the thermal conductivity of compacted bentonites. *Applied Clay Science*, Vol. 41, 3–4, p. 181–189, 2008.

EFFECT OF COASTAL ENVIRONMENT IN CLAY FACING BRICKS AND ROOF TILES

B. Sena da Fonseca

J.A.R. Simao

C. Galhano

CICEGe, U.N.L. – Fac. de Ciências e Tecnologia, Dep. Ciências da Terra, Caparica

Abstract:

Materials made from clay, such as bricks or roof tiles are widely used in recent buildings. Their severe exposure to salt spray, common in coastal environments could cause premature degradation which leads into functional, aesthetic, economical and safety problems.

The aim of this study is to analyze and quantify the physical and mechanical alteration caused by salt spray in modern facing bricks and roof tiles. For this purpose, one type of roof tile and two types of bricks were submitted to various stages of accelerated aging test by salt spray. Visually and binocular microscope inspection reveal visible degradation. The tests to evaluate the physical and mechanical characteristics of the bricks and tiles were made before, during and after the aging test, the results expose a general unfavorable trend in expected performance of the three types of materials. It was possible to identify the main intrinsic characteristics of the materials that have led to their degradation.

Key Words: Clay bricks, roof tiles, alteration, salt spray, accelerated ageing test

Introduction

Clay has been widely used as a raw material in constructions and buildings materials since 8000 BC (Houben & Guillaud, 1994). Roof tiles in houses coverings have been an ancient tradition and their current use allows the maintenance of the architectural tradition in many countries. On the other hand, their diverse forms, parts and accessories allowed varied aesthetic effects (Garcez, 2009). The facing bricks industry is currently the most interesting and their applications suffered a technological revolution in terms of manufacturing and material types (Parras, 1997). These bricks have a diverse range of characteristics, depending on the raw materials and manufacturing methods and procedures (Lucas, 2003). For a long time, the degradation of these and other construction materials have been a concern, largely due to lack knowledge of physical, chemical and mechanical properties of the materials in use and also to a disregard of environmental conditions where they will be implemented. This could, affect in long term, the materials correct performance and cause aesthetic, functional and security problems (Moser, 1999). It is well known that its rehabilitation, maintenance or substitution implies high financial costs.

The salt deterioration has been widely studied in recent and historic structures and in most of the cases the attack occurs due through rising dampness from soils (Ottosen *et al.*, 2007; Rorig-Dalgaard *et al.*, 2012). Another source of harmful salts is the sea spray, which is common in coastal areas in all parts of the world.

According to the Intergovernmental Panel on Climate Change (IPCC, 2001), every year are release into the atmosphere 3300Tg of small particles in salt spray form. These particles are produced due to wave breaking, by three mechanisms (Leeuw, 1999). During the wave breaking process, air bubbles are introduced in the water column and then return to the surface a bubble film is formed which separates the interior of the air exterior. When the film breaks produces small droplets. From other side, the bubble could collapse immediately after reaching the surface. Also when the wind speed is higher than 9 m/s, particles are naturally released from the sea.

The transport of these particles to onshore is a current situation and occurs through wind and turbulent sea that causes displacement of air. Sometimes the sea influence in the atmosphere is enhanced up to 20 km into the shoreline (Silva *et al.* 2007).

Therefore, it is obvious the interaction between the salt spray from sea and building materials, particularly in clay bricks and tiles. The degradation caused by this salt spray has been widely studied through short-term test in laboratory. Their effect is well known in several construction materials, e.g. in ornamental rocks (Galembeck *et al.*, 2008; Silva & Simão, 2009). Efflorescence resulting by salt deposits has been observed in old clay ceramic facades and results in harmful types of damaging like exfoliations or spalling (Kuchitsu *et al.*, 2000; Brocken & Hijland, 2004; Lubelli *et al.*, 2004). Due to severe exposure to sea spray, this type of degradation has also been found in several recent buildings. As an example, figure 1 show a photo taken in a building facade from a Portuguese coastal city (Sesimbra) a few years after its construction, where deterioration is clearly evident.

Studies in modern bricks are scarce and the physical, mechanical deterioration arising this interaction is poorly known. This work aims the study of the alterations of modern ceramic materials when subjected to a coastal environment, under the action of the salt spray. Physical, mechanical and aesthetical modifications in tiles characteristics were monitored by several laboratory tests.



Figure 1 - Aspects of clay bricks degradation due to sea spray action, Sesimbra, Portugal.

Materials and methods

For this investigation, three different kinds of commercial fired clay ceramics were used. The samples were taken from three different Portuguese manufacturing industries. A roof tile, “lusa” (RT), manufactured with a maximum firing temperature of 1035°C and two types of facing bricks, a extruded unglazed brick (EU), class A1b (EN 14411) and a extruded glazed brick (EG), class A1a (EN 14411), subjected to a firing temperature of 1230°C and 1220°C, respectively. Additionally, the granulometric study of RT and EG raw materials has been carried out by wet sieving for the fraction >0.063mm and with a SEDIGRAPH analyser the fraction <0.0063mm, the supplier of EU did not provide the respective raw material.

In the absence of a European standard for accelerated aging test for salt mist in building bricks and tiles, the EN 14147 for natural stone was followed with appropriate adjustments to the material under study. To this aging test salt spray chamber (ASCOTT S120T) was used. The salt solution contained 1 part of NaCl to 9 parts of distilled water. The spray cycles were comprised by 8 hours of active spray and 16 hours of drying at 40°C (\pm 2°C). The process was repeated for a total of 60 cycles. Each 10 cycles, the samples were removed from the chamber and weighed before and after 10 days of desalination by immersion in distilled water.

Every 20 cycles, the samples surface was visually inspected with a binocular microscope and the alterations registered by microphotography.

After 20, 40 and 60 cycles, six samples were used to water absorption and apparent porosity determination (EN 10545-3) and five samples were used to obtain the modulus of rupture (EN 10545-4). For the assessment of the main mineralogical composition, samples of the tiles were cut after and before the 60 cycles and submitted to x-ray diffraction.

Results and discussion

Raw Material

Figure 2 shows the granulometric analysis of the raw material corresponding to RT and EU. The analysis of the granulometric curves revealed that the average size of particles in the raw material of RT is higher than the average particle size in the EU.

The curve corresponding to the EU also reveals a smoothness slope indicating a higher range of sizes. The curve corresponding to RT has a uniform granulometry, the fraction between 30 μm and 3 μm has more than 50% of the material. In RT it is also possible to identify a 3% of particles larger than 500 μm .

The raw materials of both materials reveal the existence of a significant fraction $<0.3 \mu\text{m}$, outside the detection range of the equipment, 6.78% in RT and 26.99% in EU.

It is consensual that the raw material granulometry has a major role in the sintering process and in the overall properties of the clay ceramics (Sokolar & Swetanova, 2010). By this reason, and independently from other manufacturing parameters, with the obtained data it is possible to make a relative prediction in tiles characteristics. The suitable uniform granulometry in raw material of RT gives a good compactness which produces tiles with low porosity. However, the low average size of particles in EU raw material could contribute to smaller pores and lower percentage of porosity in terms of the total volume.

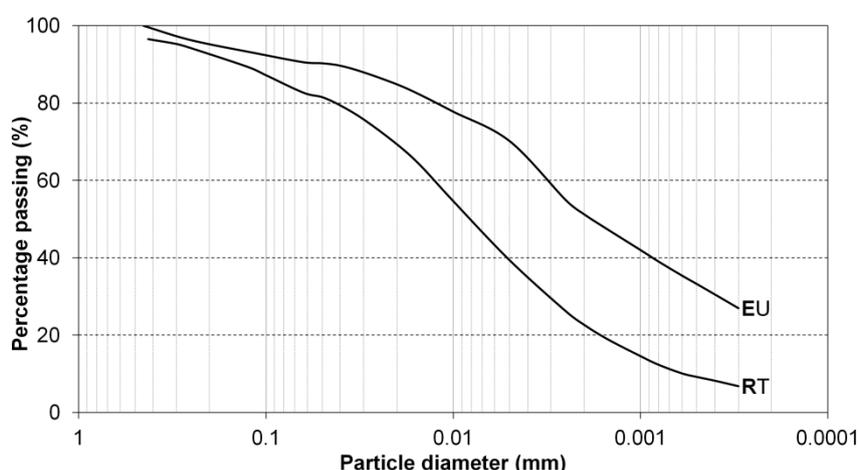


Figure 2 - Granulometric curves of raw materials.

Clay ceramics

Weight variation and visual inspection

Figure 3 and 4 represents the cumulative mass weight variations before and after samples washing, respectively. In general, it is possible to see higher values of mass variation for RT and smaller to EG. The higher or lower presence of salts within the porous structure, and the efflorescence on the surface, are reflected in the obtained mass values. This is an indication that the porosity must be greater in RT, smaller in EG and intermediate in EU.

Regarding the slopes of the curves in figure 3 is possible to see in the RT curve a constant slope up to 50 cycles, which means an equivalent mass gain in all sets of 10 cycles. In the final 10 cycles, the slope of the curve decreases, which indicates a smaller increase in mass compared to the other cycles.

The curve in figure 3 corresponding to the EU, show different slope in the first 20 cycles, having a smaller mass gain in the next sets of cycles.

The curve of EG, shows a flattening pattern, revealing a lower penetration of salts. Due to the previous interpretation in granulometric characteristics of raw material and because it is glazed, the mass gain evolution was the expected.

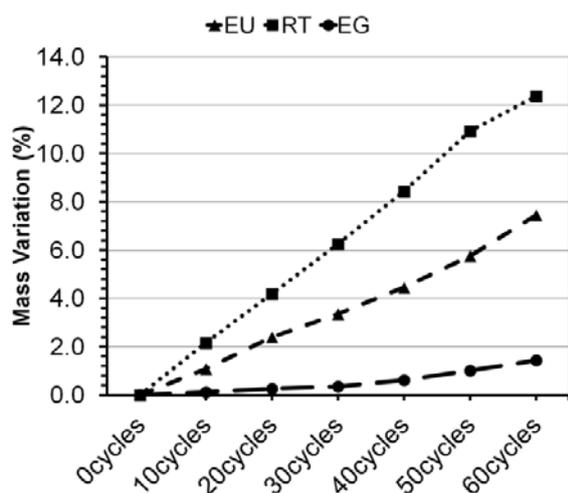


Figure 3 - Evolution of cumulative samples weight (before wash).

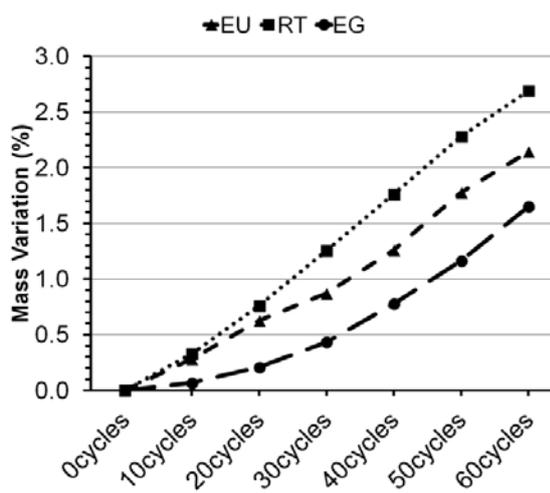


Figure 4 - Evolution of cumulative samples weight (after wash).

Figure 4 exhibits the mass gain in the three materials after their immersion in distilled water. This emphasizes the difficult of fired clay materials desalination. The results may be explained by tiles distinct porosity characteristics (porometry, quantity, geometry and its connections to the exterior). These aspects can slow down the ingress of the distilled water and reduce the effectiveness of desalination.

After all cycles, efflorescences and salt deposits were macroscopically identified in RT, the efflorescences in EU and EG have distinct morphology (figure 5).

After desalination at 10 cycles, the RT samples show dark and white pitting and after 30 cycles, the density of this pitting becomes higher.

Through observation with the aid of a binocular microscope, other aspects may be identified. Before the desalination of the ceramics is possible identify different morphologies of NaCl salts on the surface. These different forms of crystallization have been already identified and studied (Arnold & Kueng, 1985; Arnold & Zehnder, 1985; Silva & Simão, 2009). The crystal habits are related to the type of material surface and the degree of solution saturation (Arnold & Zehnder, 1985). On RT surface, salt crusts, individual and well formed crystals and bristly efflorescences were identified. The two firsts morphologies are due to surface deposition of salts. In the case of the bristly efflorescences (some needlelike) it appears that crystals start their formation in the interior of the pores growing upwards to the surface. During this growth, the crystal applied stress into the surrounding material leading to their spalling and loss of material (figure 6).

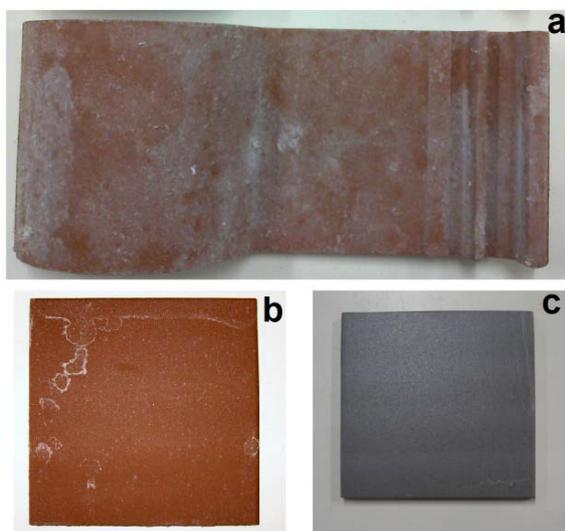


Figure 5 – Aspects of NaCl deposition on the samples: RT a); EU b); EG c).



Figure 6 – Needlelike efflorescence causing spalling.

Crystals with such aspect were not present in EU and EG. Possibly due to its low porosity some of the salts and crystallization forms are trapped inside the less accessible pores however, superficial formation of individual crystals were detected.

Throughout the various stages, some samples spots were monitorized. The sequence of figure 7 shows the evolution of a RT spot over the cycles. The visual differences are obvious, the cluster of quartz crystals cause spalling of ceramic matrix. This interface, matrix/quartz, seems to be a preferential zone to salt solution penetration and crystals growth.

Due to minerals susceptible to volume increase, different aspects of degradation were also detected. Salt can penetrate in cleavage planes of mica, such muscovite. Salt can also lead to oxidation of iron oxides like hematite. These volume increases result in visible pitting and spalling. Also the calcite, eventually present in the raw material, becomes lime by the firing temperature and when in contact with water converts into portlandite that present larger volume, causing the white pitting normally described as lime blowing (Elert *et al.* 2003). For the EU and EG the changes are not so evident. However, there are slight changes in areas near quartz crystals and loss of matrix material in the final stage of the aging test.

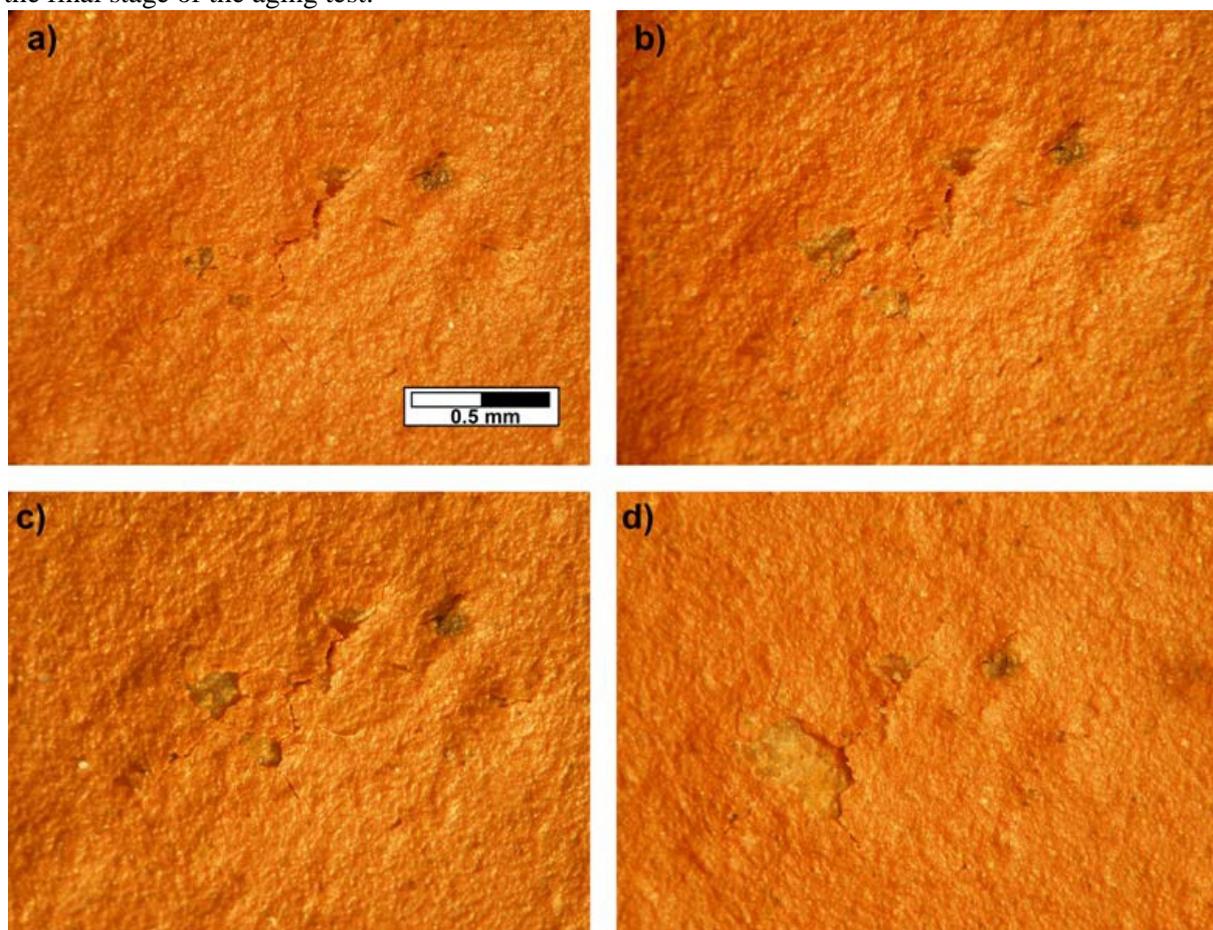


Figure 7 – Evolution at binocular microscopy of a RC spot.

X-ray diffraction

Analysing x-ray diffractograms of the three materials, stands out a peak corresponding to quartz, it is one of the most abundant mineral in all tiles. In the case of RT, hematite is also present, giving to this material its reddish colour. In EU and EG diffractogram it is possible to identify mullite in their composition. The presence of this mineral is usual to occur in ceramics manufactured with higher firing temperatures.

The x-ray diffraction analyses after the 60 cycles of salt spray in tiles are similar when compared with the first diffractograms. However, in the case of RT, there is an important new peak, corresponding to the mineral halite, which is present due to the action of the salt spray. In the case of EU and EG diffractograms, also occurs the presence of halite but with lower intensity peaks.

Resulting from minerals alteration there were no new peaks, thus there is no significant chemical alteration in the three ceramics.

Apparent porosity and water absorption

In first hand, the knowledge of apparent porosity and water absorption could give important clues in the evaluation of the bricks tendency to degrade upon salts action. The penetration and the residence of salts in the materials depend mainly on their pore structure (Benavente *et al.* 2003). Considering the relative values of water absorption for the three materials in figure 8, it can be seen that RT present the higher water absorption value, about 6.0%, followed by EU, 1.4% and finally by EG, 0.15%.

For RT the curve shows that, after 10 cycles, the water absorption suffered a significant increase of more than 0.5%. In the following cycles, the values are quite stable.

Concerning EU, after 20 and 40 cycles, the value increases linearly, to 1.9% and 2.2% respectively, with a stabilization in the last 20 cycles.

For the EG, the water absorption present a slight increase along all the 60 cycles. This seems to indicate that in clay bricks and tiles with lower water absorptions the effect of sea salt needs more time to a stabilize.

The values of apparent porosity in figure 9, can be relate with those described to water absorption. The evolution of the apparent porosity along the cycles was equivalent to the presented in figure 8 for water absorption.

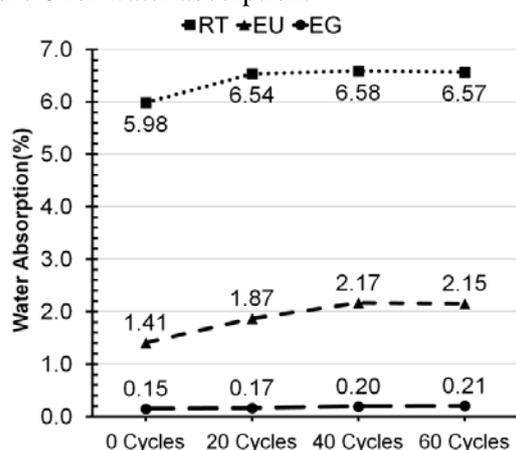


Figure 8 – Water absorption evolution of the three materials.

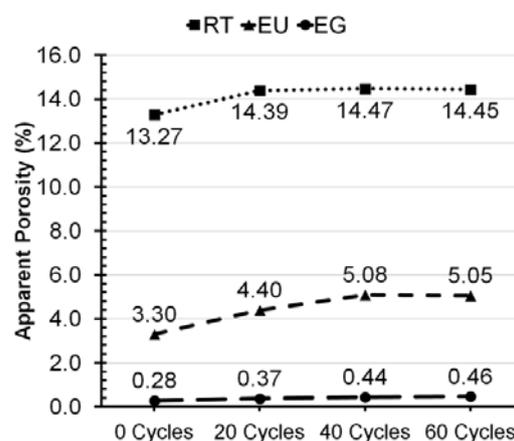


Figure 9 – Apparent porosity evolution of the three materials.

The increasing values in these two properties are explained by the, already mentioned, action of NaCl crystals growth. This cause superficial material loss and internal micro-cracks, which produce additional available voids for more incoming water in water absorption test.

Comparing the differences between the initial and final values of the three materials, it is easily detected that the samples with lower initial value have a higher rate of changes. In the case of water absorption RT, EU and EG exhibit decreases of 10, 52 and 40%, when compared to the initial state. For apparent porosity these values are 9, 53, 54%, respectively. From this point of view, the lesser porous material showed a high rate of degradation. Thus, the rate of physical damage by salt crystallization is also function of other parameters, such as the moisture supersaturation, magnitude of the repulsive forces between the salts and the confining pore walls, rates of elements supply, water evaporation and the pore size (Scherer, 2004). A higher amount of salts within the material is not sufficient to cause more damage and appears that ceramic materials with smaller pore size (EU and EG) are more favorable to supersaturation and contact between pore walls.

Modulus of rupture

The mechanical strength of fired clay materials is determined by the granulometric properties and mineralogical composition of the raw material. The manufacturing processes such as moulding features and firing temperature have also an important role in the ceramics final behavior. Figure 10 displays the modulus of rupture of the three types of samples. The RT, as expected, is the material that has the lower initial value of 19.2 MPa. On the other hand, the EG have the highest value of 51.5 MPa. The mechanical strength of the reference samples is consistent with the previous data, namely,

the firing temperature, granulometric composition and the apparent porosities. The vitrified phase results from the fusion of the various components of the clay and is connected to the maximum firing temperature. Usually, higher vitreous phase result in higher mechanical strength.

With the evolution of cycles, the mechanical strength decay in all the three materials and after all the aging cycles. The exception is the RT, after 20 cycles of salt spray the strength decrease to 17.1 MPa, in the next 20 cycles the value increase and is followed by a decrease. Show a final value of 18.1 MPa, which represents a 6% of loss in strength after the aging tests. This could be explained by a partial filling of the pores which causes an increase in compressive strength. It is known that under flexural load exists a compressive zone that could contribute to an increase in the overall flexural strength. However, to confirm this explanation it is required further investigation. The final value of modulus of rupture in EU was 23 MPa, which is a significant decrease after the aging tests, this represents a decrease of approximately 30%.

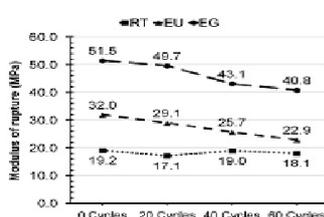


Figure 10 – Evolution of rupture modulus with the aging cycles.

In EG the action of the first 20 cycles results in a decrease of 2 MPa. However, the main loss of strength occurred in the followed 20 cycles where a value of 43 MPa was obtained. At the end of the aging tests, the modulus of rupture was approximately of 41 MPa, which corresponds to a total decrease of 20%. The effect of the salt spray had a clearly negative influence on the strength properties of these materials.

It is possible to note that the ceramics with higher initial mechanical strength had a higher rate of degradation. Thus, analogous interpretation to that made for the case of water absorption and apparent porosity can be performed. Elert *et al.* (2003) studied the deterioration by salt crystallization in clay bricks and concluded that relatively large pores are favorable to durability, which confirms these results and interpretations.

Conclusion

The salt spray action in modern facing bricks and tiles cause degradation. This degradation is expressed through physical and mechanical alterations.

The method for evaluating the degradation through weight loss, recommended by EN 14147, is not so appropriate for ceramics. From this perspective, the water absorption, apparent porosity and flexural tests proved to be suitable for quantifying physical and mechanical degradation.

Through visual inspection it was possible identify alteration forms in RT and EU. It was found that minerals with higher hardness (e.g. quartz) than the ceramic matrix and with dimensions close or greater than 0.5 mm on the tiles surface enhance the loss of material by spalling and pitting. When subjected to the salts crystallization their presence can be important to material performance by two reasons: An interface between the crystal and the ceramic matrix which facilitates the penetration of salt solution and a preferred zone for crystal growth. Different strengths between adjacent materials create heterogeneity. The stresses generated by crystal growth in these two materials results into a failure, usually of the ceramic matrix.

The minerals susceptible to swelling near the surface are also an important source of degradation.

The evolution of water absorption and apparent porosity were not favorable to the expected performance in the three materials. The salt spray cycles, produced an increase in water absorption and apparent porosity.

It was concluded by quantifying the mechanical properties, the existence of degradation at different scales between samples. In the two of the most important characteristics, water absorption and modulus of rupture, the ceramics with higher changes are the facing bricks. The strength showed a decrease of more than 20% and water absorption increased more than 40%. These observations allowed conclude that large pores have a minor role on the degradation, while a more severe damage may occur in the ceramics with a large amount of smaller pores.

References:

- Arnold, A., Kueng, A. – Crystallization and habits of salt efflorescences on walls I: Methods of Investigation and Habits. Félix G (ed.) Proc. 5th Int. Congress on the Deterioration and Conservation of Stone, Lausanne, (1985): 255–267.
- Arnold, A., Zehnder, K. – Crystallization and habits of salt efflorescences on walls II: Conditions of crystallization. Félix G (ed.) Proc. 5th Int. Congress on the Deterioration and Conservation of Stone, Lausanne, (1985): 269–277.
- Benavente, D., M.A. García del Cura, *et al.* Salt influence on evaporation from porous building rocks. *Const. Build. Mater.* 17(2) (2003): 113-122.
- Brocken H., Nijland T. - White efflorescence on brick masonry and concrete masonry blocks, with special emphasis on sulfate efflorescence on concrete blocks. *Constr Build Mater* 18 (2004): 315–323.
- Cardell, C., Delalieux, F., Roumpopoulos, K., Auger, F., Van Grieken, R. - Salt-induced decay in calcareous stone monuments and buildings in a marine environment in SW France. *Const Build Mater*, 17 (2003):165-179.
- Elert, K., G. Cultrone, *et al.* Durability of bricks used in the conservation of historic buildings - influence of composition and microstructure. *Journal of Cultural Heritage* 4(2) (2003): 91-99.
- EN 10545-3:1995 -Ceramic tiles - Part 3: Determination of water absorption, apparent porosity, apparent relative density and bulk density. CEN.
- EN 10545-4:2004 - Ceramic tiles. Part 4: Determination of modulus of rupture and breaking strength. CEN.
- EN 14147:2003 – Natural stone test methods: Determination of resistance to aging by salt mist. CEN.
- EN 14411:2006 - Ceramic tiles. Definitions, classification, characteristics and marking. CEN.
- Galembeck, T., Simão, J., Silva, Z., Artur, A.C. - Effects of salty solutions in dimension stones from Southeast Brazil. II International Congress of Dimension Stone. Carrara, Italy: 517, 2008.
- Garcez, N. – Sistema de inspeção e diagnóstico de revestimentos exteriores de coberturas inclinadas. Master Thesis, IST, Lisboa, 2009.
- Houben, H., Guillaud, H. – Earth Constuction: A comprehensive guide. ITDG Publishing. London, 1994.
- IPCC - Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). *Climate Change 2001: The Scientific Basis. Cambridge University Press.* UK, 2001.
- Kuchitsu N., Ishizaki T., Nishiura, T. - Salt weathering of the brick monuments in Ayutthaya, Thailand. *Engin Geol* 55 (1–2) (2000), : 91–99.
- Leeuw, G. – Sea Spray Aerosol Production Form Waves Breaking in the Surf Zone. *J. Aerosol Sci.*30 (1999): 63-64.
- Lubelli B., Hees R., Groot C. - The role of sea salts in the occurrence of different damage mechanisms and decay patterns on brick masonry. *Constr Build Mater*,18 (2004): 119–124.
- Lucas, J. – Azulejos ou Ladrilhos Cerâmicos. Descrição geral, exigências normativas, classificação funcional, LNEC, Lisboa, 2003.
- Moser, K. – Towards the practical evaluation of service life – Illustrative application of the probabilistic approach. *Durability of Build Mater and Compon*, 8 (1999): 1319-1329.
- Ottosen, L.M, Pedersen A.J., Inge, R. Electrochemical desalination of salt infected limestone masonry of a historic warehouse. *Proceedings of Structural Faults and Repair*, 2012.

- Ottosen, L.M, Pedersen A.J., Inge, R. Salt-related problems in brick masonry and electrokinetic removal of salts. *Journal of Building Appraisal* 3 (3) (2007): 181–194.
- Parras, J. – *Mineralogía y propiedades cerâmicas de pizarras paleozoicas de ciudad real*. Ed. Uni. Castilla-La Mancha, 1997.
- Rörig-Dalgaard, I., L. M. Ottosen, *et al.*. Diffusion and electromigration in clay bricks influenced by differences in the pore system resulting from firing." *Const Build Mater* 27(1) (2012): 390-397
- Scherer, G. W.. Stress from crystallization of salt. *Cement and Concrete Research* 34(9) (2004): 1613-1624.
- Silva Z.S.G., Simão J.A.R. - The role of salt fog on alteration of dimension stone. *Constr Build Mater*, 23 (2009): 3321–3327.
- Silva, B.; Rivas, T., García-Rodeja E., Prieto, B. – Distribution of ions of marine origin in Galicia (NW Spain) as a function of distance from sea. *Atmosp Environ*, 41 (2007): 4396-4407.
- Sokolar, R. and L. Smetanova, Dry pressed ceramic tiles based on fly ash–clay body: Influence of fly ash granulometry and pentasodium triphosphate addition. *Ceramics International* 36(1) (2010): 215-221.
- Zeza F, Macri F. - Marine aerosol and stone decay. *Sci Total Environ*, 167 (1995): 123-43.

CHARACTERISATION OF ALKALINE LIPASE FROM AN ARCTIC YEAST STRAIN *RHODOSPORIDIUM BABJEVAE* BD19

Marcin lukaszewicz, Prof.

Sławomir Jablonski, Msc

Anna Krasowska, PhD

University of Wrocław/ Faculty of Biotechnology,
Department of Biotransformation, Wrocław, Poland

Abstract:

The yeast strain isolated from freshwater from Arctic Archipelago of Svalbard was identified biochemically by API C 20 aux test, genetically by 18S rRNA and ITS sequencing as *Rhodospiridium babjevae* BD19. *R. babjevae* secreted into the culture medium lipase with the activity 67.7 U/ml after 144 hours of growth on Tween 80 as a sole of carbon source and 115.5 U/ml after 144 hours on inorganic ammonium sulphate as a sole nitrogen source. The molecular mass of purified lipase was 37.6 kDa and optimum esterase activity was at pH 9.0 and 30 °C. The chain length specificity of the enzyme was determined using p-nitrophenyl fatty acids esters as substrates and the highest hydrolysis rate was with p-nitrophenyl caprylate. The *R. babjevae* BD19 lipase seems to have unique 1-position specificity when hydrolyzing triolein.

Key Words: *Rhodospiridium babjevae*, Arctic yeast, lipase, esterase

Introduction

Lipases are capable of catalyzing a large number of reactions with high regio- and enantioselectivity (Schmid and Verger, 1998). They catalyze not only hydrolysis, but also esterification, acylation or transesterification reactions (Reis, et al., 2009). For example, hydrolytic reactions of lipases showed potential applications in degrading oil and fats in wastewater generated by dairy industries, producers of edible oils and fat refineries (Mendes and Castro, 2005). Acylation of carbohydrates is used for production of non-ionic surfactants for food and pharmaceutical industries (Nakamura, 1997). Transesterification reactions catalyzed by lipases have a prominent significance in the oil industry for tailoring vegetable oils (Gupta, et al., 2003).

The majority of lipases used for biotechnological purposes are isolated from bacteria and fungi (Schmid and Verger, 1998, Arpigny and Jaeger, 1999). Compared to other microorganisms, yeast enzymes are less used in commercial applications and only some species have been studied for potential industrial applications (Brizzini, 2006).

Extreme environments can serve as a reservoir of microorganisms secreting enzymes with novel, useful properties. Recent investigations showed many important qualities of lipases secreted by psychrophilic microorganisms (Joseph, et al., 2008). While the Antarctic yeast *Candida antarctica* is the most extensively studied microorganism due to its lipase (Kirk and Christensen, 2002), Brizzio and collaborators (Brizzio, et al., 2007) isolated new basidiomycetous yeasts characterized by their cold-active enzymes from glacial and subglacial waters. Among other yeast strains, they found *Rhodospiridium babjevae* with lipolytic activity.

Lipases isolated from different sources have a wide range of properties in terms of positional and substrate specificity, thermostability or pH optimum, and catalyze a number of reactions used for development of specific industrial applications such as organic synthesis, detergent additive or food processing. Thus, isolation of new strains secreting lipases which catalyze unique reactions will enable chemical industries to catalyze new compounds (drugs) or reduce the costs of existing technologies.

Screening of 132 morphologically distinct bacteria and yeasts isolated from freshwater from Arctic Archipelago of Svalbard for lipase secretion resulted in the isolation of ten species having high

lipase activity (Krasowska, et al., 2007, Krasowska and Łukaszewicz, 2011). In previous works we have identified *Pseudomonas fluorescens* strain producing novel biosurfactants (Janek, et al., 2011) and characterized their properties (Janek, et al., 2012). In this report we have identified a novel yeast strain *Rhodosporidium babjevae* BD19, optimized medium and growth conditions for secretion of extracellular lipase, partially purified the enzyme and characterized its specificity and activity.

Materials and methods

Isolation of yeast strain and culture conditions

The strain BD19 was isolated from freshwater from the Arctic Archipelago of Svalbard.

The strain was grown in 300 ml flasks containing 50 ml LB medium (5 g/l yeast extract, 10 g/l bacto-tryptone, 10 g/l NaCl). Flasks were inoculated with 5 ml of overnight pre-culture and incubated for 24 h at 28 °C with agitation (180 rpm).

The effect of different oils as carbon sources on lipase production by BD19 strain was investigated on a mineral medium (Papaparaskevas, et al., 1992) supplemented with 2 g/l of tested carbon source (glucose, Tween 80, rape, sunflower, flax or olive oil). For optimization of nitrogen source, BD19 strain was cultivated on the same mineral medium containing 20 g/l Tween 80 and supplemented with 10 g/l of tested substrate (yeast extract, bacto peptone, ammonium sulfate, urea or ammonium chloride).

Identification of the yeast BD19 strain

The pure culture of strain BD19 was identified by API C 20 aux test (BioMérieux, Marcy l'Etoile, France).

Moreover, the 18S rRNA gene was sequenced to support the biochemical identification. Genomic DNA was extracted with UltraClean™ Microbial DNA Isolation Kit (MoBio®) according to manufacturer's instructions and 18S rRNA gene fragment was amplified with the primers: EucA 5'-AACCTGGTTGATCCTGCCAGT-3' and EucB 5' TGATCCTTCTGCAGGTTACCTAC3'. The amplification conditions were as follows: 94 °C for 10 min., 30 cycles, 94 °C for 1 min., 59 °C for 1 min., 72 °C for 2 min., and final synthesis at 72 °C for 5 min. ITS fragments were amplified with the primers: ITS1 5' TCCGTAGGTGAACCTGCGG 3' and ITS4 5' TCCTCCGCTTATTGATATGC 3'. The PCR conditions were as follows: 94 °C for 10 min., followed by 30 cycles, 94 °C for 1 min., 48 °C for 1 min., 72 °C for 1 min., and final synthesis at 72 °C for 5 min. The PCR products were separated by agarose gel electrophoresis and purified for sequencing using a GeneMATRIX PCR/DNA Clean-up Purification kits (EURx, Gdansk, Poland). The sequences obtained were compared to rDNA sequences from the GeneBank (<http://www.ncbi.nlm.nih.gov/BLAST/>). ITS fragments obtained from GeneBank database were aligned with ClustalW (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>) and the phylogenetic tree was computed with Jalview 2.4.0.b2 using the neighbor-joining method.

Purification of lipase

The *R. babjevae* BD19 was cultured with 20 g/l of Tween 80, which enabled phase separation (Albertsson, 1986). Cells were removed by centrifugation at 4000 g for 10 minutes. To 100 ml of supernatant 37.6 g ammonium sulphate was added (50% saturation at 20 °C). After 30 min of gentle agitation ensuring solubilization and equilibrium of the partition, phase separation was accomplished by centrifugation at 25 200 g for 20 minutes 4 °C. The upper phase was collected and dialyzed against distilled water overnight. Samples were withdrawn for lipase assay (Gupta, et al., 2002) and total protein determination (Bradford, 1976).

Esterase assays

Lipase activity (Gupta, et al., 2002) was assayed spectrophotometrically with modifications using p-nitrophenyl palmitate (p-NPP) as substrate. Briefly, the reaction buffer contained Tris HCl 50 mM pH=9.0, arabic gum 0,9 mg/ml and Triton X100 4 mg/ml. To 10 ml of the reaction buffer 1 ml of p-NPP solution (3 mg/ml) in isopropanol was added before the test and 230 µl aliquots were dispensed on a 96-well plate. Reaction was started by addition of 20 µl of enzyme solution. The reaction was conducted at 23 °C for 30 minutes, and the absorbance at 410 nm was measured every 5 minutes. One unit was defined as the amount of the enzyme releasing 1 nmol of free p-nitrophenol per minute.

The chain length specificity of the enzyme was determined using p-nitrophenyl esters (Sigma) of acetate (C₂), butyrate (C₄), caproate (C₆), caprylate (C₈), decanoate (C₁₀), laurate (C₁₂), palmitate (C₁₆) and stearate (C₁₈) as substrates at pH 8.8 and at 23 °C.

Effect of pH and temperature on enzyme activity

The impact of pH was determined by measuring lipase activity at 30 °C in 0.5 M Tris HCl buffer at pH ranging from 7 to 10. The optimal temperature of the enzyme was determined by measuring its activity in a temperatures range (25-50 °C) in 0.05 M Tris HCl buffer, pH 9.0.

Positional specificity

Positional specificity was determined by analyzing lipolytic products of triolein by thin layer chromatography (TLC) on silica gel 60 plate (Merck, Darmstadt, Germany). The total reaction mixture (1.5 ml) containing 33

incubated with 33 μ l of triolein at 30 °C with shaking at 800 rpm. Aliquots of 200 μ l were removed at 0, 2, 4 and 24 hours and the products were extracted with 0.5 ml of diethyl ether. The extracted samples were applied to the silica gel-60 TLC plates for product analysis. The plate was developed in a chamber with the solvent mixture chloroform/ acetone/acetic acid (95:4:1). The hydrolysis products were visualized using saturated iodine vapor and compared with standards (oleic acid, 1-oleylglycerol, 1,2-sn-diolelylglycerol, 1,3-diolelylglycerol and triolein) from Sigma-Aldrich Sp. z o.o., Poland.

Results and Discussion

Identification of *Rhodosporidium babjevae* BD19

Cells of the isolated strain were oval, yeast-shaped, redish and aerobic. The strain BD19 was arabinose-, xylose-, idonitol-, galactose-, sorbitol-, metyl-glucopyranoside-, sucrose-, and raffinose utilization-positive and xylitol-, inositol-, cellobiose-, lactose-, maltose-, trehalose-, and melesitose utilization-negative. These biochemical results were not sufficient for classification to the genus. Comparative sequence analysis of the 18S rDNA gene (968 bp) in the GeneBank database revealed that the strain BD19 had 99% homology to *Rhodotorula glutinis*, *Rhodotorula graminis* and *Rhodosporidium babjevae*.

Rhodotorula species are ubiquitous in human environment and have been isolated from different environments such as soil (Mok, et al., 1984), fruit juice (Tournas, et al., 2006) or sea water (Nagahama, et al., 2001, Loureiro, et al., 2005). In 1967 Banno first described *Rhodosporidium toruloides* as a telomorph in the life cycle of *R. glutinis* (Banno, 1967). Several following studies indicated phenotypic and genetic heterogeneity in *Rhodotorula* gender (Yamazaki and Komagata, 1981, Hamamoto, 1987, Vancanneyt, 1992, Gadanho and Sampaio, 2002).

To further classify strain BD19 the Internal Transcribed Spacer (ITS) region of the rDNA from BD19 was PCR amplified, sequenced and compared with GeneBank database. The ITS regions are adequate to resolve the relationship between closely related genera (Libkind, et al., 2008). Neighbor-joining analysis of the data was carried out with the program Neighbor of the PHYLIP package.

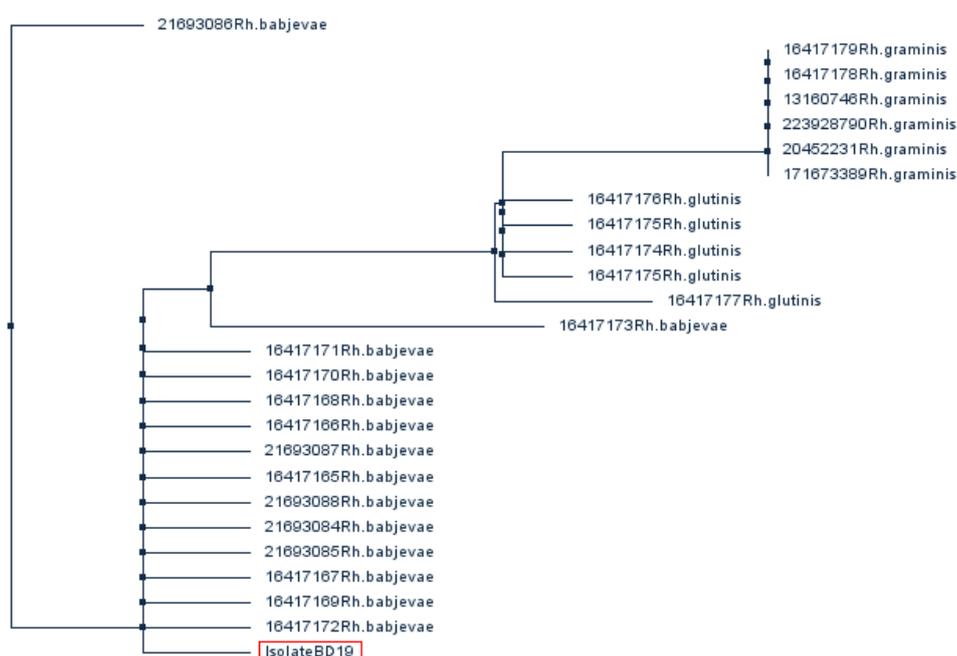


Fig.1 Phylogenetic tree of *R. glutinis*, *R. graminis*, and *R. babjevae* obtained by neighbor-joining analysis of ITS region of rDNA. The GI number of sequence is placed before the species name

The sequence obtained from BD19 strain (GeneBank JN003620) was placed together with sequences from microorganisms classified as *R. babjevae* (Fig. 1). The nucleotide sequence of ITS region from BD19 strain was identical to twelve other *R. babjevae* sequences included in the phylogenetic tree.

The influence of carbon and nitrogen sources on lipase production

In general, extracellular lipase production by yeast depends greatly on the medium composition (Papaparaskevas, et al., 1992, Darvishi, 2009, Kim, et al., 2009). With *R. babjevae* BD19 strain, the highest lipase activity (67.7 U/ml) was obtained after 144 hours of growth on Tween 80 as a sole of carbon source. The activity was 18, 45.5, 41 and 55% smaller when BD19 strain was grown on carbon sources such as olive, sunflower, rape and palm oils, respectively (Table 1). The lowest lipase activity was obtained using glucose as the sole carbon source (10.5 U/ml after 144 h incubation). In contrast to BD19 strain induced to lipase production by water-soluble emulsifier (Tween 80), Papaparaskevas and collaborators (Papaparaskevas, et al., 1992) found the highest activity of lipase secreted by *R. glutinis* cultured on olive oil (29.3 U/ml). In addition, the activity dramatically decreased when other substrates such as palm oil (2.4 U/ml), sunflower oil (2.0 U/ml) or Tween 80 (1.8 U/ml) were used (Papaparaskevas, et al., 1992). In this respect *R. babjevae* BD19 lipase production was much less influenced by different carbon sources such as esters of fatty acids.

Table 1. Effect of carbon source on lipase secretion by *R. babjevae* BD19. The strain was grown in a medium with yeast extract as a nitrogen source and different carbon sources at neutral pH. The lipase activity in culture supernatants was measured after 72 and 144 hours of incubation at 30 °C ±SD, n=3.

Carbon source	Activity [U/ml]	
	72 hours	144 hours
Glucose	1.16 ± 0.24	10.5 ± 1.16
Tween 80	47.69 ± 4.91	67.7 ± 16.9
Olive oil	41.57 ± 10.12	55.7 ± 9.06
Sunflower oil	43.96 ± 7.14	36.9 ± 9.74
Rape oil	51.9 ± 3.78	39.9 ± 0.07
Palm oil	17.79 ± 3.62	30.2 ± 4.51

The highest lipase yield (115,4 U/ml) produced by *R. babjevae* BD19 was obtained after 144 hours of incubation at 30 °C on inorganic nitrogen source (ammonium sulphate as a sole nitrogen source). When the strain was grown 144 hours on yeast extract, bacto peptone, ammonium chloride or urea as nitrogen sources, lipase activities were 65, 32, 27 and 99% lower, respectively (Table 2). Lipase secreted by *R. babjevae* BD19 after 72 hours of incubation on yeast extract, bacto peptone, ammonium sulphate, ammonium chloride or urea as carbon sources was accordingly 29.4-, 22-, 31-, 37- and 8-fold more active than lipase secreted by *R. glutinis* after 6 days of incubation (Papaparaskevas, et al., 1992).

Generally, microorganisms provide high yields of lipase when organic nitrogen sources are used, with the exception of *R. glutinis* which requires inorganic nitrogen source such as ammonium phosphate for high lipase production (Papaparaskevas, et al., 1992).

Table 2. Effect of nitrogen source on lipase production. *R. babjevae* BD19. The strain was grown in cultivation medium with Tween 80 as carbon source and different nitrogen sources at neutral pH. The lipase activity in culture supernatants was measured after 72 and 144 hours of incubation at 30 °C ±SD, n=3.

Nitrogen source	Activity [U/ml]	
	72 hours	144 hours
Yeast extract	44.10 ± 2.77	40.73 ± 4.71
Bacto peptone	26.57 ± 0.85	37.1 ± 2.71
Ammonium sulphate	37.66 ± 1.35	115.4 ± 5.44
Ammonium chloride	37.23 ± 4.54	84.0 ± 2.57
Urea	12.38 ± 4.75	1.3 ± 0.14

Purification of the enzyme

A two-phase-system purification procedure was used to characterize extracellular lipase from B19 strain. This procedure increased the specific activity of lipase 3.84-fold with a yield of 88.38% (Table 3). The lipase secreted into the medium had a molecular mass of 37.6 kDa by SDS-PAGE.

Table 3. Purification of the *R. babjevae*. BD19 extracellular lipase.

	Total activity [U]	Total protein [mg]	Specific activity [U/mg]	Purification fold	Yield [%]
Culture filtrate	5575.46	4305.40	1.29	1.00	100.00
Partitioning in two- phase system	4927.36	991.79	4.97	3.84	88.38

Effect of pH and temperature on enzyme activity

Lipases secreted by microorganisms mostly have optimum of activity in a weakly alkaline environment and at temperatures ranging from 30 to 40 °C (Choo, et al., 1998, Yu, et al., 2007). Lipase secreted by *R. glutinis* had optimum activity at 35 °C and pH 7.5 (Papaparaskevas, et al., 1992).

The highest activity of lipase from *R. babjevae* BD19 was obtained with pH 9.0 at 30 °C; it decreased by 30% when pH was increased to 10.0 and by 70% when pH decreased to 7.0 (Fig.2).

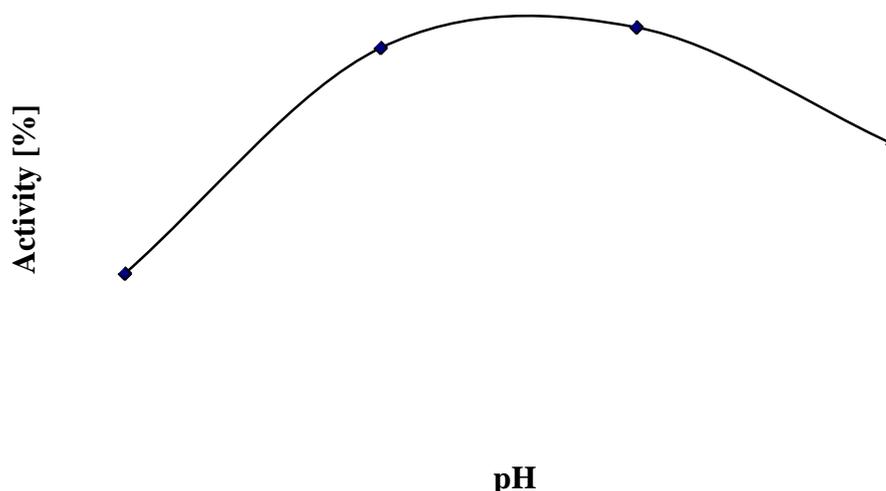


Fig. 2. Influence of pH on lipase activity. Data are given as means \pm SD, n=3.

The optimum temperature for enzyme activity was 30 °C; it decreased with rising temperature, with 80% loss of original activity at 50 °C (Fig 3). Fungal lipases are generally active at 40-50 °C (Hiol, et al., 2000). The optimum pH was higher and optimum of temperature was lower than that for lipase-producing *R. glutinis* (pH 8.0 at 35 °C) (Papaparaskevas, et al., 1992), *R. mucilaginosa* (pH 7.0-8.0 at 40 °C) (Zimmer, et al., 2006) and *R. pilimanae* (pH 4.0 and 7.0 at 45 °C) (Muderhwa, et al., 1986).

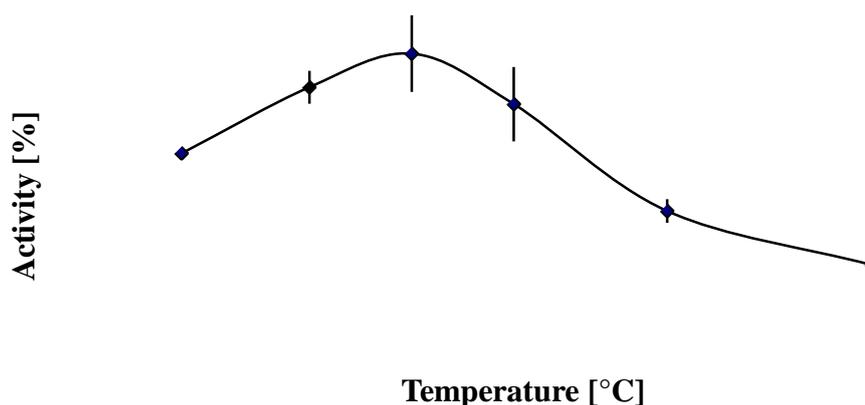


Fig 3. Influence of temperature on lipase activity. Data are given as means \pm SD, n=3.

Substrate specificity with p-nitrophenyl esters

The substrate specificity of partially purified *R. babjeviae* BD19 lipase towards p-nitrophenyl esters with alkyl chain length ranging from C2 to C18 was determined (Fig. 4). The lipase showed the highest activity with the use of p-nitrophenyl caprylate ester (C₈). Specific activity decreased gradually as the alkyl chain was shortened or lengthened. When p-nitrophenyl acetate (C₂) was used as substrate no significant activity was measured.

Grognum and collaborators (Grognum and Reymond, 2004) classified lipases according to their specificity as dependent on the chain length. They suggested classification into 3 groups: the short-chain (C₄ \pm C₆), long-chain (C₁₀ \pm C₁₆), and middle-chain (C₈ \pm C₁₀) reactivity. According to such classification *R. babjeviae* BD19 lipase has the highest specificity towards middle chain fatty acid esters. Lipase of *R. glutinis* was very active with nitrophenyl butyrate (C₄) and nitrophenyl laurate (C₁₂) (Hatzinikolaou, et al., 1999); *Rhodotorula minuta* lipase had high activity towards p-nitrophenyl butyrate (C₄) in lecithin emulsion (Cinelli, et al., 2006). Lipases which utilized natural methylthioesters of volatile fatty acids (C₃-C₈) have considerable importance for flavor industry; thus, e. g. S-methyl butanetioate is an important constituent of dairy aromas (Rajendran, et al., 2009).

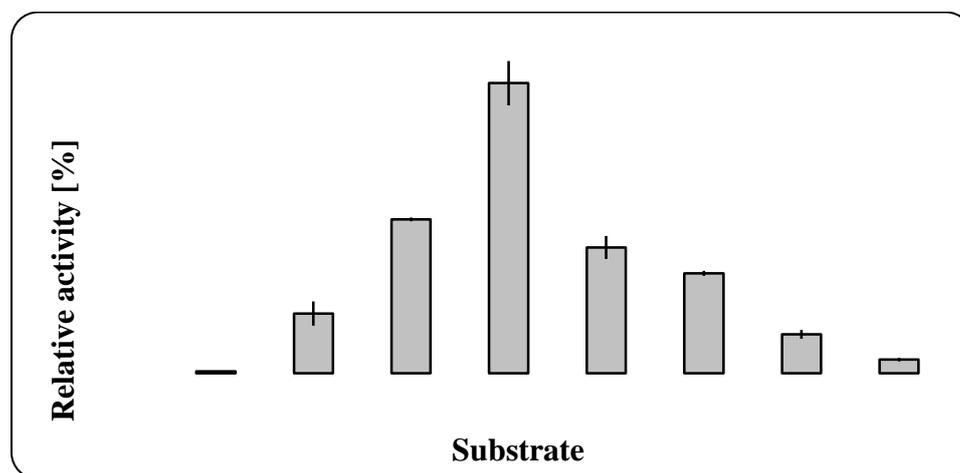


Fig. 4. *R. babjeviae* BD19 lipase substrate specificity. Activity towards different p-nitrophenyl esters was determined at 30 °C and pH=9.0. Data are given as means \pm SD, n=3

Positional specificity

Marcae and Hammond (Marcae, 1985) systematized lipases according to their regio-specificity. The first group of lipases is not specific and released fatty acids from all three positions. The second group released the fatty acids specifically from the outer 1 and 3 positions of acylglycerols. Microbial lipases show usually 1,3-positional specificity, releasing three products from triolein: 2-monoacylglycerol, 1,2- (or 2,3) – diacylglycerol and free fatty acid (Zimmer, et al., 2006,

Aloulou, et al., 2007). Lipases with specificity towards 2-position are very rare in nature. Only few lipases have 2-position specificity, such as lipase from *Geotrichum* sp. FO401B (Ota, et al., 2000), lipases from *Geotrichum candidum* ATCC 34614 (Sugihara, et al., 1994), or lipase from *Aspergillus niger* NCIM 1207 (Mhetras, et al., 2009).

Triolein was used as a substrate to investigate positional specificity of purified lipase. The separation of triolein hydrolysis products was performed by thin-layer chromatography (Fig. 5). The major product end product of the *R. babjevae* BD19 lipase was 1,2-diolein and prolonged incubation for 24 h strongly indicates that the enzyme has 1-position specificity.

Lipase from *R. babjevae* BD19 seems to be unique in its positional specificity because it cleaved triolein only once in 1-position, releasing only two products: 1,2 (or -2,3)-diolein and free fatty acid.

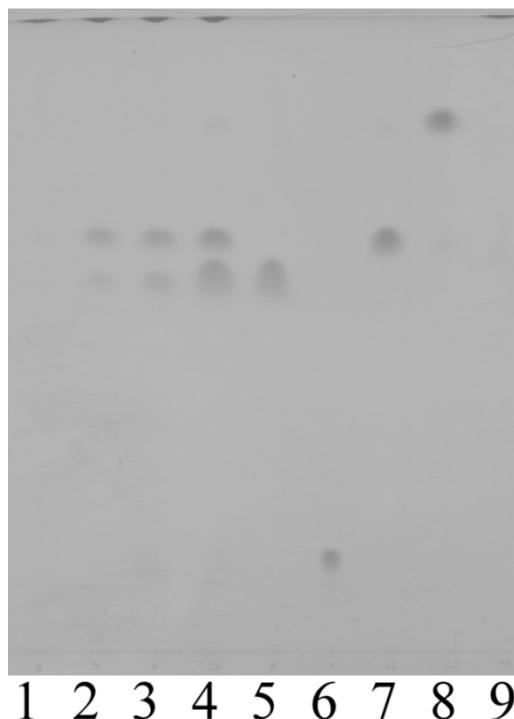


Fig. 5 Thin-layer chromatography of the hydrolysis products of triolein catalyzed by *R. babjevae* BD19 lipase. Lane 1: 0 h, Lane 2: 2h, Lane 3: 4 h, Lane 4: 24 h, Lane 5: oleic acid, Lane 6: 1-oleylglycerol, Lane 7: 1,2-sn-dioleoylglycerol, Lane 8: 1,3-dioleoylglycerol, Lane 9: triolein.

Conclusions

Biochemical and genetic characterization of the BD19 strain from Arctic fresh water led to its classification as *R. babjevae*.

The highest concentration of the *R. babjevae* BD 19 extracellular lipase in the medium was obtained after 144 hours on Tween 80 as a sole of carbon source and after 144 hours on inorganic nitrogen source (ammonium sulphate as a sole nitrogen source).

The purified lipase had a molecular mass of 37.6 kDa.

The maximum activity of the lipase was obtained at pH 9.0 and at 30 °C.

The lipase had the highest hydrolytic activity towards 8-carbon alkyl chain substrate - p-nitrophenyl caprylate.

The lipase is unique in comparison to other described enzymes since it hydrolyzed triolein to 1,2-diolein and in this respect.

Acknowledgments

We thank prof. J. P. Sampaio for phylogenetic analysis of *Rhodospiridium babjevae* strain. This work was financial supported by grants from the Polish National Centre for Science

N N302 640940, Polish National Centre for Research and Development KB/48/13639/IT1-B/U/08 and grant EU POIG.01.01.02-00-016/2008.

References:

- Albertsson PA. Partitioning of Cell Particles and Macromolecules, Third Edition, John Wiley & Sons. 1986.
- Aloulou AJ, Rodrigues A, Puccinelli D, Mouz N, Leclaire J, Leblond Y and Carriere F. Purification and biochemical characterization of the LIP2 from *Yarrowia lipolytica*. Biochim. Biophys. Acta. 2007: 1771, 228-237
- Arpigny JL and Jaeger KE, Bacterial lipolytic enzymes: classification and properties. Biochem J. 1999: 343, 177-183
- Babu JW, Pramod R and George T, Cold active microbial lipases: some hot issues and recent development. Biot. Adv. 2008: 26, 457-470
- Banno I, Studies of the sexuality of *Rhodotorula*. J. Gen. Appl. Microbiol. 1967: 13, 169-196
- Bradford M, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 1976: 72, 248-250
- Brizzio S, Turchetti B, de Garcia V, Libkind D, Buzzini P, and van Brook M, Extracellular enzymatic activities of basidiomycetous yeasts isolated from glacial and subglacial waters of northwest Patagonia (Argentina). Can. J. Microbiol. 2007: 53, 519-525
- Buzzini P, and Vaughan-Martini A, Yeast biodiversity and biotechnology in: Yeast handbook, Biodiversity and ecophysiology of yeast, Ed. C.A. Rosa and G. Peter, springer-Verlag Berlin, 2006: 533-559
- Choo D, Kurihara T, Suzuki T, Soda K and Esaki N, A cold-adapted lipase of an Alaskan psychrotroph *Pseudomonas* sp. Strain B11-1: gene cloning and enzyme purification and characterization. Appl. Evn. Microb. 1999: 64, 486-491
- Cinelli G, Cuomo F, Hochkoeppler A, Ceglie A, and Lopez F. Use of *Rhodotorula minuta* live cells hosted in water-in-oil macroemulsion for biotransformation reaction. Biotechnol. Prog. 2006: 22, 689-695
- Darvishi F, Nahvi I, Zarkesh-Esfahani H, and Momenbeik F, Effect of plant oils upon lipase and citric acid production in *Yarrowia lipolytica* yeast. J. Biomed. Biotech. 2009: 56, 29-43
- Gadanho M and Sampaio JP, Polyphasic taxonomy of the basidiomycetous yeast genus *Rhodotorula*: *Rh. glutinis* sensu stricto and *Rh. dairenensis* comb.nov., FEMS Yeast Res. 2002: 2, 47-58
- Grognux J and Reymond JL, Classifying Enzymes from Selectivity Fingerprints. Chem. Biochem. 2004: 5, 826 – 831
- Gupta R, Rathi P and Bradoo S, Lipase mediated upgradation of dietary fats and oils. Crit. Rev. Food Sci. Nutr. 2003: 43, 635-644
- Gupta N, Rathi P and Gupta R, Simplified para-nitrophenyl palmitate assay for lipases and esterases. Anal. Biochem. 2002: 311, 98-99
- Hak-Ryul K, In-Hwan K, Ching TH, Kwang-il K and Beom-Soo S, Production of a Novel Cold-Active Lipase from *Pichia lynferdii* Y-7723. J. Agric. Food Chem. 2010: 58, 1322-1326
- Hamamoto M, Sugiyama J and Komagata K, DNA-DNA reassociation studies of strains in the genera *Rhodospiridium* and *Rhodotorula*. J. Gen. Appl. Microbiol. 1987: 33, 57-73
- Hatzinikolaou DG, Kourentzi E, Stamatis H, Christakopoulos P, Kolisis FN, Kekos D and Macris BJ, A novel lipolytic activity of *Rhodotorula glutinis* cells: production, partial characterization and application in the synthesis of esters. J. Biosci. Bioeng. 1999: 88, 53–56
- Hiol A, Jonzo MD, Rugani N, Druet D, Sarda L and Comeau LC, Purification and characterization of an extracellular lipase from thermophilic *Rhizopus oryzae* strain isolated from palm fruit. Enz. Microbiol. Technol. 2000: 26, 421-430
- Janek T, Łukaszewicz M, Rezanka T and Krasowska A, Isolation and characterization of two new lipopeptide biosurfactants produced by *Pseudomonas fluorescens* BD5 isolated from water from the Arctic Archipelago of Svalbard. Bioresource technology 2010: 101, 6118-6123
- Janek T, Łukaszewicz M and Krasowska A, Antiadhesive activity of the biosurfactant pseudofactin II secreted by the Arctic bacterium *Pseudomonas fluorescens* BD5. BMC microbiology 2012: 12, 24
- Kirk O and Christensen MW, Lipases from *Candida antarctica*: unique biocatalysis from a unique origin. Org. Process Res. Dev. 2002: 6, 446-451
- Krasowska A, Dąbrowska B and Łukaszewicz M, Isolation and characterization of microorganisms from Arctic archipelago of Svalbard. J. Biotechnol. 2007: doi:10.1016/j.jbiotec.2007.07.718

- Krasowska A and Łukaszewicz M, Izolacja, identyfikacja oraz aktywność proteolityczna i lipolityczna mikroorganizmów arktycznych. *Acta Scientiarum Polonorum: Biotechnologia* 2011: 10, 3-12
- Libkind, D., M. Gadanho, M. van Broock, and J. P. Sampaio. Studies on the heterogeneity of the carotenogenic yeast *Rhodotorula mucilaginosa* from Patagonia, Argentina. *J. Basic Microbiol.* 2008: 48, 93-98
- Liu Z, Chi Z, Wang L and Li J, Production, purification and characterization of an extracellular lipase from *Aureobasidium pullulans* HN2.3 with potential application for the hydrolysis of edible oils. *Biochem. Eng. J.* 2008: 40, 445-451
- Loureiro STA, Queiroz-Cavalcantiet MA, de Neves RP and Passavante JZO, Yeast isolated from sand and sea water in beaches of Olinda, Pernambuco State, Brazil. *Brazil J. Microbiol.* 2005: 36, 333-337
- Macrae AR and Hammond RC, Present and future applications of lipases. *Biotech Genet Eng. Rev.* 1985: 3, 193-219
- Mendes A and Castro A, Effect on the enzymatic hydrolysis of lipids from dairy wastewater by replacing gum arabic emulsifier for sodium chloride. *Braz. Arch. Biol. Tech.* 2005: 48, 135-142
- Methras NC, Bastawade KB and Gokhale DV, Purification and characterization of acidic lipase from *Aspergillus niger* NCIM 1207. *Biores. Technology.* 2009, 100, 1486-1490
- Mok WY, Luizao RCC, do Socorro Barreto M, da Silva M, Teixeira F and Muniz EG., Ecology of pathogenic yeast in Amazonian soil. *Appl. Environ. Microbiol.* 1984: 47, 390-394
- Muderhwa JM, Ratomahenina JR, Pina M, Graille J and Galzy P, Purification and properties of vthe lipases from *Rhodotorula pilimanae*. *Appl. Microbiol. Biotechnol* 1986: 23, 348-354
- Nagahama T, Hamamoto M, Nakase T, Takami H and Horikoshi K, Distribution and identification of red yeasts in deep-sea environments around the northwest Pacific Ocean. *Antone van Leeuwenhoek* 2001: 80, 101-110
- Nakamura S, Using sucrose esters as food emulsifiers. *Oleochemicals.* 1997: 8, 462-468
- Ota Y, Sawamoto T and Hasuo M. Tributyrin specifically induces a lipase with a preference for the sn-2 position of triglyceride *Geotrichum* sp. FO401B. *Biosci. Biotechnol. Biochem.* 2000: 64, 2497-2499
- Papaparaskevas P, Christakopoulos P, Kokos D and Macris BJ, Optimizing production of extracellular lipase from *Rhodotorula glutinis*. *Biotechnol. Lett.* 1992: 14, 397-402
- Rajendran A, Palanisamy A and Thangavelu V, Lipase catalyzed ester synthesis for food processing industries. *Braz. Arch. Biol. Tech.* 2009: 82, 207-219
- Reis P, Holmberg K, Watzke H, Leser ME and Miller R, Lipases at interfaces: a review. *Adv. Col. Inter. Sci.* 2009: 147-148, 237-250
- Sharma R, Chisti Y, Banerjee UCh, Production, purification, characterization and applications of lipases. *Biotech. Adv.* 2001: 19, 627-662
- Schmid RD and Verger R, Lipases: interfacial enzymes with attractive applications. *Angew. Chem. Int. Ed. Engl.* 1998: 37, 1608-1633
- Sugihara A, Shimada Y, Nakamura M, Nagao T and Tominaga Y, Positional and fatty acid specificities of *Geotrichum candidum* lipases. *Protein Eng.* 1994: 7, 585-588
- Tournas VH, Heeres J and Burgess L, Moulds and yeasts in fruit salads and fruit juices. *Food Microbiol.* 2006: 23, 684-688
- Yamazaki M and Komagata K, Taxonomic significance of electrophoretic comparison of enzymes in the genera *Rhodotorula* and *Rhodosporidium*. *Int. J. Sys. Bacteriol.* 1981: 31, 361-381
- Yu M, Lange S, Richter S, Tan T and Schmid R, High-level expression of extracellular lipase Lip2 from *Yarrowia lipolytica* in *Pichia pastoris* and its purification and characterization. *Prot. Exp. And Pur.* 2007: 52, 255-263
- Vancanneyt M, Coopman R, Tytgat R, Berny JF, Hennebert GL and Kersters K, A taxonomic study of the basidiomycetous yeast genera *Rhodosporidium* Banno and *Rhodotorula* Harrison based on whole-cell protein patterns, DNA base compositions and coenzyme Q types. *J. Gen. Appl. Microbiol.* 1992: 38, 363-377
- Weete JD, Microbial lipases. *Food Sci. Technol.* 1998: 88, 641-664
- Zimmer Ch, Platz T, Cadez N, Giffhorn F, Kohring GW, A cold active (2R,3R)-(-)-di-O-benzoyl-tartrate hydrolyzing esterase from *Rhodotorula mucilaginosa*. *Biotech Relevant Enzymes Prot.* 2006: 73, 132-140

DEPOPULATION OF TRADITIONAL MINING REGIONS IN CENTRAL AND EAST EUROPE: CASE STUDY OF THE UPPER- SILESIAN BASIN (POLAND) AND THE DONETSK BASIN (UKRAINE)

Marcin Rechlowicz, M.A, PhD Candidate

Maria Tkocz, PhD., Professor

Economic Geography Department, Faculty of Earth Sciences, University of Silesia

Abstract:

The article compares demographic processes taking place in traditional industrial European regions: Upper-Silesian and the Donetsk Basins, which are in their declining stage of development. During communism, a high population concentration was characteristic, currently, migration and population decline are leading to depopulation. Depopulation in the Upper-Silesian Basin is determined mainly by labour migration to the European Union (EU) and by sub-urbanization, whereas, in the Donetsk Basin, it is due to severe population decline intensified by migration. The main reasons are poor economies of the regions, connected with the collapse of numerous industrial facilities and the lack of alternative employment.

Key Words: Coal basin, restructuring of economy, depopulation, migration, population decline

Introduction

Political and economic transformations in Central and Eastern Europe have had a crucial influence on the demographic situation in this part of the European continent. According to the report entitled "The impact of European demographic trends on regional and urban development", issued within the framework of the Hungarian Presidency of the Council of the European Union (EU) (2011), population decline is mostly typical for the new EU member states and the eastern part of Germany. A similar situation is characteristic in the countries of the former Soviet Union. According to the authors, the main cause of population decline lies in political and economic transformations taking place in this part of Europe. The economy restructuring connected with the process of its adjustment to the needs of the local as well as the global markets, forced the liquidation of obsolete facilities, which lead to the reduction of employment and to unemployment, which had been unknown in this part of Europe before 1989. The adaptation to new market conditions was exceptionally difficult in old industrial regions of mono-functional structure of production based on traditional branches of heavy industry. The process of restructuring in such regions, as proved by the experience of Western European countries, is extremely difficult and long-lasting (Hamilton, 1984; Riley, Tkocz, 1998). It is most frequently connected with a total or partial liquidation of the worked-out open cuts, technologically obsolete steelworks and processing plants connected with them. In their place, new industries are introduced and services are developed (Rachwal, 2011; Tkocz, 2006). These, however, are not able to compensate for the loss of jobs in traditional industries, hence, the vast unemployment and emigration of population. The Silesian Voivodeship constitutes an example of such a region in Poland, with the Upper-Silesian Basin in its central part, while in Ukraine, these are the Donetsk Oblast and the Luhansk Oblast, within which the Donetsk Basin is located. Both regions in the times of the communist economy were characterized by intense growth of industry, mining in particular, and mass inflow of workforce (Alforov, 2010, Runge, 2008, Sarzan, 2000, Tkocz, 2001). During the period of economy transformation in Poland, after 1989, and after gaining independence by Ukraine in 1991, in both regions the process of industry restructuring was initiated. It resulted in the liquidation of factories, which did not meet the new market requirements and which applied obsolete and environmentally unfriendly technologies and employed an excessive number of workers (Matykowski, Strykiewicz, 1992; Rechlowicz, Tkocz, 2011). These regions found themselves in the

final stage of their life (Steiner, 1985) and their further development relied on the introduction of new industries and innovative activities. Both regions rapidly faced the crisis phenomena of unemployment, an income decrease in households and a dramatic reduction in the standard of living (Deacom, 2000). The consequence of these phenomena is the migration of population, leading to a steady population decline and an alarming phenomenon of urban shrinking reflected in depopulation of centres, outflow of specialized staff and young people (Krzysztofik, et al., 2011; Mykhnenko, et al., 2010). Even though the two regions were in the past characterized by similar development determinants, in new economic conditions their situations are diverse. This is connected with the fact that Poland has been a member of the European Union since 2004 and its economy had to be adjusted to EU requirements, especially in the limitation of coal mining. The Ukrainian economy is not subject to these restrictions, but Ukraine endeavours to access the EU.

The article aims at a comparison of demographic transformations which took place in the last twenty years in the area of the Silesian Voivodeship and the Donetsk and Luhansk Oblasts in Ukraine, that is in areas intensely developed with vast population inflow, which constituted industrial cores in both countries. An attempt to explain the causes of the observed phenomena was made, including local specifics, in particular. Moreover, contemporary demographic processes were analysed thoroughly and they, to a large extent, reflect the period of the second demographic transition previously observed in Western Europe, which is characterized by a change in the family model, postponed decisions to marry and a decline in reproduction. In order to compare the demographic transformations in various regions, traditional cartographic methods were applied as well as the Webb's typology (1963).

Demographic Changes In The Researched Regions

Ever since the beginning of socio-economic changes in Central and Eastern Europe, i.e. from 1988/89 to 2010/11¹ it may be stated that in most countries of this part of Europe, population was characterized by a downward trend (The impact 2011, fig.3, p.22). High intensity of the depopulation phenomenon is also typical for majority of former Soviet republics, including Belarus (-6,4%) and Ukraine (-10,7%) (fig.1) and is perceived as a consequence of transformations which took place following the collapse of communism (Fassmann, 1994). Poland was one of the few countries where population showed a minimal increase (0,8%). At the same time, the population of the Silesian Voivodeship declined by nearly 300 thousand people (from 4.9 million to 4.6 million - that is by 6.1%), in the Donetsk Oblast by 900 thousand (from 5.3 million to 4.4 million - that is by 17.0 %), in the Luhansk Oblast by almost 600 thousand (from 2.9 million to 2.3 million - that is by 20.7%). The Silesian Voivodeship is the second area in Poland, following the Lodz Voivodeship, as far as the intensity of the depopulation phenomenon is concerned. In Ukraine, apart from the above mentioned oblasts, a high index of population decline is also present in the oblasts of Chernihiv (-21.5%) Sumy (-17.9%), Kirovohrad (-17.1%), and Zhytomyr (-17.1%). Depopulation in the analysed regions is much more severe than in other regions of Poland and Ukraine.

An intensive population inflow to the Upper-Silesian and the Donetsk Basins during communist times was accompanied by the development of urbanization processes, therefore, the researched basins in Poland and Ukraine are characterized by a high index of urbanization.

The percentage of urban population in the Donetsk Oblast amounted to 90.5%, in Luhansk 86,7%, and in the Silesian Voivodeship 78,0% (fig.2). Such high indexes result from the development of cities and urban agglomerations where industry constituted the economic basis. Apart from high urbanization and industrialization, these regions are distinguished by a high degree of population concentration. It is mostly visible in the case of the Donetsk and the Luhansk Oblasts, where over 75% of population live on 10% of the regions' areas, whereas in the case of the Silesian Voivodeship it is only half of its population. The high concentration of population in the area of the Donetsk Basin was influenced by various factors, the most important being:

- prevailing significance of heavy industry, especially coal mining, in the economic structure of the Donetsk Basin, and thereby constant attractiveness of large industrial cities as places of residence,

¹ Polish statistical offices publish data as of the last day of the year (31 December) whereas Ukrainian offices - as of the first day of the year (1 January). Therefore, the article adopted double dates, for example, data for the year 2010/11 refers to population in Poland on 31 December 2010 and 1 January 2011 in Ukraine.

- a much stronger influence of political and economic transformations on the collectivized farming in the times of the Soviet Union and in effect, depopulation of villages and towns of the Donetsk Basin as areas with no perspectives from an economic development point of view, and also deterioration of living conditions for inhabitants of villages and towns (hampered access to trade and services, including medical services),
- poverty preventing mobility, which is the reason why a large group of inhabitants are doomed to live in houses and apartments they received during the times of the Soviet Union, even in situation when there are no perspectives to find jobs in their home towns,
- weaker suburbanization, in comparison to the Silesian Voivodeship, which is an effect of the lack of attractive locations for residence outside urban agglomerations.

The high concentration of population in the Silesian Voivodeship was influenced by strong industrialization of its central part, which lasted continuously since 1740, when the first coal mine was opened. A consequence of it was the development of large and medium-sized cities comprising the Katowice conurbation, the largest metropolitan type area in Poland.

Several factors had influence on the changes in population during the times of the economic transformation of both countries, that is from the beginning of the 1990s. Firstly, social and economic transformations, connected with the transition from the centralized model of the country and the command-and-quota economy to capitalism and, above all, de-industrialization. Participation of industry in the formation of Ukraine's national income in the period of 1991-2009 decreased from 45.7% to 30.0% (Mykhnenko, 2010), in the Silesian Voivodeship from 37% to 29%. The impact of this factor was especially strong in urban centres dominated by large industrial plants. The largest population decline (by over one quarter) was observed in those cities where traditional branches of industry serving as the source of income for its inhabitants, were liquidated or significantly limited their employment. This refers to such centres as: Stakhanov, Kirovsk, Krasnyi Luch and Pervomaik in the Luhansk Oblast and Shakhtarsk, Yenakieve, Krasnyi Lyman, Torez, Kostiantynivka and Snizhne in the Donetsk Oblast, where the decline amounted to 25-33%. These cities are located along the border between the two oblasts and create one of the largest depopulation areas in Ukraine (Gentil, Marcinczak, 2012). Apart from the above mentioned centres, it is co-formed by five land units in the Luhansk Oblast, i.e. the raions of Popasna, Perevalsk, Kreminna, Antratsyt, and Sverdlovsk with equally high population decline (Mykhnenko, et al., 2010).

The outlook is slightly different in the case of the Silesian Voivodeship, where the change of residency and urban centres depopulation were equally influenced by the restructuring of traditional branches of industry as well as migration connected with the return of population to their home towns and the economic migration, initially mainly to Great Britain and Ireland, but after access to the EU, also to other countries which progressively opened their job markets. Nevertheless, the top of the list are the cities where at least several large industrial plants were liquidated, mainly coal mines and old steelworks, i.e. Katowice, Chorzow, Ruda Slaska, Piekary Slaskie and Bytom. Population decline is approximately half, in comparison to Ukraine and amounts to about 15-17%.

On the other hand, there are centres where industry restructuring happened only on a small scale and until present, a majority of industrial plants operating in the socialist times are still open; they are not public property any longer though, they have been privatized. A flagship example of such cities are large centres of metallurgy and steel working, such as Alchevsk (-10.0 %) in the Luhansk Oblast, Mariupol (-9.7 %) in the Donetsk Oblast and Dabrowa Gornicza (-6.7 %) in the Silesian Voivodeship, where the largest steelworks in Poland, constructed in the 1970s, was sold to a world leader in steel production - Lakshmi Mittal (presently Arcelor Mittal).

A relatively low population decline characterizes regional and sub-regional centres, such as Bielsko-Biala (-1.9 %) or Luhansk (-12.5 %).

Another type of cities consists of centres where the liquidation of traditional branches of industry was accompanied by the establishment of modern technology plants, which was conducive for the reduction of population outflow. Gliwice (-8.5 %), in the Silesian Voivodeship, constitutes the best example of such a centre, where the liquidation of a mine and an ironworks coincided with the opening of a General Motors factory located within the Katowice Special Economic Zone. Unfortunately, there is a lack of such examples in the Donetsk Basin, due to numerous barriers and continued vast risk for foreign investors who would want to locate new factories in Ukraine, but also with delay in implementing economy boost solutions, such as economic zones or preferential

treatment for investors. The economic situation connected with the high indebtedness of the country is also an unfavourable factor, mainly for the International Monetary Fund (IMF) and Russian Gazprom, which delivers gas to Ukraine.

Characteristics Of Demographic Transformations In 2010

The information presented above lead to a conclusion that considerable depopulation, resulting from various reasons, is present in all researched regions. To find out the reasons, the classic method of J.W. Webb (1963) was applied. Tab.1 presents demographic changes in 2010, with the division not only on downstream administrative units, but also the types of these units. Municipal units have been identified - in the case of Poland, these are township districts, and in the case of Ukraine - mis'krady, and land units - respectively - country districts and raions. The feature that diversifies the researched areas the most is undoubtedly natural migration of population. In respect to it, the the most advantageous is the situation in the Silesian Voivodeship, where natural population increase was observed in a large part of the administrative units, whereas in the Donetsk and the Luhansk Oblasts, all the researched administrative units experienced natural population decline. The dominance of depopulation types E and F clearly points this out. According to Mykhnenko (2010), the birth rate in 2002 decreased to one of the lowest in the world, amounting to 0.9 live births per 1000 women. In both oblasts, natural decline is additionally deepened by migration losses (F). The situation in the Silesian Voivodeship is different; here, the largest group of units is characterized by the predominance of migration outflow over the natural decline (G). The presence of areas with migration inflow and population growth is favourable (B,C). In contrast to the Ukrainian oblasts, all types distinguished by Webb (tab.1) are represented in the Silesian Voivodeship, which indicates that demographic processes are more diversified in this region, in comparison to Ukraine.

Table 1 Demographic transformations in the analysed area in 2010, according to Webb's types

Type of transformation	The Silesian Voivodeship		The Donetsk Oblast*		The Luhansk Oblast	
	No of units	Population [thousands]	No of units	Population [thousands]	No of units	Population [thousands]
A	1	141,4	0	-	0	-
B	4	395,3	0	-	0	-
C	8	989,3	0	-	0	-
D	2	244,9	0	-	0	-
E	2	222,4	18	1 337,6	12	605,6
F	2	330,3	24	3 013,4	19	1 637,4
G	11	1 591,7	2	35,6	1	48,3
H	6	720,6	0	-	0	-
Total	36	4 635,9	44	4 386,6	32	2 291,3

* - excluding Krasnyi Lyman and the Krasnozymanskyi raion

Source: own study based on Central Statistical Office 2012, Central Statistical Office of Donetsk Oblast 2012, Central Statistical Office of Luhansk Oblast 2012

Spatial differentiation of the birth rate index in 2010 was presented in figure 3. The first point to mention is that the general value of the birth rate for the Silesian Voivodeship equalled 10.3 of live births per one thousand inhabitants, while in the case of the Donetsk and the Luhansk Oblasts, it was 9,3‰ and 9,1‰ respectively.

Similarly to the change in population in the period of 1988/89 – 2010/11, the differentiation of the birth rate also shows spatial regularity, especially in the Silesian Voivodeship. The lowest, disadvantageous values were observed in districts located in the northern part of the region, whereas the highest values of birth rate were in the southern part. In the Donetsk and the Luhansk Oblasts, a high concentration of low values of the birth rate may be observed in the southern part of the border areas in both oblasts.

A big problem for traditional mining regions in Central and Eastern Europe is a high number of deaths and shorter average lifespan, in comparison to other regions. It is especially apparent in the Donetsk Basin, where factors such as high pollution of natural environment by obsolete industrial plants, insufficient medical services or social problems, mainly alcoholism, are superimposed (Flaga, 2006).

Table 2 presents the number of administrative units in specific value groups for death index, calculated as the number of deaths per one thousand inhabitants. A big discrepancy between the

situation in the Silesian Voivodeship and the Donetsk and the Luhansk Oblasts is noticeable, making it basically impossible to compare the phenomenon applying one scale of values in the arrangement of all analysed regions.

Table 2 The value of the death indicator in the analysed regions in 2010

Region	The value of the death indicator [number of deaths / 1000 inhabitants] in the lower level units [number of units]							
	<10	10-11	11-12	12-13	13-14	14-15	15-20	>20
The Silesian Voivodeship	19	5	7	5	0	0	0	0
The Donetsk Oblast	0	0	1	1	0	2	34	6
The Luhansk Oblast	0	0	0	0	1	1	25	5
Total	19	5	8	6	1	3	59	11

Source: own study based on Central Statistical Office 2012, Central Statistical Office of Donetsk Oblast 2012, Central Statistical Office of Luhansk Oblast 2012.

In eleven units, the number of deaths exceeds 20‰, which is rare in Europe. This refers only to the Donetsk Basin, and mainly to areas with a low level of urbanization.

A difficult position of the Ukrainian society, resulting from a profound and long-lasting economic crisis, is reflected in certain behaviour, which might lead to an increase in death rate. Lack of employment and, as a consequence, lack of means of support lead to suicide attempts. In the beginning of 1990, the suicide rate was 20.6 per 100 thousand inhabitants, reaching its highest level in 1996 with the value of 29.9. In the following period of 5 years (until 2000), the value of this index remained high. In 2010, it was 19.9, in the Donetsk Oblast - 22.8, and in the Luhansk Oblast - 25.8. Such high suicide rates are unheard of in Poland (16,9 in 2009) or other countries of the EU. Another alarming phenomenon, which comes as a consequence of not being able to cope with difficulties, is alcohol abuse. The dimension of it is reflected in the number of road accidents caused by drunkenness. In the beginning of 1990, this index was 10.4 per 100 thousand inhabitants, reaching its peak in the periods of 1995-1997 and 2002-2005, when it oscillated at the level of 19-22 deaths per 100 thousand. In 2010, it amounted to 9.8 in Ukraine, 12.9 in the Donetsk Oblast and 13.7 in the Luhansk Oblast (Rettinger, 2011, p.57).

The Webb's typology (tab.1) shows that migrations have a lesser influence on the shaping of demographic changes in the researched regions. Only in the case of the Silesian Voivodeship the dependence between the type of administrative unit and the migration balance is clearly visible. In 2010, all township districts observed negative migration balance values (from -0,6 in Chorzow to -8,7 in Jastrzebie Zdroj), while in country districts, it was only in Zawiercie (-0,1). All other country districts had positive migration balances (up to +6,3 in the Bielsko-Biala district). The Donetsk and the Luhansk Oblasts do not reveal such a dependence.

Nevertheless, the analysis of spacial differentiation of the migration balance index in 2010 allows an attempt to create an outline of potential migration trends, which, to a large extent, occur concurrently with sub-urbanization. It is mostly visible in large cities (or conurbations) neighbouring rural areas (fig.4)

This phenomenon is less apparent in the Donetsk Basin, but even here, areas where large cities with negative migration balance neighbour rural areas of positive balance were observed in 2010 (fig.4).

In this respect, the reverse phenomenon is noteworthy. In several cases in the Donetsk Oblast, cities with positive migration balance are adjacent to rural areas with negative balance. It is hard to draw simple conclusions with relation to it due to the fact that migrations happened at larger distances, including inter-regional and international ones.

Conclusions

Traditional industrial regions in Poland and Ukraine, whose genesis is connected with coal mining and the related intense development of other industry branches, especially of low level processing industries, are presently in their decadent stage of development. In the past, they were characterized by an excessive concentration of population, particularly in cities, which is still sustained, hence the contemporary population decline may be treated as positive for the regions. The main reason for depopulation is emigration and natural outflow. The first factor determines the process of population decline in the Upper-Silesian Basin, whereas the second one in the Donetsk

Basin. Heavy population outflow in the Upper-Silesian Basin is linked to employment perspectives beyond Polish borders, due to its membership in the EU since 2004 and the intensifying suburbanization process, reflected by migration and population growth in the southern part of the Silesian Voivodeship. Population decline in the Donetsk and the Luhansk Oblasts is considerably higher and more alarming due to the fact that it is determined by a deep natural outflow intensified by migration. The main reason is the poor economic state of regions connected with the collapse of many industrial plants and also the lack of alternative work places. Still, the continuity and stability in the functioning of industrial plants representing traditional branches of industry has a positive impact on the demographic state, keeping the population of some centres at a relatively constant level. In conclusion, it may be added that demographic processes observed in highly industrialized regions, to a large extent reflect the processes observed previously in highly industrialized European countries.

References:

- Alforov, M. A., Migratsyi naselennya do mist Donbasu u peryod povoyennoi vidbudovy, *Novi storinky istoriyi Donbasu. Zbirnyk statyeyi*, 19, 192-201, 2010.
- Central Statistical Office, 2012, Internet resources, www.stat.gov.pl (accessed: 04.07.2012).
- Central Statistical Office of Donetsk Oblast, 2012, Internet resources, www.donetskstat.gov.ua (accessed: 04.07.2012).
- Central Statistical Office of Luhansk Oblast, 2012, Internet resources, www.lugastat.lg.ua (accessed: 04.07.2012).
- Deacon, Bob, Eastern European Welfare States: the Impact of the Politics of Globalisation, *Journal of European Social Policy* 10, 2 :146-161, 2000.
- Fassmann, Heinz, Transformation in Ostmitteleuropa. Eine Zwischen Nilanz, *Geographische Rundschau*, 12:685-691, 1994.
- Flaga, Małgorzata, Procesy demograficzne oraz ich uwarunkowania w zachodnich obwodach Ukrainy w okresie przemian ustrojowych, (Demographic Processes and their Conditionings in Western Regions of Ukraine in the Period of Political System Transformations), Wyd. UMCS, Lublin, 2006.
- Gentile, Michael and Szymon Marcińczak, No More Work For Stakhanow: Migrants and Stayers in the Depopulating Donbas, Ukraine, *Urban Geography* vol.3:3- 15, 2012.
- Hamilton, FElan, "Industrial Restructuring: an International Problem, *Geoforum* 15,3: 28-39, 1984.
- Krzysztofik, Robert, Jerzy Runge and Iwona Kantor-Pietraga, Paths of Shrinkage in the Katowice Conurbation. Case Studies of Bytom and Sosnowiec Cities, *Wydział Nauk o Ziemi, Uniwersytet Śląski, Sosnowiec*, 2011.
- Matykowski, Roman and Tadeusz Stryjakiewicz , The Adaptation of Industry and its Labour Force to the Changing Economic System in Poland, *Erdunke* 46:149-157, 1992.
- Mykhnenko, Vlad, Dmytro Myedvyedyev and Larysa Kuzmenko, Urban Shrinkage in Donetsk and Makiiivka the Donetsk Conurbation, Ukraine, The University of Nottingham School of Geography, Nottingham, 2010.
- Rachwał, Tomasz, Transformations of the Employment Structure as an Expression of the Transformation of Polish Industry against a Background of the European Union, *Bulletin of Geography. Socio-economic series*, 15: 521-529, 2011.
- Rechłowicz, Marcin and Marcia Tkocz, Directions in Development of the Old Industrial Region in the 21st Century. The Example of the Upper Silesia Industrial District, *Charkowski Nacjonalnyj Uniwersytet, Charków*, 192-195, 2011.
- Rettinger, Renata, "The Impact of the Economic Crisis on the Demographic Situation in Ukraine" in Ziolo Z., Rachwał T., ed., *Changes in Local and Regional Structures of the Service Sector During Economic Crisis*, Commission of Geography of Industry PTG, 18: 51-58, 2011.
- Riley, Ray and Maria Tkocz, Coal Mining in Upper Silesia under Communism and Capitalism, *European Urban and Regional Studies*, 5, 3:217-235, 1998.
- Runge, Jerzy, Population Transformations in Traditional Economic Regions of Central Europe. Structural Approach, *Bulletin of Geography. Socio-economic series*, 10: 63-74, 2008.
- Sarzan, A. O., *Sotsjalno-ekonomitshni ta polititshni protsesi v Donbasi (1945-1998)*, Stalker, Donetsk, 2000.
- Steiner, Michael, Old industrial Areas: A theoretical approach, *Urban Studies*, 5:387-398, 1985.

The Impact of European Demographic Trends on Regional and Urban Development” Synthesis report, Hungarian Presidency of the Council of the European Union, Budapest, April 2011.

Tkocz, Maria., *Restrukturyzacja przemysłu regionu tradycyjnego (Restructuring of Industry in a Traditional Region)*, Wydawnictwo Uniwersytetu Śląskiego, Katowice, 2001.

Van Zon H., *The Future of Industry in Central and Eastern Europe*, Avebury, Aldershot, 1996.

Webb J.W., *The Natural and Migrational Components of Population Changes in England and Wales 1921-1931*, *Economic Geography*, 39,2, 1963.

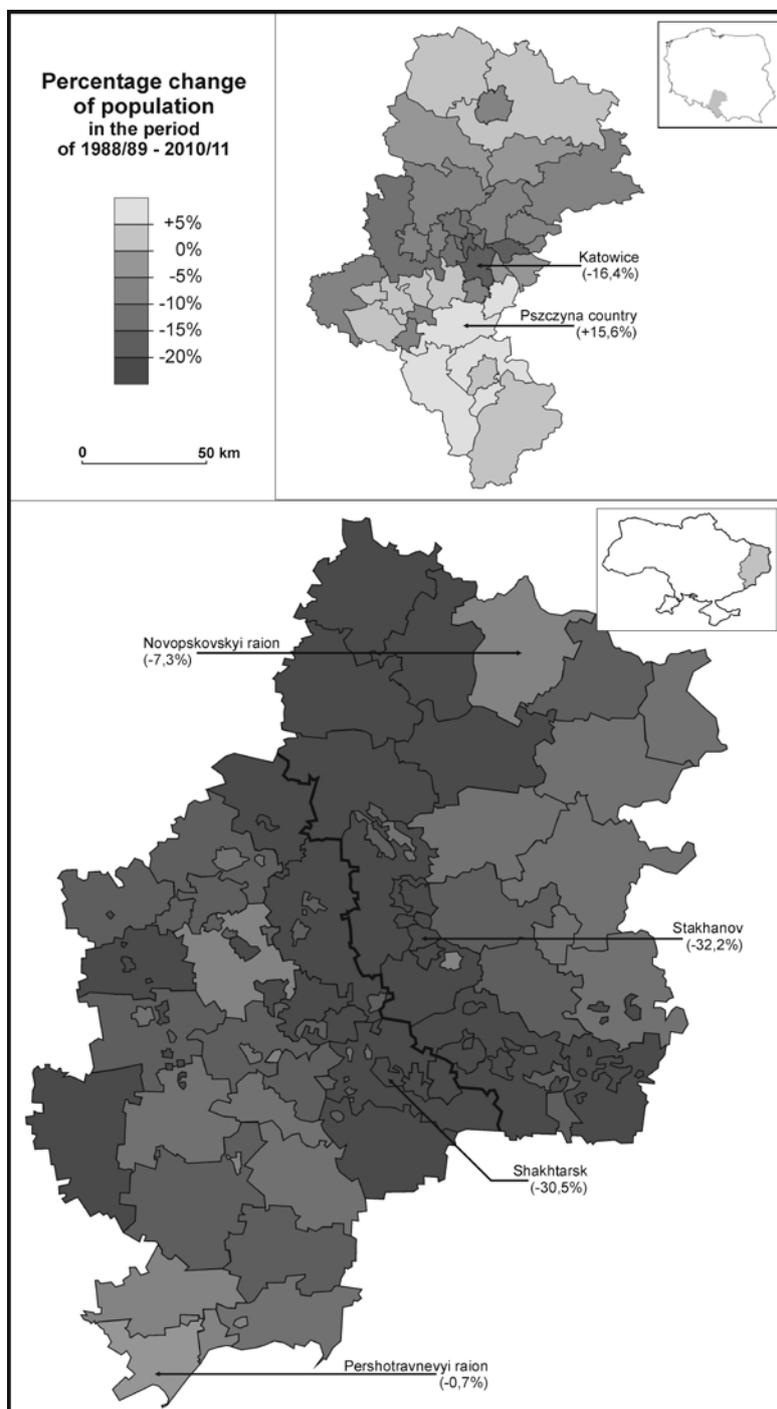


Fig. 1 Population changes in the examined regions in the period of 1988/89 – 2010/11 (own study).

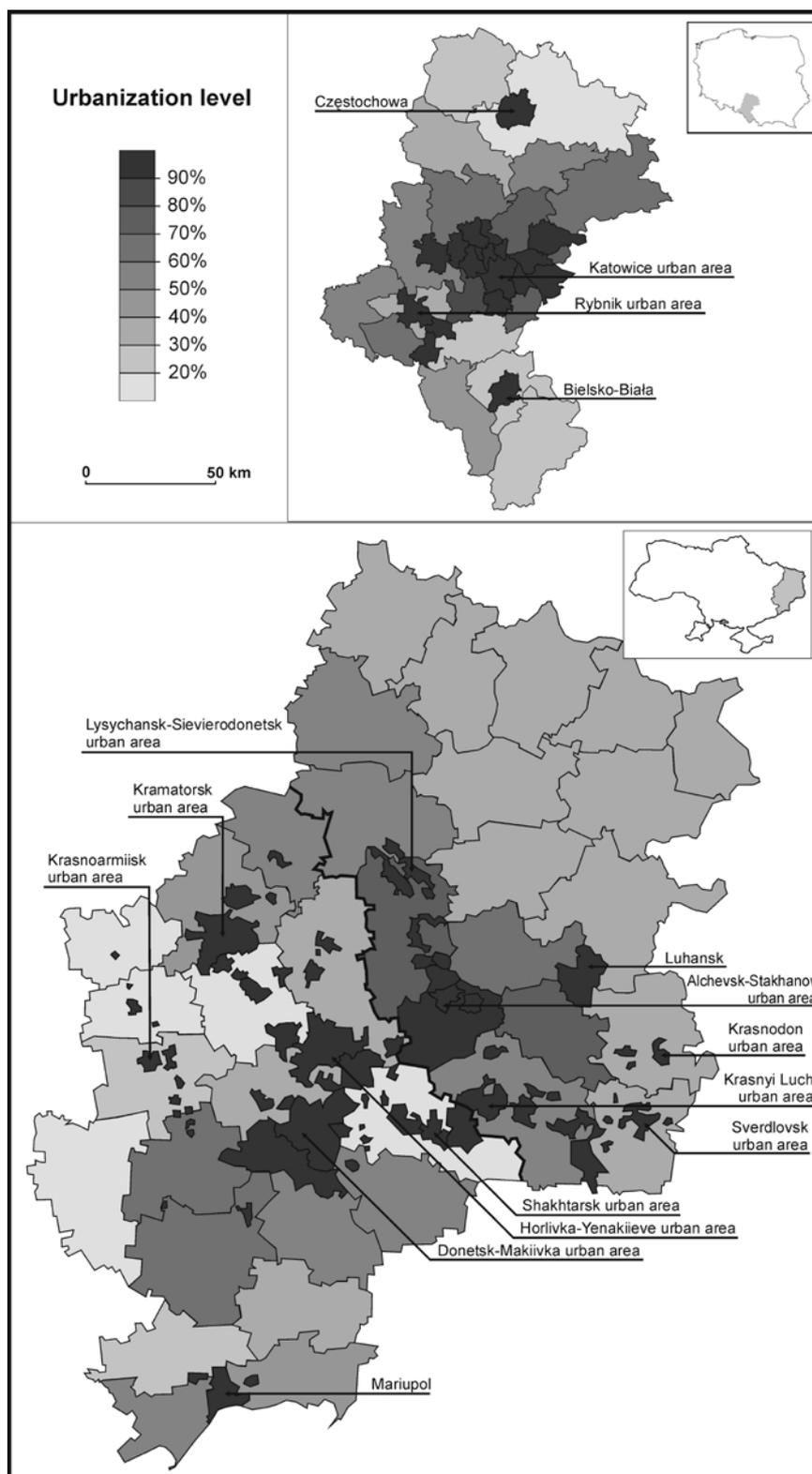


Fig. 2 The spatial diversification of the urbanization factor in 2010/11

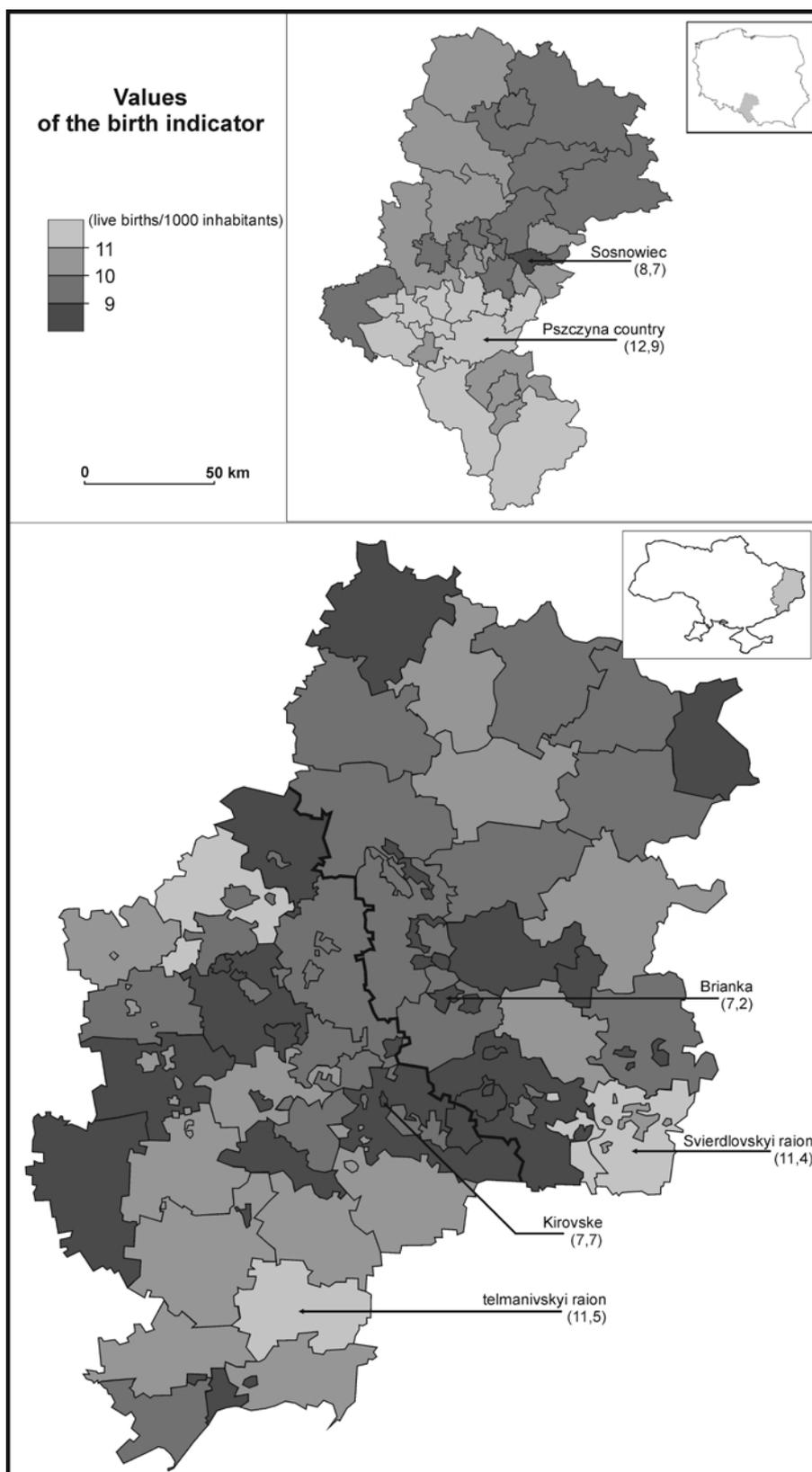


Fig. 3 The spatial diversification of the birth indicator in 2010 (own study).



Fig. 4. Spatial diversification of the migration balance indicator in 2010 (own study).

THE SOUTHERN TELLIAN SERIES AT NUMMULITES (NE ALGERIA): STRATIGRAPHY AND STRUCTURAL CONSEQUENCES

Mohamed Chadi

ENS Constantine, Algeria

El hadj Youcef Brahim

Rami Djeflal

Department of earth and universe sciences Batna University, Algeria

Abstract:

The synthesis of stratigraphic and structural data of Constantinois limestone massifs has allowed the refining of the southern Tellian series at nummulites stratigraphy and proposing a structural model.

The most significant sites and less disturbed by tectonics, are the Taxas syncline and southern flank of Guerioum massif, exactly at the Djebel Ras Rihane.

Also, at the level of Taxas syncline, and surmounting the hard ground which ended the Aptian neritic limestone, the marly levels had provided Cenomanian microfauna: *Favusella washitensis*, *Rotalipora appenninica*, *R. cushmani*, *R. brotzeni*, *Hedbergella* sp., and *Praeglobotruncana stephani*.

On the southern reverse of Djebel Ras Rihane, at the level of Chaabet Ras Chiboub notch, we can observe, on the hard ground that terminates the Aptian neritic limestones, a clayey Cenomanian over one hundred meters of thickness.

The samples from these clays have provided many *Hedbergella*, *Rotalipora brotzeni*, *R. cushmani*, *R. globotruncanoides* and *Praeglobotruncana* gr. *Stephani*.

The top of these clays has provided Coniacian foraminifera and revamped Cenomanian *rotalipora*.

The study of the stratigraphic series of southern neritic massifs allowed the refining of our predecessor's results.

Also the yellow marls stratigraphically surmounting the terminal Aptian- basal Albian through a hard ground, represents the Cenomanian-Turonian.

Microscopic analysis of samples from this hard ground highlights sedimentological phenomena that attest the emersion of Constantinois platform during the terminal Aptian- basal Albian.

In the Constantinois limestone massifs, "the southern Tellian units at nummulites " represent the normal marly cover of neritic limestones.

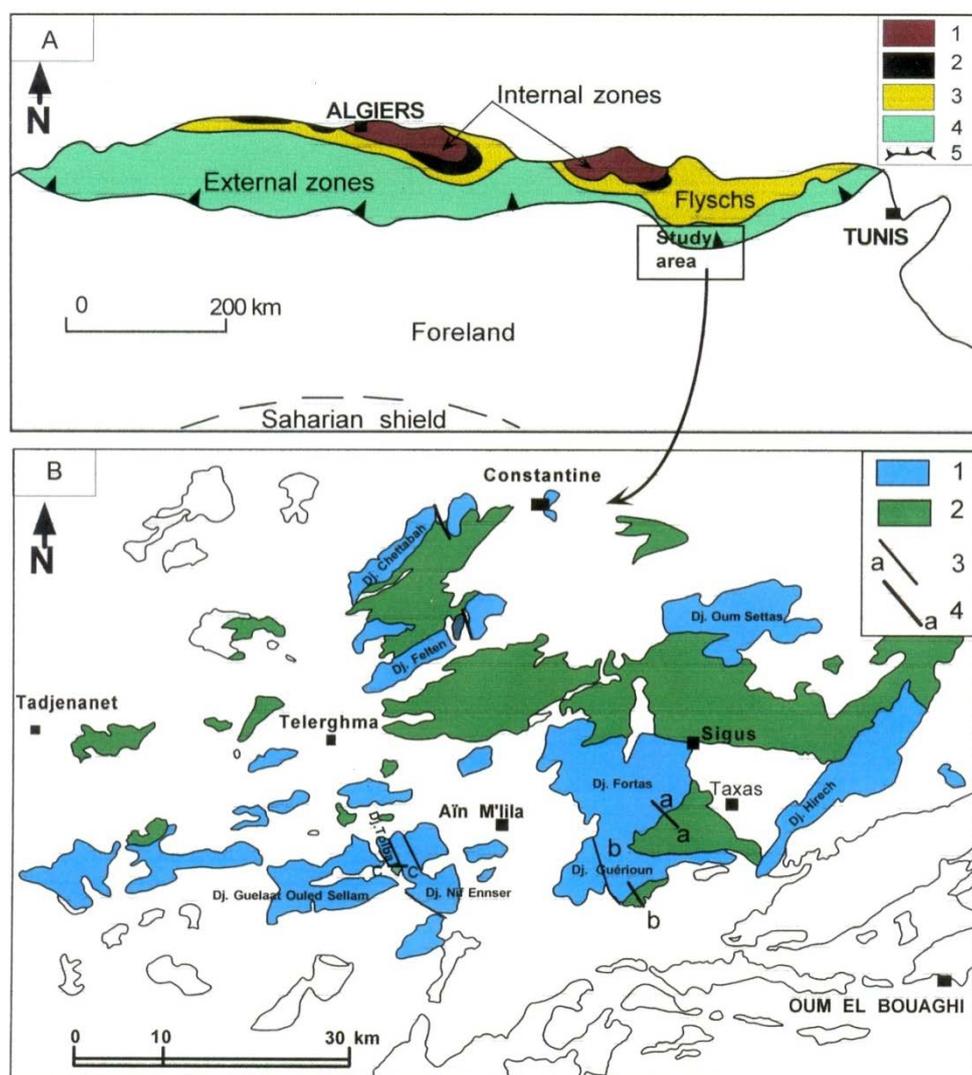
Key Words: Constantinois limestone massifs, Albo-Aptian, southern Tellian units, eastern Algeria

Introduction

The southern units at nummulites are located in the external zones of the Eastern Algeria Alpine Belt (Fig.1: A & B).

Located only in the southern confines of the Constantinois neritic mole, the stratigraphy of these units contains the Senonian and Eocene formations rich with nummulites and represents Eocene coverage of the Peni-Tellian series, thrust on the Constantinois neritic platform Vila (1980).

The objectives of this paper are: the study of the stratigraphic Tellian series at nummulites and to establish structural relationships with their substratum.



A :Location of the study area in the frame of the Alpine belt

1-Socle Kabyle , 2-Dorsale Kabyle, 3-Flysch domain, 4-External zones, 5-Limit of Alpine belt

B : Location of Southern tellian series at Nummilites 1-Carbonate platform, 2- Southern tellian series at Nummilites, 3-Fault,4-Cross sections :

a---a :Cross section of Taxas syncline

b---b:Cross section of the southern reverse of Dj.Ras Rihane

c---c :Cross section of the valley between Dj.Tolba and Dj. Guelaat Ouled Hadj

Stratigraphic study

Cross section of Taxas syncline

Over the hard ground, and when conditions of outcrops allow it, we can observe (Fig. 2):

- Gray marls alternating with some levels of micritic limestones. The samples have delivered Cenomanian microfauna with: *Favusella washitensis*, *Rotalipora appenninica*, *R. cushmani*, *R. brotzeni*, *Hedbergella* sp., and *Praeglobotruncana stephani*.

- These marls are overlain by marly-limestones and marls which have delivered rich Santonian microfauna with: *Globotruncana linneiana*, *G. sigali*, and *G. fornicata*. At the base of these marls, in the Santonian microfauna there are revamped Cenomanien *Rotalipora*.

- Gray marls, often yellowish by alteration, the samples have delivered Campanian microfauna with: *Globotruncana linneiana*, *G. fornicata*, *G. elevata* and *G. stuatiformis*.

- Bar of bedded marly-limestones, very rich in inoceramids reported by Voûte (1967) to the Campanian

- Black marls, the base still Campanian, but the top reaches the Maastrichtian with *Globotruncana stuatii* and *G. contusa*.

- Marly-limestones, well bedded, forming a marked bar by topography. This bar is very rich in inoceramids; it was reported by Voûte (1967) to the Maastrichtian.

- Black Maastrichtian marls at the base and Paleocene at the top, the pelagic microfauna is quite rare, but benthic foraminifera are abundant.

- Limestones and marly-limestone beds 0,2 to 0,5 m, light in patina, black on the break. The flint nodules are common especially in the lower part. These limestones are rich in nummulites (Late Ypresian).

- A thick series of marls with intercalations of lumachellic oyster beds. The samples have delivered rich Lutetian microfauna. Vila (1980) quotes: *Globigerina soldadoensis*, *G. linaperta*, *G. yeguaensis*, *Globorotalia boweri*, *G. topilensis*.

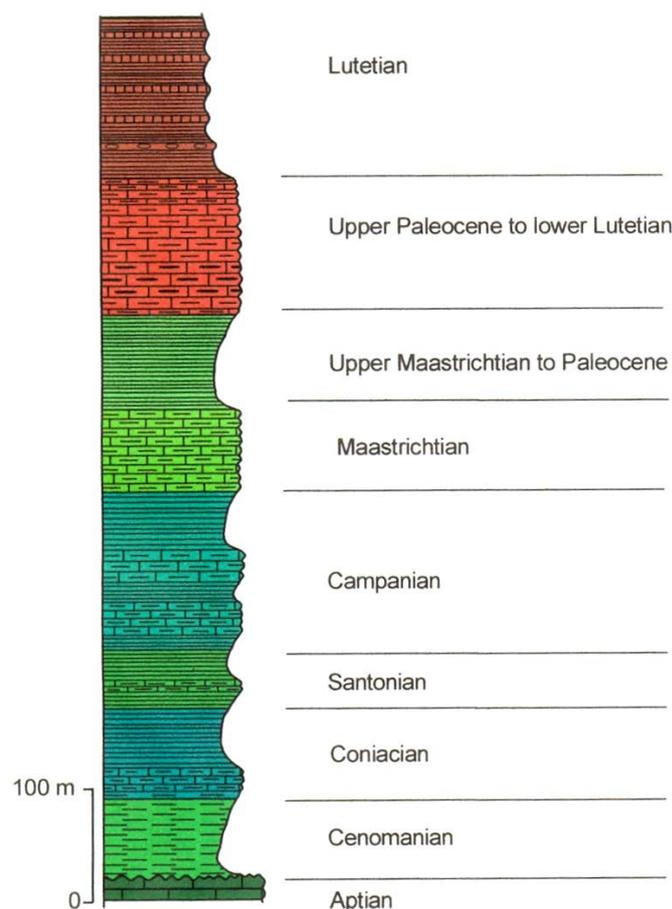


Fig. 2: Lithostratigraphic column of Texas syncline

Cross section of the southern reverse of Dj.Ras Rihane:

On the southern reverse of Djebel Ras Rihane, at the level of Chaabet Ras Chiboub notch, we can observe, on the hard ground that terminates the Aptian neritic limestones (Fig. 3):

-A clayey Cenomanian over one hundred meters of thickness, the samples from these clays have provided many *Hedbergella*, *Rotalipora brotzeni*, *R. cushmani*, *R. globotruncanoides* and *Praeglobotruncana* gr. *Stephani*; the terminal Cenomanian at *Rotalipora reicheli* is missing.

- The top of these clays has provided Coniacian foraminifera and revamped Cenomanian *Rotalipora*.

- One hundred meters of yellowish beige marls alternating with decimetric marly-limestone beds which have provided Coniacian microfauna at the base and the Santonian at the top.

-Above, come marls topped by a marly-limestone bar and marls again containing Campanian microfauna.

- Marls and marly limestone bar topped by black Maastrichtian marls.

-Black Paleocene marls.

-Phosphatic limestones at nummulites with flint, indicating the Ypresian.

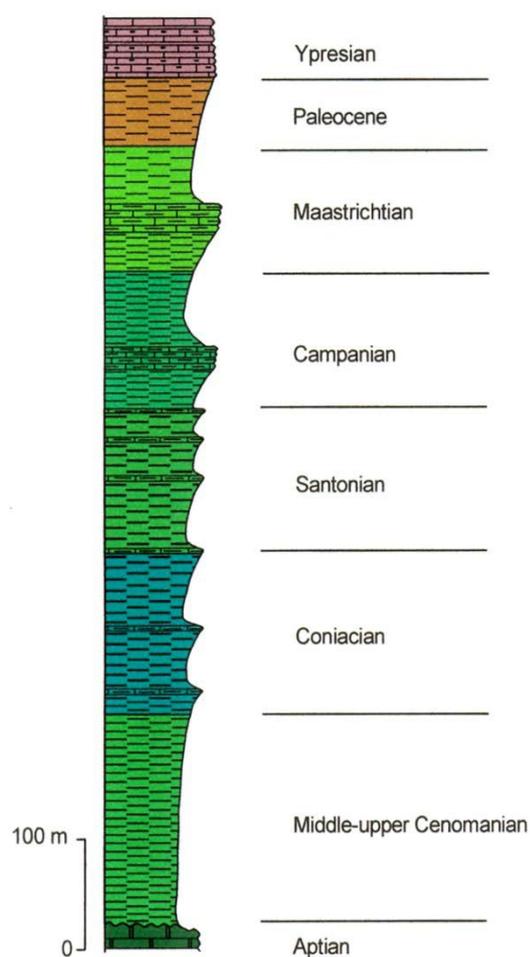


Fig.3: Lithostratigraphic column of Dj.Ras Rihane southern flank

The valley between Djebel Tolba and Djebel Guelaat Ouled Hadj

Over the gray lightened limestones at Miliolids of the Terminal Aptian which are ended by a hard ground (Fig. 4), we observe a nodular micritic limestone at algae and *Calcisphaerulidae*

This hard ground is surmounted by micritic limestones with conglomeratic appearance and marls where urchins were recognized, gender of *Hemiaster batnensis*.

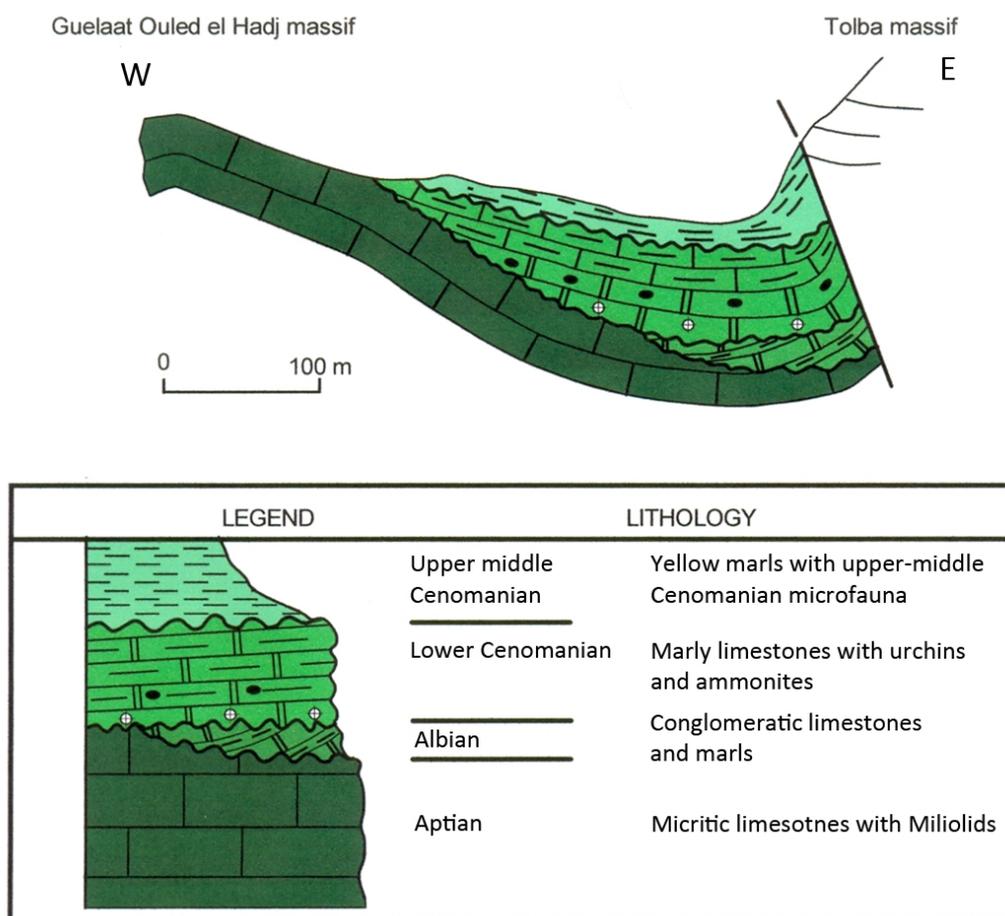


Fig.4: Geologic cross-section revealing the existence of neritic lower Cenomanian

- Above, ten meters of phosphatic limestones at Ammonites intercalating with conglomeratic levels and ending with a hard ground.

- Yellow marls finishing this series, the microfauna from these marls indicate the upper-middle Cenomanian.

To our knowledge this valley is the only place in the southern neritic massifs where Ammonites indicating the Cenomanian were collected.

These Ammonites were dated by C.Meister (Natural History Museum of Geneva). It is *Mantelliceras saxbii*. Sharpe (1856).

For Reymont Richard & Bengston (1986), it is the lower Cenomanian of the zone at Mantelli.

For Kennedy, Juignet & Wright (1986), it is the lower Cenomanian and more precisely it is the middle part of the lower Cenomanian.

In the same place, Bär (1957) noted the presence of pyritic indeterminate Ammonites and urchins including *Hemiaster batnensis* Coquand, well known in the Aurès Cenomanian.

The hard ground on the top of the lower Cretaceous

When conditions of outcrops allow it, we can observe this hard ground in all massifs of Constantinois Mountains (photo 1).

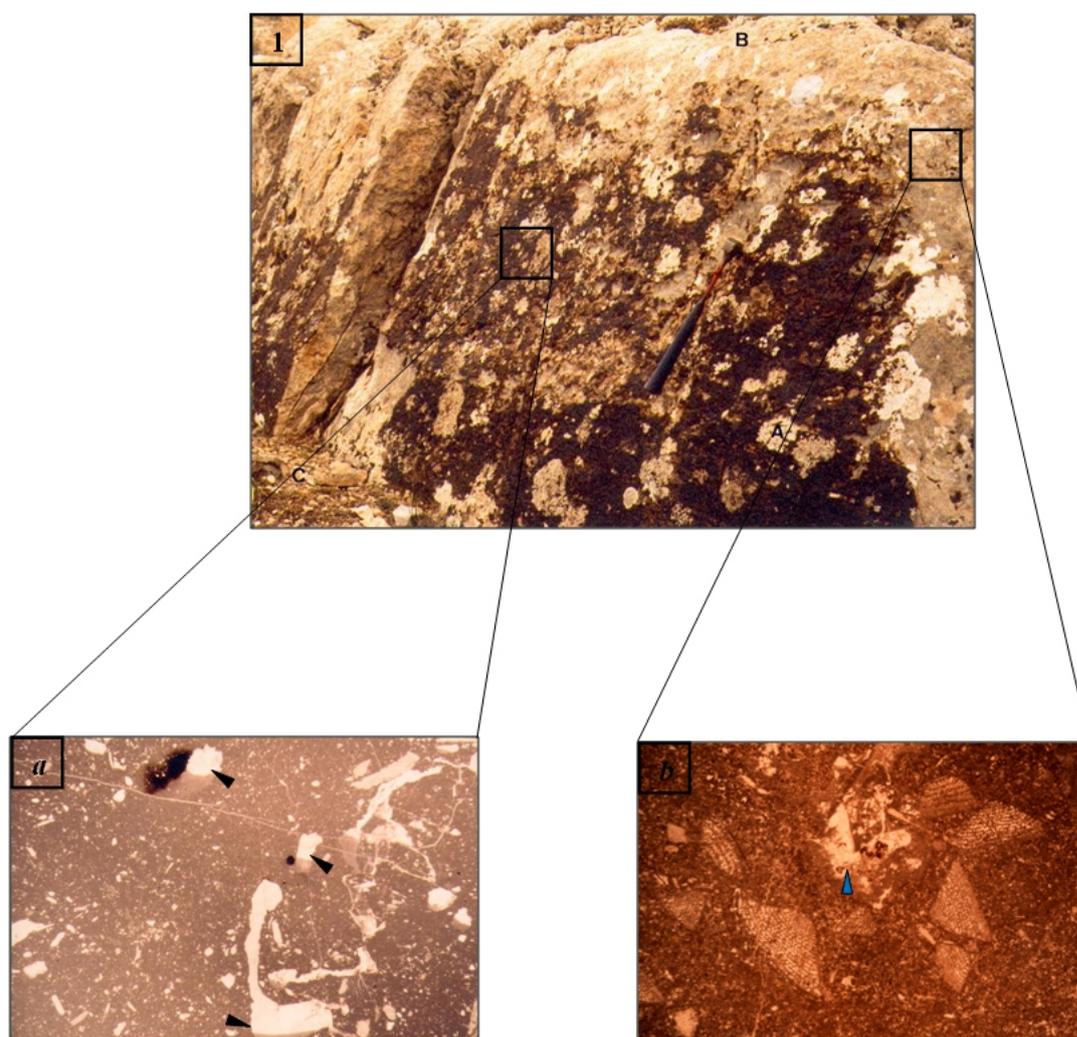


Photo 1: Macroscopic appearance of the hard ground on the top of Apto-Albian limestones.

A. Hard ground, B. Apto-Albian limestones, C. Yellow Cenomanian marls

a and b: Microscopic appearance of the hard ground on the top of Apto-Albian limestones.

Black arrows : Geopetal Figures, Blue arrow : Dissolution

Microscopic observation of thin sections from samples of this hard ground shows many small cavities partially or completely filled by internal sediment (*a* & *b*). This internal sediment consists of micrite and microcrystals of pyrite. These small cavities came from the dissolution of shells and shell debris of rudists when they are in a sedimentary environment where aragonite is unstable (ex: vadose zone). A cavity takes birth and can be filled with micrite. Residual vacuum is filled with calcite in mosaic; these diagenetic phenomena are interpreted as emersion marks.

This hard ground, where are localized dissolutions, of sedimentary origin is indisputable and cannot be mechanical as claimed by Vila (1980). It reflects an emersion of the neritic mole at that time Chadi (1991).

Conclusion

The study of the stratigraphic series of southern neritic massifs has allowed the refining of our predecessor's results. Yellow marls stratigraphically overlaying on the terminal Aptian-basal Albian through a hard ground, representing the middle Cretaceous. The presence of Ammonites remove all ambiguity about the age of these marls, dated Cenomanian-Turonian by Bär (1957) and which Vila (1980) reports to the "Miocene revamping the Cenomanian".

Microscopic analysis of samples from the hard ground surface has highlighted sedimentological phenomena of the emersion of the Constantinois platform during the terminal Aptian-basal Albian. The lower Cenomanian, conglomeratic and discordant, exists only in the Valley between Djebel Tolba and Djebel Guelaat Ouled el Hadj.

The Senonian shows lacuna in its base, also at the Taxas syncline, the Coniacian that surmounts the middle Cretaceous whereas it is the Santonian in the southern reverse of Djebel Guerioum.

These variations in thickness and these lacunas may be the result of socle structure legacy during the tectonic events that began at the Albian and which continued during the late Cretaceous.

Acknowledgement

We are most grateful to Jean-Pierre MASSE for his help in determination of fossils and microscopic analyzes.

References:

- Bär C. B. Etude géologique de la feuille 1/50.000° d'Aïn M'lila. *Publ. Serv. Carte géol. Algérie*, nouv. série, n°9, 249 p., 55 fig., 2pl.h.t(1957).
- Chadi M. Géologie structurale des Monts d'Aïn M'Lila (Algérie orientale). Thèse de l'Univ. de Nancy I, 67 fig., 10 pl., 1 pl. h.t., p. 191(1991).
- Kennedy W.J., Juignet P. and Wright C.W. A revision of the ammonite faunas of the Type Cenomanian. 3. Mantelliceratinae. *Cretaceous Research*, 7, 19-62. *Acta Geologica Polonica* (1986).
- Reymont Richard A. and Bengston P. Events of the Mid-Cretaceous. Ed. Pergamon Press. (1986).
- Sharpe.D. Description of the fossil remains of mollusca found in the chalk of England. Part III Cephalopoda. *The paleontographical Society p 45, pl XX (=Ammonites saxbii), p40 (pars.), pl XVII, fig 4 (non Ammonites mantelli Sowerby)* (1856).
- Vila J. M. La chaîne alpine d'Algérie orientale et des confins algéro-tunisiens. *Thèse ès-Sciences*, Paris VI, 2 tomes, 665 p (1980).
- Voûte C. Essai de synthèse de l'histoire géologique des environs d'Aïn Babouche et des régions limitrophes. *Publ. Serv. Carte géol. Algérie*, nouv. série, n° 36, t.1 255 p., t. 2, 136 p., 3 cartes, 41 pl (1967).

PEDOSTRATIGRAPHY, PEDOLOGICAL AND GEOCHEMISTRY OF KASHMIR LOESS: IMPLICATIONS FOR CHEMICAL WEATHERING HISTORY AND PALEOCLIMATIC RECONSTRUCTION

Rakesh Chandra
Ishtiaq Ahmad

Department of Earth Sciences, University of Kashmir, Srinagar, Jammu and Kashmir, India

Abstract:

Integrated pedological and geochemical study of the Quaternary Loess-Paleosols sediments of the Kashmir Valley was carried out in order to reconstruct their chemical weathering, paleoclimatological conditions and source. Pedological and micromorphic features of these paleosols indicate that these are weak to moderately developed. It also indicates that both loess deposition and pedogenic processes were taking place simultaneously during either phase of the loess/soil formation. These sediments are generally enriched with Fe₂O₃, MgO, MnO, TiO₂, Y, Ni, Cu, Zn, Th, Sc, V and Co while contents of SiO₂, K₂O, Na₂O, P₂O₅, Sr, Nb and Hf are lower than the UCC. Al₂O₃ is slightly higher than the UCC. However, CaO and U show large variations. Rb is generally similar to UCC whereas Ba is slightly lower than the UCC. Chondrite normalized REE patterns are characterized by moderate enrichment of LREEs, relatively flat HREE pattern ($Gd_{CN}/Yb_{CN} = 1.93$ to 2.30) and lack of prominent negative Eu anomaly ($Eu/Eu^* = 0.73$ to 1.01 , average = 0.81). The weathering indices suggest that these sediments are experiencing weak to moderate degree of weathering and not subjected to potash metasomatism. On the basis of these proxies it is inferred that the climate of Kashmir Valley for the recent past fluctuated between cold arid to warm semi-arid.

Key Words: Loess, Paleosols, weathering, Karewas, Kashmir

Introduction

Pleistocene loess deposits in the Kashmir Valley of the North Western Himalaya are the thickest and most extensive in Indian subcontinent. These loess deposits represent an important archive of terrestrial paleoenvironmental changes. Stratigraphically, the loess-paleosol sequence of the Kashmir Valley fall into two broad types: a thick sequence and a thin sequence. A thick sequence is about 21m thick and lie toward the southwest of the Kashmir Valley along the Pir-Panjal flank. This sequence consists of ten paleosols and three loess horizons (Ahmad, 2012). The thin sequence lies toward the northeastern part of the Kashmir Valley and contains four paleosols. This sequence is equivalent to the top part of the thicker sequence (Ahmad, 2012). Singhvi et al. (1987) proposed that the Kashmir loess sequence extended back to the 350ka B.P. However, Gupta et al. (1991) proposed much shorter framework and concluded that the base of the loess sequence is approximately 200ka B.P. In this paper we present the pedostratigraphy as well as pedological and geochemical characterization of the Kashmir Loess-Paleosol sequence. An attempt has also been made for paleoclimatic reconstruction.

Geological setting

Kashmir valley has the morphological characteristics of an intermountain basin and is located on a nearly horizontal Nappe sheet (Wadia, 1976). The valley is flanked by the Himalayas to the northeast and Pir-Panjal Range (Panjal Thrust) in the southwest. These mountainous ranges comprise metamorphosed Paleozoic and Mesozoic marine sediments and effusive rocks (Farooqi and Desai, 1974). The valley preserves the record of past 4 M.Y in which the sedimentation is controlled by the tectonic events. The valley possesses almost complete stratigraphic record of rocks of all ages ranging from Archean to Recent (Fig.1).

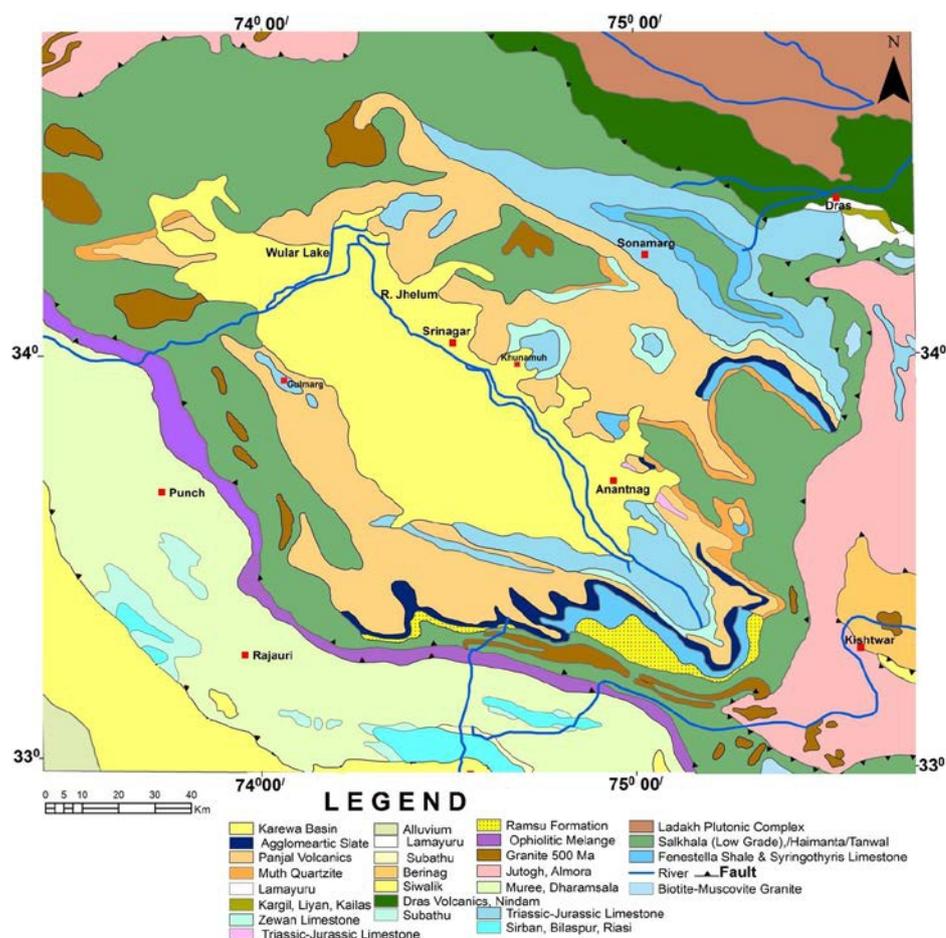


Fig.1. Geological map of Kashmir Himalaya (modified after Thakur and Rawat, 1992).

The Panjal Volcanic series and the Triassic limestone form the main geological formations and are underlain by the Archean metasedimentaries of Salkhala Formation (Fig.1). The Panjal Volcanic series of Permo-Carboniferous age is divisible into two well marked horizons, the lower agglomeratic slate and the upper Panjal lava flows (Bhatt and Zainuddin, 1979). These oldest rocks are found around the northwestern extremity of the Kashmir valley and portions of the Pir-Panjal range. Exposures of Triassic rocks comprise alternate thick dark grey limestone and shally-arenaceous impure limestone. The other rocks of lesser distribution include Dogra Slates, Cambro-Silurian, Zewan Formation and Muth-Quartzites. The Precambrian to Mesozoic basement rocks in turn are overlain by Plio-Pleistocene sediments constituting the Karewa Group, which in turn capped by the loess sediments of Dilpur Formation.

Lithostratigraphy of Kashmir Loess-Paleosols

Three representative loess-paleosol containing sections have been selected for detailed study (Fig.2). On the basis of detailed field observations, four loess horizons with seven embedded paleosol profiles have been identified at Dilpur Village section. Karapur Village section contains ten paleosol profiles and three loess horizons. However, at Burzahom Village section four paleosol profiles have been found. Key macromorphological features of these sediments are illustrated in Fig.3. The paleosol profiles show low organic matter contents and weak to moderately developed illuvial clay pedofeatures, which suggests subtle climatic changes that affect relative rates of material supply and weathering rates. This further suggests that these paleosol profiles are formed when both loess deposition and pedogenic processes were taking place simultaneously during both stadial to interstadial phases, representing cold arid to warm semi-arid climate. Therefore, it is inferred that these paleosol profiles do not each represent a complete interglacial period (Gardner, 1989). The paleosol profiles DS3, KS4 and BS4 are relatively well developed and record maximum thickness, which represent warm semi-arid climate. The parent loess horizons are mostly absent at the base of these paleosol profiles (Fig.3). This indicates relative stable land surface conditions when the

pedogenic processes become more dominant and the parent loessic material transformed into illuvial ('B_t') horizon (Yakimenko et al. 2004). The lithological characteristics such as organic activities, illuvial pedofeatures and granularity of soil are relatively weakly developed at Dilpur village section as compared to Karapur and Burzahom village sections. However, the different types of calcretes such as CaCO₃ coating, infilling, platy concretions and nodules are relatively well developed at Dilpur Village section than Karapur Village section. Lithological characters reveal that the Dilpur Village section experienced relatively arid climatic conditions as compared to the Karapur and Burzahom Village sections. The Dilpur Village section is lower in altitude than Karapur and Burzahom

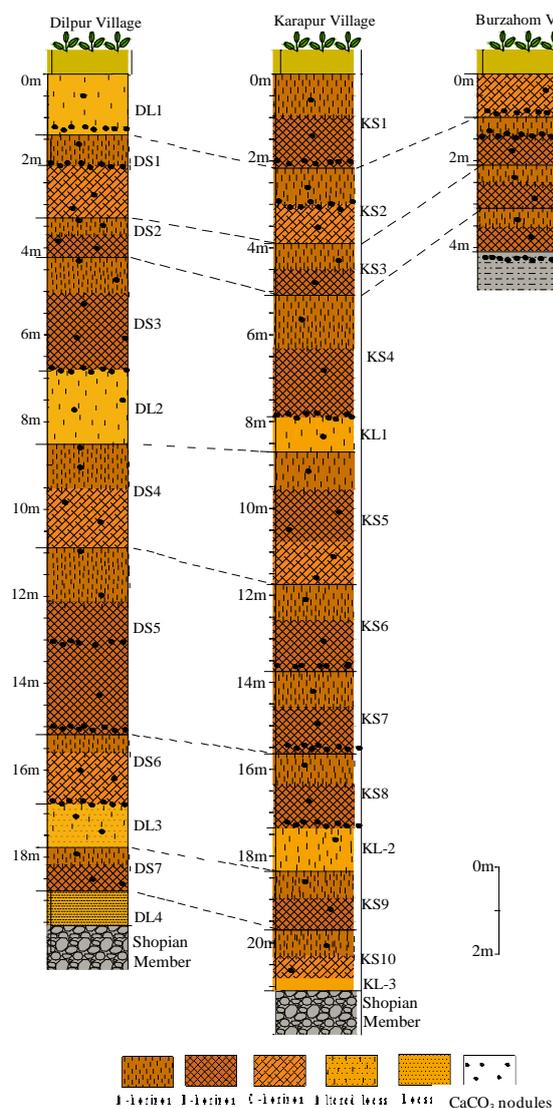


Fig.2. Lithostratigraphic correlation of Kashmir Loess-Paleosol sediments at Dilpur, Karapur and Burzahom Village sections.

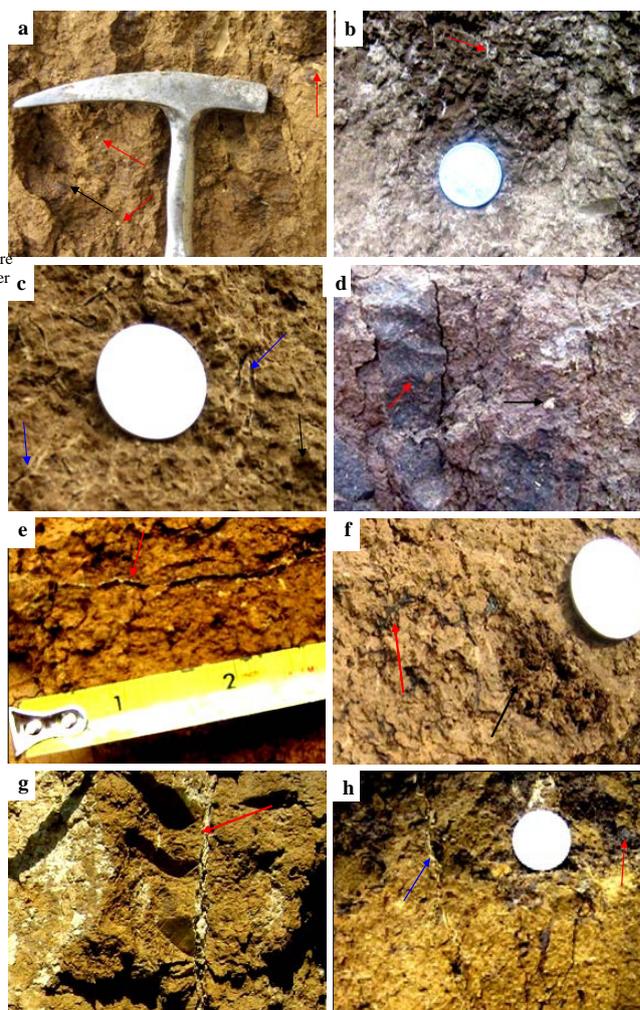


Fig.3. Photograph showing (a) clay coating (black arrows) and calcareous nodules (red arrows) in paleosol profile DS5, (b) granular structures and root traces coated with CaCO₃ in paleosol profile DS7, (c) root traces coated with CaCO₃ in paleosol profile DS3, (d) clay coating (red arrow) and nodules of CaCO₃ (black arrow) in paleosol profile KS9, (e) root traces coated with CaCO₃ (arrow) in paleosol profile KS4, (f) burrows (red arrow) and root traces (black arrows) in paleosol profile KS7, (g) platy concretions in BS2 paleosol profile, (h) clay coating (red arrows) and platy concretion of CaCO₃ (blue arrow) in BS3 paleosol profile.

Village sections and hence has high potential for evapo-transpiration. Hence, it experiences relatively dry climatic conditions than the Karapur and Burzahom Village sections. Therefore, the soil-water balance of these locations differs which affects the rate of pedogenesis (Bronger et al. 1987). This suggests that the local geographical conditions also played vital role in the pedogenic modification of these sediments (Bronger et al. 1987). Overall these loess-paleosol sediments show similar type of lithological characteristics although there is difference in degree of maturity of soil at different sections.

Micromorphology

The micromorphological study of three loess-paleosol containing sections of Kashmir Valley has been carried out to determine the pedogenic processes and climatic conditions prevailing during their development. Key micromorphological features of these loess and paleosols sediments are illustrated in Fig.4. The description refers to the most representative and typical micromorphological features of the about 150 thin sections. Presented here in detail are only the most typical pedofeatures with clear environmental implications. Among a large variety of microstructures, massive or apedal microstructures, channel microstructures, peds microstructures, spongy structure, textural pedofeatures, calcitic pedofeatures, basic mineral components and organic matter are the most common in the studied loess-paleosols sediments.

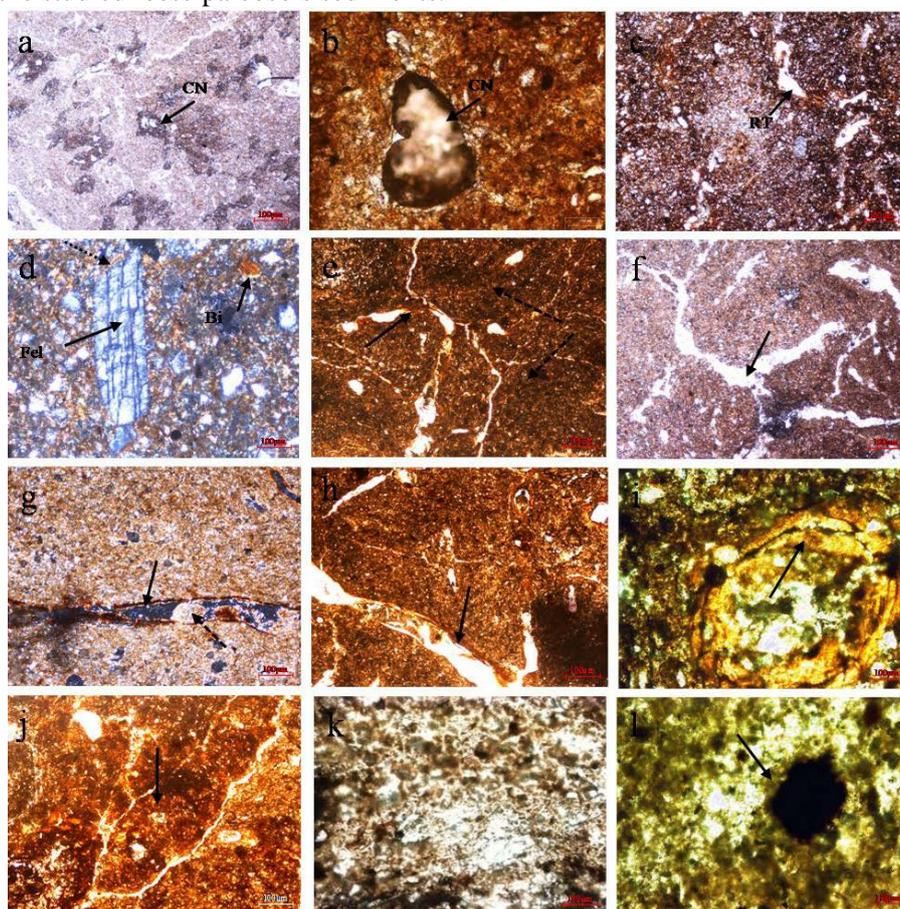


Fig.4. Photomicrographs from Kashmir Loess-Paleosols at Dilpur, Karapur and Burzahom Village sections illustrating some key micromorphological features of the Kashmir Loess-Paleosols sediments. (a-b) massive structures with CaCO_3 nodules (CN) in D-L1 and K-L1 loess horizons respectively, (c) massive structures with disseminated organic matter and fine root traces (RT) in K-S2 horizon, (d) massive structures with thin clay coating around the skeleton grains (dotted arrow), feldspar (Fel) and partially altered biotite (Bi) in paleosol profile K-S10, (e) peds (dotted arrows) and limpid yellow brown clay coating along the channels (solid arrow) in paleosol profile D-S5, (f) fluid conducting channels coated with CaCO_3 in paleosol profile B-S1, (g) ferruginized plant remains (solid arrow) and CaCO_3 infillings (dotted arrow) in paleosol profile K-S5, (h) channels showing thick ferruginous clay coatings in paleosol profile D-S3, (i) rounded channels with alternate layers of thick limpid yellow clays in paleosol profile D-S3, (j) platy pedal microstructures in paleosol profile K-S5, (k) spongy microstructures in paleosol profile D-S2, and (l) spongy microstructures with Fe/Mn oxides in D-L3 loess horizon.

Massive microstructures are generally observed in loess (Fig.4a,b) and weakly altered 'B_w' horizons of the sequences (Fig.4c,d), showing homogeneity and small porosity in the groundmass. Thin clay coating is also observed around the skeleton grains in 'B_w' horizons (Fig.4d), with finely disseminated organic matter assimilated with the groundmass (Fig.4c). The paleosols possessing these structures imply their subjection to only weak soil forming processes. However, channels mainly result from root growth (Zarate et al. 2000) and by the fluid migration within the profile (Fig.4e,f). These channels sometimes contain ferruginized plants remains with coatings, hypocoatings and infilling of secondary CaCO₃ (Fig.4g). This type of microstructure is generally found in weakly developed paleosol profiles (Li et al. 1992). In contrast, channels in the moderately developed paleosols profile are relatively large and more abundant (Fig.4h,i). All the paleosols show weak to moderate pedality along with large channel and platy peds microstructures (Fig.4e,j). These peds are formed by shrink-swell activities due to fluctuations in water saturation because of seasonal wetting and drying conditions (Kemp and Zarate, 2000). Spongy microstructure mainly results from strong biological activity. This is observed in the surface horizons of paleosol profiles DS2 and KS3 and loess horizon DL3 (Fig.4k,l). It is typically associated with surface (A) horizons of soils, so the presence of spongy microstructures in the lowermost 'C' horizons (Fig.4l) of the profile supports the accretionary nature whereby the loess was modified by bioturbation (and probably leaching) processes as it accumulated (Kemp et al. 2003). Rates of loess deposition eventually diminished to such an extent that a stable land surface was established and pedogenic processes became more dominant (Kemp et al. 2003). The soils with spongy microstructures can be regarded as weakly developed steppe soils in semi-arid environment (Zhengtang et al. 1996). Thick microlaminated clay coatings along large channel voids and thin clay coatings along planar voids or channel walls are common (Fig.4e). The secondary CaCO₃ nodules show clear boundaries to the groundmass (Fig.4b). The depth distribution of these features provides the basis for the modification of the horizon nomenclature and a reconstruction of the pedosedimentary events. Quartz is the dominant mineral in both paleosols and the parent loess material followed by feldspar (Fig.4d). The overall lack of coarse material suggests these sediments are mostly from the distant and uniform source region suggesting large provenance with variable geological settings which apparently have undergone weak to moderate recycling processes (Ahmad and Chandra, 2013).

Geochemical characterization

For chemical analysis, 46 samples of Kashmir loess-paleosols were analyzed and average composition is presented in Table 1. Major and trace elements were determined by

Sample	DL1	DS1	DS3	DS4	DS6	DS7	KS1	KS2	KL1	KS7	KL2	KS10	KL3	BS1	BS4
SiO ₂	51.4	55.17	60.57	56.27	54.05	57.71	58.72	62.97	54.5	62.60	58.21	59.96	63.18	53.9	60.73
Al ₂ O ₃	12.58	13.76	16.47	14.01	14.01	15.43	15.77	16.08	13.95	16.53	15.65	16.13	13.8	13.50	16.30
Fe ₂ O ₃	4.87	5.26	6.78	5.34	5.13	5.81	6.04	6.18	5.34	6.39	5.72	7.02	6.83	5.10	6.756
MnO	0.08	0.08	0.11	0.08	0.08	0.09	0.09	0.10	0.07	0.11	0.091	0.10	0.07	0.08	0.08
MgO	2.59	2.77	2.55	2.73	2.69	2.52	2.62	2.52	2.58	2.36	2.65	2.40	2.62	2.52	2.59
CaO	8.6	6.55	1.31	6.36	6.83	4.06	3.06	1.51	6.33	1.09	3.87	1.97	3.08	7.02	1.43
Na ₂ O	0.87	0.94	0.92	0.96	0.81	0.78	1.00	1.15	0.81	0.88	0.97	0.92	1.36	0.80	0.91
K ₂ O	2.68	2.79	3.11	2.78	2.71	2.92	2.87	2.89	2.73	2.89	2.87	2.43	1.86	2.92	2.64
TiO ₂	0.6	0.63	0.78	0.64	0.63	0.72	0.71	0.78	0.65	0.79	0.69	0.91	0.99	0.64	0.94
P ₂ O ₅	0.128	0.12	0.10	0.11	0.12	0.10	0.14	0.12	0.12	0.11	0.139	0.11	0.12	0.12	0.09
LOI	14.81	11.41	7.95	10.70	12.25	9.7	9.26	7.10	11.74	7.84	8.33	7.08	5.73	12.5	7.73
Total	99.20	99.51	100.6	100	99.31	99.85	100.3	101.4	98.88	101.6	99.19	99.08	99.65	99.14	100.2

Rb	100.9	112.4	152.0	112.3	106.5	125.8	127.2	140.4	108.4	141	128.1	105.6	68.06	114.7	110.8
Sr	142.3	143.5	112.5	136.6	146.6	155.7	116.8	114.0	146.3	106.6	135.6	133.6	149.8	139.5	108.2
Y	24.40	27.21	32.65	25.87	25.97	33.45	32.56	33.65	28.49	33.73	31.12	34.68	31.61	26.42	37.59
Zr	162.1	180.9	208.1	182.2	178.9	201.2	204.2	223.4	179.6	230.2	207.6	234.4	205	172.1	295.1
Nb	14.30	14.61	18.25	14.48	14.55	16.39	17.17	18.67	15	18.41	16.73	19.41	19.88	14.45	19.36
Ba	374.1	406.1	557	412.1	393.1	471.1	474.1	519.2	391	508.9	478.1	433.2	306.3	395.4	447.8
Ni	41.82	44.56	58.67	44.95	43.19	46.56	53.11	50.97	48.2	55.24	43.47	45.61	35.84	42.40	45.41
Cu	43.24	45.96	52.84	44.19	43.11	45.95	47.39	46.76	47.3	49.50	48.31	45.76	42.52	49.34	45.90
Zn	74.43	78.13	102.3	75.37	75.87	84.61	88.58	85.68	85.1	100.3	86.08	86.56	71.21	89.33	77.70
Ga	12.06	13.17	17.53	12.89	11.89	14.40	14.94	16.27	12.6	15.95	14.18	15.72	13.44	11.92	14.66
Pb	17.05	20.22	23.64	20.35	19.79	20.44	22.33	21.38	19	21.02	20.31	18.99	13.44	21.44	21.69
Th	10.96	13.70	16.41	12.74	12.97	17.36	16.73	18.69	12.5	15.33	16.90	12.73	8.02	11.26	15.94
U	2.98	3.84	6.45	5.06	7.35	8.48	5.49	7.53	5.64	5.78	4.81	6.62	3.82	0.949	4.41
Sc	12.39	12.97	14.99	13.24	13.29	15.39	13.81	14.39	13.44	-	-	-	-	13.93	16.66
V	189.3	89.33	100.8	99.56	93.26	110	107.3	112.6	99.91	-	-	-	-	103.7	135.5
Co	15.04	16.09	18.25	16.5	15.93	17.34	17.56	19.4	17.27	-	-	-	-	18.21	21.45
Hf	2.09	2.32	2.72	2.5	2	3	2.64	3.675	3	-	-	-	-	1.87	3
La	34.13	37.12	41.51	37.87	37.18	41.59	40.18	45.36	38.13	-	-	-	-	42.63	49.4
Ce	66.28	72.87	82.1	72.72	72.36	79.68	78.04	87.01	69.42	-	-	-	-	79.55	93.28
Pr	7.26	8.10	9.19	8.21	8.01	9.25	9.015	10.21	8.26	-	-	-	-	8.91	10.28
Nd	26.71	30.07	33.72	30.05	28.92	33.74	34.44	39.69	32.5	-	-	-	-	31.91	39.43
Sm	5.5	6.7	6.89	6.75	6.99	7.21	6.79	7.26	6.72	-	-	-	-	6.51	7.77
Eu	1.33	1.54	1.61	1.57	1.42	1.8	1.58	1.65	1.55	-	-	-	-	1.55	1.82
Gd	4.83	5.05	5.72	5.2	5.07	6.25	5.62	6.16	5.36	-	-	-	-	5.57	6.61
Tb	0.71	0.73	0.84	0.74	0.73	0.93	0.80	0.87	0.76	-	-	-	-	0.82	0.94
Dy	3.85	4.02	4.64	4.05	4.09	5.02	4.59	5.12	4.34	-	-	-	-	4.22	4.99
Ho	0.96	0.99	1.16	1	0.99	1.25	1.10	1.19	1.07	-	-	-	-	1.05	1.3
Er	2.2	2.21	2.60	2.31	2.27	2.83	2.48	2.67	2.3	-	-	-	-	2.45	2.94
Tm	0.37	0.37	0.43	0.38	0.37	0.47	0.41	0.45	0.4	-	-	-	-	0.4	0.47
Yb	2	2	2.30	2.02	2.02	2.44	2.18	2.24	2.05	-	-	-	-	2.14	2.30
Lu	0.26	0.25	0.3	0.25	0.26	0.32	0.29	0.31	0.26	-	-	-	-	0.29	0.33
CIA	68.58	69.02	71.97	69.11	71.44	72.93	71.07	69.89	71.28	73.16	71.3	74	68.02	69.87	73.55
CIW	81.46	81.49	84.43	81.36	84.01	85.73	82.71	80.88	83.96	84.95	83.06	84.15	75.51	83.57	84.46

Table 1. Average geochemical composition of loess-paleosol sediments of Kashmir valley, India. using X-Rays Fluorescence (XRF) spectrometer and REE with some trace elements were determined by ICP-MS at Wadia Institute of Himalayan Geology (WIHG), Dehradun, India. The accuracy of the analytical method was established using two internationally recognized standard reference materials: MAG-I and MAG (R.V). Loss on ignition (LOI) was calculated as a percentage of dry weight after the samples were ignited at 950°C for 24 hour.

To evaluate the nature of the Kashmir Loess-Paleosol sediments and chemical behavior of major oxides during pedogenic process, the concentration of major oxides is also plotted in Upper Continental Crust (UCC) normalized spider diagrams (Fig.5). SiO_2 and Al_2O_3 wt% show constant proportions and generally follow the similar trend to that of UCC, PAAS and NASC. This suggests warm semi-arid climatic conditions because these elements remain stable under semi-arid climate (Guo, 2010). However, the UCC normalized Fe_2O_3 , MnO and MgO wt% show slightly higher concentration than the UCC but follow similar pattern to that of PAAS. This suggests that the ferromagnesian minerals are least affected by the pedogenic modification.

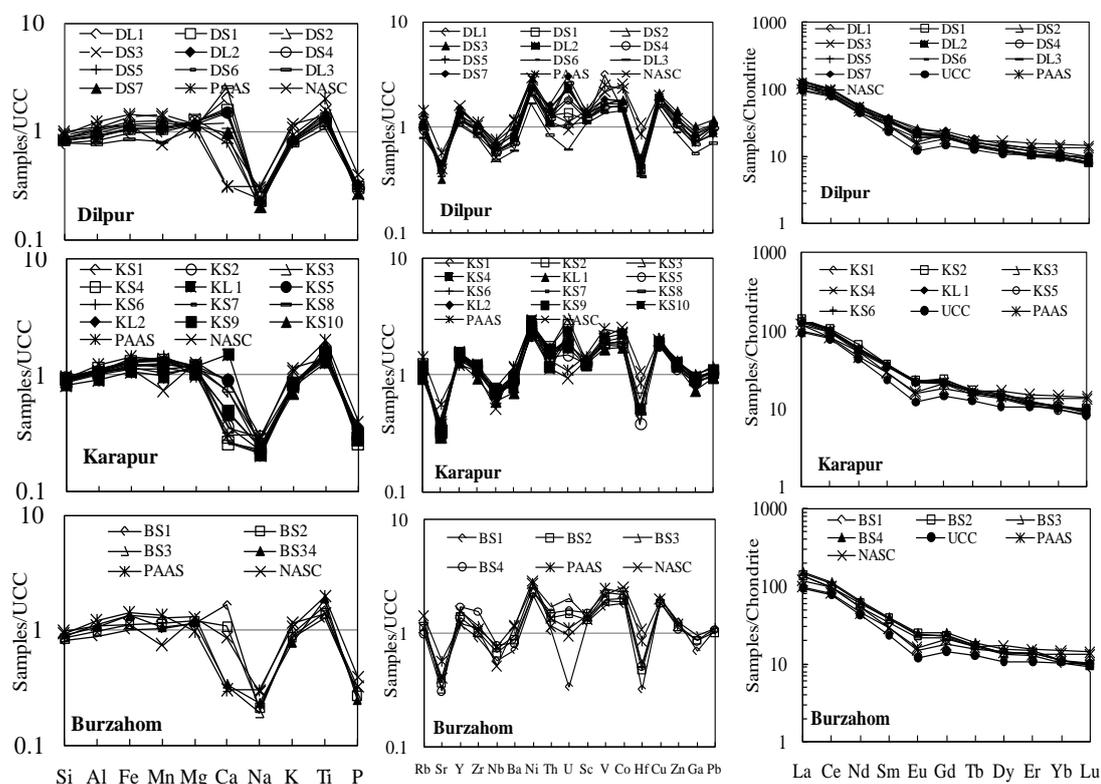


Fig.5. UCC normalized spider diagrams for major oxides and trace elements composition of Kashmir Loess-Paleosol sediments at Dilpur, Karapur and Burzahom Village sections, (PAAS and UCC values after Taylor and McLennan, 1985; NASC values after Gromet et al. 1984). Also shown are the chondrite normalized REE patterns for these sediments. All the major oxides are as oxides wt%.

All the three sections show almost similar concentration of K_2O wt%, which is generally lower than the UCC, PAAS and NASC. The negative anomalies of Na_2O followed by K_2O wt% (Fig.5) indicate relatively more alteration of plagioclase feldspar than the K-feldspar. However, CaO wt% shows wide range of variation in its concentrations with respect to UCC. It is well known that the CaO contents of loess varies greatly and shows both positive and negative anomalies on UCC normalized spider diagrams (Gallet et al. 1998). The wide range of variations in CaO wt% may be argued for high LOI, which ranges from 5.71 to 14.8 wt% (Honda et al. 2004). The lower values of CaO wt% relative to PAAS indicate an intense mobilization during post depositional processes, whereas high values indicate that these sediments are relatively less mature than the PAAS (Mahjoor et al. 2009). TiO_2 wt% contents are significantly higher than the UCC. However, it is slightly lower than the PAAS. The enrichment of TiO_2 wt% is generally ascribed to the presence of Ti-bearing phyllosilicates (biotite and chlorite). These minerals generally reside in fine clay sediments. It also suggests significant contribution from the mafic source rocks. The positive correlation between ΣREE and Al_2O_3 contents ($r = 0.81$) also indicates that the REE are mostly concentrated in the fine clay fraction. This suggests significant proportion of clay minerals in the Kashmir Loess-Paleosol sediments.

To evaluate the nature and chemical behavior of various trace elements during pedogenic process, the average concentration of trace elements is also plotted in Upper Continental Crust (UCC) normalized spider diagrams (Fig.5). The Sr shows negative correlations ($r = -0.68$) with SiO_2 (wt%) which indicates that the sediments are depleted in carbonate minerals. SiO_2 show positive correlation with Ba ($r = 0.66$) and Rb ($r = 0.50$) suggesting their robustness during pedogenesis. The Rb and Sr show different geochemical behavior during pedogenesis. The relatively higher concentration of Rb indicates that mica and K-feldspar are not strongly affected by the pedogenesis. The weathering of plagioclase decreases the concentration of Sr because Sr is more mobile than the Ba.

Rb/Sr ratio shows higher concentration in paleosols than the altered loessic layers. Ba/Sr ratio also follows the similar trend (Fig.6). This ratio has been considered as related to the leaching intensity (Gallet et al. 1996). High peaks of Rb/Sr and Ba/Sr in paleosols are the result of land surface stability during the warm and wet periods which accelerate pedogenesis (Gallet et al. 1996). The diagram also used to correlate the lithostratigraphy of Kashmir Loess-Paleosol sediments. The pedocomplex between DL1–DL2 (0m–8.5m) at Dilpur is stratigraphically equivalent to the KS1–KL1 (0m–8.7m) at Karapur. The whole Burzahom section is stratigraphically equivalent to this pedocomplex. The pedocomplex between DL2–DL3 represents another warm-wet period. It is stratigraphically equivalent to KL1–KL2. The other small peaks represent the minor fluctuations in precipitation conditions. The increase in concentration of CaO (wt%) upward suggests the increasing aridity during the close of Pleistocene Period. Similar patterns are also observed for $\text{Al}_2\text{O}_3/\text{SiO}_2$, $(\text{Al}_2\text{O}_3 + \text{Fe}_2\text{O}_3)/(\text{Na}_2\text{O} + \text{K}_2\text{O} + \text{MgO} + \text{P}_2\text{O}_5)$, $\text{Al}_2\text{O}_3/(\text{CaO} + \text{Na}_2\text{O} + \text{K}_2\text{O})$, $\text{Fe}_2\text{O}_3/\text{CaO}$ and $\text{Fe}_2\text{O}_3/\text{Al}_2\text{O}_3$ (not included here). Further, the ratios of immobile elements such as La/Co, Zr/Y and Zr/Hf, show no correlation with Al_2O_3 (-0.49 , 0.0029 , 0.075 respectively) and CIA values (-0.57 , -0.0366 and 0.051 respectively) which suggest that these elements are resistant to chemical weathering. The whole Burzahom section is stratigraphically equivalent to this pedocomplex.

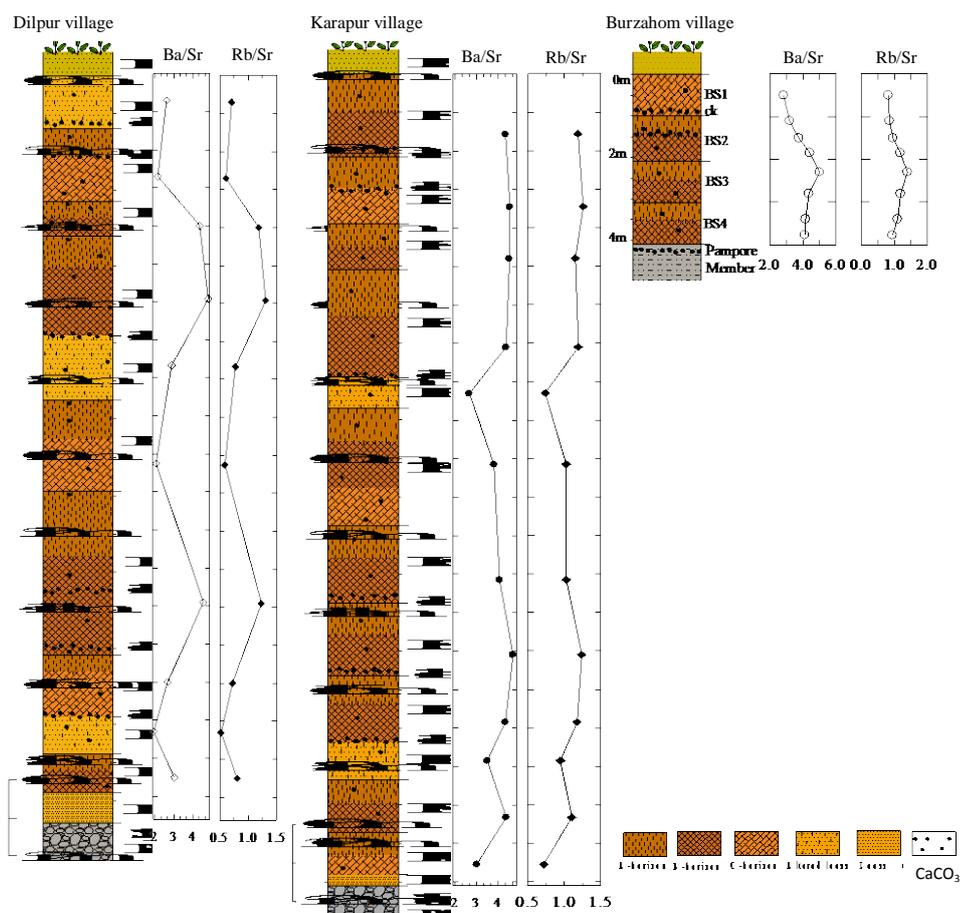


Fig.6. Showing variations of Ba/Sr (ppm) and Rb/Sr (ppm) with stratigraphic depth in Kashmir Loess-Paleosol sediments at Dilpur, Karapur and Burzahom Village sections.

The pedocomplex between DL2–DL3 represents another warm-wet period. It is stratigraphically equivalent to KL1–KL2. The other small peaks represent the minor fluctuations in precipitation conditions. The increase in concentration of CaO (wt%) upward suggests the increasing aridity during the close of Pleistocene Period. Similar patterns are also observed for $\text{Al}_2\text{O}_3 / \text{SiO}_2$, $(\text{Al}_2\text{O}_3 + \text{Fe}_2\text{O}_3) / (\text{Na}_2\text{O} + \text{K}_2\text{O} + \text{MgO} + \text{P}_2\text{O}_5)$, $\text{Al}_2\text{O}_3 / (\text{CaO} + \text{Na}_2\text{O} + \text{K}_2\text{O})$, $\text{Fe}_2\text{O}_3/\text{CaO}$ and $\text{Fe}_2\text{O}_3/\text{Al}_2\text{O}_3$ (not included here). Further, the ratios of immobile elements such as La/Co, Zr/Y and Zr/Hf, show no correlation with Al_2O_3 (−0.49, 0.0029, 0.075 respectively) and CIA values (−0.57, −0.0366 and 0.051 respectively) which suggest that these elements are resistant to chemical weathering.

The Chondrite normalized REE patterns are plotted in Fig.5. These are characterized by moderate enrichment of LREEs, relatively flat HREE pattern ($\text{Gd}_{\text{CN}}/\text{Yb}_{\text{CN}} = 1.93$ to 2.30), lack of prominent negative Eu anomaly ($\text{Eu}/\text{Eu}^* = 0.73$ to 1.01, average = 0.81) and variable amount of $\sum\text{REE}$. The Eu and Ce anomalies in the samples are determined according to: $\text{Eu}/\text{Eu}^* = (\text{Eu}_{\text{CN}}) / \{(\text{Sm}_{\text{CN}}) \times (\text{Gd}_{\text{CN}})\}^{0.5}$ and $\text{Ce}/\text{Ce}^* = (\text{Ce}_{\text{CN}}) / \{(\text{La}_{\text{CN}})^{0.666} \times (\text{Nd}_{\text{CN}})^{0.333}\}$. Eu anomaly ranges between 0.73 and 1.01 (average = 0.81). In contrast, nearly half of the samples show positive Ce anomaly and it ranges from 0.92 - 1.04 (average = 0.99). The $\text{La}_{\text{CN}}/\text{Yb}_{\text{CN}}$ ratio of the studied samples do not correlates with the weathering indices (CIW vs $\text{La}_{\text{CN}} / \text{Yb}_{\text{CN}}$; $r = 0.090$ and PIA vs $\text{La}_{\text{CN}} / \text{Yb}_{\text{CN}}$; $r = 0.14$). Further, the absence of correlation between Eu/Eu^* vs Al_2O_3 (−0.29) and Eu/Eu^* vs CIA (−0.38) indicates that chemical weathering did not fractionate LREE from HREE. This lack of evidence of intense weathering at the source depicted by the LREE/HREE ($\text{La}_{\text{CN}}/\text{Yb}_{\text{CN}}$) ratios suggests that the REEs are not subjected to weathering (Cai et al. 2008). Hence, REE pattern of the studied samples is mainly inherited from the source provenance.

Weathering Intensity

To know the extent of pedogenesis, Chemical Index of Alteration (CIA), proposed by Nesbitt and Young (1982) have been calculated. This index used to determine the proportion of primary minerals and the transformation of feldspars to secondary clay minerals relative to the fresh parent material. Hence CIA value provides an accurate measurement of degree of weathering and it can be obtained by using molecular proportions: $\text{CIA} = (\text{Al}_2\text{O}_3 / \text{Al}_2\text{O}_3 + \text{CaO}^* + \text{Na}_2\text{O} + \text{K}_2\text{O}) \times 100$. CaO^* represents the CaO in silicates bearing minerals. The CIA value of Kashmir Loess-Paleosol sediments ranges from 67.13 to 75.27 (Table-I). This narrow and restricted range of CIA value shows moderate degree of weathering, suggesting dry and cold climate during deposition (Nesbitt and Young, 1982). According to Taylor and McLennan (1985), the moderate weathering suggests that the effect of weathering had not advanced to the stage where alkali and alkaline earth elements are substantially removed from the soil. Chemical index of weathering (CIW) proposed by Harnois (1988) also used to determine the degree of weathering and it can be obtained by using molecular proportions. $\text{CIW} = (\text{Al}_2\text{O}_3 / \text{Al}_2\text{O}_3 + \text{CaO}^* + \text{Na}_2\text{O}) \times 100$. CIW values of these samples range from 79.93–88.63, suggesting moderate degree of weathering. The consistency of CIA with CIW values indicate that the degree of pedogenesis operated on the Kashmir Loess-Paleosol is moderate. Ternary A–CN–K diagram $\{\text{Al}_2\text{O}_3 - (\text{CaO}^* + \text{Na}_2\text{O}) - \text{K}_2\text{O}\}$ also used to deduce the weathering trend (Nesbitt and Young, 1982). In this diagram (A = Al_2O_3 ; CN = $\text{CaO}^* + \text{Na}_2\text{O}$; K = K_2O), the loess-paleosol sediments plot above plagioclase-potash feldspar line (Fig.7), and clustered close to the PAAS and NASC and fall intermediate between A–CN and A–K lines, which show weak to intermediate removal of Ca and Na (Bugge et al. 2008). The plots do not exhibit any inclination towards the K apex indicating that the loess-paleosol sediments were not subjected to potash metasomatism (Moosavirad et al. 2010).

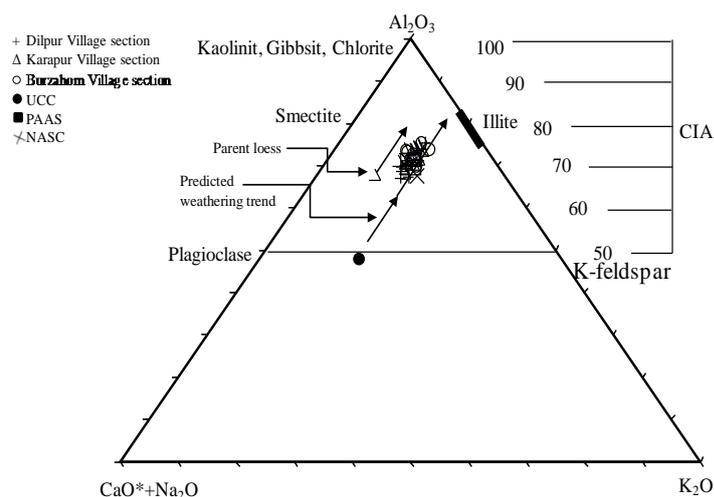


Fig. 7. A–CN–K ternary diagram for the Kashmir Loess-Paleosol sediments (after Nesbitt and Young, 1984), also plotted are the average UCC, PAAS (Taylor and McLennan, 1985) and NASC (after Gromet et al., 1984) values as well as some rock forming minerals important in silicate rock weathering; shown at the side is the CIA scale. The Kashmir Loess-Paleosol sediments fall closer to moderately weathered minerals.

Therefore, use of CIA index in weathering studies assumes that this index is a measurement of the amount of the chemical weathering. However, other factors that may affect the CIA value and need to be taken into account include sediment provenance and post-depositional processes that lead to K^+ addition (e.g. diagenetic illitization and metasomatism). Sedimentary sorting can significantly influence the chemical composition of terrigenous sediments due to grain size and mineral sorting (Bauluz et al. 2000). For instance, aluminum is concentrated in the clays, hence the larger the transport (i.e. distal regions), the finer the sediments and the higher the Al concentration (Soreghan and Soreghan, 2007). There is also a tendency of larger grain sizes to concentrate feldspars, which leads to lower CIA values (Zimmerman and Bahlburg, 2003). Therefore, the use of the CIA as a weathering index, however, can be limited by the inheritance of clays from sedimentary rocks in the source area. However, in this study it reveals that these sediments are enriched in rock forming minerals with significant proportion of clays, indicating that CIA value to some extent is affected by these clays. In addition, the A-CN-K diagram (Fig.7) also indicates that these loess-paleosol sediments are not subjected to potash metasomatism. Therefore, weathering intensity inferred by these proxies indicating moderate degree of weathering, probably suggest combined result of weathering and grain size effect due to transportation processes. Hence, on the bases of these geochemical observations it is proposed that the Kashmir Loess-Paleosol sediments experienced weak to moderate degree of weathering.

Conclusion

Integrated micromorphological and geochemical study of the Quaternary Loess-Paleosols sediments of the Kashmir Valley revealed that all these sediments are characterized by similar pedofeatures irrespective of horizon types suggesting syndepositional origin of loess deposits and weak to moderate weathering. It further suggests that loess deposition and pedogenesis is likely competing processes and neither stop completely during either phase of the loess/soil formation. Chemically, these sediments also show similar compositions and alteration history. Only mobile elements Ca, Na, P and Sr are depleted in these sediments. Chondrite normalized REE patterns are characterized by moderate enrichment of LREEs, relatively flat HREE pattern ($Gd_{CN}/Yb_{CN} = 1.93$ to 2.30), lack of prominent negative Eu anomaly ($Eu/Eu^* = 0.73$ to 1.01, average = 0.81) and variable amount of total REE ($\Sigma REE = 156.1$ to 226.43). The weathering indices (CIA, CIW and A-CN-K diagram) and others elemental ratios suggest that these sediments experienced weak to moderate degree of chemical weathering and not subjected to potash metasomatism during diagenesis. Ratios of various major and trace elements suggest that Middle to Late Pleistocene period is defined by several episodes of pedogenic activity representing warm arid to semi arid climatic conditions in the valley. This is further supported by the clay mineralogical study of these sediments which indicates that these sediments are enriched with smectite with lower concentration of mixed-layered chlorite + kaolinite (c+k) and traces of illite clay minerals (not included here).

Acknowledgements

The authors are thankful to Head of Department, Department of Earth Sciences, University of Kashmir, Srinagar, J&K, India and Director, Wadia Institute of Himalayan Geology (WIHG),

Dehradun, India for providing laboratory and analytical facilities for generating micromorphological and geochemical data.

References:

- Ishtiaq Ahmad 2012. Geochemical study of loess-paleosol Quaternary sediments of Karewa Basin with reference to paleoclimate of Kashmir valley, J&K, India. Ph.D. Thesis, Department of Earth Sciences, University of Kashmir, India.
- Ishtiaq Ahmad and Rakesh Chandra. Geochemistry of loess-paleosol sediments of Kashmir Valley, India: Provenance and weathering. *J. Asian Earth Science* 66, 73-89, 2013.
- B. Bauluz, M. J. Mayayo, C. Fernandez-Nieto, and J. M. G Lopez. Geochemistry of Precambrian and Paleozoic siliciclastic rocks from the Iberian Range (NE Spain): implications for source-area weathering, sorting, provenance and tectonic setting. *Chemical Geology* 168, 135–150, 2000.
- M. I. Bhat, and S. M. Zainuddin. Origin and evolution of Panjal volcanics. *Himalayan Geology* 9, 421-461, 1979.
- A. Bronger, R. K. Pant, A. K. Singhvi. Pleistocene climatic changes and landscape evolution in the Kashmir Basin, India: Paleopedologic and Chronostratigraphic studies. *Quaternary Research* 27, 167-181, 1987.
- B. Buggle, B. Glaser, L. Zoller, U. Hambach, S. Markovic, I. Glaser, and N. Gerasimenko. Geochemical characterization and origin of Southeastern and Eastern European loesses (Serbia, Romania, Ukraine). *Quaternary Science Reviews* 27, 1058–1075, 2008.
- I. A. Farooqi, and R. N. Desai. Stratigraphy of Karewas, Kashmir, India. *Jour. of the Geol. Soc. of India*, 1974.
- S. Gallet, B. Jahn, and M. Torii. Geochemical characterization of the Luochuan loess-paleosol sequence, China, and paleoclimatic implications. *Chemical Geology* 133, 67-88, 1996.
- S. Gallet, B. Jahn, B. Van Vliet-Lanoe, A. Dia, E. A. Rossello. Loess geochemistry and its implications for particle origin and composition of the upper continental crust. *Earth and Planetary Science Letters* 156, 157–172.1998.
- Z. T. Guo. Loess geochemistry and Cenozoic paleoenvironments. *Geoche. News*, 143, 1-10, 2010.
- R. Gardner. Late Quaternary loess and paleosols, Kashmir Valley, India. *Zeitschrift fur Geomorphologie Supplement* 6, 225-245, 1989.
- L. P. Gromet, R. F. Dymek, L. A. Haskin, R. L. Korotev. The North American Shale Composite: its compilation, major and trace element characteristics. *Geochimica et Cosmochimica Acta* 48, 2469–2482, 1984.
- S. K. Gupta, P. Sharma, N. Juyal, and D.P. Agrawal. Loess-paleosol sequence in Kashmir: correlation of mineral magnetic stratigraphy with the marine paleoclimatic record. *Journal of Quaternary Science* 6, 3-12, 1991.
- M. Honda, S. Yabuki, and H. Shimizu. Geochemical and isotopic studies of aeolian sediments in China. *Sedimentology* 51, 211–230, 2004.
- L. Harnois. The CIW index: A new chemical index of weathering. *Sedimentary Geology* 55, 319–322, 1988.
- R. A. Kemp, and M. A. Zarate. Pliocene pedosedimentary cycles in the Southern Pampas, Argentina. *Sedimentology* 47, 3-14, 2000.
- R. A. Kemp, P.S. Toms, J. M. Sayago, E. Derbyshire, M. King, and L. Wagoner. Micromorphology and OSL dating of the basal part of the loess–paleosol sequence at La Mesada in Tucumán province, Northwest Argentina. *Quaternary International* 106/107, 111–117, 2003.
- J. J. Li, J. J. Zhu, J. C. Kang, F. H. Chen, X. M. Fang, D. F. Mu, J. X. Cao, L. S. Tang, Y. T. Zhang, and B. T. Pan. Comparison of Lanzhou loess section of last glacial epoch with Antarctic Vostok ice core. *Science in China Bulletin* 35, 476-487, 1992.
- A. S. Mahjoor, M. Karimi, and A. Rastegarlar. Mineralogical and geochemical characteristics of clay deposits (Central Iran) and their applications. *Journal Applied Science* 9, 601-614, 2009.
- S. M. Moosavirad, M. R. Janardhana, M. S. Sethumadhav, M. R. Moghadam, and M. Shankara. Geochemistry of lower Jurassic shales of the Shemshak Formation, Kerman Province, Central Iran: Provenance, source weathering and tectonic setting. *Chemie der Erde-Geochemistry* 2011.
- H. W. Nesbitt, and G. M. Young. Early Proterozoic climates and plate motion inferred from major element chemistry of lutites. *Nature* 299, 715-717, 1982.

A. K. Singhvi, A. Bronger, R. K. Pant, W. Sauer. Thermoluminescence dating and its implications for the chronostratigraphy of loess-paleosol sequences in the Kashmir Valley, India. *Chemical Geology* 65, 45-56, 1987.

M. J. Soreghan, and G. S. Soreghan. Whole-rock geochemistry of upper Paleozoic loessite, western Pangaea: implications for paleoatmospheric circulation. *Earth and Planetary Science Letters* 255, 117-132, 2007.

S. R. Taylor, and S. M. McLennan. *The Continental Crust: Its composition and evolution*. Blackwell, Oxford, 312pp, 1985.

V. C. Thakur, and B. S. Rawat. *Geological map of the Western Himalaya*. Published under the authority of the Surveyor General of India. Printing Group of Survey of India, 101 (HLO), 1992.

E. Yakimenko, S. Inozemtsev, S. Naugolnykh. Upper Permian paleosols (Salarevskian Formation) in the central part of the Russian Platform: Paleoecology and paleoenvironment. *Revista Mexicana de Ciencias Geologicas* 21, 110-119, 2004.

M. Zarate, R. A. Kemp, M. Espinosa, and L. Ferrero. Pedosedimentary and paleoenvironmental significance of a Holocene alluvial sequence in the southern Pampas, Argentina. *The Holocene* 10, 481-488, 2000.

G. Zhengtang, N. Fedoroff, L. Dongsheng. Micromorphology of the loess-paleosol sequence of the last 130 ka in China and paleoclimatic events. *Science in China D* 29, 468-477, 1996.

U. Zimmerman, and H. Bahlburg. Provenance analysis and tectonic setting of the Ordovician clastic deposits in the southern Puna Basin, NW Argentina. *Sedimentology* 50, 1079-1104, 2003.

LATE CRETACEOUS SYNSEDIMENTARY TECTONIC IN EASTERN ATLAS SAHARAN (NORTH EAST OF ALGERIA)

Benmansour Sana , MAA
University of Batna/Algeria

Abstract:

This study focuses on the relations between sedimentary and tectonic, in the compression context of the terminal Cretaceous of Oriental Atlas Saharan (Tébessa). Paleogeography of the upper Senonian is influenced by the paleo-structure of the platform which determines the nature of the deposits and their geometry. From the tectonic point of view, the activity of the compressive phase eo-alpine increases during this period. This device shows a sedimentary and tectonic instability that accompanies the terminal Cretaceous sedimentation which is confirmed by the presence of synsedimentary structures ("Landslide" slumps, normal faults). The correlation between the East Atlas Saharan and West during the Maastrichtian shows a differentiation in the geometry of deposits. It reflects the variation in subsidence and sedimentation rates between these two areas.

Key Words: Terminal cretaceous, correlation, synsedimentary tectonic, landslide

Introduction

The late cretaceous is one of the most significant periods in the geological history of the Eastern Atlas Saharan in view of the important tectonic and sedimentary events that affected the region at that time.

The present study focuses on the characterization of the sedimentary of late cretaceous deposits under control of tectonic, located in the area of Tebessa, in the north eastern part of Algeria.

The selected site is located in the wilaya of Tebessa, about 17 km from the main town of the wilaya and a 3 km away from the town of Hammamet (Fig.01).

Materials and methods

The section studied is approximately 280 m thick, which are composed mainly of limestone. The survey of the geological section was supplemented by measurements of bedding plans, fault plans, stylolithiques joints. In order to make stereographic projections, We used stéreonet Aug 2000.

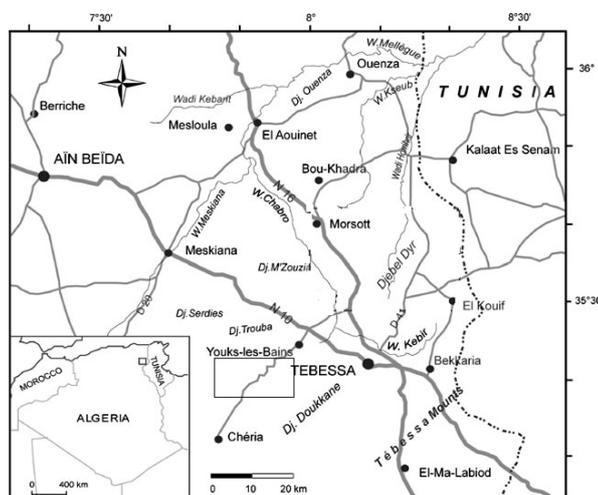


Fig. 1 Geographical location of the site

**Stereographic projection
Bedding plans**

N ^o	Bedding plans	
	Direction	Pendage
01	100	NE 15
02	120	20 NE
03	110	NE 15
04	146	16 NE
05	130	NE16
06	120	NE 15
07	55	NE 20
08	115	NE 30
09	175	NE 15
10	160	NE 30
11	160	NE 52
12	150	NE 12
13	150	NE 20
14	140	NE 10
15	130	13 NE
16	125	10 NE
17	125	15 NE
18	145	15 NE
19	127	10 NE
20	90	20 NE

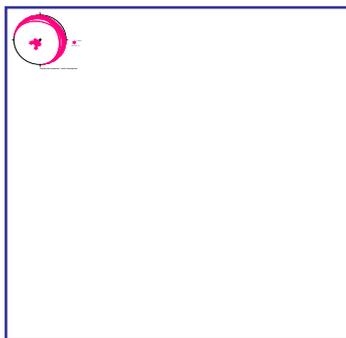


Fig.2A: Stereogram strata plans

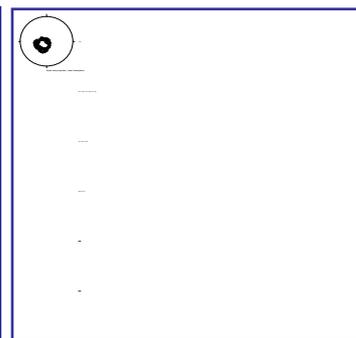


Fig.2B: Stereogram density strata plans

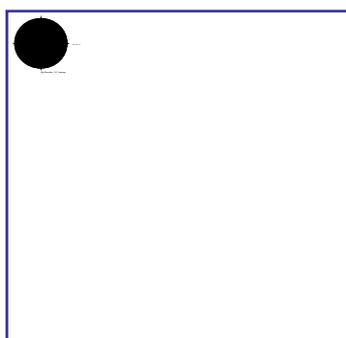


Fig.2C: Stereogram dips strata plans

Faults plans

N ^o	Faults plans	
	Direction	Pendage
01	25	NW 40
02	95	SW dec dex10
03	165	15 SW dec dex
04	95	15 SW dec dex
05	160	10 SW dec dex
06	175	SW dec dex20
07	130	SW dec dex5
08	30	SE30
09	40	20 SE dec sen
10	65	invF. SE67
11	55	SE 25
12	95	SW5
13	160	SW10
14	10	SE90
15	113	SW10
16	120	dex ec d SW5
17	100	SW64
18	160	SW20
19	175	NE25
20	166	NE45
21	118	NE60



Fig.3A: Stereogram strata plans

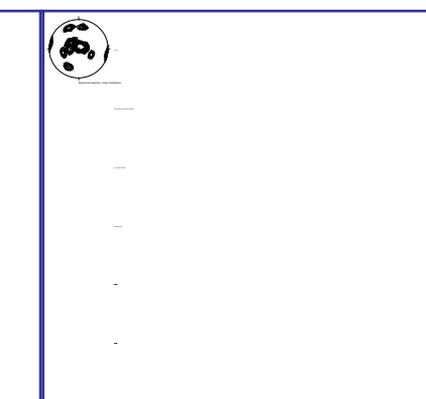


Fig.3B: Stereogram density strata plans

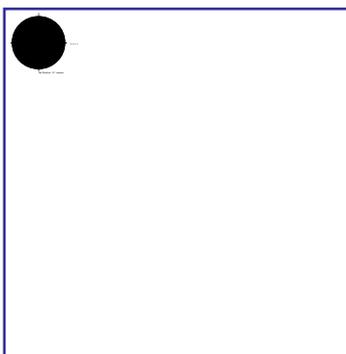


Fig.3C: Stereogram dips strata plans

Diaclasis plans

diaclasis plans		
N°	Direction	Pendage
01	45	SE90
02	155	NE85
03	70	SE85
04	180	SE85
05	127	NE85
06	45	SE85
07	60	SE90
08	165	NE90

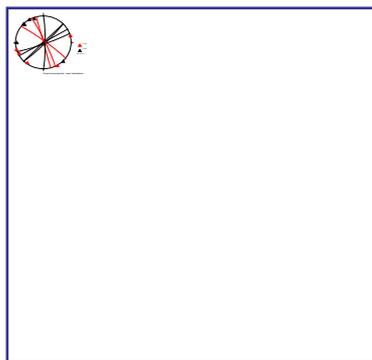


Fig.4A: Stereogram strata plans

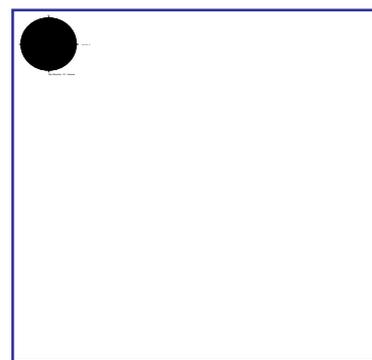


Fig.4B: Stereogram density strata plans

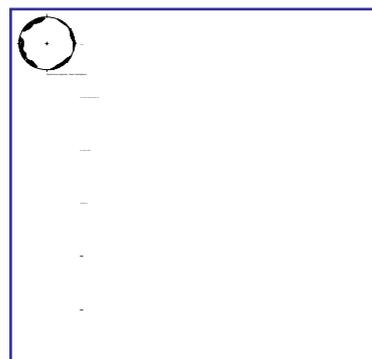


Fig.4C: Stereogram dips strata plans

Stylolites

stylolithes		
N°	Direction	Pendage
01	118	40 SW
02	25	75 SE
03	30	60 SE
04	120	35 SW
05	40	80 SE

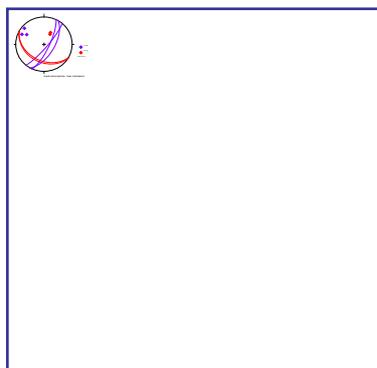


Fig.5A: Stereogram strata plans

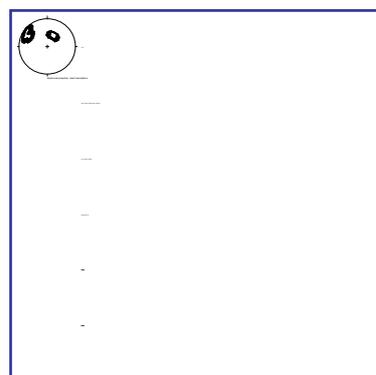


Fig.5B: Stereogram density strata plans

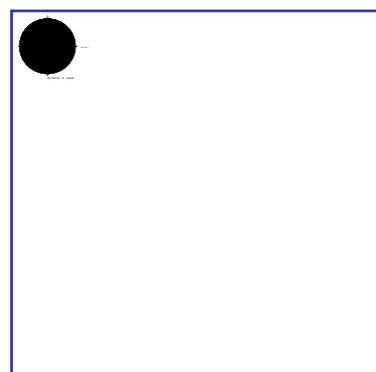


Fig.5C: Stereogram dips strata plans

Interpretation of stereographic projections

Stereograms (Fig: 2A, 3A, 4A, 5A) are projections of Schmidt, lower hemisphere with fault plans, bedding plans, and diaclasis plans, stylolites (thin lines) and ridges (small arrows, centrifugal normal cheeks). The extension direction reconstructed by various methods (Angelier, 1984) is indicated by the arrows on the edges of large diagrams.

- The density stereograms are Dimetrevic projections (Fig. 2B, 3B, 4B, 5B). Histograms (black) (Fig. 2C, 3C, 4C, 5C) are projections of Wulf summarize the distribution of dip directions.

A-Plans strata: The stereogram of bedding plans (Fig: 2A, 2B, 2C) shows the main direction:

D: N 130 ° E, **P:** 10 ° NE. That direction does not correspond to the Atlas phase (NW / SE), or the Alpine stage (N-S). We can consider that this direction has undergone a deformation resulting from the combination of Alpine tectonics and paleotectonic (accident base) which would lead to virgation of Dj. Gaâga.

B-faults Plans: From stereogram of fault planes (Fig. 3), three mean directions are mentioned:

- **D1:** N 40 ° E, P: 20 ° SE (black thin lines);
- **D2:** N 120 ° E, P: 5 ° W (red lines);
- **D3:** N 175 ° E, P: 40 ° S (blue lines).

The first two directions correspond to steps (corresponding to the first and second recesses sinistral to dextral offsets). They are the result of shortening NS direction (phase alpine) The third direction is normal faults, they are always the result of a NS shortening (phase alpine).

C-joints Plans: From stereogram plans joints (Fig. 4), two main directions are determined:

- **D1:** N 50 ° E, P: 85 ° SE (red thin lines);
- **D2:** N 150 ° E, P: 85 ° NE (black thin lines);

These two directions result of shortening direction N-S (Alpine phase).

D-joints stylolites: In the stereogram plans stylolites (Fig. 25), mentions two main directions: **D • 1:** N 30 ° E, P: 80 ° SE (blue thin lines) corresponds to the direction of tectonic stylolites (perpendicular to the stratification). They are the result of shortening NW / SE (Phase Atlas).

- **D 2:** N 125 ° E, P: 25 ° SW (red thin lines). This direction represents Stratiform stylolites (diagenetic).

4 - The synsedimentary tectonics: In order to show the tectonic / sedimentation relationship, compression context Cretaceous of Eastern Atlas Saharan and specifically in the area of Gaâga (Tebessa) was studied and analyzed synsedimentary structures recorded. Measures levied on land allow us to highlight the existence of slumps (Plates 6, 7, 8,9) and growth faults (Fig:6,7,8,9). The slip plans are observed slumps oriented NE. Synsedimentary normal faults are steering NW / SE. These directions correspond to the episode extensional phase Laramide compressive direction NW / SE.

Conclusion

In conclusion, it appears that the region of Gaâga results from the combined action of several tectonic style and variable extensions. Management stratification plans (N 130 ° E) of Dj. Gaâga does not correspond to the Atlas phase (NW / SE), or the Alpine stage (N-S). We can consider that this direction has undergone a deformation resulting from the combination of Alpine tectonics and paleotectonic (accident base) which would lead to virgation of Dj. Gaâga. Accidents caused by Atlas phase and Alpine phase, show very different directions, and their distribution is not homogeneous throughout the cut. The generated Alpine phase form:

- 1 - normal faults direction N 175 ° E;
- 2 - the sinistral offsets (N 40 ° E);
- 3 - of dextral offsets (N 120 ° E);
- 4 - diaclasis network combined (N 50 ° E and N 150 ° E).

Atlas Phase is represented in the form of tectonic stylolites. Their direction is N 30 ° E. The stratiform stylolites (diagenetic) are thus determined. Their direction is N 125 ° E. The synsedimentary tectonics, well represented, has generated slumps and normal faults. These structures correspond to extensional episode of the compressive phase of the Maastrichtian.

References:

Angelier J. (1984).-Tectonic analysis of fault slip data sets. *Journal of Géophysical Research*.
Benmansour S. (2009).- Etude tectono- sédimentaire du crétacé terminal de l'atlas saharien oriental ; p 56-62.

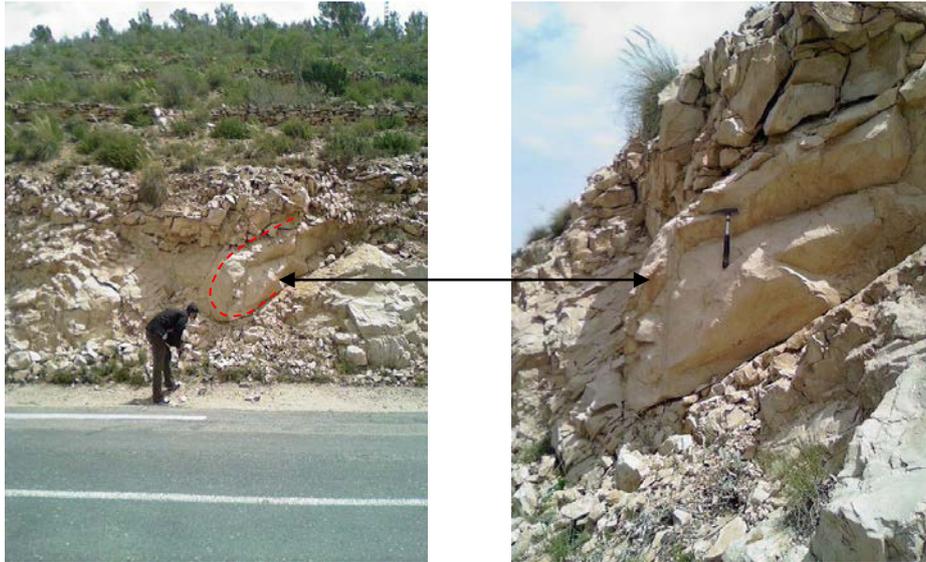


Fig. 6: Slump



Fig. 7: slump



Fig. 9: Normal synsedimentary fault

GEOLOGICAL AND GEOTECHNICAL CHARACTERISTICS OF THE SOILS IN THE REGION OF SÉTIF

Djenba Samir

Department of Earth Sciences University of Batna

Abstract:

Given its strategic location, the wilaya of Setif is the crossroads between the IS and the center of one side and the coastal towns of South and other side. it occupies the second round of a national demographic point of view and a large socioeconomic pole, hence the expansion into new land for urban or industrial activities or for the inevitable.

The majorises of this expansion is focused on the meridian Sétif El Eulma due to the topography and other predisposing infrastructure already interesting recently reinforced by the East-West. Several geotechnical studies have been done on this area has been chosen four different sites geographically distant to characterize these deposits.

Éternitaire three sites on the east-west highway, it is the points of interchanges and toll point Near the town of El eulma and agglomeration Ain sfiga and near the town of Ain Arnet, the fourth site is the foundation of a wastewater treatment plant Bazer Sakhra.

The approach in this study is geological soil description of Setif, we also tested through several geotechnical brought to the region, to characterize these soils by hand is an in situ testing and secondly by testing the laboratory.

This study has shown that the Setif high plains generally consist of deposits by fluviolacustres Mioplioquaternaire age, are from top to bottom of arable brown silty clays, limestone crusts villafranchiennes and red clay.

Geotechnical data have shown that the soils are classified into two categories:

O 2 meters sand gravel BC, inferior to 2 meters considered stiff clays Class B.

Key Words: Soil mioplioquaternaire, plasticity, sétif

Introduction

The area of study is located in the East of Algeria (Fig.1).

Setif high plains are vast deposits composed mainly by mioplioquaternaire. Several geotechnical studies have been made on this area.

We chose four different sites geographically distant to characterize these formations.

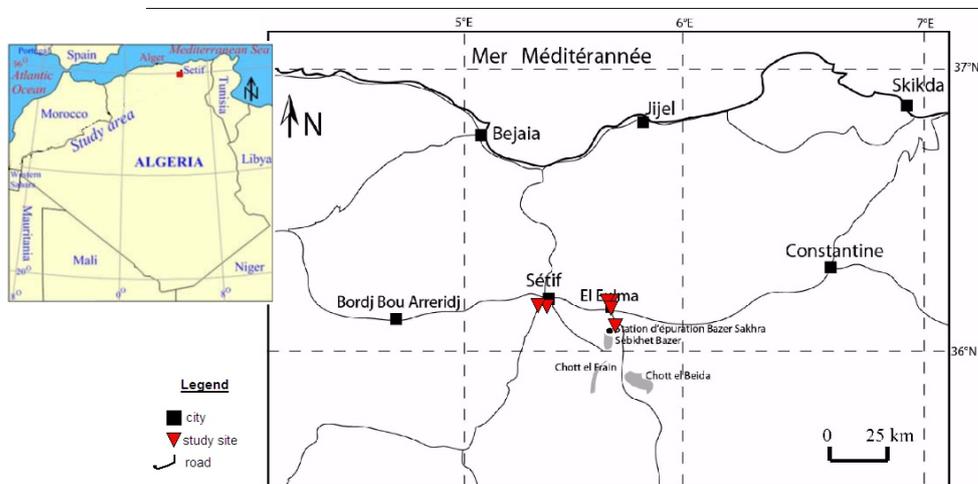


Fig.1 position of study sites

Éternitaire three sites on the east-west highway, it is the points of interchanges and toll points Near the city of El eulma and agglomeration Ain sfiga and near the town of Ain Arnet, the fourth site is the foundation of a wastewater treatment plant Bazer Sakhra.

In this paper we puts review the geological soil Sétif, moreover on trial through several points of geotechnical this region, characterize these soils by in situ tests and other laboratory tests by . Opportunities for future correlations between these parameters are underway we will later.

We began the study site by site and concluding with a comparison between the pressuremeter different sites.

2-Geology of the area:

Setif high plains, which reports the study area, are between the Saharan Atlas to the south and the north dome Setif. From a geological point of view they belong to the external domain of the Alpine orogeny of eastern Algeria (Fig. 2).

And local geology is summarized as follows:

Soils are generally sandy to clayey in texture and mostly classified as Aridisol and are calcareous. Mineralogically, most of the soils are dominated by kaolinite, illite, smectite, and chlorite—typical for most arid and semi-arid soils.

Rocks and unconsolidated deposits in the area can be divided into three geologic units: (1) upper Cretaceous (Senonian); (2) Eocene; and (3) Mio-Plio-Quaternary. Senonian (upper Cretaceous) is generally found in the northern part of the study area. Senonian units are composed of Santonian-Campanian formation and upper Senonian formation. These formations consist of various rocks with differing compositions including limestone and marl of about 550m thick. Eocene units are composed of Ypresian–Lutetian formation (Fig.2). Eocene rocks consist of a succession of marine, limestone, and silt of about 80m thick. The Mio-Plio-Quaternary is a heterogeneous continental detrital sedimentation;

The miopliocène is composed of sands, silts and clays, reddish gravel and conglomerates;

A calcareous crust of variable thickness up to 1 m doubled sometimes attributed to villafranchien;

Finally, a recent Quaternary unknown are:

Alluvium and old arable land: arable land are installed on veneers loamy brown soil and support little or no limestone.

Current or recent alluvial deposits: it is sand, silt and gravel developed along the wadis.

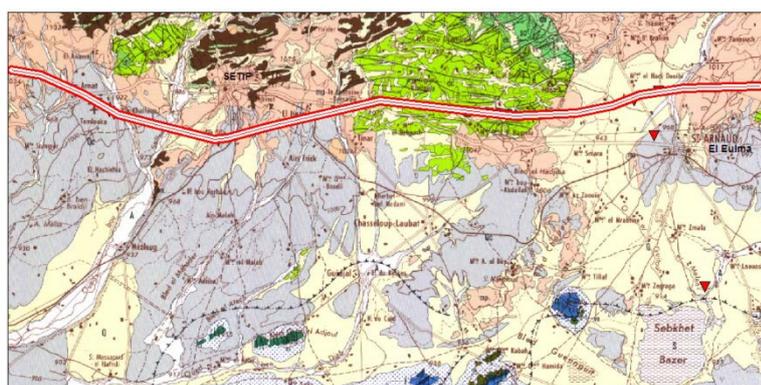


Fig2: Geology map of the studied area and site location

Legend :

	Quaternary		autoroute
	Calcareous crust		site location
	Mio-Pliocene continental		
	Ypresian and Lutetian (limestone)		
	Upper Senonian (carbonte)		
	Santonian-Campanian		
	Upper Jurassic (limestone)		

Station Bazer Sakhra

Station wastewater studied is located àBazer Sakhra agglomeration distant 5 km south-east of the city of El Eulma (Wilaya de Setif) (Fig 1). the site area is about 11 hectares.

a. organization

Recognition in situ soil was established on the basis of:

12 core drilling rig using a DB850. Depth of core drilling up to 25 m. tests SPT (Standard Penetration Test) 13 pressuremeter tests carried every 2 meters in auger drilling ϕ 63 mm conducted to 23m;

5 open wells depth of 3 m each, made using a hydraulic shovel. Finally samples were paraffined and not subject to the testing laboratory.

Synthesis of results

These operations have enabled us to recognize the nature of the different layers of soil and review of logs and core drilling augers let us distinguish:

- Soil cover: in all core holes, topsoil 0.40m thick overcomes the modern alluvium.
- The modern and recent alluvium: they are represented by brown silty clays and clays yellowish black. These alluvial deposits are developed to a depth of 2.00 to 3.00 m.
- The old alluvium: clay are gravelly to clayey and have a thickness of 0.30 to 1.50 m.
- Substratum: we meet the substratum clay Miopliocène age from 4 m.

We note the presence of water in our core drilling at shallow depths between 1.55 and 1.90 m. Pressuremeter test results obtained show that the pressuremeter characteristics are regular and tight. Layers tested can be divided into two parts:

Deposits of top: soils cover, modern alluvium, old alluvium, and the first two meters of the underlying clay

Underlying clay formations that proved moderately compact fig1.

c-analyses of physical results

The fine fraction (0.08 mm) remains dominant with over 85% of the granular composition consisting mainly of clay and silt. Regarding therefore a soil fines content is 85% and the three granular fractions are 36% clay, 35% silt and 14% for sands, we consider the point of view of granularity, the soil in depth is clay. The mean and standard deviation of these components are shown in Table 2. The water content ranges between 14 and 26.5% with an average of 20.3%. Degree of saturation between 92.3 and 100% except item SC9 3 m depth where a $S_r = 20.3\%$, this same point has a elevated dry density compared to all this that seems to be gravelly sand. Such thresholds characterizing a saturated state.

The dry density value measured on the same samples takes place in a range of 1.5 to 1.8 with an average of 1.71 t / m³. This leads to a wet density of 2 t/m³.

According to the standard geotechnical (XP P 94-011), the clay is in the family of dense soils.

According to the plasticity chart (Fig. 3), it is indeed a highly plastic clay (At), of hard consistency.

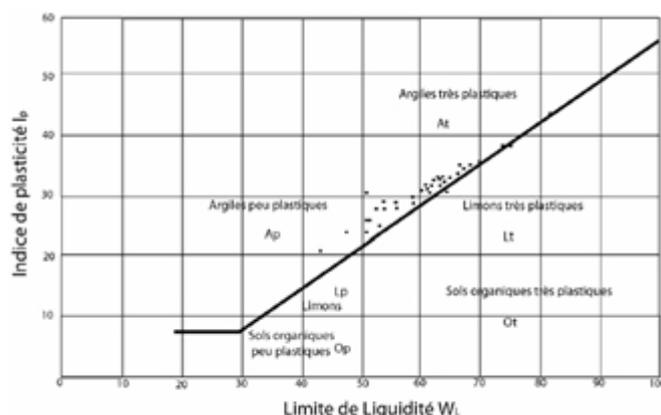


Fig 3:plasticity diagramm

d-analyzes mechanical results:

One shear test, unconsolidated undrained gives us a value of friction angle equal to 8 ° with a cohesion of 0.95 bar. They check the quality of the clay. It therefore appears that the soil tested has good cohesion.

after test data œnométriques ;the soil is classified in the category of normally consolidated soil to consolidated compressible ,non-swelling and swelling ,compressible medium in a few points.

For clays miopliocène:

The fine fraction (0.08 mm) remains dominant with over 91% of the granular composition consisting mainly of clay and silt, the water content ranges from 17.6 to 32.1% with an average of 23 4% 1, the degree of saturation is 89 to 100% which means a saturated soil.

The dry density value measured on the same samples takes place in the range from 1.4 to 1.8 with an average of 1.63 t / m 3. Which leads to a wet density of 2 t/m3. According to the standard geotechnical (XP P 94-011), the clay is in the family of dense soils.

According to the plasticity chart (Figure 5), it is indeed a highly plastic clay (At).

For shear testing was a slight decrease in cohesion provided drained;

The shear tests, unconsolidated undrained gives us a value of friction angle equal to 9 ° with a cohesion of 0.72 bar. They check the quality of the clay. It therefore appears that the soil tested has good cohesion. Test data œnométriques classify the soil in the category of normally consolidated soil to consolidated compressible compressible medium non-swelling and swelling in few points.

station Ain Sfiha

Sfiha Ain is a small town south of Sétif, two studies specific to the highway here on the subject, it is a highway junction and a point of payment.

