Contribution of TRMM 3B42 Data to Improve Knowledge on Rainfall in the Kayanga/Geba River Basin (Republic of Guinea, Senegal and Guinea-Bissau)

Saly Sambou
Honore Dacosta
Faculté des Lettres et Sciences Humaines, Département de Géographie, Université Cheikh Anta Diop (UCAD), Dakar, Sénégal
Abdoulaye Deme
Laboratoire LSAO "Laboratoire des Sciences de l'Atmosphère et de l'Océan", Université Gaston Berger (UGB), Saint-Louis, Sénégal
Ibrahima Diouf
NOAA Center for Weather and Climate Prediction 5830 University Research Court, College Park, Maryland, USA

Abstract
        The use of Tropical Rainfall Measuring Mission (TRMM) data is an option for counteracting challenge of the lack of ground based observations, particularly in Kayanga/Gêba. This paper undertakes validation of monthly TRMM rainfall estimates before using it to understand the spatial and temporal variability in the Basin. This validation based on application of statistical study, made it possible to obtain interesting results with correlation coefficients varying from 0.92 to 0.96 and Nash indices close to 1. The analysis of the seasonal rainfall pattern shows consistence with ground based observations. The study of the annual cycle reveals that their interannual variability is similar to that of ground based observations. Finally, the interpolation of average monthly rainfall in the basin highlights the North-South rainfall gradient, which shows that the South is wetter than the North, with differences more pronounced in August and September.

Keywords: Rainfall estimate, TRMM, rainfall variability, Kayanga/Gêba River, transboundary basin

Introduction

The analysis of the spatial and temporal rainfall variability requires long and complete time series. However, the study of the rainfall in the tropical regions is usually hindered by the challenge of insufficient data, due to sparse ground based observation network and many gaps in the historical time series (Dubreuil and al., 2004; Carbonnel and Hubert, 1992; Koudou and al., 2015). In addition, the access to the available data at a finer resolution is difficult due to high costs. At the Kayanga/Geba River watershed level, the ground measurement network reveals a very uneven spatial distribution with some sites older than others and a more or less incomplete data report. This leads to the existence of high variability of the historical data. Rainfall are very little studied because of the transboundary nature of the Basin and gaps in the recent years of the ground monitoring network imply difficulties to study the rainfall over the basin; for both Senegalese and Bissau-Guinean stations. This fragmentary information does not facilitate a spatio-temporal tracking of the rainfall in the Basin. Recently, the availability of satellite rainfall estimates is an alternative option which however, requires historical knowledge that will serve as a reference for the validation of their quality. Among these various estimated rainfall data, we have data from TRMM Spectroradiometer, the first satellite devoted to precipitation measurements (Simpson and al. 1996).

The TRMM Spectroradiometer satellite was launched on November 27, 1997 and it was supported by the cooperation between the United States of America and Japan in order to measure rainfall and energy exchanges in the tropical and subtropical regions (NASA, 2007). Its method is based on the microwaves retro diffusion by rainfall and ice. It is one of the most famous satellite products and the data is easily accessible. The interest of these data is that they respond to the real-time rainfall problem with a wide spatial coverage from latitudes 50° S up to 50° N, and in a hourly time steps (Huffman and al. 2007).

This paper attempts to validate the TRMM 3B42 dataset in order to assess its comparability with ground based rainfall observations. The results will inform the decision for use in filling gaps in the ground based observation time series and consequently its use in improving knowledge on the spatiotemporal distribution of the rainfall in the Kayanga/Geba river Basin. The satellite rainfall estimates under investigation is validated with the ground based observations on a monthly and annual timescale.

Data and methods
Presentation of the study area

Located between the Republic of Guinea (1.3%), Senegal (34.3%) and Guinea-Bissau (64.4%), the watershed of the Kayanga/Geba River covers an area of about 12440 km² (Figure 1).

Starting from the western part of the Badiar plateau in the Republic of Guinea, the River flows towards a North-West direction in a distance of around 10 kilometers. In Senegal, the River follows a Westward and then a South-West direction until Guinea-Bissau in a distance up to 150 km. The River is so-called Rio Geba as from Guinea-Bissau region (Brunet-Moret, 1970). The South Sudanese type of climate is characterized by a rainy season from May to October and a dry season covering the rest of the year. The latitudinal shape of the Basin has rainfall contrasts with annual rainfall average about 1000 mm in the North and 1500 mm in the South. Monthly mean temperatures range from 23.7° C in December to 31.7° C in May. Monthly relative humidity peaks in August at 82.0% and then drops to 48.1% in April. The sunstroke also varies between 5.89 hours (in August) and 9.04 hours (in April). In this Basin, agriculture remains the main activity and source of income of populations who are predominantly rural. Other socio-economic activities such as livestock, inland fishing, arboriculture, market gardening and trade are also developed there. However, they are subject to hydro-rainfall variations. Consequently, two hydro-agricultural dams have been erected in the Senegalese part in order to ensure a better water resources control for irrigated agriculture in the rainy and dry seasons. In Guinea-Bissau, the country has a very ancient rice tradition which highly depends on rainfall. Small hydro-agricultural amenities have also been built to irrigate small perimeters using motor pumps from the reservoirs.
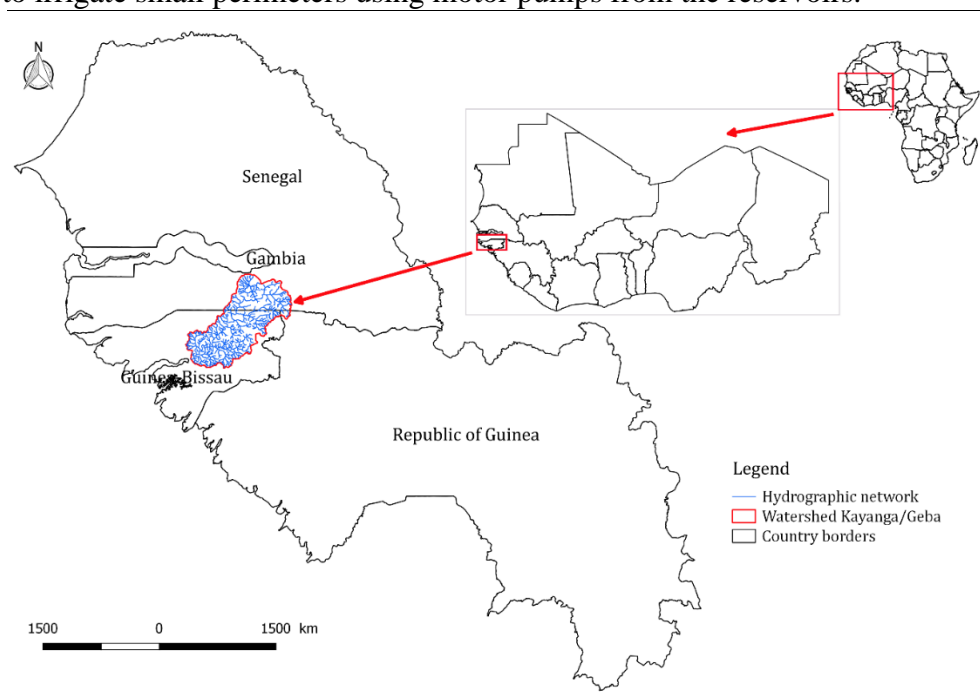

