Flood Routing Model Using Genetic Expression Programing

Onen Fevzi
Ok Oral Safak
Engineering Faculty, Civil Engineering Department, Dicle University
Diyarbakır, TURKEY

Abstract

A change depending on the time of the flood wave moving in a stream using flood routing approach is examined. Flood routing of flood discharge along the river with their account and calculating the changes in the water level of flood protection structure size is determined to safety. The aim of this study, Sutculer flood event will be modeled by Genetic Expression Programing (GEP) method. The GEP method makes use of few hydrologic parameters such as inflow, outflow, and time. Simulation results indicate that the proposed a predictive model is an appropriate for the flood routing. Case study is presented to demonstrate that the GEP model is an alternative in implementation of the Muskingum model.

Keywords: Flood Routing, Muskingum Model, GEP

Introduction

The damages caused by floods in terms of loss of life, property and economic loss due to disruption of economic activity are very high. Flood peak values are required in the design bridges, culvert waterways, spillways for dams, and estimation of scour at a hydraulic structure. Flood routing is important in the design of flood protection measures in order to estimate how the proposed measures will affect the behavior of flood waves in rivers so that adequate protection and economic solutions can be found. Flood routing is used in flood forecasting, flood protection, reservoir design, and design of spillway and outlet structures.

In the past few years, the applications of artificial intelligence methods have attracted the attention of many investigators. Many artificial intelligence methods have been applied in various areas of civil, geotechnical and environmental engineering. Ferreira (2001) suggested gene-expression programming as a new adaptive algorithm for solving problems. Sivapragasam et al. (2008) used genetic programing approach for flood routing in natural channels. Chu (2009) predicted the Muskingum flood routing model using a neuro-fuzzy approach. Azamathulla et al. (2011) used gene-expression programming for the development of a stage-discharge curve of the Pahang River. Karahan (2012) predicted Muskingum flood routing parameters using spreadsheets. Onen (2014) predicted penetration depth in a plunging water jet using soft computing approaches. Onen (2014) predicted scour around a side weir in curved channel using GEP. Karahan et al. (2015) presented a new nonlinear Muskingum flood routing model incorporating lateral flow. Luo et al (2016) presented evaluation and Improvement of Routing Procedure for Nonlinear Muskingum Models. In recently, Bagatur and Onen (2016) have presented development of predictive model for flood routing using genetic expression programming

The objective of this current study is to develop a model for prediction of flood routing in natural channels using GEP method. The performance of the models is evaluated by two goodness-of-fit measures, namely the root-mean-square error (RMSE) and the determination coefficient (R^2). The used GEP model approach is evaluated using hydrograph example and discussed with the observed results.

Genetic expression programming

Gene expression programming (GEP) is an algorithm based on genetic algorithms (GA) and genetic programming (GP). This algorithm develops a computer program encoded in linear chromosomes of fixed-length. The main aim of GEP is to develop a mathematical function using a set of data presented to GEP model. For the mathematical equation the GEP process performs the symbolic regression by means of the most of the genetic operators of GA. The process starts with the generation of the chromosomes of a certain number of individuals (initial population). Then these chromosomes are expressed and the fitness of each individual is evaluated against a set of fitness cases. Then, the individuals are selected according to their fitness to reproduce with modification. These new individuals are subjected to the same developmental processes such as expression of the genomes, confrontation of the selection environment, selection, and reproduction with modification. The process is repeated for a certain number of generations or until a good solution is found (Ferreira, 2001, 2004, 2006).

The two main elements of GEP are the chromosomes and expression trees (ETs). The chromosomes may be consisted of one or more genes which represents a mathematical expression. The mathematical code of a gene is expressed in two different languages called Karva Language such as the language of the genes and the language of the expression trees (ET). The GEP genes composed of two parts called the head and tail. The head includes some mathematical operators, variables and constants and they are used to encode a mathematical expression. Terminal symbols which are variables and constants are included in the tail. If the terminal symbols in the head are inadequate to explain a mathematical expression, additional symbols are used. The flowchart of GEP is given in Fig. 1 (Onen, 2014).

