Salami, A. W. et. al (2016). Probabilistic Analysis of Peak Daily Rainfall for Prediction purposes in Selected Areas of Northern Nigeria. NJTR Vol 11; 54-59

Probabilistic Analysis of Peak Daily Rainfall for Prediction purposes in Selected Areas of Northern Nigeria.

Salami,* A. W. [1], Aremu[1] A. S., Ayanshola,[1] A. M. , Abdulkadir,[1] T. S. and M. K. Garba[2] [1]Department of Water Resources and Environmental Engineering, University of Ilorin, P.M.B 1515, Ilorin, Nigeria. [2]Department of Statistics, University of Ilorin, P.M.B 1515, Ilorin, Nigeria

Abstract

In this study, probability analysis was performed on peak daily rainfall data in order to predict rainfall interval values and to determine the best fit functions in some parts of Nigeria. The selected towns are Kaduna, Kano, Yola, Jos, Damaturu and Maiduguri. The obtained peak daily rainfall values were subjected to Gumbel, Log-Gumbel, Normal, Log-Normal, Pearson and Log-Pearson probability distributions. Mathematical equation for probability distribution functions were established for each town and used to predict peak rainfall. The predicted values were subjected to goodness of fit tests such as Chi-square, Correlation Coefficient, Coefficient of Determination and Errors of Estimates to determine how best the fits are. The model that satisfies the tests adequately was selected as the best fit model. The study revealed that the peak rainfall at Kaduna, Jos, Kano, Yola and Damaturu are best fitted by log-Gumbel, while log- Pearson distribution is suitable for predicting peak rainfall in Maiduguri. The result also shows that the occurrences of peak daily rainfall depth of 100 mm and above are rare in the selected areas.

Key words: Peak rainfall, Probability distribution models, Return period, Probability interval and goodness of fit tests
Email: salami_wahab@unilorin.edu.ng
Received: 2015/10/30
Accepted: 2016/03/28
DOI: http://dx.doi.org/10.4314/njtr.v11i1.8

Introduction

Many rivers and streams in Nigeria are ungauged and the most readily available hydrological data is rainfall. Peak rainfall depth of certain return periods are required in convolution with synthetic unit hydrograph ordinate determined from river catchment characteristics data for use to estimate peak flow for the design of hydraulic structures. This necessitates a need for developing probability distribution models which can be used for the prediction of rainfall depth in Nigeria. When using frequency analysis, an assumed probability distribution is fitted to the available data to estimate the magnitude corresponding to return periods and the appropriate distribution models that represent the data are chosen. According to Warren *et. al.* (1972), Viessman *et. al.* (1989) and Mustapha and Yusuf (1999), the choice of probability distribution model is almost arbitrary as no physical basis is available to rationalize the use of any particular function and the search for proper distribution function has been the subject of several studies.