Figure 1: Location of the Kayanga/Geba River watershed

Data
Observed data

The data come from various sources: the National Agency of Civil Aviation and Meteorology (ANACIM), the Organization for the Development of the Gambia River (OMVG), the Agricultural and Industrial Development Company (SODAGRI) and the Montpellier HydroSciences laboratory. Table 1 provides an inventory of the available data. The rainfall network (Figure 2) of Kayanga/Geba Basin consists of up to 20 stations: 9 stations located inside the area of study and 11 located around the border. Most of the Guinea-Bissau stations set up in the 1950s stopped recording in the mid-1980s and 1990s with the exception of the Gabu synoptic station for which the series runs until 2005. Consequently, there is no recent rainfall data over Guinea-Bissau due to obsolete meteorological equipment, and demotivated weather observers. Also, some Senegalese stations recently set up have gaps in the series yet, less important than those mentioned previously. Thus, in this study only station data covering all or part of the 1998-2014 period are considered. They will serve as a reference for validating the quality of TRMM 3B42 data.

Table 1: Inventory of available rainfall data

Country	Station	Elevation (m)	Latitude (North)	Longitude (West)	Start of Record	End of Record
Guinea	Koundara	76	12° 29' 00"	-13° 18' 00"	January 1928	December 2000
Gambia	Basse	22	13° 18' 37"	-14° 13' 24"	January 1942	December 2014
Senegal	Bonconto	42	12° 58' 00"	-13° 57' 00"	January 1978	December 2006
	Linkering	56	12° 58' 00"	-13° 44' 00"	July 1944 January 1991	December 1959 December 1998
	Dabo	43	12° 48' 00"	-14° 29' 00"	January 1975	December 2008
	Kolda	8	12° 53' 00"	-14° 58' 00"	January 1922	December 2014
	Pakour	65	12° 43.2'	-13° 59.4'	January 1985	December 2011
	Kounkane	33	12° 56' 00"	-14° 05' 00"	January 1963 July 1964	March 1964 December 2011
	Fafacourou	27	13° 06' 00"	-14° 56' 00"	January 1962 January 1997	December 1995 December 1998
	Simenti	48	13° 02' 58"	-13° 17' 53"	January 1968	December 2008
	Velingara	38	13° 09' 00"	-14° 06' 00"	January 1932	December 2014
Guinea-Bissau	Pirada	90	12° 40' 00"	-14° 10' 00"	January 1950 January 2001	December 1997 August 2003
	Bafata	43	12° 10' 00"	-14° 40' 00"	January 1950	December 1994
	Sonaco	25	12° 24' 00"	-14° 29' 00"	January 1950 January 1997	December 1992 December 1997
	Gabu	63	12° 17' 00"	-14° 14' 00"	January 1941	December 2005

263

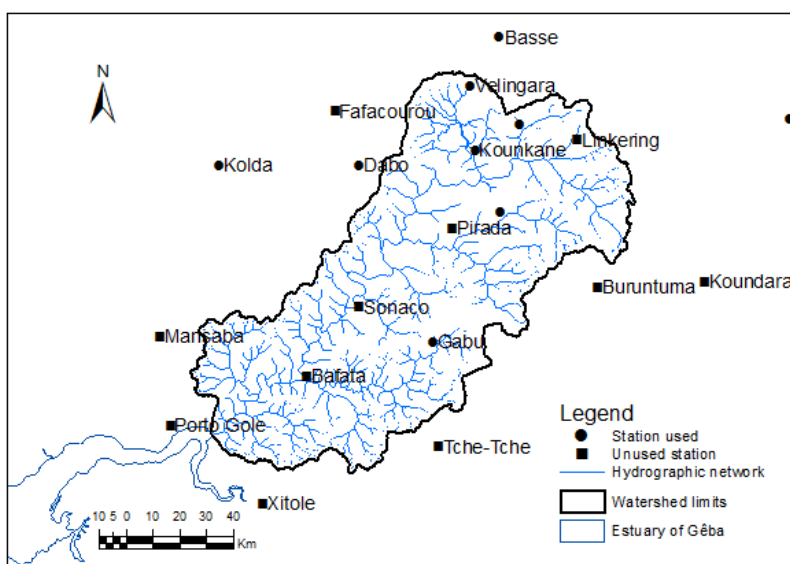| Tche-Tche	34	11° 55' 30"	-14° 12' 51"	September 1978	December 1989
Buruntuma	100	12° 28' 00"	-13° 40' 00"	January 1950	December 1992
Xitole	30	11° 44' 00"	-14° 49' 00"	January 1969	July 1995
Mansaba	43	12° 17' 32"	-15° 10' 32"	January 1950	December 1986
Porto Gole	10	11° 58' 00"	-15° 08' 00"	January 1950	December 1986

