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

O.C. Izinyon\(^1,2\), 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 25 years 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 has very important for such cases. The number of parameters are large [7]. 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° 51’E and Latitude 6° 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

<table>
<thead>
<tr>
<th>Station</th>
<th>R.Okhuwan at Ugonoba</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td></td>
</tr>
<tr>
<td>Longitude</td>
<td>5° 51' E</td>
</tr>
<tr>
<td>Latitude</td>
<td>6° 19' N</td>
</tr>
<tr>
<td>Drainage Area (Km²)</td>
<td>245</td>
</tr>
<tr>
<td>Coefficient of variation</td>
<td>0.5114</td>
</tr>
<tr>
<td>Coefficient of skew (C_s)</td>
<td>-0.3192</td>
</tr>
<tr>
<td>Record length (yrs)</td>
<td>20</td>
</tr>
</tbody>
</table>

*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]

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Probability density function and/or Cumulative distribution function</th>
<th>Range</th>
<th>Moments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generalized Extreme Value (GEV)</td>
<td>$F_x(X) = \exp\left{-\left[1 - \frac{k(x - \xi)}{\alpha}\right]^{\frac{1}{k}}\right}$</td>
<td>$\sigma_x^2$ exist for $-0.5 &lt; k$</td>
<td>$\mu_x = \xi + \left(\frac{\alpha}{k}\right)\left[1 - \Gamma(1 + k)\right]$</td>
</tr>
<tr>
<td></td>
<td>$y = \left[1 - \frac{k(x - \xi)}{\alpha}\right]^{\frac{1}{k}}$ for $k \neq 0$</td>
<td>$f_x = \left(\frac{k}{\alpha}\right)^{\frac{1}{k}}\left(1-y\right)^{\frac{1}{k}-1}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$f_x = \left(\frac{k}{\alpha}\right)^{\frac{1}{k}}\left(1+y\right)^{-\frac{1}{k}}$ for $k &lt; 0$</td>
<td>$F_x(x) = 1 - \left(1 + \frac{\alpha}{k}\right)^{\frac{1}{k}}$</td>
<td></td>
</tr>
<tr>
<td>Generalized Logistic (GLO)</td>
<td>$y = \exp\left[-\left(\frac{x - \xi}{\alpha}\right)\right]$</td>
<td>$y = \exp\left[-\left(\frac{x - \xi}{\alpha}\right)\right]$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$f(x) = \left(\frac{k}{\alpha}\right)^{\frac{1}{k}}\left(1+y\right)^{-\frac{1}{k}}$ for $k &lt; 0$</td>
<td>$F_x(x) = \left[1 - \frac{\alpha}{k}\right]^{-\frac{1}{k}}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$f(x) = \left(\frac{k}{\alpha}\right)^{\frac{1}{k}}\left(1+y\right)^{-\frac{1}{k}}$ for $k &gt; 0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generalized Pareto (GPA)</td>
<td>$y = \exp\left[-\left(\frac{x - \xi}{\alpha}\right)\right]$</td>
<td>$y = \exp\left[-\left(\frac{x - \xi}{\alpha}\right)\right]$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$f(x) = \left(\frac{k}{\alpha}\right)^{\frac{1}{k}}\left(1+y\right)^{-\frac{1}{k}}$ for $k &lt; 0$</td>
<td>$F(x) = 1 - \left[1 - \frac{k(x - \xi)}{\alpha}\right]^\frac{1}{k}$</td>
<td>$\mu_x = \xi + \frac{\alpha}{k}\left[1 - \Gamma(1 + k)\right]$</td>
</tr>
<tr>
<td></td>
<td>$f(x) = \left(\frac{k}{\alpha}\right)^{\frac{1}{k}}\left(1+y\right)^{-\frac{1}{k}}$ for $k &gt; 0$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

See Ahmad et al (1988) for $\sigma^2_x$.
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 the Generalized Pareto distribution (GPA) is useful for modeling events 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:

\[
\beta_r = E \{ X[F_2(x)^r] \} \quad (1)
\]

Where \( \beta_r \) is the rth order of PWM and \( F_2(X) \) is the cumulative distribution function (CDF) of \( X \).

