Ethiopian Journal of Environmental Studies & Management 8(4): 423 – 432, 2015.
ISSN:1998-0507 doi: http://dx.doi.org/10.4314/ejesm.v8i4.7
Submitted: February 23, 2015 Accepted: May 01, 2015

SELECTION OF AN APPROPRIATE INTERPOLATION METHOD FOR RAINFALL DATA IN CENTRAL NIGERIA

***BILEWU, S.O. AND SULE, B.F.**
Department of Water Resources and Environmental Engineering,
University of Ilorin, Ilorin, Nigeria

Abstract
Interpolation technique can be used to establish the rainfall data at the location of interest from available data. There are many interpolation methods in use with various limitations and likelihood of errors. This study applied five interpolation methods to existing rainfall data in central Nigeria to determine the most appropriate method that returned the best prediction of rainfall at an ungauged site. The methods include the inverse distance wieight method and four variants of universal kriging (spherical,exponential,gaussian and power). The Guassian model of Kriging yielded the least root mean square error for monthly rainfall interpolation and is therefore recommended for use on monthly rainfall data in Nigeria.

Key Words: *Rainfall, Boxplots, Kriging, Interpolation,Variograms, Gaussian*

Introduction

Nigeria has a land size of about 924,000km^2 and the World Meteorological Organization (WMO) has suggested a tolerable gauge network density of 1 gauge per 3000km^2 and 1 gauge for 1000km^2 for flat and mountainous areas respectively in the tropical zones of the world (Reddy, 2005). This varies for different parts of the world. The area to be covered, for example, by Indian standards is smaller than those recommended by the WMO. While Nigeria has a National Water Policy, the area of coverage for a rain gauge is not specified, thus the WMO recommendation is adopted for Nigeria. In the Sahelian region of West Africa, it is approximately 1 rain gauge per 10,000km^2 (Panthou, 2012). Considering the number of rainfall gauging stations currently fully operational in Nigeria, the density is about one gauging station to about 40,000km^2.

While the density of a rain gauge network is important, the spread or arrangement of the gauges over the area is also of importance. In this case, a fair spread must be proven as it has been established that even where there is a large density of gauges, a large error may occur if the pattern estimation of the distribution is not well done (Kutiel and Kay, 1996). The provision of the financial outlay required to install rain gauges which will simultaneously satisfy the requirement of volume measurement and pattern estimate is not feasible in Nigeria for now.

Most of central Nigeria lies in the savannah where the terrain is relatively flat, even though with occasional occurrence of inselbergs. Adopting the WMO tolerable

*Corresponding Author: Bilewu, S.O.
Email: bilewuk@yahoo.com

density of one rain gauge per 3000km^2, the area ought to have at least 210 stations. The number of functional stations in the area is far less than this and thus interpolation of rainfall data would be necessary in certain situations where data are lacking for planning and design purposes.

The lack of hydrological data as a result of progressive decline in ground based observation network is largely due to lack of funds to provide and maintain gauging stations (Oman and Edwards, 2007) and ubiquitous ungauged basins incidences has made it difficult for detailed drainage basin studies and adequate water resources assessment, hence the need for better understanding of hydrological processes and their quantifications and usage. Lack of adequate data has necessitated the use of various statistical tools for flood estimation and predictions (Olukanni and Salami, 2008). Studies have also shown that most of the existing hydro metrological stations in Nigeria are not in proper locations and cannot be said to be representative of the area they are meant to cover (Ologunorisa, 2009).

Interpolation is the estimation of the rainfall value at an unmeasured location from observed values at surrounding locations. Depending on the nature of data, there are various interpolation methods used in hydrology to generate information for specific sites and can be classified into three main groups: deterministic (Thessien Polygon, Inverse Distance Weight (IDW), Splines and Linear Regression), probabilistic (Kriging) and other methods. A comparison of the different methods has been carried out in different parts of the world and the result varies. However, in most of the cases, the probabilistic approach is more acceptable. The choice of interpolation method for rainfall data has been found to depend on the quality of valid measurements, nature of the rains in the study area and the quality of observations (Ly *et al*., 2013). The Ordinary Kriging has been found suitable for the Island of O'ahu in Hawaii, USA (Mair and Fares, 2011) and in South Africa (Coulibaly and Becker, 2007).The Gaussian model of Kriging achieved the best results in the Ourthe and Ambleve catchments of Belgium (Ly *et al*., 2013) and the Ningxia region of China (Hao and Chang, 2013), while the exponential model showed a consistent performance in Switzerland (Naoum and Tsanis, 2004). IDW, a deterministic method, had the highest efficiency in the Brisbane area of Australia (Knight *et al*., 2005).

