Nigerian Journal of Technology (NIJOTECH)
Vol. 33. No. 1, January 2014, pp. 10 – 18
Copyright© Faculty of Engineering,
University of Nigeria, Nsukka, ISSN: 1115-8443
www.nijotech.com
http://dx.doi.org/10.4314/njt.v33i1.2

L-MOMENTS APPROACH FOR FLOOD FREQUENCY ANALYSIS OF RIVER OKHUWAN IN BENIN-OWENA RIVER BASIN IN NIGERIA

O.C. Izinyon[1,*], J.O.Ehiorobo[2]
[1,2] DEPARTMENT OF CIVIL ENGINEERING, UNIVERSITY OF BENIN, BENIN CITY, NIGERIA
E-mail addresses: *[1] izinyon2006 @yahoo.com, [2] jeffa_geos@yahoo.com*

Abstract
The flood frequency analysis of River Okhuwan at Ugonoba site in Benin Owena River Basin in Nigeria was carried out using annual maximum stream flow series for 20 years (1989-2008). The objective of the study was to determine the best fit probability distribution model applicable to the site from amongst three selected 3-parameter distribution models namely: Generalized Extreme Value distribution (GEV), Generalized Logistic distribution (GLO) and Generalized Pareto distribution (GPA) whose parameters were estimated by the method of L-moments. The best fit distribution was selected based on result of four goodness-of-fit tests applied to the candidate distributions. The results indicate that GPA is the best fit distribution model for analyzing the annual maximum flood series at the study site for return period less than or equal to 25 years but for return periods between 25years and 200 years, GLO or GEV is recommended in that order.

Keywords: Flood frequency Analysis, L-moments, Probability distribution, Return period, Parameter estimation

1. Introduction

Floods are one of the most costly natural hazards that have ravaged different parts of the world [1], destroying human lives and properties. The need for preventive action to reduce the unnecessary costs, economic losses and danger arising from overflow of water has become imperative and has continued to engage the attention of water resources engineers and hydrologists. Reliable estimation of the magnitude and frequency of flood peak discharge is important for determination of flood risk, management of water resources and proper design of hydraulic structures like dams, spillways, culverts and irrigation ditches [2, 3]. The estimation of the design event should therefore be accurate to avoid excessive damage and loss of human lives which could arise from the underestimation of flood potential. The magnitude of flood peak discharge and associated exceedence probability can be estimated by various approaches including flood frequency which is based on statistical inference. Because hydrological phenomena like storm precipitation, low flow and annual flood maxima are characterized by variability, randomness and uncertainty, the analysis of these hydrological data is amenable to statistical inference [4].

Flood frequency analysis is a statistical technique widely used for predicting future events at different recurrence intervals. In flood frequency analysis, how often a flood will occur is estimated by analyzing the flow data from a stream location and a probability distribution is selected to fit the observed data. Many probability distributions have been suggested in the hydrological and statistical literature to model hydrological phenomena such as extreme events but no particular model is considered superior for all practical applications hence available models are screened based on problem to be solved and the nature of available data [4]. The choice of a distribution is more difficult in regions with short records as is the case in developing countries because conventional moment statistics are both highly biased and highly variable in small samples [5] and yet design floods with high return periods are to be estimated by extrapolation of the fitted distribution. The decision regarding which distribution to use is based on comparison of the suitability of several candidate distributions. The part of a sample that lies near the mean of a distribution can often be

*** Corresponding author, Tel: +234-8035038239**

described well by a variety of distributions; but the values estimated for a large return period or in the tail of the distribution as well as for very small cumulative probabilities by the individual distributions can differ significantly from one another. According to [6], the prediction of return periods that do not greatly exceed the length of hydrological records is less sensitive to the choice of distribution functions but estimates of events with high return periods is dependent on the behavior of the tail of the fitted distribution and since hydraulic design is often based on estimates of large recurrence interval events, it is important to determine them as accurately as possible. Therefore, the selection of appropriate distribution is very important for such cases. The more parameters a distribution has, the better it adapts to the sample data including the tail of the distribution but a large sample size is required to give accurate parameter estimates hence [6] recommends not more than two or three parameter distributions for at-site frequency analysis as is the case of this study. It is recommended by [4] on a general note that a mathematical distribution having three parameters is preferred because they make the distribution match available data more consistently.