In GEP method, the main operators are the selection, transposition, and cross-over (recombination). The chromosomes are modified to get better fitness score for the next generation by means of these operators. At the beginning of the model constructions, the operator rates which are specified show a certain probability of a chromosome. In common, recommended mutation rate is ranging from 0.001 to 0.1. Furthermore, recommended transposition operator and cross-over operator are to be 0.1, and 0.4, respectively.

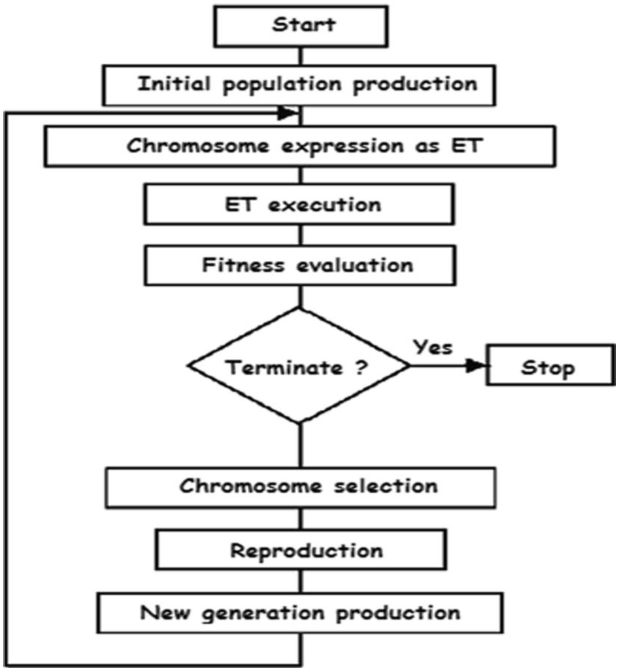

Figure 1. Genetic-expression programming (GEP) algorithm

To generate the mathematical function for the prediction of flood routing was the main aim of development of GEP models. For that reason, a development of GEP model was realized. The GEP model has two input parameters (inflow and time).

There are five major steps in preparing to use gene expression programming, and the selection of the fitness function is the first step. For this problem, it is measured the fitness f_i of an individual program i by the following expression:

$$f_i = \sum_{j=1}^{c_i} (M - |C_{(i,j)} - T_j|) \qquad (1)$$

Where M=range of selection; $C_{(i,j)}$ =value returned by the individual chromosome i for fitness case j (out of C_t fitness cases); and T_j=target value for fitness case j. If $|C_{(i,j)} - T_j|$(the precision) is less than or equal to 0.01, then the precision is equal to zero, and $f_i=f_{max}=C_iM$. For our case, we used an M=100 and, therefore, f_{max}=1,000. The advantage of this kind of fitness function is that the system can find the optimal solution for itself (Ferreira, 2001).

The second major step consists in choosing the set of terminals T and the set of function F to create the chromosomes. In this problem, the terminal set consists obviously of the independent variable, i.e. $Q = f\{I,T\}$. The choice of the appropriate function set is not so obvious, but a good guess can always be done to include all the necessary functions. In this case, we used the four basic arithmetic operators (+, -, *, /), and some basic mathematical functions (1/x, x^2, $x^{1/2}$).

The third major step is to choose the chromosomal architecture, i.e. the length of the head and the number of genes. The fourth major step is to choose the linking function. And finally, the fifth major step is to choose the set of genetic operators that cause variation and their rates. It is used a combination of all genetic operators (mutation, transposition, and recombination) with parameters of the optimized GEP model (Guven and Gunal, 2008).

This major step is to choose the chromosomal architecture, i.e. the length of the head and the number of genes. After several trials, length of the head, $h = 8$, and three genes per chromosome were found to give the best results for GEP models. The sub-ETs (genes) of GEP were linked by multiplication. Finally, a combination of all genetic operators was used as the set of genetic operators. Parameters of the training of the GEP models are given in Table 1.

