EFFICIENT FLOOD FORECASTING FOR THE OPERATION OF HYDRAULIC STRUCTURES IN A TYPICAL RIVER BASIN

[1]Aiyelokun Oluwatobi, [1]Ojelabi, Akintunde, [1]Malomo, Samson and [1]Agbede Oluwole
[1]Department of Civil Engineering, University of Ibadan, Ibadan, Nigeria
Correspondent: aiyelokuntobi@gmail.com

ABSTRACT: This study evaluated the adequacy of various probability distribution models for efficient flood forecasting which are applicable for proper operation of hydraulic structures. A 31-year (1979-2009) hydrologic record of river Opeki gauged at Abidogun was obtained to generate an annual peak data for the study area. Empirical data was fitted into eight probability distribution models, which include Normal, Log Normal, Log Pearson Type III, Exponential, Gamma, Gumbel, Frechet and three-parameter Burr distributions. In order to determine the most suitable distribution model, results were subjected to performance evaluation based on Root Mean Square Error (RMSE) and Goodness of Fit test (GoF), while the diagnostic test was performed using D-index. The study revealed that the highest flood magnitude of 262.50 m^3/s for the study period was estimated to have a return period of six-two (62) years, with a low probability exceedence of 0.01. It was further revealed that Log Pearson Type III distribution with lowest D-index value (0.466) fitted accurately to the annual peak flood, and could be represented by the model $Q = 59.98\ln(Tr) + 86.96$. This model is adequate for proper planning in reducing flood risk in the study area.

Keywords: Probability distributions, flood risk, empirical data,

1 Introduction

Engineers and hydrologists deal with annual peak data when designing water management, irrigation and drainage systems for economic planning. Engineering designs for flood management involve the construction of minor and major hydraulic structures such as barrages, bridges, culverts, dams, spillways, road/railway bridges, urban drainage systems, flood plain zoning and flood protection projects. These constructions are designed and mechanically fit for managing and utilizing water resources to the best advantage using the records of past events [1]. Uncertainty is always present when planning, developing, managing and operating water resources and hydraulic projects, especially when dealing with risk management. It arises because many factors that affect the performance of water resources systems are not and cannot be known with certainty when a system is planned, designed, built, managed and operated [2].

The success and performance of each component of a system often depends on future meteorological, demographic, economic, social, technical, and political conditions, all of which may influence future benefits, costs, environmental impacts, and social acceptability [2]. Flood has caused tremendous losses to properties and sometime life [3], Floods events cannot be described with certainty due to their stochastic nature; hence, they cannot be properly understood using empirical data. A classical way of describing the frequency and magnitude of floods is fitting annual peak instantaneous discharge of streams or annual maximum daily rainfall of an area to probability distributions. A probability distribution is merely useful if it does not fit the data of interest accurately, which inform the need to probability distribution adequacy assessment. Therefore, within the framework of this research, were emphases placed on describing the best stochastic model for hydraulic projects in order to enhance flood risk management in Opeki river basin, using gauged data at Abidogun.

2 Description of the Study Area

The Opeki catchment is a medium to large catchment with a land area of about 980 km^2. It lies entirely within one climatic environment and a consistent geological environment of the Basement Complex of Southwestern Nigeria [4]. River Opeki is gauged just as it discharges into the Ogun River at Abidogun (Figure 1). Opeki catchment is located in Oyo State, Nigeria and has Opeki as its major river. It lies between longitudes 3^o 15' and 3^o 30 'E and between latitudes 7^o 20' and 7^o 54'N. The Opeki catchment falls within the humid tropical climate with distinct wet and dry seasons. The wet season, which is usually double peaked, starts in March and lasts till November. The catchment rises to an altitude of about 460 m above sea level in the northern part of the basin around Awaiye and slopes southwards to about 135 m above sea level at Abidogun, at the mouth of the catchment. The axial length of the basin is about 73 km and its form factor and basin circularity ratio are 0.2 and 0.8 respectively, indicating a long and narrow basin [5]. The drainage pattern in the catchment is dendritic and a drainage density of 1.97 km^{-1} for the basin indicates an excellent drainage. Within the catchment are mainly the gneisses complex and minor occurrence of the Older Granites. Occurring in the upper half of the basin are the variably migmatized undifferentiated biotite and biotite hornblend gneiss with intercalated amphibolite and in the lower half, the schists, amphibolites, pegmatites and coarse porphyritic biotite and biotite- muscovite granite [6].