In flood frequency analysis, once the distribution function has been selected, the next step is to estimate the parameters of the distribution from the sample data [6] so that required quantiles can be calculated with the fitted model. A number of parameter estimation methods are available in the statistical and hydrological literature; common ones include: Method of Moments (MOM), Method of Maximum Likelihood (MML) and method of Probability Weighted Moments (PWM) and the corresponding L-moments approach [7]. MOM is a natural and relatively easy method but its estimates are usually inferior in quality and are generally not efficient for distributions with large number of parameters (3 or more) because higher order moments are more likely to be highly biased in relatively small samples. The MML is considered an efficient parameter estimation method because it provides the smallest sampling variance of the estimated parameters and estimated quantiles but it suffers serious disadvantage of frequently giving biased estimates with small samples especially if the number of parameters are large [7]. The method of Probability Weighted Moments (PWMs) and the corresponding L-moments is a variant of the method of moments which provides a different way to summarize the statistical data set [8]. L-moments offer significant advantages over ordinary product moments especially for environmental data set in the following respects [7]:

- L-moments ratio estimators of location, scale and shape are nearly unbiased regardless of the probability distribution from which the observation arose[8]
- L-moment ratio estimators such as L-CV, L-skewness and L-kurtosis exhibit lower bias than conventional product moment ratios especially for highly skewed samples.
- L-moment ratio estimators L-CV and L-skewness do not have bounds which depend on sample size as do ratio estimators of CV and skewness.
- L-moment estimators are linear combinations of the observations and thus are less sensitive to the largest observations in a sample than product moment estimators which square or cube observations. Hence in general, L-moments are more robust to extreme values in the data and therefore enable more secure inference to be made from small samples about an underlying probability distribution [8].

The main focus of this paper is the at-site flood frequency analysis of River Okhuwan at Ugonoba in Benin Owena River Basin in Nigeria using annual maximum stream flow series for 20 years (1989-2008). The objective of the study was to determine the best fit probability distribution model from among three selected 3-parameter distributions namely: Generalized Extreme Value distribution (GEV), Generalized Logistic distribution (GLO) and Generalized Pareto distribution (GPA). The parameters of each distribution were estimated using the method of L-moments while the best fit distribution was selected by application of statistical goodness of fit tests to candidate distributions.

2.0. Materials and Methods

2.1. The study Area

The study site is located within Longitude 5^0 51'E and Latitude 6^0 19'N in Edo state in the south south geopolitical region of Nigeria as shown in Figure 1. It is in the humid tropical climatic region and is characterized by two distinct seasons ("wet and dry"). Some characteristics of the study site and data collected are given in Table 1.

Table1: Salient characteristics of the study site and sample data

Station		R.Okhuwan at Ugonoba
Location	Longitude	5^0 51'E
	Latitude	6^0 19'N
Drainage* Area(Km^2)		245
Coefficient of variation		0.5114
Coefficient of skew(C_s)		-0.3192
Record length(yrs)		20

*Source [24]

2.2: Probability distribution models selected for the study

The probability distribution models selected for the study are: the Generalized Extreme Value distribution (GEV), Generalized Logistic distribution (GLO) and Generalized Pareto distribution (GPA). The Generalized Extreme Value distribution spans the three types of extreme value distribution for maxima; the Gumbel and Generalized Extreme Value distributions are widely used for flood frequency analysis around the world [9].

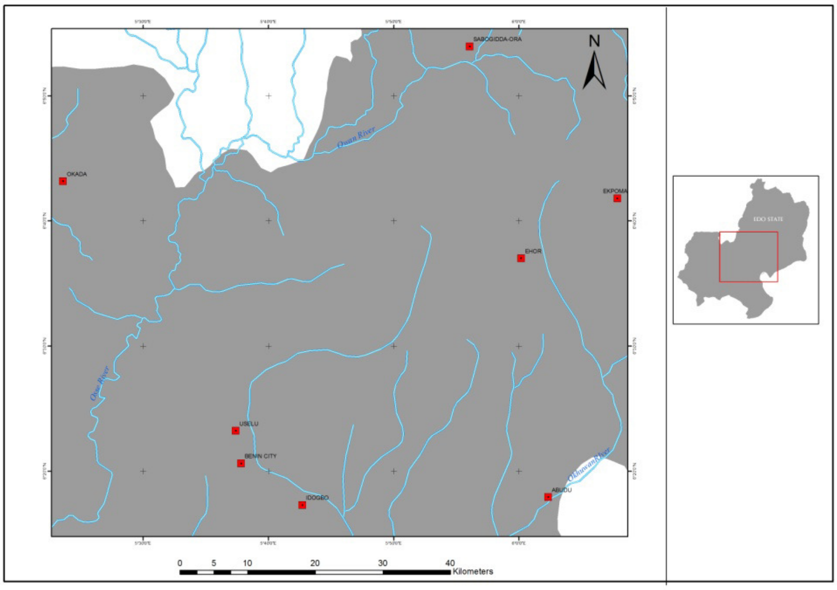

Figure 1: Map of Edo state Nigeria showing location of study River in the Benin Owena River Basin.