In total three holes pressuremeter up to 30 m deep core drilling four mechanical jusqu'un 25 m from two wells and shovel 3.5 m intact samples and redesign were also taken at different depths for tests laboratory.

Synthesis of results

From the surface to a depth of 30 can be distinguished (fig: 4)

- Land cover: Land végétale 30 to 45 cm,
- The modern and recent alluvium: they are represented by brown silty clays and clays yellowish black.
- The old alluvium: clay are gravelly to clayey in some places covered by crusts. These Quaternary formations have a thickness of 3 m
- Substratum: we find the clay substratum Miopliocène age from 3 m. it is occasionally very plastic red clay encrusted gravel overlying clays.

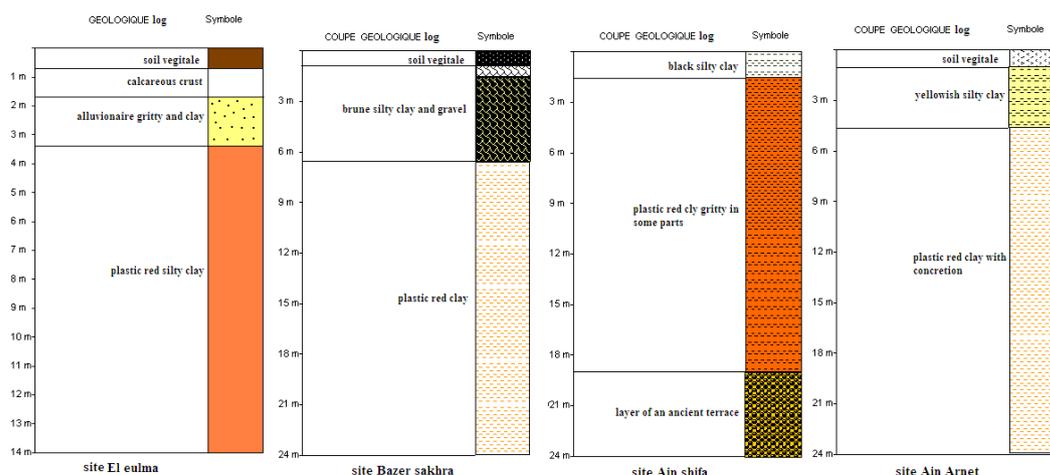


Fig 4: site lithological's colonne

Pressuremeter tests give mean values for the first four meters to the underside become higher with 21m net values, reflecting the well stratigraphy determined by core drilling.

c-analyzes of the results of laboratory tests:

The fine fraction (0.08 mm) remains dominant with over 85% of the granular composition consisting mainly of clay and silt. We therefore, a point of view granularity, the soil in depth is clayey (tab 5a-b), its density is 1.45 to 1.78 t/m³,

the water content ranges between 17 and 30% with an average of 23.6%. Degree of saturation between 84 and 96% thresholds characterizing such a saturated state.

According to the diagram of plasticity, it is indeed a highly plastic clay (At), of hard consistency.

Mechanical characteristics give a value of friction angle equal to average 22.5 ° and cohesion of 0.126 bar to 0.929 bar respectively classifying them the soil medium to stiff consistency.

test data œnométriques classify the soil in the category of consolidated normally consolidated compressible medium compressible, non-swelling swelling.

station Ain Arnet

Located west of Setif, two studies specific to the highway here on the subject, it is a highway junction and a point of payment.

In total three pressuremeter up to 25 m deep, three core samples from mechanical jusqu'un 20 m and four wells shovel 3.5 m deep, intact samples were intact and not also taken at different depth for laboratory tests.

Synthesis of results

From the surface to 25 m depth can be distinguished:

- Land cover: Land végétale 30 to 45 cm,
- The modern and recent alluvium: they are represented by brown silty clays black to yellowish.

- crust, clay, gravel and clay tuff these Quaternary formations have a thickness of 3 m
- Substratum: we meet the substratum clay Miopliocène age from 3 m. it is occasionally very plastic red clay crusted.

Pressuremeter tests give mean values for the first four meters to the underside become higher with net values to 21m deep, well reflecting the stratigraphy determined by core samples.

c-analyzes of the results of laboratory tests:

Deposits of the surface are characterized by variable percentages (91% 15to 0.08 mm)

This fine fraction has a plasticity index of 16 to 30%.

Clays underlying the fine fraction (tab 2) remains dominant with over 83% of the granular composition consisting mainly of clay and silt. We therefore consider a granularity point of view, the soil in depth is clay, its dry density of 1.81 to 1.99 t/m³,

The water content ranges between 21 and 34% with an average of 25.5%. Degree of saturation between 84 and 96% thresholds characterizing such a saturated state.

The plasticity range is very large, it is indeed a highly plastic clay to hard consistency.

The mechanical properties provide a value quite variable friction angle of 8 ° to 44 ° and a cohesion of 0.44 bar to 0.926 bar respectively classifying them the soil medium to stiff consistency.

According to data from oedometer tests the soil is classified as normally consolidated soil on consolidated compressible quite compressible, non-swelling swelling.

Tab 1a: means values of the physical characteristics of clays miopliocene								
site	(Wn %)	γd t/m ³	(Sr %)	2mm	80μ	WL%	WP%	IP%
Bazer sakhra	23,47	1,63	95,62	91,54	84,15	60,73	29,57	31,66
Ainsfiha	23,66	1,59	92	91	84,83	72,64	34,32	38,35
Ain Arnet	25,5	1,54	91,75	88,85	83,14	75,33	35,98	40,61
El eulma	21,66	1,64	95	87	82	60,48	30,2	30,06

Tab 1b: means values of the mechanical characteristics of clays miopliocene							
site	Pc	Cc	Cg	Cuu (Bars)	φuu°	Ccd (Bars)	φcd°
Bazer sakhra	1,214	0,105	0,045	0,72	9,16	0,52	16,5
Ainsfiha	2,993	0,213	0,05	0,504	22,5		
Ain Arnet	2,785	0,186	0,034	0,621	26		

Tab 2: means values of the physical & mecacanical characteristics of quaternary												
site	2mm	80μ	WL%	WP%	IP%	(Wn %)	γd t/m ³	(Sr %)	Pc	Cc	Cg	
Bazer sakhra	85,53	65,84	59,21	28,97	30,18	20,3	1,71	86,44	1,03	0,09	0,02	
Ainsfiha	97,5	91,5	66,69	31,69	35	23,5						
Ain Arnet	59,88	58,12	43,11	21,28	21,83							
El Eulma	49	35	39,66	18,63	21,03							

station El eulma

Located 27 km EAST of Setif, three studies, two for highway and headquarters CASNOS.

In total 17 points pressuremeter 8m to 14 m depth, core sampling DE19 m deep wells and two shovel 3.5 m deep, intact samples were intact and not also taken at different depths for testing the laboratory.

Synthesis of results

Lithological same configuration as the other sites.

c-analyses of the results of laboratory tests:

We have a single core sampling and two wells, testing of physical properties have been made.

From top to bottom in meeting:

alluvial deposits, or crusts with sand fraction values of 31 to 67% and the clay fraction of 18-52%. The plasticity index of 20% A22.

-red clays have a water content of 20-24%, a saturation of 96%, the fine fraction is quite present with an average of over 82% a plasticity index of 25 to 35% (TAB7). and end a dry density of 1.6 to 1.65 t/m³, it is indeed a highly plastic clay saturated more or less dense. pressuremeter tests gave the mean values of limit pressure of 23 bar for first meter and 17 to 18 bar for two meters follows then increased with depth, it clearly explains the lithology of the region is generally villafranchiennes calcareous a flagstone near the surface are then tuff surmounting silty clay reddish clay.

We comparing the different profiles presiometriques (fig4) all sites shows that they generally characterize three successive horizons from top to bottom, it is recent alluvium, and crusting villafranchien red clay with lenses of conglomerate Mioplicène;

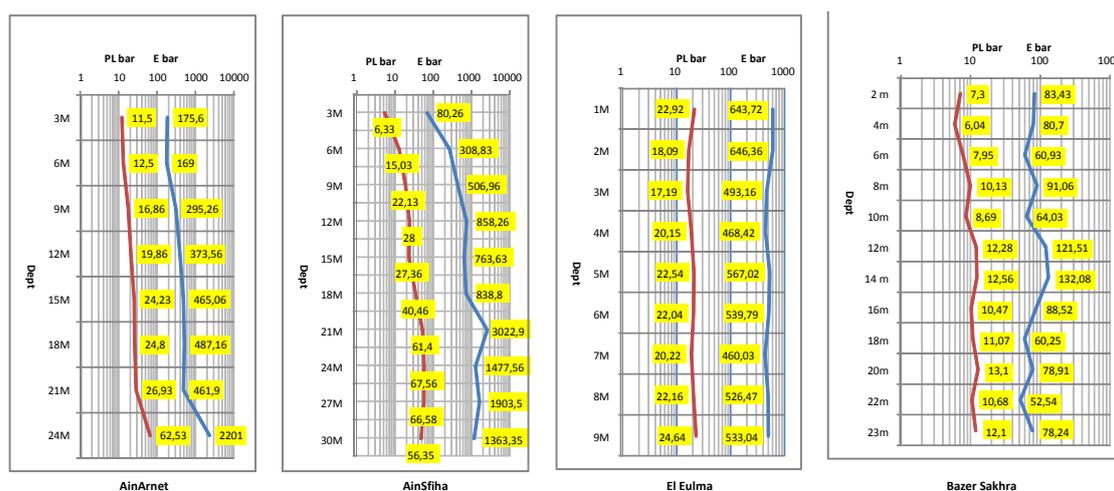


Fig7 :Parameter values of pressiometric test

The values of the pressure limits of El eulma, Ain Ain and Sfiha Arnet are homogeneous in the first ten meters and grow significantly in depth and Ain Ain Sfiha Arnet this is due to the state of this highly plastic clay and the presence conglomeratic beds also very dense.

low values of Bazer Sakhra (relative to other sites) is probably due to the presence of groundwater close to the surface.

Conclusion

The soils of the Setif region generally consist of three layers: a brown clay silt of Quaternary age, a calcareous crust thickness up to 1 meter villafranchien age and finally red clay encrusted in some places with depth, gravel and sands beveling mioplicène age, these deposits have been disposed in an environment fluiviolaustre.

Geotechnical studies made at different points in the region are character appear moderately dense plastic clay slightly compressible, however it is reported that clays in the region of Ain Arnet have a character swelling.

Pressure limits to them and they reflect the lithology encountered in core samples logs.

References:

- Anonymous soil study report. L.T.P.Est, Sétif. Unpublished report.
- Anonymous report study and realization of sewage wastewater from the city of El Eulma.
- Chenaf A: Missions geotechnical geotechnical feasibility study for the realization of sewage wastewater El Eulma (Bazer Sakhra Site, Wilaya of Sétif), memory ingeniorat 2007 universite Batna.
- AFNOR NF P 94-500 STANDARD (2006). French Association for Standardization.
- Philipponat G. and Hubert B., (2002). Foundations and earthworks. Eyrolles, 300p.
- Savornin J. (1947). Detailed geological map of Algeria 1:50,000, sheet Navarin with instructions.
- Schlosser F. (1997). Elements of soil mechanics. Presses de l'ENPC, Paris. 280 p.
- Vila J.M. (1977). Detailed geological map of Algeria 1:50,000, sheet Bir el Ahrech (eg Navarino) el Eulma and Setif, with explanatory notes.

TOWARD INDONESIA'S AGROINDUSTRIES COMPETITIVENESS: THE CASE OF BIOETHANOL DEVELOPMENT FROM SUGARCANE BASED INDUSTRIES

Gita K. Indahsari

Graduate School of Management and Business, Bogor Agricultural University, Indonesia

Rudi Wibowo

Agricultural Economics Professor at Jember University,
Lecturer of Business Management At Bogor Agricultural University

Abstract:

World oil price has been fluctuating throughout the years, however the trend continues to increase. Due to limited sources of oil production and high demand for world oil, it is predicted in ten years from 2010, world oil price could even reach as high as triple digit per barrel. Countries around the world had continuously tried to produce alternative energy, including Indonesia. The Indonesian government had tried to develop this industry, with roadmaps and targets for alternative energy production, reduce dependency on fossil fuel until the next ten years to come. The objective of this paper is to study the bioethanol industry developed by PTPN X. The bioethanol industry is still at its early stage in Indonesia. Problems to develop this industry includes finding the most potential raw materials used, availability of the industry, stage of competition with other uses of the raw materials until end product and marketing. From many potential raw materials, sugarcane is one of the most potential source of bioethanol. PTPN X case study shows how bioethanol can be competitive if there are supporting factors such resources, technology and company size. The comparative advantage of PTPN X is the ability of PTPN X to incorporate problems from upstream to downstream of the supply chain of bioethanol. Methods such as the EPC, ANP, Diamond Porter Model and case study of PTPN X are used to demonstrate the potential of sugarcane as an alternative energy and the ability of PTPN X to develop a competitive agroindustry.

Key Words: Competitiveness, bioethanol, PTPN X

Introduction

As Indonesia became the world's fourth most populous country in the world in 2011, the demand for fossil fuel continues to increase. Indonesian have been used to the cheap oil price, government policy which heavily subsidized this sector had become a boomerang for the government, causing the country to be a net importer for fuel. In 2005, Indonesia along with other countries in the world started to find other alternative energy sources, such as biofuel. As a renewable energy from plants, biofuel could be categorized into biodiesel and bioethanol. Bioethanol are eco-friendly energy which are derived carbohydrate and sugar components from plants, such as sugarcane, sweet sorghum, cassava, sweet potato and also sorghum. A study carried out by Indahsari *et al* (2012) shows that the current bioethanol industry in Indonesia has been stagnant. Problems arise from upstream until downstream of the bioethanol supply chain. Potential problems include availability of raw materials, competition with other plants, marketing and consumption.

Finding the most potential type of raw materials for bioethanol is also not an easy challenge, because the plant needs to contain either starch or sugar to later be fermented. Potential raw materials have direct competition use with other industries. Examples are cassava which can also be used for tapioca, corn for feed and fed, sugarcane with black ketchup industries. Thus, bioethanol products could also be processed as ethanol, such as alcohol, either for the pharmacy uses or for the beverages. Ethanol export products are also priced higher. Therefore, in order to develop the bioethanol industry,

there should be other factors to induce the development. An example is that the company does not only produce mainly bioethanol but could also the waste for other uses, such as fertilizer.

Considering the depletion of oil reserves of Indonesia, the government through Government Regulation No. 05/2006 targeted to be able to replace the 1.48 billion liters of gasoline with bioethanol for the period 2007-2010. where the figure percentage was planned to be increased by ten percent in 2011-2015, and 15 percent in 2016-2025 (Ircham, 2008). Figure 17 shows previous government plans to develop the bioethanol sector in Indonesia. The plan was targeted for 25 years of program development which had commenced since 2000 (BPPT, 2000). Bioethanol was part of the renewable energy plan which was concluded in the general energy selection by the government. Figure 1 shows the country's national plan to reduce dependence on imported fossil fuel.

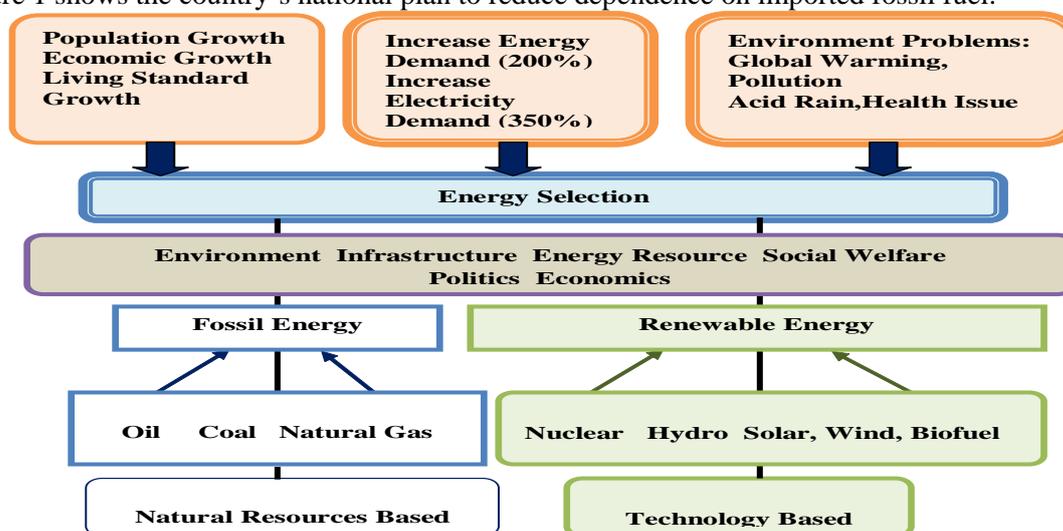


Figure 1 National Energy Plan, 2000 - 2025 (BPPT, 2005)

Figure 1 shows how potential renewable energy, including bioethanol, is. In order to fulfill demand for energy, government had planned to develop the renewable energy. It has been predicted that fossil fuel in Indonesia will be depleted in 10 to 20 years to come, while imported fossil fuel has increased to 30 percent and continues to grow. Therefore, the government has tried to develop other alternative energy strategy.

In 2005, the Indonesian government has tried to develop the bioethanol industry. Figure 2 shows the government's target for national energy mix for the year 2025. The government intends to increase the portion of alternative energy from 0.1% to 0.2% to 0.4% in 2025. Bioethanol as a part of the biofuel program was targeted at 1.335% for the whole program in 2025, which compared to other energy program, is high (solar energy 0.020%, wind power 0.028%, fuel cell 0.000%). Only nuclear target energy power which has a higher percentage value (1.993%) compared to biofuel. In 2006 Indonesian government policy to reach 2 percent of biofuel in the National energy consumption by the year 2010 with 5.29 million kilolitres at target and at least 5 percent by the year 2025 with 22.26 million kilolitres

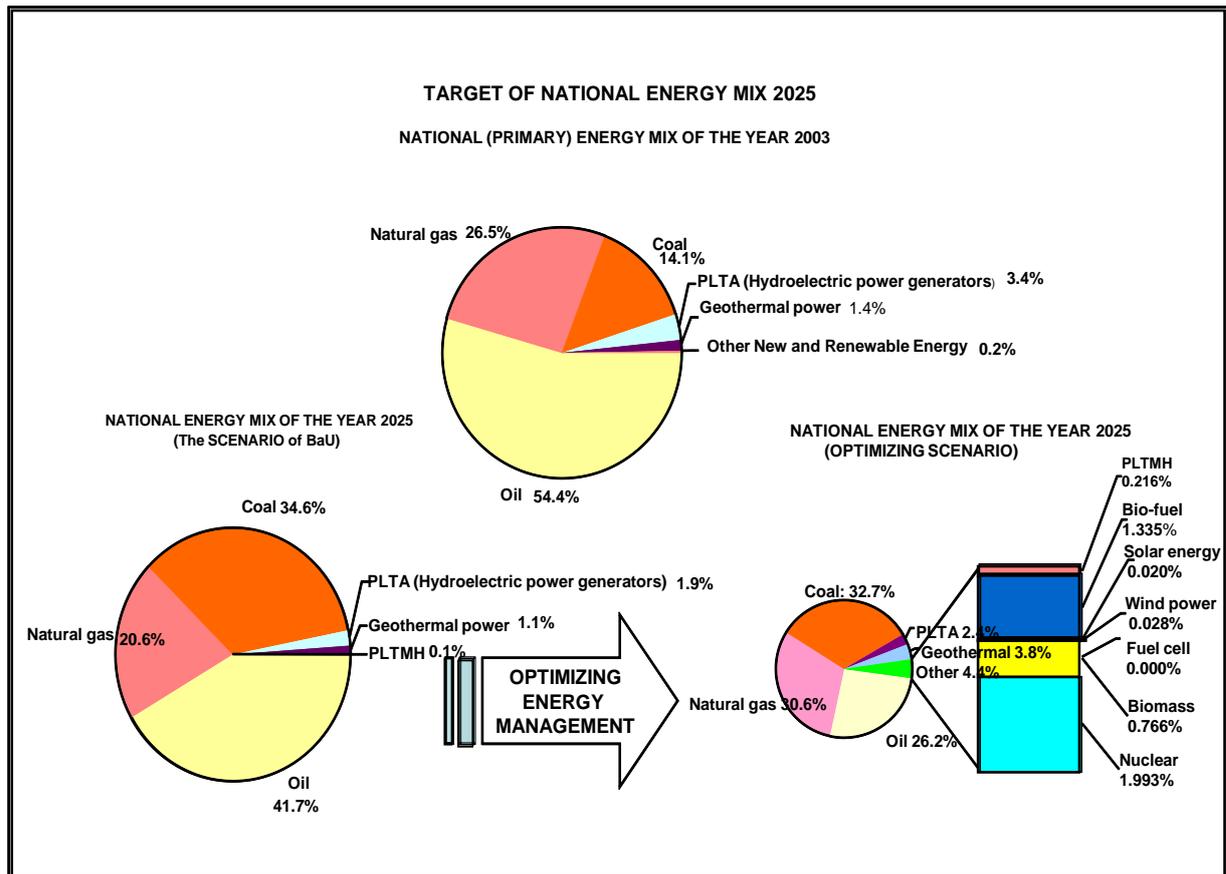
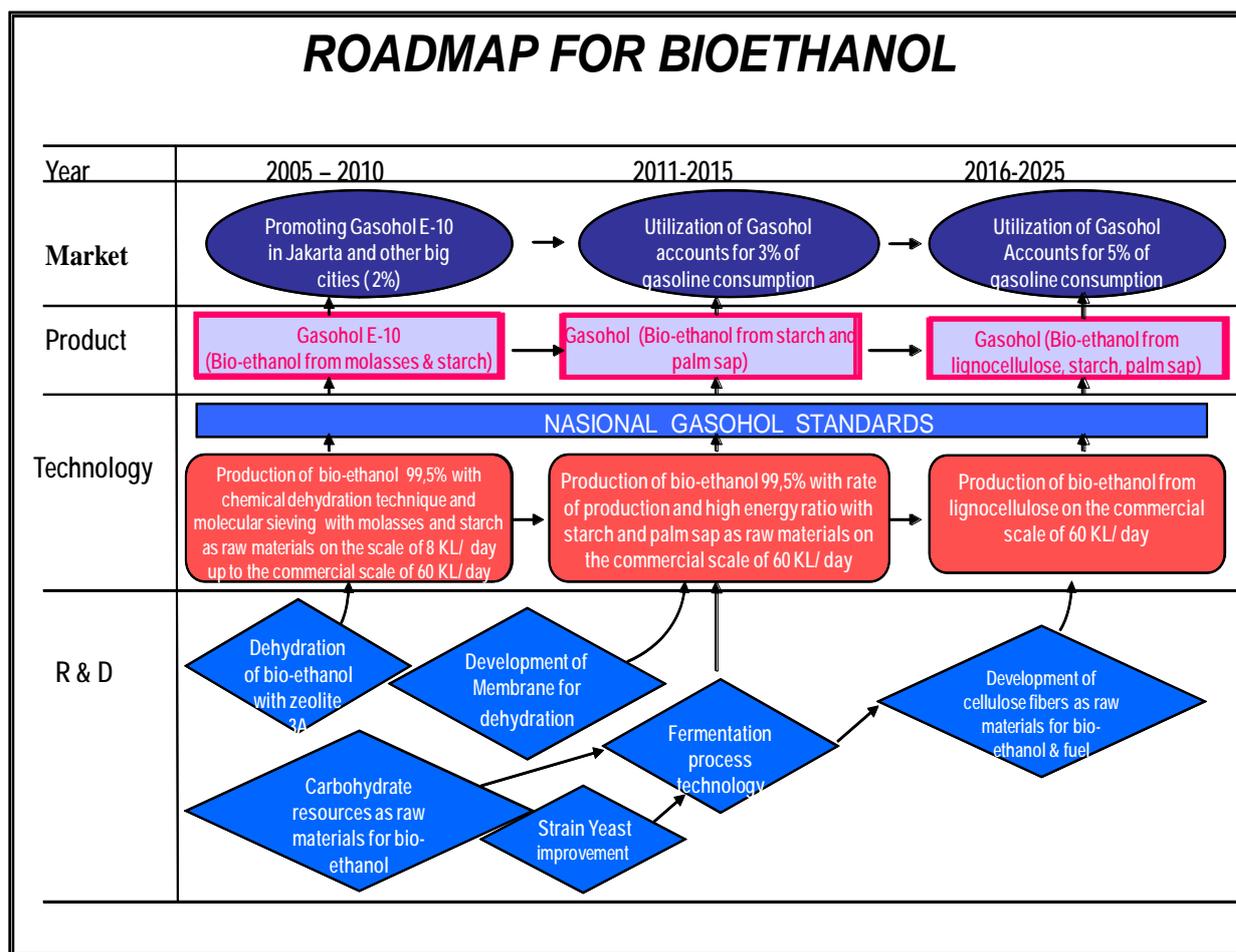


Figure 2 Government's Target for National Energy Mix 2025 (ESDM, 2006)

The government had planned to increase bioethanol uses until the year 2005, from 1% per year to around 5% in the year 2025. The government had also planned to increase the number of gasoline station to sell bioethanol from East Java, to whole part of Java island and later to Sumatra. The aim is to introduce and familiarized people with bioethanol as a substitute for fossil fuel.

The road map for gasohol or mixture between bioethanol and gasoline can be seen in Figure 3. The government had planned to increase the number of bioethanol plants from 17 plants in 2006 with a capacity of 60kL per day to a target of 25 pants in the year 2016.



55

Figure 3 Indonesian Government Roadmap for Bioethanol (ESDM, 2006)

In 2008, the government planned to utilize 5 to 5.5 million hectares of cassava and 750 thousand hectares of sugar cane to increase the productivity of these alternative fuels. The government expected to add a minimum budget of Rp 100 billion, of which the interest subsidy of Rp 1 trillion will be allocated for the farmers as well as Rp 2 trillion for the initial capital of establishing financing institutions for biofuel development (Ircham, 2008)

The government is aware of the need for a replacement of fossil fuels, as proven by the various policies that have been issued by the government to support the development of national biofuel industry. Therefore it is important to identify factors that influence the competitiveness of industry in the development of biofuel in order to avoid the worst possibilities when Indonesia would have to import biofuels in the future.

Data And Methodology

Study conducted for a year in 2010 and also in 2011. This study uses primary and secondary data collected from the experts, interview with various sources, including private companies, government officers, association and Universities. The secondary data were used to determine the availability of the raw materials before and after it is used as bioethanol as well information related to the bioethanol industry.

Porter Diamond Model

According to Porter (1991), the competitiveness of an industry could be developed through the interaction of several factors, such as the demand factor condition, firm's strategy, structure and rivalry, related and supporting industries and the demand factor (Figure 4). The Diamond Porter Model investigates the strength and weakness of an industry and how competitive is the industry. However, opportunities and chances, along with the government play important roles to the overall competitiveness of the industry.

Figure 4 describes how the Diamond Porter Model will be used to determine the competitiveness of the bioethanol industry, related to PTPN X.

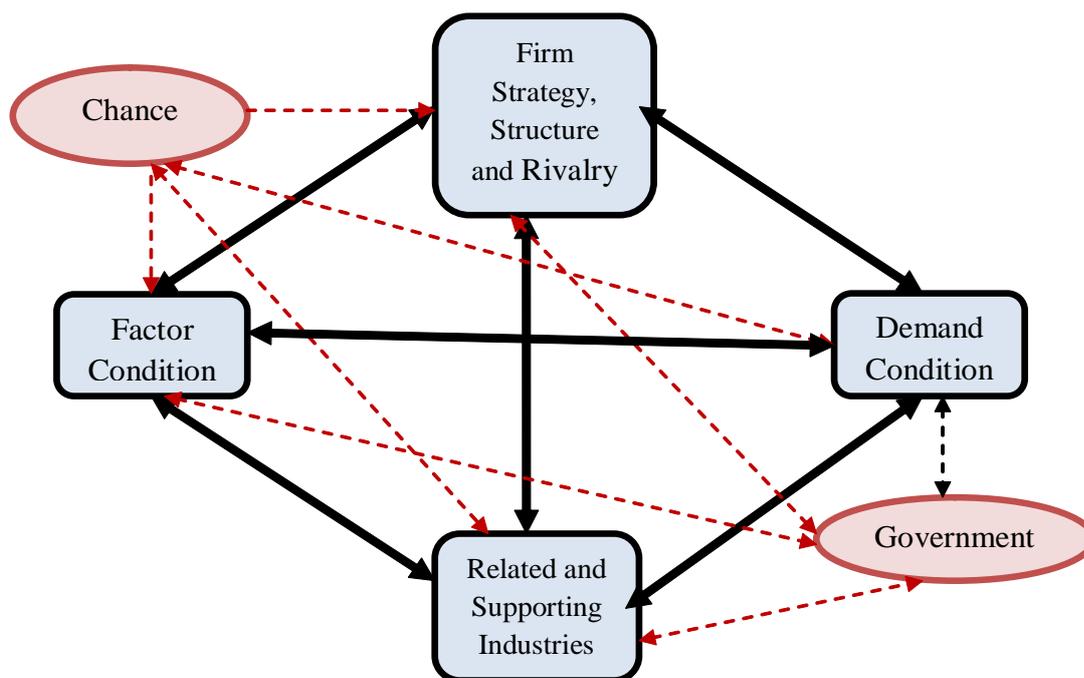


Figure 4 Diamond Porter Model (Porter, 1991)

Exponential Pair wise Comparison (EPC)

The Exponential Pair wise Comparison or EPC is a method in order to make a decision based on different alternatives (Manning, 1984 in Marimin 2008). The score of the priority alternative become very big due to the exponential function, causing the alternative decision to become significant. EPC structure is as follow:

$$VA_i = \sum (\text{Value}_{ij})^{\text{Crit}_j}$$

VA_i = value of i alternative
 Value_{ij} = value of i alternative for j criteria
 i = 1,2,3,4,...,n; n = number of alternatives
 j = 1,2,3,4,...,m; m = number of criterias
 Crit_j = importance level of j criteria; Criteria $j > 0$

EPC method has been widely used to make decisions based on alternatives, developing the agro industry dairy products (Canny, 2001), tuber agro industry products (Marimin and Sutyono, 2002), potential raw materials for the traditional drink or *jamu* (Kusnandar, 2002), agricultural system arrangement and choices of agricultural commodities (Syarifuddin *et al*, 2004), determination of supply allocation for superior product in vegetables supply chain (Marimin and Hadiguna, 2007).

Canny (2001) used the EPC method combined with Analytic Hierarchy Process (AHP) and Interpretative Structural Modelling (ISM) to construct and develop a decision support system (DSS) on the agroindustry dairy product based on the post harvest enterprises in Indonesia. Factors affecting the industry are own company's potential. The EPC method was also used in Syarifuddin *et al* (2004) studies to determine area land zone in Central Sulawesi to develop a competitive agricultural system arrangement and superior commodities arrangement. The study results seven primary zones for agricultural systems such as cacao, shallot, cattle and marine fishery.

Analytic Network Process (ANP)

Analytic is a form of the work analysis, which means the separating of any material or abstract entity into its constituent elements. Analysis is the opposite of synthesis, which involves putting together or combining parts into a whole (Forman and Selly, 2007). Saaty (2005) developed the Analytic Network Process or ANP as an extension of the Analytical Hierarchy Process. Many decision-making problems cannot be structured hierarchically because they involve interactions of

various factors, with high-level factors occasionally depending on low-level factors (Saaty, 1996; Lee and Kim, 2000).

Results And Discussions

As a country with fertile lands, Indonesia has many potential plants as a source of renewable energy however, the number of sources that had been identified is still limited. Table 1 represents different types of bioethanol sources related to other industries. Potential sources of raw materials are related to several factors namely land, cultivation techniques, low production cost, shorter growing period, mechanical gathering, high ethanol-transforming rate, government support and infrastructure (Lee *et al.*, 2011).

Table 1 Bioethanol Raw Materials and Their Functions

Types of Plants	Biomass Part That is Processed	Bioethanol Productivity (Litre/Ha/Year)	Current Function
Cassava	Fresh Tuber	4,500	Food/Industry/ livestock feed
Sweet Potato	Fresh Tuber	7,800	Food
Sugar cane	Fresh stems molases	5,000 1,000	Food Industry
Corn	Dry Corn Seed Flour	5,000-6,000	Food/Industry/livestock feed
Sweet Sorghum	Fresh Stems	5,500-6,000	Limited livestock feed
Sago	Sago flour	4,000-7,000	Limited Food Conservation

Source: Indahsari, 2012.

A number of potential bioethanol source namely sugarcane, cassava, sweet potatoes, corn, sago and sweet sorghum were studied within limited time available. The criteria used include (1) plant productivity level obtained from the Ministry of Agriculture (2010); (2) plant adaptation ability to marginal land conditions level of fertilizer; (3) plant adaptation ability to weather conditions, such as water requirement, direct sun, and humidity level; (4) continuity, ability and easiness for the plant to be generated or stored; (5) required technology for cultivation; and (6) infrastructure and government support, such as policy or subsidy that had already been imposed or developed to certain plant types (intensive and extensive programs).

Data collected will then be scored and weighted according to the Exponential Pairwise Comparison (EPC) Matrix as in Table 2:

Table 2 Exponential Matrix Pairwise Comparison (EPC)

Alternative Decisions	Decision's Criteria						Value	Priority
	1	2	3	-	-	-		
1	RK ₁₁	RK ₂₁	RK ₃₁				TN ₁	P ₁
2	RK ₁₂							
-								
N	RK _{1n}					RK _{mn}	TN _n	P _n
Degree of Importance	TKK ₁	TKK ₂	TKK ₃			TKK _m	TN	P

Source: Manning (1984) in Marimin (2008)

The calculation formula of the total value of each alternative decision are as follows:

Rk_{ij} = the relative degree of importance of j criteria on i decision, which can be expressed in ordinal scale (1,2,3,4,5)

TKK_j = the degree of importance of decision's criteria, which is expressed in weight

n = the number of alternative decisions

m = the number of decision's criteria.

Indonesia has more than 13 potential raw materials to produce bioethanol (Media data, 2008), however, only certain plants are considered to have potential to be developed. Table3 shows the result from EPC method to find three potential raw material plants for bioethanol, the three potential raw

materials are sugarcane, cassava and corn. Value of weights were determined from literature reviews and experts' opinion.

Table 3 Results for Potential Bioethanol Raw Materials Based on EPC

Type of Raw Materials	Productivity	Ability to Adapt to Weather	Ability to Adapt to Land	Continuity	Infrastructure and Technology	Government Support	EPC
Sugarcane	336000	900	800	140.00	100.00	100.00	101028.00
Cassava	245000	800	900	160.00	66.67	77.78	73724.22
Corn	188200	600	700	100.00	55.56	66.67	56628.89
Sweet Potato	21500	400	300	60.00	33.33	33.33	6542.00
Sago	100	300	200	133.33	22.22	22.22	113.33
Sweet Sorghum	100	100	100	20.00	11.11	11.11	57.33
Weight	0.3	0.1	0.1	0.2	0.1	0.2	1

Table 3 show that sugarcane, cassava and corn have been widely cultivated in Indonesia. The data was obtained from the Ministry of Agriculture however, because of insufficient data for productivity level of sago and sorghum, assumption was made to less than 1 million per year. Cassava is also considered as a high producing starch with available year round yield (Kuiper *et al*, 2007). Cassava tubers can also be chipped, dried and stored for utilization during periods of lean supply (Rañiola *et al*, 2009). On the other hand, high level of sugarcane yield and availability considered to correlate to with the amount of land area endowed by the companies (Media Data, 2008). In addition, technical support provided for producing sweet sorghum, such as cultivation, is still inadequate, noting that sweet sorghum is newly introduced crop (Sirappa, 2003).

In terms of adaptability to land condition, cassava is considered to be highly tolerant to extreme stress and even to marginal land conditions (Kuiper *et al*, 2007). On the other hand, sweet sorghum, as non native crop to Indonesia, will take some times for its adaptation to local weather condition. Similarly, data to support sago, sweet potato and sweet sorghum are also limited. Cassava's ability to adapt to weather condition is proven by its drought-resistant characteristic and even requires minimum crop maintenance (Rañiola *et al*, 2009), while sweet potato indicates its low resistant to diseases.

Production potentials of the three crops, in terms of their ability to regenerate, produce and sustainability sweet sorghum is constrained by its lack of cultivation and breeding experiences, hence the yield stability of the crop is not being considered as bioethanol source crops, as opposed to sugarcane (Köppen *et al*, 2009). In this instance, cassava is so far being viewed as the easiest crop to cultivate. Corn production is also constrained by its hybrid status, while sweet sorghum is considered the most difficult to reproduce, among other because the seeds still rely on imported.

Infrastructure includes government support to expand crop productivity. In this instance, sugarcane has been well supported by the government, including marketing system to enhance self sufficiency in white sugar production to meet domestic consumption. Equal and intensive supports are also given to corn and cassava development, while sweet sorghum as indicated earlier needs more studies before the government provides recommendation to farmers.

Table 4 shows the comparison between type of bioethanol sources related to energy produce. Sugarcane has the highest amount of bioethanol produced per litre per year per ha compared to corn, cassava, sweet potato, sweet sorghum and sago.

Table 4 Comparison of Bioethanol Energy Crops

Type of Raw Materials	Ratio Of Biomass to Bioethanol	Bioethanol (litre/yr/ha)
Corn	1-6	400-2,500
Cassava	10-50	2,000-7,000
Sugarcane	40-120	3,000-8,500
Sweet Potato	10-40	1,200-5,000
Sweet Sorghum	20-60	2,000-6,000
Sago	na*	na*

Source: Prihandana and Hendroko (2007)

Note: na* or not available

Expanding agriculture land area through the extension of commodities in unfavorable agroecology will result in non-optimal production and productivity, with higher risk and more input (Kartono, 2006). Therefore, it is important to identify the type of resources and raw materials that could be used to develop the bioethanol industry competitiveness. Sugarcane produces the highest bioethanol (3,000 to 8,000 litre/year/ha) followed by cassava (2,000 to 7,000 litre/year/ha) and sweet sorghum (2,000 to 6,000 litre/yr/ha). About 60% of world bioethanol production comes from sugarcane and 40% are from other crops (Balat and Balat, 2009). Sweet potato still produces almost twice bioethanol production (lt/yr/ha) as corn, while data for sago is still unavailable.

The government is carrying on-going research to develop superior varieties for sugarcane and cassava. In addition, supports has continued to further develop infrastructure to- in the bioethanol system, including investment tax deduction allowance such as supporting capital seed for biofuel fund at the amount of US\$ 220 million and provide incentives for farmers by subsidy interest rate as much as US\$ 110 million per year (Media Data, 2008). Again, supporting data for sago and sweet potato have been very difficult to obtain, mostly because of research done in these material sources are inadequate. Sago is more developed locally by farmers in Maluku and Papua. Moreover, even sweet sorghum has the potential to produce bioethanol, this new crop has its limitation for mass cultivation due to the fact that sweet sorghum's is not Indonesia origin (Prihandana and Hendroko, 2007). Figure 5 shows the productivity level for sugarcane, cassava and corn has increased in six years.

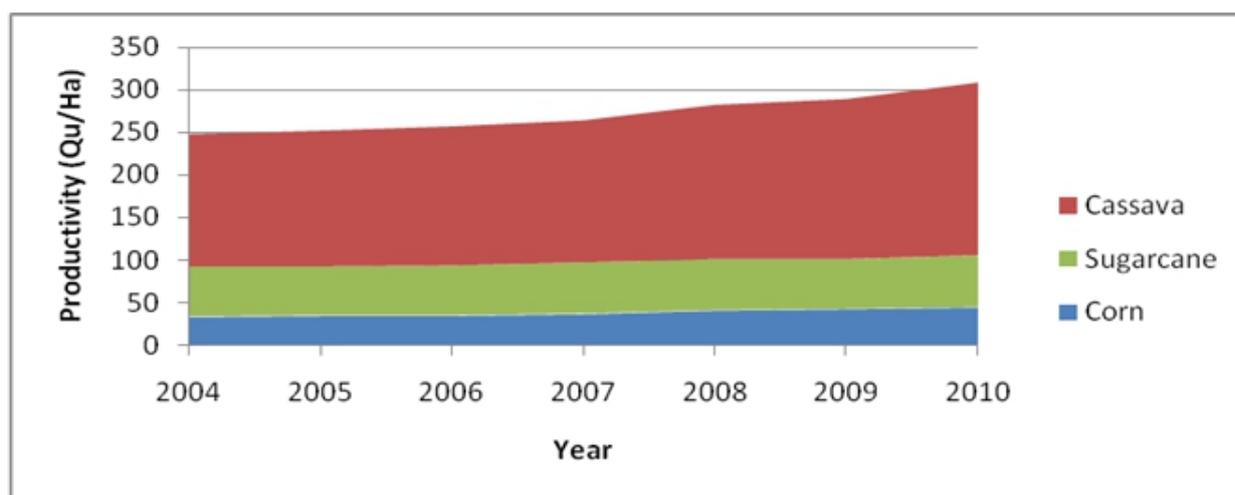


Figure 5 Comparison of Productivity Levels Between Cassava, Corn and Sugarcane (BPS, 2011)

Figure 6 compares production growth between corn, cassava and sugarcane. Corn has negative production growth rate in 2006 and has the tendency to decrease compared to production growth of cassava and sugarcane in 2010.

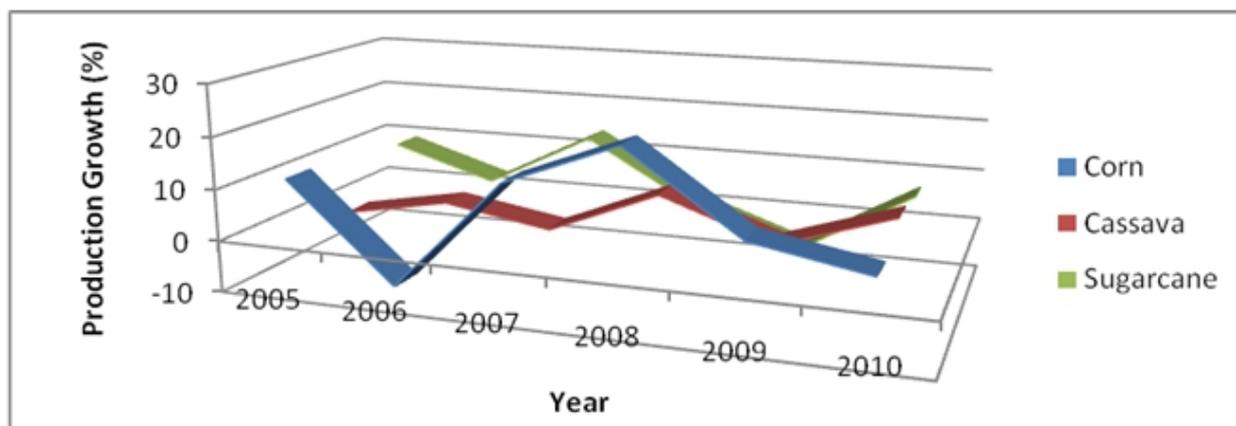


Figure 6 Comparison of Production Growth Rate (%) Between Corn, Sugarcane and Cassava (BPS, 2011)

The results either from EPC shows that sugarcane has better potential than cassava or corn as a bioethanol source, which is consistent with ANP results. Figure 6 shows type of raw materials for bioethanol and their level level of difficulties to develop these crops as material sources. The figure clearly shows that corn has the most problem as a crop material source for bioethanol followed by cassava, while sugarcane has the least problem.

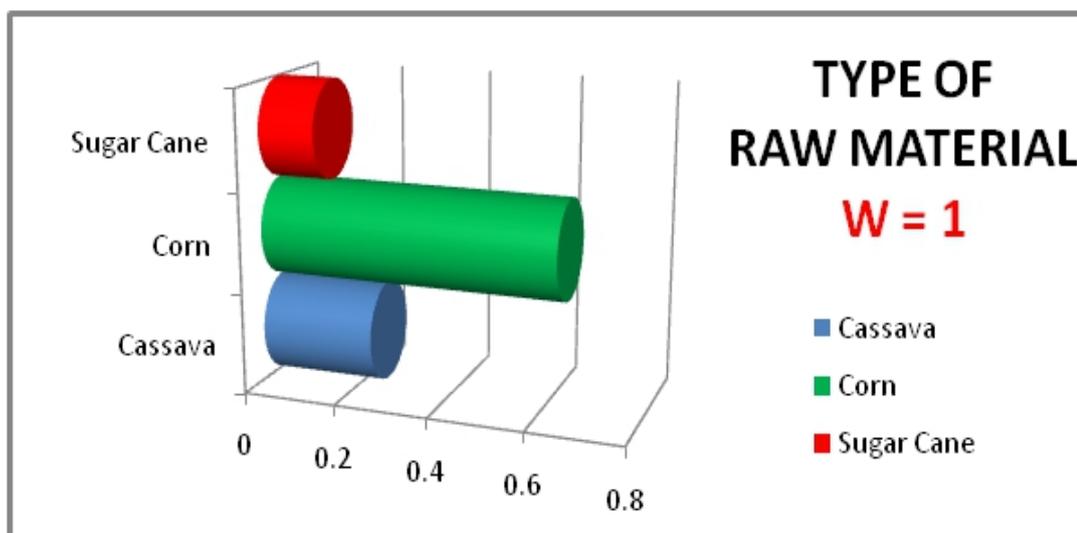


Figure 7 Types of Raw Materials and Level of Difficulties to produce them

ANP results shown in Table 5 indicates that corn has the most problem (0.63) compared to cassava (0.242) and sugarcane (0.123). According to Azhari (2010) the problem of raw materials is related to their unavailability as raw materials, due to some factors such as not enough supporting industry on processing and end results, insufficient capital and high production cost problems. The Normalize values shows how each factor affects each other within the same cluster, while the limiting values show the interaction between clusters. Total values shows the whole picture of the problem in the bioethanol industry in Indonesia.

Table 5 Normalize and Limiting Values For Type of Raw Material and Problems

Problems	Normalize Values	Limiting Values
Type of Raw Material Choice Difficulties:		
Cassava	0.242	0.009
Corn	0.635	0.025
Sugar Cane	0.123	0.005

The EPC and ANP results, indicate consistency of sugarcane being the most potential source for bioethanol. Therefore, the use of sugarcane (molasses) by PTPN X implies that the company is using the most potential material for developing bioethanol.

The Diamond Porter Model illustrates the current condition of the bioethanol industry related to PTPN X. Each factor affects each other and how the bioethanol industry can be developed. PTPN X is using sugarcane as the main source to develop the bioethanol industry, which is consistent with the finding from EPC and ANP. Figure 8 shows the Diamond Porter Model for PTPN X.

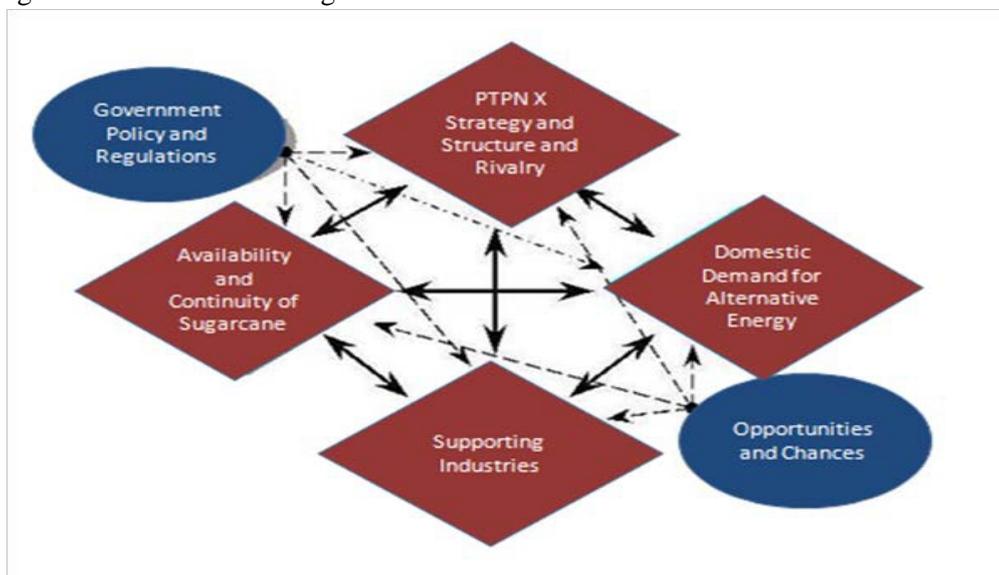


Figure 8 Development of Bioethanol by PTPN X (Adapted from Porter's Diamond Model Theory)

The factor demand for bioethanol includes how potential the market is. The end of cheap oil and the volatility in world oil price has cause bioethanol as an attractive alternative substitute for fossil fuel. Zuurbier and van de Vooren, (2008). Furthermore, the large number of Indonesian population increases the demand for fuel, hence, increase the demand for transportation which in turn to influence the demand for fossil fuels. Motorcycles represented the largest increase in transportation mode being three folds in 2010 as there were in 2000, while the amount of private cars increased two-fold since 2000 (BPS, 2011).

Factors of supporting industries include availability of current technology, economics of scale and infrastructure. Availability of support from other firms which focus on core capabilities and activities while identifying their weakness to support business (Porter, 1980). In their studies, Chan and Reiner (2011) found that many bioethanol industries development have horizontal and vertical integration. Reasons include reducing the risks of price volatility in the market, enhance security of supply, and also reducing costs. PTPN X is developing an integrated system to maximize bioethanol production. Other companies, such as bioethanol industries in India or in Brazil had been able to integrate different products while producing bioethanol. Bagasse was used to produce steam and electricity (Concuelo *et al*)

Advanced technology system is effective in producing higher bioethanol production while at the same time produce electricity. Private companies in Indonesia have also been applying this, by using electricity generated in the bioethanol process, to generate electricity for internal uses, which in turn cutting back the cost to the company. Waste from processing produces fertilizer and even fiber for producing paper. PTPN X is determine to develop the bioethanol industry as a whole with integrating other products as well.

The opportunities to develop bioethanol relies mostly on obligation or incentives given to bioethanol demand and consumption. Salvo and Huse (2011) found that consumer will use bioethanol when consumers are more concern with environment regardless of the price differences between bioethanol and fossil fuel. Therefore, educating consumers to use more eco-friendly energy alternative is important to develop the industry. Consumers are also concern with effects on vehicles engines when blending bioethanol with fossil fuel (Salvo and Huse, 2011). In Brazil, automotive companies are required to adjust the machines to facilitate the bioethanol mixture. Recently, along with the

commitment of the National Oil Company, Petrobras, automobile engines have been developed to receive any mixture of ethanol and fossil fuel (0 to 100% ethanol), called 'flex fuel vehicles'. The US Alternative Motor Fuel Acts of 1998 credits automobile producers when producing automobiles that are capable to blend bioethanol 85% to petroleum based gasoline of 15% (Zhang and Wetzstein, 2008). Bioethanol has lower energy content than fossil fuel, implying that bioethanol perform less miles per gallon compared to fossil fuel, however results in similar power and performance (Anderson, 2008). Therefore, bioethanol is more efficient and as effective as fossil fuel.

Indonesian government has produced policies and regulations to support the development of bioethanol, which provide energy security, reduce the greenhouse gas emissions while increasing and diversifying income of farmers and rural communities as a potential factors to support bioethanol industry (Zuurbier and van de Vooren, 2008). Presidential Regulation No. 5 of 2006 contains several directions of the national energy policy, such as the composition of a balanced source of energy of 54% petroleum, 26% natural gas and 14% coal. In 2025, which was expected that there would be a reduction in the contribution of oil to 20%, of natural gas to 30%, of coal to 2% , and increase of renewable energy (biomass, water, wind, solar and nuclear) to more than 5%, of geothermal to more than 5% , and of vegetable fuels (biofuels) equal to 5% in every national energy needed (BPPT, 2008). The Presidential Instruction No. 5 of 2006 mentioned the supply and use of biofuel. Therein, it has been instructed to a number of Ministries and local governments to take any measure in encouraging the supply and use of Bioethanol (Krisnamurthi, 2007).

The roadmap for gasohol or mixture between bioethanol and gasoline can be seen in Figure 9. The government had planned to increase the number of bioethanol plants from 17 plants in 2006 with a capacity of 60kL per day to a target of 25 pants in the year 2016.

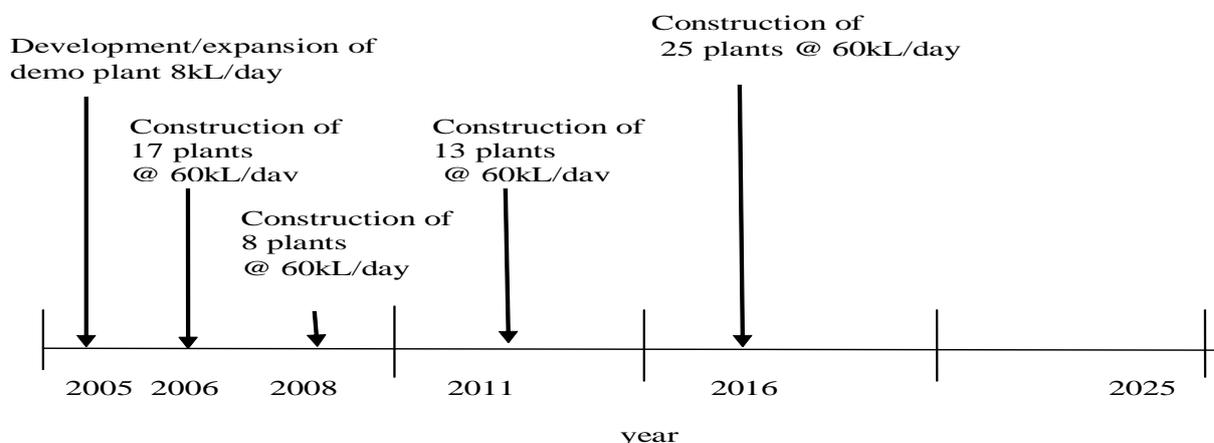


Figure 9 Government's Plan for Bioethanol (ESDM, 2006)

Concluding Remarks

Large scale bioethanol industry such as the one being operated by the PTPN X has a comparative advantage compared to others being run by small scale companies, from the viewpoints of integrating upstream through downstream factors, conditional to government policy to support the production of molasses. For example, sugar companies are still producing single final product, that is white sugar with molasses as side product. The main reason for the PTPN X being able to develop an integrated bioethanol industry with reduced costs is that electricity produced can be internally utilized while also producing steam and fertilizer.

Favorable conditions for the PTPN X to use sugarcane as the most potential raw material is its adoption of vertical integration which ensure optimum supply of the stock and technology used. In addition, the government has also been trying to stimulate the bioethanol industry, which proven that opportunity and support are equally important.

It is important that government policies related to the scale of bioethanol industry must be followed up with government and financial support for research and development, equipments and facilities, land use and market system. Further studies should be conducted on factors affecting cluster concentration development.

References:

- Anderson ST. The Demand for Ethanol as a Gasoline Substitute. *Journal of Transport Economics and Policy* 36 (1):1-26. Accessed on June, 2012. 2008.
- Azahari DH. Seminar Pengembangan Industri Biofuel (Tantangan Baru untuk Sektor Pertanian). Bogor. Pusat Analisis Sosial Ekonomi dan Kebijakan Pertanian. Badan Litbang Pertanian, Departemen Pertanian. 2008
- Badan Pusat Statistik (BPS). Data on Consumption, Production, Import and Export of Fuel. 2011
- BPPT. Energy. Where Do We Go From Here? 2005
- Chan JH and Reiner D. Vertical Integration in Growing Industry: Security of Supply and Market Access in Fuel-Ethanol Value Chain. *Strategic Management Journal* 5: 171-180. Accessed on June, 2012..
- Canny AH. *Rekayasa Sistem Pengembangan Agroindustri Susu Berbasis Usaha Lepas Panen Susu*. Dissertation. Bogor Agricultural University. 2001
- ESDM (Ministry of Energy and Mineral Resources). Blueprint of National Energy Management 2005-2025. www.esdm.go.id. Accessed May 2009 at 22.35 WIB. 2006
- Forman E and Selly MA. *Decision by Objectives*. George Washington University. Expert Choice Inc. Washington DC 2007.
- Irham M. Bioetanol Dongkrak Pamor Singkong. *Jurnal Bogor*. 2008
- Indahsari GK. *2Strategy to Enhance Competitiveness of Indonesia's Bioethanol Industry*. Bogor Agricultural University. *Dissertation*. 2012
- Kartono G. 2006. Role of AEZ (Agroecology Zone) and LQ (Location Quotient) in Extension of Horticulture in East Java. *Bulletin Teknologi dan Informasi Pertanian BPTP Jawa Timur* 9:23-32. 2006
- Köppen S, Reinhardt G and Gärtner S. Assesment of Energy and Greenhouse Gas Inventories of Sweet Sorghum for First and Second Generation Bioethanol. *FAO Environmental and Natural Resources Service Series No.30*. 83pp. 2009
- Kuiper L, Ekmekci B, Hamelinck C, Hettinga W, Meyer S and Koop K. *Bioethanol from Cassava*. Paper ECOFYS. Netherlands. 44pp. 2007
- Kusnandar. *Rekayasa Model Pengembangan Industri Kecil Jamu*. Universitas Sebelas Maret. Surakarta. 2002
- Lee SR, Shin HH, Jeong JH, Hwang KT and Kim TY. Effect of Ethanol Concentrations and Extraction Time on Acanthoside-D and Total Polyphenol Contents and Antioxidant Activities in Ethanol Extracts of Eleuthero. *Journal of Medicine Plant Reserch* 5(24): 5700-5705. 2011
- Lee JW and Kim SH. Using Analytic Network Process and Goal Programming for Interdependent Information System Project Selection. *Computers and Operations Research* 27:367-382. 2000.
- Manning WA. *Decision Making: How a Microcomputer Aids the Process* Portland State University. 1984
- Marimin dan Sutiyono A. *Model Sistem Manajemen Ahli Perencanaan Investasi* 2002.
- Produk Agroindustri Komoditas Umbi-umbian*. Proceeding, Komputern dan Sistem Intelijen (Kommit). Jakarta. 10 pages.
- Media Data. *Biofuel*. Media Data Publ. Jakarta. 2008.
- Ministry of Agriculture. peluang Agribisnis Menjadi Sumber Devisa Negara yang Utama. Direktorat Jenderal Pengolahan dan Pemasaran Hasil Pertanian Kementerian Pertanian Republik Indonesia. <http://pphp.deptan.go.id> accessed February 2009 (21.05 WIB). 2008.
- Consuelo L, Pereira F and Ortega. Sustainability Assessment of Ehtanol Production from Sugarcane. 1st International Workshop. *Advanced in Cleaner Production*. The Diamond Porter Model investigates the strength and weakness of an industry and how competitive is the industry. Accessed June 7th 2012.
- Porter ME. *Toward a Dynamic Theory of Strategy*. Strategy Management. New York: The Free Press. 1991
- Prihandana R, Noerwijat K, Adinurani PG, Setyaningsih D, Setiadi S dan Hendroko R. *Bioetanol Ubi Kayu: Bahan Bakar Masa Depan*. AgroMedia Pustaka. Tangerang. 194 hal. 2007
- Rañiola RFJ, Demafelis RB, Rosario EDS and Batalier BG. Enhancing the Viability of Cassava Feedstock for Bioethanol in the Phillippines. *Journal ISSAS Vol. 2*. Pp 147-158. 2009.
- Saaty TL. *Theory and Applications of the Analytic Network Process*. Decision Making with Benefits, Opportunities, Costs and Risks. RWS Publications, PA 2005.

RESPONSES OF RECOMBINANT INBRED LINES OF COWPEA [(*VIGNA UNGUICULATA* (L.) WALP)] TO *STRIGA GESNERIOIDES* INFESTATION IN GHANA

Aaron T. Asare

Isaac K. A. Galyuon

Department of Molecular Biology and Biotechnology,
University of Cape Coast, Cape Coast, Ghana

Francis K. Padi

Cocoa Research Institute of Ghana (CRIG), New Tafo-Akim, Ghana

Emmanuel P. Otwe

Department of Molecular Biology and Biotechnology,
University of Cape Coast, Cape Coast, Ghana

J.F. Takrama

Cocoa Research Institute of Ghana (CRIG), New Tafo-Akim, Ghana

Abstract:

Cowpea [*Vigna unguiculata* (L.) Walp.] production in West Africa is constraint by *Striga gesnerioides* parasitism associated with 83-100% yield losses which warrants development of resistance varieties against the parasite. An exotic resistant genotype, IT97K-499-35 developed by International Institute of Tropical Agriculture (IITA) was crossed with SARC-LO2, a local susceptible genotype. SSR-1 and C42-2B markers previously mapped in the region of *S. gesnerioides* resistance loci were used to amplify genomic DNA of advance recombinant inbred lines (RILs) of this cross. The responses of the RILs to *Striga* infection in pot culture and a field trial conformed to segregation ratio of resistance to susceptible genotypes of 1:1 among F₇ progenies suggesting monogenic dominant inheritance of the resistance. The markers SSR-1 and C42-2B mapped in the region of the resistance locus, but presented as dominant markers, with amplification only in resistant genotypes. The selective efficiency of SSR-1 (92.6%) was better than that of C42-2B (85.7%). In field trials, growth and morphology of susceptible genotypes were adversely affected by *S. gesnerioides* which resulted in significant ($P \leq 0.05$) reduction in seed yield compared to resistant genotypes. The resistant RILs identified in the current work would be further evaluated in multi-location trials prior to their release to farmers for cultivation.

Key Words: Cowpea, Marker-assisted selection, *Striga*, Recombinant inbred lines

Introduction

Cowpea [*Vigna unguiculata* (L.) Walp] is one of the most economically important indigenous African grain legumes with enriched protein as source of food for both human and animal nourishment and a major crop in regional trade within West and Central Africa (Langyintuo *et al.*, 2003). The relatively high protein content of cowpea makes it an essential supplement to the diet of many Africans (Bressani, 1985) who consume high carbohydrate but low in protein cereals, root and tuber crops (Omoigui, 2007). Besides, cowpea is also a valuable commodity that provides income for farmers and fixes atmospheric nitrogen to restore soil fertility for succeeding cereal crops growing in rotation with it. West and Central Africa produce 69% of the world production (Langyintuo *et al.*, 2003). However, a major biological constraint to increase production in smallholder farms is the infection by the parasitic weed, *Striga gesnerioides* (Willd) Vatke (Ehlers and Hall, 1997). Cowpea yield losses associated with *S. gesnerioides* range from 83 to 100% (Cardwell and Lane, 1995; Emechebe *et al.*, 1991)

The extent of damage to cowpea by *S. gesnerioides* infection is related to the close parasitic association between the host and the parasitic weed. Seed germination in *S. gesnerioides* occurs in response to specific stimulants exuded by host roots in the soil (Muller *et al.*, 1992). The extremity of the radical modified into haustorium (Okonkwo and Nwoke, 1978), that attaches and penetrates the vascular tissues and establish vascular connections (Ba, 1983) to derive water, minerals and organic compounds from the cowpea for the development of the parasite (Graves *et al.* 1992). Indeed, no single method is adequate to control the parasite, however, host plant resistance appears to have the potential to effectively and economically control the parasite since it is affordable to resource-poor farmers (Omoigui *et al.*, 2007) as well as being environmentally friendly. Breeding for resistance to *S. gesnerioides* has led to development of some resistant cowpea cultivars (Parker and Polniaszek, 1990; Aggarwal, 1991).

Development of cowpea host plant resistance to *S. gesnerioides* requires the application of phenotypic and genotypic diagnostic protocols to screen a population segregating for resistance to the parasite. Indeed molecular markers for identification and selection of *Striga*-resistant genotypes have been developed for most of the races of the parasite prevalent in West Africa. However, the differential virulence of races of *S. gesnerioides* on cowpea genotypes (Lane *et al.*, 1994; Singh, 2002) has serious implication to breeding and selection procedures. Hence, the need to use race specific markers to complement conventional breeding methods to identify cowpea resistant genotypes. Until now, seven races of *S. gesnerioides* have been identified based on host differential response and genetic diversity analysis within the cowpea growing regions of West Africa (Lane *et al.*, 1996). These races are designated as SG1 (Burkina Faso), SG2 (Mali), SG3 (Nigeria and Niger), SG4 and SG4z (Benin), SG5 (Cameroon) and SG6 (Senegal). According to Botanga and Timko (2005), race formation in cowpea-*Striga* association is largely a result of host-driving selection, because the parasite is autogamous with floral features that make occurrence of out-crossing very low. Identification of race-specific responses in cowpea is relevant for the development of target resistant genotypes.

Several race-specific resistance genes have been identified and located to linkage groups 1 and 6 (LG1 and LG6) of the current cowpea genetic map (Ouédraogo *et al.* 2001 and 2002). The genetics of cowpea *Striga*-resistance varies according to the biotype of the parasite; however, it is inherited mainly as a single nuclear dominant gene (Singh and Emebeche, 1990; Atokple *et al.*, 1993; Lane *et al.*, 1993; Moore *et al.*, 1995; Toure *et al.*, 1997; Carsky *et al.*, 2003). A few reports, however, have pointed out that the resistance is conferred by two independent dominant genes (Dubé, 2000) or a recessive single gene (Toure *et al.*, 1997). The implication of this variation in resistance is that reliable screening protocols are required to select recombinant inbred lines for the resistance or susceptibility to the parasite. Therefore, the focus of the current work was to screen recombinant inbred lines of the cross between *Striga*-resistant exotic line and locally adapted variety susceptible to the parasite to select resistant lines as basis for developing better adapted varieties.

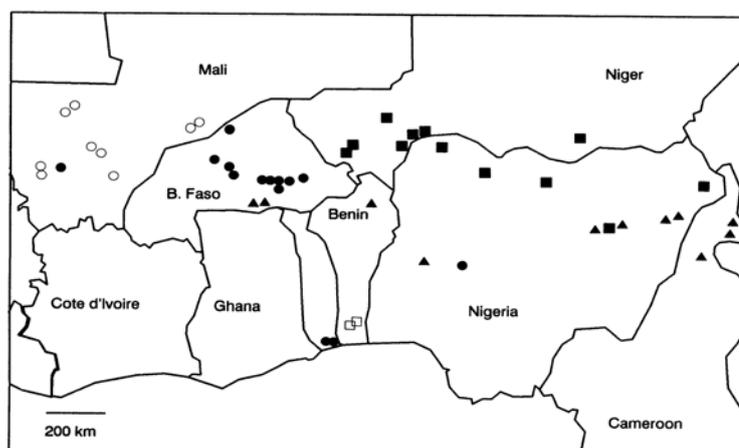


Fig. 1. Distribution of races of *Striga gesnerioides* in West Africa: closed circles, race 1; open circles, race 2; closed squares, race 3; open squares, race 4, and triangles, race 5.

Source: Lane *et al.* (1996)

Materials and Methods

Plant culture and DNA extraction

Cowpea seedlings of 98 recombinant inbred lines (F₇) derived from a cross of IT97K-499-35 (resistant parent) × SARC-LO2 (susceptible parent) were raised in plastic pots filled with sandy loam soil in the Botanical Garden of the School of Biological Sciences of the University of Cape Coast. Three cowpea seeds were sown in each plastic pot and maintained with regular watering at two-day intervals or when necessary. Three young leaves for each potted cowpea were harvested at 14 days after sowing (DAS) seeds, labeled and frozen in liquid nitrogen.

The frozen dried leaves were homogenized and total genomic DNA isolated from each sample using the plant DNAzol® ES (MRC Inc, Cincinnati, OH) as per manufacturer's protocol with slight modification. The homogenate was transferred into 2ml Eppendorf tubes containing 750µL DNAzol® ES reagent and after vigorous shaking, 750µL chloroform was added. The mixture was centrifuged at high speed (12,000rpm for 10minutes) and supernatant was transferred to a new microfuge tube. Absolute ethanol (750 µL) was added to precipitate the DNA and the DNA pellet recovered by centrifugation at low speed (6000 rpm for 5minutes). Precipitated DNA was rinsed gently with 70 % ethanol, air-dried and suspended in 200 µL TE buffer (pH 8.0) and stored in a fridge at -4 °C overnight. The suspended DNA was incubated at 65 °C for 30 minutes in a water bath. The bottom of centrifuge tube containing DNA was tapped gently and centrifuged at 6000 rpm for 5minutes. The DNA was quantified with a spectrophotometer (Thermo Scientific, Wilmington, DE) and intactness of DNA was checked by resolving 1µL genomic DNA in 0.8 % (w/v) agarose gel. A DNA working solution of 100 ng/µL was prepared and stored in a freezer at -20 °C until use.

Bulked Segregant Analysis

The Bulk Segregant Analysis (BSA) following Michelmore *et al.*, (1991) and Boukar *et al.* (2004) was employed to assess the segregation patterns and select more informative molecular markers for genomic analysis of the RILs. Equivalent amounts of genomic DNA from 10 resistant and 10 susceptible F₇ plants respectively from the population derived for the cross between IT97K-499-35 and SARC-LO2 were pooled to form resistant and susceptible bulks. Both bulks were used along with the parents to and screened with three sequenced characterized amplified region (SCAR) primers 61R, 61RM2 and C42-2B and one microsatellite primer, SSR-1. The ethidium bromide-stained gels were visualized on a UV transilluminator (M-15 UVP Upland, CA 91786 USA) and photographed using a digital camera. The highly polymorphic informative primers were selected and further used to analyze the population to determine the association of the markers with cowpea resistance or susceptibility to *S. gesnerioides*.

Polymerase chain reaction (PCR) analysis

Each PCR reaction mixture contained 8 µL Bioneer AccuPower® TLA PCR PreMix desolved in Molecular Grade Distilled water (MGDw), 0.5 µL of forward primer, 0.5 µL reverse primer and 1 µL genomic DNA of sample were added to make up 10 µL total volume. The PCR amplifications were performed in an Eppendorf Mastercycler (Techne TC-512) comprising an initial denaturation at 94 °C for 1 min followed by 35 cycles of denaturation for 5 min, annealing at 55 °C or 60 °C for 1 min, extension at 72 °C for 1 min and end with final extension at same temperature for 5 min. The PCR products were resolved for 1 h at 120 V on 2% (w/v) Agarose gel in 1 × TAE buffer using a horizontal gel electrophoresis apparatus (Model V16.2 or V16; Gibco BRL, Gaithersberg, MD, USA). The gels were stained with ethidium bromide and visualized on a UV transilluminator (M-15 UVP Upland, CA 91786 USA) and photo-documented with a digital camera. The size of DNA bands in base pairs was determined using the 1 kb DNA standard ladder (Invitrogen, Carlsbad, CA, USA).

Pot culture screening of cowpea against *Striga-gesnerioides* infection

The pot culture screening method used by Botanga and Timko (2005) was employed to assess the response of recombinant inbred lines of cowpea derived from IT97K-499-35 (resistant) × SARC-LO2 (susceptible) with a local check GH3684 (resistant) to *Striga gesnerioides* infection in the glasshouse of SARI, Tamale-Nyankpala in 2009. Each pot (17 × 11cm) filled with garden was inoculated with about 1000 seeds of *Striga gesnerioides* from the Upper East Region of Ghana. Four seeds of cowpea were sown per pot in 3 replications. The seedlings were thinned out and two were maintained per pot at 2 weeks after germination. The soil was kept moist by watering regularly every two days or when necessary. Destructive sampling was carried out at 8 weeks. The plant-soil mass

was removed from each pot, immersed into a bucket of water, and gently agitated to loosen the soil mass. The roots were washed thoroughly free of soil and examined using hand lens for presence of necrotic hypersensitive lesions, attachment of *Striga gesnerioides* and tubercles. Plants that favoured attachment, healthy development and emergence of *Striga gesnerioides* were classified as susceptible and those that appeared free from infection, without any attachment were categorized as candidate resistant genotypes.

Response of cowpea breeding lines to *S. gesnerioides* infection in field trial

Nighty-eight (98) recombinant inbred lines (RILs) selected by pot screening and molecular markers for resistance and susceptibility to the parasite were composed into a field trial. The trial was conducted under rain-fed sandy soil conditions from 3rd August- 31st October 2011 at the Savannah Agriculture Research Institute (SARI), Manga Station in Bawku located within latitude 11° 11' 0, longitude 10° 40' N and at an altitude of 249 m above sea level in a *Striga* seed-infested field (hotspot). with a local cowpea line GH3684 with resistance to *Striga* was included as a check. A 20 m × 40 m plot of land was ploughed, harrowed and ridges were constructed. The field was divided into two blocks at 1.5 m apart.

Three seeds of each cowpea breeding line were randomly sowed at 40 cm within row and 60 cm between rows. The cowpea seedlings were thinned to one per stand at 14 days after germination, allowing 12 cowpea stands per breeding line in each row arranged in a randomized complete block design. The local cowpea accession, GH3684 included in the test trial was also used as border plant for the entire set-up. Cultural practices, including hand weeding, were carried out at 3 and 6 weeks after sowing seeds, and insecticide (Dimiprid[®] 200 SL at 35ml/15L knapsack capacity) sprayings were carried on 4 and 7 weeks old plants. The number of *S. gesnerioides* emerged per plot was recorded for each cowpea breeding line. The effects of the parasitic stress on growth and morphology of cowpeas caused by *S. gesnerioides* were observed from 6 to 8 weeks of the growth period. The weight of hundred seeds was determined. The data were subjected to analysis of variance (ANOVA) using Minitab 16.2.2 statistical software. The differences between mean values were assessed using the least significant difference (LSD) at 5% level of significance.

Results

The pot test on F₇ RILs of advanced cowpea progenies derived from a cross of IT97K-499-35 (resistant parent) × SARC-LO2 (susceptible parent) expressed a segregation ratio of 1R:1S ($\chi^2 = 0.003$; P = 0.995). This ratio conforms well to those of SSR-1 and C42-2B molecular markers in the cowpea genome and further confirmed by the response of the cowpea genotypes to *S. gesnerioides* infestation on the field (Plates 1 and 2; Table 1). The susceptible genotypes of cowpea had *Striga* seedlings attached to the roots after 45 days of germination or germinated *Striga* seedlings emerged on the surface of the soil. These cowpea plants expressed varied symptoms with age of culture due to *Striga*-parasitic stress including stunted growth, leaf necrosis, chlorosis, senescence, defoliation and reduced size of young leaves. The susceptible cowpea plants also had reduced flowering and pod formation as well as poor rooting and nodulation. Contrary, the resistant recombinant inbred lines (RILs) of cowpea had normal growth and development without *Striga* attachment or emergence comparable to the resistant parental genotype IT97K-499-35 and the local genotype cowpea accession, GH3684 used as a check.

The details of *Striga* emergence and degree of parasite infection for the individual RILs in the field trial are presented in Table 1. There were significant (P ≤ 0.05) differences in the rate of emergence of *S. gesnerioides* among the genotypes of the F₇ RILs of cowpea. The emergence of *S. gesnerioides* ranged from 1 – 13.0 plants per plot (Table 1). The RILs that were not associated with *Striga* emergence and had the SSR-1 or C42-2B maker were observed as resistant genotypes. However, the cowpea RILs that were associated with *S. gesnerioides* emergence on the field or in the pot, and devoid of the SSR-1 and C42-2B resistance marker in the genome were susceptible genotypes. On the whole, the cowpea RILs differed in their growth responses with corresponding 100 seed yield under *Striga* infestation in the field.

The mean 100-seed dry weight for susceptible RILs was significantly (P ≤ 0.05) lower (15.2g) than that (16.6g) for the corresponding resistant RILs. The influence of gene segregation, coupled with the *Striga* infection on 100-seed dry weight apparently followed the regular distribution pattern for a continuous variation (Fig. 2). The 100 seed weight ranged from 3.2g in the *Striga*-susceptible progeny UC96-05 to 23.0g in the *Striga*-resistant progeny UC96-223. On uninfected plots, the

susceptible parent (SARC-LO2) had large seeds with higher 100-seed weight (26.8g) (data not shown) compared to seeds from infected plots (19.8g). The 100-seed weight of the resistant parent (IT97K-499-35) was greater (16.2g) compared to the resistant local genotype GH3684 with a 100-seed weight of 13.3g.

Bulk segregant analysis in addition to the pot and field screening data revealed that C42-2B and SSR-1 markers were more informative in distinguishing resistant from susceptible bulks. SSR-1 and C42-2B markers produced single bands of 150bp and 180bp PCR products, respectively, with amplification only in resistant genotypes (Plate 1 and 2). The marker segregation efficiency of SSR-1 was better (92.6 %) than that of C42-2B (85.7%) in identifying resistant cowpea genotypes among the RILs.

Twelve RILs (UC96-36, UC96-46, UC96-50, UC96-85, UC96-171, UC96-177, UC96-191, UC96-264, UC96-290, UC96-333, UC96-357, and UC96-357) which had the C42-2B marker for resistance were found to be susceptible. Similarly, thirteen others (UC96-07, UC96-10, UC96-38, UC96-47, UC96-48, UC96-64, UC96-113, UC96-189, UC96-206, UC96-209, UC96-229, UC96-243 and UC96-274) with both SSR-1 and C42-2B markers were also susceptible in field trials indicating crossing-over between the marker locus and the gene locus for resistance.

Table 1. Reaction of cowpea RILs derived from a cross of IT97K-499-35 × SARC-LO2 to *S. gesnerioides* infection

Genotypes	Phenotypic reaction				Genotypic reaction	
	Pot Test	Field Trial	100-seed weight/g	Striga Emergence/Plot	SSR-1 Marker 150bp	C42-2B Marker 180bp
IT97K-499-35	R	R	16.0	0	+	+
SARC-LO2	S	S	19.8	13	-	-
UC96-02	S	S	12.2	7	+	-
UC96-03	R	R	16.4	0	+	+
UC96-05	S	S	3.2	3	-	+
UC96-07	S	S	15.8	2	+	-
UC96-08	S	S	18.1	6	-	+
UC96-10	R	S	13.4	2	+	-
UC96-11	R	R	18.8	0	NA	NA
UC96-12	S	S	13	11	-	-
UC96-17	S	S	16.5	1	-	-
UC96-19	S	S	16.1	5	-	-
UC96-20	R	R	17.6	0	+	-
UC96-23	R	R	16.9	0	-	-
UC96-24	R	R	20.1	0	+	+
UC96-25	R	NA	-	0	+	-
UC96-30	R	NA	-	0	-	+
UC96-32	R	R	19.9	0	+	+
UC96-33	R	R	19.0	0	+	+
UC96-36	S	S	16.0	1	-	+
UC96-37	S	S	12.1	1	-	+
UC96-38	R	S	14.2	10	+	+
UC96-39	S	S	-	-	-	-
UC96-44	S	R	16.0	0	-	-
UC96-46	S	S	17.3	4	-	+
UC96-47	S	S	21.2	3	+	+
UC96-48	R	S	16.1	2	+	+
UC96-50	S	S	13.2	12	-	+
UC96-52	S	S	16.6	1	-	-
UC96-56	R	R	18.9	0	+	+
UC96-60	R	R	16.9	0	-	+

R: Resistant, S: Susceptible, +: Presence of marker, -: absence of marker or product, NA: Not applicable

Table 1. Continued

Genotypes	Phenotypic reaction				Genotypic reaction	
	Pot Test	Field Trial	100-seed weight/g	Striga Emergence/ Plot	SSR-1 Marker 150bp	C42-2B Marker 180 bp
UC96-64	S	S	20.9	3	+	+
UC96-73	S	S	-	2	-	-
UC96-76	S	S	20.2	1	-	-
UC96-77	R	R	17.8	0	+	+
UC96-80	R	R	15.1	0	+	+
UC96-85	S	S	14.7	4	-	+
UC96-86	R	R	14.6	0	+	+
UC96-99	S	S	14.6	1	-	-
UC96-113	R	S	16.5	2	+	+
UC96-122	R	R	16.7	0	+	+
UC96-128	R	R	14.1	0	+	+
UC96-139	R	R	15.6	0	+	+
UC96-141	R	R	16.5	0	+	+
UC96-144	R	R	-	0	+	+
UC96-145	S	S	10.9	3	-	-
UC96-148	R	R	16.3	0	+	+
UC96-151	R	R	16.6	0	+	+
UC96-153	R	R	17.7	0	+	+
UC96-154	R	R	21.4	0	+	+
UC96-168	R	R	16.9	0	-	+
UC96-171	S	S	18.4	4	-	+
UC96-173	R	R	16.2	0	+	+
UC96-174	S	S	-	-	-	-
UC96-177	S	S	18.4	2	-	+
UC96-178	R	R	16.5	0	-	+
UC96-186	R	R	-	0	-	-
UC96-189	S	S	-	2	+	+
UC96-191	R	S	16.8	4	-	+
UC96-194	R	R	16.0	0	+	+
UC96-198	R	R	15.9	0	+	+
UC96-199	R	R	16.6	0	+	+
UC96-200	R	R	16.5	0	+	-

R: Resistant, S: Susceptible, +: Presence of marker, -: absence of marker or product

Table 1. Continued

Genotypes	Phenotypic reaction				Genotypic reaction	
	Pot Test	Field Trial	100-seed weight/g	Striga Emergence/ Plot	SSR-1 Marker 150bp	C42-2B Marker 180bp
UC96-204	R	R	17.9	0	+	-
UC96-206	R	S	14.1	2	+	+
UC96-209	R	S	16.0	3	+	+
UC96-211	S	S	14.0	5	-	-
UC96-212	R	R	13.1	0	+	+
UC96-216	S	S	7.3	4	-	-
UC96-221	R	R	16.9	0	+	+
UC96-222	R	R	16.4	0	+	+
UC96-223	R	R	15.8	0	+	+
UC96-226	R	R	23.0	0	+	+
UC96-227	R	R	16.4	0	+	+
UC96-229	R	S	16.0	1	+	+
UC96-231	R	R	16.6	0	+	+
UC96-236	R	R	17.1	0	+	+
UC96-239	R	R	18	0	+	+
UC96-241	R	R	21.3	0	+	+
UC96-242	R	R	13.9	0	+	+
UC96-243	R	S	13.7	1	+	+
UC96-244	R	R	-	0	+	+
UC96-247	R	R	17.5	0	+	+
UC96-253	R	R	15.3	0	+	+
UC96-264	S	S	10.0	3	-	+
UC96-270	R	R	-	0	+	+
UC96-274	R	S	-	2	+	+
UC96-275	R	R	18.1	0	+	+
UC96-276	S	S	13.9	2	NA	NA
UC96-288	R	R	15.8	0	+	+

R: Resistant, S: Susceptible, +: Presence of marker, -: absence of marker or product, NA: Not applicable

Table 1. Continued

Genotypes	Phenotypic reaction				Genotypic reaction	
	Pot Test	Field Trial	100-seed weight/g	Striga Emergence/ Plot	SSR-1 Marker 150bp	C42-2B Marker 180bp
UC96-290	S	S	17.6	3	-	+
UC96-292	R	R	13.5	0	+	+
UC96-318	R	R	17.2	0	+	+
UC96-321	R	R	16.2	0	+	+
UC96-323	R	R	17.8	0	+	+
UC96-328	R	R	15.6	0	+	+
UC96-329	R	NA	-	-	+	+
UC96-333	R	S	14.6	1	-	+
UC96-346	R	R	18	0	+	+
UC96-352	R	NA	-	-	+	+
UC96-353	S	S	16.2	1	-	+
UC96-357	R	S	13.9	1	-	+
GH3684	R	R	13.2	0	+	+

R: Resistant, S: Susceptible, +: Presence of marker, -: absence of marker or product, NA: Not applicable

Fig.2. Variation in 100-seed dry weight of F₇ recombinant inbred lines of cowpea from a cross of IT97K-499-35 × SARC-LO2

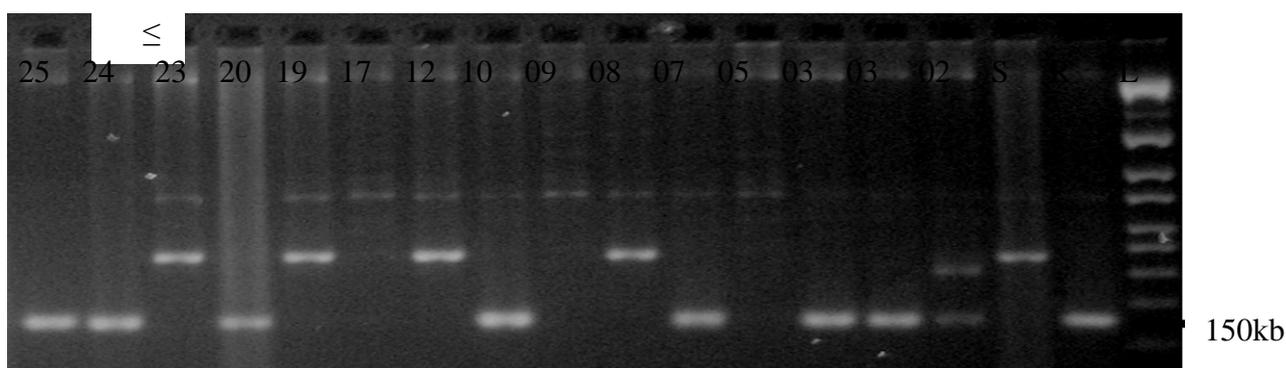
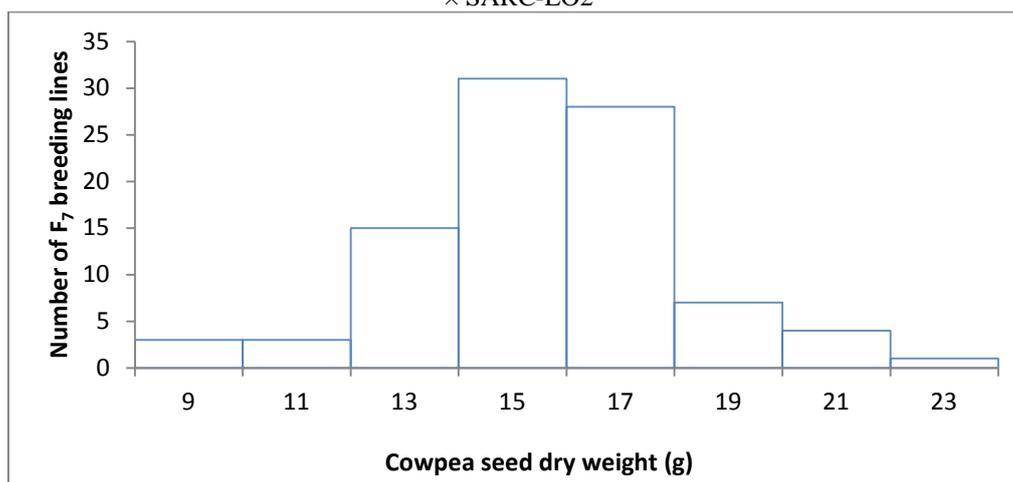
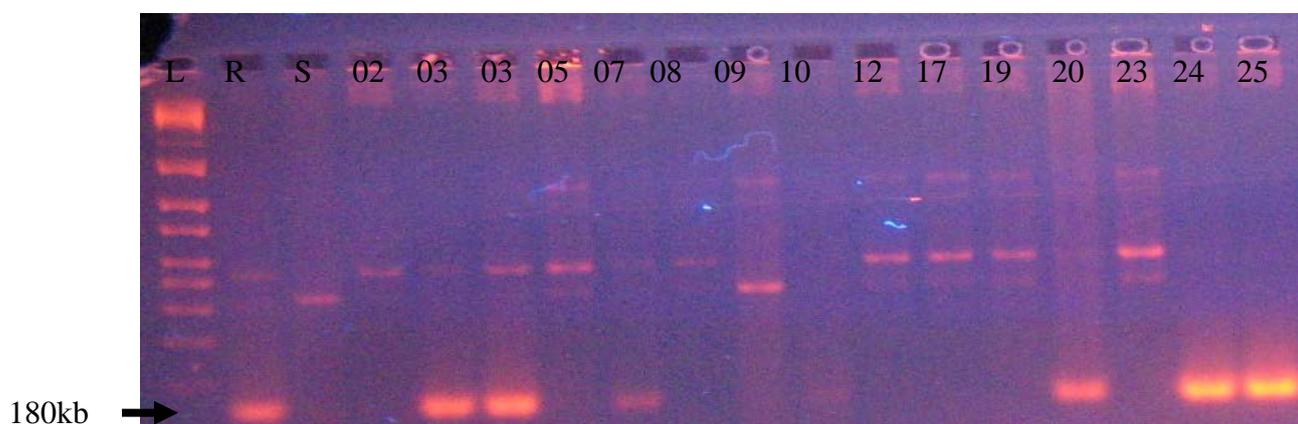


Plate 1: DNA bands from PCR amplification products of SSR-1 for some F₇ RILs of cowpea (derived from IT97K-499-35 × SARC-LO2) resolved in 2 % Agarose gel stained with ethidium bromide. The presence of 150bp band indicates resistant genotype and absence of this band indicate susceptible genotype. L represents the standard 1kb ladder.



Plates 2: DNA bands obtained from PCR amplification products of the SCAR primer C42-2B among some F₇ RILs derived from IT97K-499-35 × SARC-LO2 resolved in 2 % Agarose gel stained with ethidium bromide. Resistant lines have the 180bp band. L represents the standard 1kb ladder

Discussion

Significant effort has been made to identify natural sources of genetic resistance within cowpea and to select and breed for improved lines with resistance to *S. gesnerioides* (Singh and Emechebe, 1997; Singh *et al.*, 2002). However, the use of most resistant varieties is limited due to concerns about the potential adaptability and small or medium seed size as found in variety IT97K-

499-35 (Omoigui, *et al.*, 2007). IT97K-499-35 is a derivative from B301, local landrace from Botswana, which produces small seeds but is a multi-race resistant genotype to both *S. gesnerioides* and *Alectra vogelii* (Singh, 2002). Earlier inheritance studies indicated that the nature of resistance to *S. gesnerioides* races SG1, SG2, SG3 and SG4 in some cowpea genotypes is monogenic dominant (Singh and Emechebe, 1990; Atokple *et al.*, 1993; Moore *et al.*, 1995). The observed segregation ratio of 1R:1S in the present study is what is expected in a RIL population for a trait controlled by monogenic inheritance.. .

The SSR-1 and C42-2B primers distinguished between resistant and susceptible cowpea genotypes with different discriminating power. Indeed, the SSR-1 and C42-2B markers were found to co-segregate with *S. gesnerioides* race 3 or SG3 resistance gene (Li and Timko 2009; Omoigui, *et al.*, 2009). Both primers identified resistant cowpeas by amplification of the band in only resistant genotypes. According to Omoigui *et al.* (2009) C42B-2B identified resistant lines with a single band while the susceptible lines had no band. In the current study, the 150bp SSR-1 marker was more efficient at 92.6% discriminating ability compared to that of C42-2B (85.7%). The implication is that SSR-1 might be closer to the *S. gesnerioides* race specific-resistant gene (SG3) than C42-2B as applied to the unknown race of *S. gesnerioides* in Ghana from the Upper East Region.

Crop yield losses due to stress imposed by *S. gesnerioides* can range from 83 to 100% (Aggarwal and Ouédraogo, 1989; Alonge *et al.*, 2005; Cardwell and Lane, 1995; Emechebe *et al.*, 1991) depending on the extent of damage and level of infestation. The observed stunted growth, leaf necrosis, chlorosis, senescence, defoliation, reduced size of young leaves, poor flowering and poor pod formation as well as poor rooting and nodulation from pot culture to field trial emphasized the devastating effects of *Striga* parasitism on the crop. The significantly ($P \leq 0.05$) low average 100-seed dry weight (15.2g) among susceptible genotypes compared to that of resistant genotypes (16.6g) could be due to the parasite-induced damages giving rise to yield loss. The decrease in seed weight might have resulted from reduced seed size and/or a direct effect of reduction in photosynthesis and translocation of photosynthates to sink due to *Striga* stress. Competition between parasite and host for solutes and water coupled with lower rate of photosynthesis in the leaves may retard root and shoot growth and, consequently, yield. The fact that dry seed weight of the progenies demonstrated a seemingly continuous variation, which closely conform to the normal distribution curve, suggest the existence of potential high yielding genotypes for cultivation in both *Striga* prone and non-*Striga* prone areas.

The current field trial involving F_7 progenies attracted low emergence of *S. gesnerioides* per plot (1-13) compared to data obtained for other studies (Carsky *et al.*, 2003; Kamara *et al.*, 2008). The resultant low parasite emergence might have been influenced by poor rainfall during the trial or low density of seeds present in the soil. However, the identification of susceptible and resistant RILs conformed to the selection procedure by Singh and Emechebe (1990) and confirmed with the presence or absence of distinct markers associated with *S. gesnerioides* resistance. The combined conventional and DNA marker technology used facilitated selection of 59.2% genotypes of the RILs with the same resistance traits as the parent IT97K-499-35 which compared well with the local resistant check (GH3684) in both pot and field assessments.

Conclusions

The segregation ratio of resistance to susceptible cowpeas of 1:1 observed among the F_6 progenies indicates that inheritance of resistance to the race of *Striga* in Ghana is monogenic. Resistant genotypes were identified with the presence of single distinct DNA bands of 150bp and 180bp for SSR-1 and C42-2B, respectively, which were absent in susceptible genotypes. SSR-1 was more efficient (92.6%) than C42-2B (85.7%) in discriminating between resistant and susceptible genotypes. The small number of crossover events between the markers and the gene indicates reliability of these markers in improving cowpeas for *Striga* resistance in Ghana. The stress imposed by *S. gesnerioides* infection resulted in significant ($P \leq 0.05$) yield reduction compared to resistant genotypes due to reduced vegetative and reproductive growth. The resistant RILs identified in the current study would have to be further evaluated for release to farmers to cultivate.

The authors wish to thank Mr. Francis Kusi of the Savanna Agriculture Research Institute, Manga station in Bawku for the role played to phenotype the cowpea genotypes. We are also grateful to the Global Diversity Trust of Rome, Italy and the University of Cape Coast, Ghana for co-sponsoring this

work. The design, execution and interpretation of the research remain wholly the responsibility of the authors.

References:

- Aggarwal, V.D. and Quédraogo, J.T. (1989). Estimation of cowpea yield loss from *Striga* infection. *Trop. Agric.* **66**: 91-92.
- Aggarwal, V.D. (1991). Research on cowpea- *Striga* resistance at IITA. In: *Combating Striga in Africa*. Kim, S. K. (Ed.) IITA, Ibadan, Nigeria. pp. 90-95.
- Alonge, S.O., Lagoke, S.T.O. and Ajakaiye, C.O. (2005). Cowpea reactions to *Striga gesnerioides*: Effects on growth. *Crop protection* **24**: 565-573.
- Atokple, I.D.K., Singh, B.B. and Emechebe, A.M.. (1993). Independent inheritance of *Striga* and *Alectra* resistance in cowpea genotype B301. *Crop Sci.* **33**: 714-715.
- Ba, A.T. (1983). Biologie Du parasitisme chez deux Scrophulariacées tropicales: *Striga hermonthica* (Del.) Benth. Et *Striga gesnerioides* (Willd.) Vatke. In: Touré, M., Olivier, A., Ntare, B.R., Lane, J.A. and St-Pierré, C.A. (1997). Inheritance of resistance to *Striga gesnerioides* biotypes from Mali and Niger in cowpea [*Vigna unguiculata* (L.) Walp]. *Euphytica* **94**: 273-278.
- Botanga, C.J. and Timko, M.P. (2005). Genetic structure and analysis of host and non-host Interactions of *Striga gesnerioides* (witchweed) from Central Florida. *Phytopathology* **95**:1166-1173.
- Boukar, O., Kong, L., Singh, B.B., Murdock, L. and Ohm, H.W. (2004). AFLP and AFLP derived SCAR markers associated with *Striga gesnerioides* resistance in cowpea. [*Vigna unguiculata* (L.) Walp.]. *Crop Science* **44**: 1259-1264.
- Bressani, R. (1985). Nutritive value of cowpea. In: *Cowpea Research, Production and Utilization*. Singh, S.R. and Rachel, K.O. (Eds), pp 353- 360. Wiley and Sons, Chester, UK.
- Cardwell, K.F. and Lane, J. A. (1995). Effects of soils, cropping system and host phenotype on incidence and severity of *Striga gesnerioides* on cowpea in West Africa. *Agric. Ecosyst. Environ.* **53**: 253-262.
- Carsky, R.J., Akakpo, C., Singh, B.B., and Detongnon, J. (2003). Cowpea yield gain from resistance to *Striga gesnerioides* in Southern Benin. *Experimental Agriculture* **39** (3): 327-333.
- Dubé, M.P. (2000). *Study of inheritance to Striga gesnerioides on cowpea genotypes HTR and Wango-1*. Masters Dissertation, University of Laval, Canada.
- Ehlers, J.D. and Hall, A.E. (1997). Cowpea (*Vigna unguiculata* L. Walp.). *Field Crops Res.* **53**: 187-204.
- Emechebe, A.M., Singh, B.B., Leleji, O.I., Atokple, I.D.K. and Adu, J.K. (1991). Cowpea *Striga* problems and research in Nigeria. In: Kim, S.K. (Ed.) *Combating Striga in Africa*. Proc. Int. Workshop, Ibadan, Nigeria. 1998. IITA, Ibadan, Nigeria. pp. 18-28.
- Graves, J. D., Press, M.C., Smith, S. and Stewart, G.R. (1992). The carbon canopy economy of the association between cowpea and the parasitic angiosperm *Striga gesnerioides*. *Plant Cell Environ.* **15**: 283-288.
- Kamara, A.Y., Chikoye, D., Ekeleme, F., Omoigui, L.O. and Dugje, Y. (2008). Field performance of improved cowpea varieties under conditions of natural infestation by parasitic weed *Striga gesnerioides*. *Int. J. Pest Management* **54** (3): 189-195.
- Langyintuo, A.S., Lowenberg-DeBoer, J., Faye, M., Lambert, D., Ibro, G., Moussa, B., Kergna, A., Kushwaha, S., Musa, S., and Ntougum, G. (2003). Cowpea supply and demand in West and Central Africa. *Field Crops Research.* **82**: 2215-231.
- Lane, J.A., Bailey, J.A., Butler, R.C. and Terry, P.J. (1993). Resistance of cowpea [*Vigna unguiculata* (L.) Walp] to *Striga gesnerioides* (WILD) vatke, a parasitic Angiosperm. *The New Phytologist* **125**: 405-412.
- Lane, J.A., Moore, T.H.M., Cild, D.V. and Cardwell, K.F., (1996). Characterization of virulence and geographic distribution of *Striga gesnerioides* on cowpea in West Africa. *Plant Dis.* **80**: 299-301.
- Lane, J.A., Moore, T.H.M., Cild, D.V., Cardwell, K.F., Singh, B.B. and Bailey, J.A. (1994). Virulence characteristics of a new race of the parasitic angiosperm *Striga gesnerioides* from Southern Benin on cowpea. *Euphytica* **72**: 183-188.
- Li, J. and Timko, M.P. (2009). Gene-for gene resistance in *Striga*-cowpea associations. *Science* **325** (5944): 1094.

- Michelmore, R.W., Paran, L. and Kesseli, R.V. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. *Proc. Natl. Acad. Sci.* **88**: 9828-9832.
- Moore, T.H.M., Lane J.A., Child, D.V., Arnold, G.M., Bailey, J.A. and Hoffmann, G. (1995). New sources of resistance of cowpea to *Striga gesnerioides*, a parasitic angiosperm. *Euphytica* **84**: 165:174.
- Muller, S., Hauck, C. and Schildknecht, H. (1992). Germination stimulants produced by *Vigna unguiculata* Walp cv Saunders Upright. *J. Plant Growth Regul.* **11**: 77-84.
- Okonkwo, S. N. C. and Nwoke, F. I. O. (1978). Initiation, development and structure of the Primary haustorium in *Striga gesnerioides* (Scrophulariaceae). *Ann. Bot.* **42**: 445-463.
- Omoigui, I.O., Timko, M.P., Ishiyaku, F.S., Ousmane, B., Muranaka, B.S., Kamara, A.Y. and Yeye, M.Y. (2009). Molecular characterization of cowpea breeding line for *Striga* Resistance using FTA technology. *African Crop Science Conference, Uganda* **9**: 527-530.
- Omoigui, L. O., Kamara, A. Y., Massawe, F.S., Ishiyaku, M. F., Boukar, O., Alabi, S.O. and Ekeleme, F. (2007). Evaluation of cowpea genotypes for their reactions to *Striga gesnerioides* in the dry savanna of northern Nigeria. *African Crop Science Conference Proceedings*, **8**: 273-278.
- Ouédraogo, J.T., Maheshwari, V., Berner D.K., St-Pierre, C.-A, Belzile, F.J. and Timko, M.P. (2001). Identification of AFLP markers linked to resistance of cowpea (*Vigna unguiculata* L.) to parasitism by *Striga gesnerioides*. *Theor. Appl. Genet.* **102**: 1029-1036
- Ouédraogo, J.T., Tignegre, J.B., Timko, M.P. and Belzile, F. J. (2002). AFLP markers linked to resistance against *Striga gesnerioides* race 1 in cowpea (*Vigna unguiculata* L.). *Genome* **45**: 787-793.
- Parker, C. and Polniaszek, T.I. (1990). Parasitism of cowpea by *Striga gesnerioides*: Variation in virulence and discovery of a new source of host resistance. *Ann. Appl. Biol.* **116**: 305-311.
- Singh, B.B. (2002). Breeding cowpea varieties for resistance to *Striga gesnerioides* and *Alectra vogelii*. In: *Challenges and opportunities for enhancing sustainable cowpea production*. Fatokum, C.A. (Ed.). pp. 154-166. IITA, Ibadan, Nigeria.
- Singh, B.B. and Emechebe, A.M. (1990). Inheritance of *Striga* resistance in cowpea genotype B301. *Crop Sci.* **30**: 879-881.
- Singh, B.B. and Emechebe, A.M. (1997). Advances in research on cowpea *Striga* and *Alectra*. In: *Advances in cowpea research*. Singh, B.B., Mohan, R., Dashiell, K.E and Jackai, L.E.N. pp.215-224. IITA-Jirca, Ibadan, Nigeria.
- Singh, B.B., Ehlers, J.D., Sharma, B. and Freire-Filho, F.R. (2002). Recent progress in cowpea breeding. In Fatokum, C.A., Tarawali, S.A., Singh, B.B., Kormawa, P.M. and Tamo, M. (eds). *Challenges and opportunities for enhancing sustainable cowpea production*. IITA. Ibadan, Nigeria. pp 22-40.
- Touré, M., Olivier, A., Ntare, B.R., Lane, J.A. and St-Pierré, C.A. (1997). Inheritance of resistance to *Striga gesnerioides* biotypes from Mali and Niger in cowpea [*Vigna unguiculata* (L.) Walp]. *Euphytica* **94**: 273-278.

DETERMINANTS OF HOUSEHOLD FOOD SECURITY IN THE SEKYERE-AFRAM PLAINS DISTRICT OF GHANA

Robert Aidoo
James Osei Mensah
Thomas Tuffour

Department of Agricultural Economics, Agribusiness and Extension,
Kwame Nkrumah University of Science and Technology, Kumasi-Ghana

Abstract:

The study was carried out to assess the determinants of household food security among rural households in Sekyere-Afram Plains District of Ghana. Data were collected from 100 randomly selected households through the use of structured questionnaires. Analytical techniques employed included descriptive statistics to analyze the characteristics of respondent households, and logistic regression model to examine the determinants of food security among the households surveyed. Among the variables considered in the model, household size, farm size, off-farm income, credit access, and marital status were found to significantly influence household food security. Consistent with *a priori* expectation, larger households were found to be food insecure compared with households with smaller sizes, *ceteris paribus*. Also, consistent with findings from previous empirical studies, farm size, off-farm income and credit access were found to have significant positive effect on household food security. The study, therefore, recommended improved access to credit for rural households and diversification of household economic activities to include off-farm income generating businesses to improve food security at the household level.

Key Words: Household Food security, Logistic regression, Ghana

Introduction

Food security and insecurity are terms used to describe whether or not households have access to sufficient quality and quantity of food. Food security issues gained prominence in the 1970s and have since been given considerable attention. Food security is perceived at the global, national, household and individual levels. Food security at global level does not guarantee food security at the national level. Moreover, food security at the national level does not guarantee food security at the household or even the individual level. Ghana's economy until recently had been strongly dominated by the agricultural sector. The agricultural sector is the second largest contributor to the country's GDP after the service sector, yet it is characterized by low productivity and limited competitiveness. This is because the sector is dominated by subsistence and smallholder production units applying mostly basic and low level technology (Duffour, 2010). It is a source of livelihood for majority of Ghana's population. It is the source of raw materials for industries; a major foreign exchange earner and also the main source of food security for the country. Agriculture is the main employer of rural households. It is the small scale farmers who produce for incomes, food and employment.

Notwithstanding the enormous contribution of agriculture to the country, there are problems of food insecurity especially among producers. Food security issues in the country are also affected by both local and international issues. Some local challenges are linked to policies and their implementation, and exclusion of the poor and those affected in decision making. International constraints like policies of donors, trade arrangements, liberalization of the economy, and activities of the extractive industries affect the sector.

In 2000, world leaders committed themselves to the Millennium Development Goals (MDGs) and one aim of the MDGs is to eradicate poverty and hunger. The target is "to reduce by half the proportion of people who suffer from hunger" by 2015. Over 800 million people in the world are food insecure (Gyamfi, 2006). Ghanaians are no exception even though Ghana is endowed with numerous

natural and human resources. Ghana's overall performance in terms of agricultural production and productivity remains inadequate and has failed to make progress on the food security front. Ineffective production techniques, low yielding varieties, inadequate supply of water, among others, are part of the constraints to the achievement of household food security (Gyamfi, 2006).

Food insecurity in Ghana is concentrated in the rural areas. Majority of the Ghanaian rural population chronically suffer from mass poverty in more severe situations than the urban dwellers. In 2009, according to the report by Comprehensive Food Security and Vulnerability Analysis (CFSVA), 19% of rural households were food insecure as compared to 10% of urban households. Under-nourishment and malnutrition are common in rural Ghana and very large proportion of peasant farmers live under the absolute poverty line. Moreover, lack of means of production, and large family size (majority of which are dependants) are the main characteristics of Ghanaian peasant farmers at present. This is why the issue of food insecurity has become the concern of many academicians, political leaders and other professionals today.

Majority of the research works that have been done so far on the issues related to food insecurity in Ghana are very general and consider the problem from national or regional points of view. While aggregate data are generally available at the national level, little work has been done to understand the food security problem at the household level in specific locations/districts. Having national food balance data is not sufficient to understand the food security dynamics in the country. Most agricultural production comes from millions of rural households. Despite the increasing global concern of improving food security, the nature and extent of food security at the household level in rural areas is not well documented. The purpose of this study was, therefore, to investigate the critical determinants of food security in rural areas in Ghana using Sekyere-Afram plains District as a case study.

The main objectives of the study were to:

- Determine the food security situation among households in the study area, and
- Examine the principal determinants of household food security in the study area.

Literature Review

Most of the world's poorest countries are in Africa and many of these face chronic poverty and food insecurity. Agriculture, of which 85-90 per cent is rain-fed in Sub-Saharan Africa, accounts for 35 percent of the region's gross national product (GNP), 40 percent of exports and 70 percent of employment (World Bank, 2000). Clover (2003), Smith (2007), Babatunde et al. (2007), Swaminathan (2008), Oriola (2009), Fayeye and Ola (2007) are some of the works that have examined food security in developing countries. The authors argue that domestic policies in many developing countries have contributed very marginally to food security especially in Africa, and that, despite the growing global food production, hunger, malnutrition and famine are prevalent in many developing countries. From their analysis it is evident that improvement in food production in Sub-Saharan Africa will boost per capita GDP, raise purchasing power and access to food. Their major conclusion is that research is needed on new technologies that are output-driven, ecologically friendly, acceptable and affordable to the resource-poor farmers. Finally, they argue that good governance and stable political governance system will provide an essential and enabling environment for food security in Sub-Saharan Africa.

According to CFSVA (2009), 5 percent or 1.2 million of the Ghanaian population, of whom the majority reside in rural areas, is food insecure. Ghana has seen the number of people living in poverty reduced significantly; in 2005/2006 the share of the population living in poverty was calculated at 28.5% down from 39.5% in 1998/1999. This makes Ghana one of the few countries that are on track to meet Millennium Development Goal (MDG) before the target year of 2015.

Definitions and Concepts of Food Security

Food security is a concept that has evolved over time. As much literature has spiraled, many definitions and conceptual models on household food security have been presented (Smith et al., 1992). There are approximately 200 definitions and 450 indicators of food security (Hoddinott, 1999). In Africa, food crisis in the early 1970s stimulated a major concern on the part of the international donor community regarding supply short falls created by production failures due to drought and desert encroachment (Maxwell, 1992). In 1983, FAO analysis focused on food access, leading to a definition

based on the balance between the demand and supply side of the food security equation: “Ensuring that all people at all times have both physical and economic access to the basic food that they need” (FAO, 1983). In the World Bank (1986) report, *Poverty and Hunger*, this concept of food security is further elaborated in terms of: ‘access of all people at all times to enough food for an active, healthy life.’

At the 1996 World Food Summit 182 nations agreed and adopted a still more complex definition: ‘Food security, at the individual, household, national, regional and global levels. Food security is achieved when all people, at all times, have physical and economic access to sufficient, safe and nutritious food to meet their dietary needs and food preferences for an active and healthy life’ (FAO, 1998). This definition integrates stability, access to food, availability of nutritionally adequate food and the biological utilization of food. As a result, a synthesis of these definitions, with the main emphasis on availability, access, and utilization, serves as working definition in projects of international organizations.

Food security Components

Common to most definitions of food security are the elements of availability, access, utilization and stability or sustainability.

Availability

In this context, availability refers to the physical existence of food, be it from own production or on the markets. On national level food availability is a function of the combination of domestic food stocks, commercial food imports, food aid, and domestic food production, as well as the underlying determinants of each of these factors. Use of the term availability is often confusing, since it can refer to food supplies available at both the household level and at a more aggregate (regional or national) level. However, the term is applied most commonly in reference to food supplies at the regional or national level (Riely *et al.*, 1999).

Access

Access emphasizes on having sufficient resources to obtain appropriate foods for a nutritious diet. It is the way different people can obtain the available food. Normally, we access food through a combination of home production, stocks, purchase, barter, gifts, borrowing or food aid. Food access is ensured when communities and households and all individuals within them have adequate resources, such as money, to obtain appropriate foods for a nutritious diet (Riely *et al.* 1995). Access depends normally on; income available to the household, the distribution of income within the household, the price of food, and other factors worth mentioning are individuals’ access to market, social and institutional entitlement/rights (*ibid.*).

Utilization

Utilization has a socio-economic and a biological aspect. If sufficient and nutritious food is both available and accessible the household has to make decisions concerning what food is being consumed (demanded) and how the food is allocated within the household. In households where distribution is unequal, even if the measured aggregate access is sufficient some individuals may suffer from food deficiency.

Stability

Stability or sustainability refers to the temporal dimension of nutrition security (i.e. the time frame over which food security is being considered). In much of the food security literature, a distinction is drawn between chronic food insecurity—the inability to meet food needs on an ongoing basis—and transitory food insecurity when the inability to meet food needs is of a temporary nature (Maxwell and Frankenberger, 1992).

Determinants of Food security

Factors that affect household food security in various developing countries especially in Africa have been documented in some literature and these factors or determinants are most often than not location-specific (i.e. different study areas were found to have variant attributes as food security determinants with some attributes recurring). The study conducted in Nigeria by Oluwatayo (2008) using probit model found out that sex of household head, educational level, age and income have positive influence on food security whereas household size has negative influence on household food security. Study by Sikwela (2008) in South Africa using logistic regression model showed that per aggregate production, fertilizer application, cattle ownership and access to irrigation have positive

effect on household food security whereas farm size and household size have negative effect on household food security.

Babatunde et al. (2007) is another detailed work on food insecurity in Nigeria. The study utilized a three-stage random sampling technique to obtain a sample of 94 farm households and a cross sectional data in year 2005. Using the recommended calorie required approach; the study revealed that 36 per cent and 64 per cent of the households were food secure and food insecure respectively. The Shortfall/Surplus index showed that the food secure households exceeded the recommended calorie intake by 42 per cent, while the food insecure households fell short of the recommended calorie intake by 38 per cent. A logit regression model estimated showed that household income, household size, educational status of household head and quantity of food obtained from own production were found to determine the food security status of farming households in the study area.

Methodology

Sampling technique

With the creation of the Sekyere Afram Plains District from the old Sekyere East District, the District now covers an estimated area of 3,500.59 square kilometers and has 106 communities and settlements of varying sizes. A two stage random sampling technique was employed in this study. Random sampling was used to select four communities and 25 households per community. Both primary and secondary data were collected through personal interviews with the use of structured questionnaires. The questionnaire used covered the personal characteristics of the farmers, land acquisition, credit access, crops grown, livestock number and household assets such as cutlass, hoe, pick axe, building, etc. Also included in the questionnaire was the USDA Household Food Security Scale (Revised in March 2000) which was used to measure the food security status of households.

Analytical model

The binary logit model was used to investigate the determinants of household food security among the rural households surveyed. The USDA Household Food Security Scale (Revised in March 2000) was used to disaggregate the households into food secure and food insecure households. The dependent variable in this case, food security, was a binary variable which took a value of one if a household was found to be food secure, and zero if otherwise.

A variety of models can be used to establish the relationship between the potential determinants and food security. The study employed the logit model in line with earlier researchers. Following Bogale (2009), the cumulative logistic probability model can be econometrically stated as:

$$P_i = F(Z_i) = \frac{1}{1 + e^{-(\alpha + \sum \beta_i X_i)}} \dots \dots \dots (1)$$

Where:

P_i = the probability that an individual is being food secure given X_i

X_i = a vector of explanatory variables

α & β = regression parameters to be estimated.

e = the base of the natural logarithm

For ease of interpretation of the coefficients, a logistic model could be written in terms of the odds and log of odd. The odds ratio is the ratio of the probability that a household would be food secure (P_i) to the probability of a household not being food secure ($1 - P_i$).

That is:

$$\frac{P_i}{1 - P_i} = e^{Z_i} \dots \dots \dots (2)$$

Taking the natural logarithm of the equation yields:

$$\ln \left(\frac{P_i}{1 - P_i} \right) = Z_i = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m \dots \dots \dots (3)$$

If the error term, ϵ_i is taken into account the equation becomes:

$$Z_i = \alpha + \sum_{i=0}^m \beta_i X_i + \epsilon_i \dots \dots \dots (4)$$

In this study the explanatory variables used in the model included:

X_1 = Age of Household Head (AGE) in years

X_2 = Gender of Household Head (GEND) – Male =1, Female = 0)

X_3 = Marital Status of household Head (MSTAT) – Married =1, Otherwise=0)

X_4 = Household size

X_5 = Education level of Household Head (Number of years of formal education)

X_6 = Farm size (FSZ) in acres

X_7 = Off-farm income activity (If yes =1, otherwise= 0)

X_8 = Credit Access (if yes =1, otherwise= 0)

X_9 = Fertilizer Application (if yes =1, otherwise= 0)

X_{10} = Remittance (GHC)

The parameters of the logistic regression model were estimated using the maximum likelihood approach.

Results And Discussion

Respondent's socio-economic characteristics

Respondents' socio-economic characteristics are presented in Table 1. Male household heads constituted majority (61%) of the sampled people. The modal age was between 40 and 50 years, indicating that a typical farmer interviewed was economically active. There were more married household heads (74%) than those divorced (12%), widowed (9%) and single (5%). As high as 86% of households surveyed had more than four (4) members implying that the average household in the study area had a large size. Meanwhile, distribution of household heads by education revealed that majority of them had a junior high school education (42%). About 62% of household heads interviewed were engaged in off-farm income-generating activities.

Table 1: Household Distribution by socio-economic characteristics

Household Characteristics	Frequency	Percentage
Gender		
Male	61	61
Female	39	39
Age		
<20	2	2
21-30	11	11
31-40	20	20
41-50	31	31
51-60	27	27
61-70	7	7
>70	2	2
Marital status		
Single	5	5
Married	74	74
Divorced	12	12
Widowed	9	9
Household size		
1-3	14	14
4-6	41	41
7-9	35	35
10-12	9	9
>12	1	1
Education level		
None	18	18
Primary	26	26
JHS/Middle school	42	42
SHS/O'Level/A'Level	9	9
Tertiary	5	5
Off-farm income		
Yes	62	62
No	38	38

Source: Field survey, 2011.

Food security status of Households

As depicted in Table 2, majority of the households surveyed (79%) were found to be food insecure and only 21% were food secure.

Table 2: Food security status of the households

Food security status	Frequency	Percentage
Food secure	21	21.00
Food insecure	79	79.00
Total	100	100.00

Source: Field Survey, 2011.

Determinants of Food Security

Table 3 below provides the parameter estimates for the logit model. From the maximum likelihood estimates of the model, the Pseudo R² was 0.67 which implies that about 67% of the likelihood of a household being food secure is strongly explained by the independent variables. The marginal effects of the independent variables were estimated because they are very important for policy and decision making.

Among the 10 variables considered in the model, five were found to have significant impact on household food security. They included marital status, household size, farm size, off-farm income activity and credit access. With the exception of marital status, all the explanatory variables had the expected signs.

Table 3: Parameter estimates of determinants of household food security

Variable	Coefficients	Odds ratio	z-values	p-values
Constant	1.871572			
Age	-0.079428	0.9236445	-1.61	0.107
Gender	-0.149916	0.8607803	-0.13	0.897
Marital status	-2.900491	0.0549962	-1.74*	0.082
Household size	-1.387542	0.2496882	-2.84***	0.004
Education	0.5689745	1.766455	0.95	0.340
Farm size	0.9084391	2.480448	1.86*	0.063
Off farm	3.979105	53.46917	2.58**	0.010
Credit access	2.258334	9.56714	1.82*	0.068
Fertilizer	2.109547	8.244506	1.46	0.145
Remittance	-1.463556	0.231412	-0.96	0.339
Pseudo R²	= 0.6747			
Log Likelihood function	= -16.7147414			
Observations	= 100			

*, **, *** Significant at 10%, 5% and 1% respectively

Marital status was significant at 10% and the coefficient indicated that households headed by unmarried people are more likely to be food secure than those headed by married people. This finding was contrary to findings by Haliu et al. (2007) in Ethiopia and Kaloi et al. (2005) in Uganda, it is quite reasonable. The odds ratio in favour of food security decreases by the factor 0.0549962, and it may be due to the fact that households with married people as heads may have larger household sizes and thus many mouths to feed.

Household size had a negative and significant relationship with food security at the 1% significant level, implying that the probability of food security decreases with increase in household size. The odds ratio in favour of food security decreases by the factor 0.2496882 as the household size is increased by one member. An increase means more people to feed and indirectly reduces income per head, expenditure per head and per capita food consumption. The likely explanation is that in an area where households depend on less productive agricultural land, increasing household size results in increased demand for food. This demand, however, cannot be matched with the existing food supply from own production and this ultimately end up with the household becoming food insecure. This outcome is consistent with the outcome of a study conducted by Sikwela (2008) in Zimbabwe.

Farm size was positively and significantly related to the probability of a household being food secure. Farm size is significant at 10%. The odds ratio in favour of food security increases by the factor 2.480448 when the area under cultivation is increased by one acre. According to Van Der Veen (2010), food production can be increased extensively through expansion of areas under cultivation. With large farm size households can produce more and also diversify. This outcome is consistent with

the finding from a research conducted by Bogale (2009) in Ethiopia. It is however contradictory to the finding by Sikwela (2008) who found farm size to be negatively related to food security in Zimbabwe.

The coefficient of off farm income was positive indicating that there is a positive relationship between off farm work and food security. Off farm work is significant at 5%. Households diversify their incomes by working as daily labourers, petty traders, artisans, and by working as daily construction labourers. Off-farm income generating activities have a paramount significance to diversify the sources of farm households' livelihoods. It enables farmers to modernize their production by giving them opportunity to reduce the risks of food shortage during periods of unexpected crop failures. Income from these off-farm activities is also invested in agriculture to increase production and food availability at the household level.

Credit access was found to be significant and positively related to food security in the study area. Farmers' access to credit will increase the food security status of his household by the factor 9.56714. This may be due to the fact that households which have the opportunity to receive credit would build their capacity to produce more through the use of improved seeds and the adoption of improved technologies. This finding is also consistent with the findings of Bogale (2009) in his study in Ethiopia.

The coefficients of age, gender, education, fertilizer and remittances were not significant in explaining the food security status of households in the study area at the 10% level.

Conclusion And Recommendations

The study has shown that majority (79%) of households in the Sekyere-Afram Plains District were food insecure during the period of the survey. Consistent with *a priori* expectation and findings from previous studies, farm size, off-farm income activity and credit access were found to significantly influence household food security in the study area positively. However, household size was found to influence food security negatively at the household level. Also, households headed by married people were found to have a higher probability of becoming food insecure compared to their counterparts headed by unmarried people. Contrary to *a priori* expectation, educational level of the household head was found to be statistically insignificant in explaining the food security situation at the household level.

In the light of the findings from the study, it is recommended that efforts to improve access to credit by farmers and the promotion of off-farm activities as alternative livelihood options should be pursued by both local and central government structures in Ghana to improve the household food security situation in rural Districts. Policies that will make micro-credit from government and non-governmental agencies accessible to rural farmers will go a long way in addressing their resource acquisition constraints and eventually improving household food security in the country.

References:

- Bickel, G., Nord, M., Price, C., Hamilton, W., and Cook, J. (2000): *Guide to Measuring Household Food Security*, Revised 2000. U.S. Department of Agriculture, Food and Nutrition Service, Alexandria VA.
- Bogale, A. and Shimelis, A.(2009): Household level determinants of food insecurity in rural areas of Dire Dawa, Eastern Ethiopia. *African Journal of Food and Agriculture, Nutrition and Development*. Volume 9, No 9
- Duffour, K. (2010). *Budget Statement and Economic Policy of the Government of Ghana for the 2011 Financial Year*. Pp 49.
- Gebrehiwot, T., Van der Veen, A. (2010): *Effect of policy interventions on food security in Tigray, Northern Ethiopia*. University of Twente Faculty of Geo-Information Science and Earth Observation, Netherlands. Working Papers Series -Paper 6
- Gebrehiwot T.(2008): *Rural Food Security in Tigray, Ethiopia: Policy Impact Evaluation* International Institute for Geo-information Science and Earth observation. Enschede, The Netherlands.
- Hailu, A. and Regassa, N.(2007): *Correlates of Household Food Security in Densely Populated Areas of Southern Ethiopia: Does the Household Structure Matter?* Department of Rural Development and Family Sciences, Hawassa University, Ethiopia.

- Gyamfi S.(2006) : *Food security in Ghana: the development and diffusion of appropriate technologies*, Ghana Statistical Service.
- Kaloi,E., Tayebwa, B. and Bashaasha, B. (2005): Food security status of households in Mwingi District, Kenya .Department of Agricultural Economics and Agribusiness, Makerere University, Kampala, Uganda. *African Crop Science Conference Proceedings*, Vol. 7. pp. 867-873.
- Riely, F., Mock, N., Cogill B, Bailey, L, and Kenefick, E.(1995): Food security indicators and framework for use in the monitoring and evaluation of food aid programmes.
- Sikwela,M. M (2008): *Determinants of Household Food security in the semi-arid areas of Zimbabwe: A case study of irrigation and non-irrigation farmers in Lupane and Hwange Districts*. Thesis for the degree of Master of Science in Agriculture. Department of Agricultural Economics and Extension. University of Fort Hare, Republic of South Africa.

MOLECULAR IDENTIFICATION OF A DESTRUCTIVE PHYTOPATHOGENIC FUNGUS IN TOMATO FIELDS OF IRAN

Parissa Taheri

Atena Pourmahdi

Department of Crop Protection, Faculty of Agriculture, Ferdowsi University of Mashhad, Iran

Abstract:

The necrotrophic fungus *Rhizoctonia solani* is one of the most important, soil-borne pathogens with a wide host range. This pathogen is the casual agent of crown rot, root rot and damping off in tomato producing areas. The purpose of this research was to characterize the *Rhizoctonia* spp. isolates obtained from tomato at different taxonomic levels, to investigate genetic diversity and pathogenicity of these isolates. Sampling was carried out in the main tomato growing regions of Khorassan-Razavi Province during 2011-2012. Morphological characteristics such as the hyphal diameter and conformation, presence of dolipore septum, the number of nuclei per hyphal cell, the size of monilioid cells and sclerotia were investigated for the 39 *Rhizoctonia* spp. isolates obtained from tomato plants. Additionally, 7 *R. solani* isolates obtained from potato were investigated in this research. In total, 46 *Rhizoctonia* spp. isolates were obtained which consisted of 41 *R. solani* isolates and 5 isolates belonged to binucleate *Rhizoctonia* spp. Morphological characterization revealed variability among the isolates. The anastomosis grouping on slide was primarily used for determining anastomosis groups (AGs) and intraspecific groups (ISGs). The PCR-RFLP method was used to confirm the identified AGs and determine the ISGs. In this method, the ITS1, 5.8S, and ITS2 regions of rDNA were amplified using the species-specific primers of *R. solani* including RS1 and RS4. Then the PCR product was restricted using various restriction enzymes such as *Mun* I, *Mse* I, *Hinc* II and *Ava* II. The AG and ISG of each isolate was determined by comparing the restriction pattern of each *R. solani* isolate obtained in this study with the pattern of standard isolates. The PCR-RFLP analysis revealed that *R. solani* isolates belonged to AG2-1, AG3 PT, AG4 HG I, and AG4 HG II with the number of 4, 15, 13, and 9 isolates respectively.

Key Words: Molecular Identification, Destructive Phytopathogenic Fungus, Tomato Fields

Introduction

Tomato foot and root rot, caused by *Rhizoctonia solani* (teleomorph: *Thanatephorus cucumeris*), is one of the most destructive diseases in some of the tomato production areas worldwide [1,2]. To date, this fungal species is subdivided into 14 anastomosis groups (AGs) designated as AG 1 through 13 and bridging isolate (BI) group [3,4]. Several AGs of *R. solani* such as AG 2-1 [5], AG 3 [5,6], and AG 4 HG I [2,7] have been shown to be pathogenic on tomato, the most frequently reported being AG 3. Knowledge about the prevalence and distribution of different AGs is important, since sensitivity to chemical control treatments and probably to other control strategies is varying among AGs [8]. In addition to the similarity of disease symptoms, distinguishing the various *Rhizoctonia* species in culture is difficult due to the lack of stable morphological characters on which to base a definitive classification of the genus *Rhizoctonia* and species assigned to it. Also, identification of the intraspecific groups of various AGs based on anastomosis grouping on a slide is not accurate.

Foot rot symptoms can be characterized by soft rot of the seedlings near the soil surface. Because of the high variability in the *Rhizoctonia* populations, its wide host range, and living as a soil borne pathogen with long-term survival in soil, management of diseases caused by this fungus is difficult [9].

The objectives of the present study was to characterize *Rhizoctonia* spp. associated with tomato diseases in Iran by morphological characterization and hyphal anastomosis in conjunction with molecular techniques.

Materials And Methods

Fungal isolates. Samples from each of the geographically defined tomatrice-growing areas in Iran were collected using transect sampling by walking through the field diagonally. Per field, 10 samples were collected all along the path of the diagonal. Plant tissues of tomato root, crown and leaves with the disease symptoms were surface-disinfected with 0.5% sodium hypochlorite solution for 2 min and rinsed three times with sterile distilled water. Then, pieces of sheath or leaf blade were dried on sterilized filter paper, placed on a petri dish containing acidified water agar (PH 4.5) with 10% lactic acid (AWA), and incubated at 28°C in the dark. After 2 to 3 days, cultures were examined microscopically for hyphal characteristics typical of *Rhizoctonia* spp. . All plated samples readily yielded *Rhizoctonia* spp., and a hyphal tip of each isolate was subcultured onto AWA for further purification. Isolates were transferred to potato dextrose agar (PDA) test tube slants and maintained at 28°C. Following sufficient growth and production of sclerotia, culture tubes were kept at 4°C for short-term storage. Tester strains of *R. solani* were used for AG typing on glass slides. For long-term storage of the isolates, one of the following methods was used: (i) culturing the fungus on sterile barley grains or (ii) storing lyophilized fungal mycelium or sclerotia at 4°C.

DNA extraction. Cultures were grown at room temperature without shaking in 125-ml Erlenmeyer flasks containing 75 ml of potato dextrose broth (PDB), 24 g/liter (Difco Laboratories, Detroit, MI). Mycelium was harvested after 5 days, preceding sclerotium formation. Lyophilized tissue was pulverized in liquid nitrogen. Total genomic DNA was extracted by the hexadecyltrimethylammonium bromide method according to Zolan and Pukkila (46). Total DNA was dissolved in 50 to 200 µl of Tris-EDTA (TE) buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) depending on the size of the DNA pellet, and quantified by spectrophotometry. Dissolved DNA was stored at -20°C until used.

PCR-RFLP of rDNA-ITS region. RFLP of rDNA-ITS was employed to clarify whether isolates of *R. solani* AG1 obtained from rice belonged to subgroup IA, IB, or IC. Tester strains of AG1-IA (CSKA) obtained from rice, and of AG1-IB (B19) and AG1-IC (BV17), both obtained from sugar beet, were included for comparison. Genomic DNA of the isolates was used for PCR amplification of the ITS region using a pair of primers, RS1 (5'-CCTGTGCACCTGTGAGACAG-3' and RS4 (5'-TGTCCAAGTCAATGGACTAT- 3'), using the reaction conditions described by Guillemaut et al. (10). Restriction polymorphisms were detected using two discriminating restriction enzymes, *MseI* and *MunI* (MBI Fermentas, Vilnius, Lithuania) (10). Typical restriction enzyme reactions consisted of 7 µl of PCR product, 10 units of restriction enzyme, 2 µl of reaction buffer, 0.2 µl of bovine serum albumin, and H₂O to a total volume of 20 µl. Reactions were incubated for 3 h at 37°C (or according to manufacturer's recommendations), then stopped by incubation at 70°C for 5 min.

Results and Discussion

The rDNA-ITS region, including ITS1, 5.8S, and ITS2 for 99 isolates of *R. solani* and the tester isolates of *R. solani* was amplified and digested using two discriminating restriction enzymes (*MunI* and *MseI*) (Figures 1 and 2). The three ISGs of *R. solani* AG4 (HG I and HG II) differed from each other in restriction sites generated with restriction enzymes used in this study. The smaller bands were 0.12 and 0.06 kb for the *R. solani* isolates, respectively. In both cases (digestion with *MunI* or *MseI*), restriction fragments less than 0.05 kb were not taken into consideration because they would not be clearly resolved by electrophoresis. The whole PCR-RFLP procedure was repeated twice with similar results and it was capable of separating the *Rhizoctonia* isolates belonging to various taxonomic groups.

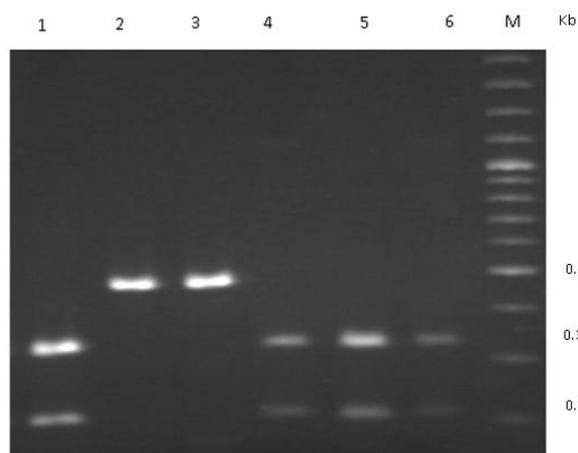


Figure 1. Restriction pattern of *Mun I*

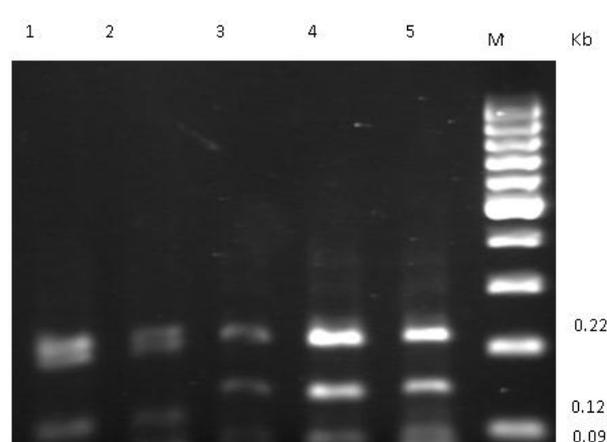


Figure 2. Restriction pattern of *Mse I*

References:

- [1] Solanki MK, Robert AS, Singh RK, Kumar S, Pandey AK, Srivastava AK, et al. Characterization of mycolytic enzymes of *Bacillus* strains and their bio-protection role against *Rhizoctonia solani* in tomato. *Curr Microbiol* 2012. <http://dx.doi.org/10.1007/s00284-012-0160-1>.
- [2] Taheri P, Tarighi S. The Role of pathogenesis-related proteins in tomato-*Rhizoctonia solani* interaction. *J Bot* 2012. <http://dx.doi.org/10.1155/2012/137037>.
- [3] Carling DE, Kuninaga S, Brainard KA. Hyphal anastomosis reactions, rDNA-internal transcribed spacer sequences, and virulence levels among subsets of *Rhizoctonia solani* anastomosis group-2 (AG-2) and AG-BI. *Phytopathology* 2002;92:43e50.
- [4] Sneh B, Burpee L, Ogoshi A. Identification of *Rhizoctonia* species. St. Paul, MN: The American Phytopathological Society Press; 1991.
- [5] Misawa T, Kuninaga S. The first report of tomato foot rot caused by *Rhizoctonia solani* AG-3 PT and AG-2-Nt and its host range and molecular characterization. *J Gen Plant Pathol* 2010;76:310e9.
- [6] Charlton ND, Carbone I, Tavantzis SM, Cubeta MA. Phylogenetic relatedness of the M2 double-stranded RNA in *Rhizoctonia* fungi. *Mycologia* 2008;100:555e64.
- [7] Kuramae EE, Buzeto AL, Ciampi MB, Souza NL. Identification of *Rhizoctonia solani* AG 1-IB in lettuce, AG 4 HG-I in tomato and melon, and AG 4 HG-III in broccoli and spinach, in Brazil. *Eur J Plant Pathol* 2003;109:391e5.
- [8] Champion C, Chatot C, Perraton B, Andrivon D. Anastomosis groups, pathogenicity and sensitivity to fungicides of *Rhizoctonia solani* isolates collected on potato crops in France. *Eur J Plant Pathol* 2003;109:983e92.
- [9] Taheri P, Tarighi S. Riboflavin induces resistance in rice against *Rhizoctonia solani* via jasmonate-mediated priming of phenylpropanoid pathway. *J Plant Physiol* 2010;167:201e8.
- [10] Guillemaut C, Herman VE, Camporota P, Alabouvette C, Molard MR, and Steinberg C. Typing of anastomosis groups of *Rhizoctonia solani* by restriction analysis of ribosomal DNA. *Can. J. Microbiol* 2003;49:556-568.

MODEL POLICY DESIGN FOR THE BEEF CATTLE RANCH DEVELOPMENT IN SOUTH SULAWESI

Machmud Achmad, PhD Candidate

School of Management and Business, Bogor Agricultural University, Indonesia

Sri Hartoyo

Lecturer at Graduate School of Management and Business, Bogor Agricultural University, Indonesia

Bustanul Arifin, Professor

Lampung University (UniLa), Indonesia

Muhammad Said Didu

Researcher and Engineering at Agency of Assessment and Application of Technology, Indonesia

Abstract:

The main objective of this study is to examine the existing condition of the ranchers, namely economic and social characteristics in related to the development of beef cattle. It examines factors that influences of policies on developing cattle beef and while designing integrated planning systems in South Sulawesi. The method for this study includes qualitative and quantitative approaches using survey method to obtain primary data. Data were analyzed and presented in the form of descriptive tables graphs with analytical tools policy.

Results were obtained from choices of recommendations, elaborated in the form of strategy and policy implications based on expert opinions (expert judgment), Internal and External Matrices with the SWOT analysis and the QSPM. Results show that current position for beef cattle in South Sulawesi is in grow and build. Therefore, by analyzing all the factors from SWOT, four strategies were developed to support the beef cattle. Using the QSPM, results further suggests that the best strategy to develop is the first strategy.

This strategy recommends that in order to develop the beef cattle in South Sulawesi, there should be an integrating system between the beef cattle rancher and crop (rice and corn). The strategy also implies that in order to achieve this goal, technology improvement and innovation play important keys, specially technology for the feed processing system.

Key Words: Beef Cattle, IFE EFE matrix, SWOT and QSPM, South Sulawesi

Introduction

Animal husbandry is part of the agricultural sector that represents the important sub-sector to support the people's economy. Livestock commodities are prospectful to be developed, an example would be the Beef cattle industry which are among the producers of protein-rich meat. This is because of the advantages of local resources that exist in different parts of Indonesia. Dwiyanto and Priyanti(2006) noted several problems in the development of beef cattle in Indonesia which are: (1)lower livestock productivity, (2)limited availability of local cows, (3)lack of productive human resources and lower level of knowledge, (4) non-continuous availability of feed, especially in dry seasons, (5) less optimal farm system, and (6) inefficient marketing.

The South Sulawesi Provincial Office of Livestock Services will launch a program entitled *A Movement for the Targeted 1 Million Cattle Population to 2013* in support the national program for self-sufficiency in beef for 2014. This is consistent with the South Sulawesi Provincial Office of Livestock Services, as the vision jurisdiction of for the main supplier of beef cattle and cows. The vision implies that livestock sector can increase population and genetic quality of cattle. Furthermore, cows are expected to have similar quality as local resources. The Bali cattle, which has been developed in South Sulawesi have been proven to be well-adapted to the local agro-climatic conditions. Furthermore, the Brahman and Limousin cattle were also raised to support the increase in number of beef cattle and cow population.

Beef cattle farming in Indonesia serve a very strategic functions, especially in providing employment opportunities for rural communities, as a producer of meat for human well-being and to meet customer needs, all of which intended to improve the people's quality of life and of intellectual (Samosa 2006). Therefore, empowering rural communities through community-based beef cattle breeding is a necessary, particularly in the province of South Sulawesi. The achievement of Accelerated Beef Sufficiency Program (P2SDS) 2014 requires more support to establish steps in order to implement the program. This is the reason for the author to conduct the study.

Beef cattle ventures in Indonesia generally took the form of traditional ranches and serves only as a sideline, which thus carried out on less optimal basis. This contributes relatively small income to farming households. Such non-optimal beef cattle farming also relates, to the lack of labor employed, green fodder, capital, and marketing. Thus placing farmers in a non-bargaining position and explains the less optimal revenue from marketing.

The policy of beef cattle development has long been enforced by the government. A study by Nasution (1983) indicated that for cattle development efforts two policies has been implemented; extensification and intensification. The former put emphasize on the increase in number of cattle supported by the procurement of improved quality cows, disease prevention, business counseling and coaching, credit assistance, procurement and improvement of feed quality, marketing and partnership building with the stakeholders. The latter was implemented by increasing the economically viable production, supported by a particular combination through the use of superior cows and feed, as well as good management.

The study aimed to examine the existing characteristics of the farmers, i.e. their social and economic characteristics in relation to the development of beef cattle farming, to determine factors influencing the policy of beef cattle farming, and to design the policy of beef cattle farming development in South Sulawesi. The study is intends to contribute the development of science, particularly in management discipline and beef cattle business; to contribute data and information and ideas in relation to sustainable farm resource management planning; to provide inputs for policy makers, both at the central and local governments, and to develop the management plan for beef cattle development in order to fulfill the national food security needs.

Development of beef cattle was collaboratively undertaken by the government, communities (small scale farmers), and the private sectors. The government sets the rules, facilitates and oversees the supply and availability of products, quantity and quality, to meet the requirements of *halal*, safe, nutritious, and health concern. Private and public play their role in bringing about the adequacy of livestock products through the production, importation, processing, marketing, and distribution of beef cattle (Bamualim *et al.* 2008).

According to Hartono (2009), in order to increase the income and to protect the people's breeding farm, there are some policies measures take that should place. They include tariff, subsidies, capital, institutional, and maintenance system. Agribusiness policy analysis is a process which synthesizes information recommendations to design agricultural policy option. Policies analysis is a process to synthesize mixed information, derived from the research, mass media or laws which then formulate further a policies to encouraging the advancement of agriculture and plantation. Therefore, it requires a comprehensive formulation to make it one of the driving forces of economic growth in Indonesia based on the economic results of the agricultural sector (Nugrohadi, 2009).

For animal husbandry in particular, livestock agribusiness is an integrated and comprehensive system of livestock management that covers all activities ranging from manufacturing and distribution of livestock production facilities (saproak), production activities (breeding), storage and processing, and distribution and marketing of farm products, supported by the supporting institutions (Rahardi and Hartono, 2000). According to Saragih (2000), livestock-based agribusiness is a system that includes four subsystems; up stream, on-farm, down-stream agribusiness and supporting institutions.

Data And Methodology

This study uses two types of data which are primary and secondary data. Primary data were obtained using questionnaires, interviews and focus group discussions. Secondary data were collected by conducting literature review to support, complement, and enhance primary data.

Descriptive analysis is used to obtain in-depth overview of the study object. To help explain the results of this analysis, the information will be presented in the form of labels, pictures, or matrix, according to the results obtained. The descriptive analysis in this study is used to describe the results

of interviews and questionnaires on beef cattle development policies. Analysis of the above data is processed using SPSS software.

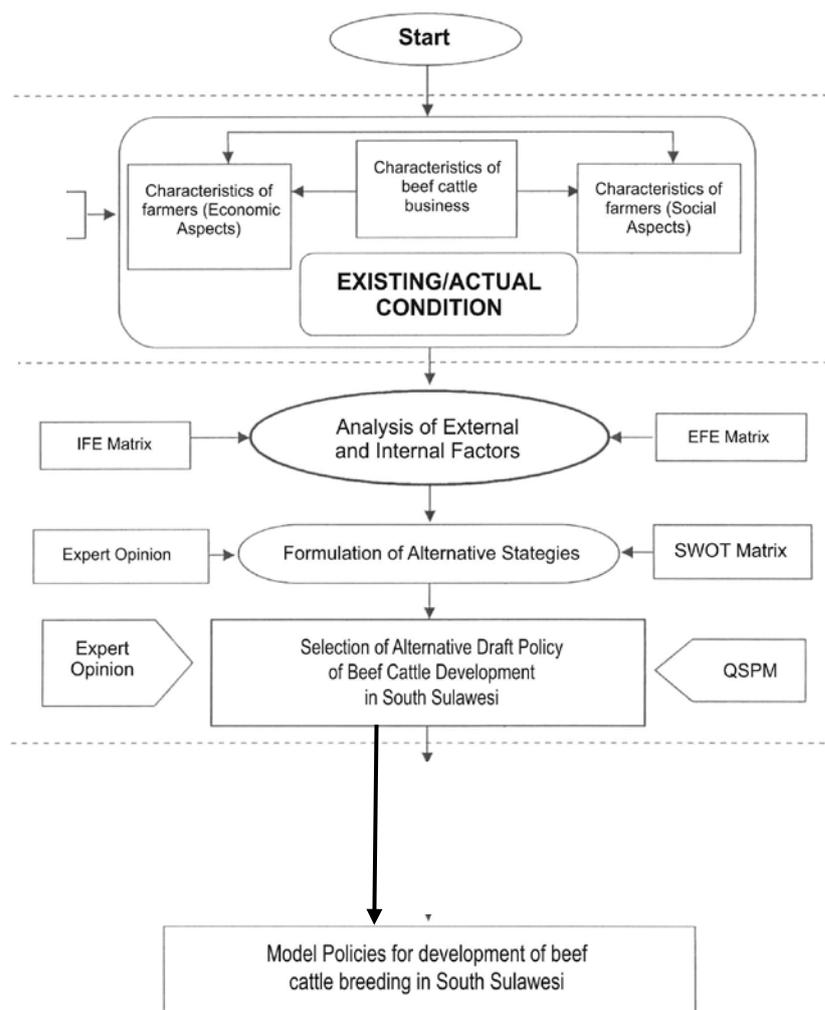


Figure 1. Research Implementation Flowchart

External Factor Evaluation (EFE) Matrix

External Factor Evaluation (EFE) Matrix is used to evaluate the external factors that determine the success of a company in competition. The relevant external data are collected for analysis. These factors are related to economic, social, cultural, demography, environment, politic, government, legal, technology, and competition in industrial markets where the company is located, and other relevant external data. This is important because external factors have direct or indirect influence on the company.

Internal Factor Evaluation (IFE)

Internal Factor Evaluation (IFE) Matrix is used to determine the internal factors that affects the competitiveness of a company, which deals with the strengths and weaknesses deemed important. Data and information on the internal aspects of the company can be obtained from companies functions, such as aspects of management, finance, human resources (HR), marketing, information systems, production and operation.

IE Matrix

Internal-External (IE) matrix is used to evaluate external factors (opportunities and challenges) and internal factors (strengths and weaknesses) that results in a matrix form consisting of 9 cells that in principle can be grouped into 3 main strategies, namely:

1. Growth strategy is the company's growth (cells 1, 2 and 3) or diversification (cells 7 and 8).

2. Stability strategy is a strategy adopted without changing the direction of the predetermined strategy.
3. Retrenchment strategy is an effort to minimize or reduce the work done by the company (cells 3, 6, and 9).

The Internal-External (IE) Matrix is illustrated in Figure 2 below:

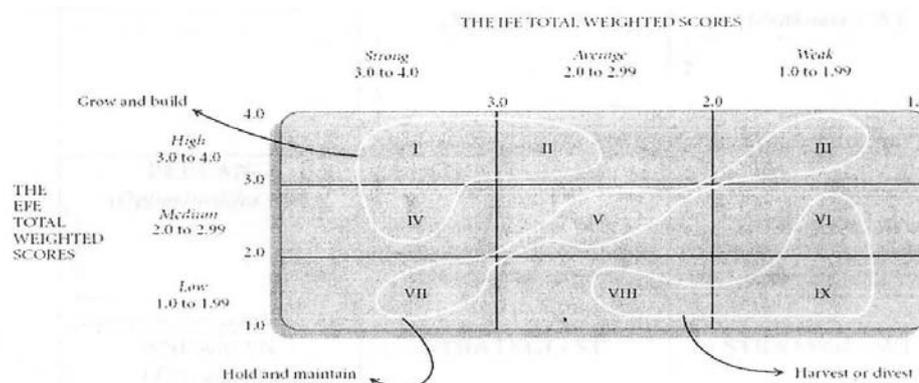


Figure 2: Internal-External (IE) Matrix (David, 2009)

IE matrix is based on two key dimensions: EFI total weighted score is on the horizontal axis and the EFE total weighted score is on the vertical axis. Of the total weighted score, each division IE Matrix at the corporate level can be composed, on the horizontal axis of IE Matrix, the EFI total weighted score from 1.0 to 1.99 indicates a weak internal position; score from 2.0 to 2, 99 is considered moderate, and score of 3.0 to 4.0 considered as strong. This hold true for vertical axis, where the EFE total weighted score of 1.0 to 1.99 considered as low; 2.0 to 2.99 as moderate, and 3.0 to 4.0 as high.

SWOT Analysis

SWOT Analysis is a powerful analytical instrument if used accordingly. "SWOT" stands for Strengths, Weaknesses, Opportunities, and Threats. The strength and weaknesses can be found in an organizational body, including the particular business unit, while opportunities and threats are environmental factors that confront organizations, companies or business units in question (Siagian, 2004).

Analysis instrument used in this phase is the matrix of SWOT (Strengths-Weaknesses-Opportunities Threats). Matrix combines the opportunities and threats that are being faced can be further tailored according to the existing strengths and weaknesses to produce SO alternative strategy, and WO, WT, and ST strategies.

SWOT matrix scheme consists of nine cells; where there are four cells of main factors (external and internal), four cells of strategy, and one cell that is always left empty (Figure 3).

	Strengths (S)	Weaknesses (W)
	1 n-th (obtained from EFE Matrix)	1 n-th (obtained from EFE matrix)
Opportunities (O)	S-O Strategies	WO Strategies
1 n-th (obtained from IFE matrix)	Develop strategies that use the strengths for addressing threats	Develop strategies that minimize the weaknesses to take advantage of an opportunity
Threats (T)	ST Strategies	WT Strategies
1 n-th (obtained from IFE matrix)	Develop strategies that use the strengths to cope with the threats	Develop strategies that minimize the weaknesses to cope with the threats

Figure 3. Schematic illustration of SWOT Matrix (David, 2009)

SWOT matrix preparation steps are described as follow

- a) Each of these external factors (opportunities and threats) and internal factors (strengths and weaknesses) of the EFE and IFE matrix are included in the SWOT matrix.
- b) Using in-depth discussions with experts, adjustments are then carried out between external and internal factors to produce and to formulate several alternative drafts of policy of on farms beef cattle development in South Sulawesi. These are intended to: match internal strengths with external opportunities (SO strategies); match internal weaknesses with external opportunities (WO strategy); match internal strengths with external threats (ST strategy) and to match internal weaknesses with external threats (WT strategy)

QSPM (*Qualitative Strategic Planning Matrix*) Analysis

QSPM (Qualitative Strategic Planning Matrix) analysis indicates objectively best alternative strategies that begins by determining the key success factors of the external environment with External Factor Evaluation (EFE) Matrix and finding the key success internal factors to companies with Matrix Internal Factor Evaluation (IFE) as the input of stage I. The next step is adjustment or phase II, which is analyze strengths, weaknesses, opportunities and threats using the SWOT matrix, analysis of internal and external with Internal-External (IE) Matrix and positions the company under investigation in one of the existing quadrants. QSPM uses input from the analysis of stage I and the results of phase I and phase II analysis adjustment to determine objectively among alternative strategies existed or referred to as phase III.

Furthermore, David (2009) noted that QSPM is an instrument that allows strategists to evaluate alternative strategies objectively, based on the internal and external key success factors that have been identified. Like other instruments of strategy formulations, strategy formulation using QSPM method requires good intuitive judgment.

The decision stage is a stage to determine which strategy is feasible and the best alternative strategies, using Quantitative Strategic Planning Matix (QSPM) or Quantitative Strategic Planning Matrix. QSPM uses the analysis results of the input and matching phases. QSPM main components are: a) the key factors, b) alternative strategies, c) weights, attractiveness Score (AS), e) total attractiveness score (TAS), and f) sum total of Attractiveness Score. Attractiveness score is obtained by determining the numeric score indicating the relative attractiveness of each strategy in particular alternative set. Attractiveness scores were assigned to each strategy to determine the relative attractiveness of one strategy over the other. On the other contrary, the total attractiveness score is the score indicating a relative attractiveness of each alternative strategy that considers the impact of external and internal factors in that line. The higher the total score of attractiveness, the more attractive the alternative strategies will be.

Results And Discussions

Analysis Results for Internal and External

External Factor Evaluation (EFE Matrix)

The Matix External Factor Evaluations (EFE) uses the weighting scoring system to identify the value opportunity weight and threat for beef cattle producers in South Sulawesi. In Table 1, results show that government policy and programs has the highest value weight score of 0.082, compared to other external factors, meaning that government plays an important role in supporting the development of beef cattle program in South Sulawesi. The total weight score which is 3.487 is higher than average score of effective strategy of 2.5, implies that the strategies used are already effective by using existing opportunity while minimizing threat. This is consistent with Nugroho (2006) findings which states that the development of animal husbandry as a part of agricultural development will be associated with the reorientation of agricultural development policy. Animal husbandry development has new paradigms, namely alignment to people in general, responsibility delegation, structural change, and people empowerment. Therefore, it is necessary to formulate strategies and policies that are comprehensive, systematic, integrated—both vertically and horizontally—competitive, sustainable and decentralized.

Table 1 EFE Matrix (*External Factor Evaluations*)

External Strategies Factor	Bobot	Rating	Score
Opportunities			
Improved knowledge of ranchers and technology development	0.072	3.88	0.280
Government support to develop beef cattle	0.082	4.00	0.328
Develop opportunities to work	0.061	3.88	0.236
Availability of sufficient transportation system	0.058	3.13	0.182
High potential feed availability	0.077	3.98	0.335
Crop intensification with hay being used	0.074	3.93	0.304
Increasing trend for beef consumption and demand	0.067	3.50	0.234
Mutual benefit for partnership	0.065	3.50	0.228
High number of population	0.055	3.75	0.207
Opportunities Totally			2.335
Threats			
Availability of imported product with higher beef quality	0.063	3.00	0.188
High transportation cost	0.060	2.75	0.179
Global marketing toward free marketing	0.065	3.25	0.178
Low bargaining position of beef cattle rancher which reduces economic scale	0.064	2.75	0.207
Tendency of people with higher education to find work else where	0.061	2.63	0.168
Low level of cattle safetyness	0.077	2.63	0.203
Threats Totally			1.123
Total			3.457

Internal Factor Evaluation (IFE matrix)

Internal Factor Evaluations (IFE) matrix is being used to determine the weight value for strength and weakness for internal factors affecting the development of beef cattle in South Sulawesi. From Table 2, it can be seen that the highest weight score is 0.666 which implies that factors that are effective are innovation, technology, maintenance and facilities. These factors are important internal factors which are effective to develop the beef cattle in South Sulawesi to support the food estate program. The result also shows that the sum of total weight score is 2.603. Therefore, it can be concluded that the strategy of developing beef cattle in South Sulawesi has been effective in using the strength and minimizing weakness factors which had contributed to negative impact.

Table 2 IFE Matrix (*Internal Factor Evaluations*)

Internal Strategies Factor	Bobot	Rating	Skor
Strength			
Availability of local cattle which are adaptive and productive (Bali cattle, cross PO)	0.059	2.75	0.162
Availability of land and paddy/corn waste used for feed	0.064	3.00	0.191
Availability of cattle facilities and maintenance	0.066	2.67	0.176
Government support from central, district or local consisting developing programs (such as GOS, SIPT, PPBP, PUTKATI, etc)	0.065	2.83	0.185
High interest for cattle beef development from society	0.056	2.50	0.140
Availability of cattle beef institutions as a leader	0.052	2.58	0.135
Availability of flexibility technology innovations	0.066	2.75	0.182
Availability of cattle beef groups	0.054	2.75	0.148
Strength Totally			1.582
Weakness			
Financial limitation for cattle beef ranchers	0.064	1.92	0.123
Limitation of local cattle breed	0.060	2.08	0.125
Uncoordination between institutions involved	0.050	1.67	0.084
Low knowledge and limitations for human resources	0.063	1.58	0.100
Unavailability support from financial institutions	0.058	1.67	0.096
Limitations of supporting institutions related to cattle beef	0.055	1.92	0.105
Unoptimized and inconsistency of programs related	0.067	1.58	0.106
Uncoordinated marketing system	0.053	1.50	0.080
Rancher's mindset where cattle are as if saving product	0.048	1.67	0.081
Weakness Totally			1.021
Total			2.603

Suryana (2009) in his studies also stated that in order to enhance the role of beef cattle as meat suppliers and livestock income sources, it is advisable to apply an intensive maintenance system with an improved feed management and improved quality of cattle with disease control. Improvement of reproduction was conducted by IB and early weaning of calves to shorten the calving interval. As for the improvement of genetic quality of the female calves, it is suggested to keep them in the breeding area for subsequent use as grading up cattle. Increased interest and motivation of cattle ranchers to expand their business can be facilitated through incentives in production.

Mapping of Internal External Matrix (IE matrix)

The EFE matrix and IFE matrix which have been completed using the weighting scoring system, are then being integrated into the external and internal matrices. These matrix shows that the mapping or position for beef cattle development in South Sulawesi, considering the strength and weaknesses factors involved. Figure 4 shows that based on weighted score and the evaluation of internal and external factors, the position for cattle beef development in South Sulawesi lies at the position of grow and build. The position of 2.603 and 3.457 is the area were based on external and internal factors, the government policy should be aimed at intensive programs, such as market penetration, market development and developing products. The other government policy should also aimed at integration programs such as backward integration, forward integration and product integration (David, 2009). In addition policy aimed at beef cattle development in South Sulawesi for food security program is highly required.

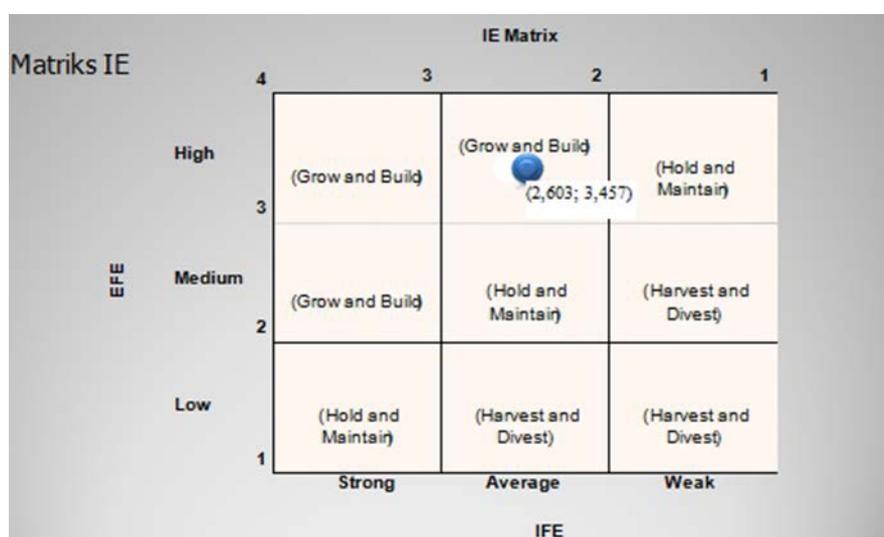


Figure 4. Mapping of Internal External Matrix for beef cattle in South Sulawesi

An alternative incentive government policy that can be implemented for product development in South Sulawesi, is product diversification, which are programs involving processing meat, such as beef bacon, beef jerky, beef *abon*, sausages or other tradition processed meat products. These type of programs could help improved local ranchers wellbeing and income. Integrative government policies could also develop an increase beef cattle program in South Sulawesi. An example would be a policy that integrates beef cattle with crops. The program involves using corn and rice waste, which are potential for cattle feed, using improved technology. Other government policies should focus in order to increase the development of beef cattle, with other stakeholders such as government, companies, ranchers or other local rancher groups.

Formulating Alternative Strategies

In order to formulate alternative strategies based on external and internal factors, these alternative strategies are being formulated using the SWOT method. Expert opinions are being used to formulate these strategies, which were then being analyzed based on factors of strength, weakness, opportunity and threat in the South Sulawesi faced by beef cattle development. The results can be seen on Table 3.

Results using SWOT Matrix show four alternatives developing policy for cattle beef in South Sulawesi. These 4 which are considered to be potential policies to be developed by ranchers in South Sulawesi, are as follow:

1. Developing an integration strategy involving ranchers and crops (rice and corn). This can be achieved using potential raw materials for feed with government support (central government, province or local) by using new developed technology or innovation specially for feed processing.
2. Increasing product and cattle quality and meat through several technology innovations, product diversification in order to increase market target and increasing the role of safe guards to guarantee the safety of cattle.
3. Optimizing government role by supporting programs which increase agribusiness potential with an integration system, increasing knowledge and ability of ranchers by training programs and assistance.
4. Limiting cattle import or products while increasing the partnership with financial institutions to increase rancher's finance.

Table 3 SWOT Matrix strategy development formulation for cattle beef in South Sulawesi

<p style="text-align: center;">SW</p> <p style="text-align: center;">OT</p>	<p>Strength</p> <ol style="list-style-type: none"> 1) Availability of local cattle which are adaptive and productive (Bali cattle, cross PO) 2) Availability of land and paddy/corn waste used for feed 3) Availability of cattle facilities and maintenance 4) Government support from central, district or local consisting developing programs (such as GOS, SIPT, PPBP, PUTKATI, etc) 5) High interest for cattle beef development from society 6) Availability of cattle beef intitutions as a leader 7) Availability of flexibility technology innovations 8) Availability of cattle beef groups 	<p>Weakness</p> <ol style="list-style-type: none"> 1) Financial limitation for cattle beef ranchers 2) Limitation of local cattle breed 3) Uncoordination between intitutions envolved 4) Low knowledge and limitations for human resources 5) Unavailability support from financial intitutions 6) Limitations of supporting institutions related to cattle beef 7) Unoptimized and inconsistency of programs related 8) Uncoordinated marketing system 9) Rancher's mindset where cattle are as if saving product
<p>Opportunity</p> <ol style="list-style-type: none"> 1) Improved knowledge of ranchers and technology development 2) Government support to develop beef cattle 3) Develop opportunities to work 4) Availability of sufficient transportation system 5) High potential feed availability 6) Crop intensification with hay being used 7) Increasing trend for beef consumption and demand 8) Mutual benefit for partnership 9) High number of population 	<p>SO Strategies</p> <p>SO 1. (S1,S2, S3,S4,O1,O2) Developing an integration strategy involving ranchers and crops (rice and corn). This can be achived using potential raw materials for feed with government support (central government, province or local) by using new developed technology or innovation specially for feed processing</p>	<p>WO Strategies</p> <p>WO 1. (W3, W4, W5, O1, O2,O6) Optimizing government role by supporting programs which increase agribusiness potential with an integration system, increasing knowledge and ability of ranchers by training programs and assistance</p>
<p>Threat</p> <ol style="list-style-type: none"> 1) Availability of imported product with higher beef quality 2) High transportation cost 3) Global marketing toward free marketing 4) Low bargaining position of beef cattle rancher which reduces economic scale 5) Tendency of people with higher education to find work else where 6) Low level of cattle safetytness 	<p>ST Strategies</p> <p>ST 1. (S1,S2,S3, S7, T1,T4,T6) Increasing product and cattle quality and meat through several technology innovations, product diversification in order to increase market target and increasing the role of safe guards to guarantee the safetytness of cattle</p>	<p>WT Strategies</p> <p>WT. 1. (W1,W2, W3, W5, T1,T5,T6) Limiting cattle import or products while increasing the partnership with financial institutions to increase rancher's finance.</p>

Choosing Priority Strategy

Table 4 shows the results for choosing the priority strategy.

Table 4. QSPM Results for Choosing Priority Strategy

Strategies Factor	Strategies								
	Bobot	Strategy I		Strategy II		Strategy III		Strategy IV	
		AS	TAS	AS	TAS	AS	TAS	AS	TAS
Strength									
Availability of local cattle which are adaptive and productive (Bali cattle, cross PO)	0.059	4	0.236	3	0.177	4	0.236	4	0.236
Availability of land and paddy/corn waste used for feed	0.064	4	0.254	4	0.254	4	0.254	4	0.254
Availability of cattle facilities and maintenance	0.066	4	0.264	3	0.198	3	0.198	3	0.198
Government support from central, district or local consisting developing programs (such as GOS, SIPT, PPBP, PUTKATI, etc)	0.065	4	0.261	4	0.261	4	0.261	4	0.261
High interest for cattle beef development from society	0.056	4	0.225	3	0.168	3	0.168	3	0.168
Availability of cattle beef intitutions as a leader	0.052	4	0.208	4	0.208	3	0.156	3	0.156
Availability of flexibility technology innovations	0.066	4	0.264	4	0.264	4	0.264	4	0.264
Availability of cattle beef groups	0.054	3	0.162	4	0.216	3	0.162	3	0.162
Weakness									
Financial limitation for cattle beef ranchers	0.064	4	0.256	3	0.192	3	0.192	3	0.192
Limitation of local cattle breed	0.06	4	0.240	4	0.240	4	0.240	4	0.240
Uncoordination between intitutions envolved	0.05	4	0.202	4	0.202	3	0.151	3	0.151
Low knowledge and limitations for human resources	0.063	4	0.253	3	0.189	3	0.189	3	0.189
Unavailability support from financial intitutions	0.058	3	0.173	4	0.230	3	0.173	3	0.173
Limitations of supporting institutions related to cattle beef	0.055	3	0.164	4	0.219	3	0.164	3	0.164
Unoptimazed and inconsistency of programs related									
Uncoordinated marketing system	0.067	4	0.267	4	0.267	4	0.267	4	0.267
Rancher's mindset where cattle are as if saving product	0.053	3	0.160	3	0.160	4	0.213	4	0.213
Weakness	0.048	3	0.145	3	0.145	3	0.145	3	0.145
Opportunities									
Improved knowledge of ranchers and technology development	0.072	4	0.290	4	0.290	4	0.290	4	0.290
Government support to develop beef cattle	0.082	4	0.328	4	0.328	4	0.328	4	0.328
Develop opportunities to work	0.061	4	0.244	2	0.122	2	0.122	2	0.122
Availability of sufficient transportation system	0.058	3	0.174	3	0.174	3	0.174	3	0.174
High potential feed availability	0.077	4	0.307	4	0.307	4	0.307	4	0.307
Crop intensification with hay being used	0.074	4	0.295	4	0.295	4	0.295	4	0.295
Increasing trend for beef consumption and demand	0.067	3	0.201	3	0.201	3	0.201	3	0.201
Mutual benefit for partnership	0.065	3	0.195	4	0.260	3	0.195	4	0.260
High number of population	0.055	2	0.110	2	0.110	3	0.166	2	0.110
Threats									
Availability of imported product with higher	0.063	3	0.188	3	0.188	4	0.250	3	0.188

beef quality									
High transportation cost	0.060	3	0.179	3	0.179	3	0.179	3	0.179
Global marketing toward free marketing	0.065	2	0.130	3	0.194	3	0.194	3	0.194
Low bargaining position of beef cattle rancher which reduces economic scale	0.064	3	0.191	3	0.191	3	0.191	3	0.191
Tendency of people with higher education to find work else where	0.061	3	0.183	3	0.183	2	0.122	3	0.183
Low level of cattle safety	0.077	2	0.155	3	0.232	4	0.310	3	0.232
Total			6.901		6.844		6.757		6.687

QSPM (*Quantitative Strategic Planning Matrix*) is made in order to choose and determine which strategy is the best to recommend to cattle beef development in South Sulawesi. Based on SWOT Matrix analyses, results, these strategies are then being chosen to be implemented to the real condition of beef cattle.

The formulation of these strategies are based on interviews of experts. Values given by experts are then being, to find the whole score of all criteria. The highest score will be the priority strategy (Table 1 and Table 2). From all four strategies, based on QSPM matrix results (Table 1 and Table 2) shows that the first strategy or priority will be **Strategy 1**. The value from QSPM, based on Total Alternative Strategy (TAS) is **6,901**. This priority strategy includes ; increasing and developing a model between cattle beef and crop (paddy and corn), based on the potential availability of feed. This model should also be supported by government (central, district, local) using advanced technology which aimed at feed processing innovations. Table 4 shows these results

These findings are consistent with results from Ananto (2011) which suggested that strategies should be integrated in order to succeed. Furthermore in his research, concluded that the prerequisites for the fulfillment of an integrated plan for self-sufficiency in beef production include: (1) trading system which is conducive to the creation of value-added livestock industry, (2) defined policy of sectoral farm program, and (3) the availability of budget for cow-calf operation, breeding and the development of the breeding areas. As for the implementation phase, the institutions that play most significant roles are Coordinating Ministry for Economic Affairs, together with the Ministry of Agriculture and Ministry of Commerce (Ananto, 2011)

Concluding Remarks

The study shows that the condition of cattle beef development in South Sulawesi is promising. The IFE and EFE matrices, concludes that based on weighted score and the evaluation of internal and external factors, the position for cattle beef development in South Sulawesi lies at the position of grow and build. Therefore, government policies should focus on increasing and developing cattle beef programs, incentives and supporting policies.

Furthermore, based on interviews with experts on strength, weakness, opportunity and threat factors, four strategies were then being formulated in order to develop the cattle beef in South Sulawesi. Thus from these four strategies provided, QSPM matrix results shows that the first strategy or priority will be **Strategy 1**. This strategy has the highest value compared to the other strategies, based on Total Alternative Strategy (TAS). This priority strategy includes ; increasing and developing a model between cattle beef and crop (paddy and corn), based on the potential availability of feed. This model should also be supported by government (central, district, local) using advanced technology which aimed at feed processing innovations.

This study recommends that the government should play more active role, specially institutions that are involved directly with the cattle beef programs. There should be an integrated, coordinated and consistency within the program. The institutions related to cattle beef rancher should be more aggressive in recruiting and maintaining ranchers. However, the study also recommends that the government should continue with existing programs that had been applied.

References:

- Ananto N, Eriyatno, Marimin, Arief D. Integration For Planning Livestock Development Model. 2011
- Amar AL. Grass Feed Availability Strategy To Develop Cattle Beef In South Sulawesi. *Journal for Agriculture Research and Development* 24 (1): 72–79.. 2008
- Bautista, M Romeo. Agricultural-Based development : A Sam Perspective On Central Vietnam,

Economic Journal, XXXIX,p.112-132. 2001

Daryanto A.. Competition Dynamics For The Livestock Industry. IPB Press. Bogor. 2009

David FR. . *Strategic Management: Concept 10th edition*. 2005.

Gittinger, J. P.. *Agriculture Projects Economic Analyze*. 2nd Edition. UI Press, Jakarta.. 1986

Hendricson J. Interactions in integrated US agricultural systems: *The past, present and future Renewable Agriculture and Food Sistems*: 314-324. 2008

Nugroho BA.. Livestock Agribusiness Development Support For The Productive Economics (Study at North Sulawesi). *Journal For Agriculture Research and Development* (16): 62–72.. 2006.

Suryana. Development Of The Cattle Beef Business Agribusiness Based With The Partnership System. *Journal For Agriculture Research and Development* 28 (1) : 29-37. 2009

DETERMINATION OF FARMERS' GOAL STATEMENTS PRIORITIES AND FACTORS EFFECTIVE ON DECISION (AN APPLICATION OF FUZZY PAIRWISE COMPARISON ANALYSIS)

Ahmet Sahin, Assoc. Prof.

University of Kahramanmaraş Sütçü Imam, Faculty of Agriculture,
Department of Agricultural Economics, Kahramanmaraş, Turkey

Cihat Gunden, Assoc. Prof.

Dr. Murat Cankurt, Assis. Prof.

Dr. Bulent Miran, Prof.

Ege University, Faculty of Agriculture,
Department of Agricultural Economics, Bornova-Izmir, Turkey

Dr. Ibrahim Yildirim, Prof.

University of Yuzuncu Yil, Faculty of Agriculture,
Department of Agricultural, Economics, Van, Turkey

Abstract:

The major objective of this study was to determine the priorities given by farmers to the previous selected goal statements and the factors which affect their decisions. Data of 2005 year was provided from 162 farmers settled in 10 villages in İzmir, Turkey. The goal priorities and the effective factors were determined using Fuzzy Pairwise Comparison and Logistic Regression Analyses, respectively.

Fuzzy Pairwise Comparison Analysis showed that risk minimization had the highest ranking in farmers' preferences with weight of 0.53 followed by profit maximization with 0.48 weights among four selected goal statements. Logistics Regression Analysis indicated that farmers who tended to take higher risks were those who involved in crop production, and adopted contractual production and input storage strategies while risk averter farmers consisted of farmers who were susceptible to natural conditions and preferred crop diversification.

Key Words: Farmers' Goals Priorities, Fuzzy Pair-wise Comparison, Logistic Regression

Introduction

Profit maximization, sustainable profit, cost minimization and risk minimization are among the major aims of farmers during their production and marketing activities.

Determining which aim has the priority is important in terms of efficient resource use and strategies to be selected (Van Kooten et al., 1986; Başarır and Gillespie, 2003; Gunden and Miran, 2007). The risk degree levels that the farmers are ready to take have a key role to reach a desired profit level. Despite of difficulty of obtaining clear responses, various methods have been applied in order to determine the farmers' risk degree level (Karahan, 2002; Başarır, 2002; Gunden, 2005; Miran, 2005; Şahin, 2008).

Determination of farmers' goal hierarchies is useful in terms of production planning and sustainable profitability. Studies regarding the goal preferences and ordering of farmers are abundant. Smith and Captick (1976) reported the farmers' goals of 'staying in business' and 'increasing farm size' as the highest and lowest ranks, respectively among 10 goals. 'to be my own bosses', 'selling through the free market' and 'can express myself' were reported as the most important ones among 11 selected goal statements (Kliebenstein et al., 1980). Van Kooten et al., (1986) reported the 'profit maximization' and 'reducing the farm dept' as the highest ranks among eight selected goal statements using fuzzy pair-wise comparison method. Walker and Schubert (1989) classified the farm families in terms of family values, roles, characteristics and decision making process into two groups namely environmentally effective and efficient entrepreneurs and expressed the first and the second groups in the mood of 'keeping the family farm' and 'business-oriented and profit maximization', respectively.

The major objectives of farmers were reported as ‘maximize gross margin’, ‘minimize working capital’, ‘minimize hired labor’, ‘minimize management difficulty’ and ‘minimize risk’ using a multi-criteria methodology (Sumpsi et.al. 1996). Aromolaran and Olayemi (2000), who used pair-wise comparison method, reported the major objectives of farmers as ‘farm household food security’, ‘limited use of external inputs’ and ‘maximum gross farm income’. The goal hierarchies of beef and dairy farmers were compared by Başarır and Gillespie (2003) using pair-wise comparison. The most important goals were cited as ‘maintain and conserve land’ and ‘avoid years of loss / low profit’ among seven selected goal statements. Among the seven selected goal statements for Turkish farmers using fuzzy pair-wise comparison method the most important goals were cited as ‘preserving their lands’ and ‘paying their depth’ (Günden and Miran, 2007).

The major aim of this study was to determine the farmers’ goal hierarchies and preferences among previously determined goal statements and the factors affecting farmers’ decisions.

Methodology and data

Fuzzy Pair-Wise Comparison

Fuzzy theory began with a paper on “fuzzy sets” by Zadeh in 1965. Fuzzy set theory is an extension of crisp set theory (Tanaka, 1997). Fuzzy sets are sets with boundaries that are not precise. Thus, fuzzy sets describe ranges of vague and soft boundaries by degree of membership (Lai and Hwang, 1994). The membership in a fuzzy set is a matter of a degree (Klir and Yuan, 1995). Fuzzy set is characterized by a membership function, which is allowed to choose an arbitrary real value between zero and one.

Fuzzy pair-wise comparisons were first used by Van Kooten et.al., (1986) to study farmers’ goal hierarchies for use in multiple-objective decision making. The first step of fuzzy pair-wise comparison approach is data collection by using a unit line segment as illustrated in Figure 1. Two goals, profit (P) and cost (C) are located at opposite ends of the unit line. Farmers are asked to place a mark on the line to indicate the degree of their preferred goal. A measure of the degree of preference for goal P over C , r_{pc} , is obtained by measuring the distance from the farmer’s mark to the P endpoint. The total distance from P to C equals 1. If $r_{pc} < 0.5$, goal C is preferred to P ; if $r_{pc} = 0.5$, the farmer is indifferent between P and C ; and if $r_{pc} > 0.5$, then goal P is preferred to C . $r_{pc} = 1$ or $r_{pc} = 0$ indicates absolute preference for goal P or C . For example, if $r_{pc} = 1$, then goal P is absolutely preferred to C (Van Kooten et al, 1986).

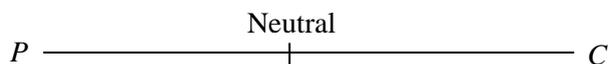


Figure 1. Fuzzy method for making pair-wise comparison between goals P and C

This study employed four farmers’ goal statements. The number of pair-wise comparisons, λ , was calculated as follows:

$$\lambda = n * (n - 1) / 2 \quad (1)$$

Where n = the number of goals. Thus, a farmer made six pair-wise comparisons in a personal interview.

In the second step of fuzzy pair-wise comparison method, for each paired comparison (i, j), r_{ij} ($i \neq j$) was obtained. r_{ij} ’s values was collected directly from the farmer. Also r_{ij} ($i \neq j$) is a measure of the degree by which the farmer prefers goal i to goal j and $r_{ji} = 1 - r_{ij}$ represents the degree by which j is preferred to i . farmers’ fuzzy preference matrix R with elements was constructed as follows (Van Kooten et al., 1986).

$$R_{ij} = \begin{cases} 0 & \text{if } i = j \forall i, j = 1, \dots, n \\ r_{ij} & \text{if } i \neq j \forall i, j = 1, \dots, n \end{cases} \quad (2)$$

Finally, a measure of preference, μ , was calculated for each goal by using farmers’ preference matrix R . The intensity of each goal was measured separately by the following equation:

$$\mu_j = 1 - \left(\sum_{i=1}^n R_{ij}^2 / (n - 1) \right)^{1/2} \quad (3)$$

μ_j has a range in the closed interval [0,1]. The larger value of μ_j indicates a greater intensity of preference for goal j. As a result, farmers' goals were ranked from most to least preferable by evaluating the μ values.

To analyze farmers' preferences derived from fuzzy pair-wise comparison, nonparametric statistical tests were used (Başarır and Gillespie, 2003). Friedman test was used to establish whether the goals were equally important within a block, which farmers' goal is ranking according to his/her preferences. Since four goals were presented to farmers, each row included four values which are the degree of goals exposed from a farmer. The null hypothesis was that there would be no difference in preferences over the goals among farmers. Alternatively, at least one goal would be preferred over the others.

Logistic Regression

Logistic regression extends the linear model to problems in which the response is either a category or a binomially distributed count (Agresti, 1996). Logistic regression allows one to predict a discrete outcome, such as group membership, from a set of variables that may be continuous, discrete, dichotomous, or a mix of any of these.

Logistic regression predicts the probability that the dependent variable event will occur given a subject's scores on the independent variables. The predicted values of the dependent variable can range from 0 to 1. If the probability for an individual case is equal to or above some threshold, typically 0.50, then our prediction is that the event will occur. Similarly, if the probability for an individual case is less than 0.50, then our prediction is that the event will not occur.

Logistic regression is more flexible than the other techniques. It has no assumptions about the distributions of the predictor variables; the predictors do not have to be normally distributed, linearly related, or of equal variance within each group in logistic regression. Unlike multiple-regression analysis, which also has distributional requirements for predictors, logistic regression cannot produce negative predicted probabilities (Hosmer and Lemeshow, 2000).

Logistic regression analysis is especially useful when the distribution of responses on the dependent variable is expected to be nonlinear with one or more of the independent variables. In logistic regression, there is a (binary or dichotomous) response of interest, and the predictor variables are used to model the probability of that response. Since the model produced by logistic regression is nonlinear, the equations used to describe the outcomes are slightly more complex than those for multiple regressions (Tabachnick and Fidell, 2001).

In this study, binary response variable was of interest. Independent variables denoted by the vector $x = \{x_1, x_2, \dots, x_p\}$. The conditional probability that the outcome is present was calculated with the following formula.

$$P(Y = 1 / x) = \pi(x) \quad (4)$$

The logit of the multiple regression model was given by

$$g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (5)$$

in which case the logistic regression model was

$$\pi(x) = [1 + \exp(-\beta_0 - \beta_1 x)]^{-1} \quad (6)$$

where the outcome variable, $\pi(x)$, is the probability of having one outcome or another based on a nonlinear function of the best linear combination of predictors with two outcomes (Gujarati, 1989).

This linear regression equation created the logit or log of the odds:

$$\log\left[\frac{\pi}{1-\pi}\right] = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} \quad (7)$$

That is, the linear regression equation is the natural \log (\log_e) of the probability of being in one group divided by the probability of being in the other group. The procedure of estimation that leads to the least squares function under linear regression model, when the error terms are normally distributed, is maximum likelihood, and the goal is to find the best linear combination of predictors to maximize the likelihood of obtaining the observed outcome frequencies. Maximum likelihood estimation is an iterative procedure that starts with arbitrary values of coefficients and determines the direction and size of change in the coefficients that will maximize the likelihood of obtaining the observed frequencies.

Four goal statements were selected and presented to farmers to respond.

a) Profit Maximization: The producers is supposed to be optimistic regarding production and marketing conditions. Profit maximization requires the producer take high risks in his activities.

b) Reasonable Profit: The producer is supposed to be pessimistic and he is satisfied with the highest income from the worst alternative.

c) Cost Minimization: A given production quantity is aimed with minimum cost. The producer minimizes his probable regrets.

d) Risk Minimization: This goal statement assumes that the producer is risk averter. He wants to obtain a guaranteed profit with minimum risks.

Data

The data of 2005 year was obtained from 162 farmers in 10 selected villages of İzmir, Turkey. (The third biggest city in terms of population suited near Aegean Sea) by personal interviewing through means of a structured questionnaire. The sample size was determined by simple random sampling method (Newbold, 1995).

The variables used in the analyses and their descriptions are presented at Table 1. Base statistics for general characteristics of farmers and farms is given at Table 2.

Results and discussions

Goal hierarchies of Farmers

Fuzzy pair-wise comparison model showed that the risk minimization was farmers' most important goal with the weight of 0.53 followed by profit maximization with 0.48 weights (Table 3). Friedman test was significant, which indicated the producers' priority of some goal statements over the others. Kendall's W test showed that farmers' agreement in goals ranking was weak. The priority goal statements found in this study was in line with some previous reported results, which indicated the highest farmers' goal statements as risk minimization and gross margin maximization (Sumpsi et al, 1996; Başarır and Gillepsie, 2003; Aromolaran and Olayemi, 2000; Walker and Schubert, 1989). However, the highest ranking goal statements of this study differed from the farmers' goal statements reported as 'to be my own bosses, 'selling through the free market' (Kliebenstein et al. 1980) and 'preserving our lands' and 'paying our depth' (Günden and Miran, 2007).

The groups of age, education level, risk attitude of farmers and land size differed significantly in terms of reasonable profit, cost minimization, risk minimization and reasonable profit/risk minimization goal statements of farmers, respectively (Table 4).

Factors Effective on Goal Statements

The logistic regression analysis results indicated that crop production had negative effect on risk minimization; natural conditions had positive effects on profit maximization and reasonable profits and negative effect on cost minimization and risk minimization; contractual production had positive effect on cost minimization and risk minimization and negative effect on profit maximization; production planning had positive effect on cost minimization; crop diversification had negative effect on cost minimization and risk minimization; input storage had positive effect on profit maximization and negative effect on cost minimization and risk minimization; product processing had positive effect on risk minimization; suggestions regarding agricultural activities had positive effect on cost minimization and finally chance had positive effect on profit maximization and negative effect on cost minimization (Table 5).

The farmers who involved in crop production; and who adopted contractual production and input storage strategies tended to take a higher risk to reach profit maximization. The risk averter farmers were those who considered the natural conditions seriously; and who adopted the crop diversification and thus satisfied with a comparable low profit.

Cost minimization goal was preferred by farmers who believed in product planning; who took into consideration the suggestions made regarding the agricultural activities; and who considered the chance as a big player in agricultural production.

Conclusion

The highest and lowest priority of farmers was risk minimization and cost minimization, respectively, which indicates that the farmers are ready to afford high costs to reach profit maximization goal as long as they are protected against risks stemmed from production and marketing activities.

To alleviate the production risks, contractual production is seen as a useful tool, which infers that policies to expand the contractual farming in various fields are expected. In addition, further product processing is another measure, which the farmers in the region are ready to adopt. Agribusiness is already at high concentrated level in the region compared to most of other regions in Turkey. However, it can still be improved in terms of relatively disadvantages sub-sectors.

Taking into consideration the susceptibility of farmers against risks, the policy makers is required to expand the scope of agricultural insurance, which a significant step was initiated under Agricultural Insurance Pool application in 2005 through Agricultural Insurance Code No. 5363. Among the cited aims of Insurance Pool are extending the insurance coverage; incentive for participation to reinsurances; coordination of activities of insurance companies and government subsidies; encouraging the participation in insurance and prevent unfair competition in the prices.

References:

- Agresti, Alan. (1996), An introduction to categorical data analysis. John Wiley and Sons, Inc. 287 p.
- Aromolaran, A.B. and Olayemi, J.K. (2000), Analysis of factor affecting the preference intensity of farmers for selected farm production objectives. *African Development Review*, 12(1): 114-127
- Başarır, A. and Gillespie, J.M. (2003), Goals of beef cattle and dairy producers: a comparison of the fuzzy pair-wise method and simple ranking procedure. Selected Paper for Southern Agricultural Economics Association Annual Meeting, February 1-5, Mobile, AL.
- Başarır, A. (2002), Multidimensional goals of farmers in the beef cattle and dairy industries. PhD Dissertation, Louisiana State University, Department of Agricultural Economics and Agribusiness, Louisiana.
- Gujarati, D. (1989), Essentials of econometrics, McGraw-Hill, New York. 553 p.
- Günden, C. (2005): A study on developing flexible and applicable individual production plans for region and farm size level by farmers via fuzzy goal programming, The Case of Torbali, Izmir, PhD Dissertation, Ege University Faculty of Agriculture, Department of Agricultural Economics, İzmir.
- Günden, C., Miran, B. (2007), A research on the determination of farmers' objectives hierarchy using fuzzy pairwise comparison. *The Journal of Agricultural Faculty of Akdeniz University* 20, 183-191.
- Hosmer, D.W., Lemeshow, S. (2000), Applied logistic regression, John Wiley&Sons New York, Inc. 375 p.
- Karahan, Ö. (2002), A research on the attitudes of farmers towards risk: case of Aegean region farmers, (PhD Dissertation), Ege University Faculty of Agriculture, Department of Agricultural Economics, İzmir.
- Kliebenstein, J. B., Barrett, D. A., Hefferman, W. D. and Kirtley. C. L. (1980), An analysis of farmers' perceptions of benefits received from farming. *North Central Journal of Agricultural Economics*, 2(2): 131-36
- Klir, G.J. and Yuan, B. (1995), Fuzzy sets and fuzzy logic: theory and application. Prentice Hall, New Jersey, 574 p.
- Lai, Y.J. and C.L. Hwang. (1994), Fuzzy mathematical programming: methods and applications. Springer-Verlag, Berlin, 301 p.
- Miran, B. (2005), Applied business planning, Ege University Faculty of Agriculture, Department of Agricultural Economics,, İzmir, 341 p.
- Newbold, P. (1995), Statistics for business and economics. Prentice-Hall International, New Jersey, 867 p.
- Smith, D. and Capstick, D.F. (1976), Establishing priorities among multiple management goals. *Southern Journal of Agricultural Economics*, 8(2): 37-43.
- Sumpsi, J.M., Amador, F. and Romero, C. (1996), On farmers' Objectives: A multi- criteria approach. *European Journal of Operational Research*, 96: 64-71.
- Şahin, A. (2008), Farm planning in risk conditions: game theory approach (Ph.D.Thesis) Ege University, Institute of Natural and Applied Sciences, Department of Agricultural Economics, İzmir.
- Tabachnick, B.G., Fidell, L.S. (2001), Using multivariate statistics, John Wiley&Sons New York, Inc. 980 p.
- Tanaka, K. (1997), An introduction to fuzzy logic for practical applications. Springer - Verlag, New York, 138 p.

Van Kooten, G.C., Schoney, R.A. and Hayward, K.A. (1986), An alternative approach to the evaluation of goal hierarchies among farmers. *Western Journal of Agricultural Economics*, 11(1): 40-49.

Walker, L. and Schubert, J. (1989), Fitting farm management strategies to farm style. *Canadian Journal of Agricultural Economics*, 37: 747-54.

Zadeh, L.A. (1965), Fuzzy sets. *Information and Control*, 8: 338-358.

Table 1: Variables used in the analyses and their descriptions

Variables	Unit	Description
Age	year	Age of farmer in years
Education	year	Farmer's years of schooling
Farm size	categorical	Total cultivated land 1: <30, 2: 31-100, 3: 101+ decar
Crop production	dummy	1: Yes 0: No
Risk attitude	dummy	1: Risk averter 0: Risk lover
Agree with agricultural production affected by natural conditions	scale	Five point scale; 1: Strongly disagree, ..., 5: Strongly agree
Agree with contractual production	scale	Five point scale; 1: Strongly disagree,, 5: Strongly agree
Agree with agricultural planning	scale	Five point scale; 1: Strongly disagree,, 5: Strongly agree
Agree with crop diversification	scale	Five point scale; 1: Strongly disagree, ..., 5: Strongly agree
Agree with storage inputs before production period	scale	Five point scale; 1: Strongly disagree,, 5: Strongly agree
Agree with processing agricultural production	scale	Five point scale; 1: Strongly disagree, ..., 5: Strongly agree
Apply suggestion for agr. activities	scale	Five point scale; 1: Strongly disagree, ..., 5: Strongly agree
Agree with agricultural production affected by chance	scale	Five point scale; 1: Strongly disagree, ..., 5: Strongly agree

Table 2. Base Statistics for General Characteristics of Farmers and Farms

Variables	Mean	Standard deviation	Min	Max	No. Respondents*	%
Age (years)	44.92	11.87	18.00	76.00		
0-35					44	27.16
36-55					89	54.94
56+					29	17.90
Education (years)	6.09	2.37	3.00	15.00		
Farm size (decar)	81.21	76.88	6.00	400.00		
0-30					56	34.57
31-100					65	40.12
101+					41	25.31
Crop production (dummy)						
1: Yes					94	58.02
0: No					68	41.98
Risk attitude (dummy)						
1: Risk averter					105	64.81
0: Risk lover					57	35.19
Agree with agricultural production affected by natural conditions (scale)	4.13	1.41	1.00	5		
Agree with contractual production (scale)	1.65	1.24	1.00	5		

Agree with agricultural planning (scale)	1.99	1.23	1.00	5		
Agree with crop diversification (scale)	3.20	1.53	1.00	5		
Agree with storage inputs before production period (scale)	2.11	1.50	1.00	5		
Agree with processing agricultural production (scale)	1.25	0.62	1.00	5		
Apply suggestion for agr. activities (scale)	2.31	1.27	1.00	5	237	62.37
Agree with agricultural production affected by chance (scale)	1.69	1.29	1.00	5	143	37.63

Table 3: Base statistics for farmers' goal statements

Farmers' goal statement	Mean	Standard deviation	Minimum	Maximum
Minimize risk	0.53	0.27	0.10	0.90
Maximize profit	0.48	0.29	0.10	1.00
Making suitable profit	0.42	0.17	0.07	0.90
Minimize cost	0.39	0.14	0.10	0.86

Friedman $p < 0.01$
Kendall's W = 0.02

Table 4: The relationship of some variables with farmers' goal statements

Variable	Number of Farmers	Profit Maximization	Reasonable Profit	Cost Minimization	Risk Minimization
Age					
0-35	44	0.49	0.46*	0.40	0.48
36-55	89	0.46	0.42*	0.38	0.55
56+	29	0.52	0.36*	0.40	0.53
Education					
0-5	125	0.49	0.41	0.38**	0.52
6-11	22	0.38	0.42	0.46**	0.58
12+	15	0.48	0.48	0.36**	0.47
Land					
0-30	56	0.45	0.38**	0.40	0.57*
31-100	65	0.51	0.46**	0.38	0.48*
101+	41	0.47	0.41**	0.39	0.54*
Risk attitude					
Risk averter	105	0.46	0.42	0.39	0.56*
Risk lover	57	0.52	0.42	0.38	0.47*

* $p < 0.10$ Kruskal Wallis ** $p < 0.05$ Kruskal Wallis

Table 5: Logistic regression analysis results for farmers' goal statements.

Variables	Profit Maximization		Reasonable Profit		Cost Minimization		Risk Minimization	
Constant	-6.41871 (0.257548)	***	-5.84728 (0.156957)	***	-5.12439 (0.12924)	***	-4.78136 (0.216977)	***
Crop Production	0.15988 (0.124622)		0.104228 (0.0759481)		-0.0470497 (0.0625361)		-0.181877 (0.10499)	*
Natural Conditions	0.112058 (0.0410837)	***	0.045668 (0.0250376)	*	-0.0664898 (0.0206161)	***	-0.0982707 (0.0346118)	***
Contractual Production	-0.119112 (0.0568892)	**	-0.0202147 (0.03467)		0.0740089 (0.0285474)	**	0.104584 (0.0479275)	**
Production Planning	-0.025062 (0.0616745)		0.0014818 (0.0375863)		0.0650997 (0.0309488)	**	0.000465879 (0.051959)	
Crop Diversification	0.0431444 (0.0393456)		0.0339616 (0.0239784)		-0.0331744 (0.0197439)	*	-0.0727777 (0.0331476)	**
Input Storage	0.132403 (0.0751163)	*	0.0417103 (0.0457781)		-0.111119 (0.0376939)	***	-0.152523 (0.0632833)	**
Product Processing	-0.132448 (0.108173)		-0.0908649 (0.065924)		0.0895682 (0.0542822)		0.214727 (0.0911329)	**
Suggestion for Agricultural Activities	0.0942634 (0.0584918)		0.0228888 (0.0356466)		-0.0646614 (0.0293516)	**	-0.0052092 (0.0492777)	
Chance	0.0877417 (0.045103)	*	-0.022686 (0.0274871)		-0.080873 (0.0226331)	***	-0.0551131 (0.037998)	
χ^2	13.0588	***	4.45749	***	2.6085	***	11.1684	***
R-sq	0.0584614		0.0240567		0.201726		0.0280553	

*** Significant at 1% level, ** Significant at 5% level, * Significant at 10% level

BREAK EVEN ANALYSIS OF BROILER PRODUCTION IN THE ACCRA-TEMA AND KUMASI AREAS

Mahama, E.A.

Agricultural Research Centre, University of Ghana, Legon

Andah, E.K.

Amegashie, D.P.K.

Mensah-Bonsu, A.

Dept of Agricultural Economics & Agribusiness, University of Ghana, Legon

Abstract:

The paper examines the profitability of broiler production, using the break even concept. The break even quantities, net present values and internal rates of returns have been estimated for 20 selected poultry farms. Poultry farms located in Kumasi area broke even in the production of broiler while about 38.5 percent of the farms in Accra-Tema area produced below their break even levels. Only 30 percent of the farms registered positive net present value. Under an interest rate scenario of 25 percent about 70 percent of all farms, would be profitable (break even). The results suggest that lower interest rate is essential for more farms to break even (be competitive) and stay in the poultry business.

Key Words: Break even, interest rate, installed capacity, broiler production, Accra – Tema and Kumasi areas

Introduction

Poultry production, especially broiler production is a big business throughout the world today. It has become a standard form and source of cheap protein. It is also a competitive industry in the developed world and fast becoming so in the less developed countries. With continued population growth, urbanization, income growth and changes in lifestyles and food preferences, it was projected that meat demand in the developing world will double between 1995 and 2020 to 190 million tons and increase by 25 percent in developed countries to 122 million tons (Delgado et al, 1999). Demand for meat will grow faster than cereals in the developing world by 2.8 percent per year for meat compared with 1.8 percent for cereals (Pinstrup-Andersen et al, 1997). It has estimated that the demand for poultry meat in the developing world is expected to increase fastest, at an average of 3.6 percent, compared with 2.8 percent for beef and 2.3 percent for pork (pig meat). The demand for livestock products in Ghana is estimated to be increasing at a rate of about 5 to 7 percent per annum (Government of Ghana, 1995).

The early 1970s witnessed the emergency of poultry farming as a major source of meat and profitable enterprise in Ghana (Tachie-Mends, 1992). The poultry industry started to decline in 1981 when prices of inputs began to shoot up and many farmers could not afford the cost of production. By 1987, however, production of poultry had picked up as a result of increased demand due to increase in population (Tachie-Mends, 1992). Most of the current meat import is poultry parts, with a substantial content of turkey tails and chicken wings (Government of Ghana, 1992). Both frozen poultry and beef are imported from European Union countries. These imports are heavily subsidized, in the European Union countries, therefore, depressing domestic prices in importing countries such as Ghana.

Whereas producers are complaining of competition from subsidized imported frozen chicken, high cost of inputs and lack of credit facilities as hindering the local poultry industry, consumers are equally dissatisfied with the expensive poultry meat produced locally. These issues raise the following research questions: Is broiler production in Ghana really profitable? What is the break-even level for broiler production for the firms in the industry? What is the rate of return on the firms' investment?

The primary objective of the study therefore is to assess the profitability of broiler production in Accra-Tema and Kumasi areas. The specific objectives are to determine the break even levels for broiler production and estimate the rate of return on the investment for the poultry firms in the two study areas.

The economy of Ghana depends on agriculture and to ensure growth in the general economy, the agricultural sector needs to be improved. The poultry industry has an important role in this respect. If the broiler meat sub-sector is well developed, it could help the economy conserve foreign exchange used to import poultry products and provide employment for the youth of Ghana. The estimation of the break-even point is useful because it shows at what quantity to produce without losing. If the break-even quantity of birds is close to the installed capacity, then the problems of being able to produce at the most profitable level would be crucial. The two areas have been chosen for the study because a large proportion (about 89 percent) of poultry farms in Ghana are located in these areas (MOFA, 1996), and they provide ready market for poultry products.

Previous related studies have given mix conclusion about the performance of the poultry industry. Using the enterprise costing method, Boa-Amponsem (1988) recommended that for efficient broiler production, overhead cost such as management salaries, office expenses and electricity expenses should not exceed 15 percent of the variable cost. Applying the comparative cost analysis for two modern poultry production technologies and implications for policy under Nigeria's Structural Adjustment Programme, Akinyodoye, and Pingpoh (1992) concluded that poultry production in Nigeria is financially but not economically profitable given the prevailing high cost of imported feed. Maham, et al (1990) in an assessment of the cost of producing cockerel and estimating a break-even point in India, observed that the fluctuations in the price of chickens and the unpredictable supply of broiler chicks encouraged broiler farmers in Tamil Nach, India to take up cockerel production.

Methodology

Data: The period within which the study was conducted, covered June 1996 to May 1997. In all 20 farms were selected including 13 farms from Accra/Tema area and 7 farms from the Kumasi area. Each farmer provided information on three production cycles. A total of 60 production cycles (3 cycles of 20 farms) were used for the analysis. The data collected included information reported in the income and expenditure statements or the profit and loss accounts, production accounts; and general information on the assets of the farm. Some secondary data collected included the banks lending rates.

Method of Analysis: Financial analysis requires identification of all revenues and costs streams over the life of the project. For the present study, the costs include the fixed and variable costs. The fixed costs constitute the capital outlay of the firms. The fixed costs include land, structures and buildings, machinery, insurance, payment on interest and rent. For the fact that fixed inputs decline in value or productivity as the annual use of the machinery, equipment, and buildings increase (or are not used up completely in one season), their annual depreciated values were estimated and used in undertaking the undiscounted break even analysis. The depreciated values of the fixed inputs were summed to give the total fixed cost of a production cycle.

The variable cost can be divided into production cost, overhead costs and in some cases research and development cost. For the purpose of this study, the production cost for broiler production included the costs of feed, veterinary drugs, wood shavings, day old chicks, fuel for vehicles, electricity and water expenses as well as other energy such as charcoal and gas, and salaries of personnel directly involved in the production process. Overhead costs included the costs of indirect materials, indirect labour (management and office staff), office expenses and maintenance of vehicles. These costs were summed up to give total variable cost (TVC). The total cost (TC) of a production cycle therefore is the sum of the total fixed cost (TFC) and total variable cost (TVC).

Total revenue is the cash value realized from the sales of broilers during the accounting period (a production cycle). To arrive at the sales revenue, the total quantity of broilers sold (Q) was multiplied by the unit price (P) of the product. Thus, $TR = Q \times P$, where TR is total revenue.

Break-Even Analysis (Undiscounted): Break-Even (B/E) analysis normally assists in establishing the relationship between fixed costs, variable costs, sales revenue and profit. The point at which sales revenue and total cost are equal is termed the break-even point. The break-even quantity is derived from the costs and revenue of the farms. It is assumed that poultry farms in Kumasi and Accra areas operate in a competitive local market and therefore they are price takers. At the Break-Even point the total cost of production is equal to the total revenue.

$$\text{Thus, TR} = \text{TC} \quad (1)$$

$$\text{At the equilibrium level, TR} = \text{QE} \times \text{PE} \quad (2)$$

$$\text{Therefore QE} \times \text{PE} = \text{TVC} + \text{TFC} \quad (3)$$

$$\text{Then, QE} = (\text{TVC} + \text{TFC})/\text{PE} \quad (4)$$

where, QE is Break-Even output (quantity) and PE is the average market selling price for the area. The break-even analysis was carried out for the individual farms, Accra/Tema farms, Kumasi farms and then for all the farms in the two areas combined. The simple average of the three production cycles for each farm was used for the analysis.

Estimation of the Net Present Value and Internal Rate of Return: The net present value (NPV) is given by the present (discounted) value of benefit minus present (discounted) value of cost (Gittinger, 1982). In calculating the financial NPV, the prevailing bank interest rate of 45 percent per annum during the period of the data collection (1996/97) was used as the discount rate. Each of the farms was expected to generate a positive NPV, denoting the profitability of the broiler operation. The internal rate of return (IRR) method was applied to determine the financial interest rate which will equate the present value of the benefit to the present value of the cost and therefore drives the NPV to zero (a break-even point). The IRR was thus used to estimate the interest rate that would make the firms break even under the given production environment. The annual net income from broilers was calculated for each farm and projected over an assumed project life period of 25 years. The net income was given by the total revenue less the total variable costs.

Also, the economic net present value and rate of return were estimated after adjusting the values used for the financial analysis. Direct transfer payments such as taxes paid by the poultry farmers, direct subsidies and credit transactions including loans receipts and repayments, insurance, interest on loans were omitted when the financial values were converted to economic values. The exchange rate effect for imported inputs were not adjusted because the exchange rate in the economy is determined on a daily basis by the commercial and development banks in a unified manner based on the supply and demand situation. Similarly, the interest rate used to discount the economic values was not adjusted. In relation to labour in the poultry industry, the wages of labour used for the financial analysis were also used for the economic analysis, because of the unemployable nature of unskilled labour in the economy.

Results And Discussion

The results of the break even analysis are presented in Tables 1, 2 and 3. The estimated mean unit selling price for broiler in the Kumasi area was ₵6,900 (Table 3). At the estimated mean selling price level, the estimated break even quantity and percentage change in production from the break even level for the different farms in the Kumasi area are presented in Table 1. All farms in the Kumasi area produced at levels above their break even levels, with output ranging between 2.6 percent and 35.5 percent above the break even point. The implication is that the farmers in the Kumasi area broke-even in the production of broilers. Three of the farms, representing 42.9 percent, produced above 60 percent of their installed capacity, while another three (42.9 percent) farms produced below 20 percent of their installed capacity. On the average a farm in Kumasi area produced 20.5 percent above the break even level, while utilizing about 50.2 percent of the installed capacity (Table 3).

Table 1: Break-Even Analysis for Broiler Meat Production in Kumasi area

FARM	1	2	3*	4*	5	6	7
Sales (₵'000')	4140	2967	34255	29015	3356	3425	5698
Total variable cost(₵000)	3614	2608	29524	18820	2772	2958	4429
Total Fixed cost (₵000)	54.7	30	9336	3324	38.2	26	40
Installed Capacity(birds)	4000	2000	10000	6000	2500	500	2000
Quantity produced (birds)	625	392	6981	3768	435	475	878
Unit price(₵'000')	6.5	7.2	4.9	7.7	8.0	7.2	6.5
Capacity utilized (%)	15.6	19.6	69.8	62.8	17.4	95.0	43.9
Break Even Quantity (birds)	532	382	5632	3209	407	432	648
Percentage increase	17.48	2.62	23.95	17.42	6.88	9.95	35.49

Source: Calculated from field data, 1996/97

Note: *Farms 3 and 4 sell dressed birds in kilo weight only.

For the Accra-Tema area, the mean unit price for broiler was estimated as $\text{¢}7,300$ (Table 3). The estimated break even quantity and percentage change in production from the break even level for the different farms in the Accra-Tema area are presented in Table 2. Five of the farms, representing 38.5 percent, produced below their break even levels. Five farms, representing 38.5 percent, produced above 60 percent of their installed capacity, seven (53.8 percent) farms produced between 30 and 60 percent of their installed capacity while one farm (7.7 percent) produced below 30 percent of its installed capacity. On the average a farm in Accra-Tema area produced 18.7 percent above the break even level with the utilization of about 59.1 percent of the installed capacity (Table 3). For the two study areas combined, about 55.03 percent of the installed capacity was used to produce 20.5 percent above the firms' break even level (Table 3).

Table 2: Break-Even Analysis for Broiler Meat Production in Accra-Tema Area

FARM	1	2	3	4	5*	6	7	8	9	10	11	12	13
Sales(¢0000)	23756	2529	6408	4083	27044	3667	15656	12296	2970	28033	6383	3040	2721
TVC (¢000)	15563	2078	6476	3555	23400	2863	9136	10959	2878	20198	4677	2774	2352
TFC (¢000)	239.3	39	982	83	3894	189	140	250	38	2356	100	35	55
Installed capacity	4000	700	3000	1000	5000	1500	4000	3500	500	5000	2000	1000	500
Qty produced	2852	281	961	578	3698	380	2037	2049	398	3781	890	467	380
unit price (¢000)	8.3	9	6.7	7.3	7.4	7.7	9.8	6	6.3	7.7	7.2	6.5	5.1
Capacity utilized (%)	71.3	40.1	32.0	57.8	74.0	25.3	50.9	58.5	79.6	75.6	44.5	46.7	76.0
B/E Quantity	2165	290	1019	498	3739	418	1271	1535	399	3090	654	385	330
Percentage increase (%)	31.73	-3.10	-5.69	16.06	-0.10	-9.09	60.27	33.49	-0.25	22.36	36.09	21.30	15.15

Source: Calculated from field data, 1996/97.

* Farm 5 sells dressed broiler in kilo weight only.

Table 3: Break-Even Analysis for Broiler Production, Mean Values for Study Areas

VARIABLE	ACCRA-TEMA	KUMASI	All Farms
Sales (€'000')	10,660	11,837	11,072
TVC (€'000')	8,224	9,246	8,582
TFC (€'000')	646	1,836	1,067
Installed capacity(birds)	2438	3,857	2935
Qty Produced(birds)	1442	1936	1615
Unit price (€'000)	7.3	6.9	7.1
Capacity utilized (%)	59.1	50.2	55.0
B/E Quantity(birds)	1215	1606	1359
Percentage increase on: B/E Quantity (%)	18.68	20.55	18.84

Source: Derived from Tables 1 and 2.

The capital outlay per farm and the calculated expected annual net income, the net present value and the internal rate of return for each of the farms in Accra-Tema and Kumasi areas are presented in Table 4.

Table 4: Capital Outlay, Net Income, Net Present Value (NPV) and Internal Rate of Return (IRR) for Broiler Meat Production in the Study Areas

F A R M	Capital Outlay* (€'000)	Net Income (€'000) /Year	NPV at 45% (€'000)	IRR (%)
ACCRA-TEMA				
1	10,000	24110	43,572.42	241.14
2	3,800	1256	(1,009.17)	33.20
3	9,800	(5888)	(22,883.14)	-
4	7,000	2341	(1,798.30)	33.58
5	191,600	13439	(161,738.54)	4.92
6	12,000	3655	(3,878.59)	30.48
7	40,000	10501	(16,666.78)	26.34
8	25,000	8240	(6,690.72)	33.11
9	3,500	(1508)	(6,850.78)	-
10	11,500	12650	16,608.30	110.00
11	6,000	6263	7,916.39	104.40
12	7,000	1952	(2,662.66)	28.04
13	3,900	1316	(975.85)	33.86
KUMASI				
1	10,800	1540	(7,378.12)	13.95
2	6,300	1020	(4,033.56)	15.95
3	42,800	(5030)	(53,976.66)	-
4	15,000	31360	54,681.92	209.05
5	6,800	4492	3,181.22	66.12
6	3,000	2800	3,221.60	93.39
7	11,000	3780	(2,600.84)	34.42

Source: Calculated from Field Data, 1996/97.

* The capital outlay of each farm was revealed to reflect 1996/97 prices.

NPV = (Annuity Factor x Net Income per year) – Capital outlay

where, Annuity Factor = $[(1 + r)^t - 1] / r(1 + r)^t$; r = interest (discount) rate and t = 25 years

Net income per year = total revenue per year less total variable costs per year

IRR = $r_L + [(r_H - r_L) * NPV^+ / (NPV^+ + |NPV^-|)]$

where, r_L = the (low) discount rate that gives the positive net present value (NPV⁺)

r_H = the (high) discount rate that gives the negative net present value (NPV⁻)

To reduce estimation error and obtain a more accurate IRR the interval $r_H - r_L$ was not more than 5.

The estimated results show that at an interest rate of 45 percent that prevailed during the survey time in 1996/97 only six farms, representing 30 percent of the farms, registered positive net

present value. This means that most of the farms were likely lose their invested capital and would find it difficult to continue their operation. Consequently, farms that continue to record negative net present value would either go in for more bank loans or fold up. A rise in product price relative to input price would help struggling farms to recover production cost. The farms that recorded positive net present value included 42.9 percent and 23.1 percent of farms located in Kumasi and Accra/Tema, respectively, suggesting that a farm located in Kumasi is more likely to record positive profit than one located in Accra/Tema area.

The internal rate of return (IRR) analysis, similarly, shows that only the six farms (30 percent of the farms) registered IRR above the interest rate of 45 percent (Table 4). The summarized results for the estimated IRR presented in Table 5 shows that if the banks' interest rate was reduced from 45 percent to 33 percent, then 25 percent additional farms, made up of 30.8 percent and 14.3 percent of farms in Accra-Tema and Kumasi, respectively, would break even. The estimated IRR results suggest that should the interest rate drop further to about 25 percent, then about 76.9 percent of farms in Accra-Tema area would be breaking even. Under the 25 percent interest rate scenario about 70 percent of all farms would be profitable (break even). The results suggest that lower interest rate is essential for more farms, particularly those located in Accra-Tema area, to become competitive and stay in the poultry business.