Gary and Robert (1971) studied the normal, log-normal, square-root-normal and cube-root-normal frequency distributions of meteorological data for Texas. In this study, precipitation data conform to the square-root-normal distribution, while evaporation and temperature data conform to all of the frequency distributions tested. Evaporation, temperature and precipitation data were further fitted to the Gumbel extreme-value and log-Pearson type III distributions and precipitation data fitted the log-Pearson type III distribution more adequately than Gumbel distribution, while both evaporation and temperature data conform well to Gumbel distribution. Salami (2004) studied the flow along Asa River and established probability distribution models for the prediction of annual flow regime. The peak flows conformed to Gumbel extreme value type I. Ogunlela (2001) studied the stochastic analysis of rainfall event in Ilorin using probability distribution functions. It was concluded that the log-Pearson type III distribution best described the peak daily rainfall data for Ilorin. Salami and Abdullahi (2003) fitted various probability distribution models to daily peak values of meteorological variables such as rainfall, temperature, sunshine, humidity, evaporation and wind speed, to evaluate the model that is most appropriate for the prediction of these variables. The analysis revealed that Gumbel probability distribution model best fitted humidity, temperature, evaporation and sunshine, while Log-Pearson distribution model best fitted rainfall and wind speed. Salami and Egharevba (2008) developed probability distribution models for forecasting flood level along River Niger and its major tributary; River Benue, in Nigeria. Nine gauging towns were selected along the river channel; annual peak flood levels were selected and subjected to five different probability distribution to determine the best fit functions for each gauging town. The study revealed that peak flood levels at Idah, Yola, Garua and Dadinkowa are best fitted with Gumbel probability distribution model, while at Jebba and Lokoja, the best fit probability model was Log-Gumbel. Best fit probability model was Log-Normal at Ibi, while at Makurdi and Onitsha best fit probability model was Log-Pearson. Salami and Yusuf (2009) established best fit probability distribution models for peak values of meteorological variables in Ibadan and its environs. It was observed that Gumbel probability distribution models best fit rainfall, humidity and evaporation. Olukanni and Salami (2008) fitted probability distribution functions to reservoir inflow at hydropower dams in Nigeria. It was reported that the stream flow for Kainji and Shiroro hydropower dams are best fitted with log-Pearson type III model, while for Jebba the best fit model was log-normal.

This paper presents the determination of the best fit probability distribution and interval probabilities for occurrence of peak daily rainfall data for some selected towns in the Northern part of Nigeria. Some of these towns are shown in figure 1.

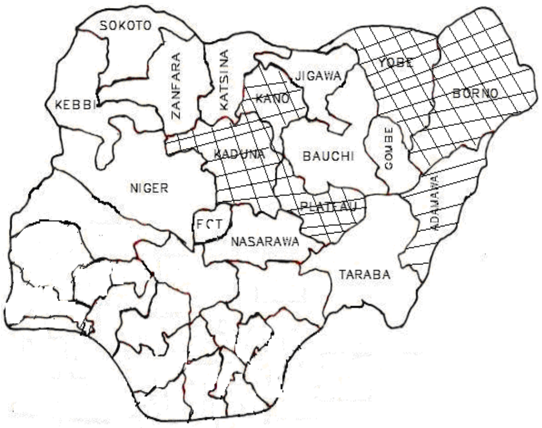

Figure 1 Map of Nigeria showing the location of source of rainfall data in Northern Nigeria

Materials and Methods

Data Collection and Analysis

Rainfall data for selected areas were obtained from the Department of Meteorological Service, Oshodi, Lagos, Nigeria. The rainfall data spanned from 1978 to 2013 and peak daily values were extracted for the purpose of analysis. The data was ranked according to Weibull's plotting position and the corresponding return period was estimated. Ranked data were evaluated with six methods of probability distribution functions to determine best – fit functions. The methods include: Gumbel (EVI type1), Log-Gumbel (LG), Normal (N), Log-Normal (LN), Log-Pearson type III (LP_3) and Pearson (P) probability distribution models. Some statistical goodness of fit tests and errors of estimates were used for the selection of the best fit models.

Probability distribution analysis was carried out in accordance with standard procedure (Warren *et. al.* 1972; Viessman *et. al.* 1989; Mustapha and Yusuf, 1999 and Topaloglu, 2002). Mathematical expressions obtained for each function were used to predict the peak rainfall data based on the estimated return periods and were also used in performing statistical tests (goodness of fit tests) for selection of the best fit models.

Statistical tests

Acceptability and reliability of fitting models (probability distributions) were tested by using statistical tests (goodness of fit tests) and errors of estimates. Statistical tests include chi-square (χ^2), probability plot coefficient of correlation (r) and coefficient of determination (R^2). Also for the purpose of comparison of results from different probability distribution models, two errors of estimates were taken into consideration. They are Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). The statistical tests were carried out in accordance with standard procedure (Chowdhury and Stedinger, 1991; Adegboye and Ipinyomi, 1995; Dibike and Solomatine, 1999; Murray and Larry, 2000). Best probability distribution model is selected based on the performance of the goodness of fit tests and the values of RMSE and MAE.