Table 2: Commonly used frequency distributions [13]

Distribution	Probability density function and/ or Cumulative distribution function	Range	Moments
Generalized Extreme Value (GEV)	$F_x(X) = exp\left\{-\left[1-\frac{k(x-\xi)}{\alpha}\right]^{\frac{1}{k}}\right\}$ when 0< k, x <$\left(\xi+\frac{\alpha}{k}\right)$; k < 0, $\left(\xi+\frac{\alpha}{k}\right)$< x	σ_X^2 exist for $-0.5 < k$	$\mu_x = \xi + \left(\frac{\alpha}{k}\right)[1-\Gamma(1+k)]$ $\sigma_x^2 = \left(\frac{\alpha}{k}\right)^2\{\Gamma(1+2k) - [\Gamma(1+k]^2\}$
Generalized Logistic (GLO)	$\gamma = \left[1-\frac{k(x-\xi}{\alpha}\right]^{\frac{1}{K}}$ for k ≠ 0 $f_x = \left(\frac{1}{\alpha}\right)\left[\frac{\gamma^{(1-x)}}{(1+\gamma)}\right]^2$ $F_x(x) = \frac{1}{1+\gamma}$	$y = \exp\left[-\left(\frac{x-\xi}{\alpha}\right)\right]$, $fork = 0$ for k< 0, $\xi + \frac{\alpha}{k} \le x < \infty$ for 0< k, -∞< $x \le \xi + \frac{\alpha}{k}$	$\mu_x = \xi + \frac{\alpha}{\left[\frac{1}{k} - \frac{\pi}{\sin(k\pi)}\right]}$ See Ahmad *et al* (1988) for σ_x^2
Generalized Pareto (GPA)	$f_x(x) = \frac{1}{\alpha}\left[1-\frac{(x-\xi)}{\alpha}\right]^{\frac{1}{(k-1)}}$ $F_x(x) = 1 - \left[1-\frac{k(x-\xi)}{\alpha}\right]^{1/k}$	for k < 0, $\xi \le x < \infty$ for $0 < k, \xi \le x \le \xi + \frac{\alpha}{k}$ (γ_x exist for k > −0.33)	$\mu_x = \xi + {}^{\alpha}/(1+k)$ $\sigma_x^2 = {\alpha^2}/(1+k)^2(1+2k)$ $\gamma_x = \frac{2(1-k)(1+2k)^{\frac{1}{2}}}{(1+3k)}$

The Gumbel distribution is a special case of a Generalized Extreme Value distribution corresponding to k = 0.

The Generalized Logistic distribution was introduced to the main hydrological literature by [10] and has been proposed as the distribution for flood frequency analysis in the UK [11]. The Flood Estimation Handbook (FEH) recommends fitting the Generalized Logistic distribution (GLO) to annual maximum flow series and it is used in the UK for estimating return periods in the range of 2 to 200 years [12]. The 3 parameters (location, scale and shape factor) of the distribution may be estimated from the L-moments of the data set. The Generalized Pareto distribution (GPA) is useful for modeling events like daily rainfall and all floods above a moderate threshold that exceed a specified lower bound at which the density function has a maximum (k < 1). The probability density functions and or cumulative distribution functions, range and moments for the selected distributions are given in Table 2.

2.3: Description and basic theory of method of L-moments for flood frequency analysis

The method of L-moments is extensively presented in [10] but in summary it is a modification of the probability weighted moments (PWMs) method presented in [14].