Figure 2: Spatial distribution of meteorological stations in the Kayanga/Geba Basin

TRMM 3B42 data

The TRMM satellite platform has microwave sensors (TMI, TRMM Microwave Imager), radar (PR, Precipitation Radar) and visible and infrared scanners (VIRS, Visible and Infrared Scanner) (Simpson et al. 1996). It offers many products estimating precipitation, among those TRMM 3B42 which presents characteristics more in adequacy with the needs of this study. This product comes from a series of algorithms that convert the observed data by different sensors into rainfall data. According to Jobard (2001), TRMM rainfall estimates 3B42, combining microwave and infrared data from geostationary satellites in addition to field data, prove to be more efficient. The data are in three-hour scale for a spatial resolution of 0.25° x 0.25° and available from January 1998 to the present. Although, the satellite provides important information compared with data points (rain gauges). In addition to the spatial distribution, they make it possible to know the spatiotemporal variability of the rains with fine scales. This, in the Kayanga/Geba Basin with a low density of rain gauges and sometimes incomplete surveys, is an important asset. The data can be found on the site:

http://disc.sci.gsfc.nasa.gov/precipitation/documentation
/TRMM_README/TRMM_3B42_readme.shtml/

Methods

The use of TRMM 3B42 data necessarily involves assessing its quality, first looking for statistical relationships, then assessing rainfall distribution during the year and finally their interannual variability by referring to the rainfall observed at the ground based weather stations. The validation of these operations makes it possible to characterize their spatial variability on a monthly scale.

Validation of TRMM 3B42 data

The TRMM 3B42 data available every 3 hours are pre-processed to compose daily, monthly and yearly totals. The validation follows a statistical approach based on their comparison with the small number of reference measurements rain gauges. The identified gaps on certain series make it unsafe to consider this validation on a daily scale but on a monthly and yearly scale. Quantitative and qualitative statistical criteria proposed by Laurent and al. (1998) are used in this study to verify the quality and reliability of satellite data with reference to ground observations. According to Ebert and al. (2007), they are very useful for quantifying errors and evaluation of estimates. Thus, for a reference dataset v_i (of estimation e_i) comprising n values with $i = (1, n)$ whose average is defined by $\bar{v}$ ($\bar{e}$) and the standard deviation by σ_i (σ_e), we have the correlation coefficient (R), the bias (B), the Root Mean Square Error (RMSE) and the Nash index (I):

$$R = \frac{(v_i - \bar{v})(e_i - \bar{e})}{n\sigma_i \sigma_e}$$

The correlation coefficient R (without unit) is the ratio between the covariance of observations and estimates and the product of their standard deviations. It gives an indication of the intensity and the direction of the linear relationship between the two variables (see note below).

The bias is the difference between the ground observations average and estimates by satellite. It provides an indication of the underestimation or overestimate of rainfall. The formula for computing bias is given by

$$B = \bar{e} - \bar{v}$$

The Root Mean Square Error (RMSE) is the square root of the ratio between the sum of the squared differences of the observations and estimates of the number of values n. Expressed in mm, it indicates the average magnitude of the error, while giving more weight to important errors (Delahaye, 2013) (see note below). The formula for computing RMSE is given by

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(e_i - v_i)^2}$$

The Nash-Sutcliffe index (1970) is a widely used criterion in hydrology to evaluate the quality of a model. This gives a squared difference ratio between the observations and estimates for reporting a gap squared on observations. If I = 1, the estimates are prefect and if I = 0, the estimates are equal to the average of the reference values (Arvor and al. 2008). The formula for computing Nash-Sutcliffe index is given by

$$I = 1 - \frac{RMSE^2}{\sigma_v^2}$$

Rainfall regime

The quality of the TRMM 3B42 estimates is also analyzed by considering how the rainfall changes during the year. The objective is to see if the satellite finds the annual cycle of rainfall at each station. For this, a calculation of monthly rainfall average was carried out for the two variables for all the selected stations. Their presentation in graphical form makes it possible to appreciate their quality.

Interannual variability

The validity of TRMM is also tested on a yearly scale. Thus, for each year, the annual amount is performed for both the estimates and the observations. In a context of high interannual variation of the rainfall, the aim is to verify a possible coincidence between the wetter years and the drier years of the two variables at all the stations.

Spatial distribution of monthly rainfall

Knowing the high spatial variability of rainfall in West Africa, TRMM estimates validated on a monthly scale with reference to observations were interpolated over the entire field of study for the period 1998-2014. Although, they are validated on only 9 stations with observed data, this interpolation relies on all the stations of the basin, since they belong to the same climatic zone. The period June-October because the rainy saison is mainly concentrated in these months. Thus, the average rainfall of these months were spatialized using the Krigeage interpolation method of the Surfer 11 software. It is a probabilistic method that estimates the values at the points not sampled by a combination of the data (Baillargeon, 2005). The observed values are weighted by a structure function that is derived from the data itself. This takes into account distances, values and correlations. The estimated value at one point is considered to be the product of an underlying process (Labat, 2012).