Deterministic methods show a continuous surface by only using the geometric characteristics of point observations. Probabilistic methods permit the inclusion of the variance in the interpolation process and to compute the statistical significance of the predicted values. Other methods consist of applications that are specially developed for meteorological purposes by applying a combination of deterministic and probabilistic methods.

Study area

Nigeria is located in the tropical zone of West Africa between latitude 4°N and 14°N and longitude 2°2′E and 14°30′E and has a total area of 923,770 km^2 (Barbour, 1982).The country's North – South extent is about 1050 km and its maximum East – West extent is about 1150 km. Land cover ranges from thick mangrove forest and dense rain forest in the South to a near desert condition in the North Eastern corner of the country. Three broad ecological zones are distinguished: the northern Sudan savanna, the guinea savanna or middle belt and the southern rain forest. Based on temperature, Table 1 presents agro-ecological zones in a North–South succession except the mountainous zone which is found at the border with Cameroon and Plateau zone at the centre of the country.

Table 1: Ecological Zones of Nigeria

Zone Description	% of Country Area	Annual Rainfall (mm)	Monthly Temperature (^oC)		
			Min.	Normal	Max.
Semi-Arid	4	400 -600	13	32 – 33	40
Dry Semi humid	27	600 – 1000	12	21 – 31	49
Sub humid	26	1000 – 1300	14	23 – 30	37
Humid	21	1100 – 1400	18	26 – 30	37
Very humid	14	1120 - 2000	21	24 – 28	37
Ultra humid (flood)	2	>2000	23	25 – 28	33
Mountainous	4	1400 - 2000	5	14 – 29	32
Plateau	2	1400 - 1500	14	20 - 24	36

(Source: FAO, Water Profile of Nigeria, 2008)

Methodology

Available Rainfall Data

Precipitation data for the study was sourced from the Nigeria Metrological Agency (NIMET) facilities located at twenty-one stations including airports across Nigeria. NIMET also provides weather information for the aviation industry. Table 2 gives a summary of the locations of gauge stations and rainfall data obtained.

Table 2: Summary of Locations of Rainfall Gauging Stations and Records obtained

S/No	LOCATION	LONGITUDE (^oE)	LATITUDE (^oN)	ELEVATION (m.a.s.l)*	PERIOD OF DATA	NO of YEARS
1.	Potiskum	11.07	11.71	475	1951 - 2005	55
2.	Maiduguri	13.16	11.85	300	1952 - 2005	54
3.	Kano	8.51	11.99	479	1952 - 2005	54
4.	Kaduna	7.35	10.50	612	1960 - 2010	51
5.	Jos	8.85	9.88	1238	1960 - 2010	51
6.	Minna	6.50	9.62	254	1960 - 2010	51
7.	Yola	12.48	9.20	163	1952 - 2005	54
8.	Numan	12.47	9.23	151	1977 - 1989	23
9.	Zing	11.67	9.00	700	2001 - 2010	10
10.	Abuja	7.53	9.08	536	1982 - 2010	29
11.	Lafia	8.57	8.50	290	2000 - 2010	11
12.	Makurdi	8.53	7.73	113	1960 - 2010	51
13.	Ilorin	4.58	8.50	305	1960 - 2010	51
14.	Ibadan	3.97	7.37	200	1960 - 2010	51
15.	Osogbo	4.62	7.80	317	1960 - 2010	51
16.	Akure	5.08	7.25	335	1980 - 2010	31
17.	Ondo	4.83	7.08	277	1960 - 2010	51
18.	Abeokuta	3.32	7.05	067	1981 - 2010	30
19.	Lagos	3.45	6.42	034	1960 - 2010	51
20.	Benin	5.62	6.34	080	1953 - 2005	53
21.	Calabar	8.32	4.95	099	1952 - 2005	54