L-moments and probability weighted moments are used to summarize theoretical distributions and observed samples thereby making them liable for use in parameter estimation, interval estimation and hypothesis testing [15]. L-moments and L-moment ratios are however more convenient than probability weighted moments because they are more easily interpretable as measures of distribution scale and shape [16]. A summary of the theory of PWMs is presented in [14] and defines them as:

$$\beta_r = E\,\{X[F_S(x)]^r\} \quad (1)$$

Where β_r is the rth order of PWM and $F_S(X)$ is the cumulative distribution function (CDF) of x

Unbiased sample estimators of β_i of the first four PWMs for any distribution can be computed as follows [10]:

$$\beta_0 = \frac{1}{n}\sum_{j=1}^{n} X_{(j)} \;; \quad (2a)$$

$$\beta_1 = \sum_{j=1}^{n-1}\left[\frac{(n-j)}{n(n-1)}\right] X_{(j)} \quad (2b)$$

$$\beta_2 = \sum_{j=1}^{n-2}\left[\frac{(n-j)(n-j-1)}{n(n-1)(n-2)}\right] X_{(J)} \quad (2c)$$

$$\beta_3 = \sum_{j=1}^{n-3}\left[\frac{(n-j)(n-j-1)(n-j-2)}{n(n-1)(n-2)(n-3)}\right] X_{(j)} \quad (2d)$$

Where, $X_{(j)}$ represents the ordered stream flows with $X_{(1)}$ as the largest stream flow and $X_{(n)}$ as the smallest.

For any distribution, the first four L-moments ($\lambda_1, \lambda_2, \lambda_3$ *and* λ_4) expressed as linear combinations of PWMs are given by [10] :

$$\lambda_1 = \beta_0 \quad (3a)$$

$$\lambda_2 = 2\beta_1 - \beta_0 \quad (3b)$$

$$\lambda_3 = 2\beta_2 - 6\beta_1 + \beta_0 \quad (3c)$$

$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \quad (3d)$$

The L-moments ratios used for expressing the parameter estimates are given as:

$$\text{L–coefficient of variation (L-CV)} = \tau = \frac{\lambda_2}{\lambda_1} \quad (4)$$

$$\text{L–skewness } (\tau_3) = \frac{\lambda_3}{\lambda_2} \quad (5)$$

$$\text{L–kurtosis } (\tau_4) = \frac{\lambda_4}{\lambda_2} \quad (6)$$

The sample estimates of L-moments (l_1, l_2, l_3 *and* l_4) are calculated by replacing β_0, β_1, β_2 *and* β_3 in equation 3 with b_0, b_1, b_2 *and* b_3 respectively [12]

The L-moments λ_1 and λ_2, the L-CV, τ and L-moment ratios τ_3 *and* τ_4 are the most useful quantities for summarizing probability distribution. The equations which relates sample moment parameters and population parameter estimates for different distributions using L-moments are given in [10].

2.4. Data used and Analysis method

The instantaneous discharges for the River Okhuwan flow gauging station at Ugonoba for 20 years period (1989–2008) were obtained from the daily flow records of Benin Owena River Basin Development Authority [17] and then analyzed to obtain the annual maximum discharge series which are given in Table 3.

The annual maximum discharges were ranked in descending order of magnitude and the associated probability weighted moments, that is, b_0, b_1, b_2 *and* b_3 for the sample data were obtained using equations 2a, 2b, 2c and 2d respectively [12].The sample estimates of L-moments (l_1, l_2, l_3 *and* l_4) were calculated by replacing $\beta_0, \beta_1, \beta_2$ *and* β_3 in equation 3a, 3b, 3c and 3d with the obtained b_0, b_1, b_2 *and* b_3 respectively while the associated L-moment ratios(i.e., L-CV, L-skewness and L-kurtosis) were computed using equations (4), (5) and (6) respectively. The cumulative probability of non -

exceedence, $F\ (Q_i)$of each event was computed using the Weibull formula given by [18]:

$$F\ (Q_i) = \frac{m}{n+1} \quad (7)$$

Table 3: Annual Peak discharge (m³/s) at River Okhuwan at Ugonoba [17]

Year	Maximum flow(m³/s)
1989	25.6
1990	23.0
1991	22.5
1992	16.0
1993	12.32
1994	22.5
1995	64
1996	75.28
1997	78.40
1998	82.9
1999	70.0
2000	67.0
2001	86.0
2002	86.52
2003	66.0
2004	90.3
2005	95.84
2006	97.6
2007	102.4
2008	125

Where m is the rank, n is the number of annual maxima in the record or the sample size and Q_i is the i^th^ element of a sample of annual maximum flows arranged in descending order of magnitude. The observed annual maxima were then subjected to flood frequency analyses with the selected three 3- parameter probability distribution functions namely: Generalized Extreme Value distribution (GEV), Generalized Logistic distribution (GLO) and Generalized Pareto distribution (GPA). The parameters of the fitted distributions were estimated by L-moment method using procedures given in section 2.5. The adequacies of the fitted distributions were evaluated by four methods of goodness of fit namely: Root Mean Square Error (RMSE), Relative Root Mean Square Error (RRMSE), Mean Absolute Deviation Index (MADI) and Correlation Coefficient (CC) as explained in subsection 2.6 and on the basis of the test results the best fit distribution was selected and then utilized for flood quantile estimation.