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

\[
\beta_0 = \frac{1}{n} \sum_{j=1}^{n} X(j) ; \quad (2a)
\]

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

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

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

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 \((\lambda_1, \lambda_2, \lambda_3 \text{ and } \lambda_4)\) 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–moment ratios used for expressing the parameter estimates are given as:

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

\[
L–\text{skewness (L-skew)} = \frac{\lambda_3}{\lambda_2} \quad (5)
\]

\[
L–\text{kurtosis (L-kurt)} = \frac{\lambda_4}{\lambda_2} \quad (6)
\]

The sample estimates of L–moments \((l_1, l_2, l_3 \text{ and } l_4)\) are calculated by replacing \(\beta_0, \beta_1, \beta_2 \text{ and } \beta_3\) in equation 3 with \(b_0, b_1, b_2 \text{ and } b_3\) respectively [12].

The L–moments \(\lambda_1 \text{ and } \lambda_2\), the L–CV, \(\tau\) and L–moment ratios \(\tau_3\) and \(\tau_4\) 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, \(b_0, b_1, b_2, \text{ and } b_3\) for the sample data were obtained using equations 2a, 2b, 2c and 2d respectively [12].

The sample estimates of L–moments \((l_1, l_2, l_3 \text{ and } l_4)\) were calculated by replacing \(\beta_0, \beta_1, \beta_2 \text{ and } \beta_3\) in equation 3a, 3b, 3c and 3d with the obtained \(b_0, b_1, b_2 \text{ and } b_3\) 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]:

\[
F(Q_i) = \frac{m}{n+1}
\]

(7)

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

<table>
<thead>
<tr>
<th>Year</th>
<th>Maximum flow(m(^3)/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989</td>
<td>25.6</td>
</tr>
<tr>
<td>1990</td>
<td>23.0</td>
</tr>
<tr>
<td>1991</td>
<td>22.5</td>
</tr>
<tr>
<td>1992</td>
<td>16.0</td>
</tr>
<tr>
<td>1993</td>
<td>12.32</td>
</tr>
<tr>
<td>1994</td>
<td>22.5</td>
</tr>
<tr>
<td>1995</td>
<td>64</td>
</tr>
<tr>
<td>1996</td>
<td>75.28</td>
</tr>
<tr>
<td>1997</td>
<td>78.40</td>
</tr>
<tr>
<td>1998</td>
<td>82.9</td>
</tr>
<tr>
<td>1999</td>
<td>70.0</td>
</tr>
<tr>
<td>2000</td>
<td>67.0</td>
</tr>
<tr>
<td>2001</td>
<td>86.0</td>
</tr>
<tr>
<td>2002</td>
<td>86.52</td>
</tr>
<tr>
<td>2003</td>
<td>66.0</td>
</tr>
<tr>
<td>2004</td>
<td>90.3</td>
</tr>
<tr>
<td>2005</td>
<td>95.84</td>
</tr>
<tr>
<td>2006</td>
<td>97.6</td>
</tr>
<tr>
<td>2007</td>
<td>102.4</td>
</tr>
<tr>
<td>2008</td>
<td>125</td>
</tr>
</tbody>
</table>

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. \( \xi \) is the location parameter, \( \alpha \) is the scale parameter and \( K \) is the shape parameter.

\[
\begin{align*}
\text{GEV} & \quad X(F) = \xi + \frac{\alpha}{K} \left[ 1 - (-\ln F)^{\frac{1}{K}} \right] \\
\text{GLO} & \quad X(F) = \xi + \frac{\alpha}{K} \left[ 1 - \left( \frac{1-F}{F} \right)^{\frac{1}{K}} \right] \\
\text{GPA} & \quad X(F) = \xi + \frac{\alpha}{K} \left[ 1 - (1-F)^{K} \right]
\end{align*}
\]