Table 5: Possible Interest Rate Charges and Percentage of Farms Breaking Even

Interest Rate (%)	Percentage of farms breaking even at the specified rates		
	Accra/Tema	Kumasi	All Farms
45	23.1	42.9	30.0
33	53.8	57.1	55.0
25	76.9	57.1	70.0

Source: Deduced from Table 4

Economic Rate Of Return

The estimated economic net present value and the economic rate of return are presented in Table 6. Though the necessary economic adjustments were carried out, the economic net present values were estimated at the commercial and development banks' interest rate of 45 percent.

Table 6: Net Present Value at 45 percent and the Economic Rate of Return Accra-Tema and Kumasi Areas

Area	Capital Outlay	Income/Year	NPV at 45%	ERR (%)
Accra-Tema	331,700	99,200	(76,900)	29.9
Kumasi	95,700	86,200	65,860	90.1
All Farms	427,400	185,400	(15,441)	43.5

Source: Calculated from Field Data, 1996/97.

Table 6 shows that generally the farms in the Kumasi area recorded a positive economic net present value, while farms in Accra-Tema area recorded a negative economic net present value. The overall net present value for the economy was negative, suggesting that broiler production is not economically viable in the country. The estimated economic rates of return suggest that, economically it is worth investing in the broiler meat production in the Kumasi area but probably not in the Accra-Tema area. The estimated economic rate of return suggests that if broiler meat production is to be economically viable, then a reduced economic interest rate of about 30 percent must be available and accessible to the poultry farms.

Conclusions

The study attempted to examine profitability of broiler production, using the break even concept. Farms located in Kumasi area broke even in the production of broiler with about 42.9 percent producing above 60 percent of their installed capacity, while about 38.5 percent of the farms in Accra-Tema area produced below their break even levels. About 38.5 percent of the farms in Accra-Tema area produced above 60 percent of their installed capacity. Only 30 percent of the farms, including 42.9 percent and 23.1 percent of farms located in Kumasi and Accra-Tema, respectively, registered positive net present value. The implication is that a farm located in Kumasi is more likely to record positive profit than one located in Accra-Tema area. Under an interest rate scenario of 25 percent about 70 percent of all farms, including 76.9 percent in Accra-Tema area and 57.1 percent in Kumasi area, would be profitable (break even). The results suggest that lower interest rate is essential for more farms to break even (be competitive) and stay in the poultry business.

References:

- Akinyodopye, V.O. and Pingpoh, D.P. 1992. Comparative cost analysis of two modern poultry production technologies and implications for policy under Nigeria's Structural Adjustment Programme. *Quarterly Journal for International Agriculture*, Vol. 31 (3): 262 - 268.
- Boa – Amponsem, K.O. 1988. *Nutrition and Husbandry, Research and Development of the Poultry Industry in Ghana*, Accra, pp. 4, 8.
- Delgado, C., Rosengrat, M., Steinfeld, H., Ehui, S., and Courbois, C. 1999. *Livestock to 2020: The next food revolution, 2020 vision for food, agriculture, and the environment*, Discussion Paper 28, International Food Policy Research Institute, Washington, D.C.
- Gittenger, P. 1982. *Analysis of agricultural projects*. The Johns Hopkins University Press, Baltimore.
- Maham, B. Selvan, S. P., Balakrishnan, D. and Shanmugam, T. R. 1990. Economy of cockerel production in Namakkai, *Poultry Adviser*, 23 (9): 27 - 31.
- MOFA. 1996. *National livestock census*. Ministry of Food and Agriculture, Accra.
- Pinstrup – Anderson, P., Pandya – Lorch, R., and Rosegrant, M.W. 1997. *The world food situation: Recent developments, emerging issues and long – term prospects*. 2020 Vision Policy Report, International Food policy Research Institute, Washington, D.C.
- Tackie – Mends, C. K. B. 1992. *Steps in rearing chicken*. Asempa Publishers, Accra.
- Government of Ghana. 1995. *Annual report of the policy planning, monitoring and evaluation department*, Ministry of Food and Agriculture, Accra, pp. 1 – 7.
- Government of Ghana. 1992. *National livestock services project, staff appraisal report*.

THE ROLE OF SOCIO-CULTURAL IN RELATIONSHIP MARKETING: CASE FINDINGS ON FARMERS IN WEST JAVA

Muchlis Ahmady

Graduate School of Management and Business, Bogor Agricultural University, Indonesia

Ujang Sumarwan

Department of Family and Consumer Sciences and Graduate School of Management and Business,
Bogor Agricultural University, Indonesia

Budi Suharjo

Department of Mathematics and Graduate School of Management and Business,
Bogor Agricultural University, Indonesia

Agus Maulana

Graduate School of Management and Business, Bogor Agricultural University, Indonesia

Abstract:

This paper studies the revelation of human relationships that drive the marketing activities of agricultural products, the producers (farmers and ranchers), collectors, dealers and the modern market, which focuses on the producers of agricultural products (vegetables, fruits and chicken pieces) in West Java Jakarta and surrounding areas, research only limited investigating the relationship between sellers and their suppliers, not the relationship between the seller (the traditional markets, modern market) with the end consumer.

The research was conducted in West Java and Jakarta and its surrounding areas with the implementation in November 2010 to January 2011. Site selection is done on purpose (purposive) on the grounds in West Java and Jakarta and its surroundings have a sufficiency objects to be studied about the manufacturer, supplier and retailer for agricultural commodities.

In this study involving 205 respondents as a source of information that consists of four business groups of the modern market by 18 respondents (8.79%), vegetables by 84 respondents (40.97%), and fruits totalling 50 respondents (24, 39%), and the chicken pieces as much as 53 respondents (25.85%). In data collection conducted through face-to-face interviews, each respondent was instrumental in accordance with its position in the business chain, such as manufacturer, suppliers or retailers. Especially for the respondent retailers, they are managers or management of a modern market primarily responsible for the procurement of commodities.

Analytical methods to be used include Descriptive Analysis and Structural Equation Model (SEM). Descriptive analysis aims to obtain an overview of the characteristics of the respondent (respondent relations and behavioural profile of research). Processing is done using frequency analysis and cross tabulation analysis, while SEM is used to validate the built model.

The results revealed that long-chain trading system that occurs not only because of economic factors alone, but discovered the existence of socio-cultural involvement in decision-making system of trading activity. Trading system that occurred more informal, in the sense that there are no rules as a guide, the factor of trust and commitment between actors trading system is crucial in encouraging the emergence of shared values that motivated the early emergence of cooperation. Performance of the three aspects will be the basis for decision-making as well as a selection for the establishment of long-term cooperation of mutual benefit both financially and socially. Future research should be pursued to investigate the direct influence of relational benefits for the trust, commitment and shared values. Therefore, this research theme will be more developed and interesting to do further research if there are other important factors that could affect the partnership.

Key Words: Relationships marketing, commitment, trust, cooperation, shared values, Social Cultural.

Introduction

Indonesia is a hugely diverse nation made up of over 17,500 islands with 33 provinces, which are home to over 300 ethnic groups, each province has its own language, ethnic make-up, religions and history. Most people will define themselves locally before nationally. In addition there are many cultural influences stemming back from difference in heritage. Due to the diverse nature of Indonesian society there exists a strong pull towards the group, whether family, village or island. People will define themselves according to their ethnic group, family and place of birth. The family is still very traditional in structure and family members have clearly defined roles and a great sense of interdependence

Java is the most densely populated island and is divided into three provinces with 61% of the population of Indonesia, in which West Java is the province with the highest population of 41 million (18.1%) and East Java provinces with 37.5 million inhabitants (15.8%), while the province of Central Java, 32.4 million people (13.6%). As the province with the largest population in Indonesia, West Java, a region with diverse cultures and local languages or dialects, and is characterized in contrast to the identity of the two urban communities who mostly live in the Greater Jakarta and traditional peoples living in the countryside.

It reflects a fraction of the facts of cultural diversity in Indonesia to the characteristics and customs that differ from one another. How they interact and what kind of social relationships occur and then may continue the business relationship is an interesting thing to note. In Indonesian people, quite a lot of social value (social capital) as the culture of mutual aid, institutional outcomes, various forms of local knowledge (local wisdom) possessed all the ethnic and may be developed as part of the culture of the modern economy.

Social capital has been proven by history as an important mechanism in their efforts to achieve economic growth and equitable society. Therefore, the economic empowerment of the people and role of social capital becomes very important to note (Mawardi, 2007). Agribusiness market has vast dimensions, both as a site of interaction between supply and demand of agriculture products (goods or services), the transactions and agreements ranging from the value, quantity, product specifications, mode of delivery, acceptance and payment, as well as the transfer of ownership of goods. Agribusiness marketer is someone who is looking for products or services then offer or sell it to another party or to the final consumer, on this basis agribusiness marketers can serve as a buyer and / or seller.

Some researchers say the study of agricultural trade system from the standpoint of economic science only reveal the shape and structure of the marketing chain, prices and fluctuations, cost and margin trade system, the integration of markets, market efficiency, marketing effectiveness, transmission rates, and others. Meanwhile, human activity is involved in moving the little researched and written, that is the actor who drive the flow of goods from the downstream to upstream.

Farming communities as major stakeholders in the process of production in the agricultural sector has not been much enjoyed encouraging results, especially from the aspect of income and welfare. Although the government has given great attention to this sector, it is thought caused there is no price guarantee of agricultural products, while agribusiness development by a partnership between employers and farmers / groups farmers have not developed as expected.

The extent to which the role of social aspects in colouring behaviour and marketing relationships between actors in the agricultural sector, and what factors are the key to the occurrence of a partnership and cooperation as well as a guarantee of continuity is not widely expressed quantitatively. Socio-cultural differences can be accounted for in the five dimensions of culture (Hofstede, 2001). These dimensions consist of: (1) individualism / collectivism, (2) uncertainty avoidance, (3) power distance, (4) masculinity / femininity, and (5) long-versus short term orientation. In accordance with the environmental survey done then focus on the first dimension is the dimension of individualism / collectivism describes the relationship between individual and collectively that can be found within a given society. While this relationship is reflected in how people live together, it is also related to social norms and value system related programming will therefore affect the individual's mental and guide their behaviour. While individuals who are part of the culture with a more collectivist system and direct their behaviour toward value oriented collective goals. Therefore, an attempt to determine patterns of relatedness (causal models) among agribusiness relationship based marketing approach becomes very crucial. This is not only intended to reveal the values that become

the key marketing relationships between actors, but also the values that a guarantee sustainability and future development of quality relationships.

Relationship Marketing or Customer Relationship is an approach that involves all actors in a marketing system, as well as reviewing various values of the reference or the key factor that links between actors in the system. In implementation, relationship-based marketing effort tracing the involvement of all actors and activities related to the marketing mix, where the consumer or customer is placed at the centre of the overall marketing activity (Palmatier et al. 2006, Brady and Roehm, 2007, Gregoire et al. 2009, and Fournier, 2002).

In Indonesia, relationship marketing in agribusiness case has not been widely studied, especially if not be said to be very rare. Meanwhile, the entities which act a role in moving (marketing) of agricultural commodities including growers, collectors, wholesalers, dealers, wholesalers and retailers (Kuma'at, 1992). According to Irawan et al. (2001), lack of harmony of the relationship between agribusiness that occur frequently due to the absence of equality of bargaining power between actors that have an impact on the imbalances of profit margin obtained as well as the difficulty for farmers get the fair price. The problems that arise in the marketing system of agricultural commodities, among others: marketing activities or trading system that has not run efficiently, so the absence of a fair wage.

Therefore, efforts to determine the pattern of agribusiness linkages with relationship marketing approach and how the socio-cultural influences become very crucial, given the close relationship and the relationship between prices, production and trading system, each has a crucial role and influence each other. This is not only intended to reveal the values that become the key marketing relationships between actors, but also the values of sustainability guarantee and the future development of relationships quality. Therefore, in the second part we submit the data and methodology used, then we submit the results and discussion in section three and in the end we give concluding remarks.

Data And Methodology

The research was conducted in West Java and Jakarta and surrounding areas, site selection is done on purpose (purposive) on the grounds in West Java and Jakarta and its surrounding have been studied regarding the adequacy of the object that producers and traders (collectors, wholesalers, dealers, wholesalers and retailers) in agricultural commodities. The study involved 205 respondents as a source of information. Respondents are decision makers or the company management (producers of agricultural commodities or livestock) and traders of fruits, vegetables, and chicken broilers.

Conceptual framework used in this study based on the model Morgan and Hunt (1994) theory of trust and commitment as home base, then from the concept model was developed by some of the results of previous studies such as (1) Trust and Commitment in Collaborative Business Relationships involving the Company in New Zealand (Palakshappa and Gordon, 2005), (2) Understanding the role of relational capabilities in relation to the buyer and seller (Smirnova and Kushch, 2006), (3) belief in the relationship between private buyers and sellers in the B2B (Akroun and Akroun, 2007), (4) the conceptual model for the retailer and vendor relationships based on trust and commitment theory (Smith, 2008), and (5) a direct relationship of trust and commitment among the sellers, distributors, and customers (Wang, 2009).

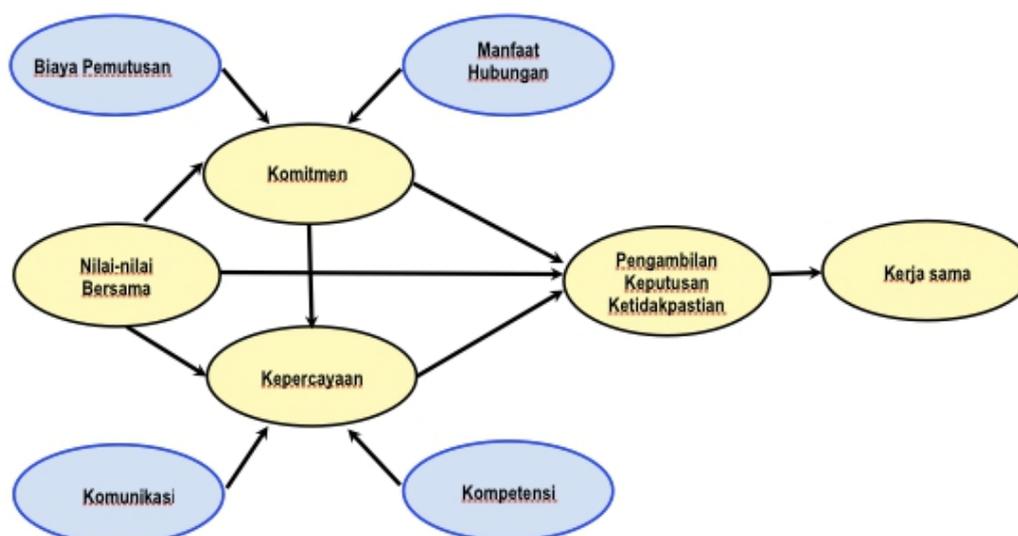


Figure 1. Conceptual Model Research

From the model above can be explained that: trust and commitment are the foundation of a business relationship. This implies a power base to stimulate the emergence of shared values that will strengthen the decision-making uncertainty. A bond of trust and commitment between partners relating signifies that would work in a mutually beneficial long term and build a positive foundation to a productive relationship. When viewed the relationship between trust and commitment fully it will increase the chances of a long-term cooperation or a strategic partnership to achieve a goal that can deliver value and profit for the company and its partners.

Trust is the basic attitudes and expectations of both parties who will partners. The trust not the things that appeared out of nowhere, but it was built by the competence and communication abilities of the parties to the partnership. This is consistent with the findings of Malecki and Tootle (1996) which says that trust requires a personal relationship that goes beyond an individual contact, and reinforced by face-to-face relationship, while the benefits of the relationship and termination costs to build a commitment, which requires long-term benefits that a relationship should be maintained and enhanced in order to fulfil the promises given.

Indicator variables are used to explain each latent variable and its source in detail described in appendix 1. There are 33 indicators variables measured by ordinal scale with 5 strata, where each stratum has a label to reinforce the strata (Lickert).

In accordance with the conceptual model that has been described, it is to predict the value of the model coefficients that indicate the degree of influence of each latent exogenous variable to endogenous latent, as well as the magnitude of the contribution of each indicator variable on its latent variable used rule Structural Equation Model (SEM). SEM is basically aimed to confirm the conceptual model of the characteristics of the population (Suharjo and Suwarno, 2002). SEM analysis is an analysis based on Confirmatory Factor Analysis (CFA), a method that combines correlation analysis, regression analysis, traffic analysis and factor analysis (Suharjo, 2002).

Modelling carried out in accordance with the modelling step Suharjo (2007), as follows:

- Development model, at this stage is to analyse the principle of causal relationships between exogenous and endogenous variables, as well as checking the validity and reliability study.
- Convert the cross diagram, this diagram serves to show the flow of a causal relationship between exogenous and endogenous. To see the flow of causal relationships created several models were then tested using SEM to obtain the most appropriate model with the criteria Goodness of Fit, and convert it into Eel Cross Diagram.
- General model equation:

$$\begin{aligned} \text{Measurements: } y &= \Lambda_y \eta + \varepsilon \\ x &= \Lambda_x \xi + \delta \\ \text{Structural: } \eta &= \mathbf{B}\eta + \Gamma\xi + \zeta \end{aligned}$$

- Determining the type of input matrix, in this study used the covariance matrix, because the purpose of the analysis is to test a model that has a theoretical justification.
- Identification of Model, problems that often arise in the structural model is the estimation of parameters can be Un Identified or Under Identified causes estimation process does not produce a unique probe and the model can not be trusted. Symptoms appear a problem of identification include: There are standard error of the parameter estimators that are too large, the inability of the program presents a matrix of information that should be presented, parameter estimation can not be obtained, figures emerged a range of negative error, and there was a high correlation -1 or 1.
- Criteria for Determination of Suitability Model, to obtain valid analytical results required several assumptions in parameter estimation and hypothesis testing. Assumption of parameter estimation and hypothesis testing include, among observational units are independent, the data are a random sample of the population and the pattern of all linear relations among variables. SEM analysis is very sensitive to the distribution of data so that major deviations to the multi normal will affect the Chi-Square test. Suitability of the model the indices mentioned above must meet the following criteria;

Size suitability	The suitability of acceptable levels
<i>Chi-square</i> (χ^2)	The smaller the better. Used small <i>chi-square</i> value so that Ho : $\sum = \sum(\theta)$, is not rejected
Probability	$\geq 0,05$
AGFI	The higher value more better. AGFI $\geq 0,90$ is <i>good fit</i> , $0,80 \leq \text{AGFI} < 0,90$ is a <i>marginal fit</i>
RMR	Residual average between the observed matrix and the estimation result. $\text{RMR} \leq 0,05$ <i>good-fit</i>
RMSEA	Average difference per <i>degree of freedom</i> is expected to occur in the population and not in the sample. $\text{RMSEA} \leq 0,08$ is a <i>good fit</i> , $\text{RMSEA} \leq 0,05$ is a <i>close fit</i> .

- Interpretation and Modification Model, if the model is good then the model can be interpreted, but if not, whether modifications need to be done. Modifications can be made when there is a significant change with the support of empirical data.
- Additionally Construct Reliability and Construct Validity also calculated respectively with the following formula:

$$CR = \frac{\sum \lambda_i^2}{\sum \lambda_i^2 + \varepsilon_i} \quad \text{And} \quad CV = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \varepsilon_i}$$

- The software used in SEM analysis is LISREL 8.7. While the methods used in estimating the coefficient is unweight least squares (ULS).

Results And Discussion

This section describes the pattern of linkages that occur among the factors thought to play a role in building a relationship based on the marketing of agricultural commodities. Initial stage is to conduct sub model analysis of relationship marketing. In terms of the SEM stage is referred to as the reliability of testing the measurement model. The goal is to determine the ability of each indicator in explaining the latent variables. To determine whether a set of indicators capable of explaining the latent variables, will be used reliability and construct validity of measurement. In this regard Hair et al. (2006) states that a construct is said reliable if the construct reliability (CR) not less than 0.7, or the construct validity (CV) is less than 0.5. While Joreskog et al. (1996) provide a simpler requirement, namely the value of the variable indicator coefficient (lambda) is the standardized value of minimum 0.5.

Based on structural models have been proposed, then the coefficients of the model to determine the value or contribution of each latent exogenous variables to endogenous latent coefficient estimate done by using the rules Unweight Least Square (ULS). The results of the model coefficient estimates that states the influence of each latent variable to the latent endogenous and exogenous contributions to the latent variable indicators presented respectively in Figure 2 below:

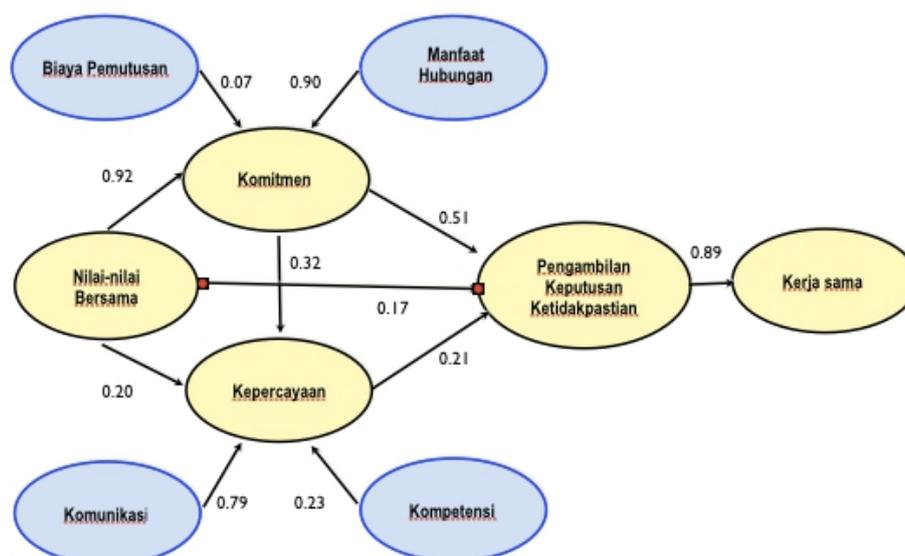


Figure 2. Coefficient Structural Equation Model relationship marketing

The results of estimation and testing of the structural model shows that the value p -value = 0.000 and RMSEA = 0073. That is, the overall empirical model can be accepted in accordance with the criteria required by Joreskog et al. (1996), although the P -value less than 0.05. From this it can be concluded that the coefficients of the model can be used as a probe the contribution or influence of latent exogenous variables to endogenous latent.

To elucidate the structural model will be discussed in stages with the expectation relationship or causal processes that occur between the latent variables can be explained in detail as follows:

Shared values

Sets of aspects that underlie the establishment of cooperate between the entities. One of the main aspects of shared values, according to Fukuyama (1995) is a trust. Aspects of trust that exist on both sides who will partner will deliver them to formulate goals and benefits of cooperation are compatible or agreed upon. In this study, shared values represented by three indicator variables, namely:

- a. Enhance the image of the company (U1)
- b. Enhance the mutual benefit financially (U2)
- c. Enhance mutual trust with (U3)

The three indicators used to describe the shared values have a relative different degree of agreement. In a separate effort to increase mutual trust has a level of agreement (agree and strongly agree) the highest, while enhancing the image of the company or business entity is relatively low. Although information on the level of agreement of each indicator can be read independently, but explains the latent variables can not be done independently because of the correlation aspect between the indicator variables to each other. In addition, sequential effects are contained in a set of indicators in explaining the latent variables. Therefore, to clarify the contribution of each variable used indicator measurement model based on Confirmatory Factor Analysis (Joreskog et al.1996).

To determine the performance or contribution of each of these indicators in explaining its latent variable, and then use the rules of Unweight Least Squares (ULS) in the SEM is intended to estimate the value of the coefficient of the measurement model. Results of analysis of the measurement model of shared values indicate that the three indicators used have a level of reliability (CR = 0.64) and adequate validity, although slightly below 0.7, but the value of CV = 0.51 and the coefficients of the indicator exceeds 0.5. The results of this analysis show that trust is an indicator of shared values that are important as a basis for establishment of cooperation, in which the value of the coefficient is 0:57 and the significance of test results of T with values above 1.96 T test ($\alpha = 0.05$). This study shows that the mutual benefit financially have a greater contribution as a foundation of cooperation between parties with a coefficient value of 0.67. While for the coefficient of image enhancement, the amount is equivalent to a 0:57 or belief variables.

Results of the measurement model test show and prove that trust is not enough to be the main basis of cooperation between parties. In cooperation efforts or business, other major aspects of the hidden even more important is the expectation of mutual financial benefit. This means that the mutual advantage to have a stronger position than the trust. What is interesting here is that there is a high emotional relationship between farmers and suppliers. For example, at once time farmers need funds while the situation was still entering the growing season and not yet in production, then the supplier easily provide loans without any doubt will be paid back at harvest time, this relationship is different from what is called “pengijon” (the plant have been purchased prior to harvest so that the course is very detrimental to farmers).

Based on these findings we can conclude that efforts to build trust and positive image of the working party are two aspects that are important, but the shared values associated with efforts to increase financial gain has a stronger role. Finally, this construct was obtained from the general picture that the establishment of cooperation between farmers, suppliers and outlets (modern and traditional market) in the business of agricultural commodities (vegetables, fruit and chicken pieces), along the major and essential to be achieved between the parties those involved is a desire for mutual benefit against the backdrop of efforts to foster mutual trust and positive image of the parties involved in such cooperation.

Relationship commitment

According to the journal is used then there are two perspectives on the perspective of the attitude of commitment (affective commitment) (Kumar et al, 1995) is a commitment to produce effect on others, and behavioural commitment (Morgan and Hunt, 1994) is more towards the commitment of the behaviour being assessed.

In practice, committed relationship is strongly associated with the values that are the result of positive evaluation of the relationship that has happened in the past. Accordingly, in this study the relationship of commitment shown by three indicator variables as follows:

- a. The growth of mutual trust that the higher (U12)
- b. Generate interest for long-term partner (U13)
- c. Generate interest to enhance cooperation in the supply of another product (U14)

Level agreement of the indicator variables suggest that the relationship's commitment to partner long-term interests and the growing confidence is a strong indicator in explaining the relationship commitment. While the interest to enhance cooperation in the provision of other products is relatively lower. By using the CFA analysis to find out the contribution of each indicator variable in explaining the relationship of commitment.

The results of the measurement analysis model and testing of the coefficient relationship commitment, it appears that the three indicators reflect these constructs have high levels of reliability and validity are very reasonable with their respective values above 0.73 and 0:51. T test results of the model coefficients showed significant results even though the value of U14 is less than 0.5. The results of this analysis indicate that a possible interest in long-term cooperation is a very strong indicator of the exercise of the commitment by both the parties work together. This is indicated by the high value of the coefficient of the indicator is 1.0. These results prove that the relationship-based marketing that occurs in agricultural commodities currently in favourable conditions in which each party involved in co-operation has been running in earnest agreement has been agreed. Not surprisingly, all the parties desire to continue the cooperation that has existed so far.

Looking at these facts interest for cooperation in the long run actually contains several meanings, among others, is because of the habits of farmers and suppliers who have established long before the existence of emotional ties of kinship and mutual trust that has grown as well as a sense of comfort in the attempt. It is also common in relationships between suppliers and retailers in traditional markets. Here the emotional bond also has a major role of the bond business because of relationships that occur more influenced by informal factors.

Trust

Melanie and Christian (2007) explain that trust is an important factor in business relationships such as corporate networks, alliances, or partnerships that involve elements such as interpersonal and inter-organizational competence, benevolence, integrity, and communication. The most important factor of trust is a positive experience that has been made in private, personal relationships that develop in accordance with time and become a bridge of communication in the relationship. Beliefs

related to commitment. It is based on the fact that in assessing the trust is often based on consistency of related parties in carrying out what has been agreed. In fact, Smith and Barclay (1997) confirmed the trust based on reputation and reinforced by the experience and the interaction between related parties. Based on the above, in this study of confidence shown by the three variables as follows:

- a. During the partnership, consistently carrying out a cooperative agreement (U27)
- b. During the partnership, consistently to fulfil what was promised (U28)
- c. During the partnership, consistently creates a feeling of trust (U29)

Estimation results of the measurement model coefficients, coefficients of test results and the value of its CR and CV. Seen that all three of these indicators have a level of reliability and validity are very reasonable with their respective values above 0.80 and 0.60. The results of this analysis indicate that the construct of trust is reflected strongly by the U28 indicator for partnership-related parties are always trying to keep what has been promised by the coefficient of 0.87. This is in line with what has been affirmed by Guilbert and Fenneteau (1998) that initial perceptions of trust can only be demonstrated through the behaviour of the parties concerned, which is reflected through the effort to keep what has been agreed.

On the other hand U27 indicator reflects the trust is consistency in carrying out cooperation with the coefficient of 0.77. This may mean that within a certain time each party seeks to establish a consistent relationship marketing practice what has been a mutual agreement. These results are in accordance with the said Gabarro, (1978) that belief starts from the weakest level (perceptions of trust), then gradually grew and strengthened through a consolidated relationship or accuracy of the fulfilment of what has been promised. It is understood that the U29 indicator for partnership-related parties is always cause a sense of trust has the lowest coefficient (0.55) because this indicator is the result of the previous two indicators.

Thus it can be said that the awakening of a trust will only be successful after going through a sense of trust that grows in stages and is followed by verification of the accuracy of what has been promised.

Decision making uncertainty

The absence or lack of commitment and trust is a burden to any of the parties that partner and can lead to dismissal. Element in the long-term relationship commitment and trust are considered to follow the rules of the stability and cumulative (Ford, 1990). But in reality sometimes the initial perception of the assumption or belief is often not evident after the cooperation takes place. That is, there are elements of uncertainty that are not anticipated at the beginning of the on-going cooperation. Based on this, in this study indicated the uncertainty of decision-making with three variables as follows:

- a. Decision to be taken, always take into account or anticipate the factors that are unexpected and can disrupt cooperation (U32)
- b. Decisions that take into account is always good or bad effects caused to the business environment (U33)
- c. Decisions are often faced with incomplete information (U34)

From interviews about the level of agreement on the indicators above, obtained the following results, in relationship marketing in agriculture, the parties always take into account the impact on the business environment (U32), has the highest level of agreement (96.1 %). In addition, approximately 95.2% of the parties who work together to anticipate the unexpected factors that can disrupt the working relationship (U33). The interesting thing is almost half of the actors in this relationship-based marketing (47.8%), in its decisions making, often do not have complete information (U34). All three indicators are considered to be capable of reflecting the latent variables, the results of the test statistic T showed significant results and in general construct has a level of reliability and validity are very reasonable with each of the CR and CV values above 0.68 and 0.50.

CFA analysis results obtained from the parties involved in cooperation in the field of agriculture stating that the decision will be taken into account is always good or bad effects caused to the business environment (U33) with a coefficient value of 0.99. It is becoming norm, especially in agriculture is very vulnerable to climate change conditions. The result of this partnership the parties take into account or anticipate the factors that are unexpected and can disrupt cooperation (u32). Meanwhile, the contribution of U34 indicator that is often faced with decisions on incomplete

information has the lowest coefficient. This could mean that partnering parties have fairly complete information related to the factors required in cooperation.

Cooperation

In many cases of cooperation between parties is due to shared goals and mutual interests. Morgan and Hunt (1994) found that commitment and trust is important to ensure the emergence of cooperation in a partnership relationship. Even, Bodi Schubert (2010) confirmed a partnership or cooperation is good and successful innovation will support the company's competence, and the search for creative solutions from the organization. It can be considered as wholly or partially the fulfilment of a reciprocal relationship between goals and expectations set together and received by each member to support the relationship (usually long term) are oriented on the efficiency and development. In this study the cooperation shown by the two variables as follows:

- a. During the partnership, always try to promote mutually beneficial cooperation (U30)
- b. Partnerships that have been awakened, plan to encourage long-term cooperation established (U31)

The results of the parameter model measurement estimates show that both indicators have high levels of reliability and validity are very reasonable with their respective values above 0.81 and 0.68. The results of this analysis indicate that the partnership has been built always focuses on the mutually beneficial cooperation (U30), and seeks to be a partnership that extends into a long-term cooperative relationship (U31), has the equivalent position in the construct of working together to build each with a coefficient of 0.82. That is, in relationships marketing in the agricultural fields studied, there are motivations of the parties who partner to constantly strive to find harmony in order to gain mutual benefits in the long run.

When referring to the findings of Bodi and Schubert (2010), this advantage can not be interpreted to benefit financially, but also on non-financial benefits such as increased competence of the company's innovation, and the search for creative solutions for the challenges of organizations in the future to ensure the work same value over the long term.

Concluding Remarks

The findings indicate that the pattern of linkages that occur in agricultural products is not only based on the purchase and sale transactions (transactional relationship) only, in which aspects of benefit financially be the main objective, but also involves psychological aspects that transform some aspect of your financial goals into a more emotional and social. The results showed the main factors that initiate the occurrence of relationship marketing in general is not solely based on the value of rupiah that is the agreement of transfer of ownership of the commodities traded, but also based on commitment and trust. This occurs because the purchase and sale transactions are generally not made in cash, but through the payment term (payment period).

This research also showed that the pattern of trading that occurs not only based on rational desire to obtain short-term benefit financially from the parties relating to business sustainability, but the relationship is also influenced by shared values which includes a sense of comfort, image enhancement and holistic long-term benefits.

References:

- Akrout, H., Akrou, W. *Interpersonal Trust in B2B Buyer/Supplier Relationship: An Empirical Study of French Industrial Sector*, The paper was published at the 23rd IMP conference in Manchester, UK. 2007,
- Brady, M K, Roehm, M.L, Consumer responses to performance failures by high equity brands, *Journal of Consumer Research*, Vol. 34, No. 4 (December 2007), pp. 537-545, 2007
- Bodi, A, Schubert. *Analyzing Relationship Success*, WIP Paper for 26th IMP Conference. 2010
- Fafchamps, M.. *Trade and social capital, Human Capital, Exports, Wages* University of Oxford, 2007,
- Ford, D. *Understanding Business Markets: Interaction, Relationships and Networks*, IMP Group, Academic Press, London. 1990.
- Fournier, S. Secrets of Customer Relationship Management: It's All about How You Make Them Feel, *Journal of Services Marketing*, Vol. 16(7), pp 700 - 703. 2002,
- Fukuyama, F. *Trust: Social Virtues and the Creation of Prosperity*. NY: Free Press. 1995,
- Gabarro, J J. The Development of Trust, Influence, and Expectations in Interpersonal Behavior. *Communications and Understanding in Relationships, Athos and Gabarro*, eds. Englewood Cliffs, New Jersey: Prentice-Hall. 1978.

- Grégoire, Y., Tripp, T., Legoux, R. When customer love turns into lasting hate, The effects of relationship strength and time on customer revenge and avoidance. *Journal of Marketing*, Vol. 73(6): pp.18–32. 2009.
- Guibert, N., Fenneteau, H. *Buyer-Seller Relationships: Towards a Dynamic Classification of Trust Antecedents*, 13th IMP Group Conference Proceedings, Lyon. 1998.
- Hair, Jr., J. F., Black, W. C., Babin, B. J., Anderson, R. E., Tatham, R. L. *Multivariate Data Analysis* (6th ed.). Upper Saddle River, NJ: Pearson Prentice Hall. 2006.
- Hofstede, G. *Culture's Consequences: Comparing Values, Behaviours, Institutions, and Organizations Across Nations*, Thousand Oaks, CA: Sage Publications. 2001.
- Irawan, B., Nurmanaf, R., Hastuti, E. L., Muslim, C., Supriatna, Y., Darwis, V., *Kebijaksanaan Pengembangan Agribisnis Komoditas Unggulan Hortikultura.*, *Jurnal Pusat Penelitian dan Pengembangan Sosial Ekonomi Pertanian*, Badan Penelitian dan Pengembangan Pertanian. Bogor. 2001.
- Jöreskog, K., Sörbom. *LISREL 8: User's Reference Guide*, Chicago: Scientific Software International. 1998.
- Kuma'at, R. *Sistem Pemasaran Sayuran Dataran Tinggi di Provinsi Sulawesi Utara*, Thesis MS –FPS IPB, Bogor. 1992.
- Kumar, N., Sheer, L.K., Steenkamp, J.E.M. The effects of perceived interdependence on dealer attitudes, *Journal of Marketing Research* 32, pp. 348– 356. 1995,
- Mawardi, M. Peranan Sosial Kapital Dalam Masyarakat., *Komunitas, Jurnal Pengembangan Masyarakat Islam*, Volume 3, Nomor 2, 5-14. 2007,
- Malecki, E J., Tootle, D M. The Role of Networks in Small Firm Competitiveness, *International Journal of Technology Management*, 11(1), pp. 43-57. 1996.
- Melanie, F., Christian, F. The Role of Trust in European Food Chains: Theory and Empirical Findings, *International Food and Agribusiness Management Review* Volume 10, Issue 2, 2007, pp. 141-164, 2007.
- Morgan, J., Hunt, S D. The Commitment -Trust Theory of Relationship Marketing, *Journal of Marketing*, 58 (July), pp. 20-38, 1994.
- Palmatier, R W., Dant, R P., Grewal, D., Evans, K R. Factors influencing the effectiveness of relationship marketing: a meta-analysis, *Journal of marketing*, 70 (October), pp. 136-53. 2006.
- Palakshappa, N., Gordon. *Trust and Commitment in Collaborative Business Relationships involving New Zealand Firms*. ANZMAC 2005 Conference: Business Interaction, Relationships and Networks. 2005.
- Smirnova, M., Kushch, S. *Understanding the Role of Relational Capabilities in Buyer-Seller Relationships*, The paper was published at the 22nd IMP- conference in Milan, Italy. 2006.
- Smith, A., Barclay, D. The effects of organizational differences and trust on the effectiveness of selling partner relationships. *Journal of Marketing*, 61, pp. 3–21. 1997.
- Smith, D. *A Conceptual Model for Retailer-Vendor Relationships Based on Commitment-Trust Theory*, 16th EDAMBA Summer Academy Soreze, France July 2008.
- Suharjo, B., Suwarno *Model persamaan structural, teori dan aplikasinya*, IPB Press, pp 11-32, 2002.
- Suharjo, B *Perbandingan metode pendugaan parameter dalam pemodelan persamaan structural*, *Jurnal Matematika IPB*, 2007.
- Wang, J S. Trust and relationship commitment between direct selling distributors and customers, *African Journal of Business Management*, Vol.3 (12), pp. 862-870. 2009.

Appendix 1. Indicators of Latent Variables and Structural Models

Variabel Latent	Variabel Indikator	Scale	Reference
Commitment	X1 Mutual trust	Likert 1-5	Canning dan Lloyd (2001) Smith (2008), Wang (2009) Anvari dan Salmiah (2010)
	X2 Long term partner		
	X3 Cooperation		
Relationship Termination Cost	X4 Impact on Supply	Likert 1-5	Morgan dan Hunt (1994) Ganesan (1994)
	X5 Impact on Profits		
	X6 Impact on Trust		
	X7 Impact on Business Continuity		
Relationship Benefit	X8 Sense of Comfort	Likert 1-5	Palakshappa dan Gordon (2005) Baglieri, Croom dan Secchi (2006)
	X9 Customer Satisfaction		
	X10 Corporate Performance		
	X11 Increased Net Income		
Trust	X12 Consistent	Likert 1-5	Akrouit dan Akrouit (2007) Smith (2008) Wang (2009)
	X13 Fulfill the Promise		
	X14 Sense of Trust (Confidence)		
Competency	X15 Infrastructure	Likert 1-5	Doney dan Cannon (1997) Morgan dan Hunt (1994)
	X16 Human Resources		
	X17 Knowledge		
	X18 Experience		
	X19 Capability		
Communication	X20 Regular Formal Meeting	Likert 1-5	Palakshappa dan Gordon (2005)
	X21 Regular Informal Meeting		
	X22 Communication of Ideas		
	X23 Communication of Problem		
	X24 Up to Date Information		
	X25 Reliable Information		
Share Values	X26 Image	Likert 1-5	Palakshappa dan Gordon (2005) Smith (2008)
	X27 Benefits		
	X28 Belief		
Decision Making Uncertainty	X29 Anticipation	Likert 1-5	Bloombaeck (2006)
	X30 Taking into account the impact		
	X31 Decision Making		
Cooperation	X32 Profitable	Likert 1-5	Walter et al. (2000), Canning dan Lloyd (2001) Palakshappa dan Gordon (2005), Smith (2008)
	X33 Long Term Planing		

Appendix 2. Coefficient Models And Values Construct Reliability and Validity

Variabel Latent		Variabel Indicator	Lamda	Error	T test	CR	CV
Commitment	X1	Mutual trust	0.53	0.72	5.50	0.73	0.51
	X2	Long term partner	1.00	0.00	6.99		
	X3	Cooperation	0.37	0.63	4.42		
Relationship Termination Cost	X4	Impact on Supply	0.94	0.12	10.14	0.8	0.61
	X5	Impact on Profits	0.64	0.60	8.18		
	X6	Impact on Trust	0.48	0.77	6.44		
	X7	Impact on Business Continuity	0.97	0.06	10.59		
Relationship Benefit	X8	Sense of Comfort	0.59	0.65	6.93	0.81	0.59
	X9	Customer Satisfaction	0.85	0.28	8.59		
	X10	Corporate Performance	0.50	0.75	6.13		
	X11	Increased Net Income	0.78	0.39	7.26		
Trust	X12	Consistent	0.77	0.41	10.58	0.8	0.60
	X13	Fulfill the Promise	0.87	0.24	11.87		
	X14	Sense of Trust (Confidence)	0.55	0.70	7.70		
Competency	X15	Infrastructure	0.69	0.53	9.73	0.76	0.54
	X16	Human Resources	0.71	0.49	10.14		
	X17	Knowledge	0.64	0.59	8.95		
	X18	Experience	0.57	0.67	7.85		
	X19	Capability	0.51	0.74	6.87		
Communication	X20	Regular Formal Meeting	0.53	0.72	4.86	0.60	0.30
	X21	Regular Informal Meeting	0.27	0.93	2.64		
	X22	Communication of Ideas	0.83	0.31	6.45		
	X23	Communication of Problem	0.57	0.68	5.96		
	X24	Up to Date Information	0.31	0.90	3.83		
	X25	Reliable Information	0.30	0.91	3.72		
Share Values	X26	Image	0.57	0.67	6.57	0.64	0.51
	X27	Benefits	0.69	0.52	7.37		
	X28	Belief	0.57	0.68	6.52		
Decision Making Uncertainty	X29	Anticipation	0.70	0.50	7.69	0.68	0.50
	X30	Taking into account the impact	0.99	0.02	9.10		
	X31	Decision Making	0.12	0.99	1.71		
Cooperation	X32	Profitable	0.82	0.32	8.73	0.81	0.68
	X33	Long Term Planing	0.82	0.32	8.72		

TARAKAN STRUCTURAL MODEL OF IMPLEMENTATION OF THE PREFERENCE PROGRAM CITY GAS UNDER ANY THEORY OF PLANNED BEHAVIOUR

Tubagus Haryono

Graduate School of Management and Business, Bogor Agricultural University, Indonesia

Ujang Sumarwan, Prof Dr Ir

Department of Family and Consumer Sciences and Graduate School of Management and Business,
Bogor Agricultural University, Indonesia

Asep Saefuddin, Ass Prof

Departement of Statistic, Faculty of Mathematics and Natural Sciences,
Bogor Agricultural University, Indonesia

Abstract:

Oil and Gas play very important role in Indonesian Economy as a source of domestic energy, feedstock for national industry and source of government income. Therefore the government of Indonesia through Law No. 22 of 2001 and Presidential Regulation No. 5 of 2006 trying to utilize alternative energy to reduce fuel subsidies including the utilization of natural gas. To support the government's plan, this study need to formulate the strategy of the Gas City development in Indonesia by identifying the factors that influence consumer's perceptions, attitudes and intentions in using natural gas, which refers to the model proposed by Icek Ajzen (1991) which is known as The Theory of Planned Behaviour.

Based on statistic examination toward γ parameter (Gamma) are able to conclude that there is a positive and significant influence from attitude toward behaviour, power of controlling behaviour factor and perceived risk toward the intention to use gas. But the attitude toward behaviour, give the most influence for decision. Trust to control of power factor give a positive and significant influence, and it happened to trust toward perceived risk. But subjective norm toward intention bring a negative influence. In addition to, the demography factor is including a significant factor in forming respondent decision for using natural gas. The respondent in higher economic status, have more independent opinion (attitude toward behaviour, the consequence evaluation, control beliefs, power of control behaviour) to determine using or not using natural gas. The suggestion which come from government and most of user is turn out them to influence using the natural gas. Likewise to the respondent in middle and low economic status but they still also have been thinking about the perceived risk that might be come when using the gas.

Key Words: city gas, intention, Theory of Reasoned Action, Theory of Planned behaviour

Introduction

Indonesian Law No. 22 year 2001 on Oil and Gas, said that the implementation of oil and gas should be able to ensure the efficiency and effectiveness of the availability of oil and gas, both as a source of energy and as industrial raw materials or as a source of state revenue. Further in Presidential Regulation No.5 Year 2006 on National Energy Policy is to be taken step-by-step use of alternative energy as a substitute fuel in order to reduce fuel subsidies, and including the use of natural gas.

Based on data sourced from the Directorate General of Oil and Gas, Ministry of Energy and Mineral Resources, in 2011, natural gas production averaged 8.43 BSCFD (Billion Standard Cubic Feet per Day or billion standard cubic feet per day). Of the total production, which is used in the country at 41.2% or BSCFD of 3.47, while those exported by 53% or by 4.47 BSCFD. While the rest of 5.8% or 0.49 BSCFD is loses. Portions export of natural gas are still great, in addition due to the obligation to fulfil contractual commitments that have been signed in the past, also due to the export

price of gas is much higher than the price of gas to the domestic market, so the producers prefer to export natural gas of the selling in the country.

Increased use of natural gas in addition will reduce kerosene consumption, it will also reduce the consumption of LPG, so the LPG which was originally use to meet the needs of urban areas will be used as additional allocations to meet the needs in areas that do not have natural gas distribution network. The areas that have the potential sources of natural gas supply, such as Samarinda, Balikpapan, Tarakan, Bontang and shoves a region which has a gas production field, or who skipped transmission pipelines or areas that already have a natural gas distribution pipelines need to be prioritized town gas.

Energy needs from day to day continues to increase, but this condition is not offset by the availability of energy resources, especially petroleum energy sources (fuel), coal and natural gas. The main factors that determine the level of demand for fuel, is economic growth, population, fuel prices, and the pattern of energy use in the past. Therefore, to ensure the security of energy supply in the country and to support sustainable development, the Government set a National Energy Policy as a guide in the management of national energy. The policy stipulated in Presidential Regulation. 5 year, 2006, on National Energy Policy.

On the other hand we are more gas reserves than oil reserves, but in fact the use of petroleum has been widely spread throughout the archipelago. While the use of gas, only spread to several cities and provinces only. Therefore, the city government communicate the formation of gas, which aims to improve the utilization of gas for our country.

Increased use of natural gas in the country, besides aiming at providing cheap energy and environment-friendly society, can also reduce the financial burden of the state in subsidizing energy, particularly oil. Providing for the needs of the natural gas with natural gas selling prices are lower than the fuel subsidy, it will encourage people to switch from using natural gas to fuel subsidies. Fuel Subsidy burden state finances have been so, especially in 2008, when crude oil prices break the USD140/barrel. At that time, the financial burden of the state to achieve energy subsidies IRD 223 billion, consisting of a fuel subsidy by the IRD 139.1 trillion and electricity subsidies for IRD 83.9 trillion. The realization of fuel subsidy (fuel) in 2011 based on sources from the Directorate General of Oil and Gas Ministry of Energy and Mineral Resources mentioned reach 165.2 trillion or IRD 127.4 per cent above the target set in budget change 2011 of IRD 129.7 trillion.

Increased use of natural gas with larger portions and more widely in various areas in Indonesia is expected to provide added value (added value) and a greater increase in macro economic side. Through integrated planning and analysis is expected to be produced a draft strategic plan and in the development of gas distribution pipeline infrastructure and facilities to serve the household consumers and small customers in various regions in Indonesia.

Research Objectives.

This study aimed to identify the presence of the influence of each of the latent exogenous variables, endogenous latent variable, by means of statistical tests on the parameters γ (Gamma) that can ultimately be one of the corner stone Gas City development strategy formulation in Indonesia. Intention model examined in this study refers to the model of Theory of Plan Behaviour proposed by Icek Ajzen (1991).

Novelty.

Based on the facts uncovered, the novelty aspect of this research is the application concepts of Theory of Planned Behaviour in the program grow the town gas. This study also proves the confidence to provide an immediate effect in the control of behaviour to accept, but found a correlation between the variables of trust with subjective norms and attitudes on behaviour, as well as the correlation between the variables attitude toward behaviour with subjective norm.

Reasonable Behaviour Theory and the Theory of Planned Behaviour

An important step in the history of the theory of attitude has developed through the Theory of reasoned Action and its successor, the Theory of Planned Behaviour (Armitage and Conner, 2001). These theories are based on the idea that the behaviour of jointly dependent on the intention and behavioural control (Ajzen, 1971; 1991). However, empirically, the operationalization of the theory of reasoned attitudes (Fishbein and Ajzen, 1975) states that the link between behavioural intention and attitude toward the behaviour. The intention is not only dependent on the attitude, but also on

subjective norms or perceived social pressure given by others, such as parents and good friends, to perform or not perform a behaviour.

The theory of planned behaviour (Ajzen, 1991) is a continuation of the theory of reasoned attitude through two variables controlling behaviour, i.e. the attitude toward behaviour and subjective norm. Fishbein and Ajzen joined forces to explore how to predict behaviours and outcomes. They assume that individuals are usually quite rational and systematic use of information made available to them. One considers the implications of their actions before they decide to engage or not engage in a particular behaviour (Ajzen and Fishbein, 1980, p 5). The theory of reasoned attitude is a theory that can predict and understand the behaviour and outcomes through attitude toward behaviour and subjective norm, this framework is known as the Theory of Reason Action.

According Sumarwan (2011) theory of reasoned attitude is the development of theories that emphasize the attitude towards an attitude object, the theory linking between the beliefs and attitudes of consumers with the intention behaviour (Sumarwan, 2011). The theory of reasoned attitudes began to become part of the social sciences, Ajzen and other researchers realized that this theory is not adequate and it has some limitations (Godin and Kok, 1996).

The main difference between the theory and the theory of reasoned behaviour attitude is the addition of a third determinant of behavioural intention, perceived behavioural control or perceived behavioural control.

Perceived behavioural control, are additional variables that exist in the theory of planned behaviour. It is assumed to have a direct impact on intention and behaviour (Ajzen, 1991). Behavioural control refers to the availability of the necessary opportunities and resources such as time, money, and the cooperation of others, have a direct impact on behaviour. Since this process is often difficult to measure, perceived behavioural control is usually used as an approach to controlling the actual behaviour (Eagly and Chaiken, 1993). It is often assumed that the two are correlated (Garling et al, 1998).

Perceived behavioural control is determined by two factors that control beliefs or control beliefs and perceived power or perceived power. Perceived behavioural control suggests that motivation is influenced by how difficult behaviours that are considered, as well as perceptions of how individuals can succeed, or cannot do activities. If someone holds strong control beliefs about the factors that would facilitate the behaviour, then the individual will have a major impact on perceived control behaviour. Instead, the person will have a low perception of control if he holds strong control beliefs that inhibit behaviour. This perception may reflect past experience, anticipating the upcoming situation, and attitudes influence the norms that surround the individual (Mackenzie and Jurs, 1993).

Trust as an element in Consumer Behaviour.

According Simamora (2003) belief is a descriptive thought that a person has about something, trust can be knowledge, opinions or simply believes and this belief will form the image of the product and brand

Some journals that discussed the theory of planned behaviour as developed by Mazzocchi et al (2004) put his beliefs as an additional predictor of consumer behaviour. Darby and Karni (1973) explains that even though the trust is a certainty, but it is similar to the knowledge and attention needs to be placed on the individual's perceived risk in combination with certain patterns of behaviour. Integration of perceived risk and trust within the framework of the theory of planned behaviour were also considered affected individual characteristics of different generating new development model called SPARTA II. SPARTA is an acronym for baseline variables suspected to establish consumer behavioural intentions, namely wiki norm or subjective norm, perceived behavioural control or perceived behavioural control, attitude toward behavioural or attitude toward the behaviour, perceived risk or perceived risk, and alia or variable sub demographics (Mazzocchi et al, 2004).

Confidence hypothesized effect on the perceived risk in particular, which only directly affects consumer intentions.

In another study Pavlou and Chai (2002) explain that trust is hypothesized as a faith that does not directly affect attitudes and intentions through perceived behavioural control, although the journal Pavlou and Chai (2002) also investigated the involvement of culture in influencing intentions.

Assessment Research Accomplished

Many researchers have contributed his thought in the realm of theory of planned behaviour. Some of the researchers like McKnight, Choudhury, Kacmar (2002), Kim, Ferrin and Rao (2007) and Dierks (2007).

McKnight, Choudhury, Kacmar (2002) in his journal entitled *The Impact of Initial Consumer Trust on Intentions to Transact With a Web Site: Faith Development Model* explains that the research conducted to develop and test a model of consumer confidence in online business (e-commerce vendors). Building trust is the key to consumer sales of web-based, strong beliefs influence consumer intentions to transact with a seller over the web. Confidence allows consumers overcome perceptions of uncertainty and risk, and link three critical behaviours that follow; receive offers advice from the seller, the seller circuitry sharing personal information, and purchase run website seller. Confidence is defined as a multi-dimensional with two interrelated components called Belief in belief itself (perception of competence, kindness and integrity of the seller) and the belief in the intention-willingness to depend (a decision that remains of the seller). Three factors are proposed to build consumer confidence: certainty structural (consumer perceptions of web security), owned by the seller's reputation and perception of the quality of the website. In this case the quality of the website and the reputation of being a major factor for fostering consumer confidence to overcome the negative perception of them. The study also suggested that perceived risk negatively impact online consumer intention to want to interact with web-based sellers.

Kim, Ferrin and Rao (2007) in his journal entitled *A Trust-based consumer decision-making model in electronic commerce: The Role of Trust, Perceived Risk, And Their Antecedents* explained that the purpose of the research to develop a theory that explains the basic process of decision-making by consumers when deciding to buy from a particular website, and test the proposed model using structural equation modelling the behaviour of buying through the Internet by using data collected through a web research, and consider the impact of these models. Results showed that consumer confidence and acceptable risk have a strong impact on the consumer to make a purchase decision. The tendency of consumers to trust the web reputation, concern for privacy and security, the quality of information from the Internet and corporate reputation has a strong effect on consumer confidence in the Internet. Interestingly, the guarantee of the three parties did not give a great effect on consumer confidence.

Dierks (2007) in his journal entitled *Does trust influence consumer behaviour* Dierks explains that examines the role of trust as a determinant of consumer behaviour in Germany. The empirical analysis suggests that the impact of belief on consumer behaviour at this time may have been overlooked. Although there is the essential element of trust is incorporated with a deeper understanding of the behaviour of consumerism. The results showed that trust plays an important role when the product to the attention of the market.

Research Location and Time.

Tarakan city is selected with an area 250.80 km² including the town close to the source of gas that consists of 4 (four) District; West Tarakan, Middle Tarakan, East Tarakan and north Tarakan. Interviews were conducted on January 19 to March 31, 2012. While the data processing began in April 2012 to complete.

Research design.

Research was conducted by using descriptive survey approach by distributing questionnaires to the public (respondents) aged 20 to over 55 years. Respondents were asked to fill out about preferences, personal factors, socio-economic, educational, and factors related to the use of natural gas as a fuel for household activities.

Research Methods.

Study used the descriptive method approach through the distribution of questionnaires to survey respondents aged 20 and over 55 years as many as 416 respondents. The variables used are the people who use or do not use natural gas as the dependent variable. While the independent variables thought to affect consumer intentions are perceived, the level of interest and demographics (personal, socio-economic, educational). Data processing is performed using AMOS 18 of SPSS software through analysis approach SEM (Structural Equation Modelling).

Sampling Techniques.

Sampling technique used in this study was multistage random sampling. Sampling started from the district till the neighbourhood by using purposive sampling. Research took a sample of 416 respondents aged 20 and over 55 years were randomized in Tarakan City. Prior to sampling, the study mapped the first sub-villages in two sub categories, namely village fed by natural gas pipelines, each taken 100 respondents / village temporarily represent this in the Karang Balik Village and Sebengkok Village.

While the village is not flowed gas pipelines represented by region West Tarakan and Central Tarakan and took 50 respondents each districts randomly. In addition, the use of secondary data obtained by the method of documentation or direct quotes from various sources through library research (library research). Studying and examining in the form of books, journals, or papers with a view to obtaining the theories and concepts related to the problem studied do the research literature.

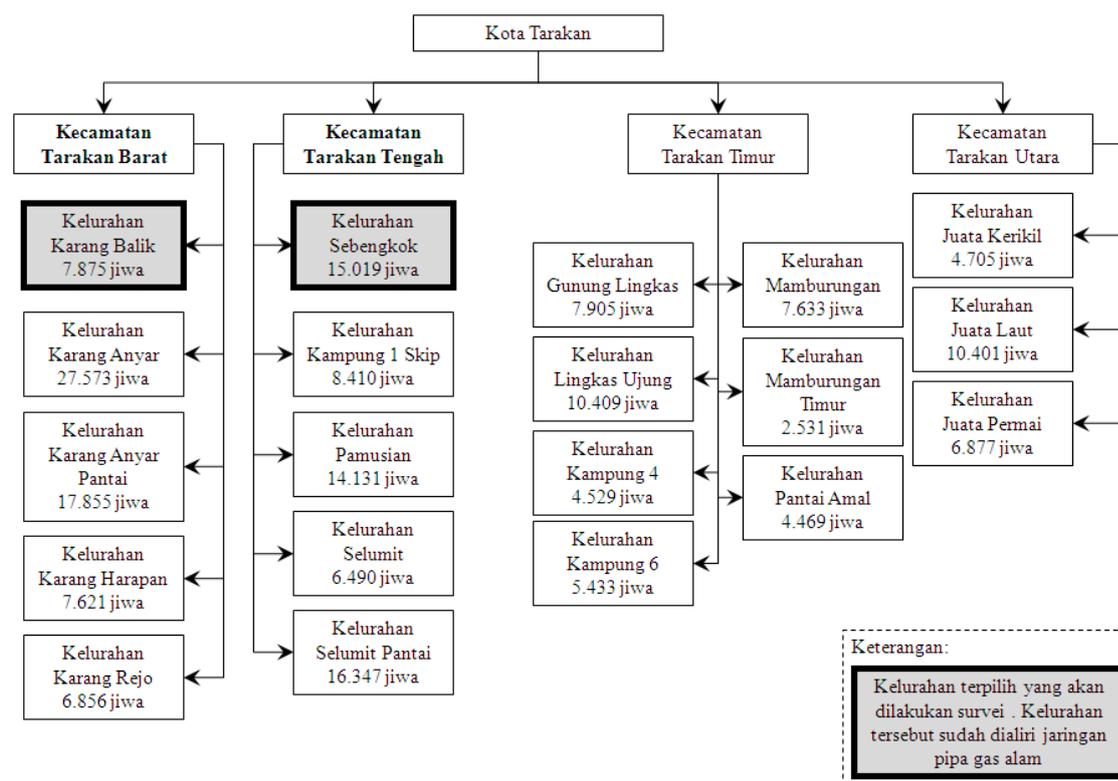


Figure 1. Sampling Technic

Descriptive analysis.

In data processing, a descriptive analysis is intended to transform the raw data set into more concise information that is easy to understand. Thus, the purpose of the analysis is descriptive is to simplify the data into a form that is more easily read and interpreted. Descriptive analysis consists of two stages. The first stage is to tabulate data on the respondents, and the second stage is to interpret the data from the tabulation. The descriptive analysis in this study is used to simplify the demographic data, perceptions and consumer behaviour.

Research variables.

Consumer Intentions consists of the intention to use and not use natural gas. Measurement intention or interest of consumers is done using a scale of measurement interval. Variable to determine the consumer intention consists of: consumer intention to use / not using natural gas (behaviour intentions, Attitude Toward Behaviour: According to Gagne and Briggs (1974) in Aiken, SC (2002), attitude is an internal state (internal state) that affect individual's choice of action to the object, person or event, wiki norm: factors outside the individual that contains a person's perception of whether others will approve or disapprove of a behaviour that is displayed (Baron and Byrne, 2000) or a person's perception of the social pressure to display or not display the behaviour (Ajzen, 2005), perceived control behaviour: perceptions about the ease or difficulty to display the behaviour of use /

continue to use natural gas. Demographics: Demographics aimed to describe the grouping of respondents into categories based on age, education level, monthly income, monthly expenses.

Method of Data Analysis

The method used in this study using SEM analysis approach (Structural Equation Modelling / Structural Equation Models. Reasons for using SEM analysis techniques in this research involve the measurement and analysis of the latent variables. Latent variables are variables that are not measured directly but measured through indicators where these indicators are measured directly / asked directly to the respondent as an object of research. In terms of SEM, the indicators that directly measured from observations of objects are often referred to as manifest variables or observed variables. Equation model of Structural Equation Modelling (SEM) allows examining a set of relationships between one or more independent variables and the dependent variable is either continuous or discrete. Basically SEM is a combination of factor analysis and path analysis. Structural model so that a formula based on the model and hypotheses are as follows:

$$\eta_1 = \gamma_{1.1} \cdot \xi_1 + \gamma_{2.1} \cdot \xi_2 + \beta_{2.1} \cdot \eta_2 + \beta_{3.1} \cdot \eta_3$$

$$\eta_2 = \gamma_{3.2} \cdot \xi_3$$

$$\eta_3 = \gamma_{3.3} \cdot \xi_3$$

While the similarities between the indicator variables and the latent variables are as follows:

$$\xi_1 = \lambda_{1.1} X_1 + \lambda_{2.1} X_2$$

$$\xi_2 = \lambda_{3.2} X_3 + \lambda_{4.2} X_4$$

$$\eta_2 = \lambda_{5.2} X_5 + \lambda_{6.2} X_6$$

$$\eta_1 = \lambda_{7.1} X_7$$

Where:

ξ_1 = Attitude toward the behavior

ξ_2 = Subjective norm

ξ_3 = Confidence

η_1 = Intention

η_2 = perceived behavioural control

η_3 = perceived risk

X_1 = Confidence behaviour

X_2 = Evaluation of consequences

X_3 = normative belief

X_4 = motivation to comply with

X_5 = power factor control

X_6 = Confidence control

X_7 = Demographics

Structural Model

In this study is a latent variable is the variable Attitude Toward Behaviour, Subjective Norm, Perceived Behavioural Control, Trust, Perceived Risk and Intention to Use Natural Gas (Intention). Latent variable is measured through indicators that can be directly measured through scores on each object of research. In the structural equation model (SEM) are two important parts, namely: Measurement Model and Structural Model. Hair et.al in the Imam Ghozali (2004) propose structural equation modelling stage and into seven steps: Development of model-based theory, the path diagram, line diagram conversion into structural equation, selection of input matrices and estimation techniques on the proposed model, Assessing the structural model identification, estimation model evaluation criteria Goodness-of-fit, interpretation and modification / re specification model. Basically, the seven steps are systematic steps to test the fit between the hypothesized model is based on the theoretical framework of the model estimates are calculated based on data from the measurement of the object of research. The degree of fit between the hypothesized models with a model based on sample data obtained indicated the suitability index model (GFI / Goodness of Fit Index).

While the reliability of the measurement of latent variables can be calculated from the estimated value of the standardized loading factor Construct Reliability view using the following formula

$$Construct\ Reliability = \frac{\left(\sum_{i=1}^n \lambda_i\right)^2}{\left(\sum_{i=1}^n \lambda_i\right)^2 + \left(\sum_{i=1}^n \delta_i\right)}$$

Where,

λ = Loading factor

δ = measurement error = $1 - \lambda^2$

Limitation reliability coefficient used is the standard coefficient of reliability of Guilford and Fruchter (1972) where values: reliability coefficient > 0.9 otherwise very reliable, 0.7 - 0.9 otherwise reliable, 0.4 - 0.7 otherwise quite reliable, 0.2 - 0.4 otherwise less reliable, the reliability coefficient < 0.2 otherwise unreliable.

Based Output AMOS 18 program, the results of testing the research model can be described in the following chart:

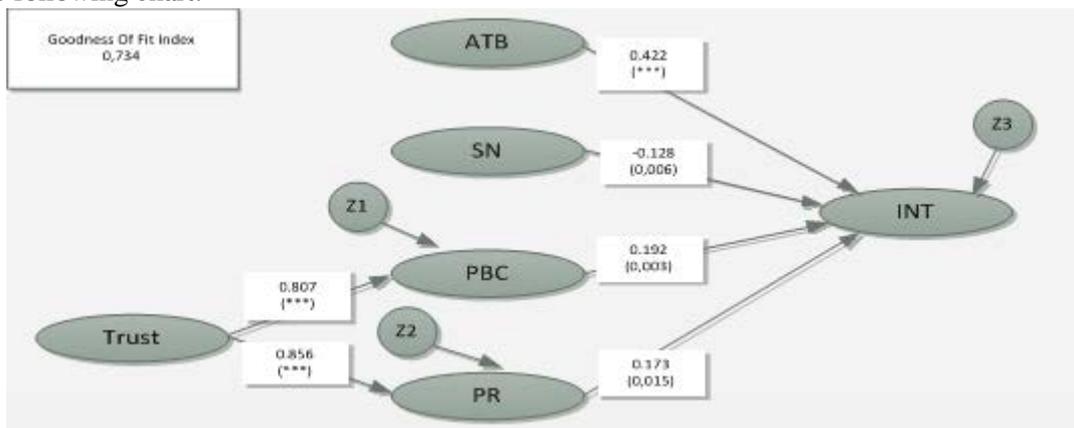


Figure 2. Structural Equation Modelling

Based on the structural model diagram above looks all parameters have estimated p-value below 0.05, thus all the parameters are expressed significant. But the value of GFI above structural models is relatively low (0.734); in the sense that it is possible to do re specification the model in order to obtain a better structural models fit the data.

Based on the modification index values AMOS output results showed no significant correlation relationship between the exogenous latent variables, namely: a correlational relationship between the latent variables Subjective Norms with confidence, correlational relationships between latent variables Attitudes Toward Behaviour with confidence, Correlational relationship between latent variables Attitudes Towards Conduct by Subjective Norms. Modification / re-specification of the hypothetical model can be described as follows:

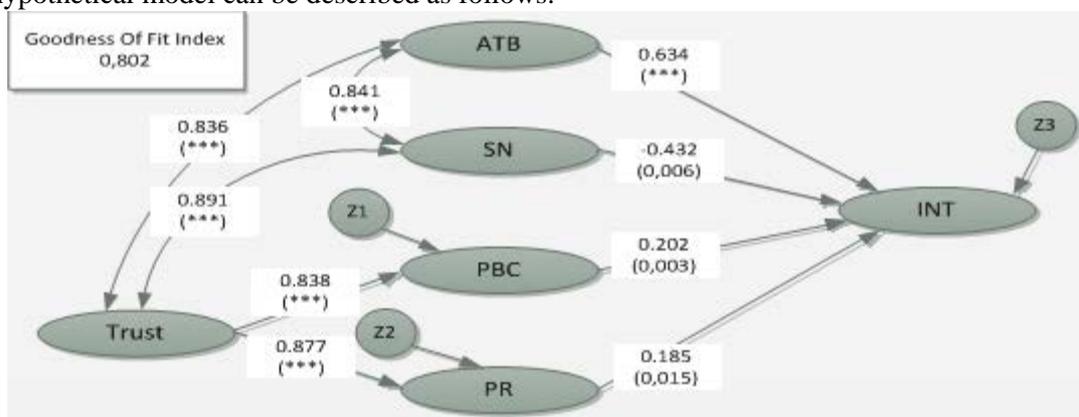


Figure 3. Modification Model

The model that has been in the re-specification of the above has a value greater than 0.802 GFI GFI value of the previous model. It shows the results of the model fit to the data re specification

more research. Furthermore, these models serve as the best model to make the interpretation of the relationship between variables.

Parameter γ (Gamma) represents the amount of influence of each of the latent endogenous latent exogenous, If the value of the parameter γ (Gamma) equals 0 (zero) means there is no effect of these factors on preference and if the parameter γ (Gamma) is not equal to 0 (zero) means there is the influence of these factors on preferences To test the hypothesis statistically tested against the values of each parameter γ (Gamma). The results of hypothesis testing on the value of the gamma parameter indicate p-value less than 0.05 chance of making mistakes.

Interpersonal Latent Variables

Based on the model already looks di re-specification correlational relationship between exogenous latent variables. The results of the statistical test for correlations between parameters of latent exogenous variables conclude that: There is a positive relationship between latent variables Subjective Norms with Confidence, There is a positive relationship between latent variables Attitudes Toward Behaviour with Confidence, There is a positive relationship between latent variables Attitudes Toward Behaviour Subjective Norms. Based on the output values obtained AMOS influence of each variable on the variable Using Natural Gas Interests that are shown in the following table:

Table 1. The Direct and Indirect Variable Intention Of Using Natural Gas.

Variable	Effect of Intention		Total Effect
	Direct	Indirect	
Attitudes toward behavioural	0.6340	0.0000	0.6340
Subjective norm	-0.4320	0.0000	-0.4320
Confidence	0.0000	0.0032	0.3320
Socioeconomic level	0.2850	0.0000	0.2850
The power factor	0.2020	0.0000	0.2020
Perceived Risks	0.1850	0.0000	0.1850

Based on the table above, we see that the variable Attitude Toward Behaviour has the greatest influence on Intentions Using Natural Gas that is equal to 0.634. While the influence of Perceived Risk variable is equal to 0.185, this variable has the smallest influence than other variables. Variable belief though not directly affect Intention Using Natural Gas, but remains an important variable that affects the variables Intentions Using Natural Gas as it can influence indirectly through a variable power control and Risk Factors in feel, while demographic factors in this case the level of socio-economic effect on the intention of using natural gas 0.285.

Attitudes of Respondents by Demographics (The Economic Status)

Demographics are a factor that has a significant role in shaping one's intention in using natural gas. Demographics are reviewed in this study is the economic status of the respondent-level factors. Based on the role of demographic factors (level of economic status) on the factors that influence the intention shown in the following table.

Table 2. Role of Demographic Factors Against Latent Factors

		A, B	C1, C2, D, E
Attitudes toward behavioural	Behavioural belief	79.0%	79.5%
	Evaluation of consequences	91.7%	89.0%
Subjective norm	Normative beliefs	71.8%	61.6%
	Motivation comply	44.8%	66.6%
Controlling behaviour	Confidence control	94.4%	85.9%
	The power factor control	94.4%	85.1%
Confidence	Confidence	83.3%	65.5%
Perceived Risks	Perceived Risks	76.2%	63.5%

The table above shows the responses of the respondents in answering agree or strongly agree to each of the latent factors based on the level of economic status of respondents. The level of economic status in this study was divided into 6 categories such as A, B, C1, C2, D, and E each have a different status levels i.e. from high to low in order. Factor A has a higher level than B, B is higher level than the C1 and so on up to the level of E. How to measure each level of the economic status was calculated by considering three measures, namely the classification based on the observation of

the surveyor, the number of facilities or property owned and expenditures each month. In this analysis created two groups of respondents, the respondents who have a high level of economic status (level A and B) and respondents with high medium and low economic status (level C1, C2, D and E).

In the table it can be seen that respondents in the high-level economic status tend to have a degree of independence or autonomy in determining the opinions or decisions that are high by respondents at the middle and lower economic status. The level of agreement among respondent on the high level of economic motivation variable comply substantially lower (44.8%) compared with the group of respondent medium and low economic level (66.6%). On the other hand, the level of agreement of respondents on the high level of economic factors attitude toward the behaviour (behavioural beliefs (79%) and evaluation of the consequences (91.7%) and behavioural control (control beliefs (94.4%) and power factor control (94.4%) appear higher than the group of respondents in the middle and lower economic level. It is clear that the respondents in the higher economic status into account the personal opinions of attitude and personality they have and come from the perception of being digested through a process to determine the respondents have confidence in the decision, in this case whether to use natural gas or not.

As has been reviewed that respondents in the high economic level are less likely to have the motivation to follow the views of others in determining the decision to use natural gas. This is supported from the facts on the ground about the agreement of respondents was quite low in receiving influence from parents / husband / wife (49.2%), advice from friends (49.2%) and advice from TV, radio and magazines (47.6%). Yet another interesting fact is seen when a high economic level respondents prefer to hear an appeal from the government (84.5%) and the advice of most people (61.9%). This suggests that this group of respondents will use the cognitive role in making decisions, i.e., following the government's appeal as a driver program or policy makers and decided to digest the information that is circulating in the community.

Table 3. Role Of Motivation Factors Demographics Comply

	A, B	C1, C2, D, E
Prompts most people	61.9%	65.4%
Influence of parents / husband / wife	49.2%	64.2%
Advice from friends	49.2%	65.9%
Appeal of the Government	84.5%	74.6%
Prompts from TV, radio and magazines	47.6%	64.1%

Conclusion

Based on the results of statistical tests on the parameters γ (Gamma) can be concluded that there is a positive and significant effect of Attitudes Toward Behavioural latent variables, latent variable Power Factor Control and Perceived Risk of latent variables to latent variable intention to use natural gas. But the attitude toward the behaviour latent variables showed the greatest influence in decision-making respondents in the use of natural gas. In addition there is a positive and significant effect of the latent variables to the latent variables Confidence Power Factor Control, and a positive and significant effect of the latent variables Belief latent variable Perceived Risk. But found a negative and significant effect of the latent variables to the latent variables subjective norm intention using natural gas.

Demographic factors such as economic status, age, gender, and level of education is also a significant factor in influencing the decision to use the respondent or not using natural gas. Respondents to the economic status of the general attitude to own independence to make decisions in the use of natural gas, while respondents in the middle and lower economic status is still worrying the risk that may occur when using natural gas. The appeal of the government is still giving a big contribution in the decision influencing respondents. Generally a young age respondents is a negative response to the program, while concerns still perceived by respondents with full-time jobs. Male respondents also had the same concerns as compared to female respondents. Level of education also affects the attitudes of respondents, mainly related to the low or the high gas prices and concerns about emerging risks.

References:

- Ajzen, I. Attitudinal vs. normative messages: An investigation of the differential effects of persuasive communications on behavior. *Sociometry*, 34, 263–80. 1971.
- Ajzen, I. The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179–211. 1991.
- Ajzen, I., dan Fishbein, M. *Understanding Attitudes and Predicting Social Behavior*. Englewood-Cliffs, NJ: Prentice-Hall. 1980.
- Armitage, C. J., dan Conner, M. Efficacy of the theory of planned behavior: A meta-analytic review. *British Journal of Social Psychology*, 40, 471–499. 2001.
- Darby, M.R. dan E. Karni. Free Competition and the Optimal Amount of Fraud. In: *The Journal of Law and Economics* 16 (April): 67-88. 1973.
- Dierks, Leef H. Does Trust Influence Consumer Behavior? Christian-Albrechts-Universität zu Kiel, 2007.
- Eagly, A. H., dan Chaiken, S. *The Psychology of Attitudes*. Fort Worth, TX: Harcourt Brace. 1993.
- Fishbein, M. dan I. Ajzen. *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*, Massachusetts: Addison- Wesley Publishing Company, Inc. 1975.
- Godin, G., dan Kok, G. The theory of planned behavior: A review of its applications to health-related behaviors. *American Journal of Health Promotion*, 11, 87–98. 1996.
- Kim, Dan J., Donald L. Ferrin, H. Raghav Rao. A Trust-based Consumer Decision-making Model in Electronic Commerce: The Role of Trust, Perceived Risk, and Their Antecedents. 2007.
- Mackenzie and Jurs.. *Theory of reasoned action/theory of planned behavior*, University of South Florida. 1993
- Mazzocchi, M., A.E. Lobb dan B.W. Traill. A Strategy for Measuring Trust in Food Safety Information: A Literature Review. University of Florence Working Paper Series on Trust No. 18. Florence University Press, Florence. – (2005): Causal Model Estimation Results. TRUST-Project Deliverable No. 9. Florence University Press, Florence. 2004.
- Mcknight, D. Harisson, Vivek Choudhury, Charles Kacmar. *The Impact of Initial Consumer Trust on Intentions to Transact With a Web Site: a Trust Building Model*. Elsevier Science B.V. 2002.
- Pavlou, Paul A dan Lin Chai. *What Drives Electronic Commerce Across Cultures? A Cross-Cultural Empirical Investigation of The Theory of Planned Behavior*. 2002.
- Simamora, Bilson. Membongkar Kotak Hitam Konsumen. Jakarta : Gramedia Pustaka Utama. 2003.
- Sumarwan, Ujang. Perilaku Konsumen. Jakarta : Ghalia Indonesia. 2011.

PREVALENCE OF POLYGLANDULAR AUTOIMMUNE SYNDROME TYPE III IN A GROUP OF ADULTS WITH THYROID DISEASES AND DIABETES MELLITUS

Gherbon Adriana MD, PhD

Assistant Profesor, Department of Physiology and Immunology University of Medicine and Pharmacy Victor Babes 2A Eftimie Murgu Square, Timisoara, Romania

Abstract:

Polyglandular autoimmune syndrome (PAS) is made up of a group of autoimmune disorders of the endocrine glands. Polyglandular autoimmune syndrome type III (PAS III) comprises autoimmune thyroiditis, immune-mediated diabetes mellitus, celiac disease, hypogonadism, myasthenia gravis, sarcoidosis, Sjogren syndrome, rheumatoid arthritis, gastric neoplasia, and malabsorption. The purpose of this study is to determine the prevalence of PAS type III in a group of adults with thyroid diseases and diabetes mellitus (DM). The studied group was of 350 cases with an age between 17-79 years.

The group of adults was subdivided according to the type of changes in the glycemic balance in 2 subgroups: DM type 1 represented by 60 cases (17.14%) and DM type 2 represented by 290 cases (82.86%).

The methods of investigation were represented by clinical, imaging, biochemical, hormonal and immunological parameters.

Key Words: Diabetes Mellitus, Autoimmune Thyroid Disease, Polyglandular Autoimmune Syndromes, Adults

Introduction

Polyglandular autoimmune syndrome (PAS) is made up of a group of autoimmune disorders of the endocrine glands [10]. There are 3 autoimmune polyglandular syndromes:

Polyglandular autoimmune syndrome type I (PAS I) is an autosomal recessive disorder caused by a mutation in the short arm of chromosome 21, characterized by the triad: muco-cutaneous candidiasis, hypoparathyroidism and Addison's disease.

The symptoms and signs appear in childhood; candidiasis is usually the first sign, followed usually by hypoparathyroidism and Addison disease [1, 12]. DM type 1 occurs in less than 4% of affected children, but increases to 12% by adults.

Polyglandular autoimmune syndrome type II (PAS II) [3, 13] is the most common endocrinopathy. Occurs in adult life and affects mostly women. The same patient has two or more of the following conditions: Addison's disease, Graves's disease, autoimmune thyroiditis, DM type 1, primary hypogonadism, myasthenia gravis and celiac disease. Most disorders are associated with the following HLA: A1, B8, DR3 (DQA1 * 0501, DQB1 * 0201) and DR4 (DQA1 * 0301, DQB1 * 0302). The autoimmune syndrome disorders present usually a long prodromal phase and the antibodies are present prior to the development of the disorder.

Polyglandular autoimmune syndrome type III (PAS III) [2] is a **PAS II** syndrome, but without the adrenocortical involvement. It comprises a group of autoimmune disorders characterized by severe glandular insufficiency. A quarter of the patients with hypo functional glands present other endocrine diseases as well. This syndrome is associated with diseases as: organ-specific autoimmune diseases (celiac disease, hypogonadism, and myasthenia gravis), organ-nonspecific or systemic autoimmune diseases (sarcoidosis, Sjogren syndrome, and rheumatoid arthritis), other diseases (gastric carcinoid tumor, malabsorption due to exocrine pancreatic deficiency), and may be classified into the following 3 subcategories:

- PAS III A – Autoimmune thyroiditis with immune-mediated diabetes mellitus
- PAS III B - Autoimmune thyroiditis with pernicious anemia

- PAS III C - Autoimmune thyroiditis with vitiligo and/or alopecia and/or other organ-specific autoimmune disease

Autoimmunity, environmental factors, and genetic factors are the 3 major factors that should be considered in the physiopathology of PAS III.

Autoimmune disease affecting a single endocrine gland is frequently followed by impairment of other glands, resulting in multiple endocrine failures. The identification of circulating organ-specific auto antibodies provided the earliest and strongest evidence for the autoimmune pathogenesis of polyglandular failure syndromes [2].

Some studies show that environmental precipitators of autoimmunity might play a role in polyglandular autoimmunity. Viral infection may exaggerate the ongoing immune response and precipitate glandular failure (ex. the role of congenital rubella infection in etiopathogenesis of type 1 diabetes mellitus and hypothyroidism) [10].

PAS III, as well as PAS II, is associated with HLA class II genes with apparently distinctive HLA alleles for each. The underlying non-HLA genes of PAS III remain to be further defined genetically. PAS III is often observed in individuals in the same family, suggesting that its inheritance could be an autosomal dominant trait with incomplete penetrance. [6, 9, 15].

Multigenetic involvement in the development of the individual components of PAS III has been proved. For example, DM type 1 is linked to several loci in non-HLA genomic regions. Furthermore, autoimmune thyroiditis also is polygenic [4].

Family and population studies showed that the PAS III A has a strong genetic background. Several gene variations present in both autoimmune thyroiditis and DM type 1 have been identified by whole genome and candidate gene approaches [4].

About PAS III epidemiology, the exact international prevalence is unknown. The morbidity and mortality of PAS III is determined by the individual components of the syndrome. PAS III typically is observed in middle-aged women but can occur in persons of any age; it is more common in females than in males and no racial or ethnic difference in its frequency has been reported [2].

Characteristic for PAS III is the absence of adrenal insufficiency. Once adrenocortical insufficiency develops, such patients are reclassified as having PAS II. The involvement of multiple glands may be apparent at the time of initial presentation, but, more commonly, individual glandular failure develops sequentially. No specific sequence exists by which the individual glandular failures develop. The clinical symptoms of PAS III are a constellation of manifestations of endocrine gland failures that comprise the syndrome [10].

Material And Method

Method

Investigated Population

350 people with DM (307 F and 43 M), aged between 18 and 79 years represented the studied group.

Depending on glyceimic balance, the group was divided into:

- the group with DM type 1 – 60 (17, 14%) (55 F and 5 M)
- the group with DM type 2 – 290 (82, 86%) (252 F and 38 M)

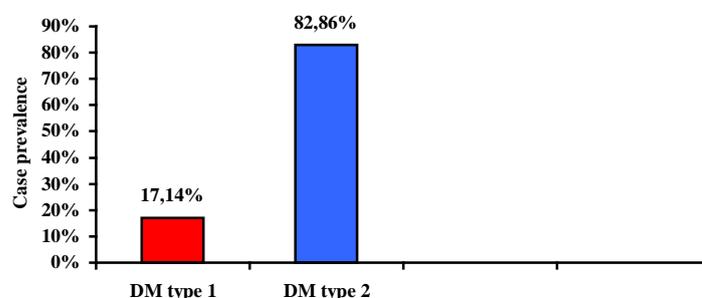


Fig. 1. Cases classification depending on glyceimic balance

Methods Of Investigation

The methods of investigation were represented by clinical data - case history, current status, imagistic- thyroid ultrasound, biochemical - for glyceimic balance: fasting blood glucose, glycosylated

hemoglobin, investigation of the thyroid gland: TSH, FT₄, FT₃, thyroid antibodies, investigation of the adrenal gland: ACTH, 21-hydroxylase antibodies, gonadotropins: FSH, LH and appropriate sex hormones (testosterone, estradiol), investigation of celiac disease: antitissue transglutaminase antibodies, investigation of pernicious anemia: complete blood count with mean cell volume and vitamin B₁₂ levels, parietal cell and anti-intrinsic factor antibodies.

Determination of plasma glucose was performed by enzyme technique with glucosooxidasis. Normal values were taken between 70 - 110 mg%; diabetes mellitus - values equal or over 126 mg%, impaired glucose tolerance - values between 110 - 125 mg% and the OGTT at 2 h between 140 - 200 mg% and impaired fasting glucose - values between 110 - 125 mg% and OGTT at 2 h under 140 mg%.

Determination of HbA1c was achieved through the DiaStat for measuring HbA1c reported to the total HbA.

To determine the TSH level in plasma, the free fraction of triiodotironin (FT₃), and the plasma free fraction of thyroxin (FT₄) were performed a quantitative method ARCHITECT; witch is an immunological method, Chemilumnescent Microparticle Immunoassay (CMIA). Normal values were following: TSH = 0.465-4.68 Miu/ml, FT₃ = 3.69 -10.4 pmol/l, FT₄ = 10-28.2 pmol/l.

To obtain the level of cortisol was performed the technique IMMULITE / IMMULITE 1000, an imunometric method, in solid phase, competitive, of chemiluminescent, Immuno Chemilumino Enzymometric assay (ICEM). It was considered normal: a.m. 5-25 microgram/dl.

FSH level was measured quantitatively by the ARCHITECT method; a Chemilumnescent Microparticle Immunoassay. Reference values: determined with ARCHITECT test.

Table I. The reference values for FSH

Population field	mIU/ml
Women:	3.35 – 21.63
- Follicular phase	4.97 – 20.82
- Ovulating phase	1.11 – 13.99
- Luteal phase	2.58 – 150.53
- Postmenopausal	
Men	1.37 – 13.58

Testosterone was determinate by ELISA method. The references values are depending by age and gender:

Adults:

- men: 0.019-0.145 nmol/L;
- women in fertile period: <0.014 nmol/L;
- pills: 0.001-0.0069 nmol/L;
- postmenopausal: 0.0003-0.0058 nmol/L.

Estradiol was determinate by immunochemical with electrochemiluminiscent detection method (ECLIA). The references values are depending by age and gender, and at women also with the menstrual cycle period and pregnancy.

Table II. The reference values for estradiol

Age and gender	References values (pmol/L)
Adults – Women • Follicular phase	46.0-607
• Ovulating phase	315-1828
• Luteal phase	161-774
• postmenopausal	<18.4-201
– Men	28.0-156
Pregnancy (first quarter)	789 – 15781
Children (1-10 years) • girls	22.0-99.1
• boys	<18.4-99.1

The immunological parameters were represented by autoimmune thyroid markers - antibodies (antiTPO and antiTg antibodies).

To determine serum levels of antiTPO antibodies it was used the kit AxSYM antiTPO, an immunological method (Microparticle Enzyme Immunoassay) (MEIA). Normal values: antiTPO antibodies <35 IU/ml.

To determine serum levels of antiTg antibodies it was used the kit AxSYM antiTg, a MEIA method as well (Microparticle Enzyme Immunoassay). Normal values: antiTg antibodies <55 IU/ml.

To determine 21-hydroxylase (anti 21-OH antibodies) antibodies level it was used the radioimmunodetermination method combined with a technique of immunoprecipitation, based on human 21-OH marked with I 125 reacting with the antibodies anti 21-OH from the samples test and forming immune complexes that precipitated with the solid-phase of protein A. Normal range: <1 IU/ml

ACTH was determinate by immunoassay with chemiluminescent detection method.

Antitissue transglutaminase antibodies were determinate by ELISA method.

References values: IgA, IgG : <10 U/mL: negative; ≥10 U/mL: positive.

Vitamin B₁₂ levels were determinate by immunochemical with electrochemiluminiscent detection method (ECLIA). References values: 191-663 pmol/L (for european population).

Parietal cell antibodies were determinate by indirect immunofluorescence. References values: negative.

Anti-intrinsic factor antibodies were determinate by ELISA method.

References values: < 6 U/mL: negative.

Determination of complete blood count was achieved with automatic method: electric impedance method. Normal values: erythrocytes = 4-5.5 mil/mm³ (men: 4.9 ± 0.7 mil/mm³, women: 4.3 ± 0.6 mil/mm³), leucocytes = 5000 – 9000 mil/mm³, plateled = 150000 – 350 000/mm³, hematocrit (Ht): men 45 ± 7%, women 42 ± 5%, hemoglobin (Hb): men: 15 ± 2 g/dl, women: 14 ± 2 g/dl.

Constants and red cell indices are calculated automatically, depending on the values of Hb, Ht and red blood cells (RBC) count. Normal values: mean corpuscular volume (MCV) = 80-100 fl, mean corpuscular hemoglobin concentration (MCHC) = 32-36 g Hb/100 ml erythrocytes, mean corpuscular hemoglobin (MCH) = 27-32 pg.

Thyroid ultrasound was performed in all cases and allowed us to measure thyroid volume, thyroid study and the changes in parenchyma's density.

An increased density, uniform, characterizes normal thyroid parenchyma easily distinguished from the neck muscles that are hypo dens.

Inflammatory processes and autoimmune pathology appears hypo dens. The scale was assessed as being discreet +, moderate ++ and marked +++.

In the autoimmune thyroid disease the parenchyma of the gland appears hypo dens.

Chronic autoimmune thyroid disorder appears with a hypoeogenity of the parenchyma and normal or increased thyroid volume.

Results And Discussion

In the group of adults 17.14% had DM type 1 and 82.86% had DM type 2.

In the group with DM type 1 the main endocrine immune combinations were represented by DM type 1 with autoimmune chronic thyroiditis (ACT). Other endocrine immune associations were represented by autoimmune ovarian insufficiency and the nonendocrine disorders as vitiligo, alopecia, Biermer anemia (Tab. III).

Table III. Prevalence of endocrine autoimmune disorders in the studied group

Associations	Subject group	
	No.	%
DM type 1	60	
DM type 1 + ACT	31	51.66%
DM type 1 + ACT + decalvant pelad	2	3.33%
DM type 1 + ACT + vitiligo	3	5%
DM type 1 + ACT + autoimmune ovarian failure	6	10%
DM type 1 + ACT + vitiligo + Biermer anemia	1	1.66%
DM type 2	290	
DM type 2 + ACT + vitiligo	6	2.06%

In the group of adults with DM type 1 the first immunopathy was DM type 1, present in 24 of the cases and was associated with ACT in all 24 cases. In 2 cases, thyroid disorder and DM type 1 were detected at the same time. In 17 cases thyroid disorder preceded the DM type 1 (Tab. IV).

Tracking the association with autoimmune ovarian insufficiency led to determine the levels of FSH, which was increased > 25 IU/l in 6 cases. Primary ovarian insufficiency (early menopause) usually occurs before the age of 40 years (in the absence of iatrogenic causes) and its clinical signs are secondary amenorrhea and hypergonadotropism with hypoestrogenemia.

Autoimmune ovarian insufficiency (AOI) is usually associated with other autoimmune pathology such as diabetes mellitus type 1, Addison's disease, ACT, vitiligo, etc.; its diagnosis is difficult and it is usually based on the exclusion of other possible causes of primary ovarian insufficiency and the notice of autoimmune etiology [2].

Also estrogen therapy in autoimmune ovarian insufficiency may increase the risk of cardiovascular disease [11].

Endocrine immunopathys may be linked to a variable incidence of systemic organ-specific nonendocrine disorders.

In 4 cases it was associated vitiligo, which occurred before the onset of endocrine immunopathys. Pelade decalvant appeared in 2 cases and also preceded the onset of autoimmune endocrinopathys. One case in the group of adults with autoimmune endocrine diseases had associated more than one nonendocrine autoimmune disorder respectively Biermer anemia and vitiligo.

Table IV. Range (years) between the onsets of immunopathys in adults with type 1 diabetes

Period of time	No. of cases	Media \pm SD (age)
Onset DM – Onset ACT	24	22.29 \pm 12.42
Onset ACT - Onset DM	17	2.47 \pm 1.94

In adults, the average time between the onset of type 1 DM and ACT was 22.29 ± 12.42 years.

In 17 cases the first immunopathy was ACT, followed at a distance of 2.47 ± 1.94 years by DM type 1.

All the patients with DM type 1 and ACT were women, with the median age 43 ± 18.95 years and the average age of onset 30.46 ± 22.94 years. It was no thyroid familial history disease found.

In the study group, all patients had DM type 1 clinically manifest, all being treated with insulin in different therapeutic schemes

Prevalence of ACT in DM type 2 was 26.55% (77 patients, 69 F and 8 M). In adults with DM type 2 PAS type III was found in 6 (2.06%) cases, of which all have ACT + vitiligo. The first disease was vitiligo follow by ACT after 2-5 years. All cases were females. The median age was 55 ± 14.14 years and the average age of onset was 53.5 ± 12.02 years. DM type 2 appears after ACT, at 5-10 years, possible because of replacement therapy for thyroid disease. It was no thyroid familial history disease found. At all patients with DM type 2 the treatment was diet.

Also in these cases it is useful to determinate the antibodies for diabetes because these patients may be latent autoimmune diabetes in adults (LADA) types and in time requires insulin.

The prevalence of PAS III is unknown. It is more often met in women than by men; usually by middle-aged women but it can occur in people of any age. Death is determined by the individual components of the syndrome [2].

So, in our study, the prevalence of PAS type III was 14% (100% F and 0% M, $p < 0.001$, $X^2 = 52.69$).

PAS type III prevalence in DM type 1 was 71.66% (100% F and 0% M, $p < 0.001$, $X^2 = 67.01$) and 2.06% in DM type 2 (100% F and 0% M, $p = 0.01$, $X^2 = 6.06$).

In the case of DM type 1 we have PAS type III A, and in the case of DM type 2 only PAS type III C.

Significant differences regarding PAS prevalence were found between the group with DM type 1 and 2 (71.66% vs. 2.06%, $p < 0.001$, $X^2 = 200.01$).

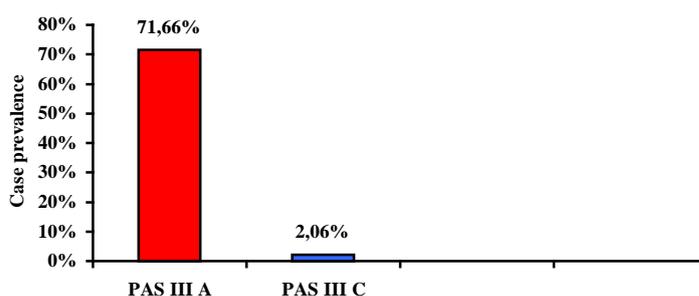


Fig. 2. PAS prevalence in study group

In general, in the first stage of PAS antibodies levels are elevated. In the second stage the disease is sub clinical and in the third stage becomes clinically manifested.

By ACT, 24 cases were with euthyroidism and 96 cases with hypothyroidism (76 clinical cases and 20 cases sub clinical). 24 cases did not require treatment; the remaining had substitution treatment with thyroid hormones. AntiTPO antibodies were present by 27 of cases of ACT and DM type 1, the remaining cases presenting insignificant values.

Ideal is to determine the presence of antibodies, especially in DM type 1, because they may be present by subjects without clinical symptoms. If their levels are raised, it is good to monitor annual the TSH level and if it is normal it is recommended to doze antithyroid antibodies by intervals of 2-3 years [5, 14].

Also, if the disease is autoimmune, the patient should be investigated for other autoimmune associations of endocrine or nonendocrine nature.

A study in Czech Republic on 51 patients with DM type 1 showed that it is associated with autoimmune thyroid diseases, with Addison's disease and celiac disease.

The authors recommend finding the specific antibodies for each disease, to diagnose the disease in the initial phase, and to prevent the complications that will affect the quality of the patients' life [7].

If DM type 2 is present it is recommended to evaluate TSH levels, and if it is normal, to repeat this evaluation every 5 years.

If pre-existing thyroid pathology is present it is recommended to evaluate plasma glucose levels annually.

Conclusion

The prevalence of PAS type III in the study group was 14%; all the patients with this were middle-aged women.

PAS type III has prevailed in the group with type 1 diabetes (71.66%) due to autoimmune origin, part of the polyglandular autoimmune syndrome (PAS) type III A.

In the case of DM type 2 the prevalence of PAS III was only 2.06%, and the type was PAS III C. In these cases the patients with autoimmune disease may be type latent autoimmune diabetes in

adults (LADA). So, if we have a patient with two or more autoimmune disease, we must investigate this for another possible autoimmune disease.

Many disorders involved in PAS present a long prodromal phase, characterized by the presence of characteristic antibodies for each disorder in part, before the clinical manifestations.

The treatment of patients with PAS involves early identification of all components.

The treatment of PAS is currently the treatment of each component of endocrine disorder (usually through hormone substitution therapy).

Isohormonal therapy has "immunomodulatory" capacities (hormone produced by the target organs may be able to influence autoimmunity).

Associations of specific autoimmune endocrinopathys require specific management. Controversial discussions are described in the literature on the effectiveness of thyroxin in patients with positive antibodies, but with euthyroidism or sub clinical hypothyroidism [8].

Some show a significant reduction of the TSH and of the anti-TPO antibodies in patients with autoimmune thyroiditis and euthyroidism after 1 year of treatment with thyroxin [8].

The PAS classification is not final. This may change over time, with the onset of new endocrine disorders or associations with new autoimmune determination.

References:

- Aldasouqi SA, Akinsoto OPA, Jabbour SA. Polyglandular Autoimmune Syndrome, Type I, In *Endocrinology* (electronic book), 2006, pag. 1 – 18
- Aung K, Salmon M. Polyglandular Autoimmune Syndrome, Type III, In *Endocrinology* (electronic book), 2006, pag. 1 - 18
- Betterle C, Lazzarotto F, Presotto F. Autoimmune polyglandular syndrome Type 2: the tip of an iceberg? *Clin Exp Immunol*, 2004; 137(2): 225 – 33
- Dittmar M, Kahaly GJ. Genetics of the autoimmune polyglandular syndrome type 3 variant. *Thyroid*. Jul 2010;20(7):737-43.
- Eisenbarth GS, Gottlieb PA. Autoimmune polyendocrine syndromes. *N Engl J Med*, 2004; 350: 2068 – 2079
- Fourati H, Mahfoudh N, Abida O, Kammoun A, Mnif F, Haddouk S, et al. HLA-DRB1/DQB1 susceptibility for autoimmune polyglandular syndrome type II and III in south of Tunisia. *Ann Endocrinol (Paris)*. Jun 2011; 72(3):232-8
- Gonem S, Wall A, De P. Routine screening for thyroid disease at patients with diabetes mellitus: is it worthwhile? *Endocrine Abstracts*, 2007; 13: P 317
- Hosseini GR, Tuttle M., Baskin HJ, Fish LH, Singer PA, McDermot MT. Subclinical thyroid dysfunction: a joint statement on management from the American Association of Clinical Endocrinologists, The American Thyroid Association, and the Endocrine Society. *J Clin Endocrinol Metab*, 2005; 90: 581 – 585
- Huber A, Menconi F, Corathers S, et al. Joint genetic susceptibility to type 1 diabetes and autoimmune thyroiditis: from epidemiology to mechanisms. *Endocr Rev*. Oct 2008; 29(6):697-725
- Kahaly GJ. Polyglandular autoimmune syndromes. *Eur J Endocrinol*. Jul 2009;161(1):11-20.
- Kalantaridou SN, Braddock DT, Patronas NJ, Nelson LM. Treatment of autoimmune premature ovarian failure. Case report. *Human Reproduction*, 1999; 14 (7): 1777 – 1782
- Myhre AG, Halonen M, Eskelin P. Autoimmune polyendocrine syndrome type 1 (APS I) in Norway. *Clin Endocrinol (Oxf)*, 2001; 54(2): 211 – 7
- Sivarajah S, Fan CY, Akinsoto OPA. Polyglandular Autoimmune Syndrome, Type II. In *Endocrinology* (electronic book), 2006
- Tunbridge WMG, Vanderpump MPJ. Population screening for autoimmune thyroid disease. *Endocrinol Metab Clin North Am*, 2000; 29: 239 – 253
- Villano MJ, Huber AK, Greenberg DA, et al. Autoimmune thyroiditis and diabetes: dissecting the joint genetic susceptibility in a large cohort of multiplex families. *J Clin Endocrinol Metab*. Apr 2009;94(4):1458-66.
- White RD and Harris GD. "Birds of a Feather Flock Together": Type 1A Diabetes and Other Autoimmune Disease States. *Clin. Diabetes*, 2006; 24 (1): 40 – 43

INFLUENCING FACTORS IN MMR IMMUNISATION DECISION MAKING

Marie C. Hill, MHM, PgDip, BSc (Hons), RN.
Senior Lecturer, City University London

Carol L. Cox, PhD, MSc, MA Ed, P G Dip Ed, BSc (Hons), RN, FAHE
Professor of Nursing, Advanced Clinical Practice and Associate Dean Research and Enterprise,
School of Health Sciences, City University London

Abstract:

Immunisation of children is not a straightforward process for parents.

The United Kingdom [UK] has a structured immunisation programme which continues to evolve and develop in meeting the demand to improve and control preventable infectious diseases.

Following a measles outbreak in Quebec, Canada in 1989 it was suggested that a single Measles, Mumps and Rubella [MMR] vaccine did not provide enough seroprotection to ensure herd immunity levels of 95%. Therefore the UK introduced a second dose of the vaccine in 1996.

Factors influencing parental decision making in relation to uptake of the MMR vaccine are the prevention of disease, the consequences of contracting infectious diseases, perceived pain and the media.

Practice Nurses are a credible source of information that parents actively seek to inform their decision making in relation to the immunisation of their children.

Immunisation decision making is not a straightforward process for parents. Many factors influence parental decision making on whether they immunise their child with the Measles, Mumps and Rubella [MMR] vaccine. The feasibility study described in this article provides insight into influencing factors associated with decisions regarding the immunisation of children by parents. The study findings suggest the practice nurse is a credible source of information parents actively seek to inform decision making. At a time when the incidence of measles and mumps is rising in the United Kingdom [UK] the provision of appropriate information by the Practice Nurse has the potential to increase uptake of the MMR vaccine.

Key Words: MMR Vaccine, Immunisation, Influencing Factors, Practice Nurses, Decision Making

Introduction

Immunisation is a proven tool for controlling and eliminating life-threatening infectious diseases. It is estimated to avert between 2 and 3 million deaths worldwide annually (World Health Organisation [WHO], 2010a). After clean water, immunisation is the most effective public health intervention in the world for saving lives and promoting good health (Health Protection Agency [HPA], 2010a). The WHO credits immunisation as one of the most cost-effective health investments in the world today (WHO, 2010b).

This article presents insight into influencing factors associated with decisions regarding the immunisation of children by parents. It begins by providing an account of the background to the resurgence of measles and mumps in the UK and the role of health professionals involved in the administration of the MMR vaccine. This article describes the perspective of five parents whose children have been immunised with the MMR vaccine. In particular, it explicates the rationale, aims, design, methodology, results and discussion associated with a feasibility study which provides the foundation for a larger study that will discern influencing factors in parental decision making associated with the MMR vaccine.

Resurgence of Measles and Mumps

The UK has a structured childhood immunisation programme, which continues to evolve and develop in meeting the demand to improve and control preventable infectious disease (Salisbury et al,

2006). The overall aim of the childhood immunisation programme is to protect all children against the preventable diseases of: diphtheria, pertussis, tetanus, polio, haemophilus influenzae type b, meningococcal serogroup C, measles, mumps, rubella and pneumococcus (Salisbury et al, 2006). Many of these vaccines are combined vaccines. One of these combination vaccines is the MMR vaccine, which is recommended to be administered at 13 months and again at approximately 4 years of age (Salisbury et al, 2006). The success of any immunisation programme, such as the childhood immunisation programme, is dependent on meeting 'herd immunity' levels to prevent local outbreaks and epidemics of the diseases it is targeting. Therefore, it is crucial that the uptake of vaccines such as the MMR vaccine meets the recommended herd immunity level of 95% of the targeted population in order to prevent disease outbreaks of measles, mumps and rubella (Salisbury et al, 2006).

Measles has been a notifiable disease in England and Wales since 1940 (Salisbury et al, 2006; Asaria and MacMahon, 2006; Kassianos, 2001). A single monovalent measles vaccine was introduced into the childhood immunisation programme in the UK in 1968 (Salisbury et al, 2006). However, due to low uptake of the vaccine an interruption of measles transmission over the next 20 years was not possible. Subsequently there were between 50,000 and 100,000 notifications annually (Salisbury et al, 2006). In an attempt to increase uptake of the measles vaccination, the trivalent MMR vaccine was introduced in October 1988 to replace the measles vaccine (Salisbury et al, 2006) and to reduce mortality rates (Jansen et al, 2003; Miller 1985). Following a measles outbreak in Quebec, Canada in 1989, it was suggested that a single MMR vaccine did not provide enough seroprotection to ensure herd immunity levels of 95% (Salisbury et al, 2006; Kassianos, 2001). Effectiveness of a two dose schedule has been demonstrated in Finland and the United States of America [USA] (Salisbury et al, 2006). Therefore, a second dose of the vaccine was added to the UK childhood immunisation programme schedule in October 1996 (Salisbury et al, 2006). Initially, the addition of a second MMR vaccine demonstrated a reduction in outbreaks of measles and rubella from 1996 - 1998 (HPA, 2011a). Despite the reduction in the incidences of measles and rubella, notifications of mumps continued to rise during this timescale with 94 confirmed cases of mumps in 1996 to 121 cases in 1998 (HPA, 2011b). In 1998, Wakefield et al published a paper in *The Lancet* regarding side effects of the MMR vaccine. Reaction by the public to the article was much like the response of the public to the pertussis vaccine. There was considerable professional and public anxiety regarding the safety and efficacy of the pertussis vaccine. Subsequently, major epidemics of pertussis occurred in 1977 - 1979 and again in 1981 - 1983 with over 68,000 notifications and 14 deaths (Salisbury et al, 2006; Baker, 2003). Similar to the reactions to the pertussis vaccine, the publicity generated by the Wakefield article led to scepticism amongst parents about the safety of the MMR vaccine and its alleged association between the vaccine, autism and Crohn's Disease (Yarwood et al, 2005). Consequently, this concern manifested itself in a decrease in the uptake of MMR vaccinations in the UK from 95% in 1993 (Kaye et al, 2001) to 80% in 2003-2004 (NICE, 2010). The decrease is despite a retraction by the majority of the researchers in an article published in *The Lancet* refuting an association of any of the syndromes described in the original study (Murch et al, 2004). Not surprisingly, the decline in vaccination rates of MMR has led to a resurgence of measles and mumps in particular. In 1998, the confirmed cases of the diseases were: measles n = 42 (HPA, 2010b); mumps n = 67 (HPA, 2010c) and rubella n = 28 (HPA, 2010d). The HPA's data for confirmed cases of these diseases in 2009 was: measles n = 876 (HPA, 2012a); mumps n = 5695 (HPA, 2012b) and rubella n = 9 (HPA, 2012c). The increases in the infectious diseases of measles and mumps notably, is despite widespread campaigns to promote the vaccine (Lamden and Gemmell, 2008) and the evidence on the vaccine's safety and effectiveness (Taylor et al, 1999; Farrington et al, 2001; Dales et al, 2001; Kaye et al, 2001; Madsen et al, 2002; Taylor et al, 2002; Madsen and Vertergaard, 2004; Honda et al, 2005).

Health Professionals Involvement in the Administration of the MMR Vaccine

Several factors have been cited in the literature that influences immunisation decision making. These have been grouped as themes that have emerged from the literature (Table 1). One of these factors is the relationship with health professionals. Research supports the influence of the General Practitioner and Health Visitor on parental immunisation decision making relating to the MMR vaccine (Harrington et al, 2000; Petrovic et al, 2001; Evans et al, 2001; Smailbegovic et al, 2003 and McMurray et al, 2004). However, this literature is silent in relation to the Practice Nurse who is the principal health professional involved in the administration of the childhood immunisation

programme (Maconachie and Lewendon, 2004).

Table 1 Immunisation Factors Influencing Parental Decision Making

The principal aim of this feasibility study described in this article has been to ascertain influencing factors on parental immunisation decision making. This study has sought to:

- i) Explore the factors that influence parental decision making on whether to immunise their child with the MMR vaccine
- ii) Ascertain who the parent seeks information from to influence their immunisation decision making on the MMR vaccine.

Design

The research design developed for this pilot study was descriptive first level exploratory research employing a Modified Grounded Theory method for data analysis and synthesis as described by Burnard (1991). It was determined, following a literature review of the subject on parental decision making associated with immunisation of their children with the MMR vaccine, that a small scale feasibility study should be undertaken prior to initiating a large scale study. Findings from the study could provide rationale for a more substantive study.

Sample

The sample was drawn from five parents of varying socioeconomic backgrounds and education (Table 2) whose children had received their MMR vaccine. The first author approached three general practices in London to participate in the study. All agreed to be involved. Each general practice was based in a different National Health Service [NHS] Trust. There were specific inclusion and exclusion criteria for this study. All parents had to be fluent in the English language; have had their child immunised within the previous 12 months; be registered in one of the general practices agreeing to take part in the research and be aged between 18 - 45 years of age. Only parents who had parental responsibility were recruited, which Griffith (2008) defined as enabling an individual to make decisions in a child's life, which would include consenting on behalf of a child for medical treatment. Four (80%) were female and one (20%) male. The mean age of participants was 32 years. The use of pseudonyms was deemed appropriate to protect the identity of all participants. The ethnic backgrounds of the parents were not homogenous. Demographics of the participants are shown in Table 2.

Table 2 Participant Demographics

Participants	Ethnicity	Socio-economic Group* * The National Statistics Socio-economic Classification	Gender	Education	Number of children and ages of children
Jo	Ghanaian	L 14 (Unemployed)	Female	Secondary education	3. 18 months, 4

		for more than one year)		until 16 years.	years and 6 years
Sharon	British White	7 L13 (Sales Assistant)	Female	Secondary education until 15 years.	1. 18 months.
Ulrika	Turkish	2 (Mathematician)	Female	University graduate	1. 14 months.
Cheryl	British Black	L 15 (Full time student)	Female	University graduate	1. 18 months.
Mark	British White	2 (Accountant)	Male	University graduate	2. 18 months.

Methodology

A Modified Grounded Theory (Burnard, 1991) approach was employed in this feasibility study. An exact sample size is difficult to estimate at the beginning of a Grounded Theory project because this can only be determined at the time of theoretical saturation. Theoretical saturation means no new data can add to the emerging theory; thus all avenues of exploration have been exhausted. Theoretical saturation may occur in as little as three or four interviews or as many as twenty to thirty (Creswell, 1998; Morse, 2000). In modified Grounded Theory methodology small samples are employed. The sample size ($n = 5$) was therefore, considered appropriate for a feasibility study (van Teijlingen and Hundley, 2001).

Data were gathered through semi-structured interviews conducted by the first author, who obtained written consent from each participant prior to undertaking interviews. Ethical approval to undertake the study was granted by the Research Ethics Committee (REC) on 23/06/2010. REC reference number: 10/HO703/24. All five interviews occurred in July 2010 in three different general practices in East London. The interviews were digitally recorded and transcribed by the first author. Each interview was conducted face-to-face and lasted between 30 - 40 minutes. The first author asked each of the participants the same questions, which are shown in Table 3 to strive for consistency in the questions posed and to minimise the potential for bias that asking participants different questions could cause.

Table 3 List of questions

<ul style="list-style-type: none"> • What has influenced your decision to immunise your child with the Measles, Mumps and Rubella (MMR) vaccine? • Who has influenced this decision? • Which health professional do you consult for immunisation advice and queries? • Why do you consult this health professional for immunisation advice? • Can you explain if this health professional has influenced your decision on whether to vaccinate your child with the MMR vaccine? • If you had a question about the MMR vaccine, who would you discuss this with and why? • Did you ask the Practice Nurse (PN) about the MMR vaccine? Did this influence your decision on immunizing your child with the MMR vaccine? • If you did not ask the PN information about the MMR vaccine, why did you not approach the PN for information?
--

Results and Discussion

Three themes were synthesised from the data. These are shown in Table 4.

Table 4 Emergent Themes from Synthesised Data

- | |
|--|
| <ul style="list-style-type: none"> • Factors influencing immunisation decision making • Sources of information • Professional role and status |
|--|

Factors influencing immunisation decision making

The MMR controversy has generated considerable academic interest. A number of studies have used surveys, interviews and focus groups to investigate parents' attitudes to, and decisions about, the vaccine (Pareek and Pattison, 2000; Evans et al., 2001; Petts and Niemeyer, 2004; Poltorak et al, 2005; Casiday et al, 2006; Hilton et al, 2007). These studies have generated broadly consistent insights into parental decision making about the MMR vaccine, which has been shown to be influenced by various factors (Skea et al, 2008). In support of these findings, participants in this study listed a range of factors that influenced their decision making. These are shown in Table 5. Interestingly, these findings have been supported in previous research (Evans et al, 2001; Hilton et al, 2007).

Table 5 Factors Influencing Immunisation Decision Making

<p>The prevention of disease The consequences of contracting infectious diseases Perceived pain The media.</p>

Participants also identified pain as a primary factor influencing their decision as to the number of vaccines their children should receive.

"The fact of nine injections to put my son through. Nine injections, she [practice nurse] made that very clear to me... and that was also a big influence. Am I going to put my child through nine injections and put myself through nine injections?" (Ulrika)

"Why let him endure the pain of having three separate injections, when he can have the one altogether and get it over and done with?" (Cheryl)

This finding reflects the findings published in the UK and Ireland highlighting parental distress and concern associated with the administration of childhood vaccines (Harrington et al, 2000; Meyerhoff et al, 2001; Bedford and Lansley, 2007). A cross sectional survey by Bedford and Lansley (2007) revealed that the majority of parents (n = 98/105) considered it less distressing for their child to have fewer injections, as revealed by two of the mothers interviewed in this feasibility study. Likewise, survey results (n = 294/296) revealed that parents had strong preferences for limiting the number of injections at one visit in order to minimise their distress at their children receiving multiple injections (Meyerhoff et al, 2001). In addition, Harrington et al's (2000) in depth interviews of 23 mothers found that many of the mothers interviewed experienced severe emotional distress at the prospect of inflicting the pain of immunisation on their infants. However, in this qualitative study, mothers acknowledged that they perceived the pain being inflicted on their infant as short lived.

As noted in Table 5, the media were cited as an influence on parental decision making:

"It was around the time, that there was the news and the media attention around the fact that Tony Blair's children were going to have the vaccines separately... Because in your mind you think, what's good enough for Tony Blair's son is good enough for my child." (Cheryl)

This participant's experience reveals how this media report made her consider separate vaccines for her son. While this participant did consider the possibility of separate vaccines, the media did not influence her final decision to immunise her son with the trivalent MMR vaccine. Literature has shown that it has been more common for parents to view the media negatively (Hackett, 2008; Hilton, 2007). Trust in the media has been reduced subsequent to sensational headlines leading many parents to view journalists as scaremongers (Hackett, 2008).

The majority of participants rated the importance of the practice nurse in either informing or influencing their decision making.

"Yes...when I come for the immunisation, she [practice nurse] will always tell me how important it is." (Jo)

"I did speak to ... [Practice Nurse]. That's one of the nurses' here. She has informed my decision." (Cheryl)

"Well, I can just say, straight away that there is no other health professional that has given us any talk, other than the practice nurse. She's the one we have only spoken to about it. I didn't think about anyone else." (Mark)

Mark identified a factor that was discrete from other parents' immunisation influencers namely, the effect of vaccine preventable infectious disease. Mark's twin daughters were born with congenital bilateral deafness and attended their local audiology clinic. Some of the children who attended the audiology clinic had acquired deafness due to contracting vaccine preventable diseases such as measles and mumps, which were part of the childhood immunisation programme:

"...You sit around together and each person talks about their child and how they got deafness. I was shocked! None of them have had any of the jabs whatsoever! And all wish they had now. So, if anyone is in any doubt about having these jabs, go and speak to these people. See the horrible results and consequences!" (Mark)

Experiencing an infectious disease such as measles or observing the negative impact of this disease in terms of morbidity and mortality was a strong motivator for parents of completely immunised children (McMurray et al, 2004). Likewise, because of Mark's experience, he was a strong advocate of immunisation.

A factor not identified by parents in previous research on influences to their immunisation decision making was the incidence of measles. This pilot suggests how local outbreaks of measles influenced some parents to immunise their child with the MMR vaccine.

"A couple of years ago, you know, there was an outbreak of measles. People weren't having their kids immunised. I just think it is best to have all their immunisations, rather than just leave it." (Sharon)

"Measles has been quite common in the area lately. So, I thought that rather than me posing the risk of him being able to catch it, I'd rather protect him". (Cheryl)

"The practice nurse was a bit concerned about the poor take up of vaccinations around here. I think it's an epidemic waiting to happen around here. 100%. I think I looked it up and it was 52% take up around here. It was horrendous! So obviously that influenced our decision even more." (Mark)

These parents were aware of local incidences of measles, which helped galvanise their resolve to immunise their children with the MMR vaccine. With the increasing incidence of measles and mumps, practice nurses are key to ensuring that parents are informed about local outbreaks of vaccine preventable diseases and the consequences of not immunising their children.

Sources of information

This theme relates to where participants accessed information and/or how they rated this information. Whilst health professionals have been identified as valuable sources of information (Smailbegovic et al, 2003), this study suggests that the practice nurse is rated as a source of information that participants rely on and actively consult for information advice. None of the participants commented that the information provided by the practice nurse was biased, which has been a finding relating to other health professionals (Evans et al, 2001; Petrovic et al, 2001; Smailbegovic et al, 2003). This finding suggests that the group of practice nurses who were consulted by the participants in this study were unlike the health professionals identified in earlier research as being "unable to give impartial advice" (Casiday et al, 2006: 183).

It is suggested from the findings that participants did not rely on a single source of information. Many relied on multiple sources, such as NHS websites; family members; Parent Fora; health professionals and NHS leaflets. This is a finding supported in previous research (Smailbegovic et al, 2003; Daniels 2002). Some participants commented on the helpfulness of leaflets which informed their decision making. This finding has been supported by findings from the research of Casiday et al (2006) and Gellatly et al (2005).

"I asked the nurse... and I went home and I asked my sister, who is a doctor (general practitioner) ... and they (general practitioners) are my first point and if they are busy then it is the nurse and she would provide me with the information that I need. I also use the internet." (Ulrika)

"I mean the Red Book contains a hell of a lot of information. I have had a lot of time sitting around the hospital for eight weeks and I have read the thing front to back... and also having a quick look on the NHS sites... Wikipedia. I did want to have a look at what the risks were, which is obviously what I had seen in the articles." (Mark)

"It would have been reading the NHS website, probably, I think. That would have been number one. Number two would have been the nurse. Number three, not the nurse or the doctor. Number three would have been the health visitor." (Cheryl)

The narrative of these participants suggests the various sources of information accessed to reach their decision. It suggests that parents seek out the practice nurse as a source of information, which can assist in influencing their decision making.

Professional role and status

Analysis of the narratives revealed how all participants identified professional roles and status of health professionals within the primary care setting. Parents attributed varying levels of knowledge to health professionals such as the general practitioner and practice nurse. An association with greater knowledge was linked to professional status.

"More enlightened...I think, because he is a doctor: he has studied." (Jo)

"A nurse is a nurse, but a doctor is a doctor...I don't know, you just feel more safe. He is a doctor." (Sharon)

Cheryl identified the practice nurse as the most important health professional when seeking immunisation advice and related this to the experience that the practice nurse had in dealing with immunisation matters.

"With kids and vaccinations, it is probably the nurse because they give the vaccinations in the practice. The doctor doesn't do that... It does make the nurse more important for me." (Cheryl)

Considering that one of the many roles that the practice nurse undertakes in general practice is administering the childhood immunisation programme (Crawford, 1997; Yarwood and Bozoky, 1998; Hampson 2002; Maconachie and Lewendon, 2004; Drennan and Goodman 2007) and who additionally has been identified as the principal immuniser in some areas (Maconachie and Lewendon, 2004), this was not a surprising finding.

The majority of participants in this study consulted with the practice nurse. However, Sharon didn't and questioned why she would need to consult with the practice nurse.

"I have never been to see the practice nurse. So I just don't feel why in what way I would want to go and speak to the nurse." (Sharon)

This statement may suggest that Sharon was unaware of the role and the services provided by a practice nurse.

Conversely the advice of the general practitioner and practice nurse were actively sought and valued by other participants as evidenced by:

"It is mainly the general practitioner. And they are my first point and if they are busy then it is the nurse. And she would provide me with the information that I need. Because I trust them; because I know what they are doing and I can ask them anything. And they will give me the honest answer. And that is what I am after." (Ulrika)

"A couple of times we have been to and that is a complete nightmare!....Getting in the car shooting off, whereas here you can just walk over. It is so much more relaxed. She [Practice Nurse] has told us as much as possible. This has saved a lot of hassle for us." (Mark)

Previous research reflects how health professionals (i.e. the general practitioner and the health visitor) have been perceived as trustworthy and how this has been a factor influencing immunisation decision making (Hackett, 2008; Mixer et al, 2007; McMurray et al, 2004; Daniels, 2002). Lack of a trusting relationship with health professionals has been shown to have an adverse effect on immunisation decision making (Austin et al, 2008). It has been interesting to find that in this study, the Health Visitor has not been identified as an influencing factor in decision making. This study's findings reflect the trusting relationship that Ulrika had with her general practitioner and practice nurse and how this guided her decision making. Related to trustworthiness was parental confidence in the ability of the general practitioner and practice nurse to answer queries concerning immunisation. A trusting relationship with a health professional has been cited as having a positive influence on parents immunisation decision making (Mixer et al, 2007). Therefore, the trusting relationship that these participants have with their practice nurses can be viewed as facilitating

decision making regarding childhood immunisation.

Mark chose to consult with the practice nurse because of easier access to immunisation services and because the practice nurse provided the information he needed to make an informed decision regarding the immunisation of his children. The input of the practice nurse as a sought after source of information on immunisation matters and the practice nurse's ability to be attuned to parental concerns both assists and informs parents in terms of influencing parental immunisation decision making.

Limitations

The sample size (n = 5) in this feasibility study may be considered insufficient to achieve saturation. Notwithstanding, it is evident from the participants narrative that the practice nurse plays a significant role in relation to influencing parental decision making associated with uptake of the MMR vaccine. Research has shown that theoretical saturation may occur in as little as three or four interviews or as many as twenty to thirty (Creswell, 1998; Morse, 2000) and for a feasibility study is appropriate (van Teijlingen and Hundley, (2001),

The sample did not include participants whose first language was not English; principally due to the additional cost of using interpreters. This may have elicited different viewpoints to those identified by the five participants, who had fluency in English. Further research will involve participants whose first language is not English and who will require the use of an interpreter to discern their views on factors that influence parental immunisation decision making.

Conclusion

This study has identified a new factor that influence parents' decision making, which have not been determined from previous research. This factor concerns risks associated with local incidences of infectious diseases such as measles from not receiving the MMR vaccine.

The findings from this study suggest practice nurses assist in influencing parents in relation to decision making associated with immunising their children.

There is a need for further research to compare the views of parents who have had their children immunised with the MMR vaccine and those who have not. A wider range of participants in terms of ethnicity and language should be sought to explore whether there are different perceptions regarding the influence of the practice nurse; dependent on ethnicity. At a time when the incidence of measles and mumps are increasing in England and Wales, the role of the practice nurse as the largest group of primary care nurses in promoting the uptake of the MMR vaccine cannot be overlooked.

References:

- Asaria, P and MacMahon, E. (2006). Measles in the United Kingdom: can we eradicate it by 2010? *British Medical Journal*; 333(7574):890 – 895.
- Austin, H. Campion-Smith, C. Thomas, S. and Ward, W. (2008). Parents' difficulties with decisions about childhood immunisation. *Community Practitioner*; 81(10): 32 - 35.
- Baker, J. P. (2003). The pertussis vaccine controversy in Great Britain, 1974 – 1986. *Vaccine*; 21(25-26); 4003 – 4010.
- Bedford, H. and Lansley, M. (2007). More vaccines for children? Parents' views. *Vaccine*; 25(45): 7818 – 7823.
- Burnard, P. (1991). A method of analysing interview transcripts in qualitative research. *Nurse Education Today*; 11(6): 461 - 466.
- Casiday, R. E. Cresswell, T. Wilson, D. and Panter-Brick, C. (2006). A survey of UK parental attitudes to the MMR vaccine and trust in medical authority. *Vaccine*; 24(2): 177 - 184.
- Crawford, M. (1997). Childhood vaccines. *Practice Nurse*; 13(9): 513 – 514, 516.
- Creswell, J. (1998). *Qualitative enquiry and research design: choosing among five traditions*. Thousand Oaks, California: Sage.
- Dales, L. Hammer, S. J. and Smith, N. J. (2001). Time trends in autism and in MMR immunization coverage in California. *Journal of the American Medical Association*; 285(9): 1183 – 1185.
- Daniels, J. (2002). Parents found it difficult to obtain objective information and advice about the combined mumps, measles, and rubella vaccine. *Journal of Evidence Based Nursing*; 5(4): 122.
- Drennan, V. and Goodman, C. (2007). *Oxford Handbook of Primary Care and Community Nursing*. Oxford: Oxford University Press.

- Evans, M. Stoddart, H. and Condon, L. Freeman, E. Grizzell, M. and Mullen, R. (2001). Parents perceptions on the MMR immunization: a focus group study. *British Journal of General Practice*; 51(472): 504 - 510.
- Farrington, C. P. Miller, E. and Taylor, B. (2001) MMR and autism: further evidence against a causal association. *Vaccine*; 19(27): 3632 – 3635.
- Friederichs, V. Cameron, J. C. and Robertson, C. (2006). Impact of adverse publicity on MMR vaccine uptake: a population based analysis of vaccine uptake records for one million children, born 1987 - 2004. *Archives of Diseases in Childhood*; 91(6): 465 - 468.
- Gellatly J. McVittie, C. Tiliopoulos, N. and Hackett, A. J. (2005). Predicting parents' decisions on MMR immunisation: a mixed method investigation. *Family Practice*; 22(6): 658 - 662.
- Griffith, R. (2008). Consent and children: The law for children under sixteen. *British Journal of School Nursing*; 3(6): 281 - 284.
- Hackett, A. J. (2008). Risk, its perception and the media: the MMR controversy. *Community Practitioner*; 81(7): 22 - 25.
- Hampson, G. (2002). *Practice Nurse Handbook*. 4th ed. Oxford: Blackwell Science.
- Harrington, P. M. Woodman, C and Shannon, W. F. (2000). Low immunisation uptake: Is the process the problem? *Journal of Epidemiology and Community Health*; 54(5): 394.
- Hilton, S. Pettigrew, M. and Hunt, K. (2007). Parents' champions vs. vested interests: Who do parents believe about MMR? A qualitative study. *BMC Public Health*; 7: 42. doi:10.1186/1471-2458-7-42.
- Honda, H. Shimizu, J. and Rutter, M. (2005). No effect of MMR withdrawal on the incidence of autism: a total population study. *Journal of Child Psychology and Psychiatry*; 46(6): 572 – 579.
- HPA. (2012a). *Measles notifications (confirmed cases) England and Wales 1995 - 2011* by quarter*. Available from: <http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/Measles/EpidemiologicalData/measDataNotConfirmed/> [Accessed 21/01/2012].
- HPA. (2012b). *Mumps notifications (confirmed cases), England and Wales, 1995 - 2011* by quarter*. Available from: <http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/Mumps/EpidemiologicalData/mumpsDataQuarter/> [Accessed 21/01/2012].
- HPA, (2012c). *Rubella notifications (confirmed cases), England and Wales, 1995 - 2011* by quarter*. Available from: <http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/Rubella/EpidemiologicalData/rubNotconfqtr/> [Accessed 21/01/2012].
- HPA. (2011a). *Measles notifications (confirmed cases) England and Wales 1995 - 2011* by quarter*. Available from: <http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/Measles/EpidemiologicalData/measDataNotConfirmed/> [Accessed 23/10/2011].
- HPA. (2011b). *Mumps notifications (confirmed cases), England and Wales, 1995 - 2011* by quarter*. Available from: <http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/Mumps/EpidemiologicalData/mumpsDataQuarter/> [Accessed 23/10/2011].
- HPA. (2010a). *Vaccination. Immunisation*. Available from: <http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/VaccinationImmunisation/> [Accessed 18/10/2010]
- HPA. (2010b). *Measles notifications (confirmed cases) England and Wales 1995 - 2009* by quarter*. Available from: <http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/Measles/EpidemiologicalData/measDataNotConfirmed/> [Accessed 06/05/2010].
- HPA. (2010c). *Mumps notifications (confirmed cases), England and Wales, 1995 - 2009* by quarter*. Available from: <http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/Mumps/EpidemiologicalData/mumpsDataQuarter/> [Accessed 06/05/2010].
- HPA. (2010d). *Rubella notifications (confirmed cases), England and Wales, 1995 - 2009* by quarter*. Available from: <http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/Rubella/EpidemiologicalData/rubNotconfqtr/> [Accessed 06/05/2010].
- Impicciatore, P. Bosetti, C. Schiavio, S. Pandolfini, C. and Bonati, M. (2000). Mothers as Active Partners in the Prevention of Childhood Diseases: Maternal Factors Related to Immunization Status of Preschool Children in Italy. *Preventive Medicine*; 31(1): 49 - 55.
- Jansen, V. A. A. Stollenwerk, N. Jensen, H. J. Ramsay, M. E. Edmunds, W. J. and Rhodes, C. J. (2003). Measles Outbreaks in a Population with Declining Vaccine Uptake. *Science*; 301(5634): 804.
- Kai, J (1996) Parents difficulties and information needs in coping with acute illness in preschool children: a qualitative study. *British Medical Journal*; 313(7063): 987 - 990.

- Kaye, J. A. del Mar Melero-Montes, M. and Jick, H. (2001). Mumps, measles and rubella vaccine and the incidence of autism recorded by general practitioners: a time trend analysis. *British Medical Journal*; 322(7284): 460 – 463.
- Kassianos, G (2001). *Immunisation: Childhood and Travel Health*. 4th Ed. Oxford: Blackwell Science Limited.
- Lamden, K. H. and Gemmell, I. (2008). General practice factors and MMR vaccine uptake: structure, process and demography. *Journal of Public Health*; 30(3): 251 - 257.
- Maconachie, M. and Lewendon, G. (2004). Immunising children in primary care in the UK - what are the concerns of principal immunisers? *Health Education Journal*; 63(1): 40 - 49.
- MacDonald, H. Henderson, R. and Oates, K. (2004). Low uptake of immunisation: contributing factors. *Community Practitioner*: 77(3): 95 - 100.
- Madsen, K. M. and Vestergaard, M. (2004). MMR vaccination and autism: what is the evidence for a causal association? *Drug Safety*; 27(12): 831 – 840.
- Madsen, K. M. Hvid, A Vestergaard, M Schendel, D. Wohlfahrt, J. Thorsen, P. Olsen, J. and Melbye, M. A. (2002). Population-Based Study of Measles, Mumps, and Rubella Vaccination and Autism. *The New England Journal of Medicine*; 374(19): 1477 - 1482.
- McGreevy, D. (2005). Risks and benefits of the single versus the triple MMR vaccine: how can health professionals reassure parents? *The Journal of The Royal Society for the Promotion of Health*; 125(2): 84 - 86.
- McMurray, R. Cheater, F. M. Weighall, A. Nelson, C. Schweiger, M. and Mukherjee, S. (2004). Managing controversy through consultation: a qualitative. Study of communication and trust around MMR vaccination decisions. *British Journal of General Practice*; 54(504): 520 - 525.
- Meyerhoff, A. S. Weniger, B. G. and Jacobs, J. (2001). Economic value to parents of reducing the pain and emotional distress of childhood vaccine injections. *Pediatric Infectious Diseases Journal*; 20(11): S57 – S62.
- Miller, C. L. (1985) Deaths from measles in England and Wales, 1970–83. *British Medical Journal (Clin Res Ed)*; 290(6466): 443 –444.
- Mixer, R. E. Jamrozik, K. and Newsom, D. (2007). Ethnicity as a correlate of the uptake of the first dose of mumps, measles and rubella vaccine. *Journal of Epidemiology and Community Health*; 61(9): 797 – 801.
- Morse, J. (2000). Determining sample size (Editorial). *Qualitative Health Research*; 10(1): 3 - 5.
- Murch, S. Anthony, A. Casson, D. Malik, M. Berelowitz, M. Dhillon, A. Thomson, P. Valentine, A. Davies, S. and Walker-Smith, J. (2004). Retraction of interpretation. *The Lancet*; 363(9411):750.
- NICE. (2010). *Guidance on differences in the uptake of immunisations (including targeted vaccines) in people younger than 19 years*. Available from: National Institute for Health and Clinical Excellence Public Health Guidance: London. <http://www.nice.org.uk/nicemedia/live/12247/45497/45497.pdf> [Accessed 06/05/2010].
- Pareek, M and Pattison H.M. (2000). The two-dose measles, mumps and rubella (MMR) immunisation schedule: factors affecting maternal intention to vaccinate. *British Journal of General Practice*; 50(461): 969 - 971.
- Petrovic, M. Roberts, R. and Ramsay, M. (2001). Second dose of measles, mumps, and rubella vaccine: questionnaire survey of health professionals. *British Medical Journal*; 322(7278): 82 - 85.
- Petts, J. and Niemeyer, S. (2004). Health risk communication and amplification: learning from the MMR vaccination controversy. *Health, Risk & Society*; 6(1): 7 - 23.
- Poltorak, M. Leach, M. Fairhead, J. Cassell, J. (2005). 'MMR talk' and vaccination choices: An ethnographic study in Brighton. *Social Science & Medicine*; 60(3): 709 - 719.
- Salisbury, D. Ramsay, M. and Noakes, K. (2006). *Immunisation against infectious disease*. 3rd ed. London: DH.
- Skea, Z. C. Entwistle, V. A. Watt, I. and Russell, E. (2008). Avoiding harm to others' considerations in relation to parental measles, mumps and rubella (MMR) vaccination discussions – An analysis of an online chat forum. *Social Science and Medicine*; 67(9): 1382 - 1390.
- Smailbegovic, M. S. Laing, G. J. and Bedford, H. (2003). Why do parents decide against immunization? The effect of health beliefs and health professionals. *Child: Care, Health and Development*; 29(4): 303 - 311.
- Taylor, B. Miller, E. Lingam, R. Andrews, N. Simmons, A. and Stowe, J. (2002). Measles, mumps

and rubella vaccination and bowel problems or developmental regression in children with autism population study. *British Medical Journal*; 324(7334): 393 – 396.

Taylor, B. Miller, E. Farrington, C. P. Petropoulos, M-C. Favot-Mayaud, I. Li, J. and Waight, P A. (1999). Autism and measles, mumps and rubella: no epidemiological evidence for a causal association. *The Lancet*; 53(916): 2026 – 2029.

Van Teijlingen, E. and Hundley, V. (2001). The importance of Pilot Studies. *Social research Update*: 35; 1 – 4. Available from: <http://sru.soc.surrey.ac.uk/SRU35.html> Accessed 29/10/2012

Wakefield, A. J. Murch, S. H. Anthony, A. Linell, J. Casson, D. M. Malik, M. Berelowitz, M. Dhillon, A. P. Thomson, M. A. Harvey, P. Valentine, A. Davies, S. E. and Walker-Smith, J.A. (1998). Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive development disorder in children. *The Lancet*; 351(9103): 637 - 641.

WHO, (2010a). *Immunisation*. Available from: <http://www.who.int/topics/immunization/en/> [Accessed 18/10/2010].

WHO, (2010b). *Measles*. Available from: <http://www.who.int/mediacentre/factsheets/fs286/en/> [Accessed 06/05/2010].

Wilson, T. (2000). Factors influencing the immunization status of children in a rural setting. *Journal of Pediatric Health Care*; 14(3): 117 - 121.

Yarwood, J, Noakes, K. Kennedy, D. Campbell, H. and Salisbury, D. (2005). Tracking mothers' attitudes to childhood immunisation 1991 - 2001. *Vaccines*; 23(48 - 49): 5670 - 5687.

Yarwood, J. and Bozoky, Z. (1998). Prevention through immunisation: take the evidence-based approach. *Practice Nurse*, 16(4): 216, 218 - 219.

PREDICTING NURSES' TURNOVER INTENTIONS BY DEMOGRAPHIC CHARACTERISTICS, PERCEPTION OF HEALTH, QUALITY OF WORK, AND WORK ATTITUDES

Mahmoud Al-Hussami, D.Sc., Ph.D., Epidemiologist

Department Head and Assistant Professor, The University of Jordan

Muhammad Darawad, RN, PhD

Assistant Professor, Faculty of Nursing, University of Jordan, Amman-Jordan.

Ali Saleh, RN, Ph.D

Assistant Professor, Faculty of Nursing, University of Jordan, Amman-Jordan.

Ferial Ahmed Hayajneh, RN, PhD

Associate Professor, The University of Jordan, Faculty of Nursing

Abstract:

Aim: The purpose of this paper is to examine the impact of demographic variables, organizational commitment levels, perception of health, and quality of work on turnover intentions.

Methods: A self-reported cross-sectional survey design was used to collect data from Jordanian registered nurses who were working between June 2011 and November 2011.

Results: the findings showed strong effects of the quality of work, perception of health, and normative organizational commitments on turnover intentions.

Conclusion: This study sheds the light on the important work outcomes in healthcare organizations. Increasing nursing quality of work and normative organizational commitment are good strategies for reducing turnover intentions.

Key Words: Nurses; Turnover Intentions; Work Attitudes

Introduction

Nursing turnover is a major problem that makes management of nursing workforce a challenge for nursing leaders. The ultimate goal of nursing turnover research is to improve the quality of patient care, which is a major concern of healthcare administrators and policy makers. Turnover of healthcare staff negatively influence health care costs¹. It was estimated that minimum cost of healthcare staff turnover loss of more than five percent of the total annual operating budget including hiring, training, and productivity loss¹.

As many other countries, Jordan experienced a national nursing shortage^{2,3}. Al-Maaitah and Shokeh² estimated a projected nursing shortage of female registered nurses (RNs) of 2,572 in 2012, while there will be a surplus of male RNs of 2,026 in the same year. On the other hand, Al-Maaitah and Shokeh² reported the calculated turnover rate from the years of 2003 to 2007 as of 32.1% among nurses with PhD, MSN, BSC, Midwifery, and Associate Degree. The highest nursing turnover of the five degrees was reported for BSN nurses (35.9%). Other researchers⁴ studied 21 Jordanian hospitals to identify RNs turnover rate. Out of the total sample of 2126 RNs 779 nurses (36.6) left their positions during study period.

The situation of nursing shortage in Jordan would be worsening especially if the trend of nursing turnover rate continues as before. So, there is a need to identify factors that affects turnover and turnover intentions. Identifying such factors may help nursing administrators and policy makers to retain nursing staff. Critical review of Jordanian nursing literature revealed a lack of research studies concerning nursing turnover especially in identifying factors that may lead nurses to leave their positions.

Turnover intention seems to be affected by a number of variables. Several studies examined the effect of demographic characteristics such as age, gender, and marital status on nurses' intention

to leave their organizations. Younger nurses were found to have higher level of turnover intention⁵⁻⁹. Other researchers reported higher level of turnover intention among specific age groups such as 25-44¹⁰, and 30-44¹¹. Also, literature suggests different demographic characteristics among nurses who were planning to leave their organizations including male nurses^{5, 11, 12}, single nurses⁵, and highly educated nurses⁹. Mrayyan¹³ examined the predictors of nurses' intent to stay in Jordanian hospitals. Mrayyan¹³ found age to be positively associated with nurses' intent to stay ($P < 0.001$).

For decades, the concept of commitment was the area of research interest for many researchers in the field of organizational behaviors. Morrow¹⁴ described organizational commitment as one of the different forms of work commitment. Other researchers described organizational commitment as a complex and multifaceted construct, and conceptualized it into three components including; affective, continuance, and normative commitment^{15, 16-17}

Meyer, Stanley, Herscovitch, and Topolnytsky¹⁷ conducted a meta-analysis to assess the relationships of the three components of commitment, affective, continuance, and normative with different work related behaviors. Meyer et al. found that the three components of organizational commitment were negatively associated with turnover intentions. In addition, one of the recommendations of the meta-analysis conducted by Meyer et al. is the need to examine the concept of organizational commitment across cultures to get in depth understanding of the concept globally.

Up to our knowledge, there is no literature evidenced the effects of organizational commitment on turnover intentions among Jordanian nurses. Researchers examined the effects of organizational commitment on nurses' turnover intentions in different countries such as United States^{8, 18}, Canada¹⁹, Taiwan^{20, 21}, and Australia²². The consistent negative relationship between organizational commitment and nurses' turnover intention was evidenced through those studies^{8, 18-22}.

A well state of physical and psychological health of nurses is a prerequisite for providing a quality of nursing care. Several studies examined the effect of aspects of employees' physical health on turnover (low back pain and disability) were significant predictors of elderly healthcare professionals' turnover in Denmark²³, and nurses who reported higher levels of musculoskeletal problems of the neck/ shoulder or knees, were more expected to leave nursing in Sweden²⁴. In addition, studies from different countries reported that higher levels of nursing turnover intention was associated with higher level of stress²⁵⁻²⁷, nurse burnout¹¹, and emotional exhaustion²⁸. In Jordan, researchers examined the effects of job stress on nurses' intent to stay at work in 206 nurses²⁹. Abualrub and Al-Zaru²⁹ found that job stress was negatively associated with nurses' intent to stay.

Underpayment of nurses is a major cause of Jordanian nurses' migration especially to Arab Gulf region. This was evidenced by the continuous claims by Jordanian Nursing Council to improve nursing work conditions and to increase nurses' salary. In her discussion regarding nursing shortage in Jordan, AbuAlRub³ emphasized that one of the causes of moving students away from nursing is the lower social status of nursing evidenced by low salary. Internationally, inconsistent results were found regarding the relationship between pay and nursing turnover intentions. Several studies reported direct or indirect negative effects of pay on nurses' intention in different countries such as China¹⁰, UK^{12, 30}, and Canada¹⁹. In contrast, other research studies highlighted that pay had not a significant relationship with nursing turnover intention in Singapore²⁵ and Australia²⁶.

Nursing quality of work is a complex phenomenon³¹. Researchers found quality of work to be a factor that can helps in decreasing nursing turnover intention^{6, 22, 27, 32-36}. Other researchers reported significant negative relationships between nursing turnover intention and different areas of nurses' quality of work such as satisfaction with control and responsibility, scheduling, extrinsic rewards⁵, supervision⁹, and workload, extent to which the nurses liked to work, and colleagues¹⁸

Reviewing Jordanian nursing literature revealed a dearth of literature regarding the relationship between satisfaction of work related issues and turnover intention. Abu AlRub, Omari, and Al-Zaru³⁷ examined the relationships between social support, quality of work and intent to stay among Jordanian nurses. The results revealed that nurses with higher level of quality of work, reported higher levels of intent to stay at work. On other hand, Al-Ma'aitah, et al.³⁸ found that negative predictors of turnover intentions among female nurses regarding satisfaction work related issues were kind of work they did, physical work conditions, and career future, while negative predictors of turnover intentions among male nurses were satisfaction with hospital identification and career future.

Purpose

The purpose of this paper is to examine the impact of demographic variables, organizational commitment levels, perception of health, and quality of work on nursing turnover intentions.

Research Question

What are the multiple correlations between a set of five predictors (age, quality of work, perception of health, organizational commitment, and pay) and the outcome, the nurses' turnover intentions?

Study Design

A self-reported cross-sectional survey design was used to collect data from Jordanian registered nurses who were working between June 2011 and November 2011. The use of self-reported questionnaire eliminates the effect of a researcher on participants which allows them more freedom to answer the questions honestly and openly³⁹.

Setting and Population

Jordan healthcare system is divided into governmental, military, university, and private institutions. In the governmental sector, the Ministry of Health operates 27 hospitals, accounting for 37% of all hospital beds; the Military Royal Medical Services run 11 hospitals, providing 24% of all beds; the University Hospitals account for 3% of total beds; and the Private Sector provides 36% of all hospital beds, distributed among 60 hospitals. Only those hospitals with a capacity of 300 beds and above and having medical, surgical, emergency room, and critical care units were approached. Therefore, 11 hospitals including; 6 governmental, 2 university, and 2 private were eligible settings for the current study.

Sample

Participants were randomly selected from eight hospitals in three clusters of Jordanian hospitals that are stratified as governmental, university, and private hospitals using simple random sampling technique. Two hundred and thirteen registered nurses (RN) accepted to participate out of three hundred RNs invited in the study. Participants were recruited from the population of RNs' who met the eligibility criteria. The eligible subjects were RNs from both genders who have acquired a Bachelor, or Master's degree in nursing with at least one year of experience in acute healthcare settings.

Power Calculation

The statistical software G*Power V.3⁴⁰ showed that the required sample size was 159 nurses. This figure was arrived at by using compromised $\beta = 0.80$, $\alpha = 0.05$ (2-tailed) and effect size = 0.3 (medium effect). Although these figures were needed, more numbers included (213) to produce significant and reliable findings and to compensate for incomplete questionnaires.

Instrument

A self-administered questionnaire was used in English language to collect the data about Jordanian nurses' turnover intention. Questionnaire of 57 items was subjected to validation process by researchers and expert nurses (n=15) that assessed the level of comprehensiveness, clarity, avoidance of ambiguity, and content validity. This involved circulating the draft items until there was consensus on content, order, and wording. As a result, four items were modified as not properly understood by three evaluators.

A pilot study was then conducted using this questionnaire among a sample of 20 nurses after an access to nurses was sought from the director of nursing in university hospitals. Fifteen completed questionnaires were received. Some items were re-worded to add more clarity and then the questionnaire was revised to combine similar items and to remove misleading or repeated items. Thus, the questionnaire was produced whose content validity was assessed by expert panel consists of four expert nurses who are nurses managers and having 5 years of experience in nursing; two PhD holders who have published work on management and leadership. The reliability of the final questionnaire was assessed using internal consistency (Cronbach's alpha test) ($\alpha = 0.87, 0.76, 0.74, 0.90, 82$, respectively for organizational commitment, perception of health, perception of feeling regarding pay, quality of work, intention to leave the organization). The nurses who have been involved in the pilot study had reported no corrections with the wording, length, and format of the questionnaire and they were not included as part of the main study.

The questionnaire was divided into six parts: Part one included professional and situational related characteristics such as; gender, age, marital status, years of experience, yearly income, and

working area. Age, years of experience, and yearly incomes were measured as ratio variables. However, gender, marital status, and working area were measured as dichotomous variables. Part two measured organizational commitment by a 23 item index called Organizational Commitment Questionnaire (OCQ) developed by Meyer, Allen, and Smith ⁴¹ with an estimated Cronbach's alpha .85 ⁴². The participants were asked to indicate their agreement on a seven Likert scale from strongly disagree to strongly agree. Part three measured nurses' perception of health by two items taken from the health related items used by Dalton and Mesch ⁴³. These items are: "The job I have now probably affect my physical health." "The job I have now probably affects my mental health." The scales ranged from 1 = very badly, to 5 = very positively.

Part four measured perception of feeling regarding pay by two items from Eisenberger et al.'s ⁴⁴ Survey of Perceived Organizational Support" (SPOS) scale. This 36-item instrument was developed to measure employees' perceptions of organizational support. The two-selected items are specifically designed to explore employees' perceptions of their feelings regarding fairness in pay and measured on a seven-point Likert scale (1 = strongly disagree, to 7 = strongly agree).

Part five, the quality of work was measured by a 20 item index called Minnesota Satisfaction Questionnaire (MSQ) short-form, developed by Weiss et al. ⁴⁵ with an estimated Cronbach's alpha .91 ⁴⁶⁻⁴⁹. The MSQ, a self-reported instrument consists of 20 items that sample job satisfaction on 20 scale areas, is an often used and widely researched job satisfaction measure. It was derived from the Minnesota Studies in Vocational Rehabilitation and measured on a on a seven-point Likert scale (1 = strongly dissatisfied, to 7 = strongly satisfied). Part six, the dependent variable, intention to leave the organization was measured by three items following Mobley et al. ⁵⁰ definition. The respondents were asked to indicate their agreement with the following three items on a seven-point scale: "I think a lot about leaving the organizations." "I am actively searching for an alternative to this organization." "As soon as it is possible, I will leave the organizations."

Ethical considerations

Ethical approval was sought and granted from the Research and Ethics Committee at Faculty of Nursing/ University of Jordan and the research and Ethics Committee at each hospital involved in the study. Furthermore, detailed information about the objectives of the study was contained through the questionnaire cover letter, and returning the questionnaire was considered an implied consent. Participants were instructed that participation is voluntary and information provided will be kept anonymous, that is, no names or other identifiers will be collected on any of the instruments used. Data will be kept in the researchers' office for five years under lock and key. After this period of time, all data will be shredded.

Data collection methods

A detailed explanation of the aims and procedure of the study was given to the nurse administrators, head nurses, and charge nurses at participating hospitals. A list of an estimated number of available nurses was prepared from the selected hospitals one day before hospital visit. At the time of data collection, questionnaires were distributed and handed to nurses by the researchers and by assistance of the departments' managers and the charge nurses at all shifts. Each questionnaire had a cover letter explaining the nature of the study, aims, the way of completion, and a return envelop. Self completed questionnaires were then handed over together in a large envelope to the researchers.

Data analysis

Based on a 57-item questionnaire, nurses' responses were summed up in total scores of the organizational commitment, perception to health, perceptions of feelings regarding pay, quality of work, and turnover intention, where then calculated in mean scores and standard deviations. Furthermore, correlation between items was measured using Pearson's test. Additionally, hierarchical multiple regression analysis was used to estimate the probability of recorded variables. All statistical procedures was performed using the Statistical Package for the Social Sciences (SPSS-17) and produced at $\alpha=0.05$ significance level (2-tailed).

Results

Of the original sample (300), 213 participants returned the questionnaires giving a response rate of 71%. The study population ranged from 22 to 52 years old, 57.7% females (n = 123) and 42.3% males (n = 90). For the purpose of describing the years of nursing experience for the study sample it was recoded from continuous into categorical variable and varied from less than two years (25%), between 2-6 years (50%), and more than 6 years (25%). The majority of the study sample

(82%) was employed at the baccalaureate level and 18% were employed at the masteral level. Among the respondents, 53.1% (n = 113) were single, 45.5% (n = 97) were married, and 1.4% (n = 3) were divorced.

Table 1 presents correlations of the study variables. Correlations between turnover intention and normative commitment and quality of work were significantly negative ($r = -0.200, p \leq .01$; $r = -0.193, p \leq .01$, respectively). However, the correlations between turnover intention and age, experience, income, affective commitment, continuous commitment, income, and pay were not statistically significant. On the other hand, the perception of pay and age, experience, income, affective commitment, continuous commitment, normative commitment, quality of work, and perception of health was all significantly positive (table 1). Moreover, the quality of work and the levels of organizational commitment were strongly significantly positive.

Table 1. Pearson Correlations of Turnover Intention (N= 213)

	Age	Exp	Inco	Affect	Cont	Norm	Qua	Heal	Pay	Turn
Age	1									
Experience	.92**	1								
Income	.69**	.77**	1							
Affective	.45**	.46**	.38**	1						
Continuous	.50**	.50**	.40**	.59**	1					
Normative	.44**	.46**	.40**	.63**	.69**	1				
Quality of Work	.44**	.43**	.31**	.50**	.47**	.62**	1			
Perception of Health	.16*	.14*	.06	.27**	.17*	.27**	.28**	1		
Perception of Pay	.23**	.22**	.14*	.23**	.33**	.48**	.41**	.28**	1	
Turnover Intention	-.05	-.06	-.01	-.03	-.10	-.20**	-.19**	-.134	-.11	1

* Correlation is significant at $\alpha=0.05$ (2-tailed), ** Correlation is significant at $\alpha=0.01$ (2-tailed)

The demographic variables (income, age, and years of experience) were entered in the first model, and R^2 change was 0.01, which was not significant ($p = 0.570$). The addition of the affective commitment and continuous commitment did not make any significant changes ($R^2= 0.01, P= 0.99, R^2= 0.017, P= 0.41$, respectively). In the fourth model, the addition of the normative commitment added 4% to the variance. Also, the single variables in model five and model seven added 4% and 0.2%, respectively, to the variance. Thus, the only statistically significant variables were normative commitment, perception of health, and quality of work (Table 2).

Table 2 Seven-step multiple hierarchal regression analysis of predictors of turnover

	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
Model 7 Variables	β	p-value	β	p-value	β	p-value	β	p-value	β	p-value	β	p-value
Age	.067	.710	.067	.712	.085	.641	.068	.701	.044	.803	.048	.786
.623												
Experience	-.213	.298	-.213	.302	-.195	.344	-.166	.413	-.166	.803	-.167	.404
.623												
Income	.106	.333	.106	.335	.107	.559	.121	.262	.123	.162	.147	.171
.237												
Affective Commitment			.000	.999	.052	.211	.159	.093	.048	.192	.050	.220
.117												
Continues Commitment					-.115	.211	.029	.778	.048	.632	.050	.623
.591												

Normative Commitment									
.019									
Perception of Health									
.003									
Perception of Pay									
.919									
Quality of Work									
.031									
R ²	.010 <0.570	.010 <0.999	.017 <0.211	.059 <0.003	.093 <0.006	.094	<0.648		
.111 <0.31									
Adjusted R ²									
.075									
R ² Change									
0.02	0.00	.00	.01	.04	.03		.00		

Discussion

It was found in the literature that younger nurses had higher level of turnover intention⁵⁻⁹. On the other hand, Mrayyan¹³ found age to be positively associated with nurses' intent to stay. The findings of this study were inconsistent with previous research reports. It was expected that nurses throughout their years should have higher salaries and promotions that prevent them to leave their jobs. In contrast, the present study found that age had no effects on turnover intention and this could be related to study sample. The study sample did not include older nurses (22-52 years old) compared to previous studies. For instance nurses' age ranged from 20-65 years old in Simon et al. study⁶, and 22-61 years old in Delobelle et al. study⁹.

The present study showed that the quality of work is the clearest and most consistent determinant of turnover intention. The findings here thus support the notion that quality of work plays a critical role in the employee's decision to leave the organization⁵¹. Although nurses' perceptions of health was not found to be associated with turnover intention in this study, it showed a predictive power to turnover intention when it was entered in regression analysis with other independent variables. Study findings showed the positive perception of health to be a predictor of increased level of turnover intention. One explanation might be that when nurses perceive their health positively, they will be more able to obtain better job offers and leave their organizations.

Commitment levels are considered the main predictor of turnover intention. Sage⁵² found in that regard organizational commitment was strongly related to the aggregated duration of voluntary absence. According to the literature^{53, 54} the work-related commitments can increase performance, reduce turnover, and benefit both the employee and the organization. Among the three components of commitment, findings showed that normative commitment was the only negative predictor of nurses' turnover intention. Since normative commitment reflects an individual's feeling of responsibility to remain in the organization¹⁶, this is considered a logical finding. In this study, normative commitment had the least predictive power of turnover intention among the significant predictors of turnover intentions. Despite that organizational commitment is considered an important factor related to turnover intention^{8, 17-22}, the findings of the present study did not show a relationship between affective neither continuous commitment and turnover intention.

Pay was found not a significant predictor of turnover intention evidenced by study data. The relationship between pay and turnover intention is not apparent and that is supported by inconsistent findings of previous literature^{10, 12, 19, 25, 26, 30}. Based on the findings of this study, it seems that pay is not a motivator for Jordanian nurses to leave their organizations.

A few limitations of this study should be noted. First, all variables in this study were measured with self-reports, thus the problem of common method was unable to recall events happened in the past. Concerning the condition of nurses during the data collection procedure, nurses were found to be very busy during their shifts, and the data collectors reported the need to visit the same unit many times to find nurses who are able to get 20 minutes to answer the questions of the study instruments.

It is true that this study is not the first study to address turnover intentions. Though, it is the first to address the relationship of work attitudes with turnover intention in Jordan. The gained

knowledge of this study is more relevant to nursing because the study asked nurses themselves to state their perception from different views. Also, this study asked the nurses to report their turnover intentions, unlike most of the nursing turnover studies that only report the official turnover rates. Doing so should give nursing administrators an insight about nurses' future decisions, and respite them sometime to take steps before nurses' turnover intention develops to be actions toward leaving nursing.

This study came to fill out a significant gap in the nursing literature regarding the nurses' turnover intention particularly in countries other than the western countries. Measuring turnover intentions among the Jordanian nurses was identified as another gap in the nursing literature that needed to be filled out. Also, this study filled out the gap and evaluated the turnover intentions among Jordanian nurses in the three healthcare sectors.

To expand the current findings, there are some possible moderators for future research. For example, Lee and colleagues'⁵⁵ meta-analysis found negative relationship between commitment levels and turnover intention. Employees are more likely to leave their work because of lower professional identity or family support. Furthermore, it is recommended to replicate this study among nurses in Jordan and other developing countries but with a larger sample size.

Implications for Practice

Nursing administrators are highly required to set the policies that are capable to accomplish nurses' desires of more contribution, and particularly when it comes to their quality of work and work attitudes. Setting job motivations seemed to be the most important step for the nursing administrators to start with as many nurses may leave their organizations. Healthcare policy makers, particularly in Jordan, need to know that without immediate actions, more nurses will be leaving the profession, and the previous turnover rates will be dangerously replaced. More respect to nurses' desires of contributing to the work attitudes and quality of work, more incentives and other retaining strategies, and obtaining a periodic feedback from the nurses are only examples of those immediate actions that are required with necessity.

Conclusion

The findings of this study shed lights on the important work outcomes in healthcare organizations. Increasing nursing quality of work and organizational commitment are good strategies for reducing turnover intentions. This paper illustrated the usefulness of such research as well as proposing directions for future work.

Acknowledgements

The authors express their appreciation to all registered nurses who participated in this study. Also thanks extended to the University of Jordan for funding this research.

References:

- Waldman JD, Kelly F, Arora S, Smith HL. The shocking cost of turnover in health care. *Health Care Management Review*. 2004;29(1):2-7.
- Al-Maaitah R, Shokeh D. *The nursing workforce in Jordan: A policy oriented approach*. Amman: The Jordanian Nursing Council;2009.
- AbuAlRub RF. Nursing shortage in Jordan: What is the solution? *Journal of Professional Nursing*. 2007;23(2):117-120.
- Hayajneh YA, AbuAlRub RF, Athamneh AZ, Almahzoomy IK. Turnover rate among registered nurses in Jordanian hospitals: An exploratory study. *International Journal of Nursing Practice*. 2009;15(4):303-310.
- El-Jardali F, Dimassi H, Dumit N, Jamal D, Mouro G. A national cross-sectional study on nurses' intent to leave and job satisfaction in Lebanon: Implications for policy and practice. *BMC Nursing*. 2009;8(3). <http://www.biomedcentral.com/1472-6955/8/3>. Accessed June 15, 2011.
- Simon M, Muller BH, Hasselhorn HM. Leaving the organization or the profession - A multilevel analysis of nurses' intentions. *Journal of Advanced Nursing*. 2010;66(3):616-626.
- Kuokkanen L, Leino-Kilpi H, Katajisto J. Nurse empowerment, job-related satisfaction, and organizational commitment. *Journal of Nursing Care Quality*. 2003;18(3):184-192.
- Ingersoll GL, Olsan T, Drew-Cates J, DeVinney BC, Davies J. Nurses' job satisfaction, organizational commitment, and career intent. *Journal of Nursing Administration*. 2002;32(5):250-263.

- Delobelle P, Rawlinson JL, Ntuli S, Malatsi I, Decock R, Depoorter AM. Job satisfaction and turnover intent of primary healthcare nurses in rural south africa: A questionnaire survey. *Journal of Advanced Nursing*. 2011;67(2):371-383.
- Moon Fai C, Luk AL, Sok Man L, Siu Ming Y, Iat Kio V. Factors influencing Macao nurses' intention to leave current employment. *Journal of Clinical Nursing*. 2009;18(6):893-901.
- Estryn-Behar M, Bij, Oginska H, et al. The impact of social work environment, teamwork characteristics, burnout, and personal factors upon intent to leave among European nurses. *Medical Care*. 2007;45(10):939-950.
- Barron D, West E. Leaving nursing: An event-history analysis of nurses' careers. *Journal of Health Services Research & Policy*. 2005;10(3):150-157.
- Mrayyan MT. Predictors of hospitals' organizational climates and nurses' intent to stay in Jordanian hospitals. *Journal of Research in Nursing*. 2008;13(3):220-233.
- Morrow PC. Concept redundancy in organizational research: The case of work commitment. *Academy of Management Review*. 1983;8(3):486-500.
- Meyer JP, Allen NJ. A three-component conceptualization of organizational commitment. *Human Resource Management Review*. 1991;1(1):61-89.
- Allen NJ, Meyer JP. The measurement and antecedents of affective, continuance and normative commitment to the organization. *Journal of Occupational Psychology*. 1990;63(1):1-18.
- Meyer JP, Stanley DJ, Herscovitch L, Topolnysky L. Affective, continuance, and normative commitment to the organization: A meta-analysis of antecedents, correlates, and consequences. *Journal of Vocational Behavior*. 2002;61(1):20-52.
- Lynn MR, Redman RW. Faces of the nursing shortage: influences on staff nurses' intentions to leave their positions or nursing. *The Journal of Nursing Administration*. 2005;35(5):264-270.
- Lum L, Kervin J, Clark K, Reid F, Sirola W. Explaining nursing turnover intent: Job satisfaction, pay satisfaction or organizational. *Journal of Organizational Behavior*. 1998;19(3):305-320.
- Chang C, Du P, Huang I. Nurses' perceptions of severe acute respiratory syndrome: Relationship between commitment and intention to leave nursing. *Journal of Advanced Nursing*. 2006;54(2):171-179.
- Liou S, Cheng C. Organisational climate, organisational commitment and intention to leave amongst hospital nurses in Taiwan. *Journal of Clinical Nursing*. 2010;19(11-12):1635-1644.
- Parry J. Intention to leave the profession: Antecedents and role in nurse turnover. *Journal of Advanced Nursing*. 2008;64(2):157-167.
- Faber A, Giver H, Strøyer J, Hannerz H. Are low back pain and low physical capacity risk indicators for dropout among recently qualified eldercare workers? A follow-up study. *Scandinavian Journal of Public Health*. 2010;38(8):810-816.
- Fochsen G, Josephson M, Hagberg M, Toomingas A, Lagerström M. Predictors of leaving nursing care: A longitudinal study among Swedish nursing personnel. *Occupational and Environmental Medicine*. 2006;63(3):198-201.
- Fang Y. Turnover propensity and its causes among Singapore nurses: An empirical study. *International Journal of Human Resource Management*. 2001;12(5):859-871.
- Cowin L. The effects of nurses' job satisfaction on retention: An Australian perspective. *Journal of Nursing Administration*. 2002;32(5):283-291.
- Applebaum D, Fowler S, Fiedler N, Osinubi O, Robson M. The impact of environmental factors on nursing stress, job satisfaction, and turnover intention. *Journal of Nursing Administration*. 2010;40(7-8):323-328.
- Spence Laschinger HK, Leiter M, Day A, Gilin D. Workplace empowerment, incivility, and burnout: Impact on staff nurse recruitment and retention outcomes. *Journal of Nursing Management*. 2009;17(3):302-311.
- AbuAlRub RF, Al-Zaru IM. Job stress, recognition, job performance and intention to stay at work among Jordanian hospital nurses. *Journal of Nursing Management*. 2008;16(3):227-236.
- Collins K, Jones ML, McDonnell A, Read S, Jones R, Cameron A. Do new roles contribute to job satisfaction and retention of staff in nursing and professions allied to medicine? *Journal of Nursing Management*. 2000;8(1):3-12.
- American Nurses Association. *Nursing quality indicators: Definitions and implications*. Washington, DC: American Nurses Publishing; 1996.

- McCarthy G, Tyrrell MP, Lehane E. Intention to 'leave' or 'stay' in nursing. *Journal of Nursing Management*. 2007;15(3):248-255.
- Gregory DM, Way CY, LeFort S, Barrett BJ, Parfrey PS. Predictors of registered nurses' organizational commitment and intent to stay. *Health Care Management Review*. 2007;32(2):119-127.
- Han GH, Jekel M. The mediating role of job satisfaction between leader-member exchange and turnover intentions. *Journal of Nursing Management*. 2011;19(1):41-49.
- de Moura GR, Abrams D, Retter C, Gunnarsdottir S, Ando K. Identification as an organizational anchor: How identification and job satisfaction combine to predict turnover intention. *European Journal of Social Psychology*. 2009;39(4):540-557.
- Tsai Y, Wu S. The relationships between organisational citizenship behaviour, job satisfaction and turnover intention. *Journal of Clinical Nursing*. 2010;19(23/24):3564-3574.
- AbuAlRub RF, Omari FH, Al-Zaru IM. Support, satisfaction and retention among Jordanian nurses in private and public hospitals. *International Nursing Review*. 2009;56(3):326-332.
- Al-Ma'aitah R, Cameron S, Horsburgh ME, Armstrong-Stassen M. Predictors of job satisfaction, turnover, and burnout in female and male Jordanian nurses. *The Canadian Journal of Nursing Research*. 1999;31(3):15-30.
- Polit D, Beck C. *Nursing Research: Principles and Methods*. 8th ed. Philadelphia: Lippincott Williams and Wilkins; 2008.
- Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*. 2009;41(4):1149-1160.
- Meyer JP, Allen NJ, Smith CA. Commitment to organizations and occupations: Extension and test of a three-component conceptualization. *Journal of Applied Psychology*. 1993;78(4):538-551.
- Feather NT, Rauter KA. Organizational citizenship behaviours in relation to job status, job insecurity, organizational commitment and identification, job satisfaction and work values. *Journal of Occupational & Organizational Psychology*. 2004;77(1):81-94.
- Dalton DR, Mesch DJ. On the extent and reduction of avoidable absenteeism: An assessment of absence policy provisions. *Journal of Applied Psychology*. 1991;76(6):810-817.
- Eisenberger R, Huntington R, Hutchison S, Sowa D. Perceived organizational support. *Journal of Applied Psychology*. 1986;71(3):500-507.
- Weiss DJ, Dawis RV, England GW, Lofquist LH. *Manual for the Minnesota Satisfaction Questionnaire*. Minneapolis: Minnesota Studies in Vocational Rehabilitation, University of Minnesota Industrial Center; 1967.
- Ben-Bakr KA, al-Shammari IS. Organizational commitment, satisfaction, and turnover in Saudi Organizations: A predictive study. *Journal of Socio-Economics*. 1994;23(4):449-456.
- DeMato DS, Curcio CC. Job satisfaction of elementary school counselors: A new look. *Professional School Counseling*. 2004;7(4):236-245.
- Hancer M. Job satisfaction of restaurant employees: An empirical investigation using the Minnesota Satisfaction Questionnaire. *Journal of Hospitality and Tourism Research*. 2003;27(1):85-100.
- Hirschfeld RR. Does revising the intrinsic and extrinsic subscales of the Minnesota Satisfaction Questionnaire short form make a difference? *Educational & Psychological Measurement*. 2000;60(2):255-270.
- Mobley W, Griffeth R, Hand H, Meglino B. Review and conceptual analysis of the employee turnover process. *Psychological Bulletin*. 1979;86:493-522.
- Spector P. *Job satisfaction: application, assessment, causes, and consequences*. Thousand Oakes, CA: Sage; 1997.
- Sage A. Employee absenteeism, organizational commitment, and job satisfaction: Another look. *Journal of Vocational Behavior*. 1998;52(2):156-171.
- Brewer AM, Lok P. Managerial strategy and nursing commitment in Australian hospitals. *Journal of Advanced Nursing*. 1995;21(4):789-799.
- Reilly NP, Orsak CL. A career stage analysis of career and organizational commitment in nursing. *Journal of Vocational Behavior*. 1991;39(3):311-330.
- Lee K, Carswell JJ, Allen NJ. A meta-analytic review of occupational commitment: Relations with person- and work-related variables. *The Journal of Applied Psychology*. 2000;85(5):799-811.

THE GENDERED KNOWLEDGE OF HIV/AIDS AMONG URBAN UNIVERSITY STUDENTS IN SOUTHERN ETHIOPIA

*Marianne Paul, RN, BScN, MScN,
Basanti Majumdar, PhD,
Noori Akhtar-Danesh, PhD,
Sheryl Boblin PhD
Diana Sherifali, PhD*
McMaster University, Canada

Abstract:

The HIV/AIDS epidemic continues to claim millions of lives worldwide. Africa alone represents the majority of HIV/AIDS cases, where the young aged 16-30, are at greatest risk. The research suggests a gender dichotomy in HIV/AIDS knowledge levels between male and female students in Ethiopia. The purpose of this research is to explore the gendered knowledge of HIV/AIDS among a group of Ethiopian male and female post-secondary students. An observational study was conducted among 126 female and 101 male college and university students. The HIV Knowledge Questionnaire was used to measure participants' knowledge. The statistical t-test demonstrated no difference between male and female HIV knowledge scores ($t=0.6$, $df=225$, $p=0.4$). The study findings demonstrate the call to address the knowledge gap in HIV sexual transmission among a vulnerable population. Future research is needed such that HIV/AIDS awareness campaigns can be specifically tailored to the student population within the Ethiopian context.

Key Words: HIV/AIDS, Ethiopia, students, gender, and knowledge

Background: HIV/AIDS and Ethiopian Students

The Ethiopian government recently stated the urgency in addressing HIV/AIDS as it targets Ethiopia's young working force (Lifson et al., 2012). Urban youths are increasingly susceptible to HIV transmission due to earlier sexual onset, an increase in multiple sexual partners and decreased condom use (Buseh, Glass, McElmurry, Mkhabela, & Sukati, 2001). A study conducted in Eastern Africa reported younger adults were 60% less likely to use a condom in comparison to their older counterparts (Hladik, Shabbir, Jelaludin, Woldu, Tsehaynesh, & Tadesse, 2006). The correlation between high-risk behavior and HIV/AIDS knowledge are strongly related in the literature (De Visser & Smith, 2001; Taffa, Klepp, Sundby, & Bjune, 2002; Ukwuani, Tsui & Suchindran, 2003). Recent studies among Ethiopian high school and university students highlight decreased usage of condoms and the lack of HIV transmission knowledge (Alene, Wheeler, & Grosskurth, 2004; Yerdaw, Nedi, & Enquoselassie, 2002). A cross-sectional study among high school students revealed that only 41% of female students and 44% of male students were educated on the various modes of HIV transmission (Alene et al., 2004). Correspondingly, the national survey of Ethiopia reported a similar trend, whereby 28.7% of young men and only 15.8% of young women were knowledgeable on the varied modes of horizontal HIV transmission (HIV/AIDS Prevention Control Office [HAPCO], 2008).

Several factors such as lower literacy rates, contextual gender roles, and lower school enrollment have been linked to the lower HIV/AIDS knowledge among young women in relation to young men in Ethiopia (UNICEF 2006; Ukwunai et al., 2003). Researchers have attributed differing HIV/AIDS knowledge levels due to delineated gender roles between men and women in the societal roles of labor, power and relationships (Berhane et al., 2001; Hadley et al., 2007; Ukwunai et al., 2003). Within Ethiopia there are clearly defined gender roles, which socially separate men and women in the division of labor and society. For example, in the rural regions of Ethiopia, men are typically responsible for raising livestock and trading agriculture commodities, while women stay local and tend to the domestic duties (Berhane et al., 2001). As such, it is customary for men to assume leadership roles socially and commercially, while women assume more subservient domestic

duties and often play a minimal role in major decision-making (Hadley, Lindstrom, Tessema, & Belachew, 2007). Hadley et al. (2007) reinforce this statement by demonstrating an increased school enrolment rate among male students opposed to females. Over time, these trends have resulted in a 50% literacy rate among males in comparison to only a 26% literacy rate among females (World Health Organization [WHO], 2009).

Due to societal norms young girls in Ethiopia are more susceptible to HIV/AIDS than boys due in part to early onset of sexual debut, early marriage customs and sexual abuse (Berhane et al. (2001). A cross-sectional survey reported older men marrying young girls in efforts to secure a virgin bride (Molla, Berhane, & Lindtjorn, 2008). Molla et al. also found men were twice as likely to have multiple sexual encounters in comparison to their female counterparts.

Theoretical Framework

The Social Theory of Gender and Power helps us understand the relationships of labor, power and cathexis and their influences on gender. According to Connell (1987) the main assumptions of this theory consider gender to be socially constructed. The theory comprises three interrelated dimensions; labor; power; and; cathexis. All three interact to further elucidate gender relations in any institutional or societal context (Maharaj, 1995). According to Wingood and DiClemente (2000) the three theoretical dimensions are seen as both overlapping and independent from each other. The theoretical dimension of labor considers the dichotomous relationship between employment practices of men and women within particular societies. The second dimension of power illustrates the gendered differences regarding authority and control, which can arise at a societal and institutional level. Lastly, the dimension of cathexis explains the affective and sexual relationship between men and women and the social norms that surround them (Connell, 1987). The Social Theory of Gender and Power was utilized to elucidate the results of the study and to further understand the students' gendered knowledge towards HIV/AIDS.

The goal of this study was to identify the levels of knowledge towards HIV/AIDS between young male and female Ethiopian students. Based on the literature, the main hypothesis generated states a significant difference between the mean scores of HIV knowledge between the male and female students.

Methods

An observational study design was used in order to measure the knowledge towards HIV transmission among male and female post-secondary students. Ethics approval was obtained from both Canadian and Ethiopian institutions. Consent forms were translated in Amharic (Ethiopia's official language). Language experts in both English and Amharic validated the content of the consent form. The consent forms outlined the complete anonymity of the questionnaires and the right of every participant to withdraw at anytime during the study.

The study took place in a college and university setting in a southern city in Ethiopia. All participants provided a written and informed consent. The sampling method was based on non-probabilistic sampling (Levin, 2006). Student outreach volunteers were chosen to help recruit participants to take part in this study. All participants and volunteers received a gift of remuneration for their time. The significance level of the independent t-test result was set at the 0.05 level. Data were analyzed using SPSS version 17 statistical software.

Data Collection Process and Instruments

In order to mitigate or diffuse potential power imbalances between students and professors, the study procedures and questionnaires were described in the absence of professors. This procedure facilitated students to leave the classroom at any point in time without the presence of school authority figures (Visser, 2008). Due to the sexual content and sensitive nature of HIV/AIDS, two separate classrooms were provided for the male and female students to fill out the questionnaires.

The HIV knowledge questionnaire (HIV-KQ-18) formulated by Carey and Schroder (2002) was utilized to measure knowledge of HIV transmission. It is an 18-item, self-administered questionnaire, which measures the general knowledge of transmission, prevention and treatment of HIV/AIDS. Respondents were asked to indicate whether the statements were true or false. See Table 1 for the questionnaire and correct answers. The HIV-KQ-18 is internally consistent ($\alpha=0.91$) and a achieved satisfactory test-retest reliability among a pilot control group after one week ($r=0.83$) (Carey

& Schroder, 2002). All questionnaires were translated from English to Amharic. The questionnaires were translated back to English in order to ensure consistency from the original English version of the questionnaire (Sobel & Kugler 2007). The study questionnaires were administered to one male volunteer and one female volunteer to ensure face validity (Haynes, Richard & Kubany, 1995). All parties agreed on the scales' appropriate comprehension and proper use of the Amharic language.

Results

A total of 227 participants, 101 males and 126 females completed the questionnaires. The students at both sites were similar in age, culture and geographic background as measured by a demographic questionnaire. All participants were between the ages of 18 and 30 years and spoke Amharic and English. The mean age among females was 20.0 (SD=2.2), while male participants were slightly older with a mean age of 20.4 (SD=2.3). As illustrated in Table 2, 78.2% of the male participants were between the ages of 18-21 and 84.9% of female students were between the ages 18-21. All participants stated Amharic as their first language of choice. Oromo was the second language and English was the third most spoken language (males 15.0% and females 10.0%). Male students reported a higher frequency of speaking English in comparison to female students (37.6% and 23.0%). Similarly, more male students stated speaking two or more languages (55.4%) in comparison to females (29.4%).

When students were asked if they were sexually active, 51.2% of males reported yes, whereas only 21.7% of females stated they were sexually active. A large percentage of both male and female students reported their relationship status as single (males 82.2%, females 60.3%). Only 13.9% of male students stated they were in a relationship, whereas 31.0% of female students stated they were in a dating relationship.

The HIV-KQ-18 results were listed in order of highest correct responses, reflecting the correct knowledge of HIV transmission between male and female students. Question five (Q5), "Showering, or washing one's genitals/private parts, after sex keeps a person from getting HIV" (False), had the highest percentage of correct answers among male students (91.1%). Similarly, males scored high (90.1%) for both Q2, "A person can get HIV by sharing a glass of water with someone who has HIV" (False) and Q8, "There is a vaccine that can stop adults from getting HIV" (False) (see Table 2). Female students' high mean score was also achieved on Q5 (92.1%) and Q8 (90.5%). However in contrast to the male students, females achieved the second highest (91.3%) on Q14, "Having sex with more than one partner can increase a person's chance of being infected with HIV" (True). Females scored higher on Q1, "Coughing and sneezing do not spread HIV" (True) (90.5%) in comparison to male students (81.2%).

The HIV-KQ-18 results were illustrated by lowest incorrect responses, outlying the specific differences in scores between male and female students. The male students scored the lowest (13.9%) on Q4 "A women can get HIV if she has anal sex with a man" (False). Likewise, the male students scored lower (27.7%) on Q12 "A natural skin condom works better against HIV than does a latex condom" (False) and on Q9 (39.6%), "People are likely to get HIV by deep kissing putting their tongue in their partner's mouth, if their partner has HIV" (False). Similarly, females scored the lowest on Q4 (5.6%). Females also scored lower on Q12 in comparison to male students (27.0%). The third lowest score among females was on Q18 "Using Vaseline or baby oil with condoms lowers the chance of getting HIV" (False) (37.3%). Females also scored comparably low on Q9 (41.3%). Both male and female groups scored lower on sexual transmission knowledge of HIV. As noted above both groups scored low on Q1, Q3 and Q4. Notably, both groups scored lower on Q17 "a person can get HIV from oral sex" (True). Only 41.6% of males and 40.5% of females answered correctly on Q17 (see Table 2).

The independent t-test for the difference in total mean HIV-KQ-18 scores in male and female students did not reveal a statistically significant difference ($t=0.6$, $df=225$, $p=0.4$). Therefore, we cannot accept our initial hypothesis.

Discussion

The demographic data revealed a gender dichotomy regarding English literacy among male and female students. The male students' responses indicated that they were almost twice as likely to utilize English as a working language in comparison to female students. This result corresponds with

recent literature, which found higher rates of English literacy among male high school students in comparison to their female counterparts (Rose, 2003). Researchers have denoted higher rates of media exposure as a potential predictor in English literacy among male Ethiopians (Bahta & Utsumi; 2004; Geldof, 2007). Further studies are needed to empirically assess the interaction between HIV knowledge and the utilization of media devices among young male and female Ethiopians.

The demographic data demonstrated variance in the sexual status between male and female students. For instance, the majority of male students stated they were sexually active in comparison to females. A larger proportion of male students stated they were single in comparison to female students. In contrast, female students reported a higher incidence of being in a dating relationship compared to male students. These results corroborate with the current literature, which states that gender differences exist in reported sexual behavior among post-secondary students residing in urban settings (Adamu, Mulatu, & Haile, 2003; Astatke, Black & Serpell, 2000; Molla, Berhane, & Lindtjörn, 2008; Mulatu, Adamu, & Haile, 2000; Wouhabe, 2007). To that end, these reports align with our theoretical framework, the Social Theory of Gender and Power. In particular, Connell's theoretical dimension of cathexis outlines the power disparity in gendered relationships. Similarly, the cathexis role among young Ethiopians continues to denote risky sexual behavior among male Ethiopians, which may adversely affect the HIV rates among female Ethiopians (Adamu et al., 2003; Astatke et al., 2000; Molla et al., 2008). These findings corroborate the demographic data, which found male students to have higher rates of sexual activity and being single in comparison to female students.

Although the study did not find a statistically significant difference in HIV knowledge scores between male and female students these results contradict the initial literature review and hypothesis generated. The majority of the literature review demonstrated a gender dichotomy in the HIV knowledge levels among young Ethiopians. However, when examined closely, the literature focused on the rural context of HIV knowledge rather than an urban perspective. The selected studies and governmental statistics from the review were largely based on the agrarian populations of Ethiopia (Alene et al., 2004; HAPCO, 2008). According to Berhane et al. (2001), the rural population starkly differs on various social and health indicators in comparison to the urbanized centers of Ethiopia. A recent demographic survey demonstrated a 100% gross enrollment rate (GER) for young boys and girls residing in urban dwellings (Rose, 2003). Rose also reports a sizable gender difference in enrollment rate between rural areas in comparison to urbanized cities. This difference in education may help explain the differences in HIV/AIDS knowledge between genders in rural versus urban settings.

The initial literature review suggested lower HIV knowledge levels among rural Ethiopian females in comparison to their male counterparts (Alene et al., 2004; HAPCO, 2008). In contrast, the study results demonstrated no significant difference in HIV knowledge levels among the male and female tertiary students. These conflicting results could be associated with the contextual differences in socioeconomic status among rural populations and the specific population of college and university students. Several studies have compared HIV knowledge levels among urban youths and found similar results (Harding, Anadu, Gray, & Champeau, 1999; Nachega, Lehman, Hlatshwayo, Mothopeng, Chaisson, & Karstaedt, 2005; Tavoosi et al., 2004).

The results from the HIV-KQ-18 questionnaire demonstrated a low average knowledge score for male and female students. Particularly, both male and female students answered considerably lower on items regarding the specific modes of HIV transmission. The majority of male and female students believed HIV is transmitted by deep kissing (Q9). Similarly, a large percentage of students also scored low on vertical transmission items. For instance, the majority of both male and female students agreed that HIV is not transmitted by oral or anal sex (Q4 and Q17). These study results correspond with the current literature, which demonstrates a gap in knowledge among Ethiopian youth regarding the varied modes of HIV transmission (Alemu et al., 2004; Cheri, Mitkie, Ismail, & Berhane, 2005; Yerdaw et al., 2002).

Limitations

The study design denoted certain study limitations. The sampling strategy employed a non-probabilistic sampling frame, by recruiting participants as a convenient sample. This form of sampling frame introduces a volunteer bias among participants (Hulley et al., 2001). The

questionnaires were self-administered; therefore, social desirability cannot be excluded in the systematic biases of the study. There also exist limitations in the study design. An observational study cannot determine cause or effect, rather associations or correlations of inferences (Benson & Hartz, 2000). Although, the current study results yield important cues as to the HIV/AIDS knowledge of tertiary students, the intent is not to generalize, as the sample was not representative among the Ethiopian student population.

Implications and Conclusion

These findings illustrate the need to have further observational studies that assess the specific gendered knowledge of HIV/AIDS among urban post-secondary students. Future studies directed towards an urban student population would help compare and contrast the findings from this study and highlight specific areas of potential research.

These results are insightful in directing the content and delivery of HIV/AIDS educational resources. Because both genders scored similarly on the HIV knowledge questionnaire, the actual content of the materials for male and female students could be the same. For instance, the educational content regarding the various sexual modes of HIV transmission could be the same regardless of the student group. The methods of delivery, however, could be reflective of cultural and social norms: separate sessions could be conducted with men and women to encourage open discussion. Program evaluation and cost-analysis research could then explore the cost-effectiveness of providing the same educational content for both male and female post-secondary students. The implications for nursing practice should focus on public health campaigns, particularly on the various modes of HIV transmission (anal and oral). These initiatives should target the specific population of post-secondary students and the various misconceptions of HIV sexual transmission.

References:

- Adamu, R., Mulatu M.S., & Haile, S.I. (2003). Patterns and correlates of sexual initiation, sexual risk behaviors, and condom use among secondary school students in Ethiopia. *Ethiopian Medical Journal*, 41(2), 163-77.
- Alemu, S., Abseno, N., Degu, G., Wondmikun, Y., & Amsalu, S. (2004). Knowledge and attitude towards voluntary counseling and testing for HIV: A community based study in northwest Ethiopia. *Ethiopian Journal of Health Development*, 18(2), 82-90.
- Alene, G. D., Wheeler, J. G., Grosskurth, H. (2004). Adolescent reproductive health and awareness of HIV among rural high school students, North Western Ethiopia. *AIDS Care*, 16(1), 57-68.
- Benson, K., & Hartz, A. J. (2000). A comparison of observational studies and randomized, controlled trials. *The New England Journal of Medicine*, 342, 1878-1886.
- Berhane, Y., Gossayea, Emmelinb, M., & Hogberg, C. (2001). Women's health in a rural setting in societal transition in Ethiopia. *Social Science & Medicine*, 53(11), 1525-1539.
- Bahta, A., & Utsumi, H. (2004). *Gender studies and internet portal in the universities: Lessons for developing countries*. Retrieved from <http://www.ictes2004-gstit.edu.et/>
- Buseh, A. G., Glass, L., McElmurry, J., Mkhabela, M., & Sukati, N., A. (2001). Primary and preferred sources of HIV/AIDS and sexual risk behavior information among adolescents in Swaziland, Southern Africa. *International Journal of Nursing Studies*, 39, 525-538.
- Carey, M. P., & Schroder, K. E. E. (2002). Development and psychometric evaluation of the brief HIV knowledge questionnaire. *AIDS Education Prevention*, 14(2), 172-182.
- Connell, R. W. (1987). *Gender and Power*. Cambridge: Polity Press.
- De Visser, R., & Smith, A. (2001). Relationship between sexual partners influences rates and correlates condom use. *AIDS Education and Prevention*, 13, 413-427.
- Geldof, M. (2007). ICT for low-literate youth in Ethiopia: The usability challenge. *Lecture Notes in Computer Science*, 4560, 67-76. doi: 10.1007/978-3-540-73289-1_8
- Hadley, G., Lindstrom, D., Tessema, F., & Belachew, T. (2007). Gender bias in the food insecurity experience of Ethiopian adolescents. *Social Science & Medicine*, 66(2008), 427-438.
- Harding, A. K., Anadu, E. C., Gray, L. A., & Champeau, D. A. (1999). Nigerian university students' knowledge, perceptions, and behaviours about HIV/AIDS: Are these students at risk? *The Journal of the Royal Society of the Promotion of Health*, 119, 23-31.

- HIV/AIDS Prevention Control Office (2008). *Report on progress towards implementation of the UN declaration of commitment on HIV/AIDS*. Retrieved on February 12th, 2010 from UNAIDS.org
- Hladik, W., Shabbir, I., Jelaludin, A., Woldu, A., Tsehaynesh, M., & Tadesse, W. (2006). HIV/AIDS in Ethiopia: Where is the epidemic heading? *Sexually Transmitted Infections*, 82, 32-35.
- Hulley, S. B., Cummings, S. R., Browner, W. S., Grady, D. B., & Newman, T. B. (2001). *Designing clinical research*. (3rd ed). Philadelphia, PA: Lippincott Williams & Wilkins
- Lifson, A. R., Demissie, W., Tadesse, A., Ketema, K., May, R., Yakob, B., Metekia, M. et al. (2012). HIV/AIDS stigma-associated attitudes in a rural Ethiopian community: Characteristics, correlation with HIV knowledge and other factors, and implications for community intervention. *International Health and Human Rights*, 12(6), 1-8.
- Levin, K. A. (2006). Study design III: Cross-sectional studies. *Evidence-Based Dentistry*, 7, 24-25.
- Maharaj, Z. (1995). A social theory of gender: Connell's "gender and power". *Feminist Review*, 49, 50-65.
- Molla, M., Berhane, Y., & Lindtjørn, B. (2008). Traditional values of virginity and sexual behaviour in rural Ethiopian youth: results from a cross-sectional study. *BMC Public Health*, 8(9), 1-10. doi:10.1186/1471-2458-8-9
- Molyneux, S., & Geissler, W., P. (2008). Ethics and the ethnography of medical research in Africa. *Social Science & Medicine*, 67(2008), 685-695.
- Mulatu, M. S., Adamu, R., & Haile, S. I. (2000). Psychosocial and contextual determinants of past and intended condom use among Ethiopian secondary school students. *Northeast African Studies*, 7(1), 85-108.
- Nachega, J. B., Lehman, D. A., Hlatshwayo, D., Mothopeng, R., Chaisson, R. E., & Karstaedt, A. (2005). HIV/AIDS and antiretroviral treatment knowledge, attitudes, beliefs, and practices in HIV-infected adults in Soweto, South Africa. *Journal of Acquired Immune Deficiency Syndrome*, 38(2), 196-201.
- Rose, P. (2003). *Can gender equality in education be attained? Evidence from Ethiopia*. Background paper prepared for the Education for All Global Monitoring Report 2003/4. Retrieved from www.worldbank.org
- Statistical Package for Social Science. (2009). *SPSS base 17.0 for Macintosh user's guide*. SPSS Inc., Chicago.
- Taffa, N., Klepp, K. I., Sundby, J., & Bjune, G. (2002). Psychosocial determinants of sexual activity and condom use intention among youth in Addis Ababa, Ethiopia. *International Journal of STD & AIDS*, 13, 714-719.
- Tavoosi, A., Zaferani, A., Enzevaei, A., Tajik, P., & Ahmadinezhad, Z. (2004). Knowledge and attitude towards HIV/AIDS among Iranian students. *BMC Public Health*, 4(17), 1-16
- Ukwuani, F. A., Tsui, A., & Suchindran, C. M. (2003). Condom Use for Preventing HIV Infection/AIDS in Sub-Saharan Africa: A Comparative Multilevel Analysis of Uganda and Tanzania. *Journal of Acquired Immune Deficiency Syndromes*, 34(2), 203-213.
- Wingood, G. M., DiClemente, R. J. (2000). Application of the theory of gender and power to examine HIV-related exposures, risk factors and effective interventions for women. *Health Education & Behavior*, 27(5), 539-565.
- World Health Organization. (2009). *Country cooperative strategy at a glance*. Retrieved August 15th, 2010 from <http://www.who.in/countries/eth/en/>
- Wouhabe, M. (2007). Sexual behavior, knowledge and awareness of related reproductive health issues among single youth in Ethiopia. *African Journal of Reproductive Health*, 11(1), 14-21.
- Yerdaw, M., Nedi, T., Wheeler, B., & Grosskurth, H. (2004). Assessment of awareness of HIV/AIDS among selected target groups in and around Addis Ababa, Ethiopia. *African Journal of Reproductive Health*, 6(2), 30-38.
- Yizengaw, T. (2007). Implementation of cost sharing in the Ethiopian higher education landscape: Critical assessment and the way forward. *Higher Education Quarterly*, 61(2), 171-196.

Table 1: Demographic Results among Male and Female Students

Variable	Total (n, %)	Male, n (%)	Female, n (%)
Sex	227(100.0)	101 (44.5)	126 (55.5)
Age Cohort (yrs)			
18-21	186 (82.0)	79 (78.2)	107 (84.9)
22-25	33 (14.5)	17 (16.8)	16 (12.7)
26-30	8 (3.5)	5 (5.0)	3 (2.4)
*Total	227	101	126
Religion			
Orthodox	114 (53.0)	42 (42.8)	72 (61.5)
Protestant	74 (34.4)	38 (38.8)	36 (30.8)
Muslim	22 (10.2)	18 (18.4)	4 (3.4)
Catholic	5 (2.3)	0 (0.0)	5 (4.3)
Total	215	98	117
Education			
University	107 (47.1)	48 (47.5)	59 (53.2)
College	120 (52.9)	53 (52.5)	67 (46.8)
Total	227	101	126
Sexually Active			
Yes	61 (36.5)	43 (51.2)	18 (21.7)
No	106 (63.5)	41 (48.8)	65 (78.3)
Total	167	84	83
Relationship			
Single	159 (70.0)	83 (82.2)	76 (60.3)
Dating	53 (23.3)	14 (13.9)	39 (31)
Married	9 (4.0)	0 (0.0)	9 (7.1)
Divorced	2 (0.8)	0 (0.0)	2 (1.6)
Separated	4 (1.8)	4 (4.0)	0 (0.0)
Total	227	101	126
Number of Languages Spoken			
1	134 (59.0)	89 (70.6)	45 (44.6)
2	56 (24.7)	24 (19.0)	32 (31.7)
>3	37 (16.3)	13 (10.3)	24 (23.8)
Total	227	126	101

* Totals are different due to missing variables.

Table 2: Number and Percentages of Correct Answers in HIV Knowledge Items (HIV-KQ-18) among Male and Female Students

HIV Knowledge Scores Correct Answers: n(%)	Male n=101	Female n=126
1. Coughing and sneezing do not spread HIV. (T)*	82(81.2)	114(90.5)
2. A person can get HIV by sharing a glass of water with someone who has HIV. (F)*	91(90.1)	110(87.3)
3. Pulling out the penis before a man climaxes keeps a woman from getting HIV during sex. (F)	71(70.3)	82(65.1)
4. A woman can get HIV if she has anal sex with a man. (T)	14(13.9)	7(5.6)
5. Showering, or washing one's genitals/private parts, after sex keeps a person from getting HIV. (F)	92(91.1)	116(92.1)
6. All pregnant women infected with HIV will have babies born with AIDS. (F)	78(77.2)	104(82.5)
7. People who have been infected with HIV quickly show serious signs of being infected. (F)	83(82.2)	108(85.7)
8. There is a vaccine that can stop adults from getting HIV. (F)	91(90.1)	114(90.5)
9. People are likely to get HIV by deep kissing, putting their tongue in their partner's mouth, if their partner has HIV. (F)	40(39.6)	52(41.3)
10. A woman cannot get HIV if she has sex during her period. (F)	86(85.1)	97(77.0)
11. There is a female condom that can help decrease a woman's chance of getting HIV. (T)	86(85.1)	104(82.5)
12. A natural skin condom works better against HIV than does a latex condom. (F)	28(27.7)	34(27.0)
13. A person will not get HIV if she or he is taking antibiotics. (F)	81(80.2)	96(76.2)
14. Having sex with more than one partner can increase a person's chance of being infected with HIV. (F)	88(87.1)	115(91.3)
15. Taking a test for HIV one week after having sex will tell a person if she or he has HIV. (F)	72(72.3)	102(81.0)
16. A person can get HIV by sitting in a hot tub or a swimming pool with a person who has HIV. (F)	70(69.3)	90(71.4)

CHANGES LYMPH FLOW, HEART AND TRANSCAPILLARY EXCHANGE FLUID AND PROTEINS REFLEX-HUMORAL INFLUENCES

Myrzakhanov N.

Myrzakhanova M.N.

Eurasian National University named after L. Gumilev, Kazakhstan Dipartimento di Biologia,
Universita di Genova, Italy

Abstract:

This report presents the results of a study lymph, lymph pressure, heart, arterial and venous pressure, state transcapillary exchange of fluid and protein in normal and correlation analysis. In addition, data on changes in the heart and lymph flow if the pressure in the lymphatic vessels and with intramuscular kallekreina. To study the reflex influences of lymphatic vessels to the lymph vessels and the heart of chronic experiments were carried out in 26 sheep with lymph-venous anastomosis and the catheter according to our methodology. We have also revealed the presence of interceptive reflexes and set intralymphatics reflex effect of receptor lymph vessels in the heart. Thus, these data suggest that each region of the body and the body as a whole is determined by multiparametric lymph circulation relationship between different functional systems (lymph, blood, heart activity, the state of biological barriers), providing homeokinesis.

Key Words: Lymph Formation, Perfusion, Intralyuminalny Pressure, Transcapillary Exchange

1. Introduction

Existing concepts of the processes lymph formation, lymph circulation and mechanisms of their regulation primarily composed on the data obtained in dogs and small laboratory animals (1, 2, 3, 4). So far are single (5) research conducted of large mammals, including domestic animals. There is evidence that the experiments carried out on mammals with unique morphological and physiological homeostatic organization will bring some clarity to the knowledge of the general laws of lymph hemodynamics.

This report presents the results of a study lymph flow, lymph pressure, heart, arterial and venous pressure, state transcapillary exchange of fluid and protein in normal and correlation analysis. In addition, data on changes in heart activity and lymph flow with increasing pressure in the lymphatic vessels and with intramuscular injection kallekreina.

2. Materials And Methods

To study the reflex influences of lymphatic vessels to the lymph vessels and the heart of chronic experiments were carried out in 26 sheep with lymph-venous anastomosis and the catheter according to our methodology (6).

Stimulation of receptors perfused lymphatic vessel increased pressure produced by the introduction of warm physiological solution through leading the catheter with a closed deferent catheter (with increasing intralyuminalny pressure). The pressure in the mesenteric lymph vessel increased to 5-8, and in the cervical and liver - to 3-6 mm mercury column.

Duration of exposure to receptors ranged from 30 seconds to 2 minutes.

The value of the maximum and minimum blood pressure was determined by the method (7), and venous - water manometer. The pressure in the lymphatic vessels also measured water manometer by connecting it to a tee area lympho-venous anostomoza. Lymph flow velocity was measured in ml / min through the collection of lymph in about 5-10 minutes, and sometimes drops the lymph flowing from the cannula, which is recorded on the tape elektrokimografa.

For preventing clotting of blood and lymph intravenously administered heparin rate of 1500-2000 ME per 1 kg of body weight.

Blood was taken from the carotid artery, jugular and common mesenteric vein before the morning feeding ("fasting") and it was determined by the total protein content (refractometric) and hematocrit. By arterio-venous difference of protein and hematocrit were determined character transcapillary exchange fluid and plasma proteins (8). Pancreatic kallekrein (firm Vantrop, England) was injected intramuscularly at a rate 40ed, previously dissolved in prolongators.

Accounting for volumetric flow rate of colon and cervical lymph heartbeat, salivation and total protein content made prior to drug administration and at 15, 30, 60, 75 and 90 minutes after injection kallekreina.

All digital material processed by variation statistics methods(9).

3. Results

Experiments have shown that in the rest of the intestinal lymph flow and cervical lymph ducts varies considerably from animal to animal (table.1). However, these variations in terms of lymph flow per 1 kg of body weight practically smoothed and average 0,732 ml/min/kg in the case of intestinal and 0,009 ml/min/kg in respect of cervical lymph.

Conjugacy changes listed in Table 1 of indicators identified in their correlation analysis. Found significant ($P < 0,001-0,05$) positive correlation between the rate of lymph flow from the intestinal lymphatic trunk and minimal blood pressure, or pressure in the common mesenteric vein, or pressure in the lymph vessels. The correlation between the current lymph from cervical of lymph vessel and the above indicators were below the limits of reliability first threshold. Phonocardiogram analysis showed that the short duration of the sheep, and especially long pauses more volatile than the duration of the first and second tones. Long pause less long in sheep with rapid pulse, and conversely, longer in sheep with a rare heart rate.

Found that systolic rate in sheep is directly related to frequency heart rate. Calculations to determine the permeability of the blood vessels based on the testimony the total protein and hematocrit showed that both in the vessels of the systemic circulation, and in the vessels of the intestinal region, the transition is not only fluid but also proteins in the main direction of the fabric → blood, and (in lesser extent) in the direction of the fabric ← blood. Less common state of equilibrium (table.2).

The data in Table 2 show that the loss of fluid and protein is much higher in the vessels of the systemic circulation than in the vessels of the intestinal region. Correlation analysis did not reveal any relationship between systole and movement of water and protein in the direction of the blood → tissue. However, was a significant relationship ($P < 0.05-0.001$) between the loss of fluid from the blood vessels and lymph pressure. Similar connection speed in the intestinal lymph flow and cervical lymph vessels could not be identified (table.1,2).

Table 1

The value of lymph flow, arterial, venous and intra lymphatic pressure in normal sheep

Sheep	lymph flow, ml/min		lymph pressure, mm. water column		arterial pressure, mm. mercury column		pressure in the common mesenteric vein, mm. water column
	intestina 1	cervical	intestinal	cervical	intestinal	cervical	
1	0,44	0,09	38	20	120	40	10
2	0,82	0,19	12	8	110	45	9
3	0,90	0,25	20	13	107	57	10
4	1,00	0,22	36	17	110	55	9
5	1,30	0,31	10	16	130	60	14
6	1,40	0,26	68	21	104	62	16
7	1,50	0,36	42	18	108	62	18
8	1,56	0,37	43	20	120	65	18
9	1,60	0,57	48	24	110	65	10
10	1,80	0,85	40	23	104	70	20
M±	1,23	0,34	36	18	121	58	13

Table 2
Indicators of capillary permeability in healthy sheep in normal

sheep	Loss of fluid and protein per 100 ml of arterial blood			
	In vessels of the great circle		In the vessels intestinal region	
	for fluid, ml	for protein,%	for fluid, ml	for protein,%
1	+8	+0,11	-14	-0,10
2	-8	+0,18	-11	-0,06
3	-25	-0,23	-0	-0,16
4	-19	+0,06	-11	-0,24
5	-35	-0,3	-20	-0,12
6	-35	-0,69	-16	-0,18
7	-20	-0,31	-4	-0,19
8	-11	-0,40	-11	-0,21
9	-24	-0,21	-11	-0,13
10	-30	-0,43	-4	-0,08
M±	-23 3,3	-0,34 5,5	-10,2 1,6	-14,4 1,8

Note: Phase (+) represents the water pressure and protein in the direction of tissue-blood (-) blood-tissue and (0) the equilibrium state.

In chronic experiments on 23 animals were studied reflex influences from one to the other lymph lymphatic vessels and heart function (ECG). Increasing intraluminal and end pressure in the lymphatic vessels - a receptor in most of the experiments was accompanied by inhibition of lymph flow in the other lymphatic vessels and slowing of heart rate. If prior to stimulation of the lymph flow from cervical lymphatic vessel was $0,32 \pm 0,07$ ml/min, in a period of increased pressure in the mesenteric lymph vessel he stopped equal $0,09 \pm 0,04$ ml/min ($n=10$, $P<0,01$). This is not by the nature the reaction was observed by the intestinal lymphatic trunk during stimulation of baroreceptors of the cervical lymph trunk. Changes in cardiac function with increasing pressure in the lymphatic vessels (the cervical, mesenteric, hepatic) expressed significant elongation ($P < 0.01$) of the cardiac cycle ($P = P$), the of unreliable change intervals PQ, PR, TP.

In experiments with intramuscular introduction kallekreina 3 sheep intestinal catheters, cervical lymph ducts and the parotid gland found that in response to the administration of the drug is increased lymph separated, increased heart rate and profuse salivation.

The effect of the lymph flow and heart rate were the highest for 60 minutes, and saliva for 30 minutes (table. 3). Correlation analysis revealed a significant association between intestinal, cervical lymph flow and pulse rate ($P<0,001$). Analysis of the total protein content showed that 45-60 minutes is a noticeable increase in his intestinal (in 3,2-5,9%) and in the cervical (at 7,9-9,9%) in the lymph (table.3).

Table 3
Change in lymph flow, pulse, salivation and total protein with intramuscular injection kallekrina

Indicators of		Time, min						
		norm	15	30	45	60	75	90
Lymph flow, ml/min	intestina l	1,90±0,08	1,22±0,17	0,78±0,11	1,58±0,35	2,40±0,50	2,30±0,35	1,06±0,12
	cervical	0,46±0,09	0,78±0,18	0,61±0,23	0,70±0,11	1,46±0,28	1,49±0,34	1,51±0,11
Total protein,%	intestina l	3,34	3,32	3,36	3,45	3,54	3,36	3,26
	cervical	2,91	2,56	2,57	3,14	3,20	2,54	2,76
Pulse beats per minute		44,0±1,4	42,0±1,6	51,0±0,08	62,0±2,2	60,0±1,6	52,0±2,1	46,0±0,8

Salivation, ml/min	0,65±0, 07	1,30±0, 09	1,46±0, 15	1,34±0, 02	1,07±0, 05	0,82±0, 06	0,60±0, 08
--------------------	---------------	---------------	---------------	---------------	---------------	---------------	---------------

4. Conclusion

Our results show that in normal sheep, there is some association between changes in lymph circulation and blood circulation and transcapillary exchange of fluid and plasma proteins. This relationship was not always clear, which confirms the presence or absence of correlations, the degree of reliability. Apparently, we have identified a positive relationship between the amount having withdrawn from the blood vessels and the pressure of the fluid in the lymphatic vessels in the absence of any correlation between the systole and release fluid from the blood vessels and plasma proteins can be explained. It is known that the pressure of the lymph, which provide its current depends not only on the pressure within the lymphatic vessels, and a number of other supporting factors (10). As for the lack of correlation between the systole and transcapillary exchange of fluid and proteins, there is evidence that the formation of lymph depends not only on the hydrostatic forces, but also on the level of permeability of biological barriers (10). This is also demonstrated by our experiments with intramuscular introduction kallekreina, which caused not only a rapid heart rate, and increased vascular permeability, which is consistent with the literature (11).

We have also revealed the presence of interceptive reflexes and set intra lymphatic reflex effect of lymphatic vessels on cardiac function.

It is known that gematolymphatics balance is maintained and constantly maintained, if there is a balance between transport of fluid from the blood into the tissues and venous and lymphatic return(10).

Comparison of the nature of reflex influences from the lymphatic vessels to the lymph vessels and the heart strongly suggests that the increase in pressure in the lymph collectors, causes inhibition of cardiac activity, apparently, reduce filtration of blood capillaries and thus has a regulating effect gematolymphatics equilibrium body. The same idea is confirmed by experiments with the introduction of pancreatic kallekreina when in 3 cases after a sharp increase in cervical lymph flow (15 min.) reported slowing of heart activity, it is very obvious due to the reduction of intestinal lymph flow for 30 minutes. The same seems to indicate we have identified in these experiments is a direct positive correlation between the cervical and intestinal lymph flow ($P<0,01$).

Thus, these data suggest that each region of the body and the body as a whole is determined by multiparametric lymph circulation relationship between different functional systems (lymph, blood, heart activity, the state of biological barriers), providing homeokinesis.

References:

- Zhdanov D.A. General anatomy and physiology of the lymphatic system. L., 1952.
- Kohanina M.I. Reflex effect of certain receptors on the lymph flow internal organs. Tr. Institute of Physiology. Kaz SSR Academy of Sciences. 1965, 6, 101-267.
- Myrzahanov N.M. Reflex effect on the current, the protein and the ionic composition of the intestinal lymph. Diss. candidate. Tartu.,1974.
- Rusnyak I., Foldi M., Szabo D. Physiology and pathology of lymph circulation. Budapest, 1957.-S. 856.
- Aynson X.X. Physiological features of the selective permeability of the capillaries to the protein molecules and the possibility of influencing it. AN ESSR, Biol., 1972. - B. 21. - In 4. - S. 295-299.
- Myrzahanov N. Methodical recommendations for the laboratory practical concepts to students "lymph and lymph circulation" course of normal physiology of farm animals. Semipalatinsk, 2010.
- Saraykin I.M. Method definitions and indicators of arterial blood pressure in sheep. Tr. Saratov Veterinary Institute, 2006,6,180-183.
- Kaznacheev V.P. Role of the nervous system in the permeability of blood capillaries in some physiological and pathological state.. Diss. candidate. Novosibirsk,1953.
- Plohinsky N.A. Guide to Biometrics for livestock. Moscow, 2009.
- Kupriyanov V.V. Gematolymphatics balance as an indicator of the state of the microcirculation. Cardiology, 1974,14,2,97-100.
- Meleva N.S., Sukhodola V.D. Influence kallekreina the salivary glands on the composition of bile. Physiology journal USSR, 2011,65,5,747-750.

HRD MECHANISMS IN HEALTH CARE SECTOR IN J&K: A COMPARATIVE STUDY

Dr. S.A. Mufeed
Qurrat.A. Hamdani

The Business School, University of Kashmir

Abstract:

Human Resource Development (HRD) is a framework for the expansion of human capital within an organization. It is a combination of training and education that ensures the continual improvement and growth of both the individual and the organization. HRD mechanisms fall in one of the elements of HRD climate, which measure the extent, to which the HRD mechanisms are implemented. Various HRD mechanisms such as training & development, career planning, career development, career counselling, performance appraisal system, organization development, quality of work life, workers participation in management are being used in various organizations for building HRD climate. Research on HRD mechanisms at National and International level, particularly in Health care sector is limited. Taking into consideration the limited study on Health care sector, the present study was undertaken on two hospitals-SKIMS & SMHS in the state of J&K with special focus on Performance Appraisal system (PAS) and Training and Development (T&D).