*m a s l = metres above sea level

Other Data Sources

As a further attempt at ensuring data quality, other sources of rainfall data consulted include the Tropical Rainfall Measuring Mission (TRMM) which is a joint mission between the National Aeronautics and Space Administration (NASA) of the United States and Japan Aerospace Exploration Agency (JAXA). It is designed to monitor and study tropical rainfall and has rainfall data measured by remote sensing from 1998 to date. Also consulted was the Climate Research Unit of the University of East Anglia in the United Kingdom which has a collection of historical climatic data from 1900 to the present that covers the globe.

Data Preparation

Boxplots and Double Mass Curves were used to check for consistency while principal component analysis (PCA) was used to confirm the nature of rainfall distribution. It has been shown that the communality values derived from PCA of rainfall records can be used to derive the best representative stations from a given group of stations (Basalirwa *et al*., 1993).

Decision on which data to use was made using the instrument of Boxplots. Inconsistencies can arise in the rainfall data at a rain gauge station if conditions relevant to recording of rainfall at the station changes significantly during the period of record. This would be felt from the time a significant change took place. Some causes of inconsistency will include shifting of rain gauge to a new location, neighbourhood of the station is undergoing significant changes, errors in observation from a certain date and change in ecosystem due to natural calamities such as forest fires, landslides etc. The double mass curve (DMC) method was used to check for data inconsistency. This method is also useful in checking arithmetical errors in transforming rainfall data from one record to another. A DMC is a plot on a graph paper of the cumulative figures of one variable against the cumulative figures of another variable or against the cumulative computed values of the same variable for a concurrent period of time (Searcy and Hardisson, 1960).

PCA identifies pattern and is used to reduce a large set of data to a smaller set with minimum loss of information. This mathematical procedure allows the derivation of principal components from a set of possibly related variables. A detailed description of PCA will be found in Jolliffe (2002).

Method of Interpolation

Cross validation remains one of the best methods to check the efficiency of an interpolation (Robinson and Metternicht, 2006). The choice of interpolation method to use was narrowed down to the Inverse Distance Weight and the different variants of kriging by cross validation. These models were used elsewhere and have proved efficient (Nusret and Dug, 2012). The Arcgis version 10.2 software was also useful in the process. A good way to select the best method is to calculate the Root Mean Square Error (RMSE) for each method and the one with the least error is considered to be the best interpolation method suitable for the data (Ghazal et al, 2013). A cross validation of the methods was thus carried out and the best method of interpolation was considered as the one that gave the least difference between the measured and the predicted value.

Results and Discussion

Available Data

Table 3 shows summary of descriptive statistics of rainfall measurement from the gauging stations. Rainfall variation between May and October is relatively stable, an indication that rainfall pattern across the country in these months may be fairly consistent. The low variation coefficients also indicate that the monthly data are less dispersed from station to station. Apart from the month of May, the positive values of the coefficients of skewness will usually be an indication that most of the rainfall data are generally lower than the mean. This is due to very high and very low values for the extreme south and extreme north of the country respectively.

The correlation matrix and PCA eigen values at a significance level of 0.05 in Tables 4 and 5 respectively show a high correlation between the gauging stations at 95% Confidence Interval. Furthermore, the first PC explained 93% of the data variance and thus the stations can be classified together. It is an indication that the normal seasonal variation of rainfall in central Nigeria is homogenous and this is dominated by the rain brought by the south west trade wind. The component loadings of each station in the first principal component are in Table 6.