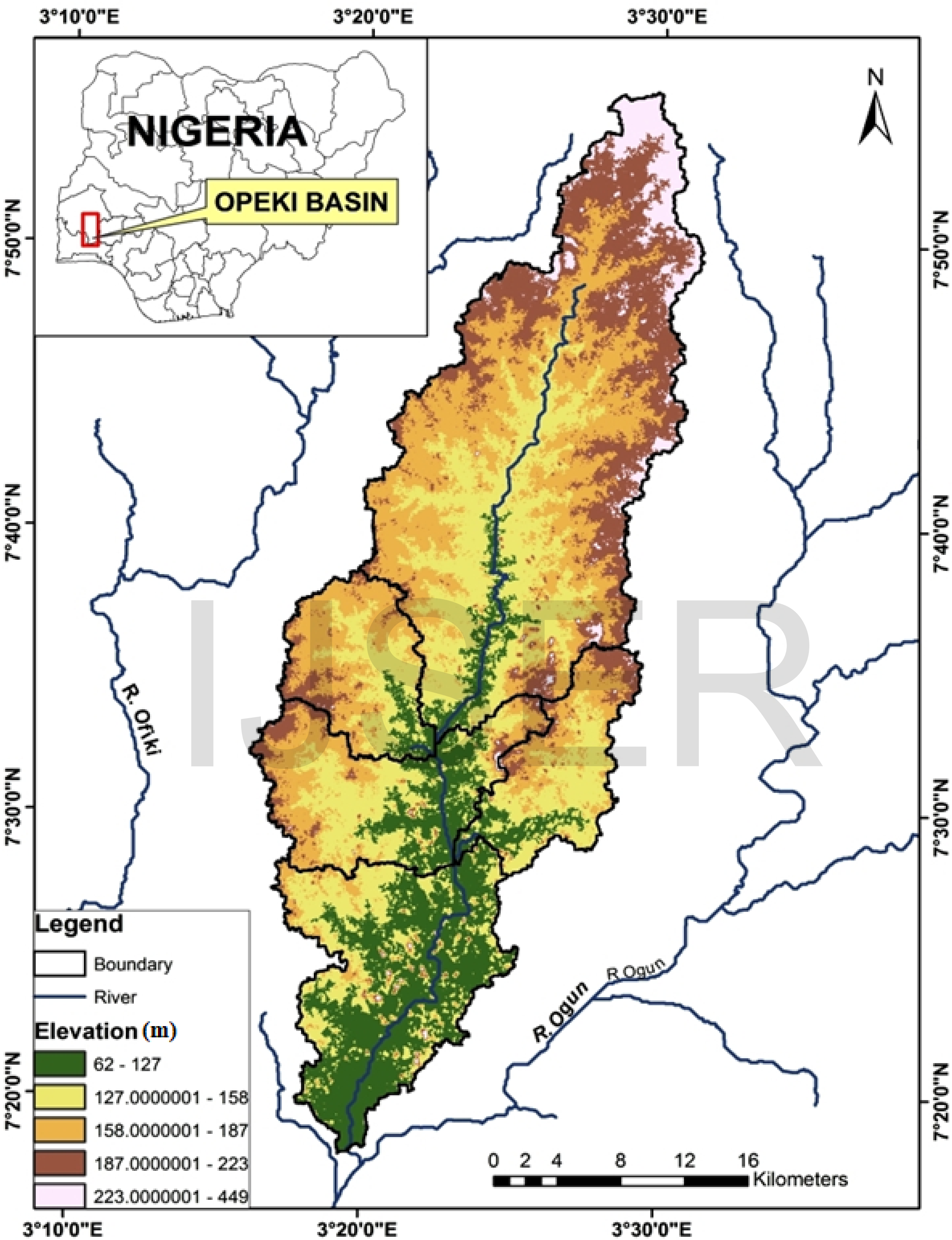

Fig. 1 Drainage Map of the Study Area

3 Materials and Methods

A 31-year (1979-2009) hydrologic record of river Opeki gauged at Abidogun was obtained from the Ogun-Oshun River Basin Development Authority (OORBDA), Abeokuta Nigeria to generate an annual peak data for the study area. The peak flows of years under study were selected and arranged in descending order of magnitude to form an annual maximum series and the probabilities that the ranked annual maximum will be equaled or exceeded in any year were determined by the Hazen's plotting position.

3.1 Plotting Position

The Hazen plotting position was selected for the estimation of the flood's return periods of the river basins and it is represented by equation:

$$Tr = 2n/(2m-1) \quad (1)$$

Where m is the order or rank while n is number of years of study.

3.2 Probability Distribution Models

Various probability distributions were chosen for the study based on their simplicity, superiority, and popularity in the literatures for frequency analysis of extreme events. More so, the maximum-likelihood estimation procedure was employed being recommended for flood estimation of large to moderate samples [2].

The probability density function (PDF) of eight probability distributions used for the study is presented in this sub-section. Here, μ_x and σ_x are the mean and standard deviation of the series of the annual peak flow; μ_y and σ_y are the mean and standard deviation of the log-transformed series of annual peak flow α, β and γ or κ are the scale, location and shape parameters respectively [7].

3.2.1 Normal Distribution (NOR)

For a symmetrically distributed data, the most appropriate distribution of continuous variable is the normal distribution which is also called the Gaussian distribution [8]. The probability density function (PDF) of this distribution model according to Chow *et al.* [9] is given by

$$f(x, \mu_x, \sigma_x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}, \ x, s_x > 0 \quad (2)$$

3.2.2 Log Normal Distribution (LN2)

Large numbers of hydrological continuous variable random variables tends to be asymmetrically distributed. It is advantageous to transform the distribution to a normal distribution by taking the logarithms of the annual maximum discharges [8]. The probability density function (PDF) under this distribution is given as

$$f(\mathrm{x}, \mu_y, \sigma_y) = \frac{1}{\sigma_y x\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu_y}{\sigma_y}\right)^2}, \mathrm{x}, s_y > 0 \qquad (3)$$

In this flood analysis, the logarithms of the annual maximum discharges were taken to base 10.

3.2.3 Log Pearson Type (III) Distribution (LP3)

This is referred to as the three parameter fit. Due to its performance in stochastic flood frequency analysis [10]. The probability model is given as

$$f(\mathrm{x}; \alpha, \beta, \gamma) = \frac{1}{\alpha x \Gamma\beta}\left[\frac{\ln(x)-\gamma}{\alpha}\right]^{\beta-1} e^{-\left(\frac{\ln(x)-\gamma}{\alpha}\right)} \qquad (4)$$

3.2.4 Exponential (Exp)

The Exponential distribution is a commonly used distribution. It has a fairly simple mathematical form with the probability density function (pdf) as follows:

$$f(\mathrm{x}, \alpha, \kappa) = \frac{1}{\alpha} e^{-\left(\frac{x-\kappa}{\kappa}\right)}, \mathrm{x}, \alpha > 0 \qquad (5)$$

3.2.5 Gamma (GAM)

The gamma probability distribution describes the number of events in Poisson process; it assumes the sum of independent and identical exponentially distributed random variables. The probability density function (PDF) under this distribution is given as

$$f(\mathrm{x}, \alpha, \beta) = \frac{1}{\alpha^{\beta}\Gamma\beta} x^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)}, \mathrm{x}, \alpha, \beta > 0 \qquad (6)$$

