Simple to Use Microsoft Excel Template for Estimating the Parameters of Some Selected Probability Distribution Model by Method of L-Moment

Idowu Rudolph Ilaboya, Ebierin Akpoebidimiyen Otuaro
1. Department of Civil Engineering, Faculty of Engineering, PMB 1154, University of Benin, Benin City, Edo State, Nigeria
2. Department of Civil Engineering, Faculty of Engineering, Delta State University, Abraka

Abstract
The focus of this research was to design a simple to use Microsoft excel algorithm that will aid in the estimation of the parameters of generalized extreme value probability distribution (GEV), generalized logistics probability distribution (GLO) and generalized pareto probability distribution (GPA), calculate the predicted rainfall/discharge based on L-moment procedures and compute the quantile estimates at various return periods. The algorithm was designed based on the underlying mathematics of L-moment and has the capacity to handle forty (40) year’s annual maximum series of either rainfall or discharge data which must first be ranked in ascending order of magnitude. Basic descriptive statistics such as the sample mean, variance, standard deviation, skewness, kurtosis, coefficient of variation have been built into the algorithm. Other exciting features include; the computation of Probability weighted moment parameters (b0, b1, b2 and b3), L-Moment values (ƛ1, ƛ2, ƛ3 and ƛ4), L-Moment ratio values (T2, T3 and T4), and goodness of fit statistics (RRMSE, RMSE, MAE, MADI and PPCC). Others include; the shape (k), scale (α) and location (ξ) parameters of GEV, GPA and GLO probability distributions. To test the performance of the algorithm, forty (40) year’s annual maximum rainfall data from Benin City was used. Basic time series analysis such as test of normality, test of homogeneity and outlier detection was conducted to ensure that the data used are adequate and suitable. Results obtained revealed that generalized logistics probability distribution GLO was the best fit distribution model for analyzing the annual maximum rainfall series at the study site. The predicted rainfall quantile magnitude (Qt) based on the GLO model ranges from 425.877mm at 2years return period to 762.759mm at 200 years return period. The coefficient of determination (r2) for the observed versus predicted rainfall based on the best fit model was observed to be 0.9793. It was thereafter concluded that L-moments and L-moment ratios are useful summary statistics for analyzing rainfall data.

Keywords: L-moments, probability distribution, normality test, goodness of fit statistics, coefficient of variation.
DOI: 10.7176/CER/11-9-05
Publication date: October 31st 2019

1.0 Introduction
Estimation of extreme flood discharge of known return period is paramount in the design of hydraulic structures such as culverts, dams, bridges and drainage systems. Owing to the stochastic nature of the hydrologic phenomena that governs extreme flood discharge, it is fundamental that we investigate most hydrologic processes such as rainfall and droughts by simply analyzing their records of observations (Ehiorobo & Izinyon, 2013). Effective analysis and determination of extreme flood discharge requires the use of statistical frequency analysis or fitting of probability distribution to the series of recorded annual maximum discharge (AMD) (Vivekanandan, 2015; Sharma & Singh, 2010). One of the widely used statistical frequency analysis methods is univariate frequency analysis technique. Univariate frequency analysis is widely used for analyzing hydrologic data, including rainfall characteristics, peak discharge series and low flow record of observations. Univariate frequency analyses are primarily employed in the estimation of exceedance probabilities and variable magnitudes. The basic assumption is that the data to be used must be satisfactorily homogeneous otherwise; the estimated probabilities or variable magnitude will be inaccurate (ECOST, 2012). The versatility of statistical frequency analysis makes it the most commonly used procedure for the analysis of flood data. Due to its wide application in the analysis of flood data, univariate statistical frequency analysis is sometime designated flood frequency analysis (FFA). A number of probability distributions such as Generalized Extreme Value (GEV), Generalized Pareto (GPA), Generalized Logistics (GLO), Gumbel distribution and Normal distribution are used in flood frequency analysis (Hosking & Wallis 1997). To employ any type of probability distribution for flood frequency analysis, the parameters of the distribution must first be estimated. Different types of probability distribution parameters estimation methods exist, namely; least square regression method (LSR), maximum likelihood estimation (MLE), method of moments (MOM) and method of L-moment. Of the four parameters estimation method, method of moments (MOM) and method of L-moment are widely used owing to their high level of sensitivity to rainfall and runoff data (Ahmad et al., 2011). Method of moments (MOM) has found a wide range of applicability in recent time based on it’s used for the determination of parameters of different probability distributions. In some cases, it is always very difficult.
to assess the exact information about the shape of a distribution that is conveyed by its third and higher order moments (Landwehr, 1979). In addition, for small sample size, the numerical values of sample moments can be very different from those of the probability distribution from which the sample was drawn (Ehiorobo & Izinyon, 2013). On account of these limitations of MOM, alternative approach such as L-moments (LMO) was introduced to accurately estimate the parameters of probability distributions. L-Moment is a dramatic improvement over conventional product moment statistics for characterizing the shape of a probability distribution and estimating the distribution parameters, particularly for environmental data where sample sizes are commonly very small (Hosking, 1990; Izinyon & Ehiorobo, 2015).