2.5. Parameter estimation using L- moments

Estimates of the parameters of the selected distributions were obtained following the L-moment procedures using the respective equations given in Table 4. ξ is the location parameter, α *is* the scale parameter and K *is* the shape parameter.

Table 4: L-Moment Parameter estimates for selected probability distribution functions [1].

Distribution	Quantile Function	Parameter Estimates
GEV	$X\ (F\ ) = \xi + \frac{\alpha}{K}\{1 - (-\ln F)^K\}$	$\alpha = \frac{l_2\, K}{\Gamma(1+K)\Gamma(1-2^{-K})}$ $\xi = l_1 + \frac{\alpha(\Gamma(1+K) - 1)}{K}$ $K = 7.8590C + 2.9554C^2$ $C = \frac{2}{3+\tau_3} - \frac{\ln 2}{\ln 3}$
GLO	$X(F) = \xi + \frac{\alpha}{K}\left\{1 - \left(\frac{1-F}{F}\right)^K\right\}$	$\alpha = \frac{l_2}{\Gamma\ (1+k)\ \Gamma(1-K)}$ $\xi = l_1 + \frac{(l_2 - \alpha)}{K}$ $K = -\tau_3$
GPA	$X(F) = \xi + \frac{\alpha}{K}\{1 - (1-F)^K\}$	$\alpha = l_2[(K+1)(K+2)]$ $\xi = l_1 - l_2(K+2)$ $K = \frac{4}{\tau_3 + 1} - 3$

The computation processes in subsections 2.4 and 2.5 were facilitated using MS EXCEL programming.

2.6. Evaluation of performance of fitted probability distributions

The adequacy of the selected probability distribution models in fitting the observed peak discharge data were evaluated by goodness of fit tests or criteria. The methods are Root Mean Square Error (RMSE), Relative Root Mean Square Error (RRMSE), Mean Absolute Deviation Index (MADI) and Probability Plot Correlation Coefficient (PPCC). The first three methods assess the fitted distribution applied to a site by summarizing the deviation between observed discharges and predicted discharges while the last gives the correlation between the ordered observations and corresponding fitted quantiles determined by a plotting position [5]. The result of the tests enabled ascertaining how sufficiently close a given distribution fits the observed data and hence the choice from the candidate distributions the one that best fits the observed data.

2.6.1. Root mean square error (RMSE)

The root mean square error of a distribution fitted to the observed discharge data at a site is the square root of the sum of the squares of the differences between the observed and predicted values. It is computed using the equation:

$$RMSE = \left(\frac{\sum (x_i - y_i)^2}{(n-m)}\right)^{\frac{1}{2}} \quad (9)$$

where x_i , i=1,--,---,n are the observed values and y_i, i=1,--,---,n are the values computed from the assumed probability distributions, the number of parameters estimated for the distribution is denoted by m.

2.6.2. Relative root mean square error (RRMSE)

The relative root mean square error (RRMSE) provides a better picture of the overall fit of a distribution. It calculates each error in proportion to the size of observation thereby reducing the influence of outliers which are common features of hydrological data [19].

It is defined as

$$RRMSE = \left(\frac{\sum \left(\frac{x_i - y_i}{x_i}\right)^2}{(n-m)}\right)^{\frac{1}{2}} \quad (10)$$

2.6.3. Mean absolute deviation index (MADI)

The MADI is calculated by [1]:

$$MADI = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i - y_i}{x_i}\right| \quad (11)$$

Wherex_i are the observed values, $y_i$the predicted values and N the number of data points. The smaller the value obtained for MADI is for a distribution, the more fitting it is for the actual data [1]. Thus the distribution with smaller values of MADI indicates that it is more fitted to the observed data.

2.6.4. Probability plot correlation coefficient (PPCC)

The probability plotting correlation coefficient (PPCC) is a measure of the linearity of the probability plot [20]. It gives the correlation between the ordered observations and corresponding fitted quantiles determined by a plotting position [5]. PPCC is defined mathematically as [5]:

$$PPCC = \frac{\sum[(X_i - \bar{X})(y_i - \bar{y})]}{[\sum(x_i - \bar{x})^2 \quad \sum(y_{i-}\bar{y})^2]^{1/2}} \quad (12)$$

Where $\bar{x}$ and $\bar{y}$ represents the mean values of the observed and predicted quantiles respectively

A value of PPCC near 1 suggests that the observed data could have been drawn from the fitted distribution at a site.