Table 3: Basic Statistics of Monthly Rainfall data from Gauging Stations

MONTH	AVERAGE MONTHLY RAINFALL (MM)			COEFFICIENT OF	
	Minimum	Median	Maximum	Variation	Skewness
January	0	3.2	29.1	1.30	1.67
February	0	3.9	46.2	1.13	0.72
March	0.2	26.2	158.2	0.99	0.90
April	7.2	96.9	225.8	0.58	0.31
May	31.7	154.0	272.6	0.38	-0.14
June	78.7	181.9	404.9	0.40	1.38
July	158.7	192.4	438.5	0.30	1.77
August	108.3	209.7	396.0	0.32	0.54
September	97.5	215.0	407.6	0.31	0.93
October	12.4	123.4	328.2	0.63	0.73
November	0	7.9	163.6	1.58	2.67
December	0	1.2	32.7	1.4	1.66

Data consistency and Criteria for Comparison

Data quality influences the interpolation method selection (Ly *et al*., 2013) and plots of the DMC for the stations listed in Table 2 show some satisfactory consistency. There was no obvious break in slope which will be an indication of inconsistency and thus no rainfall data adjustment was done. It is also reasonable to conclude that there was no change in gauge locations, type, environment and climate as to significantly affect data consistency at any of the stations. The data was then employed to test the IDW method and the different variants of Kriging though cross - validation and the resulting RMSE is shown in Table 7.While no single method fitted every month, the *Guassian* semi-variogram model of Kriging yielded the least error for most of the months and on the total average.

Table 4: Correlation matrix of annual rainfall at the gauge Stations

	ABK	AKR	MKD	ILR	ABJ	LAF	LKJ	MNA	JOS	KAD	OSB	IBD	OND	ZNG	NUM	YLA
ABK	1.000	0.970	0.895	0.960	0.821	0.882	0.917	0.819	0.816	0.773	0.945	0.965	0.978	0.867	0.827	0.827
AKR	0.970	1.000	0.947	0.985	0.890	0.936	0.974	0.892	0.846	0.847	0.971	0.987	0.990	0.912	0.892	0.892
MKD	0.895	0.947	1.000	0.951	0.971	0.990	0.988	0.976	0.942	0.960	0.874	0.911	0.932	0.987	0.979	0.980
ILR	0.960	0.985	0.951	1.000	0.886	0.935	0.971	0.894	0.824	0.842	0.968	0.972	0.967	0.918	0.894	0.894
ABJ	0.821	0.890	0.971	0.886	1.000	0.981	0.956	0.995	0.948	0.984	0.819	0.858	0.889	0.985	0.980	0.981
LAF	0.882	0.936	0.990	0.935	0.981	1.000	0.986	0.988	0.957	0.970	0.858	0.897	0.927	0.994	0.991	0.991
LKJ	0.917	0.974	0.988	0.971	0.956	0.986	1.000	0.963	0.914	0.930	0.917	0.948	0.954	0.970	0.962	0.962
MNA	0.819	0.892	0.976	0.894	0.995	0.988	0.963	1.000	0.945	0.989	0.814	0.851	0.885	0.990	0.992	0.992
JOS	0.816	0.846	0.942	0.824	0.948	0.957	0.914	0.945	1.000	0.957	0.725	0.799	0.856	0.961	0.958	0.958
KAD	0.773	0.847	0.960	0.842	0.984	0.970	0.930	0.989	0.957	1.000	0.742	0.789	0.845	0.979	0.990	0.991
OSB	0.945	0.971	0.874	0.968	0.819	0.858	0.917	0.814	0.725	0.742	1.000	0.985	0.962	0.832	0.795	0.795
IBD	0.965	0.987	0.911	0.972	0.858	0.897	0.948	0.851	0.799	0.789	0.985	1.000	0.977	0.871	0.836	0.836
OND	0.978	0.990	0.932	0.967	0.889	0.927	0.954	0.885	0.856	0.845	0.962	0.977	1.000	0.910	0.883	0.883
ZNG	0.867	0.912	0.987	0.918	0.985	0.994	0.970	0.990	0.961	0.979	0.832	0.871	0.910	1.000	0.991	0.992
NUM	0.827	0.892	0.979	0.894	0.980	0.991	0.962	0.992	0.958	0.990	0.795	0.836	0.883	0.991	1.000	1.000
YLA	0.827	0.892	0.980	0.894	0.981	0.991	0.962	0.992	0.958	0.991	0.795	0.836	0.883	0.992	1.000	1.000