Key Words: HRD Climate, Performance Appraisal System, Training and Development

Introduction

Human Resource Management is a means for improving efficiencies in terms of better productivity, reduction of costs, better generation of internal resources, better profits and better customer service. Human Resource is considered to be very crucial in the organization's well-being. Organizations will have to build global capabilities such as the ability to seamlessly move talent, ideas and information around the world to create products and services at an optimum pace with quality above par. They will have to build an employee infrastructure for hiring, maintaining, training and developing that takes a global perspective which is made possible through the Human Resource Development (HRD) activities. Human Resource Development (HRD) department is, thus, should be a continuous process to ensure the development of employee competencies, dynamism, motivation and effectiveness in a systematic way. Experiment on 14 different organizations, 7 public and 7 private, were analyzed by T V Rao using the framework of HRD linkage and the results revealed that Human Resource Development (HRD) as a function has evolved in India indigenously from the year 1975 when *Larsen & Toubro (L&T)* conceptualized HRD as an integrated system and decided to separate it from the personnel function. Since then, most organizations have started new HR departments or redesigned their personnel and other departments as HRD departments. The organization was looking at the PAS, but it was felt that it may not achieve objectives unless accompanied by other sub systems like potential appraisal, T&D. In *Crompton Greaves LTD (CGL)*, PAS was identified as priority area. CGL started thinking about introducing HRD systems in 1979-80. *L&T Construction Group (ECC)* identified several issues facing them; one of the issues which was prioritized first was performance appraisal. A new appraisal system called as performance analysis and development system (PADS) has been developed and is being implemented. *Jyoti LTD*, initiated HRD with the attempts to redesign the PAS. *TVS Iyenger & Sons* divided its personnel function into personnel administration and HRD. The organization focused on manpower planning, job rotation, training, potential appraisal, career planning, job enrichment etc. *Voltas Ltd* while reviewing that reasons for its poor performance, came to the conclusion that inadequate attention was paid to the development of human resources. Voltas revised their PAS from time to time. The HRD unit at *Sunderam Fastners Ltd(SFL)* has been working on different HRD mechanisms. Research was conducted for introducing new PAS. Training needs were identified through PAS. *Bharat Earth Movers (BEML)* initiated HRD in 1978 & 1979 focussing on PAS, performance counselling, training,

career development & potential appraisal exercises. *BHEL's* sudden dip in production was a matter of concern for the organization and an OD department was formed in 1981. The organization started introducing a new PAS. *Bank of Baroda (BOB)*, with the help of consultant introduced a separate HRD department to look after manpower planning, training & PAS. HRD department also strengthened the training function. *SBI* started HRD practices by introducing a development oriented performance appraisal system and simultaneously started working on other areas as job rotation, training, career dev, feedback & counselling. HRD mechanisms like T&D & survey feedback were initially used in *State Bank of Patiala (SBP)*. Later on, job rotation, PAS, career planning & manpower planning were paid attention to. It was in 1958 that *Indian Oil Corporation (IOC)* set up HRD departments in each unit. Today, there are high expectations from HRD. Human Resource Development is the frameworks for helping employees develop their personnel and organizational skills, knowledge and abilities. Human Resource Development includes such opportunities as employee training, employee career development, performance management and development, coaching, succession planning, key employee identification, tuition assistance and organization development.

HRD being an important function of HRM has been incorporated in almost all the organizations either on a small or large scale. Research on 20 organizations reveals no such organization which survives without a HRD department. Sooner or later every organization has realized the importance of such a department which would take care of the developmental aspect of their HR. 40 organizations from all areas were studied on the aspect of HRD. *Lanco Global Systems*, having IT as the nature of activity was established in 1999 and HRD was introduced by the chairman and MD, who are of the opinion that HRD is like a flower in bloom to be experienced. Narayan Murty initiated HRD in *Infosys*-an IT service and IT consulting organization, established in 1981, where ongoing training programs meet the different learning need of employees in specific areas of technology, management, leadership, cultural and communication skills and other soft skills. *Toyota* established in 1948, believes that developing people extends beyond vehicle manufacturing, helps people improve the quality of life. Steve Job realized the importance of HRD and helped *Apple* employees to be specialists in every role. *Face book*, a recently established (2004) social networking site focuses on developing employees to increase employee satisfaction, which in turn has led to increased productivity in terms of increased revenue every year. *Google*, dealing in internet and computer software, encourages highly trained employees. *Honda* (1948), an automotive and aviation concern, focuses on innovation, value creation, customer satisfaction by increasing dedication on the part of workers. Sanjay Bali (VP-HR) in *Samsung* is of the belief that employees are its most important asset. The organization which was established in 1995, believes that by giving opportunities for growth, employees will perform to their fullest potential. *Mahindra Satyam* an 1987 established IT service, IT consulting and software service organization has trained employees thus believing in HRD aspect.

Hyundai believes in "quality", and has developed quality HR and products. *Northern Trust*, though an old organization, hasn't yet focussed much on HRD. The place is not progressive but values creativity. The top management at *Hero Group* initiated HRD practices, focussing on innovation and value creation. *IKEA*- established in 1943 and a dealer in furniture has reduced employee turnover. Performance of duties is in accordance to the schedule. The organization enables people to grow. The vision of the leaders of yesteryears at *SIMA* (South Indian Mills Association)-1933 led to the development of the HRD centre. Industries realized that the solution of all problems is one necessary aspect i.e. HRD. At *Royal Bank of Canada*, establishment of HRD has led to fewer turnovers. The organization commits to structured performance management system that continually empowers employees to succeed. *Wipro* (1980) provides fewer opportunities in terms of HRD to the employees. *Arvind* (1931) dealing in garments, furnishes highly skilled and trained employees, focuses on effective performance. *Pepsi* (1989) focuses on HRD aspect in a different way by giving early responsibility, risk taking and participation in management. The organization is geared to churn people to develop them. *LG*-a 1958, mobile equipments and digital appliances company lacks a proper HRD area. Most employees are thrown into positions without proper training or orientation, are bombarded with work, given additional responsibilities in a short span of time.

Review Of Related Literature

Jane & Robin (2007), conducted a research on International Performance Appraisal policies practices and processes in Australian subsidiaries of health care MNC's. The findings suggested that the host country manager's experience of the appraisal often resulted in the manager perceiving limited opportunities for career development and advancement in the company. The findings showed that there was complete dissatisfaction with the feedback and follow up. A total of 78% were dissatisfied with the feedback and 95% experienced no follow up after the feedback. Only a third of the HCMs believed their current appraisal was beneficial to them. The study revealed three core problems that represented major influences on the PA process for HCMs. Firstly the HCM's appraisal was found to be limited because of the firms overwhelming short term emphasis on sales and profit termed 'bottom line'. Secondly the HCM was disadvantaged by being solely dependent on the remote relationship with his/her supervisor, who rarely had any commitment to the HCM's appraisal because his/her primary concern was the sales and profits of the business. Thirdly, the appraisal was found to be influenced by the international strategy and the structure adopted by the parent MNCs. An investigation was conducted on development of competence based management and performance assessment system for academic management by *Pooja & Suri (2010)*, the paper explored the various competencies and their relevance to educational institution. The study was aimed to explore the relevance of the competencies from the industrial sector in education sector. The study provided the base for the competency model for the faculty members. *Claudia, Isabel & Van (2000)*, conducted the research on Managing the Performance of Family Physicians in the Portuguese National Health System. The paper explored the awareness, among health system managers at various levels, of problems with performance, as well as their perception of what was being done and what could be done to improve it. This study showed that performance management was a poorly developed part of the management armamentarium of public sector managers in the Portuguese NHS. Health services managers appeared aware of the need to find mechanisms for performance management at the institutional level, but showed little concern for performance management at in the individual or sub-institutional level. Moreover, they apparently focused on evaluation of process and structure, and made little or no mention of assessing production of health.

A survey on the case of health sector regarding using performance management to meet changing citizen needs in the Mauritian context by *Hemant & Needesh (2006)*, a survey was carried out among Public sector officers in the Mauritian Health Sector to find out how far they took into consideration the needs of citizens when it came to analyzing their performance and setting performance targets, and finding the rationale for such choices. From the analysis of the present situation in the Mauritian Public Health Sector, however, it seemed that not all conditions were present for the implementation of performance management in Mauritian public hospitals with a view to improving quality of services dispensed. There were different piece-meal solutions that had been implemented so as to meet the needs of citizens. Performance management in the case of Mauritian hospitals, nevertheless, was still a farfetched idea because irrespective of how good a technique it is, there was a fear to implement it due to sheer shortage of staff who still had to satisfy the customers to the 'extent possible'. Performance management or not, there were indications that, at least, there had been continuous improvement in quality of service and reduction in the time taken to provide services to citizens in Mauritian hospitals by increasing the number of doctors and other staff and equipment including new types of specialized care previously not possible to provide locally.

An assessment of performance, management in the health care industry covering public-private sector hospitals was carried by *Aykut, Mikail & Kuralay (2011)*, and the findings suggested that Performance appraisal was either never carried out in the public hospital or carried out once a year at most; whereas it was conducted in the private hospital once-twice a year mostly by the hospital management. The most common performance appraisal method in both hospitals was —Comparison Methods. Rewarding mechanisms was mostly ignored in both the public and the private hospitals. According to the results of the questionnaire administered to 78 doctors employed in the public and private hospitals, the most common problem in performance management was the non-objective character of performance appraisal. Between the hospitals, only the problem of different working styles of individuals differed. Other problems did not differ from one to the other. According to the results of the questionnaire administered to 78 doctors employed in the public and private hospitals, it was observed that only favouritism, among the mistakes made by hospitals in performance appraisal,

was different in the two hospitals. It was determined that this mistake was more widespread in the public hospital. Effects of other mistakes in appraisal did not differ from one hospital to the other. Finally, it could be stated that it is necessary to constantly perform monitoring and correcting activities in order for a hospital to manage performance. For this, the hospital management should be fully committed to the practice of performance management system. The performance management project could not continue unless this goal is fully guaranteed.

A study on executive MBA students in Bangladesh regarding employee perception of performance appraisal politics was conducted, which investigated the relationship of employees' Perceptions of Performance Appraisal Politics (POPAP) with the employees' organizational outcomes such as job satisfaction, organizational commitment and turnover intention among the full time working (executive) MBA students in Bangladesh. The results of this study on the 67 executive MBA students selected from the four business schools in Bangladesh confirmed that the employees' POPAP for punishment motive significantly reduced their job satisfaction and organizational commitment and encouraged them to quit the organization. Contrarily, the employees' POPAP with a motivational motive, increased job satisfaction and organizational commitment and reduced the turnover intention of the employees. In order to facilitate the discussion, sub group analysis was conducted based on the gender, age, type of employees, and experience (in years). At first, t-test was conducted to see whether employees' POPAP for motivational motive, punishment motive, satisfaction and commitment level and turnover intention differed based on their gender. No significant difference was observed between male and female in this regard. The employees' POPAP for the punishment motive significantly reduced the employees' job satisfaction and organizational commitment, and encouraged them to quit the organizations. Conversely, the employees' POPAP for the motivational motive, increased the job satisfaction and organizational commitment, and reduced the turnover intention of the employees. These findings signified that the employees favoured those kinds of performance appraisal politics that benefit them.

Motivation, Performance and Satisfaction among University Teachers was compared in Public and Private Sectors in Pakistan and Malaysia. The research aimed to determine the difference in performance, achievement motivation and job satisfaction of teaching faculties of selected private and public sector higher educational institutions of Pakistan and Malaysia, which are countries representing South/South-East Asia in cross-cultural perspective. The study revealed that private and public sector management were playing significant role in managing teaching faculty of educational institutions of Pakistan and Malaysia. There did not seem cross-cultural difference between Pakistan and Malaysia when performance and job satisfaction of teaching faculty of private sectors is concerned; however, both private and public sectors educational management of Malaysia seemed to be doing vital job in developing more achievement motivation in their teaching faculty than Pakistani management. Moreover, performance of public sector Malaysians teachers appeared to be more than Pakistan public sector teachers, although it was not in the low range. Nevertheless, public sectors teachers of Pakistan were more satisfied with their jobs than private sector Pakistani teachers. *Jawahar & Stone, (1997)*, surveyed the Effects of Appraisal Purpose, Perceived Consequences, and Rater Self Monitoring on Leniency of Ratings and Decisions in the education sector. This study investigated one avenue through which appraisal purpose influences ratings and related personnel decisions. Support for hypotheses suggested that ratings and decisions were more likely to be distorted as the severity of consequences increased. *Weiyang, Yinmin, Mu Hu & Xiumei (2009)*, surveyed the Performance evaluation of inpatient service in Beijing: a horizontal comparison with risk adjustment based on Diagnosis Related Groups. Evaluation was undertaken at all at large public hospitals of general acute care in Beijing.

Melkidezek, Eustace, Mwangi & Naboth (2008), conducted a cross sectional study involving a sample of 448 hospital workers regarding motivation of health care workers in Tanzania: A study of Muhimbili National Hospital. A research examined the experiencing performance appraisal in a trust hospital. Article firstly examined the changing nature of employee management under PA, before it investigated the contemporary usage of PA and the effects on women. This was illustrated with research, gathered from a case study in the Midlands. The article also examined the changing focus of PA as a means through which the marginal and not so marginal performer could be controlled. Analysis focused on the use of subjective images of 'women', through PA, for creating functionally

flexible workers in a 'quality' environment. *Bharati & Kirti (2011)*, researched to bring about a linkage in Balanced Card to Performance Management System a Process Model.

Sorush (2000) expressed the same point in this field. The research outcome also showed that the performance appraisal results had a little effect on increasing the motivation level. The research also showed that in the present appraisal processes, there was a low or medium level in the justice, the concreteness of the method standards and the authorities control on the appraisal trend. A study of the various performance management systems adopted by select Indian private sector organizations was conducted by *Bindu & Ashish (2011)*, this paper tried to find out the performance management systems adopted by select Indian Private Sector Companies and the kind of measures they used to measure their performance. This research found that all the companies were using both financial as well as non-financial measures to measure their performance and most of the managers were satisfied with their existing performance management systems. The main objective of this research was to find out the different types of Performance Management systems that were used by the Indian Private sector Companies. A whole lot of Performance Indicators were revealed by the companies surveyed consisting of both financial and non-financial ones and all Companies were using a whole lot of financial as well as Non-financial Measures in every functional aspect be it customers, environment, employees and even internal control. The most widely used performance management system as per this survey was the 360 degree appraisal followed by TQM, and Activity Based Costing. This showed that organizations were more oriented towards employee appraisal than the whole performance of the organization. Most of the Managers were happy with their existing performance management system and very few were thinking of changing their existing PMS. The one's wanting to change their PMS said that they would like to go for the contemporary PMS like the Balanced Scorecard. *Manoharan, Muralidharan & Deshmukh (2009)*, surveyed employee Performance Appraisal using Data Envelopment Analysis. This study supported the ideas that rating formats needed re-examination with a focus on computer based models as an alternative to traditional rating methods. Earlier adopted methods had seldom identified and quantified the individual factors for inefficiency whereas DEA could overcome these shortfalls.

A cognitive analysis on moderators of sex bias in the performance appraisal process was conducted and the study was designed to analyze both processing characteristics and situational moderators which operate to influence sex bias in performance evaluations. More specifically, this research was designed to address the influence of rate sex on the recall and the ratings assigned in performance appraisal. An attempt was made to determine which, or if both, recall and rating stages in the process were susceptible to an influence of rate sex. An additional purpose of this study was to identify the situational factors which acted to moderate the influence of rate sex in performance evaluation. This study provided findings which enhanced the understanding of how sex bias operated to influence performance evaluations. These results identified contrasting levels of sex bias susceptibility among different points in the appraisal process. These findings suggested that raters were able to remember performance incidents without the interference of sex bias. Hence, processes preceding recall such as observation and encoding may not be subject to sex bias. These findings suggested that sex incongruence did play a role in the sex bias process, however not in the way that was predicted. The results of this study suggested that sex bias in evaluations was eliminated once the gender of the sex-incongruent rate was unique. These findings suggested that cognitive as opposed to motivational reasons might underlie sex bias in evaluations.

A research carried on the match between motivation and performance management of health sector workers in Mali, showed that the main motivators of health workers were related to responsibility, training and recognition, next to salary. These could be influenced by performance management (job descriptions, supervisions, continuous education and performance appraisal). The results showed the importance of adapting or improving upon performance management strategies to influence staff motivation. The study revealed that the main motivators for health workers in all eight professional categories were related to recognition or appreciation, responsibility and training. Although salaries and incentives were important factors for health workers and should not be neglected, the study did show that gains in motivation could be made by giving greater responsibility to staff, by holding staff responsible and by improving mechanisms for recognition. Suresh in his assessment of appraisal of management for lending decisions concluded that the personality traits of an entrepreneur were manifested by his/her behaviour characteristics. It is however clear that many of

these traits had an important bearing on the borrowers capacity and propensity to repay loans, making it imperative for lending institutions to assess the management of a firm for these traits before taking lending decisions. In determining the research group performance, *Janet & Andrew(2002)*, concluded that the study of high achieving, high impact research groups found a number of factors to be central in determining their success, namely strong leadership, finding, motivating and retaining talent, strategies of related diversification, strongly linked theory and practice, network connectedness. *Bard (2007)*, in an exploration of how the employee-organization relationship affected the linkage between perception of developmental human resource practices and employee outcomes, covered 64 local savings banks in Norway and the results showed that four indicators of the EOR (perceived organizational support, effective organizational commitment, and procedural and interactional justice) moderated the relationship between perception of developmental HR practices and individual work performance.

John, Stephen, Cherrie (2002), focused on a study to empirically examine the current purpose of performance appraisal in ten different countries and regions in Asia, North America and Latin America. The paper also examined how the respondents believed the purposes of appraisals should ideally be practiced. Research by *Debasish & Amir (2008)* evaluated the position and performance of Human Resources of National Thermal Power Corporation (NTPC) Limited with the help of its human resource accounting information both during the pre and post liberalization periods and made a comparison between them. The study revealed that the company achieved a very high profile in the performance of its human organization in the post liberalization period, combating efficiently in the process all the obstacles that emanate as a result of liberalization, globalization and competitiveness. Although there was a declining trend in the relative importance of non executives in the company, an overall increasing trend in the ratio of HR to TR indicating the growth importance of HR on the part of the company was noticed in the post liberalization period under study. The overall performance of HR of NTOC Ltd. improved notably in the post liberalization period. A better consistency in the performance of HR of the company during the post liberalization period was also revealed in the study. The company was able to make its human organization stronger with full of skilled personal having higher productivity in the post liberalization period. The net outcome of all the performance measures used in this study confirmed a remarkable improvement in the performance of HR during the post liberalization period. The operating performance of the company was positively as well as significantly associated with the performances of its executives and non executives only in the post liberalization period. Another notable outcome of the study was that the joint influence of the performances of executives and non executives of the company on its operating performance was also very significant during the post liberalization period. In fact the company was able to achieve a very high profile regarding the performance of its human organization in the post liberalization era by combating quite efficiently all the obstacles emanating from liberalization, globalization and competitiveness.

A study on 360 degree Performance Appraisal System in Reliance Life Insurance, Udumalpet by *Vijaya & Umamaheshwari (2009)*, was carried with the objective to study the effectiveness of an employee's performance based on 360 degree performance appraisal system in Reliance Life Insurance. To evaluate the performance of the employee from the perspective of immediate supervisors and from the perspective of customers and to suggest suitable measures to improve the performance of the individuals based on the analysis. The researcher concluded that the 360 degree performance appraisal system helped to identify training needs, performance of employees, and determination of rewards/incentives and steps to promote communication from the perspective of the employees themselves, superiors and from the customers.

Kumar (2000), focused on the neglected activity i.e. Performance Appraisal of Senior level Managers. The findings indicated that the managers of the companies were by and large aware of the importance of HRD. However, most of them did not fully understand the techniques of HRD and the linkage between HRD and performance appraisal. Of the 20 companies taken up, one did not favour the study in its organization on the ground that it did not have any formal appraisal system while other company did not want the study in the organization due to obvious reasons. Of the balance 18 companies studied, 8 did not have any formal appraisal system. Among the remaining 10, in 8 companies the form for assessing all level managers was the same. Only two companies were having somewhat satisfactory to good systems for managers including a separate format for assessing the

performance of senior level managers. It was observed that there was lack of Performance planning, counselling and feedback, non inclusion of potential appraisal, inappropriate factors of assessment, too many factors of assessment, lack of transparency, limited use of appropriate data, non availability of appraisal manual, non consideration of the views of the executives in designing a suitable appraisal system, even the management of some of the companies were not satisfied with the operation of the system but were continuing with the same system. A fact which emerged clearly was that the proper and formal appraisal of performance of senior level managers was practically non-existent.

A study of power sector in India- Benchmarking a tool for enhanced performance, *Geetika & Neeraj (2007)*, concluded that benchmarking was essentially an extrapolation of human tendency. It was a natural phenomenon and therefore could be implemented with little effort. Also the virtuous circle of continuum performance introduced earlier in the study actually works. Benchmarking was an important tool to track performance over a period of time and to identify realistic targets and priorities for action. *Steven, Guy, Arindam (2010)* investigated the relationship between firm strategy and the use of performance measures in executive compensation. Analysis showed that there was an increased emphasis on sales in the determination of executive compensation for firms pursuing a cost leadership strategy, which sought to achieve their competitive advantage through low price and high volume. In contrast, there was a decreased emphasis on accounting measures in firms pursuing a differentiation strategy, which required investments in brand recognition and innovative products, investments that were subject to unfavourable accounting treatment. These results indicated that compensation committees linked executive rewards to firm strategy.

Sanjay (2010), studied the performance evaluation of selected top ten mutual funds in India on the basis of last one year's return. *Manish & Smita(2007)*, study investigated the relationship between facets of the performance appraisal and perceptions of procedural fairness of human resource practices with perceptions of the effectiveness of the performance appraisal system among a sample of 250 managerial personnel from three large scale manufacturing organizations. Results revealed that all the system and process facets of performance appraisal had positive relationship with the perceived effectiveness of the performance appraisal system (PAS) with the exception of "system complexity". Results also revealed that while the procedural fairness of the human resource practices suppressed the negative relationship between "system complexity" and effectiveness of the PAS, it emerged as a moderator of the positive prediction of effectiveness of the PAS by 'system openness'. Findings revealed that the process facets had a strong relationship with the perceived relationship with the perceived effectiveness of the PAS. All the three dimensions of the process, that is, multiple inputs, session planning and session feedback were significantly related with perceived effectiveness of the PAS. As expected the three process facets were also found to be strongly related with procedural fairness of the organizations human resource practices. Another study examined the relationship between the organization's performance appraisal 'system' and 'process' facets, members' perception of the effectiveness of the performance appraisal system and organizational commitment. The study was conducted on a sample of managerial personnel in a manufacturing organization in North India. Results showed that the process facets, namely session planning and multiple inputs, positively predicted the perceived effectiveness of the performance appraisal system. Results further showed that the process facets-'multiple inputs', 'session feedback' and 'session planning' were positive predictors of 'affective' organizational commitment while 'continuance organizational commitment was positively predicted by both the system facets-'system commitment' and the process facets-'session planning' and 'session feedback'. Results also showed that with the exception of 'system complexity', all the system as well as process facets were significantly correlated with affective commitment and continuance commitment. Only the process facets emerged as significant predictors of affective organizational commitment. Stepwise regression analysis of performance appraisal facets with continuance organizational commitment showed that the process facets-session feedback and session planning positively predicted 62% and 3% variance in continuance organizational commitment.

An appraisal of Performance appraisal in the Indian Scenario, *Amit (2006)*, discussed the key limitations of present appraisal system in Indian Industry. It further discussed the results of a study conducted in a large construction organization regarding the effectiveness of present performance appraisal system. The results were compelling, revealing extreme dissatisfaction among employees regarding the system. *Gary, Yvonne, Rafik (2000)*, examined factors associated with the board's

adoption of a formal process for evaluating the performance of the corporation's chief executive officer (CEO) The sample was drawn from the hospital industry. The study revealed that more competitive the market, the more likely the hospital boards were to adopt a formal CEO performance evaluation process. Managerial perceptions of various dimensions of an effective Performance Appraisal System (PAS) were investigated through a survey conducted in a North Indian synthetic yarn-producing organization. The data revealed that managers perception and expectations of issues related to the nature of the appraisal process, guidance and counseling sessions, assessment of performance in key result areas, and ratings by superiors and feedback presentation. These varied greatly at all levels, thereby posing challenges for the HRD department in designing an effective performance appraisal system. Managers also perceived the existing PAS to be only moderately effective.

Vikash & Abha(2002), designed a Performance Appraisal System for shop floor employees. The study was conducted in Tata Bearing Division at Kharagpur, West Bengal from July 2001 to December 2001. Study on employees opinion on 360 degree feedback system, *Gunavathy & Vidhya(2005)*, the study served as a dip stick survey to understand the employees views on the 360 appraisal system prevalent in the organization from the following perspectives, understanding of the 360 degree feedback system, views on the 360 degree feedback system prevalent in the organization, perceived merits of the 360 degree feedback system, perceived de merits of the 360 degree feedback system. A study on Organizational Climate vis-à-vis organizations, *Souvik (2005)*, developed a scale to measure performance appraisal climate (PAC). Overall organizational climate had a direct bearing on the present utility of performance appraisal. Majority of the extension personnel perceived existing PAC and organizational climate as below average or poor. Out of twelve PAC dimensions, nine dimensions were found to be significantly related to overall organizational climate of State Department of Agriculture while seven dimensions were significantly related to overall organizational climate of the Non Government Organizations (NGOs). The PAC dimensions namely participation, performance standards, superior-subordinate relationship and employee acceptance together constituted more that 50% of total variation in overall organizational climate.

June, Cynthia & Martin (2000), carried out the research on Effects of training method and learning style on Cross Cultural training outcomes and the results indicated that cross cultural attitude and trainee reaction were more positive when the training method matched trainees learning styles than when it did not. *Nguyen, Truong & Dirk (2010)*, the analysis indicated that the relationship between training and firm performance might be mediated by employee knowledge and attitude. Furthermore, capital investment or organizational strategy does moderate the training performance relationship. *Stephen, Craig, Barbara & Kenneth (2006)*, surveyed that the level of perceived training needs varied dramatically by job category and health department type. When comparing aggregate training needs, public health workers with greater day to day contact indicated a greater need for training than their peers who did not such as those working in administrative positions. A series of studies was undertaken to establish the training and development needs of nurses and midwives working within a variety of contexts in Indonesia, with the ultimate aim of enhancing care provision within these domains. The majority of health delivery was undertaken by nurses and midwives educated to secondary school level only, with the higher, more specialist qualifications being delivered by educators with restricted clinical experience. This, together with a context of high demands for health care, restricted resources and limited equipment and facilities, inevitably presents a real challenge for the Indonesian health system. The vast majority of nurses and midwives (60%) had inadequate training and preparation for the role, which created the potential for substandard care delivery. An intensive in-service training program improved the performance of village midwives in three out of five key skill areas. A comparison of the two ratings on any item provided an assessment of the training need associated with it, in that tasks considered to be highly crucial but not well-performed had a training implication, while those items for which criticality and performance were rated similarly had little training requirement. The occupational roles of the midwives varied significantly by province, indicating regional service delivery distinctions, but very little difference in the roles of hospital and community midwives.

One of the research on the importance of human resources management in health care: a global revealed that the relationship between human resources management and health care was extremely complex, particularly when examined from a global perspective. The research and analysis

have indicated that several key questions must be addressed and that human resources management could and must play an essential role in health care sector reform. The various functions of human resources management in health care systems of Canada, the United States of America, Germany and various developing countries have been briefly examined. The goals and motivations of the main stakeholders in the Canadian health care system, including provincial governments, the federal government, physicians, nurses and allied health care professionals, have been reviewed. The possibility of a major change in the structure of Canadian health care was also explored, specifically with regard to the creation of a two-tier system. A comparative analysis between Japan and Malaysia-Education and Training in the Auto Manufacturing Industry. A comparison of education and training practices of the Japanese companies with those of locally owned firms in the auto manufacturing industry revealed a number of similarities and differences. The degree of similarity was high in the fields of on-the-job training and off the job training. It was low in such areas as self-development and intensity of training. The main difference here between the two countries was that Japanese employees were given constant and consistent training throughout their career. Malaysian enterprises, however, carried out little training for its employees, at least not until recently. A comparison of education and training practices of the Japanese companies with those of locally owned firms in the auto manufacturing industry revealed a number of similarities and differences. The degree of similarity was high in the fields of on-the-job training and off the job training. It was low in such areas as self-development and intensity of training. The main difference between the two countries was that Japanese employees were given constant and consistent training throughout their career. Malaysian enterprises, however, carried out little training for its employees, at least not until recently.

Sanjeev & Neha (2005), conducted a study of effectiveness of training as a tool for learning and knowledge in BPO sector and the result concluded that training was effective in true sense only when it was evaluated in terms of actual learning, knowledge increase and on the job confidence thereby meaning that the training evaluation must go beyond just evaluating the immediate reaction of the trainees to the training or the effectiveness of the training materials and the methods used. *Altarawneh,(2009)*, examined the training and development evaluation in Jordanian Banking organizations, the study findings revealed, that although the majority of the organizations evaluated their training program, usually there was an absence of systematic and effective procedures for evaluation, and most of the organizations relied on external providers to evaluate their T&D program. This study showed that the T&D evaluation stage in the Jordanian banking industry was not in a better situation than in other Arab organizations. Also, there was a lack of sustainable evidence that the T&D evaluation stage in the Jordanian banking industry was likely to be conducted differently than when the activity was undertaken in other Arab organizations.

Mohamed (2004), measured training effectiveness at training institutes and the results indicated significant positive reactions to the training program. They also indicated significant increase in trainee's skills and knowledge as a result of the training. A comparative study of education officers training in two different cultures by *Hood (2003)* concluded that the effectiveness of staff development was dependent upon appropriate organizational behaviour to support externally provided programs. When there was positive correlation between the commitments of immediate superiors to their subordinate staff development, a collaborative "no-blame" culture was operating within that organization. A multi-country comparative analysis on benchmarking training and development practices by *Ellen, Colette, Kevin & Michael (2002)*, indicated that there was no universal practices across all countries studied, they didn't indicate significant similarities in practices within country clusters. The common practices founded within these clusters were believed to be influenced by cultural values and industry trends. Similarity of training and development across country clusters was indeed influenced by industry trends and cultural characteristics.

Assessment of management training needs of agricultural research managers by *Manikandan & Anwer (2008)*, revealed that an analysis of the major requirements for training managers and administrators in agriculture suggested that a training program should aim at enhancing their capability to understand specific situations, to orient action, and to use effectively a problem solving approach. A research on training needs of extension specialists, concluded that the training needs of the ESs pertaining to all major areas showed be fulfilled by imparting in-service/refresher training, after every two years, particularly in the deficient items. A research carried out on training program for self employment-trainee perception on its impact concluded that it could serve as a stepping stone

for her to start a small scale business. The training also gave them a feeling of self esteem and confidence in their innate abilities. A study on Executives Training Philosophies, concluded that “Pursuit of Learning”: this particular philosophy reflected the attitude that training could improve an organization by showing people how to pull together. The second philosophy was “justification” and the underlying attitude with this philosophy was that people would eventually recognize the contributions of training; they considered that trainer was like a teacher who battles against ignorance. The third dominant philosophy was “achievements integration” which reflected the attitude that organizations were changing and training teaches people how to cope with change. The least preferred training philosophy was “escapism” which reflected the underlying attitude that there was no real future for trainers. The present study revealed that the executives strongly believed that continuous learning is vital and would help in improving their performance in organizations.

Impact of training programs in HRD activities in PSU, a study by *Chimum (2011)*, revealed that the employees were aware of the training facilities offered by the company. It also explained that training had impact on workers attitude, which meant that employees felt to work for the benefit of the organization if they found training for their benefit. A study in Jindal Steel and power Ltd by *Suman(2008)*, on training and development need analysis for ushering change revealed that training was needed for a number of reasons like opportunities, strengths, new directions, problems, impending change and career development. Training needs identification and evaluation, a study of Indian organizations by *Prarthana & Pooja(2007)*, revealed that all the organizations studied had a fairly progressive HR set up and a training set up and seemed to have realized the importance of training for employees and the organization. Employment status of displaced workers: an effect of training, a case study by *Tulika (2006)*, showed that the types of training which had been given to the displaced workers was not much effective for the workers to get the appropriate job. *Bharat, Samanta, Sandhya & Vijender (2006)*, Management training needs of Agricultural scientists of Indian Council of Agricultural Research showed that the prominent training needs pertained to IT and computer applications, statistical application for data analysis, communication and presentation skills, technical and scientific writing, research project management, evaluation and impact assessment. Evaluation of executive training at NLC Ltd-a case study by *Selvam & Panchalan (2003)*, revealed that training program of the respondent’s organization was generally effective. *Anuradha (2003)*, studied the designing of a human resource development training for a development organization. Burnout and training satisfaction of medical residents in Greece-a study by *Pavlos, Nikolaos & Athanasios(2010)*, study provided data on the prevalence of burnout syndrome among Greek residents and their dissatisfaction with residency training and job insecurity, both of which were associated with burnout subscales. The present study indicated that the gradual limitation of working hours would be a short-term measure that might not significantly alleviate resident burnout levels if it is not supplemented by concurrent long-term reforms of residency training in combination with novel patterns of care management, stress reduction programs and other systemic interventions. Relationship between training and learning was seen by *Elena (2001)*. *Scott, David, James & Mark (2000)*, studied the selection and training for integrated manufacturing, the moderating effects of job characteristics and the findings indicated that aspects of integrated manufacturing were positively associated with higher levels of selection and training for technical and problem solving skills.

Kailash, Sunita & Prasad, evaluated training effectiveness and customer satisfaction in Tata Steel and the results revealed that organizations must pay attention to the functioning of in-house training establishment in order to make them more effective so that the organizational as well as individual objectives are fulfilled. The results showed a decline in the satisfaction level of participants and their immediate supervisors. Training effectiveness and transfer of learning on the job, was also moderate. Training had its impact to improve performance but participants felt that training was not related to career advancements. *Singh & Joe (2002)*, undertook a study to assess how private club managers perceived the relative effectiveness of alternative training methods to attain specific types of training objectives. It indicated that one to one training was the preferred method to attain all objectives except interpersonal skill and development. Executive training and development was studied by *Savita (2002)*. A study on training motivation and participation *Patil & Meenakshi (2005)*, was conducted to identify factors that influenced the training participation of employees. Training as an effective HRD technique in banking sector-an opinion survey by *Raju (2005)*, was conducted on the various aspects of training viz Management attitude, selection process, quality of training, impact

on individuals, impact on productivity and post training assessment in different categories of banks in Coimbatore.

Sample Organizations At A Glance

Since hospitals are a widespread and an important sector of the society and this sector continuously deals with both, its own human resource and humans from outside the organization, to be managed and handled tactfully. So the study on this sector is important. The sample organizations include; 1) SKIMS Soura Srinagar & 2) SMHS Hospital Srinagar.

Sher-i-Kashmir Institute of Medical Sciences Srinagar. The Institute was partially commissioned on 5th December 1982. Sher-i-Kashmir Institute of Medical Sciences is a post graduate Institute for training, research and patient care. With this objective, various committees appointed by the Government of Jammu and Kashmir identified the specialties in which postgraduate and post doctoral courses would be undertaken. By an act of Legislature on 19th August 1983, Institute of Medical Sciences was granted a deemed University status. Sher-i-Kashmir Institute of Medical Sciences was conceived with the following objectives: To provide facilities of specialized medical care and particularly develop super specialties that would provide tertiary health-care. To provide need oriented education in medical sciences and clinical research. To develop a referral linkage between the primary, secondary and tertiary health-care Institutions of the State to achieve an optimum health delivery system. The State Government under the 5th and 6th plan grants earmarked for the Institute, as the planning commission approved of it as a plan project, provided the funds for construction and equipping the Institute.

Shri Maharaja Hari Singh (S.M.H.S) Hospital, Srinagar. The S.M.H.S Hospital was established in 1948 at Karan Nagar in the heart of Srinagar city as a Government-owned General Hospital to provide patient care to needy patients. Primary and secondary medical care in the major specialties of Medicine, Surgery and Gynaecology was made available to the general public on OPD and Inpatient basis. In 1959, after the inauguration of Government Medical College on the same campus, S.M.H.S hospital became a major teaching institution imparting undergraduate and postgraduate medical training in all allied specialties of Medicine, Surgery and Gynaecology. In the 1970's, as medical services became more specialized and differentiated, separate hospitals were established for specialties like Obstetrics and Gynecology and Orthopaedics etc. The shifting of the concerned departments coincided with the expansion and differentiation of the major specialties of Medicine and Surgery. Currently patient care services are offered in the specialties of Medicine, Surgery, Anaesthesiology, Radio diagnosis and Imaging, Ophthalmology, ENT, Dermatology and Radiotherapy. Facilities also exist for the Super specialties of Cardiology, Gastroenterology, Neurology, Urology, Cardiothoracic surgery and Paediatric surgery. These are supported by well organized and fully equipped diagnostic, therapeutic and support service departments. There is a separate Laboratory services block which undertakes routine, specialized and super specialized investigation.

Objectives Of The Present Study

- 1) To study the existing status of HRD Mechanisms in the sample study organizations i.e.(a)SKIMS Soura Srinagar (b)SMHS Srinagar,
- 2) To analyze the impact of existing HRD Mechanisms in the sample study hospitals,
- 3) To examine the perceptual difference in the opinion of Medical, paramedical and ministerial/supporting staff and
- 4) To draw conclusions and to provide result oriented guidelines and suggestions to the sample study hospitals for improvement of their existing HRD Mechanisms.

Hypothesis

In consonance to the above objectives, the hypotheses formulated for the present research are as under:

1. HRD Climate in the sample study organizations is not satisfactory.
2. HRD Mechanisms in the hospitals is not satisfactory.
3. There is a difference in the perception of medical and Para medical staff towards HRD Mechanisms in the two organizations.

	Staff	(A) SKIMS			(B) SMHS			A+B	
		Total HR	Served	Responded	Total HR	Served	Responded	Total Served (A+B)	Total Received
1	Medical	766	50	50	181	50	50	100	100
2	Para Medical	1519	66	54	512	75	71	141	125

Research Approach And Design

The tool used for obtaining the information was a 'Structured Non Disguised Questionnaire', a questionnaire was designed keeping in view both major and minor objectives of study. A close format questionnaire was used. Closed format questions offer many advantages in time and money. Questionnaire was used to measure the organizational ethos. This questionnaire is proposed by Dr. Udai Pareek. It consisted of two Sections and 39 statements. Sec "A" consisted of 38 statements, the rating for each to be done on a scale from 5 to 1, (5)=Almost always true, (4)=Mostly true, (3)=Sometimes true, (2)=Rarely true, (1)=Not at all true. Sec "B" sought for suggestions and name, gender, age, pay scale, designation of the respondent.

Simple Random Sampling Method (SRSM) was used to cover employees from Medical, Administrative, Supportive and Technical cadres in the sample selected organizations. A sample of 100-125 was targeted from each organization covering -**Doctors** including HOD's, Professors, Associate Professors, Asst Professors, Lecturers-**Officers** including Administrative Section, Materials Management, Library etc-**Engineers** including Civil, Electric, Mechanic, Architect-**Nurses** covering Superintendents, Nursing aids-**Others** covering Operators, Attendants, Technicians, Technologists, Drivers, Gardeners and other fourth class employees. Thus the total sample for the study was 225.

Questionnaire Served and Response Rate:

Data Analysis And Interpretation

Testing of hypothesis (H1)

TABLE 1: Existing status of HRD Climate in the sample selected organizations (No. 225).

St · N o	Statements	SKIMS (No. 104)			SMHS (No. 121)		
		M.S	S.D	%age	M.S	S.D	%age
1	The top management in health care sector goes out of its way to make sure that employees enjoy their work.	2.57	1.172	39.25	3.01	1.165	50.25
2	The top management on this organization believes the human resources are an extremely important resource and that they have to be treated more humanly.	3.34	1.252	58.5	3.50	1.089	62.5
3	Development of the subordinates is seen as an important part of their job by the managers/officers in health care sector.	3.08	1.243	52	3.07	1.148	51.75
4	The personnel policies in this organization facilitate employee development.	3.05	1.257	51.25	3.28	1.192	57
5	The top management in health care sector is willing to invest a considerable part of their time and other resources to ensure the development of employees.	2.60	1.153	40	2.93	1.270	48.25
6	Senior officers/executives in this organization take active interest in their juniors and help them to learn their job.	2.94	0.974	48.5	3.23	1.230	55.75
7	People in the health care sector lacking confidence in doing their job are helped to acquire competence rather than being left unattended.	2.84	1.208	46	3.09	1.072	52.25

8	Managers in this organization believe that employee's behavior can be changed and people can be developed at any stage of their life.	2.7 7	1.192	44.25	3.1 3	1.008	53.2 5
9	People in this organization are helpful to each other.	3.2 1	1.076	55.25	3.4 7	1.148	61.7 5
10	Employees in the health care sector are very informal and do not hesitate to discuss their personal problems with their supervisors.	2.9 3	1.225	48.25	2.9 4	1.227	48.5
11	The psychological climate of the health care sector is very conducive for any employee interested in developing himself by acquiring new knowledge and skills.	2.9 2	1.086	48	2.9 4	1.280	48.5
12	Seniors guide their juniors and prepare them for future responsibilities/roles that they are likely to take up.	3.2 2	1.343	55.5	3.3 7	1.253	59.2 5
13	The top management in the health care sector makes efforts to identify and utilize the potential of employees.	3.0 8	1.252	52	2.9 6	1.158	49
14	Promotion decisions in this organization are based on the suitability of the promotee rather than on favoritism.	3.0 2	1.393	50.5	3.1 2	1.345	53
15	There are mechanisms in this organization to reward any good work done or any contribution made by employees.	2.4 9	1.344	37.25	2.8 2	1.265	45.5
16	When an employee in the health care sector does good work his supervising officers take special care to appreciate it.	2.7 6	1.326	44	3.1 3	1.271	53.2 5
17	Performance Appraisal reports in this organization are based on objective assessment and adequate information and not on favoritism.	3.1 5	1.221	53.75	3.1 6	1.126	54
18	People in health care sector do not have any fixed mental impressions about each other.	3.0 4	1.131	51	3.1 7	1.267	54.2 5
19	Employees in this organization are encouraged to experiment with new methods and try out creative ideas.	2.8 6	1.295	46.5	2.6 5	1.346	41.2 5
20	When an employee in health care sector makes a mistake, his supervisors treat him with understanding and help him to learn from such mistakes rather than punish him or discourage him	2.9 2	1.146	48	2.9 5	1.161	48.7 5
21	Weaknesses of employees in this organization are communicated to them in a non threatening way.	2.8 9	1.088	26.89	2.9 0	1.261	47.5
22	When behaviour feedback is given to employees in health care sector, they take it seriously and use it for development	3.1 4	1.194	53.5	3.1 1	1.175	52.7 5
23	Employees in this organization take pains to find out their strengths weaknesses from their officers and colleagues.	2.8 8	1.275	47	2.9 6	1.158	49
24	When employees in health care sector are sponsored for training, they take it seriously and try to learn from the programs they attend.	3.6 4	1.131	66	3.7 9	1.156	69.7 5
25	Employees in this organization when returning from training programs are given opportunities to tryout what they have learnt	3.3 0	1.069	57.5	3.3 1	1.176	57.7 5
26	Employees are sponsored for training programs on the basis of genuine training needs in health care sector.	3.6 5	1.022	66.25	3.5 7	1.109	64.2 5
27	People trust each other in this organization.	2.9 4	1.148	48.5	3.2 0	1.100	55
28	Employees in health care sector are not afraid to discuss or express their feelings with their supervisors.	3.1 1	1.182	52.75	3.1 2	1.127	53
29	Employees in health care sector are not afraid to discuss or express their feelings with their subordinates.	3.0 2	1.132	50.5	3.1 2	1.112	53
30	Employees in health care sector are encouraged to take initiative and do things on their own without having to wait	2.4 1	1.204	35.25	2.7 2	1.097	43

	for instructions from their supervisors.						
31	Delegation of authority to encourage juniors to develop and handle higher responsibilities is quiet common in this organization.	2.58	1.103	39.5	3.06	0.951	51.5
32	When seniors in health care sector delegate authority to juniors use it as an opportunity for development.	3.30	0.954	57.5	3.40	0.988	60
33	Team spirit is of high order in this organization.	3.32	1.264	58	3.60	1.172	65
34	When problems arise in health care sector, people discuss these problems openly and try to solve them rather than keep accusing each other behind their backs.	2.84	1.167	46	3.07	1.283	51.75
35	Career opportunities are pointed out by juniors to senior officers in this organization.	2.81	1.175	45.25	2.72	1.097	43
36	The health care sector’s future plans are made known to the managerial staff to help them to develop their juniors and prepare them for future.	2.89	1.079	47.25	3.17	1.113	54.25
37	This organization ensures employee’s welfare to such an extent that the employees can save a lot of their mental energy for work purposes.	2.51	1.106	37.75	2.69	1.023	42.25
38	Job rotation in health care sector facilitates employee development.	3.57	1.260	64.25	3.25	1.362	56.25

Notes

1. Scoring Scale: Almost always true=(5),mostly true=(4),sometimes true=(3),rarely true=(2),not at all true=(1)
2. M.S=Mean Score
3. S.D=standard deviation
4. %=percentage to mean score.

Figure 1: Statement wise existing status of HRD Climate in sample study organizations

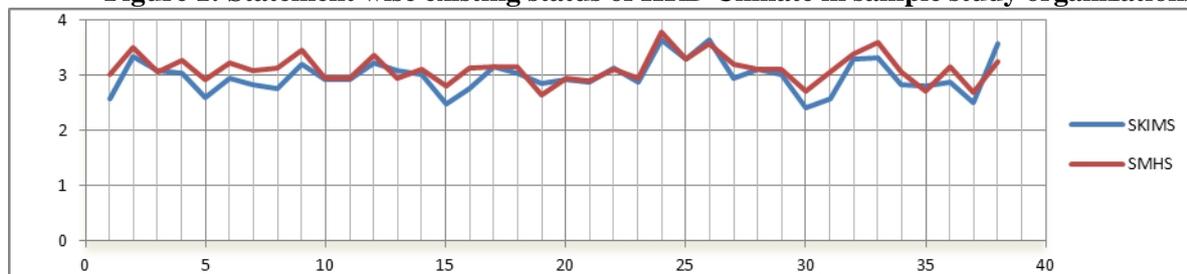


Table 1 shows the existing status of HRD Climate in SKIMS and SMHS. It is revealed from the table that HRD Climate in the two organizations is not satisfactory, though the existing Climate in the organizations is average. The overall mean score of SKIMS is 2.989(49.73%), while as the mean score of Climate in SMHS is 3.123(53.078%) which is comparatively greater than that of SKIMS. Table 1 shows that SKIMS recorded highest mean values of 3.64(66%), 3.65(66.25%), and 3.57(64.25%) against statements 24, 26 and 38 respectively i.e. “When employees in health care sector are sponsored for training, they take it seriously and try to learn from the programs they attend”, “Employees are sponsored for training programs on the basis of genuine training needs in health care sector” and “Job rotation in health care sector facilitates employee development”. On the other hand SMHS, recorded the highest mean values of 3.79(69.75%), 3.57(64.25%) and 3.60(65%) against statements 24, 26 and 33 respectively i.e. “When employees in health care sector are sponsored for training, they take it seriously and try to learn from the programs they attend”, “Employees are sponsored for training programs on the basis of genuine training needs in health care sector” and “Team spirit is of high order in this organization”. The table also reflects that the lowest mean value of SKIMS recorded against statement No 21 scored a mean score of 2.89(26.89%) viz,

“Weaknesses of employees in this organization are communicated to them in a non threatening way”. The rest of the statements scored in SKIMS show average response. Some of the statements show approximately the same mean scores for both the organizations i.e. statement No. 3, “Development of the subordinates is seen as an important part of their job by the managers/officers in health care sector” scored mean score of > 50%(3.08 & 3.07), statement No. 10, “Employees in the health care sector are very informal and do not hesitate to discuss their personal problems with their supervisors” scored mean score of >48%(2.93 & 2.94), statement No. 11, “The psychological climate of the health care sector is very conducive for any employee interested in developing himself by acquiring new knowledge and skills” scored mean score of >48%(2.92 & 2.94), statement No. 17, “Performance Appraisal reports in this organization are based on objective assessment and adequate information and not on favoritism” scored mean score of >53%(3.15 & 3.16), statement No. 20, “When an employee in health care sector makes a mistake, his supervisors treat him with understanding and help him to learn from such mistakes rather than punish him or discourage him” scored the score of >48%(2.99 & 2.95), statement No. 22, “When behavior feedback is given to employees in health care sector, they take it seriously and use it for development” scored the mean of >52%(3.14 & 3.11), a mean score of >57%(3.30 & 3.31) and >52%(3.11 & 3.12) was scored for statement No.25 and 28 i.e. “Employees in this organization when returning from training programs are given opportunities to tryout what they have learnt” and “Employees in health care sector are not afraid to discuss or express their feelings with their supervisors”. To test the level of significance, Z test was conducted and it was concluded from table 2, that Z calculated is less than Z tabulated so H1 i.e. “HRD Climate is not satisfactory in the sample study organizations”, is accepted at 5% level of significance.

TABLE 2: Existing status of HRD Climate in the sample study organizations with Z values.

St · n o	Statements	SKIMS (No. 104)		SMHS (No. 121)		M.S diff	Z value	Lev el Of Sign if
		M. S	S.D	M.S	S.D			
1	The top management in health care sector goes out of its way to make sure that employees enjoy their work.	2.57	1.172	3.01	1.165	-0.441	-2.822	0.05
2	The top management on this organization believes the human resources are an extremely important resource and that they have to be treated more humanly.	3.34	1.252	3.50	1.089	0.003	0.016	0.05
3	Development of the subordinates is seen as an important part of their job by the managers/officers in health care sector.	3.08	1.243	3.07	1.148	-0.168	-1.050	0.05
4	The personnel policies in this organization facilitate employee development.	3.05	1.257	3.28	1.192	-0.233	-1.425	0.05
5	The top management in health care sector is willing to invest a considerable part of their time and other resources to ensure the development of employees.	2.60	1.153	2.93	1.270	-0.338	-2.075	0.05
6	Senior officers/executives in this organization take active interest in their juniors and help them to learn their job.	2.94	0.974	3.23	1.230	-0.289	-1.966	0.05
7	People in the health care sector lacking confidence in doing their job are helped to acquire competence rather than being left unattended.	2.84	1.208	3.09	1.072	-0.254	-1.658	0.05
8	Managers in this organization believe that employee’s behavior can be changed and people can be developed at any stage of their life.	2.77	1.192	3.13	1.008	-0.363	-2.444	0.05
9	People in this organization are helpful to each other	3.21	1.076	3.47	1.148	-0.26	-1.740	0.05
	Employees in the health care sector are very informal	2.9	1.225	2.9	1.227	-	-0.58	

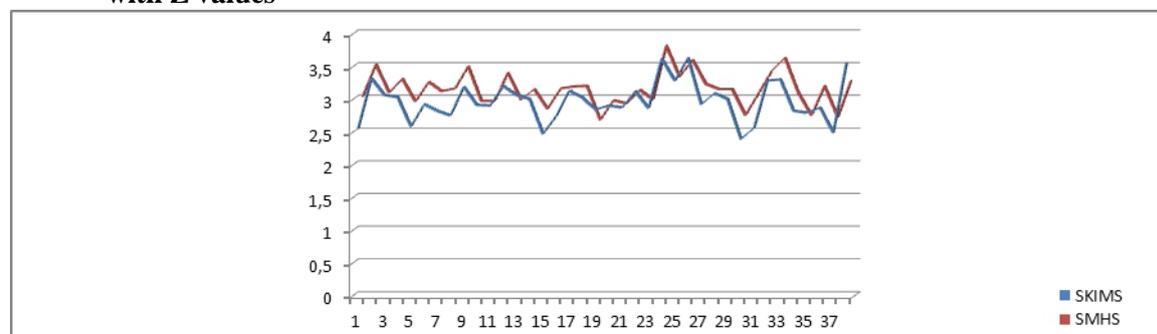
10	and do not hesitate to discuss their personal problems with their supervisors.	3		4		0.009		0.05
11	The psychological climate of the health care sector is very conducive for any employee interested in developing himself by acquiring new knowledge and skills.	2.9 2	1.086	2.9 4	1.280	- 0.019	- 0.121	0.05
12	Seniors guide their juniors and prepare them for future responsibilities/roles that they are likely to take up.	3.2 2	1.343	3.3 7	1.253	- 0.151	- 0.870	0.05
13	The top management in the health care sector makes efforts to identify and utilize the potential of employees.	3.0 8	1.252	2.9 6	1.158	- 0.118	- 0.736	0.10
14	Promotion decisions in this organization are based on the suitability of the promotee rather than on favoritism.	3.0 2	1.393	3.1 2	1.345	- 0.105	- 0.573	0.05
15	There are mechanisms in this organization to reward any good work done or any contribution made by employees.	2.4 9	1.344	2.8 2	1.265	- 0.328	- 1.883	0.05
16	When an employee in the health care sector does good work his supervising officers take special care to appreciate it.	2.7 6	1.326	3.1 3	1.271	- 0.373	- 2.149	0.05
17	Performance Appraisal reports in this organization are based on objective assessment and adequate information and not on favoritism.	3.1 5	1.221	3.1 6	1.126	- 0.003	- 0.020	0.05
18	People in health care sector do not have any fixed mental impressions about each other.	3.0 4	1.131	3.1 7	1.267	- 0.127	- 0.786	0.05
19	Employees in this organization are encouraged to experiment with new methods and try out creative ideas.	2.8 6	1.295	2.6 5	1.346	- 0.203	- 1.147	0.10
20	When an employee in health care sector makes a mistake, his supervisors treat him with understanding and help him to learn from such mistakes rather than punish him or discourage him.	2.9 2	1.146	2.9 5	1.161	- 0.027	- 0.177	0.05
21	Weaknesses of employees in this organization are communicated to them in a non threatening way.	2.8 9	1.088	2.9 0	1.261	- 0.007	- 0.040	0.05
22	When behavior feedback is given to employees in health care sector, they take it seriously and use it for development.	3.1 4	1.194	3.1 1	1.175	0.037	0.232	0.05
23	Employees in this organization take pains to find out their strengths weaknesses from their officers and colleagues.	2.8 8	1.275	2.9 6	1.158	- 0.084	- 0.516	0.05
24	When employees in health care sector are sponsored for training, they take it seriously and try to learn from the programs they attend.	3.6 4	1.131	3.7 9	1.156	- 0.141	- 0.920	0.05
25	Employees in this organization when returning from training programs are given opportunities to tryout what they have learnt.	3.3 0	1.069	3.3 1	1.176	- 0.016	- 0.106	0.05
26	Employees are sponsored for training programs on the basis of genuine training needs in health care sector.	3.6 5	1.022	3.5 7	1.109	0.084	0.584	0.10
27	People trust each other in this organization.	2.9 4	1.148	3.2 0	1.100	- 0.026	- 1.706	0.05

28	Employees in health care sector are not afraid to discuss or express their feelings with their supervisors.	3.11	1.182	3.12	1.127	-0.010	-0.064	0.05
29	Employees in health care sector are not afraid to discuss or express their feelings with their subordinates.	3.02	1.132	3.12	1.112	-0.096	-0.643	0.05
30	Employees in health care sector are encouraged to take initiative and do things on their own without having to wait for instructions from their supervisors.	2.41	1.204	2.72	1.097	-0.306	-1.991	0.05
31	Delegation of authority to encourage juniors to develop and handle higher responsibilities is quiet common in this organization.	2.58	1.103	3.06	0.951	-0.481	-3.472	0.05
32	When seniors in health care sector delegate authority to juniors use it as an opportunity for development.	3.30	0.954	3.40	0.988	-0.107	-0.822	0.05
33	Team spirit is of high order in this organization.	3.32	1.264	3.60	1.172	-0.286	-1.760	0.05
34	When problems arise in health care sector, people discuss these problems openly and try to solve them rather than keep accusing each other behind their backs.	2.84	1.167	3.07	1.283	-0.230	-1.395	0.05
35	Career opportunities are pointed out by juniors to senior officers in this organization.	2.81	1.175	2.72	1.097	0.089	0.585	0.10
36	The health care sector's future plans are made known to the managerial staff to help them to develop their juniors and prepare them for future.	2.89	1.079	3.17	1.113	-0.271	-1.847	0.05
37	This organization ensures employee's welfare to such an extent that the employees can save a lot of their mental energy for work purposes.	2.51	1.106	2.69	1.023	-0.185	-1.300	0.05
38	Job rotation in health care sector facilitates employee development.	3.57	1.260	3.25	1.362	-0.319	1.815	0.10

Notes

1. Scoring Scale : same as in table 1
2. M.S: mean score, S.D: standard deviation, M.S diff: difference between mean scores.
3. 0.05= statement is accepted at both 5% & 10% level of significance,0.10= statement is accepted at only 10% level of significance
4. H1 is accepted at 5% level of significance.

Figure 2: Statement wise existing status of HRD Climate in sample study organizations with Z values



Testing of hypothesis (H2)

Table 3: Existing status of HRD Mechanisms in the sample study organizations.

St · No	Statements	SKIMS (No. 104)			SMHS (No. 121)		
		M.S	S.D	%age	M.S	S.D	%age
13	The top management in the health care sector makes efforts to identify and utilize the potential of employees.	3.08	1.252	52	2.96	1.158	49
16	When an employee in the health care sector does good work his supervising officers take special care to appreciate it.	2.76	1.326	44	3.13	1.271	53.25
17	Performance Appraisal reports in this organization are based on objective assessment and adequate information and not on favoritism.	3.15	1.221	53.75	3.16	1.126	54
21	Weaknesses of employees in this organization are communicated to them in a non threatening way.	2.89	1.088	26.89	2.90	1.261	47.5
22	When behaviour feedback is given to employees in health care sector, they take it seriously and use it for development	3.14	1.194	53.5	3.11	1.175	52.75
23	Employees in this organization take pains to find out their strengths weaknesses from their officers and colleagues.	2.88	1.275	47	2.96	1.158	49
24	When employees in health care sector are sponsored for training, they take it seriously and try to learn from the programs they attend.	3.64	1.131	66	3.79	1.156	69.75
25	Employees in this organization when returning from training programs are given opportunities to tryout what they have learnt	3.30	1.069	57.5	3.31	1.176	57.75
26	Employees are sponsored for training programs on the basis of genuine training needs in health care sector.	3.65	1.022	66.25	3.57	1.109	64.25
38	Job rotation in health care sector facilitates employee development.	3.57	1.260	64.25	3.25	1.36	56.25

Notes

1. Scoring Scale : same as in table 1
2. M.S: mean score, S.D: standard deviation, M.S diff: difference between mean scores.

Figure 3: Statement wise existing status of HRD Mechanisms in the sample study organizations.

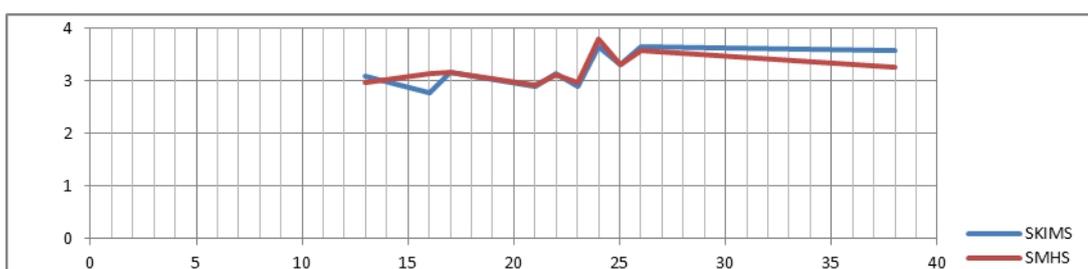


TABLE 4: Existing status of HRD Mechanisms in the sample study organizations with Z values.

St · no	Statements	SKIMS (No. 104)		SMHS (No. 121)		M.S diff	Z value	Signi fLev el
		M.S	S.D	M.S	S.D			
13	The top management in the health care sector makes efforts to identify and utilize the potential of employees.	3.08	1.252	2.96	1.158	0.12	0.736	0.10
16	When an employee in the health care sector does good work his supervising officers take special care	2.76	1.326	3.13	1.271	-0.3	-2.149	0.05

	to appreciate it.					7		
17	Performance Appraisal reports in this organization are based on objective assessment and adequate information and not on favoritism.	3.15	1.221	3.16	1.126	-0.01	-0.020	0.05
21	Weaknesses of employees in this organization are communicated to them in a non threatening way.	2.89	1.088	2.90	1.261	-0.01	-0.042	0.05
22	When behavior feedback is given to employees in health care sector, they take it seriously and use it for development	3.14	1.194	3.11	1.175	0.03	0.232	0.05
23	Employees in this organization take pains to find out their strengths weaknesses from their officers and colleagues.	2.88	1.275	2.96	1.158	-0.08	-0.516	0.05
24	When employees in health care sector are sponsored for training, they take it seriously and try to learn from the programs they attend.	3.64	1.131	3.79	1.156	0.15	-0.920	0.05
25	Employees in this organization when returning from training programs are given opportunities to tryout what they have learnt	3.30	1.069	3.31	1.176	-0.01	-0.106	0.05
26	Employees are sponsored for training programs on the basis of genuine training needs in health care sector.	3.65	1.022	3.57	1.109	0.08	0.584	0.10
38	Job rotation in health care sector facilitates employee development.	3.57	1.260	3.25	1.36	0.32	1.815	0.10
	Total	3.20		3.21				0.05

Notes

1. Scoring Scale : same as in table 1
2. M.S: mean score, S.D: standard deviation, M.S diff: difference between mean scores.
3. 0.05= statement is accepted at both 5% & 10% level of significance,0.10= statement is accepted at only 10% level of significance
4. H2 is accepted at 5% level of significance.

Figure 4: Statement wise existing status of HRD Mechanisms in sample study organizations with Z values

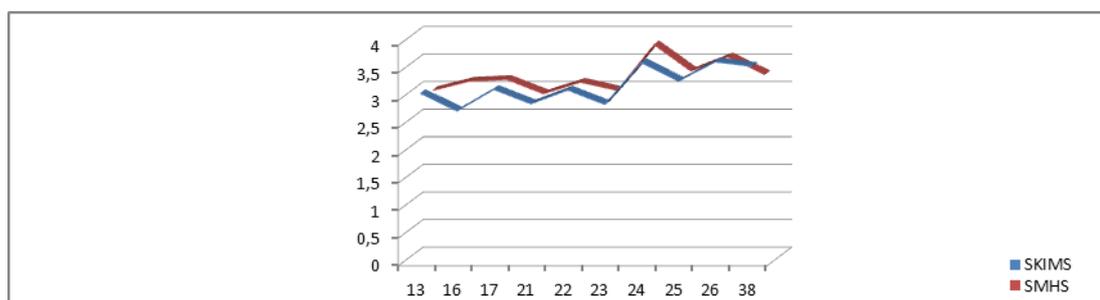


Table 3 and 4 reveals that the mean score for HRD Mechanisms in the sample study organizations falls within an average range of scale with mean > 3. While analysing T&D and Performance Appraisal system separately, it is found that the overall environment for T&D is satisfactory in the two organizations, with a mean score of 3.412(60.3%). Statement 26 i.e. “Employees are sponsored for training programs on the basis of genuine training needs in health care sector”, scores the largest mean score of 3.65(66.25%) for SKIMS and second largest score of 3.57(64.25%) for SMHS. Statement 24 i.e. “When employees in health care sector are sponsored for training, they take it seriously and try to learn from the programs they attend”, scores the highest mean score of 3.79(69.75%) for SMHS. Statement 38 scores the next highest mean score of 3.57(64.25%) and

3.25(56.25%) for SKIMS and SMHS i.e. “Job rotation in health care sector facilitates employee development”. Statement 13 scores the least mean score of 3.08 (52%) and 2.96(49%), i.e. “The top management in the health care sector makes efforts to identify and utilize the potential of employees”. On the other hand, Performance Appraisal System is found to be average in both the hospitals with a mean score of 3.06(50.15%). Statement 17, i.e. “Performance Appraisal reports in this organization are based on objective assessment and adequate information and not on favoritism”, scores the highest mean score of 3.15(53.75%) and 3.16(54) for PAS in sample study organizations. Statement 22 scores the next comparable mean score of 3.14(53.5%) and 3.11(52.75%) i.e. “When behavior feedback is given to employees in health care sector, they take it seriously and use it for development”. Least mean score of 2.76(44%) and 3.13(53.25%) is scored by statement 16, i.e. “When an employee in the health care sector does good work his supervising officers take special care to appreciate it”.

Comparing the two mechanisms, the data reveals that, though both the mechanisms are found to be average but Training and Development (3.412) is valued and practiced more than PAS (3.006). Comparing the scores of the two organizations, it is found that there lies a difference in the value of two mechanisms, with SKIMS scoring a comparatively higher score of 3.448() with respect to SMHS scoring a score of 3.376() for T&D. It reveals that SMHS does not encourage T&D to the extent SKIMS does. SMHS cores a higher score of 3.052() in comparison to SKIMS scoring 2.96() for PAS, revealing that PAS should be encouraged in SKIMS.

Using Z test for evaluating the significance level, it was inferred from the table 4, that HRD Mechanisms in SKIMS and SMHS falls under an average level. So H2 i.e. “HRD Mechanisms is dissatisfactory in the sample study organizations” is accepted at 5% level of significance.

Testing of hypothesis (H3)

Table 5: Perception of Medical and Para medical staff towards HRD Mechanisms.

St. No	SKIMS						SMHS					
	Medical(No.50)			Par Medical(No.54)			Medical(No.50)			Para Medical(No.71)		
	M.S	S.D	%age	M.S	S.D	%age	M.S	S.D	%age	M.S	S.D	%age
13	3.02	1.204	50.5	3.13	1.304	53.25	3.24	1.080	56	2.76	1.177	44
16	2.84	1.448	46	2.69	1.210	42.25	3.26	1.275	56.5	3.04	1.270	51
17	3.24	1.170	56	3.07	1.272	51.75	3.32	1.133	58	3.04	1.114	51
21	2.88	0.849	47	2.91	1.278	47.75	3.10	1.055	52.5	2.76	1.378	44
22	3.20	1.178	55	3.09	1.217	52.25	3.16	1.076	54	3.07	1.246	51.75
23	2.82	1.320	45.5	2.93	1.242	48.25	3.06	1.058	51.5	2.89	1.225	47.25
24	3.58	0.992	64.5	3.70	1.253	67.5	4.16	0.955	79	3.52	1.217	63
25	3.42	1.108	60.5	3.19	1.029	54.75	3.70	1.015	67.5	3.04	1.212	51
26	3.42	1.052	60.5	3.87	0.953	71.75	3.64	1.174	66	3.52	1.067	63
38	3.62	1.338	65.5	3.52	1.193	63	3.66	1.189	66.5	2.96	1.409	49

Figure 5: Statement wise existing status of HRD Mechanisms in sample study organizations.

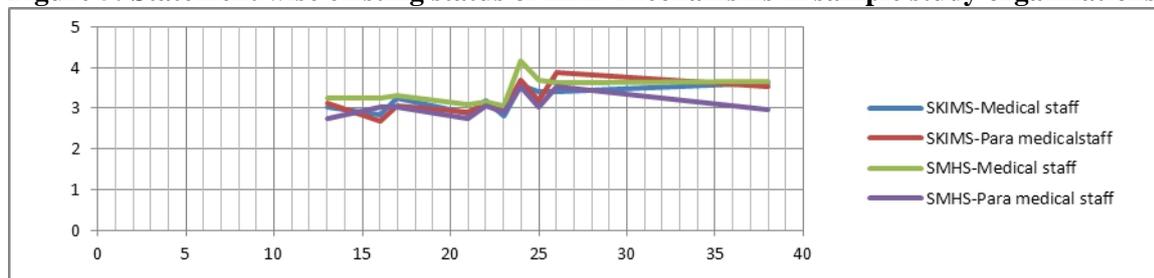


Table 5, represents the HRD Mechanisms, with respect to managerial and non managerial staff in the sample study organizations. The table represents that the perception of medical and Para- medical staff towards HRD Mechanisms, falls in the average range of scale. The overall mean values of medical group in SKIMS are 3.204(55.1%) as revealed from table 3, which falls under average range of scale. The table also reveals that the overall mean values of Para-medical staff in SKIMS also falls under an average level with the mean score of 3.21(55.25%). The mean score of medical staff in SMHS is 3.43(60.75%) which falls in a satisfactory range of scale and the inverse applies for the Para-medical staff of SMHS with a mean score of 3.06(51.5%), which indicates average level of HRD

Mechanisms. The medical and Para-medical staff of SKIMS scored less value than the medical and par- medical staff of SMHS. Medical staff of SKIMS scored the highest mean score of 3.62(65.5%) for statement 38 i.e. “Job rotation in health care sector facilitates employee development” and the medical staff of SMHS scored the highest score of 4.16(79%) for statement 24 i.e. “When employees in health care sector are sponsored for training, they take it seriously and try to learn from the programs they attend” and the same statement scored the highest score of 3.52(63%) for Para medical staff of SMHS. Statement 26 i.e. “Employees are sponsored for training programs on the basis of genuine training needs in health care sector” scored the highest score of 3.52(63%) and 3.87(71.75%) for Para medical staff in SMHS and SKIMS respectively. On the other hand, least score of 2.82(45.5%) and 3.06(51.5%) was scored by medical staff of SKIMS and SMHS respectively for statement 23 i.e. “Employees in this organization take pains to find out their strengths weaknesses from their officers and colleagues”. Mean score of 2.69(42.25%) was scored by Para-medical staff of SKIMS for statement 16 i.e “When an employee in the health care sector does good work his supervising officers take special care to appreciate it” and the Para medical staff of SMHS scored the least score of 2.76(44%) for statements 13 and 21 i.e. “The top management in the health care sector makes efforts to identify and utilize the potential of employees” and “Weaknesses of employees in this organization are communicated to them in a non threatening way”.

TABLE 6: Existing status of HRD Mechanisms for medical staff in the sample study organizations with Z values.

St no	Statements	SKIMS (No. 104)		SMHS (No. 121)		M.S diff	Z value	Leve l Of Signi f
		M.S	S.D	M.S	S.D			
13	The top management in the health care sector makes efforts to identify and utilize the potential of employees.	3.02	1.204	3.24	1.080	-0.22	-0.962	0.05
16	When an employee in the health care sector does good work his supervising officers take special care to appreciate it.	2.84	1.448	3.26	1.275	0.42	-1.540	0.05
17	Performance Appraisal reports in this organization are based on objective assessment and adequate information and not on favoritism.	3.24	1.170	3.32	1.133	0.08	-0.347	0.05
21	Weaknesses of employees in this organization are communicated to them in a non threatening way.	2.88	0.849	3.10	1.055	0.22	-1.149	0.05
22	When behaviour feedback is given to employees in health care sector, they take it seriously and use it for development	3.20	1.178	3.16	1.076	0.04	0.177	0.05
23	Employees in this organization take pains to find out their strengths weaknesses from their officers and colleagues.	2.82	1.320	3.06	1.058	0.24	-1.003	0.05
24	When employees in health care sector are sponsored for training, they take it seriously and try to learn from the programs they attend.	3.58	0.992	4.16	0.955	0.58	-2.979	0.05
25	Employees in this organization when returning from training programs are given opportunities to tryout what they have learnt	3.42	1.108	3.70	1.015	0.28	-1.317	0.05
26	Employees are sponsored for training programs on the basis of genuine training needs in health care sector.	3.42	1.052	3.64	1.174	0.22	-0.987	0.05

38	Job rotation in health care sector facilitates employee development.	2.6 2	1.338	3.66	1.189	- 1.0 4	- 0.158	0.05
	Total	3.1 0		3.43				0.05

Notes

5. Scoring Scale : same as in table 1
6. M.S: mean score, S.D: standard deviation, M.S diff: difference between mean scores.
7. 0.05= statement is accepted at both 5% & 10% level of significance,0.10= statement is accepted at only 10% level of significance
8. H2 is accepted at 5% level of significance.

Figure 6: Statement wise existing status of HRD Mechanisms among Medical staff in the sample study organizations with Z values.

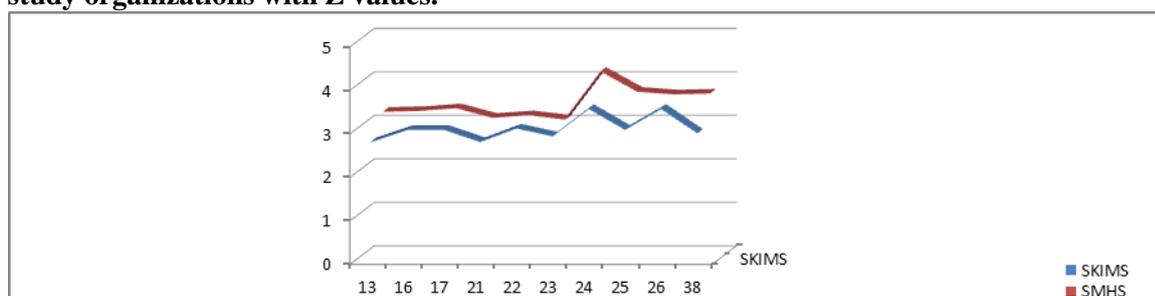


Table 7: Existing status of HRD Mechanisms among Para medical staff in the sample study organizations.

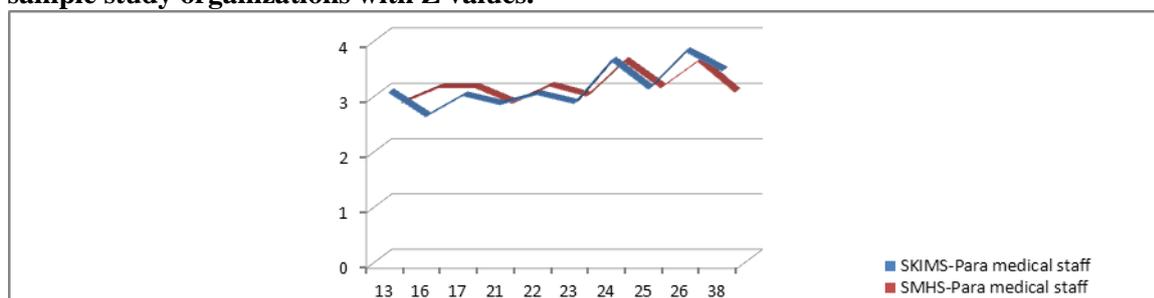
St no	Statements	SKIMS (No. 104)		SMHS (No. 121)		M.S diff	Z value	Leve l Of Signi f
		M.S	S.D	M.S	S.D			
13	The top management in the health care sector makes efforts to identify and utilize the potential of employees.	3.1 3	1.304	2.7 6	1.177	0.3 7	1.658	1.10
16	When an employee in the health care sector does good work his supervising officers take special care to appreciate it.	2.6 9	1.210	3.0 4	1.270	- 0.3 5	- 1.589	0.05
17	Performance Appraisal reports in this organization are based on objective assessment and adequate information and not on favoritism.	3.0 7	1.272	3.0 4	1.114	0.0 3	0.146	0.05
21	Weaknesses of employees in this organization are communicated to them in a non threatening way.	2.9 1	1.278	2.7 6	1.378	0.1 5	0.609	0.10
22	When behaviour feedback is given to employees in health care sector, they take it seriously and use it for development	3.0 9	1.217	3.0 7	1.246	0.0 2	0.100	0.10
23	Employees in this organization take pains to find out their strengths weaknesses from their officers and colleagues.	2.9 3	1.242	2.8 9	1.225	0.0 4	0.173	0.05
24	When employees in health care sector are sponsored for training, they take it seriously and try to learn from the programs they attend.	3.7 0	1.253	3.5 2	1.217	0.1 8	0.820	0.10
25	Employees in this organization when returning from training programs are given opportunities to tryout what they have learnt	3.1 9	1.029	3.0 4	1.212	0.1 5	0.696	0.10

26	Employees are sponsored for training programs on the basis of genuine training needs in health care sector.	3.87	0.953	3.52	1.067	0.35	1.897	0.10
38	Job rotation in health care sector facilitates employee development.	3.52	1.193	2.96	1.409	0.56	2.353	0.10
	Total	3.21		3.06				0.10

Notes

1. Scoring Scale : same as in table 1
2. M.S: mean score, S.D: standard deviation, M.S diff: difference between mean scores.
3. 0.05= statement is accepted at both 5% & 10% level of significance,0.10= statement is accepted at only 10% level of significance
4. H2 is accepted at 5% level of significance.

Figure 7: Statement wise existing status of HRD Mechanisms among Para medical staff in the sample study organizations with Z values.



Using Z test for evaluating the significance level, it was inferred from table 6 and 7 that there lies a significant difference in the perception of medical and Para medical staff of the sample study organizations. Medical staff encourages HRD Mechanisms more than Para medical staff. Hypothesis 3 is accepted at 5% and 10% level of significance.

Conclusions And Suggestions

It can be concluded from the findings that HRD Climate and HRD Mechanisms in SKIMS and SMHS is not satisfactory. An overview of the study is that,

- HRD Climate in SKIMS and SMHS is found to be average.
- There is difference in the HRD Climate of the two organizations.
- HRD Climate in SMHS is better than that of SKIMS.
- HRD Mechanisms in the sample study organizations falls within an average range of scale.
- There is a satisfactory environment for Training and Development in the two organizations.
- Training and Development is imparted only after assessing the need for it.
- Performance Appraisal System is found to be average in both the hospitals.
- Training and Development is valued and practiced more than PAS in the two organizations.
- SMHS does not encourage T&D to the extent SKIMS does.
- Medical staff of SKIMS scored less value than the medical staff of SMHS and inverse applies for the Para medical staff of the two organizations.
- Medical staff of SKIMS encourages job rotation which helps develop the employees.
- When the medical staff of SMHS is sponsored for training, they take it seriously and try to learn from the programs they attend.
- Medical staff of SMHS Hospital is less dissatisfied than medical staff of SKIMS but Para medical staff of SKIMS is less dissatisfied than Para medical staff of SMHS.

HRD Climate and HRD Mechanisms should be encouraged in both the organizations as it forms a prime element of any organization. SKIMS and SMHS should focus more on developing HRD Climate and HRD Mechanisms in the organization to ensure satisfaction of their HR so that the effectiveness and productivity of the organizations is enhanced. The top management in the health care sector should make efforts to identify and utilize the potential of employees. Both the organizations should encourage Training & Development but SMHS should focus more on this

mechanism than SKIMS. Medical staff of both the organizations should take pains to find out their strengths and weaknesses from their officers and colleagues which the research reveals is lacking in them. Good work done on the part of Para medical staff in SKIMS should be taken care of and encouraged by seniors. The top management in SMHS should make efforts to identify and utilize the potential of employees. Care should be taken while appraising the employees and in communicating the weaknesses of the employees to them in a non threatening way.

References:

- Agarwal, M., & Gupta, S. (2007). Effects of perceived procedural fairness of organizational Human Resource practices on the relationship between facets of the performance appraisal system and its effectiveness. *Indian Journal of Training and Development*, XXXVII: 2, April-June 2007, 53-67.
- Altarawneh, I. (2009). Training and development evaluation in Jordanian banking organizations. *Research and Practice in Human Resource Management*, 17(1), 1-23.
- Antonacopoulou, E.P. (2001). The paradoxical nature of the relationship between training and learning. *Journal of Management Studies* 38:3, May 2001, 327-350.
- Anuradha, R. (2003). Designing Human Resource development training for a development organization. *Indian Journal of Training and Development XXXIII: 3-4, July-December (2003)*, 99-109.
- Balsam, S., Fernando, G. D., & Tripathy, A. (2011). The impact of firm strategy on performance measures used in executive compensation. *Journal of Business Research* 64(2011), 187-193.
- Banu, C. V., & Umamaheshwari, P. (2009). A study on 360 degree performance appraisal system in Reliance life Insurance, Udumalpet. *The Icfaian Journal of Management Research*, Vol. VII, No. 7, 2009, 65-78.
- Beduk, A., Altan, M., & Sarsenova, K. (2011). Performance management in the health care industry: an application on the assessment of public-private sector hospitals. *The 2011 Barcelona Academic Conference, Barcelona, Spain 2011*, 1181-1188.
- Chakaravarty, T. (2006). Employment status of displaced workers: An effect of training. *Indian Journal of Training and Development*, XXXVII: 2, April-June (2006), 58-67.
- Chittoo, H. B., & Ramphul, N. (2006). Using performance management to meet changing citizen needs in the Mauritian context: a case of the health sector. *The International Journal of Public Sector Management*, Vol. 13, No. 1, 19-37.
- Claudia, C., Craveiro, I., & Blaise, P. (2010). Managing the performance of family physicians in the Portuguese national health system. *Human Resource for Health Development Journal (HRDJ)* Vol. 4, No. 3 September-December 2000, 184-193.
- Deshpande, B., & Arekar, K. (2011). Linking balanced card to performance management system a process model. *International Journal of Business and InfoTech*, Vol. 1, No. 1, June 2011, 109-113.
- Drost, E. A., Frayne, C. A., Lowe, K. B., & Geringer, J. M. (2002). Benchmarking training and development practices; a multi-country comparative analysis. *Human Resource Management*, Spring 2002, Vol. 41, No. 1, 67-86.
- Dwiredi, P., & Purang, P. (2007). Training needs identification and evaluation: case study of Indian organizations. *India Journal of Training and Development XXXVII: 1, January-March 2007*, 5-21.
- Garg, S. (2011). A study on performance evaluation of selected Indian mutual funds. *International Journal of Innovation Creativity and Management (IJICM)*, Vol. 1. Issue 1, January 2011, 1-11.
- Geetika, & Pandey, N. (2007). Benchmarking-A strategic tool for enhanced performance: a study of power sector in India. *The Icfaian Journal of Management Research*, Vol. VI, No. 7, 2007, 37-51.
- Ghosh, S. (2005). Organizational climate vis-a-vis performance appraisal climate of agricultural extension organizations. *Indian Journal of Training and Development XXXV:3, July-September 2005*, 93-106.
- Gunavathy, J. S., & Jayaraman, V. (2005). A study on employees opinion on the 360 degree feedback system. *Indian Journal of Training and Development XXXV:4, October-December 2005*, 52-61.
- Harvey, J., Pettigrew, A., & Ferlie, E. (2002). The determinants of research group performance towards mode. *Journal of Management Studies* 39:6, September 2002, 0022-2380, 747-774.
- Hood, P. J. (2003). A comparative study of education officer's training in two different cultures. *South Asian Journal of Management*, Vol. 10, No. 4, 25-34.

- Ibrahim, M.E. (2004). Measuring training effectiveness. *Journal of Management Research, tri annual Journal of MSA, Faculty of Management Studies, University of Delhi, Vol. 4, No. 3, December 2004*, 147-156.
- Jangid, V.K., & Chatterjee, A. (2002). Designing a performance appraisal system for shop floor employees. *Indian Journal of Training and Development XXXI:1, January- March 2002*, 79-87.
- Jawahar, I.M., & Stone, T. H. (1997). Appraisal purpose versus perceived consequences: the effects of appraisal purpose, perceived consequences and rater self monitoring on leniency of rating and decisions. *Research and Practices in Human Resource Management, 5(1)*, 33-54.
- Jian, W., Huang, Y., Hu, M., & Zhang, X. (2009). Performance evaluation of inpatient service in Beijing: a horizontal comparison with risk adjustment based on diagnosis related groups. *BMC Health services Research, 9:72, 2009*, 9-72.
- Krishnaveni, R., & Sripirabaa, B. (2008). Does Partnering and financial support have impact on alignment of training and development with strategic goals. *Indian Journal of Training and Development, 71-77*.
- Kumar, A., (2006). Women entrepreneurs in Northern India: an investigation of training related issues. *Indian Journal of Training and Development, XXXVII: 3, July-September 2006*, 61-74.
- Kuvass, B. (2007). An exploration of how the employee-organization relationship affects the linkage between perception of developmental human resource practices and employee outcomes. *Journal of management studies, j 1467-6486*, 1-26.
- Leshabari, M. T., Muhondwa, E .P. Y., Mwangu, M. A. & Mbembati, N. A. A. (2008). Motivation of health care workers in Tanzania: A case study of Muhimbili National hospital. *East African Journal of Public Health Volume 5, Number 1, April 2008*, 32-37.
- Madan, A. (2006). Appraising the performance appraisal- the Indian Scenario. *Indian Journal of Training and Development XXXVII: 3, July- September 2006*, 43-54.
- Maley, J., & Kramar, R. (2007). International performance appraisal: policies, practices and processes in Australian subsidiaries of healthcare MNC's. *Research and Practice in Human Resource Management, 15(2)*, 21-40.
- Manikandan, P., & Anwer, M. (2008). Assessment of management training needs of agricultural research managers. *The Icfaian Journal of Management Research, Vol. VII, No. 4, 2008*, 21-31.
- Manoharam, T.R., Muralidharan, C., & Deshmukh, S .G. (2009). Employee performance appraisal using data employment analysis: a case study. *Research and Practice in Human Resource Management, 17(1)*, 92-111.
- Milliman, J., Nason, S., Zhu, C., & Cieri, H. D. (2002). An exploratory assessment of the purposes of performance appraisal in North and Central America and the Pacific rim. *Human Resource Management, Spring 2002, Vol. 41, No. 1*, 87-102.
- Mishra, S. (2008). Training and development need analysis for ushering change- a study in Jindal Steel and power Ltd. *Indian Journal of Training and Development, 63-71*.
- Nair, B. & Pareek, A. (2011). A study of the various performance management system adopted by select Indian private sector organizations. *International Journal of Research in Commerce and Management, Vol. No. 2, 2011, Issue No. 6, June, ISSN 0976-2183*, 43-46.
- Nath, C. K. (2011). Impact of training programmes in HRD activities in PSU: a case study. *Indian Management Studies Journal 15(2011)*, 93-105.
- Ogunu, M. A. (2002). Evaluation of management training & development programme of Guinness Nigeria PLC. *Indian Journal of Training and Development XXXII: 1, January-March 2002*, 22-28.
- Patil, D., & Gupta, M. (2005). A study of training motivation and participation. *Indian Journal of Training and Development XXXV: 3, July-Sep 2005*, 33-48
- Poon, J. M. L., Stevens, C. K., & Gannon, M. J. (2000). Effects of Training method and learning style on cross-culture training outcomes. *Research and Practice in Human Resource Management*, 8(2), 73-97.
- Rastogi, S. (2002). Executive training and development (A comparative study of selected public and private sector enterprise). *Indian Journal of Training and Development XXXII: 1, July-Sep 2002*, 86-94.
- Selvam, M., Panchalan, R. (2003). Evaluation of executive training at NLC Ltd. - A case study. *Indian Journal of Training and Development XXXIII: 3-4, July-December 2003*, 60-70.

- Singh ,A. J., & Perdue, J. (2002). Preferred training methods for specific objectives: Survey of managers in private clubs. *Indian Journal of Training and Development XXXII: 1, Jan Mar 2002*, 47-58.
- Snell, S. A., Lepak, D. P., James, W. D., & Youndt, M. A. (2000). Selection and training for integrated manufacturing: The moderating effects of job characteristics. *Journal of Management Studies 37:3 May 2000,0022-2380*, 445-467.
- Sontakki, B. S., Samanta, R. K., Shenoy, N. S. & Reddy, P. V. (2006). Management training needs of agricultural scientists of Indian council of agricultural research. *Indian Journal of Training and Development, XXXVII: 2, April-June(2006)*, 136-148.
- Srivastava,K. B. L., Deb, S., & Prasad, A. P. (2001). Evaluating training effectiveness and customer satisfaction in Tata Steel-A case study. *Indian Journal of Training and Development*, 45-57.
- Sur, D., & Jafar, A. (2008). Evaluating performance of human resources through human resource accounting information in NTPC limited during the pre and post liberalization periods: A comparative study. *The Icfaian Journal of Management Research, Vol. VII, No. 2, 2008*, 21-35.
- Thang ,N. N., & Quang, T., & Buyens, D. (2010). The relationship between training and firm performance: A literature review. *Research and Practice in Human Resource Management, 18(1)*, 28-45.
- Tripathi, P., & Suri, R. K. (2010). Development of competence based management and performance assessment system for academic management: empirical investigation. *International Journal of Innovation and Technology, Vol. 1, No. 4, October 2010. ISSN: 2010-0248*, 357-360.
- Young, G. J., Beekun, R. I., & Stedham, Y. (2000). Board of directors and the adoption of a CEO performance evaluation process: Agency and institutional- Theory perspectives. *Journal of Management Studies, 0022-2380*, 277-294.

ARE THERE CHANGES IN ORBICULARIS ORIS MUSCLE ACTIVITY DUE TO THE CONVENTIONAL PROSTHETIC TREATMENT OF COMPLETE EDENTATION ?

S.Popsor

Claudia Soaita

University of Medicine and Pharmacy from Tîrgu Mureş, Romania

Abstract:

Introduction. It is well known that the orbicularis oris muscles, in association with m.buccinator and m. pharyngeal constrictor, form a well defined functional unit, called the „buccinator mechanism“, with important role in some orofacial functions. It is also involved in the “neutral zone” establishment, as essential feature of complete edentation treatment. The purpose of this investigation was to assess the changes in clinical and electromyographic activity of orbicularis oris muscle (OO) in complete edentulous patients before and after the insertion of dentures. Material and method. Seven patients were included in this study. Electromyographic records of the upper(OOS) and lower(OOI) orbicularis oris muscle were made using the BioEMG II (Bioresearch Assoc.Inc.) surface electromyography device, performed at rest, while whistling and sucking before and 3 days after the insertion of the complete dentures. Results. There was a significant difference between OOS and OOI activity before inserting dentures in all three conditions ($p = 0.0156$) as well as after the prosthesis insertion during sucking and whistling, a more pronounced activity of OOI being obvious. No statistically significant difference in the activity of the mentioned muscle groups before and after insertion of the dentures, in the three recording conditions was observed. Conclusion. Within the limits of the present study, we conclude that in terms of EMG at rest as well as during functional moments of sucking or whistling there are no significant changes in OOI and OOS muscle tone, before and after conventional prosthetic treatment of complete edentation.

Key Words: Orbicularis oris muscle, complete edentation, electromyography

Introduction:

Thirty five years have passed since Perkins et al.[1]based on their neurophysiological studies have shown that the m.orbicularis oris , m.buccinator and m. pharyngeal constrictor form a well defined functional unit , called the „buccinator mechanism“, with important role in orofacial function (swallowing, sucking, whistling, chewing, vowel pronunciation, kissing). In this muscle complex m.orbicularis oris in its upper (OOS) and inferior (OOI) part were electromyographically studied to assess their role in mastication [2], swallowing[3] and in some neurosenzorial abnormalities, such as deafness[4, 5].

Given the important morphogenetic role of orofacial muscles, a series of orthodontic studies [6, 7] have found an increase in electromyographic activity of m.orbicularis oris in centric occlusion and during sucking in patients treated with postural training devices. An earlier study of Gustaffson and Ahlgren [8] , investigating children with different morphologies of the lips, showed no electromyographic activity of m orbicularis oris at rest, but have noticed increased activity in chewing and swallowing in children incompetence of the lips, which was notified later by other authors [9, 10, 11] . Tryde and Tallgren [12] studying the electromyographic activity of the muscles of the lips during swallowing in dentures wearers found that OOI electromyographic activity was stronger than the OOS one.

Our aim was to assess the changes in clinical and electromyographic activity of OO muscle in complete edentulous patients before and after the insertion of dentures in three essential functional requirements (posture, whistling and sucking) in terms of total edentulous diagnostic, therapy and evolution.

Material and method

Patients. Seven patients (1F and 6 M, mean age 61 years) presented for treatment with conventional mobile prostheses were included in the study. All participants were asked for informed consent regarding the electromyographic examination as a means of evaluating the therapeutical results.

Electromyography. Electromyographic records of the orbicularis oris muscle were made using the BioEMG II (Bioresearch Assoc.Inc.) surface electromyography device, which is specifically intended to record the activity of craniofacial muscles. We used the TA-R channel for OOS and the TA-L one for OOI. The placement of the adhesive surface bipolar electrodes was made according to the anatomical disposition of muscle fibers, parallel to this, on the right side third of each of the two muscles (Fig. 1), in accordance with the protocols proposed by Barrel & Moyers [13] and Vaiman [14]. Adhesive ground electrode was attached to the patient's right shoulder.



Fig. 1. Bio EMG II device and electrodes placement

The electromyographic records were performed before and 3 days after the insertion of the complete dentures in the mouth during three conditions: rest, while whistling and sucking. For each of these conditions, from the 10 seconds of recording three sections of 2.5 seconds for the two investigated muscles, were analyzed. The mean amplitude of electromyographic signals calculated by the software device (BioPack) (Fig. 2) were listed in the tables, the results were expressed graphically and statistically interpreted.

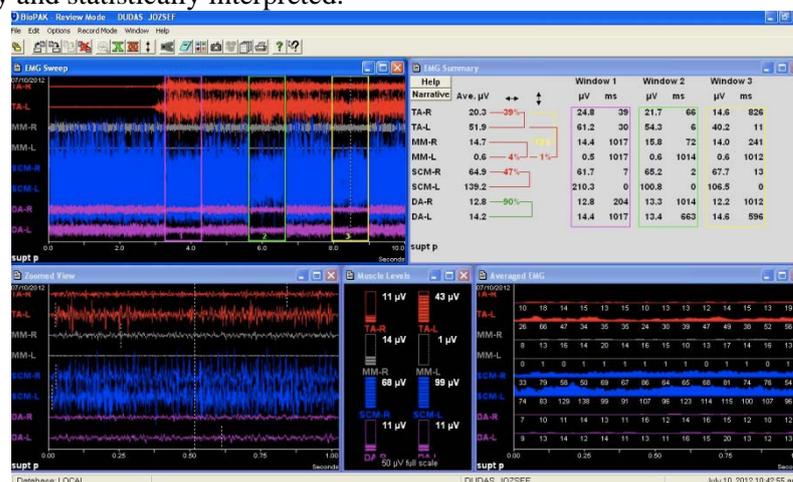


Fig. 2. The BioPack software review mode of the recordings. The TA-R channel for OOS and the TA-L one for OOI were used.

Statistics. For the statistical evaluation of the differences in electromyographic activity before and after the prosthetic treatment, the Wilcoxon matched-pairs signed rank was used. To determine the two-tailed p value at a significance level of 0.05, GraphPad Prism software version 5.0 was used.

Results

The Tables. II, III and IV present the mean amplitude of electromyographic signals picked-up from the investigated muscles of the lips in three conditions of registration. There was a significant difference between OOS and OOI activity before inserting dentures in all three conditions ($p=0.0156$). After the prosthesis insertion during sucking and whistling, a more pronounced activity of OOI was obvious. No statistically significant difference in the activity of the mentioned muscle groups before and after the insertion of the dentures, in the three recording conditions could be seen (Table I)

Table no. I. Differences in activity of both muscles in the three recording conditions before and after insertion of dentures (Wilcoxon matched-pairs signed rank test).

	Rest two-tailed p value	Sucking two-tailed p value	Whistling two-tailed p value
OOS 1 vs. OOS 2	0.9325	0.1563	0.8125
OOI 1 vs. OOI 2	0.4375	0.3750	0.0781

Table no. II. The mean amplitude of EMG signals (in microvolts) picked-up from the orbicularis oris muscle at rest (OOS = orbicularis oris superior, OOI =orbicularis oris inferior) before (1) and after (2) insertion of the dentures.

Patient	OOS 1	OOI 1	OOS 2	OOI 2
D.J	1.2	3.3	1.4	2.3
E.S	4.6	7.6	2.8	1.9
M.I	3.6	4.8	2.6	6.0
M.T	3.0	4.3	3.8	3.5
T.A	2.8	4.0	3.1	4.3
C.L	2.2	3.8	2.8	3.8
Sz.A	3.4	4.4	4.0	3.9

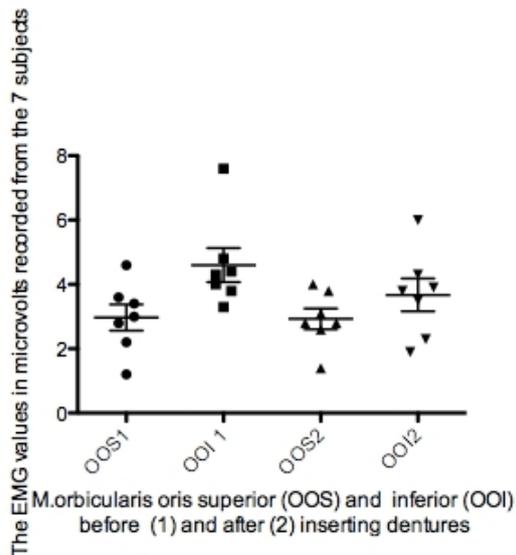


Fig. 5. Distribution of mean values of EMG signals at rest (95% CI)

Table no. III. Mean amplitude of m. orbicularis oris EMG signals (in microvolts) while sucking, before and after inserting dentures

Patient	OOS1	OOI 1	OOS2	OOI 2
D.J	22.2	77.3	20.3	51.9
E.S	39.7	68.4	37.0	66.8
M.I	53.0	88.0	44.1	51.4
M.T	38.4	62.3	22.2	50.6
T.A	42.4	78.8	38.0	66.2
C.L	38.2	65.3	44.0	77.1
Sz.A	28.9	60.4	26.0	74.3

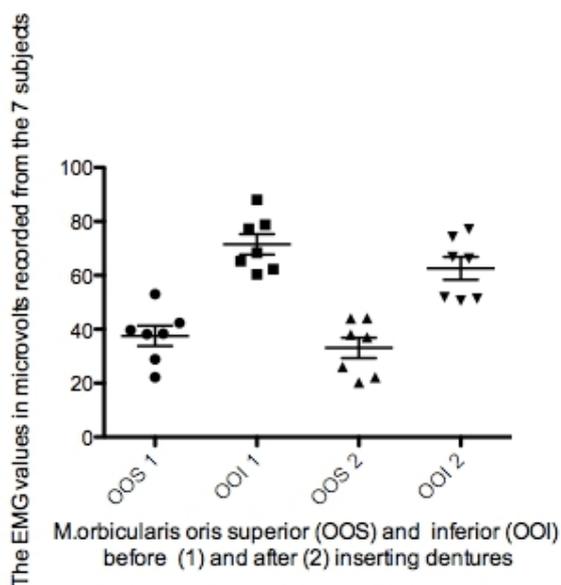


Fig.6. Distribution of EMG signals mean values during sucking (95% CI)

Table no. IV . The EMG signals mean amplitude (in microvolts) of orbicularis oris muscle while whistling, before and after inserting dentures

	OOS1	OOI1	OOS2	OOI2
Patient	17.6	60.1	11.0	35.2
E.S	34.4	82.7	41.9	82.6
M.I	37.2	45.0	43.0	53.9
M.T	49.0	61.6	45.3	50.6
T.A	22.7	61.2	20.2	56.3
C.L	36.4	82.0	34.0	70.0
Sz.A	33.0	72.2	38.0	62.8

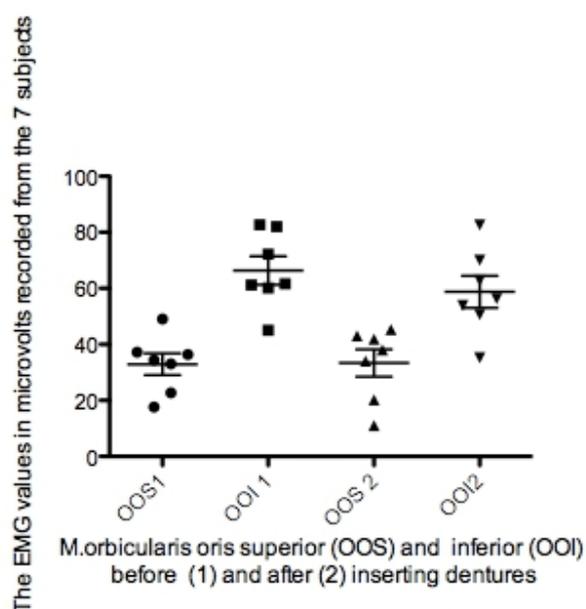


Fig. 7. Distribution of EMG signals mean values during whistling (95% CI)

Discussions

The orbicularis oris muscle is one which by its fiber orientation has a positive action to provide, the stability of the dentures. Placing the artificial teeth so as not to interfere with muscle activity and to enable the retention and stability action of the musculature, is a priority of the complete edentation treatment [15, 16, 17]. For this purpose various prosthetic techniques have been imagined [18, 19, 20, 21]. The tonicity of OO muscles can be assessed under different electromyographic examination. According to available literature, some studies have evaluated the electromyographic activity of the muscles during different syllables before and after insertion of dentures [22]. OOI muscle generally present higher mean values of electromyographic activity compared with those recorded in OOS, both before and after insertion of dentures. Electromyographic values were also higher after insertion of dentures. Redegard Ingervall [23] and have published their research on the orbicularis oris muscles before and after insertion of dentures, but their electromyographic records were achieved in centric occlusion.

The present study used three conditions for electromyographic registration proper to OO muscles functionality (rest, sucking, whistling), two of which are well known functional movements to establish the prosthetic field limits and to place de anterior artificial teeth in the neutral zone. The results showing more elevated electromyographic activity in OOI compared

with OOS in all three conditions before prosthetic rehabilitation, as well as during the whistle and while sucking with the inserted dentures are consistent with those obtained in other functional conditions, earlier presented in this paper.

The BioEMG II device used in our study is specifically destined to record the head and neck muscle activity and it is the only surface EMG system that is entirely gainless. In the available literature we found orthodontic studies using the BioEMG II electromyographic system for orbicularis oris muscles. Saccucci et al. [24] used the BioEMG II to assess the changes caused by the functional devices and found that they improved muscle contraction after treatment. The subject of the thesis developed by Daniela Maria Afonso Estavos [25] in 2011 using BioEMG II electromyographic system was to establish the relationship between the OO muscles and the irregularities in the anterior mandibular arch. She concluded that electromyographic activity of the orbicularis oris muscles can not be related to irregularities in the mandibular arch, in either functional moments. This finding is consistent with those observed by us in complete edentation, the denture insertion actually represents a change in edentulous arch. Orbicularis oris muscle activity has undergone no significant change from the original condition after prosthetic therapy in any of the recorded conditions. The relatively few patients number, however, is a limitation of our study. We also believe that further late electromyographic evaluation after 6 months would be useful, when probably the patient adaptation to dentures may manifest as changes in muscle tone.

-
- Conclusion: Within the limits of the present study, we conclude that in terms of EMG at rest as well as in functional moments of sucking or whistling there are no significant changes in OOI and OOS muscle tone, before and after conventional prosthetic treatment of complete edentation.

References:

1. Perkins R.E., Blanton P.L., Biggs N.L. Electromyographic Analysis of the Buccinator Mechanism in Human Beings, *JDR*, 1977; 56(7): 783-794.
2. Hanawa S., Tsuboi A., Watanabe M., Sasaki K. EMG study for perioral facial muscles function during mastication, *J Oral Rehabil*, 2008 Mar; 35(3): 159-170.
3. Ding R., Larson CR., Logemann JA, Rademaker AW. Surface electromyographic and electroglottographic studies in normal subjects under two swallow conditions: normal and during the Mendelsohn maneuver, *Dysphagia*, 2002 Winter; 17(1) 1-12.
4. Regalo SC, Vitti M., Moraes Maria Theresa Bagaiolo, Semprini Marisa, de Felicio Claudia Maria, de Mattos Maria da Gloria Chiarello, Hallak JEC, Moreto Santos Carla. Electromyographic analysis of the orbicularis oris muscle in oralized deaf individuals, *Braz.Dent.J*, 2005; 16(3): 1-6.
5. Regalo SC, Vitti M, Hallak JE, Siessere S., Pagnano VO, Semprini M. Electromyographic analysis of upper and lower fascicles of the orbicularis oris muscle in deaf individuals, in mandibular rest position, compared to hearers. *Electromyogr Clin Neurophysiol*, 2006; Jul-Aug 46(4): 211-215.
6. Uysal T, Yagaci A, Kara S, Okkesim S. Influence of pre-orthodontic trainer treatment on the perioral and masticatory muscles in patients with Class II division 1 malocclusion, *Eur J Orthod*, 2012 Feb;34(1): 96-101.
7. Yagaci A, Uysal T, Kara S, Okkesim S. The effects of myofunctional appliance treatment on the perioral and masticatory muscles in Class II, Division 1 patients, *World J Orthod*, 2010 Summer; 11(2): 117-122.
8. Gustaffson M, Ahlegren J. Mentalis and orbicularis oris activity in children with incompetent lips: An electromyographic and cephalometric study, *Acta Odontologica Scandinavica*, 1975; 33(6): 355-363.
9. Tosello DO, Vitti M, Berzin F. EMG activity of the orbicularis oris and mentalis muscles in children with malocclusion, incompetent lips and atypical swallowing-part II, *J Oral Rehabil*, 1999 Aug; 26(8): 644-649.

10. Siqueira VCV, Sousa MA, Berzin F, Casarini CAS. Electromyographic analysis of the orbicularis oris muscle in youngsters with Class II, Division 1 and normal occlusion. *Dental Press J Orthod*, 2011 Sept-Oct; 16(5): 54-61.
11. Souza DR, Semeghini Tatiana Adamov, Kroll LB, Berzin F. Oral Myofunctional and Electromyographic Evaluation of the Orbicularis Oris and mentalis Muscles in Patients with Class II/1 Malocclusion Submitted to First Premolar Extraction, *J Appl Oral Sci* 2008; 16(3): 226-231.
12. Tallgren A, Tryde G. Swallowing activity of lip muscles in patients with a complete upper and a partial lower denture. *J Oral Rehabil*, 1992 Jul; 19(4): 329-341.
13. Baril C, Moyers RE. An Electromyographic Analysis of the Temporalis Muscles and certain Facial Muscles in Thumb-and Finger-sucking Patients, *J Dent Res* 1960 May-June; 39: 536-552
14. Vaiman M. Standardization of surface electromyography utilized to evaluate patients with dysphagia. *Head&Face Medicine* 2007; 3: 26-31.
15. Beresin VE, Schiesser FJ. The neutral zone in complete dentures. *J Prosthet Dent* 1976 Oct; 36(4): 356-367.
16. Chandra SS. Management of a severely resorbed mandibular ridge with the neutral zone technique. *Contemp Clin Dent* 2010 Jan-Mar; 1(1): 36-39.
17. van Waas MA. Neutral zone. An anatomic space which is often neglected while fabricating removable complete dentures, *Ned Tijdschr Tandheelkd*, 2011 Nov; 118(11): 563-567.
18. Lynch CD, Allen PF. Overcoming the unstable mandibular complete denture: the neutral zone impression technique, *Dent Update* 2006 Jan-Feb; 33(1): 21-22; 24-26.
19. Demirel F, Oktemer M. The relations between alveolar ridge and the teeth located in neutral zone, *J Marmara Univ Dent Fac* 1996 Sep; 2(2-3): 562-566.
20. Gahan MJ, Walmsley AD. The neutral zone impression revisited, *Br Dent J* 2005 Mar 12; 198(5): 269-272
21. Kursoglu P., Ari N, Calikkocaoglu S. Using tissue conditioner material in neutral zone technique, *NY State Dent J*, 2007 Jan; 73(1):40-42
22. Santos M Carla, Vitti M, Mattos OF Maria, Semprini Marisa, Paranhos OF Helena, Hallak J, Regalo SC. Effect of denture quality on perioral muscle activity during speech, *Braz J Oral Sci* 2005 ; 4(14): 801-805 .
23. Ingervall B, Redegard B. An electromyographic study of masticatory and lip muscle function in patients with complete dentures, *J Prosthet Dent* 1980; 43: 266-271
24. Saccucci M., Tecco S., Ierardo G., Luzzi V., Festa F., Polimeni A. Effects of interceptive orthodontics on orbicular muscle activity: a surface electromyographic study in children, *J Electromyogr Kinesiol*, 2011 Aug; 21(4): 665-671.
25. Daniela Maria Estavos Afonso: Harmonia labial como factor de Estabilidade do Tratamento Ortodontico. Dissertação de candidatura ao grau de Mestre apresentada à Faculdade de Medicina Dentária da Universidade do Porto, 2011.

HEPATOPROTECTIVE EFFECT OF BAY LEAVES CRUDE EXTRACT ON PRIMARY CULTURED RAT HEPATOCYTES

Nahla A. Ayoub

Department of Pharmacognosy, Faculty of Pharmacy, Ain-Shams University, Cairo, Egypt

Amani N. Hashim

Sahar A. Hussein

Nesrine M. Hegazi

Heba M. Hassanein

Mahmoud A. Nawwar

Department of Phytochemistry, Division of Pharmaceutical Industries, National Research Center, Dokki, Cairo, Egypt

Abstract:

The crude extract of bay leaves (*Laurus nobilis*) grown in Egypt, was tested for its hepatoprotective activity on primary cultured rat hepatocytes against paracetamol toxicity. The Neutral Red Assay was used to assess the hepatotoxicity of the extract and was applied with a broad range of concentrations (125-1000 µg/ml) on monolayer of rat hepatocytes. It revealed that the extract exerts no toxic effect on the monolayer hepatocytes. For the evaluation of the hepatoprotective effect; different concentrations were prepared, starting from 12.5 µg/mL and increasing concentration in ascending order by dissolving in DMSO (1% maximum concentration). For each concentration, three replicates were carried out, in addition to controls which were: cell control (cells only), negative control (cells + paracetamol) and positive control (reference) (cells + sylimarin + paracetamol). 50% mortality of the hepatocytes (IC₅₀) was determined using Neutral Red Assay. *L. nobilis* extract showed hepatoprotective activity against paracetamol toxic effect at concentration of 40 µg/ml. The constitutive flavonoids of *L. nobilis* leaves were extensively studied and led to the separation and identification of 14 compounds, 10 of which have not been previously identified in *L. nobilis* and were identified using chemical, conventional and advanced spectral techniques.

Key Words: *Laurus nobilis*; Lauraceae; flavonoids; hepatoprotective

Introduction:

Laurus nobilis (Lauraceae), known in English as bay laurel, sweet bay, bay tree, true laurel, or Grecian laurel (Brown, 1956) is an aromatic evergreen tree or large shrub with green, glossy leaves, native to the Mediterranean region. It has been analyzed for alkaloids (Pech and Bruneton, 1982), essential oils (Sellami I.H. *et al.* 2011, Zola *et al.*, 1977), acylated kaempferol glycosides (Fiorin C. *et al.*, 1998), sesquiterpene lactones (Dall'Acqua *et al.* 2006), megastigmane glycosides (Marino *et al.*, 2004), (+)-catechin, (-)-epicatechin, (+)-gallocatechin, (+)-epigallocatechin and procyanidins (B2, B4, B5 and B7) (Sakar M. K. *et al.*, 1985). The plant is one of the most widely used culinary spices in all Western countries. Its leaves are traditionally used orally to treat the symptoms of gastrointestinal problems, such as epigastric bloating, impaired digestion, eructations and flatulence (Brunetton, 1999). The aqueous extract is used in Turkish folk medicine as an anti-hemorrhoidal, anti-rheumatic, diuretic, as an antidote in snake bites and for the treatment of stomach-ache (Sezik E., 1999, Baytop T., 1984).

A survey of the literature showed that no experimental data are available regarding its hepatoprotective potential. It was therefore, interesting to investigate, in the present study the possible hepatoprotective effect of *L. nobilis* leaves and to separate and identify some of their constitutive flavonoids.

Experimental:**Plant Material:**

The leaves of *L. nobilis* were collected from a private farm in El- Mansouria area, Giza, Egypt, in September 2010 and authenticated by Dr. M. El Gebali, National Research Centre (NRC), Cairo, Egypt.

Instruments and Materials:

¹H NMR spectra were measured by a Jeol ECA 500 MHz NMR spectrometer. ¹H chemical shifts (δ) were measured in ppm, relative to TMS and ¹³C NMR chemical shifts relative to DMSO-*d*₆ and converted to TMS scale by adding 39.5. spectral width = 8 kHz for ¹H and 30 kHz for ¹³C, 64 K data points and a flip angle of 45°. UV recording was made on a Shimadzu UV-Visible-1601 spectrophotometer. Paper chromatographic analysis was carried out on Whatman No. 1 paper, using solvent systems: (1) H₂O; (2) 6 % HOAc; (3) BAW (n-BuOH : HOAc : H₂O, 4:1:5, upper layer). Solvent 3 was used for preparative paper chromatography (Prep. PC) using Whatman No. 3 paper.

In vitro Bioassay on Primary Culture of Rat Hepatocytes Monolayer:**Isolation and Culture of Rat Hepatocytes Monolayer:**

Primary culture of rat hepatocytes was prepared according to (Seglen, 1976) method, modified by (Kiso, *et al.* 1983) using Wistar male rats (250-300 g) obtained from the animal house of the NRC (National Research Center, Cairo). Animal procedures were performed in accordance with the Ethics Committee of the National Research Centre and followed the recommendations of the National Institutes of Health Guide for Care and Use of Laboratory Animals (NIH Publication No. 85-23, revised 1985). At time of surgery, the rat was anesthetized by intra-peritoneal injection of 3.4mL/kg of sodium thiopental solution (0.1g/mL). A midline incision was made, the liver was exposed and the portal vein was cannulated with a needle fitted with a teflon catheter. After the teflon catheter was tied in place and the needle removed, the inferior *vena cava* was cut below the renal vein to allow blood drainage. Perfusion of the liver was started with Ca²⁺ free buffer (buffer 1), which contained HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) (10 mmole/L), NaCl (137 mmole/L), KCl (2.68 mmole/L), Na₂HPO₄ (0.7 mmole/L), D (+) glucose (10 mmole/L) and EGTA (ethylene glycol tetraacetic acid) (0.5 mmole/L), and adjusted to pH 7.45 at 37°C. The flow-rate was 11.6 mL/min. The thoracic portion of the superior *vena cava* was cannulated, and the inferior *vena cava* was tied-off above the renal vein. After perfusing the liver for 15min., recirculation of Ca²⁺-containing buffer (buffer 2), containing additionally to buffer 1, CaCl₂ (5mmole/L) and Collagenase type IV (10g/L) (pH 7.45 at 37°C) was started, adjusted at a flow-rate of 7.5 mL/min, for 10 min. After perfusion with buffer 2, the liver was dissected out of the body, placed in a beaker containing buffer 2 and transferred under sterile conditions to laminar flow and gently dispersed with two forceps into small pieces (2-3 mm). The cell suspension was then filtered through the cotton gauze into centrifuge tubes. Finally, the preparation was centrifuged at 600 rpm for 5 min. The supernatant was aspirated off and the loosely packed pellet of cells was gently re-suspended in Phosphate buffer (PB) (0.1 mole/L), which contained NaH₂PO₄ (0.2 mole/L) and Na₂HPO₄ (0.2 mole/L), adjusted at pH 7.6 at 37°C, and then diluted two-folds with distilled water. This washing procedure was repeated twice. Viability of the cells to exclude the dye, trypan blue, was determined by incubating cell suspension (0.1 mL) with trypan blue (0.9 mL) and then counting the number of the cells that excluded the dye and the number of cells whose nuclei were stained blue, using haemocytometer under light microscope (Nikon). Complete culture medium was added over the total counted cells to reach a concentration of 1 x 10⁶ cells/mL.

The culture medium was composed of RPMI-1640 medium, supplemented with 10% inactivated (56°C for 30 min) fetal calf serum (FCS) (0.05/mL), penicillin-streptomycin (0.01/mL), insulin (0.7 (g/mL) and dexamethasone (4g/mL). Inocula of 1 x 10⁶ cells/mL were seeded into plastic 96-well plates and preincubated in CO₂-Incubator at 37°C, under 5% CO₂ in air for 22-24 hrs. All buffers were freshly prepared and sterilized at 121°C for 30 min. before use.

IC₅₀ Determination of the extract:

After pre - incubation of primary culture of rat hepatocytes for 22-24 hrs, the monolayer was checked under inverted microscope (Olympus) for attachment. Then, the monolayer was washed twice with (Phosphate Buffer Saline (PBS), which contained KCl (2.68 mmole/L), Na₂HPO₄ (8.45 mmole/L), KH₂PO₄ (1.5 mmole/L) and NaCl (137 mmole/L).

In order to determine IC₅₀, different concentrations of the extract were prepared. The range of concentrations used started from 125 µg/mL followed by increasing concentrations in ascending order up to the concentration 1000 µg/mL that induced death of half the number of cells. *L. nobilis* extract was dissolved in dimethylsulfoxide (DMSO) (1% maximum concentration). For each concentration, three replicates were carried out. The plate was incubated for 2 hrs, in CO₂ incubator.

After cell incubation with the extract, cell viability was determined using neutral red assay (NR). 200 µL culture medium containing all except (PS), was added to each well together with 50 µL of a solution of 2 mg NR/mL PBS. After incubation for 3 hrs, supernatant was removed as previously described and the neutral red dye obtained was dissolved in 200 µL 50% ethanol with 1% acetic acid. The plate was then shaken for 5 min. on a microtitre plate shaker and the plate on Microelisa reader was read at 540 and 630 nm dual wavelength using automatic kinetic microplate reader (Labsystems Multiskan RC reader).

Evaluation of the effect of the extract on cultured hepatocytes was obtained by calculating the absorption of the cell viability with respect to control cells. Each experiment was carried out in triplicate to confirm validity of results. A graph was plotted with x-axis showing different concentrations of extract used, y-axis showing absorbance percentage of viable cells. IC₅₀ was graphically determined from the concentration that yielded an absorption coinciding with 50% absorbance.

Extraction, isolation and purification:

The aqueous methanolic extract of *L. nobilis* leaves (48 g) showed by preliminary two-dimensional paper chromatographic (2D-PC) screening to contain a complicated flavonoid mixture from which fourteen compounds (**1-14**) were isolated and purified through fractionation on polyamide 6s (400 g) column (150 × 5 cm) and elution with H₂O followed by H₂O-MeOH mixtures of decreasing polarities to yield eight column fractions (I-VIII). The received fractions, were individually subjected to 2D-PC and other isolatory chromatographic techniques, thus yielding ten flavonoids (**1, 4, 5-12**) isolated from this plant for the first time.

Isolation and purification of flavonoids from *L. nobilis* leaves.

Compound **1** (87 mg) was purely isolated from fraction II (eluted with 10% aq. MeOH) by repeated column fractionation of 1.8 g material over Sephadex LH-20 using H₂O for elution. Compound **2** (108.8 mg) was crystallized from an aqueous solution of 2.1 g of fraction III (eluted with 20% aq. MeOH). Compound **3** (99 mg) was purely isolated from 1.1 g of fraction IV (eluted with 30% aq. MeOH) by preparative PC, using BAW as solvent. Compounds **4** (76 mg), **5** (115 mg) and **6** (78.5 mg) were individually separated pure from a Sephadex LH-20 column of fraction V (2.2 g, eluted with 50% aq. MeOH), using n-butanol water saturated for elution. Compounds **7** (54 mg), **8** (71 mg) and **9** (57 mg) were obtained pure from 1.93 g of fraction VI (eluted with 60% aq. MeOH) by applying Sephadex LH-20 column fractionation, using n-butanol water saturated for elution. Compound **10** (31 mg) and **11** (66 mg) were obtained from 3.1 g of fraction VII (eluted with 80% aq. MeOH). Compounds **12** (74 mg) and **13** (35 mg) and **14** (20mg) were isolated from 6.2 g of fraction VIII (eluted with 90% aq. MeOH) by repeated prep. PC using 6% AcOH and BAW as solvent.

Spectral data of compounds (1, 4, 5-12) isolated for the first time from *L. nobilis* leaves.

Quercetin 3-O- α-L-rhamnopyranoside (1) was obtained as pale yellow amorphous powder of R_f values (x 100):22 (H₂O), 48 (AcOH-6), 68(BAW). It exhibited a Mr of 448 in negative ESI-MS analysis ([M-H]⁻ at m/z = 447.1). UV spectral data λ_{max} (MeOH): 259 nm, 297 nm sh., 348 nm, + NaOMe: 270 nm, 355nm, 402 nm, + NaOAc: 276 nm, 372 nm, + NaOAc + H₃BO₃: 272 nm, 383 nm, +Al Cl₃: 268 nm, 352 nm, 408 nm. ¹H-NMR spectral data (DMSO-d₆) δ (ppm), quercetin moiety: 6.17 (d, J=2.5 Hz, H-6), 6.36 (d, J=2.5 Hz, H-8), 7.25 (d, J=2.5, H-2'), 6.82 (d, J=8 Hz, H-5'), 7.25 (dd, J=2.5 and 8 Hz, H-6'), rhamnose moiety: 5.20 (Δv_{1/2} = 4 Hz, H-1''), 3.1 – 3.9 (m, overlapped with water proton resonances, H-2''- H-5''), 0.77 (d, J= 6, H-6''). ¹³C-NMR spectral data (DMSO-d₆) δ (ppm), quercetin moiety: 156.9 (C-2), 134.6 (C-3), 178.2 (C-4), 161.7(C-5), 99.19 (C-6), 164.7 (C-7), 94.15(C-8), 157.8 (C-9), 104.5 (C-10),121.2 (C-1'),115.9 (C-2'), 145.7 (C-3'), 148.9(C-4'), 116.1(C-5'), 121.6 (C-6'), rhamnose moiety: 102.2 (C-1''), 70.8 (C-2''), 71.1 (C-3''), 71.6 (C-4''), 70.5 (C-5''), 18.01(C-6'').

Kaempferol 3-O- β - glucopyranoside (4): was obtained as pale yellow amorphous powder of R_f values (x 100): 24 (H₂O), 47 (AcOH-6), 76 (BAW). It exhibited a Mr of 432 in negative ESI-MS analysis $[M-H]^-$ at $m/z = 431$. UV spectral data λ_{max} (MeOH): 266 nm, 345 nm, + NaOMe: 271 nm, 376nm, + NaOAc: 270 nm, 346 nm, + NaOAc + H₃BO₃: 270 nm, 346 nm sh., 405, +Al Cl₃: 268 nm, 340 nm sh., 385 nm. ¹H- NMR spectral data (DMSO-*d*₆) δ (ppm), kaempferol moiety: 6.186 (d, $J=2.5$ Hz, H-6), 6.38 (d, $J=2.5$ Hz, H-8), 7.7 (d, $J=2.5$, H-2', H-6'), 6.89 (d, $J=8$ Hz, H- 3', H-5'), glucose moiety: 5.26 (d, $J=8$ Hz, H-1''), 3.4-4.0 (m, H-2''- H-6''). ¹³C-NMR spectral data (DMSO-*d*₆) δ (ppm), kaempferol moiety: 158.6 (C-2), 133.6 (C-3), 179.9 (C-4), 162.5 (C-5), 100.9 (C-6), 167.0 (C-7), 95.2 (C-8), 159.3 (C-9), 106.1 (C-10), 123.1 (C-1'), 132.6 (C-2', C-6'), 116.3 (C-3', C-5'), 162.1 (C-4'), glucose moiety: 105.1 (C-1''), 73.1 (C-2''), 75.1 (C-3''), 70.3 (C-4''), 77.2 (C-5''), 61.9 (C-6'').

Quercetin 3'-O- β - glucopyranoside (5) was obtained as yellowish white amorphous powder of R_f values (x 100): 23 (H₂O), 42 (AcOH-6), 53 (BAW). It showed a Mr of 464 in negative ESI-MS, corresponding to a molecular ion $[M-H]^-$ at $m/z = 463$. UV Spectral Data λ_{max} (MeOH): 255 nm, 266 nm sh., 370 nm + NaOMe: 271 nm, 330 nm, 430 nm, + NaOAc: 260 nm sh., 272 nm, 320 nm, 340, + NaOAc + H₃BO₃ : 255 nm, 268 nm sh., 305 nm sh., 375 nm, +Al Cl₃ : 265 nm, 305 nm sh., 360 nm, 430 nm. ¹H- NMR Spectral Data (DMSO-*d*₆) δ (ppm), quercetin moiety: 6.16 (d, $J=2.5$ Hz, H-6), 6.6.4 (d, $J=2.5$ Hz, H-8), 6.86 (d, $J = 8.0$ Hz, H-5'), 7.45 (dd, $J=2.5$ and 7.5 Hz, H-6'), 7.85 (s, H- 2'), glucose moiety: 5.05 (d, $J=8$ Hz, H-1''), 3.4-4.0 (m, H-2''- H-6''). ¹³C-NMR spectral data (DMSO-*d*₆) δ (ppm), quercetin moiety: 147.4 (C-2), 137.47 (C-3), 177.36 (C-4), 162.53 (C-5), 99.3 (C-6), 165.65 (C-7), 94.42 (C-8), 158.17 (C-9), 104.55 (C-10), 124.92 (C-1'), 115.86 (C-2'), 147.67 (C-3'), 146.52 (C-4'), 116.97 (C-5'), 121.42 (C-6'), glucose moiety: 105.1 (C-1''), 73.1 (C-2''), 75.1 (C-3''), 70.3 (C-4''), 77.2 (C-5''), 61.9 (C-6'').

Quercetin 3-O- β -galactoside (6) was obtained as yellowish white amorphous powder of R_f values (x 100): 20 (H₂O), 37 (AcOH-6), 56 (BAW). It showed a Mr of 464 in negative ESI-MS, corresponding to a molecular ion $[M-H]^-$ at $m/z = 463$. UV spectral data λ_{max} (MeOH): 257 nm, 266 nm sh., 359 nm, + NaOMe: 266 nm, 227 nm sh., 408 nm, + NaOAc: 274 nm, 379 nm, + NaOAc + H₃BO₃: 268 nm, 272 nm sh., 384 nm, +Al Cl₃: 265 nm, 272 nm sh., 384 nm. ¹H- NMR spectral data (DMSO-*d*₆) δ (ppm), quercetin moiety: 6.19 (d, $J=2.5$ Hz, H-6), 6.4 (d, $J=2.5$ Hz, H-8), 7.54 (d, $J=2.5$ Hz, H-2'), 6.82 (d, $J=8$ Hz, H-5'), 7.65 (dd, $J=2$ and 8 Hz, H-6'), galactose moiety: 5.36 (d, $J=8$ Hz, H-1''), 3.2 – 3.8 (m, H-2''-6''). ¹³C-NMR spectral data (DMSO-*d*₆) δ (ppm), quercetin moiety: 156.24 (C-2), 133.33 (C-3), 177.32 (C-4), 160.57 (C-5), 98.61(C-6), 164.79 (C-7), 93.57 (C-8), 156.26 (C-9), 103.75 (C-10), 120.91(C-1'), 115.164 (C-2'), 144.89 (C-3'), 148.31 (C-4'), 115. 51 (C-5'), 121. 81(C-6'), galactose moiety: 101.79 (C-1''), 71.074 (C-2''), 72.97 (C-3''), 67.81 (C-4''), 75.54 (C-5''), 60 (C-6'').

Isorhamnetin 3-O- β - glucopyranoside (7) was obtained as yellowish white amorphous powder of R_f values (x 100): 11 (H₂O), 12 (AcOH-6), 52 (BAW). It showed a Mr of 478 in negative ESI-MS, corresponding to a molecular ion $[M-H]^-$ at $m/z = 477.3$. UV spectral data λ_{max} (MeOH): 256 nm, 300 nm sh., 354 nm, + NaOMe : 275 nm, 325 nm sh., 408 nm, + NaOAc: 275 nm, 323 nm, 371 nm, + NaOAc + H₃BO₃ : 276 nm, 295 nm sh., 355 nm +Al Cl₃ : 269 nm, 365 nm, 403 nm. ¹H- NMR spectral data (DMSO-*d*₆) δ (ppm), isorhamnetin moiety: 6.22 (d, $J=1.6$ Hz, H-6), 6.43 (d, $J=1.6$ Hz, H-8), 3.84(s, OCH₃), 7.95 (d, $J=1.6$ Hz, H-2'), 6.93 (d, $J=8.4$ Hz, H-5'), 7.56 (dd, $J=1.6$ and 8.4 Hz, H-6'), glucose moiety: 5.57 (d, $J=8$ Hz, H- 1''), 3.0-3.8 (m, H-2''- H-6''). ¹³C-NMR spectral data (DMSO-*d*₆) δ (ppm), isorhamnetin moiety: 156.3 (C-2), 133.2 (C-3), 177.3 (C-4), 161.1 (C-5), 98.7 (C-6), 164.3 (C-7), 93.6 (C-8), 156.2 (C-9), 103.9 (C-10), 55.6 (OCH₃), 120.9 (C-1'), 113.4 (C-2'), 148.4 (C-3'), 149.3 (C-4'), 115.1 (C-5'), 122.0 (C-6'), glucose: 100.8 (C-1''), 74.2 (C-2''), 76.3 (C-3''), 70.4 (C-4''), 77.3 (C-5''), 60.5 (C-6'').

Isorhamnetin 3-O- β - galactopyranoside (8) was obtained as yellowish white amorphous powder of R_f values (x 100): 11 (H₂O), 12 (AcOH-6), 52 (BAW). It showed a Mr of 478 in negative ESI-MS, corresponding to a molecular ion $[M-H]^-$ at $m/z = 477.3$. UV spectral data λ_{max} (MeOH): 256 nm, 300 nm sh., 354 nm, + NaOMe : 275 nm, 325 nm sh., 408 nm, + NaOAc: 275 nm, 323 nm, 371 nm, +

NaOAc + H₃BO₃ : 276 nm, 295 nm sh., 355 nm, +Al Cl₃ : 269 nm, 365 nm, 403 nm. ¹H- NMR spectral data (DMSO-*d*₆) δ (ppm), isorhamnetin moiety: 6.22 (d, *J*=1.6 Hz, H-6), 6.43 (d, *J*=1.6 Hz, H-8), 3.84(s, OCH₃), 7.95 (d, *J*=1.6 Hz, H-2'), 6.93 (d, *J*=8.4 Hz, H-5'), 7.56 (dd, *J*=1.6 and 8.4 Hz, H-6'), galactose moiety: 5.36 (d, *J*=8 Hz, H- 1''), 3.2 - 3.8 (m, H-2''- H-6''). ¹³C-NMR spectral data (DMSO-*d*₆) δ (ppm), isorhamnetin moiety: 156.3 (C-2), 133.2 (C-3), 177.3 (C-4), 161.1 (C-5), 98.7 (C-6), 164.3 (C-7), 93.6 (C-8), 156.2 (C-9), 103.9 (C-10), 55.6 (OCH₃), 120.9 (C-1'), 113.4 (C-2'), 148.4 (C-3'), 149.3 (C-4'), 115.1 (C-5'), 122.0 (C-6'), galactose: 101.79 (C-1''), 71.07 (C-2''), 72.97 (C-3''), 67.81 (C-4''), 75.54 (C-5''), 60.5(C-6'').

Quercetin 3-O- rutinoside (9) was obtained as yellowish white amorphous powder of R_f values (x 100): 48 (H₂O), 44 (AcOH-6), 32 (BAW). It showed a Mr of 610 in negative ESI-MS, corresponding to a molecular ion [M-H]⁻ at *m/z* = 609.1. UV spectral data λ_{max} (MeOH): 256 nm, 265 nm sh., 358 nm+ NaOMe : 268 nm, 327 nm sh., 403 nm, + NaOAc: 273 nm, 323 nm, 387 nm, + NaOAc + H₃BO₃ : 262 nm, 377 nm, +Al Cl₃ : 273 nm, 430 nm. ¹H- NMR spectral data (DMSO-*d*₆) δ (ppm), quercetin moiety: 6.0 (d, *J*=2.5 Hz, H-6), 6.2 (d, *J*=2.5 Hz, H-8), 7.4 (d, *J*=2.5 Hz, H-2'), 6.7 (d, *J*=8 Hz, H-5'), 7.5(dd, *J*=2 and 8 Hz, H-6'), glucose moiety: 5.23 (1H, d, *J* = 7.5 Hz, H-1'''), 3-3.8 (m, H-2''- H- 6'''), rhamnose moiety: 4.37 (s, H-1'''), 3-3.8 (m, H-2''''- H- 5'''), 1.10 (d, *J* = 6.0 Hz, H-6'''). ¹³C-NMR spectral data (DMSO-*d*₆) δ (ppm), quercetin moiety: 156.9 (C-2), 133.8 (C-3), 177.8 (C-4), 161.7 (C-5), 98.7(C-6), 164.6 (C-7), 93.9 (C-8), 157.1 (C-9), 104.4 (C-10), 121.6 (C-1'), 115.7 (C-2'), 145.3 (C-3'), 149.0 (C-4'), 116.7 (C-5'), 122.3(C-6'), glucose moiety: 101.2 (C-1'''), 74.59 (C-2'''), 76.9 (C-3'''), 70.5 (C-4'''), 76.4 (C-5'''), 67.5 (C-6'''), rhamnose moiety: 101.7 (C-1'''), 70.8 (C-2'''), 71.1 (C-3'''), 72.36 (C-4'''), 68.7 (C-5'''), 18.4(C-6''').

Kaempferol 3-O- rutinoside (10) was obtained as yellowish white amorphous powder of R_f values (x 100): 49 (H₂O), 47 (AcOH-6), 35 (BAW). It showed a Mr of 594 in negative ESI-MS, corresponding to a molecular ion [M-H]⁻ at *m/z* = 593.1. UV Spectral Data λ_{max} (MeOH): 252 nm, 352 nm sh., + NaOMe : 273 nm, 324 nm sh., 395 nm, + NaOAc: 273 nm, 310 nm, 385 nm, + NaOAc + H₃BO₃ : 266 nm, 355 nm, +AlCl₃ : 272 nm, 392 nm. ¹H- NMR Spectral Data (DMSO-*d*₆) δ (ppm), kaempferol moiety: 6.18 (d, *J*=2.5 Hz, H-6), 6.44 (d, *J*=2.5 Hz, H-8), 7.92 (d, *J*=2.5 Hz, H-2', H-6'), 6.82 (d, *J*=8 Hz, H- 3',H-5'), glucose moiety: 5.23 (d, *J* = 7.5 Hz, H-1'''), 3-3.8 (m, H-2''- H- 6'''), rhamnose moiety: 1.10 (1H, d, *J* = 6.0 Hz, H-6'''), 4.37 (s, H-1'''), 3-3.8 (m, H-2''''- H- 5'''). ¹³C-NMR Spectral Data (DMSO-*d*₆) δ (ppm), kaempferol moiety: 157.4 (C-2), 133.7 (C-3), 178.2 (C-4), 161.5 (C-5), 99.4 (C-6), 164.9 (C-7), 94.3 (C-8), 157.8 (C-9), 104.4 (C-10), 121.3 (C-1'), 131.7 (C-2', C-6'), 161.8 (C-4'), 115.8 (C-3', C-5'), glucose moiety: 100.9 (C-1'''), 74.59 (C-2'''), 76.9 (C-3'''), 70.5 (C-4'''), 76.4 (C-5'''), 67.5 (C-6'''), rhamnose moiety: 102.2 (C-1'''), 70.8 (C-2'''), 71.1 (C-3'''), 71.6 (C-4'''), 70.5 (C-5'''), 18.01(C-6''').

Isorhamnetin-3-O-rutinoside (11) was obtained as yellowish white amorphous powder of R_f values (x 100): 22 (H₂O), 39 (AcOH-6), 48 (BAW). It showed a Mr of 624 in negative ESI-MS, corresponding to a molecular ion [M-H]⁻ at *m/z* = 623. UV Spectral Data λ_{max} (MeOH): 252 nm, 352 nm sh., + NaOMe : 268 nm, 326 nm sh., 405 nm, + NaOAc: 269 nm, 322 nm, 360 nm, + NaOAc + H₃BO₃ : 252 nm, 353 nm sh., +Al Cl₃ : 264 nm, 298 nm sh., 355 nm. ¹H- NMR Spectral Data (DMSO-*d*₆) δ (ppm), isorhamnetin moiety: 3.97 (s, OCH₃), 6.15 (d, *J*=2.5 Hz,H-6), 6.35 (d, *J*=2.5 Hz,H-8), 6.93 (d, *J* = 8.0 Hz, H-5'), 7.5 (d, *J* = 8.0 Hz, H-6'), 7.85 (s, H- 2'), glucose moiety: 5.4 (d, *J* = 7.5 Hz, H-1'''), 3-3.8 (m, H-2''- H- 6'''), rhamnose moiety: 4.53 (s, H-1'''), 3-3.8 (m, H-2''''- H- 5'''), 1.10 (d, *J* = 6.0 Hz, H-6'''). ¹³C-NMR Spectral Data (DMSO-*d*₆) δ (ppm), isorhamnetin moiety: 157.4 (C-2), 134.3 (C-3), 161.9 (C-5), 98.9 (C-6), 165.1 (C-7), 93.8 (C-8), 157.7 (C-9), 104.5 (C-10), 122.7 (C-1'), 113.4 (C-2'), 147.2 (C-3'), 55.6 (C-3'-OCH₃), 149.6 (C-4'), 114.9 (C-5'), 121.8 (C- 6'), glucose moiety: 103.0 (C-1'''), 74.7 (C-2'''), 77.0 (C-3'''), 70.4 (C-4'''),76.2 (C-5'''), 67.3 (C-6'''), rhamnose moiety: 101.3 (C-1'''), 70.8 (C-2'''), 70.9 (C-3'''), 72.7 (C-4'''), 68.6 (C-5'''), 16.7 (C-6''').

Isorhamnetin (12) was obtained as yellowish white amorphous powder of R_f values (x 100): 0 (H₂O), 2 (AcOH-6), 73 (BAW). It showed a Mr of 316 in negative ESI-MS, corresponding to a molecular ion [M-H]⁻ at *m/z* = 315. UV Spectral Data λ_{max} (MeOH): 255 nm, 266 nm sh., 370 nm + NaOMe : 271 nm, 330 nm, 430 nm, + NaOAc: 260 nm sh., 272 nm, 320 nm, 340, + NaOAc + H₃BO₃ : 255 nm,

268 nm sh., 305 nm sh., 375 nm, +Al₃ Cl₃ : 265 nm, 305 nm sh., 360 nm , 430 nm. . ¹H- NMR Spectral Data (DMSO-d₆) δ (ppm): 3.85 (s, OCH₃), 6.2 (d, *J*=2.5 Hz, H-6), 6.45 (d, *J*=2.5 Hz, H-8), 6.95 (d, *J* = 8.0 Hz, H-5'), 7.5 (dd, *J*=2.5 & 7.5 Hz, H-6'), 7.85 (s, H- 2').

Results and Discussion:

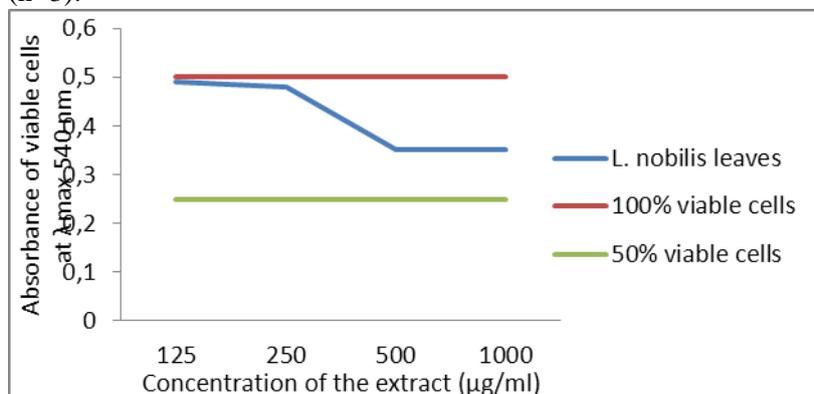
Hepatotoxicity:

The viability assay was applied with a broad range of concentrations of the studied extract of *L. nobilis* leaves (from 125-1000 µg/mL) on monolayer of rat hepatocytes. It revealed that the extract exerts no toxic effect on the monolayer hepatocyte layer, Table (1) and Figure (1).

Table (1): Hepatotoxicity of different concentrations of *L. nobilis* leaf extract.

Sample No.	Concentration µg/mL	Absorbance at (540 λ _{max}) of Neutral red for Viable Cells	Mean	S.D
<i>L. nobilis</i> leaves	125	0.436	0.49	0.07
		0.564		
		0.466		
	250	0.479	0.38	0.10
		0.382		
		0.277		
	500	0.307	0.32	0.05
		0.371		
		0.275		
	1000	0.340	0.28	0.07
		0.196		
		0.293		
Control	100% Viable cells	0.604	0.50	0.06
		0.502		
		0.495		
		0.438		
		0.478		
		0.496		

Figure (1): Viability of monolayer of rat hepatocytes after 2 hrs treatment with different concentrations of the Extract using NR Colourimetric Assay. Each point represents the Mean ± S.D (n=3).

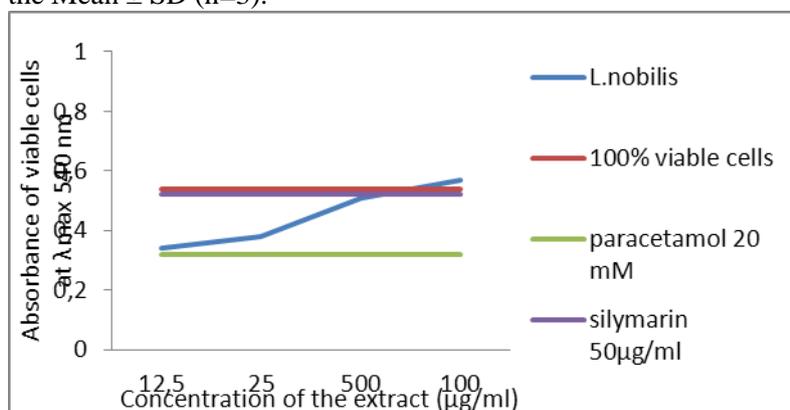


Evaluation of Hepatoprotective Activity Applying Rat Hepatocyte Monolayer:

The hepatoprotective effect of the tested extract against paracetamol toxic effect could be concluded from Table (2) and Figure (2). The *L. nobilis* leaves extract exhibited a hepatoprotective activity at 40 µg/ml.

Table (2): Hepatoprotection of different concentrations of *L.nobilis* leaves under Investigation.

Sample No.	Concentration µg/mL	Absorbance at (540 λmax) of Neutral red for Viable Cells	Mean	S.D
<i>L. nobilis</i> leaves	12.5	0.371	0.34	0.04
		0.293		
		0.360		
	25	0.380	0.38	0.06
		0.321		
		0.443		
	50	0.513	0.51	0.05
		0.548		
		0.457		
	100	0.667	0.57	0.12
		0.610		
		0.443		
Control	100% viable cells	0.563	0.54	0.03
		0.556		
		0.507		
	Paracetamol 20 mM	0.363	0.32	0.03
		0.300		
		0.309		
	Silymarin 50 µg/mL	0.582	0.52	0.06
		0.516		
		0.470		

Figure (2): Viability of Monolayer of Rat Hepatocyte after 2 hrs Treatment with Different Concentrations of the Extract Followed by Treatment with 20 mM Paracetamol for 18 hrs. in comparison with 50 µg Silymarin as Control Using N.R Colourimetric Assay. Each Point Represents the Mean ± SD (n=3).

Identification and Structure Elucidation:

Following column chromatographic fractionation of the *L.nobilis* leaf extract, 14 compounds (1–14) were isolated and identified using conventional and spectral analyses mainly NMR spectroscopy and mass spectrometry.

Chromatographic behavior, UV spectral, ESI-MS (negative mode), ¹H & ¹³C NMR data were consistent with those previously reported for, quercetin 3-O- α -rhamnopyranoside*1 (Babaei H. *et al.*, 2008), kaempferol 3-O- α -rhamnopyranoside 2 (Babaei H. *et al.*, 2008), quercetin 3-O- β -glucopyranoside 3 (Kang *et al.*, 2002), Kaempferol 3-O- β -glucopyranoside*4 (Makuch M. and Awska I., 2011), Quercetin 3'-O- glucoside*5 (Susan and Paul, 1984), quercetin 3-O- β -galactopyranoside*6 (Mahmoud *et al.* 2002), isorhamnetin 3-O- β -glucopyranoside*7 (Olennikov *et al.* 2011), isorhamnetin 3-O- β -galactopyranoside*8 (Sikorska M. and Matlawska I., 2001), quercetin 3-O- rutinoside*9 (Harborne J.B. 1994), kaempferol 3-O- rutinoside*10 (Brochado *et al.* 2003, Cardosa *et al.*, 2005), isorhamnetin 3-O- rutinoside*11 (Harput *et al.*, 2004), isorhamnetin*12 (Cao X. *et al.*, 2009), quercetin 13 (Harborne, J.B., 1994) and kaempferol 14 (Smolarz, HD. , 2002).

*Compounds isolated for the first time from *L. nobilis* leaves.

Conclusion:

L. nobilis leaves proved its capability on synthesizing and accumulating an appreciable number of flavonoid glycosides in its leaves. The extract of these leaves showed hepatoprotective activity against paracetamol toxic effects on rat hepatocytes at low concentration of 40 μ g / ml. These findings require further investigation to isolate and identify the active compounds and to unravel the underlying mechanisms of action.

References:

- Babaei H., Sadeghpour O., Nahar L., Delzar A., Nazemiyeh H., Mansouri M., Poursaied N., Asannshari S., Moghadam S., Sarker S., (2008). "Antioxidant and Vasorelaxant Activities of Flavonoids from *Amygdalus lycioides* var. horrid", *Turk. J. Biol.*, **32**: 203-208
- Baytop T. (1984). "Therapy with medicinal plants in Turkey", No. 3255. Istanbul: Publications of Istanbul University.
- Borenfreund, E, Puerner, J (1985). "Toxicity determined in vitro by morphological alterations and neutral red absorption". *Toxicol. Lett.* **24**:119-124.
- Brochado, C.D.O., Aleida, A.P., Barreto, B.P., Costa, L.P., Ribeiro, L.S., Pereira, R.L. da. C., Koatz, V.L.G. and Costa, S.S. (2003). " Flavonol robinobiosides and rutinosides from *Alternanthera brasiliana* (Amaranthaceae) and their effects on lymphocyte proliferation in vitro". *J. Braz. Chem. Soc.* **14** (3): 449.
- Brown, R.W. (1956). "*Composition of scientific words: A manual of methods and a lexicon of materials for the practice of logotechnics*". Washington, D.C.: Smithsonian Institution Press.
- Brunetton J., Pharmacognosy (phytochemistry medicinal plants). 2nd ed. Paris: Lavoisier Publishing, 1999.
- Cao X., Wei Y., and Ito Y. (2009), "Preparative isolation of isorhamnetin from *Stigma maydis* using high-speed countercurrent chromatography." *J. Liq. Chromatogr. Relat. Technol.* ; **32**(2): 273–280.
- Cardosa C.L., Silva D.H.S., Gamboa I.C. and Bolzani V.d.S.(2005). " New biflavonoid and other flavonoids from the leaves of *Chimarrhis turbinata* and their antioxidant activities". *J. Braz. Chem. Soc.*, **16**, (6B): 1353.
- Dall'Acqua, Stefano; Viola, Giampietro; Giorgetti, Michela; Loi, Maria Cecilia; Innocenti, Gabriella, (2006). "Two New Sesquiterpene Lactones from the Leaves of *Laurus nobilis*". *Chemical & Pharmaceutical Bulletin*; **54** (8), 1187 – 1189.
- Duke, J. A. (1997). "The green pharmacy: New discoveries in herbal remedies for common diseases and conditions from the world's foremost authority on healing herbs". New York, NY: Rodale Press.
- Fiorin C., David B., Fourasti I. and Vercauteren J., 1998. "Acylated Kaempferol Glycosides From *Laurus nobilis* Leaves". *Phytochemistry*, **47** (5): 821-824.
- Harput U. Sebnem , Iclal Saracu Glu and Yukio Ogihara, (2004). "Methoxyflavonoids from *Pinaropappus roseus*". *Turk. J. Chem.* **28**: 761- 766.
- Harborne, J.B. (1994), " The flavonoids: Advances in Research", Chapman & Hall.

- Heinerman, J. (1983). "The complete book of spices, their medical, nutritional and cooking uses". New Canaan, CT: Keats publishing Inc.
- Kang, Hye Won; Yu, Kwang Won; Jun, Woo Jin; Chang, Ih Seop; Han, Sang Bae; Kim, Hee Yun; Cho, Hong Yon, (2002). "Isolation and Characterization of Alkyl Peroxy Radical Scavenging Compound from Leaves of *Laurus nobilis*". Biological & Pharmaceutical Bulletin; **25** (1):102 – 108.
- Kiso Y, Tohkin M, Hikino H (1983). "Assay method for antihepatotoxic activity using galactosamine-induced cytotoxicity in primary-cultured hepatocytes". *J. Nat. Prod.*, **46**, (6): 841-847.
- Mahmoud, I.I., Marzouk, M.S.A., Moharram, F.A., Notle, J., Fobbe, R. and Saleh, M.I. (2002), "Polyphenolic constituents of *Callistemon lanceolatus* leaves", *Pharmazie*, **57**, (7): 494-496.
- Makuch M. and Awaska I. (2011), "Flavonoids from the flowers of *Aesculus Hippocastanum*", *Acta Poloniae Pharmaceutica n Drug Research*, **68** (3): 403-408.
- Marino, Simona De; Borbone, Nicola; Zollo, Franco; Ianaro, Angela; Meglio, Paola Di; Iorizzi, Maria, (2004). "Megastigmane and phenolic components from *Laurus nobilis* L. leaves and their inhibitory effects on nitric oxide production". *Journal of Agricultural and Food Chemistry*; **52**, 7525 - 7531
- NIH Guide for the Care and Use of Laboratory Animals. NIH *Publication* Number. 85-23, Revised 1985, U.S. Department of Health, Education and Welfare. Research Triangle Park, North Carolina
- Olenikov D.N., Tankhaeva L.M., Partilkhayev V.V., Rokhin A.V., (2011). "Chemical constituents of *Caragana bungei* shoots". *Brazilian Journal of Pharmacognosy*, under press.
- Pech, B. and Bruneton, (1982) "Alkaloids from *Laurus nobilis*". *J., Journal of Natural Products*, **45**: 560.
- Sakar MK, Engelshowe R (1985) "Tannin Producing Monomeric and Dimeric Substances in Bay Leaves *Laurus-Nobilis*". *Zeitschrift fuer Lebensmittel-Untersuchung und -Forschung* 180, 494-5.
- Seglen PO (1976). "Preparation of isolated rat liver cells". *Methods Cell Biol.*, **13**: 29-83.
- Sellami I.H., Wannas W.A., Bettaieb I., Berrima S., Chahed T., Marzouk B. and Limam F., (2011). "Qualitative and quantitative changes in the essential oil of *Laurus nobilis* L. leaves as affected by different drying methods". *Food Chemistry* **126**: 691–697.
- Sezik E, Yesilada E.(1999), Turkish folk medicine containing volatile oils. In: Kirimer N, Mat A, editors. *Essential oils*.
- Sikorska M. and Matlawska I., (2001). "Kaempferol, isorhamnetin and their glycosides in the flowers Of *Asclepias syriaca* L." *Acta Pol. Pharm.*, **58**: 269-272.
- Smolarz HD. (2002), "Flavonoids from *Polygonom lapatifolium* ssp.tomentosum". *Pharm. Biol.* **40**: 390-394).
- Susan K. Waage ,Paul A. Hedin (1984), "Biologically-active flavonoids from *Gossypium arboretum*", *Phytochemistry*, **23** (11): 2509–2511.
- Zola, A., Le Vanda, J. P. and Guthbrod, F. (1977), " L'Huîle Essentielle de *Laurier Noble*". *Plantes Medicinales et Phytotherapie*, **11**:241.

LARYNX PATHOLOGY AT PATIENTS WITH GASTRO-INTESTINAL TRACT DISORDERS

Toguzbaeva Dinara Yerkenovna

Candidate of Medical Science, Assistant of otorhinolaryngology department, Almaty State Institute of Doctors Advanced Training, Republic of Kazakhstan

Abstract:

The conducted clinical and instrumental study revealed that the aggregate group of the patients with chronic diseases of throat differs by increased frequency of isolated and combined disorders of gastro-intestinal tract. If a patient has such symptoms as hoarseness, tickling, scratching, burning, feeling of dryness in throat, frequent variant of gastro-intestinal tract disorders is gastroesophageal reflux disease with pharyngolaryngeal reflux. There are such severe pathologies of throat as pachydermia of vocal cords and nodules of vocal cords among more frequently met disorders at gastro-intestinal tract diseases. High probability of accompanied pathology of throat in the form of pachydermia and nodules of vocal cords at patients with gastro-intestinal tract disorders stipulates for the necessity of conducting fibrorhinolaryngoscopy in combination with fibergastroscopy.

Kew Words: Pachydermia, pathology of throat, gastro-intestinal tract disorder, interrelation, video image endoscopy

Topicality of the theme:

Long-term course of the functional diseases of throat, especially at people of voice and speech professions, leads to organic pathology of vocal apparatus. The mucous membrane of throat, larynx and gastro-intestinal tract represents non-separable whole, therefore, laryngitis in many cases develops simultaneously or subsequent to chronic gastritis or cholecystopancreatitis. One of the reasons of such symptoms occurrence is pharyngolaryngeal reflux leading to refluxate entering into the area located above the upper esophageal sphincter [O.P. Alekseyev and others, 2006; V.F. Privorotskiy, N.E. Luppova, 2000]. Lately the manifestations of gastroesophageal reflux disease outside esophagus, especially from the part of bronchopulmonary system and ENT-organs, attract more and more attention. The cause-effect relations of gastro-intestinal tract disorders with pathology of throat in spite of their frequent combination [A.D. Svishev, 2003; Rosbe K.W. et al., 2003] have been studied insufficiently fully.

Uncertainty of the problem related to adequate diagnostics of gastro-intestinal tract disorders at patients with the clinical picture of vocal dysfunction at throat pathology gave rise to the goal and determined the tasks of this work.

Goal of study – is to study interrelation between gastro-intestinal tract disorders at the patients with different vocal dysfunction.

Tasks of study: 1. To study interrelation between gastro-intestinal tract disorders and throat diseases and vocal dysfunction.

2. To define the priority of throat diseases and vocal dysfunction diagnostics methods based on complex clinical and instrumental study.

3. To conduct the comparative analysis of frequency of occurrence of gastro-intestinal tract disorders options depending on pathology of throat.

Materials and methods: The clinical and instrumental examination of 99 patients at the age between 50 and 60 years ($8,0 \pm 2,6$ years) – 62 male and 37 female with chronic diseases of throat with different disorders of gastro-intestinal tract or combination of these nosologies was held. The gastroenterological group included 39 patients with disorders of upper gastrointestinal tract according to the data of physical, endoscopic, X-ray examinations. The criterion of inclusion into otorhinolaryngologic group was presence of chronic diseases of throat in combination with vocal dysfunction. It included 60 patients followed up by the otolaryngologists with pachydermia of vocal cords (PVC, n=38) and “singing nodules” vocal cords nodules (VCN, n=22).

Fibrogastroduodenoscopy was held to evaluate the condition of mucous membrane of esophagus, stomach and duodenum as well as motor defects in the area of esophagogastric junction. The radiographic contrast study of esophagus made it possible to diagnose gastroesophageal refluxes of different degree of manifestation.

All patients were held the endoscopic study of throat and larynx by means of Karl Storz video endoscope with camera with 90⁰ angle of vision. Analysis of vocal function was held by means of Speech analyzer computer program. The patients' speech was recorded for objective study of vocal function. The patients called name, surname and patronymic, year of birth and date of study and then pronounced the vowel sounds "A", "И" and "Y" in a drawling manner. Analysis of signal was held by means of one of the options of software, this program is able to reveal the abnormalities in vocal range without specification of such acoustic parameters. This was made it possible to evaluate the presence of deviations in vocal function condition and to reveal the necessity of more detailed analysis.

In the otorhynolaryngologic group the endoscopic study was the basic one in establishing the diagnosis of chronic pathology of throat. In the course of study the endoscopic signs of pharyngolaryneal reflux were also defined: hyperemia and edema of mucous membrane of arytenoid cartilages and interarytenoid area, hyperemia of back sections of vocal folds, inflammable rollers in under folds section and pachydermia. In the course of study 93% (n=92) of patients from both groups complained about hoarseness, tickling, irritation, scratching, burning, and feeling of dryness in throat. The terms of disease were between 1 month and 20 years.

The statistical processing of obtained data was held at personal computer using Statistica 6.1, Microsoft Excel 2003 programs according to the medical statistics recommendations.

Following the results of study, the following was revealed: isolated disorder of upper sections of gastro-intestinal tract, combination of gastro-intestinal tract disorders and pathologies of throat, isolated disease of throat. The patients with combination of gastro-intestinal tract disorders and throat pathologies (n=56 – 56,6%) amounted the most part among examined patients (n=99). The isolated diseases of throat and gastro-intestinal tract were observed rarely (n=14 – 14,1% и n=14 – 14,1%, accordingly).

The analysis of the degree of throat pathology influence upon detectability of gastro-intestinal tract disorders was also of interest. The accumulative group of the patients with throat pathology included 60 patients followed up by otolaryngologists with above mentioned ENT diseases and 40 patients with initial diagnosis of reflux disease.

Pachydermias (n=33) and nodules of vocal folds (n=7) in combination with dysphonia were diagnosed at the patients with initial diagnosis of reflux disease. Pachydermia was located on vocal folds – 21 (63,7%), vestibular folds – 8 (24,2%), in interarytenoid area or near to arytenoid cartilages – 4 (12,1%). The following was revealed by means of endoscopy: hyperemia, edema of mucous membrane of arytenoid cartilages and interarytenoid area, hyperemia of back sections of vocal folds, inflammable rollers in under folds section and pachydermia n=33 (82,5%). Pachydermias had warty structure; they were similar to outgrowth or patchers. Coloring of pachidermia varied from whitish-grey to bright yellow or pink. Thus, frequency of revealing the specified pathological states of throat at the patients with gastroenterological practice was 66,7%.

The symptomatic analysis at the examined patients established certain influence of reflux disease option and gastro-intestinal tract disorder upon the clinical implications of pathology. There were regurgitation and vomiting (42,9% and 21,4%) at isolated disorder of gastro-intestinal tract. Vocal dysfunction was met among symptoms outside esophagus at the majority of patients at isolated disease of throat and combined option of reflux disease (80,0% and 66,7%, accordingly). The signs of endoesophagitis (77,8%) without evident frequency differences depending on reflux disease option were revealed at the majority of the patients who was held fiber optic esophagogastroduodenoscopy. Reflux disease with esophagitis signs (64,8%) was prevailing pathology of esophagus and without evident inflammatory changes of mucous membrane (22,2%).

The conducted acoustic analysis of voice (functionally limited under technical parameters) revealed by means of software the vocal dysfunction at all patients in different degree of manifestation. The radio graphic contrast study at 1/3 of patients didn't reveal the signs of contrast entering into esophagus (33,3%). The same study poorly reflected the level of reflux. The

gastroesophageal reflux of 3-4 degree was met insignificantly more often at combined option of reflux disease (26,9%) in comparison with gastroesophageal reflux disease (14,3%).

It is expedient to consider the revealed association of gastro-intestinal tract disorder with the diseases of throat from the point of view of cause-and-effect relationships. In particular, disorders of upper sections of gastro-intestinal tract are able to play a part of provoking factor of throat pathologies and to be its consequence by damaging the mechanisms of clearance of mucous membrane of respiratory tract upper divisions [Contencin P. et al, 1999, Wierzbicka M. et al, 2003].

The revealed patterns are likely to have interdependent nature and may reflect different pathogenetic mechanisms of ENT-pathology or degree of its influence upon gastro-intestinal tract disorders.

Conclusions:

1. The frequent option of gastro-intestinal tract disorders is gastroesophageal reflux with pharyngolaryngeal reflux.
2. The accumulative group of the patients with chronic diseases of throat differs by increased frequency of isolated and combined disorders of gastro-intestinal tract.
3. Presence of severe pathology of throat, e.g., pachydermia of vocal folds and nodules of vocal folds, is the most often pathology met at gastro-intestinal tract diseases.
4. High probability of accompanied pathology of throat in the form of pachydermia and nodules of vocal folds at the patients with gastro-intestinal tract disorders stipulates for necessity to conduct fibrorhinolaryngoscopy in combination with gastrofibroscopy.

References:

1. N.A. Agadzhanian Influence of phonatory load upon indexes of cardiorespiratory system at hypotonia dysphonia / N.A. N.A. Agadzhanian, U.S. Vasilenko, A.I. Smirnova // Bulletin of otorhinolaryngology. – 2005. No 4 – page 17.
2. T.N. Agapova. State of vocal apparatus at patients with diseases of nose and paranasal sinuses (clinical picture, diagnostics, treatment) / T.N. Agapova. Candidate of Medical Science. M., 1997. – page 130.
3. T.N. Agaphonova. Psychoemotional status at patients with functional impairment of voice / T.N. Agaphonova // Russian otorhinolaryngology. 2002 – No 1 – Page 7-8.
4. H.A. Alimetov. Nodules of vocal folds / H.A. Alimetov, Z.H. Shafiyeva // Materials of XVII meeting of Russia otolaryngologists. – Nizhniy Novgorod, 2006, page 56.
5. U.S. Vassilenko. Functional status of throat from the point of view of age / U.S. Vassilenko, G.F. Ivanchenko // Ear, throat and nose. 1977. –No 23.- page 75-79.
6. S.S. Gerassimova. Professional diseases of throat at people of voice and sheep professions (diagnostics, clinical picture, treatment, dispensary observation) / S.S. Gerassimova // Synopsis of thesis, Candidate of Medical Science – Khabarovsk, 1972. – 20
7. L.A. Zaritskiy. Practical phoniatics / L.A. Zraitskiy, V.A. Tripos, L.A. Tripos. Kiev: Higher School. – 1984 – page 186.
8. Outside esophagus manifestations of gastroesophageal reflux disease at children // Gastroenterology Saint-Petersburg, No 1-2, 2006. – M55 (coauthor N.K. Shumeiko, N.B. Serebrovskaya, N.V. Sorokina, I.S. Strizhova).
9. Influence of gastroesophageal reflux upon condition of throat at children // Materials of XVII meeting of Russia otolaryngologists. – Nizhniy Novgorod, 2006. – Page 468-469 (coauthor I.E. Pogossova, U.L. Soldatskiy, E.K. Onufriyeva, N.K. Shumeiko, N.B. Serebrovskaya, A.M. Steklov, S.F. Gasparyan)
10. Clinical peculiarities of pharyngolaryngeal reflux at children with chronic pathology of throat // Materials of VII Russian Congress “Modern technologies in pediatrics and children surgery”. – Moscow, 2008. – Page 394 (coauthor V.I. Kirilov, E.K. Onufriyeva, U.L. Onufriyeva, U.L. Soldatskiy, I.E. Pogossova).

ASSESSMENT OF THE REACTIVE HYPERAEMIA THROUGH QUANTIFICATION OF COLOUR CHANGES USING PIXEL ANALYSIS SOFTWARE

T. Tanwi, MBBS Student

Kasturba Medical College, Manipal University, Manipal, Karnataka, INDIA

S. Manjunatha M.D.

Professor of Physiology, Department of Physiology,
Kasturba Medical College, Manipal University, Manipal, Karnataka, INDIA

B. Priya, M.D student

Department of Medicine, Bangalore Medical College, Bangalore, Karnataka, INDIA

Abstract:

A sound cardiovascular system is essential for human health. There are many effective tests to measure cardiac performance, but no matching test exists to detect vascular reactivity. Hence the acute need for a sensitive, non-invasive, cost-effective screening test. The present study purports such a method of assessing variation in blood flow by quantifying colour changes visible in palm secondary to reactive hyperaemia for detection of presence of vascular lesions.

Colour changes in palms of test and control hands were recorded in healthy young volunteers (5 males, 5 females) by taking a series of photographs at 5 seconds interval in pre, during and post occlusion periods. Sphygmomanometer was used to produce arterial occlusion in test hand. The colour changes were analysed for red pixels by indigenously developed colour analysis software "Chitradeepam". Mean of values for each hand in above-mentioned periods were calculated for all subjects and corresponding values of test and control hands were compared using student's t test.

No difference was observed in baseline values of both hands. During occlusion, there was significant decrease in test hand colour in females compared to their control hand but not in males. Post occlusion, significant colour increase was observed in test hands of males compared to control. A similar trend, though statistically not significant was seen in females.

This technique can be useful in determining normal vascular response in healthy individuals through quantification of reactive hyperaemia which in turn can be a parameter to assess the decreased vascular response in Peripheral Vascular Disease.

Key Words: Vascular function, Pixel analysis software, Simple screening test

Introduction:

The cardiovascular system, along with the respiratory system, is the most important organ system for supplying nutrients to the tissues. All the metabolic activities, and in turn the "life" depends on this. Hence the integrity of both the Cardiac and Vascular component of this system needs to be maintained, as malfunction of any one or both will have disastrous effects. There have been many tests for assessment of cardiac performance, but the tests for vascular integrity and reactivity have not matched with that of cardiac function tests, in terms of accuracy, reliability, sensitivity and also ease of administration. The testing of vascular functions has gained importance as the incidence and prevalence of Peripheral Vascular Diseases (PVD) has been increasing (1).

PVD is commonly called peripheral arterial disease (PAD), which refers to the obstruction of large arteries not within the coronary or aortic arch vasculature. It is a manifestation of atherosclerosis characterized by atherosclerotic occlusive disease of the extremities and is a marker for atherothrombotic disease in other vascular beds.

Diabetes is one of the major causes of PAD, prevalence being more than three times higher in patients with Diabetes compared to non-diabetic persons and more so in chronic cases than newly identified ones at the baseline (2, 3). In people with diabetes, the risk of PAD is increased by age, duration of diabetes, and presence of peripheral neuropathy (4). Peripheral vascular disease affects 1 in 3 diabetics over the age of 50. Generally DM induced PAD affects the older age group but now-a-days there is an increased prevalence of the same in younger age groups as well because DM is reported to be increasing dramatically among young individuals. If this change in the epidemiology of DM continues, it is likely that a larger proportion of the youth population will have PAD in the future (5, 6).

If PAD is suspected, a number of tests need to be performed to detect the presence of atherosclerosis, as well as to localize areas of stenosis and to estimate the degree of the stenosis. Two types of techniques are used for the assessment of a patient with PAD: Noninvasive tests and Invasive tests. The main vascular laboratory tests are:

- ◆ Doppler velocity wave form analysis
- ◆ Angiography
- ◆ Color-assisted duplex ultrasonography
- ◆ Magnetic resonance angiography (MRA)
- ◆ Computed tomography angiography (CTA)

Each of these tests has its own advantages and disadvantages. In general the major disadvantages are high cost; difficult to carry out as they are technique intensive; non availability at the peripheral centers and even risks of adverse reaction and transmission of diseases (7).

Rationale for the present study

Data on the prevalence of diabetes induced PAD in the primary care setting are sparse, although this information is critically important as a scientific basis for developing strategies to enhance treatment of this condition and prevention of cerebrovascular and cardiovascular events in the community. Primary care is the principal target for investigation if the aim is improved population-based care. Primary care doctors play a key role, as they are the first point of contact for recognition, diagnosis and referral. Due to the availability of modern pharmacological and adjunctive therapy they are also increasingly important for the treatment of PAD.(8,9) However, there are several issues that urgently need to be addressed with new data in order to help design rational strategies to further improve the service provision and quality of care for PAD patients and of topmost priority in this regard is the large scale accessibility and availability of a non-complicated, noninvasive and less expensive method of diagnosis which can be easily and accurately handled even with minimal expertise.

The present study purports a simpler and non-invasive method of assessing the “change in the colour of the palmar-aspect of the hand with ischemia” and “reactive hyperaemia” as a diagnostic tool for the detection of presence of PAD in chronic diabetics.

Aim of the study

Qualitative and semi-quantitative assessment of the changes in the blood flow due to the phenomenon of the reactive hyperaemia by quantifying the colour changes as seen in the palm using a simple and non-invasive technique.

Objective

To observe and quantify the change in the colour of the palm with ischemia and reactive hyperaemia and compare with the other palm which is taken as control. This is done with the help of pixel analysis software “Chitradeepam” which is indigenously developed in the department of Physiology, K M C Manipal.

Methodology

Subjects

Healthy young students (5 males and 5 females) and Patients with Type 2 Diabetes Mellitus of more than 5 years’ duration, who were interested to volunteer in the study of “assessment of reactive hyperaemia through quantification of the colour changes” which was conducted in the department of physiology; Kasturba Medical College (KMC); Manipal; were selected as subjects. The study has been completed on healthy volunteers only.

Criteria for Selection of control subjects

- ◆ Healthy young adults in the age group of 20-25 years
- ◆ Subjects suffering from IDDM, HTN or any other ailment were excluded.
- ◆ Gender equality was maintained while choosing subjects.
- ◆ Subjects with fair complexion were preferred for sake of convenience.
- ◆ Subjects with any deformity and disfiguration in their palms were avoided.
- ◆ Too obese or too underweight subjects were excluded from the study.

Procedure

The experiment was done during the day time. The subject was sitting comfortably in a chair, completely relaxed in the air conditioned research laboratory and the full procedure was explained to the subjects. If willing he/she had filled in the consent form and only after that the procedure was started. After five minutes of rest, his/her blood-pressure was measured (using mercury sphygmomanometer) in the left arm and reported as systolic and diastolic-pressure. The mean of the three readings was used later for producing arterial occlusion. After five minutes of recording blood-pressure the subject was keeping the hand on the table on the area marked in such a way that both the palms would come in one frame of the camera (Sony digital camera, two mega pixel resolution was used) which was fixed on the camera stand. First four photos were taken at an interval of thirty seconds, which was used only as the base line data. After the base line recording, pressure in the left arm was increased 20 mm Hg above the systolic blood-pressure and maintained at a constant level. Photos were clicked at an interval of every thirty seconds until the subject complained of any discomfort. Whenever the subject complained of any kind of discomfort in the left hand, the pressure was immediately released completely. Photos were clicked of both the hands at an interval of around 5 seconds for a period of 2 minutes which was a total of 30 photos after the release of occlusion which corresponds with the period of the reactive hyperaemia. The cuff of the sphygmomanometer was removed and the subject was allowed to leave after five minutes. The images in JPEG format obtained were transferred to the laptop and converted to BITMAP format. Using the CHITRADEEPAM PIXEL ANALYSIS SOFTWARE for each photo the number of red pixels in each photo was counted for the left hand and compared with the right hand values. The area selected for each hand for analysis was equal. Thus two sets of values were obtained i.e. one for the control hand and one for the experiment hand. The values were incorporated into the bar diagrams using MICROSOFT EXCEL and then documented. The results were analysed and accordingly interpreted also.

Statistical Analysis

The mean of the first four photographs for each hand represented the baseline value. The mean of all the photos during occlusion, which is of variable duration, represented the occlusion value. During post occlusion, the first minute value was separately calculated by analysing first 15 photos and the mean of the 15 photos represented the change at the end of first minute after release. Then the mean all the photographs of the post occlusion, in total 30 photographs including the first 15, represented the mean change of the post occlusion period.

The percentage change in the colour was analysed for first minute post occlusion and also at the end of 2 minutes' by calculating the change by comparing that respective value with the mean baseline value of that hand.

The mean values of the test hand of all the subjects (Males and Females) were compared with the corresponding values of the control hand using Student's 't' test for paired observations and $p < 0.05$ was considered significant.

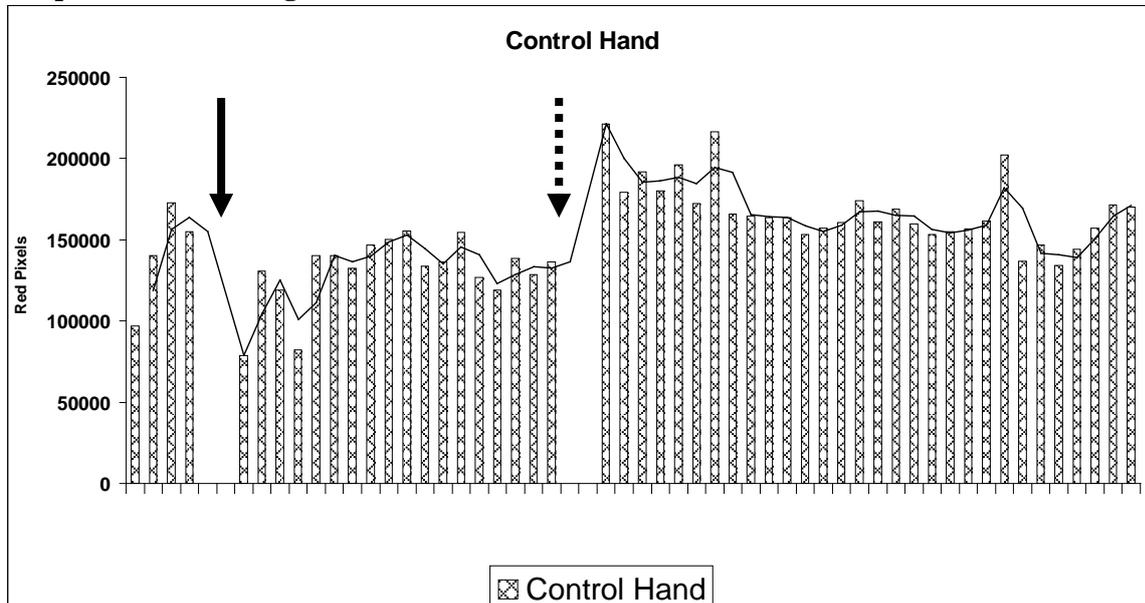
Results

The study was conducted in 10 healthy subjects who were free from any kind of diseases and selected based on the inclusion and exclusion criteria already mentioned. Out of the 10 subjects selected, 5 were males and the other 5 were females.

Before the experiment was started, the height, weight and the baseline blood pressure (reported as systolic/diastolic) of each subject was recorded. Three readings were taken for the blood pressure and the mean of the three readings was used for producing the arterial occlusion. For each subject, 3 sets of values were obtained, i.e. baseline data (4 values before occlusion), during ischemia (variable number of values) and post-ischemic period (30 values) for the control and test

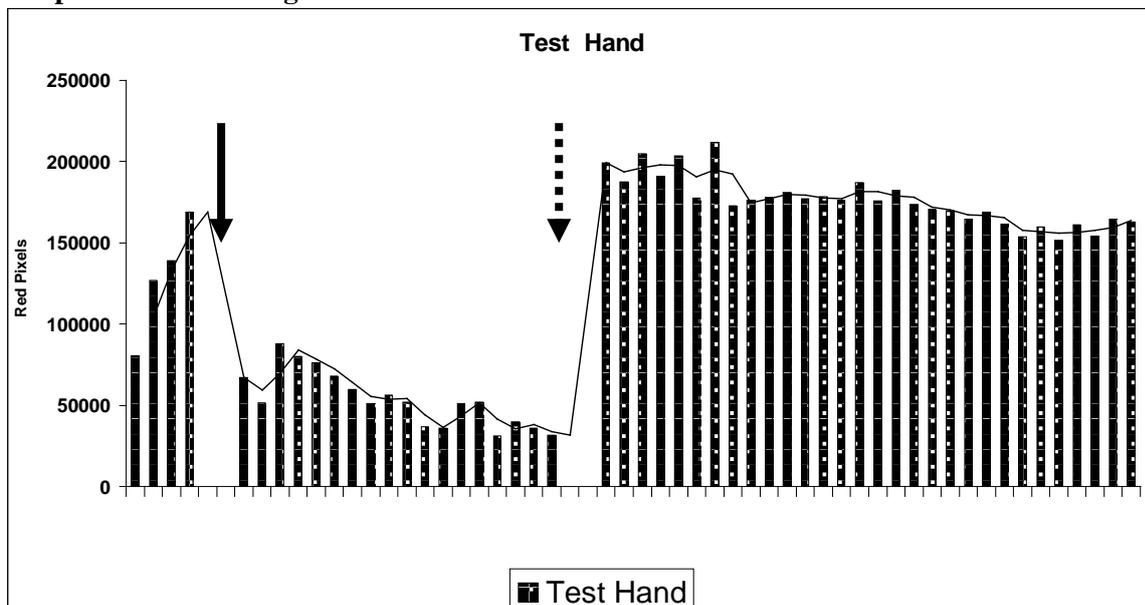
hand separately. A representative graph of an individual for the control hand, the test hand and for both hands together is given below.

Graph 1: Colour changes observed in the control hand



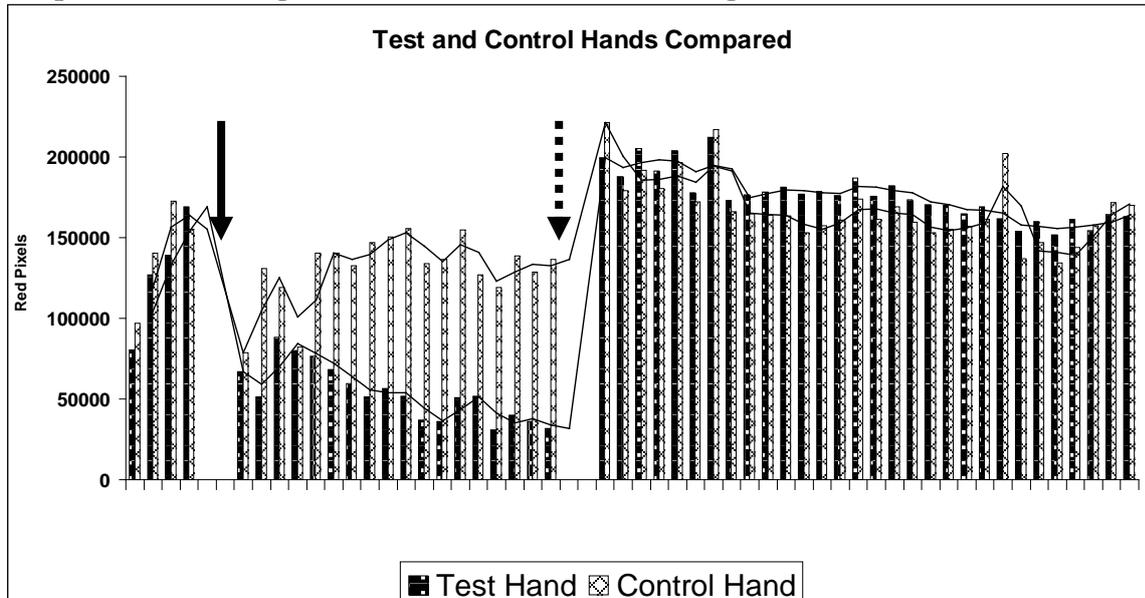
Solid Arrow– Occlusion started, Broken Arrow – Occlusion released

Graph 2: Colour changes observed in the test hand



Solid Arrow– Occlusion started, Broken Arrow – Occlusion released

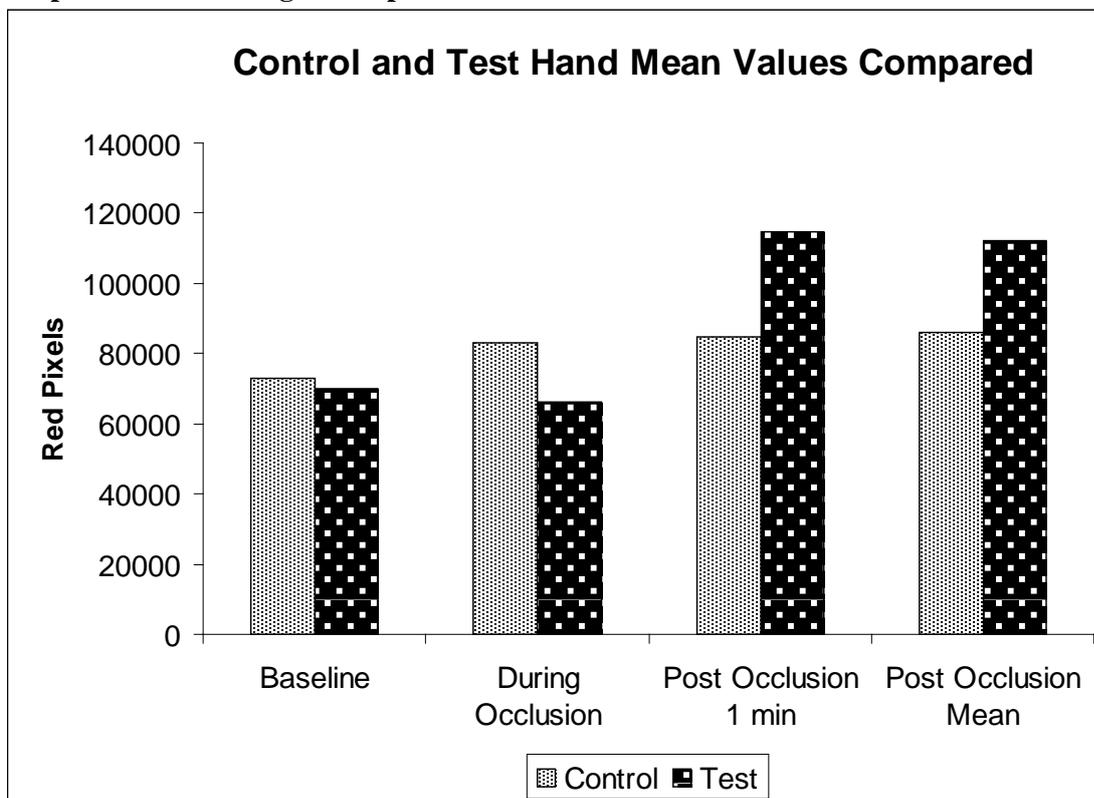
Graph 3: colour changes observed in both hands taken together



Solid Arrow– Occlusion started, Broken Arrow – Occlusion released

The mean was calculated for each of these stages separately. In the test hand, mean of the first four values (for all the subjects) served as the baseline data for that subject only. The same procedure was carried out during the period of ischemia, though this time period varied in each individual, depending up on the time for which they were able to tolerate it comfortably. During the post-ischemic period, mean was calculated for the first 1 minute (first 15 photos) and again an average of all the 30 photos in the post-ischemic period was calculated separately. Similarly the procedure was carried out in the control hand of all the subjects. A representative graph is shown below.

Graph 4: Colour changes Compared



All the mean values of each stage of control hand and test hand of all the 5 subjects in each group were tabulated. The mean and the standard deviations of those values were calculated and the paired t-test was used for the statistical-analysis. The values obtained were analyzed as shown below. Table 1 and 2: Comparison of the values obtained in the control and the test-hand taken together and analyzed using Paired student's 't'- test.

Table 1: MALE SUBJECTS

No	Topics	Control hand (mean ± SD)	Test hand (mean ± SD)	p - value
1.	Baseline	72778.3 ± 12602.07291	70105.5 ± 9010.29096	0.42168066
2.	During ischemia	83144.93953 ± 41439.19732	66203.54091 ± 20965.78187	0.37156512
3.	Post- ischemia (1min)	84710.45333 ± 25886.14198	114794.3867 ± 29444.83298	0.0195332*
4.	Post- ischemia (mean)	85911.8423 ± 24542.78987	112351.9582± 25356.95444	0.0280362*
5.	% Change Pre Vs Post1 min	21.5662395 ± 54.33920012	67.53944658 ± 58.14007126	0.0045972*
6.	% Change Pre Vs Post Final	22.58375408 ± 49.88107477	63.55933163 ± 50.84734894	0.0067888*

(SD- standard deviation, *-p < 0.05 which is taken as significant).

Table 2: FEMALE SUBJECTS

No	Topics	Control hand (mean ± SD)	Test hand (mean ± SD)	p - value
1.	Baseline	101725.9 ± 35693.69725	94673.4 ± 30530.50848	0.060679
2.	During ischemia	115187.6133± 38581.35804	68826.84 ± 31855.1013	0.0250490*
3.	Post- ischemia (1min)	120308.4933± 47522.36298	132029.1333 ± 43562.83392	0.3852659
4.	Post- ischemia (mean)	117744.1044± 44672.04738	126036.9687 ± 45781.67196	0.4807511
5.	% Change Pre Vs Post1 min	19.87884078± 26.87505435	39.52435222 ± 15.18476165	0.30651334
6.	% Change Pre Vs Post Final	16.99207658± 21.18426156	31.25447227 18.89916251	0.4444394

(SD- standard deviation, *-p < 0.05 which is taken as significant).

After analysis of all the 10 subjects, the following observations are made.

1. There are no significant differences in the colour of the baseline photos of the control hand and the test hand in both the groups.

2. During the period of ischemia, the two groups showed difference. In males, though there is a decrease in the colour of the test hand as denoted by the number of red pixels in the test hand, but the colour change is statistically not significant both when compared with the baseline data as well as with control hand. But in females, there was a significant difference between the test and the control hands during ischemia. This seems to be due to the combined effects of decrease colour in the test hand and a simultaneous increase in colour in the control hand, the increase denoting an increased blood flow during the occlusion of the test hand. **This is an interesting and novel observation.** The similar change was also seen in males.
3. In males during the first one minute of post-ischemia the test hand showed marked increase in the number of the red pixels, compared to the baseline and in the next 1 minute also, there was a significant increase in the number of the red pixels but the colour changes were less when compared to the first 1min. The average of the post-ischemic period values showed a significant increase in the number of red pixels (as indicated from the values) when compared with the baseline and also with corresponding values of the control hand. There was a similar trend in the females, though it was not to the same extent as seen in males and hence, was not statistically significant. This showed that there might be hormonal component to this effect, which needs to be further investigated.
4. The control hand also showed similar changes in the post ischemic period as that of test hand, but the magnitude was very less and statistically not significant compared to the baseline values.

Discussion and Conclusion

The transient increase in the blood flow that follows a brief period of arterial occlusion is called as the reactive hyperaemia. There are various methods which have been used to analyze the phenomenon of reactive hyperaemia. Some of these procedures are venous occlusion plethysmography, ultrasound-scanning and other methods. But most of these techniques are invasive. Moreover it requires well equipped tools and expertise to perform such techniques. Hence it is difficult to repeat these procedures over-time. Moreover the range of normal values also varied considerably, probably due to differences in the methodological-factors. Due to all these complications, the patient/subject compliance is also affected and it becomes difficult to use such techniques as a teaching tool or continue further researches in the long run.

In the present study, the change in the colour of the palmar aspect of the hand with ischemia was considered to assess reactive hyperaemia. For each subject 3 sets of values were obtained, i.e. baseline data (before occlusion), during ischemia (variable) and post-ischemic period (after the occlusion is released). As expected, there was a decrease in the number of red pixels during the period of ischemia compared to the baseline data, but the post ischemic period showed a significant increase in the number of red pixels, due to the phenomenon of reactive hyperaemia as assessed by the 'Chitradeepam' software used in the study as mentioned before. The colour change was more intense in the first 1 minute after the release of the occlusion. Though this colour change was still high, compared to the baseline in the next one minute as well, the magnitude of the change was less when compared to the initial 1 minute of the post ischemic period probably due to the fact that the blood supply was returning almost to the baseline levels. These finding are similar to the other results previously mentioned and show that the new method used in this study does compare with the other invasive methods previously used to study reactive hyperaemia.(10) Interestingly, in our study it was seen that the control hand also behaved in a similar manner, though the magnitude of the changes were much less when compared with test hand. The rapidity of the response indicates that there could be a neural component mediating this in addition to the metabolites released in the test hand, which are carried by the blood stream to bring about similar changes. This indicates that there is a scope to continue further research in this field to study the behaviour of the non occluded hand and since this method is non invasive and hence can be used extensively.

Our study also showed that there might be gender differences in the response. Male subjects showed a better vascular response to reactive hyperemia compared to females. This needs to be further looked into by studying more subjects and also the possible mechanisms.

Therefore the present study gives us a simple, non-invasive and a novel technique for the assessment of reactive hyperaemia. The test is easy to repeat overtime and the normal range of values

obtained also do not vary much as compared with the other complicated methods. So this method can be used as a teaching tool and also to continue further researches with a better patient/subject compliance in this field like analyzing reactive hyperaemia in patients with diabetes mellitus, peripheral vascular diseases like peripheral neuropathies and other disorders.

Conclusion:

In the light of the present study, it could be concluded that this is a very simple and non-invasive method with a possibility of similar sensitivity as the other procedures which are invasive and complicated being used for the assessment of reactive hyperaemia. Hence this technique should further be extensively tested in all age groups and both in healthy subjects & in patients with peripheral vascular diseases so that it can be used as a teaching tool and also a research tool.

References:

- Criqui MH, Denenberg JO, Langer RD, Fronck A. The epidemiology of peripheral arterial disease: importance of identifying the population at risk. *Vasc Med* 2:221–226,1997.
- American Diabetes Association: Peripheral arterial disease in people with diabetes (Consensus Statement). *Diabetes Care* 26:3333–3341, 2003.
- Jensen SA, Vatten LJ and Myhres HO. The association between diabetes mellitus and the prevalence of intermittent claudication: the HUNT study. *Vasc Med* 13: 239,2008.
- Selvin E, Coresh J, Wattanakit K, Sharrett R, Steffesh MW. HbA1c and Peripheral Arterial Disease in Diabetes. *Diabetes Care* 29:877–882, 2006.
- Hotu S, Carter B, Watson PD, Cutfield WS, Cundy T. Increasing prevalence of type 2 diabetes in adolescents; *Journal of Paediatrics and Child Health* 40:201-204, 2004.
- King H, Aubert RE, Herman WH. Global burden of diabetes, 1995–2025: prevalence, numerical estimates, and projections. *Diabetes Care* 21:1414–1431.
- Donnelly R, David H, London NJM. ABC of arterial and venous disease:- Non-invasive methods of arterial and venous assessment *BMJ* 320:698-701,2000.
- Curt D, Alexander S, Jens R. Allenberg , Harald D, Roman H, Stefan L, David P, Berndt S, Gerhart T, Hans-Joachim T. High prevalence of peripheral arterial disease and co-morbidity in 6880 primary care patients: cross-sectional study. *Atherosclerosis* 172 (2004) 95–105.
- Doubeni CA, Yood RA, Emani S, Gurwitz JH. Identifying unrecognized peripheral arterial disease among asymptomatic patients in the primary care setting. *Angiology*. 2006 Mar-Apr; 57(2):171-80.
- Norgren L, Hiatt WR, et al.; TASC II Working Group. Inter-Society Consensus for the Management of Peripheral Arterial Disease (TASC II). *Eur J Vasc Endovasc Surg* 2007; 33(Suppl 1): S1–75.

EMERGENCY TROLLEYS: AVAILABLE AND MAINTAINED BUT ARE THEIR LOCATIONS KNOWN?

Sawsan Ghonaimy

MBBCH, Primary FFA, MSC, Victoria Infirmary, Glasgow, UK

Yahya Khedr

Medical Student, Victoria Infirmary, Glasgow, UK

Ahmed Boraie

MRCOG, Ninewells Hospital, Dundee, UK

Abstract:

Emergency trolleys are developed and placed in strategic locations to improve the efficiency of the medical teams' response to emergencies. We conducted a survey to assess the team's knowledge of the presence and location of these trolleys in the Ambulatory Care Hospital in Glasgow, Scotland.

The results highlighted a considerable deficiency in the knowledge of these trolleys' locations in the unit.

We anticipate this problem to be much more common than expected. We suggest that similar surveys should be conducted as part of the regular audits in all units and should involve all staff involved in such emergencies.

Key Words: Emergency trolleys, difficult intubation trolley, cardiac resuscitation trolley

Introduction:

There are a number of emergency trolleys put together for easy access of specialist equipment in case of emergencies.¹ The trolleys are developed through national guidelines and local multidisciplinary team discussions. They have named individuals and a clear schedule that ensures they are maintained. Most medical units regularly audit these practices.² The location of these trolleys has to be known by the teams using it. We conducted a survey to check the percentage of doctors that know about the emergency trolleys and their locations. The study was performed in the Ambulatory Care Hospital (ACH) in the Glasgow Victoria Infirmary. The doctors surveyed worked in the Anaesthetic department in the hospital.

Objectives:

The study objectives were to determine the percentage of doctors who know about the presence and the location of the emergency trolleys in the ACH. Furthermore, we aimed to determine the percentage of the different grades of doctors who know about the presence and places of the emergency trolleys.

Methods:

A paper questionnaire was developed and distributed to the ACH Anaesthetists. Their responses were collected and transferred to an electronic spreadsheet and analyzed.

The emergency trolleys in the ACH are the Difficult Intubation trolley, the Cardiac Resuscitation trolley, the Snatch Box (containing the O negative blood), the Malignant Hyperpyrexia trolley and the Major Haemorrhage trolley.

Results:

Information was collected from 31 Anaesthetists who worked in the ACH, over a period of 3 days. They included 11 junior trainees, 4 middle graders and 16 consultants.

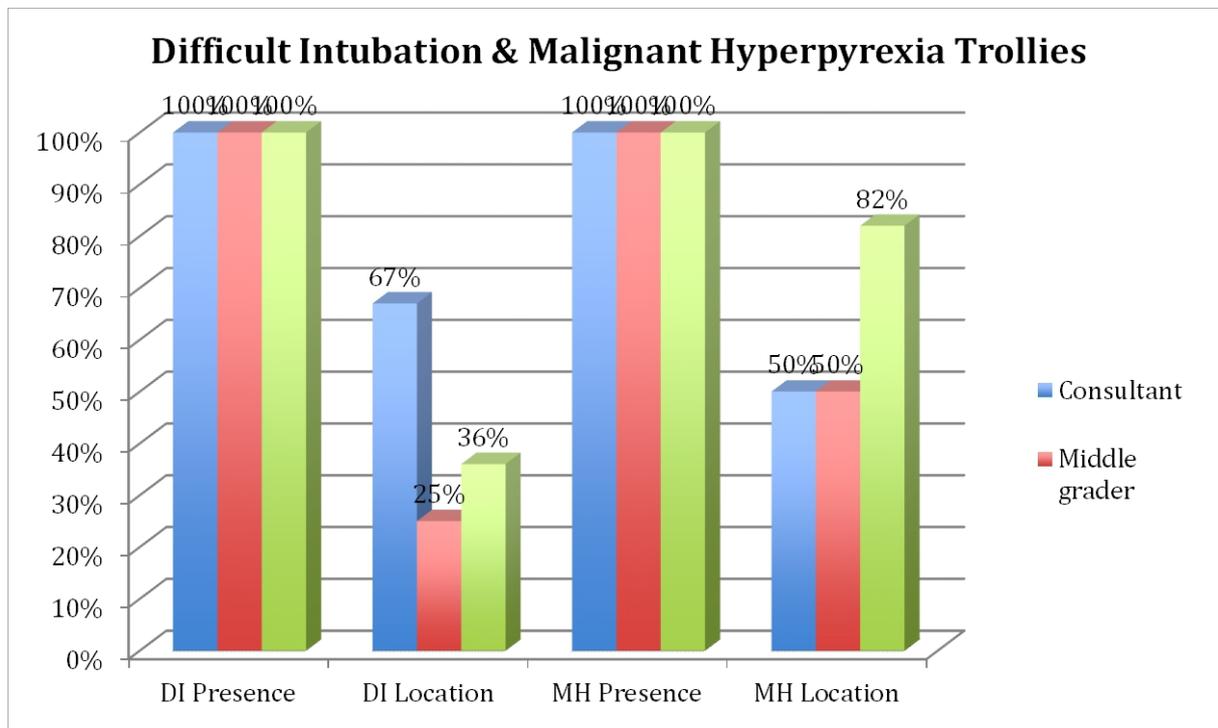
In total, 90% of all doctors knew of the presence of all the emergency trolleys, but only 48% knew of their correct locations.

As regards the overall knowledge of the presence of the emergency trolleys all 3 groups (consultants, middle graders & junior trainees) scored similarly at 92.5%, 80% and 91% respectively. As regards the overall knowledge of the location of the trolleys, the consultants scored worst at 42.5% followed by the middle graders at 45% and the junior trainees scored 58%.

The knowledge of the presence of the different types of trolleys varied.

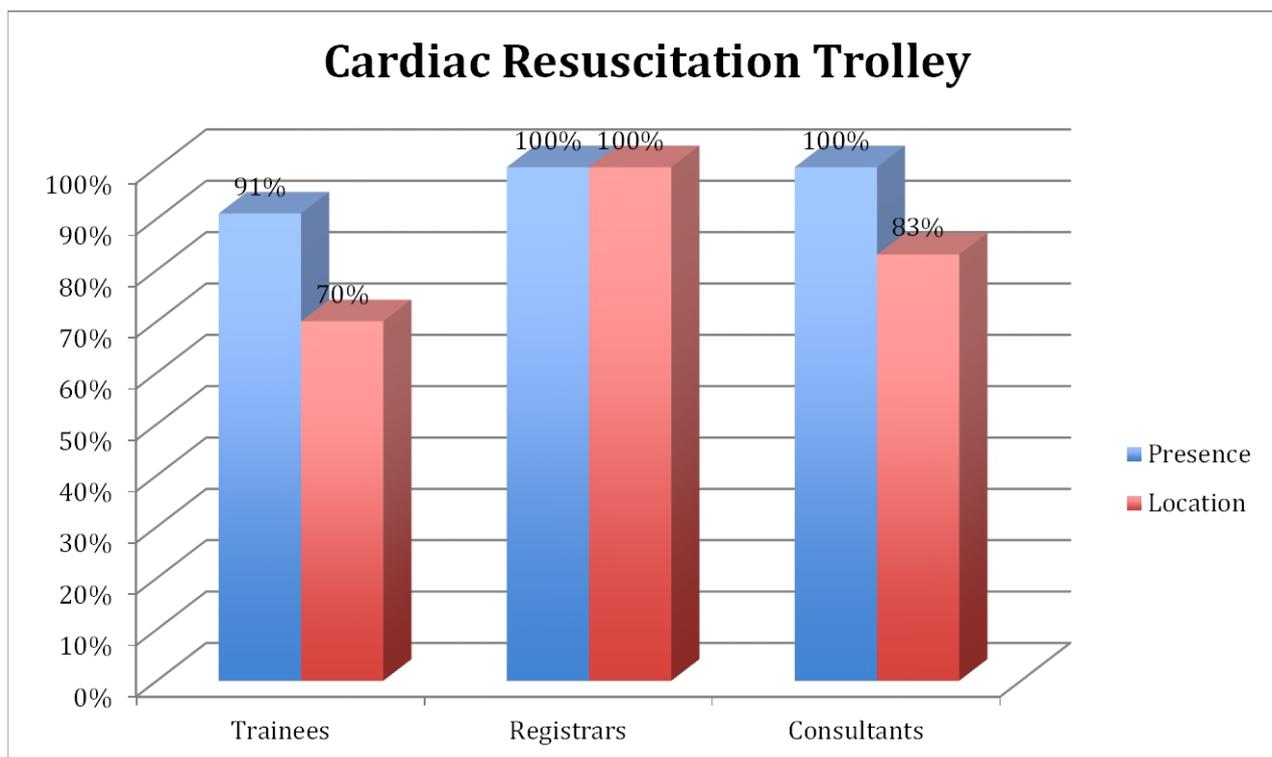
The difficult intubation (DI) and the malignant hyperpyrexia (MH) trolleys scored the highest. All doctors knew of the presence of these trolleys. The location of the trolleys was not as well known and this is highlighted in the following graph.

It is notable that more trainees (82%) were aware of the correct location of the MH trolley than



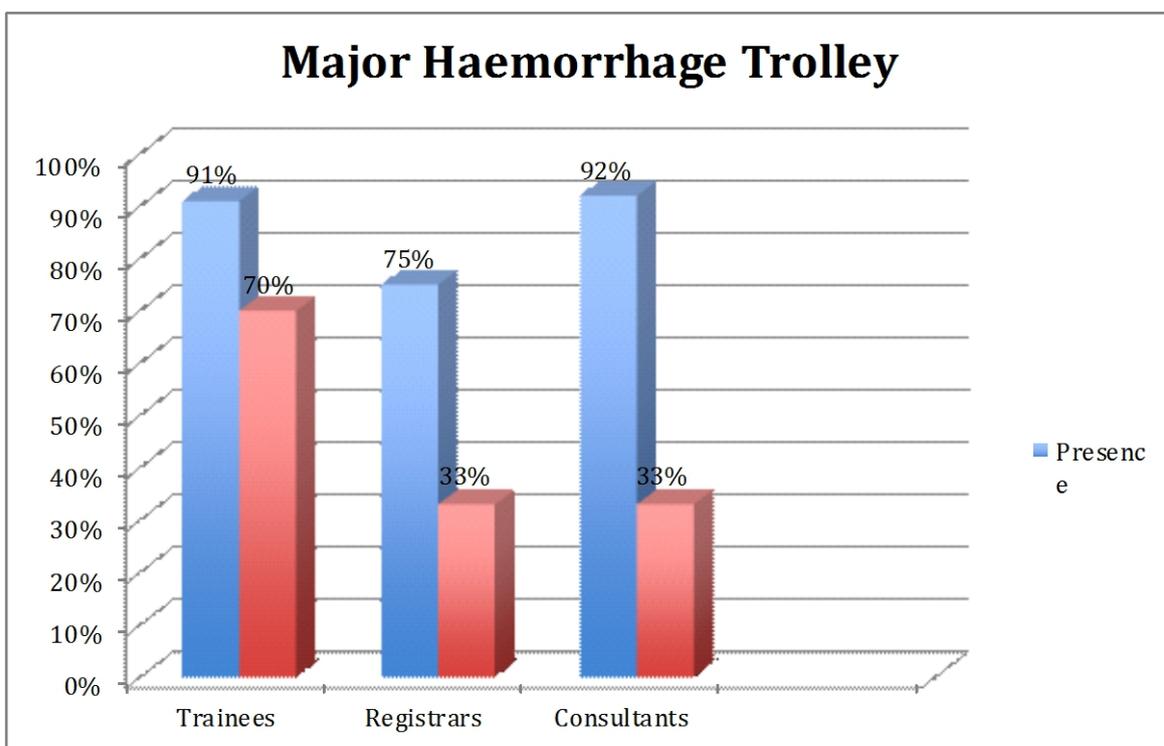
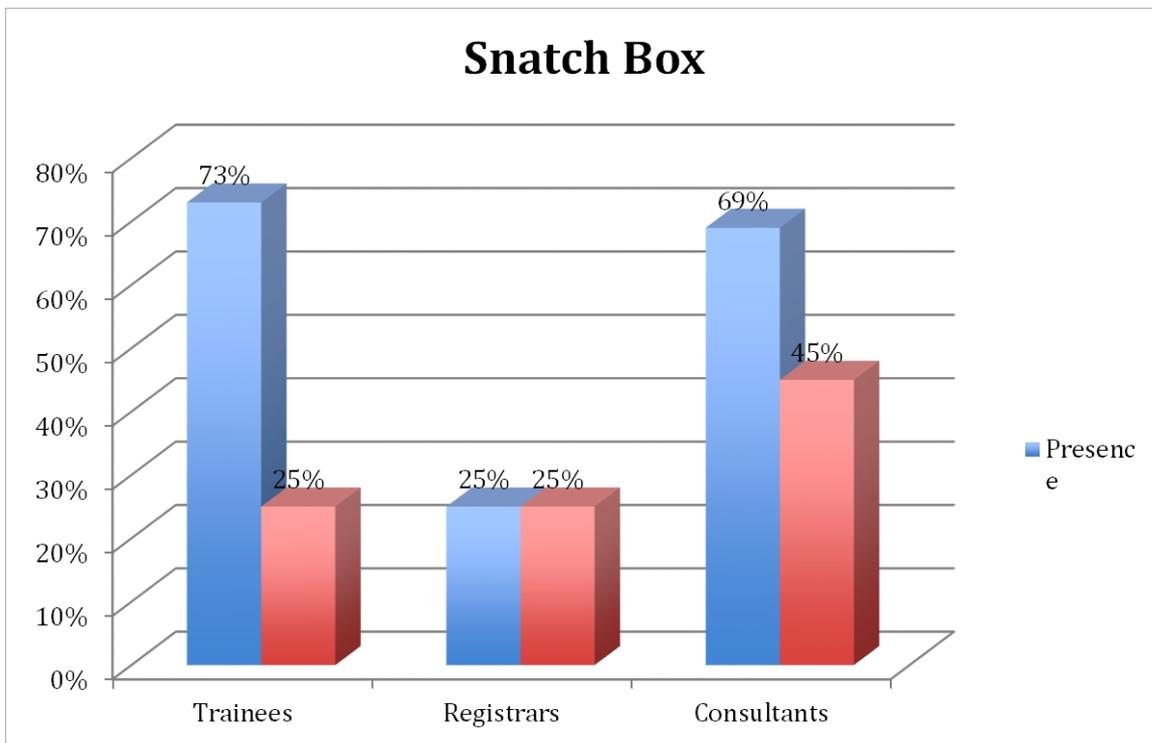
consultants (50%).

The location of the cardiac resuscitation trolley was better known, with 83% of consultants aware of its location.



As for the snatch box’s location, it scored very poorly with only 45% of consultants knowing its correct location and only 69% aware of its presence.

While 92% of consultants were aware of the availability of a major haemorrhage trolley only 33% of the consultants knew of its correct location. Once again more trainees (70%) knew the location of the trolley than consultants (33%).



Discussion:

The emergency trolleys were developed to maximize the efficiency in critical situations when seconds could make a difference for the patients' survival. The trolleys are routinely checked and if used, are restocked. This process is audited regularly. No such emphasis is placed on ensuring that all staff know the location of these trolleys.

Our survey shows clearly that a significant number of both permanent staff and trainees are ignorant of the correct location of the trolleys. In some instances the doctors were not aware of its presence at all.

It is interesting that a greater number of junior trainees were aware of the location of the trolleys than the consultants (42.5% vs. 55% respectively). This may be due to the induction programme the trainees attend at the start of their placement. Their knowledge of the trolleys' location, however, remains generally poor.

We suggest various ways to increase awareness of the presence and location of the trolleys. This includes signs and arrows highlighting their location in the hospital. We also advise a greater degree of involvement of the trainees and consultants in updating and restocking the trolleys. Local study days and emergency drills should include the location of the trolleys and not just the clinical aspect of saving the patients. Lastly, we recommend that this survey should be repeated at regular intervals.

Conclusion:

Our survey identifies an important and dangerous risk factor in the management of emergencies. We advocate the expansion of the survey to include other doctors from different disciplines as well as nursing and paramedical staff who may be called upon to retrieve those trolleys in emergencies.

References:

Dyson, E. & Smith, G.B., 'Common faults in resuscitation equipment-guidelines for checking equipment and drugs used in adult CPR'. *Resuscitation*, 2002.

Smith, A., Kinross, J., Bailey, M., Aggarwal, R., Toresen, D. & Vincent, C., 'Restocking the resuscitation trolley: how good is compliance with, 2008

DRUG SCENE CHANGES DURING TWO DECADES: SLOVAKIA 1993 – 2012

Alojz Nociar

Assistant Professor St. Elisabeth University of Health and Social Work/Department of Psychology,
Slovak Republic

Abstract:

Study presents short overview of drug situation's indicators during the first, or heroin decade of drugs, and then during the second, soft and synthetic drugs decade of drugs in Slovakia. Then the results of series of nation-wide school surveys, mapping development of licit and illicit drugs use among children and youth during nineties are outlined. The results indicated constant growth of licit and illicit drugs consumption among primary school pupils, and secondary school students until the fourth wave of nation-wide school surveys in the years 2006 – 2007. However, the next wave after four years has detected profound change in the expected pattern of further growth in the case of illicit drugs. Instead of this decline of use or at least stabilization occurred among Slovak youth. At the same time we have found growth of use in the case of licit drugs – alcohol and tobacco, more profoundly among girls. Also the use of new synthetic drugs among young people was revealed via traditional school survey.

Key Words: Licit and illicit drugs, school surveys, new synthetic drugs

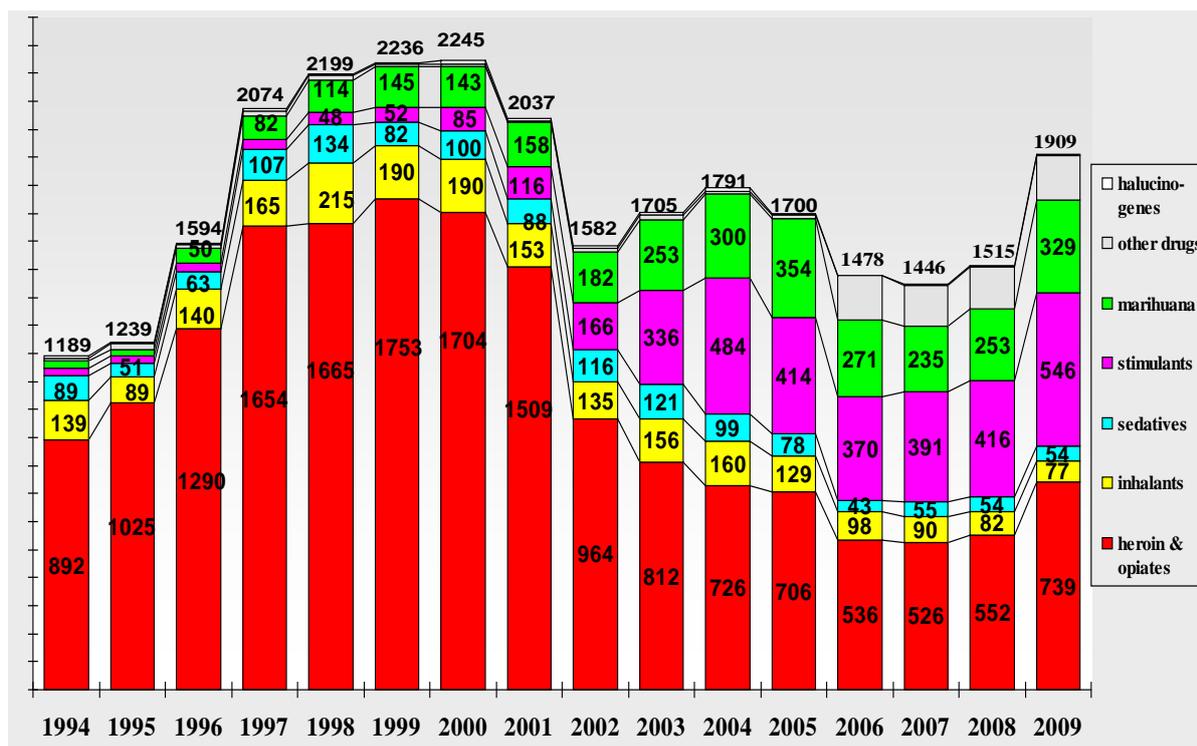
Introduction:

Slovakia became an independent state by January 1st 1993, when common state of Czechs and Slovaks ceased to exist. This event had happened three years after the fall of communist regimes in Eastern Europe. Czechoslovakia was divided by a parliamentary decision, through which the two parts of former state became two independent states: Czech and Slovak Republics. Slovakia has an area of 49 thousands km², overall number of inhabitants is 5.4 million.