3. Presentation and discussion of results

The values of the computed sample probability weighted moments (PWMs), L-moments and L-moment ratios obtained by applying equations 2, 3, 4, 5 and 6 to the observed data given in Table 3 is presented in Table 5.

The L-moments λ_1 and λ_2, their ratio ($\tau = \frac{\lambda_2}{\lambda_1}$) termed L-CV ; a dimensionless measure of variability and L-moment ratios τ_3 *and* τ_4 are the most useful quantities for summarizing probability distribution [21]. The value of λ_1(L-mean) is a measure of central tendency; λ_2(L-standard deviation) is a measure of dispersion and L-CV(τ) is the coefficient of L-variation, L-skewness (τ_3) measures whether the distribution is symmetric with respect to the dispersion from the mean and L-kurtosis (τ_4) refers to the weight of the tail of a distribution.

Table 5: Probability weighted moments and sample statistics for River Okhuwan

Parameters		Value
PWMs	b_0	65.46
	b_1	42.28
	b_2	31.07
	b_3	24.55
L- Moments	L_1	65.46
	L_2	19.10
	L_3	-1.8
	L_4	0.8
L- Moment ratios	L-CV(τ)	0.2918
	L-Skew(τ_3)	-0.094
	L-Kurtosis(τ_4)	0.0418

The negative value of L-skew indicates that the left tail is long compared to the right tail and the fact that computed L-skewness value (-0.094) lies in the range $0.05 < |L - skewness| \leq 0.150$ suggests that the observed or sample data is moderately skewed [22]. Also, L-CV value of 0.2918 indicates that the sample data is moderately variable [22]. The parameters of location (ξ), scale (α) and shape (K) of the selected distributions computed using the relevant equation given in Table 4 are as presented in Table 6.

Table 6: Estimates of probability distribution parameters at station using L-moments for Okhuwan

	K	α	ξ
GEV	0.4603	36.340	56.426
GPA	1.415	157.52	0.234
GLO	0.094	18.82	68.39

From Table 6 it can be seen that for the three distributions (GEV, GPA, GLO), the shape parameter (k) values are greater than zero indicating that the distributions have finite upper bound [10].

The estimated parameter values for the different distributions given in Table 6 were applied to their respective quantile functions defined in Table 4 and the corresponding predicted discharge results are presented in Table 7.

The predicted discharge values by the distributions were subjected to four statistical goodness-of- fit tests as described in section 2.6 in order to select the best among the candidate distributions that adequately fits the observed data at the station. The summary of the results of the goodness of fit tests are presented in Table 8 and it shows that the minimum value of RMSE, RRMSE, MADI and the value of PPCC closest to 1 is obtained by applying GPA distribution to the observed data at the station .

Table 7: Observed and the predicted peak discharge at station by different distributions

Rank	Observed discharge (m^3/s)	Predicted discharge by distributions		
		GEV	**GPA**	**GLO**
1	125	121.98	110.06	117.54
2	102.4	116.73	107.56	106.58
3	97.6	112.63	104.46	99.43
4	95.84	109.08	100.90	93.85
5	90.3	105.84	96.94	89.13
6	86.52	102.80	92.64	84.92
7	86	99.88	88.03	81.02
8	82.9	97.03	83.14	77.32
9	78.4	94.21	77.99	73.73
10	75.28	91.37	72.59	70.18
11	70	88.50	66.97	66.59
12	67	85.54	61.13	62.90
13	66	82.47	55.08	59.04
14	64	79.22	48.83	54.91
15	25.6	75.72	42.40	50.38
16	23	71.87	35.79	45.25
17	22.5	67.50	29.00	39.22
18	22.5	62.31	22.05	31.66
19	16	55.66	14.93	21.20
20	12.32	45.59	7.66	3.26

Table 8: Goodness of fit test results for the distributions at the station

Test Criteria	Distribution values		
	GEV	GPA	GLO
RMSE	29.11	8.811	10.65
RRMSE	1.316	0.23	0.7115
MADI	0.781	0.142	0.278
PPCC	0.9639	0.9685	0.2834

The overall goodness of fit of each distribution was judged using a ranking scheme by comparing the four categories of test criteria based on the relative magnitude of the statistical test results. The distribution with the lowest RMSE, lowest RRMSE, lowest MADI or highest PP CC at a station was considered the best fitting distribution [5] and was assigned a score of 3, the next best was given the score 2, while the worst was given the score 1. The overall score of each distribution was obtained by summing the individual point scores for the station and the distribution with the highest total point score was selected as the best fit distribution model for the station. The scoring scheme and the overall ranking of the distributions models at the station based on the goodness of fit tests are presented in Table 9.On the basis of the above analysis, GPA with the highest total score of 12 is selected as the best distribution for the station followed by GLO and then GEV. From Table 9 it is seen that the selected distribution outperformed the other distributions on all four test criteria.