Table 1. Parameters of the optimized GEP model

Parameter	Description of parameter	Setting of parameter
P_1	Chromosomes	30
P_2	Fitness function error type	R^2
P_3	Number of the genes	3
P_4	Head size	8
P_5	Linking function	*
P_6	Function set	+, -, *, /,1/X, $X^{1/2}$, $X^{1/3}$, X^2, X^3
P_7	Mutation rate	0.044
P_8	One-point recombination rate	0.3
P_9	Two-point recombination rate	0.3
P_{10}	Inversion rate	0.1
P_{11}	Transposition rate	0.1

The performance of GEP models is validated in terms of the common statistical measures coefficient of determination (R^2) and root-mean-square error (RMSE).

$$R^2 = \left(\frac{\sum Q_x Q_y}{\sqrt{\sum Q_x{}^2 \sum Q_y{}^2}} \right)^2 \qquad (2)$$

$$\text{RMSE} = \left[\frac{\sum (\text{Q}_\text{o} - \text{Q}_\text{p})^2}{\text{n}} \right]^{1/2} \qquad (3)$$

Where $Q_x=(Q_o-Q_{om})$; $Q_y=(Q_p- Q_{pm})$; Q_o=observed values; Q_{om}=mean of Q_o; Q_p=predicted value; Q_{pm}=mean of Q_p; and n=number of samples

Introductıon basın, gaugıng statıon and sutculer flood

Situated on the western Taurus zone of the region, east Dedegöl, located in the south Kuyuluk mountain elevations. Aksu stream flowing and bridge forming deep canyons on the Taurus belt, it reaches the Mediterranean Sea. Sutculer district's annual average rainfall is 916.7 mm. The daily maximum rainfall was measured 212 mm in September 1990.

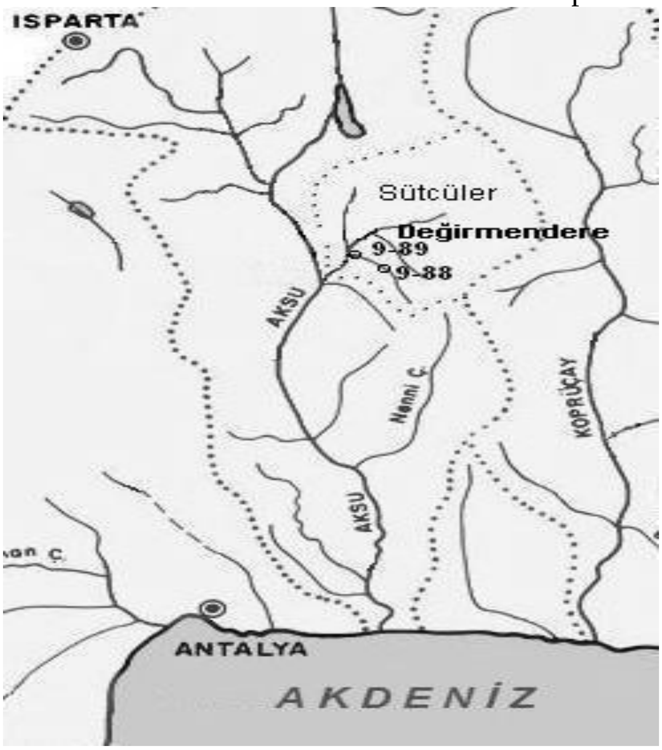

Figure 2. The Flow gauging stations in Aksu River and floodplain

Rainfall started as fraught around Sutculer near town center, and then rainfall turned to rain, and continued without interruption for 4 hours too severe. With the start of precipitation in the form of fraught, delayed flow and deposited with the conversion of rainfall caused a rapid stream. Among them 10 km away with two stations is located where Değirmendere excessive swelling results during the flood, stuck on which bridges are under water, bedside damaged farms producing fish involved and occurred four casualties (Ülke, 2006).

Local falling rainfall is 111.4 mm for 4 hours; this value corresponds to a 25-year time-intensity-iteration value. The water levels in river have found 6.00 m during floods and the discharge was reached 206 m^3/s. Flood routing calculations were performed between 9-88 and 9-89 numbered stations (Fig.2). Properties of those stations are presented in Table 2. Flow values measured in the numbered stations 9-88 and 9-89 are shown in Fig 2.