Slovakia in similar way as neighbour post-communist countries in Europe did not have real drug problem before the year 1990, as the main problem consists in tobacco and before all alcohol. After few years the problem with illicit drugs, however, became fully developed and so called "heroin epidemics" had reached its peak around the year 2000 (see Picture 1). After this heroin prevalence, indicated mainly by number of treated heroin addicts, went down and other non-opiate drugs prevailed, resulting to the occurrence of new synthetic drugs.

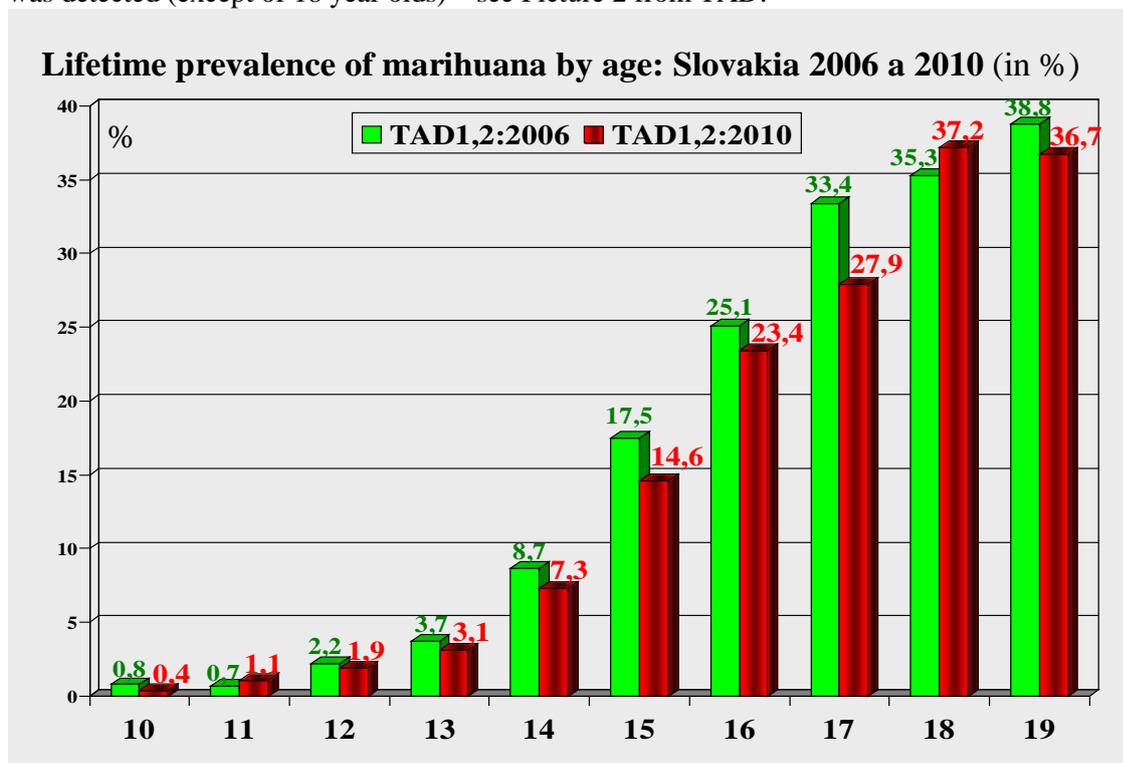
Main Text:

School surveys are part of key indicator of drug situation monitoring in population. In Slovakia the fifth wave of school surveys TAD (Tobacco–Alcohol–Drugs) took place in primary school pupils from 10-11 to 14-15 years (TAD1), in secondary school students from 15-16 to 19-20 years (TAD2), and in their teachers (TAD3). These nation-wide representative surveys were implemented in the spring of 2010 (Nociar 2010).



Picture 1. Numbers of treated drug addicts in health care facilities: Slovakia 1994-2009
Source: UVZ SR. Processing and graph: A. Nociar

One year later the fifth wave of the European school survey on alcohol and drugs (ESPAD) took place in secondary school students from 15-16 to 19-20 years. It should be said, that from years 1994-1995 until 2006-2007 TAD and ESPAD surveys demonstrated constant growth of illicit drugs use, before all marihuana. In the case of marihuana, however, in the years 2010-2011 clear decline was detected (except of 18 year olds) – see Picture 2 from TAD:



Picture 2. Results of the primary school pupils and the secondary school students by age:
Lifetime prevalence of marihuana.

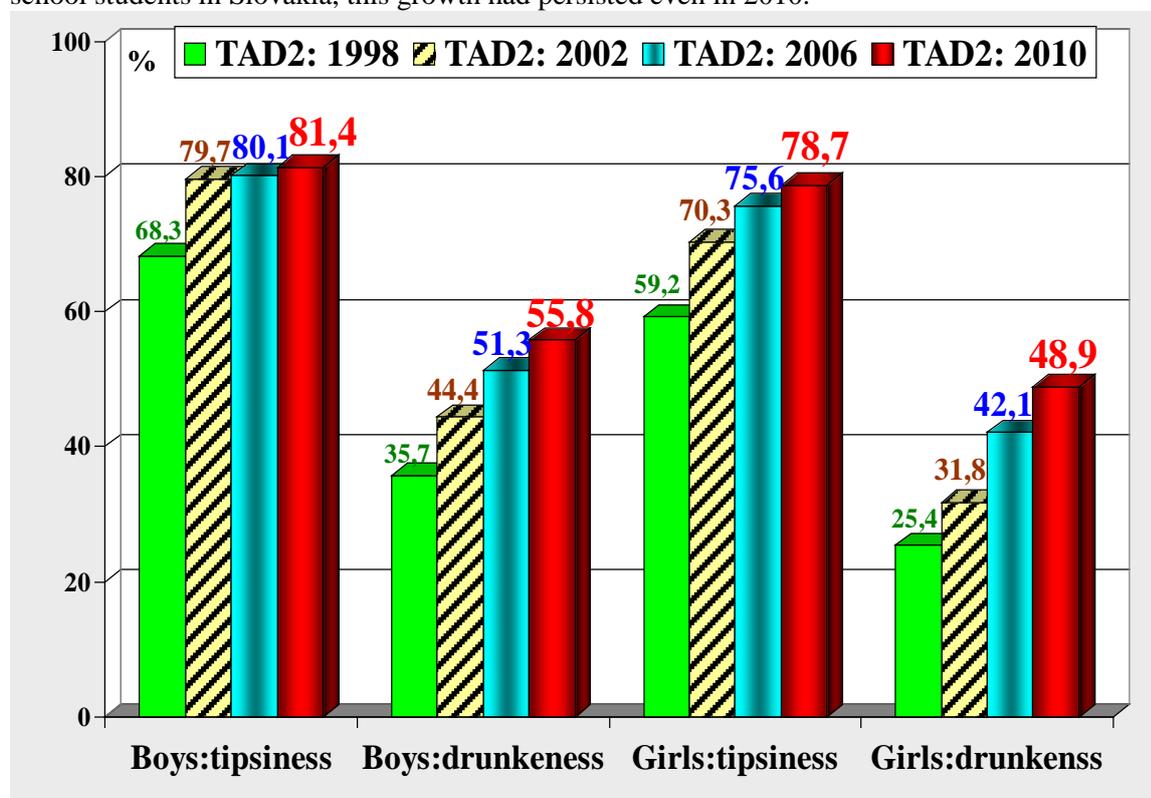
Source: UVZ SR. Processing and graph: A. Nociar

The same decline in TAD2 from 2010 had been detected practically in all remaining illicit drugs:

Table 1. Lifetime prevalence of illicit drugs – the TAD2 results from 2006 compared to the TAD2 results from 2010 (in %)

Lifetime prevalence of illicit drug	Boys		Girls		TOTAL	
	2006	2010	2006	2010	2006	2010
Marihuana	39,1	36,4	26,5	22,4	31,6	28,5
Ecstasy	6,8	5,6	5,5	3,5	6,0	4,4
Inhalants	6,0	5,4	4,4	4,1	5,1	4,7
Amphetamines	3,2	3,6	5,6	2,7	4,5	3,1
Heroin	1,0	0,6	0,6	0,4	0,8	0,5
Cocaine	1,9	2,0	1,6	1,0	1,7	1,4
LSD	3,9	2,8	1,7	1,2	2,6	1,9
Stimulants	3,8	4,7	3,5	3,0	3,7	3,7

It should be added, however, that in the case of licit drugs, i.e. tobacco and alcohol in secondary school students in Slovakia, this growth had persisted even in 2010:



Picture 3. Results of the secondary school students by gender: tipsiness and drunkenness 1 and more time through life.

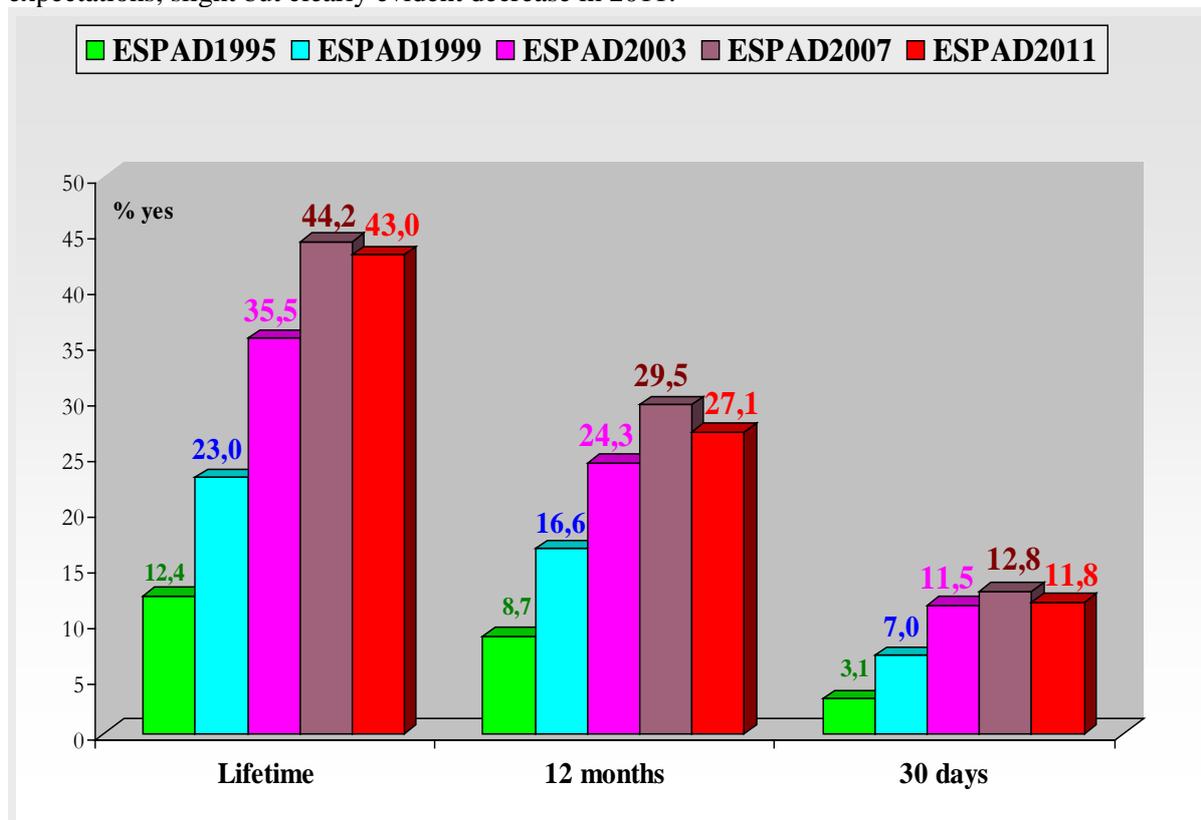
Source: UVZ SR. Processing and graph: A. Nociar

The ESPAD survey results from 2011 in majority of cases had confirmed a trend towards decrease of illicit drugs:

Table 2. Lifetime prevalence of illicit drugs – the ESPAD results from 2007 compared to the ESPAD results from 2011 (in %)

Lifetime prevalence of illicit drug	Boys		Girls		TOTAL	
	2007	2011	2007	2011	2007	2011
Marihuana	50,6	47,9	37,8	38,3	44,2	43,4
Ecstasy	8,6	7,2	6,5	5,1	7,6	6,2
Inhalants	10,6	10,0	8,6	8,3	9,6	9,2
Pervitin (meth)	8,1	7,4	6,4	6,2	7,3	6,8
Alcohol with pills for getting high	14,1	11,0	19,5	13,9	16,8	12,4
Heroin	1,0	2,4	0,6	1,4	1,7	2,0
Cocaine	2,6	4,0	2,5	2,8	2,6	3,4
LSD	6,6	7,1	5,0	5,1	5,8	6,2
Synthetic cannabinoids	-	5,4	-	2,5	-	4,1
Mephedrone	-	2,3	-	1,0	-	1,7

The ESPAD surveys from 1995 till 2007 have also demonstrated constant growth of illicit drugs use, before all in marihuana. But also in the case of marihuana there was, in contrary to our expectations, slight but clearly evident decrease in 2011:

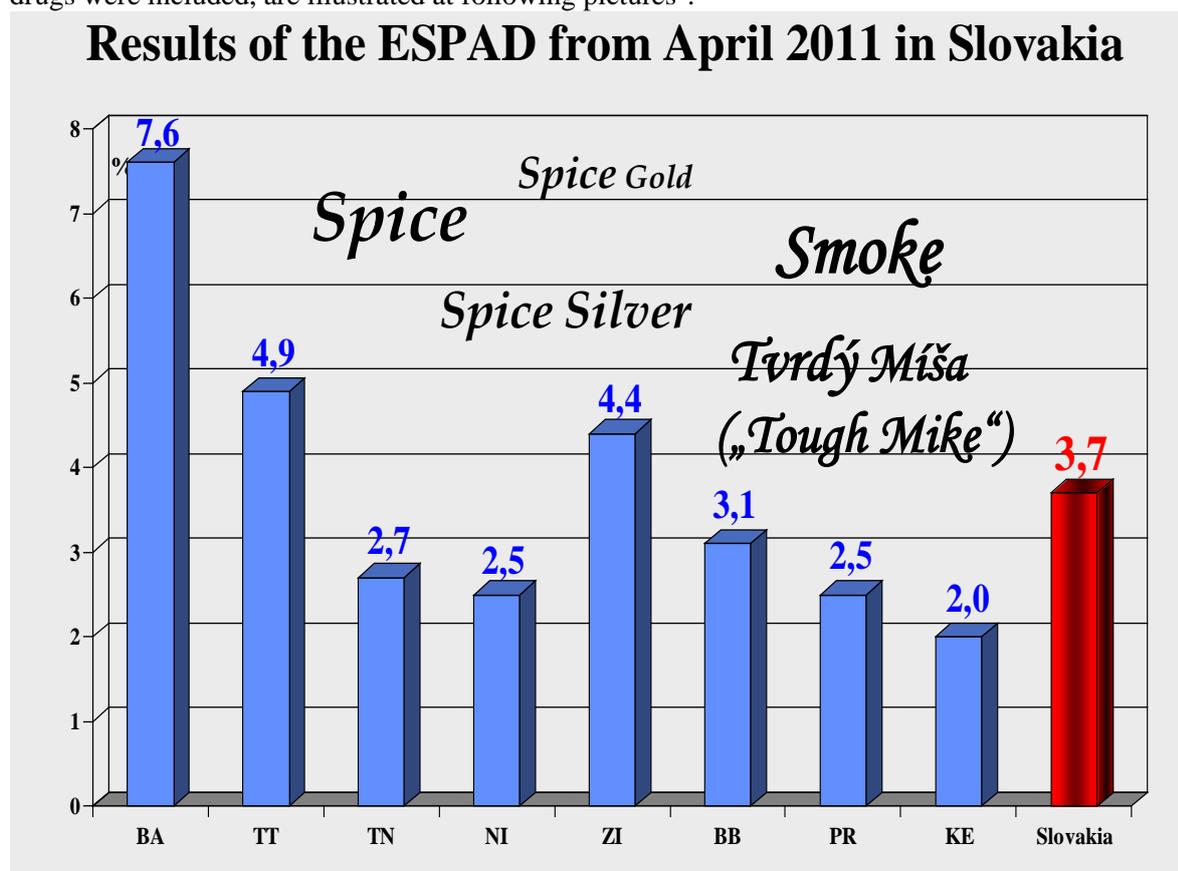


Picture 4. Results of the secondary school students in Slovakia: Marihuana prevalence
Source: UVZ SR. Processing and graph: A. Nociar

After the fifth wave of nation-wide surveys we can conclude, that drug situation is getting relatively stable, or in some instances we might see some improvement. Anyhow, before we begin to celebrate, we ought to be aware of the fact, that this decline is really small and the level of eventual stabilization is unacceptably high.

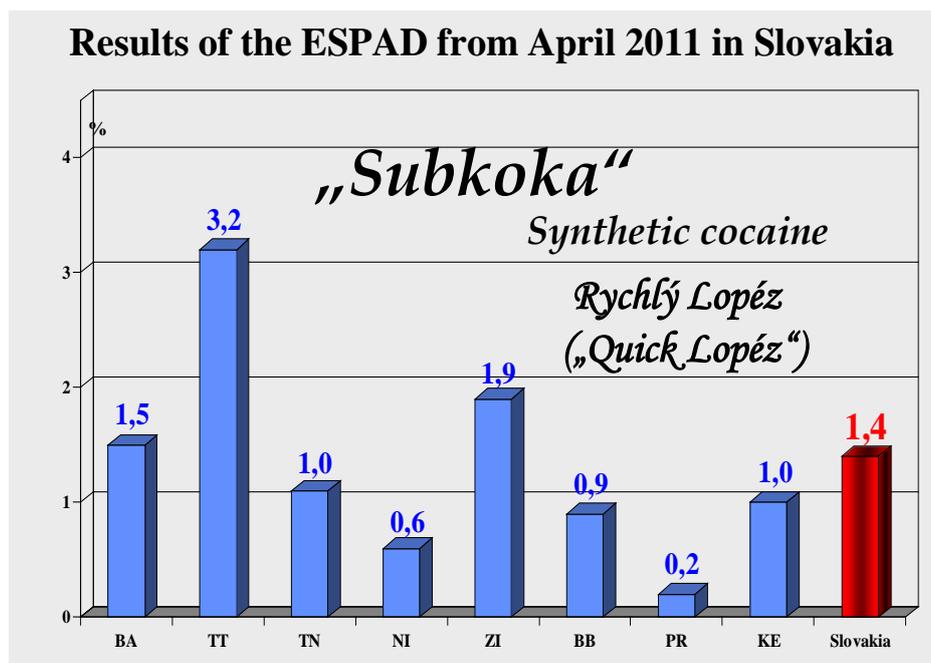
It is difficult to say anything conclusive towards causes of that decrease. What might be said is, that Slovakia have joined majority of EU member countries, where such a trend have started earlier. It is also possible, that we are again witnessing very quick changes of drug scene, rise of synthetic cannabinoids and cathinons, and dangerous mixtures of unknown composition, sold as „bath salts”, causing sometimes uncontrolled aggressive behaviour.

Therefore it was important to get at least preliminary information regarding new synthetic drugs and their possible presence in our potentially vulnerable population of young people. The results of ESPAD in the whole country and eight districts, where such a questions on new synthetic drugs were included, are illustrated at following pictures¹:



Picture 5. Results of the secondary school students in Slovakia: Lifetime prevalence of synthetic cannabinoids
Source: UVZ SR. Processing and graph: A. Nociar

¹ Abbreviations from BA to KE stand for 8 regions of Slovakia: Bratislava, Trnava, Trenčín, Nitra, Zilina, Banska Bystrica, Prešov, Kosice



Picture 6. Lifetime prevalence of mephedrone in secondary school students in Slovakia

Source: UVZ SR. Processing and graph: A. Nociar

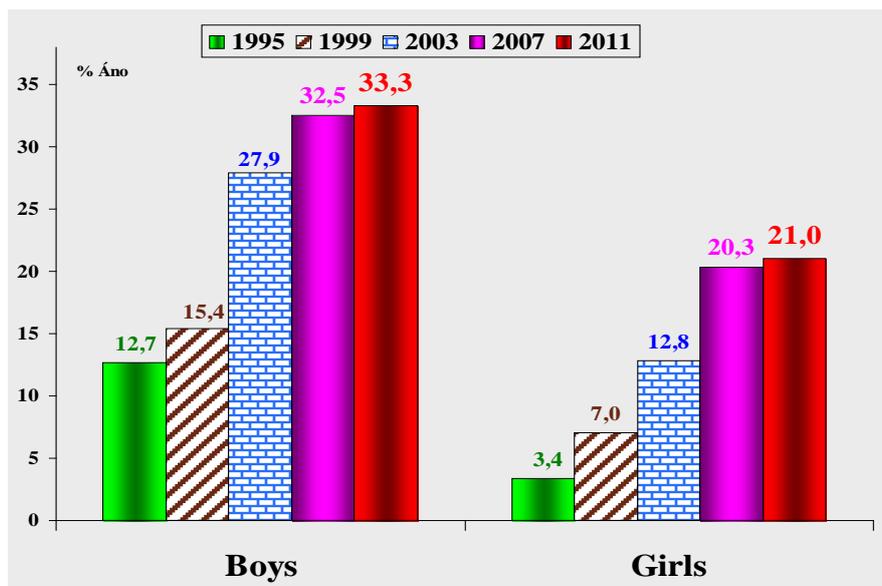
Affirmative answers to questions on new synthetic drugs are overlapped to some extent, but it might be said, that overall prevalence was around 5 %.

Reflecting further about changes on drug scene in Europe, it might be possible also, that in similar manner like after all-European “drinking binge” after the 2nd World war practically in all European countries (Smart 1989), where only after more then 20 years some decrease of per capita alcohol drinking occurred, perhaps something is happening also in the field of drugs. Following decline of illicit drug use among youth in the countries of Western Europe, such a decline, after some time shift started also in the Central and Eastern European post-communist countries, where illicit drug use also have reached its natural peak.

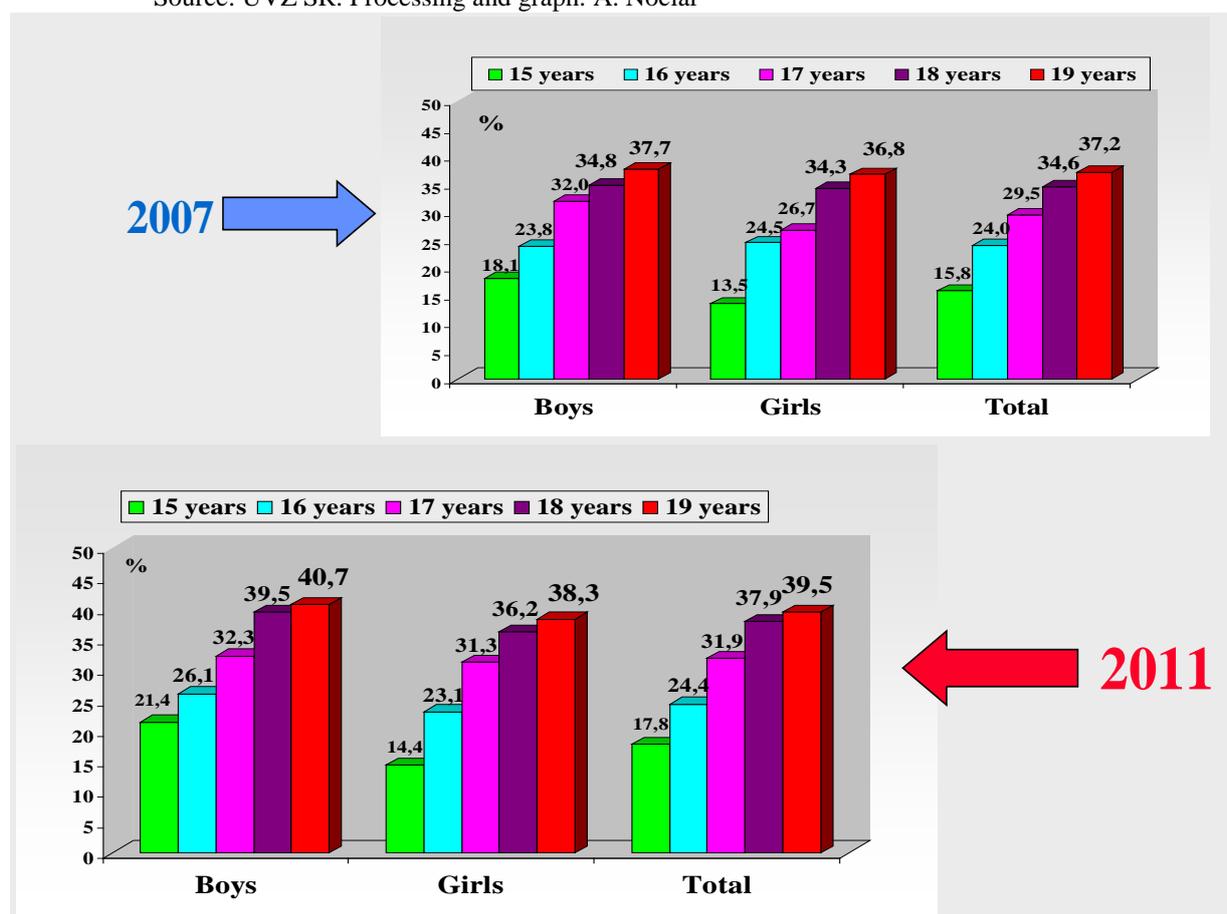
Sorry to say, but growth of abusive drinking of alcohol among youth was again confirmed, especially among girls and young women. This was true not only for hazardous and risk drinking, but also in binge drinking, i.e. drinking to get drunk as quickly as possible. Because our version of questionnaire was complemented by CAGE and ADS (Nociar 2011), some results on these topics are below (see Table 3, Picture 7). Moreover, this unfavourable trend was present also in the case of tobacco (see Picture 8):

Table 3. Results of TAD2 from 2006 compared to 2010; and the ESPAD results from 2007 compared to 2011 (in %)

T A D 2006 – 2010:	Boys		Girls		TOTAL	
	2006	2010	2006	2010	2006	2010
CAGE 3-4: problem drinking	6,8	10,2	5,1	8,8	5,7	9,4
ADS: heavy dependence signs	5,8	6,5	2,8	3,0	4,0	5,7
E S P A D 2007 – 2011:	Boys		Girls		TOTAL	
	2007	2011	2007	2011	2007	2011
CAGE 3-4: problem drinking	6,7	9,4	6,5	7,9	6,6	8,7
ADS: heavy dependence signs	5,3	6,6	3,9	4,6	4,4	5,6
Binge drinking – 5+drinks in a row: 3-times or more in 30 days)	32,5	33,3	20,3	21,0	26,3	27,6



Picture 7. Comparison of ESPAD surveys 1995-2011: binge drinking (3-5 drinks or more 3 or more times during 30 days) among secondary school students in Slovakia
Source: UVZ SR. Processing and graph: A. Nociar



Picture 8. Comparison of the ESPAD surveys 1995-2011: regular smoking (1-5 cigarettes or more per day) among secondary school students in Slovakia
Source: UVZ SR. Processing and graph: A. Nociar

Conclusion:

Except of certain decrease or stabilization of traditional drugs like marihuana, heroin, cocaine and amphetamine-like stimulants, there were detected in previous 4 – 5 years in Europe and also in

Slovakia, more and more frequent emergences on new synthetic drugs, known as „Spice“, „Legal highs“, offered via internet, or directly in so called „Smart shops“, or „Crazy shops“. EU member states as well as the EU as a whole have reacted by their own legislation changes, or common EU legislation, like in the case of mephedrone. Drug situation in Slovakia parallelizes overall situation in the EU, when after a period of traditional drug use there was swift change to new synthetic drugs, like metamphetamine and ecstasy, but more recently to synthetic cannabinoids, stimulants like mephedrone or other “Legal highs”, frequently of unclear composition.

After two decades of drug problem in Slovakia it's time to come to terms with reality – that young people today are living in the environment where a presence of drugs is no more exception. Drugs are offered via black market, or even more openly, but also in the atmosphere of more-less publicly declared opinions on illicit drugs “safety”. Moreover, we can hear that perhaps the time is coming to decriminalise, or even to legalise, at least some of them, as some voices try to declare more and more publicly.

Finally, realistically speaking, due to still high prevalence of marihuana among 15 to 19 years old, it would be reasonable to support the development of low threshold network of contact points with accessible short term counseling and treatment, targeted at so called soft drugs, including relatively new synthetic drugs and dance or disco drugs.

Anyhow, we should not forget old observation from clinical practice – that intuitive, but also empirically based conclusion: topics of so called licit, and so called illicit drugs belong to each other as a male-screw entirety, and to try to put them apart has no sense.

References:

Nociar, A.: Závěrečná správa z prieskumu TAD u žiakov ZŠ, študentov SŠ a ich učiteľov v roku 2010 [Final report from the TAD survey in primary school pupils, secondary school students and their teachers in 2010]. Bratislava, VÚDPaP [Research Institute of Child Psychology and Pathopsychology] 2010, 74 p.

Nociar, A.: Európsky školský prieskum o alkohole a iných drogách (ESPAD) v SR za rok 2011 [European school survey on alcohol and other drugs (ESPAD) in Slovakia. Final report on research task]. Závěrečná správa k výskumnej úlohe P - 117. Bratislava, VÚDPaP [Research Institute of Child Psychology and Pathopsychology] 2011, 83 p.

Smart, R. G.: Is the postwar binge ending? Crossnational trends in per capita alcohol consumption. *Brit. J. Addict.* 1989, 84, pp.743-748.

COMPARISON OF ELECTRICAL ACTIVITY OF PATELLA STABILIZER MUSCLES BETWEEN ATHLETES WITH AND WITHOUT PATELLOFEMORAL PAIN SYNDROME

Nahid Khoshraftar Yazdi

Majid Mehdikhani

Assistant Professors, Faculty of Physical education and sport science- Ferdowsi University of Mashhad, Iran

Stephan Starischka, Prof. Dr.

Institute of Sport, Dortmund University, Germany

Elke Zimmermann, Prof. Dr. med

Department of sports medicine, Bielefeld University, Germany

Abstract:

The aim of this study was to compare the electrical activity of the Tensor fascia lata (TFL) and Vastus lateralis (VL) muscles as lateral stabilizers and the Vastus medialis oblique (VMO) muscle as medial stabilizer of the patella in athletes with and without patellofemoral pain syndrome (PFPS). Eleven healthy athletes (4 women, 7 men) and 9 athletes with PFPS (3 women, 6 men) ranging in age from 20 to 30 years performed the submaximal isometric knee extension contraction initiated from 45° of flexion and an extended position to 30° and 15° angles. The activation amplitude of the TFL, VL and VMO muscles were quantified from the recorded electromyographic (EMG) signals. The results showed that Electrical activity of the VMO muscle at 30° and 45° angles for the patients was significantly greater than those of the control subjects. No significant differences were demonstrated in VMO: VL and VMO: TFL across all knee flexion angles between groups. However, the mean value of the ratio of VMO: (TFL + VL) at an angle of 30° for the subjects with PFPS was significantly greater than that of the control group. It may be assumed that the higher electrical activity of the VMO at 30° angle in PFPS patients is an effort to prevent more lateral patella tracking. Since the TFL and VL muscles produce lateral force on the patella, both of them should be studied for consideration of lateral stabilizers of the patella in the patients with PFPS.

Key Words: Patellofemoral pain syndrome, Vastus lateralis, Vastus medialis oblique, Tensor fascia lata, Patella, EMG

Introduction

Patellofemoral pain syndrome (PFPS) is the most common knee complaint found in adolescents and young adults. Researchers showed an incidence as high as one in four, and even higher, among athletes (Cowan, Bennell & Hodges, 2002; Witvrouw, Lysens, Bellemans, Cambier & Vanderstraeten, 2001). Despite this high incidence, the exact cause of these disorders remains enigmatic (Crossley, Bennell, Green, Cowan & McConnell, 2002) and PFPS remains one of the most vexatious clinical challenges in rehabilitation medicine (Wilk, Davise, Mangine & Malone, 1998).

The major complaint of patients with PFPS is retropatellar pain during activities such as running, squatting, going up and down stairs, prolonged sitting, cycling, and jumping (Witvrouw et al., 2001). A combination of factors, such as abnormal lower limb biomechanics and abnormal lateral tracking of the patella may result in increased cartilage and subchondral bone stress, subsequent PFPS and subtle patellar malalignment or more overt patellar maltracking (Fredericson & Yoon, 2006; Miller, Sedory & croce, 1997; Souza & Gross; 1991).

Several researchers showed that PFPS is associated with specific quadriceps muscle atrophy, especially in the vastus medialis oblique muscle (VMO) (Boucher, King, Lefebvre & Pepin, 1992; Lefebvre et al., 2006). Insall (1982) suggested that mechanism of abnormal lateral tracking of the patella is an imbalance in the activity of the VMO muscle relative to the vastus lateralis muscle (VL). Owings and Grabiner (2002) stated that the rationale for an important biomechanical role of the VMO

in patellar tracking is substantiated by its physiologic and biomechanical properties. They believed that due to smaller maximum contraction force and slower maximum contraction velocity of the VMO, it would be expected that VL dominates the movement of the patella unless the activation patterns of these muscles account for this difference (Owings & Grabiner, 2002).

It is assumed that the medial tracking role of the VMO counteracts the laterally directed force of the VL on the patella (Grabiner, Koh & Miller, 1991). Lateral tracking of the patella may be due to inadequate medial control from the VMO in persons with PFPS (Gilleard, McConnel & Parsons, 1998).

The other soft tissue which influences the patella is illiotibial band (ITB). The ITB, the tendon of Tensor fascia lata (TFL), is an important dynamic lateral stabilizer of the patella (Chimera, Swanik, Buz Swanik & Straub, 2004; Mosher & Jackson, 2008; Recondo et al., 2000). With regard to the role of the VL and TFL on the lateral side of the patellofemoral joint and the VMO on the medial side, it can be assumed that the TFL as well as the VL contributes in making lateral force on the patella. Therefore, in order to prevent patella tracking, the VMO should counteract this force which is produced by the TFL and VL. If the VMO is not as strong as the TFL and VL, it has to recruit more motor units to dominate the force of the TFL and VL.

Since the EMG intensity provides a reliable estimate of the volume of recruited muscle, but not necessarily of the developed force (Roberts & Gabaldon, 2008), we expect that the VMO muscle has more activity in patients than that of a control group.

Although the electromyographic (EMG) activity of the VMO and the VL was evaluated in patients with PFPS by several investigators (Boucher et al., 1992; Gilleard et al., 1998; Owings & Grabiner, 2002), none of them considered the activity of the TFL against the VMO.

The purpose of this study was to examine the electrical activity of lateral and medial stabilizer muscles of patella including the VMO, the VL and the TFL and to compare their activities in subjects with and without PFPS.

Methods

Subjects volunteered to participate in this study and were placed in an experimental group and a control group based on the presence of PFPS symptoms with no evidence of any other specific pathologic condition. All group members were athletes who were active in sports such as running, football, basketball, and handball, for more than 10 years. The control group was composed of 4 women and 7 men with a mean age of 25.1 ± 3.2 years. They were healthy and athletic, and reported no history of knee injury.

The experimental group consisted of 3 women and 6 men with a mean age of 26.3 ± 2.6 years, who had history of PFPS with duration of symptoms was more than 6 months and intensity sufficient to limit function or cause the individual to seek intervention. These symptoms consisted of retropatellar pain during physical activities such as jumping, running, squatting, and going up or down stairs. Clinical criteria include pain on direct compression of the patella against the femoral condyles with the knee in full extension, tenderness of the posterior surface of the patella on palpation, pain on resisted knee extension and pain with isometric quadriceps muscle contraction against suprapatellar resistance with the knee in 15° of flexion. These clinical signs are summarized in Table 1. Participants were excluded if they had signs or symptoms of meniscal, bursa, ligament laxity or tenderness, tenderness over the patellar tendon, illiotibial band, or pes anserinus tendons, patellar apprehension sign, patellar dislocation and previous knee surgery. The subjects did not have pain at rest, and did not have pain during a submaximal isometric contraction of knee flexion. The study received ethics committee approval from the Ruhr-Universität Bochum Ethics commission of Germany.

Table 1. Clinical signs of the patients with PFPS

Pain on direct compression of the patella
Tenderness of the posterior surface of the patella on palpation
Pain on resisted knee extension
Pain with isometric quadriceps muscle contraction

Before beginning the study, every subject signed an informed consent document.

The motor point of muscles was estimated by the following method (Rainoldi, Melchiorri & Caruso, 2004; Hermans HJ, Freiks B, 1997), (Figure 1):

Vastus Lateralis: Two anatomical landmarks (the anterior superior iliac spine (ASIS) and the superior lateral side of the patella (SLSP)) were determined. Electrodes were placed at 2/3 on the line from the ASIS to the SLSP in the direction of the muscle fibers.

Vastus medialis obliquus: Two anatomical landmarks (ASIS and the superior medial side of the patella (SMSP)) were determined. A quadriceps line was drawn from the ASIS to the SMSP. Electrodes were placed at 80% of the quadriceps line (starting from ASIS) with a medial inclination of 50°.

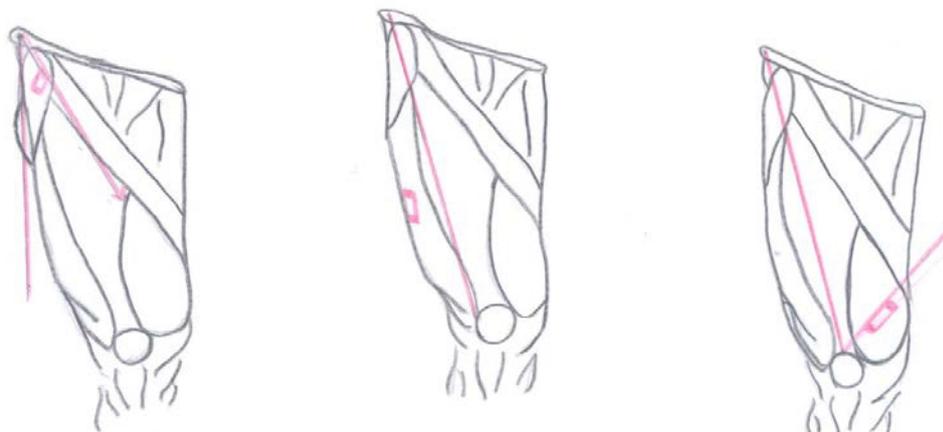


Figure 1. Schematic drawing of proper electrode positioning over the Tensor fascia lata (A), Vastus lateralis (B), and Vastus medialis oblique(C) muscles

Tensor fascia lata: Two anatomical landmarks (ASIS and greater trochanter (GT)) were determined. A line was drawn from the ASIS to the GT. Electrodes were placed at 50% of this line with an inclination of 30°.

Active bipolar Ag/AgCl surface electrodes (pre-amplification gain= 10, recording diameter, 1 mm; center-to-center distance, 20 mm) were placed on the motor point of the VL, the VMO and the TFL muscles of the tested leg. The subject's skin was prepared by shaving, then used abrasive paste for gentle local abrasion and finally cleaned with isopropyl alcohol. The resistance between the electrode pair was also measured by an EMG electrode impedance tester. The long axis of the electrodes was positioned over each muscle in the assumed direction of the underlying muscle fibers.

The protocol prescribed submaximal isometric contractions with 60% of maximal voluntary contraction (MVC) for the knee extension. A device was designed for displaying the exact degree of knee extension. We named it angle meter. Each subject was seated on a chair. The hip-trunk angle was approximately 100°. The angle meter was positioned so that knee flexion angle was 90° and the estimated center of knee joint rotation was on a level with the angle meter's axis of rotation. The subjects put their foot on the pedal of the angle meter. By moving the pedal upwards, the knee was extended and when the angle meter displayed 45°, the subjects had to maintain the position for 6 seconds (Figure 2). Afterward, they contracted isometrically again at 30° and 15° maintaining the contraction for 6 seconds at each stage. The test was repeated 3 times. A rest of 1 min was given between tests. The mean of the three measurements was used for analysis.



Figure 2. Measuring electrical activity of the muscles in different angles of knee extension

EMG raw signals were recorded by using a Noraxon Inc., Scottsdale, AZ, USA EMG system. Signals were amplified differentially (total gain= 1000; CMRR> 130db). A band-pass Butterworth filter with cut-off frequencies ranging from 15 to 500 HZ was applied. The transmitted signals were sampled at 1 kHz, input to an analog-to digital circuit (Data translation Inc., Marlboro, MA.USA, 16-bit resolution) and were stored. All signals processing was supported by the Noraxon MyoResearch XP software.

The activation signals of the VMO, VL and TFL were full-wave rectified through root mean square (RMS). Amplitude was analyzed by calculating IEMG in a window of 6 seconds and normalized by the mean value calculated within this time-window. Signals were also time normalized from 0-100% over these six seconds.

The t-test for independent variables was used to compare the activation amplitude of the VL, the VMO, the TFL muscles and their ratios between patients and control group. Statistical analysis was performed with SPSS Version 17.0, and a value of 0.05 was accepted as reflecting significance.

Results

Significant differences were identified in the activation amplitude of the VMO at 30° and 45° angles between the subjects with and without PFPS ($P= 0.05$, $P= 0.03$, respectively), (Table 2). According to the t-test analysis, VMO electrical activity in the subjects with PFPS was significantly greater than that of the subjects without PFPS. However, no significant differences were demonstrated in the activation amplitude of the VMO at an angle of 15° between the two groups ($P= 0.76$). In addition, there were no significant differences in the activation amplitude of the VL at 15° ($P= 0.61$), 30° ($P= 0.55$) and 45° ($P= 0.66$) angles between the subjects with and without PFPS. Furthermore, no significant differences were found in the activation amplitude of the TFL at 15° ($P= 0.25$), 30° ($P= 0.67$) and 45° ($P= 0.89$) angles between the two groups.

Table 2. The means and standard deviations of normalized VMO, VL and TFL IEMG value (%) for knee flexion angles in the subjects with and without PFPS

Group	Muscle	Knee Flexion Angle		
		45°	30°	15°
Control	VMO	22.3 ± 11.3	26.4 ± 11.5	72.3 ± 7.3
	VL	37.3 ± 18.1	40.7 ± 17.1	80.4 ± 8.9
	TFL	32.5 ± 16.2	45.2 ± 19.8	66.1 ± 11.3
Patellofemoral pain	VMO	34.6 ± 12.6	38.9 ± 15.3	73.6 ± 10.5
	VL	41.3 ± 21.4	46.2 ± 23.2	78.0 ± 11.0
	TFL	33.5 ± 16.2	41.7 ± 15.4	71.5 ± 8.0

There were no significant differences in the VMO: VL ratio across all knee flexion angles between the subjects with PFPS and the subjects without PFPS ($P= 0.35$, $P= 0.09$, $P= 0.07$, respectively from an angle of 15° to 45°). Similarly, no significant differences were identified in the VMO: TFL ratio at all knee flexion angles between the patients and control group ($P= 0.39$, $P= 0.06$, $P= 0.12$, respectively from an angle of 15° to 45°).

We found a significant difference in the VMO: (TFL+ VL) ratio at 30° of knee flexion between the two groups ($P= 0.02$). This ratio at 30° of knee flexion for subjects with PFPS (0.46) was significantly greater than for the subjects without PFPS (0.31), (Table 3). There was no difference in this ratio at 15° ($P= 0.90$) and 45° ($P= 0.06$) of knee flexion between the two groups.

Table 3. The VMO: VL, VMO: TFL and VMO: (TFL+VL) ratios for knee flexion angles in the subjects with and without PFPS

Group	Ratios	knee flexion angle		
		45°	30°	15°
Control	VMO: VL	0.67 ± 0.29	0.68 ± 0.22	0.90 ± 0.08
	VMO: TFL	0.79 ± 0.41	0.66 ± 0.33	1.14 ± 0.31
	VMO: (TFL+VL)	0.34 ± 0.13	0.31 ± 0.11	0.50 ± 0.07
Patellofemoral pain	VMO: VL	1.01 ± 0.51	0.98 ± 0.46	0.95 ± 0.12
	VMO: TFL	1.48 ± 1.31	0.99 ± 0.43	1.04 ± 0.17
	VMO: (TFL+VL)	0.48 ± 0.19	0.46 ± 0.14	0.49 ± 0.07

Discussion

In the current study, the electrical activity of muscles (VMO, VL and TFL) was investigated at 3 different angles of knee flexion (45°, 30°, 15°) during submaximal isometric contraction in the subjects with and without PFPS. The normalized EMG data of VMO and VL indicated that in subjects with PFPS, the activation amplitude of the VMO muscle during the contraction at 30° of knee flexion was significantly greater compared with that of the subjects without PFPS ($P= 0.05$). There was no significant difference in the activation amplitude of the VL muscle between the two groups at the same angle. However, Owings & Grabiner (2002) reported the activation amplitude of both muscles (VMO and VL) of the subjects with PFPS were significantly higher than those of the control subjects during eccentric contraction.

Following the data analysis, we found no significant differences in VMO: VL ratio across all knee flexion angles between the two groups. This finding has also been reported by previous studies (Boucher et al., 1992; Souza & Gross, 1991). Souza and Gross suggested that patients with PFPS may not differ from healthy individuals with regard to VMO: VL activation patterns. Since the TFL as well as VL, is a lateral stabilizer of the patella (Hughston, Andrews, cross & Moschi, 1976), it may be suggested to consider the TFL muscle as well as VL, when studying the ratio of lateral and medial dynamic stabilizer of the patella.

With regard to the role of the TFL/ ITB complex, which is an important lateral dynamic stabilizer of the patellofemoral joint, in particular from 0° to 30° of knee flexion (Hughston et al., 1976; Kanamiya et al., 2002), several studies suggested that TFL/ITB complex tightness may contribute to the development of PFPS (Puniello, 1993; Reider, Marshall & Warren, 1993).

On the other hand, some researchers indicated there is a correlation between hip internal rotation and PFPS (Berger, Crossett, Jacobs & Rubash, 1998; Ireland, Davis, Ballantyne & Willson, 2003). Berger et al (1998) stated that the direct correlation of combined (femoral and tibial) internal component rotation to the severity of the patellofemoral complication suggests that internal component rotation may be the predominant cause of patellofemoral complications in patients with normal axial alignment. Ireland et al. (2003) indicated that female runners, who have demonstrated significant knee valgus and hip internal rotation movements during running, are especially prone to PFPS. According to Press and Young (1998), internal rotation may be caused by a tight TFL and a weak gluteus medius. McConnel (2002) showed that decreased flexibility of TFL muscle is a contributing factor in the etiology of PFPS. Since TFL is an internal rotator of the hip and a lateral stabilizer of patella (Fredericson et al., 2000), we considered the electrical activity of this muscle. To the authors' knowledge there is no study that considers the electrical activity of the TFL muscle in the patients with PFPS.

As the TFL and the VL play the role of the dynamic lateral stabilizers of the patella, the VMO represent a dynamic medial stabilizer of the patella. This function of VMO counteracts the force of lateral stabilizers on the patella. Due to balances in lateral and medial stabilizing forces, patellar tracking is prevented. Therefore, an insufficiency of these muscles can have an effect on the patella and develop patellar malalignment and consequently PFPS.

In the present study, we found no significant differences in VMO: TFL ratio between two groups across all knee flexion angles, however, significant differences were demonstrated in VMO: (TFL+VL) ratio at 30° of knee flexion for the two groups. This ratio was greater for the subjects with PFPS (0.46) than for the subjects without PFPS (0.31). This result showed that there was no significant differences in (TFL+ VL) at an angle of 30° of knee flexion between two groups, but the electrical activity of the VMO was significantly higher in patients than that of the healthy subjects. Powers (2000) believed that increase motor unit activity of the vastus medialis appeared to be in response to meeting the increased demand of providing patellar stability. With regard to the significant differences in VMO: (TFL+ VL) ratio between patients and control group, it may be assumed that VMO as a medial stabilizer of the patella recruits more volume of active motor units thus preventing more patella tracking. Roberts and Gabaldon (2008) stated that it is generally assumed the EMG intensity provides a reliable estimate of the volume of recruited muscle, but not necessarily of the developed force. According to Roberts and Gabaldon (2008), high activation level of the VMO of the subjects with PFPS is not necessarily associated with developed force in this muscle. Powers (2000) suggested that increased motor unity activity of the Vastus medialis muscle appears to be associated with abnormal patellar kinematics in women, but it is not necessarily a cause of abnormal

patellar kinematics. It may be presumed that VMO with more activation and more recruited motor unit tries to counteract the lateral forces of VL and TFL on the patella and prevent more patella tracking.

This finding may provide evidence that PFPS is associated with a disruption in the control of the VMO during the isometric contraction at 30° of knee flexion. Furthermore, according to the results of the present study, it may be suggested to study the lateral stabilizers of the patella in the patients with PFPS, TFL as well as VL to be considered.

In the light of the results at 30° of knee flexion in the current study and another study (Orchard, Fricker, Abud & Mason, 1996) about the Iliotibial friction syndrome in runners, which shows that friction occurs at 30° of knee flexion, it may be assumed that this angle plays an important role in injuries occurring in runners.

We recognize that two limitations were present in our experimental design. The subjects' symptoms of patellofemoral pain varied in duration of pain. Nevertheless, as a group, these subjects were associated with a pattern of activation amplitude that was different from that of control subjects at angles of 30° and 45°. It may be assumed that PFPS, regardless of pain duration, effects on activation amplitude of VMO at above mentioned angles.

The other limitation that we found in our study was daily activities of the subjects which it was not possible to control them. Although all subjects were active in sports such as running, football, basketball and handball, their daily activities may be different. However, this limitation was similar for both groups.

Conclusion:

In the present study, the electrical activity of the VMO at 30° angle of knee flexion was significantly higher in the patients than that of the control group. It may be assumed that high electrical activity of the VMO in the patients is an effort for counteracting lateral force, which is produced by lateral stabilizers, and subsequently preventing patella tracking. With regard to no significant differences in VMO: VL and VMO: TFL ratios across all knee flexion angles between the two groups, and significant differences in VMO: (TFL+ VL) ratio at 30° angle of knee flexion, it is suggested that in consideration of lateral stabilizers of the patella, the TFL and VL should be assessed together.

References:

- Ahmad CS, Shubin Stein BE, Matuz D, Henry JH. (2000). Immediate Surgical Repair of the Medial Patellar Stabilizers for Acute Patellar Dislocation, A Review of Eight Cases. *Am J Sports Med*, 28(6), 804-810.
- Berger RA, Crossett LS, Jacobs JJ, Rubash HE. (1998). Malrotation causing patellofemoral complications after total knee arthroplasty. *Clinical orthopaedics and related research*, 356, 144-153.
- Boucher JP, King MA, Lefebvre R, Pepin A. (1992). Quadriceps femoris muscle activity in patellofemoral pain syndrome. *Am J Sports Med*, 20, 527-532.
- Chimera NJ, Swanik KA., Buz Swanik C, Straub SJ. (2004). Effects of plyometric training on muscle-activation strategies and performance in female athletes. *J Athl Train*, 39(1), 24-31.
- Cowan S, Bennell K, Hodges W. (2002). Therapeutic Patellar Taping Changes the Timing of Vastus Muscle Activation in People With Patellofemoral Pain Syndrome. *Clin J Sport Med*, 12(6), 339-347.
- Crossley K, Bennell K, Green S, Cowan S, McConnell J. (2002). Physical Therapy for Patellofemoral pain: a randomized, double-blind, placebo-controlled trial. *Am J Sports Med*, 30(6), 857-865.
- Fredericson M, Cookinham CL, Chaudhari AM, Dowdell BC, Oestreicher N, Sahrman SA. (2000). Hip abductor weakness in distance runners with iliotibial band syndrome. *Clinical journal of sport medicine*, 10(3), 169-175.
- Fredericson M, Yoon K. (2006). Physical examination and patellofemoral pain syndrome. *Am J Phys Med Rehabil*, 85, 234-243.
- Gilleard W, McConnel J, Parsons D. (1998). The effect of patellar taping on the onset of vastus medialis obliquus and vastus lateralis muscle activity in persons with patellofemoral pain. *Phys Ther*, 78(1), 25-32.
- Gabiner MD, Koh TJ, Miller G. (1991). Fatigue rates of vastus medialis oblique and vastus lateralis during static knee extension. *J Orthop Res*, 9(2), 391-397.

- Hermans HJ, Freiks B. SENIAM; The state of the art on sensors and sensor placement procedures for surface ElectroMyoGraphy: A proposal for sensor placement procedures. Available at: <http://www.seniam.org>. 1997
- Hughston JC, Andrews JR, cross MJ, Moschi. (1976). Classification of knee ligament instabilities, Part II, The lateral compartment. *J Bone Joint Surg (Am)* 58(2), 173-179.
- Insall J. (1982). Current concepts review: patellar pain. *J Bone Joint Surg Am*, 64, 147-152.
- Ireland ML, Davis IS, Ballantyne BT Willson JD (2003). Hip strength in females with and without patellofemoral pain. *J Orthop Sports Phys Ther*, 33(11), 671-676.
- Kanamiya T, Whiteside L, Nakamura T, Mihalko W, Steiger J, Naito M. (2002). Effect of selective lateral ligament release on stability in knee Arthroplasty. *Clin Orthop Relat Res*, 404, 24-31.
- Lefebvre R, leroux A, Poumarat G, Galtier B, Guillot M, Vanneuville G, Boucher JP. (2006). vastus medialis : anatomical and functional considerations and implications based upon human and cadaveric studies. *J Manipulative Physiol Ther*, 29(2), 139-144.
- McConnel J. (2002). The physical therapist's approach to patellofemoral disorders. *Clin Sports Med*, 21(3), 363-387.
- Miller J, Sedory D, croce R. (1997). Leg rotation and vastus medialis oblique/vastus lateralis electromyography activity ratio during closed chain kinetic exercises prescribed for patellofemoral pain. *J Athl Train*, 32, 216-220.
- Mosher TJ, Jackson J. (2008). Knee, extensor mechanism injuries: Overview of the knee extensor mechanism. Available online at: <http://emedicine.medscape.com/article/401001-overview>. (Accessed 1 May 2009).
- Orchard J, Fricker P, Abud A, Mason B. (1996). Biomechanics of iliotibial band friction syndrome in runners. *Am J Sports Med*, 24(3), 375-379.
- Owings TM, Grabiner MD. (2002). Motor control of the vastus medialis oblique and vastus muscles is disrupted during eccentric contractions in subjects with patellofemoral pain. *Am J Sports Med*, 30, 483-487.
- Powers CM. (2000). Patellar kinematics, part I: The influence of vastus muscle activity in subjects with and without patellofemoral pain. *Phys Ther*, 80(10), 956-964.
- Press JM, Young JL. (1998). Rehabilitation of patellofemoral pain syndrome. In W.B. Kibler, S.L. Herring, J.M. Press (Eds.), *Functional rehabilitation of sports and musculoskeletal injuries* (pp. 254-264). Maryland, Gaithersburg: Aspen publications.
- Puniello MS. (1993). Iliotibial band tightness and medial patellar glide in patients with patellofemoral dysfunction. *J Orthop Sports Phy ther*, 17, 144-148.
- Rainoldi A, Melchiorri G, Caruso I. (2004). A method for positioning electrodes during surface EMG recordings in lower limb muscles. *J Neurosci Methods*, 134(1), 37-43.
- Recondo J, Salvador E, Villanua J, Barrera M, Gervas C, Alustiza J. (2000). Lateral stabilizing structures of the knee: Functional anatomy and injuries assessed with MR imaging. *Radiographics*, 20, S91-S102.
- Reider B, Marshall JL, Warren RF. (1993). Clinical characteristics of patellar disorder in young athletes. *Am J Sports Med*, 9, 270-274.
- Roberts TJ, Gabaldon AM. (2008). Interpreting muscle function from EMG: lessons learned from direct measurements of muscle force. *Integrative and comparative biology*, 48(2), 312-320.
- Souza D, Gross M. (1991). Comparison of vastus medialis obliques, vastus lateralis muscle integrated electromyographic ratios between healthy subjects and patients with patellofemoral pain. *Phys Ther*, 71, 310-320.
- Wilk KE, Davise GJ, Mangine RE, Malone TR. (1998). Patellofemoral disorders: a classification system and clinical guidelines for nonoperative rehabilitation. *J Orthop sports Phys Ther*, 28, 307-322.
- Witvrouw E, Lysens R, Bellemans J, Cambier D, Vanderstraeten G. (2001). Intrinsic risk factors for the development of anterior knee pain in an athletic population. *Am J Sports Med*, 28, 480-489.

EFFECT OF LUMBAR STABILIZATION EXERCISES VERSUS PRESSURE FEEDBACK TRAINING IN LOW BACK ACHE PATIENTS

Shadab Uddin

Department of Physical therapy, Faculty of Applied Medical Sciences, Jazan University, Jazan, Kingdom of Saudi Arabia (KSA)

Fuzail Ahmed

Department of Physical therapy, Faculty of Applied Medical Sciences, majmaah University, al majmaah, Kingdom of Saudi Arabia (KSA)

Abstract:

Chronic low back pain is the most common complaints in the urban society causing to absent from the work and activity limitation. Its health, social and economic burden is hefty. Despite developments in modern medicine in general and growing knowledge of spinal diseases, problem of nonspecific low back pain remains unsolved. Although the nonspecific types of back pain affects approximately 85% of patients while 40% of low back pain patients worry that pain affects their work ability and will make them cripple, or that it underlies some serious disease (Waddell 1998). Even though there is ample evidence stating the efficacy of core stability training but not able to find any study which has compared Pressure Biofeedback training with Core Stability Training in the treatment of chronic low back pain. The aim of the present study was to compare the effectiveness of stabilizer pressure biofeedback training and core stability exercises on pain perception as measure by visual analogue scale in chronic low back pain patients.

Key Words: Back pain, core stability training, stabilizer pressure biofeedback

Introduction

Chronic low back pain is the most common complaints in the urban society causing to absent from the work and activity limitation. Its health, social and economic burden is hefty. If we look to western societies, it has one of the most human suffering, disabling and enormous economic consequences with the frequent use of medical consultation and visit to rehabilitation unit. It is estimated that 70%-80% of population of United States of America is bear on back pain at one point in their life time.¹ It is 2nd most frequent reason for visit to the physician and 5th ranking cause of hospital admission.²⁻³ In the United Kingdom back pain is the largest single reason of absence from work in 1988-1989 and is responsible for approximately 12.5% of all sick days.⁴

Despite developments in modern medicine in general and growing knowledge of spinal diseases, the problem of nonspecific low back pain remains unsolved, although the nonspecific type affects approximately 85% of patients reporting any back pain. Approximately 40% of low back pain patients worry that pain affects their work ability, that it will cripple them, or that it underlies some serious disease. Better understanding of multidimensional aspects has widened our concept of low back pain.

There are several causes of low back pain. In a mechanical model, research has implicated pain sensitive vertebral structure such as the intervertebral disc and the zygoapophyseal joints as potential sources of low back pain. Irrespective of the actual source of symptoms, it has been shown that muscles are adversely affected secondary to low back pain. This phenomenon is accepted at peripheral joints such as the knee. For example, irrespective of the knee structure which is injured, e.g. meniscus or ligament, it is commonly accepted that quadriceps function will be adversely affected. Possible neurophysiological mechanisms include pain and reflex inhibition. Muscle re-education is therefore a commonly used intervention in the rehabilitation process, and for many years exercise has been advocated in the treatment of low back pain. Many researchers have focussed on issues of muscle strength. The basis for this focus is the premise that strong abdominal and back muscles are able to provide support for the lumbar spine.⁵

Despite the common acceptance of this principle, systematic reviews have not on the whole supported general trunk muscle strengthening programs.⁶ This has led to the development of specific exercise programs designed to protect and support the damaged joints and allow healing of the injured tissues.^{7,8} Muscles can be broadly divided into two categories, local and global muscles.⁹ The local muscle system includes deep muscles that are attached to the lumbar vertebrae and are capable of directly controlling the stiffness of the lumbar segments.¹⁰ In contrast, the global muscle system encompasses larger and more superficial muscles of the trunk. Their role is to move the spine and to control larger external loads, which occur with normal daily function. Biomechanical research has demonstrated that deep, local muscles are important for controlling, protecting and supporting the joints. The muscles of the local synergy, which are important for the lumbo-pelvic region include the segmental lumbar multifidus, the transverses abdominis, the pelvic floor and the diaphragm. There is evidence that low back pain results in an alteration in function of the local muscles, which lose their protective role.^{8,10,11}

A common clinical finding in low back pain patients is decreased range of motion of the spine with increased paraspinal activity. Disturbances in neuromuscular control have also been frequently connected with chronic low back pain and considered a possible linkage between pain and disability¹². These impaired functions recover with treatment or active rehabilitation. Spinal manipulative therapy is commonly recommended for low back pain, although previous systematic reviews and practice guidelines have produced discordant findings as to the effectiveness of this therapy. A recent focus in the management of chronic low back pain patients has been the specific training of the deep abdominal (internal oblique and transversus abdominis) and lumbar multifidus muscles. The primary role of these muscles is considered to be the provision of dynamic stability and segmental control of the spine. For the treatment of chronic low back pain with radiological diagnosis of spondylolysis or spondylolisthesis, one randomized controlled trial of specific stabilizing exercises showed significant and longer-lasting reduction in pain intensity and functional disability levels than did other commonly prescribed conservative treatment programs.

Investigators have reported that individuals with back pain may have motor control deficits or errors that affect their ability to engage the muscles that stabilize the spine. Maintaining segmental control within the trunk contributes to spinal stability and reduces unnecessary movement intersegmentally. This can serve to decrease the risk of back pain by causing a reduction in tissue strain, deformation, compression, and overstretching. Clinicians maintain that the mechanisms involved in spinal stability can be linked to differences in the function of the trunk muscles¹³. The trunk muscles have been classified into 2 categories, and each one performs distinct functions. The first category is the global muscle system. The muscles in this category are larger and more superficial in comparison to other muscles. The global muscles act to transfer loads and move the spinal column as a whole. Examples include the rectus abdominis and the external oblique¹⁴.

The second category is the local muscle system. The local muscles are smaller and deeper than the global muscles. They are intersegmental and produce only small amounts of force. Local muscles aid in proprioception and postural control, which can decrease the risk of injury. The local muscles also contribute to maintaining stiffness along the spinal column. The multifidus and the transverse abdominis are two examples¹⁵.

A number of investigators have cited evidence that supports the use of stabilization exercises for enhancing spinal stability²⁰. The local muscles are said to be crucial in this mechanism. This may be because of their contribution to maintaining the position of the spine and their ability to improve trunk endurance. Core stability training is frequently used to improve spinal stability. It has been used for many years in physical therapy and has become popular in fitness settings¹⁷. It has been speculated that this method of training improves spinal stability and may assist in decreasing the risk of back pain.

Studies that have been done on core stability training demonstrate promise for its effects on the musculature of the trunk¹⁸. However, previous investigations have not been designed to explore the involvement of the local muscles, which act to stabilize the spine. In addition, the methods of analysis have typically stressed the global muscles through assessments for strength or surface EMG recordings. These measures may not adequately identify improvements in spinal stability brought on by the local muscles. Core stability training that focuses on exercises with a neutral spine may be

appropriate for targeting the specific function of the local muscles during the early phases of programming for improving spinal stability¹⁹⁻²⁰.

Even though there is ample evidence stating the efficacy of core stability training I was unable find any study which has compared Pressure Biofeedback training with Core Stability Training in the treatment of chronic low back pain. The aim of the present study was to compare the effectiveness of Stabilizer Pressure Biofeedback Training and Core Stability Exercises on pain perception as measure by visual analogue scale in chronic low back pain patients.

Methodology

A total of 30 subjects were selected for the study on the basis of inclusion and exclusion criteria. The subjects were randomly assigned to Group 1 and Group 2 each comprising 15 subjects. All the patients were recruited from the inpatient outpatient department Government Civil Hospital, Ahmadabad. Only those patients were included who were between 40-60 years of age, presented with chronic nonspecific recurrent low back pain. The duration of the symptoms was diagnosed as sub-acute and chronic according to the IASP classification of pain. Those patients who were having severe or excruciating pain, radiating pain to the legs, history of fracture, surgery, or constitutional symptoms like fever, malaise, etc indicating infection, any inflammatory conditions, radiographic changes showing cervical spinal malformations, osteoporosis, bony abnormalities, scoliotic or kyphotic spine, pregnancy, sensory impairments, vascular causes of radiating pain or neoplasms were excluded. The above stated conditions were ruled out on the discretion of a medical professional.

Study Design

The study was designed as a two group pre-test and post-test longitudinal study. Pain and Functional Disability were taken as outcome measure for this study. The experimental design included a pre-test measure of the dependent variables Pain and Disability and after 4 weeks of treatment the dependent variables were measured again. Visual Analogue Scale was used for measuring pain and Oswestry Disability Questionnaire for finding level of functional disability.

All the selected subjects were informed in detail about the type and nature of the study and were made to sign the informed consent. After taking down the demographic data the visual analogue scale and Disability scores were noted down. Then the subjects in both the groups were made to do common warm-up exercises initially. These exercises include light aerobic work in the form of exercise on bicycle for 5 minutes at moderate pace or jogging for 10 minutes.

Group 1: Core Stability Training

Core stabilization refers to the muscles that act to stabilize the lumbar spine and lumbopelvic and hip complex as well as muscles acting to control position of the head, arms, and trunk segment relative to the body's base of support.⁴¹

All the subjects in the Group 1 had undergone one month core stability training which comprises of exercises aimed at increasing spinal stability. Isolated Lumbar stabilizing muscles training: Development of the perception of the isolated isometric specific contraction of the stabilizing muscles.

Group 2: Stabilizer Bio-feedback training

The Stabilizer is a simple device which registers changing pressure in an air filled pressure cell. This allows body movement, especially spinal movement, to be detected during exercise. The unit consists of a combined gauge/inflation bulb connected to a pressure cell. The Stabilizer is used to monitor and provide feedback on body movement during exercise. The three-chamber pressure cell of the Stabilizer was placed between the part of the body and floor. It was inflated till it molds between the body part and the supporting surface. A pressure of 40 mmHg was maintained as the resting pressure of the inflated cell. Changes in body weight on the cell on any of the three compartments will register a pressure change on the gauge.

The Stabilizer biofeedback training was given for 6 days a week for four weeks. This maneuver was repeated for all the muscles responsible for lumbar spinal stability. All the activities were performed in three sets of 10 repetitions each. Each training session comprises of 45 mins.

Results

Thirty subjects, 23 males and 7 females were randomly divided into two groups; group 1 and group 2.

Table 1: Demographic Characteristics of Subjects in both the groups

Group	Visual Analogue Scale Score (n=15)		
	Baseline	After 15 Days	After 30 days
Core Stability Exercise	6.7±1	4.6±1.7	2.3±1.7
Pressure Biofeedback	6.2±1.7	3.5±1.8	0.9±1
Group	Oswestry Disability Index Score (n=15)		
	Baseline	Baseline	Baseline
Core Stability Exercise	40.4±12.5	30.6±12.6	19.8±14
Pressure Biofeedback	46.6±21	24.8±18	5.1 ± 6

Fifteen subjects were taken in each group, with a mean age and standard deviation of subjects in group 1 and 2 were 23.4±1.95 and 21.66±1.95 respectively. There were 12 males and 3 females in group1 and 11 males and 4 females in group2. Statistical analysis revealed no statistically significant differences between the groups during the baseline readings.

The outcome measures were pain, measured on Visual Analogue Scale and functional disability measured by Oswestry Disability Index. All the measurement was taken on the 1st day, 15th day and after completion of the training on 30th day. All the data were compared for both the within group and between group analysis using repeated measure ANOVA.

Effect of Training on Pain and Disability: Assessment of improvement in this experiment included two dependent variables; Visual Analog Scale (VAS) score and Oswestry Disability Index/Questionnaire Score (ODI). One-way MANOVA demonstrated a significant improvement in visual analogue scale and ODI among the three readings taken after 1st, 15th and 30th day respectively. Further analysis was performed to determine the difference in dependent variables between both the groups a post hoc analysis was performed using Bonferroni test. In Group1 it was found that there was a significant difference for visual analogue scale ($F = 33.29$, $P=0.001$), and ODI ($F= 73.19$, $P=0.001$) among all three readings. In Group 2 also it was found out that there was a significant difference for visual analogue scale ($F=20.79$, $P=0.001$) and ODI ($F= 32.56$, $P=0.001$) between the three readings. Results are presented in the following sections. (Table 5.2)

Effect of Core Stability Exercises on Pain and Disability: The group1 getting Core Stability Training has shown significant improvement in terms of decrease in pain and functional disability. The one month Core Stability Training brought 66% decrease in pain and 25% improvement on Oswestry Disability Index.

Even after 15 days of training there was a positive effect of this training which improved further with four weeks of training. Post Hoc analysis has shown that there was a significant improvement in baseline readings of visual analogue scale and ODI when the comparison was made between 1st and 15th day, 1st and 30th day and 15th and 30th.

Effect of Stabilizer Bio-Feedback Training on Pain and Disability: In Group2 also a similar improvement was seen as with the Group1. There was a significant decrease among the baseline reading measured on 1st day and the consequent readings taken on 15th and 30th day. Post hoc analysis has shown that there was improvement in pain and disability score with the training which was found statistically significant.

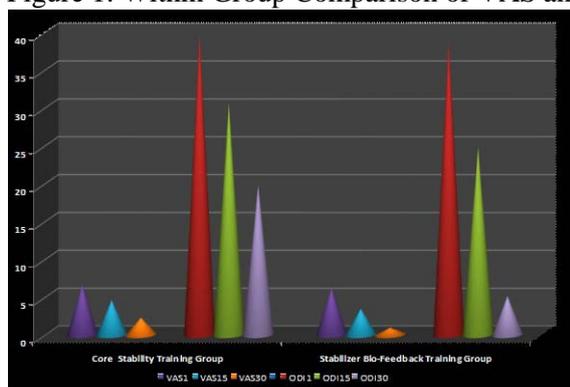
Table 2: Repeated Measure ANOVA for visual analogue scale and ODI

			Group1 (n=15)	Group2 (n=15)
Visual Analogue Scale	Day 1	M ± SD	6.69±0.9	6.2±1.7
	Day 15	M ± SD	4.6±1.7	3.46±1.8
	Day 30	M ± SD	2.27±1.7	0.94±1
	RANOVA	F	33.29	73.19
		P	0.001	0.001
Post-Hoc Analysis (P=)	1vs15	0	0.001	
	15vs30	0.002	0.001	
	1vs30	0.00	0.00	
Oswestry Disability Index	Day 1	M ± SD	40.4±12.5	46.6±21.2
	Day 15	M ± SD	30.6±12.6	24±18.2
	Day 30	M ± SD	19.8±14.6	5.13±6.1
	RANOVA	F	20.79	32.56
		P	0.001	0.001
	Post-Hoc Analysis (P=)	1vs15	0.003	0.003
		15vs30	0.003	0.003
1vs30		0.00	0.00	

Between Group Analysis

To find out, how both the groups have behaved in terms of improvement and to test the experimental hypothesis, a between group analysis was performed using independent t-test for all the variables. The findings suggested that the group getting Stabilizer Bio-feedback training have shown comparatively more improvement on both Visual Analogue Scale and Oswestry Disability Index.

Figure 1: Within Group Comparison of VAS and ODI



The result shows that for initial two weeks treatment both the group behaved similarly and improvement in pain and functional disability was seen in both the groups.

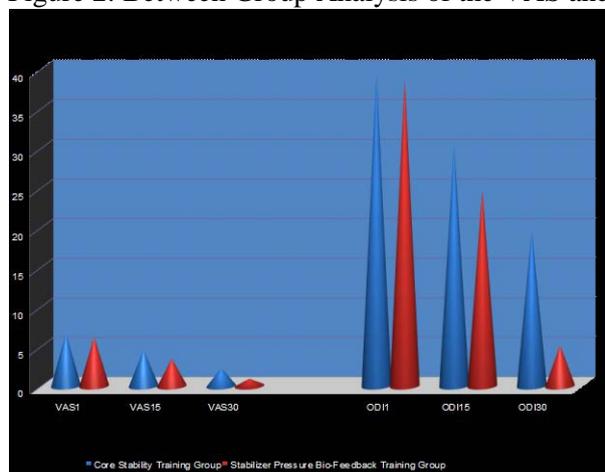
Table 3: Between group analysis of visual analogue scale and ODI among all three conditions.

Variables	Group 1 (n=15)	Group2 (n=15)	t-test	
	M ± SD	M ± SD	t	p
VAS1	6.7±0.9	6.2±1.7	0.903	0.374**
VAS15	4.6±1.7	3.5±1.8	1.738	0.093**

VAS 30	2.3±1.7	0.9±1	2.585	0.015*
ODI 1 5	39.7±12.5	38.8±16.2	0.176	0.861**
ODI 15 6	30.6±12.6	24.8±14	1.016	0.381**
ODI 30	19.8±14	5.1±6.1	28	0.001*

As evident by the results there was no statistically significant difference between the group getting Core Stability Training or Stabilizer Pressure Bio-Feedback Training at 1st day (VAS: $t = 0.903$, $p = 0.374$ ODI: $t = 0.176$, $p = 0.861$) or after 15 days of treatment (VAS: $t = 1.738$, $p = 0.093$ ODI: $t = 1.016$, $p = 0.381$). But after four weeks of treatment the result shows a statistically significant difference between both the groups and also it shows that the group getting Stabilizer Bio-Feedback training improved better on both the scales Visual Analogue Scale as well as Oswestry Disability Scale.

Figure 2: Between Group Analysis of the VAS and ODI



The results of this study support the initial hypothesis that Stabilizer Pressure Biofeedback training of the “stability” muscles of the trunk is more effective in reducing pain and functional disability in patients with chronic low back pain. Analysis of the pain and functional disability score data in the experimental group revealed that this treatment approach was more effective than the Core Stability Training. These findings support the Punjabi’s²¹ hypothesis that the stability of the lumbar spine is dependent not solely on the basic morphology of the spine, but also the correct functioning of the neuromuscular system. Also, Radebold A.²² stated that muscle recruitment and timing pattern play an important role in maintaining lumbar spine stability. Therefore, exercises specifying the isolated contraction of multifidus muscle were incorporated in the Group2, in contrast to the core stability exercises in the group1. Hence, it gets clearer that the significantly better results in the group1 are due to the proper recruitment of the specific back muscles, which was facilitated by the Stabilizer. Hodges and Richardson¹⁹ showed that the co-contraction of the transverse abdominis and multifidus muscles occurred prior to any movement of the limbs. They also showed that the timing of coordination of these muscles was very significant, and that back injury patients were unable to recruit their transverse abdominis and multifidus muscles early enough to stabilise the spine prior to movement. It hence makes it utmost important to strengthen these deep and local muscles to uproot the low back aches completely and maintains segmental stability. Therefore, it can be safely stated that in the present study, the group in which this was emphasized showed a significant improvement in terms of pain and functional ability in comparison to the other group, as evident in the result section.

By definition, the deep-trunk muscles act as 'stabilizers' and are not involved in producing movements, but instead use static or isometric contractions. Furthermore, they must act as stabilisers continuously during everyday activities as well as sport, and so require very good endurance of low-level forces. Muscle impairments are not more of strength but rather problems in motor control. This

is what was kept in mind while planning the exercise regime of the group1, which enhanced the spinal segmental support and control. The subjects were trained to selectively contract the stabilizers by the help of stabilizer biofeedback and later worked on improving the endurance in terms of static control. This form of specific training at low levels of activation supports the recent findings of Cholewicki and McGill²⁴ that only low levels of maximal voluntary contraction of the segmental muscles are required to ensure the stability of the spine in vivo. It is also consistent with assertion that motor learning and control are not simply a process of strength training, but depend on patterning and inhibition of motor neurons, with the acquisition of skills occurring through selective inhibition of unnecessary muscular activity, as well as the activation of additional motor units.

Further, Shaughnessy M²⁵ in a pilot study discovered that program of lumbar stabilization is effective in improving quality of life and functional outcome in patients with chronic low back pain. A similar study conducted by Cholewicki and McGill¹¹ revealed that lumbar stability is maintained in vivo by increasing the activity (stiffness) of the lumbar segmental muscles, and highlighted the importance of motor control to coordinate muscle recruitment between large trunk muscle and small intrinsic muscles during functional activities, to ensure stability is maintained. This concept when merged with that of Bergmark,²⁵ specifying the classification of trunk muscles into local and global muscles, has overemphasized the need to train the deep muscles of the back to provide segmental stability while directly controlling the lumbar segments. The present study on segmental stabilisation has incorporated all these theories and has come out with an exercise regimen working directly at the deep stabiliser muscles. The positive result with a significant difference in terms of pain and functional ability thus supports the hypotheses put forth by eminent researchers as stated above. In addition to this, Tesh K M²⁶ has also suggested that the muscles of the antero-lateral abdominal wall increase the stability of the lumbar region of the vertebral column by tensing the thoracolumbar fascia and by raising intra-abdominal pressure. Of the back extensor muscles, the lumbar muscles is considered to have the greatest potential to provide dynamic control to the motion segment, in its neutral zone. This study was considered important on account of the fact that patients of chronic low back pain would always seek not only a relief from pain but also the ability to perform ADL without discomfort. Hence, the patients need to be trained not only for the static control but also dynamic functional independence. With this view in mind, the exercises focused on the appropriate strengthening of the deep back muscles such that it can lead to alleviation of pain during motor tasks as well, thus aiming at complete recovery.

The most significant finding of the present study was the sustained reduction in symptoms and functional disability levels in the experimental groups at the 15th and 30th day follow up. The findings of this study support the view that a change in the motor program had occurred in both the group after the intervention, such that the automatic pattern of recruitment of the abdominals to stabilize the spine during a motor task incorporated higher levels of deep abdominal muscle activity. This appears to represent an enhanced ability, in those in the experimental group, to stabilize dynamically their spine during functional tasks. Hence it can very well be stated that stabilization exercises do appear to provide additional benefits to patients with sub acute or chronic low back pain who have no clinical signs suggesting the presence of spinal instability. Therefore, such population of chronic low back pain patients must be identified and treated with specific stabilizing exercise intervention based on motor control and motor learning in order to achieve efficient relief of excessive load from the spine, to enhance segmental stabilization and to control pain in a functional manner.

Future Research

The significant difference in results observe in this study suggests the effectiveness of both exercise regimes, but also suggests the need for a more comprehensive research for future including patients with gross instability (evident radiologically) and side-to-side differences while administering stabilization exercises. Other factors such as role of age, gender, duration of study etc. may also be included in the future studies. Use of sophisticated devices such as electromyography biofeedback units, or real time ultrasound scanners might have made this study more sensitive in terms of generalization of results.

Relevance to Clinical Practice

Since EMG was not used in this study, the results observed in this study suggest the cost effectiveness in assessing pain and disability as would be required in clinical settings (devoid of EMG and other sophisticated equipments to assess level of muscle imbalance). There are studies stating the

use of exercise regimes only (without electrotherapeutic modalities) in which patients have benefited and thus the exclusion of these does not pose to be unethical. Also, since no other modality was used, it is a cost-effective method of treatment, as would be preferred in clinical settings.

Limitations

- The long term effect of the study could not be established due to lack of follow up on pain and disability, after the thirty days of treatment.
- The home care programme taught to the patients could not be done under direct supervision.
- Apart from the clinical physical therapist palpating the transversus abdominis and multifidus muscle contraction in the subjects, there was no other means of verifying whether these muscles were recruited appropriately.

Conclusion

The findings of this trial support the view that the functional integration of Stabilizer Bio-feedback training directed at the deep abdominals and the lumbar muscles are effective in reducing pain and functional disability in patients with chronic low back pain. This supports Punjabi's hypothesis, that spinal stability is dependent on an inter play between the passive, active, and neural control systems. Specific training of the muscles considered to provide dynamics stability to the lumbar spine may act to maintain the neutral zones of the motions segment within more normal limits during functional activity. In addition, the results of this study indicate that a "Stabilizer Bio-Feedback" treatment approach directed at specific muscles activation is more effective than Core Stability Training commonly used in patients with chronic low back pain.

References:

1. Anderson GBJ. Epidemiology features of chronic low back pain. *Lancet* 1999; 354: 581-585
2. Taylor VM, Deyo RA, Cherkin DC, Keuter W. Low back pain hospitalization: recent United States trends and regional variation. *Spine* 1994; 19: 1207-1213
3. Hart LG, Deyo RA, Cherkin DC. Physician office visits for low back pain. Frequency, clinical evaluation, and treatment patterns from a U.S. national survey. *Spine* 1995; 20: 11-19.
3. Frank A. Low back pain. *BMJ* 1993; 306: 901-909
4. Koes B, Bouter L, Beckerman H. Physiotherapy Exercises and Back Pain: a blinded review, *British Medical Journal* 1991; 302: 1572-1576.
5. Richardson CA, Jull GA. Muscle control – pain control. What exercises would you prescribe? *Manual Therapy* 1995; 1: 2-10.
6. Richardson C, Jull G, Hodges P, et al; Therapeutic Exercise for Spinal Segmental Stabilization in Low Back Pain – Scientific Basis and Clinical Approach. Churchill Livingstone, Edinburgh. 1999.
7. Bergmark A. Stability of the lumbar spine. A study in mechanical engineering. *Acta Orthopaedica Scandinavica* 1989; 230(Suppl): 20-24.
8. Hides JA, Richardson CA, Jull GA. Multifidus muscle recovery is not automatic following resolution of acute first episode low back pain. *Spine* 1996; 21: 2763-2769.
9. Hodges PW, Richardson CA. Inefficient muscular stabilisation of the abdominis. *Spine* 1996; 21: 2640-2650.
10. Richardson CA, Snijders CJ, Hides JA, et al; The relationship between the transversely oriented abdominal muscles, sacroiliac joint mechanics and low back pain. *Spine* 2002; 27(4):399-405.
11. Hodges, P. and Moseley, L., Pain and motor control of the lumbopelvic region: effect and possible mechanisms, *J Electromyogr Kinesiol*, 13 (2003) 361-70.
12. Hodges, P.W., B.P. Hons, and C. Richardson. Inefficient muscular stabilization of the lumbar spine associated with low back pain: A motor control evaluation of transversus abdominis. *Spine* 21:2640–2650. 1996.
13. Nichols, J.F., D. Medina, and E. Dean. Effects of strength, balance, and trunk stabilization training on measures of functional fitness in older adults. *Am. J. Med. Sports* 3:279–285. 2001.
14. Richardson, C.A., and G.A. Jull. Muscle control—pain control: What exercises would you prescribe? *Man. Ther.* 1:2–10. 1995

15. Richardson, C., G. Jull, P. Hodges, and J. Hides. *Therapeutic Exercise for Spinal Segmental Stabilization in Low Back Pain: Scientific Basis and Clinical Approach*. Edinburgh, Scotland: Churchill Livingstone, 1999.
16. Mc Gills, S.M. Low back exercises: Evidence for improving exercise regimens. *Phys. Ther.* 78:754–765. 1998.
17. O’sullivan, P.B., G.D.M. Phyty, L.T. Twomey, and G.T. Allison. Evaluation of specific stabilizing exercise in the treatment of chronic low back pain with radiologic diagnosis of spondylolysis or spondylolisthesis. *Spine* 22:2959–2967. 1997.
18. Stanton, R., P.R. Reaburn, and B. Humphries. The effect of short-term Swiss ball training on core stability and running economy. *J. Strength Cond. Res.* 18:522–528. 2004.
19. McGill, S., A. Childs, and C. Liebenson. Endurance times for low back stabilization exercises: Clinical targets for testing and training from a normal database. *Arch. Phys. Med. Rehabil.* 80:941–944. 1999.
20. Radebold A, Cholwicki J, PasnjabiMM, et al. Muscle response pattern to sudden trunk loading in healthy individuals and in patients with chronic low back pain. *Spine* 2000;25:947-54
21. Richardson CA, Jull GA, Hodges P Hides . *Therapeutic exercise for spinal segmental stabilization in low back pain scientific basis and clinical approach*, Edinburgh, Churchill Livingstone ,1999
22. Cholewicki J, Mc Gill SM. Mechanical stability of the in vivo lumbar spine: implications for injury and chronic low back pain. *Clin Biomech* 1996; 11: 1-15.
23. Shaughnessy M, Caulfield B. Apilot study to investigate the effect of lumbar stabilization exercise training on functional ability and quality of life in patients with chronic low back pain. *Int J Rehab Res.* 2004 Dec;7(4):297-301.
24. Bergmark A . Stability of the lumbar spine. *Acta Orthop Scand* 1989;60:1-54.
25. Tesh KM, Dunn JS, Evans JH. The abdominal muscles and vertebral stability. *Spine.* 1987 Jun;12(5):501-8.

TYPE 1/TYPE 2 CYTOKINE SERUM LEVELS AND ROLE OF INTERLEUKIN-18 IN CHILDREN WITH STEROID SENSITIVE NEPHROTIC SYNDROME

Sherein A. Shalaby^{a,e}; *Howaida M. Al-Edressi*^b; *Shereen A. El-Tarhouny*^c; *Mohamed Fath EL-Bab*^{d,f}; *Mohamed A. Zolaly*^a.

^aDepartment of Paediatrics, College of Medicine, Taibah University, ^bDepartment of paediatric nephrology, Maternity and Children Hospital, ^cDepartment of Clinical Biochemistry, ^dDepartment of Physiology, College of Medicine, Taibah University, AL-Madinah AL-Munawarah, Kingdom of Saudi Arabia

^e Department of Paediatrics, ^f Department of Physiology, Faculty of Medicine, Suez Canal University, Ismailia. Egypt

Abstract:

Aims: In view of the conflicting evidence of helper T cell type 1 (Th1) or type 2 (Th2) pattern of cytokine synthesis in steroid sensitive nephrotic syndrome (SSNS) this study aims to assess type-1 /type-2 cytokines in different stages of SSNS and define the potent involvement of IL- 18

Methodology: We prospectively studied thirty children with steroid sensitive nephrotic syndrome (SSNS), aged 2–12 years. Thirty children were in active stage before treatment initiation, the same 30 children in remission still on steroids; 21/30 in remission off steroids as well. The control group included 30 healthy age and sex matched siblings. Using ELISA technique we assessed serum levels of Serum, IL-2, IFN- γ , IL-4, IL-13 and IL-18 in different stages of the disease and in controls.

Results: IL-2 levels were not significantly different in children of all disease stages of SSNS and controls ($p > 0.05$). Levels of IL-4, IL-13 and IL-18 were significantly higher in the active stage of SSNS compared with the remission stages and controls ($p < 0.05$). But, serum IFN- γ was significantly lower in children with active disease compared with remission stages and controls ($p < 0.05$). In children with SSNS, of all disease stages, serum levels of IL-18 were significantly correlated with both IL-4 and IL-13 ($r = 0.72$ and $p < 0.0001$, $r = 0.82$ and $p < 0.0001$, respectively).

Conclusion: Children with active SSNS seem to have a shift to type-2 cytokine production, and IL-18 expression is significantly correlated with this type-2 immune response.

Key Words: Type 1 cytokine, Type 2 cytokine, Nephrotic Syndrome

Introduction:

Idiopathic nephrotic syndrome (INS) is considered to be an immune-mediated disease [1]. But, how T-cells affect the course of the disease remain unanswered question. However, circulating factors were proposed to be released from activated T-cells which may affect the pathogenesis of the disease. Several cytokines are considered prime candidates for the role of mediators of INS [2].

Based on the predominant cytokines, the immune response can be functionally subdivided into type-1 and type-2. Type-1 response, which normally prevails, and produce interferon-gamma (IFN- γ) and interleukin (IL)-2 and enhance both the production of complement-fixing and opsonizing antibodies and, macrophage activation resulting into delayed type hypersensitivity (DTH), Whereas type-2 response is dominated by IL-4 and IL-13 and provide optimal help for antibody production; and promote both mast cell growth and eosinophil differentiation and activation, resulting in humoral responses [3].

Measurement of cytokine level in serum or urine in INS patients have been performed by many investigators. Increased levels in serum or urine in relapse were reported in interleukin (IL)-2 [4], soluble IL-2 receptor [5-8], and (IFN- γ) [5, 9], IL-4 [9, 10], IL-12 [11], INS has been suggested to be a Th2-dependent glomerular disease [12] depending on the evidenced association between it and atopy and allergy [13-15], many allergic disorders, as asthma, allergic rhinitis and eczema, are typically linked to the presence of a Th2 immunologic response.

Further support to the hypothesis that INS is a Th2-dependent glomerular disease comes from the elevated serum IgE and preservation of IgG4 observed in INS, [16-18]. In addition, levels of IL-4 and IL-13, which are two of the major Th2- associated cytokines, have been shown to be elevated in INS patients in relapse [9, 10, 16, 19]. . However, contradictory observations have been also reported, [19- 21].

IL-18 is an exclusive cytokine which, can stimulate both type-1 and type-2 immune responses depending on the cytokine environment [4].

Since Matsumoto and Kanmatsuse [22] elegantly demonstrated that IL-18 was involved in the active stage of the disease, and that there was a positive correlation between IL-18 production and disease activity was by, efforts have been made to identify the pathogenetic vascular permeability factors VPF(s) released from T-cells, such as cytokines, as well as to clarify the participated cells in its pathogenesis.

For the last two decades many studies investigated the cytokine pattern in SSNS with sometimes conflicting results [5–7]. So, it is still questionable is type-2 response dominated in active SSNS.

Besides, literature is lacking regard the involvement of IL-18.

The current study aims to evaluate prospectively type-1/type-2 cytokine profile in the same children with different stages of primary SSNS and to investigate the potent involvement of IL-18.

Subjects and methods:

This is a cross-sectional prospective longitudinal study; the study was carried on in Madinah Maternity and Children Hospital (MMCH) The hospital is the main maternity and children referral hospital in the Madinah Al-Munawarah, .

The study investigated thirty children (2 -12 years.) who were proved to have steroid sensitive nephritic syndrome (SSNS) and, consecutively admitted in our department during the period from August 2010 to August 2012. Active stage of SSNS was defined as increased urinary protein excretion (Albustix $\geq 2+$ for at least three consecutive days or > 40 mg/m² per hour) and serum albumin ≤ 25 g/l. Remission was defined as serum albumin concentration ≥ 35 g/l and normal protein excretion (Albustix trace or negative for at least three consecutive days or 5 mg/m² per hour) [23]

Children with SSNS were subdivided into 3 groups: group (A) Patients in the active stage (30 children), Patients in the remission stage were further subdivided into two groups:

(B1) remission phase still on steroid dose of prednisone of 40 mg/m² on alternate day regimen (30 children) and, (B2) remission phase off steroid treatment for at least 6 months (21/30 children).

Patients (30/30) were tested both in active stage and at the time of testing no patient or control was taking any immunomodulating drug (i.e. cyclosporine A, cyclophosphamide, levamisole or MMF), to avoid their effects on the they have history of recent (within the previous 6 months) infection and/or inflammatory conditions, or abnormal urinary sediments (abnormal casts or crystalluria). The control group (group C) included thirty age and sex matched healthy children aged 2-/12 years (male, 15 female, 15) with (urinary protein/ creatinine ratio (< 0.2) and, serum albumin level (4.1 \pm 4.6 mg/dl).

Ethical considerations:

The study was approved by the Ethics Committees of Madinah Maternity and Children Hospital (MMCH). All parents of the patients and the controls were informed about the study and written informed consent was obtained

Methods:

A comprehensive clinical examination was done by one of the investigators or his assistant to all individuals who consented to participate in the study. The examination aimed to elicit signs and symptoms of the disease, presence or absence of complications and risk factors for relapse (upper respiratory tract infections, fever, common cold, vaccinations.....).

The following investigations were routinely performed in all children in the laboratory of Maternity and Children Hospital.

Blood: full blood count, urea; Creatinine; liver function test; ASOT; C3/C4; Varicella titers , urine culture and urinary protein/creatinine ratio, urinalysis including glucose, in all children before commencing steroids and, hepatitis B status to exclude secondary hepatitis.

Serum cytokine assay:

Blood samples were collected from all patients and controls in the test tubes without anticoagulant under sterile conditions. Serum was separated by centrifugation at 300g for 10 min, then divided in small aliquots and stored at -80°C for future serum cytokine assessment.

Serum IFN- γ , IL-2, IL-4, IL-13 and IL-18 concentrations were measured by using quantitative colorimetric sandwich ELISA kits purchased from R&D China Co. Ltd., Shanghai. Following the manufacturer's instructions, Each cytokine sample was run in duplicate and the mean cytokine concentration was calculated

Statistical analysis:

Statistical analysis was performed using Statistical Package for the Social Sciences for Windows (SPSS version 17). Values were expressed as median and ranges. The non-parametric Wilcoxon signed-rank test was used to compare differences between study groups with paired data. For non-paired data, statistical significance was analyzed by the Mann-Whitney U test. Spearman's coefficient of correlation (r) and Regression analysis model were used to determine the correlations. $P < 0.05$ was considered to be statistically significant.

Results:

The current study investigated 30 children (17 males and 13 females) in the active stage of SSNS aged from 2 to 12 years (median = 3.52 years). The same group of children was re-investigated during remission phase still on steroid treatment; 21/30 was tested in remission phase off steroids for at least 6 months as well. All 30 patients who were studied in remission still on steroids were on the same dose of prednisone (40 mg/m² on alternate day regimen). The control group consisted of 30 healthy siblings of the patients (15 males and 15 females) aged from 2 to 12 years (median = 4.12 years). There was no statistically significant difference between the group of children with nephritic syndrome and the control group in regard to age or sex ($p > 0.05$).

Table 1 summarizes the results of serum IL-2, IFN- γ , IL-4, IL-13 and IL-18 levels and p values in all study groups. There was no significant difference in IL-2 levels between nephrotic children of all disease stages and controls ($p > 0.05$). Children with active stage had significantly lower levels of IFN- γ when compared with the two remission phases (remission on steroids, remission off steroids ($p = 0.005$, $p = 0.001$, respectively) and controls ($p = 0.007$), While there was no significant difference between controls and each of the two groups of remission ($p > 0.05$).

To the opposite children with active stage of SSNS had significantly higher levels of IL-4 when compared with the two remission phases (remission on steroids, remission off steroids and $p < 0.0001$, $p < 0.0001$, respectively). In comparison to the controls patients in both remission groups showed significantly elevated IL-4 levels ($p < 0.0001$ and $p = 0.034$ respectively). Regarding, serum IL-13 levels, they were significantly higher in the active stage of SSNS compared with the two remission groups ($p < 0.0001$, $p < 0.0001$, respectively). IL-13 levels, still elevated during both remission phases, compared with the controls ($p < 0.0001$ and $p = 0.002$, respectively). Worth notice were also the findings regarding IL-18 levels. In the active stage of SSNS, IL-18 levels were significantly higher compared with the two remission groups, and controls ($p = 0.003$, $p = 0.001$, and $p < 0.0001$ respectively), and IL-18 levels, even lower than in the active stage, remained significantly higher in both remission groups, compared with the controls ($p < 0.0001$, $p < 0.0001$, respectively).

As shown in Fig. 1a and 1b, using the linear regression analysis serum IL-18 and IL-4 levels were significantly correlated in nephrotic children in all disease stages ($r = 0.72$ and $p < 0.0001$).

In children in active stage of SSNS, serum IL-18 and IL-4 levels, were also significantly correlated ($r = 0.73$ and $p = 0.001$) Fig. 1b.

Fig. 2a illustrates the significant correlation between serum IL-18 and IL-13 levels in nephrotic children in all disease stages ($r = 0.82$ and $p < 0.0001$). In children in active stage of SSNS, serum IL-18 and IL-4 levels, were also significantly correlated ($r = 0.76$ and $p = 0.001$) Fig. 2b.

Table 1: Serum Levels, median and range values of IL-2, (IFN- γ), IL-4, IL-13 and IL-18 in patients with different stages of SSNS and in controls.

serum cytokine level	Active state Group (A) (n = 30)	Remission on steroids Group (B1) (n = 30)	Remission off steroids Group (B2) (n = 21)	Controls Group (C) (n = 30)
IL-2(pg/ml) Median Range	8.3 8- 13.5	98 7.9- 12.8	9.1 7.6 – 12.9	8.7 7.9 – 13.1
	^a P > 0.05*		^a P > 0.05	^a P > 0.05
(IFN- γ) (pg/ml) Median Range	16.25 14.76 – 29.9	24.5 13.72 – 36.9	25.9 15.28 - 35.7	21.8 13.4 – 33.3
		^a P = 0.005 ^c P > 0.05	^a P = 0.001 ^c P > 0.05	^a P = 0.007
IL-4 (pg/ml) Median Range	65 19.7 – 109.5	33.7 15.8 – 81.8	21.9 11.6 – 69.5	14.5 7.7 - 42.
		^{a, c} p < 0.0001	^a p < 0.0001 ^c P = 0.034	^a p < 0.0001
IL-13 (pg/ml) Median Range	59.9 32.5 – 160.8	28.7 11.8 – 133.7	19.7 11.1 - 96.3	12.7 10.4 – 20.6
		^{a, c} p < 0.0001	^a p < 0.0001 ^c p = 0.002	^a p < 0.0001
IL-18 (pg/ml) Median Range	1564 532.1 - 2716	1311 417 – 2245	765 321 - 1278	109 29 - 236
		^a p = 0.003 ^c p < 0.0001	^a p = 0.001 ^c p < 0.0001	^a p < 0.0001

P < 0.05 was considered to be statistically

^a p; comparing active stage with remission on steroids, remission off steroids ,and controls.

^c p; , comparing controls with remission on steroids, and remission off steroids.

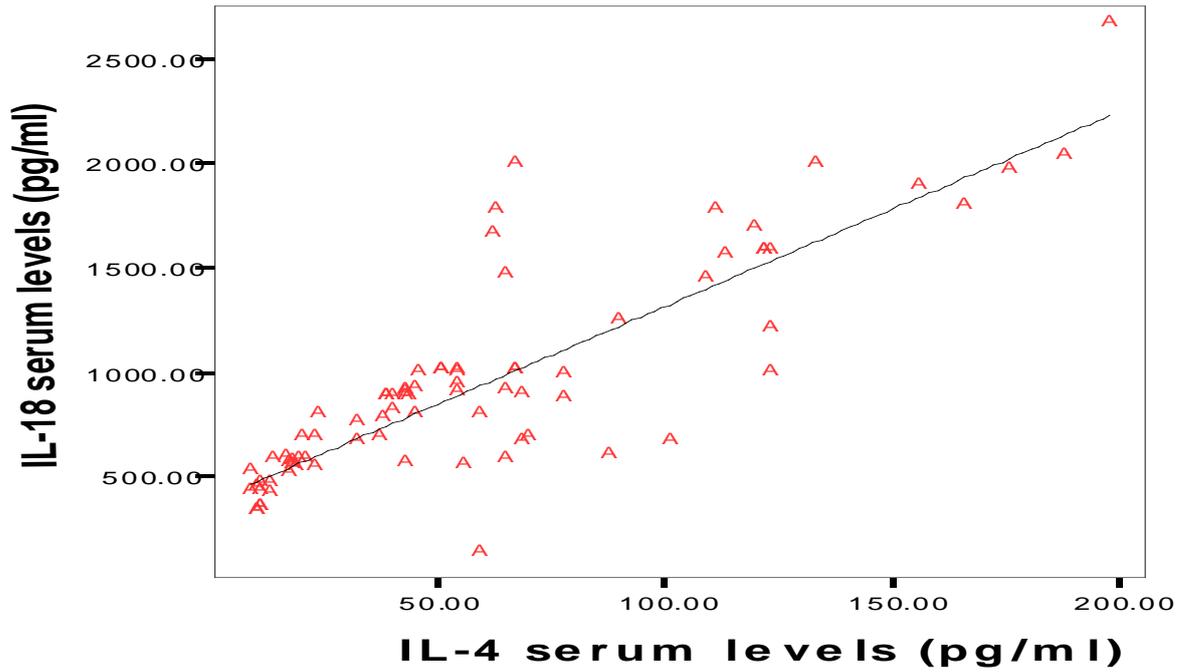


Figure 1a: Correlation of serum IL-18 and IL-4 levels in children with SSNS in all stages.

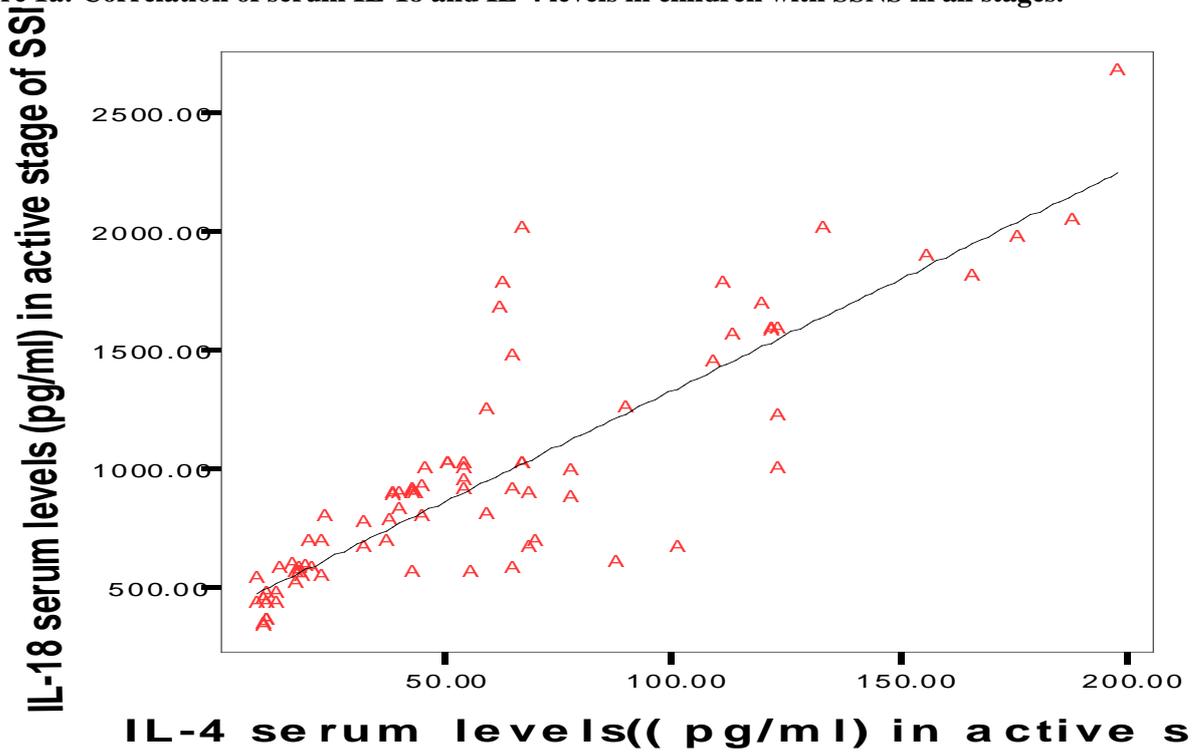


Figure 1b: Correlation of serum IL-18 and IL-4 levels in children with active stage of SSNS.

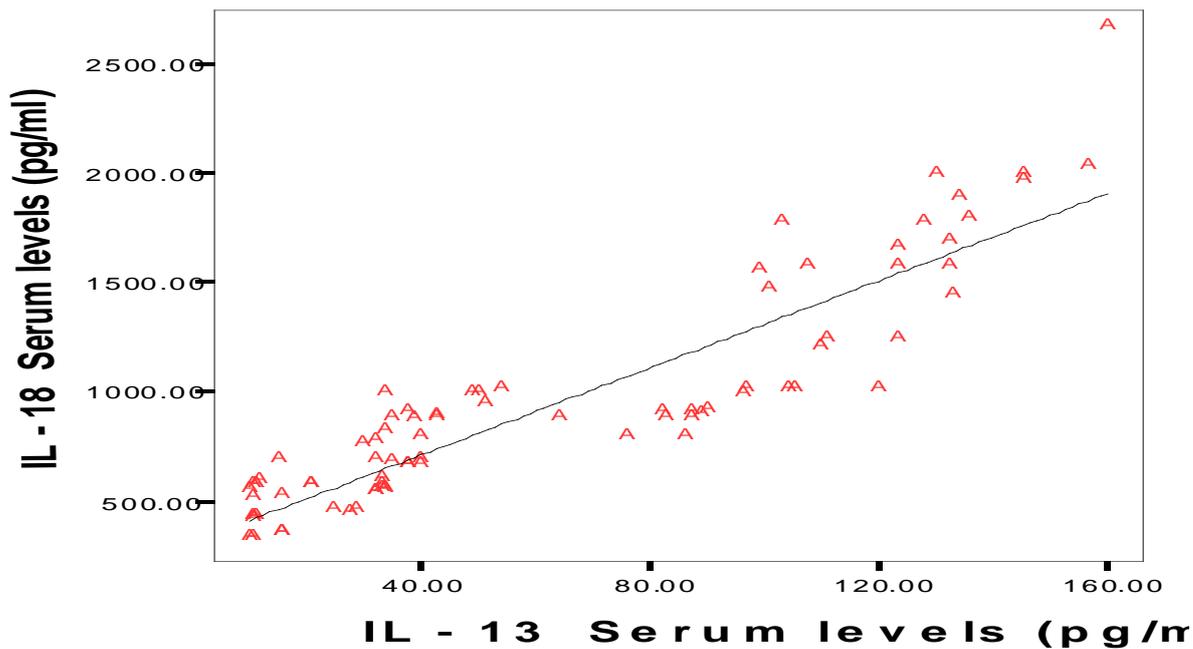


Figure 2a: Correlation of serum IL-18 and IL-13 levels in children with SSNS in all stages

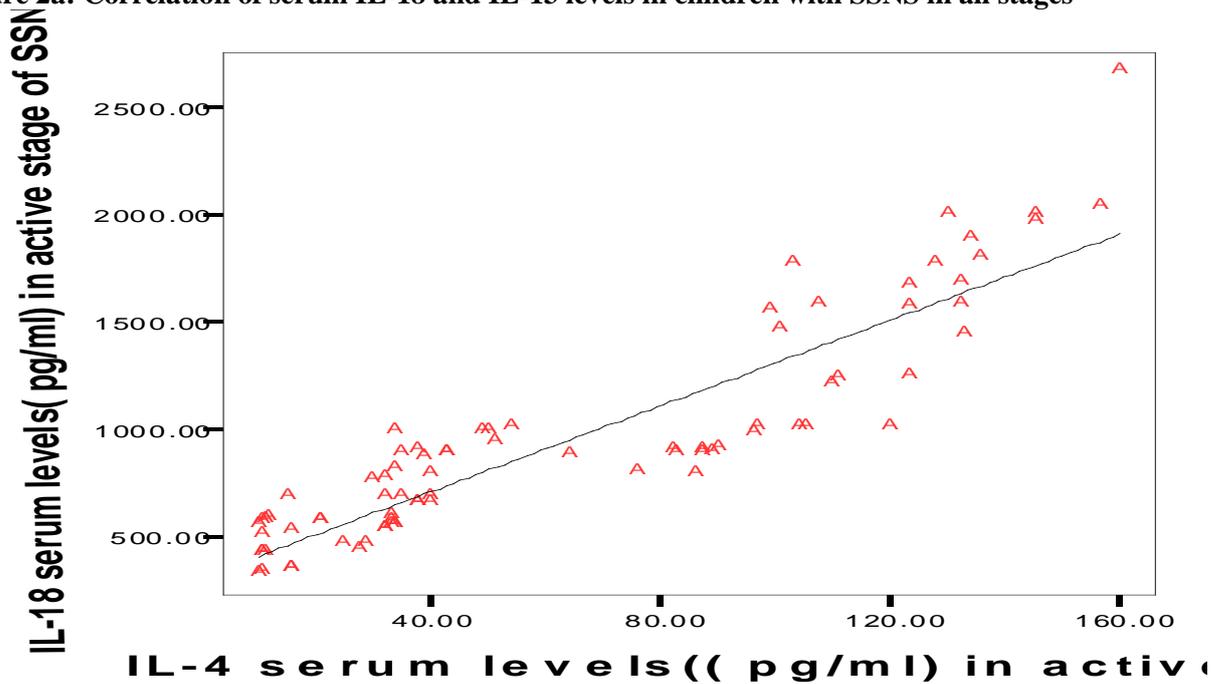


Figure 2b: Correlation of serum IL-18 and IL-13 levels in children with active stage of SSNS.

Discussion:

Studies of the type-1/type-2 cytokine patterns in the sera of patients with SSNS have generally been variable and inconsistent. This conflicting result may be due to the differences in immunologic techniques which were used to assess cytokine synthesis [24].

Our study has clearly shown that children with active SSNS showed a down-regulation of type-1 immune response. IL-2 serum levels were not significantly different in children with SSNS in all stages. This is in agreement with previous cited studies [16–19].

Printza et al. [24] reported, in agreement with our findings, no significant difference of IL-2 serum levels in children with SSNS.

On the contrary Zachwieja et al. [16] demonstrated the presence of higher intracellular expression of IL-2 using a three-color flow cytometric assay. Considering the physiological fact that increased intracellular production does not usually result in increased secretion, we can partially explain the difference between our results and Zachwieja et al. [25].

In the group of children with active, IFN- γ serum levels were lower than that in both remission groups which were not significantly different from controls. Our results are in accordance with the results of two previous studies [24, 26] which showed a decreased production of IFN- γ by stimulated peripheral blood mononuclear cells in patients with active SSNS. Cheung et al. [27] and Kaneko et al. [28], who assessed the percentage of IFN- γ producing N cells in patients with SSNS, did not find significant difference between patients and controls. The predominance of type-2 immune response in our children with active SSNS is further supported by the upregulation of the serum levels of IL-4. These results are in agreement with earlier published works by Stokowski et al. [20], Neuhaus et al. [9], and Cho et al. [10], who reported in patients with active SSNS an increased synthesis of IL-4 by stimulated peripheral mononuclear cells. Also the results of the current study are in agreement with Printza et al. [24] Kang et al. [29], and Lama et al. [30]. On the other hand, Cheung et al. [27] did not report any increase in the percentage of IL-4 producing T cells.

Further support to the shift to type-2 immune response in our children with active SSNS is demonstrated by the increase of IL-13 serum levels in all our patients with SSNS. The results of this study are in accordance with previous published studies [27, 19]. Yap et al. [19] demonstrated an elevated expression of IL-13 mRNA using a semi-quantitative reverse transcriptase PCR technique.

An experimental study on rat by Lai et al. [31] demonstrated the occurrence of podocyte injury, inducing a minimal-change like nephropathy with the over-expression of IL-13. Recently Printza et al. [24] reported that serum IL-13 levels were significantly higher in the active stage of SSNS compared with the two remission phases and that although IL-13 levels were even lower than in the active stage, they remained elevated during both remission phases, compared with the controls.

Our results highlight an important finding of increased levels of IL-18 in all disease stages, and particularly in the active stage of SSNS compared with the controls. This elevation should be greatly emphasized considering that IL-18 is a unique cytokine that can trigger both type-1 and type-2 immune responses depending on the predominant cytokines [2].

Earlier in the literature IL-18 was shown to be an IFN- γ trigger, which plays a critical role in the host defenses [32]. Recently, IL-18 has been postulated to induce IL-13 and/or IL-4 production by NK cells, mast cells and basophils [31, 32].

IL-18 is currently incriminated to play a potent role in various pathological conditions. The available data support its involvement in various diseases such as insulin dependent diabetes mellitus, rheumatoid arthritis, Chron's disease and atopy.

In collaboration with IL-12, IL-18 stimulates Th1-mediated immune responses, which play a critical role in the host defense against infection with intracellular microbes through the induction of IFN-gamma. However, the overproduction of IL-12 and IL-18 induces severe inflammatory disorders, suggesting that IL-18 is a potent proinflammatory cytokine that has pathophysiological roles in several inflammatory conditions [2]. IL-18 mRNA is expressed in a wide range of cells including Kupffer cells, macrophages, T cells, B cells, dendritic cells, osteoblasts, keratinocytes, astrocytes, and microglia. Thus, the pathophysiological role of IL-18 has been extensively tested in the organs that contain these cells. Somewhat surprisingly, IL-18 alone can stimulate Th2 cytokine production as well as allergic inflammation. Therefore, the functions of IL-18 in vivo are very heterogeneous and complicated. In principle, IL-18 enhances the IL-12-driven Th1 immune responses, but it can also stimulate Th2 immune responses in the absence of IL-12 [31].

Printza et al. [24] showed that in children with SSNS IL-18 levels were elevated in the active stage of the disease and that there was a positive correlation between IL-18 production and disease activity. Our results demonstrated that in SSNS, IL-18 was significantly correlated with IL-4 and IL-13 which are type-2 cytokines.

In conclusion, the current study reports that during the active stage of SSNS the balance is tipped in favor of type-2 cytokine pattern, and that apparently IL-18 is correlated with these type-2 cytokines.

Acknowledgement:

We are grateful to the Vice Rector of Graduate Studies & Academic Research and Deanship of Scientific Research, Taibah University, Al-Madinah Al-Munwarah, Saudi Arabia for granting the study and their support throughout the study (**grant no 674/431**)

References:

1. Van De Berg J, Weening J. Role of the immune system in the pathogenesis of idiopathic nephrotic syndrome. *Clin Sci* 2004;107:125–36
2. Nakanishi K, Yoshimoto T, Tsutsui H, Okamura H. IL-18 is a unique cytokine that stimulates both Th-1 and Th-2 responses depending on its cytokine milieu. *Cytokine Growth Factor Rev* 2001;12:53–72.
3. Webb NJ, Lewis MA, Iqbal J, Smart PJ, Lendon M, Postlethwaite RJ. Childhood steroid-sensitive nephrotic syndrome: does the histology matter? *Am J Kidney Dis* 1996; 27: 484-8.
4. Laflam PF, Garin EH. Effect of tumor necrosis factor α and vascular permeability growth factor on albuminuria in rats. *Pediatr Nephrol* 2005;21:177–81.
5. Daniel V, Trautmann Y, Konrad M, Nayir A, Scharer K. T lymphocyte populations, cytokines and other growth factors in serum and urine of children with idiopathic nephritic syndrome. *Clin Nephrol* 1997; 47: 289-97.
6. Mandreoli M, Beltrandi E, Casadei-Maldini M, Mancini R, Zucchelli A, Zucchelli P. Lymphocyte release of soluble IL-2 receptors in patients with minimal change nephropathy. *Clin* 77-82.
7. Hulton SA, Shah V, Byrne MR, Morgan G, Barratt TM, Dillon MJ. Lymphocyte subpopulations, interleukin-2, and interleukin-2 receptor expression in childhood nephrotic syndrome. *Pediatr Nephrol* 1994; 8: 135-9.
8. Kemper MJ, Meyer-Jark T, Lilova M, Muller-Wiefel DE. Combined T- and B-cell activation in childhood steroid-sensitive nephrotic syndrome. *Clin Nephrol* 2003; 60: 242-7.
9. Neuhaus TJ, Wadhwa M, Callard R, Barratt TM. Increased IL-2, IL-4 and interferon-gamma (IFN-gamma) in steroid-sensitive nephrotic syndrome. *Clin Exp Immunol* 1995; 100: 475-9.
10. Cho BS, Yoon SR, Jang JY, Pyun KH, Lee CE. Up-regulation of interleukin-4 and CD23/Fc epsilon RII in minimal change nephrotic syndrome. *Pediatr Nephrol* 1999; 13: 199-204.
11. Lin CY, Chien JW. Increased interleukin-12 release from peripheral blood mononuclear cells in nephrotic phase of minimal change nephrotic syndrome. *Acta Paediatr Taiwan* 2004; 45: 77-80.
12. Mathew JL, Kabi BC, Rath B. Anti-oxidant vitamins and steroid responsive nephrotic syndrome in Indian children. *J Paediatr Child Health* 2002; 38: 450-7.
13. Hardwicke J, Soothill JF, Squire JR, Holti G. Nephrotic syndrome with pollen hypersensitivity. *Lancet* 1959; 1: 500-2.
14. Wittig HJ, Goldman AS. Nephrotic syndrome associated with inhaled allergens. *Lancet* 1970; 1: 542-3.
15. Meadow SR, Sarsfield JK, Scott DG, Rajah SM. Steroid responsive nephrotic syndrome and allergy: Immunological studies. *Arch Dis Child* 1981; 56: 517-24.
16. Kimata H, Fujimoto M, Furusho K. Involvement of interleukin (IL)-13, but not IL-4, in spontaneous IgE and IgG4 production in nephritic syndrome. *Eur J Immunol* 1995; 25: 1497-501.
17. Yokoyama H, Kida H, Tani Y, et al. Immunodynamics of minimal change nephritic syndrome in adults T and B lymphocyte subsets and serum immunoglobulin levels. *Clin Exp Immunol* 1985; 61: 601-17.
18. Warshaw BL, Check IJ. IgG subclasses in children with nephritic syndrome. *Am J Clin Pathol* 1989; 92: 68-72

19. Yap HK, Cheung W, Murugasu B, Sim SK, Seah, CC, Jordan SC. Th1 and Th2 cytokine mRNA profiles in childhood nephritic syndrome: Evidence for increased IL-13 mRNA expression in relapse. *J Am Soc Nephrol* 1999; 10: 529-37.
20. Stachowski J, Barth C, Michalkiewicz J, et al. Th1/Th2 balance and CD45-positive T cell subsets in primary nephrotic syndrome. *Pediatr Nephrol* 2000; 14: 779-85
21. Araya CE, Wasserfall CH, Brusko TM, et al. A case of unfulfilled expectations. Cytokines in idiopathic minimal lesion nephritic syndrome. *Pediatr Nephrol* 2006; 21: 603-10.
22. Matsumoto K, Kanmatsuse K. Augmented interleukin-18 production by peripheral blood monocytes in patients with minimal-change nephrotic syndrome. *Am J Nephrol* 2001;21:20-7.
23. A report of the International Study of Kidney Disease in children. The primary nephrotic syndrome in children. Identification of patients with minimal change nephrotic syndrome from initial report to prednisone. *J Pediatr* 1981; 98:561-564
24. Printza N, Papachristou F, Tzimouli V, Taparkou A, Kanakoudi-Tsakalidou F. IL-18 correlated with type-2 immune response in children with steroid sensitive . *Cytokine*. 2008 Nov;44(2):262-8. Epub 2008 Sep 25.
25. Zachwieja J, Bobkowski W, Dodrowolska-Zachwieja A, Lewandowska- Stachowiak M, Zaniew M, Maciejewski J. Intracellular cytokines of peripheral blood lymphocytes in nephrotic syndrome. *Pediatr Nephrol* 2002;17:733-40.
26. Stefanovic´ V, Golubovic´ E, Mitic´-Zlatkovic´ M, Vlahovic´ P, Jovanovic´ O, Bogdanovic´ R. Interleukin-12 and interferon- γ production in childhood idiopathic nephrotic syndrome. *Pediatr Nephrol* 1998;12:463-6.
27. Cheung W, Wei CL, Seah CC, Jordan S, Yap HK. Atopy, serum IgE and interleukin-13 in steroid responsive nephrotic syndrome. *Pediatr Nephrol* 2004;19:627-32.
28. Kaneko K, Tuchiya K, Fujinaga S, Kawamura R, Ohtomo Y, Shimizu T, et al. Th1/Th2 balance in childhood idiopathic nephrotic syndrome. *Clin Nephrol* 2002;58:393-7
29. Kang J, Bai KM, Wang BL, Yao Z, Pang XW, Chen WF. Increased production of IL- 4 in children with simple idiopathic nephritic syndrome. *Chin Med J* 1994;107:347-50.
30. Lama G, Luogo I, Tirino G, Borriello A, Carangio C, Salsano ME. T-lymphocyte populations and cytokines in childhood nephrotic syndrome. *Am J Kidney Dis*. 2002;39:958-65.
31. Lai KW, Wei CL, Tan LK, Tan PH, Chiang G, Lee C, et al. Overexpression of interleukin-13 induces minimal-change like nephropathy in rats. *J Am Soc Nephrol* 2007; 18:1476-85.
32. Okamura H, Tsutsui H, Komatsu T, Yutsudo M, Hakura A, Tanimoto T, et al. Cloning of a new cytokine that induces IFN- γ production by T cells. *Nature* 1995;378:88.

AN ASSESSMENT OF THE SPATIAL RELATIVITY OF HEALTH, LITERACY AND SOCIO-ECONOMIC GROWTH INDEX USING GIS TECHNIQUES: A CASE STUDY ON UDUPI SOUTH WEST COAST OF INDIA, KARNATAKA

B S Maddodi¹, H N Udayashankara², B R Raghavan³, Alina E Mathews⁴

^{1, 2, 4} Department of Civil Engineering, MIT Manipal

³Professor, Coordinator Geo-informatics, Dept. of Marine Geology Mangalore University

Abstract:

GIS can be thought of as a system that digitally creates and "manipulates" spatial areas that may be jurisdictional, purpose or application-oriented for which a specific GIS is developed. Therefore, in a general sense, the term describes any information system that integrates, stores, edits, analyzes, shares and displays geographic information for informing decision making. GIS applications are tools that allow users to create interactive queries (user-created searches), analyze spatial information, edit data, maps, and present the results of all these operations. In the simplest terms, GIS is the merging of cartography, statistical analysis, and database technology. This paper presents the results of using GIS as a tool to study the health status of Udupi District. Health is a parameter that is influenced by various factors both environmental and non-environmental. The thesis, at the latter stage, emphasizes the influence of non-environmental factors such as literacy and socioeconomic index on health status of Udupi district. The results demonstrate an affirmative relation of health status with health literacy, awareness and socio-economic index of a locality, despite the well-knit health care facilities available within the district.

Key Words: Cartography, Statistical Analysis, Geographic Information, Socioeconomic Index

Introduction

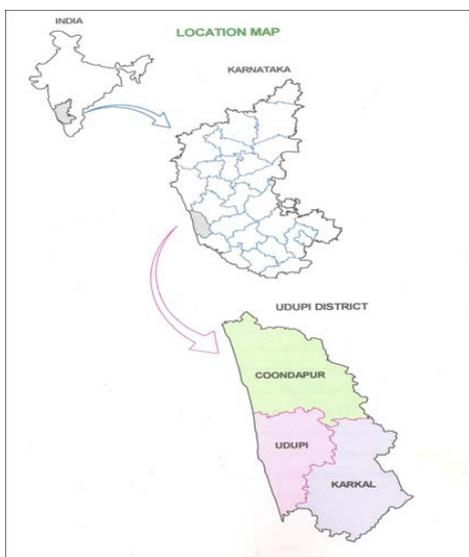
Data encompassing all the fields can be collected, stored and retrieved from such geo-based reference system. Geo-information Technology is a tool by which one can integrate the sustainable indicators easily arriving at suitable modeling and decision making, thereby helping us in evaluating the results and implementing an effective and efficient planning of the project. The health sector is an area that requires continuous monitoring and improvement and the general practice followed to achieve this is by reporting and recording statistical data, conducting studies and drawing probable interpretations directed towards improving the health status.

This is where GIS could be a breakthrough and assist in easier record keeping and interpretation of the data. The end-result would shift the interpretations from single parametric to a multi-parametric relation. Study of the various aspects influencing health can also be done to analyze the root cause of health issues and find permanent solutions. The other aspect that is dealt with in this thesis includes the influence of literacy and socioeconomic index on health. The relationship between literacy and health is complex. Literacy impacts health knowledge, health status, and access to health services. Health status is influenced by several related socioeconomic factors. Literacy impacts income level, occupation, education, housing, and access to medical care. The poor and illiterate are more likely to work under hazardous conditions or be exposed to environmental toxins.

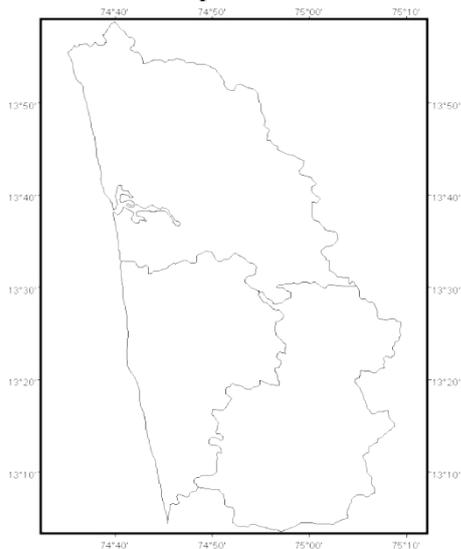
Health literacy is defined as the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions. It includes the ability to understand instructions on prescription drug bottles, appointment slips, medical education brochures, doctor's directions and consent forms, and the ability to negotiate complex health care systems. Health literacy is not simply the ability to read but requires a complex group of reading, listening, analytical, and decision-making skills, and the ability to apply these skills to health situations.

Profile Of Udupi District:

Udupi district is a fast growing region of coastal Karnataka along the west coast of India. Udupi district, as demonstrated in Figure 1, lies between longitude 74°35' E to 75°10'E and latitude 13°5'N to 14°N, and falls in the Survey of India Top sheets No.48K/9, 10, 11, 12, 13, 14, 15, 16 and 48O/2, 3, & 4. It is the newly formed district during 1997 having a total geographic area of 3575 sq. km. and population 1,177,908 (2011 census) with three taluks namely Kundapur, Karkal and Udupi,



Source: www.mapsofindia.com



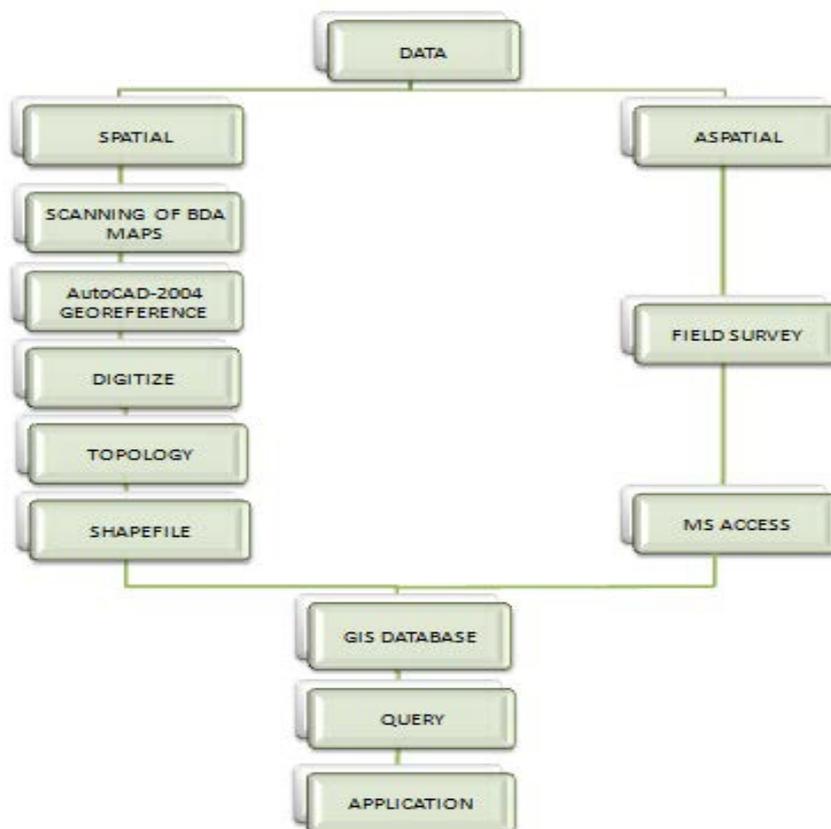
Source: www.mapsofindia.com Figure 1. Udupi District

Methodology:

The project is carried out in the following phases:

- Data Collection
- Base map preparation
- Image editing
- Input of data into individual vector layers
- Preparation of contour maps
- Preparation of intersect Map
- Query Modeling
- Extrapolation of conclusions

Figure 2: Methodology Flow Chart



Query
and

Extrapolation of Conclusions:

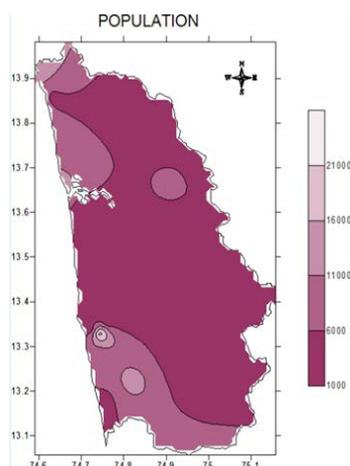
Modeling

Query modeling is done to draw conclusions about the inter-relationship of each parameter. In this step, different query equations are formed which specify certain conditions or limits to each parameter. One can then use the ARC GIS software to find the contours that satisfy the fore given conditions. This step can be used to quote questions and draw conclusions that can be further extrapolated to understand the inter-relationship between the various parameters. Extrapolation is done by studying the various results obtained, comparing with present situation and concluding to find root causes.

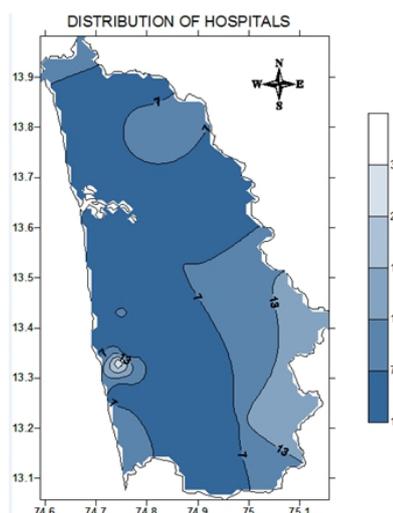
Results:

The stations selected for data collection included: Kollur, Attur, Parakala, Jadal, Kidiyoor, Manipal, Byndoor, Hebri, Brahmavar, Shirva, Shiruru, Kota, Udupi City, Karkala, Herga, Koteswara, Haladi, Paladka, Hosangadi, Someshwara, Kunjibettu, Siddapura, Shankaranarayana, Uchila, Malpe, Katpadi, Kaup, Gangoli, Kumbashi, Belur, Koni and Marvanthe. Primary and secondary data, corresponding to each of the 32 stations, of the various health, literacy and socio-economic index parameters were collected through field work. These data were then input into digitized locations on the base map as individual vector

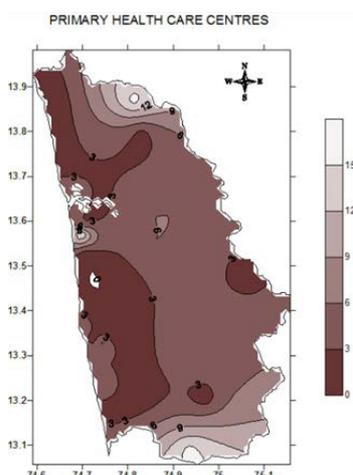
layers. This was done in a GIS environment using the ERDAS IMAGINE 9.1 software. The contour maps were then plotted (Maps 1-6) using GOLDEN SURFER 8.0 software and the following results were observed.



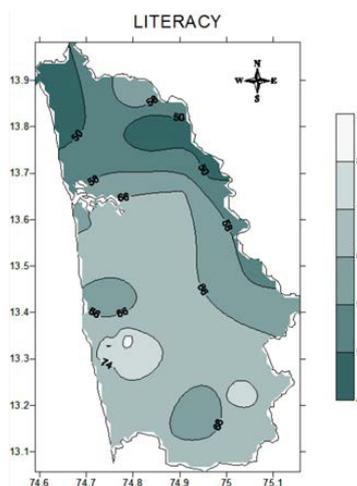
Map 1: Contour map showing distribution of population



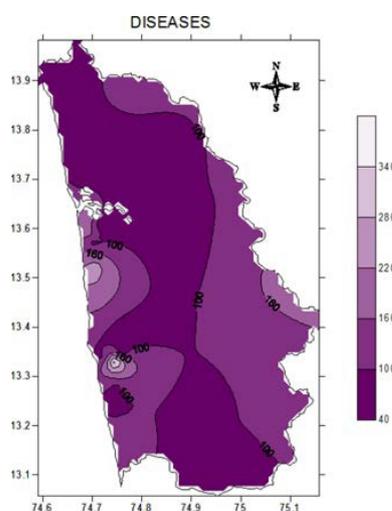
Map 2: Contour map showing distribution of hospitals and clinics



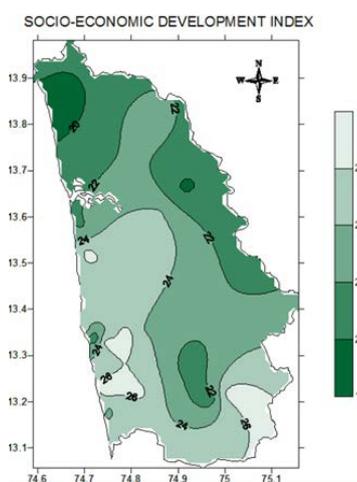
Map 3: Contour map showing distribution of primary health care centre



Map 5: Contour map showing distribution of literacy

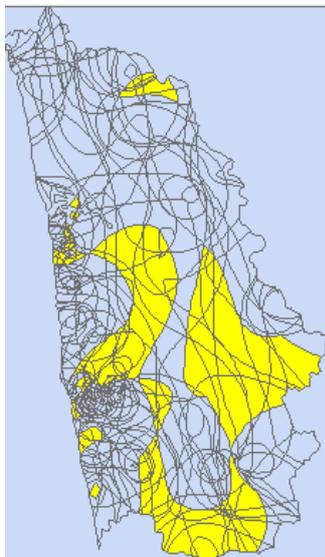


Map 4: Contour map showing distribution of diseases

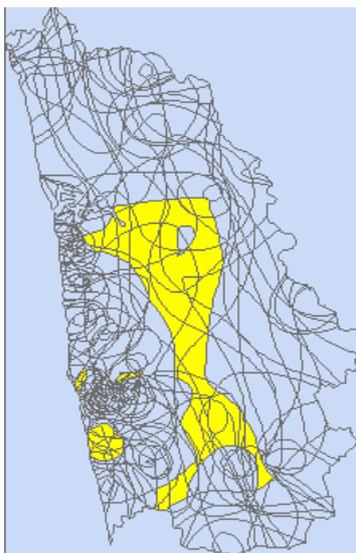


Map 6: Contour map showing distribution of socio-economic index:

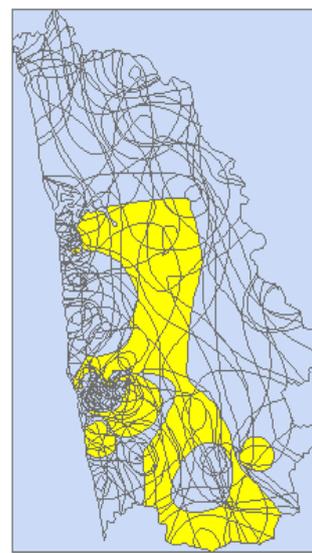
These contour maps were then intersected in ARC GIS 9 software where query modeling was carried out. Contours satisfying the conditions of query were highlighted as shown below



Map 7: Result of Query 1



Map 8: Result of Query 2



Map 9: Result of Query 3

Query 1:

"disease" = '<100' AND "literacy" = '66-74' OR "disease" = '100-160' AND "literacy" = '74-82'

Query 2

"disease" = '<100' AND "SED" = '24-26' OR "disease" = '100-160' AND "SED" = '22-24'

Query 3:

"hospitals_" = '<7' AND "literacy" = '74-82' AND "disease" = '<100' OR "phc" = '3-6' AND "literacy" = '66-74' AND "disease" = '<100'

Discussions And Conclusion:

Udupi district, as can be observed in Map 4, has a good health situation; above three quarters of the district has a low distribution of diseases. This can be explained not only by the well-knit web of hospitals, clinics and primary health care centers but also by the existing literacy and socio-economic index of the district.

However, it can be noted that certain regions of the district, mainly the eastern and coastal areas, have a higher density of diseases. At Someshwara located towards the east of the district such a behavior can be exemplified. Despite having a good distribution of hospitals and clinics the poor health status could be due to the unmistakable fall of both literacy and socio-economic index. Similarly in others areas towards the east, the lack of a fine socio-economic status and literacy could be one of the reasons that diseases are higher even when there is a good distribution of hospitals.

The coastal regions such as Kaup, Malpe, Marvanthe and Kota face health problems due to their location close to the sea. Especially during monsoon seasons and in case of a disease outbreak the coastal area is affected the most due to poor hygiene conditions. The primary health care centers located in these areas must be well equipped to heed to the health demands of such a community.

Udupi city is also observed to have a high density of diseases. This can be justified by its high population and the fact that patients, from all over the district, have a tendency to flock over to Udupi city in search of better and more advanced treatment techniques.

By comparing and carefully analyzing the results of query modeling, it can be concluded that an increased literacy rate and a fine socio-economic status can improve health conditions. Thus along with setting up a good health care system the Ministry of Health must also look into health education. Awareness must be given to the public on the importance of good hygienic conditions and a healthy lifestyle. Also, along with equipping the hospitals and clinics to meet the demands of the local public,

care must be given to equip the public with information about the most prevalent diseases and their symptoms. This way the efficiency of the health care system can be improved and health status of Udupi district can be enhanced.

References:

1. Baker DW, Williams MV, Parker RM, Gazmararian JA, Nurss J. Development of a brief test to measure functional health literacy. *Patient Educ Couns*. 1999;**38**:33–42.
2. Bennett CL, Ferreira MR, Davis TC, *et al*. Relation between literacy, race, and stage of presentation among low-income patients with prostate cancer. *J Clin Oncol*. 1998;**16**:3101–4.
3. David H. Howard PhD, Tetine Sentell PhD, Julie A. Gazmararian PhD. Impact of Health Literacy on Socioeconomic and Racial Differences in Health in an Elderly Population. DOI: 10.1111/j.1525-1497.2006.00530.x
4. DeWalt DA, Berkman ND, Sheridan SL, Lohr KN, Pignone M. Literacy and health outcomes: a systematic review of the literature. *J Gen Intern Med*. 2004;**19**:1228–39.
5. Gerard Rushton. Public Health and Spatial Analytical Tools. *Annual Review of Public Health*. Vol. 24: 43-56. DOI: 10.1146/annurev.publhealth.24.012902.140843
6. Kulldorff, Martin. Geographic Information Systems (GIS) and Community Health: Some Statistical Issues. *Journal of Public Health Management & Practice*: March 1999 - Volume 5 - Issue 2
7. Parker RM, Baker DW, Williams MV, Nurss JR. The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. *J Gen Intern Med*. 1995;**10**:537–45.
8. Sara L. McLafferty. GIS and Health Care. *Annual Review of Public Health*. Vol. 24: 25-42. DOI: 10.1146/annurev.publhealth.24.012902.141012
9. Thomas C. Ricketts. Geographical Information Systems On Public Health. *Annual Review of Public Health*. Vol. 24: 1-6. DOI: 10.1146/annurev.publhealth. 24.100901.140924

THE INFLUENCE OF REFERRAL SERVICE ON HOUSEHOLD ACCESS TO HEALTHCARE SERVICES IN ELDORET MUNICIPALITY, KENYA

Nicholas Walter Otieno Ajwang'
Eldoret, Kenya

Abstract:

The severe shortage and unfavourable distribution of health services continues to widen with observed disparities and imbalance in funding and poor distribution of human resources that continues to be limited since those in medical training have also been increasing slowly. The main objective of the study was to establish the influence of referral service on household access to healthcare services in Eldoret Municipality, Kenya. The specific objectives of the study were: to examine the influence of referral service on household access to a healthcare provider; and to examine the influence of referral services on household access to healthcare services. Results of the study will widen understanding on the referral system and access to healthcare. Survey research design was employed for the study. The unit of analysis constituted household heads in Eldoret. Stratified, systematic, purposive and convenience sampling procedures were used to arrive at a sample of 260 household heads whom were systematically interviewed concerning matters relating to access to referral care services. An interview schedule prepared especially for the study was administered to the selected household heads. Attention was paid to the factors which predispose the decision to access or not to access the available referral care services. Data was analysed using Statistical Package for Social Sciences (SPSS). On the basis of this study, a number of considerations are derived for the planning of healthcare in the country. Also a number of recommendations are given for future studies that will improve access to healthcare.

Key words: Access Africa, Eldoret, Healthcare, Household

1. Introduction

Healthcare is essential for the socio-economic development of a nation and it has been at the top of public policy agenda since Kenya's independence. Making healthcare services accessible to everyone is a duty imposed by the Kenyan constitution, yet it remains a great challenge to the existing healthcare system. Health services and policies to promote health have developed as part of modern welfare states and the form they have taken has been shaped by the social structure and institutions in different countries as well as by socio-demographic and economic factors (Allsop 1995). In Kenya, just like in other developing nations, health facilities are concentrated in the urban areas and are used mainly by higher income groups (Westendorf and Ghai 1995). Though Kenya remains, as it has always been, the best-off country in the region of East Africa (EA), it has high unequal income distributions (Ellis and Freeman 2007). The diversity and disparity of income pose a great challenge to comparable global standards for human well-being and safety. The Kenya National Bureau of Statistics (KNBS) indicates that the percentage of hard core poverty has increased in urban areas and that inequality and disparity in the country remains high (Republic of Kenya 2007).

According to Moi Teaching and Referral Hospital (MTRH), as far as financial and health sector resources are concerned, a large proportion of Kenya's population has limited access to affordable and adequate healthcare services (MTRH 2005). The Kenyan Government acknowledges in its national development plan that unfavourable distribution of health services continues to widen with observed disparities in access and affordability across the country (Republic of Kenya 2002a). It further states that disparities also exist in the distribution of medical personnel, distance to healthcare facility and expenditure allocated to rural and urban areas. In addition to this unfavourable distribution, retention of medical personnel in the public health facilities has remained a major challenge due to poor remuneration. Further, the health sector faces significant constraints due to inadequate funding and poor distribution of human resources (Republic of Kenya 2009).

The main objective of the study was to establish the influence of referral service on household access to healthcare services in Eldoret Municipality, Kenya. The specific objectives of the study were: to examine the influence of referral service on household access to a healthcare provider; and to examine the influence of referral services on household access to healthcare services. Results of the study will hopefully widen understanding on the referral system and access to healthcare.

2. Literature Review

The literature is drawn from works of various scholars and organizations of different but relevant fields of study. The literature material assembled is meant to provide the current state of knowledge on the research topic, show gaps in the area of study and to provide background information related to the research problem. At the end of the literature review, the theoretical framework and the conceptual framework used in the study are presented.

2.1 Situating the Research

Along with other countries of Eastern Africa (EA), Kenya endorsed the World Health Organization (WHO) world-wide social aim of 'Health for all by the year 2000' (WHO 1978). The WHO defines health as a 'state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity in an individual'. Kitts and Roberts (1996) observe that the WHO definition of health emphasizes the significance of the social welfare of populations and not only merely the medicalization of disease. It allows for the consideration of the complex set of social and economic components that influence the health and well-being of populations. This broad definition acknowledges the role of social structures and human economic activities to good health.

Issues in healthcare are salient not only because they touch us individually in important respects, but because it exposes many of the political, economic, social and ethical dilemmas of our time. United Nations Children's Fund -UNICEF- (1991), reports that health for all remains inaccessible due to lack of quality, unresponsiveness of services, inadequate funding and staff. In the United Kingdom, O'Reilly *et al.* (2001) in Heenan (2006) states that increasing distance from the healthcare centre reduced the likelihood of using the service. In India access to healthcare is reported to be guided by notions of efficiency rather than equity, and focus on supply, rather than demand side issues (Iyer 2005). He reports that in such a context, access to healthcare is a double edged sword. Not having it, is disempowering and amounts to denial of one's rights, but the consequences of having accessed and paid substantial sums for healthcare could also be detrimental to the economic and social well-being of the household. This cannot be farther from the truth even in Kenya.

It is acknowledged in various works that the need for treatment and cure of disease is universal in time. As a result access to healthcare services has generated considerable debate and discussion, and a number of studies have formulated frameworks or models to represent the influences on it (Andersen and Newman 1973; Penchansky and Thomas 1981; Andersen 1995; Arksey *et al.* 2003) in (Heenan 2006). Indeed, access is a shorthand term referring to the timely use of health services to achieve the best possible health outcomes (Patrick and Erickson 1993). March *et al.* (1999) reports that access to a resource refer to the opportunity available to use it. They acknowledge that these opportunities are socially constructed and constrained by structural inequalities. Healthcare service on the other hand, is broadly defined by Oleske (2001) as physician or other individual healthcare professional services, facility use, prescription use, or even the use of medical devices.

Currently, the issue of access to healthcare is critically important to households experiencing the effects of economic and health sector reform (Iyer 2005). He reports that it is now common knowledge that such reforms have benefited the rich and middle classes but not the poor. In addition, he notes that individuals may lack usual source of healthcare or may face other barriers to receiving services including financial barriers (having no health insurance or being underinsured). Scambler (2003) also found that health varies with social class. Carr (2004) observes that, poor communities typically face multiple health risks related to their location in remote areas lacking infrastructure, services and trained personnel. Further he states that health services and trained health personnel are less accessible to the poor than for the better off groups. This study takes this inquiry forward by looking at the influence of the referral system on household access to healthcare services, using data from Eldoret, Kenya. Considering the magnitude of health problems in Kenya and the global concern

for health for all, the research is not merely a curious one, but an important practical guide to understanding the referral system.

2.2 Formal Referral and Access to Healthcare

Health services are provided at different levels depending on the type of intervention the patient requires. Indeed, referral is defined as any process in which healthcare providers at lower levels of the health system, who lack the skills, the facilities, or both to manage a given clinical condition, seek the assistance of providers who are better equipped or specially trained to guide them in managing or to take over responsibility for a particular episode of a clinical condition in a patient (Al-Mazrou *et al.*, 1990). Hence individual patients need to get direct clinical services provided at lower levels of the system, before accessing upward referral which is frequently the most functional component of the health system (Hensher *et al.*, 2006), paying greatest attention to quality of care with elaborate equipment involved in diagnosis and treatment.

Early referral helps to optimize healthcare use and patient management (Wavamunno and Harris, 2005). Musgrove (2004) states that a sick or injured person can be referred “up” from a health centre or physician to a hospital and referred “down” when hospital care is no longer required. He states that there is natural hierarchy of organizations and treatment in healthcare, but there is no natural sequence like primary education followed by secondary schooling followed by university or other higher level training. In schooling, the worse results are at one level, the harder it is to proceed to the next higher one. In healthcare the exact opposite is true. The aim of this study is to determine the extent of the use that urban people are now making of available referral healthcare services available.

Referral has two faces: one, formal referral and two, self-referral. Formal referral refers to official movement to a different healthcare facility to seek further treatment, while self-referral refers to the choice of healthcare facility by the patient seeking treatment on his or her own accord. The former primarily is done within the healthcare set-up. The latter, self-referral is done by individual patients. Although self-referral has been the subject of extensive research (*ibid*), formal referral has received less attention. Studies on referrals have continuously demonstrated that it is not adhered to as it should be. Preker and Carrin (2004) found that among other factors, improving quality and increasing the referral rates would increase the utilization rates of health services. However, they report that analysis of the value of referral hospitals is bedevilled by the fact that, when judged empirically, they do not work as they are supposed to. Similarly, Kiranandana and Apairatr (1990) have also documented that people still do not follow the referral system.

The reasons for the concentration of research efforts on self-referral may be the assumption that the factors that determine formal referral are the same as those that determine self-referral. However, if the determinants of self-referral and formal referral are different, then there may be need to look more closely at existing policies and strategies in order to ensure that both types of referrals are taken into account. Moreover, if formal referral is likely to have a multiplier effect on self-referral, then the effect on overall utilization would be much greater. The study examined whether formal referral facilitated greater ability to seek specialist treatment.

In Kenya, the health sector comprises the public system, with major players including the Ministry of Health (MOH) and parastatal organizations, and the private sector, which includes private for-profit, Faith Based Organizations (FBOs), and NGO facilities (see for example, NCAPD *et al.*, 2005; Ndeti *et al.*, 2008). Further, the public health system consists of levels of health facilities: national referral hospitals, provincial general hospitals, district hospitals, health centres, and dispensaries. The public delivery system is organised in a traditional pyramidal structure (KNBS and ICF Macro, 2010). They state that the first level care is provided at dispensaries and medical clinics. The next level comprises health centres and sub-district hospitals. Third-level care is provided at district hospitals and provincial general hospitals. They further report that there are two national hospitals - Moi Referral and Teaching Hospital in Eldoret and Kenyatta National Hospital in Nairobi.

The national referral hospitals are at the apex of the healthcare system, providing sophisticated diagnostic, therapeutic, and rehabilitative services. Provincial hospitals act as referral hospitals to their district hospitals. They act as an intermediary between the national level and the districts. The district hospitals are the first referral hospitals. Health centres offer preventive and curative services. Dispensaries (lowest level) are meant to be the system’s first line of contact with patients (See Figure 1).

Institute for Policy Analysis and Research -IPAR- (2003) reports that user charges were introduced to among other reasons to improve the functioning of the referral system and rationalize utilization of health services in Kenya. But the MOH *et al.* (2004) found out that referral for complicated cases are still not working well. The referral system has also been used to determine the amount of user fees charged in healthcare outlets. In Kenya, inpatient and outpatient fees were envisaged to be higher at hospitals, lower at health centres, and almost non-existent at dispensaries to encourage the first use of lower level facilities (Republic of Kenya, 2001).

The different health systems categorize hospitals and services rendered differently discouraging access by the poor to expensive facilities. Mills (1990) reports that her analysis of available data indicated that secondary-level hospitals were typically twice as expensive per bed day as district hospitals and that tertiary hospitals were typically between twice and five times as expensive per bed day as district hospitals. Other facets is when clients are referred to another facility without any formal documentation, they risk being refused services, or having services delayed if the referral facility must assess them as totally new clients (Republic of Kenya, 2005).

The appropriate allocation of resources to referral hospitals within a national health system has long been a controversial issue in health system planning in developing countries (Hensher *et al.*, 2006). They indicate that consensus appears to be widespread that a referral hospital consume an excessive share of health budgets and that their contribution to improving health and welfare is low relative to the expenditure on these facilities. For instance, Barnum and Kutzin (1993) found Kenya to be among five countries (Belize, Indonesia, Kenya, Zambia, and Zimbabwe), that tertiary hospitals account for between 45 and 69 per cent of total public expenditure on health.

A number of studies have indicated that public hospitals in many poor countries disproportionately benefit the better off, leading their authors to argue that diverting public funds from hospitals and toward primary healthcare would be pro-poor (Castro-Leal *et al.*, 2000; Filmer *et al.*, 1997). Yet, Family Care International Kenya (2007) reports that peripheral health facilities are the most accessible, especially for the poor. In addition, the cost of care is lowest at these sites. It states that traditionally these sites have received little investment and support, and many, if not most are challenged by a crumbling physical infrastructure, shortages of skilled personnel, serious gaps in and limited referral capacity. As a result, they are often bypassed by communities in preference for hospital-level care. However, if clients are confident that they will be assisted in gaining access to higher-level facilities when needed, they may be less likely to by-pass lower-level care facilities for their health needs (NCAPD, 2005).

In developing countries the higher-level hospitals do not treat only referred patients; rather they are frequently the first point of contact with health services for many patients (Nordberg *et al.*, 1996). Perhaps the most frequent theme in research literature on referral hospitals in developing countries is the inappropriate utilization of higher-level facilities and the apparent failure of most referral systems in developing countries to function as intended (Holdsworth *et al.*, 1993). Broadly speaking, hospitals of all levels, up to and including national tertiary centres -especially in their outpatients departments- are overwhelmed by patients who could have been treated successfully at lower-level facilities, many of whom have self-referred, bypassing primary healthcare or district hospitals in the process (Sanders *et al.*, 2001). The problem of bypassing typically seems to be driven by a number of factors, including patients' perception of superior quality of care and resource availability at referral hospitals (London and Bachmann 1997); the desire to avoid delays in care if referral to a higher-level facility proves to be necessary. The current study assessed the relation between formal referral and availability of medical review services.

There is excessive and inappropriate use of referral hospitals for primary care by urban residents. The urban phenomenon of widespread bypassing and formal referral is frequently accompanied by low rates of formal referral from outlying facilities (Nordberg *et al.*, 1996). These problems have a number of negative impacts and consequences. Simple conditions are unnecessarily treated in a high-cost environment; outpatient departments are congested by patients requiring primary care, thus causing long waiting times; scarce staff time is diverted from specialized areas and into inappropriate care; and more complex cases requiring specialized care are crowded out by more urgent but less technically demanding cases that could be cared for at lower levels (Stefanini, 1994).

The World Bank (1994), reports of the phenomenon of the ill travelling past a free or subsidized public clinic (or other public facility) to get to an alternative source of care at which they pay a considerable amount for healthcare. They observe that when it is a poor person choosing to bypass a free public facility and pay for care further away, such an action is especially bothersome. It states that bypassing behaviour is not very different across income groups, as is the fact that the more severely ill tend to bypass and travel further for care than do the less severely ill. To them prices tend to deter use, and improved quality of services to increase the likelihood of a facility being chosen. Such care-seeking patterns result in higher costs to households, as well as to the health system as a whole, in that resources are not used efficiently. There is therefore a need to gain a more detailed picture of the relation between a formal referral and the preferred healthcare facility.

Hensher *et al.*, (2006), reports that particular problems may also arise in the referral system where the same doctors provide care in both public and private hospitals. Under fee-for-service arrangements, physicians may focus on their more lucrative private patients to the disadvantage of public hospital patients, refer patients with adequate insurance to their private practices and private hospitals, and transfer patients with expensive diseases or inadequate insurance to public hospitals. Patients need satisfaction and convenience in accessing healthcare. The study sought to identify the services obtained as a result of referral cases.

Studies on the accessibility of referral hospital care have repeatedly confirmed the existence of a steep distance-decay function, in countries such as Ethiopia (Kloos, 1990) and Nigeria (Lyun, 1983), indicating that individuals with a given need for a clinical service will be less likely to access that service the farther away from the referral centre they live. In Kenya, (Mbeya, 1997) found that the location of healthcare facilities far away from the population affects the level of use of health services. He found that the more accessible one is to a health facility the more one is able to make use of it. He found that most of the respondents do not visit far away facilities. Further, the respondents still frequented nearer facilities despite poor services offered. He credits this to their closeness. In addition the far away facility, is visited only when the closer facility lacks medicine or equipment necessary for treatment. The study went further and examined the influence of the referral system on household access to care.

2.3 Theoretical Framework

The study used the Health Belief Model (HBM) to test the field data. It should be noted that the term model and theory are being used interchangeably in this section of the chapter for the sake of convenience.

2.3.1 Health Belief Model

A model can be seen as a 'theory' or a set of 'hypotheses' which attempt to explain the connections and interrelationships between components of the study phenomena (Gilbert, 1993). It is a representation of the significant features of the problem under study (Tones and Tilford, 2001). Health Belief Model (HBM) was originally developed to explain why people failed to utilize health services (Hochbaum, 1958; Rosentock, 1966). It has since undergone various revisions (Becker, 1974; Janz and Becker, 1984).

The HBM centers on three beliefs that account for the variance in predispositions to adopt a recommended health practice (Becker, 1974). These include a belief that individuals will not adopt health behaviours, unless they believe they are susceptible to a disease or disorder (belief that you could have the disease and not know it in the case of undertaking screening or treatment for conditions such as hypertension), and they believe it is serious (a belief in the severity of the consequence of not taking action). The third belief influencing action is that the benefits of treatment or intervention will outweigh the costs (including social benefits and costs such as inconvenience, discomfort, or embarrassment). That is, they accept that the recommended preventive actions will be effective and that the benefits accruing from their actions will outweigh any costs or disadvantages that they believe that they will incur as a result.

According to Cockerham (1986) individuals will seek healthcare when they believe they are susceptible to disease, consider the disease to be serious and accept that benefits of their action outweigh the costs. This is incorporated as a motivational element, in that, beliefs about susceptibility and seriousness are considered to generate a level of 'perceived threat' which in turn contributes ultimately to health choices-together with beliefs about cost and benefit. The assumption of this

model is that by taking a particular action, susceptibility would be reduced, or severity would be reduced. Thus the HBM components of perceived susceptibility, perceived severity and perceived benefits of taking health action of interest aided this present study.

2.4 Conceptual Framework

The study adopted and modified the Health Delivery System in Kenya framework by Ndeti *et al.* (2008). The lowest level of care is the dispensary and the highest level of care is the national referral hospital. There are private care outlets that do offer the same level of care as the public care outlets.

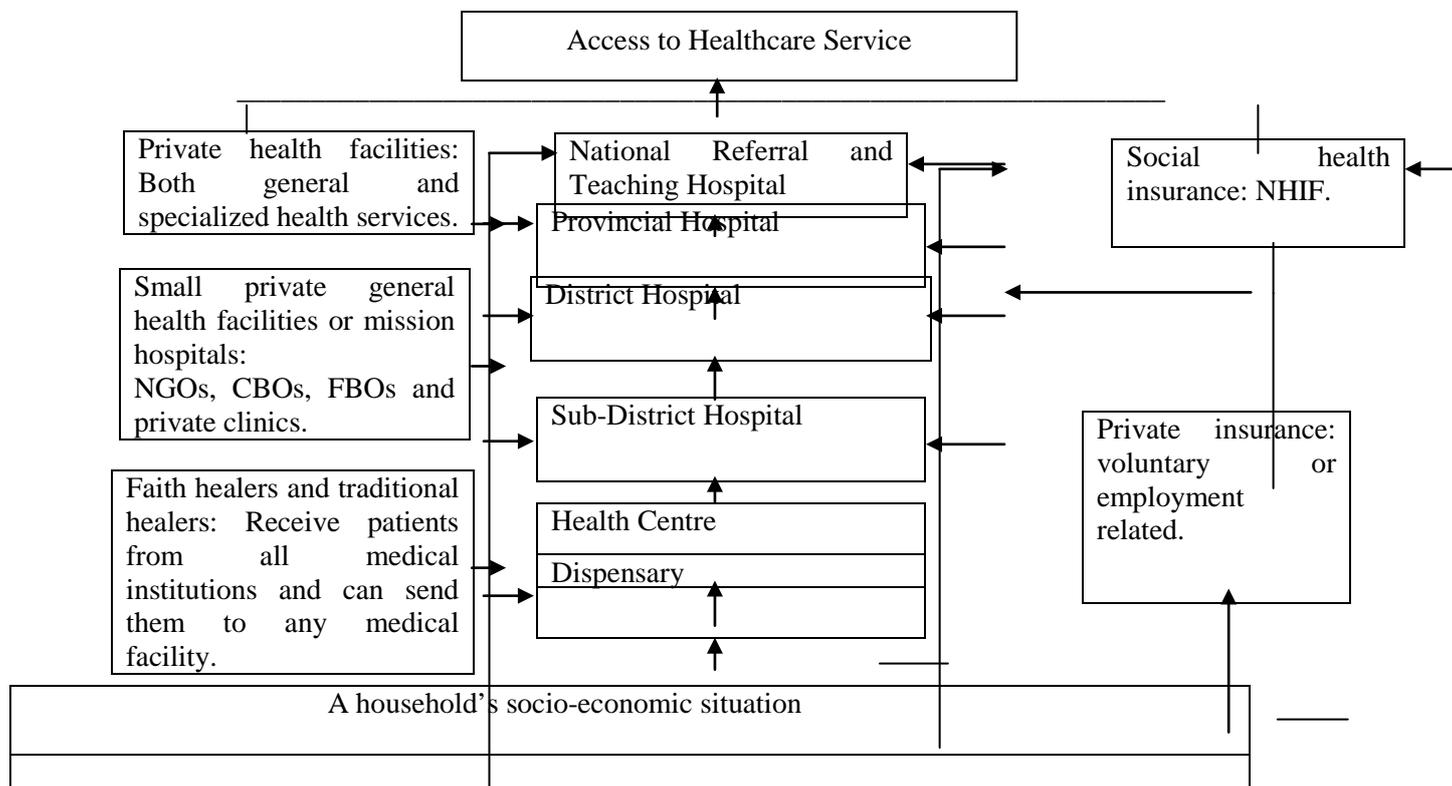


Figure 1. Conceptual Framework (Adopted and Modified from Health Delivery System in Kenya, Ndeti *et al.* 2008).

3. Methodology

3.1 The study area

The Republic of Kenya is located on the eastern coast of Sub-Saharan Africa (SSA). According to KBNS and ICF Macro, the country lies between 5 degrees north and 5 degrees south latitude and between 24 and 31 degrees east longitude (KNBS and ICF Macro, 2010). It is almost bisected by the equator. The country is bordered by Ethiopia (north), Somalia (northeast), Tanzania (south), Uganda and Lake Victoria (west), and Sudan (northwest). It is bordered on the east by the Indian Ocean. It has a total area of 582,646 square kilometres of which 571,466 square kilometers form the land area. Approximately 80 percent of the land area of the country is arid or semiarid, and only 20 percent is arable.

The study area is Eldoret Municipality, the capital and administrative centre of Uasin Gishu County, Rift Valley Province, Kenya. The map is presented in fig. 2. The town has undergone rapid economic development since 1974 and this has caused rapid population growth and physical expansion (Nyakaana 1996; Republic of Kenya 2002b). The urban population is indicated as 144,223 in Eldoret West, 80,729 in Eldoret East and 120,607 in Wareng districts all of Uasin Gishu County according to the 2009 Kenya population and housing census (Republic of Kenya, 2010). Further, the number of urban households is recorded as 88,956 in Eldoret West, 51,486 in Eldoret East and 61,866 in Wareng districts.

The towns' rapid urbanization has led to a lure for jobs, amenities and stimulation. But despite this impressive urban growth, economic gains have not benefited everyone as real incomes of people have decreased substantially and disparities between the rich and the poor have increased. Its health facilities specifically the public ones, are over utilized (Republic of Kenya 1997). The area is among the country's most 'land-rich' areas and agriculture (crop and dairy farming) is the mainstay of the economy and the vast majority of the population is composed of agricultural producers. Commerce, tourism, tertiary services and industry also form part of the economy of the county. The industries located in the county are mainly agro-based (Republic of Kenya, 2005). They include textiles, wheat, pyrethrum, milk and corn. Commercial agriculture tends to be highly vulnerable to fluctuations of the world market. Further, many people are unable to afford the private hospital charges. Hence, they resort to public facilities.

The health sector is characterised by high under five mortality rates, and its health facilities are located far apart (Republic of Kenya 2008a; Republic of Kenya 2008b). Like other urban centres in the country, it is better served with both Government and private hospitals and health centres. The municipality was selected for study due to its easy access, its relatively static population and numerous healthcare facilities (Republic of Kenya 2005).

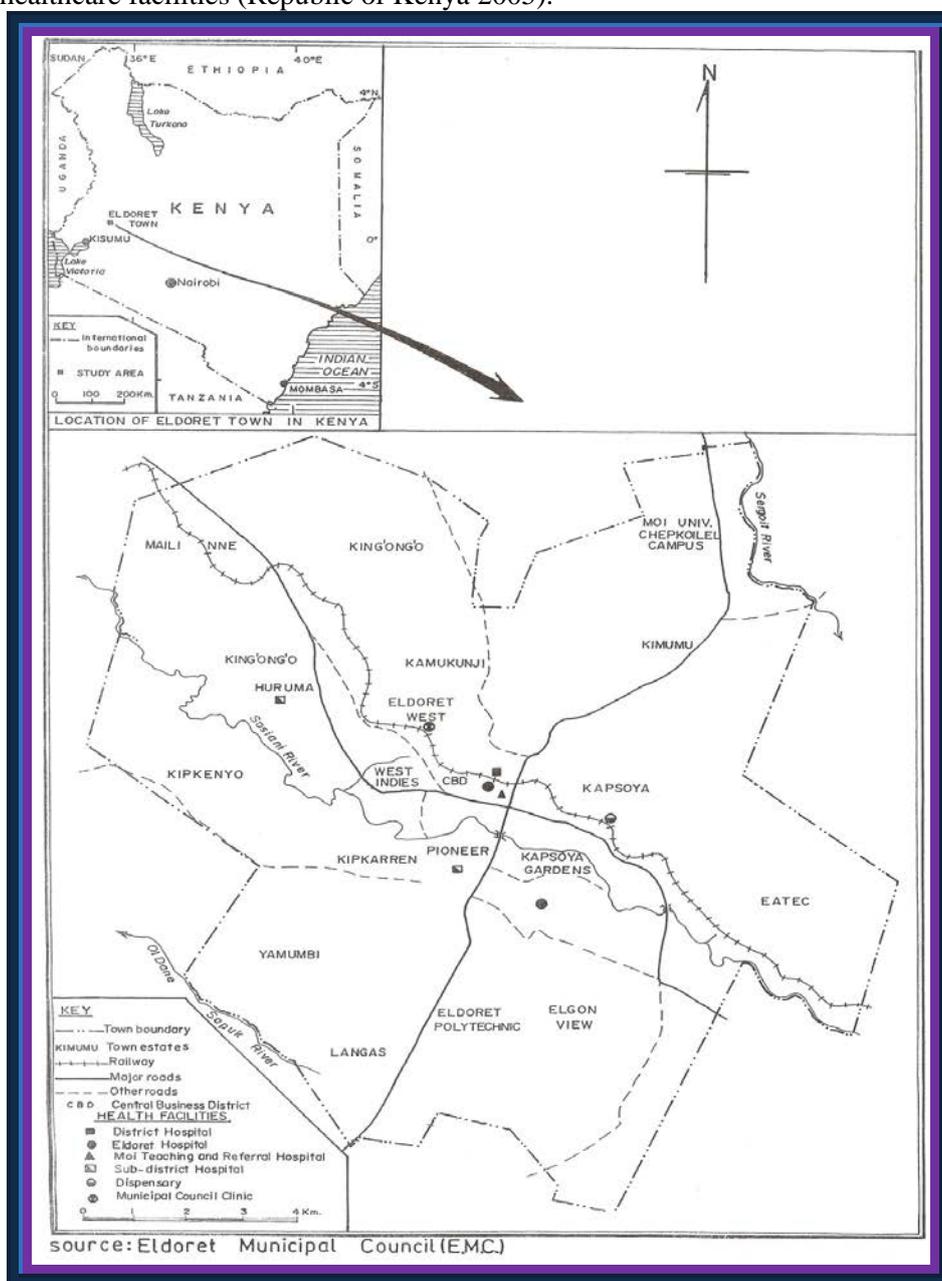


Fig. 2. Map of Eldoret Municipality (From a drawing by Eldoret Municipal Council E.M.C.).

3.2 Research design

The study employed survey research design. Data for the study was derived from primary source through quantitative method because it seeks measurement as the basis for forming generalizations concerning social reality and employs survey method which is concerned with establishing statistical associations between variables (Ridsdale 1998; Mwanje 2001). Primary data was created using a pre-tested interview-administered schedule. It was both closed-ended (structured, fixed response) for quick and easy response; and open-ended (unstructured, free response) to help learn how a respondent thinks, to discover what is really important to him or her, or to get an answer to a question with many possible or sensitive answers. The study strived to identify the various heads of households within Eldoret municipality among the urban population as the unit of analysis. A household represents the most important unit in terms of time and resources invested to attain health. In the study, the research questions for the quantitative method were predefined, so that the study tested a precise set of research questions.

3.3 Sampling Procedure

In order to obtain representative samples of households belonging to different groups, distribution across geographical location, as well as economic classes, a four-stage sampling procedure was adopted. The first stage involved obtaining representative samples of households by stratification. This divided the population into mutually exclusive sets or strata to ensure that all the classes/cadres of estates were adequately represented and no bias prevailed. The municipal residential type is low, medium and upper income. An equal number of three estates were drawn from each stratum. The estates are namely, Huruma, Kamukunji, and Langas, (low income), Pioneer, Kapsoya, and Kimumu (middle income), Elgon View, West Indies and Garden Estate (upper income).

The second step of selection involved the systematic sampling procedure. In this case, units for the study were selected directly. This procedure was used to select twenty heads of households (male or female) from each of the estates at a constant interval of ten. Using such a sampling procedure, every tenth household head was selected in a circular systematic fashion, with equal probability, after a random start. The sampling strategy allowed for substitution of the households in the field, if they were not available despite repeated attempts to contact them, or if they refused to participate in the survey. The procedure resulted in a sample size of 180 household heads from the nine estates.

In the third stage, healthcare facilities were purposely selected to obtain two Government hospitals, two private hospitals, two health centres, and two dispensaries. Purposive sampling also known as deliberate or judgmental sampling of the healthcare facilities enabled selection of all categories and cadres of healthcare service outlets located in the study area. Lastly, based on the ease of access, convenience sampling also called haphazard or accidental sampling was applied to obtain ten household heads seeking healthcare service in each facility for an in-depth interview that included patient exit interviews. This procedure generated an addition of 80 household heads. Therefore, the total sample for the study was 260 household heads.

3.4 Strengths, Limitations and Ethical Considerations of the Research

The study was faced with certain limitations but efforts were made to ensure that they were adequately addressed so as to limit their impact on the findings. The study was restricted to residents of Eldoret, Kenya. This excluded the population that commutes to the town every day for work, business and healthcare. Respondents were reluctant and suspicious at the beginning but after explaining the study objectives they were receptive and provided the information. In addition, the study was limited by the self-reporting accuracy of the participants.

Ethical issues in social science research were considered including the importance of voluntary and informed participation and the preservation of the participants' anonymity. The work aimed at projecting an ethos that encouraged trust; hence permission to carry out the study was sought from Eldoret Municipality's Medical Officer of Health (MoH), Uasin Gishu County MoH and the Uasin Gishu County Hospital Medical Superintendent. Last but not least, debriefing of participants in the research will take place at the conclusion of the study, and it will involve revealing the purposes of the research in both written and oral form. The written form will leave participants in the research experience with a tangible description of the activities they performed. An oral debriefing will be done to simplify the research findings.

4. Data Presentation and Analysis

4.1 Results

Table 1 lists the profiles of households that were slotted into this study. Household level data were merged with healthcare facility-level data. While the data collected from the households in the residential areas had twenty heads from each sampled estate, data collected from the sampled healthcare facilities contained varied number of individual respondents from the sampled estates of study. As is presented in Table 1, the survey contains information on 260 interviewed households, and data set on 1,349 persons from nine estates within Eldoret. Male household heads accounted for 79% and females 21% of all the surveyed households. The males mostly headed married families. However, due to changes in gender relations and family patterns that could be said to jeopardize the institutional underpinnings of the traditional male breadwinner family, females headed separated, divorced, widowed families, as well as families where males had migrated for various reasons including migrant labour.

From the survey, it is revealed that female adults (29.5%) were the majority followed by male adults (27.4%). In terms of age, 88.4% of the respondents were aged below 55 years while 11.6% were above 55 years. Christians were a majority contributing to 84.2% to the overall distribution of the survey. Muslims were 7%, Hindu 8%, and other religions were a minority, representing 1% of the surveyed households. Respondents from the up-market residential areas were 29%, from middle income estates 32% and those from the low income estates 39%. In terms of education majority of the respondents (87.3%) had at least attained secondary school education. The respondents with no formal education represented 2.3% of the total distribution.

Table 2 presents the information about chronic illness, illness prevalence and treatment in the household during a four week recall period preceding the survey. As is presented, the vast majority of the respondents (81.2%) had no incidences of chronic illness, while 18.8% did state presence of chronic illness in the household. A vast majority (66.5%) reported incidences of illness four weeks prior to the survey while 33.5% had no incidence of illness in the same period. A higher proportion of females in the households (82.2%) reported illness than their male counterparts. This is higher than the 51.5% distribution of females in the overall sample. Out of the households who had been ill during the four weeks preceding the survey, 62.3% had sought healthcare outside the home. Most of the households had experienced only one treatment episode (73.5%), while 25.4%, had two and above episodes of treatment four weeks prior to the survey.

Table 1 Socio-Demographic Data

Characteristics	Category	<i>f</i>	%
Household Head	Yes	260	100.0
	No	-	-
Gender of head	Male	205	78.8
	Female	55	21.2
Age (years)	18-39	160	61.5
	40-54	70	26.9
	55-70	29	11.2
	Over 70	1	0.4
Total household population size	Female Adults	398	29.5
	Male Adults	369	27.4
	Female children	297	22.0
	Male children	285	21.1
Marital Status	Single/ (Separated/ Divorced/ Widowed)	92	35.4
	Married	168	64.7

Education level	None	6	2.3
	Primary	27	10.4
	Secondary	74	28.5
	Tertiary	153	58.8
Religion	Christian	219	84.2
	Muslim	17	6.5
	Hindu	21	8.1
	Other	3	1.2
Residence	Upper income	74	28.5
	Middle income	84	32.4
	Low income	102	39.2
Household Population Size	1-5	161	61.9
	6 and above	99	38.1

SOURCE: Field Data (2012).

Table 2 Households' Illness and Treatment Characteristics

Characteristics	Category	<i>f</i>	%
Number of sick members with chronic illness	None	211	81.2
	1	45	17.3
	2	4	1.5
Presence of illness in last one month	Yes	173	66.5
	No	87	33.5
Sought treatment	At a healthcare outlet	108	62.3
	At home	65	37.7
Gender of the invalid	Male	31	17.8
	Female	142	82.2
Treatment episodes in last one month	None	2	1.2
	1	119	73.5
	2 >	41	25.4

SOURCE: Field Data (2012).

4.2 Incidence of Referral and Access to Healthcare Provider

The results in terms of the determinants of formal referral on access to healthcare are varied. Table 3 presents the basic information on referral and access to healthcare that was found in the household survey. The most important variable to be looked at in the context of the research question is formal referral and its effect on access to healthcare services. As shown in table 3, referral awareness is close to 89% over the entire sample with only about 11% reporting that they are not aware of the existence of the referral system. However, the use of the referral system is lower in Eldoret. There are variations on reasons for seeking healthcare, as only about 26.5% of healthcare cases are as a result of being referred and 73.5% are as a result of other reasons including a first contact with a service provider. However, the results show that a greater proportion of the referral cases were formal referrals with written documentation representing 73.9% of the referrals, while 26.1% indicated that they had self-referred.

With Referral cases originating from the public healthcare outlets totaling about 33% of the overall share of the referrals, the bulk of the referral cases were from the private healthcare providers

with about 67% of the referred cases coming from private healthcare outlets. It is important to note that no respondent indicated to have been referred from the national referral hospital. It might be that the facility does not refer patients down to lower cadre facilities as should be the case in a proper referral system which can be either up or down the referral system. It should be noted that most of the referrals ended up in the public healthcare outlets representing 79.7%, with private healthcare facilities receiving 20.2% in the distribution (see Table 3) of all the referred cases. The first referral level facility which is the district hospital received 20.6% of the overall referral cases in the distribution.

Table 3. Seeking a Provider as a Function of Formal Referral

Characteristics	<i>f</i>	%
Aware of existence of referral		
Yes	231	88.8
No	29	11.2
Reasons for seeking care		
Referral	69	26.5
Other	191	73.5
Referral used		
Formal referral	68	73.9
Self-referral	24	26.1
Referred from		
Public healthcare outlet	23	33.0
Private healthcare outlet	46	66.7
Referred to		
Public healthcare outlet	55	79.7
Private healthcare outlet	14	20.2
Facility preferred		
Private healthcare outlet	51	73.9
Public healthcare outlet	17	24.6
Self-treatment	1	1.4
Neighbourhood of referred case		
High Class	41	59.4
Middle class	14	20.2
Low class	14	20.2

Source: Field Data (2012).

As a matter of interest the study went further and looked at the direction of the referrals and the category (public or private) of healthcare accessed. As presented in Table 4, majority of the referrals were from the private sector to public sector at 79.6%. Referrals between private to private was 17.4%, public to public 26% and public to private 2.9%.

The overall chi-squared test for the data gives a *p*-value of 0.000 which is less than the set criterion of 0.05. It is concluded that the two variables are not independent or, put another way, there is a statistically significant difference in the proportions. This is an illustration of how statistically significant with a very strong association as indicated in Cramer's *v* of 0.737.

Table 4 Healthcare Facility Referred to

Category of Care	<i>f</i>	%
Private – Public	37	53.6
Private – Private	12	17.4

Public – Private	2	2.9
Public – Public	18	26.0
<i>Total</i>	69	100.0

$\chi^2 = 1.500$, $df = 24$, $phi = 1.475$, *Cramer's V* = 0.737, $p < 0.05$ ($p = 0.000$)

Source: *Field Data (2012)*.

4.3 Services Accessed from Referral

Table 5 below presents the various reasons why households are formally referred from one healthcare outlet to the other. The reasons given for the referrals were to seek specialist treatment, better equipment, drugs and quality care. Majority of the formal referrals are for specialist treatment at 58% followed by need for equipment at 23.2%. Drugs and quality come at 11.6% and 7.2% respectively. Chi-square test conducted for the cross-tabulation indicates that specialist treatment, drug, equipment and quality are not statistically significant. These show that there is no relationship between access to these services and formal referral.

Table 5 Services Obtained from the Formal Referral

Referral Services	<i>f</i>	%
Specialist treatment	40	58.0
Drug/medicine	8	11.6
Equipment	16	23.2
Quality	5	7.2
<i>Total</i>	69	100.0

$\chi^2 = 7.856$, $df = 12$, $phi = 0.337$, *Cramer's V* = 0.195, $p > 0.05$ ($p = 0.796$)

Source: *Field Data (2012)*.

As presented in Table 6 below, most of the referral cases ended up in the hands of a medical doctor at 68.1%. This was followed by the clinical officer at 17.4% and pharmacist at 5.8%. The nurse was consulted by 4.3% of the overall respondents in the distribution. Other medical specialist (physio-therapist, medical social workers etc) also recorded 4.3% of the respondents consulting them. Since the chi-squared test is less than 0.05 we can conclude that the variables are not independent. Put another way, there is a statistically significant difference in the proportions as $p = 0.023$. From the *phi* values it can be concluded that the relationship is positive as well.

Table 6 Specialist Consulted Based on Formal Referral

Medical Specialist	<i>f</i>	%
Doctor	47	68.1
Clinical Officer	12	17.4
Nurse	3	4.3

Pharmacist	4	5.8
Other	3	4.3
<i>Total</i>	69	100

$\chi^2 = 29.217, df = 16, phi = 0.651, Cramer's V = 0.325, p < 0.05 (p = 0.023)$

Source: Field Data (2012).

As is presented in table 7 majority of the respondents who had gone through the formal referral indicated that they had been booked for review or further consultation as a result of the referral. About 94% had been booked for review while about 6% had not been booked for review. The chi-squared value for the test item was higher than 0.05 ($p = 0.428$) indicating there is no statistical relationship between the items.

Table 7 Booked for Review from the Formal Referral

Booked for review	<i>f</i>	%
Yes	65	94.2
No	4	5.8
<i>Total</i>	69	100.0

$\chi^2 = 3.838, df = 4, phi = 0.236, Cramer's V = 0.236, p > 0.05 (p = 0.428)$

Source: Field Data (2012).

5. Discussion

5.1 Formal Referral Determinant

Results of the study showed that in general, usage of referral was very low in Eldoret municipality as previously documented elsewhere in the municipality (Nyakaana, 1996) and beyond (Kiranandana and Apairatr, 1990). Nyakaana's findings that there are fewer number of residents using the national referral hospital confirms the study. He notes that though the demand for medical facilities is on the increase in the municipality, the increase is especially cases from outside the municipality, which has outnumbered those from within the council. This characteristic can be seen in light of Hensher *et al.* (2006) view that referral hospitals consume an excessive share of health budgets and that their contribution to improving health and welfare is low relative to the expenditure on them. This can be one of the reasons why healthcare is increasingly becoming inaccessible to households (Maliyamkono and Ogbu 1999).

Only about 27% of the respondents in the overall distribution received referral healthcare in the four weeks preceding the survey. Such levels of referrals are considered low. All indicators are that most of the cases in healthcare facilities are not as a result of referrals. This evidence supports the conclusion made by Atkinson *et al.* (1999) that the smaller number of referrals may be traced to referral facilities acting as first contact provider for most households in developing countries. The low participation by households in the referral system strongly suggests that the level of adherence to the system is below the potentially attainable. London and Bachmann (1997) attributed this phenomenon to patients' perception of superior quality of healthcare and resource availability at referral hospitals, which often may be entirely well founded and rational; the desire to avoid delays in healthcare if referral to a higher-level facility proves to be necessary; and the fact that for many urban populations a referral hospital may simply be the closest health facility.

The findings indicate that the proportion of households reporting awareness of the referral services is enormous at 88.8%. The more common reason for this phenomenon is that there are varied categories of healthcare outlets in the municipality ranging from the lowest, the dispensary, to the highest, the national hospital. However, of the respondents surveyed only 26.5% had accessed

healthcare as a result of referred cases. The results also indicate that a large proportion of the referral cases are directed to the public healthcare outlets (79.7%) compared to the private outlets (20.2%). However, majority of the referrals originated from the private healthcare outlets (66.7%) compared to 33.0% from the public outlets.

Formal referral will in most cases send a patient to a public healthcare outlet. This was found to be statistically significant. Additional insights can be obtained by examining access to healthcare disaggregated by the type of specialist sought. The results suggests that for the referred households 68.1% of them accessed a doctor; either a general practitioner or a medical specialist, while 17.4% accessed clinical officers and 14.4% accessed the other lower qualified medical specialists. Analysis of the surveyed data reveals that while most of the referrals are directed to public healthcare outlets at 79.7%, this is more than the percentage of patients who prefer the facilities who are 24.6%. This finding was as well found to be positively significant. The households that preferred private healthcare facilities were 73.9%. However, about 20% were referred to these private facilities for healthcare services. It can also be reported that the use of services in a healthcare facility is not tied to the formal referral as the test statistic was found to be greater than the set criterion of 0.05.

In general, the potential of referral in accessing healthcare services in the lives of Eldoret municipal people is evident in the high incidence reported of awareness of one or more type of referral facilities by household members. Overall, awareness of the existence of referral system is 88.8% among the respondents. However, the notion of referral awareness do not necessarily entail sustained adherence to the system as indicated above. But the results of the study from the cases of referrals in the municipality, reinforces the notion that referrals increases access to healthcare. This calls for the strengthening of referral capacity in the various Government healthcare organs vested with the responsibility of providing healthcare.

Formal referral is an important entitlement that enables access to healthcare services. This observation in itself implies that for healthcare access programmes to achieve their intended purpose, they have to be designed and implemented in a manner that takes into account the referral system. About 74% of the referral cases in the survey are as a result of formal referral, meaning that cases of formal referral are prevalent compared to self-referral in the municipality. This finding is unlike that of Nordberg *et al.* (1996) and Omaha *et al.* (1998) who report that the urban phenomenon of widespread bypassing and self-referral is frequently accompanied by low rates of formal referral from outlying facilities.

Findings confirm that formal referral in Eldoret contributes to access to healthcare for the concerned households. Critical to the question of access to healthcare services is the one of access to better facility outlet and medical specialist. Formal referral has significantly improved accessibility to healthcare for households by increasing probability of visit to a higher cadre health facility. The study emphasises and shows trends toward higher rates of access to specialist care (medical specialist). Access to a medical specialist as opposed to a general physician increases with formal referral. About half of all referred persons had accessed specialist treatment in one facility or the other. Formal referral enhances proper use of specialists as well. Overuse of specialists care has often been blamed for the explosion in healthcare costs when too many patients go to specialists for common conditions that need primary healthcare physicians (Gatsonis *et al.*, 2005). Inevitably, the individuals with formal referral also have greater and better chance of being assessed faster. This has been indicated by NCAPD (2005) as one of the merits of formal referral. It reports that when clients are referred to another facility without any formal documentation, they risk being refused services, or having services delayed if the referral facility must assess them as totally new clients.

Survey findings also allude to the fact that the formal referral is closely associated with accessing the higher cadre care. Though most of the referral cases are from the private sector, the private healthcare outlets also do receive cases of referral. Referral cases from the public facilities were more and they referred cases among themselves. There is some association between formal referral and access to preferred healthcare as well. Formal referrals help to strengthen the link between the healthcare facility preference (the healthcare facilities where people would like to seek healthcare) and where they actually do seek healthcare. About 74% of the respondents preferred private healthcare facilities though only 20.2% were referred to the private healthcare outlets. In addition to improved access to private healthcare there is also increased satisfaction with the services received in the healthcare outlets they were referred to.

Respondents in the survey tended to relate formal referral with quality of healthcare. The fact that formal referral is a determinant of household access to healthcare, as a result of perceived quality of healthcare is in line with health belief models (Hochbaum, 1958; Rosentock, 1966). One of the key descriptors of health belief model is the perceived benefit of taking action; taking action toward the prevention of disease or toward dealing with an illness is the next step to accept after the individual has accepted that they are susceptible to a disease and recognised it is as serious. During the initial contact of seeking healthcare, the household chooses a facility because the associated benefits are perceived to be higher than those of alternative outlets. However, with formal referral contacts, the quality of healthcare previously received is used as a learning opportunity and the household will only return to the same provider if there is a perceived benefit.

The lack of relation between higher existence of chronic illness in a household and formal referral suggest that most of the referred cases are not as a result of chronic illness after all. This finding is surprising in view that healthcare providers at lower levels of the health system, who lack the skills, the facilities, or both to manage a given clinical condition, are supposed to seek the assistance of providers who are better equipped or specially trained to guide them in managing or to take over responsibility for a particular episode of a clinical condition in a patient (Al-Mazrou *et al.*, 1990). This kind of skills and facilities are normally located in referral outlets. Finally, though the familiarity of the respondents with referral system does not translate to high usage, the referral system is meant to improve access to healthcare among the populace and to protect the poor from the cost of medication by accessing the lower cadre facilities before moving to the upper cadre ones (see related works in Bedi *et al.*, 2004; and Owino and Were, 1997). It is evident from the study findings that strengthening of the referral system would help increase access to healthcare services for the households.

6. Conclusion and Recommendation

6.2.1 Conclusion

The data used for this particular analysis can be considered unique in terms of content, geographic coverage, and timeliness on households' access to healthcare services. The data collected was in appropriate form, in desired detail and accuracy. The research question had asked to what extent formal referral influence household access to healthcare services. Results of the study showed that in general, usage of referral was very low in Eldoret municipality as previously documented elsewhere in the municipality (Nyakaana, 1996) and beyond (Kiranandana and Apairatr, 1990). Nyakaana found that the fewer number of residents using the national referral hospital confirms the findings. He notes that though the demand for medical facilities is on the increase in the municipality, the increase is especially cases from outside the municipality, which has outnumbered those from within the council. This characteristic can be seen in light of Hensher *et al.* (2006) view that referral hospitals consume an excessive share of health budgets and that their contribution to improving health and welfare is low relative to the expenditure on them.

Only about 27% of the respondents in the overall distribution received referral healthcare in the four weeks preceding the survey. Such levels of referrals are considered low. All indicators are that most of the cases in healthcare facilities are not as a result of referrals. This evidence supports the conclusion made by Atkinson *et al.* (1999) that the smaller number of referrals may be traced to referral facilities acting as first contact provider for most households in developing countries. The low participation by households in the referral system strongly suggests that the level of adherence to the system is below the potentially attainable. London and Bachmann (1997) attributed this phenomenon to patients' perception of superior quality of healthcare and resource availability at referral hospitals, which often may be entirely well founded and rational; the desire to avoid delays in healthcare if referral to a higher-level facility proves to be necessary; and the fact that for many urban populations a referral hospital may simply be the closest health facility.

The findings indicate that the proportion of households reporting awareness of the referral services is enormous at 88.8%. The more common reason for this phenomenon is that there are varied categories of healthcare outlets in the municipality ranging from the lowest, the dispensary, to the highest, the national hospital. However, of the respondents surveyed only 26.5% had accessed healthcare as a result of referred cases. The results also indicate that a large proportion of the referral cases are directed to the public healthcare outlets (79.7%) compared to the private outlets (20.2%).

However, majority of the referrals originated from the private healthcare outlets (66.7%) compared to 33.0% from the public outlets.

Formal referral will in most cases send a patient to a public healthcare outlet. This was found to be statistically significant. Additional insights can be obtained by examining access to healthcare disaggregated by the type of specialist sought. The results suggests that for the referred households 68.1% of them accessed a doctor; either a general practitioner or a medical specialist, while 17.4% accessed clinical officers and 14.4% accessed the other lower qualified medical specialists. Analysis of the surveyed data reveals that while most of the referrals are directed to public healthcare outlets at 79.7%, this is more than the percentage of patients who prefer the facilities who are 24.6%. This finding was as well found to be positively significant. The households that preferred private healthcare facilities were 73.9%. However, about 20% were referred to these private facilities for healthcare services. It can also be reported that the use of services in a healthcare facility is not tied to the formal referral as the test statistic was found to be greater than the set criterion of 0.05.

In general, the potential of referral in accessing healthcare services in the lives of Eldoret municipal people is evident in the high incidence reported of awareness of one or more type of referral facilities by household members. Overall, awareness of the existence of referral system is 88.8% among the respondents. However, the notion of referral awareness do not necessarily entail sustained adherence to the system as indicated above. But the results of the study from the cases of referrals in the municipality, reinforces the notion that referrals increases access to healthcare. This calls for the strengthening of referral capacity in the various Government healthcare organs vested with the responsibility of providing healthcare.

Formal referral is an important entitlement that enables access to healthcare services. This observation in itself implies that for healthcare access programmes to achieve their intended purpose, they have to be designed and implemented in a manner that takes into account the referral system. About 74% of the referral cases in the survey are as a result of formal referral, meaning that cases of formal referral are prevalent compared to self-referral in the municipality. This finding is unlike that of Nordberg *et al.* (1996) and Omaha *et al.* (1998) who report that the urban phenomenon of widespread bypassing and self-referral is frequently accompanied by low rates of formal referral from outlying facilities.

Findings confirm that formal referral in Eldoret contributes to access to healthcare for the concerned households. Critical to the question of access to healthcare services is the one of access to better facility outlet and medical specialist. Formal referral has significantly improved accessibility to healthcare for households by increasing probability of visit to a higher cadre health facility. The study emphasises and shows trends toward higher rates of access to specialist care (medical specialist). Access to a medical specialist as opposed to a general physician increases with formal referral. About half of all referred persons had accessed specialist treatment in one facility or the other. Formal referral enhances proper use of specialists as well. Overuse of specialists care has often been blamed for the explosion in healthcare costs when too many patients go to specialists for common conditions that need primary healthcare physicians (Gatsonis *et al.*, 2005). Inevitably, the individuals with formal referral also have greater and better chance of being assessed faster. This has been indicated by NCAPD (2005) as one of the merits of formal referral. It reports that when clients are referred to another facility without any formal documentation, they risk being refused services, or having services delayed if the referral facility must assess them as totally new clients.

Survey findings also allude to the fact that the formal referral is closely associated with accessing the higher cadre care. Though most of the referral cases are from the private sector, the private healthcare outlets also do receive cases of referral. Referral cases from the public facilities were more and they referred cases among themselves. There is some association between formal referral and access to preferred healthcare as well. Formal referrals help to strengthen the link between the healthcare facility preference (the healthcare facilities where people would like to seek healthcare) and where they actually do seek healthcare. About 74% of the respondents preferred private healthcare facilities though only 20.2% were referred to the private healthcare outlets. In addition to improved access to private healthcare there is also increased satisfaction with the services received in the healthcare outlets they were referred to.

Respondents in the survey tended to relate formal referral with quality of healthcare. The fact that formal referral is a determinant of household access to healthcare, as a result of perceived quality

of healthcare is in line with health belief models (Hochbaum, 1958; Rosentock, 1966). One of the key descriptors of health belief model is the perceived benefit of taking action; taking action toward the prevention of disease or toward dealing with an illness is the next step to accept after the individual has accepted that they are susceptible to a disease and recognised it is as serious. During the initial contact of seeking healthcare, the household chooses a facility because the associated benefits are perceived to be higher than those of alternative outlets. However, with formal referral contacts, the quality of healthcare previously received is used as a learning opportunity and the household will only return to the same provider if there is a perceived benefit.

The lack of relation between higher existence of chronic illness in a household and formal referral suggest that most of the referred cases are not as a result of chronic illness after all. This finding is surprising in view that healthcare providers at lower levels of the health system, who lack the skills, the facilities, or both to manage a given clinical condition, are supposed to seek the assistance of providers who are better equipped or specially trained to guide them in managing or to take over responsibility for a particular episode of a clinical condition in a patient (Al-Mazrou *et al.*, 1990). This kind of skills and facilities are normally located in referral outlets. Finally, though the familiarity of the respondents with referral system does not translate to high usage, the referral system is meant to improve access to healthcare among the populace and to protect the poor from the cost of medication by accessing the lower cadre facilities before moving to the upper cadre ones (see related works in Bedi *et al.*, 2004; and Owino and Were, 1997). It is evident from the study findings that strengthening of the referral system would help increase access to healthcare services for the households.

The results of the study showed that in general, usage of referral was very low in Eldoret municipality as previously documented elsewhere in the municipality (Nyakaana, 1996). However, in terms of referrals, formal referral is higher compared to self-referral. Findings confirm that formal referral in Eldoret contributes to access to healthcare for the concerned households. Critical to the question of access to healthcare services is the one of access to better facility outlet and medical specialist. Formal referral has significantly improved accessibility to healthcare for households by increasing their probability of visit to a higher cadre health facility. The study emphasises and shows trends toward higher rates of access to specialist care (medical specialist). Access to a medical specialist as opposed to a general physician increases with formal referral.

6.2 Recommendations

Social science research aims at evolving new possibilities for action. This research finding is available for the attainment of social action and change. Therefore, the recommendation to policy-makers is to focus attention on improving access not only healthcare services, but the referral system. The study found that most households in urban areas rely on public healthcare facilities. However, majority of these are from the lower income households and without health insurance plans. An implication is that majority of the poor (who already have multiple health needs and are vulnerable) are not able to access proper healthcare in the private sector, hence their reliance on public sector outlets, which are considered to be of 'lower quality'. This is indeed a double tragedy. Consequently, provision of 'better' private healthcare is skewed towards the well off in the community.

Even as households are educated on the importance of adhering to the referral system, the referral hospitals need to move some of their services from their main hub to other parts of towns' healthcare facilities to increase their visibility, and expand access to treatment, care, and support services. Further, to be effective, referrals must be linked to referral documentation to ensure appropriate follow up of the system by clients.

Investigations are needed to better understand self-treatment form of healthcare so that policy can be designed to improve its use by the invalid, or if necessary to divert demand to other providers of healthcare. In addition there is also need to investigate further the reasons for their increased use. With the increasing prevalence of drug resistant strains of illness, there is a risk of self-treatment contributing to prevalence of diseases exacerbating problem of access to healthcare.

Finally, the work, however, is far from done. Indeed, many studies on healthcare access have been conducted in the past. In recent years, the results of various research studies have taken center stage in the popular media. More informed policies can be formulated as a consequence. It is my hope that the findings of this present study, along with the several others produced, will now stimulate greater interest in this line of inquiry. Further, the conclusions made in this study can later be verified

in different research areas taking into account of particular local circumstances, cultural beliefs, and characteristics of health systems. Hopefully, findings from this study will now be used to help focus and prioritize interventions aimed at improving the performance of the health facilities up the referral ladder.

References

- Becker, M.H. (ed.) (1974). *The Health Belief Model and Personal Health Behaviour*. San Francisco: Society for Public Health Education, Inc.
- Bedi, A.S., Kimalu, P., Kimenyi, M.S., Manda, D.K., Mwabu, G., and Nafula, N. (2004). *User Charges and Utilization of Health Services in Kenya*. The Kenya Institute for Public Policy Research and Analysis (KIPPRA). Discussion Paper Series No.32.
- Carr, D., (2004). *Improving the Health of the World's Poorest People*. Health Bulletin 1 Washington DC: Population Reference Bureau.
- Cockerham, W.C., (1986). *Medical Sociology*. New Jersey Prentice- Hall, Inc., Englewood Cliffs.
- Ellis, F., Freeman, H.D. (2007). 'Rural Livelihoods and Poverty Reduction Strategies in Four African Countries.' *Journal of Development Studies* 40:4, 1-30.
- Filmer, D., Hammer, J., and Pritchett L. (1997). *Health Policy in Poor Countries: Weak Links in the Chain*. Policy Research Working Paper 1874, World Bank, Washington, DC.
- Gilbert, N. (ed.) (1993). *Researching Social Life*. London: Sage.
- Heenan, D. (2006). "The factors influencing access to health and social care in the farming communities of County Down, Northern Ireland" *Ageing and Society* 26: 373-391.
- Hochbaum, G.M. (1958) *Public Participation in Medical Screening Programs: A Sociopsychological Study*, PHS Publication no. 572. US. Washington, DC: Government Printing Office.
- Institute for Policy Analysis and Research (2003). *Safety Nets in Kenya's Public Health Sector: More Insights into Waivers and Exemptions*. IPAR Policy Brief Vol. 9, No. 7 pp. 1-2.
- Iyer, A. (2005). *Gender, Caste, Class and Health Care Access: Experience of Rural Households in Koppal District, Karnataka, Trivandrum, Achutha Menon Centre for Health Science Studies, Sree Chitra Tirunal*. Kerala: Institute for Medical Sciences and Technology.
- Janz, N.K., and Becker, M.H. (1984). 'The Health Belief Model: A Decade Later.' *Health Education Quarterly*, 11, 1-47.
- Kenya National Bureau of Statistics (KNBS) and ICF Macro. (2010). *Kenya Demographic and Health Survey 2008-2009*. Calverton, Maryland: KNBS and ICF Macro.
- Kitts, J., and Roberts, J.H. (1996). *The Health Gap: Beyond Pregnancy and Reproduction*. Ottawa: International Development Research Centre.
- Maliyamkono, T.L., and Ogbu, O. eds. (1999). *Cost Sharing in Education and Health*. Dar es Salaam: Tema Publishers Company Ltd.
- March, C., Smyth, I., and Mukhopadhyay, M. (1999). *A Guide to Gender-Analysis Frameworks*. Oxford: Oxfam.
- Ministry of Health (MOH). (1997). *A question of survival? Review of Safe Motherhood*. Nairobi: Ministry of Health - Division of Primary Health Care.
- Moi Teaching and Referral Hospital. (2005). *Strategic Plan July 2005 - June 2010: Striving for Excellence*. Eldoret: MTRH.
- Musgrove, P., ed. (2004). *Health Economics in Development*. Washington D.C.: The World Bank.
- Mwanje, J.I. (2001). *Issues in Social Science Research: Social Science Research Methodology Series*. Addis Ababa: Organization for Social Science Research in Eastern and Southern Africa (OSSREA).
- National Coordinating Agency for Population and Development (NCPAD). Ministry of Health (MOH), Central Bureau of Statistics (CBS), ORC Macro (2005). *Kenya Service Provision Assessment Survey*. 2004. Nairobi: NCPAD.
- Ndeti, D.M., Khasakhala, L., and Omolo, J.O. (2008). *Incentives for Health Worker Retention in Kenya: An Assessment of Current Practice*. EQUINET Discussion Paper Series, 62.
- Nyakaana, J.B. (1996). *Kenya's Development Centre Policy: the Case of Eldoret an Assessment of its Implementation and Impact*. Amsterdam: Drukkerij Elinkwijn b.v. Utrecht.

- Oleske, D.M. (2001). *Epidemiology and the Delivery of Health Care Services*. New York: Kluwer Academic/ plenum publishers.
- Owino, E.G. (2005). *Preferences and Utilization of Health Services by Slum Residents: A Case Study of Mathare, Nairobi*. Unpublished MA Thesis, Kenyatta University.
- Owino, W., and Were, M. (1997). *Delivery and Financing of Healthcare Services in Kenya: Critical Issues and Research Gaps*. IPAR Discussion Paper No. 002/97.
- Patrick, D.L., and Erickson, P. (1993). *Health Status and Health Policy: Quality of Life in Health Care Evaluation and Resource Allocation*. New York: Oxford University Press.
- Preker, A.S., Carrin, G. eds. (2004). *Health Financing for the Poor People: Resource Mobilization and Risk Sharing*. Washington DC: The World Bank.
- Rosenstock, I. M. (1966). 'Why People Use Health Services.' *Milbank Memorial Fund Quarterly*, 44:94-124.
- Republic of Kenya (1997). *Uasin Gishu District Development Plan 1997-2001*. Nairobi: Govern Printer.
- Republic of Kenya (2001). *1999 Population and Housing Census*. Nairobi: Government Printer.
- _____. (2002a). *National Development Plan 2002-2008: Effective Management for Sustainable Economic Growth and Poverty Reduction*. Nairobi: Government Printer.
- _____. (2002b). *Uasin Gishu District Development Plan 2002-2008: Effective Management for Sustainable Economic Growth and Poverty Reduction*. Nairobi: Government Printer.
- _____. (2005). *Uasin Gishu District Strategic Plan 2005-2010 for Implementation of the National Population Policy for Sustainable Development*. Nairobi: NCPD 2005.
- _____. (2007). *The Planning Bulletin, July-December 2007*. Nairobi: Government Printer.
- _____. (2008a). *Eldoret West District Development Plan 2008-2012*. Nairobi: Government Printer.
- _____. (2008b). *Eldoret East District Development Plan 2008-2012*. Nairobi: Government Printer.
- _____. (2009). *End Term Review of the Economic Recovery strategy for Wealth and Employment Creation (ERS) 2003-2007*. Nairobi: Government Printer.
- _____. (2010). *The 2009 Kenya Population and Housing Census, Counting our people for the Implementation of Vision 2030 Volume 1A: Population Distribution by Administrative Units*. Nairobi: Government Printer.
- Ridsdale, L. (1998). *Evidence-Based Practice in Primary Care*. Edinburgh: Churchill Livingstone.
- Scambler, G. ed. (2003). *Sociology as Applied to Medicine 5th edition*. London: Saunders.
- Tones, K., Tilford, S. (2001). *Health Promotion: Effectiveness, Efficiency and Equity*. Cheltenham: Nelson Thornes Ltd.
- United Nations Children's Fund. (1991). *Community Financing of Primary Health Care: The Bamako Initiative in Kenya*. Nairobi: UNICEF.
- Westendorf, D.G., and Ghai, D. (ed). (1995). *Monitoring Social Progress in the 1990's*. Brookfield: Avebury.
- World Bank. (1994). *Monitoring and Evaluating Social Programmes in Developing Countries*. Washington D.C.: World Bank.
- _____. (2006). *African Development Indicators*. Washington DC: The World Bank.
- _____. (2007). *Private Voluntary Health Insurance in Development: Friend or Foe?* Washington D.C.: The World Bank.
- World Health Organization. (1978). *Primary Health care: Report of the International Conference on Primary Health Care, Alma-Ata, USSR, 6-12 September 1978*. Geneva: WHO.

FROM EPISTEMOLOGICAL TRENDS TO CLINICAL AND EDUCATIONAL PRACTICES IN SPEECH-LANGUAGE PATHOLOGY

Nicole Keating, (Sessional Professor, M.Sc.S. S-LP)
Chantal Mayer-Crittenden, (Assistant Professor, M.Sc.S. S-LP)
Michele Minor-Corriveau, (Assistant Professor, PhD. S-LP)
Manon Robillard, (Assistant Professor, M.Sc.S. S-LP)
Roxanne Belanger, (Assistant Professor, M.Sc.S. S-LP)

Laurentian University, Canada

Abstract:

Without denying the importance of disciplinary studies, it is important to identify the role of interdisciplinary research. Although there is a division of knowledge into disciplines, the movement toward interdisciplinarity demonstrates a limitless innovation that better represents the reality of contemporary society and its beliefs. According to John Locke's theory of the mind, one is born without innate ideas or concepts (Deigh 1994). We could assume that our knowledge is based on personal experiences and sensory perceptions. Understanding the human being is almost by necessity multidisciplinary because there is no single school of thought or discipline that can define the human being without taking into consideration the role of other disciplines. This paper will not only examine the role that interdisciplinarity plays in current research but it will also examine how speech-language pathology is innately interdisciplinarian and evidence-based. Clinical and educational practices will serve as examples to show how this field is inherently interdisciplinary.

Key Words: Epistemology, Educational Speech-Language Pathology, Interdisciplinarity

Introduction:

The degree of integration of the interactions between disciplines should not be seen as combinations of thought but rather as the establishment of a hierarchical distinction (Piaget 1970). According to this school of thought, multidisciplinary equates a phenomenon of borrowing from one science to another without causing changes in the original discipline. Piaget goes on to describe the interdisciplinary level as a reciprocal exchange between disciplines while the transdisciplinary level represents relationships between disciplines that would fall within a total system without borders. We notice, once again, that interdisciplinarity represents an exchange of ideas between disciplines. In order to standardize this interdisciplinary exchange, we should discuss the importance of epistemological trends in regards to scientific research.

Terminology

In order to further examine interdisciplinarity, it is essential to review the terminology associated with the study of disciplines. A summary of interdisciplinarity is defined as an area that encompasses several disciplines whereas multidisciplinary is defined as an area for several disciplines (Voyer 2000). Finally, transdisciplinarity examines an area that crosses the borders of various disciplines which represents the integration of other disciplines to support a specific study (Voyer 2000). In short, despite the lexical incoherence, we may find that interdisciplinary study means a study not only between disciplines but also a study that respects the contribution of each discipline and takes into consideration the importance of interdisciplinary collaboration in specific research. According to Peter Weingart, "Both interdisciplinarity and disciplinarity are, thus, given positive valuations for different functions: innovation on the one hand and rigor and control for error on the other. (...) They are complementary rather than contradictory: No new discovery is made without a frame of mind that allows one to distinguish between new and old, and relevant and irrelevant, and to record and remember "(Weingart 2000, p.29).

Epistemological Trends

In this section we will examine the epistemological views of several authors. Firstly, epistemology, according to Gauthier, is a prospective - which therefore concerns the future - that does not necessarily reflect current science, but rather critiques science in a constructive heuristic manner (Gauthier 2005). The philosophical interests of Gauthier go beyond the scope of logic, mathematics and physics and he notes that there is a distinction between science and philosophy. According to Gauthier, a scientific theory is "a discourse on a subject of science" (Gauthier 2005, p.19) while the philosophical discourse becomes more of a *meta* as in the study of metaphysics. It should be noted therefore that an object that qualifies as *scientific* can only be thusly defined when a scientific theory determines it so. Could philosophical discourse, then, lead to the identification of an object if it is first a thought that becomes a scientific theory? Gauthier adds that the dichotomy between theory and practice is found at all levels of scientific discourse (Gauthier 2005). If the important link between theory and practice is recognized, can we therefore conclude that Gauthier offers an interdisciplinary approach to the study of science?

Secondly, Edgar Morin also offers a logical approach when determining the pillars of classical science: order, separability and reason. Morin says that order is a deterministic and mechanistic concept of the world while separability is based on the Cartesian principle that solves a problem or a phenomenon by breaking the simple elements. He explains that reason is based on three principles: induction, deduction and identity (Morin 1999). It would be almost remiss not to mention that Karl Popper was among the first to assert himself against induction because he did not support the theory of accepting a conclusion by induction only (Popper 1972).

Morin, whose three pillars have been shaken by the developments of contemporary science later offered three theories: information theory, cybernetics and systems theory. The information theory deals with the uncertainty, the surprise and the unexpected where the information becomes what controls the energy and gives autonomy to a machine (Morin 1999), while cybernetics examines autonomous machines and hints at the appearance of artificial intelligence. This theory also includes the aspect of feedback that can have either a negative or a positive form. Negative feedback stabilizes a system and reduces the deviance while positive feedback represents a rather inflationary approach (Morin 1999). When we discuss feedback, however, we also speak of causality that, despite Morin's reference to artificial intelligence, is somewhat ironic because one of the possible foundations of causality includes the role of religion as the primary cause in the feedback cycle.

Finally, Morin introduced the systems theory, which also lays the groundwork for organizational thinking (Morin 1999). According to Morin, this theory includes the psychological aspect of gestalt because it explains that the whole is greater than the sum of its parts. According to Morin, the three theories introduce us to a world of phenomena where organization is done with order as well as with chaos (Morin 1999). Complex thinking is essentially the theory that examines research and / or opinions with uncertainty and which is able to create somewhat of an organization. The paradigm of complexity however cannot be synonymous with simplification as it requires to connect but always while making a distinction (Morin, 1999) and so the debate ensues regarding unidisciplinarity, multidisciplinarity and interdisciplinarity.

Logic seems to be a common feature between Gauthier, with his constructivist approach and Morin, with his complex thought, but how does this common trait bind these two schools of thought? Is a connection even possible? Gauthier speaks of internal logic as that which refers to a definite theory that aims to better represent individual factors of a theory (Gauthier 1991) while Morin considers logic without the appearance of contradiction. He implies that thought would lose creativity, a sense of innovation and complexity if logic prevailed. Evidently two theorists referencing the same vocabulary however, semantically, two schools of thought prevail.

Gauthier claims that philosophy, metaphysics and theology cannot be excluded a priori from the field of science (Gauthier 2005). Gauthier insists therefore that the unscientific eye must first be rejected despite the fact that it plays a fundamental role in the field of science. Morin for his part argues that complex thought can be reduced to neither science nor philosophy, but rather allows for the communication from one to another (Morin 1999). Can we therefore conclude that Yvon Gauthier acknowledges the presence of other disciplines while Edgar Morin recognizes the importance of other disciplines? According to Laflamme, Yvon Gauthier builds bridges between systematic and empiric

theory that emerge from his research of less interpretative findings (Laflamme 2008) and, perhaps, by supporting probability and statistics. Edgar Morin, however, presents a logic that is not representative of a linear function as well as a thought that supports the mosaic of different areas such that interdisciplinarity is inherent. Is it realistic to be able to convey the notion of interdisciplinarity in a clinical setting and in a field of scientific study?