Table 5: PCA Eigen values of the gauge stations

Value	PC 1	PC 2	PC 3	PC 4	PC 5	PC 6	PC 7	PC 8	PC 9	PC 10	PC 11	PC 12 -13
Eigenvalue	14.84	0.909	0.123	0.055	0.032	0.022	0.012	0.006	0.002	0	0	0
% of Variation.	92.74	5.680	0.767	0.345	0.199	0.140	0.074	0.036	0.010	0.003	0.002	0
Cumulative. %	92.74	98.42	99.19	99.53	99.73	99.87	99.94	99.98	99.99	99.99	100.0	100.0

Table 6: PCA Component Loadings and Eigenvectors (Component Score Coefficients) of the Stations

	PC1	
STATION	Component Loadings	Eigenvectors
ABK	0.925	0.240
AKR	0.969	0.251
MKD	0.992	0.258
ILR	0.964	0.250
ABJ	0.971	0.252
LAF	0.992	0.258
LKJ	0.994	0.258
MNA	0.974	0.253
JOS	0.936	0.243
KAD	0.948	0.246
OSB	0.908	0.236
IBD	0.939	0.244
OND	0.963	0.250
ZNG	0.985	0.256
NUM	0.973	0.252
YLA	0.973	0.252

All the four variants of Kriging considered gave better results than the IDW method. These results are valid to the limit of a satisfactory quality of rainfall data used. If the data is of poor quality or unreliable, the uncertainty in error estimation is increased or distorted and a wrong method may emerge as the best. This follows especially from the relatively small difference in the values of the RMSE for all the models.

A plot of the RMSE for all the models is in Figure 3. RMSE for May to July seems to be the same for the Kriging models which implies that rainfall does not vary much across locations in central Nigeria in the middle of the rainy season. Since monthly data is used for this analysis, it is reasonable to assume that the number of rainy days in a year plays a role in the propagation of interpolation errors.

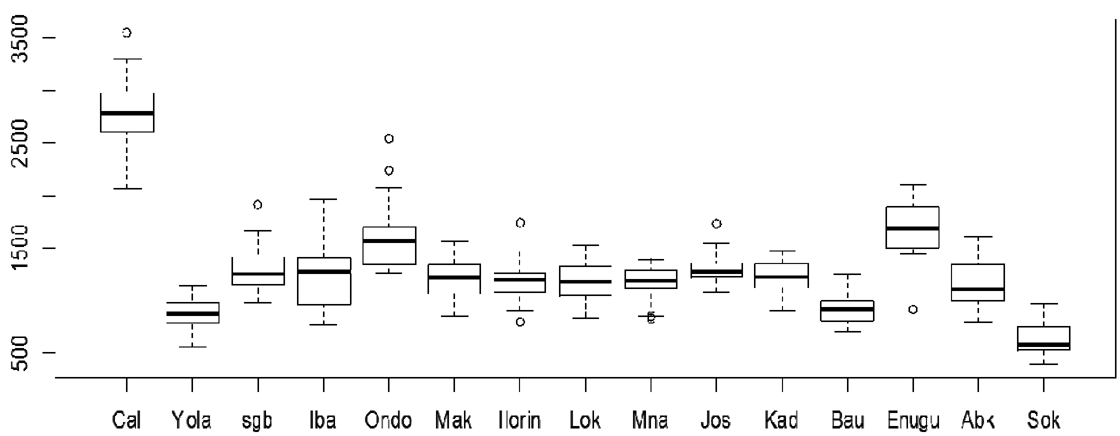

Figure 2: Boxplots of annual rainfall data at gauge locations

Table 7: Root Mean Square Errors of the Interpolation methods

		KRIGING MODELS			
MONTH	IDW	STABLE	SPHERICAL	EXPONENTIAL	GUASSIAN
January	6.988	6.697	6.703	6.736	6.672
February	7.986	4.954	5.806	6.577	5.222
March	25.770	20.206	22.435	24.014	20.206
April	31.323	19.311	23.854	26.617	19.641
May	39.229	30.551	32.237	34.728	30.551
June	68.929	65.652	66.075	65.831	65.930
July	65.201	55.687	56.003	56.353	54.465
August	51.193	33.321	42.352	45.208	34.320
September	52.440	42.177	41.552	45.676	40.880
October	51.530	40.270	44.632	47.765	38.179
November	31.787	22.175	29.314	30.748	22.175
December	6.808	4.710	5.532	6.228	4.704
Sum	439.184	345.712	376.493	396.482	342.945
Average	36.599	28.809	31.374	33.040	28.579