Table 2. Properties of Station

Station number	Elevation	Latitude	longitude	Rainfall area km^2
9-88	750	37° 28' 38.8"	30° 58' 41.40"	131
9-89	320	37° 27' 50.1"	30° 54' 29.90"	314

Studies and developing of gep models

The GEP model approaches make use of few hydrologic parameters (inflow, outflow and time) as Muskingum model. In developing of GEP model, one case study will be considered here (Fig.3). This case study is based on the inflow and outflow hydrographs exhibiting multiple peaked discharge characteristics (Viessman and Lewis, 2003). This example is hypothetical and probably does not relate to any real-life observation.

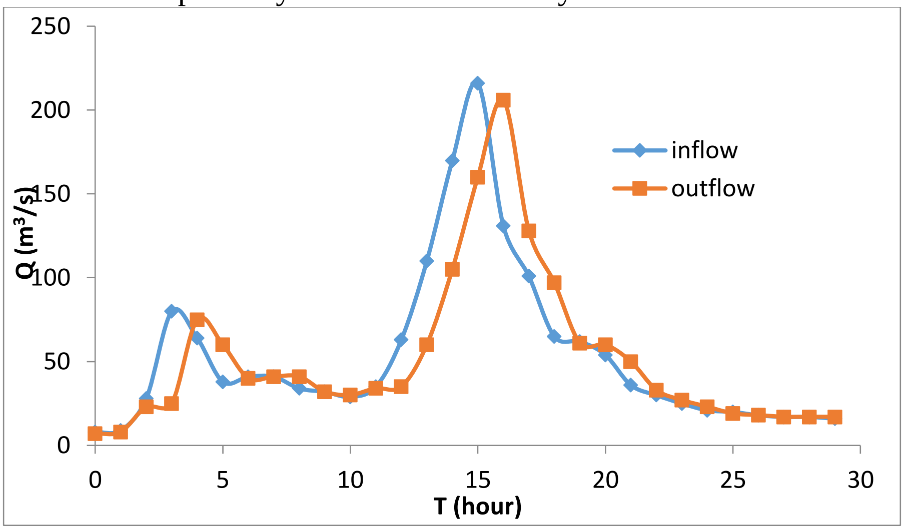

Figure 3. Input and output of Sutculer flood hydrograph

Flood routing procedures may be classified as either hydrological or hydraulic. Hydrological methods use the principle of continuity and a relationship between discharge and the temporary storage of excess volumes of water during the flood period. Hydraulic methods of routing involve the numerical solutions of either the convective-diffusion equations or the one dimensional Saint-Venant equations of gradually varied unsteady flow in open channels. The hydraulic methods generally describe the flood wave profile more adequately when compared to hydrological methods, but practical application of hydraulic methods are restricted because of their high demand on computing technology, as well as on quantity and quality of input data. In practical applications, the hydrological routing methods are relatively simple to implement and reasonably accurate. An example of a simple hydrological flood routing technique used in natural channels is the Muskingum flood routing method (Gill, 1978; Tung, 1985).

In this paper, gene expression programming (GEP) technique is evaluated as an alternative solution against to Muskingum model. Thus, GEP models will be developed without Muskingum flood routing parameters and model. The proposed models include only inflow (I), outflow (Q) and time (T) parameters as model approaches. After all the parameters are defined, the models are simulated. The powerful soft computing software package GeneXproTools 4.0 (Ferreira, 2006) was used to develop GEP-based models for flood routing prediction in this work. This program provides a compact and explicit mathematical expression for flood routing. The terminating criterion was the maximum fitness function, which in turn is a function of the root-mean-square error (RMSE). The program was run for a number of generations and was stopped when there was no improvement in fitness function value or coefficient of determination (R^2).

In the beginning of model studies, the program could not be obtained sufficient predictive model for multiple peaked hydrograph of case study. Therefore, the hydrograph was separated two single peaked hydrographs. Thus, effective models were obtained with three brackets for hydrographs.