Clinical Practice

Speech-language pathologists have a range of knowledge regarding normal development and communication disorders and master the technical evaluation and response thereto. On the one hand, the evaluation of communication disorders is done through a process of screening, identification, assessment and, most recently, diagnostic. For its part, the intervention involves the promotion, prevention, counseling, treatment, consultation, care, rehabilitation, and finally, training and rehabilitation. Speech-language pathologists work directly with clients and other individuals who regularly interact with them, including primarily members of the family, paraprofessionals, colleagues and support staff. Speech-language pathology (SLP) is a regulated health profession which must meet the requirements mandated by the governing law. In order to join the College of speech-language pathologists, a member shall, in addition to meeting the academic expectations, meet "certain aspects of local expertise, including: case law, values and ethical framework, systems and health policies" (CASLPA 2011). Speech-language pathologists must also demonstrate continued competence and respect the code of ethics that, in general, describes the core professional values.

Recently, the Canadian Association of Speech-Language Pathologists and Audiologists (CASLPA) and the College of Audiologists and Speech-Language Pathologists (CASLPO) strongly recommend the use of practices that are observable, measurable, manageable, and which are established by experts (CASLPA 2011). These guidelines are also skills in close connection with the practice that is evidence-based, otherwise known as evidence based practice (EBP). When speaking of practices that are evidence-based, they must include separate, measurable and meaningful elements that often result from a collection and data analysis. This evidence therefore eventually leads to a practice generally accepted by the clinicians and thus forms a standard applicable in daily work. The evidence based practice not only encourages clinical approaches that are the basis of formal research, but also persuades the clinicians to think in terms of EBP, which is to say that clinicians are encouraged to weigh the evidence relating to their own practice and provide measures by which to gauge performance and progress. Take for example the speech-language pathologists of the *Conseil scolaire catholique de Nouvel Ontario*, a French school-board set in a linguistic minority setting. A group of francophone clinicians working within a linguistic minority, decided to create their own screening tool instead of simply translating an English tool that did not represent normative data (Minor-Corriveau, 2012). One could deduce that they adopted a kind of logical positivism as they created their own tool, based on data collection and following an experimental reliability and validity, to meet their own needs. One might also add that they simply respected the mandate of their College, which explains that a proper assessment involves the following: "a) the identification of disorders best performed in the native language and combined with the assessment in the second language b) cultural and linguistic appropriateness c) well supported; d) an evaluation process that is natural and holistic and includes the use of non-standard methods, e) consideration of societal factors which may hinder language proficiency; f) evaluation reports that are descriptive" (CASLPA 2011). Certainly it is not realistic that a clinician develop a new assessment tool each time a different tool is inadequate but logical positivism certainly encourages evidence-based practice as logical positivism is based on the optimal situation, either the model or ideal, of research (Gauthier 2005).

The speech-language pathologist may work in several environments: a clinic that targets the pre-school population, a school board, a treatment center for children, a hospital, a rehabilitation center, a private clinic, etc. In short the environment in which the clinician performs can vary but one factor that remains constant is the client and despite this consistent factor, no two clients are alike. How can a clinician assess or act strictly based on the evidence of certain practices especially if there are no established practices for each affected population? The relationship between the client and its environment therefore becomes an essential factor.

Consider a group intervention session. Group members include the speech-language pathologist, a 5-year old boy, a 5-year old girl and another 6-year old boy. The three children have a phonological awareness delay, specifically in terms of syllabic segmentation. It is also noteworthy

that the 5-year old girl is blind and the 6-year old boy has a hearing impairment. In addition, the SLP must consider the techniques taught in the classroom so as not to confuse the message of the teacher. What factors dominate the approach to assess and intervene? Evidence-based practice, the client's needs and his or her environment or a combination of both approaches? It seems that a combination of both approaches would be the ideal practice. Knowing that technology plays a role in learning for the client with a loss of vision (Hersh and Johnson 2008), the clinician could incorporate a hearing program to explain the syllabic segmentation while the little boy who has a hearing impairment could use tokens or other visual cues to facilitate understanding. Meanwhile, the 5-year old boy could benefit from both approaches. This portrait of a speech-language session is not exceptional, but rather typical, and the speech-language pathologists must often change their approaches to accommodate as much as possible the needs of their clients. So we could see that despite the importance of the influence of logical positivism, the SLP takes into account the interaction between the subject and its environment.

Although Piaget moves away from logical positivism, his theory is largely based on observations (Piaget 1970). According to this author, it is also thought that knowledge does not develop without interaction with the environment even if it remains neutral. Take for example the role of feedback and modeling in the field of speech-language pathology. An adult is seated during a session of intervention where he works on word fluency following a brain injury. At the beginning of the session, the clinician explains the breathing techniques and strategies to consider when moments of dysfluency occur. During the course of the session, the therapist is merely carrying out the activities and allows the client to self-regulate when facing difficulties. In this scenario, the session would evidently be ineffective as the clinician does not offer any feedback or modeling of the recommended techniques. One could therefore argue that intervention and learning are not unidirectional and that the interaction between clinician and client is essential because it not only contributes to the effectiveness of the intervention but is also necessary in order to achieve therapeutic goals. The feedback provided by the SLP could also allow for learning to take place by problem solving because the client could change his techniques and self-correct in response to comments from the feedback.

Epistemology and interdisciplinary research

In epistemology, research can be referred to as soft or hard in the sense that we may or may not see the links with other sectors or disciplines (Vanpouille 2011). These external or scientific borrowings contribute to the field of interdisciplinary studies because they give rise to the transfer of methods and / or knowledge from one discipline to another. According to Franck, the unity of knowledge claims that knowledge is unique and there is a harmonization of concepts (Franck 1999). Furthermore he explains that the unity of knowledge does not require erasing disciplinary boundaries but rather that interdisciplinarity is not limited strictly to scientific borrowings. Research can thus be defined as an interdisciplinary collaboration of researchers where anyone from any discipline includes the study area without being aware of jargon and / or knowledge related to a specific discipline (Voyer 2000).

According to the Canadian Association of Speech-Language Pathologists and Audiologists, SLPs are autonomous professionals who have expertise with regards to normal development and communication disorders, and swallowing as well as in the assessment and intervention of these areas (CASLPA 2011). According to the College of Psychologists of Ontario (CPO), psychology services include, without limitation, one or more of the following: a) assessment and diagnosis of a person or group; b) interventions conducted with a person or group; c) consultation d) development and evaluation; e) supervision; f) research (The College of Psychologists of Ontario 2009). Sociology can be defined as the branch of human sciences which seeks to understand and explain the impact of social representations on human behavior: research subjects are varied but are still connected to the human environment and take into consideration human social phenomena from different angles (Touraine 2003); whereas, linguistics is the study of human language and should not be confused with the study of grammar (Laurence 2003). The linguist studies the mechanisms of language in a broad manner, however there is also a specific analysis in theoretical linguistics (i.e. phonetics, syntax, morphology, semantics) The linguist also compares the use of a language for a specific period (synchronic linguistics) against the evolution of a language (diachronic linguistics) (Rodrigues Aristar

1991). Research in linguistics is often descriptive and explains the nature of language without passing value judgments (i.e. ebonics) (Baugh 2000).

Following a brief description of the disciplines mentioned above, it is useful to emphasize the commonalities of these compared to speech and language development. First, the speech-language pathologist is interested, among other things, in the comprehension and use of language in daily life. For its part, the psychologist is interested in the cognitive processes involved in the use of language and often uses skills related to language in order to determine the cognitive abilities of an individual, while the sociologist possibly considers the interests of an individual regarding language use and / or socioeconomic factors such as education which has an impact on language development. Finally, linguists are interested in language processes included in the use of a language (also important for clinicians), as well as the omission of phonetic or morphological markers, for example, omission of the plural marker - possibly as a sociological factor. Evidently there are many commonalities between the above mentioned disciplines and the study of speech and language development and its use would be incomplete without taking into consideration various fields of study – many of which remain to be examined in this paper such as biology, neuroscience and medicine just to name a few.

Conclusion:

As mentioned above, the speech-language pathologist works with a diverse population with various communicative disorders. Language can certainly be interpreted, analyzed and studied by several disciplines and because of its complexity, language lends itself to systematic reviews at various angles depending on various fields of study. First, because language is a complex orchestration of multiple cognitive processes, it is an important field of research in areas such as psychology and speech-language pathology as we attempt to understand it with a microstructural approach. Similarly, in linguistics, we focus on the concept of language according to its linguistic components. On the other hand, since it is an essential skill in today's society, language development and the use thereof are also researched in the fields of education and sociology where there is more interest in the practical applications and its macrofunctional role. In summary, language skills are related to human development that is studied in various fields and it should be noted that a unidisciplinarian approach alone does not allow for an extensive and comprehensive review. However, the field of speech-language pathology and its interdisciplinarian nature allows for such an extensive and comprehensive review.

References:

- Aristar, Anthony Rodrigues. On Diachronic Sources and Synchronic Pattern: An Investigation Into The Origin Of Linguistic Universals. *Language* 67(1):1-33,1991.
- Baugh, John. Beyond Ebonics: Linguistic Pride and Racial Prejudice. New York: Oxford University Press, 2000.
- Canadian Association of Speech-Language Pathologists and Audiologists. Scopes of Practice, CASLPA, <http://www.caslpa.ca/francais/resources/scopes.asp>, November 6, 2011.
- College of Audiologists and Speech-Language Pathologists of Ontario. Practice Standards – Code of Ethics, CASLPO, <http://www.caslpo.com/PracticeStandards/CodeOfEthics/tabid/102/Default.aspx>, November 4, 2011.
- College of Physicians of Ontario. Scopes of Practice. CPO, <http://www.cpo.on.ca/about-the-college/index.aspx?>, November 7, 2011.
- Deigh, John. Cognitivism in the Theory of Emotions. *Ethics* 104(4):824-854, 1994.
- Franck, Robert. La pluralité des disciplines, l'unité du savoir et les connaissances ordinaires. *Sociologie et société. L'interdisciplinarité ordinaire* 31(1):129-142, 1999.
- Gauthier, Yvon. Entre science et culture : Introduction à la philosophie des sciences. Montréal: Presses de l'Université de Montréal, 2005.
- Hersh, Marion A. and Michael A. Johnson. Assistive Technology for Visually Impaired and Blind People. London: Springer-Verlag, 2008.
- Laflamme, Simon. Analyse statistique linéaire et interpretation systématique. *Nouvelles perspectives en sciences sociales : revue internationale de systématique complexe et d'études relationnelles* 4(1):141-159, 2008.
- Laurence, Stephen. Epistemology of Language. Oxford: Oxford University Press, 2003
- Minor-Corriveau, Michèle. Étude normative du développement du langage chez

l'enfant. PhD diss., Laurentian University, 2012.

Morin, Edgar and Jean-Louis Le Moigne. *L'intelligence de la complexité*. Paris: L'Harmattan, 1999.

Piaget, Jean. *L'épistémologie génétique*. Paris: Presses Universitaires de France, 1970.

Popper, Karl. *Objective Knowledge: An Evolutionary Approach*. USA: Oxford University Press, 1972.

Touraine, Alain. *Sociology without Societies*. *Current Sociology* 51(2):123-131, 2003.

Vanpouille, Yannick. *Épistémologie du corps en Staps: Vers un nouveau paradigme*. Paris: L'harmattan, 2011.

Voyer, Philippe. *L'interdisciplinarité, un défi à relever*, *The Canadian Nurse/L'infirmière canadienne* May:39-44, 2000.

Weingart, Peter and Nico Stehr. *Practising Interdisciplinarity*. Toronto: University of Toronto Press, 2000.

A DEA-SFA COMPARISON OF THE IMPACT OF ICT'S UTILIZATION

Puiu Fatulescu, PhD student
Alexandru Ioan Cuza University, Romania

Abstract:

This article has the purpose of comparing different statistical methods of measuring efficiency, data envelopment analysis and stochastic frontier analysis. An adjacent objective is to assess the impact of the information and communication technologies on today's society. The hypotheses of the article are that there is indeed a connection between the scores obtained by applying DEA method with the ones obtained by the SFA method. Also, it is hypothesized that the level of ICTs use is higher for those countries that are more developed and it is driven by a couple of key factors, chosen as input variables in the present study.

The results show that both hypotheses are valid, as the efficiency scores obtained were statistically correlated between the different methods and as the variables chosen as inputs were significant in the SFA model.

Key Words: Efficiency, health systems, DEA, SFA

Introduction:

The end of the 20th century and the beginning of the 21st have marked the highest peak of socio-economical progress that society has ever known across history. Undoubtedly, among the factors channeled towards optimization and perpetual development of the society, scientific and technologic progress has the most consistent contribution. Furthermore, present trends of creating a virtual parallel world developed from the need of cost optimization and time saving, made information technologies an essential instrument in each citizen's life.

This article has the purpose of comparing two statistical methods of measuring efficiency, data envelopment analysis (DEA) and stochastic frontier analysis (SFA). An adjacent objective is to assess the impact of the information and communication technologies (ICT) on today's society. The hypotheses of the study are that there is indeed a connection between the scores obtained by applying DEA method with the ones obtained by the SFA method. Also, it is hypothesized that the level of ICTs use is higher for those countries that are more developed and it is driven by a couple of key factors, chosen as input variables in the present study.

Subsequent the introductory part, a literature review on the connection between ICTs and education and methods of assessing will be presented in the second part. The third part will cover the methods and data used for the study following that the forth part will be dedicated to discussing the results. The study will end with a section of general conclusions and future directions for other studies.

Literature Review:

The progress of all societies depends on the efficiency with which natural, human and financial resources are being disposed of. Therefore, when assessing public sector efficiency managers take into account that the efforts to fulfill social needs can be measured, often quantified with the value of inputs (e.g. costs of raw materials, costs of human resources, costs of information), while social effects are difficult to determine and measure. Furthermore, they are difficult to be fully forecasted. Improving public sector performance is an objective with a high importance role in the agenda of each industrialized state.

It is no wonder that governmental efficiency as a whole became the subject of an increased number of papers, received key contributions from Gupta et Verhoeven (2001), Tanzi et Schuknecht (1997, 2000) and Alfonso et al. (2006). These studies measure the efficiency of public sector by connecting government spending with socio-economical indicators. Those indicators are assumed to

be in close connection with the objectives of public spending. For example, the percentage of pupils enrolled in educational units or percentage of infant mortality are indicators of substantial differences of the efficiency levels between countries, regardless of the level of development.

A study from 1998 reflected that the simple allocation of the resources, even for necessary goods and services, is not enough and it is possible that the outcomes might not be the expected ones if the budgetary institutions involved in planning, management and execution do not work at their maximum efficiency. Vicious budgetary management has often been cited as one of the reasons for which the governments from the developing countries found themselves incapable of transposing public expenses into efficient services (World Bank, 2003).

The effect of income per inhabitant can be analyzed from several points of view. From the first point of view the income could reduce efficiency by increasing relative public services costs (Baumol, 1967). From the other point of view, a higher income has been many times associated with a higher level of health and education (Afonso et al., 2006). When talking about the level of efficiency, evidences show that it can be improved by increasing scale operation. This fact is shown primarily in health and education sectors (Coelli et al., 2001; Curristine T, 2005; Dronkers, 2004). This is due to scale economies which are the result of the saving of additional marginal costs compared to the fixed costs of resources. Nevertheless, their impact over other fields of public sector like equity, quality or access to services has to be considered (Dooren et al., 2007).

The importance of environmental variables is also taken into account. Kasman's (2005) paper presents an empirical analysis of banks' efficiency in Poland and Czech Republic, taking into account the country-specific environmental factors in the second part of the analysis. The authors used Stochastic Frontier Analysis in order to measure efficiency. The results show that, without environmental variables, the cost-inefficiency scores of Czech banks are quite high compared with those of Polish banks. However, when environmental variables are included in the model, the differences between the two banking sectors decrease dramatically. Therefore, the results indicate that country-specific environmental variables are important in the definition of the common frontier (Kasman, 2005).

Researcher's interest has also been channeled towards the connection between sectoral public spending (especially for education and health) and their outcomes (Rajkumar et al. 2008). Jayasuriya et al. (2007) use panel data for provinces from Argentina and Mexico to measure the efficiency of medical and educational services.

Nowadays ICTs are an important leverage of social and economic development, both at micro and macro level. Most of the production processes rely on human resources that are not only trained according to their positions but who can use different forms of ICTs. In other words, given the extent of today society's speed in changing and adapting to never-ending needs, workers have to be more efficient or to improve their skills, might those skills be in their areas of expertise or in any interdisciplinary fields. ICTs have many advances and recently governments have started emphasizing those advances. Some of the attractive features of ICTs are flexibility, interactivity and their ability to move high amount of data in almost real time or to engage more people at once.

In the field of education, ICTs have been growing in importance for the last years, especially by the use of E-learning which started to be more and more chosen by students in higher education institutions. Compared to higher education enrollments in general, online enrollments have been growing significantly faster (Allen et al., 2008). In connection to this, Asandului et al (2008), conclude that students wish to have a wider range of options in connection to the e-technologies, which will definitely contribute to increasing students' competences. Therefore, the student will benefit from the added value of e-learning after graduating university too, by making use of the technical skills he will have acquired during the study period. Moreover, Asandului et al., (2011) emphasize the connection between computers and e-learning, stating that by the use of the e-technologies the curricula is not abandoned and that abilities and capacities are more easily formed. Moreover, the use of e-technologies has been identified as a measure to increase firm's performance and productivity growth (Clayton et al., 2003).

DEA has also been used with various occasions for assessing different aspects of the medical field, such as hospital performance (Zhu, 2002), the efficiency of public policies (Coppola et al. 2003; Miller et al, 1996; Rosko, 1990; Sherman, 1984) and the performances of cardiac surgery (Chilingerian, 1995).

Method and Data:

This study is based upon two methods of measuring efficiency: data envelopment analysis (DEA) and stochastic frontier analysis (SFA). The main difference between the two is that DEA is a nonparametric method, based on empirical observed data, whereas SFA is a parametric method who based on the observed data infers the values of the efficiency also taking into account the inefficiency factors. The two methods are shortly presented in the next paragraphs.

The economic literature that influenced directly the development of SFA is given by the total number of studies researching the efficiency of the production process from the beginning of the 1950. Thus, Koopmans (1951) suggested a definition for technical efficiency while Debreu (1951) and Shephard (1953) presented distance functions and used them to model multiple outputs on one hand and to measure radial distance of a producer from the frontier on the other hand. This was done either by extending the outputs (Debreu) or by conserving the inputs (Shephard).

The first method applied in this research, Data Envelopment Analysis (DEA), is generally accepted as one of the best in assessing the efficiency of a set of decision making units (DMU)s. First presented in 1978 and based on the paper of Farrell, the first DEA model is known in the literature as the CCR model, after its authors, Charnes, Cooper and Rhodes. In essence, DEA is a non-parametric approach, who, with the help of linear programming techniques and based on the dataset, computes an efficiency frontier on which only the most efficient DMUs are placed. The DEA model is usually input or output oriented. An output oriented DEA model is channeled towards maximizing the outputs obtained by the DMUs while keeping the inputs constant whilst the input oriented models focus on minimizing the inputs used for processing the given amount of outputs. For input-oriented models, the others DMUs not as efficient as those placed on the frontier will be given a certain efficiency score above than 0 but below 1².

Thus, by using linear programming and by applying nonparametric techniques of frontier estimation, it can be measured the efficiency of a DMU, by comparing it with an identified frontier of efficiency. The advantage is that DEA does not require any prerequisite hypotheses regarding the analytical shape of the production function.

In this study, two DEA models will be run: a constant returns to scale, input oriented and a variable returns to scale input oriented model.

The CRS model

For a given set of data, the efficiency DMU_j is measured, n times, where n represents the number of DMU to be evaluated j ranges over 1, 2, ..., n . To obtain the scores for the weights of the inputs (v_i) ($i=1,2,\dots,m$) and the weights of the output (u_r) ($r=1,2,\dots,s$), the following set of linear programming equations need to be solved:

$$\text{Max } \theta = \mu_1 y_{10} + \dots + \mu_s y_{s0} \quad (1)$$

Subject to

$$\mu_1 y_{10} + \dots + \mu_s y_{s0} \leq v_1 x_{10} + \dots + v_m x_{m0} \quad (j=1,2,\dots,n)$$

$$v_1, v_2, v_3, \dots, v_m \geq 0$$

$$\mu_1, \mu_2, \mu_3, \dots, \mu_s \geq 0,$$

where θ is the optimal objective value and it is at most 1.

The VRS model is not completely different to the one previously presented. Thus, by adding the restraint $v_1 x_{10} + \dots + v_m x_{m0} = 1$ to the model above, we obtain the input oriented VRS DEA model.

Stochastic Frontier Analysis

As Greene (1997) concluded, the function of production frontier can be described as an extension of the regression, based on the micro economical assumption that a production function is an ideal, the maximum level of output that can be attained with the use of a given set of inputs³. SFA has become a very popular, used in analyzing the production from several points of view. Similar to DEA, the development of SFA and the preceding papers is based upon Farrell's study from 1957. Among others, Aigner et. Chu (1968), Seitz (1971), Timmer (1971) or Afriat (1972) contributed to SFA's

² Output oriented and super efficiency models have different restrictions regarding the efficiency scores.

³ Greene, W. (1997). Frontier Production Functions. In Handbook of Applied Econometrics. Volume II: Microeconomics, M.H. Pesaran and Schmidt, P. (Eds.), Oxford: Blackwell, p 35

elaboration. And even though the authors' inputs differ under certain aspects, each of them contributed to identifying a production frontier, either by using linear programming or by modifying the technique of the sum of squares, having the a-priori hypothesis of nonnegative residuals.

The fundamentals of SFA were set concomitantly with the release of two papers from two continents. The paper of Meesen and van den Broeck (the MB model) was released in June 1977, and the paper of Aigner, Lovell and Schmidt (the ALS model) one month after. The second one was actually a compilation of papers very similar in context, some belonging to Aigner, some to Lovell and Schmidt.

The ASL model was the first who separated the environmental effect by two components.

$$y_i = f(x_i; \beta) + \varepsilon_i, \quad i = 1, 2, \dots, N. \quad (2)$$

where $\varepsilon_i = v_i - u_i$, $i = 1, 2, \dots, N$

where y_i is the maximum output attainable by using x_i , a (non-stochastic) input, and β is an unknown vector parameter which has to be estimated. Moreover, the first component of the inefficiency term (ε_i), v_i is a sum of random events and facts (e.g. hazard, climate, machines' performance). Observation and measurement errors also fall in its structure. This error is assumed to be independent and normally distributed with $N(0, \sigma_v^2)$. The second component u_i is actually the sum of certain factors on which the firm, the producer can have an impact on like economic and technological inefficiency, availability, producer's effort, etc. The second component is considered to be independent from the first one, also complying to the restriction $u_i \geq 0$. The second error is meant to reflect that the output of any firm can be placed under or on the frontier of production $[f(x_i; \beta) + v_i]$. If we consider that $f(x_i; \beta)$ is a log-linear Cobb Douglas⁴ function, we can rewrite (2) as

$$\ln y_i = \beta_0 + \sum_n \beta_n \ln x_{ni} + v_i - u_i \quad (3)$$

For the analysis, the study uses data on 75 countries, the first 75 United Nation countries by the Human Development Index which had available information in regards to the variables chosen in the DEA model. The number of variables was set to five, four input variables and one output variable, with no missing data for the 75 countries.

Inputs

Quality of the Educational System. The variable uses 2010- 2011 values from the World Economic Forum's Reports and surveys, and represents a weighted score of the efficiency of the Educational System. To be more precise, the subjects were asked to rate on a scale from 1 (very bad) to 7 (very good) how well does the educational system in their country cope with the needs of a competitive economy. Thus the variable is very interesting by itself, as it contains information about the overall way of functioning of the educational systems and about how it interacts with the economy in the given country.

The availability of latest technologies. This indicator, collected from the World Economic Forum, Executive Opinion Survey 2010 and 2011 editions, shows to what extent are the latest technologies available in each country, on a scale from 1 (not available) to 7 (widely available).

The impact of ICT on access to basic services. This indicator is a 2010-2011 weighted average data from the World Economic Forum, Executive Opinion Survey 2010 and 2011 editions showing to what extent are information and communication technologies enabling access to citizens for basic services (education, health, etc.), on a scale from 1 (do not enable at all) to 7 (enable access significantly).

ICT use and government efficiency. The indicator is a 2010-2011 weighted average data from the World Economic Forum, Executive Opinion Survey 2010 and 2011 editions showing to what extent has the use of information and communication technologies by the government improved the efficiency of government's services in the countries on a scale from 1 (no effect) to 7 (extensive).

Outputs

Households with a personal computer. The indicator, based on data collected from the International Telecommunication Union, ITU World Telecommunication/ICT Indicators Database 2011, marks the proportion of households with at least one computer (e.g. a laptop or a desktop computer).

⁴ Three-dimensional production function, usually used to describe the relationship between inputs and output.

Data processing and analyses were run in R, a statistical software, using SFA and Benchmarking libraries.

Results:

Table 1 summarizes the descriptive statistics of the data. The use of latest technologies ranges from 3.1 (Barbados) to 6.9 (Sweden). This is a normal trend since towards the highest values the most developed countries are identified. From the 27 European Union member states, Romania is the one having the lowest score, being placed on the 72nd position. The small coefficient of variance (14.92%) indicates that the series is homogenous.

Table 1. Descriptive statistics for the chosen variables

Descriptive Statistics	Use of Latest Technologies	Quality of Educational System	Use of Basic ICTs	Use of ICT by the Government	Percentage of PCs at home
Minimum	3.1	2.3	3.1	2.5	15.6
1st Quintile	4.85	3.3	4.3	4.05	36.3
Median	5.5	4.1	4.9	4.5	61.4
Mean	5.503	4.108	4.909	4.625	58.23
3rd Quintile	6.25	4.8	5.6	5.3	78.75
Maximum	6.9	5.9	6.3	6.4	93
Coeff. of Variance	14.92%	22.70%	15.67%	18.08%	40.88%

From the Summary table, it can be observed that the Educational System variable ranges from a minimum score of 2.3 (Dominican Republic) to 5.9 (Switzerland, Finland and Singapore). For this variable, the last country from the EU-27 was Greece. Given that the Coefficient of variance is 22.7%, it can be concluded that the values for the variable are homogenous.

The use of basic ICTs ranges from 3.1 (Lebanon) to 6.3 (Sweden). The coefficient of variance is showing a homogeneous series (15.67%) and the mean and median are very close to each-other, pointing out the symmetry of the series.

Table 2. The frequencies of the efficiency scores for the three frontier models.

Efficiency Score	DEA				SFA	
	CRS		VRS			
	No of countries	% of countries	No of countries	% of countries	No of countries	% of countries
0-0.1	11	14.67%	0	0.00%	0	0.00%
0.1-0.2	14	18.67%	0	0.00%	0	0.00%
0.2-0.3	13	17.33%	4	5.33%	2	2.67%
0.3-0.4	6	8.00%	28	37.33%	7	9.33%
0.4-0.5	7	9.33%	9	12.00%	6	8.00%
0.5-0.6	7	9.33%	9	12.00%	8	10.67%
0.6-0.7	9	12.00%	12	16.00%	5	6.67%
0.7-0.8	2	2.67%	7	9.33%	12	16.00%
0.8-0.9	0	0.00%	0	0.00%	19	25.33%
0.9-1	2	2.67%	0	0.00%	16	21.33%
1	4	5.33%	6	8.00%	0	0.00%
Total	75	100.00%	75	100.00%	75	100.00%

The results in table 2 are controversial and deserve further discussions. Thus, the CRS DEA model has an average efficiency score of 0.38 and identifies 4 countries on the efficiency frontier: Qatar, Singapore, Sweden and Switzerland. The ranking is not surprising, as these countries are

known for having high levels of development and standards of living. Also, the countries on the frontier were identified as having strong efficiency⁵ (efficiency score 1 and no inputs excess). Moreover, all the Nordic countries were identified in the top ten, which reconfirms these states' reputation of high quality public services providers. Moreover, the model places the developed countries towards the first part of the ranking and the less developed countries towards the second half. From the 27 European Union countries, Bulgaria had the lowest efficiency score (0.13). This means that Bulgaria has to decrease the utilization of inputs until 13% of the current level, to become as efficient as Qatar and Singapore (Bulgaria's set of reference) for the amount of output obtained. Except for the countries on the frontier, the others were identified as having different amounts of excess for each of the four inputs. The availability of latest technologies should improve in average by 19.81% (from 5.5 points to 6.59 points) to ensure maximum efficiency, indifferent proportions ranging from 0.016 points for Iceland to 2.59 points for Armenia. The quality of Educational System should improve in average by 2.00 points (from 4.108 to 6.108), with improvements ranging from 0.056 for Iceland to 4.439 for Venezuela. Similar, the impact of ICT on access to basic services should improve by 1.468 points, from 4.909 to 6.377, ranging from 0.04 points for Iceland to 3.312 for Venezuela. The ICT use and government efficiency should improve by 1.63 points, from 4.625 to 6.255 points to ensure efficiency, ranging from 0.068 for Iceland to 3.596 for Venezuela.

The VRS DEA model is consistent with the CRS model, especially since it is known that all DMUs CRS efficient will also be VRS efficient (Cesaro et al., 2009). Thus, besides the four efficient countries already identified by the CRS model, the VRS identified Iceland and Finland as being efficient (score of 1) who, in the CRS model, had the next scores after the efficient countries. Moreover, in average, the VRS model had efficiency scores 0.13 points higher than the CRS model and consistent rankings⁶. The average efficiency increase is also aligned with the theory that the VRS frontier is more flexible and therefore envelops the data more tightly. Out of the 27 EU members, in the VRS model Greece was identified as having the lowest score (0.31), dropping 16 positions compared to the CRS model. Regarding the inputs excess, the VRS model identified, in average, 0.334 points of inputs excess compared to the CRS model, meaning approximately 21% decrease of inputs excess.

The maximum likelihood estimates of the SFA model are shown below.

Table 3. SFA estimates and significations

	Estimate	Std. Error	Z value	Sig.
(Intercept)	2.1354	0.204	10.484	***
Latest Technologies	0.8002	0.087	9.225	***
Educational System	0.5118	0.124	4.118	***
ICT impact on basic access	0.9801	0.178	5.510	***
ICT use & Gov Eff	-0.9236	0.167	-5.536	***
sigmaSq	0.2567	0.026	9.957	***
Gamma	1.0000	0.001	1205.452	***
Log - Likelihood	-7.4792			

***-significant at 1%

The SFA model identified an average efficiency of 0.71, the highest of the three tested model (0.51 the VRS DEA model and 0.38 the CRS DEA model). However, the results obtained with the SFA model differ significantly from the ones obtained by the other two models.

Thus, table 3 encompasses the estimates, standard errors, z-values and significances for the model's parameters. As shown, all coefficients are statistically significant for the 1% threshold. However, even though we would have expected positive values for all the parameters, ICT use and government efficiency was identified as having negative influence on the efficiency score. The estimation of gamma parameter, statistically significant, suggested that inefficiency was present during the production process and therefore the traditional production average would not be an

⁵ See Hongliang et al., (2007) for more details regarding strong and weak disposability.

⁶ Only 7% of the countries had rankings that changed more than 20 positions between the 2 models.

adequate representation of the data. As a result, technical inefficiencies have significant impact on the output and the one sided error component accounts for up to 100% of the total variance. In other words, 100% of the variation in the data between countries can be considered inefficiency, the model identifying no “noise”. Sigma squared was also statistically significant suggesting that the data does not fit the conventional production function.

As it can be seen in the table above, in the CRS DEA model 50.67% of the countries have the efficiency score of less than 0.3, whilst in the SFA model 62.66% of the countries have scores higher than 0.7. The results are not surprising since DEA considers all the deviations from the frontier as being inefficiency whereas SFA takes into account the fact that random shocks which are not under the control of the DMUs can have an impact on the output amount.

Table 4. Pearson correlations coefficients

	CRS	VRS	SFA
CRS	1		
VRS	.954**	1	
SFA	.587**	.408**	1

** . Correlation is significant at the 0.01 level (2-tailed).

Table 4 shows Pearson correlation coefficients between the three models. The first important fact to notice is that all three correlation coefficients are positive and highly statistically significant for the 1% threshold. The highest correlation coefficient is obtained as expected between the CRS and VRS DEA models whilst the weakest is between the VRS and SFA models.

Similar to the present study, in the literature there are a couple of studies that have been conducted to analyze the technical efficiency scores obtained by applying the methods above. Lin et al. (2005), Theodoridis et al. (2008) obtained similar efficiency scores rankings (SFA>VRS>CRS) and similar correlation coefficients ranking (CRS-VRS>CRS-SFA>VRS-SFA).

Conclusion:

Information is the resource of the future and correct communication will be the leverage necessary to fully benefit from the use of information. The beginning of the 21st century created the needed premises so as those who possess information have a considerable advantage. One of the biggest added values of ICTs is that they provide an astonishing amount of information with a very low level of requirements and they also work as catalysts in channeling information towards those who needed. Moreover the information pool is increasing exponentially in the virtual environment, supported by the new cloud computing concept.

However, when discussing the implementation of ICT, sensitive aspects need to be taken into account. The technical infrastructure is not sufficient and might not provide the expected results unless it is correlated with Educational efforts to provide minimum level of know-how.

This study analyzed the efficiency scores obtained by applying a CRS DEA model, a VRS DEA model and a SFA model. The results showed significant correlation exists between the efficiency scores of the above methods, the CRS DEA model having the lowest average efficiency score whilst the SFA model has the highest average efficiency. Also, the secondary hypothesis that developed countries make more use of ICTs that countries under developed or developing countries was validated by the models.

Future studies should focus identifying a relationship between the statistical methods applied in the research. Also, regarding the ICTs, the author believes that ICTs’ use and efficiency is highly influenced by the domain in which they are applied, therefore the exact impact of ICTs on each domain should be studied.

Acknowledgements

This work was supported by the **European Social Fund in Romania**, under the responsibility of the Managing Authority for the Sectoral Operational Programme for Human Resources Development 2007-2013 [**grant POSDRU/CPP 107/DMI 1.5/S/78342**]

References:

1. Afonso, A., Schuknecht, L., Tanzi, V., (2006). Public sector efficiency: evidence for the new EU member states and emerging markets. Working paper no. 581, European Central Bank, Frankfurt, Germany.
2. Afriat, S.N., (1972), Efficiency estimation of production functions, *International Economic Review* 13, Oct., 568-598.
3. Aigner, D. J., C. A. K. Lovell, and P. Schmidt (1977). Formulation and Estimation of Stochastic Frontier Function Models. *Journal of Econometrics* 6, 21–37.
4. Aigner, D.J. and Chu, S.F. (1968). On Estimating the Industry Production Function. *American Economic Review*, 58(4): 826–39.
5. Allen, E., & Seaman, J. (2008). Online nation: Five years of growth in online learning. The Sloan Consortium. Retrieved October 19, 2012 from http://sloan-c.org/publications/survey/online_nation
6. Asandului, L., Ceobanu, C. (2008). E-learning in Romanian Higher Education: A study case. *Turkish Online Journal of Distance Education-TOJDE*. ISSN 1302-6488 Volume: 9 Number: 3 Article 14
7. Asandului, L., Ceobanu, C. (2011). Use of Information and Communication Technologies in the instructors' activities of teaching and assessment of students – a Case Study. *Recent Researches in Educational Technologies*. ISBN: 978-1-61804-010-7
8. Baumol, W.J., 1967. Macroeconomics of unbalanced growth: the anatomy of the urban crisis. *American Economic Review* 57, 415–426.
9. Beraldo S., Montolio D., Turati, G. *The Journal of Socio-Economics* 38 (2009) 946-956
10. Charnes, A., W.W. Cooper, and E. Rhodes, 1978, Measuring the efficiency of decision making units, *European Journal of Operational Research* 2, 429-444.
11. Cesaro, L., Marongiu, S., Arfini, F., Donati, M., Capelli, M., G. (2009). Methodology of Analysing competitiveness, efficiency and economy of scale. *Farm Accountancy Cost Estimation and Policy Analysis of European Agriculture*.
12. Chilingerian J (1995), "Evaluating Physician Efficiency in Hospitals: A Multivariate Analysis of Best Practices", *European Journal of Operational Research*, 80, 548- 574.
13. Clayton, T., Criscuolo, C., P. Goodridge, P., (2003): "E-Commerce and Firm Performance: An Assessment Using Multiple Survey Sources and Linked Data," *European Commission: Working Papers and Studies*.
14. Coelli, T., D.S. Prasada Rao and G.E. Battese (2001). *An Introduction to Efficiency and Productivity Analysis*, Kluwer Academic Publishers, Deventer, Netherlands.
15. Coppola M, Ozcan Y, Bogacki R (2003), "Evaluation of Performance of dental Providers on Posterior Restorations: Does Experience Matter? A Data Envelopment Analysis (DEA) Approach", *Journal of Medical Systems*, 27(5):447-458.
16. Curristine, T. (2005b), "Government Performance: Lessons and Challenges", *OECD Journal on Budgeting*, 5(1), pp. 127-151.
17. Debreu, G. (1951). The Coefficient of Resource Utilization. *Econometrica*, 19(3): 273–292. Fare, R., Grosskopf, S. and Lovell, C.A.K. (1985). *The Measurement of Efficiency of Production*. Boston: Kluwer Academic Publishers.
18. Dooren, Wouter van, Miekatrien Sterck and Geert Bouckaert (2007), "Institutional Drivers of Efficiency", internal literature review, *Public Governance and Territorial Development Directorate*, OECD, Paris.
19. Dronkers, J. and P. Robert (2004), "The Effectiveness of Public, Private Government dependent and Private Independent Schools: A Cross-national Analysis", *European University Institute*, Florence, Italy, www.iue.it.
20. Geys, Benny and Moesen, Wim, 2009, Measuring Local Government Technical (In)Efficiency An Application and Comparison of FDH, DEA, and Econometric Approaches, *Public Performance & Management Review*, Volume 32, Number 4, 499-513. <http://www.natpact.info/uploads/Chronic%20Care%20Compendium.pdf>
21. Greene, W. (1997). Frontier Production Functions. In *Handbook of Applied Econometrics*. Volume II: Microeconomics, M.H. Pesaran and Schmidt, P. (Eds.), Oxford: Blackwell.

22. Gupta, S., Verhoeven, M., 2001. The efficiency of government expenditure: experiences from Africa. *Journal of Policy Modeling* 23, 433–467.
23. Harvey, A. (2004). Building a learning library. *The British Journal of Administrative Management*, (3/4),26-27.
24. Kasman, Adnan, 2005, Efficiency and Scale Economies in Transition Economies : Evidence from Poland and the Czech Republic, *Emerging Markets Finance and Trade*, Volume 41, Number 2, 61-80.
25. Knowles, M. (1984). *Andragogy in action*. San Francisco: Jossey-Bass.
26. Koohang, A. & Harman, K. (2005). Open source: A metaphor for e-learning. *Informing Science: The International Journal of an Emerging Transdiscipline*, 8, 75-86. Available at <http://inform.nu/Articles/Vol8/v8p075-086Kooh.pdf>
27. Koopmans, T.C. (1951). An analysis of Production as Efficient Combination of Activities. In *Activity Analysis of Production and Allocation*, Koopmans, T.C., eds, Cowles Commission for Research in Economics, Monograph no. 13. New York.
28. Lin, L. C., Tseng, L (2005). Application of DEA and SFA on the Measurement of Operating Efficiencies for 27 International Container Ports. (2005). *Proceedings of the Eastern Asia Society for Transportation Studies*, Vol. 5, pp 592-607.
29. Miller, J.L. and Adam Jr, E.E. (1996), "Slack and performance in health care delivery", *International Journal of Quality and Reliability Management*, Vol. 13 No. 8, pp. 63-74.
30. Moallem, M. (2004). Distance learning and university effectiveness. [Review of Caroline Howard, et al.
31. Porter, L. (2003). ABCs of e-learning: Reaping the benefits and avoiding the pitfalls. *Research Library*, 50(2), 273.
32. Rajkumar, S., Swaroop, V. Public spending and outcomes: Does governance matter? *Journal of Development Economics* 86 (2008) 96-111
33. Rosko, M. D. (1990), "Measuring technical efficiency in health care organizations", *Journal of Medical Systems*, Vol. 14 No. 5, pp. 307-21.
34. Shephard, R.W. (1953). *Cost and Production Functions* (Princeton University Press)
35. Sherman, H. (1984), "Hospital efficiency measurement and evaluation: empirical test of a new technique", *Medical Care*, Vol. 22 No. 10, pp. 922-38.
36. Tanzi, V., Schuknecht, L., 2000. *Public spending in the 20th century: a global perspective*. Cambridge: Cambridge University Press.
37. Theodoridis, A.M., Psychoudakis, A. (2008). Efficiency Measurement in Greek Dairy Farms: Stochastic Frontier vs. Data Envelopment Analysis. *International Journal of Economic Sciences and Applied Research* 1 (2): 53-67
38. Timmer, C. Peter, (1971). Using a probabilistic frontier production function to measure technical efficiency, *Journal of Political Economy* 79, 776-794.
39. World Bank, 2003. *Making services work for poor people*. World development report 2003. World Bank, Washington, D.C.
40. Zhu, J. 2002, *Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets and DEA Excel Solver*, Kluwer Academic Publishers, Boston.

A MULTI-STATE MARKOV MODEL FOR PROJECTING HEALTH CARE SPENDING

Francisco Javier Blanco-Encomienda, PhD
University of Granada, Spain

Abstract:

Demographic developments have always been central to economic growth and welfare. One of the observed trends has been that people live longer and they also remain healthy for longer. Besides this change, regarding the population structure the birth rates have declined, causing the so-called 'progressive aging of the population'.

Since the aging process is important in the field of social protection, spending on health and long-term care for dependent older people is a first-order policy issue. In this respect, considerable attention has focused on projecting these health care expenditures; emphasis has been given to estimating the number of dependent older people so that government can direct resources and services efficiently and effectively. These projections can be obtained by a multi-state Markov model, which is presented in this paper.

Projections of dependency amongst the elderly provides the basis for assessing the future cost of associated health care and the necessary services in order to guarantee that requirements for the dependent population are suitably covered.

Key Words: Multi-state Markov model, health care, spending

1. Introduction

Estimating the expenditures on health care for dependent older people has been the subject of discussion since several years. In this direction, the proposed paper aims to address, on the one hand, the link between aging and dependency and, on the other hand, the methodology for projecting the future demand for health care in order to serve as a basis for economic planning, in the medium term, of the spending on required assistance.

Due to the progressive aging of the population the growth in the number of older people who need help to perform the basic activities of daily living is unstoppable. Therefore, public institutions must act and provide health and social services that satisfy the needs of this population group.

The phenomenon *Dependency* may affect the society in diverse and complex ways, especially regarding the social security systems. Besides the broadly discussed consequences on the pension insurance schemes, it is very important to determine the influence of dependency on the health care sector, since it will be a fundamental driver of health and long-term care expenditure in the coming decades.

According to Kunkel and Applebaum (1992), the aging of a population has profound effects on all aspects of the nation. From the marketing of products, to retirement and employment patterns and to social relationships between generations, aging of a population results in dramatic changes in the way a society functions. Although these demographic changes have significant effects on every aspect of life in the society, one of the greatest challenges faced by an aging society lies in its ability to provide health and social service care of high quality.

Providing high quality long-term care services for dependent people in a society in which the number of people requiring those services is increasing rapidly raises difficult public policy decisions. Debates about the best approaches to providing health and long-term care, and equally difficult issues concerning what level of resources a society can allocate to this care, have already become common in industrialized nations. Projections of the need for health and long-term care are fundamental for sound public policy. *"If our efforts to plan for our aging population are to become more intelligent*

and less feeble, the highest priority must be given to amassing data and making projections of health care needs" (Brody, Brock and Franklin, 1987).

2. Dependency

Advancing in the development and the individual autonomy of disabled and dependent people in order to get a true quality of life for themselves and their families corresponds to common sense. Since 2003 was proclaimed 'European Year of People with Disabilities' and 2007 'European Year of Equal Opportunities for All', this is an idea that has been and is being widely discussed and analyzed in different international areas (Blanco and Latorre, 2008).

In the following sections the issue of dependency is introduced; for this, firstly its concept is analyzed and then several measures adopted to address it are presented.

2.1. Concept of Dependency

As a phenomenon that affects a significant number of people in our society, Dependency (or Dependence) is of great interest to those who work and live with it, needing solid theoretical frameworks.

According to the expert group constituted in 1998 by the Council of Europe, dependency is: *"A state in which persons, by reason of lack or loss of physical, psychological or intellectual autonomy, require significant assistance or help in carrying out their usual day-to-day activities"* (p. 2).

From the above definition three basic factors that define the status of dependent people can be extracted:

- The existence of a physical, mental or intellectual limitation.
- The inability to perform autonomously the basic activities of daily living.
- The need for assistance or help from a third party.

The need for assistance derived from dependency situations is not a new issue. In every age there has been a part of the population that -because of age, illness or deficiency- has needed, more or less intensely, the attention of third parties in the development of the daily life. What has changed has been its size (due mainly to the growth in the number and proportion of older people), its social importance (no longer seen as an exclusively individual or family issue, but a problem affecting to society as a whole) and its nature (while representing a redelimitation of the objectives and functions of the Welfare State and involving new protection and funding commitments). For this reason, dependency has become both a social and familiar issue and, at the same time, opens a new field of intervention that tests the ability of society and its institutions to adapt to the new realities of social fragility (IMSERSO, 2005).

When speaking about dependency, three levels of severity are applied:

- *Degree I. Moderate dependency:* when the person needs help in order to perform various basic activities of daily living, at least once a day or when the person needs intermittent or limited support for his/her personal autonomy.
- *Degree II. Severe dependency:* when the person needs help in order to perform various basic activities of daily living two or three times a day, but he/she does not want the permanent support of a carer or when he/she needs extensive support for his/her personal autonomy.
- *Degree III. Total dependency:* when the person needs help in order to perform various basic activities of daily living several times a day or, due to his/her total loss of physical, mental, intellectual or sensorial autonomy, he/she needs the indispensable and continuous support of another person or when he/she needs generalised support for his/her personal autonomy.

Each of the degrees of dependency established above is classified into two levels, depending on the person's autonomy and on the intensity of care that is required.

2.2. Measures adopted to address dependency

The protection of people with disabilities or those who are dependent on others for their care has become the focus of increasing attention on an international scale. The United Nations' 1948

Universal Declaration of Human Rights, and the Council of Europe's *Convention for the Protection of Human Rights and Fundamental Freedoms* (1950) and *European Social Charter* (1961) were the first such international treaties to make explicit mention of people with disabilities and set out measures designed to achieve optimum support for their personal and professional wellbeing.

In the 1970s, the European Union began to focus its attention on improving the living conditions of people with disabilities, approving in 1974 the initial *Community Action Program for the Vocational Rehabilitation of Handicapped Persons*. This established a basis for cooperation between those entities responsible for this area, and outlined actions intended to establish and disseminate good practice in the field.

These texts, together with others promoted by the World Health Organization (WHO), were followed in 1982 by the United Nations' *World Program of Action Concerning Disabled Persons*. The UN also declared 1983-1992 the *Decade of Disabled Persons*, which was conceived as a vehicle for the World Program of Action. In more recent years, 2006 saw the UN introduce *The Convention on the Rights of Persons with Disabilities*, which aims to promote, protect, and ensure the full enjoyment of human rights by persons with disabilities, ensure their fundamental freedoms, and protect their innate dignity.

Also worthy of note is the current *Council of Europe Action Plan to Promote the Rights and Full Participation of People with Disabilities in Society: improving the quality of life of people with disabilities in Europe 2006-2015*. This plan considers that non-governmental support organisations – that is, those charities and voluntary groups that are exclusively devoted to helping people with disabilities – are perfectly competent and qualified to make policy in this area, and that they should therefore be consulted when making any decision that may have repercussions for the lives of the people they represent.

3. Link Between Aging And Dependency

The social development and the improvement in the quality of life that are taking place in our society are causing an increase in longevity. The extension of life expectancy is coming with a declining birth rate, hence the weight of the elderly population is increasing progressively.

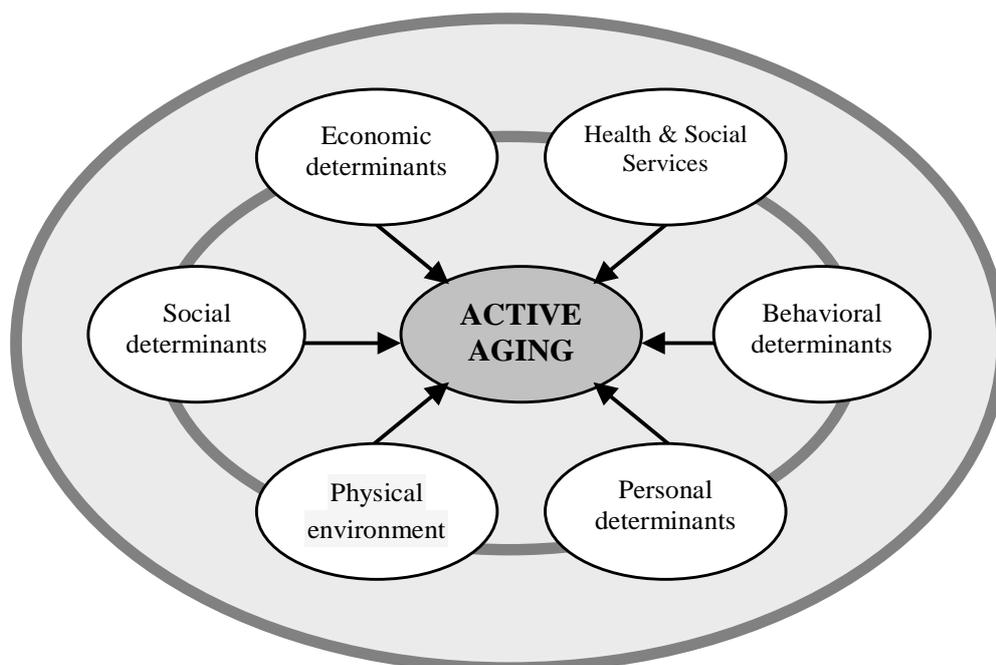
Dependency is strongly related to age. Dependency rates tend to be lowest at the youngest ages and highest for the elderly. This relationship is important because knowledge of a population structure and how this might change gives important information on the current number of dependent people and its possible trend. The strong relationship between age and dependency allows projections to be divided into demographic and dependency risk factors.

Although there have always been disabled and dependent persons (even not considered as such), it is true that the accelerated aging process that is currently experiencing the population, the high survival rates (due, largely, to medical advances) and the changes in family structure are causing social awareness for the protection of the seniors found in these conditions.

However, this aging process should not only be considered as a problem, but also as a social opportunity that must be seized. The aging population is the expression of an human achievement: live longer and live better. It constitutes one of the major social transformations produced in the last third of the last century: to advance in age and to do it actively.

Active aging is the optimizing process of opportunities for health, participation and security in order to improve the quality of life when people become older (WHO, 2002). It lets enhance physical, social and mental welfare throughout the entire life cycle of the individual. The term “active” refers not only to a physical capacity, but also a continued participation in the social, economic, cultural and civic one.

Specifically, the determinants that influence active aging are summarized in the following graphic:



Graphic 1. Determinants of the active aging

Aging is a challenge that has to be answered. It is a complex challenge because the increasing number of older people and the increasing care needs arising from this growth coincide with a crisis of informal support systems that have been responding to those needs, motivated by other two major social changes: the change in the family model and the increasing incorporation of women into the workforce (IMSERO, 2005).

The aging of the population is having a great impact in various sectors of the society: the pension system, the demand for technical assistance and social/health services, etc... For the elderly who, in varying degrees, suffer a loss of personal autonomy it is being especially significant the increase of the demand for long-term care (LTC) services. Long-term care is the personal assistance that enables dependent people to perform activities of daily living, such as eating, bathing and dressing. These services may be provided at home by family and friends, through home and community-based services such as home health care, personal care and adult day care; or in institutional settings, such as nursing or residential care facilities.

Since the challenge of social protection for dependent people is unavoidable for social policies, it is necessary that health and social services systems face it up, satisfying the personal help needs of these citizens.

In this line, in the context of aging population, to study different issues about the forecasting of the older care has become a major objective for many countries of the Organisation for Economic Co-operation and Development (see Achterberg et al., 2010; Batljan and Lagergren, 2004; Haynes, Hill and Banks, 2010; OECD, 2005).

4. Projecting Methodology

In order to estimate the cost of the health care that will be required by dependent older people, it is previously necessary to project the dependent elderly population to the year for which the forecast wants to be made. For this, there are different methods, being one of the most recurrent the projection of the prevalence rates (Siegel, 2002).

Under this method, the projected prevalence rates are applied to the total projected population. Prevalence rates indicate the proportion of people of all ages who have some kind of dependency at a given time respect to the total population.

$$D(x, i, t) = N(x, t) \cdot P(x, i, t)$$

where:

- $D(x, i, t)$ is the number of persons aged x , with a level of dependency i at time t .
- $N(x, t)$ is the total projected population aged x at time t .
- $P(x, i, t)$ is the dependency prevalence rate of level i , aged x and projected at time t .

There are two scenarios that may arise regarding the projected prevalence rates: to remain them unchanged over time (static prevalence rates) or to vary them (dynamic prevalence rates).

Although this method has been widely used (Giles, Cameron and Crotty, 2003; Harwood, Sayer and Hirschfeld, 2004; Lee and Miller, 2002; Schulz, Leidl and König, 2004), incorporating different levels of severity and with the purpose of knowing more about the transitions that may occur between several states and of relating the dependency with other factors -such as mortality-, it is necessary a methodology that, while still based on prevalence rates, is more complex and responds to issues like indicated above. Thus appear macrosimulation methodologies, such as multi-state models.

4.1. Multi-state Markov model

Multi-state models have been extensively applied in the social sciences, in particular to the analysis of longitudinal data. Hence, given the remarkable advances in statistics, these models are used in various fields such as medicine, biology, physics, economics...

A multi-state model is defined as a model for a stochastic process, which at any time occupies one of a set of discrete states. In medicine, the states can describe conditions like healthy, diseased, diseased with complication and dead. A change of state is called a transition, or an event. This then corresponds to outbreak of disease, occurrence of complication and death. It is important to recognize the difference between an event (like death) and a state (like dead). The state structure specifies the states and which transitions from state to state are possible. The full statistical model specifies the state structure and the form of the hazard function (intensity function) for each possible transition (Hougaard, 1999).

The greatest utility of these models in dealing with the issue of dependency is the possibility of projecting the number of persons who will be in a situation of dependency based on transition probabilities or rates between states, that is, if a healthy person becomes dependent, if a dependent person moves to another level of severity or whether a person, regardless of his/her health, dies.

Specifically, the projection model to be used would require the following tasks:

1. Generating baseline estimates of the level of dependency of the current older population.
2. Determining the transition rates between states.
3. Formulating assumptions about the transition rates.
4. Projecting the number of older people with need for long-term care under different scenarios.

Although all tasks play a key role in the projection methodology, the second and the third ones need to be discussed further.

Regarding the second task, the ideal method would be to obtain the transition probabilities directly from the data available, having followed a number of people and observed if they had changed over a period; then, we would work with longitudinal information. However, normally the survey data allow to calculate the prevalence rates but not to know when a transition between states occurs.

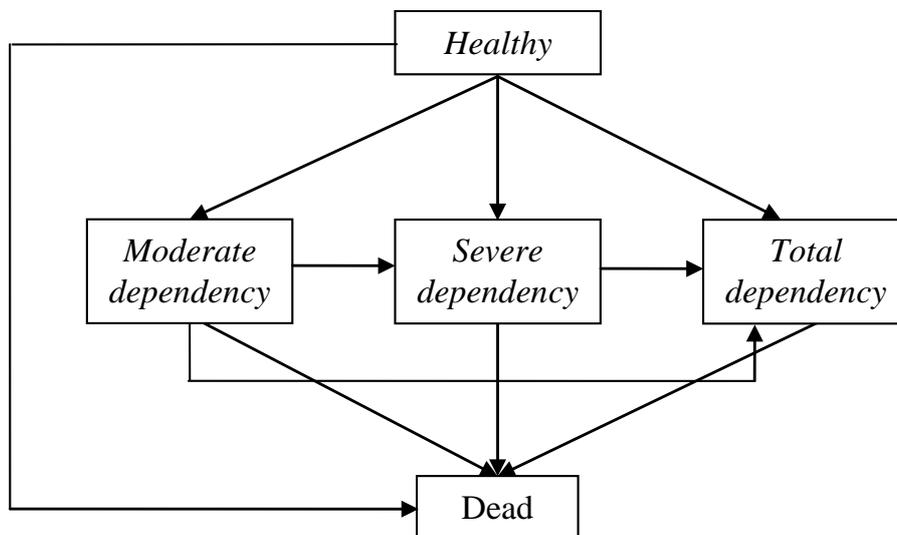
As posed Leung (2006), one possibility could be to compare the prevalence rates in two or more consecutive surveys and to calculate the maximum likelihood estimations of the probability of, after t years, a person aged x has undergone a transition from state i to state j , through the following expression:

$$P_x^{ij} = \frac{n_{x,x+t}^{ij}}{\sum_k n_{x,x+t}^{ik}}$$

where $n_{x,x+t}^{ij}$ is the number of people in state i , aged x at the moment of the first survey and in the state j , aged $x+t$ years in the date of the second survey.

The biggest disadvantage that usually appears when applying this method is that the surveys to be compared have been designed differently. Although in some cases both surveys can present common data, there are normally some conceptual differences, since different classifications of Disability, Dependency and Health could have been used. For this reason, the maximum likelihood estimation is not useful in this type of investigation. An alternative to calculate transition rates is the Markov model based on the method proposed by Sullivan (1971).

The state structure of the Dependency multi-state Markov model is shown in the following graphic:



Graphic 2. Transitions in the multi-state model

From these possible transitions it can be defined the following set: $\wp = \{a \rightarrow d_1, a \rightarrow d_2, a \rightarrow d_3, a \rightarrow m, d_1 \rightarrow d_2, d_1 \rightarrow d_3, d_1 \rightarrow m, d_2 \rightarrow d_3, d_2 \rightarrow m, d_3 \rightarrow m\}$, where the correspondence between the terms and the states is as follows:

- $a \rightarrow$ Healthy
- $d_1 \rightarrow$ Moderate dependency
- $d_2 \rightarrow$ Severe dependency
- $d_3 \rightarrow$ Total dependency
- $m \rightarrow$ Dead

We can define p_x^{ij} as the probability that a person aged x had undergone a transition from state i to state j during a year, being $i, j =$ healthy, moderate dependency, severe dependency, total dependency and dead.

Following the transition of a person aged x through a discrete-time process $\{S(x), x \in N\}$, the probabilities satisfy the Markov property:

$$\begin{aligned} \text{Prob}\{S(x+1) = s_{x+1} / S(x) = s_x \wedge S(x-1) = s_{x-1} \wedge \dots \wedge S(0) = s_0\} = \\ = \text{Prob}\{S(x+1) = s_{x+1} / S(x) = s_x\} \end{aligned}$$

According to Monteverde (2004), the Sullivan method presents two main advantages: first, the simplicity of its calculations and, second, the wide availability of the information required. Its main drawback is that the transitions between states are not observed; however, these can be estimated from the observed prevalence.

In relation to the third task, it would be necessary to make assumptions about the following transition probabilities:

- Probability that a healthy person becomes dependent.
- Probability that the health of a dependent person worsens and he/she moves to a more severe level of dependency.
- Probability of death of healthy people.

- Probability of death of dependent people. It is assumed that dependent people present an extra-mortality respect to the healthy persons.

For example, regarding to the first two transition probabilities, following the trend of recent years an improvement in the incidence of restrictions for the basic activities of daily living could be expected. In this case, we would be assuming a decrease in the number of people that move to the more severe dependency levels in two dimensions: first, fewer people move to a situation of dependency and then, less dependent people pass to more severe degrees.

Considering the different trends that transition rates may present, it is possible to determine several scenarios under which the projections can be generated. It would not deal a static scenario in which rates were assumed to remain unchanged with respect to the latest available, but a dynamic setting. In this case, different assumptions about rates may be established, assuming that each rate undergoes a change, distinguishing the way they do it: slightly, moderately or markedly.

Concretely, considering that improvements on health have the same effect on all transitions, we could establish some scenarios under which to carry out the projection of the dependent elderly population. Although for each scenario this group is expected to increase over the next years, the way in which the number of older dependents increases depends on the scenario: the more accused the assumed reductions in the components of deterioration are, the lower the projected number of dependents.

4.2. Projecting health care expenditures

Empirical evidence suggests that per-capita expenditure on health is higher for persons in situations of dependency and for those in their declining years. This expenditure increases in line with age, as the older people become, the more likely they are to find themselves in at least one of these categories (Bryant, Teasdale, Tobias, Cheung and McHugh, 2004). This, combined with the increasingly aging population, suggests that demand for care amongst the elderly dependent will become a socio-economic issue of growing concern.

To estimate the cost of care for the elderly dependent, we can draw on the number of dependents, together with the unit cost of services that may be used, such as home helps, day care centers, residential care homes and telephone-based remote care services ('telecare'). For each service, estimated cost could be calculated as follows:

- *Number of users of the service*: the projected population multiplied by the percentage of the population expected to use the service.
- *Units of service*: The number of users of the service multiplied by the units of service used by each user.
- *Total costs*: Units of service used by the entire population multiplied by the expected cost per unit of service.

On the other side, there are many different ways to assign care services to each degree of severity for the elderly dependent. For example, one of the options may only consider home care. Another alternative assignation of services, thinking that they are close to the real needs of the moderate, severe and total dependent population, may be the following: for the moderately dependent person, telecare and home help (one hour per day); for the severely dependent person, a place at a day centre and home help (one hour per day); and for the person who is totally dependent, a place in a residential home.

5. Conclusion

The aging population will be one of the most important social phenomena of the twenty-first century. It is projected that old population aged 65 and over will increase so much in the next decades and much faster than population as a whole. The most significant growth will be among the oldest seniors aged 85 or older, who have the greatest probability of losses in physical functioning. This surge will produce a similar increase in the demand for long-term care services. This trend, accompanied by the decline in informal care resources, resulted from reduced family size and increased proportion of women in the labor force, raises doubts about sustainability of the current

distribution of LTC financing and the incentives for the increased demand for long-term care without heightening budgetary strains.

For this reason, projecting the health care required by dependent elderly population and estimating their costs have become a subject of particular interest for researchers, since this provides the basis for creating predictive systems capable of identifying the context in which public and private policy needs to be shaped.

The financing framework will experience pressures in coming years not only as a result of the rising number of dependent elderly population but also because of other demographic trends, such as declining numbers of informal caregivers.

References:

- Achterberg, W. P. et al. Pain in European long-term care facilities: Cross-national study in Finland, Italy and the Netherlands. *Pain*, 148(1), 70-74, 2010.
- Batljan, I. and Lagergren, M. Inpatient/outpatient health care costs and remaining years of life-effect of decreasing mortality on future acute health care demand. *Social Science & Medicine*, 59(12), 2459-2466, 2004.
- Blanco, F. J. and Latorre, M. J. Promoción de la autonomía en situaciones de dependencia: el uso de nuevas tecnologías. *Docencia e Investigación*, 8, 1-10, 2008.
- Brody, J. A., Brock, O. B. and Franklin, T. Trends in the health of the elderly. *Annual Review of Public Health*, 8, 211-234, 1987.
- Bryant, B., Teasdale, A., Tobias, M., Cheung, J. and McHugh, M. *Population ageing and government health expenditures in New Zealand, 1951-2051*. Wellington: New Zealand Treasury, 2004.
- Council of Europe. *Recommendation No. R (98) 9 of the Committee of Ministers to member states on dependence*. Adopted by the Committee of Ministers on 18 September at the 641st meeting of the Ministers' Deputies, 1998.
- Giles, L. C., Cameron, I. D. and Crotty, M. Disability in older Australians: projections for 2006-2031. *Medical Journal of Australia*, 179(3), 130-133, 2003.
- Harwood, R. H., Sayer, A. A. and Hirschfeld, M. Current and future worldwide prevalence of dependency, its relationship to total population and dependency ratios. *Bulletin of the World Health Organization*, 82(4), 251-258, 2004.
- Haynes, P., Hill, M. and Banks, L. Older People's Family Contacts and Long-term Care Expenditure in OECD Countries: A Comparative Approach Using Qualitative Comparative Analysis. *Social Policy & Administration*, 44(1), 67-84, 2010.
- Hougaard, P. Multi-state Models: A Review. *Lifetime Data Analysis*, 5(3), 239-264, 1999.
- IMSERSO. *Atención a las personas en situación de dependencia en España. Libro Blanco*. Madrid: Ministerio de Trabajo y Asuntos Sociales, 2005.
- Kunkel, S. R. and Applebaum, R. A. Estimating the prevalence of long-term disability for an aging society. *J Gerontol.*, 47(5), 253-260, 1992.
- Lee, R. and Miller, T. An Approach to Forecasting Health Expenditures, with Application to the U.S. Medicare System. *Health Services Research*, 37(5), 1365-1386, 2002.
- Leung, E. A Multiple State Model for Pricing and Reserving Private Long Term Care Insurance Contracts in Australia. *Australian Actuarial Journal*, 12(2), 187-247, 2006.
- Monteverde, M. *Discapacidades de las personas mayores en España: prevalencia, duraciones e impacto sobre los costes de cuidados de larga duración*. Doctoral dissertation. Barcelona: University of Barcelona, 2004.
- OECD. *The OECD Health Project: Long-term care for older people*. Paris: Organisation for Economic Co-operation and Development Publications, 2005.
- Schulz, E., Leidl, R. and König, H. The impact of ageing on hospital care and long-term care. The example of Germany. *Health Policy*, 67(1), 57-74, 2004.
- Siegel, J. S. *Applied Demography: Applications to Business, Government, Law and Public Policy*. London: Academic Press, 2002.
- Sullivan, D. F. A single index of mortality and morbidity. *HSMHA Health Reports*, 86(4), 347-354, 1971.
- WHO. Envejecimiento activo: un marco político. *Revista Española de Geriatria y Gerontología*, 37(S2), 74-105, 2002.

SEEKING MEDICAL ATTENTION AMONG JORDANIAN PATIENTS WITH ACUTE MYOCARDIAL INFARCTION

Nidal F. Eshah, RN, CNS, PhD

Assistant Professor, Zarqa University, Faculty of Nursing, Zarqa, Jordan

Abstract:

Background: Significant reduction of mortality rate among acute myocardial infarction (AMI) patients could be achieved by early reperfusion. Successful reperfusion inversely related to the time from onset of AMI symptoms to treatment. Unfortunately, only about 25% of patients with AMI receive reperfusion treatment and the primary reasons for withholding this treatment was the time delay before admitting to hospital. In Jordan, as many other countries there is incomplete understanding for the dimensions of the delay problem in Jordanian community.

Purposes: To understanding factors leading to delay seeking treatment among Jordanian AMI patients. Specifically to explore the effects of sociodemographic, clinical, contextual and cognitive factors on delay; as well as identifying factors that predict delay among AMI Jordanian patients.

Methods: Comparative descriptive design was used; where convenient sampling technique used to recruit AMI patients within 24-72 hours of admission to coronary care unit. Eligible patients were invited to complete the modified Response to Symptoms Questionnaire.

Results: 150 patients were enrolled in the study. The mean delay time for Jordanian AMI patients was nine hours. Only 28% of patients were presented to the hospital within the first hour of symptoms experience. Six factors predicting patient's delay; those are: living in impoverished area, having low income, being alone at the time of symptoms onset, experienced intermittent symptoms, took medicine to relive symptoms, and worried about troubling others.

Conclusion: Jordanian AMI patients have a long delay time before they seek medical treatment. Sociodemographic, clinical, contextual, and cognitive factors were all found to be associated with delay seeking medical treatment. Findings highlighted the need to reduce the time to presentation among Jordanian AMI patients. Such reduction could be achieved by designing and implementing proper interventional programs, as well as improving public awareness of the appropriate responses to AMI symptoms.

Key Words: delay, Jordanian, myocardial infarction, responses, seeking medical attention

Introduction

Cardiovascular diseases (CVDs) considered the leading cause of death worldwide. Among CVDs, acute myocardial infarction (AMI) is the first leading cause of death in many countries (American Heart Association, 2009; Jordanian Ministry of Health, 2007). According to Robinson (1999), one third of AMI victims will not survive after that experience, and 60% of them die suddenly during the first hour after the onset of symptoms and before they reach the hospital (Chambless et al., 1997; Leslie, Fitzpatrick, & Morrison, 1996).

The care for AMI patients has shifted dramatically over the past decades. Remarkable achievements in reducing the mortality rates have been evolved through utilizing of thrombolytic therapy and primary percutaneous coronary angioplasty (PTCA) (Mckinley, Moser, & Dracup, 2000). The successful reperfusion by either thrombolytic therapy or primary PTCA inversely related to the time from symptoms onset to treatment and diminish markedly when the intervention is received more than six hours after the onset of AMI symptoms (Fibrinolytic Therapy Trialists Collaborative Group, 1994; Van't Hof et al., 1997). Actually, studies revealed that mortality rate of AMI cases can be reduced by 45% if thrombolytic is given in one hour of onset of symptoms, and by 25% if given within three hours of symptoms (Leslie, Urie, Hooper & Morrison, 2000).

Despite those benefits many patients who experience AMI do not receive either thrombolytics or primary PTCA (Zerwic, 1999), and only about 25% of patients with AMI receive thrombolytic treatment (Perry, 2001). The primary reasons for Withholding thrombolytic therapy was delay admitting to hospital more than six hours after the symptoms (Chareonthaitwee et al., 2000). Approximately one quarter to one half of all AMI patients delay longer than six hours in seeking medical treatment (Dracup & Moser, 1991; Schwarz, Schoberberger, Rieder, & Dunze, 1994).

Al-Hassan and Omran (2005) interviewed 83 Jordanian patients on the third day after AMI symptoms onset, to examine the health care seeking decision for myocardial infarction symptoms. They reported that 43% of the patients delayed the health care seeking decision more than one hour.

Till now there is incomplete understanding for treatment delay phenomenon among AMI Jordanian patients. Therefore, the current study could be a good resource to understand the delay phenomenon in a society such as Jordan. Additionally, it could guide health workers in designing and applying appropriate community interventions. The purposes of this study are: (1) to compare between the delayers and non- delayers Jordanian AMI patients based on the sociodemographic, clinical, contextual, and cognitive factors. (2) Identifying the most significant predictors of Jordanian AMI patients' delay. (3) Identifying the factors that are associated with early hospital presentation among AMI Jordanian patients.

Research Questions

1. What are the differences in sociodemographic, clinical, contextual, and cognitive factors between the delayers and non- delayers AMI Jordanian patients?
2. What are the most significant predictors of patient's delay to seek medical treatment in Jordanian AMI patients?
3. What are the sociodemographic, clinical, contextual, and cognitive factors that are associated with early hospital presentation among Jordanian AMI patients?

METHODOLOGY

Design

Comparative descriptive design was used to examine and describe differences in variables between the two groups (delayers and non-delayers patients).

Sample and Sampling Criteria

Power analysis indicated that 128 subjects would be sufficient to detect the significant differences between the two groups, with 80% power to pick up a medium effect size for two tailed t-test with a P value equal to 0.05 (Cohen, 1987). Convenient sampling technique was utilized to recruit 150 patients to participate in this study.

The target population of the current study was Jordanian patients who were diagnosed with AMI and admitted to the coronary care units (CCU), in two major hospitals in Amman. The emergency department physician made the diagnosis of AMI if two of the following criteria were present: chest pain lasting more than 20 minutes; serial electrocardiographic (ECG) evidence of AMI; or significant elevation of serum Creatin Kinase-MB.

Patients were enrolled in the study if they were diagnosed with AMI, oriented, free from neurological and mental disorders, experienced out hospital AMI, free from malignancy, and have the Jordanian nationality.

Measurements

In the current study; the modified Response to Symptoms Questionnaire (RSQ) were used. The first version of RSQ was developed by Burnett and her colleagues in 1995 (Burnett et al., 1995). The RSQ consisted of 18 items that provides information about patient delay and factors contributing to delay in six domains: (1) the context in which the AMI symptoms first appeared (i.e. where patient was when symptoms began, day of week, time of day, whom patient was with). (2) The antecedents of symptom onset (i.e. what patient was doing when symptoms occurred, how expected to anticipate the symptoms were, the level of emotional stress the patient was under). (3) Behavioral responses to the symptoms (i.e. first thing the patient did when symptoms were noticed, ease in reaching the doctor, get difficulty in transportation to the hospital). (4) Affective and emotional response to the symptoms (i.e. how anxious or upset the patient felt, comfort in seeking medical assistance, severity of pain). (5) Cognitive responses to the symptoms (i.e. symptoms attribution, perceived seriousness of the symptoms, perception of ability to control over the symptoms). (6) The response of others to patient symptoms (e.g. behavioral, emotional responses of others) (Burnett et al., 1995).

Based on growing evidence that cognitive, emotional, symptoms appraisal and social factors may be more important determinants of delay than knowledge about appropriate response to symptoms; Dracup and Moser (1997) modified, and used the modified RSQ in their study. They added sub-items to assess additional cognitive (e.g. assessment of symptoms experienced, knowledge of AMI symptoms), emotional (e.g. fear concerning consequences of symptoms, embarrassment about seeking help), and social factors surrounding the patient's decision to seek care for AMI symptoms (e.g. not wanting to trouble others) (Dracup & Moser, 1997).

Both the original and modified instrument has content validity (Dracup & Moser, 1997; Mckinley et al., 2000). For the purposes of the current study, the questionnaire was translated to Arabic language. Then pilot study was conducted, and internal consistency reliability was assessed by using cronbach's alpha ($\alpha=0.78$).

Protocol of Data Collection

Between 24 and 72 hours of admission to the CCU, eligible patients were invited to complete the modified RSQ. Patients were asked if they wanted to complete the questionnaire by themselves, or if they wanted the researcher to read the questionnaire. Patients were hemodynamically stable and free from pain when they completed the questionnaire. The researcher by himself obtained sociodemographic and clinical data from the patients and medical records.

Ethical Considerations

Patients were informed that participation is voluntary, and assured that data will be use only for research purposes and this data will be treated confidentially; and their anonymity were assured, and they can withdraw from the study at any time.

Data Analysis

The statistical package of social science (SPSS) for windows statistical software package (15th version) was used for data analysis. Descriptive statistics were used to describe the sample, summarize the patients' answers to all questionnaire's elements. Patients then classified into two groups according to the delay time: non-delayers (presented to the hospital < 6 hours) and delayers (presented to the hospital ≥ 6 hours).

To answer the first research question, and based on the level of measurement; Mann-Whitney U, Kruskal-Wallis, and independent t-test were used to compare the sociodemographic, clinical, contextual, and cognitive factors between the two groups (delayers and non-delayers patients). Then significant variables ($P<0.05$) that associate with increasing delay time, were entered into forward stepwise logistic regression to identify the most significant predictors of patients' delay.

Depending on the delay time, early presented patients with delay time (≤ 1) hour where listed, and according to the level of measurement; Spearman and Pearson's correlation coefficient were used to test the association between the (sociodemographic, clinical, contextual, and cognitive factors) and the early hospital presentation.

Results

Hundred and fifty patients were participated in the study. Participants were predominantly males (75.3%), married (94%), reside in the city (93.5%), working (53.3%), not insured (68%). Most of them were older than 45 years old (69.3%), have low monthly income (88%), non-educated (43.3%). Thirty patients (20%) had a prior history of angina, and 19 (12.6%) had a previous myocardial infarction, 52 (34.6%) have hypertension, and 48 patients (32%) have diabetes mellitus (Table 1 & 2).

The mean delay time for Jordanian AMI patients participated in the study was (9.1) hours, were as the median delay time was (4) hours. Only 28% of patients were presented to the hospital within the first hour of symptoms onset, 58.7% presented <6 hours, and 41.3% presented ≥ 6 hours. Non-delayers (presented <6 hours) had a mean delay time about (1.8) hours, and median time about (1.7) hours. Delayers (presented ≥ 6 hours) had a mean delay time about (19.4) hours, and median time about (14.5) hours (Table 3).

Differences between the Delayers and Non- Delayers

Comparing the sociodemographic, and clinical factors between the two groups revealed that individuals arrived ≥ 6 hours; have less than 200 J.D. monthly income ($P=0.00$), less than 9 years of formal education ($P=0.00$), un insured ($P=0.0003$), and they were less than 25 years old ($P=0.005$), or more than 45 years old ($P=0.005$). Additionally, the delayers were residing in camps ($P=0.009$), retired from work ($P=0.02$), sought help in a governmental hospital ($P=0.04$), and have a greater

prevalence of diabetes ($P=0.0003$). Additionally, there were no significant differences between the delayers and non-delayers based on previous history of angina, myocardial infarction, heart failure, percutaneous transluminal coronary angioplasty (PTCA), coronary artery bypass graft (CABG), or family history. Only the history of previous cardiac catheterization was significantly different, and associated with non-delay ($P=0.03$).

There were no significant differences in delay times based on the setting where the symptoms first noted. However, there were significant differences based on patient's companionship when symptoms began. In which, delayers were alone when they noticed their symptom ($P=0.03$), and non-delayers were with their spouses when they noticed their symptoms ($P=0.006$). Moreover, There were significant differences in delay times between the two groups based on the reaction of bystanders to the patient's symptoms; in which some reactions increased delay time "they tried to comfort me" ($P=0.0001$), "they suggest I rest or take medicine" ($P=0.0002$). On the other side, reactions such as "they took me to the hospital" ($P=0.00$), or "they suggested I get medical help" ($P=0.03$) lead to decrease the delay time. Delay time were also increased in patients who never told anyone about their symptoms ($P=0.01$).

Based on the cognitive and emotional responses; the following responses were associated with increasing delay time: "pretended nothing was wrong" ($P=0.00$), "tried self-help remedy" ($P=0.0003$), took some medication ($P=0.049$), attributing symptoms to muscle pain ($P=0.0001$), flu or flu-like illness ($P=0.006$), or breathing problems ($P=0.03$), thought that they have the ability to control their symptoms ($P=0.01$), waited to see if symptoms would go away ($P=0.00$), their symptoms were intermittent rather than constant ($P=0.02$), or being worried about troubling others with a request for assistance ($P=0.00$). On the other side; attributed symptoms to the heart ($P=0.00$), patients who "transported themselves or had someone transport them to the hospital" ($P=0.00$), or "told someone about their symptoms" ($P=0.006$) had decreased delay time.

Predictors of Patient's Delay to Seek Medical Treatment

Stepwise logistic regression was used to identify the most significant predictors of patient's delay. Results revealed that among the sociodemographical factors, the most significant predictors were: living in impoverished area ($P=0.001$), and having low monthly income ($P=0.003$). Were as the most significant predictor among the contextual factors was being alone at the time of symptoms onset ($P=0.005$). Among the cognitive factors; the most significant predictors were; the intermittent nature of symptoms ($P=0.001$), took some medicine to relieve symptoms ($P=0.011$), and worried about troubling others by requesting help ($P=0.043$).

Factors Associated with Early Hospital Presentation

According to the level of measurement, Spearman and Pearson's correlation coefficients were used to identify the sociodemographic, clinical, contextual, and cognitive factors that associated with early hospital presentation. Results revealed moderate positive relationships between early hospital presentation and the age group ranged between 25 and 45 years old ($r = 0.34$, $P=0.02$), and requesting care in a private hospital ($r = 0.33$, $P=0.03$).

When testing the contextual factors; results revealed moderate negative relationship with experiencing AMI symptoms at home ($r = -0.41$, $P=0.007$), and strong positive relationship with being visited friends or relatives at the time of symptoms started ($r = 0.58$, $P=0.00$).

Regarding the cognitive factors, results revealed a moderate positive relationship between early presentation and told nearby person about the symptoms ($r = 0.43$, $P=0.004$), moderate negative relationship with perception of self-ability to control the symptoms ($r = -0.37$, $P= 0.01$), and a moderate positive relationship with waiting to see if symptoms would go away ($r = 0.35$, $P=0.02$), and strong positive relationship with attributing symptoms to the heart ($r = 0.57$, $P=0.00$).

Additionally, early hospital presentation were positively associated with attributing symptom to indigestion or stomach problems ($r = 0.43$, $P=0.004$), and negatively associated with symptom attribution to the heart ($r = -0.43$, $P=0.004$).

Discussion

Inconsistent with Al-Hassan and Omran (2005) who suggested that Jordanian AMI patients demonstrated rapid arrival to emergency department; patients in the current study arrived at the hospital after a mean delay time of (9.1) hours. Although median delay time was (4) hours, this time is long as compared to western studies (Burnett et al., 1995; Dracup & Moser, 1997; Mckinley et al., 2000; Trent et al., 1995). Jordanian AMI patients who delay seeking treatment were characterized by

having low income, low educational level, uninsured, either too young or too old, reside in camps, retired from work, sought help in the a governmental hospital, and have greater prevalence of diabetes. While had previous cardiac catheterization associated with being non-delayers.

Increasing delay among lower income patients is logically expected, and consistent with the findings of (Dracup et al., 1997; Mckinley et al., 2000; and yarzebski et al., 1994). Because there is no free medical services in Jordan; so such patients are unable to tolerate the expenses of medical treatment, so they search for available alternatives; such as taking some medications, herbs, or even deny the present of symptoms.

Consistent with Dracup et al. (1997), and Mckinley et al. (2000), lower educational level associated with longer delay time. Increased delay among patients who have low educational level might be attributed to the shortage in knowledge regarding the disease. Increased delay among the uninsured patients could be explained depending on the patient's economical condition. In the light of increasing the treatment cost in Jordan, the uninsured patients and those receiving low monthly income are unable to tolerate the hospital expenses.

Regarding the age group; the current study revealed that the delayers belonging to two different age groups; those are < 25, or > 45 years. Delay among the too young patients is attributed to the patient's previous healthy condition that negatively affected patient's decision to seek medical attention. Older age also make symptoms recognition more difficult. This difficulty in symptom recognition might be attributed to the presence of co-morbidities. This finding was also consistent with (Goff et al., 1999; Mckinley et al., 2000; Ruston & Clyton, 1998; Sheifer et al., 2000).

Delayers were alone when they noticed their symptoms, and this might be attributed to the confusion associated with sever pain, in addition to lack of external motivation to seek help. On the contrary, being with the spouse at the time of symptoms onset was noticed among the non-delayers, and this finding is inconsistent with Crumlish and Hand (1999), and Dracup et al. (1995). In Jordan, usually the relationship between spouses is very close and strong; so it could be that the appearance of symptoms in a spouse, leads to a high stressful condition in the other partner.

There were significant differences between the delayers and non-delayers based on the cognitive and emotional responses. The delayers were pretended nothing was wrong, tried self-help remedy, or took some medicine, and this is consistent with Perry et al. (2001). Delay might occur in those patients because they tried to get control over their symptoms before they seek help for those symptoms. On the contrary, patients who applied effective coping actions arrived sooner. The effective coping actions described by the patients in the current study were transport them selves or had someone transport them to the hospital, or told some one about their symptoms. The last result was consistent with (Crumlish & Hand, 1999; Dracup, Moser & Eisenberg, 1995).

Consistent with the findings of Dracup and Moser (1997) and Mckinley et al. (2000), taking the choice to wait and see if symptoms would go away, thought that they have very much ability to control their symptoms, worried about troubling others with a requesting for assistance, and having intermittent symptoms. In Jordan, the family relationships are very strong; so it's expected from each member to be worry about troubling others if he or she has any health problem. Unfortunatly; this condition negatively affected requesting early medical help.

Consistent with Burnett et al. (1995), Dracup et al. (1997) and Mckinely et al. (2000), non-delayers correctly attributed their symptoms to the heart, whereas the delayers attributed their symptoms to muscle pain, or breathing problems. Patients who believed that their symptoms were cardiac in origin, made true mental representation and labeled the experienced illness as AMI, so they sought medical care faster than those who attributed their symptoms to another origin.

Conclusions

Jordanian AMI patients have a long delay time (9.1 hours) before they seek medical treatment, this delay will decrease the effectiveness of thrombolytic therapy, leading to bad prognosis. Several factors were found to be associated with delay seeking medical treatment among Jordanian AMI patients; those factors are belonging to four main categories; sociodemographic, clinical, contextual, and cognitive factors. These factors have a great effect on patient's decision to seek medical treatment. Therefore, health care workers in Jordan need to apply proper interventions to increased public awareness about AMI and the appropriate response to it's related symptoms.

Limitations

Generalizability can be restricted by using convenience sampling, and calculation of time elapsed before hospital presentation was depending on retrospective recall by the patients themselves. Additionally, the study included AMI survivor, in which died patients might have completely different characteristics.

Implications and Recommendations

The present study reinforces the need to reduce the time to hospital presentation among Jordanian AMI patients. Such reduction could be obtained by proper application of public educational programs that highlight the correct contextual and cognitive responses. Public educational contents should include cardiac risk factors, symptoms of AMI, and benefits of early hospital presentation. Appropriate coping actions should be discussed with individuals and families, and the role of emergency medical system should be reinforced in Jordanian society.

References:

- Al-Hassan, M. Omran, S. (2005). The effects of health beliefs on health care-seeking decisions of Jordanian patients with myocardial infarction symptoms, *International Journal of Nursing Practice*, 11: 13-20.
- American Heart Association. (2009). Heart Disease and Stroke Statistics-2009 Update: A Report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee, *Circulation*, 119, e21-e181.
- Burnett, R. Blumenthal, J. Mark, D. Leimberger, J. Califf, R. (1995). Distinguishing between early and late Responders to symptoms of acute myocardial infarction, *American Journal of Cardiology*, 75: 1019-1022.
- Chambless, L. Keil, U. Dobson, A. Mahonen, M. Kulasmaa, K. Rajakangas, A. Lowel, H. Tunstall-pedoe, H. (1997). Population versus clinical view of case fatality from acute coronary heart disease. Results from the WHO. MONICA project 1985-1990, *Circulation*, 96: 3849-3859.
- Chareonthaitawee, P. Gibbons, R. Roberts, R. Christians, T. Burns, R. Yusuf, S. (2000). The impact of time to thrombolytic treatment on outcome in patients with acute myocardial infarction, *Heart*, 84: 124-148.
- Cohen, J. (1987). *Statistical power analysis for the behavioral sciences*, Hillsdale, Nj: Lawrence Erlbaum Associate.
- Crumlish, C. Hand, M. (1999). Reducing patient delay in seeking treatment for acute myocardial infarction, *Medsurg Nursing*, 8 (2): 77-89.
- Dracup, K. Moser, D. (1991). Treatment seeking behavior among those with symptoms and signs of acute myocardial infarction, *Heart & Lung*, 20: 570-575.
- Dracup, K. Moser, D. (1997). Beyond sociodemographics: Factors influencing the decision to seek treatment for symptoms of acute myocardial infarction, *Heart & Lung*, 26(4): 253-262.
- Dracup, K. Moser, D. Eisenberg, M. (1995). Causes of delay in seeking treatment for heart attack symptoms, *Social Science Medicine*, 40: 370-392.
- Fibrinolytic therapy trialists (FTT) Collaborative group. (1994). Indications for fibrinolytic therapy in suspected acute myocardial infarction: over view of early mortality and major morbidity from all randomized trials of more than 1000 patients, *Lancet*, 343: 311-322.
- Goff, D. Feldman, H. McGovern, P. (1999). Pre hospital delay in patients hospitalized with heart attack symptoms in the United State: The REACT trial. Rapid early action for coronary treatment (REACT) study group, *American Heart Journal*, 138: 1046-1057.
- Jordanian Ministry of Health, Directorate of Information and Research. (2007). Information and Research Newsletter: mortality data in Jordan 2007, (Issue 1, March). Retrieved April 2, 2012, from <http://www.moh.gov.jo>.
- Leslie, W. Fitzpatrick, B. Morrison, C. (1996). Out of hospital cardiac arrest due to coronary heart disease: a comparison of survival before and after the introduction of defibrillation in ambulances, *Heart*, 75: 195-199.
- Leslie, W. Urie, A. Hooper, J. Morrison, C. (2000). Delay in calling for help during myocardial infarction: reasons for the delay and subsequent pattern of accessing care, *Heart*, 84: 137-141.
- Mckinley, S. Moser, D. Dracup, K. (2000). Treatment-seeking behavior for acute myocardial infarction symptoms in North America and Australia, *Heart & Lung*, 29(4): 237-247.

- Perry, K. Petrke, K. Ellis, C. Horne, R. Morris, R. (2001). Symptoms expectation and delay in acute myocardial infarction patients, *Heart*, 86: 91-93.
- Robinson, A. (1999). Getting to the heart of denial, *American Journal of Nursing*, 99 (5): 38-43.
- Ruston, A. Clyton, J. (1998). Patient's action during their cardiac event: Qualitative study exploring differences and modifiable factors, *British Medical Journal*, 316 (7137): 1060-1065.
- Schwarz, B. Schoberger, R. Rieder, A. Dunze, M. (1994). Factors delaying treatment of acute myocardial infarction, *European Heart Journal*, 15: 1595-1598.
- Sheifer, S. Rathore, S. Gersh, B. Weinfurt, K. Oetgen, W. Breal, J. Schulman, K. (2000). Time to presentation with acute myocardial infarction in the elderly. Association with race, sex, and socioeconomic characteristics, *Circulation*, 102: 1651-1656.
- Trent, R. Rose, E. Adams, J. Jennings, K. Rawles, J. (1995). Delay between the onset of symptoms of acute myocardial infarction and seeking medical assistance is influenced by left ventricular function at presentation, *British Heart Journal*, 73: 125-128.
- Van't Hof, A. Zijlstra, F. DeBoer, N. Liem, A. Hoorntje, J. Suryapranata, H. (1997). Patency and re infarction in late entry myocardial infarction patients treated with reperfusion therapy, *Angelology*, 48: 215-222.
- Yarzebski, J. Goldberg, R. Gore, J. (1994). Temporal trends and factors associated with extend of delay to hospital arrival in patients with acute myocardial infarction: The Worcester heart attack study, *American Heart Journal*, 128: 255-263.
- Zerwic, J. (1999). Patient delay in seeking treatment for acute myocardial Infarction Symptoms, *Journal of Cardiovascular Nursing*, 13(3): 21-32.

Table 1: Patients' characteristics according to the sociodemographical variables

Variables	Delayers (n=62) n (%)	Non-Delayers (n=88) n (%)
Age		
< 25 years	3 (4.8)	2 (2.3)
25 - < 35 years	1 (1.6)	3 (3.4)
35 - < 45 years	8 (12.9)	29 (33)
≥ 45 years	50 (80.6)	54 (61.4)
Educational Level		
< 9 years	42 (67.7)	23 (26.1)
9 – 12 years	15 (24.2)	25 (28.4)
> 12 years	5 (8.1)	40 (45.5)
Treatment Expenses		
Insured	10 (16.1)	38 (43.2)
Non-insured	52 (83.9)	50 (56.8)
Monthly income		
< 200 J.D.	33 (53.2)	31 (35.2)
200 – 400 J.D.	29 (46.8)	39 (44.3)
> 400 – 600	-	10 (11.4)
> 600 J.D.	-	8 (9.1)
Employment status		
Working	27 (43.5)	53 (60.2)
Don't working	20 (32.3)	19 (21.6)
Retired	15 (24.2)	16 (18.2)
Residency area		
City	59 (95.2)	80 (92)
Village	-	07 (8)
Camp	03 (4.8)	-
Treating hospital		
Governmental	54 (87.1)	66 (75)
Private	08 (12.9)	22 (25)
Marital status		
Single	05(6.5)	06 (5.7)
Married	57 (93.5)	82 (94.3)
Gender		
Male	41 (66.1)	72 (81.8)
Female	21 (33.9)	16 (18.2)

Table 2: Patients' characteristics according to the clinical variables

Variables	Delayers (n=62) n (%)	Non-Delayers (n=88) n (%)
History of angina Negative Positive	51 (82.3) 11 (17.7)	69 (78.4) 19 (21.6)
History of myocardial infarction Negative Positive	52 (83.9) 10 (16.1)	79 (89.8) 9 (10.2)
History of heart failure Negative Positive	60 (96.8) 2 (3.2)	82 (93.2) 4(6.8)
Previous cardiac catheterization Negative Positive	61 (98.4) 1 (1.6)	79 (89.8) 7 (10.2)
History of hypertension Negative Positive	40 (64.2) 22 (35.5)	58 (65.9) 30 (34.1)
History of diabetes Negative Positive	32 (51.6) 30 (48.4)	70 (79.5) 18 (20.5)
Family history for cardiac diseases Positive Negative	19 (30.6) 43 (68.3)	23 (26.1) 65 (73.9)

Table 3: Patients' groups with mean and median delay time for each group.

	Non-delayers	Delayers	Early presented
N (%)	88 (58.7)	62 (41.3)	42 (28)
Mean delay time	1.8 hours	19.4 hours	0.85 hours
Median delay time	1.7 hours	14.5 hours	0.5 hours

COMPARISON BETWEEN DIFFERENT CORD BLOOD STEM CELL POPULATIONS IN EFFICIENCY OF TRANSDIFFERENTIATION INTO HEPATIC LINEAGE

Hala Gabr¹, Rania Zayed¹, Ahmed Rezk Elzayat²

1Clinical Pathology Department, 2 Obstetrics and Gynecology Department,
Faculty of Medicine, Cairo University

Abstract:

Background: Cord blood is established as a source of stem cells for hemopoietic reconstitution. Cord blood transplants have been performed for more than 20 years now. However, cord blood stem cells as a source for regenerative medicine is still under trial. The availability of cord blood and its banking facilities make it a very useful source of hepatocytes for support of endstage liver disease. Cord blood contains a number of stem cell subsets: CD34+, CD133+, and mesenchymal stem cells (MSCs).

Objectives: This study was conducted to compare between these subsets in hepatocyte transdifferentiation efficiency. Hepatocyte lineage commitment was evaluated by alpha-fetoprotein (AFP) expression and albumin synthesis.

Methods: Cord blood is assayed for viability. Magnetic separation was done for CD34+ve, CD133+ve populations, MSCs were separated by culture on plastic flasks. Each cell fraction (CD34+ve, CD133+ve and MSCs) was cultured in liquid culture containing hepatocyte growth factor for 7 days.

AFP expression was done using immunocytochemistry, albumin synthesis was measured in culture supernatant using microalbumin assay kit.

Results: All three populations showed hepatocyte transdifferentiation; although with varying percentages. There was no statistically significant difference in AFP expression with MSCs showing 31% positivity, CD133+ve 30% followed by CD34+ve showing 28.8%. Also, MSCs population showed the highest albumin synthesis levels, followed by CD34+ then CD133+ cells.

Conclusion: Induction of hepatocyte-like cells is possible with all three stem cell subsets of the cord blood. However, establishment of functional hepatic cells is higher in MSCs population.

Key Words: CD 34+, CD 133+, mesenchymal stem cells, umbilical cord blood, Hepatocyte differentiation

Introduction

Liver diseases have been increasing worldwide and are considered a leading cause of death due to the prevalence of viral induced and other intractable liver diseases as primary cirrhosis and primary sclerosing cholangitis (Kim et al., 2002; Tanikawa, 1992). Most liver diseases lead to hepatocyte dysfunction with the possibility of eventual organ failure (Sellamuthu et al., 2011).

Egypt has the highest incidence of hepatitis C virus (HCV) worldwide with estimated anti-HCV antibody prevalence of 14.7% and the number of chronically infected Egyptians 9.8%, according to the most recently published Egyptian Demographic Health Survey in 2009, which was a national probability sample of the resident Egyptian population (El-Zanaty and Way 2009). HCV is a major health problem in Egypt and is the most common cause of chronic hepatitis, liver cirrhosis, hepatocellular carcinoma, and liver transplantation in the country (Nguyen and Keeffe 2005; Abdel-Aziz et al., 2000).

The need for liver transplantation worldwide is always increasing which is hindered by the shortage of donated organs leading to critical condition. In addition, liver transplantation is associated with significant morbidity and mortality to both the donor and the recipient. Therefore, increased needs for developing alternative therapies for the treatment of liver disease especially end stage liver disease are emerging and necessary (Fox and Chowdhury 2004). Cellular therapies replacing the diseased hepatic cells by stem cells are the main approach in liver directed cell therapy (Sellamuthu et al., 2011).

Stem cells sources are the bone marrow of an adult person, the peripheral blood of an adult person and the umbilical cord blood (UCB) of a newborn baby (Sellamuthu et al., 2011). As a source of stem cells for regenerative medicine, UCB has certain advantages over the bone marrow and peripheral blood as UCB has a high concentration of highly proliferative stem cells, can be easily collected without any harm to the mother or the baby and have a low rate of infection with cytomegalovirus (McAdams et al.1996, Bromeyer 1995). In addition, UCB- derived cells are more primitive than bone marrow cells which makes it more suitable cell source for cell-based therapies, regenerative medicine and tissue engineering (Lee et al., 2010).

Several stem cell subsets are present in UCB, including; CD34+ cells, CD133+cells and mesenchymal stem cells (MSCs). CD34 has been used as a human hematopoietic stem marker and most colony forming cells are present within the CD34+ population (Holyoake and Alcorn, 1994; Sutherland and Keating, 1992). CD 133 expression defines very early subset of progenitor cells, early hematopoietic progenitor cells express CD133, which is not expressed after differentiation (Walter and Dimmeler, 2002; Yin et al., 1997). MSCs are multipotent progenitors capable of differentiating towards other cell types as adipocytes, osteocytes (Kern et al., 2006), and hepatic cells (Campard et al., 2008; Lee et al., 2004) with a low immunogenicity (Aggarwal and Pittenger, 2005).

In our study, we compared the transdifferentiation potential of each of these subsets into hepatic cells. Transdifferentiation to hepatic cells was assessed by alpha-fetoprotein (AFP) expression as AFP is one of the earliest markers for endodermal differentiation (Hammer et al., 1987), hepatic metabolic function was assessed by albumin secretion as albumin production is a specific test for the presence and metabolic activity of hepatocytes (Dunn et al., 1991).

Subjects and Methods

Ten umbilical cord blood samples were obtained with oral consent of the mother and the approval of Cairo University ethical committee. Samples were collected from the umbilical vein ex utero after spontaneous delivery of the placenta following full-term vaginal delivery. Ex utero collections are less invasive, and there is better control over technique. The cord blood was collected in a bag containing citrate phosphate dextrose anticoagulant. Mononuclear cells were isolated by centrifugation of cord blood over Ficoll-Hypaque density gradient (density 1.077, Biochrom KG, Berlin).

Isolated mononuclear cells (MNC) were assessed for viability using Trypan blue dye exclusion test. Mononuclear cells were mixed with Trypan blue dye and incubated at 37°C for 5 minutes. Two hundred cells were counted using a light microscope at low power. Cells not taking the dye were counted as viable, whereas cells taking the dye were considered nonviable.

MNC were then divided into three parts for magnetic separation of CD133+, CD34+ cells and MSCs separation by culture on plastic flasks.

Magnetic separation was done using The MiniMACS™ Separation System (Miltenyi Biotec, Germany). 1×10^6 MNC were suspended in a final volume of 80 μ L MACS (Miltenyi Biotech) buffer and labeled with 20 μ L of microbeads with FITC (fluorescein isothiocyanate) conjugated mouse anti-human CD34/ CD133 antibodies (QBEND/10). The cells were mixed well and incubated at 4°C for 15 min in dark. After incubation the cells were washed thrice with 500 μ L of MACS buffer by spinning at 300xg for 10 min. The cells were resuspended in 500 μ L of buffer and used for magnetic sorting. The column was washed with 500 μ L of MACS buffer. The magnetically labeled cells were passed through the column. The cells with magnetic microbeads are retained within the column and those that are unlabelled eluted out. The eluted fraction was collected as negative fraction. The column was washed thrice with 500 μ L of MACS buffer. Then the column was removed from the magnetic field. The retained cells in the column were firmly flushed out by applying pressure on the matrix of the column by a plunger supplied with the kit. These were the positive fractions which were washed twice with MACS buffer by spinning at 300xg for 5 min and resuspended in 500 μ L of MACS buffer. Both fractions, magnetic and non -magnetic, were completely recovered.

Fluorescence-activated cell sorting analysis of the separated cells fractions; CD133+, CD34+ was done. Tubes were prepared for analysis; were samples were mixed with fluorescein isothiocyanate-conjugated mouse monoclonal antibody against CD133 and CD34 (Dako, Glostrup, Denmark) and with appropriate isotype-matched control monoclonal antibody. Cells were incubated with monoclonal antibody for 30 min at 4°C, washed once with phosphate-buffered saline, and resuspended in a small volume of phosphate-buffered saline for analysis by means of fluorescence-

activated cell sorting (FACScan) flow cytometer (Coulter Epics Elite, Miami, FL). Forward and side scatter gates were established to exclude cell debris and clumps before analysis for expression of CD133 and CD34.

MSCs separation was done by culture of the separated mononuclear cells into 25-cm² flask in complete culture media. After 24 hours of culture, non-adherent cells were washed out. Fresh medium was replaced every three days. At 70% monolayer confluence, the MSCs were enzymatically detached using 0.25% trypsin-EDTA (GIBCO BRL Grand Island, NY, USA).

Hepatocyte differentiation; 5x10⁴ cells/mL of each cell fraction (CD34+ve, CD133+ve and MSC) were cultured in Dulbecco's modified Eagle's medium (GIBCO, Sigma, St. Louis, MO) supplemented with 20% fetal calf serum (GIBCO BRL Grand Island, NY), penicillin (10,000 units/mL), streptomycin (10 mg/mL; GIBCO BRL Grand Island, NY), and 20 ng/mL hepatocyte growth factor (HGF; R&D Systems GmbH, Wiesbaden-Nordenstadt, Deutschland). Samples were incubated at 37 °C in a 5% CO₂ atmosphere for 7 days.

Cytospins and immunocytochemistry; on the 7th day, cells were harvested from cultures. Cytospins were prepared by centrifugation of the cell suspension (400xg for 10 min) and cellular pellets were applied to a glass slide. For AFP immunocytochemical testing, the cells were fixed by dipping in absolute alcohol for 2 min. After slide rehydration, blockage of endogenous peroxidases was done with 3 % H₂O₂. To reduce nonspecific hydrophobic interactions between the primary antibodies and the fixed cells, the slides were incubated with a blocking solution (1:50 normal horse serum in phosphate-buffered saline). The slides were incubated with primary antibodies in incubation buffer over night at 4 °C (Biogenex, SanRamon, CA). Then incubated with secondary antibodies for 30 min (R&D Systems) followed by incubation with one to three drops of high sensitivity streptavidin-horseradish peroxidase (HRP) conjugate for 20 min (R&D Systems). One to five drops of diaminobenzidine chromogen solution (R&D Systems) were applied for 8 min (colored precipitate localizes to the sites of antigen expression as the chromogenic substrate was converted by HRP enzyme into insoluble end product). Slides were counterstained with nuclear counterstain hematoxylin (Sigma). Visualization of the stained cells was done under a microscope using a bright-field illumination. Quantitative evaluation of alpha fetoprotein-positive cells (AFP + cells) was done using Leica Qwin 500 image analyzer computer system (England). The numbers of AFP + ve cells were counted/HPF in ten fields of each culture specimen and the mean percentages were obtained.

Albumin concentration was determined in the culture supernatant using enzyme linked immunosorbent assay kit (DRG International Inc., USA) according to the manufacturer instructions.

Results

The viability of cells was estimated by using the Trypan blue dye exclusion test. The result obtained revealed that the viability of the cells was high. Mean cell viability after mononuclear cell separation was 98.8± 1.12 viable cells.

CD34+ and CD133+ cells were isolated from UCB MNC fractions by incubation with microbeads, followed by passage through Mini- MACS columns. Fluorescence-activated cell sorting analysis with anti-CD34 and CD133 antibodies was performed to determine the percentage purity of the positive fraction. Positive selection of CD133 cells yielded 52.67 ± 11.77 CD133+ cells, while positive selection of CD34 cells yielded 51.76 ± 3.04 CD34+ cells.

We examined the expression of hepatic protein markers; albumin and AFP. Albumin and AFP are liver cell functional markers used to determine and characterize hepatic cell population. The presence of albumin is a prominent feature of mature hepatocytes as the liver is the site for albumin synthesis (Sellamuthu et al., 2011).

After 7 days in culture media for hepatocyte differentiation, the cultured cells showed positive expression of AFP with varying percentages; although statistically non- significant (p >0.05). MSCs showed 31% ± 1.65 positivity, CD133+ cells 30% ± 2.21 followed by CD34+ cells showing 28.8% ± 2.91.

Fig (1): Mean Alfa fetoprotein Expression after 7 day Culture

	Alfa fetoprotein (%)	P-value
MSCs	31 ± 1.65	>0.05
CD34+ cells	28.8 ± 2.91	
CD133+ cells	30 ± 2.21	

Albumin secretion was detected in the culture supernatant at the 7th day culture; the mean albumin level was 0.51 mg/L \pm 0.016 in MSCs culture supernatant, 0.46 mg/L \pm 0.049 in CD34+ cells culture and 0.4 mg/L \pm 0.032 in CD133+ cells culture. MSCs showed the highest albumin level in the culture supernatant compared to CD34+ cells and CD133+ cells although statistically non-significant ($p > 0.05$).

Fig (2): Mean Microalbumin Level after 7 day Culture

	Microalbumin (mg/L)	P-value
MSCs	0.5 \pm 0.016	>0.05
CD34+ cells	0.46 \pm 0.049	
CD133+ cells	0.4 \pm 0.032	

Discussion

Umbilical cord blood (UCB) is a rich source of stem cells and progenitor cells, which makes it a target for extensive experimental and clinical trials. UCB derived cells being more primitive than BM- derived cells, are more suitable source for cell based therapies and regenerative medicine (Lee et al., 2010).

Several types of cells have been addressed in the UCB including; hematopoietic stem cells, mesenchymal stem cells (MSCs) (Lee et al., 2004a, Lee et al., 2004b), unrestricted somatic stem cells (Kögler et al., 2004), cord blood derived embryonic stem cells (McGuckin et al., 2004), and umbilical derived multipotent progenitor cells (Lee et al., 2007).

It has to be addressed, which cell population in the UCB is capable of differentiation into specific cell types. In our study we focused on hepatocyte differentiation, several UCB subsets were proposed by previous studies as being capable of differentiation into hepatic like cells. It has been reported that CD 45+ subpopulation of UCB cells were capable of generating hepatocytes (Ishikawa et al., 2003), CD34+ UCB cells differentiated into hepatocytes after transplantation into mouse recipient (Di Campli et al., 2006; wang et al., 2003, Danet et al., 2002). Also, MSCs isolated from UCB are capable of differentiation into hepatocyte like cells (Lee et al., 2004c). Also, other pluripotent somatic cells can be isolated from UCB and may be responsible for hepatocyte repopulation in the liver (Kögler et al., 2004).

In our work, we studied the differentiation potential of UCB stem cell subsets; CD34+ cells, CD133+cells and MSCs into hepatic cells, comparing their ability to produce albumin and AFP. Hepatic metabolic function was assessed by albumin secretion as albumin production is a specific test for the presence and metabolic activity of hepatocytes (Dunn et al., 1991). AFP is one of the earliest markers for endodermal differentiation (Hammer et., 1987), produced primarily by the visceral endoderm of the yolk sac, as well as by hepatoblasts and more differentiated fetal hepatic cells (Chen et al., 1997, Tilghman, 1985), while it is not expressed in all adult stem cells (Hong et al., 2005, Lee et al., 2004c).

Hepatocyte differentiation was induced by hepatocyte growth factor (HGF) which was considered one of the most hepatogenic-inducing functionality (Chivu et al., 2009). HGF plays an essential role in the development and regeneration of the liver (Wang et al., 2004). In previous studies, many different cytokines were studied in order to optimize conditions required for hepatic differentiation, and it was found that only fibroblast growth factor-4 (FGF-4) and HGF were able to promote hepatocyte differentiation. FGF-4 alone induced hepatocyte differentiation, however, the degree of differentiation measured by absence of immature markers such as AFP and cytokeratin-19 was higher when cells were also treated with HGF. FGF-4 is important in initial endoderm patterning and may play a role in endoderm specification (Wells and Melton, 2000) while HGF induces differentiation of hepatocytes that are not actively proliferating (Oh et al., 2000; Hamamoto et al., 1999; Yoon et al., 1999; Miyazaki et al., 1998). In a study by Tang et al., it was reported that HGF can promote the expansion of human umbilical cord blood stem cells and induce their differentiation into liver cells (Tang et al., 2006).

Albumin secretion and AFP expression were detected in the culture supernatant at the day 7 of culture; the mean albumin level was 0.51 mg/L \pm 0.016 in MSCs culture supernatant, 0.46 mg/L \pm 0.049 in CD34+ and 0.4 mg/L \pm 0.032 in CD133+ culture. Many other groups have studied

hepatocyte differentiation from UCB and albumin expression was detected one week after the beginning of culture (Tang et al., 2006, Teramoto et al., 2005). This showed the capability of albumin-positive cells derived from UBCs to proliferate into functional hepatocyte-like cells.

The cultured cells showed positive expression of AFP with varying percentages; MSCs showed 31% \pm 1.65 positivity, CD133+ cells 30% \pm 2.21 followed by CD34+ cells showing 28.8% \pm 2.91. In the study by Tang et al, low level expression of AFP was detectable by day 7 and remained detectable up to day 35 (Tang et al., 2006). AFP is an early developmental marker gene of hepatoblasts (Hammer et al., 1987), indicating successful mRNA translation into specific proteins.

Based on our results, we believe that UCB-MSCs have a strong potential for differentiating into hepatic lineage cells in vitro and may be a promising source of cell therapy in intractable liver diseases as they can be easily isolated and expanded in quantities relevant to clinical application and can be cryopreserved for long periods without losing their stem cell properties in addition to having a broad differentiation potential (Kim et al., 2004, Zuk et al., 2002, Reyes et al., 2001). UCB-MSCs have a major advantage in that it does not require invasive procedures that could be harmful to the donor and can be used universally across the HLA barrier. MSCs can repair injured tissue by differentiating into damaged cell types, secreting appropriate cytokines and growth factors, and undergoing cell fusion (Prockop et al., 2003, Spees et al., 2003, Terada et al., 2002). In addition, MSCs possess the unique ability to suppress immune responses, both in vitro (Krampera et al., 2003, Tse et al., 2003) and in vivo (Polchert et al., 2008, Zappia et al., 2005, Ortiz et al., 2003).

In a study by Burra et al, UCB-MSCs were transplanted in necrotic mouse liver. Cells were recruited in the injured tissue and were able to engraft the liver and to regulate the inflammatory process (Burra et al., 2012).

Although, several studies had reported the ability of UCB-MSCs to differentiate into hepatic like cells, the application in clinical practice needs more thorough information on cell behavior in vivo. More investigations are needed to study the potential of frozen/cryopreserved cells to undergo proper differentiation into required cell type and its engraftment and cellular reconstitution ability.

References:

- Abdel-Aziz F, Habib M, Mohamed MK, Abdel-Hamid M, Gamil F, Madkour S, et al. Hepatitis C virus (HCV) infection in a community in the Nile Delta: population description and HCV prevalence. *Hepatology* 2000, Jul; 32(1):111-5.
- Aggarwal S, Pittenger MF. Human mesenchymal stem cells modulate allogeneic immune cell responses. *Blood* 2005, 105:1815–1822.
- Bromeyer HE. Questions to be answered regarding umbilical cord blood hematopoietic stem and progenitor cells and their use in transplantation. *Transfusion* 1995, 35: 694–702.
- Burra P, Arcidiacono D, Bizzaro D, Chioato T, Di Liddo R, Banerjee A, et al. Systemic administration of a novel human umbilical cord mesenchymal stem cells population accelerates the resolution of acute liver injury. *BMC Gastroenterology* 2012, 12:88.
- Campard D, Lysy PA, Najimi M, Sokal EM. Native umbilical cord matrix stem cells express hepatic markers and differentiate into hepatocyte-like cells. *Gastroenterology* 2008, 134:833–848.
- Chen H, Egan JO, Chiu JF. Regulation and activities of alpha-fetoprotein. *Crit Rev Eukaryot Gene Expr* 1997, 7:11–41.
- Chivu M, Dima S, Stancu C, Dobrea C, Uscatescu V, Necula LG et al. In vitro hepatic differentiation of human bone marrow mesenchymal stem cells under differential exposure to liver-specific factors. *Transl Res* 2009, 154 (3):122–132.
- Danet GH, Luongo JL, Butler G, Lu MM, Tenner AJ, Simon MC, Bonnet DA. ClqRp defines a new human stem cell population with hematopoietic and hepatic potential. *Proc Natl Acad Sci USA* 2002, 99:10441–10445.
- Di Campli C, Piscaglia AC, Pierelli L, Rutella S, Bonanno G, Alison MR, et al. A human umbilical cord stem cell rescue therapy in a murine model of toxic liver injury. *Dig Liver Dis* 2004, 36:603–613.

- Dunn JC, Tompkins RG, Yarmush ML. Long-term in vitro function of adult hepatocytes in a collagen sandwich configuration. *Biotechnol Prog* 1991, 7:237–245.
- El-Zanaty F, Way A. *Egypt Demographic and Health Survey 2008*. Egyptian: Ministry of Health (El-Zanaty and Associates and Macro International, Cairo) 2009, pp 431.
- Fox IJ, Chowdhury JR. Hepatocyte transplantation. *Amer J Transplant* 2004, 4:7–13.
- Hamamoto R, Kamihira M, Iijima S. Growth and differentiation of cultured fetal hepatocytes isolated various developmental stages. *Biosci Biotechnol Biochem* 1999, 63:395–401.
- Hammer RE, Krumlauf R, Camper SA, Brinster RL and Tilghman SM. Diversity of alpha-fetoprotein gene expression in mice is generated by a combination of separate enhancer elements. *Science* 1987, 235:53–58.
- Holyoake TL, Alcorn MJ. CD34⁺ positive haemopoietic cells: biology and clinical applications. *Blood Rev.* 1994, 8:113–24.
- Hong SH, Gang EJ, Jeong JA, Ahn C, Hwang SH, Yang IH et al. In vitro differentiation of human umbilical cord blood-derived mesenchymal stem cells into hepatocyte-like cells. *Biochem Biophys Res Commun* 2005, 330:1153–1161.
- Ishikawa F, Drake CJ, Yang S, Fleming P, Minamiguchi H, Visconti RP et al. Transplanted human cord blood cells give rise to hepatocytes in engrafted mice. *Ann NY Sci* 2003, 996:174–185.
- Kern S, Eichler H, Stoeve J, Kluter H, Bieback K. Comparative analysis of mesenchymal stem cells from bone marrow, umbilical cord blood, or adipose tissue. *Stem Cells* 2006, 24:1294–1301.
- Kim JW, Kim SY, Prk SY, Kim YM, Kim JM, Lee MH et al. Mesenchymal progenitor cells in the human umbilical cord. *Ann Hematol* 2004, 83:733–738.
- Kim WR, Brown RS, Terrault NA and El-Serag H. Burden of liver disease in the United States: summary of a workshop. *Hepatology* 2002, 36:227–42.
- Koçgler G, Sensken S, Airey JA, Trapp T, Muschen M, Feldhahn N et al. A new human somatic stem cell from placental cord blood with intrinsic pluripotent differentiation potential. *J Exp Med.* 2004, 200:123–35.
- Krampera M, Glennie S, Dyson J, Scott D, Laylor R, Simpson E et al. Bone marrow mesenchymal stem cells inhibit the response of naive and memory antigen-specific T cells to their cognate peptide. *Blood* 2003, 101:3722–9.
- Lee KD, Kuo TK, Whang-Peng J, Chung YF, Lin CT, Chou SH et al. In vitro hepatic differentiation of human mesenchymal stem cells. *Hepatology* 2004 c, 40:1275–1284.
- Lee MW, Choi J, Yang MS, Moon YJ, Park JS, Kim HC et al. Mesenchymal stem cells from cryopreserved human umbilical cord blood. *Biochem Biophys Res Commun.* 2004b, 320:268–73.
- Lee MW, Jang IK, Yoo KH, Sung KH, Koo HH. Stem and progenitor cells in human umbilical cord blood. *Int J Hematol* 2010, 92:45–51.
- Lee MW, Moon YJ, Yang MS, Kim SK, Jang IK, Eom Y, Park JS, Kim HC, Song KY, Park SC, Lim HS, Kim YJ. Neural differentiation of novel multipotent progenitor cells from cryopreserved human umbilical cord blood. *Biochem Biophys Res Commun.* 2007;358:637–43.
- Lee OK, Kuo TK, Chen WM, Lee KD, Hsieh SL, Chen TH. Isolation of multi-potent mesenchymal stem cells from umbilical cord blood. *Blood* 2004a, 103:1669–75.
- McAdams TA, Miller WM, Papoutsakis ET. Hematopoietic cell culture therapies. *Trends Biotechnol* 1996, 14:388–396.
- McGuckin CP, Forraz N, Allouard Q, Pettengell R. Umbilical cord blood stem cells can expand hematopoietic and neuroglial progenitors in vitro. *Exp Cell Res.* 2004, 295:350–8.
- Miyazaki M, Mars WM, Runge D, Kim TH, Bowen WC, Michalopoulos GK. Phenobarbital suppresses growth and accelerates restoration of differentiation markers of primary culture rat hepatocytes in the chemically defined hepatocyte growth medium containing hepatocyte growth factor and epidermal growth factor. *Exp Cell Res* 1998, 241:445–457.
- Nguyen MH, Keeffe EB. Prevalence and treatment of hepatitis C virus genotypes 4, 5, and 6. *Clin Gastroenterol Hepatol.* 2005, Oct;3(10 Suppl 2):S97-S101.
- Oh SH, Miyazaki M, Kouchi H, Inoue Y, Sakaguchi M, Tsuji T et al. Hepatocyte growth factor induces differentiation of adult rat bone marrow cells into a hepatocyte lineage in vitro. *Biochem Biophys Res Commun* 2000, 279:500–504.

- Ortiz LA, Gambelli F, McBride C, Gaupp D, Baddoo M, Kaminski N, et al. Mesenchymal stem cell engraftment in lung is enhanced in response to bleomycin exposure and ameliorates its fibrotic effects. *Proc Natl Acad Sci USA*. 2003, 100:8407–11.
- Polchert D, Sobinsky J, Douglas G, Kidd M, Moadsiri A, Reina E et al. IFN- γ activation of mesenchymal stem cells for treatment and prevention of graft versus host disease. *Eur J Immunol*. 2008, 38:1745–55.
- Prockop DJ, Gregory CA, Spees JL. One strategy for cell and gene therapy: harnessing the power of adult stem cells to repair tissues. *Proc Natl Acad Sci USA*. 2003, 100(Suppl. 1):11917–23.
- Reyes M, Lund T, Lenvik T, Aguiar D, Koodie L, Verfaillie CM. Purification and ex vivo expansion of postnatal human marrow mesodermal progenitor cells. *Blood* 2001, 98:2615–2625.
- Sellamuthu S, Manikandan R, Thiagarajan R, Babu G, Dinesh D, Prabhu D et al. In vitro trans-differentiation of human umbilical cord derived hematopoietic stem cells into hepatocyte like cells using combination of growth factors for cell based therapy. *Cytotechnology* 2011, 63:259–268.
- Spees JL, Olson SD, Ylostalo J, Lynch PJ, Smith J, Perry A et al. Differentiation, cell fusion, and nuclear fusion during ex vivo repair of epithelium by human adult stem cells from bone marrow stroma. *Proc Natl Acad Sci USA*. 2003, 100:2397–402.
- Sutherland DR, Keating A. The CD34 antigen: structure, biology, and potential clinical applications. *J Hematother*. 1992, 1:115–29.
- Tang XP, Zhang M, Yang X, Chen LM, Zeng Y. Differentiation of human umbilical cord blood stem cells into hepatocytes in vivo and in vitro. *World J Gastroenterol* 2006, 12(25):4014–4019.
- Tanikawa K. Indication of liver transplantation for hepatocellular carcinoma in Japan. *Surg Today* 1992, 22:395–400.
- Terada N, Hamazaki T, Oka M, Hoki M, Mastalerz DM, Nakano Y et al. Bone marrow cells adopt the phenotype of other cells by spontaneous cell fusion. *Nature*. 2002, 416:542–5.
- Teramoto K, Asahina K, Kumashiro Y, Kakinuma S, Chinzei R, Shimizu-Saito K et al. Hepatocyte differentiation from embryonic stem cells and umbilical cord blood cells. *J Hepatobiliary Pancreat Surg* 2005, 12:196–202.
- Tilghman SM. The structure and regulation of the alpha-fetoprotein and albumin genes. *Oxf Surv Eukaryot Genes* 1985, 2:160–206.
- Tse WT, Pendleton JD, Beyer WM, Egalka MC, Guinan EC. Suppression of allogeneic T-cell proliferation by human marrow stromal cells: implications in transplantation. *Transplantation* 2003, 75:389–97.
- Walter DH, Dimmeler S. Endothelial progenitor cells: regulation and contribution to adult neovascularization. *Herz*. 2002, 27:579–588.
- Wang PP, Wang JH, Yan P, Hu MY, Lau GK, Fan ST et al. Expression of hepatocyte-like phenotypes in bone marrow stromal cells after HGF induction. *Biochem Biophys Res Commun* 2004, 320: 712–716.
- Wang X, Ge S, McNamara G, Hao QL, Crooks GM, Nolte JA. Albumin-expressing hepatocyte-like cells develop in the livers of immune-deficient mice that received transplants of highly purified human hematopoietic stem cells. *Blood* 2003, 101:4201–4208.
- Wells JM, Melton DA. Early mouse endoderm is patterned by soluble factors from adjacent germ layers. *Development* 2000, 127:1563–1572.
- Yin AH, Miraglia S, Zanjani ED, Almeida-Porada G, Ogawa M, Leary AG et al. AC133, a novel marker for human hematopoietic stem and progenitor cells. *Blood* 1997, 90:5002–5012.
- Yoon JH, Lee HV, Lee JS, Park JB, Kim CY. Development of a non-transformed human liver cell line with differentiated hepatocyte and urea-synthetic functions: applicable for bioartificial liver. *Int J Artif Organs* 1999, 22:769–777.
- Zappia E, Casazza S, Pedemonte E, Benvenuto F, Bonanni I, Gerdoni E et al. Mesenchymal stem cells ameliorate experimental autoimmune encephalomyelitis inducing T-cell anergy. *Blood* 2005, 106:1755–61.
- Zuk PA, Zhu M, Ashjian P, De Ugarte DA, Huang JI, Mizuno H et al. Human adipose tissue is a source of multipotent stem cells. *Mol Biol Cell* 2002, 13:4279–4275.

VERBAL AND NUMERICAL APTITUDE OF GRADUATE NURSING STUDENTS AT ENTRY LEVEL⁷

Maxie Andrade

Asst Professor, Manipal College of Nursing Manipal, Manipal University

Dr. Anice George

Professor and Dean Manipal College of Nursing Manipal, Manipal, Manipal University

Dr. Shreemathi Mayya

Associate Professor, Department of Biostatistics, Manipal University

Renee Furtado

Clinical Psychologist, St. Ignatius Institute of Health and Education Centre, Honnavara

Abstract:

Nurses apply numerical skills for various functions e.g. drug/fluid calculations, estimation of traction weight, nutritional requirement estimations, estimation of incidence and prevalence in epidemiological surveys etc. Nurses are involved in emergency care and handle a variety of written and verbal orders in relation to patient care by health care providers. Thus computational and vocabulary skills are very essential for nurses. This article presents the entry level verbal and arithmetic abilities of Indian graduate nursing students.

Key Words: Math aptitude, verbal aptitude, B. Sc. Nursing programme

Introduction:

Nursing was chosen as a vocation by service minded young ladies a century ago and it was the preferred career choice by young men and women a decade (2000-2010) ago. There is a significant growth in the number of nursing educational institutions and the number of nurses in India, however India still needs nurses.¹

An important function of nurse is education and counseling. To function effectively as educators and counselors, nurses require good communication skills especially verbal ability. Nurses communicate knowledge and thus are expected to transmit information and transform people under their care. Ineffective communication results in poor nurse patient communication.

Nurses apply numerical skills for various functions e.g. drug/fluid calculations, estimation of traction weight, nutritional requirement estimations, estimation of incidence and prevalence in epidemiological surveys etc. Nurses are involved in emergency care and handle a variety of written and verbal orders in relation to patient care by health care providers. A study by Ashby reported that medical surgical nurses with 13 to 15 years of clinical practice had the highest scores in the Bayne-Bindler Medication Calculation test whereas with 1-3 years of experience had the lowest. A study by Newton et al reported that as math aptitude decreased, the number of attempts to pass the medication calculation assessment increased.²

Math aptitude was a factor associated with completing a nursing program. Newton et al. discuss that students would also need to be assessed for their reading comprehension and understanding of the English language. When a student is unable to read a question correctly and or does not understand what needs to be extracted from it, then the likelihood of successfully performing a medication calculation decreases dramatically. Thus it is essential that nursing students possess numeric and verbal ability.¹

⁷ **Acknowledgement:** Sr. Susan, Principal St Ignatius Institute of nursing, Honnavara and the Management of Nitte USHA institute of Nursing sciences, Mangalore for facilitating the conduct of the study.

This article describes the verbal and numerical aptitude of graduate nursing students of a selected University in South Canara district of Karnataka, upon admission. The need for testing arose, as the researchers in their attempt to develop a tool on aptitude towards nursing found low scores in arithmetic ability, which was assessed using a subscale with eight items requiring ratio and multiplication skills.

Methodology: The study was approved by institutional review committee. The setting was selected by simple random sampling. Cluster sampling enrolled 92 participants. The permission to conduct the study was obtained from the Principal of the nursing institute and informed consent was obtained from the participants. The participants seated three feet apart and calculators were not permitted. The tests were administered as per guidelines and time was monitored by an assistant. The students were on the fourth day of their admission to the nursing programme.

The tool is part of the revised David's Battery of Differential Abilities, available from Psycom services. Verbal ability test had two parts. Part 1: vocabulary (15 items for 4 minutes) and Part II: Understanding proverbs (9 items for 3 min and 30 sec). Verbal ability was expected to figure success in school related and academic performances. It is a predictor of occupations involving reading and writing of reports and so on. This ability is assessed by taking the sum of a person's score in Part 1 and 2. Numerical ability is a predictor of occupations involving computational skills. It had 20 items involving the tasks of addition, subtraction, division, multiplication, squaring, dealing with fractions etc. and was administered for 5 min and 30 sec. The reported reliability of numerical ability was 0.82 (split half) and 0.79 (test-retest) and of verbal ability test was 0.72 (Split half) and 0.79 (test retest). No correction for guessing was made for both the tests. The norms for interpretation gender and age-wise were used for analysis.³

Results: The sample characteristics are described in table 1. Participants were unmarried females, with science background in their pre-university education programme. Of the 92 participants, 57 (62%) opted for nursing as their first choice and 75 (81.5%) wanted to become a nurse. Of the 75 who wanted to become a nurse, 65 were rural residents and of the 57 whose first choice was nursing, 49 were rural residents. Two of the rural and three of the urban participants did not respond to these two items.

The factors which influenced the choice of nursing career were: Job opportunity abroad (78.3%); interest to help others (70.7%); interest to be in medical field (47.8%); job opportunity in India (29.3%); financial problems (26.1%); influence from nurses in the family (18.5%); as a career to actualize self (16.3%); career guidance programme (13%) and parents' influence (10.9%).

Table 1: The sample characteristics (n=92)

Characteristics	f	%
Age in years: 17	28	30.4
18	63	68.5
20	01	1.1
Religion: Hindu	14	15.2
Muslim	01	1.1
Christian	74	80.4
Budhists	03	3.3
Residence: Rural	77	83.7
Urban	15	16.3

The mean and standard deviation of the obtained scores are presented in table 2. The raw scores were transformed into STEN (Standard ten) equivalent scores, and the distribution of STEN scores is presented in table 3.

Table 2: Description of obtained verbal and numerical ability scores (n=92)

Ability	Min score	Max score	Mean	SD
Verbal Ability Part	0	7	3.3	1.64

1				
Verbal Ability Part II	0	6	2.72	1.79
Verbal Ability Part I & II	1	13	5.84	2.48
Numerical Ability	3	16	7.99	2.75

A STEN score of 9 and 10 are good predictors of success involving that particular ability. The three students who were good in numerical ability, were rural residents, moreover two of these three had poor verbal ability. Of the 73, who had poor verbal ability, 44 (91.7%) had poor numerical ability and of the 41 who had moderate numerical ability, 27 (65.9%) had poor verbal ability.

Table 3: Verbal and numerical ability of B. Sc. Nursing students (n=92)

Verbal ability			
	STEN score	Frequency	Percentage
Poor	1-3	73	79.3
Average	4-7	19	20.7
Good	8-10	0	0
Numerical Ability			
Poor	1-3	48	52.2
Average	4-7	41	44.6
Good	8-10	03	03.3

Association of verbal ability and numerical ability scores revealed that both these abilities were associated with age (exact test $p = 0.001$). The total of part I and II of verbal ability was associated with the variables: nursing was my first choice (exact test $p = 0.021$) and I wanted to become a nurse (exact test $p=0.004$)

These participants were also tested for caring abilities with a researcher developed tool using Swanson's theory of caring, the reliability of which was 0.904 (Cronbach's alpha). There was no significant relationship found between verbal, numerical and caring abilities (verbal ability and caring: $r = -0.193$, $p=0.134$; numerical ability and caring: $r = 0.019$, $p=0.891$; verbal and numerical ability: $r = 0.170$, $p=0.142$)

Discussion: Nursing is an evolving profession. Technology equipped health care today focuses on quality orientation. Nursing students of today enjoy sparing their cognitive computational skills by using mechanical computational devices such as calculators or computers. Newton argues that nursing graduates to have the holistic knowledge necessary for safe and efficient nursing practice, must have not only math aptitude but also the ability to read at the level demanded by the clinical questions being posed. However the study on the contrary found that majority of students possessed poor verbal ability though half of the participants had average numerical ability.^{2,4}

English is taught as a subject in the first year of graduate nursing programme and Indian nursing council recommends to teach nursing programmes in English. The study finding indicates that nursing students might be finding it difficult to comprehend the basic sciences taught in the first year nursing programme. The findings support the observation on number of failures in the first year of nursing compared to subsequent years of study, in most of the Indian nursing institutions. The findings also justify an opinion expressed by one of the validators of the tool developed to measure caring: "English is not the primary language in India. Hence verbal ability and reading comprehension assessments in English may not be valid measures for Indian students." Thus there is a scope to study the growth of verbal ability since admission to completion of nursing programme or to correlate the verbal ability with academic success.⁵

Newton et al report that math aptitude is the predictor of academic success, however they also insist on assessment of reading comprehension ability to identify whether the math proficiency is related to conceptual/contextual factors or are purely mathematical. Newton's study found a moderately strong correlation ($r = 0.351$ $p < 0.001$) between passing medication calculation assessment and reading aptitude, however English aptitude was not related to passing the medication calculation assessment test. The numerical ability test used in this study had no much demands on reading comprehension, as the items printed symbols for mathematical operations and were self explanatory. Thus the test assessed purely computational ability and not contextual factors.⁴

The eligibility criteria for graduate nursing programme in India include, physical fitness and a score of a minimum 50% (45% for SC/ST) in English, Physics, Chemistry and Biology subjects of pre university course. Diploma nursing programmes used to admit interested entrants with arts, commerce or science background till 2012, however since the year 2013, science background is made mandatory. Multiple regression analysis used in a Chilean study on medical students showed that the parameters with better predictive value were high school grades, biology test and mathematics academic aptitude test. Verbal academic aptitude test did not have a predictive value.⁷

The factors which influenced the nursing career choice were similar to a study by the investigator. The job opportunity attracted young men and women. Whatever could be the influencing factor for choice of career, to function effectively in a career, one need to possess a few pre requisite abilities, which can be enhanced through training. The pre requisite skills of nursing are not assessed in India including, the mathematic ability. Mathematics or statistics scores of pre university programme are not valued during admission to nursing course wherein biology and English scores are. Possibly this is an area requiring a look by the Indian Nursing Council, considering the essentiality of numerical skills in nursing. Further studies are required to identify the other essential abilities for nursing in Indian context.⁶

The nation is focussing on nurses to be effective health care providers. To be in pace with the goals of correcting growing imbalances in illness wellness patterns, nurses need to be equipped with effective and efficient nursing skills. Thus nursing organizations should identify right candidates upon admission to nursing programmes and mould them as effective nurses. Less time taken for shaping these nurses will certainly contribute in greater folds to the nation in alleviating preventable disease load.

References:

- Andrade M, George A, Roch E J. (2012) Perceptions of health care consumers, deliverers and nurse educators on nurses, nursing practice and nursing education system. *International Journal of Nursing Education*. 4(1), 39-41.
- Newton S E, Harris M, Pittiglio L, Moore G. (2009). Nursing student math aptitude and success on a medication calculation assessment. *Nurse educator*, 34 (2), 80-83
- Vohra S. (1994) *Handbook for DBDA Revised*. New Delhi: Psycom services.
- Newton S E, Moore G, Harris M, Pittiglio L. The effect of context on nursing student math aptitude. *Journal of professional nursing*. 26 (6), 341-345.
- Indian nursing Council. Syllabus for B. Sc. Nursing programme.
- Andrade M. (2012) Choice of nursing career: Pleasure or pressure. *International Journal of Nursing Education*. 4(1); 42-44.
- Garcia R E. (1995) *Rev Med Chil*.123(1), 28-36.

PRODUCTION AND DEVELOPMENT OF NUTRACEUTICALS USING *BACILLUS SUBTILIS* NCIM 2708 UNDER SOLID STATE FERMENTATION BY RESPONSE SURFACE METHODOLOGY

Md Faruque Ahmad

Department of Clinical Nutrition, Faculty of Applied Medical Sciences,
Jazan University, Jazan, Kingdom of Saudi Arabia (KSA)

Syed Aamir Ashraf

ZR Azaz Ahmad Azad

Department of Food Technology, Faculty of Engineering and Interdisciplinary Sciences,
Hamdard University, New Delhi, India

Bibhu Prasad Panda

Department of Pharmaceutical Biotechnology, Faculty of Pharmacy,
Hamdard University, New Delhi, India

Abstract:

Nutraceuticals is a broad umbrella term used to describe any product derived from food sources that provides further health benefits in addition to the essential nutritional value found in foods intense interest the public towards worth full effect not only for medicinal but also as a nutraceutical products. This research aimed to maximize production of Menaquinone-7(MK-7) by solid state fermentation process. Menaquinone-7 (MK-7) was produced by solid state fermentation process of *Glycine max* and *Phaseolus vulgaris* in presence of a co-cultured by *Bacillus subtilis*. Three separate nutritional parameters were screened by each experiment for the production of MK-7 using Plackett–Burman experimental design. Parameters were optimized by Box–Behnken design of response surface methodology for the production of menaquinone-7 (MK-7). In present study, 16.39 $\mu\text{g g}^{-1}$ MK-7 produced of *Glycine max* in optimised formula in the medium contains sorbitol 51.22g kg^{-1} , urea- 3.40g kg^{-1} and MnCl_2 -0.20g kg^{-1} . With *Phaseolus vulgaris* 31.35 $\mu\text{g g}^{-1}$ of MK-7 produced per gram of the substrate in a optimised media containing glycerol-54.05g kg^{-1} , urea-3.40g kg^{-1} and soybean extract-15.00g kg^{-1} .

Keywords: Menaquinone-7, *Bacillus subtilis*, Response surface methodology, Placket Burman Design

Introduction

When we talk of vitamin K, naturally first thing which comes on our mind is coagulant vitamin and its role in blood clotting processes. Menaquinone-7 (MK-7) has laid us a new prospective of thinking, being a highly valuable member of the vitamin–K family. Vitamin K occurs in two forms: Phylloquinone (vitamin K₁), which is present in green plants; and menaquinone (vitamin K₂), which is produced by some intestinal bacteria (Briggs and Calloway 1979; Conly and Stein 1992; Bentley and Meganathan, 1982). MK-7 is a highly bioactive homologue of Vit-K (Sato et al. 2001) MK-7 having seven isoprene units, one of the analogue of Vit-K₂ is abundant in fermented soybean (natto) (Tsukamoto, 2004) and has significant effect on preventing osteoporosis and cardiovascular diseases besides its positive effects on blood coagulation. (Gast et al. 2009; Schurgers et al. 2007; Yamaguchi et al. 1999). MK-7 is a part of the family known as Vit-K₂, and is necessary for the synthesis of blood coagulation factor, the activation of proteins involved in the building of bones and inhibition of vascular calcification (Brug et al. 2011; Rheume et al. 2012). MK-7 has an anabolic effect on bone calcification in rat femoral tissue. Zinc has been shown to enhance the effect of MK-7 in increasing bone calcium content in vitro (Ehara et al. 1996). The combined administration of zinc and MK-7 was found to have a synergistic or additive enhancing effect on bone components in the femoral tissues of

female elderly rats. (Ma et al. 2001). MK-7 may be significant in preventing osteoporosis with increasing age. More recently, in an in vitro study, it has been demonstrated that MK-7 can directly stimulate calcification in femoral metaphyseal tissue obtained from normal rats. (Sato et al. 1996; Yamaguchi et al. 1999).

Materials and Methods

Microorganisms

The cultures of *Bacillus subtilis* NCIM 2708 were obtained from National Collection of Industrial microorganism (NCIM), National Chemical Laboratory (NCL), Pune, Maharashtra, India, was maintained on slants of starch agar peptone medium at 4°C, and sub-cultured in 30-day intervals.

Preparation of seed culture and fermentation

Spore suspension of *Bacillus subtilis* NCIM 2708 were prepared from actively growing slants. The suspension was inoculated to conical flask containing nutrient broth liquid media (0.5% peptone, 1.5% beef extract, 0.5% sodium chloride, quantity sufficient in 100ml of distilled water adjusted pH to $7.2 \pm .2$). These cultures were incubated at 37°C for 24 hrs in shaker incubator at 125 RPM using solid state fermentation.

Optimization of nutrient

Glycerol, Mannitol, Sorbitol, Maltose, Yeast, soybean extract, Urea, $(\text{NH}_4)_2\text{SO}_4$, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ and MnCl_2 constituents were selected for the study. The Plackett-Burman experimental design for eleven variables i.e. nine nutritional components (independent variables) and two dummy variables, were used to evaluate the relative importance of various nutrients for a higher production of menaquinone-7. For each nutrient variable, two different concentrations i.e. high (+) and a low (-), tested. Data analysis was carried out by the standard procedure of Plackett-Burman experimental design (Plackett & Burman 1946).

Three important nutrient parameters (Sorbitol, Urea and MnCl_2) as per Plackett-Burman experimental design were selected, when *Glycine max* was the solid substrate. In case of *Phaseolus vulgaris*, three important parameters (Glycerol, Urea and Soybean extract) as per Plackett-Burman experimental design were selected for the determination of optimal value for menaquinone-7 production. An experimental design of 17 runs containing 5 central points was made according to Box-Behnken's response surface design for selected three parameters using DESIGN EXPERT 7.1.3 software (Statease Inc., USA). The relative effects of two variables on MK-7 were identified from the contour and response surface plot. Optimum values of the parameter for maximum production of MK-7 were determined by the point prediction tool of the software.

Extraction and analysis of MK-7

Extraction process of MK-7 was carried by using acetonitrile. 10gm fermented beans were extracted with 25ml acetonitrile. The mixture is shaken for 10 minutes. Transferred the solution into centrifuge tube and did the centrifugation for 5 min at 6000 rpm to get the decant solvent. Now concentrate the sample over nitrogen chamber. Transfer the sample into vials for further analysis and till that keep in refrigerator.

MK-7 were estimated by HPLC (Shimadzu, Japan) using 250 mm x 4.6 mm ID Lichrosper® 100 C₁₈ column of 5 µm particle size and 20 µl loop injector in presence of mobile phase acetonitrile and methanol 1:1 v/v with the flow rate of 1ml min⁻¹ and detection was carried out by UV detector at 254 nm. The chromatogram was analyzed by HPLC software class-VP.

Results and discussion

To identify the concentration of key nutrients influencing production of MK-7 Plakette Burman Design was used. Medium components for substrate *Glycine max* (Sorbitol, Urea and MnCl_2), While in case of *Phaseolus vulgaris* (Glycerol, Urea and Soybean extract) screened by Plakette Burman Design. An experimental design of 17 runs containing 5 central points was made according to Box-Behnken response surface design for 3 selected media parameters for each MK-7. The individual and interactive effects of nutrients variables were studied by conducting the fermentation run (Tables 1 and 2). The response was measured in terms of MK-7 production. The results of experimental data and simulated values are listed (Tables 1 and 2). Data collected for MK-7 concentration in each run were analyzed using the software DESIGN EXPERT 7.1.6 and fitted into a multiple nonlinear regression model proposes following equation (in the coded factor) for MK-7.

(Substrate *Glycine max*) Menaquinone-7 ($\mu\text{g}/100\text{gm}$) =

$$1426 + 76.87 \times A + 144.32 \times B + 87.41 \times C - 20.57 \times A \times B - 1.4725 \times A \times C - 23.15 \times B \times C - 369.786 \times A^2 - 100.04 \times B^2 + 141.21C^2$$

Where A, B and C represent sorbitol, urea and MnCl_2 respectively in g kg^{-1} .

(Substrate *Phaseolus vulgaris*) Menaquinone-7 ($\mu\text{g}/100\text{gm}$) =

$$2481.15 + 89.08 \times A + 739.6525 \times B + 14.51 \times C + 46.9225 \times A \times B - 8.4425 \times A \times C + 120.6675 \times B \times C - 415.044 \times A^2 - 128.999 \times B^2 + 11.77125C^2$$

Where A, B and C represent glycerol, urea and soyabean extract respectively in g kg^{-1} .

The effects of all nutrient parameters on MK-7 production can be compared with the help of perturbation plots Fig. 1 (a, b). The lines in the graph represent influence and sensitivity of respective factor for MK-7 production. This model resulted in response surface graphs. The response surface plots of calculated model for MK-7 production are shown in Fig. 2 (a-f). The relative effect of medium components on MK-7 production was depicted in response surface graphs. The analysis of variance of regression for MK-7 production was summarized in (Table 3).

Response Surface Methodology (RSM) proved to be a powerful tool for optimizing menaquinone-7 production by *Bacillus subtilis* NCIM 2708. Three nutrients sorbitol, urea, MnCl_2 for *Glycine max* and glycerol, urea and soybean extract for *Phaseolus vulgaris* screened by Plackett-Burman Experimental design were optimized by Box-Behnken design of RSM with help of DESIGN EXPERT 7.1.6 software. From the studied nutrient variable different ingredients have different effects on MK-7 production. In case of *Phaseolus vulgaris* all the three parameters glycerol, urea and soybean extract were positively significant factors. The proposed model equation illustrate the interaction between two factors, from the equation it was found that urea is positively interacted with glycerol and soybean extract while glycerol negatively interacted with the soybean extract. In case of *Glycine max* all the three parameters sorbitol, urea and MnCl_2 were positively significant factors. The proposed model equation illustrates the interaction between two factors. From the equation it was found that sorbitol is negatively interacted with urea and MnCl_2 while urea is also negatively interacted with the MnCl_2 . In present study, in case of *Glycine max* $16.39\mu\text{g g}^{-1}$ of menaquinone-7 production was obtained in optimized medium under solid state fermentation using *Bacillus subtilis* NCIM 2708. In case of *Phaseolus vulgaris* $31.35\mu\text{g g}^{-1}$ of menaquinone-7 production was obtained in optimized medium under solid state fermentation using *Bacillus subtilis* NCIM 2708. Menaquinone-7 production can further be increased by optimization of process parameters.

References

1. Bentley R. and Meganathan R. (1982) Biosynthesis of vitamin K (menaquinone) in bacteria. *Microbiol Rev.* 46: 241–80.
2. Briggs G. and Calloway D. (1979) Bogert's Nutrition and Physical Fitness, 10th edn, W.B. Saunders Co, Philadelphia.
3. Conly J.M. and Stein K. (1992) The production of menaquinones (vitamin K2) by intestinal bacteria and their role in maintaining coagulation homeostasis. *Prog Food Nutr Sci.* 16: 307–343.
4. Ehara Y., Tahakashi H., Hanahisa Y. and Yamaguchi M. (1996) *Res Exp Med.* 196: 171-178.
5. Brug F., Bacchetti T., Principi F., Paolo G. and Tiano L.L. (2011) Olive oil supplemented with MK-7 significantly affects osteocalcin carboxylation. *Br J Nutr.* 1: 1-5.
6. Gast G.C.M., Roos N.M.D., Sluijs I., Bots M.L. and Beulens J.W., Geleijnse J.M., Witteman J.C., Grobbee D.E., Peeters P.H. and Van der Schouw Y.T. (2009). A high menaquinone intake reduces the incidence of coronary heart disease. *Nutr Metab Cardiovasc Dis.* 19: 504-510
7. Rheaume- Bleue K. (2012) Vitamin K2 and the calcium paradox: How a Little- Known Vitamin Could save Your Life, Wiley.
8. Ma Z.J., Igarashi A., Yamakawa K., Yamaguchi M. (2001) enhancing effect of zinc and vitamin k_2 (menaquinone-7) on bone component of femoral tissue. of female elderly rats. *J Health Sci.* 47: 40-45.

9. Plackett R.L., Burman J.P. (1946) The design of optimum multifactorial experiments. *Biometrika*. 33: 305-325.
10. Sato T., Yamada Y., Ohtani Y., Mitsui N., Murasawa H. and Araki S. (2001) Production of menaquinone (vitamin K₂)-7 by *Bacillus subtilis*. *J Biosci. Bioeng.* 91: 16-20.
11. Sato T., Isobe Y., Ehara Y., Yamaguchi M. (1996) Effect of mk-7 purified from the fermented soybean (natto) in the femoral metaphysis from rat femur (in japanese). *Vitamin*. 70: 317-322.
12. Schurgers L.J., Teunissen K.J., Hamulyak K., Knapen M.H. and H. Vik and Vermeer C. (2007) Vitamin Kcontaining dietary supplements: Comparison of synthetic vitamin K1 and natto-derived menaquinone-7. *Blood*. 109: 3279-83.
13. Yamaguchi M., Taguchi H., Gao Y.H., Igarashi A. and Tsukamoto Y. (1999) Effect of vitamin K2 (MK-7) in fermented soybean (*natto*) on bone loss in ovariectomized rats. *J Bone Miner Metab.* 17: 23-29.
14. Tsukamoto Y. (2004) Studies on action of menaquinone-7 in regulation of bone metabolism and its preventive role of osteoporosis, *Bio factors*. 22: 5-19.

Table 1 Box-Behnken design for production of MK-7 by *Glycine max* with result (actual and predicted)

Run	Sorbitol (g/l)	Urea (g/l)	MnCl ₂ (g/l)	Menaquinone-7 (µg /100gm)	
				Actual	Predicted
1	35.00	2.00	0.10	674.55	714.395
2	65.00	2.00	0.10	932.5	909.295
3	35.00	3.14	0.10	1021	1044.205
4	65.00	3.14	0.10	1196.64	1156.795
5	35.00	2.71	0.00	1034.75	1031.673
6	65.00	2.71	0.00	1128.39	1188.393
7	35.00	2.71	0.20	1269.42	1209.448
8	65.00	2.71	0.20	1357.17	1360.248
9	50.00	2.00	0.00	1249.05	1212.283
10	50.00	3.14	0.00	1567.37	1547.243
11	50.00	2.00	0.20	1413.29	1433.418
12	50.00	3.14	0.20	1639	1675.768
13	50.00	2.71	0.10	1426	1426
14	50.00	2.71	0.10	1426	1426
15	50.00	2.71	0.10	1426	1426
16	50.00	2.71	0.10	1426	1426
17	50.00	2.71	0.10	1426	1426

Table 2 Box-Behnken design for production of MK-7 by *Phaseolus vulgaris* with result (actual and predicted)

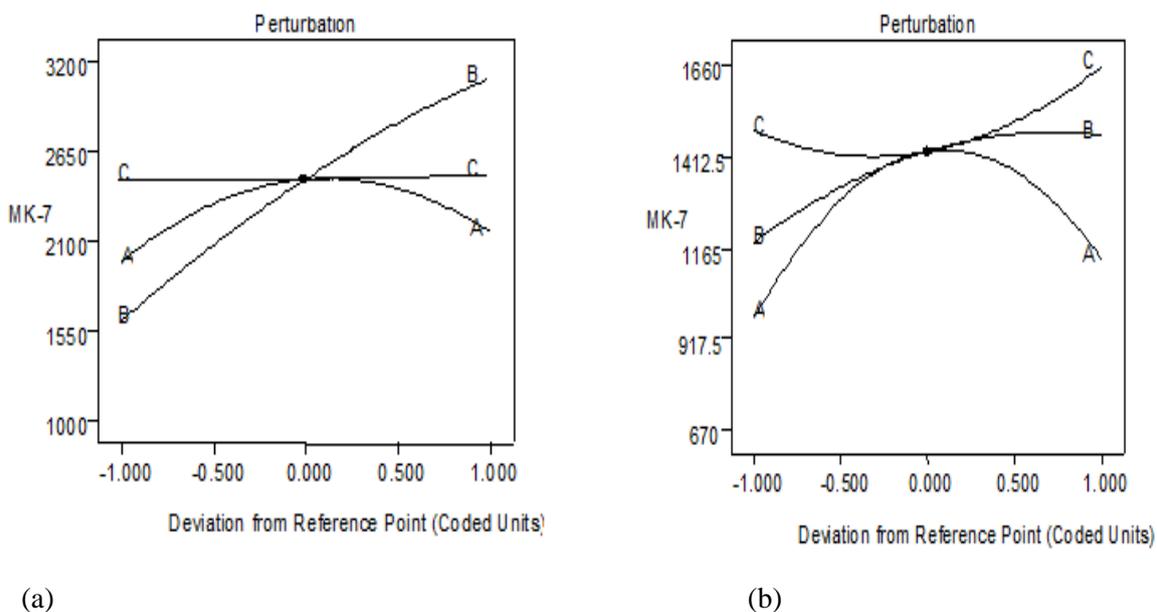
Run	Glycerol (g/l)	Urea (g/l)	Soybean Extract (g/l)	Menaquinone-7 (µg /100gm)	
				Actual	Predicted
1	35.00	2.00	10.00	1050.17	1155.298
2	65.00	2.00	10.00	1178.71	1239.613
3	35.00	3.40	10.00	2601.66	2540.758
4	65.00	3.40	10.00	2917.89	2812.763
5	35.00	2.70	5.00	1967.27	1965.845
6	65.00	2.70	5.00	2118.09	2160.89
7	35.00	2.70	15.00	2054.55	2011.75

8	65.00	2.70	15.00	2171.6	2173.025
9	50.00	2.00	5.00	1834.13	1730.428
10	50.00	3.40	5.00	2906.07	2968.398
11	50.00	2.00	15.00	1580.44	1518.113
12	50.00	3.40	15.00	3135.05	3238.753
13	50.00	2.70	10.00	2481.00	2481.45
14	50.00	2.70	10.00	2480.95	2482.15
15	50.00	2.70	10.00	2481.05	2482.15
16	50.00	2.70	10.00	2480.75	2478.15
17	50.00	2.70	10.00	2481.10	2481.19

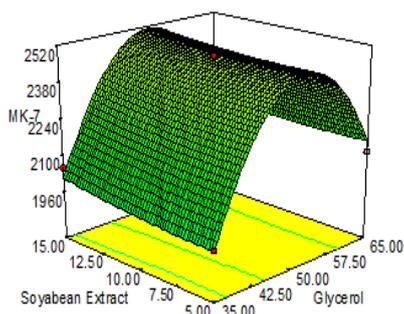
Table 3 Analysis of variance of model of nutrient parameters for MK-7 production

Parameters	<i>Glycine max</i> (MK-7)	<i>Phaseolus vulgaris</i> (MK-7)
Regression analysis of model		
Sum of squares	971075.4	5332011
Df	9	9
Mean squares	107897.3	592445.7
F-value	50.42	66.38854
*p-value	<0.0001	>0.0001
Regression analysis of Residual		
Sum of squares	14978.47	62467.4
Df	7	7
Mean squares	2139.78	8923.915
Correlation coefficient (R ²)	0.9848	0.96688

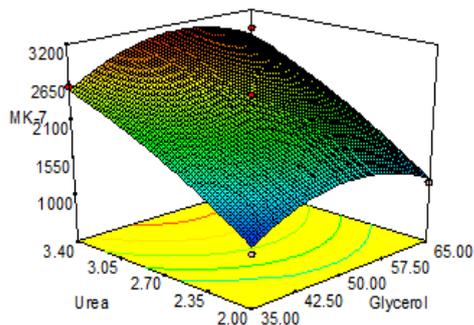
* Less than 0.0500 indicate model terms are significant.



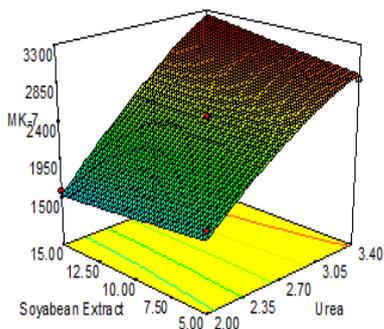
(a) (b)
Figure 1 Perturbation plot showing the effect of all nutrients parameter on MK-7 production ($\mu\text{g g}^{-1}$).
 Where A, B, and C represent glycerol, urea, and soyabean extract (Figure 1 a)
 and sorbitol, urea and MnCl_2 (Figure 1 b)



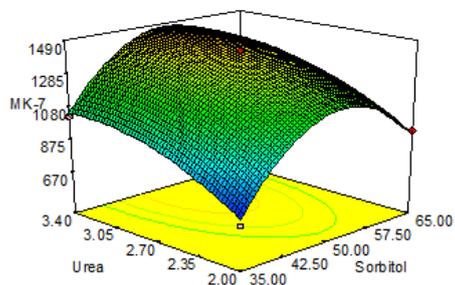
(a)



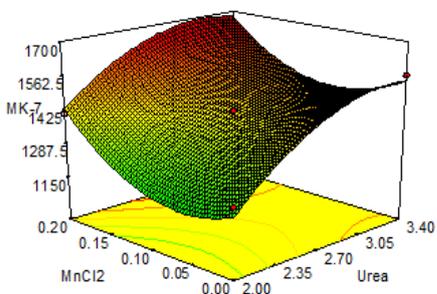
(b)



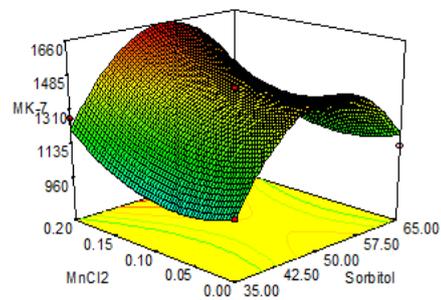
(c)



(d)



(e)



(f)

Figure 2 Response surface plots showing relative effect of two nutrient parameters for the production of MK-7 by *Phaseolus vulgaris* (a-c) and *Glycine max* (d-f) while keeping other nutrient at constant level.

SOLID LIPID NANOPARTICLES FOR TOPICAL DELIVERY OF MELOXICAM: DEVELOPMENT AND IN VITRO CHARACTERIZATION

Rawia M. Khalil

Department of Pharmaceutical Technology, National Research Center, Dokki, Cairo, Egypt

Ahmed Abd El-Bary

Faculty of Pharmacy, Cairo University, Cairo, Egypt

Mahfoz A. Kassem

Department of Pharmaceutical Technology, National Research Center, Dokki, Cairo, Egypt

Mamdouh M. Ghorab

Faculty of Pharmacy, Suez Canal University, Ismailia, Egypt

Mona B. Ahmed

Department of Pharmaceutical Technology, National Research Center, Dokki, Cairo, Egypt

Abstract:

Solid lipid nanoparticles (SLN) are colloidal carrier systems representing a promising approach as a drug delivery system for topical application. Therefore, the objective of this investigation was to develop Meloxicam Solid lipid nanoparticles (MLX SLNs) for topical delivery. The present study addresses the influence of different formulation compositions as lipid type and concentration in addition to surfactant concentration on the physicochemical properties and drug release profile of MLX SLNs. The nanoparticles were developed by modified high shear homogenization and ultrasonication technique using Geleol, Compritol 888 ATO or Precirol ATO 5 as solid core and poloxamer 188 as a surfactant. The results of the study revealed that MLX loaded SLNs showed extremely spherical shape having enriched core drug loading pattern with particle size (LD 90%) in the range of 325 to 1080 nm. A relatively high drug entrapment efficiency ranging from 61.94 to 85.33 % was obtained with zeta potential values lie between -17.6 to -38.6 mV indicating good stability. DSC examination revealed that MLX encapsulated in SLNs was in the amorphous state. According to the rheological study, all nanoparticulate systems exhibited non-Newtonian pseudoplastic flow with thixotropic behavior. In vitro release study showed a sustained release of MLX from the SLNs up to 48 h following Higuchi or zero order equations. Results of stability evaluation showed a long-term stability after storage at 4 °C for 12 months. In conclusion, SLNs with excellent physical stability, high entrapment efficiency and controlled drug release can be produced representing a promising carrier for topical delivery of meloxicam.

Key Words: Meloxicam, Solid Lipid Nanoparticles, Topical delivery, DSC, in vitro release study

Introduction

During the recent decades several studies have suggested that novel drug delivery systems based on lipid nanoparticles have the potential of increasing cutaneous drug delivery of both hydrophilic and lipophilic, compared to the other conventional vehicles (Mandawgade and Patravale, 2008; Liu et al., 2007a).

Solid lipid nanoparticles (SLNs) are colloidal carrier systems composed of a high melting point lipid/s as a solid core coated by surfactants (Mehnert and Mader, 2001; Wissing et al., 2004). Distinct advantages of solid lipid nanoparticles include negligible skin irritation, controlled release and protection of active substances (Jenning and Gohla, 2001; Mei et al. 2003; Jee et al., 2006; Muller et al., 2008). Because they are composed of physiologically tolerated, non-irritative and non-toxic lipids, SLNs seem to be well suited for use on inflamed and damaged skin (Muller et al., 2000). Moreover, the small size of the lipid particles ensures close contact to the stratum corneum increasing the amount of the drug penetrating into the mucosa or skin (Jenning et al., 2000a). These nanometer

sized particles with their solid lipid matrix may also allowing for sustained drug release (Cevc, 2004; Schafer-Korting et al., 2007).. After topical application, occlusive properties were also reported which decreases transepidermal water loss and favors the drug penetrating through the stratum corneum (Wissing and Muller, 2003; Muller et al., 2002; Wissing et al., 2001).

Meloxicam (MLX), a non-steroidal anti-inflammatory drug (NSAID), is a preferential inhibitor of cyclooxygenase-2 and has demonstrated potent analgesic and anti-inflammatory activity (Noble and Balfour, 1996). Considering the fact that most inflammatory diseases occur locally, topical application of MLX on the inflamed site can offer the advantage of delivering the drug directly producing its local effect. This occurs by avoiding gastric irritation, obtaining a substantial reduction of the systemic side effects, in addition to improvement of the patient compliance. In view of the characteristics of MLX including small oral dosage (7.5-15 mg/day), low molecular weight (354.1), lipid solubility and excellent tissue tolerability (Parfitt, 1999), it seems that there is a great need for investigating the MLX topical delivery system as an additional route for MLX administration. MLX performs very poorly in aqueous solubility and wettability, leading to difficulties in the design of pharmaceutical formulations. In order to overcome the formulation problems of MLX in addition to the barrier properties of the intact skin which limit the permeability of wide variety of pharmaceutical active agents, the development of a suitable vehicle system for optimum topical delivery of MLX is required.

Thus, the aim of the present study is to make the most benefits of solid lipid nanoparticles as drug delivery system for MLX through developing Meloxicam loaded Solid lipid nanoparticles (MLX SLNs) using high shear homogenization and ultrasonication method. Furthermore, the influence of some formulation variables on the characteristics of the MLX SLNs were also investigated.

Materials and methods

Materials

Meloxicam was supplied by Medical Union Pharmaceuticals, Abu-Sultan, Ismailia, Egypt. Geleol (glyceryl monostearate 40-55; 40-55 % monoglycerides - 30-45% diglyceride, m.p. 54.5-58.5 °C), Compritol 888 ATO (glyceryl behenate; 15-23 % monoglycerides - 40-60% diglyceride - 21-35% triglycerides, m.p. 69.0-74.0 °C) and Precirol ATO5 (glyceryl palmitostearate; 8-22 % monoglycerides - 40-60 % diglyceride - 25-35 % triglycerides m.p. 50-60 °C), kindly donated by Gattefossé, France. Poloxamer 188 (Pluronic[®] F68; a triblock copolymer of polyoxyethylene-polyoxypropylene), Methanol, Chromasolv[®] and Dialysis tubing cellulose membrane (molecular weight cut-off 12,000 g/mole) were purchased from Sigma Chemical Company, St. Louis, USA. All other chemicals and reagents used are of analytical grade.

Methods

Preparation of solid lipid nanoparticles

Solid lipid nanoparticles were prepared by a slight modification of the previously reported high shear homogenization and ultrasonication method (Mehnert and Mader, 2001; Venkateswarlu and Manjunath, 2004, 2005). Briefly, the lipid phase consisted of Geleol, Compritol or Precirol as solid lipid was melted 5 °C above the melting point of the lipid used and MLX (0.5% w/w) was dissolved therein to obtain a drug-lipid mixture. An aqueous phase consists of poloxamer 188 was heated up to the same temperature of the molten lipid phase. The hot lipid phase was poured onto the hot aqueous phase and homogenization was carried out at 25000 rpm for 5 minutes using Heidolph homogenizer (Heidolph Instruments, Germany). The resulted hot oil in water emulsion was sonicated for 30 minutes (Digital sonicator, MTI, USA). MLX loaded solid lipid nanoparticles were finally obtained by allowing hot nanoemulsion to cool to room temperature. Blank SLNs were prepared using the same procedure variables.

MLX entrapment efficiency

The entrapment efficiency percent (E.E. %), which corresponds to the percentage of MLX encapsulated within the nanoparticles, was determined by measuring the concentration of free MLX in the dispersion medium. The untrapped MLX was determined by adding 500 µl of MLX loaded nanoparticles to 9.5 ml methanol and then this dispersion was centrifuged at 9000 rpm (Union 32R, Hanil Science Industrial, Korea) for 30 minutes. The supernatant was filtered through millipore membrane filter (0.2 µm) and analyzed for unencapsulated MLX at 360 nm by using validated UV-spectrophotometric method (Shimadzu, model 2401/PC, Japan) after suitable dilution.

The E.E. % was calculated using the following equation (Hou et al., 2003; Souto et al., 2004):

$$\text{E.E. \%} = \frac{W_{\text{initial drug}} - W_{\text{free drug}}}{W_{\text{initial drug}}} \times 100$$

Where “ $W_{\text{initial drug}}$ ” is the mass of initial drug used and the “ $W_{\text{free drug}}$ ” is the mass of free drug detected in the supernatant after centrifugation of the aqueous dispersion.

Particle size analysis

Particle size analysis of MLX loaded nanoparticles was performed by Laser diffraction particle size analyzer (LD, Master sizer X, Malvern Instruments, UK) at 25°C. The LD data obtained were evaluated using volume distribution as diameter values of 10%, 50%, 90% and Span values. The diameter values indicate the percentage of particles possessing a diameter equal to or lower than the given value. The Span value is a statistical parameter useful to evaluate the particle size distribution, the lower the Span the narrower is the particle size distribution. It is calculated applying the following equation (Teeranachaideekul et al., 2007):

$$\text{Span} = \frac{\text{LD 90\%} - \text{LD 10\%}}{\text{LD 50\%}}$$

Zeta potential (ζ) and pH measurement

Zeta potential was measured in folded capillary cells using Laser Zetameter (Malvern Instruments, UK). Measurements were performed in distilled water adjusted with a solution of 0.1mM NaCl at 25 °C. The zeta potential values were calculated using the Smoluchowski equation.

The pH values of MLX lipid nanoparticles were measured at 25°C using digital pH meter (Jenway, UK).

Transmission Electron Microscopy (TEM)

The morphological examination MLX loaded SLNs was performed with TEM (model JEM-1230, Jeol, Tokyo, Japan). One drop of diluted sample was deposited on the surface of carbon coated copper grid and negatively stained with a drop of 2 % (w/w) aqueous solution of phosphotungstic acid for 30 s. Excess staining solution was wiped off by filter paper, leaving thin aqueous film on the surface. After being stained, samples were allowed to dry at room temperature for 10 minutes for investigation (Li et al., 2006).

Differential scanning calorimetry (DSC) analysis

DSC analysis was performed using Shimadzu Differential Scanning Calorimeter (DSC-50, Kyoto, Japan). About 10 mg sample was added in a 40 μ l aluminium pan which was sealed and heated in the range of 30-300°C at a heating rate of 10°C /min. An empty aluminium pan was used as reference standard. Analysis was carried out under nitrogen purge.

Rheological study

The rheological properties of the prepared lipid nanoparticles were measured using Brookfield’s viscometer (Brookfield LV-DV II+, USA). The sample (20 g) was placed in a beaker and allowed to equilibrate for 5 min. The measurements were carried at ambient temperature using the suitable spindle. The spindle speed rate was increased in ascending order from 1 to 100 rpm and then in descending order speed setting from 100 to 1 rpm with each kept constant for 10 seconds before a measurement was made.

In vitro release study

In vitro release of meloxicam was evaluated by the dialysis bag diffusion technique reported by Yang, et al. (Yang et al., 1999). The release studies of meloxicam from solid lipid nanoparticles were performed in phosphate buffer pH 5.5 and methanol (75: 25). The aqueous nanoparticulate dispersion equivalent to 2 mg of meloxicam was placed in a cellulose acetate dialysis bag and sealed at both ends. The dialysis bag was immersed in the receptor compartment containing 50 ml of dissolution medium, which was stirred in a water bath shaker at 100 rpm (Memmert GmbH, Germany) and maintained at $32 \pm 2^\circ\text{C}$. The receptor compartment was covered to prevent the evaporation of dissolution medium. A 2 ml sample of the receiver medium was withdrawn at predetermined time intervals (0.5, 1, 2, 3, 4, 5, 6, 8, 24 and 48 h) replaced by equivalent volume of fresh medium to maintain constant volume. The samples were analyzed for drug content spectrophotometrically at 360.5 nm. The data was analyzed using linear regression equations and the order of drug release from the different formulations was determined (zero order, first order or Higuchi diffusion model).

Effect of storage on particle size

MLX loaded SLN formulations were stored at 4°C for 12 months. Particle size was determined using Laser diffraction particle size analyzer (LD).

Statistical analysis

All experiments were repeated three times and data were expressed as the mean value \pm S.D. The statistical analysis of data was determined using one-way analysis of variance (ANOVA). Individual differences were evaluated using a nonparametric post hoc test. A difference of $p < 0.05$ was considered to be statistically significant.

Results and discussion**Preparation of Solid lipid nanoparticles**

SLNs have been prepared in various researches using different methods (Hu et al., 2002; Priano et al., 2007; Lva et al., 2009). In the present study, we had adopted an economical, simple and reproducible method for the preparation of SLN, i.e. homogenization followed by ultrasonication at above the melting point of the lipid (Fang et al., 2008; Chen et al., 2010). MLX loaded SLNs dispersions were composed of Geleol, Compritol 888 ATO or Precirol ATO 5 as core matrices used in different concentrations 5, 7.5 and 10 % (w/w). These lipid based carrier systems were stabilized by 0.5, 1, 2.5 and 5 % (w/w) Poloxamer 188. Meloxicam was incorporated at a constant concentration of 0.5% (w/w). The w/w percent composition of the investigated MLX SLNs is shown in table 1.

Table I. Composition of MLX SLNs (% w/w) of different lipids

Formula	Lipid		Surfactant conc. (Poloxamer 188) (%)
	Type	Conc. (%)	
SLN ₁	Geleol	5	0.5
SLN ₂			1
SLN ₃			2.5
SLN ₄			5
SLN ₅		7.5	0.5
SLN ₆		10	
SLN ₇	Compritol	5	0.5
SLN ₈			1
SLN ₉			2.5
SLN ₁₀			5
SLN ₁₁		7.5	0.5
SLN ₁₂		10	
SLN ₁₃	Precirol	5	0.5
SLN ₁₄			1
SLN ₁₅			2.5
SLN ₁₆			5
SLN ₁₇		7.5	0.5
SLN ₁₈		10	

MLX entrapment efficiency

The entrapment efficiencies of all SLN formulations are shown in table 2. The data clearly shows that all formulations possessed high entrapment efficiency (E.E. %) ranged from 61.94 ± 1.31 to 85.33 ± 1.07 %. The results might be related to the structure of the lipid which had a great influence on the capacity for drug incorporation. Lipids which form highly crystalline particles with a perfect lattice (e.g. monoacid triglycerides) lead to drug expulsion (Westesen et al., 1997). More complex lipids as Geleol, Compritol 888 ATO and Precirol ATO 5; being mixtures of mono-, di- and triglycerides form less perfect crystals with many imperfections offering space to accommodate the drugs (Muller et al., 2000). The data clearly shows that Geleol SLNs exhibited the lowest entrapment of meloxicam compared to Compritol and Precirol SLNs (Figs. 1 and 2). This can be attributed to the difference in composition and chain length of the three lipids used. The higher drug E.E % noticed with Compritol and Precirol was attributed to the high hydrophobicity due to the long chain fatty acids attached to the triglycerides resulting in increased accommodation of lipophilic drugs (Jenning and Ghola, 2000).

As observed in Fig. 1, at a constant amount of Poloxamer 188 (0.5 % w/w) increasing the lipid concentration from 5 to 7.5 % (w/w) resulted in a consequent increase in E.E. % ($p < 0.05$), while decreased upon further lipid increase to 10 % (w/w). This is in accordance with the study done by Abdelbary and Fahmy on diazepam-loaded SLN, in which they found that increasing the Compritol concentration to 10% (w/w) consequently resulted in a decrease in the amount of diazepam entrapped. A possible explanation is that during the crystallization of the lipid phase a partial expulsion of the drug on the particle surface may occur. Furthermore, the higher viscosity at the interface produced by higher lipid concentration may cause a decrease in diffusion and hence fewer lipid molecules will be carried into the aqueous phase. Therefore, the formation and stabilization of lipid aggregates at these higher concentrations are reduced (Abdelbary and Fahmy, 2009).

Table II. Physicochemical characterization of the MLX loaded SLNs

Formula	LD 10	LD 50	LD 90	Span	E.E %	Z potential (ζ) (mV)	pH
SLN ₁	235 ± 53.03	350 ± 56.57	525 ± 35.36	0.83	70.00 ± 0.82	-19.0	5.50 ± 0.68
SLN ₂	260 ± 28.28	350 ± 14.14	420 ± 14.14	0.46	65.67 ± 0.26	-20.3	6.03 ± 0.14
SLN ₃	210 ± 30.90	270 ± 42.43	325 ± 7.07	0.43	62.07 ± 2.17	-23.7	6.39 ± 0.23
SLN ₄	150 ± 14.14	230 ± 28.28	335 ± 7.07	0.80	61.94 ± 1.31	-24.7	5.60 ± 0.02
SLN ₅	235 ± 77.78	405 ± 53.03	755 ± 21.21	1.28	72.08 ± 0.60	-27.2	5.88 ± 0.02
SLN ₆	185 ± 7.07	385 ± 35.36	850 ± 28.28	1.73	67.03 ± 0.30	-38.2	5.40 ± 0.02
SLN ₇	195 ± 7.07	395 ± 21.21	680 ± 14.14	1.23	76.60 ± 0.88	-18.7	5.37 ± 0.17
SLN ₈	205 ± 3.54	405 ± 7.07	585 ± 7.07	0.94	73.18 ± 0.24	-17.6	5.37 ± 0.02
SLN ₉	215 ± 25.80	355 ± 77.78	475 ± 21.21	0.73	71.22 ± 0.56	-22.9	6.10 ± 0.05
SLN ₁₀	235 ± 20.21	300 ± 70.71	415 ± 7.07	0.60	66.74 ± 0.16	-22.9	5.92 ± 0.04
SLN ₁₁	185 ± 7.07	395 ± 21.21	1025 ± 21.21	2.13	80.96 ± 0.68	-22.8	6.33 ± 0.06
SLN ₁₂	375 ± 75.77	505 ± 70.50	1080 ± 14.14	1.40	75.88 ± 0.91	-37.8	6.21 ± 0.09
SLN ₁₃	190 ± 14.14	380 ± 56.57	675 ± 35.36	1.28	82.22 ± 0.39	-20.5	6.05 ± 0.23
SLN ₁₄	185 ± 7.07	355 ± 7.07	565 ± 7.07	1.07	79.29 ± 0.90	-23.3	6.07 ± 0.02
SLN ₁₅	155 ± 3.54	270 ± 14.14	440 ± 14.14	1.06	74.73 ± 1.68	-23.1	5.82 ± 0.29
SLN ₁₆	145 ± 7.07	220 ± 14.14	325 ± 21.21	0.82	72.94 ± 2.99	-28.9	5.69 ± 0.15
SLN ₁₇	185 ± 21.21	335 ± 7.07	1005 ± 35.36	2.45	85.33 ± 1.07	-32.1	5.77 ± 0.08
SLN ₁₈	385 ± 50.91	515 ± 33.35	1000 ± 28.28	1.19	84.12 ± 0.94	-38.6	5.37 ± 0.06

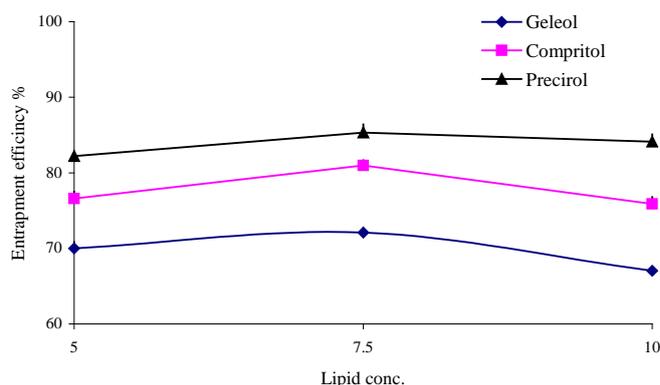


Fig.1. Effect of lipid concentration and type on the E.E% of MLX SLNs

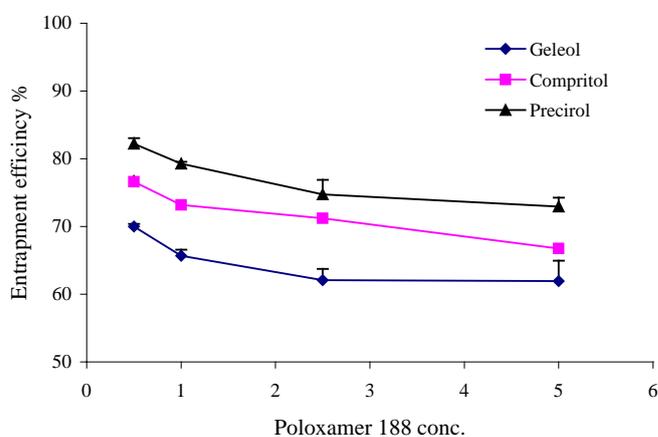


Fig.2. Effect of surfactant concentration on the E.E% of MLX SLNs

It was also evident that increasing the amount of surfactant from 0.5, 1, 2.5 to 5 % (w/w) at a constant amount of lipid (5 % w/w), resulted in a gradual significant decrease in the E.E. % of the produced SLNs ($p < 0.05$) (Fig. 2). This observed decrease in E.E. % could be explained by partition phenomenon. High surfactant level in the external phase might increase the partition of drug from internal to external phase of the medium. This increased partition is due to the increased solubilization of the drug in the external aqueous phase so more drug can disperse and dissolve in it (Rahman et al., 2010). However, in case of Geleol SLNs no further decrease in E.E. % was observed upon increasing the poloxamer 188 concentration above 2.5% (w/w) ($p > 0.05$) suggesting that an optimum concentration of surfactant was reached sufficient to cover the surface of nanoparticles effectively.

Particle size analysis

Table 2 depicts LD data of resulting MLX SLNs where LD 90% was used for comparing the influence of different parameters on the size of the nanoparticles. The LD 90% of all formulations ranged from 325 ± 7.07 to 1080 ± 14.14 nm with low Span values indicating narrow particle size distribution. The results clearly showed that there was a gradual decrease in particle size with an increase in surfactant concentration from 0.5, 1, 2.5 to 5 % (w/w) ($p < 0.5$) (Fig. 3). This decrease in size at high surfactant concentrations might be due to effective reduction in interfacial tension between the aqueous and lipid phases leading to the formation of emulsion droplets of smaller size

(Liu et al., 2007b). Higher surfactant concentrations effectively stabilized the particles by forming a steric barrier on the particle surface and thereby protect smaller particles and prevent their coalescence into bigger ones (Rahman et al., 2010).

As revealed in Figs. 3 and 4, Compritol 888 ATO showed the largest particle sizes followed by Precirol ATO 5 then Geleol SLNs. A possible explanation for the differences in sizes may be due to differences in chain lengths and viscosities of lipids used (Ahlin et al., 1998). Compritol 888 ATO (m.p. 69.0-74.0 °C) is a solid lipid based on glycerol esters of behenic acid (C22), where the main fatty acid is behenic acid > 85% but other fatty acids (C16-C20) are also present. Precirol ATO 5 (m.p. 50.0-60.0 °C) and Geleol (m.p. 54.5-58.4 °C) are composed mainly of palmitic (C16) and stearic acid (C18) > 90 %. High melt temperature resulting in higher viscosity plus the long hydrocarbon chain length of Compritol 888 ATO might result in larger particle size compared to Precirol ATO 5 and Geleol SLNs.

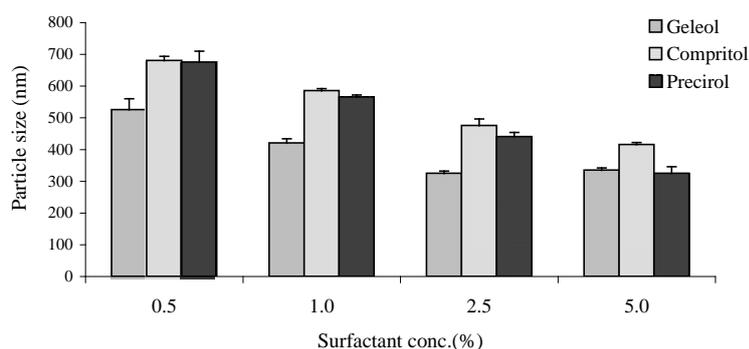


Fig.3. Effect of surfactant concentration on particle size measured by LD 90 % of different SLNs

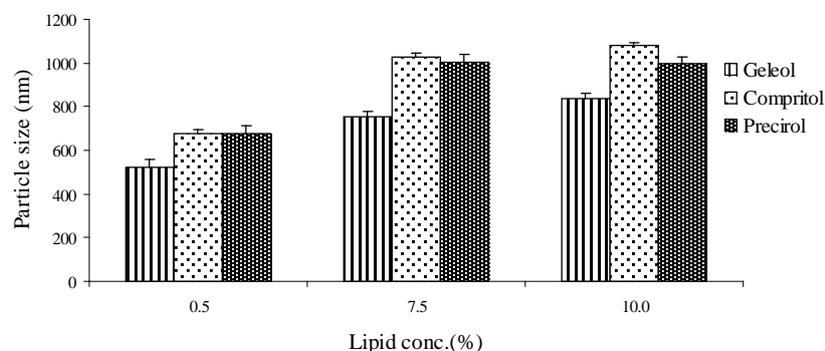


Fig.4. Effect of lipid concentration and type on particle size measured by LD 90 % of MLX SLNs

Lipid concentration seems to be one of the most important factors influencing the particle size. According to Fig. 4, increasing the lipid content from 5, 7.5 to 10 % (w/w) resulted in a subsequent increase in particle size. It was noticeable that SLN11, SLN12, SLN17 and SLN18 showed larger particle size exceeding the nanometer range. This increase in particle size may partially be related to the viscosity of the samples. The use of a low viscous lipid phase improves size reduction and enhances stability in SLN production (Manjunath et al., 2005). At higher lipid contents, the efficiency of homogenization decreases due to a higher viscosity of the sample, resulting in larger particles. Also, a high particle concentration at high lipid contents increases the probability of particle contact and subsequent aggregation (Freitas and Muller, 1998).

Zeta potential analysis (ξ) and pH measurements

Table 2 shows the measured zeta potential (ξ) values of MLX SLNs. As depicted from the table, all formulations were negatively charged, the zeta potential varied from -17.6 to -38.6 mV

indicating a relatively good stability and dispersion quality. It was noticeable that as the amount of surfactant increased in the formulation, the zeta potential became more negative. A similar finding was previously reported upon increasing Tween 80 concentration from 0.5 to 1 % which was attributed to the formation of denser surfactant film (Estella-Hermoso de Mendoza et al., 2008). Poloxamer 188 being non-ionic surfactant succeeded in the production of relatively stable dispersions. Although non-ionic surfactant could not ionize into charging group like ionic ones, but still demonstrated its zeta potential. The reason might be due to molecular polarization and the adsorption of emulsifier molecule on the charge in water, it was absorbed to the emulsifier layer of particle/water interface and electric double layer similar to ionic was formed. Liu et al. reported that poloxamer 188 was found to be one of the most effective non-ionic surfactants to avoid the formation of aggregates (Liu et al., 1996). Poloxamer 188 can provide additional steric stabilization of particles; so we can expect combined electrostatic and steric stabilization of SLN formulations (Schwarz and Mehnert, 1999; Lim and Kim, 2002).

Considering the effect of lipid type and concentration on the zeta potential (ξ) of the produced SLN formulations, the results showed no direct relationship between the type of lipid used and the measured ξ values. On the contrary, as the lipid concentration increased the zeta potential was found to be more negative. Rahman et al. reported the same observation when studying the effect of increasing Compritol amount in final formulation (Rahman et al., 2010).

The pH of different MLX SLN formulations was found to be within acceptable limits for topical application ranging from 5.37 ± 0.02 to 6.39 ± 0.23 (table 2).

Transmission Electron Microscopy (TEM)

TEM was conducted to investigate the morphology of MLX loaded SLNs. It was evident from TEM images that nanoparticles were almost spherical with smooth morphology, appeared as black dots, well dispersed and separated on the surface (Fig. 5). This description agrees with a previous observation that the use of chemically heterogeneous lipids in combination with heterogeneous surfactants favors the formation of ideally spherical lipid nanoparticles (Mehnert and Mader, 2001). Fig. 5d illustrates the presence of a very thin layer surrounding the particles which postulate a drug-enriched core model. This model can be achieved if during the lipid solidification process; the drug precipitates first, which results in a drug-enriched core covered with a lipid shell which has a lower drug concentration. This drug distribution within the nanoparticles will have its impact on in vitro drug release profile discussed later on.

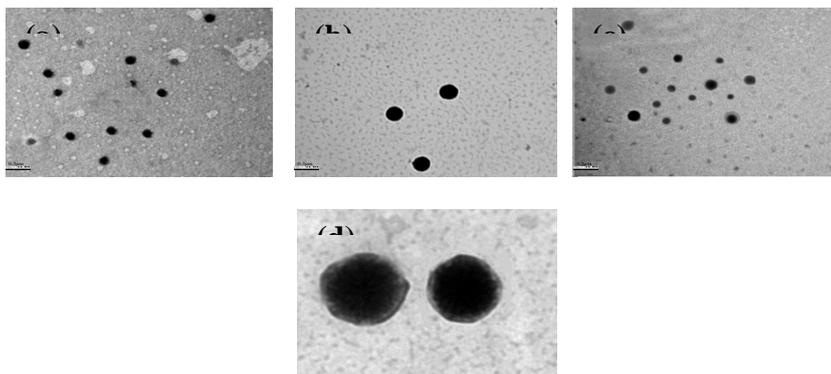


Fig.5. Transmission electron micrographs of MLX loaded SLN dispersions (a) SLN4 (b) SLN10 (c) SLN16 and (d) SLN representing the core and shell theory

Differential scanning calorimetry (DSC) analysis

DSC is a tool to investigate the melting and recrystallization behavior of crystalline material like SLNs (Liu et al., 2005). Fig. 6 shows the DSC thermograms of pure MLX, bulk lipids (Geleol, Compritol 888 ATO and Precirol ATO 5) and MLX loaded SLNs. Pure MLX showed a sharp endothermic peak at 259.54 °C corresponding to its melting point, indicating its characteristic crystalline nature. Bulk Geleol showed distinctive melting peak at 66.01 °C, while a sharp peak at 74.22 °C was observed for Compritol 888 ATO. The bulk Precirol ATO 5 exhibits a sharp endothermic event, ascribing to the melting, around 63.35 °C, with a small but well defined shoulder at 57.37 °C which might be due to the melting of α polymorphic form (Araujo et al., 2010). These sharp melting endothermic peaks of bulk lipids indicate that the starting materials were crystalline. As observed in Fig. 6, the thermograms of all investigated SLN systems did not show the melting peak of MLX around 259.54 °C indicating the conversion of crystalline MLX to the amorphous form which could be attributed to complete dissolution of the drug in the molten lipid matrix. The melting points of Geleol, Compritol 888 ATO and Precirol ATO 5 in SLN form were depressed showing slight shift to lower temperature side when compared to the corresponding bulk lipids. This melting point depression could be due to the small particle size (nanometer range), the high specific surface area, and the presence of surfactant - in other words, the depression can be attributed to the Kelvin effect (Jenning et al., 2000b). Kelvin realized that small, isolated particles would melt at a temperature lower than the melting temperature of bulk materials. In the same way, the melting enthalpy values of different lipids in SLN formulations showed drastic depression compared to their bulk lipids. These lower melting enthalpy values should suggest less ordered lattice arrangement of the lipid within nanoparticles compared to the bulk materials (Hou et al., 2003). For the less-ordered crystal or amorphous state, the melting of the substance requires less energy than the perfect crystalline substance, which needs to overcome lattice force. Lipid nanoparticles seem to loose part of their crystalline state transforming from a mixture of β' and β polymorphs to the most stable β polymorph, permitting MLX to fit in the molecular gaps. Therefore this decrease in the melting point and enthalpy values is associated with numerous lattice defects and the formation of amorphous regions in which the drug is located

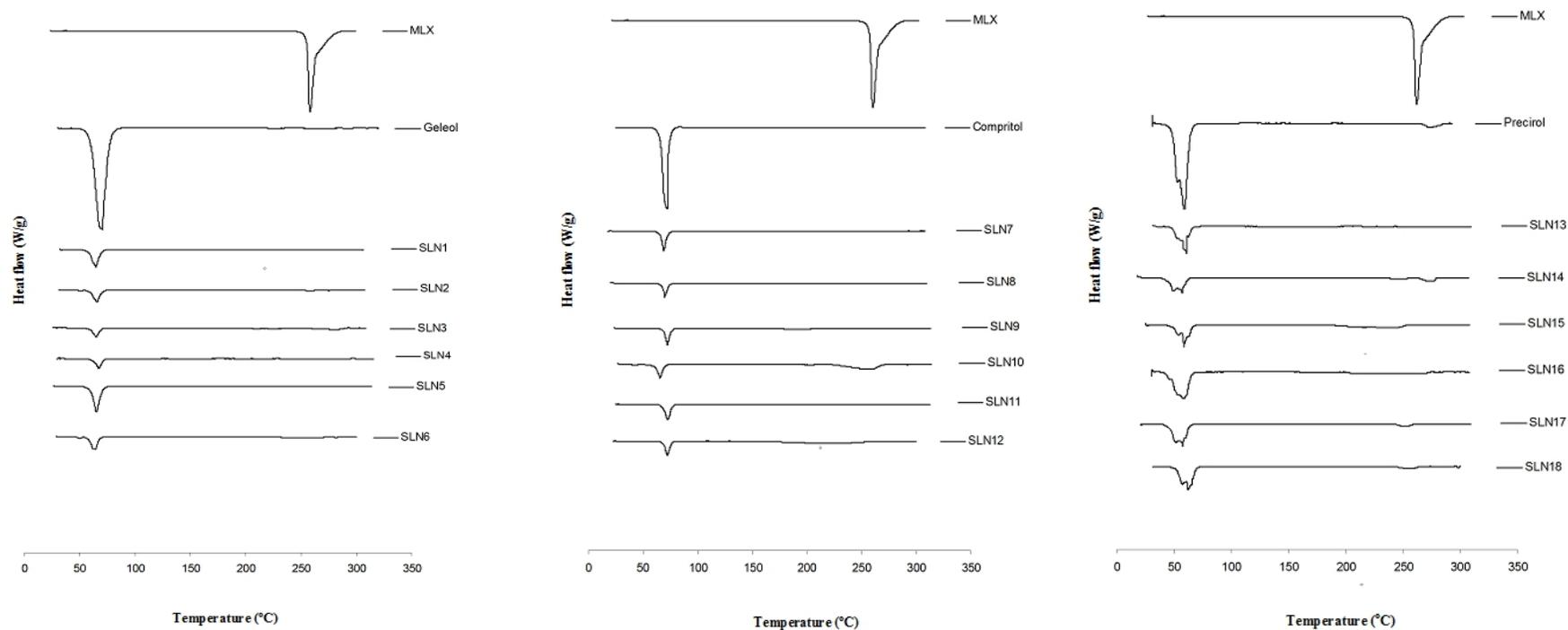


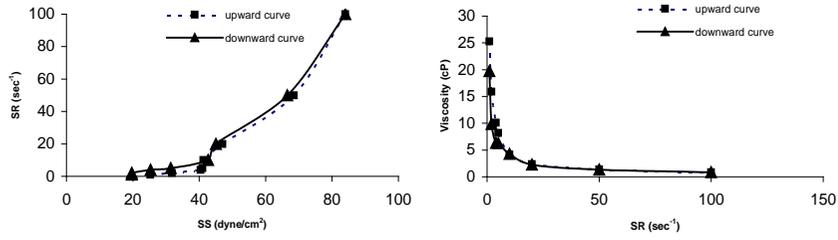
Fig.6. DSC curves of pure drug (MLX), bulk lipids (Geleol, Compritol and Precirol) and MLX loaded SLNs

Rheological Study

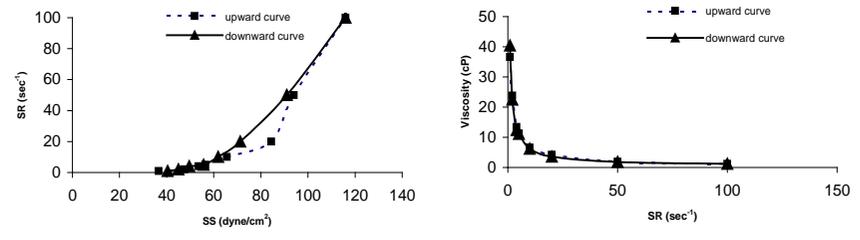
As for other disperse liquid and semisolid systems, the rheological properties of lipid nanoparticles influence their potential for dermal application in a fundamental way (Muller et al., 2002; Lippacher et al., 2001, 2002). Therefore, rheological behavior of different lipid nanoparticles formulations was studied and presented by plotting the shear stress versus shear rate (flow curves) and the viscosity versus the shear rate (viscosity curves) (Illinga and Unruh, 2004; Liu et al., 2008). The rheograms of different SLN formulations were shown in Fig. 7.

All SLN dispersions revealed non-Newtonian flow where the viscosity of non-Newtonian fluids changes according to the shear rate i.e. has no constant viscosity (49-Barnes, 1997). This flow was characterized by shear-thinning behavior in which the viscosity of the SLN dispersions decreased with the increase of shear rate. This dependent change in viscosity is a desired property in the pharmaceutical formulations due to their requirement of flexibility in topical drug delivery (50-Sheth, 2007). When the preparation is subjected to a shear force, its network structure breaks down leading to a gradual decrease in viscosity in order to spread on the skin. When the shear force is removed, the viscosity recovers slowly and the increased viscosity keeps the preparation on the skin.

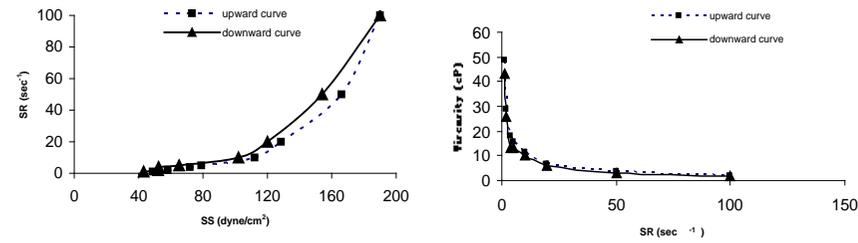
Comparing the viscosity curves of various SLN formulations, it was observed that, for each type of lipid, as the concentration of lipid increased the viscosity increased (Fig. 7). This may be related to the density of network structure, meaning that the network structure of 10 % (w/w) lipid was more viscous than that of 7.5 and 5 % (w/w) lipid, respectively, which was due to the higher amount of lipid incorporated into the system inducing an increase in the interaction between the lipid particles (Seetapan et al., 2010). In addition, the type of lipid affected the viscosity of the final product; Geleol SLNs showed lower viscosities compared to Precirol ATO 5 and Compritol 888 ATO SLNs. However, increasing the surfactant concentration did not result in considerable change in viscosity (data not represented).



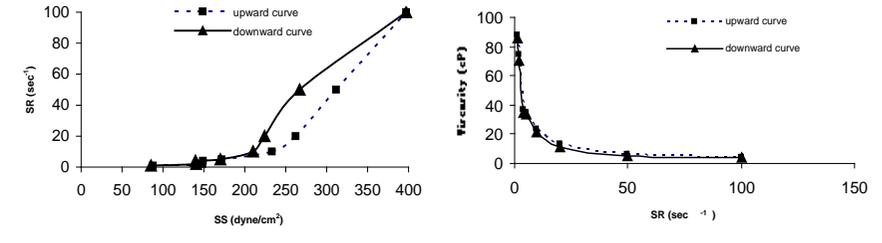
(SLN1)



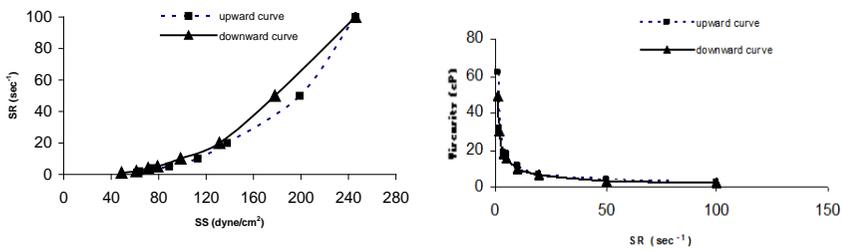
(SLN7)



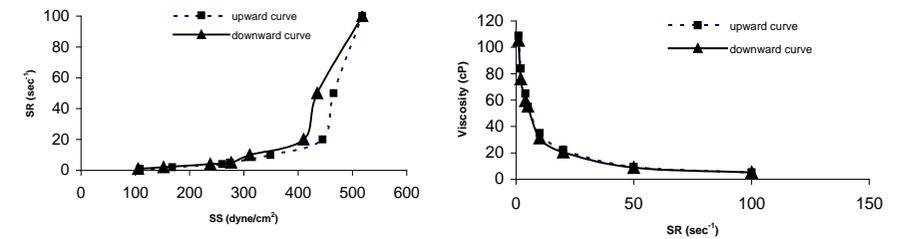
(SLN5)



(SLN11)



(SLN6)



(SLN12)

(A)

(B)

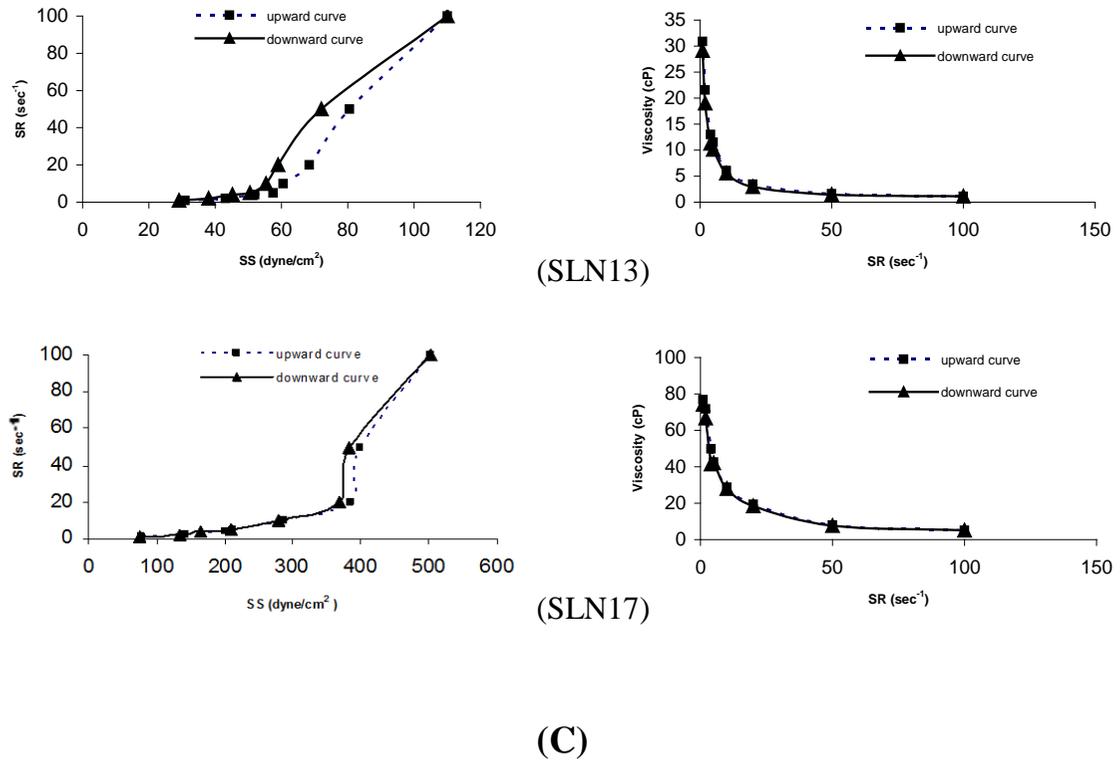


Fig.7. Rheograms of (A) Geleol SLNs (B) Compritol SLNs (C) Precirol SLNs

In vitro Release Study

To elucidate the mechanism of MLX release from SLNs for topical administration, *in vitro* release study using dialysis bag diffusion technique has been performed over 48 hrs, each sample was analyzed in triplicate and the release curves are shown in Fig. 8. The release data was fitted into zero order, first order and Higuchi equations which are widely used in determining the release kinetics of lipid nanoparticles. The release pattern of the drug from almost all SLN formulations followed Higuchi equation with some fitted to zero order equation. The previous result is in agreement with many studies which reported that drug loaded SLN provide a controlled release pattern following Higuchi's square root model (Tiyaboonchai et al., 2007; Vivek et al., 2007).

As shown in Fig. 8, MLX SLN formulations were able to release MLX in controlled manner and the percentage of MLX released up to 48 hours ranged from 43.11% to 100%. Interestingly, the amount of poloxamer 188 used had a great influence on the release pattern of SLNs. Increasing the poloxamer 188 concentration from 0.5, 1, 2.5 to 5% (w/w) led to corresponding increase in the percentage of MLX released as well as the release efficiency % (R.E. %) after 48 h, which was noticeable for all three lipids used (Figs. 8 and 9). However, in case of Geleol (SLN1& SLN2) and Precirol ATO 5 (SLN13 & SLN14) no significant increase in R.E. % was found upon increasing poloxamer 188 concentration from 0.5 to 1 % ($p > 0.5\%$). The fast or rapid release and higher release efficiency noticed at higher surfactant concentration could be explained by the partitioning effects of the drug between the melted lipid phase and the aqueous surfactant phase during particle production. During particle production by the hot homogenization technique, drug partitions from the liquid oil phase to the aqueous water phase. The amount of drug partitioning to the water phase will increase with the increase of the drug solubility in the water phase, which means with increasing temperature of the aqueous phase and increasing surfactant concentration. The higher the temperature and surfactant concentration, the greater is the solubility of the drug in the water phase so the amount of drug in the outer shell increased and released in a relatively rapid way (Zur Muhlen and Mehnert, 1998).

Concerning the type of lipid matrix, the results clearly show that among the glycerides used, the highest release was achieved with Geleol compared to Compritol 888 ATO and Precirol ATO 5. Being the lipid of highest monoglyceride content; Geleol had shown the highest release efficiency and consequently lower $t_{50\%}$, while in case of Compritol 888 ATO and Precirol ATO 5 the relatively slow release and higher $t_{50\%}$ can be attributed to the hydrophobic long chain fatty acids of the triglycerides that retain the lipophilic drug resulting in more sustained release effect (Reddy and Murthy, 2005; Kumara et al., 2007) (Figs. 8-10). Furthermore, the lower melting point of Geleol (54.5-58.5°C) may result in a higher mobility at the temperature used in the release experiment. It is well known that the melting point of colloidal structures may be lower than that of the bulk due to the influence of surface energy (Mader, 2006). A difference in release profiles caused by a difference in lipid melting points was also suggested by Paolicelli et al. in a study with ibuprofen and acylglycerols differing in melting points (Paolicelli et al., 2009). The higher amount released from Geleol particles may also reflect the smaller size of these particles as the mean diameter of Geleol nanoparticles represent the smallest of the SLN tested.

The results also pointed to the effect of lipid concentration on SLNs release profile. Increasing the lipid concentration from 5, 7.5 to 10 % (w/w) resulted in a corresponding decrease in percentage of MLX released, R.E. % after 48 h and a consequent increase in $t_{50\%}$ (Figs. 8-10). However, for Geleol and Precirol SLNs further increase of the lipid above 7.5% didn't result in a significant decrease in R.E. % ($p > 0.5\%$). This decrease in release profile observed can be attributed to the higher lipid content encapsulating the drug thus reducing drug partition in the outer phase and consequently its release in the receiver media. The release profiles of these SLNs resemble the drug enriched core model (Wissing et al., 2004). In such a model, the drug enriched core is surrounded by a practically drug-free lipid shell. Due to the increased diffusional distance and hindering effects by the surrounding solid lipid shell, the drug has a sustained release profile.

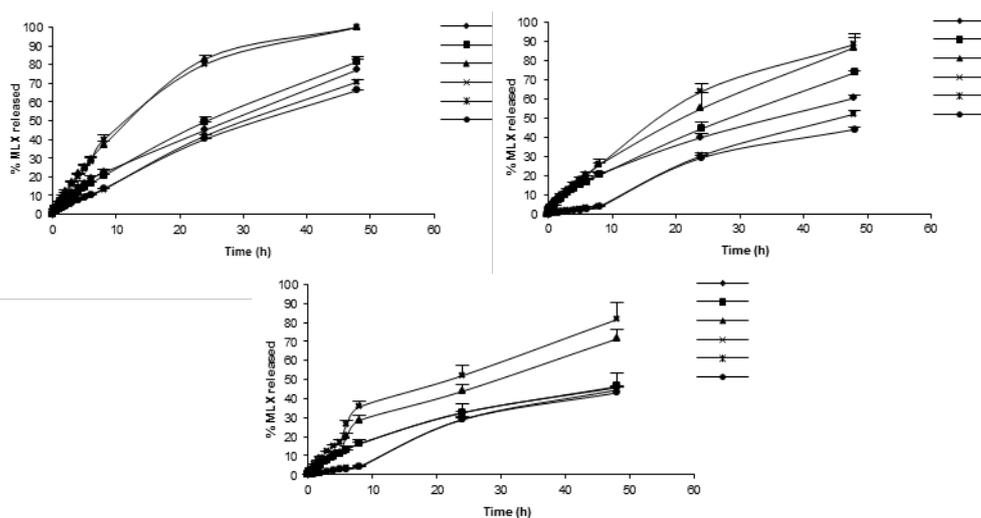


Fig.8. In vitro release percentage (%) - time profiles of MLX from SLNs. Data are expressed as the mean \pm S.D. (n = 3)

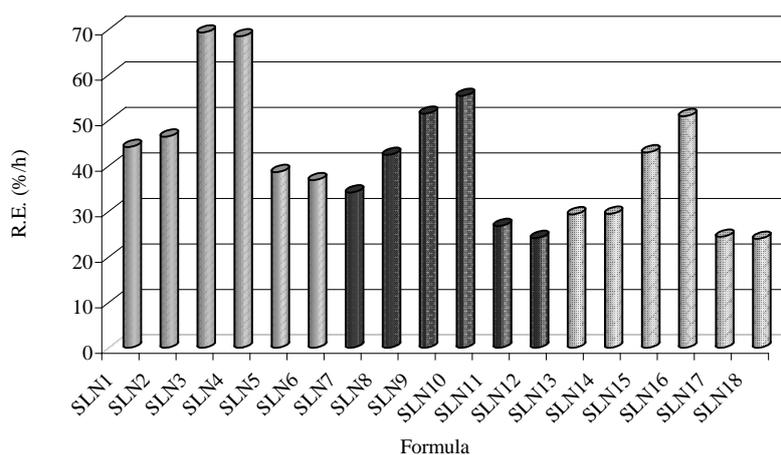


Fig.9. Release efficiency (R.E. %) of different MLX SLNs formulations

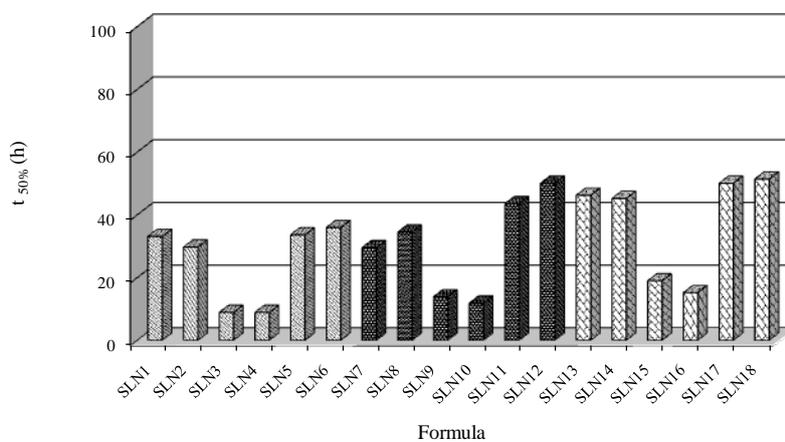


Fig.10. t_{50%} of different MLX SLNs formulations

Effect of storage on particle size

In order to evaluate the stability of lipid nanoparticles, the study of size was generally used as a characterization tool. The stability of the MLX loaded SLNs was studied for particle size after 12 months of storage at 4 °C. As previously mentioned, LD 90% was used for comparing the change in particle size of different SLNs.

As illustrated in Fig. 11, SLNs prepared from different lipids showed non significant change in particle size indicating good stability during the period of study. No obvious change of clarity or degradation was observed. All samples were in the nanometer range except for SLN11, SLN12, SLN17 & SLN18 which was probably due to higher amounts of lipid added to these formulations leading to increment in viscosity and subsequent aggregation of the nanoparticles.

The increase in particle size was observed to be more pronounced at low surfactant concentration(0.5-2.5%) for the three surfactants used , while 5% SLN showed the highest stability and particle size was almost unchanged allover the course of this investigation. This good stability might be attributed to the strong repulsion by sterical stabilization of the surfactants used at higher concentration (Freitas and Muller, 1999), emphasizing the importance of surfactant concentration for SLN stabilization.

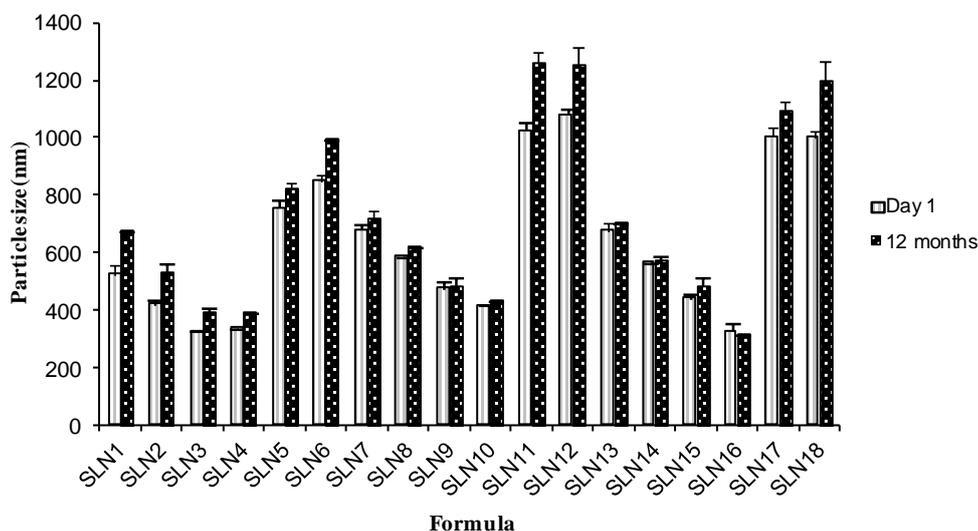


Fig.11. Particle size measured by LD 90 % of MLX SLNs after 1 day and 12 months of storage at 4 °C

It was noticeable that the role of surfactant in SLN systems varies depending on the lipid matrix used. SLN prepared with Geleol showed lower physical stability represented in higher percentage increase in particle size compared to Compritol 888 ATO and Precirol ATO 5. This could be explained by increased amount of partial glycerides like monoglycerides (40–55 % for Geleol, 15–23 % for Compritol 888 ATO and 8–22 % for Precirol ATO 5) which might be responsible for this physical destabilization. These results are in agreement with those reported by Jenning and Gohla (Jenning and Gohla, 2000).