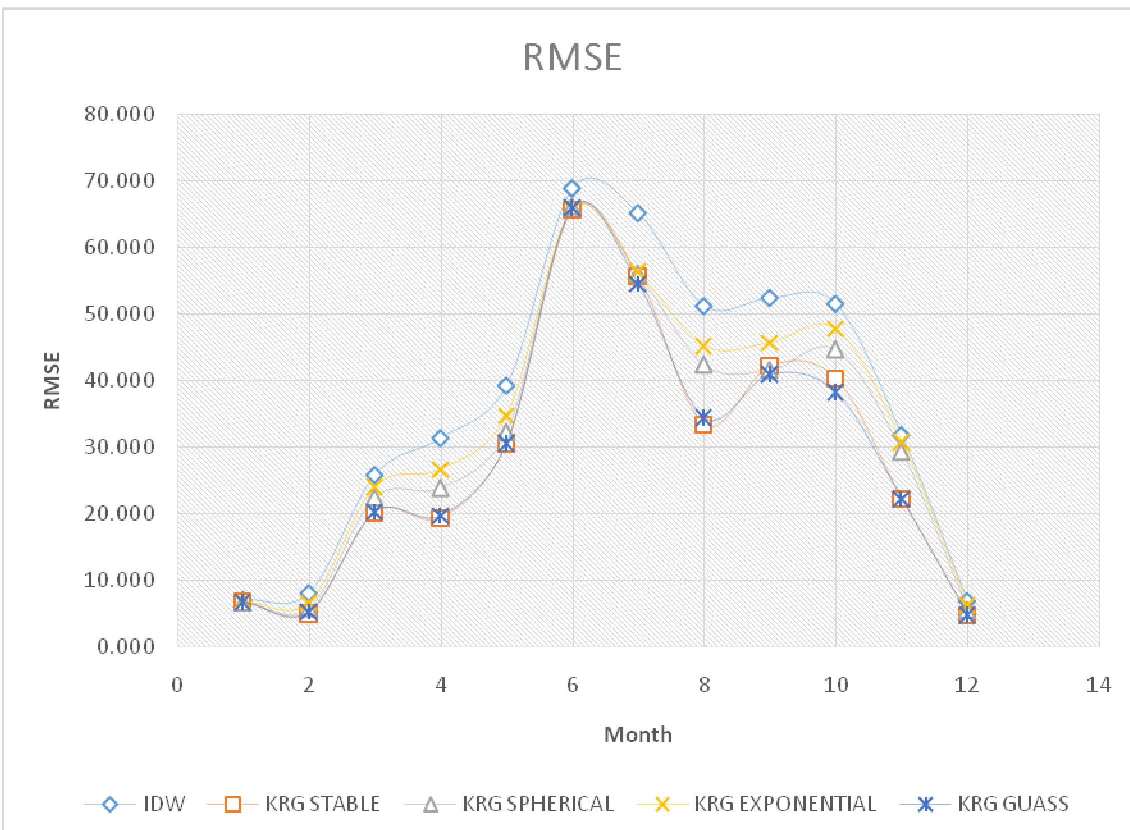

Figure 3: A plot of RMSE variation with months

Conclusion

The number of rainfall gauging stations in Nigeria is lower than the recommended density by the WMO and this makes interpolation imperative for the many ungauged catchments in Nigeria. Rainfall data collected by NIMET and the River Basin Authorities has been shown to be of a satisfactory quality and can be used for modelling studies and water resources planning. However, where the need for interpolation arises, the Guassian model of Kriging is best suited for central Nigeria.