Results
Results of TRMM 3B42 data validation

Although, the time series have unequal periods, the correlation coefficients are all close to 1. This means that estimated rainfall is quantitatively close to those observed on ground for the 9 selected stations (Table 2). Figure 3 shows the dispersion of the points around the linear regression line between the observed data and those estimated by TRMM. With the exception of Dabo station (0.79), the biases are all negative, indicating an underestimation of rainfall by TRMM compared to observations. The RMSE is lower for Kounkane and Bonconto compared to the rest of the stations. The Nash indices are appreciable, they are close to 1. Overall, these statistical validation results suggest that the estimates made by TRMM, in spite of the quantification problems, better represent the rainfall station data.

Table 2: Correlation coefficient, biases, RMSE and Nash index calculated between observed and estimated rainfall

| Station	Period	Statistical criteria			
		Correlation coefficient	Bias	RMSE	Nash index
Basse	1998-2014	0.93	-9.1	41.23	0.86
Velingara	1998-2014	0.93	-2.7	41.90	0.87
Kounkane	1998-2011	0.96	-3.2	35.74	0.92
Bonconto	1998-2006	0.95	-13.08	39.38	0.88
Kolda	1998-2014	0.94	-11.4	47.43	0.87
Dabo	1998-2008	0.93	0.79	42.36	0.87
Simenti	1998-2008	0.95	- 19.5	47.57	0.84
Pakour	1998-2011	0.94	-15.3	57.08	0.86
Gabu	1998-2005	0.92	-20.91	68.92	0.81
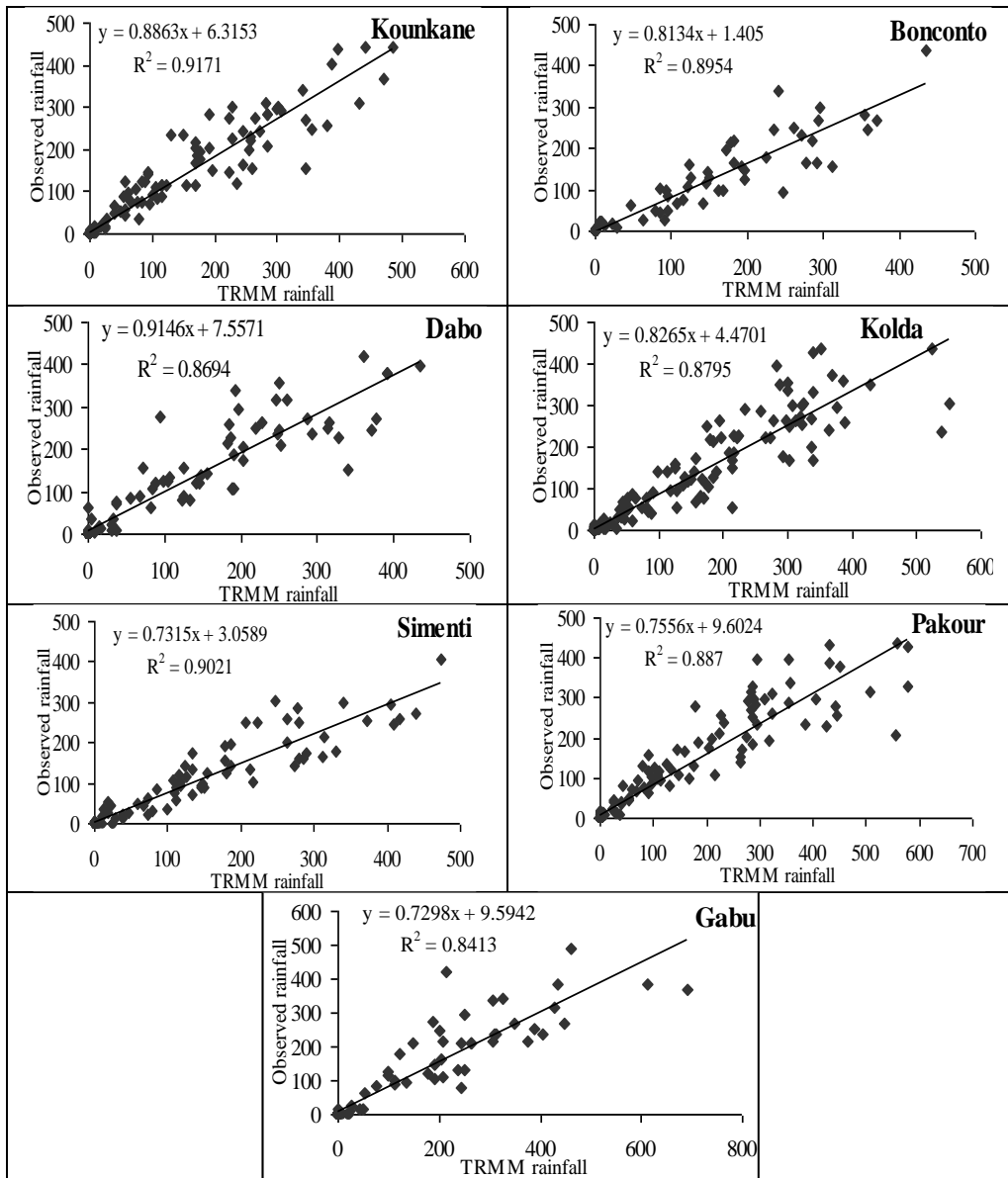
Figure 3: Regressions between observed and estimated monthly rainfall