Conclusion

In the present work, MLX-loaded SLNs were successfully prepared by high shear homogenization and ultrasonication technique.. The various physicochemical properties, rheological properties and the in vitro release behavior, were greatly affected and can be controlled by optimizing the compositional variables represented in the concentration of surfactant and lipid as well as the type of lipid used. The sustained release behaviour of MLX loaded SLNs with favourable physicochemical characteristics can form a foundation for further clinical studies using these nanoparticles for the topical delivery of meloxicam.

Acknowledgements

The authors would like to thank the National Research Center of Egypt, for the financial supporting of this research.

References:

- Abdelbary G, Fahmy RH. (2009). Diazepam-Loaded Solid Lipid Nanoparticles: Design and Characterization. *AAPS PharmSciTech*, 10, 211-219.
- Ahlin P, Kristl J, Smid-Kobar J. (1998). Optimization of procedure parameters and physical stability of solid lipid nanoparticles in dispersions. *Acta Pharm*, 48, 257-267.
- Araujo J, Gonzalez-Mira E, Egea MA, Garcia ML, Souto EB. (2010). Optimization and physicochemical characterization of a triamcinolone acetate-loaded NLC for ocular antiangiogenic applications. *Int J Pharm*, 393, 167-175.
- Barnes HA. (1997). Thixotropy - a review. *J Non-Newton Fluid Mech*, 70, 1-33.
- Cevc G. (2004). Lipid vesicles and other colloids as drug carriers on the skin. *Adv Drug Deliv Rev*, 56, 675-717.
- Chen CC, Tsai TH, Huang ZR, Fang JY. (2010). Effects of lipophilic emulsifiers on the oral administration of lovastatin from nanostructured lipid carriers: Physicochemical characterization and pharmacokinetics. *Eur J Pharm Biopharm*, 74, 474-482.
- Estella-Hermoso de Mendoza A, Rayo M, Mollinedo F, Blanco-Prieto MJ. (2008). Lipid nanoparticles for alkyl lysophospholipid edelfosine encapsulation: Development and in vitro characterization. *Eur J Pharm Biopharm*, 68, 207-213.
- Fang JY, Fang CL, Liu CH, Su YH. (2008). Lipid nanoparticles as vehicles for topical psoralen delivery: Solid lipid nanoparticles (SLN) versus nanostructured lipid carriers (NLC). *Eur J Pharm Biopharm*, 70, 633-640.
- Freitas C, Muller RH. (1998). Effect of light and temperature on zeta potential and physical stability in solid lipid nanoparticle (SLNTM) dispersions. *Int J Pharm*, 168, 221-229.
- Freitas C. and Muller, R.H. (1999) Stability determination of solid lipid nanoparticles (SLN) in aqueous dispersion after addition of electrolyte. *J. Microencapsulation*, 16, 59-71.
- Hou D, Xie C, Huang K, Zhu C. (2003). The production and characteristics of solid lipid nanoparticles (SLNs). *Biomaterials*, 24, 1781-1785.
- Hu FQ, Yuan H, Zhang HH, Fang M. (2002). Preparation of solid lipid nanoparticles with clobetasol propionate by a novel solvent diffusion method in aqueous system and physicochemical characterization. *Int J Pharm*, 239, 121-128.
- Illing A, Unruh T. (2004). Investigation on the flow behavior of dispersions of solid triglyceride nanoparticles. *Int J Pharm*. 284, 123-131.
- Jee JP, Lim SJ, Park JS, Kim CK. (2006). Stabilization of all-trans retinol by loading lipophilic antioxidants in solid lipid nanoparticles. *Eur J Pharm Biopharm*, 63, 134-139.
- Jenning V, Gohla S. (2001). Encapsulation of retinoids in solid lipid nanoparticles (SLN). *J Microencapsulation*, 18, 149-158.
- Jenning V. and Gohla S. (2000). Comparison of wax and glyceride solid lipid nanoparticles (SLN), *Int. J. Pharm.*, 196, 219-222.
- Jenning V, Gysler A, Schafer-Korting M, Gohla SH. (2000a). Vitamin A loaded solid lipid nanoparticles for topical use: occlusive properties and drug targeting to the upper skin. *Eur J Pharm Biopharm*, 49, 211-218.
- Jenning V, Thunemann AF, Gohla SH. (2000b). Characterization of a novel solid lipid nanoparticle carrier system based on binary mixtures of liquid and solid lipids. *Int J Pharm*, 199, 167-177.
- Kumara VV, Chandrasekar D, Ramakrishna S, Kishan V, Rao YM, Diwan PV. (2007). Development and evaluation of nitrendipine loaded solid lipid nanoparticles: Influence of wax and glyceride lipids on plasma pharmacokinetics. *Int J Pharm*, 335, 167-175.
- Li Y, Dong L, Jia A, Chang X, Xue H. (2006). Preparation and characterization of solid lipid nanoparticles loaded traditional chinese medicine. *Int J Biol Macromol*, 38, 296-299.
- Lim SJ, Kim CK. (2002). Formulation parameters determining the physicochemical characteristics of solid lipid nanoparticles loaded with all-trans retinoic acid. *Int J Pharm*, 243, 135-146.
- Lippacher A, Muller RH, Madder K. (2001). Preparation of semisolid drug carriers for topical application based on solid lipid nanoparticles. *Int J Pharm*, 214, 9-12.
- Lippacher A, Muller RH, Madder K. (2002) Semisolid SLN dispersions for topical application: influence of formulation and production parameters on viscoelastic properties. *Eur J Pharm Biopharm*, 53, 155-160.

- Liu F, Yang J, Huang L, Liu D. (1996). Effect of non-ionic surfactants on the formation of DNA/emulsion complexes and emulsion-mediated gene transfer. *Pharm Res* 13, 1642-1646.
- Liu J, Gong T, Wang C, Zhong Z, Zhang Z. (2007b). Solid lipid nanoparticles loaded with insulin by sodium cholate-phosphatidylcholine-based mixed micelles: Preparation and characterization. *Int J Pharm*, 340, 153-162.
- Liu J, Hu W, Chen H, Ni Q, Xu H, Yang X. (2007a). Isotretinoin-loaded solid lipid nanoparticles with skin targeting for topical delivery. *Int J Pharm*, 328, 191-195.
- Liu M, Dong J, Yang Y, Yang X, Xu H. (2005). Characterization and release of triptolide-loaded poly (d, l-lactic acid) nanoparticles. *Eur Polym J*, 41, 375-382.
- Liu W, Hu M, Liu W, Xue C, Xu H, Yang X. (2008). Investigation of the carbopol gel of solid lipid nanoparticles for the transdermal iontophoretic delivery of triamcinolone acetonide acetate. *Int J Pharm*, 364, 135-141.
- Lva Q, Yua A, Xia Y, Lia H, Songa Z, Cuic J, Caoa F, Zhaia G. (2009). Development and evaluation of penciclovir-loaded solid lipid nanoparticles for topical delivery. *Int J Pharm*, 372, 191-198.
- Mader K. (2006). In: Torchilin, V.P. (Ed.), *Solid Lipid Nanoparticles as Drug Carriers*. Imperial College Press, London, p. 187.
- Mandawgade SD, Patravale VB. (2008). Development of SLNs from natural lipids: Application to topical delivery of tretinoin. *Int J Pharm* 363, 132-138.
- Manjunath K, Reddy JS, Venkateswarlu V. (2005). Solid lipid nanoparticles as drug delivery systems. *Methods Finds Exp Clin Pharmacol*, 27, 127-144.
- Mehnert W, Mader K. (2001). Solid lipid nanoparticles: production, characterization and applications. *Adv Drug Deliv Rev*, 47, 165-196.
- Mei Z, Chen H, Weng T, Yang Y, Yang X. (2003). Solid lipid nanoparticle and microemulsion for topical delivery of triptolide. *Eur J Pharm Biopharm*, 56, 189-196.
- Muller RH, Mader K, Gohla S. (2000). Solid lipid nanoparticles (SLN) for controlled drug delivery - a review of the state of art. *Eur J Pharm Biopharm*, 50, 161-177.
- Muller RH, Radtke M, Wissing SA. (2002). Solid lipid nanoparticles (SLN) and nanostructured lipid carriers (NLC) in cosmetic and dermatological preparations. *Adv Drug Deliv Rev*, 54, S131-S151.
- Muller RH, Runge SA, Ravelli V, Thunemann AF, Mehnert W, Souto EB. (2008). Cyclosporine-loaded solid lipid nanoparticles (SLN): drug-lipid physicochemical interactions and characterization. *Eur J Pharm Biopharm*, 68, 535-544.
- Noble S, Balfour JA. (1996). Meloxicam. *Drugs*. 51, 424-430.
- Paolicelli P, Cerreto F, Cesa S, Feeney M, Corrente F, Marianecchi C, Casadei MA. (2009). Influence of the formulation components on the properties of the system SLN-dextran hydrogel for the modified release of drugs. *J Microencapsulation*, 26, 355-364.
- Parfitt K. (1999). *Martindale: The Complete Drug Reference*, 32nd ed. Pharmaceutical Press, USA, p. 52.
- Priano L, Esposti D, Esposti R, Castagna G, De Medici C, Fraschini F, Gasco MR, Mauro A. (2007). Solid lipid nanoparticles incorporating melatonin as new model for sustained oral and transdermal delivery systems. *J Nanosci Nanotechnol*, 7, 3596-3601.
- Rahman Z, Zidan AS, Khan MA. (2010). Non-destructive methods of characterization of risperidone solid lipid nanoparticles. *Eur J Pharm Biopharm*, 76, 127-137.
- Reddy LH, Murthy RSR. (2005). Etoposide-Loaded Nanoparticles Made from Glyceride Lipids: Formulation, Characterization, in Vitro Drug Release, and Stability Evaluation. *AAPS PharmSciTech*, 6, E158-E166.
- Schafer-Korting M, Mehnert W, Korting HC. (2007). Lipid nanoparticles for improved topical application of drugs for skin diseases. *Adv Drug Deliv Rev*, 59, 427-443.
- Schwarz C, Mehnert W. (1999). Solid lipid nanoparticles (SLN) for controlled drug delivery II. Drug incorporation and physicochemical characterization. *J Microencapsulation*, 16, 205-213.
- Seetapan N, Bejrappa P, Srinuanchai W, Ruktanonchai UR. (2010). Rheological and morphological characterizations on physical stability of gamma-oryzanol-loaded solid lipid nanoparticles (SLNs). *Micron*, 41, 51-58.
- Sheth B. (2007). Viscosity measurement and interpretation of viscosity data. *J Texture Stud*, 7, 157-178.

- Souto EB, Wissing SA, Barbosa CM, Muller RH. (2004). Development of a controlled release formulation based on SLN and NLC for topical clotrimazole delivery. *Int J Pharm*, 278, 71-77.
- Teeranachaideekul V, Souto EB, Junyaprasert VB, Muller RH. (2007). Cetyl palmitate-based NLC for topical delivery of Coenzyme Q10-Development, physicochemical characterization and in vitro release studies. *Eur J Pharm Biopharm*, 67, 141-148.
- Tiyaboonchai W, Tungpradit W, Plianbangchang P. (2007). Formulation and characterization of curcuminoids loaded solid lipid nanoparticles. *Int J Pharm*, 337, 299-306.
- Venkateswarlu V, Manjunath K. (2004). Preparation, characterization and in vitro release kinetics of clozapine solid lipid nanoparticles. *J Control Release*, 95, 627-638.
- Venkateswarlu V, Manjunath K. (2005). Pharmacokinetics, tissue distribution and bioavailability of clozapine solid lipid nanoparticles after intravenous and intraduodenal administration. *J Control Release*, 107, 215-228.
- Vivek K, Reddy H, Murthy RSR. (2007). Investigations of the Effect of the Lipid Matrix on Drug Entrapment, In Vitro Release and Physical Stability of Olanzapine-Loaded Solid Lipid Nanoparticles. *AAPS PharmSciTech*, 8, E1-E9.
- Westesen K, Bunjes H, Koch MHJ. (1997). Physicochemical characterization of lipid nanoparticles and evaluation of their drug loading capacity and sustained release potential. *J Control Release*, 48, 223-236.
- Wissing SA, Kayser O, Muller RH. (2004). Solid lipid nanoparticles for parenteral drug delivery. *Adv Drug Deliv Rev*, 56, 1257-1272.
- Wissing SA, Lippacher A, Muller RH. (2001). Investigations on the occlusive properties of solid lipid nanoparticles (SLN). *J Cosmet Sci*, 52, 313-323.
- Wissing SA, Muller RH. (2003). Cosmetic applications for solid lipid nanoparticles (SLN). *Eur J Pharm Biopharm*, 56, 67-72.
- Yang SC, Lu LF, Cai Y, Zhu JB, Liang BW, Yang CZ. (1999). Body distribution in mice of intravenously injected camptothecin solid lipid nanoparticles and targeting effect on brain. *J Control Release*, 59, 299-307.
- Zur Muhlen A, Mehnert W. (1998). Drug release and release mechanism of prednisolone loaded solid lipid nanoparticles. *Pharmazie*, 53, 552-555.

AUTOMATED CLASSIFICATION OF BRAIN MRI USING COLOR-CONVERTED K-MEANS CLUSTERING SEGMENTATION AND APPLICATION OF DIFFERENT KERNEL FUNCTIONS WITH MULTI-CLASS SVM

N.Rajalakshmi

Design engineer, Sanson Engineers Ltd, Coimbatore, India

V.Lakshmi Prabha

Principal, Government College of Technology, Tamilnadu, India

Abstract:

This paper proposes a hybrid approach for classification of brain magnetic resonance images (MRI) based on color converted hybrid clustering segmentation algorithm and wrapper based feature selection with multi-class support vector machine (SVM). The texture, color and shape features have been extracted and these features are used to classify MR brain images into three categories namely normal, benign and malignant. The MR images are classified by wrapper approach with Multi class Support Vector Machine classifier (MC-SVM) using color, texture and shape features. Performance of the MC-SVM classifier is compared with different kernel functions. From the analysis and performance measures like classification accuracy, it is inferred that the brain MRI classification is best done using MC- SVM with Gaussian RBF kernel function than linear and polynomial kernel functions. The proposed system can provide best classification performance with high accuracy and low error rate.

Key Words: Magnetic resonance imaging (MRI), Color-converted segmentation algorithm, PSO+K-means clustering technique, Feature extraction, Classification accuracy, multi class Support vector machine (MC-SVM)

Introduction

This is especially true for any attempt to classify brain tissues [1]. The most important advantage of MR imaging is that it is non-invasive technique [2]. The use of computer technology in medical decision support is now widespread and pervasive across a wide range of medical area, such as cancer research, gastroenterology, brain tumors etc. [3, 4]. Fully automatic normal and diseased human brain classification from magnetic resonance images (MRI) is of great importance for research and clinical studies. In the recent past, the development of Computer Aided Diagnosis (CAD) systems for assisting the physicians for making better decisions have been the area of interest. In CAD method, computer output has been used as a second opinion for radiologist to diagnose the information with confident and quicker mechanism as compared to manual diagnosis. Pathologies are clearly identified using automated CAD system [2]. It also helps the radiologist in analyzing the digital images to bring out the possible outcomes of the disease. The developments of Computer Aided Diagnosis (CAD) systems have been focused by many researchers for providing valuable information to the radiologists. Recent work [2, 5] has shown that classification of human brain in magnetic resonance (MR) images is possible via supervised techniques such as artificial neural networks and support vector machine (SVM) [2], and unsupervised classification techniques unsupervised such as self organization map (SOM) [2] and fuzzy c-means combined with feature extraction techniques [5]. Other supervised classification techniques, such as k-nearest neighbors (k-NN) also group pixels based on their similarities in each feature image [1] can be used to classify the normal/pathological T2-weighted MRI images. We used supervised machine learning algorithms

(ANN and MC-SVM) to obtain the classification of images under three categories, either normal or benign, malignant.

Methodolog

The proposed approach is shown in fig1. There are four major steps in the proposed approach: (1) Preprocessing (2) Color based K-means segmentation (3) Feature extraction: color, shape, texture features extracted from the segmented image (4) feature selection with classification.

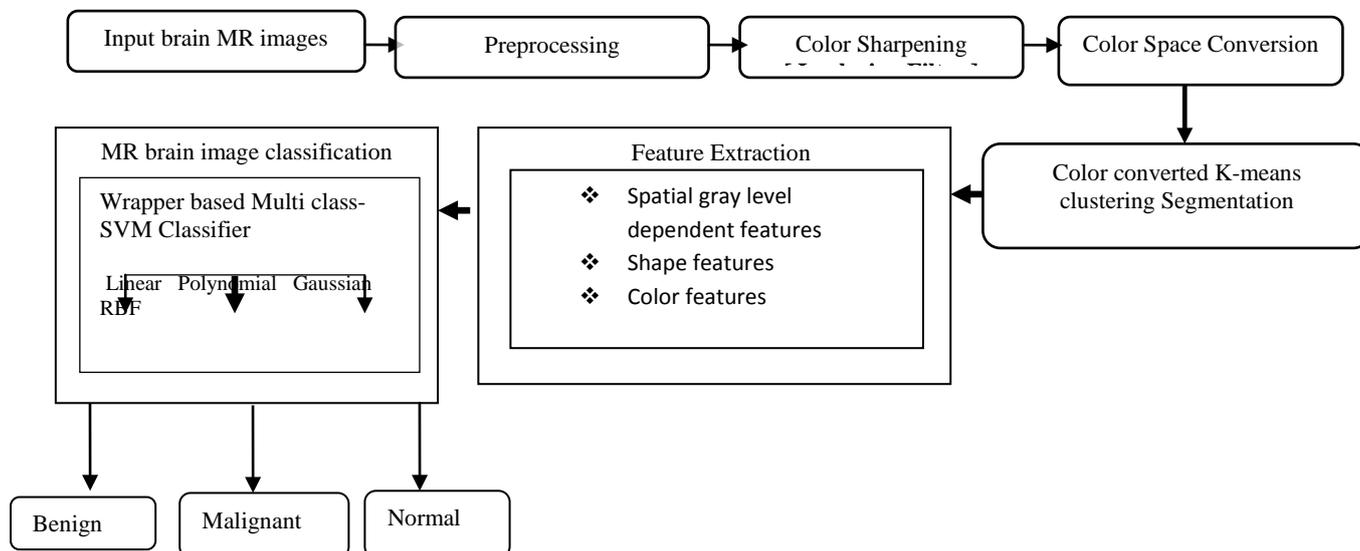


Fig1.Flowchart of proposed methodology

Preprocessing

Image denoising is a common pre-processing steps in many Magnetic Resonance (MR) image processing and analysis tasks, the goal of denoising is to remove the noise, which may corrupt an image during its acquisition or transmission, while retaining its quality. In this paper effectiveness of two denoising algorithms viz. (1) Wiener filter (2) Wavelet filter in the presence of additive white gaussian noise is compared. The Wiener filtering [6] executes an optimal tradeoff between inverse filtering and noise smoothing. It removes the additive noise and inverts the blurring simultaneously. The Wiener filtering is optimal in terms of the mean square error. In other words, it minimizes the overall mean square error in the process of inverse filtering and noise smoothing. The Wiener filtering is a linear estimation of the original image. The approach is based on a stochastic framework. Wiener method does a good job at deblurring; however, it behaves very poorly in the presence of large noise. To overcome the weakness of the Wiener filtering, Donoho and Johnstone proposed the wavelet based denoising scheme in [7].

Wavelet Thresholding approach

Steps for implementing denoising using wavelet based soft thresholding technique is as follows

- calculate two-level *haar* wavelet transform of the noisy image
- modify the noisy wavelet coefficients according to soft thresholding rule

$$w_{j,k} = \text{sgn}(w_{j,k})(|w_{j,k} - t_u|) \quad \text{if } w_{j,k} > t_u \tag{1}$$

$$= 0 \quad \text{if } w_{j,k} \leq t_u \tag{2}$$

Where Donoho [12] threshold also called Universal threshold given by: $t_u = \hat{\sigma} \cdot \sqrt{2 \log(n)}$, where n is the number of wavelet coefficients, and $\hat{\sigma} = \frac{MAD}{0.6745}$ is the estimates of the noise standard deviation. MAD denotes the Median Absolute Deviation of the wavelet coefficients in the finest resolution level. The wavelet coefficients $w_{j,k}$ above the universal threshold are updated by soft thresholding: $\text{sgn}(w_{j,k})(|w_{j,k} - t_u|)$. In practical applications, the variance of the noise is estimated by dividing the MAD by 0.6745.

- Compute the inverse *haar* wavelet transform using modified coefficients and then get denoised image.

Wavelet filter [7] removes noise pretty well in smooth regions but perform poorly along the edges. The results are compared on the basis of PSNR, SNR, and MSE. The comparisons of two denoising schemes are tabulated in Table1. It have been concluded that wavelet based techniques gives better results as compared to wiener filtering technique.

Table 1 shows comparison of two denoising schemes.Fig2 shows the output of the preprocessed image using wiener and wavelet methods. Thus the obtained results in qualitative and quantitative analysis show that this denoising algorithm based on wavelet transform outperforms the wiener method in terms of PSNR, SNR and MSE.

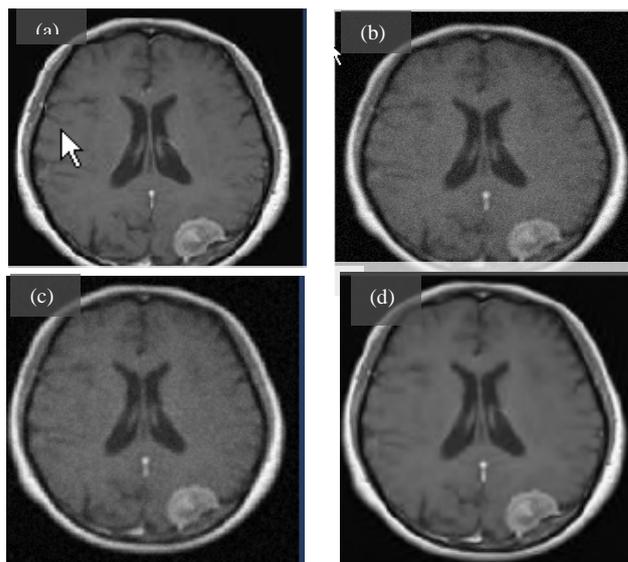


Fig2. Denoising of MR brain Image for variance=10(a)Original image(b)Noisy image(c) Denoised image with wiener filter (d) Denoised image using wavelet filter(soft threshold)

Table 1. Comparison of SNR, PSNR, MSE of denoising schemes of MR (magnetic resonance) brain image corrupted by additive white gaussian noise

Denoising schemes	Noise(σ)=10			Noise(σ)=20		
	PSNR	SNR	MSE	PSNR	SNR	MSE
Wiener Filter	20.07	10.60	639.93	19.64	10.17	706.13
Wavelet-soft thresholding	27.14	13.81	0.0034	25.16	12.61	0.0048

Color-Based Segmentation Using K-Means Clustering

Segmentation is the partition of a digital image into similar regions to simplify the image representation into something that is more meaningful and easier to analyze [8]. Pixels in the region are similar to each other with respect to some characteristic property like color, intensity or texture. In color based k-means clustering segmentation method MR gray level brain image is converted into an RGB color image by applying pseudo-color transformation a mapping function that maps a gray-level pixel to a color-level pixel by a lookup table in a predefined color map. An RGB color map contains R, G, and B values for each item. The method [9] has adopted the standard RGB color map, which gradually maps gray-level values 0 to 255 into blue-to-green-to-red color. Before transforming RGB color image into a CIE Lab color model, laplacian operator [10] has been applied to each channel in the RGB image using the equation $g(x,y)=f(x,y)+c[\nabla^2 f(x,y)]$ Where $g(x,y)$ is the new sharpened image, ∇^2 is Laplacian edge detection and $c(x,y)$ in RGB color image expressed as a vector of red, green and blue image components $c(x,y)=[R(x,y),G(x,y),B(x,y)]^T$ by computing the Laplacian of individual scalar components and the output color image appears sharpened. Color based

segmentation is significantly affected by the choice of color space. The general RGB color space gives high degree of detail, but it is not in tune with the normal human perception. L*a*b' color space is the better representation of the color content of the image and an added advantage of L*a*b' color space is that distance metric for clustering techniques continues to be Euclidean ,so we used L*a*b' color space in our work. To retrieve important features to benefit the clustering process, the RGB color space is further converted to a CIE Lab color model. The L*a*b* color space is derived from the CIE XYZ tristimulus values. The L*a*b* space consists of a luminosity layer 'L*', chromaticity-layer 'a*' indicating where color falls along the red-green axis, and chromaticity-layer 'b*' indicating where the color falls along the blue-yellow axis. This resultant features such as information in the a^*b^* layers along with gabor wavelet features which is a feature vector (texture representation) created using mean and standard deviation as the feature components with scale of 6 and orientation of 4 can extract the texture frequency and orientation information effectively. So color and gabor features are feature vectors, that will be given to the clustering process as an input. Here we are using K-Means algorithm for the clustering purpose. Use K-means to cluster the objects into three clusters using the Euclidean distance metric. And then labels every pixel in the image using the results from the clustered algorithm.

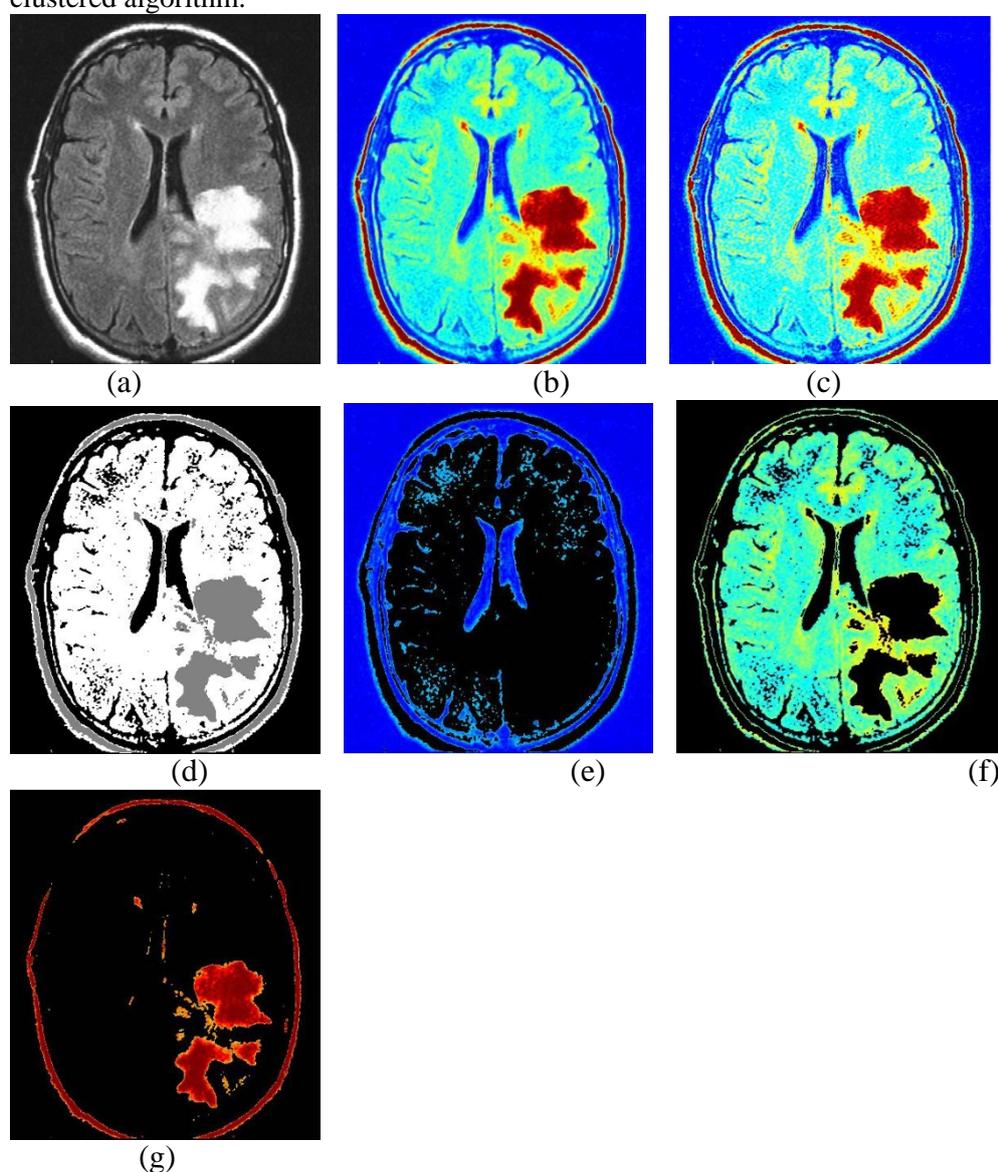


Fig 3(a) Original Gray Level Brain Image **(b)** RGB Color Space translated Image**(c)** Color sharpened image **(d)** image labeled by cluster index **(e)** objects in cluster 1**(f)** objects in cluster 2 and **(g)** final segmentation of K-means algorithm

Feature extraction

The third step of the proposed system is extracting Useful features from the image for classification purpose. It is a challenging task to extract good feature set for classification. There are many techniques for feature extraction, in this work texture, color and shape features have been considered for diagnosis.

Texture Features Extraction

A. Texture Features Extraction

Proposed system used texture feature extraction method proposed by Haralick [11], namely, the spatial gray-level dependence method (SGLDM). This well known statistical method for extracting second order texture information is based on the estimation of the second-order joint conditional probability density function [8,9] for the pixel (i, j), $P(i, j | d, \theta)$ for $\theta=0^\circ, 45^\circ, 90^\circ$ and 135° . The function $P(i, j | d, \theta)$ is the probability that two pixels which are located with an inter sample distance d and a direction θ . For an offset distance $d=1$, cooccurrence matrices are calculated for offset angles of $0^\circ, 45^\circ, 90^\circ$ and 135° . and fourteen Haralick features can be extracted.

B. Color Features

Color is one of the most widely used features. Color features can be obtained by various methods. In this work color moments [12] has been used because the Color moment method has the lowest feature vector dimension and lower computational complexity. Three Color moments features such as mean, variance and skewness are calculated for each of the Color component. Hence for each MR color space converted brain image total 9 features of Color moments are obtained

C. Shape features

Shape provides geometrical information of an object in image, it is an important visual feature and it is one of the primitive features for image content description. In this study the shape features such as Centroid, eccentricity, Euler number, solidity, convex hull, equidiameter, orientation, extrema, and extent are considered.

Feature Selection with MC-SVM

The numbers of texture, color and shape features extracted from the MR brain image can be irrelevant or redundant. Feature reduction improves classification by searching for the best features subset, from the fixed set of the original features, according to a given processing goal and a feature evaluation criterion: classification accuracy. Hence to reduce the large numbers of features to a smaller set of features in this work we used wrapper algorithm with multi-class SVM.

Multi-class support vector machine (MCSVM) classifier

Model Selection for Support Vector Machines

Model selection is a crucial part for SVMs classifier design and still an ongoing research issue. This usually involves the kernel and the corresponding parameters selection. In our proposed methodology, we have used one-against-all multi-class SVM [13, 14] (MC-SVM) with following kernel functions such as

- **Linear kernel:** $K(x_i, x_j) = 1 + x_i^T x_j$
- **Polynomial kernel:** $K(x_i, x_j) = (1 + x_i^T x_j)^P$
- **Gaussian Radial basis Function (RBF):** $\exp\left[\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right]$

In this article, we propose a feature ranking based wrapper algorithm [15] for multi-class SVM classification. Feature ranking is achieved by Information theoretic ranking criterion. In our proposed methodology, we have used one-against-all multi-class SVM (MC-SVM) with gaussian RBF kernel function which is as

follows; **Gaussian Radial basis Function (RBF):** $\exp\left[\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right]$, and we applied a grid search

with 10-fold cross validation process to select the optimal parameters in this paper. Range values for C and σ tested were $1 \leq C \leq 100$ and $0.05 \leq \sigma$ (sigma) ≤ 2 . Optimal values were experimentally determined equal to $C=10$ and $\sigma=0.5$. The procedure is summarized below

- Apply feature ranking technique on dataset using mutual information. Feature ranking is provided by Information theoretic ranking criterion. To measure the non linear

dependencies between a feature and the target, mutual information between each feature and the target is investigated in information theoretic approach.

The mutual information is defined by entropy I as:

$$I(i) = \int_{x_i} \int_y p(x_i, y) \log \frac{p(x_i, y)}{p(x_i)p(y)} dx dy \quad (33)$$

Where $p(x_i)$ and $p(y)$ are the probability densities of x_i and y , and $p(x_i, y)$ is the joint density. The criterion $I(i)$ is a measure of dependency between the density of variable x_i and the density of the target y .

- Select top k features and train MC-SVM with top k features then evaluate 10 fold cross validation accuracy.
- Select $k+i$ features and evaluate MC-SVM
- Repeat step 3 until the accuracy of $k+i \geq k$
- Set maximum classification accuracy from the wrapper approach classification.

Results and Discussion

Twenty three Scans containing normal brain, twenty five scans containing benign tumor and hundred and two scan containing malignant T1-weighted brain tumor images were obtained from hospitals in Coimbatore and have been considered for this work. The images under study were acquired using the Siemens 1.5-Tesla MR Systems. Initially MR images are subjected to preprocessing using wiener filter. The performance of this filter is compared with wavelet filter method and shown in table 1. Fig2 shows the output of the preprocessed image using wiener and wavelet transform. It is seen that the wavelet performs well. Once the preprocessing is done convert all the MRI images into a RGB color space images then laplacian sharpening operator is applied to RGB color images in order to enhance both the intensity and the edge of the images. To retrieve important features to benefit the clustering process, the RGB color space is further converted to a CIE Lab color model. Segmentation was done by color based K-means approach. The sample image considered in this work and their respective outputs after RGB color space conversion, color sharpening and K-means Clustering Segmentation are shown in fig 3,4. For the feature extraction step texture, color and shape features were estimated. In the proposed system 12 SGLDM measures with offset distance of 1 and offset angles of 0° , 45° , 90° and 135° are derived and nine color, nine shape features were extracted and used for obtaining optimized feature set. The wrapper algorithm with MC-SVM classification used to classify the input features into normal, benign or malignant. In the classification step MC-SVM with Gaussian RBF kernel is compared with linear and polynomial kernel functions. It can be concluded from the experimental results that Gaussian RBF kernel based MC-SVM is a promising technique for MRI brain image classification and give high classification accuracy with low error rate.

The performance of the proposed method has been evaluated in terms of sensitivity, specificity and accuracy. Table 2 represents the performance comparison for classifier with different kernel functions. Here total 150 images are taken for training and testing. Among 150 images the Normal category is 23 images, benign 25 images, malignant 102 images are taken for training and testing which is classified using color based K-means and MC-SVM with different kernel functions. The results show that the proposed system with Gaussian RBF give better percentage of classification while compared to MC-SVM classifier with linear and polynomial kernel functions. Table 3 illustrates the classification accuracy, sensitivity, specificity, area under curve and standard error for performing the proposed approach by using the common kernel functions including linear, polynomial and gaussian RBF. The experimental results have shown that the proposed method with gaussian RBF achieves good classification accuracy and less standard error while compared to MC-SVM classifier with linear and polynomial kernel functions. Therefore, it can be concluded that Gaussian RBF kernel based MC-SVM is a promising technique for MRI brain image classification

The Receiver Operating Characteristic (ROC) curves are plotted with respect to sensitivity and specificity. The area under the ROC (AUC) curve is an important parameter to determine the overall classification accuracy of the proposed system. Fig 4 shows the comparison of ROC plot for MC-SVM classifier with linear, polynomial, Gaussian RBF kernel functions. It has been seen that, proposed method with Gaussian RBF has the highest (0.91) AUC whereas the other method give lesser (0.89, 0.86) value. Hence, proposed method provides a higher accuracy than other method.

Classes	No. of data for training/testing	Number of correctly classified data			Percentage of correct classification		
		Color based K-means with MC-SVM(Gaussian RBF)	Color based K-means with MC-SVM(Poly)	Color based K-means with MC-SVM(Lin)	Color based K-means with MC-SVM(Gaussian RBF)	Color based K-means with MC-SVM(Poly)	Color based K-means with MC-SVM(Lin)
Normal	23/23	19	18	16	82.6	81.7	79.9
Benign	25/25	23	23	21	92.4	92.4	89.3
Malignant	102/102	99	98	96	98.7	97.1	95.1
Average					91.6	90.5	88.5

Legend: Lin - Linear; Poly - Polynomial; RBF - Radial basis function

Table 2. Performance of the classifier

Classification Technique	Kernel used	Sensitivity %	Specificity %	Accuracy%	AUC	Standard error
Color based K-means with MC-SVM(color+texture+shape features)	Linear	90.12	80.54	89.17	0.86	0.04046
Color based K-means with MC-SVM(color+texture+shape features)	Polynomial	91.87	83.92	91.7	0.89	0.03924
Color based K-means with MC-SVM(color+texture+shape features)	Gaussian RBF	92.13	84.12	92.6	0.91	0.03340

Table 3. Results of the classifier

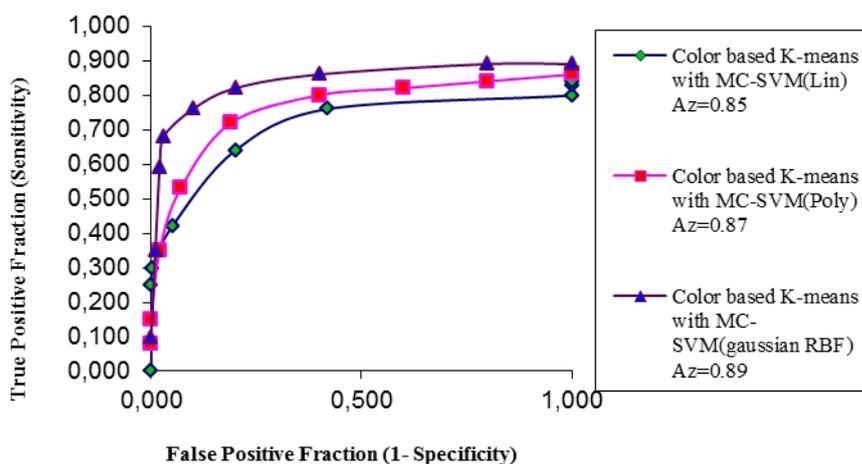


Fig 4. ROC analysis for tumor classification system

Conclusion

An improved automated classification technique using color-converted K-means clustering segmentation algorithm and wrapper algorithm with Multi-class SVM classifier with linear, polynomial, Gaussian RBF kernel functions for classifying Brain MRI as normal or abnormal (benign or malignant tumor) has been proposed and the performance is evaluated. It is concluded from the analysis that the multiple features, color converted K-means segmentation approach, the wrapper

approach MC-SVM with Gaussian RBF kernel function enhance the classification of MR brain image with normal and benign or malignant classes The proposed approach is efficient for classification of the human brain normal or abnormal (benign or malignant tumor) with high sensitivity, specificity and accuracy rates.

References

- [1] L. M. Fletcher-Heath, L. O. Hall, D. B. Goldgof, F. R. Murtagh, "Automatic segmentation of non-enhancing brain tumors in magnetic resonance images", *Artificial Intelligence in Medicine*, Vol.21, no.1, pp. 43-63, 2001.
- [2] L.M.Sandeep Chaplot, N.R. Patnaik and Jagannathan, "Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network", *Biomedical Signal Processing and Control 1*, Vol.4, no.3, pp. 86-92, 2006.
- [3] F. Gorunescu, "Data Mining Techniques in Computer-Aided Diagnosis: Non-Invasive Cancer Detection", *PWASET Volume 25* November 2007 ISSN 1307-6884, pp. 427-430, 2007.
- [4] [S. Kara, F. Dirgenali, "A system to diagnose atherosclerosis via wavelet transforms, principal component analysis and artificial neural networks", *Expert Systems with Applications*, Vol.32, no.2, pp. 632-640, 2007.
- [5] M. Maitra, A. Chatterjee; Hybrid multi resolution Slantlet transform and fuzzy c-means clustering approach for normal-pathological brain MR image segregation, *Med Eng Phys* (2007), doi:10.1016/j.medengphy.2007.06.009.
- [6] Ashok Kumar Nagawat, Manoj Gupta, Papendra Kumar and Suresh Kumar, "Performance Comparison of Median and Wiener Filter in Image De-noising", *International Journal of Computer Applications*, 12(2010) 0975-8887.
- [7] D.L.Donoho, "Denoise by softthresholding", *J. IEEE Transactions on Information Theory*. 41(1995) 613-627.
- [8] Muthusamy Madheswaran, Muthusamy Suganthi, "An Improved Medical Decision Support System to Identify the Breast Cancer Using Mammogram". *J. Med. Syst*, (2010), DOI 10.1007/s10916-010-9448-5.
- [9] Li-Hong Juang, Ming-Ni Wu, "MRI brain lesion image detection based on color converted K-means clustering segmentation", *Elsevier measurement*.43(2010) 941-949.
- [10] Ali Nosrati, Hamed Nosrati, Masoud Nosrati and Ronak Karimi, "A method for detection and extraction of circular shapes from noisy images using median filter and CHT", *Journal of American Science*.(2011) 84-88.
- [11] R.M.Haralick, I.Dinstein and K. Shanmugam, "Textural Features for Image Classification", *IEEE Trans On Systems Man and Cybernetics*.3(1973) 610-621.
- [12] Jayamala K. Patil, Raj Kumar, Color Feature Extraction of Tomato Leaf Diseases. *International Journal of Engineering Trends and Technology*, 2 (2):.72-74. 2011.
- [13] Lahouaoui lalaoui, Tayebmohamadi et al. "Support Vector Machine (SVM) and the Neural Networks for Segmentation the Magnetic Resonance Imaging", *SETIT 2009 5th International Conference: Sciences of Electronic, Technologies of Information and Telecommunications March, 2009 TUNISIA*, pp.22-26
- [14] D. Glotsos, P. Spyridonos et al, An image-analysis system based on support vector machines for automatic grade diagnosis of brain-tumour astrocytomas in clinical routine. *Medical Informatics and the Internet in Medicine* September 2005; 30(3): 179 – 193
- [15] Jianmin Jiang, Yonghong Peng and Zhiqing Wu, "A novel feature selection approach for biomedical data classification", *Elsevier Journal of Biomedical Informatics*.43 (2010), 15–23.

THE EFFECTS OF DIFFERENT DIETARY PROTEIN AND LIPID LEVELS AND OIL SOURCES ON THE GROWTH PERFORMANCE AND BODY COMPOSITION OF RAINBOW TROUT (*ONCORHYNCHUS MYKISS*, W.)

Murat Bilguven

Mersin University, Aquaculture and Fisheries Faculty, Aquaculture Department, Turkey

Ibrahim AK

Uludag University, Faculty of Agriculture, Department of Animal Science, Bursa, TURKEY

Abstract:

In this research, the effects of different dietary protein and fat level, and various oil sources on the fattening performance and body carcass composition of rainbow trout (*Oncorhynchus mykiss*) was investigated during 14 weeks.

Throughout the feeding period, it was found that dietary protein and fat level had affected the average live weight, live weight gain, feed conversion ratio, protein efficiency ratio (PER), specific growth rate (SGR) and carcass composition significantly ($P < 0,01$), although they had not effected the condition factor and the ratio of internal organs weight to total body weight.

Throughout the feeding period, it was determined that oil kinds had also affected the average live weight, live weight gain, feed conversion ratio, protein efficiency ratio (PER), specific growth rate (SGR) significantly ($P < 0,01$), while they hadn't effected the carcass composition, condition factor and the ratio of internal organs weight to total body weight.

Key Words: Rainbow trout (*Oncorhynchus mykiss*), dietary protein level, dietary fat level, oils, growth performance, carcass composition

Introduction

The biggest problem that the world will be facing in the future is going to be the ever increasing human population and the hunger or malnutrition that comes with it. Therefore, every nation on earth has to investigate and utilize its natural sources wisely and obtain the maximum yield per cultivated land or animal production unit. This necessity has consequently directed most of the research conducted in agriculture for finding new resources or developing and efficient use of existing ones.

It is obvious that food shortage caused by ever increasing human population on earth will only be alleviated by agriculture ventures that produce cost effectively. Aquaculture industry is no different than any other agriculture ventures and it will only be a part of sustainable food production if it maintains cost effectiveness and quality assured production. In parallel to the increment of intensive aquaculture production, aqua feed use in farms increased significantly in the last two decades. However, aqua feeds are expensive compared to the feeds of other farm animals and feed cost comprises of nearly 2/3 of the operating cost of any aquaculture farm during the whole production cycle.

Farmed fish require high protein diets with an optimum amount of non-nitrogenous energy sources (mainly lipid and digestible carbohydrates) in order to grow rapidly to market size in an intensive fish farm (Lovell, 1988). Fish feeds generally contain 25 to 50 % of crude protein. Salmonid feeds, however, contain crude protein between 40 to 50 % in their diets. In order to maintain these high protein levels in aqua feeds, ingredients that contain high amount of crude protein levels are extensively used in feeds and comprised of almost 50 to 75 % of the ingredients in commercial aqua feeds. Of the ingredients that are classified as high protein ingredient, fish meal is the most important

one and it generally comprises of between 25 and 65 % of the ingredients used in salmonid diets (Akiyama, 1988; Akyurt and Erdogan, 1994).

Protein is the most effectively used nutrient among other energetic nutrients in fish. Lipids and carbohydrates are used as an energy source on a lesser extent in fish. Therefore, It is required from aqua feeds that amino acids used in diets are mostly spent for depositing new proteins rather than used as an energy source to maintain basal metabolism in farmed aquatic animals. To achieve this, crude protein levels in diets are reduced and lipids and digestible carbohydrates are added to the diets to provide energy lost by reducing protein. However, It should be kept in mind that carnivorous fish species do not utilize complex carbohydrates well. Therefore, lipids remain as the only option to be considered to provide energy needs of these fish. In fact It was demonstrated that lipids as an energy source resulted in utilization of feed protein better for growth compared to carbohydrates (α -starch and dextrose) in rainbow trout (Ogino *et al.*, 1976). Similarly De la Higuera *et al.*, (1977) investigated the effects of crude lipid levels in diets on protein and lipid apparent digestibilities in the rainbow trout. Authors found that crude lipid apparent digestibility were not correlated with the crude lipid levels in diets. Dietary crude lipid levels were also found not to influence the apparent protein digestibility in the trout. They also concluded that protein use for growth had significantly increased in relation to the increment in dietary crude levels.

Luquet (1971) who investigated 4 diets containing 30, 40, 50, and 60 % crude protein and 2297, 2530, 2652 and 2839 kcal ME (Metabolizable Energy)/kg diet respectively for 36 week-grow out period in the rainbow trout found that decrement in dietary crude protein level from 60 % to 30 % increased the feed efficiency ratio from 0.98 to 1.97 whereas trout fed diets in rich of protein and energy were found to have more body crude lipid levels.

Takeuchi *et al.*, (1978a) have employed feed with a protein level of 35% and oil with a level between 5-25% in their study for the determination of optimal energy / protein ratio in the feeds of rainbow trout. Growth and feed efficiency have increased in line with the increase in the oil content of the feeds and reached the maximum level when the oil ratio was 18%. The value of PER has also increased with the increase in feed fat and it has been reported that DE/protein ratio that must be obtained should be 130 in feed in order to ensure an optimal growth of rainbow trout.

Takeuchi *et al.*, (1978b) in a study where they examined the impact of oil (5-20%) added to high-protein rainbow trout feed (54%) -casein was employed as the protein source- have found that there is not a significant difference in terms of growth rate between the groups however they have found that feed containing 20% oil has led to better results compared to the experimental feed containing 5% oil in terms of both growth rate and feed efficiency.

Reinitz *et al.*, (1978) have examined the effect of various fat levels at two different protein levels in rainbow trout diets. In this study, it was determined feed fat achieve saving of protein. Growth has been associated with feed efficiency while increase in length has been associated with energy content of the feed. High fat containing fish feed has resulted in the accumulation of higher levels of fat in carcass and accumulation of lower levels of protein in these fish compared to the control group.

Akiyama *et al.*, (1981) in an experiment made with *Oncorhynchus Keta*, have examined the relation between the need for protein in 9 and 12 °C water temperature and protein-fat requirement in ration. Researchers have reported that at both temperatures the feed containing 5% fat and 43% protein will require 38% protein when the level of oil is increased to 10%.

Kim *et al.*, (1988) have applied feed containing 9.14%, and 21.0% fat and 31.0%, 38.0 % and 44.0 % protein for 12 weeks to 10 different groups of rainbow trout the initial weight which were 46 g ad libitum. The results have revealed that the feed where fish oil is used as feed oil in feeds with low protein (38%) and high fat (21%) and has the effect of achieving saving in terms of protein.

Martins *et al.*, (2006) have also used soybean oil replacing fish oil in trout diets as much as 50% and reported that trout demonstrated better growth than that of fish in the control group. Yu and Sinnhuber (1981) Mugrditchian *et al.*, (1981), Hartfiel *et al.*, (1982), De La Hoz *et al.*, 1987, Hardy *et al.*, (1987) and Arzel *et al.*, (1993) have reported similar results to our findings obtained in this study using trout and salmon. They also found that oil sources used in their diets have not a significant effect on the feeding performance and body composition.

For all these reasons, it has been sought in this study to identify more appropriate protein and fat levels through employment of two different proteins and vegetable oils (soybean and linseed oil)

and animal fats (tallow and fish oil) in trout rations and the impact ratio of fat and protein levels and fat varieties on the composition of carcass has also been researched by conducting the nutrient analysis of the carcasses of fish at the end of the experiment.

Material and method

Fish and feeds

Juvenile rainbow trout used in this study were obtained from Istanbul University, Sapanca Freshwater Fishery Research Institute. The experiment was performed in a commercial trout farm with a total of 4800 randomly selected rainbow trout (*Oncorhynchus mykiss*, W.) (initial mean body weight, 17,3-17,7 g). Fish were stocked into 48 tanks at a rate of 100 fish per tank with 3 replication for each dietary treatment after they were counted.

Experimental diets were basically formulated to contain two protein levels at 45 and 35 CP% and two fat levels at 10 and 20 CF % on a dry matter basis. Four types of oils (soybean oil-SO, linseed oil-LO, tallow- T and fish oil-FO) were included in each of the CP and CF combination diets. Therefore a total 16 feeds were investigated in this research (Table 1). Dietary treatments were also formulated to contain between 3200 and 3750 kcal/kg DE (Digestible Energy) depending on the crude fat (CF) level of the diets (10 % vs 20 %, Table 2). Before adding to the diet mixtures, feedstuffs were ground to medium fine size (0.3 mm) and pelleted using a meat grinder. Pellet sizes were adjusted to have 3 mm diameter and 6 mm in length.

Feeding trial

The research was conducted in plastic tanks sized 200 L×50 W×60 H cm. Water was distributed with PVC pipes for each tank. The water flow rate was fixed at 2 L/min for all treatments, therefore almost 5 times of total water volume was changed with fresh water daily. Values of pH (D-51 Horiba), dissolved oxygen and water temperature (OM-51 Portable Dissolved Oxygen meter) measured periodically and detected as 7.5; 16.3-16.7; 7.9-8.2 respectively.

Fish were fed twice a day, morning (9 AM to 10 AM) and evening (5 PM to 6 PM) to apparent satiation. Utmost care was given that all the feed was consumed by the fish each feeding period. The daily feed intake was recorded throughout the experimentation. Fish were starved for 24 h before bi-weekly weigh gain measurements and batch measurements were performed for each tank in dietary treatments.

Chemical analysis and calculations

At the end of the experiment, fish weight gain, FCR, PER⁸, SGR⁹, CF¹⁰ and survival rate were calculated as given in the end note. Just before the start of the experiment, 5 fish were randomly collected and killed for the initial proximate carcass analysis. 5 fish from each treatment were also randomly selected and sacrificed and pooled for total body and carcass chemical composition analyses at the end of experiment. The chemical compositions of carcass, complete feeds and feedstuffs were measured following standard AOAC methods (AOAC, 1995). Feed, whole fish and fillet proximate composition analysis are given in Table 2 and 4.

Statistical analyses

16 combinations composed of three factors (dietary protein levels, dietary fat levels and oil types) were used in each application with three replicates of factorial experimental design. The mean final body weights in each treatment were subjected to statistical comparisons using ANOVA. All statistical analyses were carried out using the Minitab (v12) program. Results and Mean differences between treatments were tested for significance ($P<0.01$ and $P<0.05$) by Tukey's multiple range test. Results presented in Table 3 and 6 are reported as means \pm SD (n=3, 5 and 5 respectively).

$$^8 \text{ PER} = \frac{\text{live weight gain in an identified period, g}}{\text{consumed protein wit the diet in the same period, g}} \quad (\text{Hepher, 1988})$$

$$^9 \text{ SGR} = \left[\frac{\ln W_1 - \ln W_0}{t - t_0} \right] \times 100 \quad (\text{Hepher, 1988})$$

$$^{10} \text{ CF} = \frac{\text{DIV}}{t^3} \quad (\text{Brown, 1957})$$

Table 1. Biochemical Composition of Experimental Fish Diets, %.

Feed Ingredients	Feeds including 45% CP								Feeds including 35% CP							
	Feeds including 10 % CF				Feeds including 20 % CF				Feeds including 10 % CF				Feeds including 20 % CF			
	SO	LO	T	FO	SO	LO	T	FO	SO	LO	T	FO	SO	LO	T	FO
Fish Meal	31.11	31.11	31.11	31.11	31.11	31.11	31.11	31.11	31.11	31.11	31.11	31.11	31.11	31.11	31.11	31.11
Meat-Bone Meat	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
Soybeaan meal	28.22	28.22	28.22	28.22	30.00	30.00	30.00	30.00	5.88	5.88	5.88	5.88	5.00	5.00	5.00	5.00
Corn Gluten	6.22	6.22	6.22	6.22	6.69	6.69	6.69	6.69	1.89	1.89	1.89	1.89	4.16	4.16	4.16	4.16
Wheat Gluten	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
wheat midds	16.55	16.55	16.55	16.55	3.90	3.90	3.90	3.90	42.75	42.75	42.75	42.75	31.02	31.02	31.02	31.02
Soybean Oil	4.68	-	-	-	15.00	-	-	-	4.41	-	-	-	14.71	-	-	-
Linseed Oil	-	4.68	-	-	-	15.00	-	-	-	4.41	-	-	-	14.71	-	-
Tallow	-	-	4.68	-	-	-	15.00	-	-	-	4.41	-	-	-	14.71	-
Fish Meal	-	-	-	4.68	-	-	-	15.00	-	-	-	4.41	-	-	-	14.71
D. C. P.	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50	1.50
Mineral Mix. ^a	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
Vitamin Mix. ^b	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40	0.40
Vitamin C	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
DL-Methionine	0.52	0.52	0.52	0.52	0.60	0.60	0.60	0.60	0.76	0.76	0.76	0.76	0.80	0.80	0.80	0.80
L- Lysine	-	-	-	-	-	-	-	-	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
Lignobond ^c	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
Oxigard ^d	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
TOTAL	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

^a M-1 (g/kg of DM): 80.000 mg Mn, 35.000 mg Fe, 50.000 mg Zn, 5.000 mg Cu, 2.000 mg I, 400 mg Co, 150 mg Se.

^b V-221 (mg/kg or IU/ kg of DM): 4.800.000 IU Vitamin A, 800.000 IU Vitamin D₃, 12.000 mg Vitamin E, 1.200 mg Vitamin K₃, 1.200 mg thiamine, 2.400 mg riboflavine, 2.000 mg Vitamin B₆, 6 mg Vitamin B₁₂, 10.000 mg niacine, 16 mg biotin, 3.200 mg Calcium pantothenat, 400 mg folic acide, 120 mg Cholin chlorid, 20.000 mg Vitamin C.

^c This commercial product is used as pellet binder.

^d Antioxidant powder.

Table 2. Nutritional Composition of Experimental Feeds, %.

Feeds	Dry Matter	Organic Matter	Crude Protein	Crude Oil	Crude Cellulose	Nitrogen Free Matter	Crude Ash	Ca	P	Lysine ^a	Met.+ Cys. ^a	DE, KJ/kg
45CP10SO	91.10	83.21	44.98	9.98	2.68	25.57	7.89	2.96	1.96	2.29	1.50	13.39
45CP10LO	91.14	82.93	45.01	9.87	2.65	25.40	8.21	2.81	2.01	2.29	1.50	13.39
45CP10T	90.89	83.28	44.86	10.03	2.96	25.43	7.61	2.88	1.93	2.29	1.50	13.39
45CP10FO	90.13	82.57	44.91	10.03	3.01	24.62	7.56	2.91	1.90	2.29	1.50	13.39
45CP20SO	90.89	83.22	44.71	19.84	2.46	16.21	7.76	3.11	1.81	1.84	1.51	15.69
45CP20LO	91.56	84.04	45.01	20.12	2.45	16.46	7.52	3.03	1.78	1.84	1.51	15.69
45CP20T	91.75	84.22	44.96	19.96	2.61	16.69	7.53	2.98	1.76	1.84	1.51	15.69
45CP20FO	90.86	83.45	45.01	20.03	2.40	15.94	7.41	3.20	1.75	1.84	1.51	15.69
35CP10SO	91.55	84.59	34.34	9.97	1.70	38.58	6.96	2.95	2.04	2.08	1.50	13.39
35CP10LO	91.80	84.54	34.83	10.26	1.58	37.87	7.26	2.94	2.01	2.08	1.50	13.39
35CP10T	90.76	83.63	35.01	9.96	1.61	37.05	7.13	2.87	1.96	2.08	1.50	13.39
35CP10FO	90.29	83.40	35.02	9.98	1.55	36.85	6.89	2.88	1.97	2.08	1.50	13.39
35CP20SO	91.01	84.23	34.95	20.16	1.34	27.78	6.78	2.93	1.96	2.00	1.50	15.69
35CP20LO	91.08	83.99	35.03	19.89	1.43	27.64	7.09	2.97	1.94	2.00	1.50	15.69
35CP20T	90.16	83.61	34.77	20.04	1.37	27.43	6.55	3.01	1.93	2.00	1.50	15.69
35CP20FO	90.41	83.50	35.12	20.13	1.29	26.96	6.91	2.94	1.94	2.00	1.50	15.69

^a As a rate of dietary protein (N.R.C. 1981).

Results and Discussion

The results obtained in this experiment are summarized in Table 3 and 4. The effect of protein level on live weight gain was found significantly ($P < 0.01$). Our results are in agreement with Austreng and Refstie (1979), Tabachek (1986) and Heinen and Hankins (1995) findings. These authors have also demonstrated that live weight gain was affected by increasing level of protein in diets. Because fish could use protein both for supplying energy and routine metabolism function of cold water fish like trouts, need high level protein and non-nitrogenous energy sources (mainly fat and carbohydrates) in their feeds. As a matter of fact it can be seen that these carnivorous fish have consumed high level protein foods in the wild life when compared with warm water fish. However, high protein demands of trout can be reduced by addition of oil to feeds and it can be ensured that fish use protein for body mass production. Our study has also revealed that protein levels could be reduced by addition of fat to rations. The results of the experiment have shown that rations containing 35% protein and 20% fat and 45% protein and 10% fat generally leads to similar results in trout diets. As such, it has been determined that increase of the fat level in the ration has led to demonstrated in trout protein sparing effect in juvenile rainbow trout. Overall findings of this study are in line with the findings obtained by Luquet (1971), Reinitz *et al.* (1978) and Akiyama *et al.* (1981). Furthermore, It was found that live weight gain also was increased by the increment of lipid levels in diets throughout the study ($P < 0.01$). These results were supported by the findings of Takeuchi *et al.* (1978b), Tabachek (1986), Kim *et al.* (1988) and Heinen and Hankins (1995). Different oil sources than fish oil used in the experiment had no effect on body weight gain of the live fish insignificant. But the small differences observed in live weight gain among treatments could be attributed to the different digestibility of fats or fish oil included diet was more palatable to the trout. Several studies have also revealed that using vegetable oils in fish feed partly or wholly do not have negative effects on the growth performance and body composition (Luzzana *et al.*, 1994; Figueiredo-Silva *et al.*, 2005; Mourente *et al.*, 2006; Güler and Yıldız, 2011; Parpoura *et al.*, 2011).

It was determined in this study that feed intake of fish fed 35% CP and 10% CO was significantly higher than that fish in other dietary treatments ($P < 0.01$) and their FCR rates were better ($P < 0.01$). This can be explained by the fact of that fish consume more feed to fulfill their energy requirements. Furthermore, when the impact of oil varieties on feed consumption were analyzed it was thought that higher consumption of feed containing animal-derived fats than plant-derived oils is due to the fact that animal-derived fats are more palatable than the latter. The findings obtained in this research is similar to the results obtained by Luquet (1971), Takeuchi *et al.* (1978b), Austreng and Refstie (1979), Heinen and Hankins (1995), De La Hoz *et al.* (1987), Teskeredzic (1990), Akyurt and Erdogan (1993) and Arzel *et al.* (1993).

Increase in the level of oil in the ration has led to increase of PER and therefore better utilization of protein ($P < 0.01$). During the feeding period it has been determined that feeds containing fish oil containing diets provided better PER and this ratio is significantly lower in the fish fed tallow, soybean and linseed oil ($P < 0.01$). The findings of the experiment are coherent with the results obtained by Takeuchi *et al.* (1978a) De La Higuera *et al.* (1977) and Silver *et al.* (1991).

During the feeding period it has been determined that both protein and fat levels have significant effects on specific growth rate (SGR, $P < 0.01$). SGR of fish has increased in parallel with the increase of the level of protein and fat in diets. Silver *et al.* (1991) and Arzel *et al.* (1993) have also obtained similar findings to this study conducted experiments using rainbow trout.

Moreover, It also has been determined that protein level have significant effect on the body composition of fish in this experiment ($P < 0.01$). Moisture, crude fat and crude ash contents were all reduced in the carcass in parallel with the increased protein level of feed and accumulation of protein content has increased due to the crude protein levels in diets. These findings were also supported by the findings of Austreng and Refstie (1979) and Heinen and Hankins (1995). Crude oil levels in the body composition have also increased in parallel with the increased fat level of the diets whilst the content of crude protein and moisture have decreased. Findings of Luquet (1970), Reinitz *et al.* (1978), De La Higuera *et al.* (1987), Silver *et al.* (1991), Akyurt and Erdogan (1993), Arzel *et al.* (1993) and Heinen and Hankins (1995) have also supported our findings obtained in this experiment.

Researches have revealed that oil sources of feeds had no significant effect on the body proximate composition of fish. However, Mugrditchian *et al.* (1981) and Hartfiel *et al.* (1982) have also reported similar results to this experiment.

In conclusion;

*Protein sparing effect can be achieved in fish if the level of proteins in diets is reduced oil is cheaper and protein sources to be used in feed are scarce or expensive;

*By adding cheaper and digestible plant oil or animal originated fats as an energy source in trout diets, soybean oil, linseed oil, and tallow can be partially or completely substituted for the fish oil; and

*Inasmuch as the results of many studies on this subject are different from each other more research needs to be done to classify the optimal replacement levels of fish oil in the rainbow trout for the future development of trout diets.

Table 3. Average initial weight (WG), final weight, weight gain, feed intake (FI), FCR, SGR, PER and survival rate for rainbow trout fed different diets for 84 days.^a

Item	Protein Ratio		Lipid Ratio		Oil Sources			
	35 %	45 %	10 %	20 %	SO	LO	T	FO
	$\bar{X} \pm S_{\bar{X}}$	$\bar{X} \pm S_{\bar{X}}$	$\bar{X} \pm S_{\bar{X}}$	$\bar{X} \pm S_{\bar{X}}$	$\bar{X} \pm S_{\bar{X}}$	$\bar{X} \pm S_{\bar{X}}$	$\bar{X} \pm S_{\bar{X}}$	$\bar{X} \pm S_{\bar{X}}$
Trial Period, day	84	84	84	84	84	84	84	84
Total Fish Num.	2400	2400	2400	2400	1200	1200	1200	1200
Survival Rate, %	94	96	92	94	92	93	93	95
Initial Weight, g	17.54 ± 0.014	17.53 ± 0.014	17.53 ± 0.014	17.54 ± 0.014	17.57 ± 0.021	17.51 ± 0.021	17.56 ± 0.021	17.50 ± 0.021
Final Weight, g	178.15 ± 0.750 ^b	196.77 ± 0.750 ^a	179.41 ± 0.750 ^b	195.51 ± 0.750 ^a	187.65 ± 1.061 ^b	185.35 ± 1.061 ^b	184.24 ± 1.061 ^b	192.60 ± 1.061 ^a
AWG, g	160.61 ± 0.717 ^b	179.24 ± 0.717 ^a	161.88 ± 0.717 ^b	177.97 ± 0.717 ^a	170.08 ± 1.014 ^b	167.84 ± 1.014 ^b	166.68 ± 1.014 ^b	175.10 ± 1.014 ^a
FI, g	238.81 ± 1.012 ^a	214.66 ± 1.012 ^b	237.63 ± 1.012 ^a	215.84 ± 1.012 ^b	226.19 ± 1.431 ^b	226.06 ± 1.431 ^b	232.20 ± 1.431 ^a	232.20 ± 1.431 ^a
FCR	1.49 ± 0.005 ^b	1.21 ± 0.005 ^a	1.48 ± 0.005 ^b	1.22 ± 0.005 ^a	1.34 ± 0.007 ^b	1.37 ± 0.007 ^b	1.41 ± 0.007 ^a	1.28 ± 0.007 ^c
PER	1.94 ± 0.007 ^a	1.86 ± 0.007 ^b	1.72 ± 0.007 ^b	2.08 ± 0.007 ^a	1.90 ± 0.010 ^b	1.87 ± 0.010 ^b	1.82 ± 0.010 ^c	2.00 ± 0.010 ^a
SGR	2.758 ± 0.0050 ^b	2.877 ± 0.0050 ^a	2.767 ± 0.0050 ^b	2.867 ± 0.0050 ^a	2.817 ± 0.0050 ^b	2.805 ± 0.0050 ^b	2.796 ± 0.0050 ^b	2.851 ± 0.0050 ^a

Table 4. Nutritional composition of carcass in initial and final of the trial.^a

Item	Protein Ratio		Lipid Ratio		Oil Sources				
	35 %	45 %	10 %	20 %	Soybean Oil	Linseed Oil	Tallow	Fish Oil	
	$\bar{X} \pm S_{\bar{X}}$	$\bar{X} \pm S_{\bar{X}}$	$\bar{X} \pm S_{\bar{X}}$	$\bar{X} \pm S_{\bar{X}}$	$\bar{X} \pm S_{\bar{X}}$	$\bar{X} \pm S_{\bar{X}}$	$\bar{X} \pm S_{\bar{X}}$	$\bar{X} \pm S_{\bar{X}}$	
Initial	Moisture				79.41 ± 0,09				
	Crude Protein				17.36 ± 0.43				
	Crude Oil				1.95 ± 0.037				
	Ash				1.28 ± 0.011				
Final	Moisture	72.25 ± 0.092 ^a	71.74 ± 0.092 ^b	72.93 ± 0.092 ^a	71.07 ± 0.092 ^b	72.11 ± 0.130	71.85 ± 0.130	71.88 ± 0.130	72.16 ± 0.130
	Crude Protein	18.75 ± 0.054 ^b	19.84 ± 0.054 ^a	19.78 ± 0.054 ^a	18.81 ± 0.054 ^b	19.24 ± 0.077	19.44 ± 0.077	19.11 ± 0.077	19.39 ± 0.077
	Crude Oil	4.54 ± 0.015 ^a	3.68 ± 0.015 ^b	3.48 ± 0.015 ^b	4.74 ± 0.015 ^a	4.12 ± 0.021	4.11 ± 0.021	4.15 ± 0.021	4.06 ± 0.021
	Ash	1.35 ± 0.004 ^a	1.30 ± 0.004 ^b	1.33 ± 0.004	1.32 ± 0.004	1.33 ± 0.006	1.32 ± 0.006	1.33 ± 0.006	1.33 ± 0.006

^a Results are means ± SD (n=3; n=5). Averages followed by different letters in the same parameter column are significantly different ($P < 0.01$) by Tukey's test.

References:

- Akiyama, T., I. Yagisawa and T. Nose (1981). Optimum Levels of Dietary Crude Protein and Fat For Fingerling Chum Salmon (*Oncorhynchus keta*). Bull. Nat. Res. Inst. Aquaculture, 2: 35-42 pp.
- Akiyama, D.M. (1988). Soybean Meal Utilization in Fish Feeds. Korean Feed Association Conference. Seoul, Korea.
- Akyurt, I. and O. Erdogan (1994). Farklı Orijinli Lipidlerin Gökkuşluğu Alabalığı Fingerlingleri (*Oncorhynchus mykiss*) Rasyonunda Kullanılabilir Olanakları Üzerine Bir Araştırma. J. Of Veterinary and Animal Sciences. 18: 73-77. TUBITAK, Ankara, Turkey.
- AOAC (1997). Official Methods of Analysis of Association of Official analytical Chemists, International. 16th edn. AOAC International, Arlington, Virginia, USA
- Arzel, J., F.X. M. Lopez, R. Métailler, G. Stéphan, M. Viau, G. Gandemer and J. Guillaume (1994). Effect of Dietary Lipid On Growth Performance and Body Composition of Brown Trout (*Salmo trutta*) Reared In SeaWater. Aquaculture, 123: 361-375 pp.
- Austreng, E. and T. Refstie.. 1979. Effect of Varying Dietary Protein Level In Different Families of Rainbow Trout. Aquaculture, 18: 14-56 pp.
- Brown, M. E. (1957). Experimental studies on growth. In: M.E. Brown (Ed.), the physiology of fishes. Academic press, new york: 361-400 pp.
- De La Higuera, M., A. Murillo, G. Varela and S. Zamora (1977). The Influence Of High Dietary Fat Levels On Protein Utilization By The Trout (*Salmo gairdneri*). Comp. Biochem. Physiol. Vol 56A: 37-41 pp.
- De La Hoz, L., J.A. ordonez, M.A. asensio, M.I. cambero and B. Sanz (1987). Effects of Diets Supplemented With Olive Oil Bagasse or Technical Rendered Fat On The Apolar Lipids and Their Fatty Acid Composition Of Trout (*Salmo gairdneri*). ASFA. 11: 87-93 pp.
- Figueiredo-Silva, A., E. Rocha, J. Dias, P. Silva, P. Rema and E. Gomes (2005). Partial replacement of fish oil by soybean oil on lipid distribution and liver histology in European sea bass (*Dicentrarchus labrax*) and rainbow trout (*Oncorhynchus mykiss*) juveniles. Aquaculture Nutrition, 11(2):147-155 pp.
- Güler, M. and M. Yıldız (2011). Effects of dietary fish oil replacement by cottonseed oil on growth performance and fatty acid composition of rainbow trout (*Oncorhynchus mykiss*) Turkish Journal of Veterinarian Animal Science, 35(3): 157-167 pp.
- Hardy, R.W., T.M. Scott and L.W. Harrell (1987). Replacement Of Herring Oil With Menhaden Oil, Soybean Oil, or Tallow In The Diets Of Atlantic Salmon Raised In Marine Net-Pens. Aquaculture, 65: 267-277 pp.
- Hartfiel, V.W., D. Schulz und E. Greue (1982). Untersuchungen Über Die Fettverwertung Der Regenbogenforelle (*Salmo gairdneri*) III: Einsatz Von Leinöl, Maiskeimöl, Olivenöl, Schweineschmalz und Rindertalg im Vergleich Zu Sonnenblumenöl in Einer Gereinigten Futtermischung. Fette-Seifen Anstrichmittel 84: 31-33 pp.
- Hartfiel, V.W., D. Schulz und E. Greuel (1982). Untersuchungen Über Die Fettverwertung Der Regenbogenforelle (*Salmo gairdneri*) III: Einsatz Von Leinöl, Maiskeimöl, Olivenöl, Schweineschmalz und Rindertalg im Vergleich Zu Sonnenblumenöl in Einer Gereinigten Futtermischung. Fette-Seifen Anstrichmittel 84: 31-33 pp.
- Heinen, J.M. and J.A. hankins (1995). Evaluation of Two Higher Fat Diet For Rainbow Trout . J. Of Appl. Aquacult. 5 (2): 73-83 pp.
- Hepher, B. (1988). Nutrition Of Pondfishes. Cambridge University Press. Cambridge, UK, 388 pp.
- Luquet, P. (1971). Efficacité Des Protéines En Relation Avec Leur Taux D'incorporation Dans L'alimentation De La Truite Arc-en-ciel. Ann. Hydrobiol. 2: 176-186 pp.
- Luzzana, U., G. Serrini, V.M. Moretti, C.Gianesini and F. Valfre (1994). Effect Of Expended Feed With High Fish Oil Content On Growth and Fatty Acid Composition Of Rainbow Trout. Aquacult. Int., 2 (4): 239-248 pp.
- Kim, J.D., M. Pascaud and S.J. Kaushik (1988). Influence Of Dietary Lipid To Protein Ratios On Growth and Fatty acid Composition Of Muscles In Rainbow Trout. Ichthyophysiology- ACTA. 12: 7-25 pp.
- Mugrditchian, D.S., R.W. Hardy and W.T. Iwaoka (1981). Linseed Oil and Animal Fat As Alternative Lipid Sources In Dry Diets For Chinook Salmon (*Oncorhynchus tshawytscha*). Aquaculture, 25: 161-172 pp.
- Martins, D.A., E. Gomes, P. Rema, J. Dias, R. Ozório, and L.M.P. Valente (2006). Growth,

- digestibility and nutrient utilization of rainbow trout (*Oncorhynchus mykiss*) and European sea bass (*Dicentrarchus labrax*) juveniles fed different dietary soybean oil levels. *Aquaculture International*, 14: 285–295 pp.
- Minitab Inc. (1997). Minitab for Windows (v12). USA.
- Mourente, G. and J.G. Bell (2006). Partial replacement of dietary fish oil with blends of vegetable oils (rapeseed, linseed and palm oils) in diets for European sea bass (*Dicentrarchus labrax*, L.) over a long term growth study: effects on muscle and liver fatty acid composition and effectiveness of a fish oil finishing diet. *Comparative Biochemistry and Physiology Bulletin*, 145:389–399 pp.
- Parpoura, A.C.R. and M.N. Alexis (2001). Effects of different dietary oils in sea bass (*Dicentrarchus labrax*) nutrition. *Aquaculture International*, 9: 463-476 pp.
- Ogino, C., J.Y. Chiou and T. Takeuchi (1976). Protein Nutrition in Fish-VI. Effects Of Dietary Energy Sources On The Utilization Of Proteins By Rainbow Trout and Carp. *Bull. Jap. Soc. Sci. Fish.* 42 (2): 213-218 pp.
- Reinitz, G.L., L.E. Orme, C.A. Lemm and F.N. Hitzel (1978). Influence Of Varying Lipid Concentrations With Two Protein Concentrations in Diets For Rainbow Trout (*Salmo gairdneri*). *Trans. Am. Fish. Soc.* 107: 751-754 pp.
- Silver, G.R., D.A. Higgs, B.S. Dosanjh, B.A. Mckeown, G. Deacon and D. French (1991). Effect Of Dietary Protein To Lipid Ratio On Growth and Chemical Composition Of Chinook Salmon (*Oncorhynchus tshawytscha*) In Sea Water. *Fish Nutrition In Practice*, Biarritz (France). June 24-27, 1991.
- Tabachek, J.L. (1986). Influence of Dietary Protein and Lipid Levels On Growth, Body Composition, and Utilization Efficiencies Of Arctic Charr (*Salvelinus alpinus*). *J. Fish Biol.* 29:139-151 pp.
- Takeuchi, T., M. Yokoyama, T. Watanabe and C. OGInO (1978a). Optimum Ratio Of Protein To Lipid In Diets Of Rainbow Trout. *Bull. Jap. Soc. Sci. Fish.*, 44 (7): 729-732 pp.
- Takeuchi, T., T. Watanabe and C. Ogino (1978b). Supplementary Effect Of Lipids In A Protein Diet Of Rainbow Trout.. *Bull. Jap. Soc. Sci. Fish.*, 44 (6): 677-681 pp
- Teskeredzic, Z., E. Teskeredzic and M. Tmec (1989). The Influence Of Four Different Kinds Of Oil Upon The Growth Of Rainbow Trout (*Oncorhynchus mykiss*). *Proc. III. Intern. Symp. On Feeding and Nutrition In Fish*. Tobe, Aug. 28- sept. 1, 1989. 245-250 pp. Japan.
- Yu, T.C. and R.O. Sinnhuber (1981). Use Of Beef Tallow As An Energy Source In Coho Salmon (*Oncorhynchus kisutch*) Rations. *J. Fish. Res. Board Can.* 38: 367-370 pp.

DETERMINATION OF CLENBUTEROL IN MEAT SAMPLES WITH ELISA AND GC-MS METHOD

Zehra Hajrulai-Musliu
Risto Uzunov
Velimir Stojkovski
Elizabeta Dimitrieska-Stojkovik
Biljana Stojanovska-Dimzoska
Pavle Sekulovski

Faculty of Veterinary Medicine, Sts "Cyril and Methodius" University, Skopje, Macedonia

Abstract:

It is essential from the public health point of view, that the Macedonian regulatory agencies implement a monitoring surveillance program that not only uses preliminary screening methods for the determination of clenbuterol, but also uses confirmatory methods such as MS techniques. Thus, the aim of this work was to monitor the presence of clenbuterol residues in bovine meat muscle collected from the veterinarian inspectors, using screening and confirmatory methods. From 55 bovine meat samples analyzed by the ELISA test, 1 (1.8%) of presented clenbuterol residues was 1.19 µg/kg, 7 (12.7%) from 0.5 to 1.0 µg/kg, 13 (23.64%) from 0.1 to 0.5 µg/kg and 34 (61.82%) were below the limit of detection (LOD) of 0.1 µg/kg. In this study, from twenty one meat samples analyzed by GC-MS only one sample was confirmed positive for clenbuterol. The obtained results indicated that clenbuterol was still illegally used as a bovine growth promoter.

Key Words: Clenbuterol, residues, EU, ELISA, GC-MS, bovine meat

Introduction

Clenbuterol is used in human and veterinary medicine for its broncholytic and tocolytic action. In 5- to 10- fold the therapeutic doses, clenbuterol acts as a repartitioning agent to improve the performance of food-producing animals. Such a growth-promoting dose of clenbuterol influences animal growth and carcass composition by increased muscle mass and decreased fat mass, presumably through a direct effect of BAA on skeletal muscle and adipose tissue, indirect effect on many other tissues, or a combination of both. The monitoring of raw meat and poultry for drugs and chemical residues is necessary to ascertain that approved compounds are not misused and are not presenting a danger to consumers. Beta-agonists are a group of veterinary drugs that have been used illegally in some countries as they have a similar action to anabolic steroids in altering body composition (1). Beta-agonists act by impeding the uptake of adrenal hormones by nerve cells and stimulation of the cardiovascular system. When the treatment is prolonged, they also induce a redistribution of fat to muscle tissues (2). The beta-agonist clenbuterol has been implicated in several food poisoning cases in European countries (3). The use of b-agonists for growth promoting purposes in farm animals is not permitted in the European Union (EU), the United States of America (USA) and most other countries. The maximum residue limits (MRL) set by the EU are 0.1 µg/kg for muscle and 0.5 µg/kg liver and the MRL recommended by the WHO and the *Codex Alimentarius* are 0.2 µg/kg for muscle and 0.6 µg/kg for liver (4). Nevertheless, several cases of food poisoning between 1989 and 1992 were reported in Spain and France when liver containing high levels of clenbuterol were consumed (3). It is essential from the public health point of view, that the Macedonian regulatory agencies implement a monitoring surveillance program that only uses preliminary methods for the determination of clenbuterol. Although immunoassay techniques are very sensitive, the potential lack of specificity is a drawback that may result in false positives since other compounds of similar chemical structures present cross-reactivities (5-6). Therefore results of screening analysis should be confirmed by gas or liquid chromatography-mass spectrometric analysis (GC-MS) (7-9). The EU suggested the use of tandem mass spectrometry, also known as MS-MS, for the acquisition of

one parent ion and two product ions for the unambiguous GC-MS identification of the forbidden β -agonists (10). It is known that MS can provide significantly more information than standard mass spectrometry (MS) in cases where the analytes exhibit a similar primary fragmentation. This is because product ion mass spectra can be significantly different even for similar analytes, thus providing enhanced structural information with increasing selectivity of the analytes of interest (11). Additionally, MS-MS offers a number of advantages over single MS since the former provides discrimination between analytes on the basis of chromatographic properties (retention time), parent ion (MS1) and daughter ion (s) (MS2) (12). Thus, the aim of this work was to monitor the presence of clenbuterol residues in bovine meat muscle collected from the veterinarian inspectors by using screening and confirmatory methods.

Material And Methods

Reagents

Clenbuterol HCl, potassium dihydrogen phosphate, perchloric acid, bis(trimethylsilyl)-trifluoro-acetamide (BSTFA), formic acid, ethyl acetate, methanol and n-heptane were from Sigma (St. Louis, MO, USA). Trisbuffer (hydroxymethyl-aminomethane) was from Bio-Rad (Richmond, CA, USA). The immunoenzymatic test was carried out by using an ELISA kit for clenbuterol and other β 2-adrenergic agonists (r-Biopharm, Darmstadt, Germany). Sodium hydroxide and hydrochloric acid were from Merck Darmstadt, Germany). All solvents were of HPLC grade.

Samples

A total of 55 bovine meat samples were screened for the presence of clenbuterol as part of national monitoring residue plan. The samples were collected within period of 6 months as they were delivered by the authorised veterinary inspectors. Samples were kept frozen at -20 °C until analysis.

Clenbuterol analysis by ELISA (screening test)

A Ridascreen clenbuterol kit for ELISA was provided by R-biopharm (Darmstadt, Germany). Each kit contained a microtiter plate with 96 wells coated with antibodies to rabbit IgG, clenbuterol standard solutions (0, 100, 300, 900, 2700 and 8100 ng/L), peroxidase-conjugated clenbuterol, anti-clenbuterol antibody, substrate/chromogen solution, stop reagent, conjugate and antibody dilution buffer, and washing buffer. The extraction and clean-up procedures were those described by the ELISA kit manufacturer (r biopharm, Darmstadt, Germany). Minced meat samples (2 g) were homogenized for 30 min with 6 mL of methanol in 50 mL centrifuge tubes by using an Ultra-Turrax homogenizer, and centrifuged for 10 min at 4000 rpm at room temperature (20-25°C). Then 4 ml of methanolic solution was transferred into a new centrifugal vial and evaporated to dryness at 50°C. The dry residues were dissolved in 2 ml n-hexane than was added 1 ml sample dilution buffer and mixed vigorously for 30 sec. After mixing the sample was centrifuged for 10 min at 4000 rpm at room temperature (20-25°C). The aqueous (lower) phase was transferred into a new vial.

Two 20 μ L aliquots of this extract were used for the ELISA test. Data were analyzed using a special software RIDAWIN ELISA (R-Biopharm, Darmstadt, Germany). The mean absorbance values obtained for the standards and the samples divided by the absorbance value of the first standard (zero standard) and multiplied by 100 was the % absorbance. The zero standard was thus made equal to 100% and the absorbance values were quoted in percentages. The method recovery was evaluated by fortifying negative meat samples with clenbuterol standards (0.1, 0.5, 1.0 and 5.0 μ g/kg).

Validation of the ELISA method

The GC-MS method for determination of clenbuterol was validated for the recovery and repeatability. Clenbuterol-free meat samples were spiked at four concentration levels in the range from 0.1 to 5.0 μ g/kg.

Clenbuterol analysis by GC-MS (confirmatory method)

Briefly, minced meat samples (1 g) were homogenized with 1 mL of tris-buffer (50 mM, pH 8.5). Then, 2 mL of n-heptane were added, vortexed for 2 min and centrifuged at 10000 rpm for 15 min at 4 °C. The upper organic layer was discarded and the extraction was repeated. Concentrated perchloric acid (0.5 mL) was added to the meat extract, mixed for 20 min and centrifuged for 15 min

at 10000 rpm at 4 °C. The supernatant was collected in a tube containing 300 mL of 1M NaOH and mixed for 5 min. Then, 4 mL of 500 mM KH₂PO₄-buffer (pH 3) were added, the pH was adjusted to pH 6 and the extract was stored at 4 °C for 1 h. Finally, the entire supernatant was purified by solid-phase extraction on a C18 column. The C18 column was first rinsed with 6 mL of 100% methanol. After the methanol was eluted, the column was equilibrated with 2 mL of 1 M phosphate (KH₂PO₄) buffer (pH 3). Then, the sample was loaded, and the column was rinsed with 2 mL of 1 M phosphate (KH₂PO₄) buffer (pH 6). The sample was eluted with 1.5 mL of 100% methanol and the eluate was evaporated to dryness under continuous flow of nitrogen at 45 to 50 °C.

Derivatization

The sample residue was dried before derivatization. Then 50 µL of 1% TMCS/BSTFA was added, vortex, mixed and kept and heated at 80 °C for 60 min. Once the derivatization process was completed, the solution was evaporated to dryness under continuous flow of nitrogen and reconstituted in 50 µL of toluene. Finally, 2 µL of this solution were injected into the gas chromatograph. The series standard solutions were derivatized simultaneously.

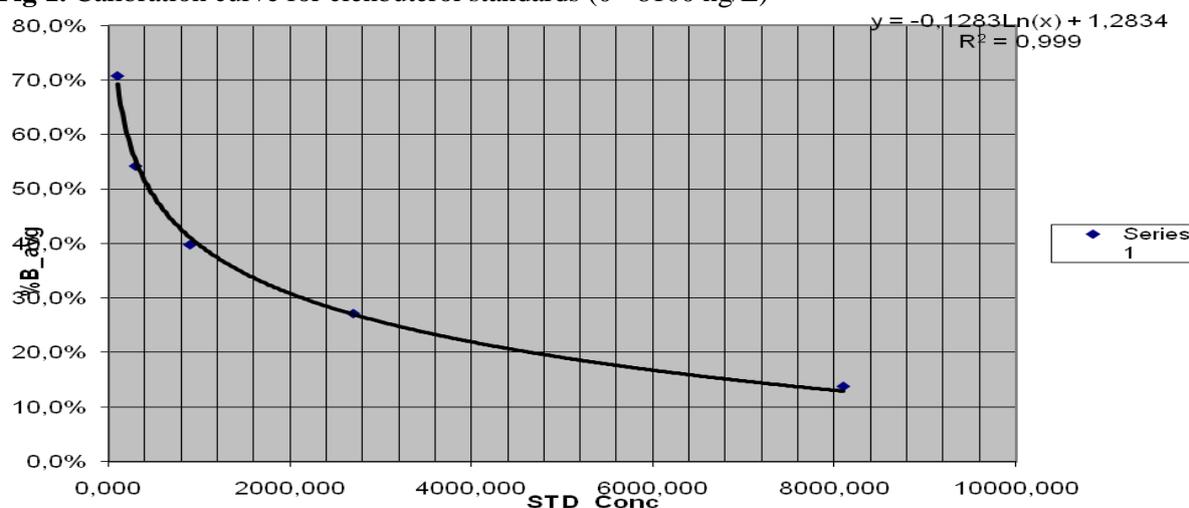
GC-MS conditions

GC-MS analysis was carried out with an Agilent 7890 A Gas chromatograph and Agilent 5975 C mass selective detector. GC was equipped with electronic pressure control unit, auto injector G 4513 A. Injection was done in splitless mode onto a 30m x 0.25 mm i.d. fused silica capillary column (HP-1 MS) with a 0.25 µm film thickness. The initial oven temperature was 70°C and was programmed to 200°C at 25°C/min, hold for 6 min and finally to 300°C at 25°C/min and hold for 5 min. Helium was used as carrier gas with a constant flow of 1.5 ml/min. Injection port and transfer zone temperatures were 250°C and 280°C, respectively. The mass spectrometer conditions were as follows: electron impact ionization voltage 70 eV for both SCAN and selected ion monitoring (SIM) mode.

Results

In this study a commercial ELISA kit was used for presumptive clenbuterol detection and quantification, however performance of the kit was not clearly described by the manufacturer and validation of the actual performance of the immunoassay with the sample matrix was carried out. The obtained calibration curves for the ELISA method in the range 100-8100 ng/L in linear and exponential form are presented on Figure 1. The curve equation $y = -0.1283 \cdot \ln(x) + 1.2834$, where y was relative absorbance (%) and x was clenbuterol concentration in ng/L, was utilized for determining clenbuterol concentration in meat samples, obtaining high regression coefficient ($R^2 = 0.999$).

Fig 1. Calibration curve for clenbuterol standards (0 - 8100 ng/L)



The ELISA method performances were confirmed through the validation experiments for the recovery and repeatability (Table 1). Good recoveries of 72.0 to 84.2% were observed at levels of 0.1 to 5.0 µg/kg of clenbuterol fortification, followed by coefficient of variation lower than 6%.

Table 1. Validation procedure for meat samples fortified with clenbuterol standard

Validation parameter	No. of replicates	Spiked concentration (µg/kg)	Determined concentration (µg/kg)	Mean recovery (%)	Coefficient of variation (%)
Recovery	6	0.10	0.072	72.0	0.82
	6	0.50	0.421	84.2	3.32
	6	1.00	0.799	79.9	1.50
	6	5.00	3.991	79.8	3.01
Repeatability	18	0.10	0.075	74.9	2.82
	18	0.50	0.389	77.8	1.42
	18	1.00	0.773	77.3	5.73
	18	5.00	4.015	80.3	1.60

55 meat samples were tested applying the validated screening method. Twenty one of them were considered to be potentially positive, and further were subjected to a confirmatory GC-MS method. On Figures 3 and 4 a typical selected ion monitoring (SIM) chromatogram for clenbuterol-free and positive meat sample, respectively, are presented.

Fig 3. Typical GC-MS chromatograms of TMS clenbuterol derivates – blank meat sample

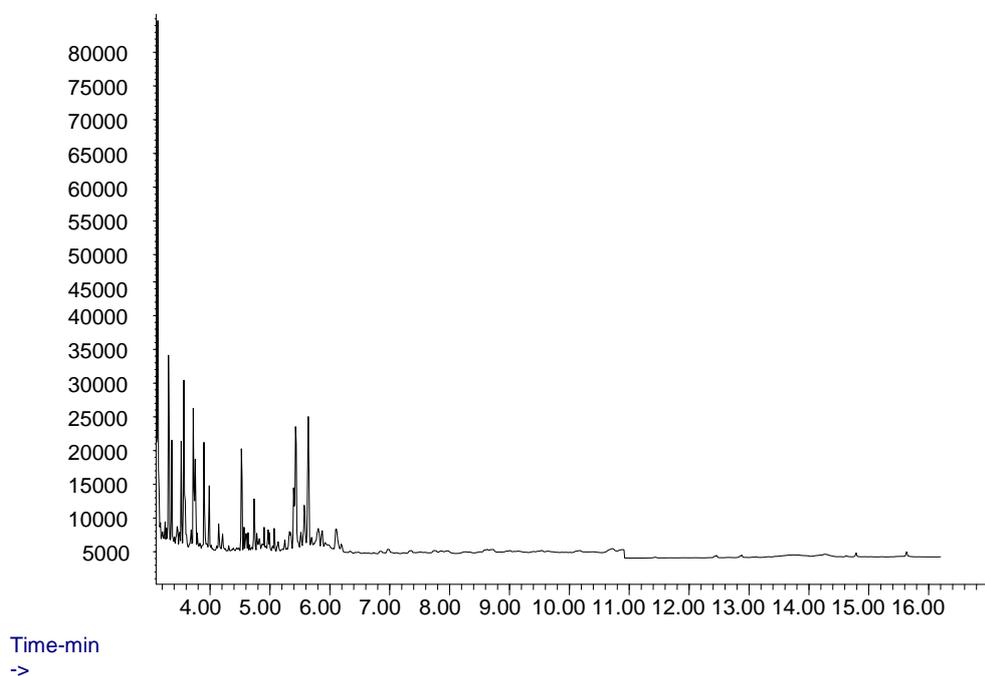
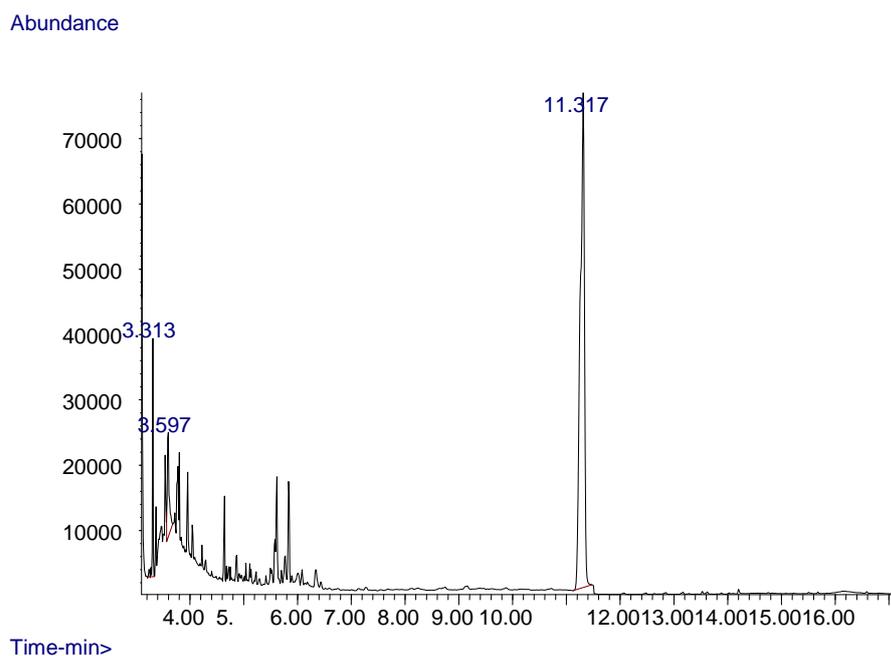
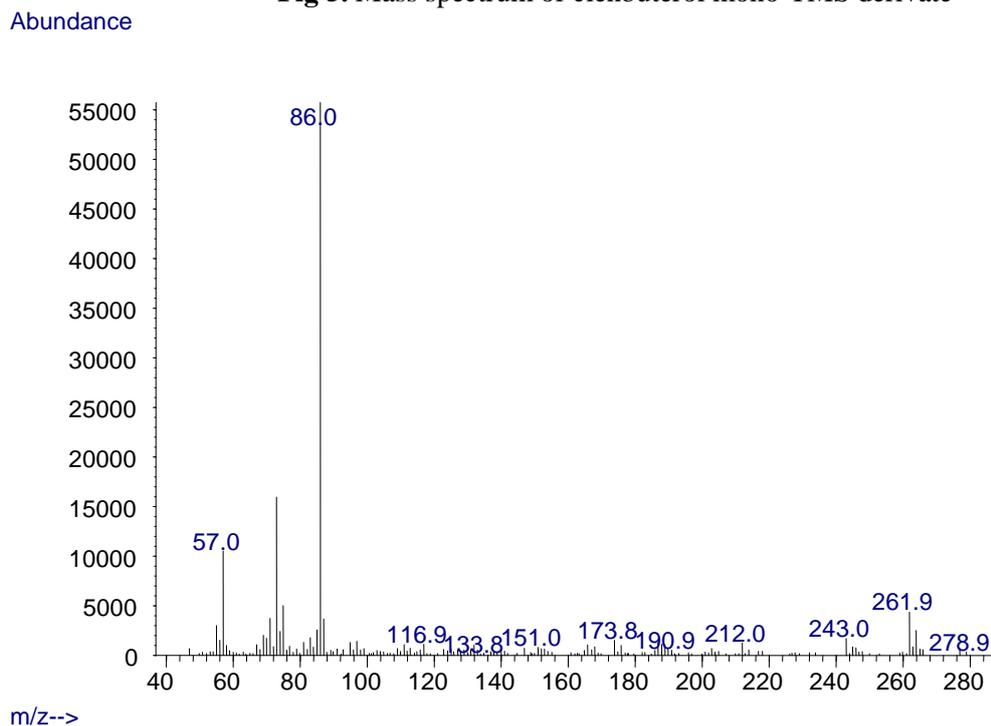


Fig 4. Typical GC-MS chromatograms of TMS clenbuterol derivates – positive meat sample



Clenbuterol was confirmed by comparison of the mass spectra of the reference standard solution of clenbuterol and analyzed samples, following the criteria laid down in Commission Decision (93/256/EEC). In electron impact mode the spectrum of the clenbuterol derivative, obtained by derivatisation with BSTFA +1 % TMCS shows the following diagnostic ions at m/z : 57, 86, 212, 243. (Figure 5).

Fig 5. Mass spectrum of clenbuterol mono-TMS derivate

Discussion

Determination of clenbuterol in bovine meat by ELISA (screening test)

The ELISA technique showed very high correlation between clenbuterol concentration (in $\mu\text{g}/\text{kg}$) and relative % of absorbance. The calibration curve constructed for the ELISA determinations followed a relationship with a highly significant ($P < 0.01$) coefficient of multiple determination ($R^2 = 0.999$). The efficiency of clenbuterol extraction by the ELISA technique was evaluated by analyzing clenbuterol fortified meat samples. Good recoveries of 72.0 to 84.2% were observed at levels of 0.1 to 5.0 $\mu\text{g}/\text{kg}$ of clenbuterol fortification (Table 1). The only study that reported clenbuterol recovery from meat products was much lower (63%) at a fortification level of 0.4 $\mu\text{g}/\text{kg}$ and determined by LC-MS-MS (13). Similarly, clenbuterol recoveries of 44-75% from urine samples spiked with 0.2 to 1.5 $\mu\text{g}/\text{kg}$ were detected by using ELISA (14). On the other hand, very good clenbuterol recoveries were (93-94%) reported previously for bovine liver samples fortified with 0.5 to 2.0 $\mu\text{g}/\text{kg}$ and analyzed by GC-MS-MS (10) or those (95-113%) reported for liver samples fortified with 1.0-2.0 $\mu\text{g}/\text{kg}$ and analyzed by ELISA (15). Therefore, for the quantitative determination of clenbuterol either in animal tissues or urine, it is important to perform recovery studies to ensure the accuracy of the method, since most important for the screening test is to obtain an acceptable low number of false negative results. A large number of false negative results were reported by the European residue control laboratories as a result of an interlaboratory study (16). On the other hand, other studies reported high rates of false positives in meat and urine samples, when the presumptive ELISA results were confirmed by GC-MS (17). It has been reported that commercial ELISA kits for the analysis of clenbuterol present cross reactivity with other β -agonists of similar structures (6).

Indeed, high cross reactivities were reported for the ELISA kit used in this study. Thus, confirmation of the ELISA test should always be carried out by chromatographic methods coupled to spectrometric methods (18). The limit of detection (LOD) for the ELISA test was at the lowest concentration of clenbuterol standard used in the standard curve (0.1 µg/kg), although the manufacturer reported a LOD of 0.3 µg/kg. The reproducibility of the technique was very good since coefficients of variation determined for clenbuterol standards and meat samples were lower than 0.6 and 2.5%, respectively. From 55 bovine meat samples analyzed by the ELISA test, 1 (1.8%) of presented clenbuterol residues was 1.19 µg/kg, 7 (12.7%) from 0.5 to 1.0 µg/kg, 13 (23.64%) from 0.1 to 0.5 µg/kg and 34 (61.82%) were below the limit of detection (LOD) of 0.1 µg/kg (Table 1). According to the maximum residue limited (MRL) recommended by the WHO and the *Codex Alimentarius* of 0.1 µg/kg, 21 (38.18%) samples would be considered presumptive positive samples. However, the samples analyzed in this study presented clenbuterol concentrations much lower than those reported for cases of food poisoning documented in several countries (19-22). Samples implicated in documented cases of food poisoning in different countries reported clenbuterol concentrations of 161 to 291 µg/kg (3), 375 to 500 µg/kg (22) and 19 to 5395 mg/kg (23) in veal liver or 450 µg/kg (21) and 800 to 7400 µg/kg (24) in meat samples. These documented cases presented clenbuterol concentrations much higher than the MRL set by the EC. For this reason, twenty one samples (38.18%) with the clenbuterol residues above LOD (0.1 µg/kg) found in this study were subjected to confirmatory analysis by GC-MS.

Identification of clenbuterol in bovine meat by GC-MS (confirmatory method)

Clenbuterol was derivatized with 1% TMCS/BSTFA, to form a mono-trimethylsilyl (TMS) derivative which gave a typical most abundant precursor ion of $m/z = 86$. This fragment ion was selected as the primary ion, and it was utilized to confirm unequivocal presence of clenbuterol at the retention time of 11.317 min (Figure 4). The other identification product ions, when the clenbuterol-TMS was subjected to MS, were fragments with $m/z = 57, 212, \text{ and } 243$. A clenbuterol-free meat sample analyzed by GC-MS showed the absence of the clenbuterol peak (Figure 3). On the other hand, a typical chromatogram of a clenbuterol fortified meat sample analyzed by GC-MS showed the clenbuterol peak at retention time of 11.317 ± 0.02 min (Figure 4).

In this study, the twenty one meat samples analyzed by GC-MS only one sample was confirmed positive for clenbuterol. For the confirmation of clenbuterol residues in the meat samples, the relative intensities of the product ions corresponded to those of the standard analyte ($\pm 10\%$). According to the European Commission Decision 2002/657/EC (18), the relative abundances of all diagnostic ions monitored for the analyte should match those of the standard analyte or from spiked samples at comparable concentrations, preferably within a margin of $\pm 10\%$. The illegal use of clenbuterol as a growth promoting agent has been well documented to have occurred in Europe, Asia, and North and Central America, spanning a time frame of nearly a decade. Much of the detected illegal use of clenbuterol in Europe have occurred in spite of the European ban of all anabolic compounds in animal production. Extensive illegal use of clenbuterol and others β -agonists have not been halted by the ban, and the problem of unsafe residues in food remains. In fact, the most recent food poisoning outbreak caused by clenbuterol residues in meat was reported in Portugal in 2005 (20). Although samples were collected from the authorized veterinarian inspectors and mostly it comes from meat samples collected from different farms and slaughterhouses in Macedonia, but still remain questionable from meat samples who are imported from the different countries in Macedonia. Thus, this initial study should be useful as a first attempt to gain some insight in the illegal use of clenbuterol by some Macedonian meat producers.

References:

- Reeds PJ, Hay SM, Dorward PM, Palmer RM: The effects of beta-agonists and antagonists on muscle growth and body composition of young rats (*Rattus* sp). *Comp Biochem Physiol*, 89, 337-341, 1988.
- Dixon SN, Tennant DR, Kay, JF: Veterinary drug residues. In, Watson DH (Ed): Safety of chemicals in foods. 1st ed. 24-29, Ellis Horward, New York, 1993.
- Martínez-Navarro, JF: Food poisoning related to the consumption of illicit beta-agonists in liver. *Lancet*, 336, 1311, 1990.
- Codex Alimentarius CL 1996/27-RVDF and ALINORM 97/31A

Council Regulation 2377/90/EEC: Laying down a procedure for the establishment of maximum residues limits of veterinary medicinal products in foodstuffs of animal origin. OJ L 224/1, 18.08.1990.

Kuiper HA, Noordam MY, van Dooren-Flipsen, MMH, Schilt, R, and Ross, AH: Illegal use of b-adrenergic agonists: European Community. *J Anim Sci*, 76, 195-207, 1998.

Cai J, Henion J: Quantitative multi-residue determination of b-agonists in bovine urine using on-line immunoaffinity extraction-coupled column packed capillary liquid chromatography-tandem mass spectrometry. *J Chromatogr B Biomed Sci Appl*, 691 (2): 357-370, 1997.

Polettini A, Montagna M, Hogendoorn EA, Dijkman E, Van Zoonen P, Van Ginkel LA: Applicability of coupled-column liquid chromatography to the analysis of b-agonists in urine by direct sample injection. I. Development of a single-residue reversed-phase liquid chromatography-UV method for clenbuterol and selection of chromatographic conditions suitable for multi-residue analysis. *J Chromatogr A*, 695, 19-31, 1995.

Whaites LX, Murby EJ: Determination of clenbuterol in bovine urine using gas chromatography-mass spectrometry following clean-up on an ion-exchange resin. *J Chromatogr B Biomed Sci Appl*, 728, 67-73, 1999.

Fiori M, Cartoni C, Bocca B, Brambilla G. The use of nonendcapped C18 columns in the cleanup of clenbuterol and a new adrenergic agonist from bovine liver by gas chromatography-tandem mass spectrometry analysis. *J Chromatogr Sci*, 40, 92-96, 2002.

Zampronio CG, Gurden SP, Moraes LA, Eberlin MN, Smilde AK, Poppi RJ: Direct sampling tandem mass spectrometry (MS/MS) and multiway calibration for isomer quantitation. *Analyst*, 127 (8):1054-1060, 2002.

De Wasch K, Brabander HD, Courtheyn D: LCMS- MS to detect and identify four b-agonists and quantify clenbuterol in liver. *Analyst*, 123, 2701- 2705, 1998.

Guy PA, Savoy MC, Stadler RH: Quantitative analysis of clenbuterol in meat products using liquid chromatography-electrospray ionization tandem mass spectrometry. *J Chromatogr B Biomed Sci Appl*, 736 (1-2): 209-219, 1999.

Sawaya WN, Lone K, Husain A, Dashti B, Saeed T: Screening for b-agonists in sheep urine and eyes by an enzyme-linked immunosorbent assay in the state of Kuwait. *Food Control* 11, 1-5, 2000.

Degand G, Bernes-Duyckaerts A, Maghuin-Rogister G: Determination of clenbuterol in bovine tissue and urine by enzyme immunoassay. *J Agr Food Chem*, 40, 70- 75, 1992.

Wolf C, Behrendt D, Conradi B, Julicher B: Results of two european interlaboratory studies for the determination of b-agonists in various matrices. *Analyst*, 123, 2719-2723, 1998.

Boyd D, O'Keefe M, Smith MR: Methods for the determination of b-agonists in biological matrices. *Analyst*, 121, 1R-10R, 1996.

Commission Decision: 2002/657/EC of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of the results, OJ L 221, 17.08.2002.

Estrada-Montoya MC, González-Córdova AF, Torrescano G, Camou JP, Vallejo-Cordoba B: Screening and confirmatory determination of clenbuterol residues in bovine meat marketed in the northwest of Mexico. *Cienc Technol Aliment*, 6(3): 130-136, 2008.

Barbosa J, Cruz C, Martins J, Silva JM, Neves C, Alves C, Ramos F, da Silveira MIN: Food poisoning by clenbuterol in Portugal. *Food Addit Contam*, 22, 563-566, 2005.

Brambilla G, Loizzo A, Fontana L, Strozzi M, Guarino A, Sporano V: Food poisoning following consumption of clenbuterol-treated veal in Italy. *J Am Med Assoc*, 278 (8): 635, 1997.

Pulce C, Lamaison D, Keck G, Bostvironnois C, Nicolas J, Descotes J: Collective human food poisonings by clenbuterol residues in veal liver. *Vet Hum Toxicol*, 33, 480-481, 1991.

Salleras L, Dominguez A, Mata E, Taberner JL, Moro I, Salva P: Epidemiologic of an outbreak of clenbuterol poisoning in Catalonia, Spain. *Public Health Rep*, 110, 338-342, 1995.

Sporano V, Grasso L, Esposito M, Oliviero G, Brambilla G, Loizzo A: Clenbuterol residues in non-liver containing meat as a cause of collective food poisoning. *Vet Hum Toxicol*, 40, 141-143, 1998.

ANTIHYPERGLYCEMIC AND LIPID LOWERING ACTIVITIES OF ETHANOLIC EXTRACT OF *ERIOBOTRYA JAPONICA* SEEDS IN ALLOXAN INDUCED DIABETIC RATS

Sabeeha Shafi
Nahida Tabassum

Department of Pharmaceutical Sciences, University of Kashmir, Hazratbal, Srinagar, Kashmir, India

Abstract:

Diabetes is a metabolic disorder characterized by hyperglycemia, hypertriglyceridemia and hypercholesterolemia, resulting from defects in insulin secretion or action or both. India is known as the diabetic capital of the world. The study of plants having antihyperglycemic and hypolipidemic activities gives a new approach in the treatment of diabetes mellitus. The study was carried out to evaluate the antidiabetic and hypolipidemic activity of ethanolic extract of seeds of *Eriobotrya japonica* (EBS) in alloxan induced diabetic albino rats. Diabetes was induced in albino rats by administration of alloxan monohydrate (120 mg/kg) by intraperitoneal route. Rats were divided into 6 groups of 6 animals each. Group I served as non-diabetic control, Group II as diabetic control, Group III^d received antidiabetic standard drug (10 mg/kg of glibenclamide) Group IV and Group V received 100 and 200 mg/kg b.w of EBS. Blood samples were analysed for blood glucose on day 1,4,7 and day 10th and lipid profile was analysed on day 10. All the values are expressed as Mean±SEM. The results were subjected to statistical analysis using one-way ANOVA followed by students t test. p<0.001 was considered highly significant. The ethanolic extract of EBS at the dose of 200 mg/kg body weight showed highly significant reduction in blood glucose and serum lipid profile levels in alloxan induced diabetic rats. It is concluded that ethanolic extract of EBS is effective in controlling blood glucose levels and in improving lipid profile in diabetic rats.

Key Words: Diabetes mellitus, blood glucose level, Lipid Profile, *Eriobotrya japonica* seeds.

Introduction

Diabetes mellitus is a major degenerative disease in the world today affecting at least 15 million people. Diabetes mellitus is associated with long term complications, including retinopathy, nephropathy, neuropathy and angiopathy and several others. It is a multifactorial disease which is characterized by hyperglycemia, lipoprotein abnormalities, raised basal metabolic rate, defect in reactive oxygen species scavenging enzymes and high oxidative stress induced damage to pancreatic beta cells. It is ranked third among the leading causes of death when its fatal complications are taken into account. Today in India alone there are more than 4.00 crore diabetics and the number is going to be around 9.00 crore by 2030. Over 7.20 lakh Indians die every year due to diabetes. People with diabetes are 2-4 times more likely to develop heart diseases¹. Efforts are ongoing to understand and manage diabetes mellitus because the disease and disease related complications are increasing day by day. In spite of presence of large number of medicines in the pharmaceutical market, remedies from medicinal plants are used with success to treat and this disease. India has 45,000 plant species and several thousand have medicinal properties. More than 800 plant species have anti-diabetic activity. There has been great demand for plant products due to low cost, easy availability and lesser side effects. For this plant materials are continuously scrutinized and explored for their effect as antidiabetic agents.

One of the plants is *Eriobotrya japonica* locally known as loquat, has been used since olden times in the ethno medicine for treating diseases. The plant is reported to possess antioxidant^{2,3,4} antiviral^{5,6} cytotoxic⁷ hepatoprotective⁸ anti-inflammatory/antitussive activity^{9,10,11}. There is dearth of reports on the antidiabetic and hypolipidemic effects of the seeds of this plant.

The present study was aimed to investigate antidiabetic and hypolipidemic activity of ethanolic extract of *Eriobotrya japonica* seeds in alloxan induced diabetic rats.

Materials And Methods

Plant Material:

The fruits of *Eriobotrya japonica* (family Rosaceae) were collected from Shalimar area of the district, Srinagar, during the months of April to June and authenticated by a plant taxonomist in the Centre of Plant Taxonomy, University of Kashmir, Srinagar. The identification was done on the basis of the characters described by Kirtikar and Basu, 1935. A sample of the plant material was deposited in the herbarium of the Department of Taxonomy, University of Kashmir under voucher specimen number 1012(KASH) dated 15-09-2008 for future reference. The seeds were separated from fruits, and dried in a well ventilated room with outside temperature ranging between 18 to 32^oC.

Preparation of the extract:

The seeds were coarsely powdered and 500 gm of the material was allowed to macerate for 48 hrs with 50% ethanol, with occasional shaking. After 48 hrs, the ethanolic extract was filtered through Whatmans filter paper. The plant material was then macerated again with fresh 50% ethanol and the filtrate obtained from the first and the second maceration was then combined and the solvent was recovered. After the recovery of alcohol, the extract was then evaporated to dryness and the yield was noted. The procedure was repeated with 500 gm of the material again. The extract was refrigerated at 4^oC for future use in experimental studies.

Phytochemical Screening:

The extract obtained was subjected to qualitative tests for identification of different constituents like tannins, alkaloids, saponins, glycosides, terpenes, phenolics, flavonoids, carbohydrates, proteins and steroids, by using simple and standard qualitative methods described by Trease and Evans¹²

Pharmacological Study:

Animals

Healthy albino rats of either sex weighing about 180-210 g were used during the study. The animals were procured from Central Animal House, IIM (Indian Institute of Integrative Medicine) Jammu & were housed in clean polypropylene cages. Before initiation of experiment, the rats were acclimatized for a period of 7 days. Standard environmental conditions such as temperature ranging from 18 to 32^oC, relative humidity (70%) and 12 hrs dark/light cycle were maintained in the quarantine. All the animals were fed with rodent pellet diet (Ashirwad Industries) and water *ad-libitum* under strict hygienic conditions. All procedures were performed in accordance to CPCSEA guidelines after approval from the Institutional Animal and Ethics Committee (IAEC) of the Department of Pharmaceutical Sciences, University of Kashmir [No. F-IAEC (Pharm.Sc) APPROVAL/2008/4 Dated Oct 23rd, 2008]

Induction of diabetes:

Alloxan monohydrate was obtained from S.D Fine Chemical, Mumbai, India. All the other chemicals used were of analytical grade and were acquired from commercial sources. A single dose (120mg/kg, b.w, i.p) of alloxan monohydrate in sterile saline was used for the induction of diabetes in rats after overnight fasting. After one hour of alloxan administration, the animals were fed standard pellets and water *ad libitum*. After 5 days of alloxan administration, animals showing blood glucose levels above 250 mg/dl were selected for the study. Extract of EBJs was administered at two dose levels 100 and 200 mg/kg¹³

Experimental design:

Rats fasted overnight for 12 hrs were randomly divided into 6 groups of 6 rats per group. The various groups were:-

Group I- Served as normal control and received only 0.2 ml of 2% aqueous gum acacia

Group II- Served as diabetic control and received only alloxan monohydrate and 2% aqueous gum acacia.

Group III- Alloxan monohydrate + Glibenclamide (10 mg/kg, p.o) and served as Standard Antidiabetic drug.

Group IV- Alloxan monohydrate + 50% Ethanolic extract of EBJs (100 mg/kg, p.o)

Group V- Alloxan monohydrate +50% Ethanolic extract of EBJs (200 mg/kg, p.o)

The treatment (p.o) of the ethanolic extract was started the same day except normal control and diabetic control groups which received only 0.2 ml of 2% aqueous gum acacia for a period of 10 days. During this period, animals in all groups had free access to standard diet and water. Body weight and blood glucose levels were estimated on 1st, 4th, 7th and 10th day of the treatment.

Sample Collection:

Blood samples were collected by pricking the tail from overnight fasted rats and blood glucose levels were estimated using One Touch Ultra glucose strips (Johnson & Johnson Ltd) on 1st, 4th, and 7th day.

Estimation of biochemical parameters:

On day 10th, blood was collected from overnight fasted rats under ether anesthesia by cardiac puncture and was kept aside for 30 min for clotting. By centrifuging the same sample at 6000 rpm for 20 min, the serum was separated and was analyzed for blood glucose^[14], total cholesterol^[15], triglycerides^[16], HDL cholesterol^[17] and LDL cholesterol^[18]

Statistical analysis:

All the values are expressed as mean±SEM. The results were subjected to statistical analysis using one-way ANOVA followed by students t test. $p < 0.001$ was considered highly significant.

Results

Phytochemical analysis:

Phytochemical analysis of the extract showed the presence of alkaloids, flavonoids and glycosides (Table-1)

Antihyperglycemic study:

Blood glucose levels showed a highly significant decrease in groups III, IV, V ($p < 0.001$) as compared to group II (Diabetic control). A highly significant increase in blood glucose levels was seen in diabetic group as compared to normal control group I ($p < 0.001$). (Table 2)

Effect of 50% ethanolic extract of *Eriobotrya japonica* seeds on biochemical parameters in alloxan induced diabetic rats.

Serum total cholesterol levels showed a highly significant decrease in groups IV and V ($p < 0.001$) as compared to diabetic control (Group II). Serum triglyceride levels showed a highly significant decrease in groups IV and V ($p < 0.01$). HDL levels showed a non significant increase in groups III, IV and V. LDL levels showed a significant decrease in groups IV and V groups. (Table 3)

Effect of 50% ethanolic extract of *Eriobotrya japonica* seeds on body weight in alloxan induced diabetic rats

Normal control animals were found to be stable in their body weight but diabetic rats showed significant reduction in body weight after 10 days. ($p < 0.001$) Alloxan mediated body weight reduction was reversed by the ethanolic extract in dose dependant fashion 100 mg/kg and 200 mg/kg b.w showed a highly significant increase in body weight ($p < 0.001$). The effect of extract at 200 mg/kg on body weight of the animals was also found statistically significant. Results are shown (Table 4)

DISCUSSION

Pancreas is the primary organ involved in sensing the organism's dietary and energetic states via glucose concentration in the blood and in response to elevated blood glucose, insulin is secreted. Alloxan is one of the usual substances used for the induction of diabetes mellitus apart from streptozotocin. It has a destructive effect of the beta cells of the pancreas^[19] Alloxan causes a massive reduction in insulin release by the destruction of beta-cells of the islets of Langerhans thereby inducing hyperglycemia^[20]. Insulin deficiency leads to various metabolic alterations in the animals viz increased blood glucose and increased lipid profile.

The results of the present study found that 50% ethanolic extract of *Eriobotrya japonica* seeds reduce the glucose level in animals made diabetic with alloxan. Alloxan has been shown to induce free radical production and cause tissue injury. The pancreas is especially susceptible to the action of alloxan induced free radical damage. In the present investigation 50% ethanolic extract of *Eriobotrya japonica* seeds demonstrated the significant antihyperglycemic activity. The results from the present study also indicate that ethanolic extract can reduce the levels of serum lipids. The antihyperglycemic effect of the ethanolic extract may be due to the enhanced secretion of insulin

from the beta cells of pancreas or may be due to increased tissue uptake of glucose by enhancement of insulin sensitivity.

Elevated plasma total cholesterol, triglycerides and LDL cholesterol are the major risk factors of cardiovascular diseases. Diabetic rats showed elevated plasma cholesterol, triglycerides and LDL cholesterol. Ethanolic extract in the dose of 200 mg/kg reduced the lipid profile along with the reduction in the blood glucose levels.

The literature reports reveal that flavonoids present in the plant extract known to possess antihyperglycemic and hypolipidemic activity. In the present investigation also the observed antihyperglycemic and hypolipidemic potential of test extract may be due to presence of similar phytoconstituents which was evident by preliminary phytochemical screening. Since many antihyperglycemic drugs do not correct dyslipidemia, the observed hypolipidemic effects of the plant extract in diabetic rats makes EBJS quite important in the management of diabetes. Since there is a strong well-established link between diabetes mellitus, dyslipidemia, obesity, hypertension and ischemic heart disease, effect of the plant extract on weight loss/gain needs to be explored on scientific base.

Conclusion

From the study, it can be concluded that the 50% ethanolic extract of *Eriobotrya japonica* seeds has beneficial effects on blood glucose levels as well as improving hyperlipidemia and other metabolic aberrations. Further pharmacological and biochemical investigations will clearly elucidate the mechanism of action and will be helpful in projecting this plant as an therapeutic target in diabetics research.

Acknowledgement

We are highly thankful to Sri Krishna Drugs Ltd., C-4 Industrial Area Uppal, Hyderabad for providing a free gift pure sample of Glibenclamide which was used as standard anti diabetic drug and also to University Grants Commission for financial assistance. The facilities provided by the Department of Pharmaceutical Sciences University of Kashmir for carrying out this work also need appreciation.

References:

1. Alberti KG, Zimmet PZ. Definition diagnosis and classification of diabetes mellitus and its complications. Part I: Diagnosis and classification of diabetes mellitus, provisional report of a WHO consultation. *Diab Med* 1998; 15: 539-553.
2. Hamada A, Yoshioka S, Takuma D, Yokota J, Cui T, Kusunose M, et al . The effect of *Eriobotrya japonica* seed extract on oxidative stress in adriamycin-induced nephropathy in rats. *Biol and Pharm Bull* 2004; 27: 1961-1964.
3. Huang Y, Li J, Cao Q, Yu SC, Lv XW, Jin Y, et al. Anti-Oxidative effect of triterpene acids of *Eriobotrya japonica* (Thumb) Lindl Leaf in chronic bronchitis rats. *Life Science* 2006; 78:2749-2757.
4. Yokoto J, Takuma D, Hamada A, Onogawa M, Yoshioka S, Kusenose M, et al Scavenging of reactive oxygen species by *Eriobotrya Japonica* seed extract. *Biol and Pharm Bull* 2006; 29: 467-471.
5. Know HK, Kang MJ, Kim HJ, Choi JS, Paik KJ, Chung HY. Inhibition of NF kappaB by methyl chlorogenate from *Eriobotrya japonica*. *Molecules and Cells* 2000; 10: 241-246.
6. Baljinder S, Seena G, Dharmendra K, Vikas G, Bansal P. Pharmacological Potential of *Eriobotrya japonica*- An overview. *Int Res J of Pharm* 2010; 1: 95-99.
7. Ito H, Kobayashi E, Takamasto Y, Li SH, Hatano T, Sakagami H, et al . Polyphenols from *Eriobotrya japonica* and their cytotoxicity against human oral tumour cell lines. *Chem Pharm Bull* 2000; 48(5): 687-693.
8. Nishioka Y, Yoshioka S, Kusunose M, Cui T, Hamada A, Ono M, et al. Effects of extracts derived from *Eriobotrya japonica* on liver function improvement in rat. *Biol and Pharm Bull* 2002 ; 25 : 1053-1057.
9. Shimizu M, Fukumura H, Tsuji H, Tanaami S, Hayashi T, Morita N. Anti-inflammatory constituents of topically applied crude drugs constituents and anti-inflammatory effect of *Eriobotrya japonica* Lindl *Chem and Pharm Bull* 1986; 34: 2614-2617.

10. Ju JH, Zhou L, Lin G, Liu D, Wang LW, Yang JS. Studies on constituents of triterpene acids from *Eriobotrya Japonica* and their anti-inflammatory and anti-tussive effects. *J of Chinese Pharmacol* 2003; 38: 752-757.
11. Banno N, Akihisa T, Tokuda H, Yasukawa K, Taguchi Y, Akazawa H, et al. Anti-inflammatory and antitumor-promoting effects of the triterpene acids from the leaves of *Eriobotrya japonica*. *Biol and Pharm Bull* 2005; 28: 1995-1999.
12. Rafia Rasool, Bashir A Ganai, Seema Akbar, Azra Kamili, et al. Phytochemical screening of *Prunella vulgaris* L An important Medicinal Plant of Kashmir. *Pak J of Pharm Sci* 2010; 23(4): 399-402.
13. Vivek SK, Suresh K, Hitesh JP, Shivkumar. Hypoglycaemic activity of *Ficus glomerata* in alloxan induced diabetic rats. 2010; 1(2), 18-22
14. Trinder P. Glucose oxidase method. *Ann Clin Biochem* 1969; 6.
15. Allain CC, Poon LS, Chan CS, Richmond W, Fu PC. CHOD-PAP method for determination of total cholesterol. *Clin Chem* 1974; 20: 470-475.
16. McGowan MW, Joseph DA, Strandbergh DR, Zak B. A Peroxidase coupled method for the colorimetric determination of serum triglyceride. *Clin Chem* 1983; 29:538-542.
17. Fringe CS, Feridley TW, Dunn RT, Owan CA. Improved determination of total serum lipids by sulphosphovanillin reaction. *Clin Chem* 1972; 18: 673-674
18. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low density lipoprotein cholesterol in Plasma, without use of the preparative ultra centrifuge. *Clin Chem* 1972; 18: 499-502.
19. Szkudelski T. The mechanism of alloxan and streptozotocin action in beta cells of the rat pancreas. *Physiol Res* 2001; 50: 536-546.
20. Lenzen S, Pantenu. Alloxan: history and mechanism of action. *Diabetologia* 1988; 31: 337-342.

Table -1. Phytochemical Results of *Eriobotrya japonica* seeds**Table 1**

S.No	Phytoconstituents	Results
1	Tannins	-
2	Alkaloids	+
3	Saponins	-
4	Glycosides	+
5	Terpenes	-
6	Phenolics	-
7	Flavonoids	+
8	Carbohydrates	-
9	Proteins	-
10	Steroids	-

Table-2 Effect of 50% ethanolic extract of *Eriobotrya japonica* (EBSJ) seeds on fasting blood glucose level (mg/dl) in alloxan induced diabetic rats

Table 2

Group	Treatment	1 st day	4 th day	7 th day	10 th day
I	Normal control 0.2 ml of 2% aqueous Gum acacia	85.07± 4.35	86.16±4.43	84.82±5.96	84.71±6.11
II	Diabetic control (Vehicle) 0.2 ml of 2% aqueous gum acacia	261.47±8.37	264.28±8.29	268.03±8.48	271.33±8.18* **
III	Alloxan monohydrate + glibenclamide (10 mg/kg)	200.37±5.25	141.18±2.43	124.52±2.00**	114.84±3.21* **
IV	Alloxan monohydrate+ 50% Ethanolic extract (EBSJ,100 mg/kg)	204.27±4.25	156.99±2.48	156.21±2.85**	155.54±2.54* **
V	Alloxan monohydrate+ 50% Ethanolic extract(EBSJ,200 mg/kg)	202.95±4.95	144.67±1.56	128.22±5.55**	120.74±6.13* **

Animal: Albino Rats

Alloxan: 120 mg/kg.i.p

Extract: p.o.

Values are Mean ±S.E.M n=6; except in Group V where n=5

**p<0.01 significant

***P< 0.001 highly significant

Groups III,IV,V, vs Diabetic Control (Group II) and Group I vs Group II on 10th day

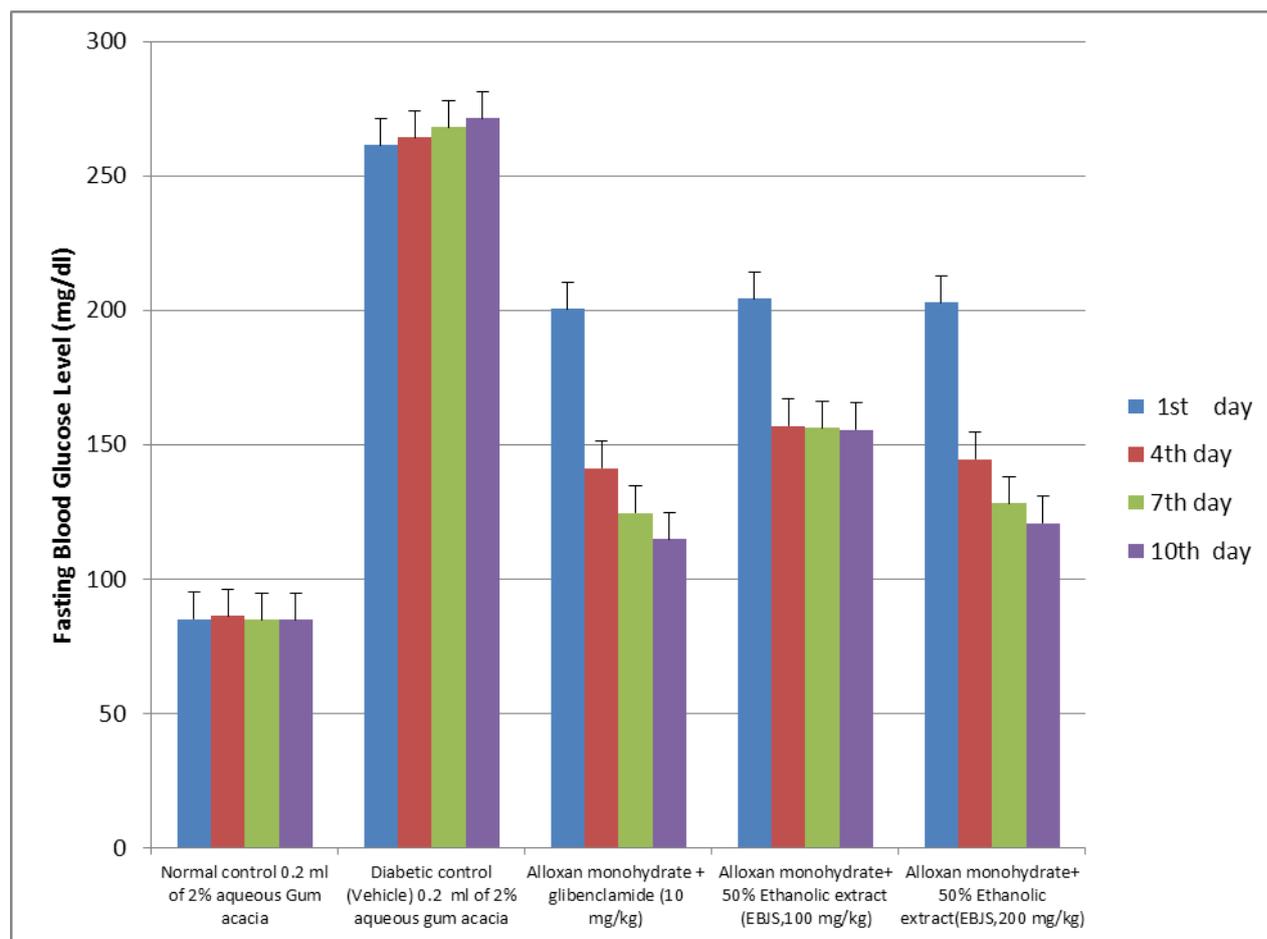


Fig 1: Effect of 50% ethanolic extract of *Eriobotrya japonica* (EBJS) seeds on fasting blood glucose level (mg/dl) in alloxan induced diabetic rats after 10 days of dosing. Each bar represents the mean \pm SEM.

Table-3. Effect of 50% ethanolic extract of *Eriobotrya japonica* seeds on lipid profile in alloxan induced diabetic rats.

Table 3

Gro uP	Treatment	Serumtotal Cholesterol mg/dl	Serumtriglyceride mg/dl	SerumHDL Cholesterol mg/dl	SerumLDL Cholesterol mg/dl
I	Normal control 0.2 ml of 2% aqueous gum acacia	85.25 \pm 8.51	80.71 \pm 9.38	24.54 \pm 6.49	52.03 \pm 3.21
II	Diabetic control 0.2 ml of 2% aqueous gum acacia	218.15 \pm 23.79**	194.56 \pm 14.99*	21.11 \pm 1.45	89.59 \pm 13.82
III	(Alloxan monohydrate +standard drug glibenclamide 10 mg/kg)	206.35 \pm 6.11*	184.30 \pm 9.68*	29.03 \pm 3.16*	85.39 \pm 7.24*
IV	Alloxan monohydrate + Ethanollic extract (EBJS, 100 mg/kg)	170.55 \pm 9.03**	127.49 \pm 12.66*	25.04 \pm 3.21*	68.91 \pm 13.34**
V	Alloxan monohydrate +Ethanollicextract (EBJS, 200 mg/kg)	121.68 \pm 14.57**	124.29 \pm 10.67**	31.29 \pm 6.12*	50.47 \pm 2.63**

Animal: Albino Rats

Alloxan: 120 mg/kg.i.p
Extract: p.o.

Value are Mean ±S.E.M: n=6 except in Group V where n =5

* p> 0.05 non significant

**p< 0.01 significant

***P< 0.001 highly significant;

Groups III,IV,V vs Diabetic Control (Group II) and Group I vs Group II on 10th day

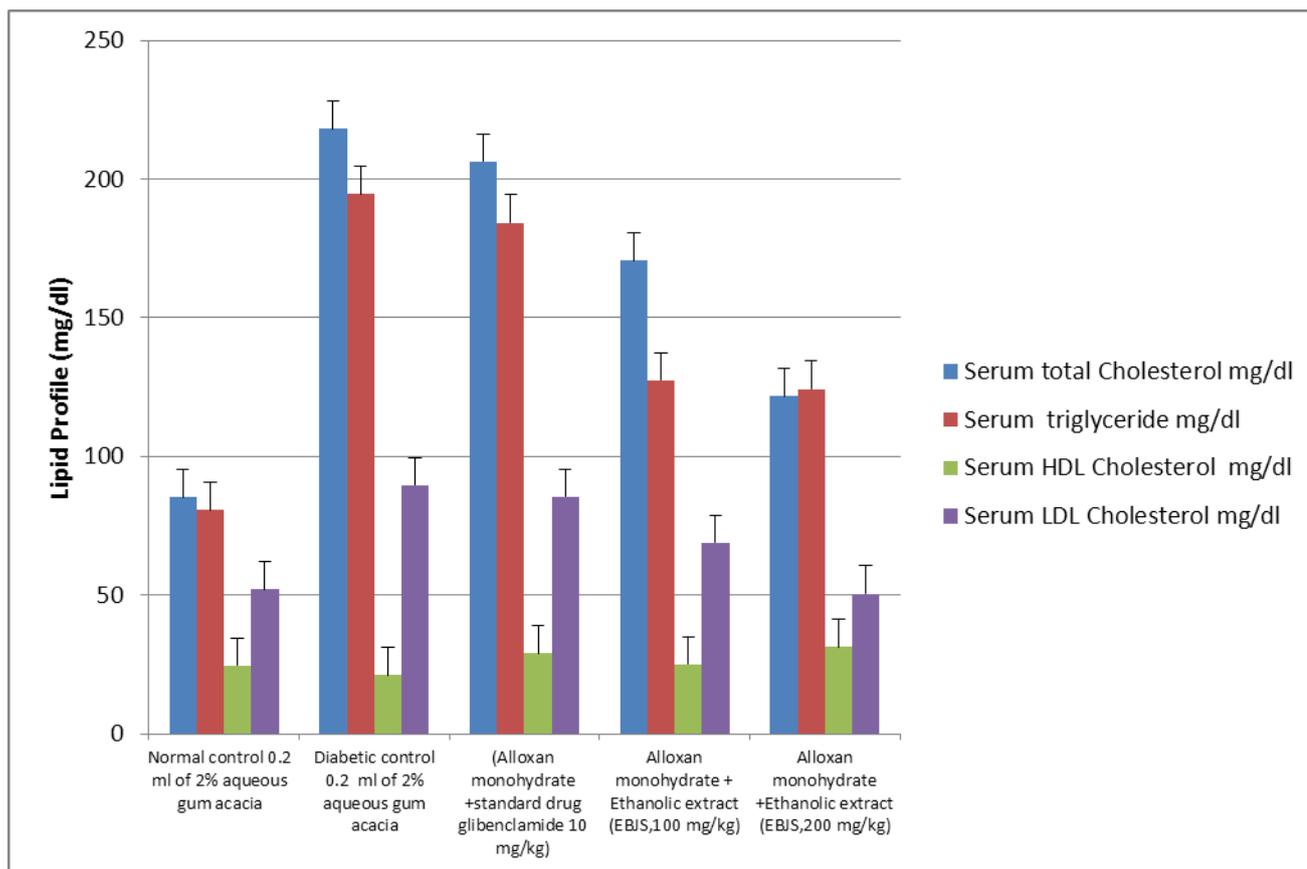


Fig 2: Effect of 50% ethanolic extract of *Eriobotrya japonica* seeds on lipid profile in alloxan induced diabetic rats after 10 days of dosing. Each data column represents the mean ± SEM.

Table 4: Effect of 50% ethanolic extract of *Eriobotrya japonica* (EBS) seeds on average body weight (g) in alloxan induced diabetic rats

Table 4

Group	Treatment	Average Body weight of the animal (g)			
		1 st day	4 th day	7 th day	10 th day
I	Normal control 0.2 ml of 2% aqueous gum acacia	228.85±6.03	229.88±6.66	233.00±7.07	222.05±4.75
II	Diabetic control 0.2 ml of 2% aqueous gum acacia	180.58±3.66** *	163.86±2.08***	152.21±3.12***	124.76±2.35***
III	Alloxan monohydrate+ standard drug glibenclamide (10 mg/kg)	183.83±3.34	173.85±3.37	152.98±4.44	137.50±3.54***
IV	Alloxan monohydrate + Ethanollic extract (EBSJ,100 mg/kg)	183.87±2.83	174.70±1.64	165.52±2.73	162.37±2.74***
V	Alloxan monohydrate + Ethanollic extract (EBSJ,200 mg/kg)	185.56±3.74	178.11±3.24***	171.43±3.41***	167.20±4.05***

Animal: Albino Rats

Alloxan: 120 mg/kg.i.p
Extract: p.o.

Value are Mean ±S.E.M: n=6 except in Group V where n =5

* p> 0.05 non significant

**p< 0.01 significant

***P< 0.001 highly significant;

Groups III,IV,V vs Diabetic Control (Group II) and Group I vs Group II on 10th day

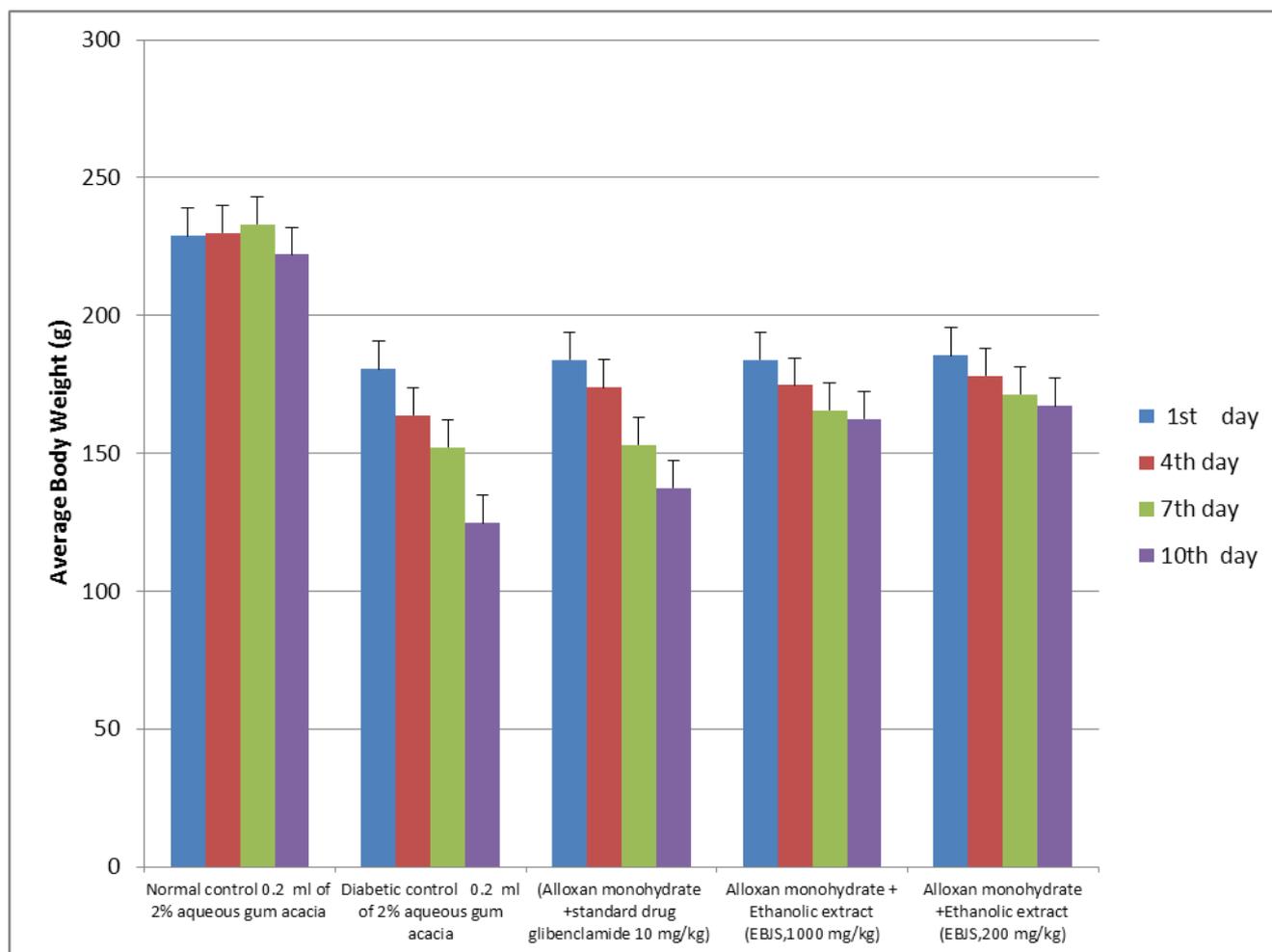


Fig 3: . Effect of 50% ethanolic extract of *Eriobotrya japonica* seeds on Average Body Weight (g) in alloxan induced diabetic rats after 10 days of dosing. Each data column represents the mean \pm SEM.

¹ Dyson, E. & Smith, G.B., 'Common faults in resuscitation equipment-guidelines for checking equipment and drugs used in adult CPR'. *Resuscitation* 55(2), 137–149; 2002. [http://dx.doi.org/10.1016/S0300-9572\(02\)00169-7](http://dx.doi.org/10.1016/S0300-9572(02)00169-7)

² Smith, A., Kinross, J., Bailey, M., Aggarwal, R., Toresen, D. & Vincent, C., 'Restocking the resuscitation trolley: how good is compliance with checking procedures'? *Clinical Risk* 14(1), 4–7; 2008. <http://dx.doi.org/10.1258/cr.2007.070008>

INFLUENCE OF E-BEAM POSTHARVEST IRRADIATION IN THE COLOUR OF FOUR EUROPEAN CHESTNUT FRUIT VARIETIES OF *CASTANEA SATIVA* MILL

Amilcar L. Antonio, Msc

Mountain Research Center (CIMO), ESA, Polytechnic Institute of Braganca, Portugal
IST/ITN Nuclear and Technological Institute, Portugal
Department of Fundamental Physics, University of Salamanca, Spain

Marcio Carocho, Msc

Albino Bento, PhD

Mountain Research Center (CIMO), ESA, Polytechnic Institute of Braganca, Portugal

Andrzej Rafalski, PhD

Centre for Rad. Research and Techn., Inst. of Nuclear Chem. and Technology, Warsaw, Poland

Begona Quintana, PhD

IST/ITN Nuclear and Technological Institute, Portugal

Abstract:

In this study we characterized the physical dimensions of four European chestnut varieties of *Castanea sativa*, 3 from Portugal (*Cota*, *Longal* and *Judia*) and 1 from Italy (*Palummina*). The typical physical dimensions of the different chestnut fruit varieties could be used to optimize the process in an irradiation preservation treatment. The chestnut fruits were submitted to an irradiation preservation treatment with electron beam, at 1 kGy, and the colour parameters (CIE L*, a*, b*) of the skins, fruit and interior (half-cutted), were monitored after irradiation treatment and along 2 months of storage. We could conclude that e-beam irradiation, for the dose of 1 kilogray, did not induce any significant change in the skin, fruit and interior's colour parameters.

Key Words: Chestnuts, *Castanea sativa* Mill., e-beam irradiation, physical parameters, colour

Introduction

In the Mediterranean countries the production of chestnut fruits is about 200 000 ton, being Portugal the third producer of *Castanea sativa* Mill varieties, with about 30 000 ton, exporting 10 000 ton, representing an income of about 14 million Euros (FAOSTAT, 2010; INE 2010). Due to phytosanitary regulations the exported fruits must be postharvest treated. Until recently, March 2010, this was done with methyl bromide (MeBr) for postharvest fumigation, now prohibited by E.U. regulations due to its toxicity for the operators and environmental concerns. Food irradiation is a well-established technology, regulated by several directives and approved by the international food and safety organizations for preservation or insect disinfestations, using gamma radiation, electron beam or x-rays (EU, 1999; FAO/IAEA/WHO, 1981). The typical dose for postharvest insects' disinfestations is lower than 1 kGy (IDIDAS, 2013). However, each irradiation process must be validated to meet the needed food quality parameters. Due to the limitation of electron penetration for e-beam irradiation, the dimensions are an important issue in the design of the process that imposes limitations in using this technology, since the authorized energy for e-beam food processing is limited to 10 MeV. Also, the throughput of the process is limited by the dimensions of the fruit. From the market point of view, the dimensions and colour are two main parameters valued by the consumer. The effect of e-beam radiation on nutritional parameters of irradiated chestnut fruits are also under validation by our research group (Carocho et al., 2012). Previous studies on chestnut fruits preservation by irradiation were done in Asian varieties and only recently in European varieties, mainly using gamma radiation (Antonio et al., 2012). Here we have studied the effect of e-beam irradiation on the colour parameters of four European chestnut fruit varieties, from Portugal and Italy.

Materials and Methods

Samples

European chestnut fruits (*Castanea sativa* Miller) of varieties “*Longal*”, “*Judia*” and “*Cota*”, harvested in October 2012, were obtained from local producers in the Northeast of Portugal. The Italian chestnuts, from a selected variety “*Palummina*”, were obtained in a local market in Naples, Italy, in October 2012. The fruits were divided in two groups, control and irradiated, weighing approximately 0.5, kg to be exposed to the radiation dose of 1 kGy - considering 0 kGy the non-irradiated, control sample. After irradiation, the samples were stored at 4 °C and the colour parameters monitored immediately after irradiation and at fixed intervals of 30 days of storage,

Irradiation

The irradiation with electrons was performed at the INCT –Institute of Nuclear Chemistry and Technology – in Warsaw, Poland, with an e-beam of 10 MeV of energy. A pulse duration of 5.5 μs, pulse frequency of 440 Hz, average beam current of 1.1 mA, a scan width of 68 cm, conveyer speed in the range 20-100 cm/min and a scan frequency of 5 Hz. The absorbed dose was 1 kGy, with an uncertainty of 4%. To estimate the dose, routine dosimeters, Amber Perspex and Gammachrome YR dosimeters (from Harwell Company, U.K.) and a standard dosimeter, Graphite Calorimeter, were used (IAEA, 2002).

Physical Parameters - Dimensions

The axial dimensions of the fruits were determined measuring the length (L), width (W) and thickness (T) of eighteen chestnut fruits, using a digital calliper with a precision of 0.01 mm. The arithmetic and geometric diameters, as well as the sphericity, were calculated using the equations (1), (2) and (3) (Mohsenin, 1986).

$$\text{Arithmetic diameter: } Da = (L + W + T) / 3 \quad (1)$$

$$\text{Geometric diameter } Dg = (L \times W \times T)^{1/3} \quad (2)$$

$$\text{Sphericity } \Phi = (L \times W \times T)^{1/3} / (\max [L,W,T]) \quad (3)$$

Physical Parameters - Colour

For colour determination, the CIE L*, a*, b* scale was used, measuring the colour parameters with a colorimeter (CR400, from Konica Minolta, Japan), calibrated with a white tile, using C illuminant and a diaphragm aperture of 8 mm. The measurements were made on the skin, fruit (after hand peeling) and fruits interior (after cutting the fruits in two halves). The colour of nine chestnuts was measured in three different points, for each dose (0 and 1 kGy), n = 27, and at each time point (0, 30 and 60 days of storage), n = 54.

Results and discussion

Physical characteristics

The axial dimensions and mass of irradiated chestnut fruits are presented in Table 1.

Table 1 - Physical characteristics of chestnut fruits, *Castanea sativa* Mill.:
mass (m), width (W), length (L), thickness (T).

	m (g)	W (mm)	L (mm)	T (mm)
<i>Longal</i>	12.17 ± 1.26	31.36 ± 1.41	36.77 ± 1.29	18.41 ± 2.07
<i>Judia</i>	16.92 ± 2.79	36.30 ± 2.41	35.82 ± 2.58	22.79 ± 3.04
<i>Cota</i>	14.65 ± 2.52	34.43 ± 2.57	35.91 ± 1.66	20.61 ± 3.18
<i>Palummina</i>	18.94 ± 3.55	36.84 ± 3.16	35.88 ± 1.94	25.95 ± 3.99

The results are expressed as mean ± std. dev (n = 18).

The four chestnut fruit varieties present similar lengths and similar widths, except for the *Longal* variety with a lower width. The lower thickness and mass is presented by the *Longal* variety and the biggest thickness and mass by the *Palummina* variety (Fig. 1).

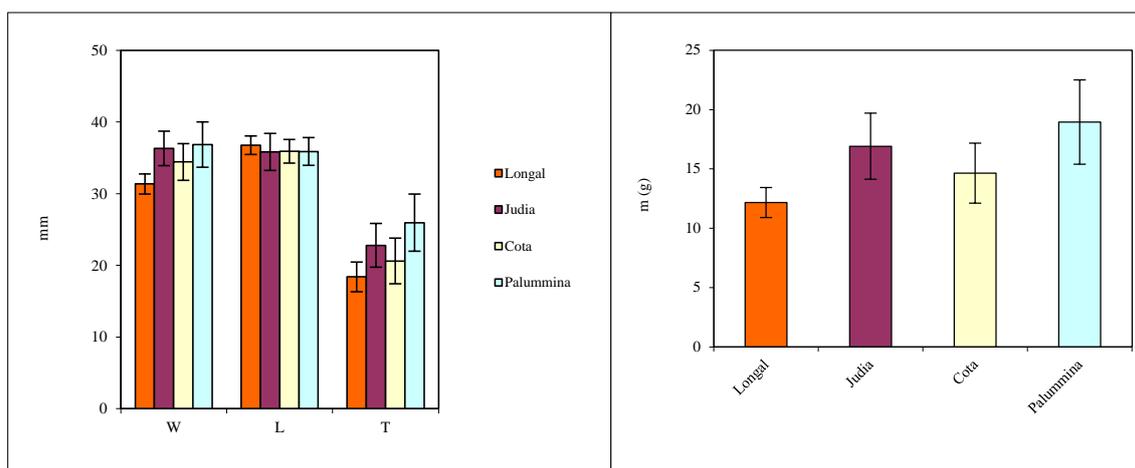


Fig.1. Dimensions and mass of four chestnut fruit varieties of *Castanea sativa* Mill.

Regarding arithmetic diameter, geometric diameter and sphericity, the lower values are presented by the *Longal* variety and the higher values by the *Palummina* variety, showing that the first were almost planar and the last almost round (Table 2).

Table 2 - Physical characteristics of chestnut fruits, *Castanea sativa* Mill.: arithmetic diameter (D_a), geometric diameter (D_g) and sphericity (Φ).

	D_a (mm)	D_g (mm)	Φ
<i>Longal</i>	28.85 ± 0.95	27.64 ± 1.16	0.75 ± 0.03
<i>Judia</i>	31.64 ± 1.81	30.87 ± 1.96	0.83 ± 0.04
<i>Cota</i>	30.32 ± 1.70	29.34 ± 1.98	0.80 ± 0.04
<i>Palummina</i>	32.89 ± 1.76	32.37 ± 2.00	0.88 ± 0.07

The results are expressed as mean \pm std. dev (n = 18).

Colour

The colour parameters L^* , a^* and b^* were measured for the non-irradiated (0 kGy) and irradiated (1 kGy) at 0, 30 and 60 days of storage at 4 °C. Here we present only the results for L^* and b^* , since for the skins a^* parameter follows the same tendency. And for the fruit and interior this value is close to zero (data not shown).

Chestnut skins present lower values for L^* (“lightness”) and b^* (“redness-yellowness”) than fruits and interior (Fig. 2).

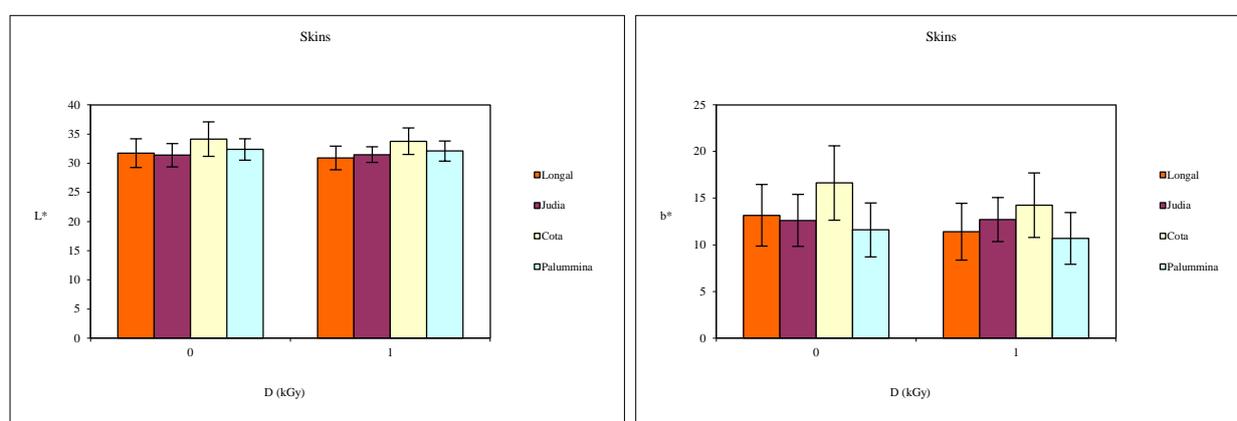


Fig. 2. Skins colour (L^* - , b^* - value) of four varieties of *Castanea sativa* Mill.

The higher value for L^* was registered for the interior (half-cut) and the higher value for b^* (“yellowness”) was displayed by the peeled fruits (Fig. 3 and Fig. 4).

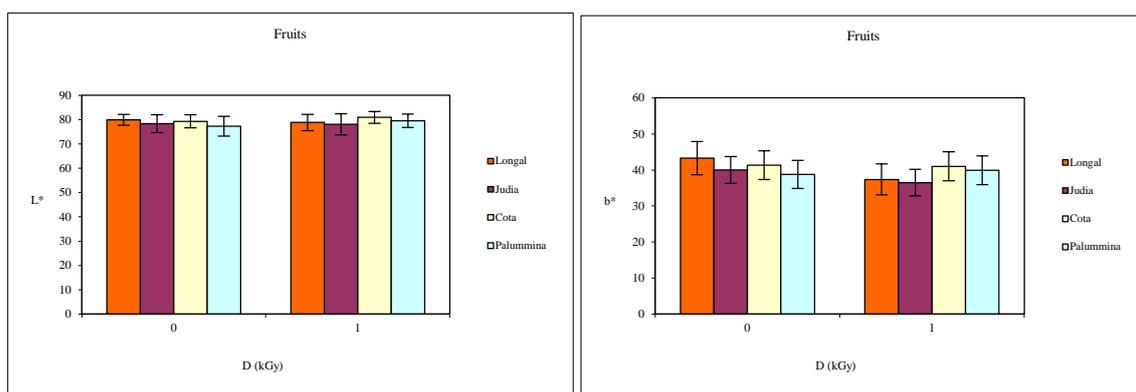


Fig. 3. Fruits colour (L^* - , b^* - value) of four varieties of *Castanea sativa* Mill.

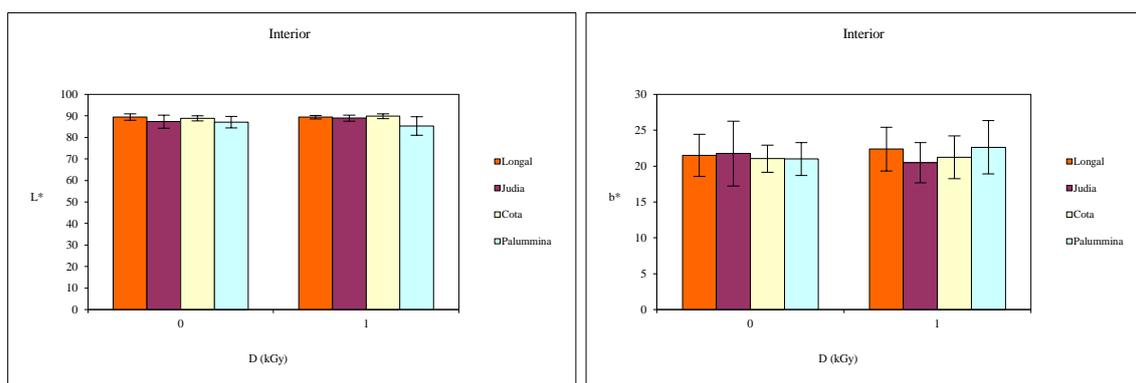


Fig. 4. Fruits interior colour (L^* - , b^* - value) of four varieties of *Castanea sativa* Mill.

From the results, we could also see that chestnut varieties present similar values for L^* and b^* , for skins, fruit and interior (half-cut), for non-irradiated and irradiated fruits (Fig. 2 - 4).

According to these results (Fig.2-4), L^* and b^* parameters showed similar values. We have followed these parameters along storage time (Fig. 5), concluding that non-irradiated and irradiated samples present similar results at each time point. With storage time, b^* -value (“yellowness”) showed an increasing tendency for irradiated and non-irradiated samples, represented in the graphs by the dashed line

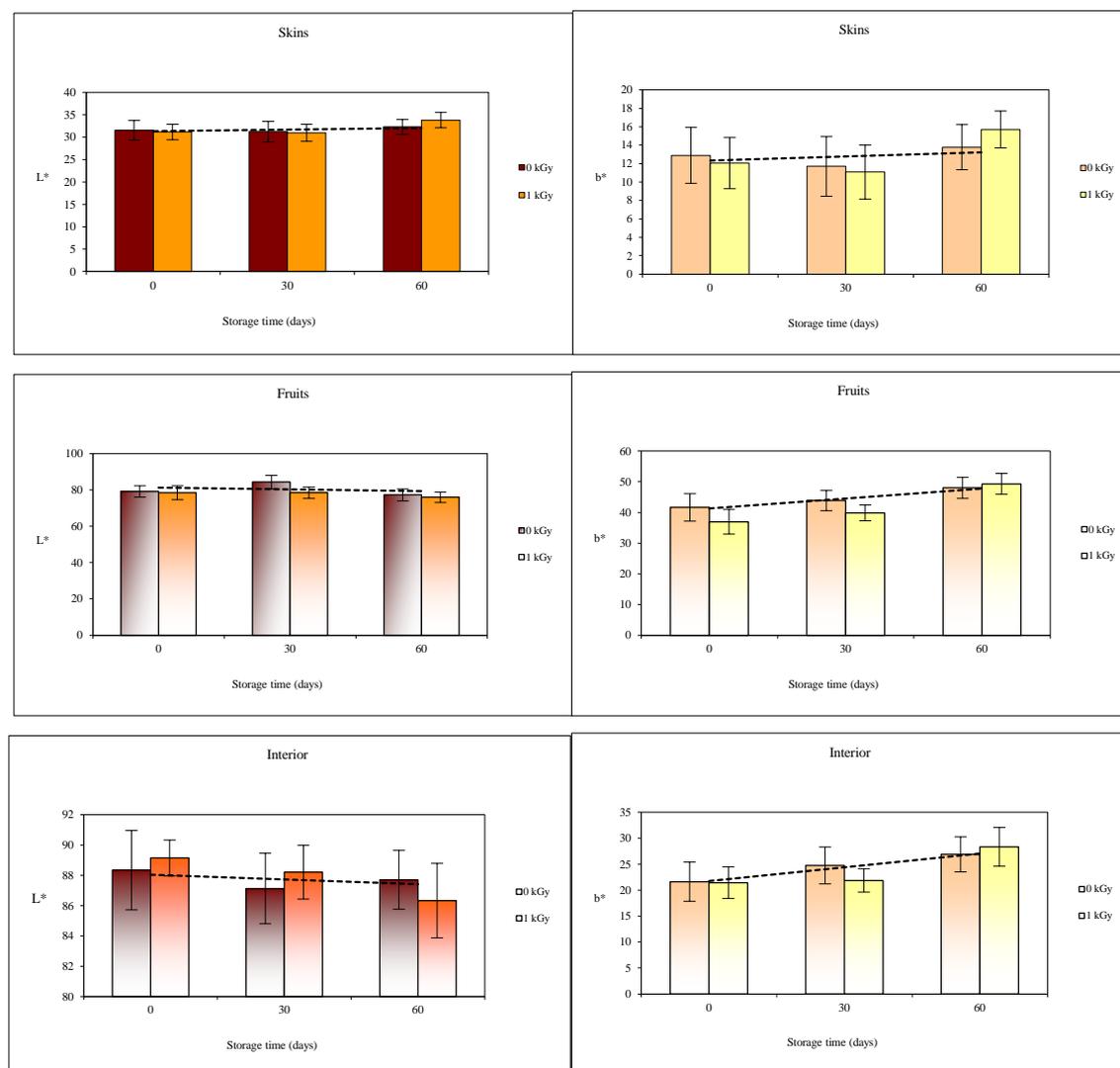


Fig. 5. Chestnuts colour of *Castanea sativa* Mill. along storage time.

Conclusion

There's a tendency in the agro-industrial units that processes these fruits to separate the chestnuts by varieties, due to their different size value in the market. In an irradiation preservation treatment, the typical physical dimensions of the different varieties could be used to optimize the irradiation process, the throughput and e-beam penetration. Up to the applied dose of 1 kilogray and 60 days of storage, e-beam irradiation did not induce any significant change in the skins, fruits and interior colour on the four chestnut fruit varieties. Since the typical dose for insect disinfestations is lower than 1 kGy, we can conclude that e-beam postharvest irradiation could be a promising alternative to the banned MeBr fumigation.

Acknowledgements

The authors thank ON.2/QREN/EU Project no.13198/2010 for financial support. A.L. Antonio thanks the Foundation for Science and Technology (FCT, Portugal) grant SFRH/PROTEC/67398/2010. They also thank Prof. A. Chmielewski, Director of the Institute of Nuclear Chemistry and Technology, for allowing e-beam irradiations.

References:

Antonio, A. L. et al, 2012. "Effects of gamma radiation on the biological, physico-chemical, nutritional and antioxidant parameters of chestnuts - A review". *Food and Chemical Toxicology* 50, 3234-3242.

Carocho, M., Barreira, J.C.M., Antonio, A.L., Bento, A., Kaluska, I., Ferreira, I.C.F.R., 2012. "Effects of Electron-Beam Radiation on Nutritional Parameters of Portuguese Chestnuts (*Castanea sativa* Mill.)". *J. Agricultural Food Chemistry* 60, 7754–7760.

EU, 1999. Directive 1999/2/EC: On the approximation of the laws of the Member States concerning foods and food ingredients treated with ionising radiation. Official Journal of the European Communities L 66/16, 13th March.

EU, 2008. "Commission Decision, 753/2008". Official Journal of the European Union L 258/68, 26th September.

FAOSTAT, 2010. Food and Agriculture Organization of the United Nations. Available from: <http://faostat.fao.org/site/339/default.aspx> . Accessed March 2013.

FAO/IAEA/WHO, 1981. Wholesomeness of irradiated food. Report of a Joint FAO/IAEA/WHO Expert Committee. Technical Report Series, 659, World Health Organization, Geneva, Switzerland.

IAEA, 2002. Dosimetry for food irradiation. Technical report series 409, International Atomic Energy Agency, Vienna, Austria, ISSN 0074–1914.

IDIDAS, 2013. International Database on Insect Disinfestation and Sterilization Available from: <http://nucleus.iaea.org/ididas/> . Accessed March 2013.

INE, 2010. Portuguese Agricultural Statistics. Instituto Nacional de Estatística I.P., Lisbon, Portugal. Available from: <http://www.ine.pt> . Accessed March 2013.

Mohsenin N. N. (1986). Physical properties of plant and animal materials. Gordon and Breach Sci. Publishers, New York, USA.

SIMULTANEOUS SEPARATION/PRECONCENTRATION OF NICKEL AND COBALT BY DISPERSIVE LIQUID-LIQUID MICROEXTRACTION PRIOR TO DETERMINATION BY FAAS

F. Sanchez Rojas, Professor

C. Bosch Ojeda, Professor

J.M. Cano Pavon, Professor

Department of Analytical Chemistry, University of Malaga, Campus Teatinos s/n, Malaga, Spain

Abstract:

In this work, an efficient microextraction method was applied to separation and preconcentration of Ni(II) and Co(II), namely dispersive liquid-liquid microextraction (DLLME). The influences of analytical parameters, including pH, extraction solvent volume, disperser solvent type and its volume, concentration of chelating agent, on the quantitative recoveries of nickel and cobalt were investigated. Under the optimized conditions, the detection limits were $1 \mu\text{g L}^{-1}$ and $3.5 \mu\text{g L}^{-1}$ for Ni and Co, respectively. The proposed method was applied to the determination of these elements in food samples with satisfactory results. The accuracy of the method was studied by analysing certified reference materials.

Key Words: Cobalt, nickel, DLLME, certified reference materials, foods, FAAS

Introduction

The impact of trace elements on environment and man's health has fostered the development of analytical techniques and instrumentation capable of measuring concentrations lower than the parts per million levels or less.

Cobalt is a natural earth element present in trace amount in soil, plants and our diets, which is an essential mineral, although the body only needs a small amount. Cobalt is an element that occurs naturally in many different chemical forms throughout the environment that has properties similar to those of iron and nickel. Small amounts of cobalt are essential for good health. Cobalt is a relatively rare element of the earth's crust, composing approximately 0.001% of the earth's crust, which is essential to mammals in the form of cobalamin (vitamin B₁₂).

Natural sources of cobalt in the environment are soil, dust, seawater, volcanic eruptions and forest fires. People are commonly exposed to small amounts of cobalt naturally present in the air they breathe, the water they drink, and the food they eat. Very small amounts of cobalt in people's diets are necessary for good health.

Cobalt is not often freely available in the environment, but when cobalt particles are not bound to soil or sediment particles the uptake by plants and animals is higher and accumulation in plants and animals may occur. The natural background air concentration of cobalt is in the order of 1 pg/m^3 , but concentrations up to 40 ng/m^3 have been reported in urban areas. In occupational settings, workers are exposed to cobalt compounds by inhalation of dusts. The main industrial use of cobalt is for the manufacture of alloys and hard metals. The concentration of cobalt in soil varies widely, generally ranging from about 1 to 40 ppm, with an average level of 7 ppm. Soils containing less than about 3 ppm of cobalt are considered cobalt-deficient because plants growing in them do not have sufficient cobalt to meet the dietary requirements of cattle and sheep. On the other hand, soils near cobalt-containing mineral deposits, mining and smelting facilities, or industries manufacturing or using cobalt alloys or chemicals may contain much higher levels of cobalt. The concentration of cobalt in surface and groundwater is generally low, between 1 and 10 ppb in populated areas; concentration may be hundreds or thousands times higher in areas that are rich in cobalt-containing minerals or in areas near mining or smelting operations.

As an essential biochemical element, cobalt is mainly stored in red blood cells with smaller amounts in kidney, liver pancreas and spleen. Research indicates that cobalt helps with the repair of the myelin sheath, increases the effectiveness of glucose transport from the blood into body cells, and increases the assimilation of iron and the building of red blood cells. Cobalt is also an important agent of Vitamin B₁₂; it increases the body's ability to absorb it. Because of its low absorption rate and high excretion rate, cobalt toxicity is not common, but an excess can lead to enlargement of the thyroid gland.

Cobalt is toxic in large amounts and chronic ingestion of Co in the daily diet can cause toxic effects. Toxicological effects of cobalt include vasodilatation, flushing and cardiomyopathy in humans and animals. The respiratory system is the main target organ on inhalation exposure to cobalt, with a higher risk of fibrosing alveolitis and lung cancer in the hard metal industry, where workers are exposed to cobalt metal mixed with tungsten carbide particles. Other target organs include the hematopoietic system, the myocardium, the thyroid gland, and possibly the reproductive system.

The adult human body contains approximately 1 mg of cobalt, 85% of which is in the form of vitamin B₁₂. Human dietary intake of cobalt varies between 5 and 50 µg/day, and most of the cobalt ingested by humans is inorganic, vitamin B₁₂ representing only a small fraction (Nordberg G. F., Fowler B. A., Nordberg M. and Friberg L. T., 2007).

On the other hand, nickel is a very abundant natural element. Nickel makes up about 0.01 to 0.02 % of the Earth's crust. It ranks about 22nd among the chemical elements in terms of abundance in the Earth's crust. Similarly to many other trace elements, nickel is ubiquitous in the environment being present in natural waters and practically all soils and foods. Food is the major source of nickel exposure, with an average intake for adults estimated to be approximately 100-300 µg per day. Individuals also may be exposed to nickel in occupations involved in its production, processing, and use, or through contact with everyday items such as nickel-containing jewellery and stainless steel cooking and eating utensils, and by smoking tobacco. Nickel is found in ambient air at very low levels as a result of releases from oil and coal combustion, nickel metal refining, sewage sludge incineration, manufacturing facilities, and other sources. (Cempel M and Nickel G., 2006)

Nickel can pose a health hazard to certain individuals. The most common health problem is called nickel allergy. Some people are more likely to develop nickel allergy than are others. People who are sensitive to nickel may develop a skin rash somewhat like poison ivy. The rash becomes itchy and may form watery blisters. Nickel is present in dozens of products. So it is easy for sensitive people to develop nickel allergy. Nickel can cause more serious health problems too. For example, people who are exposed to nickel fumes (dust and gas) breathe in nickel on a regular basis. Long term nickel exposure may cause serious health problems, including cancer.

The International Agency for Cancer Research, concluded in 1990 that nickel compounds were human carcinogens, and most animal data *in vivo* and genetic toxicology data *in vitro* before that time suggested that the insoluble particulate nickel species were the most carcinogenic; however, more recent human epidemiology and experimental data are pointing to the water-soluble nickel compounds as perhaps of equal hazard.

Several novel microextraction techniques are being developed in order to improve the quality and the sensitivity of the analytical methods. The cloud point extraction (CPE) (Ojeda C. B, Rojas F.S., 2009, 2012), the homogeneous liquid-liquid extraction (HLLE) (Ebrahimzadeh, H., Yamini, Y., Kamare, F., Shariati, S., 2007; Ghiasvand, A. R., Shadabi, S., Mohagheghzadeh, E., Hashemi, P., 2005); the liquid phase microextraction (LPME) (Ahmadi, F., Assadi, Y., Hosseini, M. R.M., Rezaee, M., 2006; Zanjani, M. R. K., Yamini, Y., Shariati, S., Jonsson, J. A., 2007), and the solid phase microextraction (SPME) (Djozan, D., Assadi, Y., Haddadi, S.H., 2001; Djozan, D., Assadi, Y., 2004) are fairly new methods of sample preparation. They are employed in the separation and preconcentration of environmental contaminants in different matrices and can solve some of the problems, encountered with the conventional pretreatment techniques. Recently, a new liquid-liquid microextraction method namely, dispersive liquid-liquid microextraction (DLLME) was proposed (Rezaee, M., Assadi, Y., Hosseini, M. R. M., Aghae, E., Ahmadi, F., Berijani, S., 2006). This method has been applied for the determination of trace organic pollutants and metal ions in the environmental samples (Ojeda C.B., Rojas F.S., 2009, 2011). DLLME is based on a ternary component solvent system like homogeneous liquid-liquid extraction and cloud point extraction. In this method, the appropriate mixture of extraction solvent and dispersive solvent is injected into an

aqueous sample rapidly by a syringe, and a cloudy solution is formed. The analyte in the sample is extracted into the fine droplets of extraction solvent. After extraction, phase separation is performed by centrifugation.

This work presents the development of a procedure for the separation/preconcentration of cobalt and nickel from diverse samples and its determination by FAAS. This procedure is based on DLLME after complexing these metal ions with 1,5-bis(di-2-pyridylmethylene) thiocarbonylhydrazide (DPTH).

Experimental

Standard solutions and reagents

Stock standard solution for Ni(II) and Co(II) (1000 mg L⁻¹) were supplied by Merck, Darmstadt, Germany. Standard solutions were prepared by appropriate dilution of the stock solutions daily.

High purity water (resistivity 18.2 MΩ cm) obtained by a Milli-Q water purification system (Millipore, Bedford, MA, USA) was used throughout this work.

2,2'-Bis(di-2-pyridinyl-methylene)-thiocarbonylhydrazide (DPTH) solution in DMF was prepared by dissolving solid reagent samples prepared and purified by the authors (Bonilla Abascal J., García de Torres A., Cano Pavón J.M. 1983).

All the other reagents including extractants and disperser solvents were analytical-grade reagents, as well as the reagents mentioned above.

Acetate buffer solution was prepared to adjust pH values for the extraction of Ni and Co.

NaCl solutions were prepared by dissolving appropriate amounts of NaCl in deionized water.

Instrumentation

Phase separation was achieved with a centrifuge Selecta Centromix in 15 mL calibrated conical tubes. A Varian Model SpectrAA 50 (Mulgrave, Victoria, Australia) flame atomic absorption spectrometer was used for the analysis with the appropriate nickel and cobalt hollow cathode lamps. The operating parameters were set as recommended by the manufacturer. Atomic absorption measurements were carried out in an air-acetylene flame. The following conditions were used: absorption line Ni: 232.0 nm; slit widths: 0.2 nm; and lamp currents: 4 mA. Absorption line Co: 240.7 nm; slit widths: 0.2 nm; and lamp currents: 7 mA.

Sample analysis

The accuracy of the method for determination of nickel and cobalt content was checked by analyzing the reference standard material TMDA 54.4, Estuarine water (CRM LGC6016), Eau de mer (CASS-5), Riverine water (SLRS-5). These samples were analyzed by standard addition method and in the case the content of these analytes were under detection limits of the method, recovery studies were carry out.

The proposed method was also evaluated by analysis of Ni and Co in several spiked food samples. The Co and Ni concentrations in all the original samples were below the detection limit. For this purpose, standard solutions containing Co and Ni were added to 0.3–1.2 g of diverse foods, and the resulting materials were mineralized by reflux digestion, then evaporate to eliminate excess of acid, adjusted pH and diluted at convenient volume. Standard addition method was used in all instances.

DLLME procedure

For DLLME under optimum conditions, 10 mL analyte solution containing variable amounts of nickel and cobalt, 2 mL acetate buffer solution, pH 5.4, 1 mL of 0.05% DPTH solution in DMF as chelating agent was placed in a 15 mL screw cap glass test tube. Then, 1 mL of methanol (as disperser solvent) and 0.3 mL of chloroform (as extraction solvent) were rapidly injected into a sample solution by using a microsyringe. A cloudy solution was formed in the test tube and separation of the phases was achieved by centrifugation at 3800 rpm for 5 min. After this process, the organic phase was sedimented in the bottom of conical test tube. After removal of the whole aqueous solution, the extraction phase was aspirated into the FAAS.

Results and discussion

In order to obtain a high preconcentration factor, the effect of different parameters affecting the complex formation and extraction conditions, such as kind of extraction and disperser solvent and their volume, pH, concentration of the chelating agent, and salt addition, was optimized. One variable at a time optimization was used to obtain the optimum conditions for the DLLME.

pH study

The separation of metal ions by DLLME involves prior formation of a complex with sufficient hydrophobicity to be extracted into the small volume of the sedimented phase, thus, obtaining the desired preconcentration. pH plays a unique role on metal chelate formation and subsequent extraction. The effect of pH on the complex formation and extraction of nickel and cobalt from samples was studied in the range of 3.0–8.0 by using acetate or phosphate buffer. The results reveal that the absorbance is nearly constant in the pH range of 4.5–6 for Ni and 5.0–5.5 for Co. Thus, the value of pH 5.4 was selected for the following experiments. Also, the influence of 0.2 M acetate buffer solution amount was investigated for variation of volume added from 1 to 4 mL. The extraction efficiency diminished between 3–4 mL. A volume of 2 mL was selected as the optimum value for subsequent work.

Effect of chelating reagent (DPTH) concentration

The effect of DPTH concentration on the absorbance was examined using increasing volumes of 0.05% DPTH from 0.5 to 2 mL. The results showed that the change of DPTH concentration in the studied range has effect on analytical signals (Figure 1), thus the volume of 1 mL, corresponding to its maximum value, was used in other experiments.

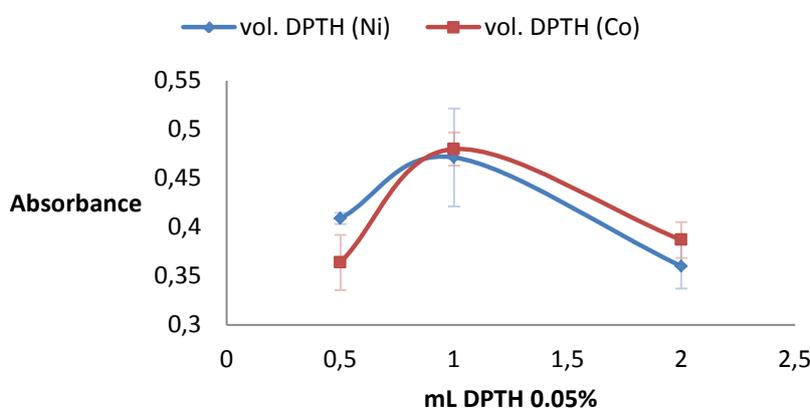


Figure 1. Influence of reagent volume

Effect of ionic strength

For investigating the influence of ionic strength on performance of DLLME, various experiments were performed by adding different amounts of NaCl (0–3% (w/v)). Other experimental conditions were kept constant. Ionic strength had a negative effect upon percent recovery and sensitivity for Ni extraction.

Effect of DLLME parameters

Effect of type and volume of extractant

Careful attention should be paid to the selection of the extraction solvent. It should have a higher density than water, extraction capability of the interested compounds and low solubility in water. Chloroform, carbon tetrachloride and dichloromethane were compared in the extraction of nickel and cobalt. Results showed that the maximum extraction recovery was obtained by using chloroform.

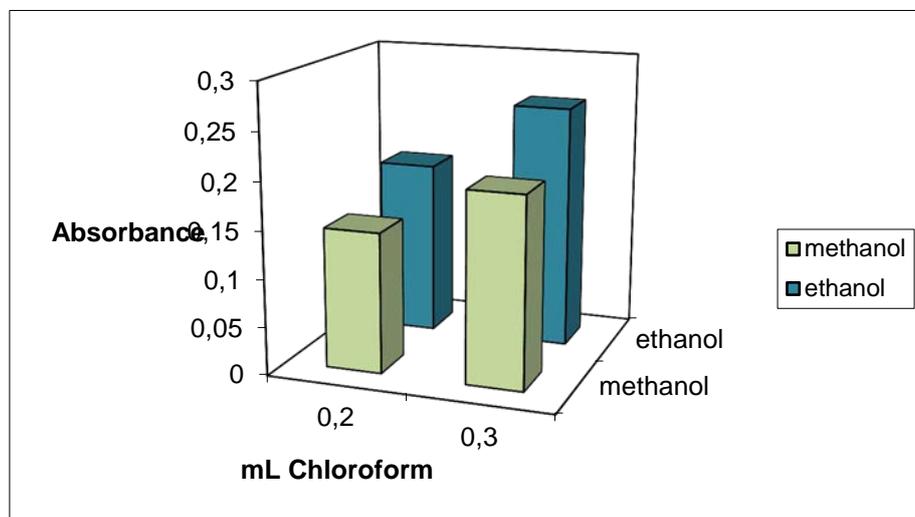
To examine the effect of extraction solvent volume, solutions containing different volumes (0.4–0.7 mL) of chloroform were subjected to the same DLLME procedure. When the volume of extraction solvent was increased, the volume available for the measurement also increased, but the enrichment factors decreased. Thereby, in the following studies, the optimum volume of 0.3 mL was selected for the extraction solvent.

Effect of type and volume of disperser solvent

The role of a disperser is dispersion of an extraction solvent into aqueous sample to make extensive contact area between them and facilitating the mass transfer of analyte from water to organic solvent which causes considerable acceleration in the extraction of analytes. Miscibility of disperser solvent with extraction solvent and aqueous phase is the main point for selection of disperser solvent. Therefore in this section the ability of ethanol and methanol was investigated. The results

(figure 2) show light differences between disperser solvents containing 0.2 and 0.3 mL chloroform (extraction solvent). As can be seen from this figure better results were obtained by using 0.3 mL chloroform and ethanol as disperser solvent. Then, extraction of analytes was carried out by using 0.5 to 1.5 mL of ethanol and in figure 3 are shown the results obtained. A 1 mL ethanol was selected for subsequent studies.

a)



b)

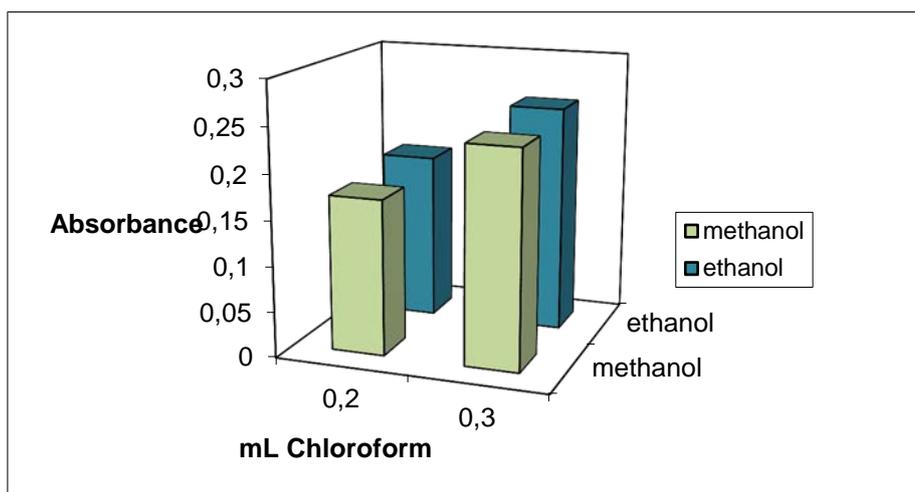


Figure 2. Influence of disperser solvent: a) Ni signal b) Co signal

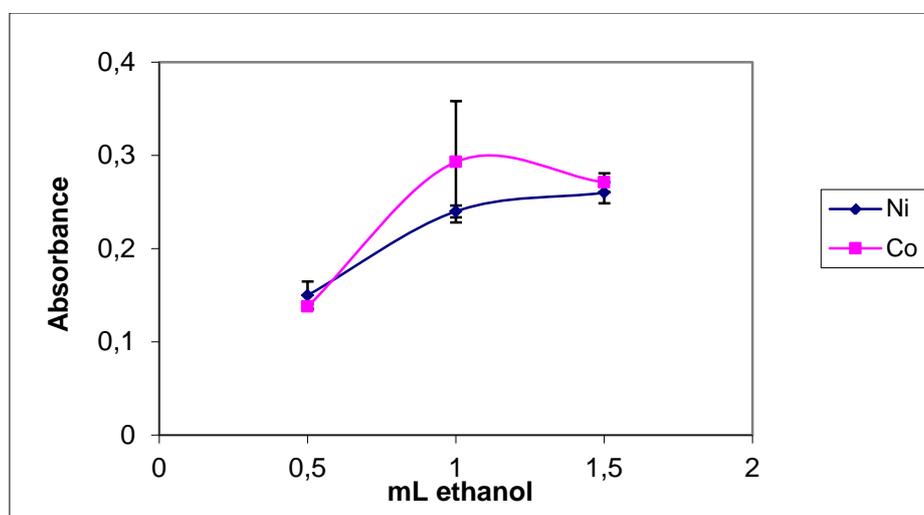


Figure 3. Influence of disperser volume

Interferences

Because DPTH is a versatile chelating agent, interferences may occur due to the competition of other heavy metal ions and their subsequent co-extraction with Co(II)/Ni(II). For this purpose, the effect of typical potential interfering ions was investigated. The tolerance limit was defined as the concentration of added ion that caused less than $\pm 5\%$ relative error in the determination of Ni and Co. About 500-fold excess of Fe^{3+} , Ca^{2+} , Ba^{2+} , Mn^{2+} , K^+ , I^- , F^- , SO_4^{2-} and HCO_3^- do not affect Ni signal and Co. Cd^{2+} , Hg^{2+} , Cr^{3+} , Ba^{2+} , Mn^{2+} and Al^{3+} do not interfere at 100-fold excess. Cu^{2+} , Bi^{3+} , Zn^{2+} can be tolerated at 50-fold excess.

Analytical figures of merit

Analytical figures of merit of the proposed DLLME-FAAS method were obtained under optimal conditions and summarized in table 1. For this purpose, calibration equation were obtained by this procedure for Ni with different amount of Co (between 10-500 $\mu\text{g L}^{-1}$); and the same manner calibration equation for Co were obtained with different amount of Ni (between 10-500 $\mu\text{g L}^{-1}$). In all instances the slopes obtained were similar.

Analysis of synthetic mixtures

In order to check for potential synergistic effects of the mixtures of two ions, various synthetic mixtures were readily resolved by using pertinent calibration graphs. The analytical results are listed in table 2. As it can be seen, mixtures of the two ions can be resolved with satisfactory results.

Table 1. Analytical figures of merit

	Ni	Co
Dynamic range ($\mu\text{g L}^{-1}$)	10-500	10-500
Regression equation	$A=0.0005[\text{Ni}]+0.0105$	$A=0.0004[\text{Co}]+0.016$
R^2	0.9987	0.9995
Detection limit ($\mu\text{g L}^{-1}$)	1	3.5
Determination limit ($\mu\text{g L}^{-1}$)	15	13.5
Precision (% RSD) Ratio 1:1 (Ni:Co) n=6	4.62	5.38
Precision (% RSD) Ratio 1:4 (Ni:Co) n=6	5.79	5.86
Precision (% RSD) Ratio 4:1 (Ni:Co) n=6	6.11	6.15
Preconcentration factor	50	40

Analysis of standard reference materials

In order to assess the accuracy and validity of the presented procedure, the method was applied to the determination of nickel and cobalt in certified reference materials (TMDA 54.4, Estuarine water (CRM LGC6016), Eau de mer (CASS-5), Riverine water (SLRS-5), which were analyzed according to the proposed method. It was found that analytical results were in good agreement with the certified values (table 3). Contents of Ni and Co in CASS-5 and SLRS-5 were under detection limits of the method, so different amounts of these ions were added to the samples for to verify if it is possible to determine Ni and Co under the proposed procedure. Goods recoveries were obtained in all cases.

Analysis of foods

In view of the application of the method to the determination of nickel and cobalt in food samples, the ability to recover these elements from different samples spiked with nickel and cobalt were investigated. All samples were arbitrarily selected and acquired from a local superstore. For this purpose, standard solutions containing different quantities of these elements were added to samples and the resulting materials were prepared as described under Experimental. Standard additions method was used in all instances and the results were obtained by extrapolation. The results of these analyses are summarised in table 4. The recoveries for the spiked samples were in the acceptable range.

Table 2. Resolution of synthetic Ni(II)-Co(II) mixtures by use of the proposed method

Ratio	Concentration added ($\mu\text{g L}^{-1}$)		Concentration found ($\mu\text{g L}^{-1}$)		Concentration added ($\mu\text{g L}^{-1}$)		Concentration found ($\mu\text{g L}^{-1}$)	
	Ni	Co	Ni	Co	Co	Ni	Co	Ni
1:1	50	50	56.8	49.25	50	50	45.5	50.2
1:2	50	100	56.8	94.25	50	100	43.0	91.2
1:4	50	200	54.8	196.75	50	200	48.0	202.2

1:10	50	500	50.8	474.25	50	500	53.0	510.2
10:1	100	50	82.8	49.5	100	50	93.0	51.0
1:1	100	100	86.8	95.0	100	100	95.5	92.0
1:2	100	200	90.8	192.5	100	200	95.5	210.0
1:5	100	500	106.8	487.5	100	500	108.0	499.0
4:1	200	50	204.8	43.75	200	50	180.5	59.6
2:1	200	100	194.8	91.25	200	100	185.5	107.6
1:1	200	200	234.8	206.25	200	200	193.0	179.6
1:2,5	200	500	234.8	506.25	200	500	210.5	511.6
10:1	500	50	476.8	48.5	500	50	395.5	51.2
5:1	500	100	478.8	98.5	500	100	433.0	115.2
2,5:1	500	200	474.8	197.25	500	200	458.0	169.2
1:1	500	500	478.8	498.5	500	500	460.5	496.2

Table 3. Analysis of four certified reference materials for the determination of Ni and Co with DLLME-FAAS method

Sample	Certified value ($\mu\text{g L}^{-1}$) ¹⁾		Found value ($\mu\text{g L}^{-1}$) ^a		% Recovery	
	Ni	Co	Ni	Co	Ni	Co
TMDA 54.4	337±15,5	309±13.5	336.2±0.36	308.3±6.29	99.76	99.77
Estuarine water. CRM LGC6016	186 ± 3	26.08*	184.9±8.9	25.84 ± 0.42	99.4	99.08
	Added*		Found ($\mu\text{g L}^{-1}$) ^a		% Recovery	
	Ni	Co	Ni	Co	Ni	Co
Eau de mer. CASS-5	20	20	18.93±0.50	22.27±1.17	94.65	111.35
	50	50	46.93±3.01	50.60±0.92	93.86	101.2
	100	100	99.27±1.63	92.60±0.92	99.27	92.6
	200	200	192.27±2.80	200.27±0.42	96.13	100.13
Riverine water. SLRS-5	20	20	21.28±1.36	19.53±1.2	106.40	97.65
	50	50	51.48±1.84	44.53±2.34	103.68	89.06
	100	100	102.76±2.94	100.87±2.81	102.00	100.87
	200	200	197.6±13.72	190.87±0.70	98.80	95.43

^a mean±SD; n=3*Added in $\mu\text{g L}^{-1}$ **Table 4.** Analytical results for Ni and Co in food samples (Avg. ± SD of three trials)

Sample	Added ($\mu\text{g g}^{-1}$)		Found ($\mu\text{g g}^{-1}$)		% Recovery	
	Ni	Co	Ni	Co	Ni	Co
Rice	9.09	9.09	9.09±0.04	10.03±1.27	100	110.34
Chick pea	8.68	8.68	8.78±0.71	8.58±0.25	102.40	98.85
Lettuce	15.81	15.81	16.19±0.45	15.81±1.40	102.40	100.00
Lentil	11.10	11.10	11.62±0.06	11.26±0.14	104.68	101.44
Apple	8.45	8.45	8.25±0.10	8.70±0.68	97.63	102.96

Conclusion

Sample preparation by DLLME is a procedure that considered inside the Green Chemistry, because of the small volumes of dissolvent employed. In this work, a simple, rapid, and sensitive DLLME preconcentration technique coupled with FAAS has been developed for the determination of cobalt and nickel in spiked food samples. All variables that influence in the formation of the complexes Ni-DPTH and Co-DPTH and then application of DLLME procedure have been optimized. Employing FAAS as detection technique, the detection limit obtained is in the order of $\mu\text{g L}^{-1}$ for both analytes that by the direct method of FAAS is impossible to obtain because of the low sensitivity that present.

To study the accuracy of the proposed method, certified reference materials have been analysed with good agreements.

Acknowledgements

The authors thank to the Ministerio de Ciencia e Innovación for supporting this study (Projects CTQ2009-07858) and also the Junta de Andalucía.

References:

- Abascal J. B., Torres A. G., and Cano Pavón J. M., *Microchem. J.*, 1983, 28, 132–136.
- Ahmadi F., Assadi Y., Hosseini M. R.M.; Rezaee, M., *J. Chromatogr. A* 2006, 1101, 307.
- Cempel M. and Nikel G., *Pol. J. Environ. Stud.* 2006, 15, 375.
- Djozan D., Assadi Y., *Chromatographia* 2004, 60, 313.
- Djozan D., Assadi Y., Haddadi, S. H., *Anal. Chem.* 2001, 73, 4054.
- Ebrahimzadeh H., Yamini Y., Kamare F., Shariati S., *Anal. Chim. Acta* 2007, 549, 93.
- Ghiasvand A. R., Shadabi S., Mohagheghzadeh E., Hashemi P., *Talanta* 2005, 66, 912.
- Nordberg G. F., Fowler B. A., Nordberg M. and Friberg L. T., Eds., “Handbook on the Toxicology of Metals,” Elsevier Inc., Amsterdam, 2007.
- Ojeda C.B., Rojas F.S., *Anal. Bioanal. Chem.*, 2009, 394, 759.
- Ojeda C.B., Rojas F.S., *Chromatographia* 2009, 69, 1149.
- Ojeda C.B., Rojas F.S., *Chromatographia* 2011, 74, 651.
- Ojeda C.B., Rojas F.S., *Microchim. Acta* 2012, 177, 1.
- Rezaee M., Assadi Y., Hosseini M. R. M., Aghaee E., Ahmadi F., Berijani S., *J. Chromatogr., A* 2006, 1116, 1.
- Zanjani M. R. K., Yamini Y., Shariati S., Jonsson J. A., *Anal. Chim. Acta* 2007, 585, 286.

POLYCYCLIC AROMATIC HYDROCARBONS (PAHs) AND POLYCHLORINATED BIPHENYLS (PCBs) IN SOILS OF AGBABU, NIGERIA

Olubunmi E. Fagbote, M. Sc.

Edward O. Olanipekun, Professor

Department of Chemistry, Ekiti State University, Ado-Ekiti, Nigeria

Abstract:

Polycyclic aromatic hydrocarbons (PAHs) and polychlorinated biphenyls (PCBs) were investigated in soils of Agbabu, Nigeria. PAHs in the samples were quantified using gas chromatograph - mass spectrometer, while gas chromatograph - flame ionization detector was used for PCBs analysis. The maximum sum of concentrations of PAHs (Σ PAH) was 0.07 ± 0.036 ppm in the dry season, and 0.69 ± 0.19 ppm in the rainy season. Apart from three stations in both seasons where PAHs were not detected, other Σ PAH concentrations obtained were higher than $1-10 \mu\text{g}/\text{kg}$, indicating anthropogenic sources. Most of the PAH concentrations obtained were within the expected limits for a rural environment. However, in the rainy season, 0.20 ± 0.04 ppm for Benzo(b)fluoranthene, 0.21 ± 0.05 ppm for Benzo (g, h, i) perylene and 0.32 ± 0.04 ppm for Indenol (1, 2, 3 - cd) pyrene exceeded the expected limit. In the dry season, there was no significant relationship ($r^2 = 0.002$) between mean concentrations and percentage soil organic carbon while there was significant correlation ($r^2 = 0.57$) in this relationship in the rainy season. The predominance of 4 - 6 ring PAHs and the ratios between pairs of individual PAHs showed that the PAHs were from pyrolytic sources. No PCB congener was detected in the samples investigated.

Key Words: Agbabu Nigeria, Polycyclic aromatic hydrocarbons, Polychlorinated biphenyls, soils

Introduction

Bitumen was discovered at Agbabu, Nigeria over a hundred years ago, preceding the discovery of oil by over 50 years. The probable reserve of bitumen and heavy oil in the entire Nigerian belt is about $120 \times 4.3\text{km}$ (Adegoke and Ibe, 1982; Oboh et al., 2006).

Agbabu is a village of about 8, 611 inhabitants in Ondo State, south-west, Nigeria; in the coordinates of $E004^{\circ}48-49^{\circ}$ and $N06^{\circ}34-36^{\circ}$ (Figure 1). Farmers at Agbabu area deal mainly in cash crops such as cocoa and kolanut; and food crops such as yam and plantain. In Nigeria, there are two seasons: rain (April to September) and dry (October to March).

Polycyclic aromatic hydrocarbons (PAHs) contain complex chemicals which include carbon and hydrogen with a fused ring structure, containing at least two benzene rings (Sexton et al., 2011). PAHs have been documented to cause several health problems (Hati et al., 2009). Some PAHs and their derivations are highly toxic. Their mutagenic or carcinogenic properties are the main risk to human health (Prycek et al., 2007). Bitumen contains very small quantities of PAHs at 4 – 6 condensed rings which are not bioavailable. The effects of PAHs on human health depend on the concentration, type and extent of exposure.

PCBs belong to a class of organic compounds having a basic chemical structure, which include two benzene (biphenyl) molecules and between 1 and 10 chlorine atoms substituted on each of the benzene molecules (ATSDR, 2001). There are 209 congeners of PCBs; there are no known natural sources of PCBs. The major sources of PCBs in Nigeria include importation of electrical transformer oils and other anthropogenic activities such as agriculture, discharge of industrial wastes into rivers and incineration (Adeyemi et al., 2009). Studies on both human and laboratory manuals provide strong evidence of the toxic potential of exposure to PCBs (Zaborski et al., 2011). The health effects include liver, thyroid, dermal and ocular changes, reduced birth weight, reproductive toxicity and cancer (Adeyemi et al., 2009).

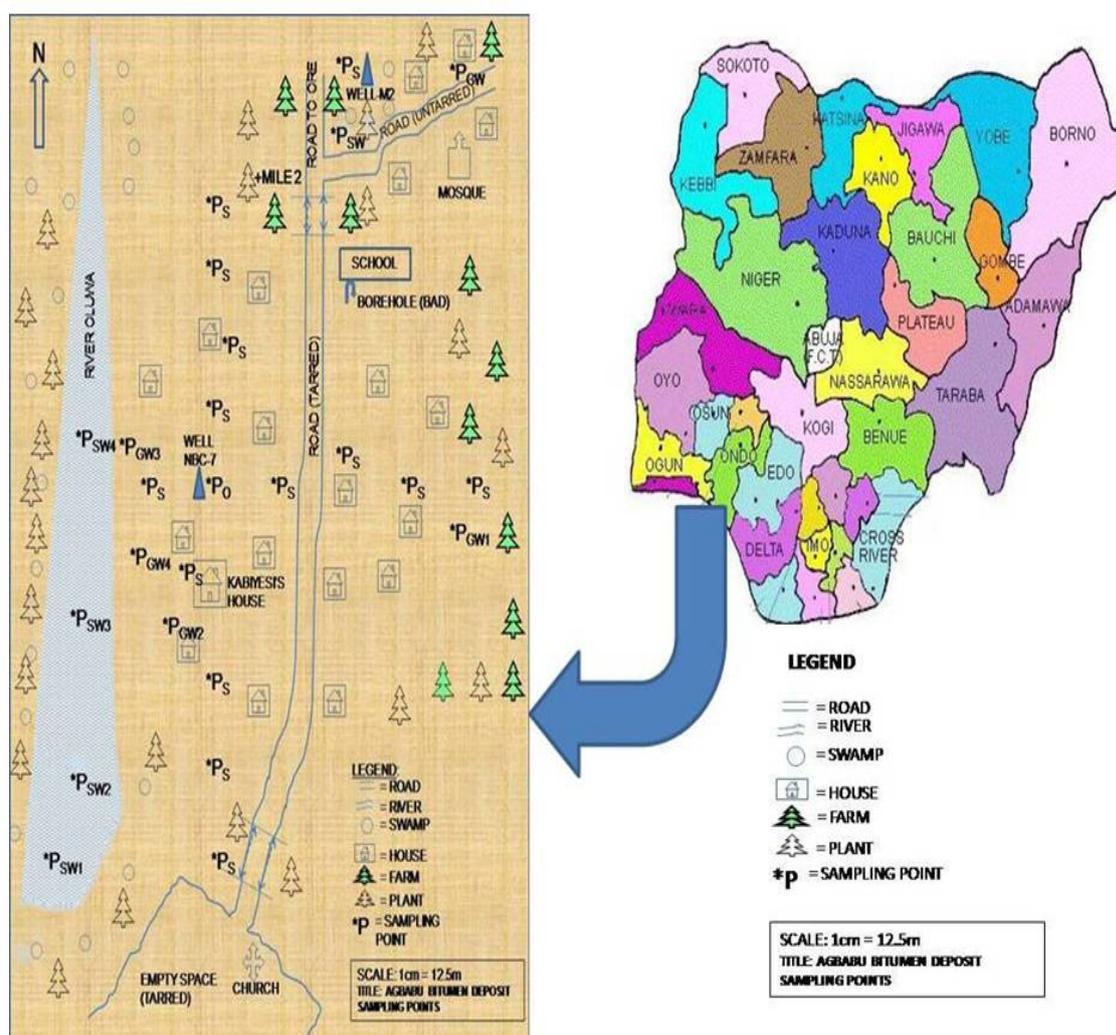


Figure 1: Map of Nigeria showing Agbabu sampling points

Distribution of PAHs in water and soil samples from the vicinity of Agbabu bitumen field has been studied using GC-FID (Abass et al., 2007). n-Alkanes distribution in soil and water samples collected near Agbabu bitumen field of south-western, Nigeria has been reported by Olajire et al. (2008). Evaluation of the status of heavy metals pollution of sediments of Agbabu bitumen deposit area, Nigeria has been investigated (Fagbote and Olanipekun, 2010a). Levels of polycyclic aromatic hydrocarbons and polychlorinated biphenyls in sediments of bitumen deposit impacted area have also been reported by Fagbote and Olanipekun (2010b). Agbabu bitumen deposit area is presently characterized with seepages. The impact of bitumen seepages on the Nigerian physical environment may be enormous (Adewole, 2010). Seasonal and depth effects on physico-chemical parameters of the soils of farm settlements with bitumen deposit have been studied and it was observed that sources of the parameters varied from one location to another, except for Mn^{2+} which had a common source which was bitumen (Fagbote and Olanipekun, 2011).

The aim of the present study was to determine the levels of PAHs and PCBs in soils of the Agbabu bitumen field using GC-MS and GC-FID, respectively. Whenever commercial exploitation of the bitumen commences, it is expected that enormous quantities of both PAHs and PCBs will be introduced into the environment through anthropogenic activities. Knowledge of the pre-exploitation status of pollution in the area will thus provide baseline data which are essential in monitoring the extent of contamination by toxic substances.

Experimental

Collection of Samples

Samples were collected from five different points in Agbabu, Nigeria during the dry (March) and rainy (August) seasons. The sampling points were geo-located with GPS to ensure consistency. Composite soil samples were wrapped with aluminium foil, kept in dry ice and stored at -20°C.

Extraction

The samples were dried in an oven at 105°C, ground, sieved with 1mm sieve and extracted with dichloromethane using a Soxhlet extractor. The extracts were concentrated using Kuderna Danish (K-D) concentrator. The extracts were analyzed for PAHs and PCBs.

Methods of Analysis

A. PAHs in the extracts were quantified using Agilent 6890N gas chromatograph coupled with mass spectrometer (Lemoine and Hoberecht, 2000).

Extract peaks generated that were above scale were diluted with methylene chloride and re-analyzed.

Detection limits are <0.01ppm for Benzo(k)fluoranthene, <0.03ppm for Benzo(g,h,i)perylene and <0.02ppm for Naphthalene, 2-Methylnaphthalene, Acenaphthylene, Acenaphthene, Fluorene, Phenanthrene, Anthracene, Fluoranthene, Pyrene, Benzo(a)anthracene, Chrysene, Benzo(b)fluoranthene, Benzo(a)pyrene, Dibenzo(a,h)anthracene, Indeno(1,2,3-d)pyrene.

GC oven program for PAHs is as follows:

Column type – 30meters x 0.25mm ID x 0.25µm film thickness silicone-coated fused-silica capillary column (DB-5 or HP-5).

Oven - On

Set point - 65°C

Initial Temp - 65 °C hold for 1 minute
65°C to 140 °C @ 15 °C/minute for 4 minutes

Final Temp - 140 °C to 290 °C @ 10 °C/minute hold for 10 minutes

To validate the procedure the following quality controls were observed:

- A solvent blank was run every day prior to analysis. After every 20 samples, a sample was run as duplicate.
- The Relative Percent Difference (RPD%) was determined as follows:

$$\%RPD = [2(D1 - D2)/(D1 + D2)] \times 100$$

where D1 = Concentration of the analyte in the first duplicate sample

D2 = Concentration of the analyte in the second duplicate sample

- A mid-point calibration was run after every 20 samples. For each sample analyzed, the percentage recovery of each surrogate was calculated
- When recovery was not within 80 – 120%, it was checked to be sure that there were no errors in the calculation, otherwise the sample was re-extracted and re-analyzed.

B. PCBs in samples were analyzed using GC 610 connected to FID (Clescery et al., 1989).

Extract peaks generated that were above scale were diluted with dichloromethane and re-analyzed.

Detection limit for PCBs is <0.01 ppm.

GC oven program for PCB is as follows:

Column type – 30meters x 0.25mm ID x 0.25µm film thickness silicone-coated fused-silica capillary column (DB-5 or HP-5)

Oven - On

Set point - 125°C

Initial Temp - 125 °C hold for 3 minutes

Final Temp - 125 °C to 270 °C @ 12 °C/minute and held for 2 minutes

The procedure was validated as in (A) above.

C. Organic carbon in the samples was determined by the wet combustion method

(Schumacher, 2002)

Results and Discussion

Tables 1 and 2 show the sum of concentrations of PAHs in the soil samples in the dry and rainy seasons, respectively. In both seasons, the sum of concentrations of PAHs was highest at AGMILE2 (0.07 ± 0.036 ppm for dry) and (0.69 ± 0.19 ppm for rainy); no PAHs were detected at T2E in the dry season, and T1S and T2N in the rainy season.

Generally, sum of concentrations of PAHs in soil samples analysed were higher in the rainy season than in the dry season; this observation is in agreement with what was reported by Teaf (2008). Some of the Σ PAH concentrations determined were within the limit of values expected for a rural environment as given in Table 4. However, sum of concentrations of PAHs obtained in the rainy season: 0.20 ± 0.04 ppm for Benzo(b)flouranthene, 0.23 ± 0.06 for Benzo(k)flouranthene, 0.21 ± 0.05 ppm for Benzo(g, h, i)perylene, and 0.32 ± 0.04 ppm for Indenol (1, 2, 3 - cd) pyrene; exceeded the recommended limits. Σ PAH concentrations in the soil samples ranged between 0.02 ± 0.015 ppm and 0.07 ± 0.036 ppm in the dry season, and 0.22 ± 0.053 ppm and 0.69 ± 0.019 ppm in the rainy season. These values are higher than Σ PAH concentrations of 1 - 10 μ g/kg, indicating that the PAHs were likely to come from anthropogenic sources (Olajire and Brack, 2005).

A significant factor that controls the concentration of PAHs in soil is total organic carbon (Opune et al., 2007). The organic carbon of soils determines the sorption of organic molecules. The linear regression analyses between the mean concentrations and percentage soil organic carbons are shown in Figures 2 and 3 for dry and rainy seasons, respectively; while the correlation coefficients for Figures 2 and 3 are given in Tables 5 and 6, respectively. The linear regression analysis showed no significant correlation ($r^2=0.002$) between mean concentrations and percentage soil organic carbon in the dry season samples. This suggests that the sites were contaminated to different levels probably due to easy degradation of PAHs. However, there was a significant correlation ($r^2=0.57$) between mean concentrations and percentage soil organic carbon in the rainy season samples. This implies that during the rainy season, degradation of PAHs was apparently slowed down and there was a significant influence of percentage organic carbon on the distribution of PAHs in the soil samples.

Table 1: Sum of concentrations and ratios of PAHs in soil (Dry season)

S/No	Name		T0 (SOIL) ppm	T2E (SOIL) ppm	T1S (SOIL) ppm	T2N (SOIL) ppm	AGMILE2 (SOIL) ppm
1	Benzo(b)flouranthene (BbF)	5 rings	0.02 \pm 0.01	ND	ND	ND	0.02 \pm 0.006
2	Benzo(k)flouranthene (BkF)		0.02 \pm 0.005	ND	ND	ND	0.02 \pm 0.01
3	Benzo(a)pyrene (PaP)		ND	ND	0.02 \pm 0.008	0.03 \pm 0.01	0.03 \pm 0.02
		Sum	0.04 \pm 0.015	0	0.02 \pm 0.008	0.03 \pm 0.01	0.07 \pm 0.036
	SumPAHs		0.02 \pm0.015	ND	0.02 \pm0.008	0.03 \pm0.01	0.07 \pm0.036
	% Org. Carbon		1.2 \pm 0.01	2.28 \pm 0.02	2.27 \pm 0.07	2.21 \pm 0.07	2.03 \pm 0.02
	RATIOS OF PAHs						
	BbF/BkF		1.00	ND	ND	ND	1.00
	BkF/BaP		ND	ND	ND	ND	0.67
	BbF/BaP		ND	ND	ND	ND	0.67

ND = Not Detectable

Table 2: Sum of concentrations and ratios of PAHs in soil (Rainy season)

S/No	Name		T0 (SOIL) ppm	T2E (SOIL) ppm	T1S (SOIL) ppm	T2N (SOIL) ppm	AGMILE2 (SOIL) ppm
1	Naphthalene (NAPH)	2-3 rings	0.01 \pm 0.003	ND	ND	ND	0.02 \pm 0.01
2	Acenaphthylene		ND	0.04 \pm 0.02	ND	ND	ND
3	Anthracene (ANT)		ND	ND	ND	ND	0.06 \pm 0.02
4	Phenanthrene (PHEN)		ND	ND	ND	ND	0.07 \pm 0.03
		Sum	0.01 \pm 0.003	0.04 \pm 0.02	0	0	0.15 \pm 0.06
5	Benzo(b)flouranthene	5	ND	ND	ND	ND	0.2 \pm 0.04

	(BbF)	rings							
6	Benzo(k)flouranthene (BkF)		ND	ND	ND	ND	ND	0.23	±0.06
7	Benzo(a)pyrene (PaP)		ND	ND	ND	ND	ND	0.11	±0.03
8	Benzo(g,h,i)perylene (BPE)		0.21 ±0.05	ND	ND	ND	ND	ND	
	Sum		0.21 ±0.05	0	0	0	0	0.54	±0.13
9	Indeno (1,2,3-cd)pyrene (INPY)	6 rings	ND	0.32 ±0.04	ND	ND	ND	ND	
	SumPAHs		0.22 ±0.05	0.36 ±0.06	ND	ND	ND	0.69	±0.19
	% Org. Carbon		2.40 ±0.08	2.38 ±0.03	2.7 ±0.14	2.95 ±0.27	2.40 ±0.02		
	RATIOS OF PAHs								
	NAPH/ANT		ND	ND	ND	ND	ND	0.33	
	NAPH/PHEN		ND	ND	ND	ND	ND	0.29	
	ANT/PHEN		ND	ND	ND	ND	ND	0.86	
	BbF/BkF		ND	ND	ND	ND	ND	0.87	
	BkF/BaP		ND	ND	ND	ND	ND	2.09	
	BbF/BaP		ND	ND	ND	ND	ND	1.82	

ND = Not Detectable

Table 3: Ratios of PAHs from literature

PAHs	RATIO OF PAH	INFERENCE
PHEN/ANT	3	Motor vehicle exhaust
PHEN/ANT	>50	From mineral oil
FLT/PYR	1	Paralytic process
FLT/PYR	>1	Pathogenic process
BaP/PaP	1	Motor vehicle exhaust
BaP/PBE	0.2 – 05	Motor vehicle exhaust
BaP/PBE	>1	Coal combustion
IND/BPE	0.37	PAH arising from traffic exhaust
PYR/BaP	<1 to 50	Diesel fuel-powered truck exhaust
NAPH/PHEN	>1	Pathogenic sources
BbF/BkF	0.8-1.1	Wood combustion
	1.1-1.5	Gasoline exhaust
	2.5-2.9	Smelters
	3.5-3.9	Coal/coke

(Olajire and Brack, 2005).