Estimation of interval probability

Comparative statistics will be used to compare the observed peak daily rainfall data with the predicted data after which the appropriateness of the best fit model will be ascertained. However, for the purpose of prediction, the established best – fit probability distribution models would be used to estimate the probability that the rainfall depth will lie in some specific range. This was computed based on $p(R_1 \le R \le R_2) = F(R_2) - F(R_1)$, where

$F(R_2)$ and $F(R_1)$ are the cumulative distribution function values at R_2 and R_1 respectively.

Results and Discussion.

Statistics of peak daily rainfall from 1978 to 2013 for selected areas are presented in Table 1. Kano has the highest mean of 75.09 mm with standard deviation of 41.92 mm while Jos has the lowest mean of 59.75mm and standard deviation of 8.83mm.

Table 1. Summary of statistics for peak daily rainfall (1978 – 2013)

Parameter	Selected areas in Northern-Nigeria					
	Kaduna	Kano	Yola	Jos	Damaturu	Maiduguri
Mean (mm)	68.77	75.09	60.49	59.75	63.81	60.02
Stdev (mm)	21.97	41.92	12.32	8.83	21.95	18.33
G	0.98	2.63	0.68	0.08	1.53	0.23
Cv	0.32	0.56	0.21	0.15	0.35	0.31
Max. (mm)	132.10	150.10	95.20	82.50	129.30	97.20
Min. (mm)	33.10	32.51	38.60	43.70	38.70	24.50

G = Skewness coefficient, Stdev = Standard deviation, Cv = Coefficient of variation

Mathematical expressions obtained from rainfall values in Table 1 using six probability distributions models (Normal, Log-Normal, Gumbel (EVI), Log-Gumbel, Log-Pearson, and Pearson) for each town are presented in Table 2. For the purpose of theoretical determination of best fit probability function, two statistical tools; Goodness of Fit tests and Error of Estimates were applied. Result of the Goodness of Fit tests is presented in Table 3 and 4. Best fit model was chosen based on the value of the ratio of calculated Chi-square to the table Chi-square ($\chi^2_{cal} / \chi^2_{tab}$), the closeness of Correlation Coefficient (r) to unity (1), and low Error of Estimate.

Table 2. Mathematical expressions for probability distributions models

S/N	Towns	Probability distribution models	Developed Equations
1.	Kaduna	Normal	$R_T = 68.77 + 21.97K_T$
		Log-Normal	$\text{Log } R_T = 1.82 + 0.13K_T$
		Gumbel (EVI)	$R_T = 68.77+ 21.97(0.78Y_T -0.45)$
		Log-Gumbel	$\text{Log } R_T = 1.82+0.13(0.78Y_T -0.45)$
		Log-Pearson	$\text{Log } R_T = 1.82 + 0.13 K'_T$
		Pearson	$R_T = 68.77 + 21.97 K'_T$
2.	Jos	Normal	$R_T = 59.75 + 8.83K_T$
		Log-Normal	$\text{Log } R_T = 1.77 + 0.07K_T$
		Gumbel (EVI)	$R_T = 59.75 + 8.83(0.78Y_T -0.45)$
		Log-Gumbel	$\text{Log } R_T = 1.77+0.07(0.78Y_T -0.45)$
		Log-Pearson	$\text{Log } R_T = 1.77 + 0.07K'_T$
		Pearson	$R_T = = 59.75 + 8.83K'_T$
3.	Kano	Normal	$R_T = 75.09 + 41.92K_T$
		Log-Normal	$\text{Log } R_T = 1.83 + 0.19K_T$
		Gumbel (EVI)	$R_T = 75.09+ 41.92(0.78Y_T -0.45)$
		Log-Gumbel	$\text{Log } R_T = 1.83+0.19(0.78Y_T -0.45)$
		Log-Pearson	$\text{Log } R_T = 1.83 + 0.19 K'_T$
		Pearson	$R_T = 75.09 + 41.92 K'_T$
4.	Yola	Normal	$R_T = 60.49 + 12.32K_T$
		Log-Normal	$\text{Log } R_T = 1.77 + 0.09K_T$
		Gumbel (EVI)	$R_T = 60.49+ 12.32(0.78Y_T -0.45)$
		Log-Gumbel	$\text{Log } R_T = 1.77+0.09(0.78Y_T -0.45)$
		Log-Pearson	$\text{Log } R_T = 1.77 + 0.09K'_T$
		Pearson	$R_T = 60.49 + 12.32K'_T$
5.	Damaturu	Normal	$R_T = 63.81 + 21.95K_T$
		Log-Normal	$\text{Log } R_T = 1.78 + 0.13K_T$
		Gumbel (EVI)	$R_T = 63.81+ 21.95(0.78Y_T -0.45)$
		Log-Gumbel	$\text{Log } R_T = 1.78+0.13(0.78Y_T -0.45)$
		Log-Pearson	$\text{Log } R_T = 1.78 + 0.13K'_T$
		Pearson	$R_T = 63.81 + 21.95K'_T$
6.	Maiduguri	Normal	$R_T = 60.02 + 18.33K_T$
		Log-Normal	$\text{Log } R_T = 1.76 + 0.14K_T$
		Gumbel (EVI)	$R_T = 60.02+ 18.33(0.78Y_T -0.45)$
		Log-Gumbel	$\text{Log } R_T = 1.76+0.14(0.78Y_T -0.45)$
		Log-Pearson	$\text{Log } R_T = 1.76 + 0.14K'_T$
		Pearson	$R_T = 60.02 + 18.33K'_T$