References

Basalirwa, C.P.K., Ogallo, L.J. and Mutua, F.M. (1993). The Design of a Regional Minimum Rainguage Network, International Journal of Water Resources Development, 9(4): 411 – 424.

Coulibaly, M. and Becker, S. (2007). Spatial Interpolation of Annual precipitation in South Africa – Comparison and Evaluation of Methods. *Water International*, 32(3): 494 – 502.

Food and Agricultural Organization (FAO) (2008). *Geography, Climate and Population of Nigeria,* Retrieved on 11/12/2010 from http://www.fao.org/nr/water/aquastat/countries/nigeria/index.stm

Ghazal, N.K., Fadhel, E., Abdu-Alkareem, K. and Zuher, A. (2013). Comparison of Three Interpolation Methodsfor the Average Monthly Temperature in the South of Iraqi Zone. *Iraqi Journal of Physics*, 11(21): 59 – 66.

Hao, W. and Chang, X. (2013). Comparison of Spatial Interpolation Methods for Precipitation in Ningxia, China. *International Journal of Science and Research*, 2(8): 198 – 184.

Jolliffe, I.T. (2002). *"Principal Component Analysis"*, Springer Series in Statistics, Second Edition, ISBN 978-0-387-22440-4.

Knight, Y., Yu, B., Jenkins, G. and Morris, K. (2005). "Comparing Rainfall Interpolation Techniques for Small Subtropical Urban Catchments", MODSIM 2005, International Conference on Modelling and Simulation, Melbourne, December 2005.

Kutiel, H. and Kay, P.A. (1996). Effect of Network Design on Climatic Maps of Precipitation, *Climate Research*, 7: 1 – 10.

Ly, S., Charles, C. and Degre, A. (2011). "Geo-statistical Interpolation of Daily Rainfall at Catchment Scale: the use of various Variogram Models in the Ourthe and AmbleveCatchments, Belgium", *Hydrology and Earth System Sciences*, 15: 2259 – 2274.

Mair, A.and Fares, A. (2011). Comparison of Rainfall Interpolation Methods in a Mountainous Region of a Tropical Island. *Journal of Hydrologic Engineering*, 16(4): 371 – 383.

Naoum, S. and Tsanis, I.K. (2004). Ranking Spatial Interpolation Techniques using a GIS - Based DSS *Global Nest International Journal*, 6(1): 1 – 20.

Nusret, D. and Dug, S. (2012). Applying the Inverse Distance weighting and Kriging Methods of the Spatial Interpolation on the Mapping the Annual Precipitation in Bosnia and Herzegovina, 2012. *Proceedings of the International Congress on Environmental Modelling and Software, Sixth Biennial Meeting*, Leipzig, Germany.

Ologunorisa, T.E. (2009). Content Analysis of Hydrometeorological Network in the Lower Benue River Basin, Nigeria. *Journal of Applied Science and Environmental Management*, 13(2): 33 – 35.

Olukanni D.O. and Salami A.W. (2008). Fitting Probability Distribution Functions to Reservoir Inflow at Hydropower Dams in Nigeria, *Journal of Environmental Hydrology*, 16: 35.

Oman, C. and Edward, R. (2007). *Strengthening Capacity for Water Resources Research in Countries with Vulnerable Scientific Infrastructure*. Published by the International Foundation of Science, Sweden.

Panthou G., Vischel, T., Lebel T., Blanchet J., Quantin, G. and Ali, A. (2012). Extreme Rainfall in West Africa: A Regional Modelling. *Water Resources Research*, 48: W08501

Reddy, P.J.R. (2008). "*A Textbook of Hydrology*", University Science Press, New Delhi.

Robinson, T. P. and Metternicht, G. (2006). Testing the Performance of Spatial Interpolation Techniques for Mapping Soil Properties. *Computer and Electronics in Agriculture*, 50: 97 – 108.

Searcy, J.K. and Hardison, C.H. (1960). "Double Mass Curves", In Manual of Hydrology, Part 1: General Surface Water Techniques, US Geological Survey, Water Supply Paper 1541 B, Washington DC. 31 – 59.