Table 4: Background concentrations of PAHs in rural soils

PAH	Conc. (x10 ⁻³ mg/kg)
Acenaphthene	1.7
Acenaphthylene	NE
Anthracene	NE
Benzo(a)anthracene	5-20
Benzo(a)pyrene	2-1,300
Benzo(b)flouranthene	20-30
Benzo(e)perylene	NE
Benzo(g,h,i)perylene	10-70
Benzo(k)flouranthene	10-110
Chrysene	38.3
Flouranthene	0.3-40
Flourene	NE
Indeno(1,2,3-c,d) pyrene	10-15
Phenanthrene	30.0
Pyrene	1-19.7

NE = Not Established.

(Jones et al., 1987; Eviassogie et al., 2006).

Table 5: Correlation coefficient table for Figure 2 (PAH versus % Org carbon in soil – Dry season)

	T0 (Soil)	T2E	T1S	T2N	AGM2	total
Sum(PAHs) (x)	0.02	0	0	0.03	0.07	0.14
% Org. Carbon (y)	1.2	2.28	2.3	2.21	2.03	9.99
x ²	0.0004	0	0	9E-04	0.005	0.01
y ²	1.44	5.1984	5.2	4.884	4.121	20.8
Xy	0.024	0	0	0.066	0.142	0.28

-0.040556267 Corr. Coeff
 0.001644811 R2

Table 6: Correlation coefficients for Figure 3 (PAH versus % Org carbon in soil – Rainy season)

	T0 (Soil)	T2E	T1S	T2N	AGM2	Total
Sum(PAHs) (x)	0.22	0.36	0	0	0.67	1.25
% Org. Carbon (y)	2.4	2.38	2.7	2.95	2.4	12.83
x ²	0.0484	0.1296	0	0	0.449	0.627
y ²	5.76	5.6644	7.29	8.703	5.76	33.18
Xy	0.528	0.8568	0	0	1.608	2.993

-0.758086081 Corr. Coeff
 0.574694506 R2

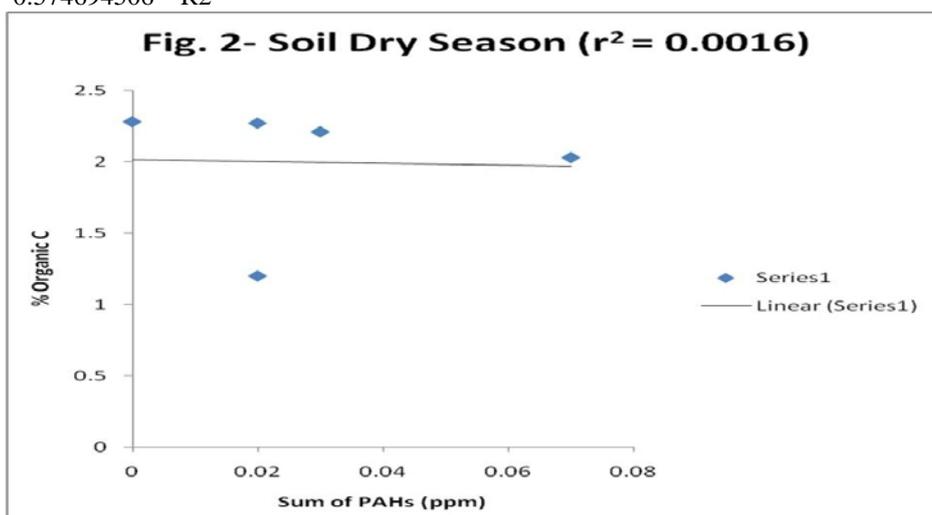


Figure 2: Mean concentration versus percentage organic carbon (Dry season)

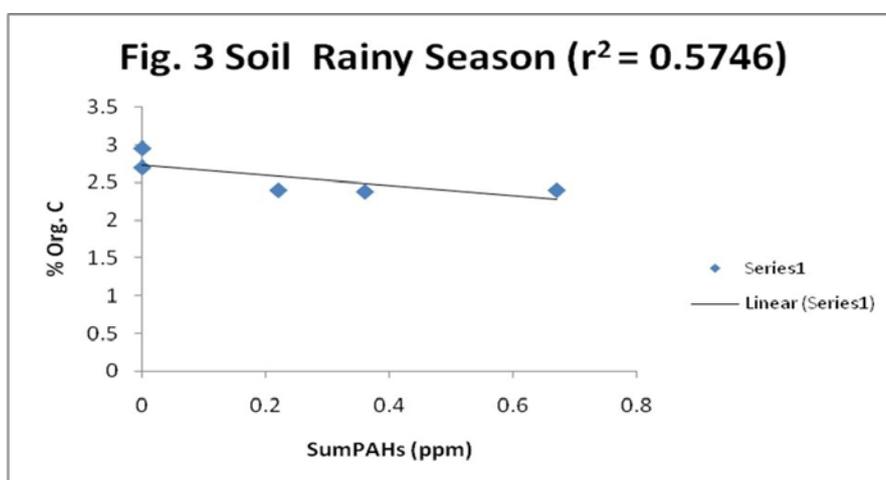


Figure 3: Mean concentration versus percentage organic carbon (Rainy season)

Sources of PAHs and PCBs

AGMILE2 (Station 5) was the most polluted with PAHs, probably because of its proximity to the road. PAHs in bitumen are not available when they are in the pure form, but when in contact with some solvents, they become mobile and can contaminate the environment. In the rainy season, naphthalene (0.02 ± 0.01 ppm), anthracene (0.06 ± 0.02 ppm) and phenanthrene (0.07 ± 0.03 ppm) were detected at AGMILE2. These PAHs with 2-3 rings were probably coming from petrogenic sources. PAHs of 2 – 3 rings were not detectable at this sampling point in the dry season, but those with 4-6 rings were determined in both seasons at this location, which showed that this group of PAHs was more predominant. On the whole, the most predominant PAHs in the Agbabu soils were those with 4-6 rings (Tables 1 and 2), indicating that they were probably formed by pyrolysis (Olajire and Brack, 2005). Several PAHs containing 4-6 aromatic rings in their structure are known to be mutagenic, carcinogenic and inducers of tumours in mammals (Nagpal, 1994).

The most significant sources of PAHs in an environment can be identified from the ratios between pairs of concentrations of individual PAHs, and the literature values, presented in Table 3, were used to characterize the PAHs in this study. The ratios of the concentrations of PAHs calculated from Tables 1 and 2 for BENZO(b) FLOURANTHENE/BENZO (k) FLOURANTHENE (BbF/BkF) were 1.0 at T0 and AGMILE2 in the dry season, and 0.87 for AGMILE2 in the rainy season. These values suggest that the PAHs were apparently from wood combustion sources. The ratio of PHENANTRENE/ANTHRACENE (PHEN/ANT) was 0.86 at AGMILE2 in the rainy season, suggesting that the PAHs were probably not from motor vehicle exhaust and mineral oil; the ratio of NAPHTHALENE/PHENANTRENE (NAPH/PHEN) was 0.29 at AGMILE2, which is indicative that the PAHs were apparently not derived from pathogenic sources. Sum of PAH concentrations above $1-10 \mu\text{g}/\text{kg}$ show that the PAHs may likely come from anthropogenic sources and not from plant synthesis and natural fires (Edward, 1987).

Conclusion

The concentrations of most of the PAHs in the soils at the study area were lower than the recommended limits and the PAHs were probably from anthropogenic sources, including pyrolysis. Higher concentrations were obtained in the rainy season, apparently due to dilution and lower rates of degradation of the PAHs. The predominance of PAHs containing 4 – 6 rings in the investigated samples suggests that they could be mutagenic and carcinogenic.

As at the time of this study, PCBs were not detectable in any of the soil samples investigated; but despite this, it was still necessary to know the pre-exploitation status of PCBs in the area. Finally, monitoring of the levels toxic substances in Agbabu and its vicinity should be conducted periodically.

References:

- Abass, A.A.; Abass, O. A.; Adejare, A. A and M. D. Ojeyemi. Distribution of Polycyclic Aromatic Hydrocarbons in surface soils and water from the vicinity of Agbabu bitumen field of south-western Nigeria. *Journal of Environmental Science and Health, Part A*. Vol. 42, No. 8, pp. 1043 – 1049, 2007.
- Adegoke, O. S and E. C. Ibe. (1982) "The tar sand and heavy crude resources of Nigeria" *Proc. 2nd International Conference on Heavy crude and Tarsands, Caracas, Venezuela, ch., 32, pp 280 – 285, 1982.*
- Adewole, M. G. Environmental implications of bitumen seep induced pollution in parts of Ogun state, south-western Nigeria. *Environmental Earth Science* Vol. 59, No. 7, pp. 1507–1514, 2010.
- Adeyemi D.; Anyakora, C and J. Uyimadu. Polychlorinated biphenyl in fish samples from Lagos Lagoon, Nigeria. *African Journal of Biotechnology*, Vol. 8, No. 12, pp. 2811-2815, 2009.
- Agency for Toxic Substances and Disease Registry, ATSDR. (2001). Polychlorinated Biphenyls". Available online at <http://www.atsdr.cdc.gov/tfact17.html>. Accessed in October 2009.
- Clescery L.S.; Greenberg, A. E and R. R. Trussels. Liquid-liquid extraction gas chromatographic method II. APHA 6630c, Standard methods for examination of water and waste water, 17th ed. American Public Health Association American water works Association, Water Pollution Control Federation, 6-170-6-181, 1989.
- Edward, N.T.J. Polycyclic Aromatic Hydrocarbon (PAHs) in the Terrestrial Environment. A review. *Journal of Environmental Quality*, Vol. 12, pp. 427-441, 1987.
- Eviasogie, P.O.; Ukepebor, E. E and U. Onoh. Distribution of Polycyclic Aromatic Hydrocarbons in rural agricultural wetland soils of the Niger Delta Region. *African Journal of Biotechnology*, Vol. 5,

No. 15, pp. 1415-1421, 2006.

Fagbote, E. O and E. O. Olanipekun. Evaluation of the status of heavy metal pollution of sediment of Agbabu bitumen deposit area, Nigeria. *European Journal of Scientific Research*, Vol. 41, No. 3, pp. 373 –382, 2010A

Fagbote, E. O and E. O. Olanipekun. Levels of polycyclic aromatic hydrocarbons and polychlorinated biphenyls in sediment of bitumen deposit impacted area. *International Journal of Environmental Science and Technology*, Vol. 7, No. 3, pp. 561-570, 2010B.

Fagbote, E. O and E. O. Olanipekun. Seasonal and depth effects on the physico-chemical parameters of the soils of farm settlements with bitumen deposit” *International Research Journal of Biotechnology*, Vol 2, No. 9, pp. 198 -212, 2011.

Hati, S. S.; Dinar, G. A.; Egwu, G. O and V. O. Ogunbujaja. Polycyclic Aromatic Hydrocarbons (PAHs) contamination of synthetic industrial essentials utilized in the Northern Nigeria. *African Journal of Pure and Applied Chemistry*, Vol. 3, No. 5, pp. 86 – 91, 2009.

Jones K.C.; Stratford, J. A and K. Waterhouse. Polynuclear Aromatic Hydrocarbons in U.K. soils, Long term temporal trends and current levels, *Trace Subst. Environmental Health*, Vol. 2, pp. 140 – 1480, 1987.

Lemoine, E. A and H. Hoberecht. Multicomponent Analyte Determination, Gas Chromatography” USEPA Method 8270, 2000.

Nagpal, N. K. (1994). Development of water quality and sediment criteria for PAHs to potent aquatic life”. Available online at <http://www.heb.pac.afo-mpo.ca/congress/1994/nagpal.pdf>. Accessed in October 2009.

Oboh, O. B.; Ilori, M. O.; Akinyemi, J. O and S. A. Adebuseye Hydrocarbon Degrading potentials of bacteria isolated from a Nigerian bitumen (Tarsand) deposit *Nature and Sciences* Vol. 4, No. 3, pp. 51-58, 2006.

Olajire A and N. Brack. Polycyclic Aromatic hydrocarbons in Niger Delta soil contamination, sources and profiles. *International Journal Environmental Science and Technology*, Vol. 2, No. 4, pp. 343 – 352, 2005

Olajire, A. A.; Olujobade, M and O. Olabemiwo. N-Alkane distribution in soil and Water samples collected near Agbabu bitumen field of South Western Nigeria. *International Journal of Environmental Studies*, Vol. 65, No. 6, pp. 769-779, 2008.

Opuene, K.; Agbozu, I. E and C. C. Iwuozor. Identification of perylene in sediments: Occurrence and diagenetic evolution *International Journal of Environmental Science and Technology* Vol. 4, No. 4, pp. 457 – 462, 2007.

Prycek J.; Ciganek, M and Z. Srinek. Clean up of extracts for nitrated derivatives of polycyclic aromatic hydrocarbons analysis prior their gas chromatography determination” *Journal Brazilian Chemical Society* Vol. 18, pp. 1125 – 1131, 2007.

Schumacher B. A. (2002). Methods for the determination of Total Organic Carbon (TOC) in soils and sediments. United States Environmental Protection Agency. Available online at <http://www.epa.gov/esd/cmb/research/papers/bs116.pdf>. Accessed in October, 2009.

Sexton, K.; Sahmas, J. J.; MacDonald, T. J.; Gowen, R. M. Z.; Miller R. P.; McCormick, J. B and S. P. Fisher-Hosh. PAHs in Maternal Unbilical cord blood from pregnant Hispanic women living in Brownsville, Texas. *International Journal Environmental Research and Public Health* Vol.8, No. 8, pp. 3365 – 3379, 2011.

Teaf, C.M. Polycyclic Aromatic Hydrocarbons (PAHs) in urban soil; A Florida Risk Assessment Perspective. *International Journal of soil, sediment and water* Vol. 1, No. 2, pp. 1-15, 2008

Zaborski, A.; Carroll, J.; Pazdro, K and J. Pemkowiak. Spatio-temporal patterns of PAHs and PCBs and HCBs in sediments of the Western Barents Sea. *Oceanologia* Vol. 53, No. 4, pp. 1005 – 1026, 2011.

COMPARATIVE STUDY OF PEAT HUMIC ACIDS BY USING UV SPECTROSCOPY

Oskars Purmalis, PhD Candidate
Maris Klavins, Dr. hab. chem, professor
University of Latvia, Latvia

Abstract:

Studies of the living organic matter humification process are essential for understanding of the carbon biogeochemical cycle. The aim of this study is to analyze relations between the properties of peat humic acids and peat humification degree using UV spectroscopy. The analysis has been done on samples of humic substances extracted from peat profiles in three ombrotrophic bogs and relations between peat age, decomposition and humification degree, botanical composition and properties of peat humic acids (elemental, functional composition) were studied.

The found variability of peat properties is less significant than differences in the properties of peat-forming living matter, thus revealing the dominant impact of humification process on the properties of peat. Correspondingly, composition of peat humic acids is little affected by differences in the composition of precursor living organic material. UV spectroscopy and used absorbance ratios can be successfully used for describing organic matter properties and diagenesis.

Key Words: UV spectroscopy, peat, humic acids

Introduction

Peat is a light brown to black organic material, which is formed under waterlogged conditions from the partial decomposition of mosses and other bryophytes, sedges, grasses, shrubs, or trees (Cocozza et al., 2003). The interest in peat properties is growing, as peat is a substance that supports and influences bog and wetland ecosystems, while peat profiles can serve as “archives” indicating conditions in past environments (Yeloff and Mauquoy, 2006; Zaccone et al., 2007). Significant amounts of organic carbon are stored in the form of peat. Therefore, peat reserves play a major role in the carbon biogeochemical cycling, which is of key importance in the context of the on-going process of climate change (Falkowski et al., 2003; Borgmark, 2005). Industrial and agricultural uses of peat are growing (Ghaly et al., 1999; Brown et al., 2000), and significant amounts of peat are mined industrially. Considering this, there is an increasing interest into studies of peat properties and their diagenesis.

In the carbon biogeochemical cycle, the transformation of living organic matter into refractory part of organic matter (humic substances, such as humic acids, fulvic acids, and humin) or humification is of key importance. Humification can be defined as the transformation of numerous groups of substances (proteins, carbohydrates, lipids etc.) and individual molecules present in living organic matter into groups of substances with similar properties (humic substances) (Francioso et al., 2003). To characterize the humification process, an important tool is the development of humification indexes that link the rate of transformation of living organic matter, development of humic substances with parameters that describe the properties of formed materials (Lu et al., 2001). Several humification indexes have been suggested to study the humification process during composting to evaluate maturity of compost (Jerzykiewicz et al., 1999; Domeizel et al., 2004) and to study soil formation processes (Zsolnay et al., 1999; Cavani et al., 2003; Ikeya and Watanabe, 2003; Rosa et al., 2005; Corvasce et al., 2006). Usually the humification process is evaluated using indirect measurements that describe structural changes occurring during the humification process. Several methods have been suggested, such as measurement of E_4/E_6 ratio that indicates the development of condensed macromolecules, the amount of organic/aliphatic carbon estimated with ^{13}C nuclear magnetic resonance spectra. Furthermore, measurements of presence of free radicals, determined

using electron spin resonance spectra and studies of fluorescence properties of humic macromolecules, have been used to describe humification processes (Milori et al., 2002). However, only a few articles study the humification process of peat (Schnitzer and Levesque, 1979; Preston et al., 1989; Hargitai, 1994; Baran, 2002; Francioso et al., 2003; Štíre et al., 2008) despite the fact that bogs and wetlands form one of the largest sources of refractory organic matter.

The aim of this study is to analyze relations between the properties of peat humic acids and peat humification degree using UV spectroscopy.

Materials and methods

Site location

In-depth study of peat composition, humification degree and peat humic acid properties was carried out in three ombrotrophic bogs located in Latvia. Full peat profiles were obtained and cut into 5-cm layers for analysis of peat properties and isolation of humic acids.

Isolation of peat humic acids

HAs were extracted and purified, using procedures recommended by the International Humic Substances Society (IHSS) (Tan, 2005).

Analysis of peat and humic acid properties

The ¹⁴C dating was done at the Institute of Geology of the Tallinn Technical University (Estonia). Carbon, hydrogen, nitrogen and sulphur concentrations in the peat and humic acid samples (elemental analysis of C, H, N, S) were determined by combustion-gas chromatography technique, using an Elemental Analyzer Model EA-1108 (Carlo Erba Instruments). Ash content was measured after heating 50 mg of each peat sample at 750 °C for 8 h. Elemental composition was corrected considering the ash content, and the oxygen amount was calculated as a difference.

The analysis of botanical composition was performed microscopically, using a Carl-Zeiss binocular microscope, thereby determining the decomposition degree (Listvan, 1975).

For estimation of carboxylic groups and total acidity, an automatic titrator TitroLine easy (Schott-Geräte GmbH) was used. The known Ca-acetate method (Tan, 2005), based on the formation of acetic acid, was used for determining the total number of carboxylic groups. To estimate the total acidity, Ba(OH)₂ method was used (Tan, 2005).

Humification degree (according to (Blackford and Chambers, 1993) and modified by (Borgmark, 2005)) was estimated as absorption at 540 nm after treatment of 1.00 g of peat sample for 1½ hrs with 25 ml of 8 % NaOH in 25 ml plastic tube in a boiling water bath (95 °C) and filtration.

UV spectroscopy

UV-Vis spectra were recorded on a Thermospectronic Helios γ UV (Thermolectron Co) spectrophotometer in a 1-cm quartz cuvette. The UV-Vis spectral ratios E_2/E_3 (Peuravuori et al., 1997) and E_4/E_6 ratio (Chen et al., 1977), i.e. the ratio of absorbance at 280 and 360, and 465 and 665 nm respectively, was determined for humic acid solutions 10 mg/l diluted in 0.05 M NaOH.

Absorbance ratios $E_{280/472}$, $E_{280/664}$, $E_{472/664}$ were calculated from UV-Vis absorbance at 280, 472 and 664 nm (Albrecht, 2011). Absorbance ratio $E_{270/400}$ were calculated from UV-Vis absorbance at 270 and 400 nm (Uyguner and Bekbolet, 2005). Absorbance ratio $E_{254/436}$ were calculated from UV-Vis absorbance at 254 and 436 nm (Shirshova et al., 2006).

Specific absorbance A_{280} (Chin et al. 1994), A_{254} and A_{436} (Shirshova et al., 2006), A_{340} (Scotta et al., 2001) were calculated based on UV-Vis absorbance at 280, 254, 436 and 340 nm, where values were normalized per mol C with units L mol⁻¹ cm⁻¹ of organic matter in solution (Chin et al. 1994).

The value of $\Delta \log K$ coefficient, $\Delta \log K = \log A_{400} - \log A_{600}$, was calculated to determine degree of humification (Fong et al., 2006).

Results and discussion

For description of molecular properties of humic acids (HA) UV-Vis spectroscopy is widely used as simple and informative method (Shirshova et al., 2006). It is known that the UV-Vis spectra of humic acids are broad, featureless and monotonously decrease with increasing wavelength, but there are UV regions and absorbances of spectra that can be used to analyse HA (Uyguner and Bekbolet, 2005). Vieyra (Vieyra et al., 2009) suggests to describe UV absorption spectra by three Gaussian-like bands which are related to electronic transition of benzene-like molecules. Very intense band centred to 180 nm due to allowed π - π^* transition. Less intense benzenoid and electron transfer bands are centred to 205 nm and 250 nm and largely are affected by the presence of polar functional

groups, such as hydroxyl, carbonyl, carboxyl, etc., whereas the benzenoid band is almost unaffected (Vieyra et al., 2009). The UV absorbance at 280 nm represents total aromaticity, because $\pi-\pi^*$ electron transition occurs in this UV region (ca. 270-280 nm) common for phenolic arenes, benzoic acids, aniline derivatives, polyenes and polycyclic aromatic hydrocarbons with two or more rings (Chin et al., 1994; Glatzel et al., 2003) and also specific absorbance A_{280} can be used to characterize humification of the peat (Chin et al. 1994). Many of these compounds are thought to be common structural subunits in humic matter and E_2/E_3 as absorbance ratios measured at 250 and 365 nm corresponds to correlation of molecular size and aromaticity (Peuravuori et al., 1997). To describe aromaticity of humic acids with UV absorbance at 280 nm (Chin et al. 1994) also absorbance at 254 nm is used (Uyguner and Bekbolet, 2005). Studies of aquatic humic acid photocatalytic degradation shows that specific absorbance A_{254} decreases during photocatalytic degradation (Eggin et al., 1997) and may represent an alternative indicative parameter for research humic acids (Lipski et al., 1999; Kumke et al., 2001; Uyguner and Bekbolet, 2005).

The relationship E_4/E_6 (the ratio absorbances at 465 nm and 665 nm) is related to the aromaticity and to the degree of condensation of the chain of aromatic carbons of the humic acids, and could be used as a humification index (Kononova 1966; Stevenson and Schnitzer 1982). Low E_4/E_6 ratio reflect a high degree of condensation of these structures while high ratios mean presence of large quantities of aliphatic structures and low quantities of condensed aromatic structures (Chin et al., 1997). This ratio also is inversely related to the degree of aromaticity, particle size, molecular weight, and acidity (Uyguner et al., 2004). Typically E_4/E_6 is larger for non-humified material by presence of proteins and carbohydrates, which increase the absorbance at the UV region of the spectrum (Vieyra et al., 2009).

Coefficient $E_{270/400}$ (the ratio absorbances at 270 and 400 nm) characterize the degradation of phenolic/quinoid core of HA to simpler carboxylic aromatic compounds (Uyguner and Bekbolet, 2005).

To analyze humification of organic matter commonly is used UV-Vis absorbance at 260-280 nm, due to lignin and quinone moieties (material at the very beginning of transformation). The absorbance at 460-480 nm reflects the organic material at the beginning of humification and the absorbance at 600-670 nm is indicative of strongly humified material with a high degree of aromatic, condensed groups (Albrecht, 2011). Therefore following absorbance ratios have been used to characterize aquatic humic material $E_{280/472}$, $E_{280/664}$, $E_{472/664}$. The $E_{280/472}$ reflects the proportion between the lignins and other materials at the beginning of humification. The $E_{280/664}$ denotes the relation between non-humified and strongly humified material. $E_{472/664}$ ratio is used to indicate the degree of condensation and polymerization of the aromatic constituents. Low $E_{280/664}$ or $E_{472/664}$ ratios reflect a high degree of aromatic condensation and indicate a higher level of organic material humification (Albrecht, 2011).

The value of $\Delta \log K$ coefficient can be used for describing degree of humification of organic matter, although this coefficient were not significant enough to use this parameter as indicator of humification degree of soil organic matter (Fong et al., 2006). On the basis of $\Delta \log K$ coefficient, humic acids can be divided into 3 types: A – humic acids of high degree of humification for which the values of $\Delta \log K$ coefficient is up to 0.6; type B – those of corresponding values between 0.6-0.8 and type C – where $\Delta \log K$ within range 0.8-1.1 (Fong et al., 2006).

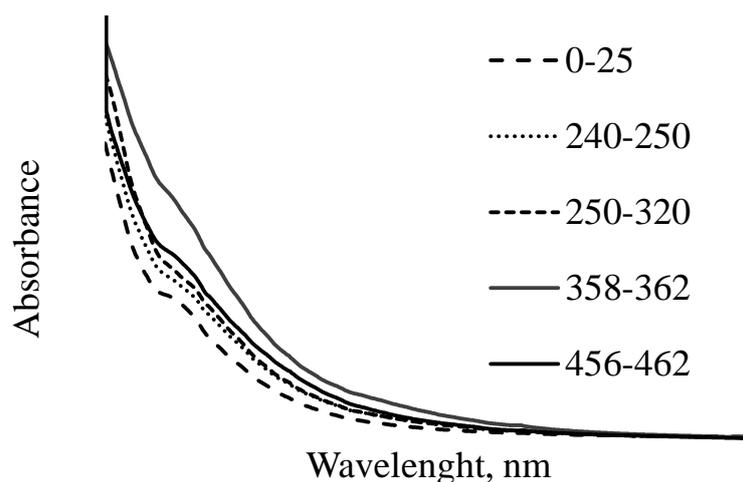


Fig. 1. UV-Vis spectra of peat humic acids from Eipurs bog depending on the sample location depth

The shape of UV-Vis spectra of humic acids revealed a monotonically decrease with increasing wavelength (Fig. 1, 2). Presence of shoulder between 240 and 290 nm are characteristic to aromatic or unsaturated compounds (conjugation of quinone and ketones) (C=C, C=O, N=N) (Chin et al., 1994). Absorption intensity increases with peat depth demonstrating nature of humification process in bogs – increasing aromaticity of isolated HAs. It depends not only on humification process, but also on peat forming biota. Bottom layers of the bog are composed of plants containing aromatic compounds (lignine) in their structure and thus UV-Vis spectra of HAs are affected.

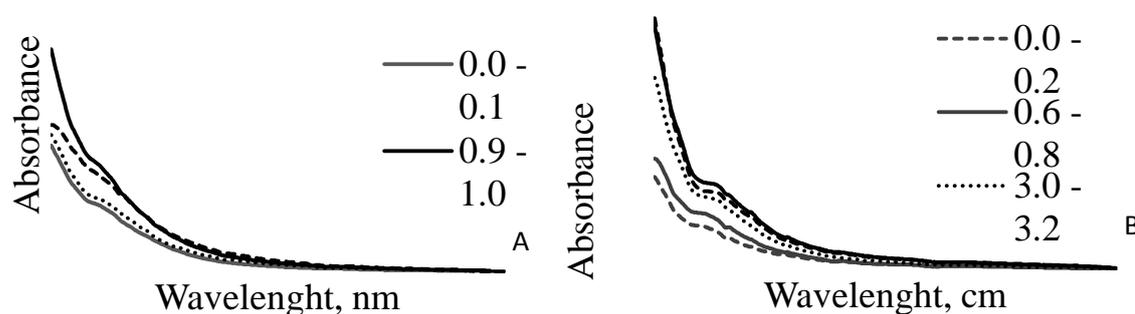


Fig. 2. UV-Vis spectra of peat humic acids from Dzelve (A) and Dizpurvs (B) bog depending on the sample location depth

There are other UV spectra absorbances and absorbance ratios: A_{436} , $E_{254/436}$ suggested by Shirshova (Shirshova et al., 2006), representing that HAs with lower A_{436} , $E_{254/436}$, E_4/E_6 , A_{254} values contain more aliphatics, carbohydrates, aromatics and/or amides compared with HAs which are relatively rich in carboxylic and phenolic groups (Shirshova et al., 2006). The results of mean A_{340} indicated that the humic acid fraction had more colour per unit weight than the fulvic or the hydrophilic fractions, which means that HAs are more aromatic than FAs (Scotta et al., 2001).

Lower E_4/E_6 ratios measured for humic acids from deeper bog layers accordingly to data in the literature (Zaccone et al., 2007) suggest a high degree of condensed aromatic systems and smaller particle sizes or molecular weights and acidity, whereas in upper layers large quantities of aliphatic structures, mainly polysaccharides probably are dominant. The E_4/E_6 correlates with $E_{472/664}$ ratio (Table 1, 2, 3) indicating the degree of condensation and polymerization of the aromatic constituents (Uyguner and Bekbolet, 2005). Typically for studied bogs lower degree of the aromatic constituents has been found for sphagnum, hypnum and sedge peat.

Table 1 Calculated UV-Vis absorbance ratios, specific absorbance A_{280} and $\Delta \log K$ of peat HAs isolated from Eipurs bog

Humic acid	A_{280}	E_4/E_6	E_2/E_3	$E_{280/472}$	$E_{280/664}$	$E_{472/664}$	$E_{270/400}$	$\Delta \log K$
Eipurs HA 0.0 - 0.25	19.99	5	3.4	10.05	47.75	4.75	5	0.81
Eipurs HA 0.25 - 0.50	18.96	7.2	3.46	10.56	71.8	6.8	5.14	0.88
Eipurs HA 0.50 - 0.70	20.16	5.86	3.34	9.67	53.86	5.57	4.8	0.85
Eipurs HA 0.70 - 1.20	18.67	6.28	3.1	8.73	51.14	5.86	4.38	0.83
Eipurs HA 1.20 - 1.35	16.74	6.86	2.96	7.62	49	6.43	4.08	0.78
Eipurs HA 1.35 - 1.70	28.22	6.23	2.84	6.97	40.77	5.85	3.81	0.79
Eipurs HA 1.70 - 1.87	28.64	6.58	2.86	7.03	40.54	5.77	3.88	0.78
Eipurs HA 1.87 - 2.20	25.16	5.92	2.83	7	39.08	5.58	3.78	0.77
Eipurs HA 2.20 - 2.30	26.20	6.36	2.86	7.38	44.27	6	3.85	0.80
Eipurs HA 2.30 - 2.40	25.35	6.4	2.92	7.59	46.3	6.1	4	0.81
Eipurs HA 2.40 - 2.50	24.17	6.77	2.94	7.62	49.11	6.44	4	0.81
Eipurs HA 2.50 - 3.20	24.23	6.2	3.09	8.12	47.1	5.8	4.18	0.80
Eipurs HA 3.20 - 3.58	29.48	8	3.09	7.45	55.9	7.5	4.14	0.82
Eipurs HA 3.58 - 3.62	36.18	6.69	2.96	6.51	41.125	6.31	3.86	0.77
Eipurs HA 3.62 - 4.10	29.95	7.22	3.05	8.77	59.44	6.78	4.31	0.86
Eipurs HA 4.10 - 4.56	28.22	6.55	3.08	9.27	56.67	6.11	4.34	0.85
Eipurs HA 4.56 - 4.62	24.45	7.2	2.76	7.55	50.6	6.7	3.78	0.88

Table 2 Calculated UV-Vis absorbance ratios, specific absorbance A_{280} and $\Delta \log K$ of peat HAs isolated from Dzelve bog

Humic acid	A_{280}	E_4/E_6	E_2/E_3	$E_{280/472}$	$E_{280/664}$	$E_{472/664}$	$E_{270/400}$	$\Delta \log K$
Dzelve HA 0.0 - 0.10	10.68	10	3.31	10	100	10	4.91	0.87
Dzelve HA 0.10 - 0.20	14.52	12	3.36	10.36	114	11	4.77	0.94
Dzelve HA 0.20 - 0.30	16.62	8	3.08	8.53	64	7.5	4.28	0.81
Dzelve HA 0.30 - 0.40	20.87	10	3.39	9.63	73.83	7.67	4.81	0.92
Dzelve HA 0.40 - 0.50	15.76	4	2.76	6.21	50	3.8	3.57	0.64
Dzelve HA 0.50 - 0.60	20.22	8.5	3.35	9.875	79	8	4.75	0.95
Dzelve HA 0.60 - 0.70	15.47	14	3.39	9.85	128	13	4.79	0.99
Dzelve HA 0.70 - 0.80	17.23	7	3.35	9.77	63.5	6.5	4.72	0.86
Dzelve HA 0.80 - 0.90	19.03	8	3.43	9.6	72	7.5	4.73	0.82
Dzelve HA 0.90 - 1.00	21.20	6.3	3.42	8.78	39.5	4.5	4.65	0.79
Dzelve HA 1.00 - 1.10	21.78	7.3	3.22	7.95	55.67	7	4.18	0.86
Dzelve HA 1.10 - 1.20	19.16	4.5	3.38	8.35	35.5	4.25	4.22	0.79
Dzelve HA 1.20 - 1.30	19.84	5	3.17	7.74	36.75	4.75	4.21	0.73
Dzelve HA 1.30 - 1.40	19.81	6.3	3.2	8.11	48.67	6	4.24	0.79
Dzelve HA 1.40 - 1.50	21.42	4.2	3.12	7.74	29.4	3.8	4.07	0.75
Dzelve HA 1.50 - 1.60	15.00	12	3.21	9.36	103	11	4.44	1.10
Dzelve HA 1.60 - 1.70	13.56	10	3.39	10.44	94	9	4.81	1.02
Dzelve HA 1.70 - 1.80	12.25	8	3.5	12.57	88	7	5	0.98
Dzelve HA 1.80 - 1.90	14.01	10	3.32	10.56	95	9	4.63	1.04
Dzelve HA 1.90 - 2.00	12.49	9	3.31	11.125	89	8	4.8	1.00
Dzelve HA 2.00 - 2.10	13.84	6	3.22	9.18	50.5	5.5	4.54	0.78
Dzelve HA 2.10 - 2.20	14.48	11	3.38	10.4	104	10	4.66	1.08
Dzelve HA 2.20 - 2.30	13.27	9	3.63	11.875	95	8	5.2	1.00
Dzelve HA 2.30 - 2.40	15.29	11	3.48	10.9	109	10	4.87	0.90
Dzelve HA 2.40 - 2.50	15.51	4.67	3.1	8.46	36.67	4.33	4.21	0.75
Dzelve HA 2.50 - 2.60	14.21	6	3.24	9.27	51	5.5	4.58	0.78
Dzelve HA 2.60 - 2.70	12.82	5.5	3.3	9.3	46.5	5	4.54	0.74
Dzelve HA 2.70 - 2.80	12.26	5.5	3.25	8.8	44	5	4.57	0.85
Dzelve HA 2.80 - 2.90	24.87	4.2	3.05	7.87	30.17	3.83	4.06	0.78
Dzelve HA 2.90 - 3.00	22.74	4	2.98	7.5	27.5	3.67	4.02	0.74
Dzelve HA 3.00 - 3.10	23.16	3	2.95	6.72	18.67	2.78	3.83	0.59

Dzelve HA 3.10 - 3.20	22.96	3.43	2.94	7.32	23	3.14	3.91	0.69
Dzelve HA 3.20 - 3.30	15.84	5	3.17	8.62	37.33	4.33	4.25	0.75
Dzelve HA 3.30 - 3.40	16.72	5	3.13	8.5	39.67	4.67	4.34	0.76
Dzelve HA 3.40 - 3.50	21.71	5.75	2.97	6.64	36.5	5.5	3.83	0.71

Ratios $E_{270/400}$, $E_{280/472}$ and E_2/E_3 (Table 1, 2, 3) strongly correlates and shows degradation of phenolic/quinoid core of HA (Uyguner and Bekbolet, 2005) as well as proportion between the lignins and other materials at the beginning of humification (Albrecht, 2011) in peat profiles. Also these ratios corresponds to detected humification index (D_{540}) and peat decomposition degree and confirm the same results as other ratios (E_4/E_6 , $E_{472/664}$) calculated from UV-Vis spectra - a high degree of aromatic condensation indicating a higher level of organic material humification in deeper layers of bog and higher proportion of material at the beginning of humification in upper layers of the bog.

Specific absorbance A_{280} indicates aromaticity of HAs and is reversely related with ratio $E_{280/664}$ which denotes the relation between non-humified and strongly humified material, but aren't directly associated with other ratios indicated aromaticity (E_2/E_3 , E_4/E_6 , $E_{270/400}$), although on average level correlates with peat decomposition degree and humification index.

Table 3 Calculated UV-Vis absorbance ratios, specific absorbance A_{280} and $\Delta \log K$ of peat HAs isolated from Dizpurvs bog

Humic acid	A_{280}	E_4/E_6	E_2/E_3	$E_{280/472}$	$E_{280/664}$	$E_{472/664}$	$E_{270/400}$	$\Delta \log K$
Dizpurvs HA 0.0 - 0.20	9.05	3	3.26	8	21.33	2.67	4.38	0.60
Dizpurvs HA 0.20 - 0.40	11.63	6	2.97	8.09	44.5	5.5	4.13	0.76
Dizpurvs HA 0.60 - 0.80	11.79	4.5	3.38	10.625	42.5	4	4.84	0.80
Dizpurvs HA 0.80 - 1.00	17.34	5	2.96	6.89	32.75	4.75	4.03	0.64
Dizpurvs HA 1.00 - 1.20	17.10	4.75	2.96	7.06	31.75	4.5	4.00	0.69
Dizpurvs HA 1.20 - 1.40	19.39	4.6	2.72	6.23	27.4	4.4	3.51	0.66
Dizpurvs HA 1.40 - 1.60	21.69	4.5	2.75	6.2	25.83	4.17	3.54	0.66
Dizpurvs HA 1.60 - 1.80	19.54	6.33	2.98	7.56	45.33	6	4.03	0.71
Dizpurvs HA 1.80 - 2.00	21.40	10	3.04	7.89	75	9.5	4.21	0.80
Dizpurvs HA 2.00 - 2.20	19.78	4	2.74	5.96	22.83	3.83	3.43	0.62
Dizpurvs HA 2.20 - 2.40	20.72	3.86	2.67	5.8	20.71	3.57	3.40	0.57
Dizpurvs HA 2.40 - 2.60	19.20	3.25	2.63	5.48	17.13	3.125	3.35	0.55
Dizpurvs HA 2.60 - 2.80	19.52	3.28	2.78	6.05	19	3.14	3.56	0.59
Dizpurvs HA 2.80 - 3.00	17.24	7	3.31	9.15	59.5	6.5	4.50	0.75
Dizpurvs HA 3.00 - 3.20	17.59	5	3.17	8.5	39.67	4.67	4.39	0.75
Dizpurvs HA 3.20 - 3.40	16.99	4.33	3.23	9.5	38	4	4.42	0.72
Dizpurvs HA 3.40 - 3.60	15.98	4.33	3.05	9	36	4	4.19	0.72
Dizpurvs HA 3.60 - 3.80	18.88	3.14	2.92	6.6	18.86	2.86	3.72	0.56
Dizpurvs HA 3.80 - 4.00	21.76	3.28	2.96	6.9	20.71	3	3.79	0.59
Dizpurvs HA 4.00 - 4.30	19.61	3.12	2.77	5.87	16.875	2.875	3.41	0.57

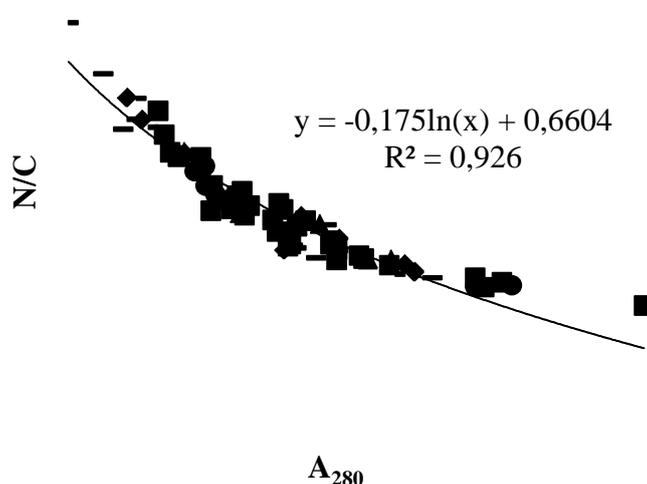
The value of $\Delta \log K$ coefficient for peat humic acids range 0.55 – 1.09 describing them as less humified than soil humic acids (Fong et al., 2006). Value of $\Delta \log K$ coefficient strongly correlates with other calculated UV absorbance ratios (E_4/E_6 , $E_{472/664}$, $E_{270/400}$, $E_{280/472}$, E_2/E_3 , and $E_{280/664}$) except specific absorbance A_{280} (Table 4). If for soil humic acids this coefficient is not significant, then for peat humic acids with lower decomposition degree and quite wide range (10-60 %) of very well characterized humification parameters within peat profiles this coefficient might be used. Moreover, $\Delta \log K$ on average level correlates with humification index – D_{540} .

Table 4 Correlations between calculated UV absorbance ratios (E_4/E_6 , $E_{472/664}$, $E_{270/400}$, $E_{280/472}$, E_2/E_3 , $E_{280/664}$, specific absorbance A_{280} and $\Delta \log K$) and decomposition, D_{540} .

	A_{280}	E_4/E_6	E_2/E_3	$E_{280/472}$	$E_{280/664}$	$E_{472/664}$	$E_{270/400}$
Decomposition, %	0.38	0.01	0.3	0.29	0.08	-	
D_{540}	0.39	0.3	0.26	0.14	-	0.28	
E_4/E_6	0.03						
E_2/E_3	0.22	0.3					
$E_{280/472}$	0.25	0.39	0.83				
$E_{280/664}$	0.11	0.91	0.45	0.59			
$E_{472/664}$	0.03	0.98	0.26	0.36	0.91		
$E_{270/400}$	0.22	0.37	0.92	0.91	0.53	0.33	
$\Delta \log K$	0.03	0.74	0.46	0.64	0.77	0.71	0.55

Significance level $p=0.05$

There are small variability of UV-Vis ratios in Dzelve and Dizpurvs peat profiles indicating other factors influencing structure of HAs without humification of organic material. In Dzelve bog peat profile from 1.5 to 2.5 m have higher E_4/E_6 , $E_{280/472}$, $E_{472/664}$ and $E_{280/664}$ ratios, and $\Delta \log K$, but lower specific absorbance A_{280} (Table 2). In this depth of bog lower decomposition degree of peat and therefore lower aromaticity and higher concentration of phenol groups are observed, but lower concentration of carboxylic groups is detected, which typically increases with depth of peat profile. Decrease of carboxylic group concentration in Dzelve bog and lower decomposition degree of peat are probably due to very intense peat growth rate in past (about 4-5 mm/year). In Dizpurvs peat profile from 2.0 to 2.8 m lower E_4/E_6 , E_2/E_3 , $E_{270/400}$, $E_{280/472}$, $E_{472/664}$, $E_{280/664}$ ratios and $\Delta \log K$ was observed, but without significant changes of specific absorbance A_{280} (Table 3). Higher amount of carboxylic groups correlates with decomposition degree of peat. It could be expected increase of specific absorbance A_{280} , and loss of phenolic groups, but due to low aromaticity depending on botanical composition of peat – Sphagnum mosses are void of aromatic substances, but no major changes in this parameter was observed.

**Fig. 3.** Correlation between specific absorbance A_{280} and N/C relation of studied bogs

All three groups of associated ratios (1 - E_4/E_6 , $E_{472/664}$; 2 - $E_{270/400}$, $E_{280/472}$ and E_2/E_3 ; 3 - specific absorbance A_{280} , $E_{280/664}$) and $\Delta \log K$ shows that aromaticity, acidity as well as degradation of phenolic structures increases with depth of peat profile and only partly have influenced from peat forming biota (Table 4). Moreover, these calculated ratios have not directly related to amount of aromatic carbon detected with ^{13}C NMR spectroscopy. These findings corresponds to investigations in literature (Chen et al., 1977; Chen et al., 2002), where was found that E_4/E_6 ratio not directly related to the relative concentration of condensed aromatic rings ratio, but have mainly governed by the particle size (or particle or molecular weight), correlated with the free radical concentration, contents of O, C, COOH and total acidity.

Most significant relation from calculated ratios have specific absorbance A_{280} and N/C ratio (Fig. 3) and corresponds to results from literature (Kalbitz et al., 1999; Kalbitz et al., 2000) and provides an initial estimate of the degree of humification because peat decomposition and humification tend to result in the decay of carbon structures and thus the relative enrichment of nitrogen.

Conclusion

Different UV ratios can be successfully used as complex parameter for describing humification of organic matter. Calculated UV ratios and specific absorbance shows that aromaticity, acidity as well as degradation of phenols increase with depth of peat profile and only partly have influenced by peat forming biota. Peat humic acids (HA) are formed in a process in which more labile structures (carbohydrates, amino acids, etc.) are destroyed, but thermodynamically more stable aromatic and polyaromatic structures emerge.

Acknowledgements

This work has been supported by the European Social Fund within the project «Support for Doctoral Studies at University of Latvia».

References:

- Albrecht, R., Petit, J.L., Terrom, G. and Périsol, C. (2011) Comparison between UV spectroscopy and nirs to assess humification process during sewage sludge and green wastes co-composting. *Bioresource Technology*, 102(6), 4495–4500
- Baran, A. (2002) Characterization of Carex peat using extinction values of humic acids. *Bioresource Technol.*, 85, 99-101
- Blackford, J.J. and Chambers, F.M. (1993) Determining the degree of peat decomposition for peat-based paleoclimatic studies. *International Peat Journal*, 5, 7-24
- Borgmark, A. (2005) Holocene climate variability and periodicities in south-central Sweden, as interpreted from peat humification analysis. *Holocene*, 15, 387-395
- Borgmark, A. The colour of climate: changes in peat decomposition as a proxy for climate change – a study of raised bogs in South-central Sweden. PhD thesis, Stockholm: Stockholm University, 2005
- Brown, P.A., Gill, S.A. and Allen, S.J. (2000) Metal removal from wastewater using peat. *Wat. Res.*, 34, 3907-3916
- Cavani, L., Ciavatta, C. and Gessa, C. (2003) Identification of organic matter from peat, Leonardite and lignite fertilisers using humification parameters and electrofocusing. *Bioresource Technology*, 86(1), 45–52
- Chen, J., Gu, B., LeBoeuf, J.E., Pan, H. and Dai, S. (2002) Spectroscopic characterization of the structural and functional properties of natural organic matter fractions. *Chemosphere*, 48, 59–68
- Chen, Y., Senesi, N. and Schnitzer, M. (1977) Information provided on humic substances by E_4/E_6 ratios. *Soil Sci. Soc. Am. J.*, 41, 352–358
- Chin, Y., Aiken, R.G. and Danielsen, M.K. (1997) Binding of pyrene to aquatic and commercial humic substances: the role of molecular weight and aromaticity. *Environ. Sci. Technol.*, 31(6), 1630–1635
- Chin, Y.P., Aiken, G. and O'Loughlin, E. (1994) Molecular Weight, Polydispersity, and Spectroscopic Properties of Aquatic Humic Substances. *Environ. Sci. Technol.*, 28(11), 1853-1858
- Cocozza, C., D'Orazio, V., Miano, T.M. and Shotyk, W. (2003) Characterization of solid and aqueous phases of a peat bog profile using molecular fluorescence spectroscopy, ESR and FT-IR, and comparison with physical properties. *Organic Geochemistry*, 34, 49–60
- Corvasce, M., Zsolnay, A., D'Orazio, V., Lopez, R. and Miano T.M. (2006) Characterization of water extractable organic matter in a deep soil profile. *Chemosphere*, 62(10), 1583–1590
- Domeizel, M., Khalil, A. and Prudent, P. (2004) UV spectroscopy: a tool for monitoring humification and for proposing an index of the maturity of compost. *Bioresource Technology*, 94(2), 177–184
- Eggins, B.R., Palmer, F.L. and Byrne J.A. (1997) Photocatalytic treatment of humic substances in drinking water. *Water Research*, 31(5), 1223–1226
- Falkowski, P., Scholes, R.J., Boyle, E., Canadell, J., Canfield, D., Elser, J., Gruber, N., Hibbard, K., Hogberg, P., Linder, S., Mackenzie, F. T., Moore, B., Pedersen, T., Rosenthal, Y. and Tan, K. H. Humic matter in soil and the environment: Principles and Controversies. New York: Marcel Dekker, 2003

- Fong, S.S., Lau, I.L., Chong, W.N., Asing, J., Faizal, M., Nor, M., Satirawaty, A. and Pauzan, M. (2006) Characterization of the coal derived humic acids from Mukah, Sarawak as soil conditioner. *J. Braz. Chem. Soc.*, 17(3), 582-587
- Francioso, O., Ciavatta, C., Montecchio, D., Tugnoli, V., Sanchez-Cortes, S. and Gessa, C. (2003) Quantitative estimation of peat, brown coal and lignite humic acids using chemical parameters, ¹H-NMR and DTA analyses, *Bioresource Technology*, 88, 189-195
- Ghaly, R.A., Pyke, J.B., Ghaly A.E. and Ugursal V.I., (1999) Remediation of Diesel-Oil-Contaminated Soil Using Peat. Energy Sources. A: Recovery, Utilization, and Environmental Effects. *Chemosphere*, 21, 785 – 799
- Glatzela, S., Kalbitz, K., Dalvac, M. and Moore, T. (2003) Dissolved organic matter properties and their relationship to carbon dioxide efflux from restored peat bogs. *Geoderma*, 113, 397– 411
- Hargitai, L. (1994) Biochemical transformation of humic substances during humification related to their environmental functions. *Environ. Internat.*, 20(1), 43-48
- Ikeya, K. and Watanabe, A. (2003) Direct expression of an index for the degree of humification of humic acids using organic carbon concentration. *Soil Science and Plant Nutrition*, 49(1), 47-53
- Jerzykiewicz, M., Drozd, J. and Jezierski, A. (1999) Organic radicals and paramagnetic metal complexes in municipal solid waste composts. An EPR and chemical study. *Chemosphere*, 39(2), 253–268
- Kalbitz, K., Geyer, W. and Geyer, S. (1999) Spectroscopic properties of dissolved humic substances—a reflection of land use history in a fen area. *Biogeochemistry*, 47(2), 219-238
- Kalbitz, K., Geyer, W. and Geyer, S. (2000) A comparative characterization of dissolved organic matter by means of original aqueous samples and isolated humic substances. *Chemosphere*, 40(12), 1305–1312
- Kononova, M.M. Soil organic matter, its nature, its role in soil formation and in soil fertility. 2nd English ed. Oxford: Pergamon, 1966
- Kumke, M.U., Specht, C.H., Brinkmann, T. and Frimmel, F.H. (2001) Alkaline hydrolysis of humic substances – spectroscopic and chromatographic investigations. *Chemosphere*, 45(6-7), 1023–1031
- Lipski, M., Slawinski, J. and Zych, D. (1999) Changes in the luminescent properties of humic acids induced by UV radiation. *J. Fluoresc.*, 9, 133–138
- Lishtvan, I.I., Korol, N.T. Basic properties of peat and methods for their determination, (in Russian), Nauka i Tekhnika, Minsk, 1975
- Lu, X.Q., Hanna, J.V. and Johnson, W.D. (2001) Evidence of chemical pathways of humification: a study of aquatic humic substances heated at various temperatures. *Chemical Geology*. 177(3-4), 249–264
- Milori, D.M.B.P., Neto, L.M., Bayer, C., Mielniczuk, J. and Bagnato, V.S. (2002) Humification Degree of Soil Humic Acids Determined by Fluorescence Spectroscopy. *Soil Sci.*, 167(11), 739–749.
- Peuravuori, J. and Pihlaja, K. (1997) Molecular size distribution and spectroscopic properties of aquatic humic substances. *Anal. Chim. Acta*, 337, 133-149
- Preston, C. M., Axelson, D. E., Lévesque, M., Mathur, S. P., Diné, H., Dudley, R. L. (1989) Carbon 13 NMR and chemical characterization of particle size separates of peats differing in degree of decomposition. *Org. Geochem.*, 14(4), 393-403
- Rosa, A.H., Simões, M.L., Camargo de Oliveira, L., Rocha, J.C., Neto, L.M. and Milori, D.M.B.P. (2005) Multimethod study of the degree of humification of humic substances extracted from different tropical soil profiles in Brazil's Amazonian region. *Geoderma*, 127(1-2), 1–10
- Scotta, M.J., Jones, M.N., Woof, C., Simon, B. and Tipping, E. (2001) The molecular properties of humic substances isolated from a UK upland peat system. A temporal investigation. *Environment International*, 27, 449–462
- Shirshova, L.T., Ghabbour, E.A. and Davies, G. (2006) Spectroscopic characterization of humic acid fractions isolated from soil using different extraction procedures. *Geoderma*, 133(3-4), 204–216
- Schnitzer, M. and Lévesque, M. (1979) Electron Spin Resonance As A Guide to the Degree of Humification of Peats. *Soil Sci.*, 127(3), 140-145
- Stevenson, I.L. and Schnitzer, M. (1982) Transmission electron microscopy of extracted fulvic and humic acids. *Soil Sci.*, 133, 179-185
- Šīre, J., Kļaviņš, M., Purmalis, O. and Meleciš, V. (2008) Experimental study of peat humification indicators. *Proc. Latv. Acad. Sci., B*, 62(1/2), 18–27

- Tan, K.H. Soil sampling, preparation, and analysis. Second Edition. New York: Taylor & Francis, 2005
- Uyguner, C.S. and Bekbolet M. (2005) Evaluation of humic acid photocatalytic degradation by UV–vis and fluorescence spectroscopy. *Catalysis Today*, 101(3), 267-274
- Uyguner, C.S., Hellriegel, C., Otto, W. and Larive, C.K. (2004) Characterization of humic substances: Implications for trihalomethane formation. *Anal. and Bioanal. Chem.*, 378(6), 1579-1586
- Vieyra, F.E.M., Palazzi, V.I., Sanchez de Pinto, M.I. and Borsarelli, C.D. (2009) Combined UV–Vis absorbance and fluorescence properties of extracted humic substances-like for characterization of composting evolution of domestic solid wastes. *Geoderma*, 151(3–4), 61-67
- Yeloff, D. and Mauquoy, D. (2006) The influence of vegetation composition on peat humification: implications for paleoclimatic studies. *Boreas*, 35, 662-673
- Zaccone, C., Miano, T.M. and Shotyk, W. (2007) Qualitative comparison between raw peat and related humic acids in an ombrotrophic bog profile. *Organic Geochemistry*, 38(1), 151–160
- Zsolnay, A., Baigar, E., Jimenez, M., Steinweg, B. and Saccomandi, F. (1999) Differentiating with fluorescence spectroscopy the sources of dissolved organic matter in soils subjected to drying. *Chemosphere*, 38(1), 45–50

SYNTHESIS OF 5-BENZIMIDAZOLYLBENZOFURAN DERIVATIVES OF EXPECTED BIOLOGICAL ACTIVITY

Fatma AbdAleem A.ElHag

Reem A.K. ElHarbi

Fatma M. Edress

Taibah University College of Applied Science,
Applied Chemistry Department, Al-Madinah Al-Munawara, Saudi Arabia

Hoda I. El Diwani

Chemistry of Natural and Microbial Products Department,
National Research Center, Dokki, Cairo, Egypt

Abstract:

The benzimidazole starting material 5-(1*H*-benzo[*d*]imidazol-2-yl)-4,7-dimethoxybenzofuran-6-ol (**2**) was synthesized. The Mannich bases **3a-k** were formed using formaldehyde and primary or secondary amines. (**2**) was subjected to Mannich reaction using 1,4-phenylenediamine or piperazine to lead to the 1,4-disubstituted benzenes **4a** or 1,4-disubstituted piperazines **4b**. The compounds were screened for their antimicrobial activities

Key Words: 5-Benzimidazolylbenzofuran Derivatives, Biological Activity

Introduction

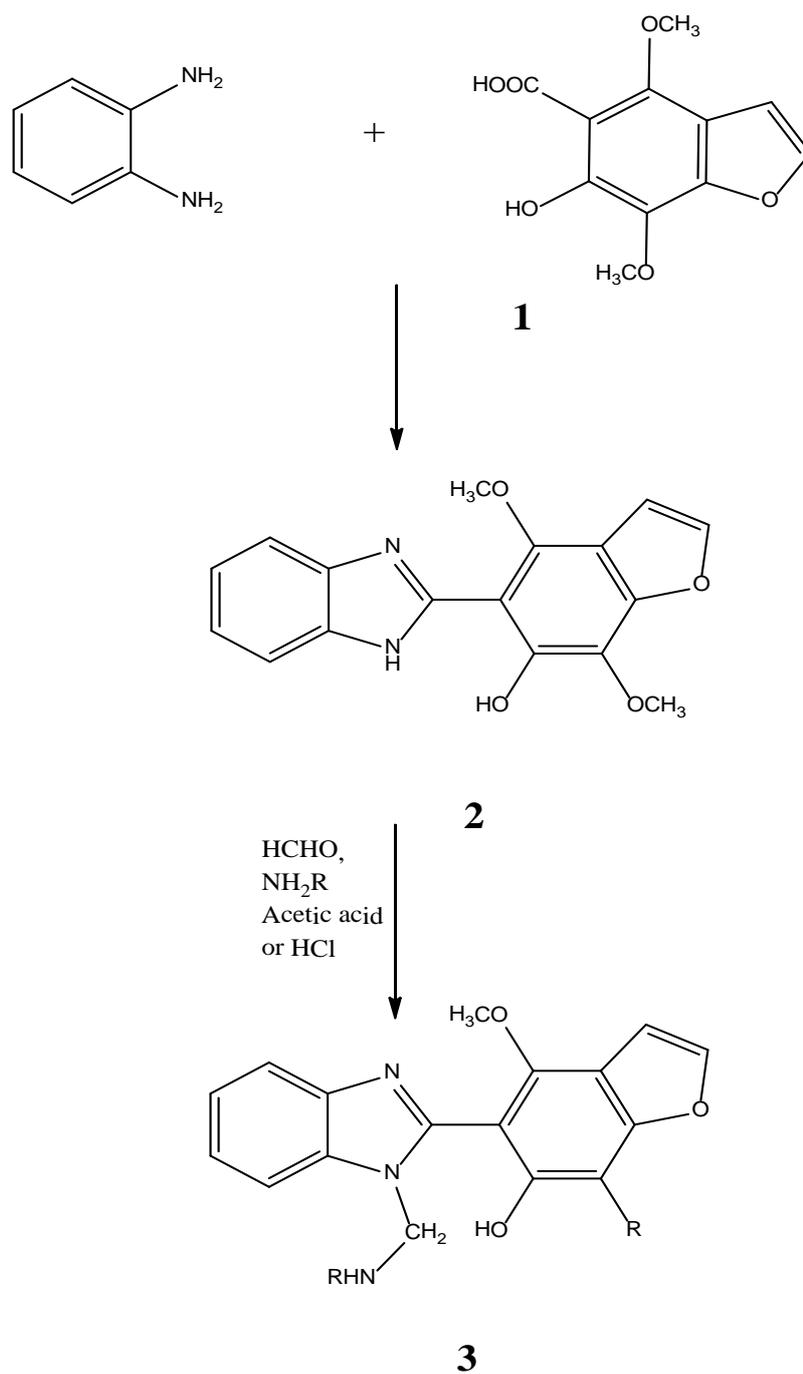
Benzimidazole derivatives are of wide interest because of their diverse biological activity and clinical applications. Several promising antitumor active agents were found to contain the benzimidazole ring system⁽¹⁻¹¹⁾. They were found to exert their antitumor activity by acting mainly as topoisomerases inhibitors⁽¹²⁻¹³⁾. Benzimidazoles belonging to the fused heterocyclic system prepared from amino acids are associated with diverse pharmaceutical activities such as antibacterial⁽¹⁴⁾, insecticidal⁽¹⁵⁾, fungicidal⁽¹⁶⁾, antimicrobial⁽¹⁷⁾, antagonist⁽¹⁸⁾, anthelmintic⁽¹⁹⁻²¹⁾ and anti-inflammatory⁽²²⁾ *etc.* also Benzofuran derivatives possess a wide range of biological activities. They have been reported to possess antimicrobial⁽²³⁻²⁶⁾, antitumor^(27, 28), anti-inflammatory⁽²⁹⁾ activity *etc.* Hence it was thought interesting to prepare the combined molecule having benzimidazole-benzofuran with expected biological activity.

Chemistry

The benzimidazole starting material 5-(1*H*-benzo[*d*]imidazol-2-yl)-4,7-dimethoxybenzofuran-6-ol (**2**) was synthesized using Philips' ⁽³⁰⁾ method starting from *o*-phenylenediamine and 6-hydroxy-4,7-dimethoxybenzofuran-5-carboxylic acid (**1**)

The Mannich bases **3a-k** were formed by the reaction of compound **2** with formaldehyde and 4-chloro-2-nitroaniline or 4-fluoroaniline, 2,6-dichloroaniline, *n*-propylamine, dimethylamine, ethylamine, isopropylamine, piperidine, 2-aminobenzoic acid, or 2-hydroxyaniline respectively. (c.f. scheme 1).

When (**2**) was subjected to Mannich reaction using 1,4-phenylenediamine or piperazine led to the 1,4-disubstituted benzenes **4a** or 1,4-disubstituted piperazines **4b**. (c.f. scheme 2).

Scheme 1:

3a, R= 4-chloro2-nitrophenyl

b, R= 4-chlorophenyl

c, R= 4-fluorophenyl

d, R= 2,6-dichlorophenyl

e, R= n-propyl

f, R= dimethyl

g, R= ethyl

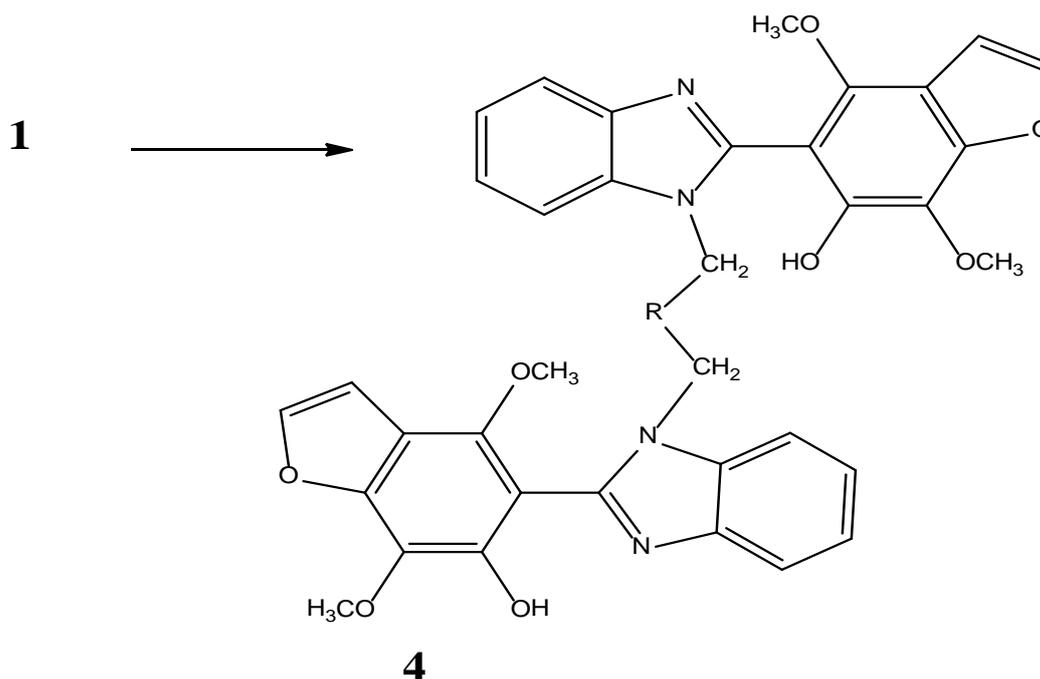
h, R= isopropyl

i, R= piperidyl

j, R= 2-carboxyphenyl

k, R= 2-hydroxyphenyl

Scheme 2:



4a, R=1,4-diaminophenyl

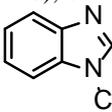
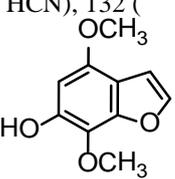
b, R=piperazinyl

Experimental

NMR of the synthesized compounds were not performed because the great difficulty in solubility

Table1: Spectral data of compounds 2, 3a-k, 4a,b

No.	IR (KBr) $\nu_{\max}/\text{cm}^{-1}$	MS, m/z (%)	m.p.	Color
2	3389 (br. NH & OH), 2939, 2842 (aliph. C-H stretching), 1660 (C=N), 1623 (C=C), 1592 (C=C)	328 (M+H ₂ O), 310 (M), 309 (M-1), 280 (M-HCHO), 245 (base peak), (M-HCHO-H ₂ O-NH ₃)	240 ^o C	Black
3a	3614 (OH), 3263(OH & NH), 2943, 2839 (aliph. C-H), 1620(C=N), 1508, 1481 (C=C), 920 (C-Cl), 806 (C-NO ₂)	495 (M), 497 (M+2), 456 (M-NO+1), 172 (C ₆ H ₄ ClN ₂ O ₂), 155 (171-16-1)	>350 ^o C	Dark brown
3b		413 (M-HCl-H), 393 (413-18), 367, 389, 319, 281, 42 (base peak)	>350 ^o C	Dark brown
3c	3421 (OH & NH), 2921 (C-H stretching), 1640 (C=N), 1618 (C=C), 1482.	439 (M), 440 (M+1), 438, 437, 383, 355 (base peak), 239, 211	>350 ^o C	Dark brown
3d	3407 (OH & NH), 2939, 2840 (C-H stretching), 1650 (C=N), 1611 (C=C), 1482, 678 (C-Cl)	483 (M), 442 (483-Cl), 417 (447-HCHO), 389 (417-CO), 63 (base peak)	>350 ^o C	Dark brown
3e	3409 (NH & OH), 2939, 2841 (aliph. C-H	367 (M-CH ₃ +1), 339(M-n-propyl+1), 294 (M-n-propyl-NH-CH ₂ -17+1), 276,	>350 ^o C	Dark brown

	stretching), 1655 (C=N), 1617 (C=C)	265		
3f	3418 (NH & OH), 2938 (aliph. C-H stretching), 1660 (C=N), 1616,1482 (C=C)	369 (M+2), 368(M+1), 367(M), 353(M-CH ₂), 119 (base peak).	>350 °C	Dark brown
3g	3407, 3224 (NH & OH), 2939 (aliph. C-H stretching), 1655 (C=N), 1614, 1513 (C=C).	368 (M-CH ₃ +2H ⁺), 277, 204, 190, 57 (base peak)	>350 °C	Reddish brown
3h	3412, (NH & OH), 2971, 2937, 2846 (aliph. C-H stretching), 1640 (C=N), 1616, 1500 (C=C).	382 (M+1), 381 (M), 369 (M-CH ₂ +1), 286(base peak), 204, 189, 161, 119...	>350 °C	Dark brown
3i		407 (M), 311 (M-piperidyl-2), 78, 63(base peak).	>350 °C	Dark brown
3j	3402, (OH, NH ₃ ⁺), 2937, (aliph. C-H stretching), 1700 (shoulder C=O), 1616 (C=N), 1612 (COO ⁻), 1481(C=C).	460 (M+1), 459 (M), 414, 397, 249, 121 (base peak)	>350 °C	Dark brown
3k	3614 (OH), 3139 (NH & OH), 2958, 2750, 2592 (aliph. C-H stretching), 1616 (C=N) 1523, 1446 (C=C).	401 (M), 254, 244, 230 (base peak), 122,109.....	>350 °C	Dark brown
4a	3409 (NH & OH), 2943, 2839 (aliph. C-H stretching), 1660 (C=N), 1620, 1512 (C=C)	752 (M), 108 (base peak)	>350 °C	Dark brown
4b	3394 (NH & OH), 2947 (aliph. C-H stretching), 1635 (C=N), 1508 (C=C)	730 (M+1), 729 (M), 672 (M-2CHO), 654 (M-2CHO-18), 627 (M-2CHO-18-  CH ₂ +1), 193 (), 112 (base peak), 84 (piperazine)	>350 °C	Dark brown

Biological Activity

The antimicrobial activity of material against *E. coli* was investigated as a pathogenic bacteria. Bacteriological Tests were performed in lysogeny broth (LB) (Bertani, G. (2004)) medium on solid agar plates and in liquid systems. The materials were shown to be an effective bactericide or bacteriostatic. The results show effect of 32 materials on growth of pathogenic bacteria by Growth measurements methods

Material and Methods

Bactericidal test

Escherichia coli strain was grown on (lb) broth for 24 hour were cultured on 20 ml on 50 ml falcon of lb broth by add 10 microliter from culture and supplemented with 0.3 gram from tested materials . material- free lb broth cultured under the same conditions were used as a control > the falcons was incubation in shaker incubator for 24 hour in 30 ° at 150 rpm >>>the table show O.D (optical density) for control and other treatments at wave 600

Counting plates

The treatments was cultured after 24 hour incubation on lb agar plates by pour plate method to counting survival bacteria to indicator the effect of material on bacteria. The culture was diluted to 10⁻⁷ and then put 50 microliter on plate and then incubated for 24 hour and 30 °

Results and Discussion

The table show the Comparison between the control and other treatment on optical dencity and Counting plates and it showe the materials 3b was Active growth and materials 3a and 4b was made Percentage of death 66 % ,and material 3f was made percentage of death 33%.

Table 2: O.D for culture supplemented with 10 materials .and show plat counting result

a Material	O.D	Percentage of death %	Plate count
Cont	0.6	0	1205 cell
2	0.5	16	1020
3a	0.2	66	400
3b	0.6	X	1170
3e	0.5	16	1172
3f	0.4	33	900
3g	0.5	16	1100
3i	0.6	0	1205
3j	0.5	16	1006
4a	0.5	16	1010
4b	0.2	66	204

Acknowledgement

The authors would like to thanks Taibah University KSA ,Scientific Researches Deanship for Financial support under research number 431/430 , thanks for faculty of science, department of chemistry. And Dr. Sahar A. Zaki Head, Enverionmental Biotechnology Department Gentetic Engineering and Biotechnology Research Institute City of Scientific Research and Technological Applications.

References:

- 1-Shadia A Galal, Ahmed S. Abdelsamie, Mireya L. Rodriguez, Sean M. Kerwin, Hoda I. El Diwani. European Journal Chemistry, 1 (2) (2010) 67-72.
- 2- Abdel-Mohsen, H. T., Ragab, F.A.F., Ramla, M. M., El Diwani, H. I., European Journal of Medicinal Chemistry, 4(6), 2010, 2336-2344.
- 3- Shadia A. Galal, Khaled H. Hegab, Ahmed M. Hashem, Asmaa A. Magd-El-Din, Nabil S. Youssef and Hoda I. El Diwani . European Journal of Medicinal Chemistry, 45(7), 2010, 3035-3046.
- 4- Katritzky, A. R.; Dobchev, D. A.; Fara, D. C. and Karelson, M. *Bioorg. Med. Chem.*, **13** (2005) 6598-6608.
- 5- S. A. Galal, K. H. Hegab , A. S Kassab, M. L. Rodriguez, S. M. Kerwin, A. A. El-Khamry and H.I. El Diwani. European Journal of Medicinal Chemistry. 44(4), 2009, 1500-1508.
- 6- El-Naem, Sh. I.; El-Nzhawy, A. O.; El-Diwani, H. I. and Abdel Hamid, O. *Arch. Pharm. Pharm. Med. Chem.*, **1** (2003) 7-17.
- 7-Jordan, A. D.; Vaidya A. H.; Rosenthal, D. I.; Dubinsky, B.; Kordik, C. P.; Sanfilippo, P. J.; Wu, W. N.; Reitz,A. B. *Bioorg. Med. Chem. Lett.* **2002**, 12, 2381.
- 8-Orjales, A.; Mosquera, R.; Labeaga, L.; Rodes, R. *J. Med. Chem.* **1997**, 40, 586.

- 9-Evers, D. L.; Komazin, G. ; Shin D.; Hwang, D. D.; Townsend, L. B.; Drach, J. C. *Antiviral Res.* **2002**, *56*, 61.
- 10-Corbett, A. H.; Guerry, P.; Pflieger, P. and Osheroff, N. *Antimicrob. Agents Chemother.*, **37** (1993) 2599-2605.
- 11- Velaparthi, U.; Liu, P.; Balasubramanian, B.; Carboni, J.; Attar, R.; Gottardis, M.; Li, A.; Greer, A.; Zoeckler, M.; Wittman, M. D. and Vyas, D. *Bioorg. Med. Chem. Lett.*, **17** (2007) 3072-3076.
- 12- Alper, S.; Arpacı, Ö. T.; Aki, E. Ş. and Yalcin, I. *Farmaco*, **58** (2003) 497-507.
- 13- Sun, Q.; Gatto, B.; Yu, C.; Liu, A.; Liu, L. F. and LaVoie, E. J. *J. Med. Chem.*, **38** (1995) 3638-3644.
- 14- Kiyoshi O, *Japan Kokai*, 10, 024; *Chem Abstr.* 1978, **88**, 3318g.
- 15- Werner D, Walfgang B, Ingebarg H, Hans S and Willum B, *Ger Offen.* 2612315, *Chem. Abstr.*, 1978., **88**, ,79698.
- 16- Hans W, Gnenter D and Paul H, Ger. (East) 127636. *Chem Abstr.* **88**, 136680 w.
- 17- Pegfei X, Ximping Y, Shaozn W and Ziyl Z, *Indian J Chem.*, 1998, **127**, 3713.
- 18- Neustadt B R and Smith E M, *PCT Int. Appl.* W00032578, (*Chem Abstr*, 2003, **133**, 309589).
- 19- Puratchikody A, Sivagothi V, Zaswanth A, Rutkmani and K, Nalul M, *Indian J Heterocycl Chem.*, 2001, 241.
- 20- Pandey V K and Raj N, *Curr Sci.*, 1980, **53**, 256.
- 21- Sengupta A K, Gupta A A and Chandra V, *Indian Drug*, 1985, **18**, 233.
- 22- Evans D, Hicks T A, Willamsan W R N, Aavidisian W, Meacock S C R and Kitchen E A, *Chem Abstr.*, 1986, **125**, 2476894.
- 23- Manna K and Agarwal Y K, *Bioorg Med Chem Lett.*, 2009, **19(10)**, 2688- 2692.
- 24- Alejandro Urzúa , Marcos C. Rezende, Carolina Mascayano and Loretta Vásquez, *Molecules*, 2008, **13(10)**, 882-891.
- 25- Aruna Kumar D B, Prakash G K, Kumarasamy M N, Nandheswarappa B P, Sheringara B S and Mahadevan K M, *Indian J Chem.*, 2007, **46B**, 336- 343.
- 26- Nalan Gundogdu K, Kadriya B, Yagmur T, Umit U and Seref D, *Eur J Med Chem.*, 2006, **41**, 6516
- 27- Hayakawa I, Shioya R, Agatsuma T, Furukawa H, Naruto S and Sugano Y, *Bioorg Med Chem Lett.*, 2004, **14**, 455- 458.
- 28- Galal S A, Abd El All, Abdullah M M and EL- Diwani HI, *Bioorg Med Chem Lett.*, 2009, **19(9)**, 2420- 2428.
- 29- Lourdes Santana, Marta Teijeira, Eugenio Uriarte, Carmen Teran, Belen Liñares, Rosa Villar, Reyes Laguna and Ernesto Cano, *Eur J Pharm Sci.*, 1999, **7**, 161-166
- 30- Philips, A.M., *J. Chem. Soc.*, (1931) 1143.

PHYSICOCHEMICAL SENSORS FOR THE DETERMINATION OF SILVER

*Wagiha , H. Mahmoud
Fatma Abdul Aleem EL-hag
Reem A.K. AL –Harbi*

University, Faculty of Applied Science, Al-Madinah Al-Munawarah, KSA

Abstract:

A novel silver sensor based on 3,7- dimethoxy -6-hydroxy benzofuran-5- carboxylic acid as an ionophore for ion-selective polymeric membrane electrode toward Ag^{+1} ions was examined. The sensor showed a near Nernstian response for Ag^{+1} ions over a concentration range 1×10^{-6} to 1.0×10^{-2} mol L^{-1} with a slope of 61 ± 0.4 mV per concentration decade in an acidic solution (pH 5). The limit of detection was 0.47 mg mL^{-1} . It had a response time of < 20 s and can be used for at least 3 months without any divergence in potentials. The influence of plasticizer as well as the amount of lipophilic anionic site additive in the sensing membrane was discussed. It was shown that membrane electrodes formulated with the ionophore and appropriate anionic additive exhibited enhanced potentiometric response toward Ag^{+1} over all other cations tested. Validation of the assay method revealed good performance characteristics, including long life span, good selectivity for Ag^{+1} ions over a wide variety of other metal ions, long term response stability, and high reproducibility. The sensors were used for direct measurement of silver content in different rocks collected from different geological zones. The results agreed fairly well with data obtained using atomic absorption spectrometry.

Key Words: Psysichemical Sensors, Silver Determination

Introduction

'Heavy metals' is a general collective term applying to the group of metals and metalloids with an atomic density greater than 6 g /cm^3 . Although it is only a loosely defined term it is widely recognized and usually applied to the elements such as silver, nickel, iron, lead, copper and zinc which are commonly associated with pollution and toxicity problems. Unlike most organic pollutants, heavy metals occur naturally in rock-forming and ore minerals and so a range of normal background concentrations is associated with each of these elements in soils, sediments, waters and living organisms. On an annual basis, significant quantities of various heavy metals are produced from the mining of their respective ores [1]. Industrial uses of metals and other domestic processes (e.g. burning of fossil fuels, incineration of wastes, automobile exhausts, smelting processes and the use of sewage sludge as landfill material and fertilizer) have introduced substantial amounts of potentially toxic heavy metals into the atmosphere and into the aquatic and terrestrial environments. Electrochemical sensors are simple, sensitive and selective devices for real-time monitoring of analytes of interest when properly designed. The development of chemical sensors for determination of heavy metals in the environment is of great importance from the ecotoxicological point of view [2]. For selective recognition, an efficient molecular receptor which has high potential as a complexing agent for a target heavy metal is required as sensing element [3]. Nowadays many compounds have been designed and synthesized to form remarkably stable and selective complexes with transition metal ions such as Schiffbases [4], podands [5], cyclams [6] and calixarenes [7]. Among them crown ethers containing nitrogen and sulfur donor atoms (i.e. azathiacrown ethers) are of special interest as they exhibit extremely high affinities toward heavy metal ions such as Ag^{+} , Hg^{2+} and Pb^{2+} [8–10]. Their selectivities and coordination abilities with transition metal ions depend on ring cavity size, ligand rigidity, the nature of donor atoms and their disposition [11]. In recent years, many azathiacrown ethers have been synthesized and successfully utilized for highly selective binding with

heavy metals in diversal processes such as ion transports through artificial and natural membranes, liquid-liquid phase-transfer reactions and isotope separations [12–14]. Ion-selective electrodes (ISEs) are most frequently used potentiometric sensors for heavy metals due to high selectivity, good precision, low cost and simplicity. In recent years, the quest for improved lower detection limits of polymeric membrane ISEs has reinvigorated the search for better molecular receptors [3]. Several thiacrown ethers based on calixarenes have been explored as ionophores for heavy metals with lower detection limits [15,16]. Many crown ethers have been utilized as Ag⁺-ionophores [18,19]. However, most of them suffer from problems of complicated synthesis procedures and poor selectivities. However, so far there has been no report on the characterization of 3,7-dimethoxy -6-hydroxy benzofuran-5- carboxy drazide as ionophores for the polymeric membrane ISEs for silver with improved lower detection limits. In the present work, we have designed potentiometric plasticized poly (vinyl chloride) membrane electrodes using this compounds as Ag⁺-ionophore in terms of their selectivity coefficients and detection limits for Ag⁺. The selectivity towards Ag⁺ is tunable by altering the composition of donor atoms, ring size and ligand geometry [17]. The electrodes based on this ionophore show excellent affinities to Ag⁺. The effects of flexibilities of these ligands on the response of polymeric membrane Ag⁺-ISEs have been studied.

Experimental part

Equipments

All potentiometric measurements were made at 25 ± 1 C with an Orion digital pH /mV using 3,7- dimethoxy -6-hydroxy benzofuran-5- carboxylic acid membrane sensor in conjunction with an Orion double junction Ag/AgCl reference electrode (Model 90-02) filled with 10 % W/W KNO₃ in the outer compartment, Orion combination pH electrode was used for pH adjustment.

Chemicals and Reagents

Ionophore (3,7- dimethoxy -6-hydroxy benzofuran-5- carboxylic acid) was prepared according to the recommended procedure. Analytical reagent grade chemicals , and bidistilled water were used for preparing all aqueous solutions. High molecular weight carboxylated poly vinyl chloride (PVC-COOH), *o*-, nitropheny octyl ether (*o*-NPOE), Dibutyl sebacate (DBS), Dioctyl phthalate (DOP) and (THF) tetrahydrofuran were obtained from Fluka. Stock solutions (10⁻¹ M) of different metal salts (chloride, sulfate, nitrate) were prepared form reagent grade chemicals.

Preparation of 3,7- dimethoxy -6-hydroxy benzofuran-5- carboxylic acid

0.01mmole khellin and 50 ml of 30% of H₂O₂ was mixed and stirred for 30 min. , the resulting compound was completely dissolved. The undissolved residue was Filtered, and the filtrate was acidified by hydrochloric acid to precipitate 3,7- dimethoxy -6-hydroxy benzofuran-5- carboxylic acid and the product was crystallized from ethanol. The structure(Fig. 1) was comfermed by elemental analysis ,I. R spectrophotometry ,HNMRspectra and mass spectra.

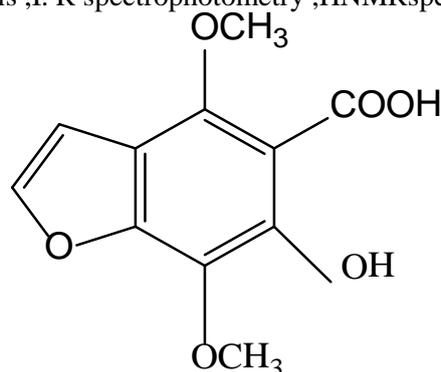


Fig.1 structureormula of 3,7 dimethoxy -6-hydroxy benzofuran-5- carboxylicacid

Construction and characterization of the sensor

PVC-membrane sensor based on 3,7dimethoxy -6-hydroxy benzofuran-5- carboxylic acid as ionophore was prepared as described elsewhere [18] using 3 mg of the ionophore, 130 mg dibutyl sebacate(DBS) as plasticizer, 65 mg PVC as matrix, 1 mg of tetra dedocyl methyl ammonium chloride (TDMAC) as lipophilic anionic site additive and about 3 ml of THF as solvent. The cocktail in a Petri- dish of a diameter 3 cm was mixed well and left to stand overnight at room temperature.

Working sensors were assembled by cutting 5 mm diameter disks of the prepared membrane and mounting these disks into glass electrode bodies (ISE-561, glasblaseri Moller, Zurich). A solution consisting of equal volumes of 10^{-2} M AgNO_3 and KNO_3 is adjusted to pH 5.5, and used as internal filling solution. Before use, the membrane sensors were conditioned overnight in 10^{-2} M AgNO_3 solution. A standard stock solution of 10^{-1} M AgNO_3 was prepared. A series of diluents concentrations prepared with buffer acetate pH = 5 and another series of diluents concentrations prepared by 10^{-3} KNO_3 of pH = 5. The potentials were plotted against concentration for two different buffer solutions for the Ag^+ sensor based on ionophore, to clarify the effect of type of buffer on the sensor.

EMF measurements

All measurements were carried out at 25°C with cells of the following type: $\text{Ag}/\text{AgCl} / \text{KCl}$ (0.1 M) / 0.1 M KNO_3 / sample solution // sensor membrane // internal filling solution / AgCl / Ag . Sensors calibrations were carried out by measuring the potential of 10^{-6} – 10^{-2} M AgNO_3 solutions starting from the low to high concentration. The sensor in conjunction with a double junction Ag/AgCl reference electrode was immersed in 50 ml beaker containing 1 ml of 10^{-1} M AgNO_3 and 9 ml 10^{-3} KNO_3 of pH = 5 and more diluted solution from 10^{-6} – 10^{-2} M were prepared by the same way. The potential change for each concentration was recorded. A calibration curves were constructed by plotting the potential change against logarithm $\text{Ag}(\text{I})$ concentration. The lower detection limits was taken at the point of intersection of the extrapolated linear segments of the silver (I) calibration curve. Sensor life span was examined by repeated monitoring of the slope of silver calibration curve.

The effect of membrane composition

Studies incorporating varying weight of ionophore (20 mg, 15 mg, 10 mg and 5 mg), 190 mg PVC and 350 mg of DBS with 5 ml THF in 50 mm diameter ground glass. The potential of AgNO_3 solutions starting from the low to the high concentrations were carried and the potentials were plotted as a function of logarithm silver ion concentrations. Two different types of PVC, PVC-COOH were used in the preparation of two different membranes with the same ionophore and (DBS) as plasticizer. The membranes were prepared as described above and the slope, linear range and the lower limit of detection of each membrane were determined.

Effect of pH and dynamic response time

The effect of pH was tested by measuring the potential at pH values ranged 2-10 in $\text{Ag}(\text{I})$ ions solutions of 10^{-4} – 10^{-3} M. The pH was adjusted by NaOH or HNO_3 dilute solutions. The corresponding mV readings were recorded after each addition. The pH dependences on silver sensors were examined by plotting the change of potential against pH values. Dynamic response times of the electrodes were measured in constantly stirred solution of varying $\text{Ag}(\text{I})$ concentrations during 180 sec period for each concentration, and examined by plotting potentials as a function of time.

Selectivity of membrane sensor

Selectivity coefficients $K_{\text{AgNO}_3}^{\text{POT}}$ of AgNO_3 sensor were measured using the separate solutions method [19]. In this method the potentials of the cell containing Ag ionophore sensor in conjunction with double junction Ag/AgCl reference electrode were measured independently for each of two separate solutions for all concentrations. Then, the activates that correspond to the same sensor potential were used to determine $K_{\text{AgNO}_3}^{\text{POT}}$ and can be calculate selectivity coefficient using equation.

$$-\text{Log } K^{\text{pot}} = E_{\text{M}} - E_{\text{Ag}} / s$$

Where E_{Ag} and E_{M} are the response potentials of the sensor for silver nitrate ion and interference M, respectively at 10^{-3} M and S is the sensor slope ($\text{mV}/\text{decade}^{-1}$) potentials are recorded for concentration from low to high analytical application.

Analytical Application

Direct Potentiometric determination of silvers

These sensors in conjunction with a double junction Ag/AgCl reference electrode were immersed in a fixed volume of AgNO_3 with fixed concentration and 0.1 M of NaCl were successively added and the potential change for each addition was recorded. This process repeated three times with

different concentration of Ag^+ in a fixed 10 ml volume. Titrations were plotted for three sensors based on ionophore and the equivalent volumes were determined.

Results and Discussion

Sensor characteristics

Direct potentiometry with chemical sensors provides a selective, precise and fast method for monitoring of various metal ions including silver at low concentrations. Silver - ISEs have been intensively studied and extensively used for the formation of coordination networks based on the binding of silver. In preliminary experiments, the new sensitized ligand combined with various plasticizers used to prepare PVC membrane selective for Ag ions. The performance characteristics were evaluated under many different variables such as type of PVC, plasticizer (solvent mediator), the influence of pH, effect of internal solution concentration of the electrode Potentiometric response of sensor based neutral ionophore is greatly influenced by the polarity of the membrane medium, which is in turn defined by the dielectric constants of the major membrane components. Silver PVC-COOH matrix membrane incorporating three different plasticizers having dielectric constants 4-24 namely DBC, DOP and NPOE. The results indicate that, Performance characteristics of membrane incorporating PVC-carboxylated with (DBS) plasticizer and 10^{-2} M internal solution showed high sensitivity, wide linear response range and near Nernstian slope at pH 5 as tabulated in table1. The results showed that the sensor displays a linear response for Ag^+ over the concentration range 1×10^{-2} - 1×10^{-6} M with a detection limit 5.0×10^{-6} M and a calibration slope of 61 ± 0.2 mV/decade, the slope of sensor is near to the theoretical Nernstian value. In the light of the soft heavy metal ions displaying great affinity for soft coordination centers like nitrogen, sulfur or oxygen. We try to explain the properties of the electrodes in terms of coordinating ability of the O atoms in ionophore. The coordination interaction between electron donors and acceptors cause the electrode to respond to metal ion.

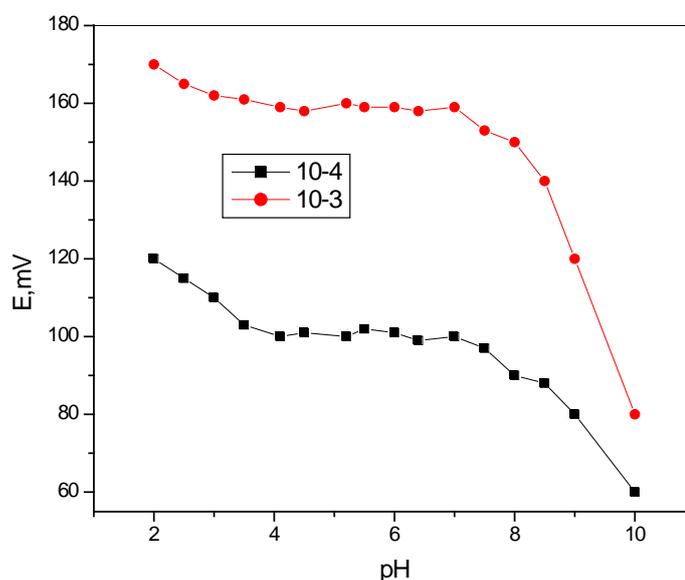


Fig 1. Effect of the pH on the potential of the proposed sensor

Parameter	Silver sensor
Slope, (mV/decade)	61.1
Working concentration range (M)	10^{-6} - 10^{-2}
Lower limit of detection, (M)	5×10^{-6}
pH range	4-7
Correlation coefficient, (r)	0.99994
Standard deviation	0.94868

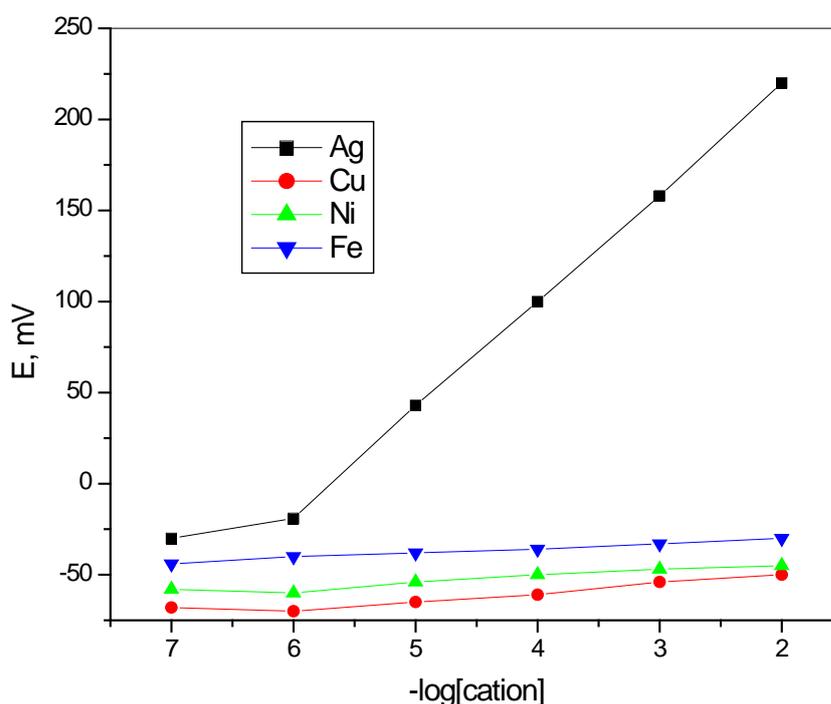
Table 1. Performance characteristics of silver sensor

Selectivity of Sensor

The most important characteristic of a membrane sensor is its response for the primary ion in the presence of other cations. This is measured in terms of the potentiometric selectivity coefficients (Robert et al., 2000) with 10^{-3} M concentration level of alkali, alkaline earth and transition metal ions. The results obtained (Table 2.9) (Fig. 2.17) (Fig. 2.18) (Fig. 2.19) For the ionophore (I) (II) (III) respectively.

Potentiometric selectivity coefficients of Ag-sensor (II) membrane based sensor for some common cations were evaluated using the separate solutions method (SSM), (Ma and Hassan, 1982). It can be seen from Table (2.9) that (I) (II) (III) sensors exhibits high selectivity towards Ag^+ ions over many of the ions under investigation. The selectivity was measured for 10^{-3} M concentration of the analyte and the interferents.

Fig. 2. Comparison of the response characteristics of silver sensor with silver and other cations



Conclusion

Silver sensor based on 3,7-dimethoxy-6-hydroxy benzofuran-5-carboxylic acid as an ionophore was constructed. The sensor showed response for Ag^+ ions over a concentration range 1×10^{-6} to 1.0×10^{-2} mol L⁻¹ with a slope of 61 ± 0.4 mV per concentration decade in an acidic solution (pH 5). The limit of detection was 0.47 mg mL⁻¹. It had a response time of < 20 s and can be used for at least 3 months without any divergence in potentials. The influence of plasticizer in sensing membrane was discussed. It was shown that membrane electrodes formulated with the ionophore and appropriate anionic additive exhibited enhanced potentiometric response toward Ag^+ over all other cations tested. Validation of the assay method revealed good performance characteristics, including long life span, good selectivity for Ag^+ ions over a wide variety of other metal ions, long term response stability, and high reproducibility. The sensors were used for direct measurement of silver content in different rocks collected from different geological zones. The results agreed fairly well with data obtained using atomic absorption spectrometry.

Acknowledgement

The authors would like to thank Taibah University KSA, Scientific Researches Deanship for Financial support under research number 431/430, also thanks for faculty of science, department of chemistry.

References:

- [1] US Geological Survey, 2004. Minerals commodity summaries. Available from: <<http://minerals.usgs.gov/minerals/pubs/commodity>>.
- [2] Notten, M.J.M. Oosthoek, A.J.P. Rozema, and Aerts, J. R. *Ecotoxicology* 15 (2006) 295.
- [3] Bakker, E. and Pretsch, E. *Angew. Chem. Int. Ed.* 46 (2007) 5660.
- [4] Borisova, N.E. Reshetova, M.D. and Ustynyuk, Y.A. *Chem. Rev.* 107 (2007) 46.
- [5] Bünzli, J.C. Piguet, G. C. *Chem. Rev.* 102 (2002) 1897.
- [7] Nolan, E.M. and Lippard, S.J. *Chem. Rev.* 108 (2008) 3443.
- [8] Homden, D.M. Redshaw, C. *Chem. Rev.* 108 (2008) 5086.
- [9] M.B. Gholivand, M. Rahimi-Nasrabadi, M.R. Ganjali, M. Salavati-Niasari, *Talanta* 73 (2007) 553.
- [10] Li, X.G., Ma, X.L. Huang, M.R. *Talanta* 78 (2009) 498.
- [11] Gupta, V.K. Singh, A.K. Khayat, M.A. and Gupta, B. *Anal. Chim. Acta* 590 (2007) 81.
- [12] Fernández-Fernández, M.C. Bastida, R. Macías, A. Pérez-Lourido, P. and Valencia, L. *Inorg. Chim. Acta* 362 (2009) 471.
- [13] Ocak, M. Alp, H. Kantekin, H. Karadeniz, Ocak, H. *J. Incl. Phenom. Macrocycl. Chem.* 60 (2008) 17.
- [14] Cui, J. and Zhang, L. *J. Hazard. Mater.* 158 (2008) 228.
- [15] Domash, V.I. Sharpio, T.P. Zabreiko, S. A. and Sosnovskaya, T.F. *Russ. J. Bioorg. Chem.* 34 (2008) 318.
- [16] Szigeti, Z. Malon, A. Vigassy, T. Csokai, V. Grün, A. Wygladacz, K. Ye, N. Xu, C. Chebny, V.J. Bitter, I. Rathore, R. Bakker, E. and Pretsch, E. *Anal. Chim. Acta* 572 (2006)
- [17] Gupta, V.K. Pal, M.K. and Singh, A.K. *Anal. Chim. Acta* 631 (2009) 161.
- [18] Mahmoud, W. H. *Analytica Chimica Acta* 436 (2001) 199.
- [19] Umezawa, Y. Buhlmann, P. Umezawa, K. Tohoda, K. and Amemiya, S. *Pure Appl. Chem* 72(2000) 1851.

CHEMOINFORMATICS AND ITS APPLICATIONS

Wenli XU

Min LING

Jing HU

Ying HUANG

Jia LI, PhD

Jianhua YAO, PhD, Professor

Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, China

Abstract:

With development of computer science and network, computer and software are getting more and more applications in scientific research, production and service areas. *In silico* is an expression used to mean “performed on computer or via computer simulation”.

Chemoinformatics is one of *in silico* technologies employed in researches relating to Chemistry. It is an application of informatics methods to solve chemical problems concerned with molecular design, synthesis design and structural identification. It is an interdisciplinary combining of Computer science, Mathematics and Chemistry. In principle, three methods are included in Chemoinformatics: data-based, logic-based and principle-based.

These years, Chemoinformatics has been applied not only in chemical research but also in domains relating to Chemistry, i.e., Drug Discovery, Pesticide Design, Environment Protection, Material Design, Traditional Chinese Medicine (TCM), Food Safety and etc., which relate to Chemistry. Chemoinformatics will be applied in more and more domains relating to Chemistry because “Chemistry is (almost) everywhere and in everything”.

Herein, two works: 1. modern mode of chemical research or other works relating to chemical research is proposed, which is combined by thinking, experiment and *in silico*; 2. *in silico* platform for pesticide design will be presented.

Key Words: Chemoinformatics, *In silico*, Chemistry

Introduction

Chemistry is a branch of physical science and focused on studying the composition, properties and behavior of matter. In general, it includes 23 sub disciplines: 1) Acid-base reaction theories; 2) Alchemy; 3) Analytical chemistry; 4) Astrochemistry; 5) Biochemistry; 6) Crystallography; 7) Environmental chemistry; 8) Food chemistry; 9) Geochemistry; 10) Green chemistry; 11) Inorganic chemistry; 12) Materials science; 13) Molecular physics; 14) Nuclear chemistry; 15) Organic chemistry; 16) Photochemistry; 17) Physical chemistry; 18) Radiochemistry; 19) Solid-state chemistry; 20) Stereochemistry; 21) Supramolecular chemistry; 22) Surface science and 23) Theoretical chemistry. [1] And it is (almost) everywhere and in everything [2]. In Chemistry, the three subjects: molecular modeling, synthesis design and structural identification, are considered as three chemical problems.

Chemoinformatics (also termed as cheminformatics and chemical informatics) is an application of informatics methods to solve the chemical problems concerned with molecular design, synthesis design and structural identification [3]. It is an interdisciplinary [4] combining of Computer science, Mathematics and Chemistry. In principle, three methods are included in Chemoinformatics: data-based, logic-based and principle-based. They are used at different stage and different levels in chemical study.

Data-based means that people retrieve data from database. It is usually employed at the beginning of study.

Logic-based means that people retrieve knowledge from normalized data and do predictions based on knowledge. In general, it is employed to decide if compounds should be synthesized or extraction.

Principle-based means that people do the study by quantum chemical calculation. The calculation results are often used in mechanism study.

Herein, two works: 1. modern mode of chemical research or other works relating to chemical research is proposed, which is combined by thinking, experiment and *in silico*; 2. *in silico* platform for pesticide design will be presented.

Modern Mode of Chemical Research

According to the history of Chemistry, ancient Egyptians pioneered the art of synthetic "wet" chemistry up to 4,000 years ago; ancient people did works relating to Chemistry by 1000 BC. [1] In chemical research, computer started to be used in 1940s. At present, most chemists prefer traditional mode: thinking based on experiences and experiments (synthesis, extraction or separation).

The modern mode of chemical research showed in Figure 1, is development of traditional mode and combined by thinking, experiment and *in silico*. *In silico* is an expression used to mean "performed on computer or via computer simulation" [5]. Chemoinformatics is one of *in silico* technologies employed in researches relating to Chemistry.

In chemical researches, molecular design, synthesis design and structural identification, are often considered.

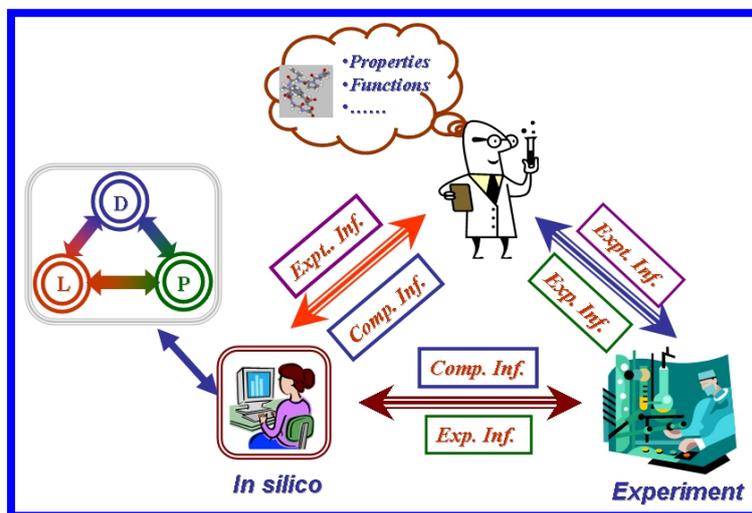


Figure 1. Modern Mode of Chemical Research

Molecular design is the application of all techniques leading to the discovery of new chemical entities with specific properties required for the intended application [6]. In traditional mode, people do molecular design and synthesis design mainly based on experiences and experiments. However, in modern mode, people do molecular design and synthesis design not only based on experiences and experiments, but also on database, knowledge base and quantum chemical calculation. The process of molecular design (showed in Figure 2) can be expressed as: 1. Inspiration (proposing a new chemical entity with specific properties based on experiences); 2. Virtual design (Search information in databases; Prediction of properties based on chemical structures) 3. Synthesis or extraction.

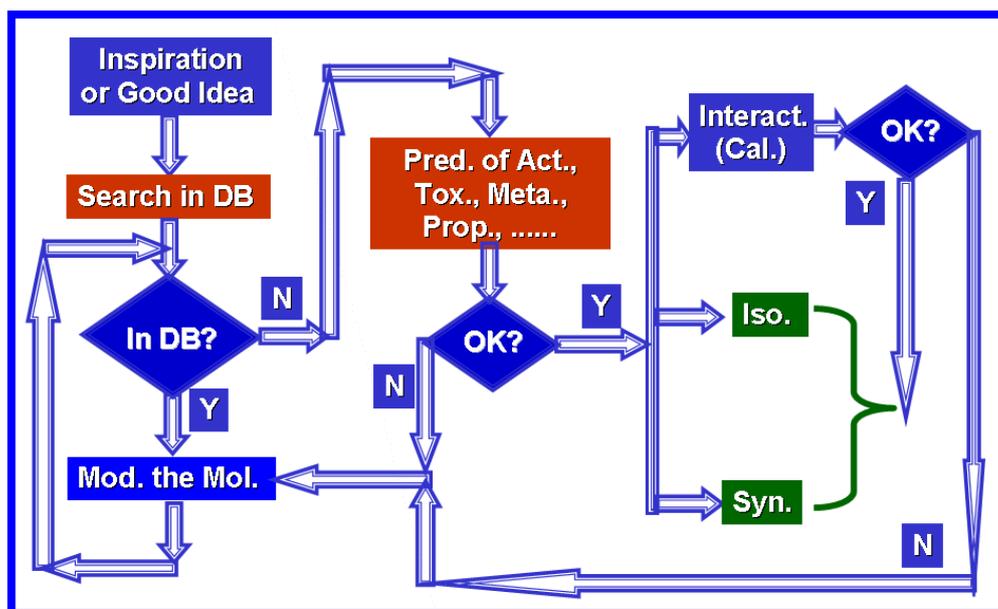


Figure 2. Flowchart of Molecular Design (*in silico*)

The work flow of synthesis is: 1. proposing a product; 2. deducing its candidate reagents based on retrosynthetic analysis [7]; 3. repeating 2. The flow will be stopped when the reagents are available.

Chemoinformatics Platform for Pesticide Design

The pesticide was defined by Food and Agriculture Organization (FAO) in 2002 and its definition is as followed: Pesticide means any substance or mixture of substances intended for preventing, destroying or controlling any pest, including vectors of human or animal disease, unwanted species of plants or animals causing harm during or otherwise interfering with the production, processing, storage, transport or marketing of food, agricultural commodities, wood and wood products or animal feedstuffs, or substances which may be administered to animals for the control of insects, arachnids or other pests in or on their bodies. The term includes substances intended for use as a plant growth regulator, defoliant, desiccant or agent for thinning fruit or preventing the premature fall of fruit, and substances applied to crops either before or after harvest to protect the commodity from deterioration during storage and transport [8]. It means that pesticide is related to Chemistry, environment, human health and etc. In pesticide design, molecular design is one part of most important works. Bioactivity, toxicity and reactivity of compounds which will become pesticide should be paid attention.

In silico platform (showed in Figure 3) for pesticide design has been applied in pesticide design.

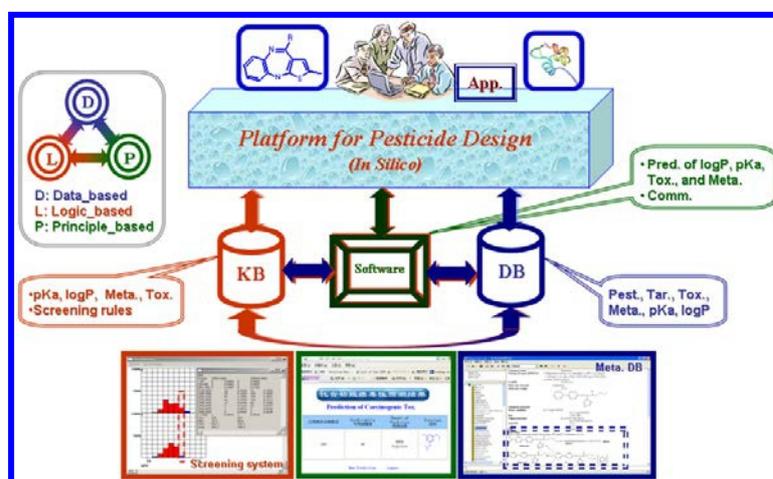


Figure 3. *In silico* Platform for Pesticide Design

This platform is supported by software, database and knowledge base as followed: 1) Prediction System of Carcinogenic Toxicity (CISOC-PSCT) [9, 10] is used to predict carcinogenic toxicity of compounds based chemical structures; 2) Prediction System of Mutagenic Toxicity (CISOC-PSMT) [11, 12] is used to predict mutagenic toxicity of compounds based chemical structures;; 3) Prediction System of Acute Toxicity”(CISOC-PSAT)[13] is used to predict acute toxicity of compounds based chemical structures; 4) Prediction System of logP (CISOC-logP) [14, 15] is used to predict the ratio of concentrations of a compound in a mixture of two immiscible phases at equilibrium according to the chemical structure. The ratio is the measure of hydrophilic ("water-loving") or hydrophobic ("water-fearing") for a chemical substance. 5) Prediction System of pKa”(CISOC-pKa) [16] is used to predict the acidity constant, or acid-ionization constant of a compound. The constant is a quantitative measure of the strength of an acid in solution; 6) Metabolism Knowledge Base Management System (CISOC-MetaKBMS) [17] is used to retrieve metabolism knowledge from normalized metabolism data. And the knowledge in the base can be used to predict metabolites of a chemical substance; 7) Receptor Information Search System (CISOC-WEBPESTRMS) [18] is used to search receptor or ligand information; 8) Web-based Agriculture Chemicals Database (in house) [19] is used to search information about commercial pesticides.

Before synthesis or extraction, scientists propose a compound, then they browse information in databases, do bioactivity, toxicity and reactivity predictions by the prediction systems. Quantum chemical calculation results are used to study chemical behaviors. They will have a list of candidates based on predictions, calculations and experiences.

Conclusion

The two examples mentioned in the front show: 1) the role of *in silico* in chemical research and pesticide design: increasing efficiency of work, decreasing pollution and costs; 2) people should pay attention on study and application of interdisciplinary.

Acknowledgement

This work was financially supported by MOST (973 Program, No: 2010CB126103; Key Program, No: 2011BAE06B05) and National Natural Science Foundation of China (No: 21072216).

References:

- [1] <http://en.wikipedia.org/wiki/Chemistry>.
- [2] Arnon Shani, *Chem. Edu. Inter.*, Vol. 4, No. 1, 2003.
- [3] J., Gasteiger, T., Engel (Eds): *Chemoinformatics: A Textbook*. John Wiley & Sons, 2004.
- [4] <http://en.wikipedia.org/wiki/Interdisciplinary>.
- [5] http://en.wikipedia.org/wiki/In_silico.
- [6] Compendium of Chemical Terminology Gold Book, Version 2.3.2, 2012-08-19, International Union of Pure and Applied Chemistry, pp950.
- [7] E. J., Corey; X-M., Cheng, *The Logic of Chemical Synthesis*. New York: Wiley, 1995.
- [8] International Code of Conduct on the Distribution and Use of Pesticides, Food and Agriculture Organization of the United Nations, Rome, 2002.
- [9] Quan. Liao, Jianhua Yao, et al, SAR and QSAR in Environmental Research, 2004, 15(3), 217-235.
- [10] CISOC-PSCT, Registration Number (China): 022239.
- [11] Quan Liao, Jianhua Yao, et al, *Molecular Diversity*, 2007, 11, 59-72.
- [12] CISOC-PSMT, Registration Number (China): 042160.
- [13] CISOC-PSAT V1.0, Registration Number (China): 0232654.
- [14] Quan Liao, Jianhua Yao, et al, *Molecular Diversity*, 2006, 10, 301-309.
- [15] CISOC-LogP, Registration Number (China): 046394.
- [16] pKa prediction system (CISOC-pKa)V1.0, Registration Number (China): 0154688.
- [17] CISOC-MetaKBMS V1.0, Registration Number (China): 0154319.
- [18] CISOC-WEBPESTRMS V1.0, Registration Number (China): 0411338.
- [19] <http://202.127.145.116/pesticide/index.php>.

PHYTODEPURATION PROCESS FOR THE RECYCLING OF WASTEWATER IN DAIRY

Lydia Ferrara

Chair of Hydrology Chemistry, Department of Pharmacy,
University of Naples "Federico II" Via Domenico Montesano, Naples, Italy

Abstract:

The problem of saving water is very important for the need to rationalize the use of the water resource. We have tested a method founded on phytodepuration for waste water of cheese factory, for to recover and recycle the water used in food industry, a sector characterized by high consumption of water.

The phytodepuration represents an innovative method for the disposal of wastewater. Some plants have a natural capacity to absorb and / or degrade toxic substances and contaminants to the environment, for the presence of a rhizosphere microflora, which is able to metabolize and make them available for their growth.

In the laboratory we developed a little system for phytodepuration using a plastic basin in which were placed some plants of *Cyperus papyrus* immersed in a quantity of wastewater remained constant over time. We performed the analysis before the process and then a distance a few months to verify the variation of the parameters most important pollutants. At the end of our experience, approximately one year, it was possible to draw the following conclusions: the *Cyperus papyrus* proved effective to lessen the concentration of organic substances: the value of COD was lower than the initial wastewater; the concentration of phosphorus, of zinc was decreased and the concentration of chlorides also, which remained constant until the end of the experience. The plants after the stage of adjustment is not showing signs of suffering and therefore can be considered suitable for such use.

Key Words: Wastewaters, phytodepuration, *Cyperus papyrus*

Introduction

The problem of disposing of industrial waste is a subject of constant disputes because of the power polluting the environment. In fact very often you can see streams, rivers and lakes are polluted so as to make it almost impossible to aquatic life and seriously affect the health of the surrounding cultivated areas. The blame for this degradation is commonly attributed to industrial discharging their effluents into waterways regardless of the ecological damage that can provide. It is created in such a way that environmental degradation is depleting the water resources, which until a few years ago seemed unlimited, whether because the surface waters contain higher amounts of pollutants, or because their pollution through the soil, has achieved in many cases even the underground water reserves. Whether the discharges were repaid a little doses in surface water bodies, they could be slowly degraded by aerobic microorganisms present, supported by the presence of dissolved oxygen. Otherwise the pollution load is always very high and the oxygen is insufficient for the oxidation of organic matter for the survival of aquatic species. In this case are involved the anaerobic microorganisms that establish a process of decay with the development of unpleasant odors while simultaneously appears the phenomenon of eutrophication.

The Italian legislation on waste water was low and confused until 1976, when he was adopted and published the law 319/76, known as "Legge Merli". that was subsequently supplemented by the "Technical Standards" and amended by Law 650/79, which provided subsidies for the construction of sewage treatment plants enlarging the powers of municipalities as responsible for most of the tasks performed until that point by the provinces. In 1980 and 1981 were published some ministerial

directives for the regions to which were committed to regulate discharges into sewers and installations for sewage treatment.

The law Merli and all its amendments have been replaced by the D. Lgs. 152/99, currently in force, which sets maximum limits for a number of parameters that characterize a water discharge such as color, the concentration of lead, sulphides, organic solvents and other substances. These values vary depending on the water body to which it is intended the treated water, in particular, are more stringent for discharges to surface water in relation to the wastes and can be modified by the regions according to local need.

The dairy industry in Campania is a highly productive in terms of turnover and places of employment, but the pollution caused by wastewater from the product is superior to twice that of domestic waste. The milk during the process of dairies production suffers a slow pasteurization followed by the addition of rennet, which determines the coagulation of casein. It is observed as a clear separation between the solid mass, used for the subsequent production of mozzarella cheese, and whey, which is further treated to produce ricotta. After removal of the ricotta the whey waste still contains substances of very high nutritional value, and could be an additional source of income considering other possible uses. The use of the whey as food for animals, particularly pigs, is hampered by the limited pig farms in the Campania territory, unable to absorb the large amount of whey produced each day. With regard to human nutrition the whey has very good properties so it is recommended not only to athletes but also to children, women in menopause and all those people who need a further dose of mineral salts. However till now not yet developed a practical and economic method for its use as an integrator, and various applications are proposed for disposal of this waste to particularly high costs

An innovative method for the disposal of this particular wastewater rich in organic matter is the phytodepuration

Materials and methods

AOAC methods for determination of COD, organic matter, chlorides, Conductivity meter Crison micron C.M 2200 (Crison instruments Srl Lainate-MI) pHmeter Φ 50 at 20°C BeckmanCoulter (Cassina De' Pecchi – Milano Italia)

In laboratory was set up a simple system for the purifying using *Cyperus Papyrus*, plant typical of Mediterranean regions, which grows naturally along rivers, whose purifying capacity have been highlighted recently by several authors. The benefits that make this species interesting for the purpose of purification are easy availability, cultivation, breeding, excellent placement and environmental landscape, good adaptation to different climatic conditions.

We analyzed three samples of wastewater coming respectively from the manufacture of mozzarella (I, II, III) whose characteristics are reported in Table 1. The variability of the values can be attributed to the composition of the milk to start., The sample to purify was obtained by mixing equal parts of the three wastewater

In a rectangular basin of 30 x70 cm well exposed to sunlight have been placed on a layer of 10 cm of soil 20 young shrubs and 10 liters of drinking water. The water level was maintained constant throughout the experimental phase, while stirring the water itself was performed every day to avoid the formation on the surface of a layer of fatty substances. The wastewater was analyzed before being subjected to treatment taking into account parameters such as COD, organic matter, chlorides, pH, conductivity as indices of assessment of purifying plant, as well as total phosphorus and mineral elements.

After a week were initially added 10 ml of serum every 20 days after which we took a water sample to evaluate the parameters of pollution. The values of these parameters are given in the following graphs (Figs 1,2,3)

These initial tests have allowed to evaluate the resistance of the plant in the presence of wastewater. afterwards we increased the rate of serum added, 100 ml every 20 days, and took a sample of which we explored the various parameters. The values are given in the following graphs (Figs 4,5,6)

Samples	I	II	III
pH	3.90	3.98	4.01
Conductivity mS/cm	16.70	17.78	17.77
Dry weight g/L	42.68	36.6	68.0
Organic g/L	13.52	12.2	11.2
COD mg/L	8000	7218	6627
Na ⁺ mg/L	370.0	395.0	550.0
K ⁺ mg/L	875	600	962.5
Cl ⁻ mg/L	2481	2127	2481
Total phosphorus mg/L	870	750	800
%Total protein mg/L	12.37	10.42	10.93
Lactose mg/L	4.6	4.3	4.4

TAB 1 Analysis of wastes

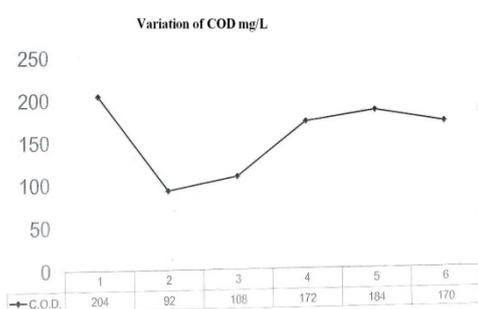


Fig. 1 COD variations

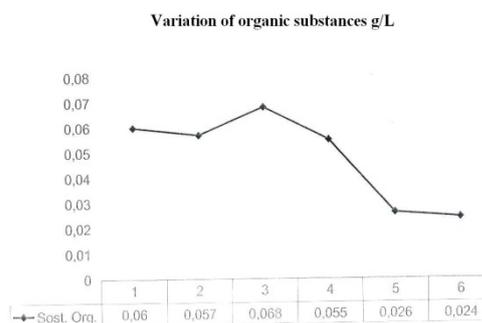


Fig.2 Organic substances variations

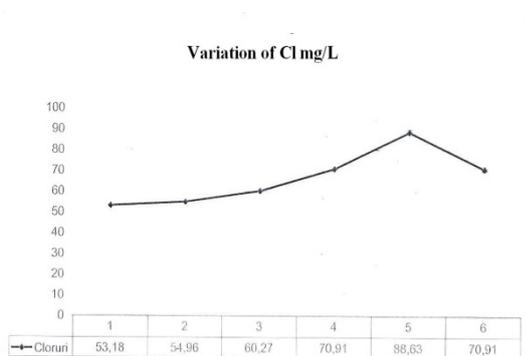


Fig 3 Cl variations

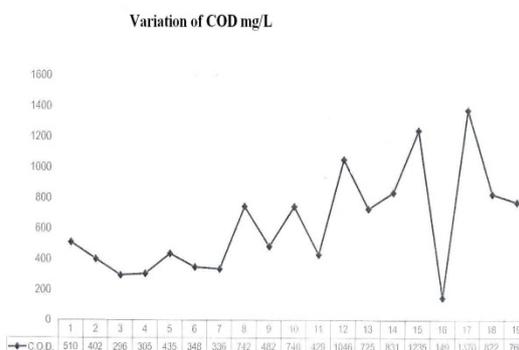


Fig 4 COD variations

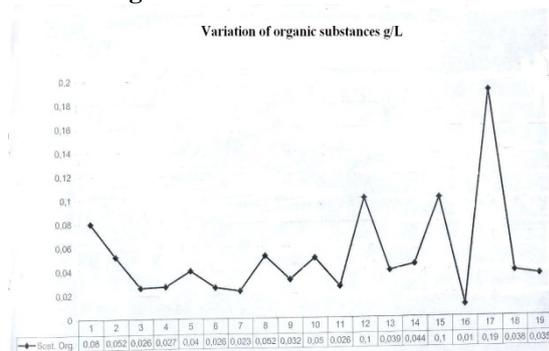


Fig 5 Organic substances variations

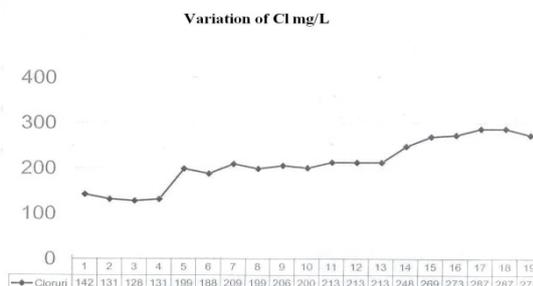


Fig 6 Cl variations

Conclusion

The techniques of phytodepuration can be successfully used for various applications, as installations for purification of waste water, or industrial wastewater depuration, o the rehabilitation

of contaminated sites: their use in any case requires a preliminary assessment to check the adequacy of the chosen system

For each class of contaminants is necessary to discriminate which of the phytodepuration techniques currently known is best suited..

We can be choose between phytoextraction, based on absorption of contaminants radical; phytotransformation, in which contaminants are degraded by the metabolism of the plant; phyto-stimulation or plant-assisted bioremediation based on the stimulation of biodegradation by microbial activity in the plant root zone; phytostabilization, using plants to reduce the mobility and migration of contaminants in soil.

The choice of plants to decontaminate soils and waters, based on the natural capacity of plants to absorb, accumulate and /or degrade, due largely to stimulation of the rhizosphere microflora, molecules in the environment in which they live. The results of research and field testing have demonstrated the applicability of such systems to a broad group of contaminants, including many metals, radionuclides and organic solvents such as chlorinated, pesticides, insecticides and explosives.

The plant species used are usually aquatic plants or highly hydrophilic since those systems are used to purify water. The process occurs through the cooperative growth of macrophytes and micro-organisms associated with them. The plants absorb the nutrients (mainly inorganic sals) present in water to be purified through the roots ; the radical development of the plant species used acts as a coupling for micro-organisms, whose activity is favored by the release of atmospheric oxygen that, absorbed by equipment of the plants is transferred to the roots and released into the surrounding environment. In practice it has a wet ecosystem in which various components plants, microorganisms, soil, solar radiation, contribute to the removal of pollutants

On the process of plant prepared in our laboratory we can make some considerations. After an initial phase during which the plant selected has shown the difficulty of adaptation, as shown in the graphs by a descent of the initial values of the parameters considered, we could record a good performance with the advance of time.

In particular, the activities of the plant compared to the COD and organic substances is very high, thanks to the constant oxygenation provided by the continuous movement in the basin

For the chlorides, however the activity of plant was initially reduced, until to reach an almost constant trend.

Note that there are no such phenomena of eutrophication due to the ability by this plant to absorb phosphorus, which is responsible with the nitrogen of this phenomenon.

References:

Official Methods of Analysis of A.O.A.C International 1990

Forgione P., Ferrara L., Schettino O. The use of *Eichornia crassipes* for the improvement of some characteristics of water *Bollettino della Società italiana di biologia sperimentale*. 1982; 58(5):241-5.

Ferrara L., Forgiane P, Schettino O... The use of aquatic plants in the correction water pollution from chemical agents, absorption of zirconio and cadmium . *Bollettino della Società Italiana di Biologia Sperimentale* . LXI (9) 1985 ,1343-8

Ferrara L., Forgione P:, Schettino O., Rullo V. Purification of waste water from olive-oil pressing plants using hydrophytes. Preliminary study on the use of *Eichornia crassipes* *Bollettino della Società italiana di biologia sperimentale*. 05/1989; 65(4):303-9.

Ribeyre F., Boudou A.,. "Experimental study of inorganic and methylmercury bioaccumulation by four species of freshwater and sediment contamination sources". *Ecotoxicology and environmental safety*, 28, 270-286, 1994

Ayaz-Akca. "Treatment of wastewater by natural systems". *Environ Int.*, 26 (3): 189-195, 2001

ZINC DIPEPTIDE COMPLEX ([Zn(II)-GLY-TYR]⁺)-NINHYDRIN REACTION IN PRESENCE OF GEMINI SURFACTANTS: A KINETIC STUDY

Mohd. Akram, PhD

Dileep Kumar, M.Sc.

Kabir-ud-Din, PhD

Department of Chemistry, Aligarh Muslim University, Aligarh, India

Abstract

Kinetics of the title reaction has been carried out by studying the reaction spectrophotometrically at 80 °C and pH 5.0 in presence of cationic gemini (alkanediyl- α , ω -bis(dimethylhexadecylammonium bromide)) surfactants. The reaction followed first-order kinetics in [Zn(II)-Gly-Tyr]_T⁺ complex and fractional-order kinetics in [ninhydrin]_T. Typical rate constant (k_{ψ}) increase and leveling-off regions (like conventional monomeric surfactants, e.g., CTAB) are found with gemini surfactants. Additionally, the gemini surfactants also produced a third region of increasing k_{ψ} at higher concentrations. The data are interpreted in terms of *pseudo*-phase model (assuming the association/incorporation of both the reactants at the micellar surface).

Key Words: Micellar catalysis, Zinc-glycyltyrosine complex, Cationic surfactants, Ninhydrin

Introduction

Among natural forces, the hydrophobic-lipophilic effect is one of the most important and necessary forces for the formation of ordered assemblies of amphiphilic molecules, e.g., association colloids, vesicles, biological membranes, etc. [1,2]. Surfactant molecules self associate to form micelles when concentration in a solution reaches a threshold value known as the critical micelle concentration (cmc).

The dimeric (dicationic) gemini surfactants are attracting a lot of interest in several fields (e.g., academic and industrial worlds). They are made up of two amphiphilic moieties connected by a spacer group. Due to their unique properties, they differ from their amphiphilic analogs containing only one head group and one hydrocarbon chain in characteristic features, such as markedly low cmc, high surface activity, low Krafft temperature, unusual rheological properties, multifarious aggregate structures, better wetting ability, and so on.

Protein is an important chemical substance in our life and one of the main targets of all medicines in organism. Confinement of peptides at a surface can induce structure formations and the use of peptidic lipids has been proposed in such a context, aiming at induced structure and functionality of such assemblies that may not be present or may organize differently in free solution [3]. Peptides have been extensively used as a protein model compound recently in the studies of the thermodynamic properties of protein because they contain more complex structures and more components of protein than amino acids. Amino acid side chains of proteins also play essential roles in molecular recognition, information transfer and catalysis.

Ninhydrin is a well known fingerprint-developing agent with interesting forensic, biochemical, pharmacological, and biomedical properties, is extensively used for the identification/detection of amino acids, peptides, amines, and amino sugars in analytical chemistry [4]. Metal-amino acid complexes have been studied in detail from the view points of (1) evaluating their binding constants and speciation, (2) their role in colorimetric estimation, (3) models for metalloproteins, and (4) increasing stability and sensitivity of detection of amino acids, in general, and latent finger prints, in particular [5]. The study of the condensation reactions of the dipeptide-metal

complexes with ninhydrin in the presence of micelles may be a better model than studies in water from which to draw conclusions concerning the mechanism of the development of latent fingerprints by ninhydrin. The mechanism of action of an enzyme during the development of latent fingerprints was not ascertained. One possibility is that the enzyme produces amino acids by hydrolysis of proteins present in the fingerprint [6].

The investigation, therefore, concerns the reaction between $[\text{Zn(II)-Gly-Tyr}]^+$ complex and ninhydrin in gemini micellar media.

Experimental

N, N-dimethylhexadecylamine (Fluka, >95.0%), 1,6-dibromohexane (Fluka, >97.0%), 1,5-dibromopentane (Fluka, >98.0%), 1,4-dibromobutane (Fluka, >98.0%), ethyl acetate (HPLC and spectroscopy grade, 99.0%) and ethanol absolute (Merck, 99.8%) Zinc sulfate heptahydrate (Merck, 99.0%), Gly-Tyr (SRL, 99.0%), ninhydrin (Merck, 99.0%), acetic acid (Merck, 99.0%), sodium acetate (Merck, 99.0%), were used as received. Distilled and deionized water was used throughout. The acetate buffer of pH 5.0 was prepared by mixing 30 cm³ of 0.20 mol dm⁻³ acetic acid and 70 cm³ of 0.20 mol dm⁻³ sodium acetate [7]. This buffer solution was used as a solvent for preparing all stock solutions. A LI-digital ELICO 122 pH meter in conjunction with a combined electrode was used for pH measurements.

The dimeric gemini surfactants (alkanediyl- α,ω -bis(dimethylhexadecylammonium bromide)) was synthesized by refluxing the corresponding α, ω -dibromoalkanes ($s = 4, 5, 6$) with *N, N*-dimethylhexadecylamine (molar ratio 1:2.1) in dry ethanol with continuous stirring at 80 °C for 48h. The progress of the reaction was monitored using the TLC technique. Finally, the solvent was removed under vacuum from the reaction mixture and the solid thus obtained was recrystallized several times from ethyl acetate to obtain the compound in a pure form. After recrystallizations, the surfactant was characterized by ¹H NMR. All the values obtained were satisfactory, which indicate that the surfactants are well purified [8].

Solutions of the reactants (1:1 molar) were taken in a graduated standard flask, boiled for 2 min, and heated in a controlled manner at 90 °C for 2h. After completion of the reaction, the flask was brought to room temperature and loss in volume, if any, was maintained by the buffer (pH = 5.0). As a results, a colored complex was formed, which was then stored in dark. This Zn(II)-glycyltyrosine complex was used for further kinetic experiments.

All the UV-visible spectra (Fig. 1) of the solution were recorded after the completion of the reaction between $[\text{Zn(II)-Gly-Tyr}]^+$ (3.0×10^{-4} mol dm⁻³) and ninhydrin (6.0×10^{-3} mol dm⁻³) in aqueous and gemini micellar media. The absorbance increases with increase in surfactant concentration. No shift in λ_{max} (400 nm) was detected in presence of micelles. The results show that the reaction product is same in presence of micelles as that of aqueous solution. In order to determine the composition of the reaction product formed, Job's method of continuous variations was employed in the presence of gemini micelles. It is found that one mole of ninhydrin reacts with one mole of $[\text{Zn(II)-Gly-Tyr}]^+$ complex to give the product.

A 250 ml, three-necked, round-bottomed flask equipped with a double surface water condenser to prevent evaporation and N₂ inlet tube, containing $[\text{Zn(II)-Gly-Tyr}]^+$, surfactant and solvent were taken, immersed in an oil-bath thermostated at desired temperature. The reaction was charged with addition of required volume of thermally equilibrated ninhydrin solution. The progress of the reaction was followed spectrophotometrically by pipetting out aliquots at definite time intervals and measuring the absorbance of the reaction product at 400 nm (Fig. 1) using UV-vis spectrophotometer. The *pseudo*-first-order rate constants (k_p) were calculated up to 80% completion of the reaction by using a computer program [9].

The cmc values of surfactants in the absence and presence of reactants were obtained from the break points of nearly two straight lines of the specific conductivity vs. [surfactant] plots [10]. The experiments were carried out at 30 °C and 80 °C under different experimental conditions, i.e., solvent being water, water + ninhydrin, water + $[\text{Zn(II)-Gly-Tyr}]^+$, water + ninhydrin + $[\text{Zn(II)-Gly-Tyr}]^+$ and respective cmc values are:

(a) 16-6-16 ($\times 10^3$): 0.043, 0.039, 0.046 and 0.048 (at 30 °C); 0.058, 0.060, 0.057 and 0.054 mol dm⁻³ (at 80 °C).

(b) 16-5-16 ($\times 10^3$): 0.034, 0.033, 0.047 and 0.048 (at 30 °C); 0.055, 0.053, 0.057 and 0.053 mol dm⁻³ (at 80 °C).

(c) 16-4-16 ($\times 10^3$): 0.032, 0.031, 0.041 and 0.039 (at 30 °C); 0.043, 0.044, 0.049 and 0.050 mol dm⁻³ (at 80 °C).

Results and Discussion

To see the effect of pH on the rate, kinetic runs were carried out under varying condition of pH from 4.0 to 6.0, at fixed [ninhydrin] = 6.0×10^{-3} mol dm⁻³, [Zn(II)-Gly-Tyr]⁺ = 3.0×10^{-4} mol dm⁻³, temperature = 80 °C and [gemini] = 30×10^{-5} mol dm⁻³. It was found that the value of rate constant increases up to pH = 5.0 and then becomes almost constant. So, all subsequent kinetic measurements were made at pH = 5.0.

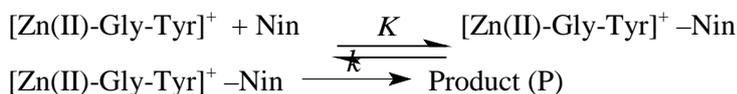
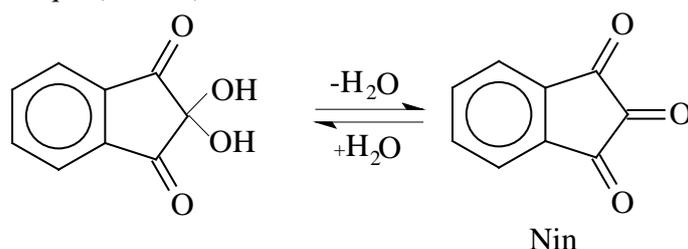
In order to understand the effect of [Zn(II)-Gly-Tyr]⁺ on rate, the experiments were carried out under *pseudo*-first-order conditions of [ninhydrin] \gg [Zn(II)-Gly-Tyr]⁺ in the range of $2\text{--}4 \times 10^{-4}$ mol dm⁻³ of [Zn(II)-Gly-Tyr]⁺ at constant [ninhydrin] of 6×10^{-3} mol dm⁻³, [surfactant] of 30×10^{-5} mol dm⁻³ and an analytical pH 5.0 at 80 °C (Table 1). The rate constant was invariant with [Zn(II)-Gly-Tyr]⁺ which follows first-order kinetics in [Zn(II)-Gly-Tyr]⁺ complex (Eq. (1)),

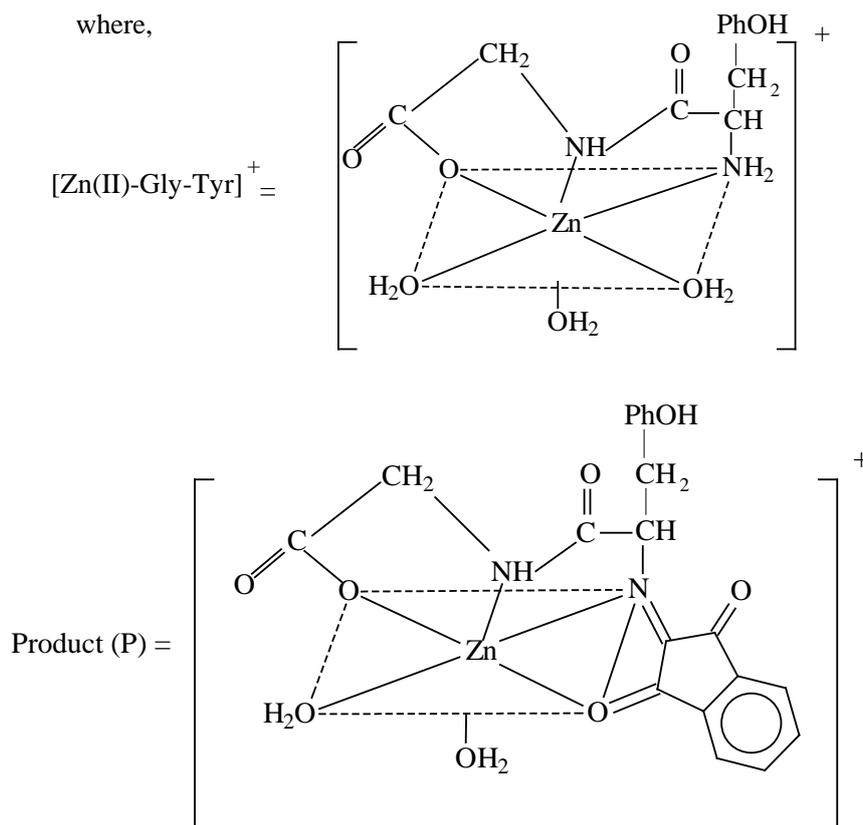
$$\frac{d[P]}{dt} = k_p [\text{Zn(II)-Gly-Tyr}]^+ \quad (1)$$

The effect of [ninhydrin] upon the rate of formation of product was studied within the [ninhydrin] range of 6×10^{-3} mol dm⁻³– 40×10^{-3} mol dm⁻³ at 80 °C and in the presence of 30×10^{-5} mol dm⁻³ surfactant and 3×10^{-4} mol dm⁻³ [Zn(II)-Gly-Tyr]⁺ (Table 1). The rate constants–[ninhydrin] profiles pass through the origin which indicates fractional-order kinetics in [ninhydrin].

On the basis of the above results and previous observations, the mechanism shown in Scheme 1 has been proposed for the reaction of [Zn(II)-Gly-Tyr]²⁺ complex with ninhydrin. The lone pair electrons of amino group are necessary for nucleophilic attack on the carbonyl group of ninhydrin [11]. In complex, the lone pair is not free, therefore, nucleophilic attack is not possible. The reaction, therefore, proceeds through condensation of coordinated amino group to the coordinated carbonyl group within the coordination sphere of zinc(II). The coordination of both ninhydrin and Gly-Tyr occurs with the same zinc metal ion (template mechanism [11]) (Scheme 1).

Activation parameters are believed to provide useful information regarding the environment in which chemical reactions take place. In order to learn more about the micro-environments of submicroscopic assemblies, the effect of temperature on reaction rate, a series of kinetic runs were carried out within the temperature range 70 to 90 °C at fixed [ninhydrin] (6.0×10^{-3} mol dm⁻³), [Zn(II)-Gly-Tyr]⁺ (3.0×10^{-4} mol dm⁻³) and pH (5.0) in the gemini micellar media (16-*s*-16 = 30×10^{-5} mol dm⁻³, *s* = 4, 5, 6). The values of activation parameters were evaluated with the help of linear least squares regression technique (Table 1).



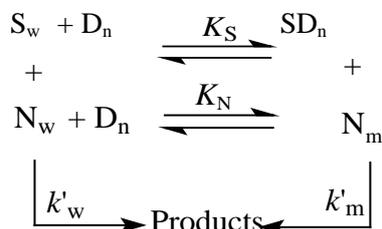


Scheme 1 $[\text{Zn(II)-Gly-Tyr}]^+$ -ninhydrin reaction mechanism.

Investigations in presence of gemini surfactants disclose that the rate of formation of product is first-order in $[\text{Zn(II)-Gly-Tyr}]^+$ and fractional-order in [ninhydrin] (just as in aqueous medium), confirms formation of the same reaction product in both the media. To find out the effect of gemini surfactants on the reaction rate, experiments were performed at varying concentrations of geminis at constant [ninhydrin] ($6.0 \times 10^{-3} \text{ mol dm}^{-3}$), $[\text{Zn(II)-Gly-Tyr}]^+$ ($3.0 \times 10^{-4} \text{ mol dm}^{-3}$), at temperature (80°C) and pH (5.0) (Table 2). It has been found that the rate constant increases firstly (part I, [geminis] below their cmc), and then becomes almost constant up to definite concentration (part II, up to $400 \times 10^{-5} \text{ mol dm}^{-3}$) (the characteristics of part I and part II are as conventional surfactant micelles). In the last, k_{p} -values increase sharply (part III) (Fig. 2).

In part I, rate constant (k_{p}) should remain constant because [geminis] are below the cmc but increase in the observed rate constant values may be due to presence of premicelles or preponement of micellization by $[\text{Zn(II)-Gly-Tyr}]^+$ and ninhydrin [12]. Regarding part II, the values of rate constant remain almost constant up to $400 \times 10^{-5} \text{ mol dm}^{-3}$ of surfactants. The gemini micelles show better environment over the corresponding monomeric single chain surfactants. This could be because of presence of spacer in geminis, which decreases the amount of water in aggregates providing different microenvironment (less polar), thus causing the value of rate constant increases [13]. It has been mentioned that due to proximity of positive charges in gemini surfactants anion binding at the surfaces is increased at the expense of binding of water molecules [14]. The k_{p} -values for all the three geminis at all concentrations follow the order $16\text{-}4\text{-}16 > 16\text{-}5\text{-}16 > 16\text{-}6\text{-}16$ and having same behavior (Fig. 2). It was found several times that the best results were noticed by 16-4-16 gemini surfactant among the series 16-*s*-16 (*s* = 4, 5, 6) [15]. It is well known that, to minimize its contact with water molecule, a spacer longer than the 'equilibrium' distance between two $-\text{NMe}_2$ head groups (the 'equilibrium' distance occurs at *s* = 4 in 16-*s*-16 geminis) tends to loop towards the micellar interior [8]. *s* > 4, increased looping more and will make much more wet to the Stern layer in comparison to *s* = 4, resulting decrease in the rate constant value (k_{p}). Thus, these findings are in agreement with the earlier results that on increasing amount of water content, the reaction environment behaves as an inhibiting effect [16].

In the range III, with the increasing [geminis]: $400-3000 \times 10^{-5} \text{ mol dm}^{-3}$ the value of rate constants (k_{ψ}) increase slowly. Fast increment in rate of reaction occurs at higher [geminis]; this is probably associated with change of micellar structure. This is in conformity with the ^1H NMR studies of the surfactants [13,17]. Thus, at higher [geminis], increase in k_{ψ} -value happens due to changes in the aggregate morphology that provides different reaction environments (less polar).



Scheme 2 Ninhydrin-[Zn(II)-Gly-Tyr]⁺ reaction in gemini micelles.

The catalytic behavior of geminis surfactants can be rationalized in terms of *pseudo*-phase model (Scheme 2) proposed by Menger and Portnoy [18] for the incorporation/association of reactant into the micellar phase, the most successful appears to be that of Bunton [19] and Romsted [20], who suggested an equation,

$$k_{\psi} = \frac{k_w [N]_T + (K_S k_m - k_w) M_N^S [D_n]}{1 + K_S [D_n]} \quad (2)$$

M_N^S being the molarity of ninhydrin bound to the micellar headgroups, $[D_n]$ represents the micellized surfactant (= [surfactant] – cmc), and N_m and SD_n are micellized ninhydrin and [Zn(II)-Gly-Tyr]⁺ complex, respectively. k_w and k_m are the second-order rate constants, referring to aqueous and micellar pseudo-phases, respectively. To confirm the validity of rate Eq. (2), the rate constants were calculated by substituting the values of k_m and K_S in Eq. (2) and comparing with the observed k_{ψ} values (Table 2). The close agreement between the observed (k_{ψ}) and calculated ($k_{\psi\text{cal}}$) values supports the validity of the proposed mechanism.

Activation parameters (Table 1), when compared with those obtained in aqueous medium, show that the presence of cationic micelles decreases the enthalpy of activation (ΔH^\ddagger) which indicates that the transition state is highly solvated in the micellar system. The decrease in entropy of activation (ΔS^\ddagger) shows that the transition state is well structured in the case of micelles.

Conclusions

Effect of gemini surfactants was carried out on the kinetics of the title reaction at 80 °C and pH 5.0. Comparing the values with those obtained in aqueous and CTAB, it was found that gemini micelles catalyze the reaction more. An important point to be noted is that, at present reaction conditions, a smaller amount of surfactants, below the cmc values, was sufficient to accelerate the reaction rate than that of pure water.

Acknowledgements

Authors are thankful to UGC (India) for financial assistance (to DK) and BSR UGC-Faculty Fellowship (to KU).

References:

1. Tanford, C. The Hydrophobic Effect: Formation of Micelles and Biological Membranes. 2nd ed.; New York: Wiley, 1980.
2. Fendler, J.H. Membrane Mimetic Chemistry. New York: Wiley Interscience, 1982.
3. (a) Yu, Y.C.; Tirrell, M.; and Fields, G.B. J. Am. Chem. Soc. 1998, 120, 9979.
(b) Forns, P.; Lauer-Fields, J.L.; Gao, S.; and Fields, G.B. Biopolymers. 2000, 54, 351.
4. McCaldin, D.J. Chem. Rev. 1960, 39, 60.
5. Joullie, M.M.; Thompson, T.R.; and Nemeroff, N.H. Tetrahedron. 1991, 47, 8791.
6. Everse, K.E.; and Menzel, E.R. J. Forensic Sci. 1986, 31, 446.
7. Britton, H.T.S., Hydrogen Ions. vol.1; London: Chapman and Hall, 1942.

8. Zana, R.; Benraou, M.; and Rueff, R. *Langmuir*. 1991, 7, 1072.
9. (a) Kabir-ud-Din; Salem, J.K.J.; Kumar, S.; Rafiquee, M.Z.A.; and Khan, Z. J. *Colloid Interface Sci.* 1999, 213, 20.
(b) Kabir-ud-Din; Salem, J.K.J.; Kumar, S.; Rafiquee, M.Z.A.; and Khan, Z. *Colloids Surf. A*. 2000, 168, 241.
10. Mukerjee, P.; and Mysels, K. J. *Critical Micelle Concentrations of Aqueous Surfactant Systems*. Washington, D. C.: Superintendent of Documents, 1971.
11. Rafiquee, M.Z.A.; Khan, Z.; and Khan, A.A. *Trans. Met. Chem.* 1994 19, 477; *Indian J. Chem. A*. 1994, 33, 1018; Khan, Z.; Gupta, D.; Khan, A.A. *Int. J. Chem. Kinet.* 1993, 25, 437.
12. Cerichelli, G.; Mancini, G.; Luchetti, G.; Savelli, G.; and Bunton, C.A. *Langmuir*. 1994, 10, 3982.
13. Brinchi, L.; Germani, R.; Gorracci, L.; Savelli, G.; and Bunton, C.A. *Langmuir*. 2002, 18, 7821.
14. Menger, F.M.; Keiper, J.S. ; Mbadugha, B.N.A.; Caran, K.L.; and Romsted, L.S. *Langmuir*. 2000, 16, 9095.
15. Kabir-ud-Din, and Siddiqui, U.S. *Colloid J.* 2010, 72, 14.
16. Kabir-ud-Din; Salem, J.K.J.; Kumar, S.; Rafiquee, M.Z.A.; and Khan, Z. J. *Colloid Interface Sci.* 1999, 215, 9.
17. (a) Kabir-ud-Din; and Fatma, W. J. *Phys. Org. Chem.* 2007, 20, 440.
(b) Kabir-ud-Din; Fatma, W.; Khan, Z.A.; and Dar, A.A. *J. Phys. Chem. B*. 2007, 111, 8860.
18. Menger, F.M.; and Portnoy, C.E. *J. Am. Chem. Soc.* 1967, 89, 4698.
19. Bunton, C.A. *Catal. Rev. Sci. Eng.* 1979, 20, 1.
20. Mittal K.L., ed. *Micellization, Solubilization and Microemulsions*. vol. 2; New York: Plenum Press, 1997.

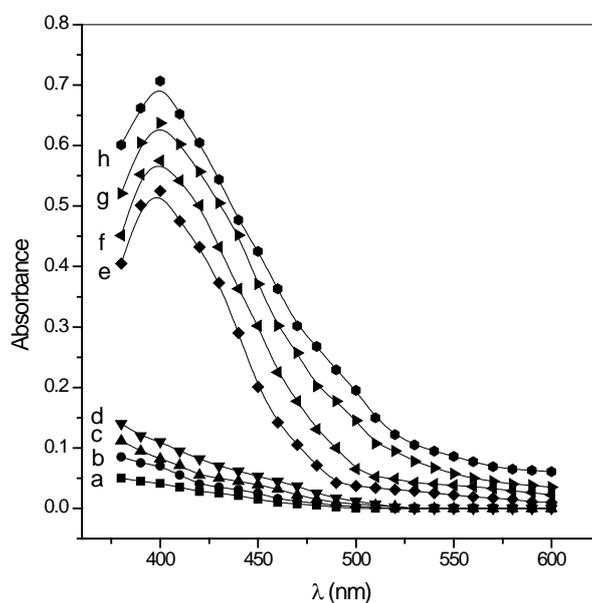


Fig. 1. Absorption spectra of the reaction product of ninhydrin with $[\text{Zn(II)-Gly-Tyr}]^+$ in the absence and presence of gemini surfactants: (a) Immediately after mixing the reactants, (b) Same as solution (a) in presence of 16-6-16, (c) Same as solution (a) in presence of 16-5-16, (d) Same as solution (a) in presence of 16-4-16, (e) After heating solution (a) at 80 °C for 2h, (f) After heating solution (b) at 80 °C for 2h, (g) After heating solution (c) at 80 °C for 2h, (h) After heating solution (d) at 80 °C for 2h. *Reaction conditions:* $[\text{16-}s\text{-16}] = 30 \times 10^{-5} \text{ mol dm}^{-3}$ ($s = 4, 5, 6$), $[\text{ninhydrin}] = 6.0 \times 10^{-3} \text{ mol dm}^{-3}$, $[\text{Zn(II)-Gly-Tyr}]^+ = 3.0 \times 10^{-4} \text{ mol dm}^{-3}$, temperature = 80 °C and pH = 5.0.

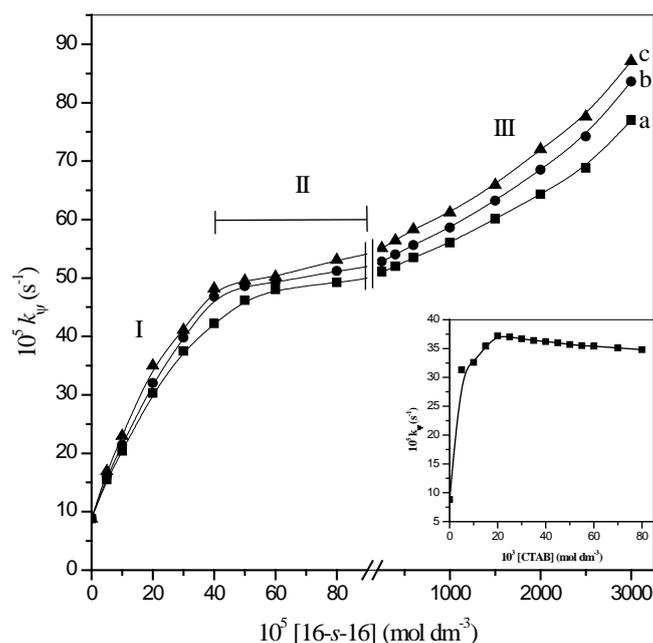


Fig. 2. Effect of varying [geminis] on the reaction rate (k_p) for the reaction of ninhydrin with $[\text{Zn(II)-Gly-Tyr}]^+$: (a) 16-6-16, (b) 16-5-16, (c) 16-4-16 and (Inset) for variation of [CTAB]. Reaction conditions: [ninhydrin] = $6.0 \times 10^{-3} \text{ mol dm}^{-3}$, $[\text{Zn(II)-Gly-Tyr}]^+ = 3.0 \times 10^{-4} \text{ mol dm}^{-3}$, temperature = 80°C and pH = 5.0.

Table 1 Dependence of rate constants (k_p) on $[\text{Zn(II)-Gly-Tyr}]^+$, [ninhydrin] and temperature 80°C at pH 5.0 in presence of gemini surfactants ($30 \times 10^{-5} \text{ mol dm}^{-3}$).

$10^4 [\text{Zn(II)-Gly-Tyr}]^+$ (mol dm^{-3})	$10^3 [\text{ninhydrin}]$ (mol dm^{-3})	temp. ($^\circ\text{C}$)	$10^5 k_p \text{ (s}^{-1}\text{)}$			
			16-6-16	16-5-16	16-4-16	
2.0	6	80	37.3	39.8	40.9	
2.5			37.6	39.4	41.2	
3.0			37.5	39.8	41.1	
3.5			37.7	39.3	41.1	
4.0			37.3	39.3	40.9	
3.0	6	80	37.5	39.8	41.1	
			10	40.0	44.7	47.2
			15	44.5	48.1	50.0
			20	47.3	51.5	53.1
			25	49.5	53.7	55.7
			30	51.1	55.2	57.5
			35	52.2	56.6	58.4
			40	52.8	57.3	59.2
3.0	6	70	30.3	31.1	32.7	
			75	34.8	35.3	36.7
			80	37.5	39.8	41.1
			85	49.1	50.0	50.9
			90	61.5	63.1	72.1

Activation parameters, rate, and binding constants

	E_a (kJ mol^{-1})	ΔH^\ddagger (kJ mol^{-1})	$-\Delta S^\ddagger$ ($\text{JK}^{-1} \text{ mol}^{-1}$)	$10^3 k_m^a$ (s^{-1})	K_S^a ($\text{mol}^{-1} \text{ dm}^3$)	K_N^a ($\text{mol}^{-1} \text{ dm}^3$)
Aqueous	115.1	112.2	307.0	-	-	-
16-6-16	36.2	33.3	307.7	61.0	110.0	66.2
16-5-16	34.5	31.6	308.1	33.0	104.0	70.9
16-4-16	29.9	27.0	308.6	43.0	96.0	63.6

^aAt 80°C

Table 2 Effect of varying [geminis] on the reaction rate for reaction of ninhydrin ($6.0 \times 10^{-3} \text{ mol dm}^{-3}$) with $[\text{Zn(II)-Gly-Tyr}]^+$ ($3.0 \times 10^{-4} \text{ mol dm}^{-3}$) at constant pH 5.0 and 80 °C and their comparison with calculated values ($k_{\Psi\text{cal}}$).

$10^5 [16\text{-s-16}]$ (mol dm^{-3})	16-6-16			16-5-16			16-4-16		
	$10^5 k_{\Psi}$ (s^{-1})	$10^5 k_{\Psi\text{cal}}$ (s^{-1})	$\frac{k_{\Psi} - k_{\Psi\text{cal}}}{k_{\Psi}}$	$10^5 k_{\Psi}$ (s^{-1})	$10^5 k_{\Psi\text{cal}}$ (s^{-1})	$\frac{k_{\Psi} - k_{\Psi\text{cal}}}{k_{\Psi}}$	$10^5 k_{\Psi}$ (s^{-1})	$10^5 k_{\Psi\text{cal}}$ (s^{-1})	$\frac{k_{\Psi} - k_{\Psi\text{cal}}}{k_{\Psi}}$
0	8.8	-	-	8.8	-	-	8.8	-	-
5.0	15.5	-	-	16.0	-	-	16.9	-	-
10.0	20.4	20.5	0.00	21.4	16.2	+0.24	23.0	20.3	+0.12
20.0	30.3	30.1	+0.01	32.0	31.4	+0.02	35.0	34.4	+0.02
30.0	37.5	37.9	-0.01	39.8	40.2	-0.01	41.1	41.2	0.00
40.0	42.2	41.8	+0.01	46.8	46.0	+0.02	48.2	48.8	-0.01
50.0	46.2	45.6	+0.01	48.6	49.3	-0.01	49.5	49.5	0.00
60.0	48.0	48.0	0.00	49.2	49.6	-0.01	50.2	50.3	0.00
80.0	49.2	48.5	+0.01	51.2	51.5	-0.01	53.1	53.4	-0.01
100.0	50.0	49.4	+0.01	52.0	51.8	0.00	54.2	53.8	+0.01
250.0	51.0	50.4	+0.01	52.8	53.5	-0.01	55.1	55.5	-0.01
400.0	52.0	51.7	+0.01	54.0	54.4	-0.01	56.4	56.2	0.00
600.0	53.5	53.0	+0.01	55.6	56.1	-0.01	58.3	58.5	0.00
1000.0	56.0	-	-	58.6	-	-	61.2	-	-
1500.0	60.1	-	-	63.2	-	-	65.9	-	-
2000.0	64.3	-	-	68.5	-	-	72.0	-	-
2500.0	68.8	-	-	74.2	-	-	77.6	-	-
3000.0	77.0	-	-	83.6	-	-	87.1	-	-

PARASITIC INFECTIONS OF THE COMMON CARP *Cyprinus carpio* FROM LESSER ZAB RIVER IN KURDISTAN REGION, IRAQ

Kamaran S. Mama

Shamall M. A. Abdullah

Dep. Biology, Coll. Education, Univ. Salahaddin, Erbil, Iraq

Abstract

A total of 150 common carp *Cyprinus carpio* were collected from Lesser Zab river, near Alton Kupri, south of Erbil city, Kurdistan region, Iraq, during the period from August 2010 until the end of May 2011. The fishes were examined for ecto- and endoparasites. The study revealed the existence of nine species of parasites. These included two species of protozoans (*Ichthyophthirius multifiliis* and *Trichodina reticulata*), four species of monogeneans (*Gyrodactylus baikalensis*, *G. elegans*, *G. gobioninum* and *G. vicinus*), one species of each of trematode (*Diplostomum spathaceum*), cestode (*Bothriocephalus acheilognathi*) and crustacean (*Lernaea cyprinacea*). The record of *G. gobioninum* in the present study is considered as the first record in Iraq, and *T. reticulata* and *G. baikalensis* in Kurdistan region. Also, *C. carpio* is regarded as a new host for *T. reticulata* in Iraq.

Keywords: Common carp, Parasite, Lesser zab river, Kurdistan region, Iraq. received

Introduction

The common carp *Cyprinus carpio* Linnaeus, 1758 belong to Order Cypriniformes, Family Cyprinidae (Froese and Pauly, 2011). This fishes was introduced for the first time into Iraq in 1955 from Holland and Indonesia into Al- Zaaferaniya fish farm at Baghdad city (Hamed, 1960). In Kurdistan region, the first batch of these fishes was cultured in Dokan lake in 1967. As recently as 2004, two million fingerlings of this species were distributed by FAO to the Dokan hatchery to enhance production, and now spread to many farms in the region (Coad, 2010).

Due to common carp's adaptation to a wide range of climatic and geographical condition, many of parasites have been found in wild and domestic carp (Tekin-Özan *et al.*, 2008). As a result of the development of carp culture industry and the high prevalence of this species in all main rivers, lakes and marshes In Iraq, the parasites of *C. carpio* had a great deal and much attracted the attention of biologists (Mhaisen, 1993). The importance of fish parasites is directly related to the importance of the fishes that they may infect and result in different types of injuries and damages (Amlacher, 1970).

The present paper is the second part of a large scale study concerning with the parasites infected *C. carpio* from Lesser Zab river in Kurdistan region, Iraq.

Materials and Methods

A total of 150 *C. carpio* were collected from Lesser Zab river near Alton Kupri town in the north-east of Iraq, between latitudes 34°-36° south to north and longitude 43°-46° west to the east, 45 km south of Erbil city. The fish were collected by fishermen by using cast net and gill nets, during the period from August, 2010 until the end of June 2011. Fish were transported a live in a cool box with pond or local river water to the laboratory of parasitology, College of Education/ Scientific Departments, of Biology, University of Salahaddin. The fishes were identified according to Froese and Pauly (2011).

In the laboratory, the fishes were examined externally and internally for parasites. Skin, fins, gills and buccal cavity smear were prepared by slight scraping and examined under a light compound microscope. Whole eyes were removed then the lens was dissected and removed from each eye then inspected under dissecting microscope for parasites. For the study of the internal parasites, the fishes were dissected from the ventral side. The body cavity, stomach, intestine, spleen, liver, kidneys, heart, muscles, swim bladder and gonads were separated and examined carefully under a dissecting microscope for the presence of parasites or cysts (Amlacher, 1970).

Methods used for collecting, fixing, staining and mounting of the parasite specimens were as follows:

Protozoa: For the study of the protozoan, most of them were identified in a direct way without any technique, by preparing smears of mucus which was scraped from the skin, fins and gills, onto a slide and spreaded carefully with a cover slip and examined, then important measurement were taken for each protozoan.

Monogenea: Skin, fins and gills smears were prepared and microscopically examined. Care was taken to isolate and flatten the specimens, which were then stained with aqueous neutral red and permanent slides were prepared with glycerin-gelatin (Kritsky *et al.*, 2004).

Digenea and Cestoda: Live worms were washed in 0.6% saline solution and fixed in 5% hot formalin, stained with haematoxylin or acetocarmine, then cleared in xylene and mounted in Canada balsam (Scholz, 1989).

Crustacea: Specimens were cleared with 85% lactic acid, and permanent slides were prepared by using jelly glycerin (Kim, 2004).

Photos were taken with Sony Syber Shot Digital camera model DSC-T9, 7.2 mega pixels. The figures were drawn by using a Camera Lucida (Drawing tube). Measurements of parasite were made with an Olympus ocular micrometer. The detected parasites were identified according to their morphology. Parasites were identified according to Bykhovskaya-Pavlovskaya *et al.* (1962), Gussev (1985) and Pugachev *et al.* (2010).

Results and Discussion

A total of 150 common carp *C. carpio* were collected and inspected for parasites from Lesser Zab river. The present study showed the existence of nine species of parasites including two species of protozoans (*Ichthyophthirius multifiliis* and *Trichodina reticulata*), three species of monogenetic trematodes (*Gyrodactylus baikalensis*, *G. elegans*, *G. gobioninum* and *G. vicinus*), one species of each of metacercariae of digenetic trematode (*Diplostomum spathaceum*), cestode (*Bothriocephalus acheilognathi*), and crustacean (*Lernaea cyprinacea*). The distribution of the parasites, their location on or in the fish host body, the prevalence and mean intensity of infection are summarized in Table (1). The following is an account on the description and measurements of these parasites, especially those which were recorded for the first time in Kurdistan region and in Iraq:

***Ichthyophthirius multifiliis* Fauquet, 1876**

This ciliated protozoan was found on the skin, fins and gills of *C. carpio* with prevalence of 2.66%, (Table 1). This parasite has been reported for the first time in Iraq from *Mugil dussmien* (Herzog, 1969). After that, it was reported from 23 different fish hosts including *C. carpio* (Mhaisen, 2012). In Kurdistan region, it was reported from different species from Darbandikhan lake, Lesser Zab river, Greater Zab river and Ainkawa fish hatchery (Abdullah, 2005; Abdullah and Mhaisen, 2006; Al-Marjan and Abdullah 2009).

***Trichodina reticulata* Hirschmann et Partsch, 1955**

This ciliated protozoan was isolated from the skin and fins of *C. carpio* with prevalence of 3.33% (Table 1). Medium to large disc-shaped trichodinid diameter of body 55-70 μm , adhesive disc 48-55 μm in diameter, denticle ring 30-38 μm . Number of denticles 22-26. Inner margin of denticle blade curved and roughly angular on outside margin, with squared distal end, length of blade 6-7 μm . Central part oblong, with oblong to rounded overlapping end. Projection of central part is invisible. Ray with little taper, tip blunt to square; length of ray 5-7 μm . Length of span 14-16 μm (**Fig. 1**).

This parasite was recorded for the first time in Iraq by Jori (2006) on the gills of *Silurus triostegus* from Al-Hammar marshes in Basrah city. No more hosts are known for this species in Iraq. So, *C. carpio* is now considered a new host for this ciliated protozoan in Iraq and the present study represents the first record of *T. reticulata* in Kurdistan region.

***Gyrodactylus baikalensis* Bogolepova, 1950**

This species was recovered from the skin of *C. carpio* with a prevalence of 3.33% (Table 1). Small, elongated worms. Body length 0.40-0.62 mm and width 0.13-0.20 mm. The anterior end of the body bilobed and each lobe has a head organ. Eye spots are absent. The anterior end of the body bilobed and each lobe has a head organ. Eye spots are absent. The haptor armament consists of 16 hooklets and two median hooks without outer roots but with two bars (ventral bar with a membrane and a dorsal bar). Total length of hooklets 0.028-0.032 mm. Total length of median hooks 0.06-0.07

mm, main part 0.047 mm, point 0.032 mm. Size of ventral bar 0.006-0.008 X 0.025-0.028 mm, membrane 0.005-0.008 mm. Size of dorsal bar 0.002 X 0.015-0.018 mm (Fig. 2).

The parasite was recorded for the first time in Iraq from the same host from Suwairah and Latifiyah (Salih *et al.*, 1988). After that, it was reported from eight different fish hosts (Mhaisen, 2012), but there are no any previous records for *G. baikalensis* in Kurdistan region.

***Gyrodactylus elegans* Nordmann, 1832**

This trematode was found on skin and gills of *C. carpio* with a prevalence of 2% (Table 1). This species was reported for the first time in Iraq on the skin *C. carpio* and *Liza abu* from Al-Zaafarany fish farm (Ali and Shaaban, 1984). After that, it was reported on 22 fish species in Iraq (Mhaisen, 2012). Also, it was reported on *C. carpio* from Lesser Zab river, Darbandikhan lake and Ainkawa fish hatchery in Kurdistan region (Abdullah and Mhaisen, 2004; Abdullah, 2005; Al-Marjan and Abdullah, 2009) respectively.

***Gyrodactylus gobioninum* Gussev, 1955**

This parasite was found on the skin of *C. carpio* with a prevalence of 1.33% (Table 1). Small worms. Body length 0.30-0.35 mm and width 0.05-0.09 mm. Length of hooklets 0.022-0.025 mm. Total length of median hooks 0.047-0.050 mm, main part 0.037 mm, point 0.026 mm, inner root 0.016 mm. Size of ventral bar 0.005-0.006 X 0.015-0.017 mm, membrane 0.009-0.014 mm. Size of dorsal bar 0.002 X 0.015-0.017 mm (Fig. 3).

The present specimen agreed very well with Pugachev *at el.* (2010) specimens except that the present parasite is larger, besides it was detected from a different fish host (*C. carpio*). While, Pugachev *et al.* (2010) found it on the fins of *Gobio gobio*, *Romanogobio albipinnatus* and *Abbottina rivularis* in the Danube, Elbe, Older rivers and lake Khanka. Since no previous report about recording of this species is available in Iraq, the present record represents the first record of *G. gobioninum* in Iraq.

***Gyrodactylus vicinus* Bykhovskii, 1957**

This trematode was detected on the skin and gills of *C. carpio* with a prevalence of 4.66% (Table 1). Previously, *G. vicinus* was reported in Iraq on the skin of *C. carpio* from Al-furat fish farm in Babylon province and Lesser Zab river in Kurdistan region (Al-Zubaidy, 1998; Abdullah and Mhaisen, 2004). The present monogenetic trematodes are greatly in accordance with the specimens which were recorded previously in Iraq.

***Diplostomum spathaceum* (Rud., 1819)**

This metacercariae was found in the eye lens of *C. carpio* with a prevalence of 1.33% (Table 1). Abdullah (1990), who recorded *D. spathaceum* for the first time in Iraq from eye lens of *B. luteus*, *C. macrostomum* and *C. carpio* from Dokan Lake. According to Mhaisen (2012) a total of 26 fish host species were so far known for *D. spathaceum* in Iraq.

***Bothriocephalus acheilognathi* Yamaguti, 1934**

Many specimens of larvae and adults of this cestode were found in intestine of *C. carpio* with a prevalence of 2% (Table 1). This parasite was recorded for the first time in Iraq by Khalifa (1982) from *C. carpio* in some fish farms in Baghdad. Later, it was reported from 19 fish hosts in Iraq (Mhaisen, 2012).

***Lernaea cyprinacea* Linnaeus, 1758**

This crustacean was found on the skin and gills of *C. carpio* with prevalence of 2.66% (Table 1). The anchor worm was firstly recorded in Iraq on seven fish species in Al-Zaafarany fish farm south of Baghdad (Al-Hamed and Hermiz, 1973). Later, it was reported from 24 fish host species from many fish farms as well as from many inland water bodies in Iraq (Mhaisen, 2012). In Kurdistan region, it was reported from five species of fishes namely: *Barbus barbulus*, *B. grypus*, *B. luteus*, *Cyprinus carpio* and *Leuciscus Lepidus* from Dokan, Darbandikhan lakes and Ainkawa fish hatchery (Abdullah, 1990; 2005; Al-Margan, 2007).

References:

- Abdullah, S. M. A. (1990). Survey of the parasites of fishes from Dokan lake. M. Sc. Thesis, Coll. Sci., Univ. Salahaddin: 115pp. (In Arabic).
 Abdullah, S. M. A. (2005). Parasitic fauna of some freshwater fishes from Darbandikhan lake, North of Iraq. J. Dohuk Univ., 8(1): 29-35.

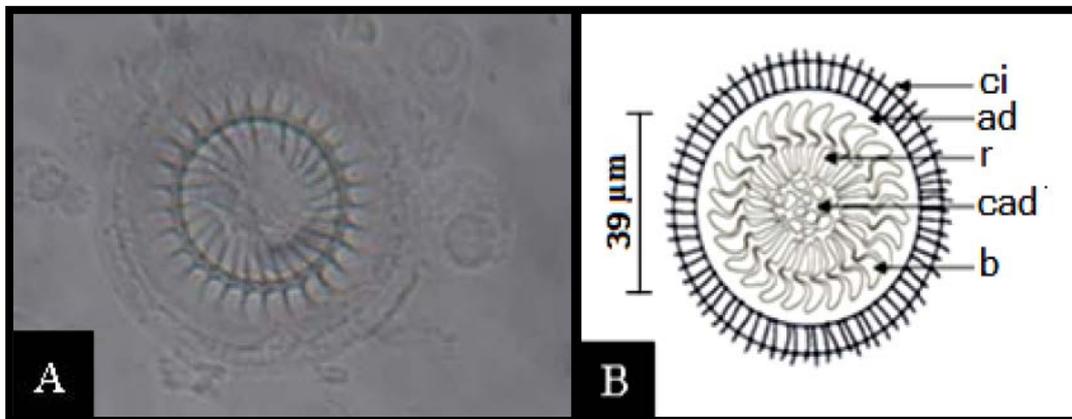
- Abdullah, S. M. A. and Mhaisen, F. T. (2004). Parasitic infections with Monogenetic trematodes on fishes of Lesser Zab and Greater Zab Rivers in northern Iraq. *Zanco*, 16(4): 43-52.
- Abdullah, S. M. A. and Mhaisen, F. T. (2006). Parasitic infections with Protozoa and Crustacea on fishes of Lesser Zab and Greater Zab rivers, north of Iraq. *Proc. 4th Sci. Conf. Coll. Vet. Med., Univ. Mosul / 20-21 September 2006*, 1: 51-58.
- Al-Hamid, M. I. and Hermiz, I. (1973). Experiments on the control of anchor worm (*Lernaea cyprinacea*). *Aquaculture*, 2: 45-51.
- Ali, M. D. and Shaaban, F. (1984). Some species of parasites of freshwater fish raised in ponds and Tigris-Al-Tharthar canal region. *Seventh Sci. Conf., Iraqi Vet. Med. Ass., Mosul*. (Abstract).
- Al-Margan, K. S. N. (2007). Some ectoparasites of the common carp (*Cyprinus carpio*) with experimental study of the life cycle of the anchor worm (*Lernaea cyprinacea*) in Ainkawa fish hatchery, Erbil province. *M. Sc. Thesis, Sci. Educ. Coll., Univ. Salahaddin*: 99pp.
- Al-Marjan, K. S. N. and Abdullah, S. M. A. (2009). Some ectoparasites of the common carp (*Cyprinus carpio*) in Ainkawa fish hatchery, Erbil province. *Duhok, J. Duhok Univ. (Special Issue)*, 14(1): 102-107.
- Al-Zubaidy, A. B. (1998). Studies on the parasitic fauna of carps in Al-Furat fish farm, Babylon province, Iraq. *Ph. D. Thesis, Coll. Sci., Univ. Babylon*: 141pp. (In Arabic).
- Amlacher, E. (1970). *Textbook of fish diseases (Engl. Transl.)*. T. F. H. Publ., Jersey city: 302pp.
- Bykhovskaya-Pavlovskaya, I. E.; Gusev, A. V.; Dubinia, M. N.; Izyumova, N. A.; Smirnova, T. S.; Sokolovskaya, I. L.; Shtein, G. A.; Shul'man, S. S. and Epshtein, V. M. (1962). *Key to parasites of freshwater fish of the U. S. S. R. Akad. Nauk, S. S. S. R., Moscow*: 727pp. (In Russian).
- Coad, B. W. (2010). *Freshwater fishes of Iraq*. Pensoft Publisher, Sofia: 275 pp+ 16 plates.
- Froese, R. and Pauly, D. (2011). *FishBase. World web electronic publication*. www.fishbase.org, version (2/2011).
- Pugachev, O. N.; Gerasev, P. I.; Gussev, A. V.; Ergens, R. and Khotenowsky, I. (2010). *Guide to monogenoidea of freshwater fish of Palaearctic and Amur regions*. Ledizioni Ledipublishing, Milano: 567pp.
- Gussev, A. V. (1985). Parasitic metazoans: Class Monogenea. In: Bauer, O. N. (Ed.). *Key to the parasites of freshwater fish fauna of the U.S.S.R.* Nauka, Leningrad, 2: 1-424. (In Russian).
- Hamed, M. I. (1960). *Carp culture in Iraq*. *Iraqi J. Agric. Sci.*, 1(2):14-23. (In Arabic).
- Herzog, P. H. (1969). *Untersuchungen über die parasiten der süßwasserfische des Irak*. *Arch. Fischereiwiss.*, 20(2/ 3): 132-147.
- Jori, M. M. (2006). Parasitic study on the asian catfish *Silurus triostegus* (Heckel, 1843) from Al-Hammar marshes, Basrah, Iraq. *M. Sc. Thesis, Coll. Educ., Univ. Basrah*: 192pp.
- Khalifa, K. A. (1982). Occurrence of parasitic infection in Iraqi an fish ponds. *Second Arab. Conf. Biol. Sci., Arab Biol. Union Fes*. 17-20 March, 1982 (Abstract).
- Kim, I. H. (2004). Copepodid stages of *Ergasilus hyponesi* Yamaguti, (Copepoda, Poecilostomatoida, Ergasilidae) from a brackish lake in Korea. *Kor. J. Biol. Sci.*, 8: 1-2.
- Kritsky, D. C.; Panndey, K. C.; Agrawal, N. and Abdullah, S. M. A. (2004). Monogenoids from the gills of spiny eels (Teleostei: Mastacembelidae) in India and Iraq, proposal of *Mastacembelocleidus* gen. n., and status of the Indian species of *Actinocleidus*, *Urocleidus* and *Haplocleidus* (Monogenoidea: Dactylogyridae). *Folia Parasitol.*, 51: 291-298.
- Mhaisen, F. T. (1993). A review on the parasites and diseases in fishes of ponds and farms of Iraq. *Iraqi J. Vet. Sci.*, 6(2): 20-27. (In Arabic).
- Mhaisen, F. T. (2012). *Index-catalogue of parasites and disease agents of fishes of Iraq*, Unpuble. (Personal communication).
- Salih, N. E.; Ali, N. M. and Abdul-Ameer, K. N. (1988). Helminthic fauna of three species of carp raised in ponds in Iraq. *J. Biol. Sci. Res.*, 19(2): 369-386.
- Scholz, T. (1989). Amphilinida and Cestoda, parasites of fish in Czechoslovakia. *Acta Sci. Nat. Brno.*, 23(4): 1-56.
- Tekin-Özan, S.; Kir, İ. and Barlas, M. (2008). Helminth parasites of common carp (*Cyprinus carpio* L., 1758) in Beyşehir lake and population dynamics related to month and host size. *Turk. J. Fish. Aquat. Sci.*, 8: 201-205.

Table (1): The distribution of parasites on the gills of *C. carpio* (150) from Lesser Zab river.

Parasites	No. of infected fish	Prevalence (%)	Mean intensity	Site of infection
<i>Ichthyophthirius multifiliis</i>	4	2.66	10	Skin, Fins & Gills
<i>Trichodina reticulata</i> ** α	5	3.33	12	Skin & Fins
<i>Gyrodactylus baikalensis</i> **	5	3.33	4.4	Skin
<i>Gyrodactylus elegans</i>	3	2	5.66	Skin & Fins
<i>Gyrodactylus gobioninum</i> *	2	1.33	3	Skin
<i>Gyrodactylus vicinus</i>	7	4.66	5.28	Skin & Fins
<i>Diplostomum spathaceum</i>	2	1.33	4	Eye lens
<i>Bothriocephalus acheilognathi</i>	3	2	5	Intestine
<i>Lernaea cyprinacea</i>	4	2.66	3	Skin & Gills

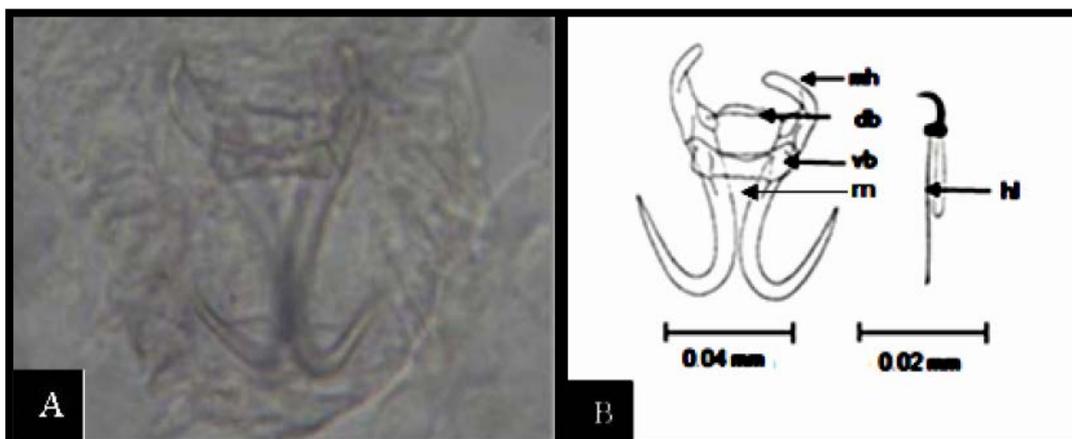
* First record in Iraq.

** First record in Kurdistan region.

 α New host record in Iraq.**Fig. (1): *Trichodina reticulata*.**

A- Photomicrograph (500X); B- Camera lucida drawing.

ad= adhesive disc; b= blade; cad= central of adhesive disc; ci= cilia; r= ray.

**Fig (2): *Gyrodactylus baikalensis*.**

A- Photomicrograph of the haptor (500X); B- Camera lucida of the haptor.

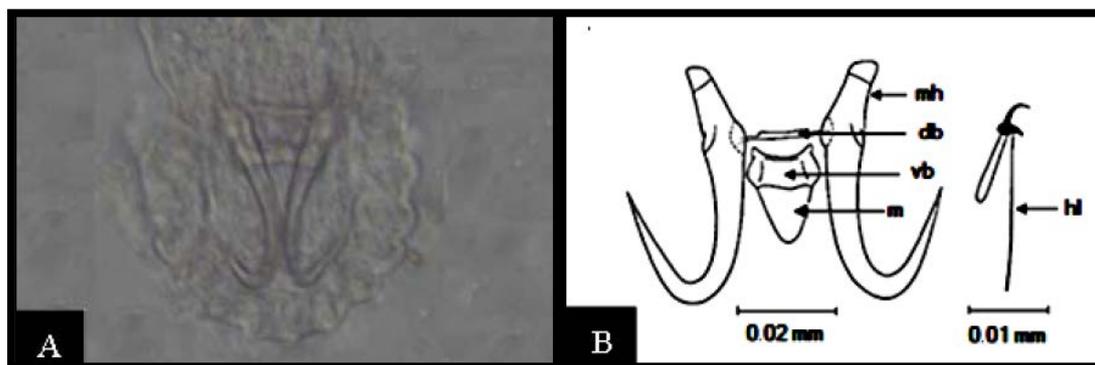


Fig. (3): *Gyrodactylus gobioninum*.

A- Photomicrograph of the haptor (500X); **B-** Camera lucida of the haptor.
db= dorsal bar; **hl**= hooklet; **m**= membrane; **mh**= median hook; **vb**= ventral bar.

PREVALENCE AND COMPARISON BETWEEN THE EFFICACY OF DIFFERENT TECHNIQUES FOR DIAGNOSIS OF *TOXOPLASMA GONDII* AMONG WOMEN IN ERBIL PROVINCE-IRAQI KURDISTAN

Narmin R. Hamad

College of Science, Salahaddin University

Mohammed A. Kadir

College of Medicine, Kirkuk University

Abstract

Introduction: Toxoplasmosis is one of the most common social and gynecological problems in Iraq.

Materials and Methods: This study investigated the prevalence of *Toxoplasma gondii* among females attending the gynaecological out patient's clinics in Maternity and Rizgary hospitals, several health centers and some private clinics in Erbil-Iraqi Kurdistan, from 16th July 2007 till 15th July 2008, by using different diagnostic methods.

Serological diagnosis was done using LAT, ELISA-IgM and IgG tests and 2ME method, for investigation of *T. gondii* seroprevalence, in relation with residency, marital status, age, education, occupation, gestation, frequency of abortion and infertility.

Results: The rate of *Toxoplasma* seropositivity, was higher using LAT 54.46%, followed by ELISA IgG 37.5%, 2ME 14.29% and ELISA IgM 9.13%, and higher rates were recorded among rural inhabitants in comparison with those living in urban areas. The rate of *Toxoplasma* seropositivity was highest among married women, and the age group 47-57 years revealed highest infection rate. The higher rates were indicated among the housewives than the employers.

The pregnant women have higher seroprevalence rate than non-pregnant patients. Women with more than three abortions showed the higher rate of seropositive antibodies.

Conclusions: Significant difference was recorded between the techniques and seropositivity of *T. gondii* antibodies which was higher by LAT than ELISA-IgG, ELISA IgM and 2ME.

Keywords: *Toxoplasma gondii*, Prevalence, Diagnostic methods, Erbil-Iraq

Introduction

Toxoplasmosis is a major public health problem, with a high socioeconomic impact in terms of human suffering including cost of caring for sick, mentally retarded and blind children (1). The parasite is an extremely successful pathogen, responsible for significant morbidity and mortality, especially in congenitally infected and immuno-compromised individuals (2).

Toxoplasmosis can be diagnosed serologically by several tests that depend on the demonstration of *Toxoplasma* antibodies IgM and IgG in serum (3). These tests include the dye test, immunofluorescent study test, latex agglutination test, the enzyme-linked immunosorbent assay (ELISA) and complement fixation test. In recent years, the PCR (polymerase chain reaction) amplification of parasite ribosomal DNA was developed for the detection of single organism in the tissue samples (4).

The study was planned to determine the prevalence of toxoplasmosis in women in Erbil governorate and its relationship with the number of abortion and infertility and to compare between the efficacy of different direct examination and serological tests for diagnosis of *Toxoplasma* seropositivity.

Materials And Methods

Time and location

Three hundred and three female patients were enrolled in the study, between 16th July 2007 and 15th July 2008, for detection of *Toxoplasma gondii* infection, they were attended to some private

clinics, gynecological out patient's clinics in Maternity and Rizgary hospitals and several health centers in Erbil Governorate.

Collection of specimens

Blood samples were withdrawn, from the patients and controls, centrifuged and sera were collected kept at 4-8°C for about 24-48 hours, if longer period was needed, then they were stored in a deep freezing at -20°C for performing serological tests.

Serological tests

Four serological tests were used for detection of specific antibody to *Toxoplasma gondii* by the use of Latex Agglutination Test (LAT), IgM and IgG antibodies using ELISA (Enzyme-Linked Immunosorbent Assay) technique and 2 Mercapto Ethanol (2ME).

a. Latex agglutination test (LAT) is performed using toxoplasmosis Latex Test Kits from Plasmatec Laboratory Products Ltd (U.K.).

b. The *Toxoplasma* IgM ELISA and IgG ELISA were done, using the Genesis Diagnostics (UK) and BioCheck (CA) *Toxoplasma* IgM kits.

c. 2Mercapto-Ethanol (2ME) done by using the complete kit of 2-Mercapto Ethanol from BioMerieux.

Statistical analysis

The statistical analysis was conducted using the software program Statistical Program Social System (SPSS version 13.0). Independent-Samples t-test was used for the biochemical tests for determination of the significant variations. The other remaining data were analyzed using Chi-square test (5).

Results

Relationship between prevalence of *Toxoplasma gondii*-Ab and residency of the patients using different laboratory methods is shown in Table (1). The rate of *Toxoplasma* seropositivity was higher using LAT 54.46%, followed by ELISA IgG 37.5%, 2ME 14.29% and ELISA IgM 9.13%. Higher rates in the subjects were recorded, in almost all the tests, from rural villages than urban areas, but statistical difference was insignificant.

Table 1: Relationship between prevalence of *Toxoplasma* Ab and residency of the patients using different laboratory methods.

Residency	Latex agglutination			ELISA-IgM			ELISA-IgG			2ME		
	×	∞	%	×	∞	%	×	∞	%	×	∞	%
Urban	190	103	54.2	137	13	9.49	137	50	36.5	63	7	11.1
Rural	113	62	54.8	71	6	8.45	71	28	39.4	35	7	20.0
Total	303	165	54.4	208	19	9.13	208	78	37.5	98	14	14.2
	$\chi^2=0.012$ df=1 P>0.01			$\chi^2=0.061$ df=1 P>0.01			$\chi^2=0.172$ df=1 P>0.01			$X^2=1.452$ df=1 P>0.0		

The overall marital status-adjusted *T. gondii* seroprevalence, as shown in Table (2), was found among married community higher than single (unmarried) and others (widow, divorced, and separated relationships). Using LAT, the highest seroprevalence rate was among married patients 54.95% followed by single and others (widow, divorced and separated women) 50% and 33.33% respectively, while only married females show seropositivity for toxoplasmosis using ELISA-IgG, 2ME and ELISA-IgM 38.81%, 14.58% and 9.45% respectively.

Table 2: Seroprevalence of *Toxoplasma gondii* according to the marital status of the patients using different laboratory methods.

Marital status	Latex agglutination			ELISA-IgM			ELISA-IgG			2ME		
	exam	No. +ve	%	exam	No. +ve	%	exam	No. +ve	%	exam	No. +ve	%
Single	4	2	50.00	2	0	0	2	0	0	1	0	0
Married	293	161	54.95	201	19	9.45	201	78	38.81	96	14	14.5
Others	6	2	33.33	5	0	0	5	0	0	1	0	8
Total	303	165		208	19		208	78		98	14	0

Others = Widow, Divorced and Separated Women.

Considering the age group and its relation with the distribution of seropositive *Toxoplasma* antibodies, women of the age group 47-57 years had the highest percentage 66.67% of the positive results of *Toxoplasma* Ab, followed by the age groups 25-35 years 56.93%, 14-24 years 56.48% and 36-46 years 43.64% using LAT. The highest seropositivity rate of *Toxoplasma* antibodies by ELISA-IgM was 13.43% in the age group of 14-24 years, followed by 8.25% and 4.76% in the age groups 25-35 and 36-46 years respectively. While in ELISA-IgG, the age group 47-57 years showed the highest positive rate 50% followed by 14-24 years 46.27%, 25-35 years 34.02% and 36-46 years 30.95%. Using 2ME method, only the age groups 14-24 and 25-35 years were positive for toxoplasmosis 19.15% and 11.9% respectively (table 3).

Table 3: Seroprevalence of *Toxoplasma gondii* according to the age groups using different laboratory methods.

Age (year)	Latex agglutination			ELISA-IgM			ELISA-IgG			2ME		
	exam	No. +ve	%	exam	No. +ve	%	exam	No. +ve	%	exam	No. +ve	%
14-24	108	61	56.48	67	9	13.43	67	31	46.27	47	9	19.15
25-35	137	78	56.93	97	8	8.25	97	33	34.02	42	5	11.90
36-46	55	24	43.64	42	2	4.76	42	13	30.95	9	0	0
47-57	3	2	66.67	2	0	0	2	1	50.00	-	-	-
Total	303	165		208	19		208	78		98	14	

Table (4) reveals seroprevalence of toxoplasmosis according to the educational status using various laboratory methods and a statistical association was not found between *T. gondii* seroprevalence and the education level of the women but the disease is more prevalent among women with school education and illiterate patients almost in all the tests and lower seroprevalence was present between women had college education.

Table 4: Seroprevalence of *Toxoplasma gondii* according to the educational status using different laboratory methods.

Educational status	Latex agglutination			ELISA-IgM			ELISA-IgG			2ME		
	exam	No. +ve	%	exam	No. +ve	%	exam	No. +ve	%	exam	No. +ve	%
Illiterate	60	33	55.00	41	4	9.76	41	15	36.59	13	3	23.08
School	194	110	56.70	128	13	10.16	128	54	42.19	72	11	15.28
University	49	22	44.90	39	2	5.13	39	9	23.08	13	0	0
Total	303	165		208	19		208	78		98	14	
	$\chi^2=2.206$ df=2 P>0.01			$\chi^2=0.934$ df=2 P>0.01			$\chi^2=4.464$ df=2 P>0.01			$\chi^2=3.045$ df=2 P>0.01		

Table (5) illustrates the seroprevalence of toxoplasmosis according to the occupation using four laboratory techniques and the study revealed that the prevalence of toxoplasmosis increases among the housewives than the employers, but was significantly different by LAT. The higher positive infection rates were investigated among the housewives 57.79% by using LAT, 39.13% ELISA IgG, 16.05% 2ME and 10.56% ELISA IgM, than the employers 40.68%, 31.91%, 5.882% and 4.26% respectively.

Table 5: Seroprevalence of *Toxoplasma gondii* according to the occupation using different laboratory methods.

Occupation	Latex agglutination			ELISA-IgM			ELISA-IgG			2ME		
	x a m	o + >	%	x a m	o + >	%	x a m	o + >	%	x a m	o + >	%
Housewife	244	141	57.79	161	17	10.56	161	63	39.13	81	13	16.05
Employer	59	24	40.68	47	2	4.26	47	15	31.91	17	1	5.882
Total	303	165	54.46	208	19	9.13	208	78	37.50	98	14	14.29
	$\chi^2=5.607^*$ df=1 P<0.01			$\chi^2=1.742$ df=1 P>0.01			$\chi^2=0.808$ df=1 P>0.01			$\chi^2=1.186$ df=1 P>0.01		

* Significant

Table (6) clarifies the significant relationship between seropositive *Toxoplasma* antibodies in regard to gestation using various laboratory methods. Higher seroprevalence of toxoplasmosis, using different serological tests, was found in pregnant women, in whom 66.02% were positive for LAT, 48.39% of the women were IgG-positive, indicating previous maternal infection, 14.29% for 2ME and 11.29% were positive for anti-*Toxoplasma* IgM antibody.

Table 6: Frequency of *Toxoplasma* seropositive Ab among pregnant and non-pregnant women using different laboratory methods.

Gestation	Latex agglutination			ELISA-IgM			ELISA-IgG			2ME		
	No. examined	No. +ve	%	No. examined	No. +ve	%	No. examined	No. +ve	%	No. examined	No. +ve	%
Pregnant	103	68	66.0	62	7	11.2	62	30	48.3	42	6	14.2
Non-pregnant	190	93	48.9	139	12	8.63	139	48	34.5	54	8	14.8
	$\chi^2=8.544^*$						df=1			P<0.05		

* Significant

Regarding the frequency of abortion, the higher rate of seropositive *Toxoplasma* antibodies was found in sera of women with more than 3 abortions by the diagnostic tests (Table 7). The distribution of *Toxoplasma* seropositivity in women with more than 3 abortions was 100% rate of seropositivity using 2ME method, 75% by LAT, 50% for IgG antibodies and 0% for IgM type antibodies specific for *T. gondii*, but statistical analysis appears no significant difference.

Table 7: Relationship between number of abortion and seropositive *Toxoplasma* Ab using different laboratory methods.

Number of abortion	Latex agglutination			ELISA-IgM			ELISA-IgG			2ME		
	No. examined	No. +ve	%	No. examined	No. +ve	%	No. examined	No. +ve	%	No. examined	No. +ve	%
Single	63	30	47.62	46	5	10.87	46	17	36.96	18	2	11.11
Double	29	15	51.72	20	1	5.00	20	8	40.00	8	2	25.00
Triple	14	8	57.14	11	1	9.09	11	5	45.45	8	2	25.00
> 3	8	6	75.00	2	0	0	2	1	50.00	1	1	100.0
Total	114	59	51.75	79	7	8.86	79	31	39.24	35	7	20.00
	$\chi^2=2.326$ df=3 P>0.01			$\chi^2=0.794$ df=3 P>0.01			$\chi^2=0.388$ df=3 P>0.01			$\chi^2=5.139$ df=3 P>0.01		

Table (8) explores the relationship between *Toxoplasma* infection and infertility by determination the level of anti-*Toxoplasma gondii* antibodies in sera of women from the studied group

and the results revealed that the positive result of *Toxoplasma* infection was higher 52.33% using LAT, followed by ELISA-IgG 39.62%, 2ME 15% and ELISA-IgM 9.433% respectively.

In LAT and ELISA tests (IgG and IgM type antibodies), the positive serum samples for *Toxoplasma* infections were more prevalent in women suffering from primary infertility 53.19%, 46.43% and 14.29% than secondary infertility 51.28%, 32% and 4% respectively, while in 2ME, the seropositivity rate in women with secondary infertility was slightly higher than those with primary type of infertility 15.79% and 14.29% respectively. Statistical difference was insignificant.

Table 8: Seroprevalence of *Toxoplasma gondii* according to the infertility of the patients using different laboratory methods.

Infertility	Latex agglutination			ELISA-IgM			ELISA-IgG			2ME		
	No. examined	No. +ve	%	No. examined	No. +ve	%	No. examined	No. +ve	%	No. examined	No. +ve	%
Primary	47	25	53.19	28	4	14.29	28	13	46.43	21	3	14.29
Secondary	39	20	51.28	25	1	4.00	25	8	32.00	19	3	15.79
Total	86	45	52.33	53	5	9.433	53	21	39.62	40	6	15
	$\chi^2=0.350$ df=1 P>0.01			$\chi^2=1.635$ df=1 P>0.01			$\chi^2=0.484$ df=1 P>0.01			$\chi^2=0.018$ df=1 P>0.01		

Table (9) shows comparison between efficacies of using four laboratory methods in determination of seropositivity of *Toxoplasma gondii* antibodies in 98 patients. LAT showed the highest rate 48.98% followed by 45.92%, 16.33% and 14.29% with ELISA-IgG, ELISA-IgM and 2ME respectively.

Table 9: Comparison between efficacies of using four laboratory methods in determination of seropositivity of *T. gondii* Ab in 98 patients.

Test	No. +ve	%
Latex agglutination	48	48.98
ELISA-IgM	16	16.33
ELISA-IgG	45	45.92
2ME	14	14.29
$\chi^2=47.331^*$ df=3 P<0.01		

* Significant

The sera of 303 women were examined by Latex agglutination test, 165 were seropositive, and the rate of infection with toxoplasmosis was 54.46%. Out of 208 serum sample tested by ELISA-IgM and IgG, 19 and 78 were seropositive respectively with the rate of seropositivity 9.134% and 37.5% respectively. While among 98 cases tested by 2ME method, only 14 were positive and the rate was 14.29%. The Latex agglutination method revealed the high rate of infection followed by ELISA-IgG, 2ME and ELISA-IgM (Table 10).

Table 10: Comparison between efficacies of using four laboratory methods in determination of seropositivity of *T. gondii* Ab in 303 patients.

Test	No. examined	No. +ve	%
Latex agglutination	303	165	54.46
ELISA-IgM	208	19	9.134
ELISA-IgG	208	78	37.50
2ME	98	14	14.29
$\chi^2=132.7^*$ df=3 P<0.01			

* Significant

Discussion

The rate of *Toxoplasma* seropositivity was higher in subjects from rural villages than urban areas, with insignificant statistical difference.

Several studies determined a higher prevalence in rural regions in comparison with urban areas like: Al-Griari (6) in Diyala, Kadir and Khana (7) in Sulaimani and Ali (8) in Kalar. While

others did not observe any difference between rural and urban inhabitants like: Al-Kaysi (9) in Baghdad; Al-Jubori (10) in Kirkuk.

The overall marital status-adjusted *T. gondii* seroprevalence, was found among married community higher than single (unmarried) and others (widow, divorced, and separated relationships). Using LAT, the highest seroprevalence rate was among married patients 54.95% followed by single and others (widow, divorced and separated women) 50% and 33.33% respectively, while only married females show seropositivity for toxoplasmosis using ELISA-IgG, 2ME and ELISA-IgM 38.81%, 14.58% and 9.45% respectively. The reasons for higher incidence in married females may be related to soil exposure since they are more involved with child rearing and women with three children were found to have significantly higher seroprevalence of *T. gondii* infection compared with other women (11).

Considering the age group and its relation with the distribution of seropositive *Toxoplasma* antibodies, women of the age group 47-57 years had the highest percentage 66.67% of the positive results of *Toxoplasma* Ab, using LAT. The highest seropositivity rate of *Toxoplasma* antibodies by ELISA-IgM was 13.43% in the age group of 14-24 years, While in ELISA-IgG, the age group 47-57 years showed the highest positive rate 50%. Using 2ME method, only the age groups 14-24 and 25-35 years were positive for toxoplasmosis 19.15% and 11.9% respectively. The highest rate of *Toxoplasma* seropositivity among the age groups 47-57 years respectively are in agreement with the study of Al-Jubori (10) in Kirkuk, he recorded rate of seropositivity 80% in the age group 51 and above.

Statistical association was not found between *T. gondii* seroprevalence and the education level of the women but the disease is more prevalent among women with school education and illiterate patients almost in all the tests and lower seroprevalence was present between women had college education because increased knowledge results in awareness, which consequently results in changes in risky behavior and decline in infection rates, such finding was evident from results of several studies indicated no significant differences were observed between educated and uneducated women with toxoplasmosis like: Al-Griari (6) in Diyala, Ertug *et al.* (12) and Sert *et al.* (13) in Turkey, While Mohammed (3) found the occurrence of the disease was higher among uneducated people than educated ones. In United States it was found that the seroprevalence was significantly higher among those with education below college level in United States (14).

The study revealed that the prevalence of toxoplasmosis increases among the housewives than the employers, but was significantly different by LAT. The higher rate of infection among the housewives than employers may be due to direct contact with infection sources through handling contaminated raw meat or vegetables; poor hand and kitchen hygiene habits, nutrition habits; drinking raw milk or unfiltered water which was found to increase risk of *Toxoplasma gondii* seropositivity for lower and middle socioeconomic populations (15).

Significantly, higher seroprevalence of toxoplasmosis, using different serological tests, was found in pregnant women, in whom 66.02% were positive for LAT, 48.39% of the women were IgG-positive, indicating previous maternal infection, 14.29% for 2ME and 11.29% were positive for anti-*Toxoplasma* IgM antibody, and the frequency of *T. gondii* IgM immunoglobulin detected in the current study may be not attributed to old infection-a possibility suggested to Gras *et al.* (16) who declared that in up to 27% of pregnant women IgM immunoglobulin levels persist for more than two years, making it difficult to pinpoint the timing of infection.

In comparison of the present findings with results of other studies performed in Iraq and other countries, the prevalence of anti-*Toxoplasma* antibodies was found to be much lower than that recorded by Al-Jubori (10) in Kirkuk 93.22% and 61.02% for pregnant and non-pregnant women respectively using the same techniques, and Al-Griari (6) in Diyala, she found the rate of seropositivity among pregnant women 70.4%; 63% in Mosul (17); but the present rate was higher than that recorded by Kadir *et al.* (18) 36.67% in Kirkuk

The higher rate of seropositive *Toxoplasma* antibodies was found in sera of women with more than 3 abortions by the diagnostic tests. The distribution of *Toxoplasma* seropositivity in women with more than 3 abortions was 100% rate of seropositivity using 2ME method, 75% by LAT, 50% for IgG antibodies and 0% for IgM type antibodies specific for *T. gondii*, but statistical analysis appears no significant difference. These results are in consistence with the findings of other investigators like: Abdel Hafez *et al.*, (19) in Jordan, who found that the prevalence of IgG seropositivity in women with

habitual abortions is two times higher than that in women with normal pregnancies 58.2% vs 26.1% which means that the infection may have long lasting effects which produce multiple abortions. Sahwi *et al.* (20) in Egypt estimated that chronic and not acute toxoplasmosis, most probably is a significant cause of repeated abortion and it may be due to the reactivation of chronic infection of the uterus, when no parasitic medication is given after the first abortion, causing rupture of the tissue cyst as a result of the enlargement in the size of the uterus during pregnancy causing liberation of the organisms and fetal infection and damage leading to pregnancy wastage. The present findings also strongly support the conclusion of several researchers in Iraq and other countries in their investigations that there is no relationship between *Toxoplasma* seropositivity and number of abortion (7, 8, 12 and 21).

In LAT and ELISA tests (IgG and IgM type antibodies), the positive serum samples for *Toxoplasma* infections were more prevalent in women suffering from primary infertility 53.19%, 46.43% and 14.29% than secondary infertility 51.28%, 32% and 4% respectively, while in 2ME, the seropositivity rate in women with secondary infertility was slightly higher than those with primary type of infertility 15.79% and 14.29% respectively. Statistical difference was insignificant, which may be related to sample size or strain of the parasite, but are in agreement with the investigations of several studies like: Janitschke *et al.* (22) in Germany, recorded 49% rate of seropositivity for *Toxoplasma*-IgG antibodies and 6.78% IgM antibodies.

The sera of 303 women were examined by Latex agglutination test, 165 were seropositive, and the rate of infection with toxoplasmosis was 54.46%. Out of 208 serum sample tested by ELISA-IgM and IgG, 19 and 78 were seropositive respectively with the rate of seropositivity 9.134% and 37.5% respectively. While among 98 cases tested by 2ME method, only 14 were positive and the rate was 14.29%. The Latex agglutination method revealed the high rate of infection followed by ELISA-IgG, 2ME and ELISA-IgM.

The compound 2 mercapto-ethanol (2ME) is used as reductant of sulphur dioxide which binds the five units of IgM antibody and thus leads to its destruction and lysis. The addition of this compound to serum, the IgM will lose its ability to agglutinate antigen while IgG antibodies will remain unaffected, therefore it is possible to diagnose the chronic form of infection (23).

The prevalence of positive 2ME in this study 14.29% was higher than that recorded in Kirkuk 7.83% (10) and lower than that reported in Erbil city 54.45% (24). These findings could be explained by the fact that the group examined consisted of healthy persons, and IgG positive persons were infected with latent toxoplasmosis without a persistence of IgM antibodies after acute infection in the past.

It is recommended to screen young females of reproductive age before marriage for *Toxoplasma* antibodies using ELISA (IgM and IgG) and treat them.

References:

- 1- Roberts, T.; Murrell, K.D. and Marks, S. Economic losses causal by foodborne parasitic disease. *Parasitol. Today*, 1994, 10: 419-423.
- 2- Elsheikha, H.M. Congenital toxoplasmosis: Priorities for further health promotion action. *Pub. Health*, 2008, 122(4): 335-353.
- 3- Mohammed, H.H. Serological detection of antitoxoplasma antibodies in apparently healthy blood donors in Nineveh province. MSc Thesis, College of Medicine, Mosul Univ., 2012.
- 4- MacPherson, J.M. and Gajadhar, A.A. Sensitive and specific polymerase chain reaction detection of *T. gondii* for veterinary and medical diagnosis. *Can. J. Vet. Res.*, 1993, 57: 45-48.
- 5- Milton, J.S. and Tsokos, J.O. Statistical methods in the biological and health sciences. McGraw-Hill international book. Company, 1983.
- 6- Al-Griari, A.J.A. A seroepidemiological study of toxoplasmosis in Diyala province/Iraq. M.Sc. Thesis, College of Education, Diyala Univ., 2007.
- 7- Kadir, M.A. and Khana, T.Q. Seroprevalence of Toxoplasmosis among women in Sulaimani City. *Tikrit J. Pharmac. Sci.*, 2009, 5: 1-10.
- 8- Ali, S.E. Seroprevalence of toxoplasmosis among women in Kalar-Garmian region, Iraqi Kurdistan. M.Sc. Thesis, College of Education/Kalar, Sulaimania Univ., 2008.

- 9- Al-Kaysi, A.M.J. Toxoplasmosis among random sample of Iraqi men and premature infants with certain immunological aspects. men and premature infants with certain immunological aspects. MSc. Thesis, College of Science, Al-Mustansiriyah Univ., 2001.
- 10- Al-Jubori, A.M. Serological study of Toxoplasmosis in Kirkuk Province. M.Sc. Thesis, College of Health and Medical Technology, Foundation of Technical Education, Baghdad Univ., 2005.
- 11- Ramsewak, S.; Gooding, R.; Ganta, K.; Seepersadsingh, N. and Adesiyun, A.A. Seroprevalence and risk factors of *Toxoplasma gondii* infection among pregnant women in Trinidad and Tabago. *Rev. Panam Salud Publica/Pan. Am. J. Public Health*, 2008, 23(3): 164-170.
- 12- Ertug, S.; Okyay, P.; Turkmen, M. and Yuksel, H. Seroprevalence and risk factors for *toxoplasma* infection among pregnant women in Aydin province, Turkey. *BMC Public Health*, 2005, 5: 66-71.
- 13- Sert, M.; Ozbek, S.; Paydas, S. and Yaman, A. Is there any relationship between *toxoplasma* infection and reactive arthritis ? *J. Postgrad. Med.*, 2007, 53(1): 14-16.
- 14- Jones, J.L.; Kruszon-Moran, D.; Wilson, M.; McQuillan, G.; Navin, T. and McAuley, J.B. *Toxoplasma gondii* infection in the United States: Seroprevalence and risk factors. *Am. J. Epidemiol.*, 2001, 154(4): 357-365.
- 15- Bahia-Oliveira, L.M.; Jones, J.L.; Azevedo-Silva, J.; Alves, C.C.; Orefice, F. and Addiss, D.G. Highly endemic, waterborne toxoplasmosis in north Rio de Janeiro state, Brazil. *Emerg Infect. Dis.*, 2003, 9: 55-62.
- 16- Gras, L.; Gilbert, R.E.; Wallon, M.; Peyron, F. and Cortina-Borja, M. Duration of the IgM response in women acquiring *Toxoplasma gondii* during pregnancy: implications for clinical practice and cross-sectional incidence studies. *Epidemiol. Infect.*, 2004, 132: 541-548.
- 17- Abdulla, B.A. and Omer, F.H. Exploration of the relationship between toxoplasmosis and brucellosis in high risk pregnancies in Mosul/Iraq. *J. Tikrit Pure Sci.*, 2007, 12(1): 39-43.
- 18- Kadir, M.A.; Ghalib, A.K. and Othman, N.F. The significance of modified Latex agglutination with 2-Mercapto-Ethanol(2ME) test in differentiation between acute and chronic toxoplasmosis. *Assiut. Med. J.*, 2007, 31(1): 1-6.
- 19- Abdel Hafez, S.K.; Shbeeb, I.; Ismail, N.S. and Abdel Rahman, F. Serodiagnosis of *Toxoplasma gondii* in habitually aborting women and other adults from North Jordan. *Folia Parasitol.*, 1986, 33: 7-13.
- 20- Sahwi, S.Y.; Zeki, M.S.; Haiba, N.Y.; Elsaid, O.K.; Anwar, M.Y. and AbdRabbo, S.A. Toxoplasmosis as a cause of repeated abortion. *J. Obstet. Gynecol.*, 1995, 21(2): 145-148.
- 21- Razzak, A.H.; Wais, S.A. and Saeid, A.Y. Toxoplasmosis: the innocent suspect of pregnancy wastage in Duhok, Iraq. *East. Mediterr. Health J.*, 2005, 11(4): 625- 632.
- 22- Janitschke, K.; Stoffels, G.; Schleyer, K. and Kentenich, H. Screening for *Toxoplasma* infection before infertility treatment. *Geburtshilfe und Frauenheikunde*, 2002, 62(5): 465-469.
- 23- Desmots, G. and Remington, J.S. Direct agglutination test for diagnosis of *Toxoplasma* infection: method for increasing sensitivity and specificity. *J. Clin. Microbiol.*, 1980, 11: 562-568.
- 24- Bakir, H.M. Seroprevalence of *Toxoplasma gondii* agglutinin in sera of women with history of abortion and control in Erbil city. M.Sc. Thesis, college of Medicine, Salahaddin Univ., 2002.