Seasonal regime analysis

From a qualitative point of view, the TRMM estimates correctly reproduce the seasonal rainfall regime of stations of Kayanga/Geba Basin, distinguishing the rainy season and the dry season (Figure 4). We note a good coherence between the months, with a peak in August which is the wettest month at all used stations, followed by September. This seasonal comparison also helps to notice that the main difference between estimates and observations is an underestimation of rainfall by TRMM.
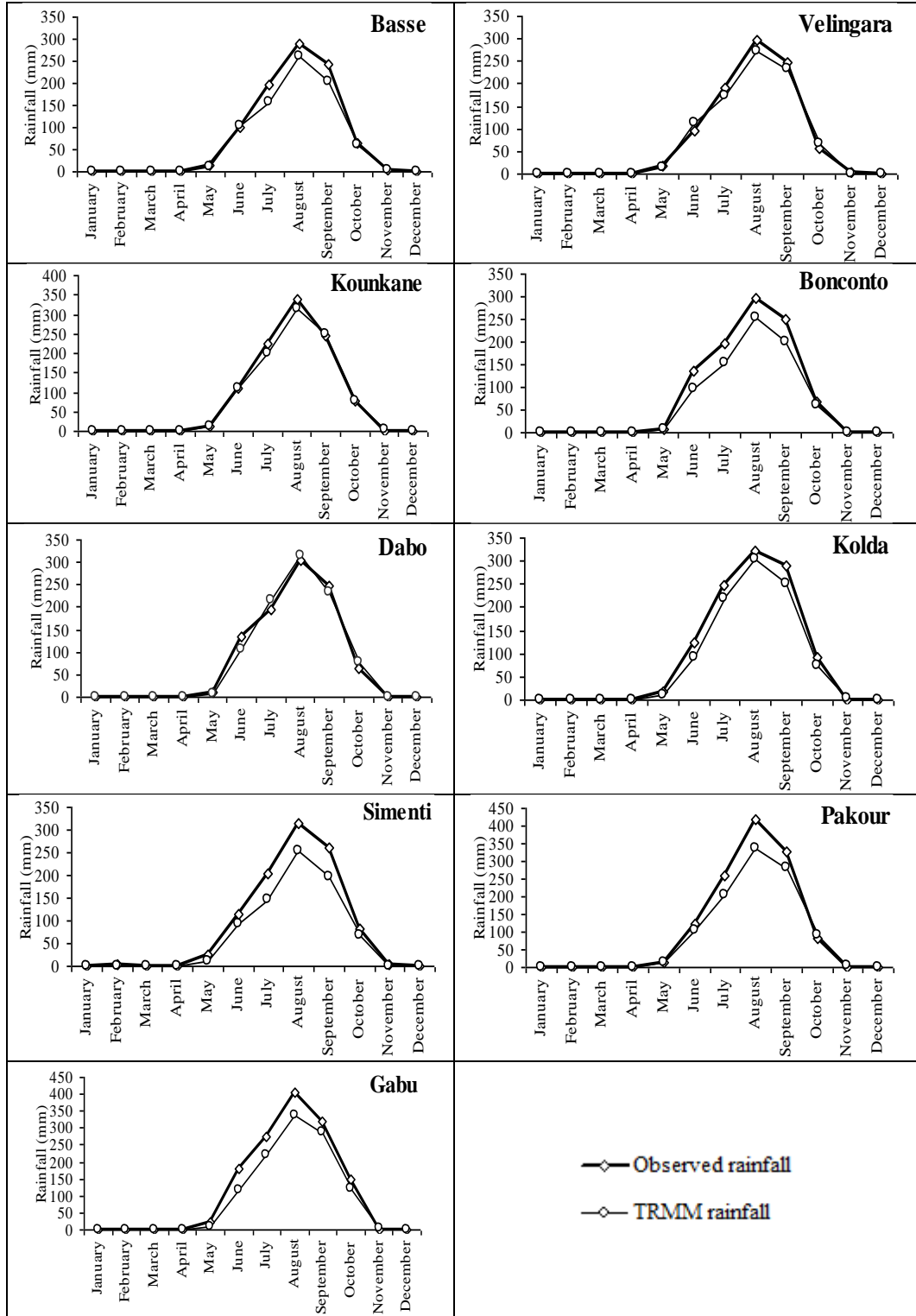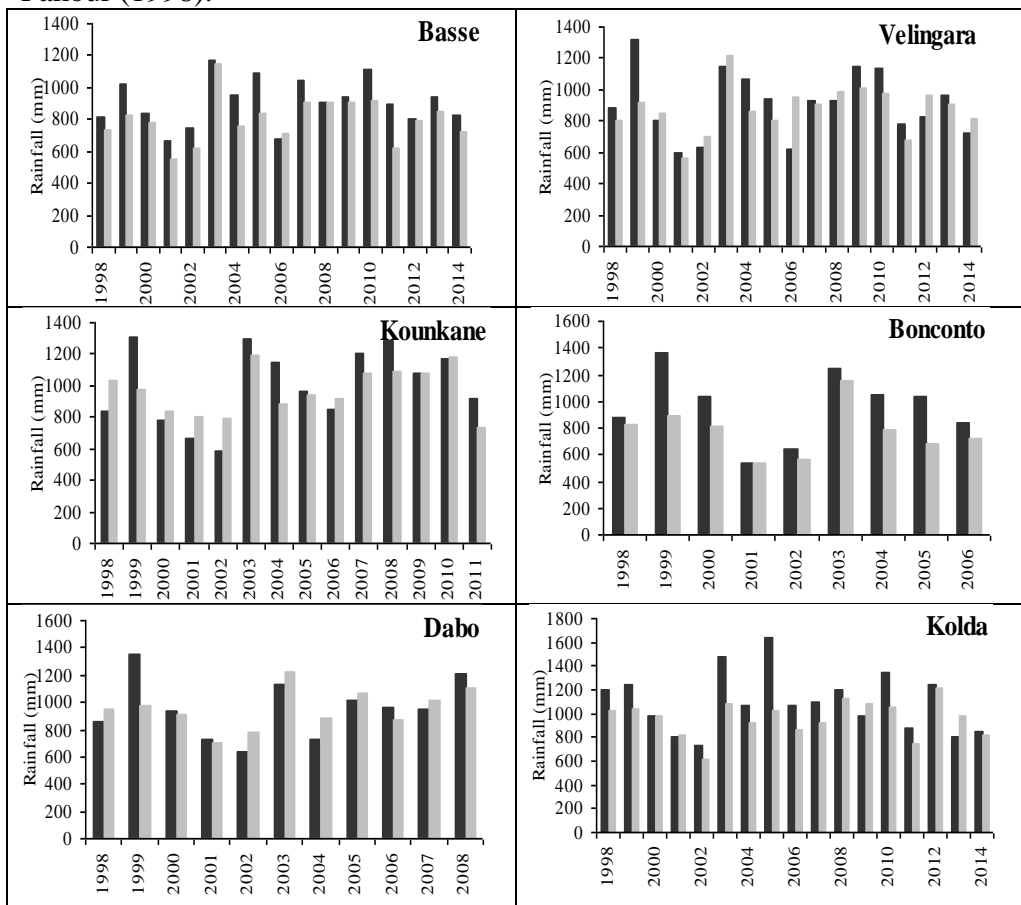
Figure 4: Distribution of monthly average observed and estimated rainfall