Where:

R_T = Peak rainfall depth of certain return periods, T (mm)

K_T = Coefficient in Normal and Log-Normal distributions, which is a function of probability (p) of exceedence

Y_T = Reduced variate in Gumbel and Log-Gumbel distributions which is a function of probability (p) of exceedence
K'_T = Coefficient in Pearson and Log-Pearson distributions, which is a function of probability (p) of exceedence and Skewness coefficient (G).

Table 3. Results from Gumbel, Log-Gumbel, and Pearson Probability distribution models

Towns	Probability distribution models														
	Gumbel					Log-Gumbel					Pearson				
	χ^2	RMSE	MAE	R^2	r	χ^2	RMSE	MAE	R^2	r	X^2	RMSE	MAE	R^2	r
Kaduna	3.84	3.12	0.52	0.98	0.99	0.02	0.03	0.01	0.97	0.99	41.42	8.64	1.44	0.84	0.92
Jos	3.24	2.24	0.38	0.93	0.97	0.01	0.02	0.01	0.90	0.95	97.08	12.83	2.14	0.02	0.15
Kano	95.86	17.52	2.92	0.823	0.91	0.04	0.05	0.01	0.94	0.97	339.88	33.11	5.52	0.36	0.76
Yola	3.21	2.24	0.37	0.97	0.99	0.01	0.02	0.01	0.94	0.97	29.28	6.91	1.15	0.68	0.83
Damaturu	12.24	5.35	0.89	0.94	0.97	0.01	0.02	0.01	0.98	0.99	21.28	6.53	1.09	0.91	0.96
Maiduguri	10.67	3.87	0.65	0.96	0.98	0.17	0.09	0.02	0.60	0.91	29.12	7.02	1.17	0.85	0.92

Table 4. Results from Log-Pearson, Normal, and Log-Normal Probability distribution models