The simplified analytical form of the proposed GEP model is expressed for first and second single peaked outflow hydrograph of case study as respectively:

$$Q = \left[\left(1/\left(I^{0.5} - T\right) - T - 6.59\right)\right]\left[I - 2I/\left(761T - I - 9.41\right)\right]\left[\left(2I + T - 9.98\right)^{0.5} + \left(10.46/\left(T - 9.98\right)\right)\right] \quad (8a)$$

$$Q = \left[0.74 - 1/0.74I\left(T - 5.1\right)\right]\left[126.1 + 0.15I^{3} - I\right]\left[4.11 - I^{1/3} + 16.89/\left(I + 2.90\right)\right]^{1/3} \quad (8b)$$

where I and Q are the amounts of inflow and outflow respectively at time T. In case study, Eq. (8a) is valid for $0<T\leq 10$ and Eq. (8b) is valid for $10<T\leq 29$.

As seen from Figure (Fig.4) GEP model performs extremely well in routing the multi-peaked hydrograph for case study

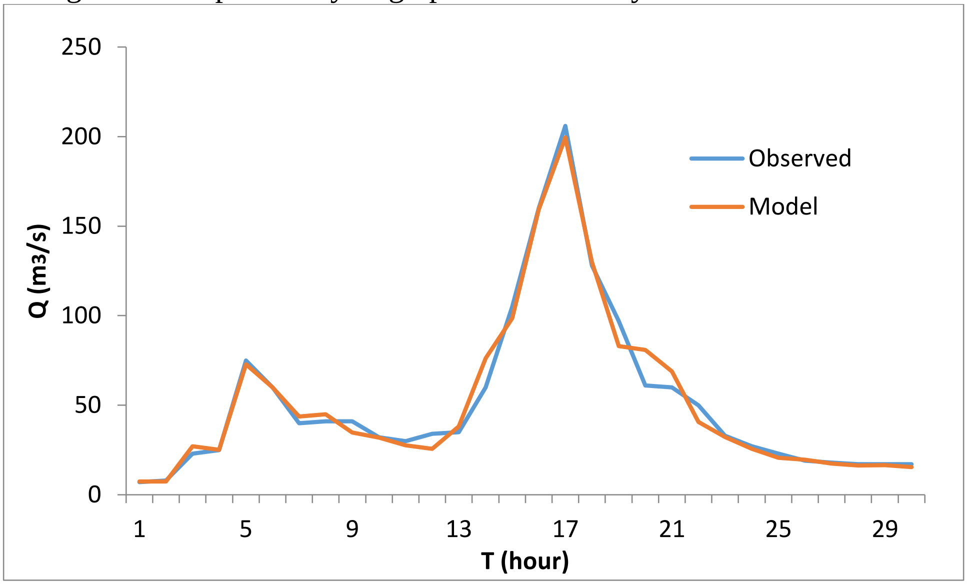

Figure 4. Comparison of observed and predicted outflow values for case study

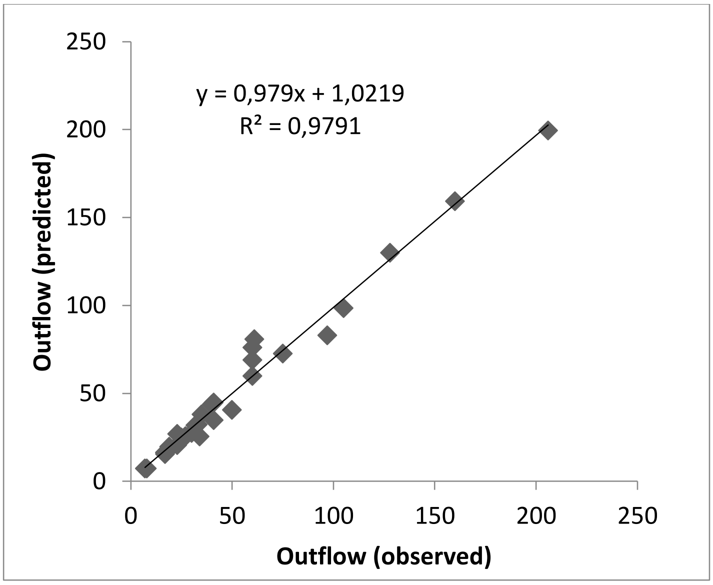

Figure 5. Comparison of observed values versus predicted outflow values for case study

The proposed GEP approach gives good results (R^2=0.979 and RMSE=6.56) compared to the existing predictor for case study (Fig.5). Peak is predicted accurately (199.56 m^3/s) and without Muskingum model.