Interannual variability analysis

Despite the fact that the periods are very short (8 to 17 years) to be representative of the interannual variability of rainfall in the Kayanga/Geba Basin, there are significant fluctuations from one station to another. Figure 5 shows that TRMM estimates perfectly the interannual rainfall variability for all stations. We find a good agreement between observed and estimated rainfall either during the wet years or the dry years. The years 1999 and 2003 are wetter while 2001 and 2002 are drier. Moreover, we can see that the underestimation of TRMM rainfall is more obvious in 1999 (almost all the stations), 2003 (Kolda, Simenti, Pakour and Gabu), 2005 (Kolda) and 2007 (Pakour) which are wet years. However, for some stations like Basse (2006), Velingara (2000, 2002-2003, 2006, 2008, 2012, 2014), Kounkane (1998, 2000-2002, 2006), Dabo (1998, 2002-2005, 2007), Kolda (2000-2001, 2009, 2013) and Pakour (2002, 2005), there is a slight overestimation of rainfall by TRMM; but a good coherence between estimates and observations at Basse (2008), Kounkane (2009-2010), Bonconto (2001), Simenti (1998) and Pakour (1998).
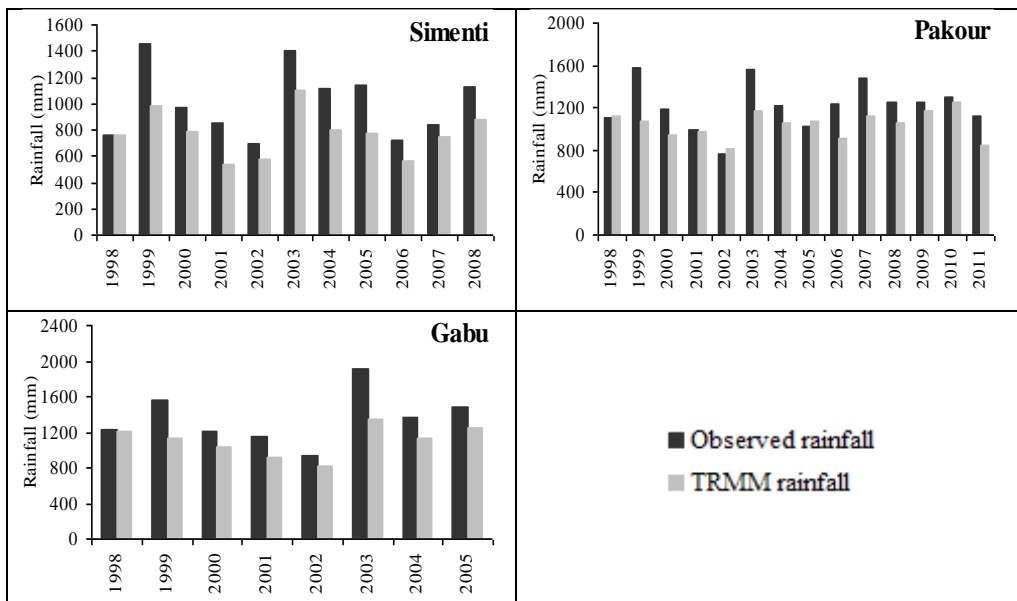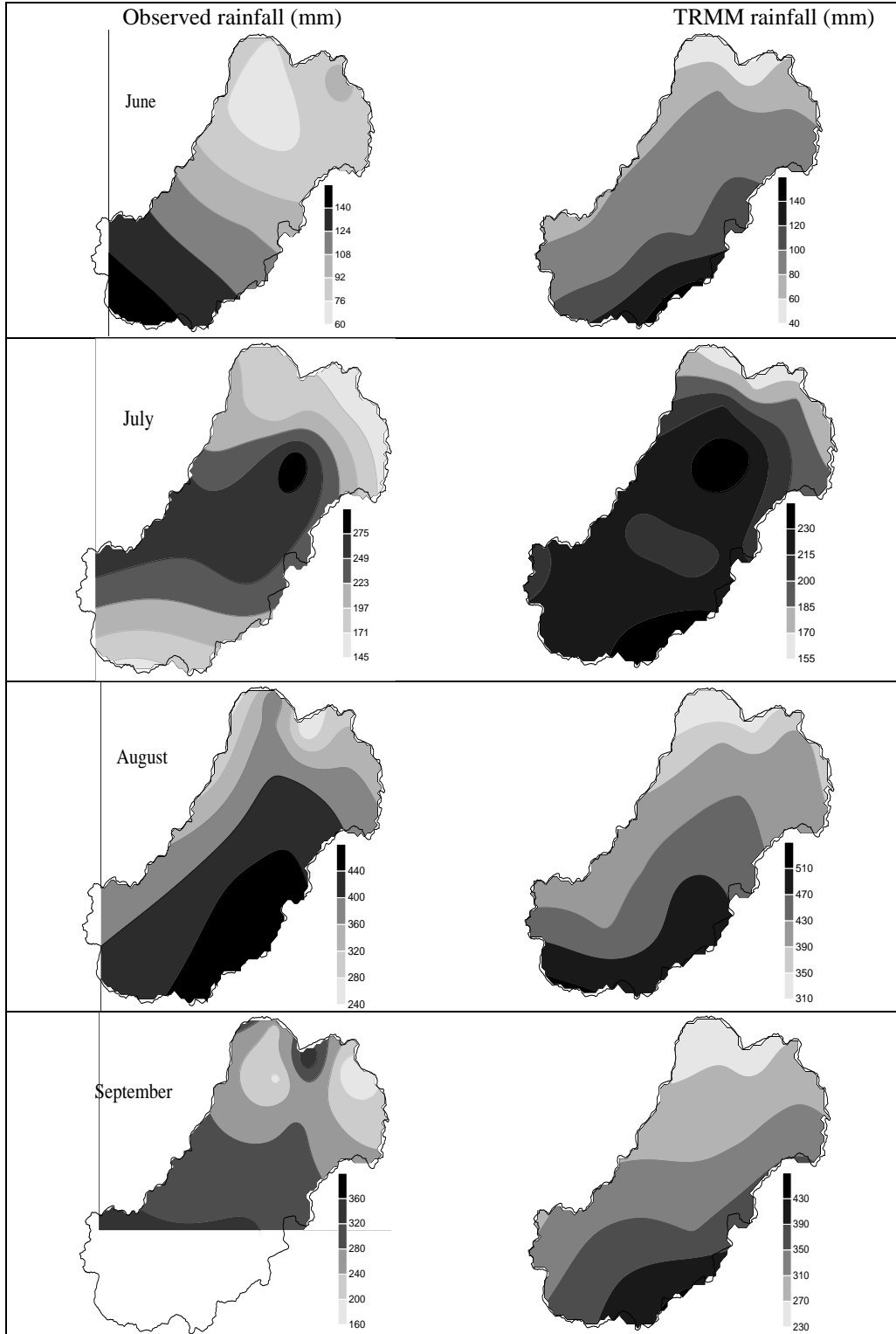
Figure 5: Interannual variability of observed and estimated rainfall