Towns	Probability distribution models														
	Log-Pearson					Normal					Log-Normal				
	χ^2	RMSE	MAE	R^2	r	χ^2	RMSE	MAE	R^2	r	X^2	RMSE	MAE	R^2	R
Kaduna	0.05	0.05	0.01	0.86	0.93	38.80	8.77	1.46	0.84	0.92	0.05	0.05	0.01	0.87	0.94
Jos	0.03	0.04	0.01	0.59	0.87	17.24	5.97	1.00	0.53	0.84	0.04	0.05	0.01	0.56	0.85
Kano	0.36	0.14	0.03	0.45	0.72	20.88	31.40	5.23	0.42	0.69	0.32	0.14	0.02	0.50	0.74
Yola	0.04	0.05	0.01	0.74	0.87	21.72	6.00	1.00	0.76	0.87	0.04	0.04	0.01	0.77	0.88
Damaturu	0.03	0.04	0.01	0.91	0.96	213.51	19.49	3.25	0.19	0.55	0.04	0.04	0.01	0.90	0.96
Maiduguri	0.05	0.05	0.01	0.89	0.95	28.62	6.95	1.16	0.85	0.93	0.06	0.05	0.01	0.86	0.93

The best fitted probability distribution model for each town is presented in Table 5. For example based on the statistical analysis performed on peak rainfall data at Kaduna town (Table 3 and 4), Log-Gumbel was selected because the value of the ratio of calculated chi-square to the table chi-square (0.0004) is less than one. Also the model gave value of correlation coefficient (r=0.99, R^2 = 0.97) very close to one which indicates that the model is strong and there is close linearity between the observed and predicted peak rainfall. The error of estimate (RMSE = 0.03 and MAE = 0.01) is also very low. The most acceptable and reliable fitting model for Jos, Kano, Yola, and Damaturu is the Log-Gumbel distribution model while for Maiduguri is Log-Pearson model.

Table 5. Goodness of fit tests and the selected model for the peak rainfall

S/N	Towns	Best fit distribution model	RMSE	MAE	$\chi^2_{cal}/\chi^2_{tab}$	R^2	r
1.	Kaduna	Log-Gumbel	0.03	0.01	0.0004	0.97	0.99
2.	Jos	Log-Gumbel	0.02	0.01	0.0002	0.90	0.95
3.	Kano	Log-Gumbel	0.05	0.01	0.0007	0.94	0.97
4.	Yola	Log-Gumbel	0.02	0.01	0.0002	0.94	0.97
5.	Damaturu	Log-Gumbel	0.02	0.01	0.0002	0.98	0.99
6.	Maiduguri	Log-Pearson	0.05	0.01	0.0009	0.89	0.95

For the purposes of comparison, the observed and the predicted rainfall data from the best fit probability distribution model for each town is presented in Figures 2-7. The graphs confirm the appropriateness of the selected models because they are identical to the observed yearly trend.