The best fit distribution model and the other two distributions were used to obtain estimates of quantile (Q_T) for a range of return periods (2years, 5years, 10years, 25 years, 50 years, 100years and 200 years) of hydraulic and hydrologic relevance which results are presented in Table 10.

Table 9: Scoring and ranking scheme for distributions at River Okhuwan

Test Criteria	Distribution scoring		
	GEV	GPA	GLO
RMSE	1	3	2
RRMSE	1	3	2
MADI	1	3	2
PPCC	2	3	1
Total sore	5	12	7
Rank	3rd	1st	2nd

The results presented in Table 10 and show that though GPA is the best fit probability distribution model in comparison with GEV and GLO models, the GPA distribution is upper bounded at a return period of 25 years; that is to say in other words that the distribution produces very minimal differences in quantile estimates at return periods beyond 25 years. This suggests that GPA may not provide reliable estimates at return periods beyond 25 years hence for higher return periods it is safer and thus advised that the relatively higher quantile values predicted by GLO and GEV in that order be used for design purposes considering the challenge posed by the short length of data in extrapolating quantile magnitudes for return periods beyond the range of data length. For return periods greater than 40 years, the quantile estimates at the station predicted by the three distributions should be used with caution and at best as preliminary estimates because of the short length of sample data used for the study (20 years). This is in conformity with the prescription of Flood Studies Report (FSR) [23] that satisfactory at-site quantile estimates can be achieved for return period (T) of up to $T = 2N$ years (N being the length of Annual Maximum Series (AMS) in years) from an AMS set of N years length and where $T \gg 2N$, FSR recommends that a regional flood frequency analysis by index flood approach be carried out.

4. Conclusions and recommendations

Arising from the results of this study, the following conclusions and recommendations are made.

(1) L –moments and L –moment ratios are useful for summarizing statistical properties of hydrological data and can be used for estimation of distribution parameters and selection of best fit distribution.

(2) The selected probability distribution models(GLO, GEV and GPA) can be utilized to predict the flood quantile magnitudes (Q_T) at the station for return periods of interest in hydraulic and hydrologic assessments and design.

(3) The selected probability distribution models (GLO, GEV and GPA) have shape parameter (k) values greater than zero and hence have finite upper bound.

(4) The best fit probability distribution model for analyzing the annual maximum flood series at the site is Generalized Pareto distribution (GPA) for lower return periods (up to 25 years) but for higher return periods (between 25 and 200 years), it is safer to use the relatively higher values predicted by GLO and GEV in that order.

(5) For return periods greater than 40 years, the values of quantile estimates at the station predicted by the three distributions should be used with caution and at best as preliminary estimates because of the short length of sample data used for the study (20 years). This is especially important as there is no general guidance for extrapolating distribution beyond twice the available record length [5].

Table 10: Quantile estimates for station for selected return periods using different distributions

Return Period ,T(years)	2	5	10	25	50	100	200
F = (1-1/T)	0.5	0.8	0.9	0.96	0.98	0.99	0.995
$X(F) = Q_T (m^3/s)$ GPA	69.81	100.14	107.27	110.38	111.12	111.39	111.49
$X(F) = Q_T (m^3/s)$ GEV	68.68	95.80	107.35	117.26	122.27	125.87	128.48
$X(F) = Q_T (m^3/s)$ GLO	68.39	92.85	105.75	120.09	129.73	138.62	146.87

(6) In view of the challenge posed by short record length and data gaps to reliable flood quantile estimation, it is recommended that efforts should be intensified nationally in the collection of hydrological data

(7) It is further recommended that the density of stream gauging networks in river basins in the country should be increased to enhance national water resources planning and development.

References

[1] Ahmad, U. N., Shabri, A. and Zakaria. Z.A, "Flood frequency Analysis of Annual Maximum Flows Using L- Moments and TL-Moments Approach", *Applied Mathematics Sciences,* Vol. 5, Number.5, 2011, pp. 243 -253.