Where

\[
\begin{align*}
\alpha &= \frac{l_2 K}{\Gamma(1+K)\Gamma(1-2^{-K})} \\
\xi &= l_1 + \frac{\alpha(\Gamma(1+K) - 1)}{K} \\
K &= 7.8590 \xi + 2.9554 \xi^2 \\
\xi &= \frac{2}{3 + \tau_3} - \ln \frac{2}{\ln 3} \\
\alpha &= \frac{l_2}{\Gamma(1+K)\Gamma(1-K)} \\
\xi &= l_1 + \frac{(l_2 - \alpha)}{K} \\
K &= -\tau_3 \\
\alpha &= l_2[(K+1)(K+2)] \\
\xi &= l_1 - l_2(K+2) \\
K &= \frac{4}{\tau_3 + 1 - 3}
\end{align*}
\]
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:

\[ RMSE = \left( \frac{\sum (x_i - y_i)^2}{n - m} \right)^{1/2} \]  

(9)

where \( x_i, i=1,\ldots,n \) are the observed values and \( y_i, i=1,\ldots,n \) 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

\[ RRMSE = \left( \frac{\sum \frac{(x_i - y_i)^2}{x_i}}{n - m} \right)^{1/2} \]  

(10)

2.6.3. Mean absolute deviation index (MADI)

The MADI is calculated by [1]:

\[ MADI = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{y_i - y}{x_i} \right| \]  

(11)

Where \( x_i \) 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]:

\[ PPCC = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \]  

(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 \( \lambda_1 \) and \( \lambda_2 \), their ratio \( \tau = \frac{\lambda_2}{\lambda_1} \) termed L-CV ; a dimensionless measure of variability and L-moment ratios \( \tau_3 \) and \( \tau_4 \) are the most useful quantities for summarizing probability distribution [21]. The value of \( \lambda_1 \) (L-mean) is a measure of central tendency; \( \lambda_2 \) (L-standard deviation) is a measure of dispersion and L-CV(\( \tau \)) is the coefficient of L-variation, L-skewness (\( \tau_3 \)) measures whether the distribution is symmetric with respect to the dispersion from the mean and L-kurtosis (\( \tau_4 \)) refers to the weight of the tail of a distribution.
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 $0.05 < |L - \text{skewness}| \leq 0.150$ 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 ($\xi$), scale ($\alpha$) and shape ($k$) of the selected distributions computed using the relevant equation given in Table 4 are as presented in Table 6.

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].

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

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PWMs</td>
<td></td>
</tr>
<tr>
<td>$b_0$</td>
<td>65.46</td>
</tr>
<tr>
<td>$b_1$</td>
<td>42.28</td>
</tr>
<tr>
<td>$b_2$</td>
<td>31.07</td>
</tr>
<tr>
<td>$b_3$</td>
<td>24.55</td>
</tr>
<tr>
<td>L-Moments</td>
<td></td>
</tr>
<tr>
<td>$L_1$</td>
<td>65.46</td>
</tr>
<tr>
<td>$L_2$</td>
<td>19.10</td>
</tr>
<tr>
<td>$L_3$</td>
<td>-1.8</td>
</tr>
<tr>
<td>$L_4$</td>
<td>0.8</td>
</tr>
<tr>
<td>L-Moment ratios</td>
<td></td>
</tr>
<tr>
<td>L-CV($\tau$)</td>
<td>0.2918</td>
</tr>
<tr>
<td>L-Skew($\tau_3$)</td>
<td>-0.094</td>
</tr>
<tr>
<td>L-Kurtosis($\tau_4$)</td>
<td>0.0418</td>
</tr>
</tbody>
</table>

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 PPCC 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 (2 years, 5 years, 10 years, 25 years, 50 years, 100 years 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

<table>
<thead>
<tr>
<th>Test Criteria</th>
<th>GEV</th>
<th>GPA</th>
<th>GLO</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>RRMSE</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>MADI</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>PPCC</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Total score</td>
<td>5</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>Rank</td>
<td>3rd</td>
<td>1st</td>
<td>2nd</td>
</tr>
</tbody>
</table>

The results presented in Table 10 and show that although 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 \( T > 2N \), 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

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


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