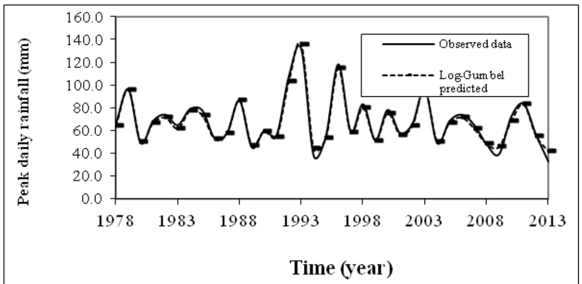

Figure 2 Comparison of observed and predicted data for Kaduna

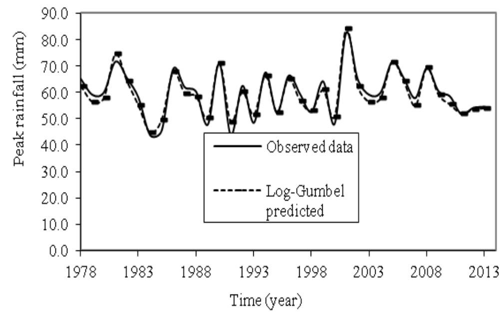

Figure 3 Comparison of observed and predicted data for Jos

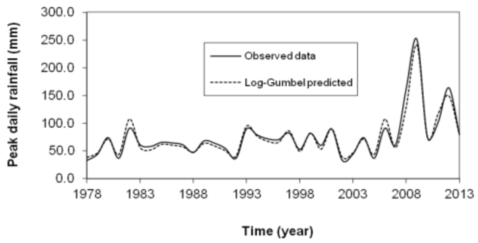

Figure 4 Comparison of observed and predicted data for Kano

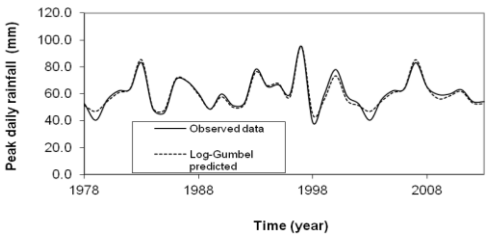

Figure 5 Comparison of observed and predicted data for Yola

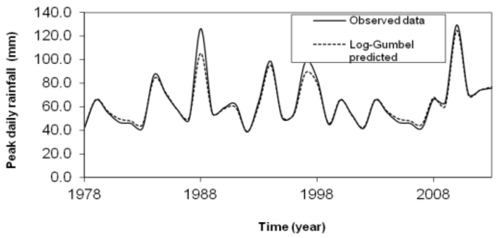

Figure 6 Comparison of observed and predicted data for Damaturu

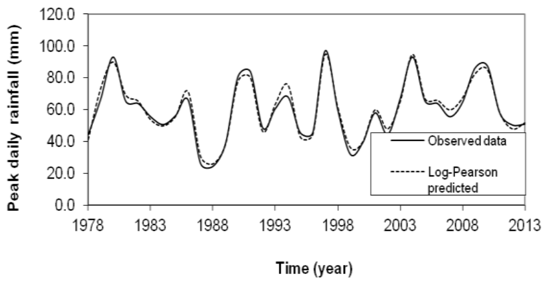

Figure 7 Comparison of observed and predicted data Maiduguri

The interval probabilities estimated for each town based on the best fit distribution model is presented in Table 6. For all the towns, interval probabilities are very low for intervals involving peak daily rainfall depths of 90 mm and above. This suggests that maximum daily rainfall depths of magnitudes greater than 100 mm are rare in those towns. For example, probability that peak rainfall of magnitude between 90 mm and 110 mm will occur varies from 3.0% to 14.0%, peak rainfall values between 110 mm and 130 mm varies from 0.30% to 5.0% while peak rainfall values between 130 mm and 150 mm varies from 1.0% to 3.0% in all the selected areas. This will assist the farmers and water resources planners to know the rainfall distribution pattern for effective planning.

Best fit model developed for each town can be useful for prediction purposes when considering various return periods. Arising from this study, the predicted maximum daily rainfall depths for return periods of 5, 10, 20, 50, 100 and 200 years in the selected towns is illustrated in Table 6. The predicted quartile estimates can find useful application in hydrological planning and design.

Table 6 Interval probabilities for peak daily rainfall (mm) (1978 - 2013).

$p(R_1 \le R \le R_2)$	Kaduna	Kano	Yola	Jos	Damaturu	Maiduguri
$p(30 \le R \le 50)$	0.21	0.21	0.18	0.13	0.22	0.36
$p(50 \le R \le 70)$	0.42	0.35	0.64	0.77	0.48	0.30
$p(70 \le R \le 90)$	0.20	0.20	0.17	0.08	0.19	0.21
$p(90 \le R \le 110)$	0.14	0.10	0.03	0.02	0.07	0.06
$p(110 \le R \le 130)$	0.02	0.05	0.003	0.003	0.03	0.03
$p(130 \le R \le 150)$	0.02	0.03	0.001	0.001	0.01	0.002

Table 6. Quartile rainfall estimates (in mm) for various return periods

Town	Best-Fit Model	Reccurence interval (year)					
		5	10	20	50	100	200
Kaduna	Log-Gumbel	81.99	98.27	116.90	146.37	173.22	204.87
Jos	Log-Gumbel	65.00	71.95	78.29	87.35	94.81	102.89
Kano	Log-Gumbel	92.79	120.13	153.89	212.04	269.63	342.54
Yola	Log-Gumbel	68.00	77.12	86.33	99.91	111.48	124.32
Damaturu	Log-Gumbel	75.69	90.46	107.32	133.90	158.04	186.43
Maiduguri	Log-Pearson III	85.90	97.26	108.94	125.31	138.61	152.87

Conclusion.