3.2.6 Extreme Value Type-1 (EV1)/ Gumbel

The Gumbel distribution also referred to as the extreme value type I distribution [11] has two forms, one is based on the smallest extreme (minimum case), and the other is based on the largest extreme (maximum case). In this study, the maximum case is used. Its pdf is given by

$$f(\mathrm{x}; \alpha, \beta) = \frac{e^{-(x-\beta)/\alpha} e^{-e^{-(x-\beta)/\alpha}}}{\alpha}, -\infty < \mathrm{x} < \infty, \alpha > 0 \qquad (7)$$

3.2.7 Frechet (EV2)/ Weibull

The Weibull distribution, also known as extreme value type III distribution, is still a two-parameter distribution with parameters α and β. The pdf is given by

$$f(x;\alpha,\beta) = \frac{\beta}{\alpha}\left(\frac{\alpha}{x}\right)^{\beta+1} e^{-(\alpha/x)^{\beta}}, \quad -\infty < x < \infty,\ \alpha > 0 \qquad (8)$$

The Weibull distribution is a versatile distribution that can take on the characteristics of other types of distributions, based on the value of the shape parameter, β.

3.2.8 Three-Parameter Burr Distribution

The three-parameter Burr distribution is a very flexible distribution family that can express a wide range of distribution shapes. The Burr distribution includes, overlaps, or has as a limiting case many commonly used distributions such as Gamma, Lognormal and Loglogistic ones. It has two asymptotic limiting cases: Weibull and Pareto type I. Due to different values of its parameters covering a broad set of skewness and kurtosis, the Burr distribution can fit a wide range of empirical data in various fields such as hydrology, meteorology, and finance. The pdf for Burr distribution is given by:

$$f(x;\kappa,\alpha,\gamma) = \frac{\gamma\kappa\left(\frac{x}{\alpha}\right)^{\alpha-1}}{\alpha\left(1+\left(\frac{x}{\alpha}\right)^{\alpha}\right)^{\kappa+1}} \qquad (9)$$

3.3 Performance Evaluation

This study employed the use of two statistical procedures for evaluating the performance of the distributions. They include correlation coefficient and root mean square error (RMSE). The RMSE is expressed by the equation given by O'Donnell [12] as:

$$RSME = \left[n^{-1} \sum_{i-1}^{n} (p - o)^2 \right]^{0.5} \qquad (10)$$

Where

RMSE is root mean square error (m^3/s),

P is predicted discharges under each distribution (m^3/s),

Q is observed discharges (m^3/s), and *n* is as previously defined.

ISSN 2229-5518

3.4 Goodness of Fit (GoF) Test

In order to check for the adequacy of fitting of the probability distributions to the recorded annual peak data, two goodness of fit test was applied for the study. GoF tests include Anderson-Darling (A^2) and Kolmogorov-Smirnov (*KS*).

3.4.1 Anderson-Darling Test

The Anderson-Darling (A^2) test compares an observed Cumulative Distribution Function (CDF) to an expected CDF. This method gives more weight to the tail of the distribution than KS test, which in turn leads to the A^2 test being stronger, and having more weight than the KS test. The test rejects the hypothesis regarding the distribution level if the statistic obtained is greater than a critical value at a given significance level (α) [13]. The significance level most commonly used is α=0.05, producing a critical value of 2.5018. This number is then compared with the test distributions statistic to determine if it can be rejected or not. The AD test statistic (A^2) is:

$$A^2 = n - \frac{1}{n}\sum_{i=1}^{n}(2i-1).\left[lnF(x_i) + \ln\left(1 - F(x_{n-i+1})\right)\right] \quad (11)$$

3.4.2 Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov (KS) test statistic is based on the greatest vertical distance from the empirical and theoretical CDFs. Similar to the AD test statistic, a hypothesis is rejected if the test statistic is greater than the critical value at a chosen significance level. For the significance level of α=0.05, the critical value calculated is 0.12555 [13]. The samples are assumed to be from a CDF F(x). The test statistic (KS) is:

$$KS = \max\left(F(x_i) - \frac{i-1}{n}, \frac{1}{n} - F(x_i)\right) \quad (12)$$

3.5 Diagnostic Test

The selection of a most suitable probability distribution for estimation of PFD is performed through D-index, which is defined as:

$$\text{D-index} = \left(\frac{1}{\bar{u}}\right)\sum_{i=1}^{6}\left|x_i - x^*_i\right| \quad (13)$$

Here, $\bar{u}$ is the average value of the recorded annual peak data, x_i is the i^{th} sample of the first six highest values in the series of annual peak data and x_i* is the corresponding estimated value by probability distribution. The distribution having the least D-index is considered as the better suited distribution for estimation of PFD [14].

4 Results and Discusion

4.1 Stage-Discharge Relationship

The rating curve for stage – discharge relationship of river Opeki gauged at Abidogun is shown in Fig 2. Generally, flooding increased with increase in stage for the thirty (31) years of observation. The high co-efficient of determination of 0.993 suggests the ability for the model $Q = 5.252H^{2.142}$ to predict accurate values of discharge for corresponding stages.

4.2 Flood Frequency Analysis

The return periods estimated from the Hazen plotting position of each of the peak flows of the ranked years between 1979 and 2009 is presented in table 1. The highest flood magnitude of 262.50 m^3/s for the study period was estimated to have a return period of six-two (62) years, with a low probability of been equaled or exceeded of 0.01. High floods between 252.64 m^3/s and 230.98 m^3/s are also expected to occur once in 4.1 to 20.7 years with a probability of exceedence of 0.24 to 0.05 respectively in the study area. Moderate flood magnitude of 179.47 m^3/s to 100.68 m^3/s was found to have a return period of 3.6 to 2.1 years, while floods with low magnitude of 95.32 m^3/s to 34.26 m^3/s is expected to occur every 2.0 to 1.1 years. Furthermore, least flood of 16.37 m^3/s has a return period of 1 year, and a high probability of exceedence. It could also be observed in the table 1 that as the magnitude of floods increases their return period increases, while the probability of exceedence decreases. This goes to show that floods with great magnitude are not frequently experienced in the basin area; however, such floods have high risk when they occur.