Mapping monthly rainfall average estimated by TRMM

Figure 6 shows the spatial distribution of average monthly observed rainfall (column 1) and estimated by TRMM (column 2) on the Kayanga/Geba Basin for the period 1998-2014. The inequality of the observed series and the existence of numerous gaps make the interpolation biased. From June to August, we see empty places (South-West of the Basin) that are not taken into account by the interpolation whereas in September and October it is the entire southern part of the Basin that is concerned. This situation is an obstacle for the spatial knowledge of the rainfall in the Basin. Nevertheless, the estimated data by TRMM and interpolated help by their regularity and good spatial resolution to overcome the gaps in the observations. The maps highlight, on the one hand the great spatial variability, and on the other hand the unequal distribution of rainfall heights between the selected months. They also highlight the North-South rainfall gradient, due to the latitudinal extension of the Basin, with monthly rainfall varying between 40 and 140 mm in June, 155 and 230 mm in July, 230 and 430 mm in September, 10 and 80 mm in October, North and Southeast. In August moreover, they vary between 310 mm in the North to 510 mm in the Southwest. The South of the basin remains wetter than the North. This opposition between North and South is particularly marked in August and September with differences more pronounced than in June, July and October. This spatial variability is consistent with what is observed during the northward migration, at surface and altitude, of the meteorological equator, which is recorded by ground gauges.
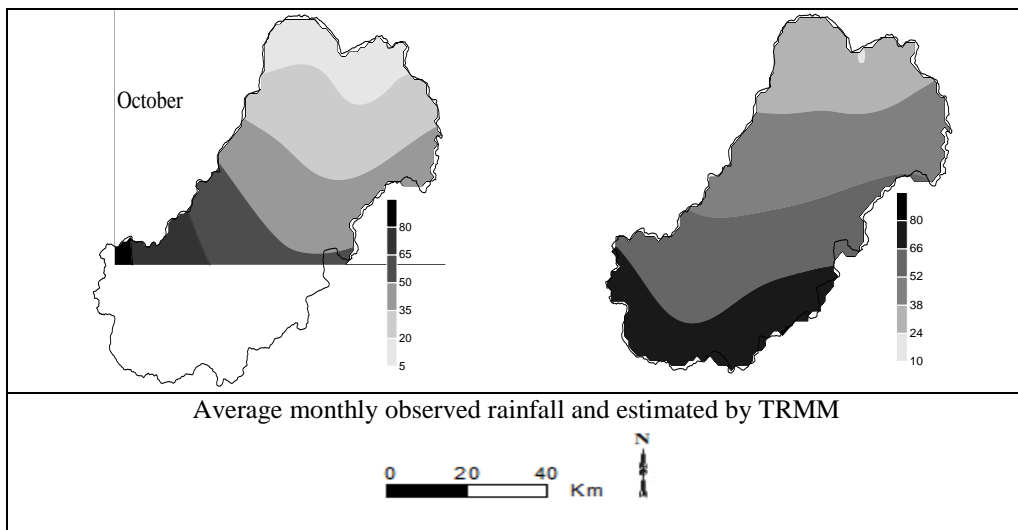
Average monthly observed rainfall and estimated by TRMM

Figure 6: Spatial distribution of average monthly observed and estimated rainfall (1998-2014) from June to October in the Kayanga/Geba Basin

Conclusion

The insufficiency of observed rainfall data and the high variability of the historical time series in the Kayanga/Geba Basin, led to the use of data estimated by the satellite TRMM which proved interesting. Many authors consider that these difficulties are the main reasons for the small number of studies on the evolution of rainfall particularly in areas concerning several countries as the Kayanga/Geba Basin.