[2] Abida, H. and Elliouze, M. "Probability distribution of flood flows in Tunisia", *Hydrology and Earth System Sciences*, Vol.12, 2008, pp.703 – 713.

[3] Thorvat, A.R and Mujumdar, M.M "Design flood estimation for Upper Krishna basin though RFFA", *International Journal of Engineering, Science and Technology*, Vol.3 Number.6, 2011, pp.5252 -5259.

[4] World Meteorological Organization (WMO), *Guide to Hydrological Practices Volume II: Management of Water Resources and Application of Hydrological Practices, WMO No.168,* World Meteorological Organization, Geneva. Switzerland, 2009.

[5] Abdul Karim, M. and Chowdhury, J.U. "A comparison of four distributions used in flood frequency analysis in Bangladesh", H*ydrological Sciences Journal,* Vol.40, Number 1. (February), 1995, pp.55-66.

[6] World Meteorological Organization (WMO), *Manual on Low flow Estimation*, Operational Hydrology ReportNo.50, World Meteorological Organization, Geneva. Switzerland, 2008.

[7] Central Water Commission of India *Development of Hydrological Design Aids (Surface water) under Hydrology Project II:* State of the Art Report submitted by Consulting Engineering Services (India) in Association with HR Wallingford, 2010.

[8] Hosking, J.R.M."L-moments: analysis and estimation of distributions using linear combinations of order statistics", *Journal of Royal Statistical society of Britain*, Vol.52, Number.1, 1990, pp.105 -124.

[9] Cunnane, C., *Statistical distributions for frequency analysis*, Operational Hydrology Report, World Meteorological Organization (WMO), Geneva. 1989.

[10] Hosking, J.R.M and Wallis, J.R. *Regional flood frequency analysis: An approach based on L-moments*, Cambridge University Press, New York, 1997.

[11] Robson, A. and Reed, D. *Flood Estimation Handbook, Volume 3, Statistical Procedures for flood frequency estimation*, Institute of Hydrology, Wallingford, Oxford shire, United Kingdom. 1999.

[12] Chadwick, A., Morfett, J. and Borthwick, M. *Hydraulics in Civil and Environmental Engineering,*Spon Press, London.2004.

[13] Stedinger, J.R., Vogel, R.M. and Foufoula- Georgiou. E. "Frequency analysis of extreme events", In *Handbook of Applied Hydrology*, D.R Maidment (Ed.),pp. 1-66 McGraw – Hill, New York,1993.

[14] Greenwood, J.A., Landwehr, J.M., Matalas, N.C. and Wallis, J.R. "Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form". *Water Resources Research*, Vol.15, Number.5, 1979, pp. 1049 – 1054.

[15] Vogel, R.M., McMahon, T.A.andChiew, F.H.S. "Flood flow frequency model selection in Australia", *Journal of Hydrology*, *Amsterdam,* Vol.146, (1993b), pp.421 – 449.

[16] Hosking, J.R.M. "The four parameter kappa distribution", *IBM Journal of Research and Development*. Vol. 38, Number.3, 1994, pp. 251 – 258.

[17] Benin OwenaRiver Basin Development Authority(BORBDA),Daily flow records of River Okhuwan at Ugonoba. Hydrology unit, BeninOwena River Basin Development Authority, Benin City. 2012.

[18] Ojha, C.S.P., Berndtsson, R and Bhunya, P., *Engineering Hydrology*, Oxford University Press, New Delhi, India, 2008.

[19] Tao, D.V., Nguyen, V.T. and Bourque, A., "On selection of probability distributions for representing extreme precipitations in Southern Quebec", *Annual conference of the Canadian Society for Civil Engineering*, Montreal, Quebec, Canada. June 5 – 8, 2008, pp 1-8.

[20] Filliben, J.J "The probability plot correlation test for normality", *Technometrics*, Vol.17, Number 1, 1975, pp.111 – 117.

[21] Maleki- Nezhad, H., "Regional flood frequency analysis using L-moments approach". *7th International Congress on Civil Engineering*, Tehran Iran, May 8-10, 2006, pp.1-7.

[22] *L-RAP Users Manual,* MGS Software LLC, USA. 2011

[23] Natural Environmental Research Council (NERC) (1975). Flood Studies Report, Vol.1-5, London.

[24] Benin Owena River Basin Development Authority (BORBDA) (2007). Benin Owena River Basin Authority (1999 – 2000) Hydrological Year Book (Vol. V), Benin City, Nigeria.