The frequency distribution of the annual peak data is presented in fig 3. The Fig shows that approximately 80% of the annual peak floods have magnitudes less than 250 m^3/s. While the remaining 20% of the annual peak floods have greater magnitude of floods greater than 250 m^3/s. Implying that high floods are not be expected frequently.

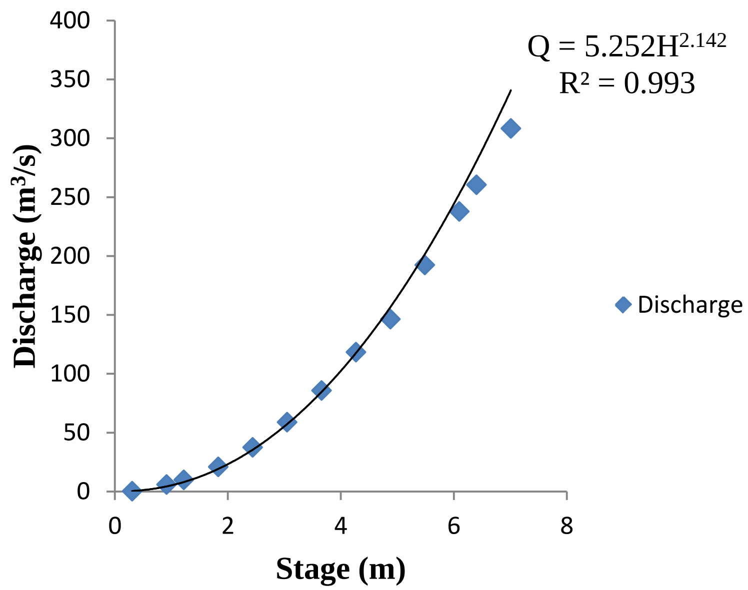

Fig. 2: Rating Curve of Opeki River at Abidogun

Table 1: Frequency Analysis of River Opeki Flood Data (1979-2009)

Year	order	Discharge	Return Period	Probability of Exceedence
2006	1	262.50	62.0	0.02
1998	2	252.64	20.7	0.05
1985	3	247.35	12.4	0.08
2008	4	246.47	8.9	0.11
2000	5	243.85	6.9	0.15
1990	6	239.52	5.6	0.18
2005	7	239.52	4.8	0.21
1980	8	230.98	4.1	0.24
1987	9	179.47	3.6	0.27
1984	10	138.02	3.3	0.31
1991	11	119.46	3.0	0.34
1993	12	110.14	2.7	0.37
1994	13	105.63	2.5	0.40
1989	14	101.22	2.3	0.44
2007	15	100.68	2.1	0.47
1981	16	95.32	2.0	0.50
1995	17	92.19	1.9	0.53
1999	18	92.19	1.8	0.56
2001	19	92.19	1.7	0.60
2009	20	92.19	1.6	0.63

1988	21	90.64	1.5	0.66
2002	22	90.64	1.4	0.69
1986	23	90.13	1.4	0.73
1992	24	79.23	1.3	0.76
2004	25	79.23	1.3	0.79
1979	26	71.33	1.2	0.82
1996	27	63.44	1.2	0.85
1982	28	61.33	1.1	0.89
2003	29	44.09	1.1	0.92
1997	30	34.26	1.1	0.95
1983	31	16.37	1.0	0.98

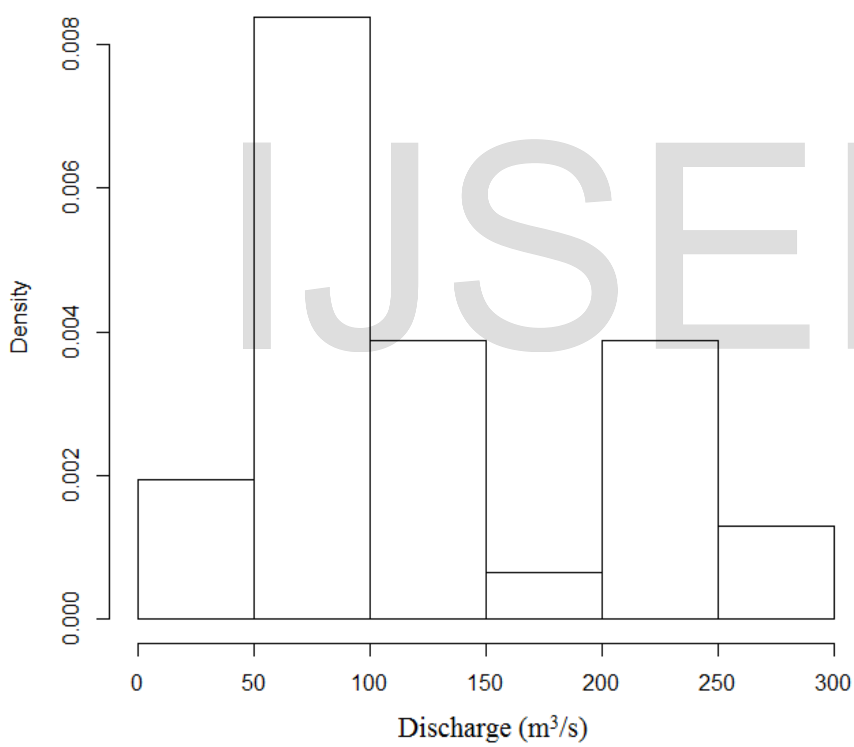

Fig. 3: Frequency Distribution of Annual Peak Data

As shown in table 2, a mean annual peak data of 129.10 m^3/s was computed for the study period, while the low standard deviation of 75.53 m^3/s indicates that the floods were not largely dispersed from their mean. Furthermore, the statistics of Skewness with a value of 0.7 shows that data collected is asymmetrical with a long tail to the right, while the statistics of Kurtosis with a value of 2.04 indicates that the flood distribution is more peaked than a normal distribution.