The approach of using quantitative and qualitative statistical calculation led to the validation of these estimated data with reference to the actual observed data, on a monthly timescale. The correlation coefficients range from 0.92 (Gabu) to 0.96 (Kounkane), with near-negative biases reflecting an underestimation of rainfall by the satellite and Nash indices close to 1. Their evolution during the year corresponds perfectly to the seasonal rainfall regime stations of the Kayanga/Geba Basin, differentiating the rainy season and the dry season. Despite the short estimation period, the annual variability reproduces the interannual variability of observed rainfall at all selected stations, with good agreement between wet and dry years. The major difficulty is the quantification of what is expressed by an underestimation of rainfall by TRMM. The interpolation of average monthly estimated rainfall from June to October over the period 1998-2014 made it possible to compensate for the shortcomings of the observations and the inequality of the time series. It highlighted the North-South rainfall gradient characteristic of rainfall in the Kayanga/Geba Basin. TRMM data could solve major obstacles meet by researchers for the reliable quantitative documenting of the rainfall for the last two decades in Africa. The

availability of new complete dataset on a daily timescale could allow validation of the estimates to be underaken to establish for example the onset, duration and cessation of the rainy season, which are necessary for the success of the crops in the Basin. Also, the characterization of daily rainfall from the fraction method could help to better understand the daily variability and consequently explain the annual variation of rainfall.

Note

The strength of the linear association between a given reference dataset v_i and an estimation e_i is quantified by the correlation coefficient. Given a set of observations (v_1, e_1), (v_2, e_2),... (v_n, e_n), the formula for computing the correlation coefficient is given by

$$R = \frac{1}{N} \sum_{i=1}^{N} \frac{(v_i - \bar{v})}{\sigma_v} \frac{(e_i - \bar{e})}{\sigma_e}$$

With $\quad \sigma_v = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (v_i - \bar{v})^2}$; and $\quad \sigma_e = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (e_i - \bar{e})^2}$

The Root Mean Square Error (RMSE) is the standard deviation of the residuals (prediction errors). Residuals are a measure of how far from the regression line data points are; RMSE is a measure of how spread out these residuals are. In other words, it tells you how concentrated the data is around the line of best fit. Root mean square error is commonly used in climatology, forecasting, and regression analysis to verify experimental results.

References:
1. Arvor, D., Dubreuil, V., Ronchail, J., Simões Penello Meirelles, M. (2008). *Apport des données TRMM 3B42 à l'étude des précipitations au Mato Grosso*. Annales de l'Association Internationale de Climatologie vol. 5, 49-69.
2. Baillargeon, S. (2005). *Le krigeage : revue de la théorie et application à l'interpolation spatiale de données de précipitations*. Mémoire de maîtrise, Faculté des Sciences et de Génie, Université Laval, Québec ; p. 137.
3. Carbonnel, J-P., Hubert, P. (1992). *Pluviométrie en Afrique de l'Ouest soudano-sahélienne. Remise en cause de la stationnarité des séries*. Pp 37-51. In *L'aridité : une contrainte au développement. Caractérisation, réponses biologiques, stratégies des sociétés*. Ed. ORSTOM; p. 601.
4. Delahaye, F. (2013). *Analyse comparative des différents produits satellitaires d'estimation des précipitations en Amazonie brésilienne*. Thèse, Université Rennes 2 ; p. 228.

5. Dubreuil, V., Jallet, A., Ronchail, J., Maitelli, G. (2004). *Estimation des précipitations par télédétection au Mato Grosso (Brésil).* Annales de l'Association Internationale de Climatologie, vol. 1, 133-156.

6. Ebert, E. E., Janowiak, J. E., & Kidd C. (2007). *Comparison of near-real-time precipitation estimates from satellite observations and numerical models.* Bull. Am. Met. Soc. 88, 47 – 64.

7. Huffman, G. J., Adler, R. F., Dolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., & Wolff D. B. (2007). *The TRMM multi-satellite precipitation analysis: Quasi-global, multi-Year, combined-sensor precipitation estimates at fine scale.* Journal of Hydrometeorology, 8, 38-55.

8. Jobard, I. (2001). *Status of satellite retrieval of rainfall at different scales using multi-source data.* MEGHA-TROPHIQUES 2nd Scientific Workshop, 2-6 juillet 2001, Paris, France, p. 10.

9. Koudou, A., Kouamé, K. A., Niamké, K. H., Kouamé, K. F., Saley, M. B., Adja M. G. (2015). *Contribution de l'Analyse en Composantes Principales à la régionalisation des pluies du bassin versant du N'zi, Centre de la Côte D'Ivoire*. Rev. Ivoir. Sci. Technol., 26, 156-172.

10. Labat, J-Ph. (2012). *Utilisation de méthodes d'interpolation pour la régularisation de données d'observation. Eléments méthodologiques*. Atelier SOMLIT- Analyse de série long terme, Talence, Juin 2012; p. 58. (http://jeanphilippe.labat.free.fr/oceano/Document_pedagogique/Somlit/interp_Talence_2012_V2.pdf)."

11. Laurent, H., Jobard, I., & Toma, A. (1998). *Validation of satellite and ground-based estimates of precipitation over the Sahel.* Atmospheric Research. 47-48, 651-670.

12. NASA, (2007). *Tropical Rainfall Measuring Mission (TRMM).* Senior Review Proposal; p. 50.

13. Nash, J. E., & Sutcliffe, J. V. (1970). *River flow forecasting through conceptual models. Part I - A discussion of principles.* Journal of Hydrology, 10 (3), 282–290. Simpson, J., Kummerow, C., Tao, W., & Adler R. (1996). *On the tropical rainfall measuring mission (TRMM).*Meteorology and Atmospheric Physics, 60(1), 19-36.