Various probability distribution functions were fitted to peak rainfall data at six different towns in Northern part of Nigeria in order to evaluate the model that is most appropriate. The best model was selected based on the statistical goodness of fit tests and value of error estimates. The Log-Gumbel model was found to be appropriate for predicting peak rainfall depth in Kaduna, Jos, Kano, Yola and Damaturu, while Log-Pearson model was suitable for Maiduguri. The establishment of the best fit probability distribution functions and estimation of interval probabilities for peak daily rainfall depth would be a useful guide in the prediction of peak rainfall at the selected towns in Nigeria.

References

Adegboye, O. S. and R. A. Ipinyomi (1995). Statistical tables for class work and Examination. Tertiary publications Nigeria Limited, Ilorin, Nigeria, pp. 5 – 11.

Chowdhury, J U and Stedinger, J R. (1991). Goodness of fit tests for regional generalized extreme value flood distributions. Water Resource. Res., 27(7): 1765 – 1776.

Dibike, B. Y. and D. P. Solomatine (1999). River flow forecasting using Artificial Neural Networks. Paper Presented at European Geophysical Society (EGS) XXIV General Assembly. The Hague, The Netherlands, 1 – 11.

Gary, E O and A. C. Robert (1971). Applications of Climatology and Meteorology to hydrologic Simulation. Issues 38 of Technical report, Texas A & M University Water Resources Institute.

Murray, R S and J. S. Larry (2000). Theory and problems of statistics. Tata Mc Graw – Hill Publishing Company Limited, New Delhi, pp. 314 – 316.

Mustapha, S and M. I. Yusuf (1999). Textbook of Hydrology and Water Resource, Jenas Prints and Publishing Co., Abuja, Nigeria, pp. 164 – 184.

Olukanni, O O and A. W. Salami (2008) Fitting Probabilty Distribution functions to Reservoir Inflow at Hydropower Dams in Nigeria. Journal of Environmental Hydrology, the Electronic Journal of the International Association for Environmental Hydrology.16: 1-7.

Ogunlela, A O (2001). Stochastic Analysis of Rainfall Events in Ilorin, Nigeria. Journal of Agricultural Research and Development. 1: 39 -50.

Salami, A W (2004), Prediction of the annual flow regime along Asa River using probability distribution models. AMSE periodical, Lyon, France. Modelling C-2004. 65 (2): 41-56.

Salami, A W and A. A. Abdulahi (2003). Evaluation of effects of Jebba Dam on Hydro-Meteorological Variables at Jebba and its environs. Technical report submitted to the Department of Civil Engineering, University of Ilorin, Ilorin., Nigeria.

Salami, A W and N. A. Egharevba (2008). Determination of Suitable Probability Distribution Models for Peak flood level series of selected gauging towns in the Niger and Benue Rivers. Journal of Engineering and Technology (JET) Bayero University, Kano. 3(1), 57- 74, 2008.

Salami, A W (2004). Application of Probability Distribution models on studying the Characteristics of Meteorological variables in Ibadan and its environs. Journal of Engineering and Technology (JET) Bayero University, Kano. Nigeria. 4(2); 7 – 14.

Topaloglu, F (2002), Determining Suitable Probability Distribution Models for Flow and Precipitation Series of the Seyhan River Basin. Turk Journal of Agric.26: 189 – 194.

Viessman, W, Krapp, J W, and T. E. Harbough (1989), Introduction to Hydrology. Harper and Row Publishers Inc., New York. pp. 675 – 695.

Warren, V, Terence, E H, and W. K. John (1972), Introduction to hydrology.Educational Publishers, New York, pp. 106 – 141.