Table 2: Statistical Summary of Annual Peak Data of Opeki River Basin (1979-2009)

Parameters	Mean (m3/s)	Minimum (m3/s)	Maximum (m3/s)	Standard Deviation (m3/s)	Skewness	Kurtosis
Value	129.10	16.37	262.5	75.53	0.70	2.04

4.3 Estimation of Design Floods Using Probability Distribution Models

Based on table 3, it could be noted that the Exponential distribution gave higher estimates for return period between 10-year and 100-year consistently when compared with corresponding values of other seven distributions for the data under study. It could also be observed that the Three-Parameter Burr Distribution gave the highest estimates for return period of 200-year and above. Fig. 4 shows the plots of recorded and estimated flood discharge for different return periods obtained using the eight flood distributions for river Opeki gauged at Abidogun.

Table 3: Design Flood Estimates Given by eight Frequency Distributions for Opeki River

Return	Estimated Design Flood m^3/s							
Period (yr)	NOR	LN	LP3	EXP	GAM	EV1	EV2	3BD
2	129.13	107.73	114.41	89.49	114.67	115.54	119.57	108.27
5	191.64	184.17	184.77	207.78	184.71	179.02	189.04	175.71
10	224.33	243.77	230.78	297.27	230.37	221.04	229.68	234.65
25	259.19	328.69	287.28	415.67	286.31	274.14	257.57	334.4
50	281.71	398.71	327.98	505.06	326.58	313.53	306.4	433.58
100	301.69	474.34	367.49	594.54	365.63	325.64	334.81	560.56
200	320.51	556.06	406.08	684.83	403.75	391.6	361.34	723.77
500	342.97	674.19	456.03	802.33	453.06	442.99	394.08	1013.75
1000	358.73	771.74	493.16	891.81	489.7	481.84	417.4	1307.69

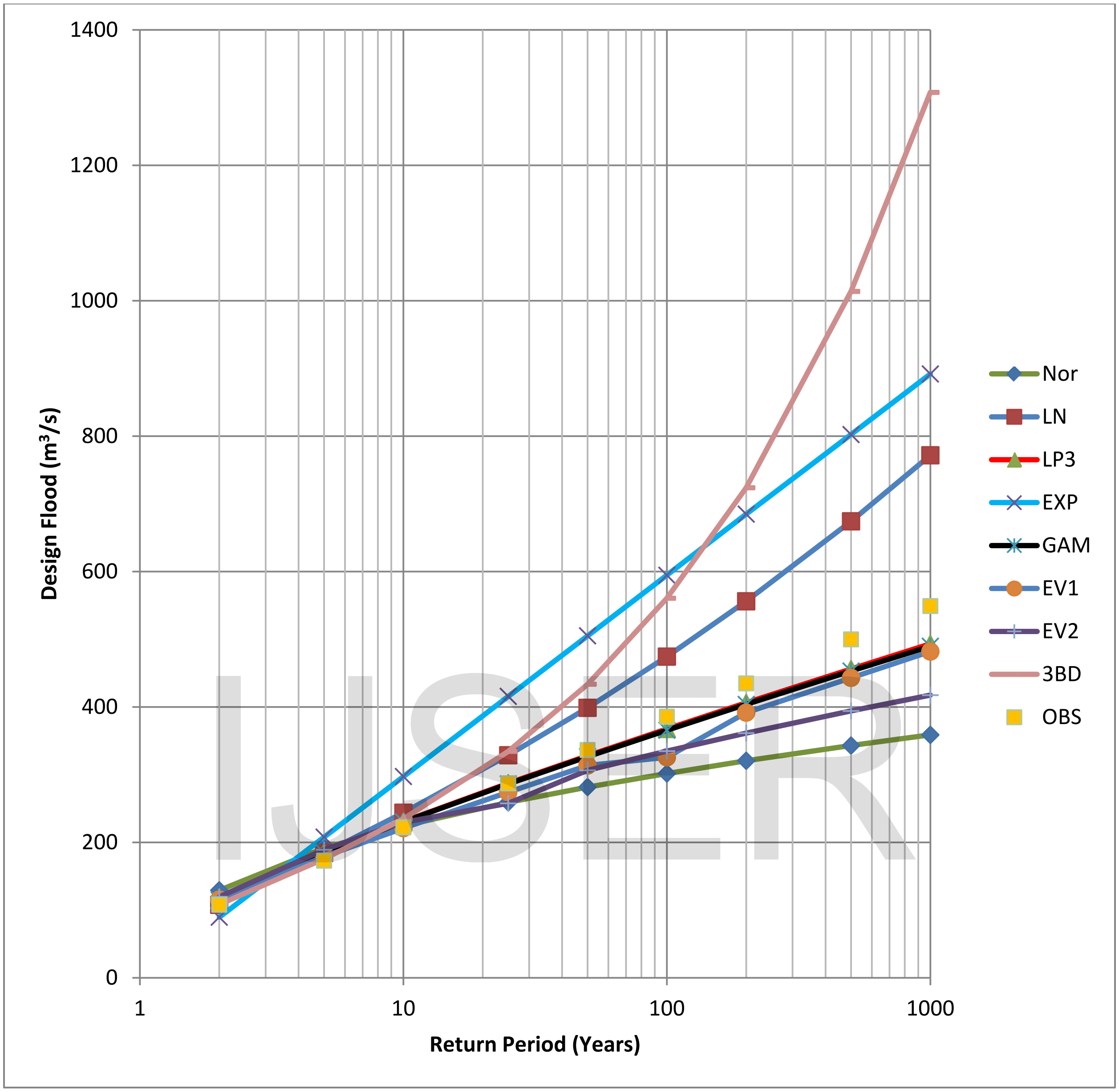

Fig. 4: Plots of Recorded and Estimated Flood Discharge using eight Distributions