Conclusion

This study demonstrates the potential of the GEP model for flood routing in natural channels. Therefore, the GEP approach can be used to derive a new model for the prediction of flood routing in natural rivers. The proposed GEP models are tested for the data sets given in literature and it has been shown that the model results are good agreement with the observation values. The comparison shows that the model expressions have the least root mean square error and the highest coefficient of determination. The GEP model predicts the outflow, with an R^2=0.979 and RMSE=6.56 for case study. The study suggests that GEP techniques can be successfully used in modeling flood routing from the available observed data.

References:

Azamathulla H.M, Ghani A.A, Leow C.S, Chang C.K, ‘’Gene-Expression programming for the development of a stage-discharge curve of the Pahang river,’’ Water Resour. Manage, Vol. 25, 2011, pp. 2901-2916.

Bagatur and Onen, ‘’Development of Predictive Model for Flood Routing Using Genetic Expression Programming,’’ Journal of flood risk management, 2016, DOI: 10.1111/jfr3.12232 (in press).

Chu, H. J. (2009), The Muskingum flood routing model using a neuro-fuzzy approach. KSCE Journal of Civil Engineering, *13*(5), 371-376.

Ferreira C, ‘’Gene expression programming: A new adaptive algorithm for solving problems,’’ Complex Systems Vol. 13 (2), 2001, pp. 87-129.

Ferreira C, ‘’Gene expression programming and the evolution of computer programs,’’ in: L.N. de Castro. F.J. Zuben (Eds.). Recent Developments in Biologically Inspired Computing. Idea Group Publishing, 2004, pp. 82–103.

Ferreira C, ‘’Gene-expression programming: Mathematical modeling by an artificial intelligence,’’ Springer. Berlin, 2006,

Gill MA, ‘’Flood routing by the Muskingum method,’’, Journal of Hydrology Vol. 36, 1978, pp. 353–363.

Guven A, Gunal M, ‘’Genetic programming approach for prediction of local scour downstream hyraulic structures,’’ J Irrig and Drain Eng Vol. 134(2), 2008, pp. 241-249.

Karahan, H. (2012). Predicting Muskingum flood routing parameters using spreadsheets. *Computer Applications in Engineering Education*, *20*(2), 280-286.

Karahan, H., Gurarslan, G., & Geem, Z. W. (2015). A new nonlinear Muskingum flood routing model incorporating lateral flow. *Engineering Optimization*, *47*(6), 737-749.

Luo, J., Yang, X., Xie, J. (2016). Evaluation and Improvement of Routing Procedure for Nonlinear Muskingum Models. *International Journal of Civil Engineering*, *14*(1), 47-59.

Sivapragasam, C, Maheswaran, R, Veena, V, ‘’Genetic programing approach for flood routing in natural channels,’’ Hydrol. Process, Vol. 22, 2008, pp. 623-628.

Onen F, ‘’Prediction of penetration depth in a plunging water jet using soft computing approaches,’’ Neural Computing and Application, Vol. 25, 2014, pp. 217–227.

Onen F, ‘’GEP prediction of scour around a side weir in curved channel,’’ Journal of Environmental Engineering and Landscape Management Vol. 22 (3), 2014, pp. 161–170.

Tung Y.K, ‘’River flood routing by nonlinear Muskingum method,’’ Journal of Hydraulic Engineering, ASCE Vol. 111, 1985, pp. 1447–1460.

Ülke, A, ‘’Muskingum Metodu Kullanılarak Taşkın Ötelenmesi,’’ Yüksek Lisans Tezi, Süleyman Demirel Üniversitesi, Fen Bilimleri Enstitüsü, 2003.

Viessman W, Lewis G.L, ‘’Introduction to hydrology,’’ Pearson Education. Inc., Upper Saddle River. New Jersey. USA, 2003.