4.4 Quantitative Analysis

4.4.1 Analysis Based on Goodness of Test (GoF)

In order to assess the fit of probability distributions to the recorded annual peak data series, Goodness of Fit test statistics for eight distributions were computed and are given in Table 4.17. For the study, degrees of freedom for all distributions were considered as thirty (30). A critical value of 2.50 and 0.242 at 0.05 significant level for Anderson-Darling and Kolmogorov-Smirnov (*KS*) test respectively was used in assessing the fit of the eight distributions. As shown in table 4.17, the computed values of Anderson-Darling for Normal and Exponential distributions are greater than 2.5 hence they were found to be unsuitable to fit the empirical data, while the remaining seven distributions are considered suitable. It could also be observed in table 4 that the computed values of Kolmogorov-Smirnov statistics of Normal and Exponential distributions is greater than the theoretical values of 0.242 at 5% level of significance, and hence at this level, Normal and Exponential distributions are not found suitable to fit the recorded annual peak data of Opeki river, while other distributions are considered suitable.

Table 4: Computed Values of Anderson-Darling and Kolmogorov-Smirnov (*KS*)

Distributions	Anderson-Darling (A^2)	Kolmogorov-Smirnov (*KS*)
Normal (NOR)	2.32	0.246
Log Normal (LN2)	1.052	0.142
Log Pearson Type (III)	1.245	0.17
Exponential	3.171	0.281
Gamma	1.259	0.171
Extreme Value Type (I)	1.46	0.179
Extreme Value Type (II)	1.482	0.196
Three Parameter Burr	1.017	0.154

4.4.2 Performance Evaluation of Models

The performance of the models was evaluated using correlation coefficient and root mean square error (RMSE). It could be observed in table 5 that Exponential distribution has the highest correlation coefficient of 0.999, followed by Gamma and Log Pearson Type (III) distributions with a value of 0.998 respectively. This implies that the three distributions had high positive linear co-variation with the empirical distribution of the floods in the study area. The table further shows Log Pearson Type (III) distributions had the lowest RMSE of 29.92 m^3/s, this provides information on the short-term performance which is a measure of the variation of predicated floods around the recorded flood data. Which implies the Log Pearson Type (III) distribution is more accurate in estimating design floods in the study area.

Table 5: Computed Values of Correlation Coefficient and RMSE

Distributions	Correlation Coefficient	RMSE
Normal (NOR)	0.982	97.47
Log Normal (LN2)	0.996	109.93
Log Pearson Type (III)	0.998	26.92
Exponential	0.999	202.01
Gamma	0.998	28.72
Extreme Value Type (I)	0.997	39.44
Extreme Value Type (II)	0.992	65.59

Three Parameter Burr	0.957	327.54

4.4.3 Diagnostic Analysis

For the selection of a most suitable distribution for estimation of design floods Opeki River basin, the D-index values of eight probability distributions were computed and given in Table 6. It could be observed in the table that the D-index value given by Log Pearson Type (III) distribution is found to be minimum when compared to the corresponding indices of other distributions. On the basis of the diagnostic test results, Log Pearson Type (III) distribution has been identified as better suited for estimation of peak flood discharge for Opeki River.

Table 6: Indices of D-index for Eight Probability Distributions

Distributions	D-index
Normal (NOR)	1.889
Log Normal (LN2)	2.135
Log Pearson Type (III)	0.466
Exponential	4.21
Gamma	0.504
Extreme Value Type (I)	0.791
Extreme Value Type (II)	1.266
Three Parameter Burr	5.65

4.5 Qualitative Analysis

A qualitative assessment of the goodness of fit was ascertained from the plot of the recorded and estimated flood discharge by suitable probability distribution. Fig 5 shows the plots of recorded and estimated flood discharge given by Log Pearson Type (III) distribution for Opeki within a confidence limit of 95%. As depicted in the Fig 5, it can be observed that the recorded annual peak data falls within the confidence limits of the estimated values given by the selected distribution. In addition, the percentage of error in estimated design floods using Log Pearson Type (III) distribution is 2.0%. This implies that the percentages of uncertainty in flood estimation for the sites are within the acceptable tolerance limit of $\pm 10\%$ as recommended by Ang and Tang [15].

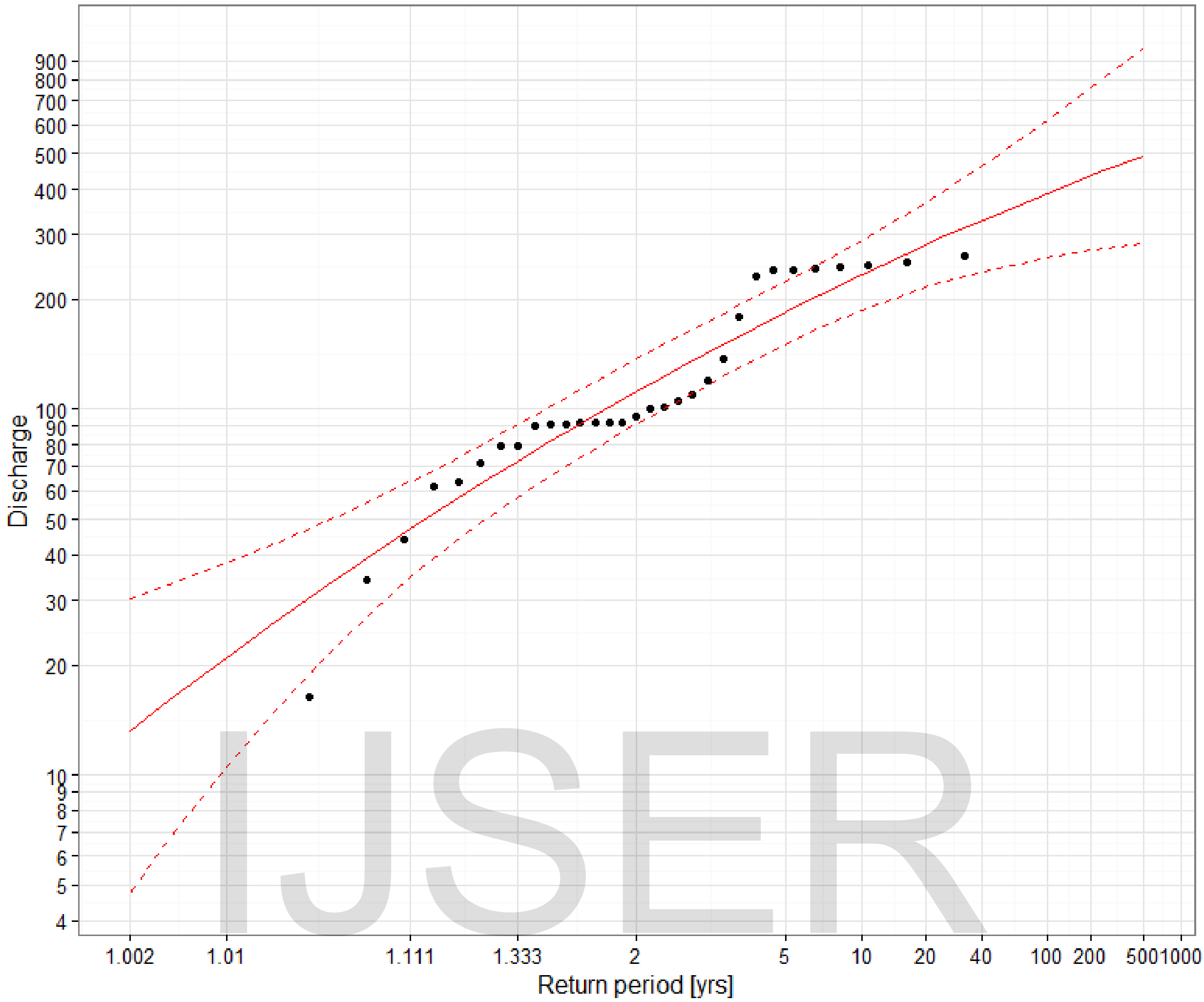

Fig. 5: Plots of Recorded and Estimated Flood given by the Selected Distributions with Confidence Limits at 95% Level

4.6 Discussion of Findings

Water resources management project either long-term involving planning, design and construction, or short-term involving maintenance of new and existing facilities is sensitive to climate variability [16]. As a result, divers methodologies have been used for the design and constructions of sustainable hydraulic structures. Flood frequency analysis is one of those methodologies embraced by civil, hydrologic and hydraulic engineers for decreasing flood damages and economic losses. Hydrological data such as flow rate is used in engineering for the design of hydraulic structures to mitigate flooding [1]. In principle estimating the frequency of a given magnitude event by using an empirical distribution function is possible, but in practice where too few data are available, the empirical distribution produced is handicapped in estimating the frequency of occurrence of events larger than the maximum records. As a result, it is logical to fit the empirical data into a theoretical frequency distribution [3]. This informed the fitting of 31-year instantaneous peak data of Opeki River into Normal, Log Normal, Weibull, Gumbel, EV1, EV2, Log Pearson Type III and Three Parameter Burr Probability Distributions.

The analysis of the empirical data revealed that flooding increased with increasing stage for the thirty (31) years of observation. The relationship between floods and river stage was represented by $Q = 5.252H^{2.142}$, having a high co-efficient of determination of 0.993 which suggested that the model could give accurate values of discharge for corresponding stages.

Flood frequency analysis of the study area showed that as the magnitude of floods increased, their return periods also increased, while their probability of exceedence decreased. Which goes to show that floods with great magnitude are not frequently experienced in the basin area; however, such floods have high risk when they occur. Return period of each discharge were computed through the use of Hazen plotting position, the estimated flows of the selected return periods of 2, 5, 10, 25, 50, 100, 200, 500 and 1000 years fitted into the eight distributions also show an increase with discharge. It was observed that the Exponential distribution gave higher estimates for return period between 10-year and 100-year consistently when compared with corresponding values of other seven distributions for the data under study, while the Three-Parameter Burr Distribution gave the highest estimates for return period of 200-year and above. In order to assess the adequacy of the models, quantitative assessment involving GoF tests as well as qualitative test involving GoF plots were employed. Based on the quantitative assessment by both Anderson-Darling and Kolmogorov-Smirnov (*KS*) test, Normal and Exponential distributions were found to be unsuitable to fit the empirical data of Opeki River.

The performance of the models was further evaluated using correlation coefficient and root mean square error (RMSE). Exponential distribution has the highest correlation coefficient of 0.999, followed by Gamma and Log Pearson Type (III) distributions with a value of 0.998 respectively, which implied that the three distributions had high positive linear co-variation with the empirical distribution of the floods in the study area. Log Pearson Type (III) distributions had the lowest RMSE of 29.92 m^3/s. Hence, it performed more accurately in estimating design floods in the study area. For the selection of a most suitable distribution for estimation of design floods Opeki River basin, the D-index values of eight probability distributions were computed, on the basis of diagnostic test results, Log Pearson Type (III) distribution was identified as better suited for estimation of peak flood discharge for Opeki River.

Qualitative assessment of the goodness of fit was ascertained from the plot of the recorded and estimated flood discharge by Log Pearson Type (III) distribution, which showed that the recorded annual peak data falls within the confidence limits of the estimated values given by the selected distribution.

The percentage of error in estimated design floods using Log Pearson Type (III) distribution was 2.0%. This implies that the percentages of uncertainty in flood estimation for the sites are within the acceptable tolerance limit of ±10% as recommended by Ang and Tang [15].

This study further confirmed the adequacy of combining Log-Pearson type III distribution and Hazen plotting position in modeling design floods within the Opeki River basin. This corroborates with many other studies such as Fasinmirin and Olufayo [17], Adeboye and Alatise [18] and Ewemoje and Ewemooje [19]. This finding institutionalize that Hazen plotting positions fits well with the log-Pearson type III distribution when used for designing hydraulic structures along rivers in Southwest Nigeria.

5 Conclusion

An important problem faced by engineers is the choice of a frequency distribution function for the fitting of extreme flood series in a region or country, since there is no general agreement in applying a particular distribution for flood frequency analysis for different region or country. An in-depth assessment of eight models was performed using quantitative and qualitative methods. The following facts emerged from the study:

1. Floods of high magnitudes are not expected to occur frequently; however, such floods have high risk when they occur.
2. The Log Pearson type III distribution in combination with Hazen plotting position was found to be most suitable in flood prediction in Opeki River basin. Having the least value of the D-index diagnostic test and low Root Mean Square Error (RMSE).
3. The relationship between the design flood and the return period based on the Log Pearson type III distribution is represented by $Q = 59.98\ln(Tr) + 86.96$ and could be used to estimate design floods of any return periods. A coefficient of determination of 0.996 suggests that the model is adequate in predicting design floods in the study area. This can assist hydrologist and civil engineers in planning for flood regulation and designs of hydraulic structures.

References

[1] A.O. Agbede and O. Abiona, O, "Plotting Position Probability Fittings to Lagos Metropolitan Precipitation: Hydrological Tools for Hydraulic Structures Design in Flood Control," *International Journal of Pure and Applied Sciences and Technology*, vol. 10, no. 1, pp. 37-43, 2012.

[2] UNESCO, "Concepts in Probability, Statistics and Stochastic Modelling", Water Resources Systems Planning and Management Report, 2005.

[3] O.A Agbede and O.O. Aiyelokun, "Establishment of a Stochastic Model for Sustainable Economic Flood Management in Yewa Sub-Basin, Southwest Nigeria", *Civil Engineering Journal*, vol. 2, no. 12, pp. 646-655, 2016.

[4] O.A. Idowu, and O. Martins, "Hydrograph Analysis for Groundwater Recharge in the Phreatic Basement Aquifer of the Opeki River Basin, Southwestern Nigeria", *Asset Series, B*, vol. 6, no.2, pp.132-141, 2007.

[5] R.E. Horton, “Drainage Basin Characteristics”, Trans. Am. Geophys. Union 13, 1932.

[6] A. Jones, and C. Hockey, 1964. Geology of Southwestern Nigeria. Report. Vol. 21, no.108, 1964.

[7] A.R. Rao, and K.H. Hamed, “Flood frequency analysis”, CRC Publications, New York, 2000.

[8] K. Tilahun, “The Characterization of Rainfall in the arid and semi-arid regions of Ethiopia”, *Water SA*. Vol. 32, no.3, pp. 429-439. 2006.

[9] V.T. Chow, D.R. Maidment and L.W. Mays, “Applied Hydrology”, McGraw-Hill, New York, 1988.

[10] S.S. Sumioka, D.L. Kresh and K.D. Kasnick, “Magnitudes and Frequencies of Flood in Washington”, US Geological Survey, Water Resources Investigation Report, 97-4277 W.A, vol. 15, 1997.

[11] S. Coles, “An Introduction to modelling of extreme values”, Springer, London, 2001.

[12] T. O’ Donnell, “A Direct three-parameter Muskingum procedure incorporating lateral inflow”, *Journal of Hydrological Sciences*, vol. 30, no. 4, pp. 494-495, 1985.

[13] N. Millington, D. Samiran and S. P. Simonovic, 2011. “The Comparison of GEV, Log-Pearson Type 3 and Gumbel Distributions in the Upper Thames River Watershed under Global Climate Models”, Water Resources Research Report, 2011.

[14] U.S. Water Resources Council (USWRC), “Guidelines for Determining Flood Flow Frequency”, Bulletin 17A, U.S Geological Survey, Washington, D.C, 1981.

[15] A.H-S, Ang and W.H. Tang, “Probability Concepts in Engineering Planning and Design, Vol II- decision, risk and reliability”, John Wiley & Sons, Inc., New York, 1984.

[16] O. Aiyelokun and A. Olodo, “Forecasting One-Decade Ahead Minimum Temperature and Relative Humdity for Water Resources Management in Lower Niger”. *Journal of Water Security*. Vol 3. pp 1-7, 2017.

[17] J.T. Fasinmirin and A.A. Olufayo, “Comparison Of Flood Prediction Models For River Lokoja Nigeria”, Department of Agricultural Engineering, Federal University of Technology, Akure, Nigeria, 2006.

[18] O.B. Adeboye, and M.O. Alatise, “Performance of probability distributions and plotting positions in estimating the flood of river Osun at Apoje Sub-basin, Nigeria”. *Agricultural Engineering International: the CIGR Ejournal*, Manuscript LW 07 007. vol. IX, 2007.

[19] T.A. Ewemoje, and S.O Ewemooje, "Best distribution and plotting positions of daily maximum flood estimation at Ona River in Ogun-oshun river basin, Nigeria", *Agricultural Engineering International: CIGR Journal*. Vol. 13, no. 3, pp. 1-13, 2013.