Although, L-moment is more reliable compared to other conventional methods, but the underlying equations needed to obtain the L-moment parameters are usually very complex and requires an in-depth knowledge of mathematics. In addition, most of the tools (software) that can be employed in the determination of L-moment parameters such as L-RAP, DHI software are not only very expensive to acquire, they are also not user friendly. The implications are that so many researchers are not able to do much in this area of studies since they are either not able to purchase the needed software or are unable to solve the complex mathematical equations required to compute the L-moment parameters. The purpose of this research therefore, was to design an algorithm using Microsoft Excel that will help estimate the parameters of three probability distributions namely; Generalized Extreme Value (GEV), Generalized Logistics (GLO) and Generalized Pareto (GPA) using L-Moment procedure.

2.0 L-Moment Theory and Statistics

L-moments can be obtained by considering linear combinations of the observation in a sample of data that has been arranged in ascending order. Consider measurement of the shape of a distribution, given a small sample drawn from the distribution. Denote by \( X_{1:1} \leq X_{2:n} \leq \ldots \ldots X_{n:n} \) (Hosking & Wallis, 1997; Eregno, 2014).

The basic steps in the determination of L-Moment statistics are described below;

**Step One: Computation of probability weighted moments of distribution (PWMs)**

Probability weighted moments is needed for the calculation of L-moment. The data must first be ranked in ascending order of magnitude, thereafter; the following equations proposed by Cunnnane, 1989 can thus be applied

\[
b_0 = \frac{1}{N} \sum_{j=1}^{n} X_{(j:n)}
\]

\[
b_1 = \frac{1}{N} \sum_{j=2}^{n} X_{(j:n)} [(j-1)/(n-1)]
\]

\[
b_2 = \frac{1}{N} \sum_{j=3}^{n} X_{(j:n)} [(j-1)(j-2)/[(n-1)(n-2)]
\]

\[
b_3 = \frac{1}{N} \sum_{j=4}^{n} X_{(j:n)} [(j-1)(j-2)(j-3)/[(n-1)(n-2)(n-3)]
\]

Where;

\( X_{(1)} \) represent the ranked annual maximum series in which \( X_{(1)} \) is the smallest precipitation or stream flow data and \( X_{(n)} \) is the largest. The parameters \( (b_0, b_1, b_2 \text{ and } b_3) \) can easily be determined by using the developed Microsoft excel algorithm.

**Step Two: Computation of L-Moment Values**

L-moment values are easily calculated in terms of probability weighted moment (PWMs). In particular, the first four L-moment values are given as follows (Hosking & Wallis, 1997).

\[
\lambda_1 = L_1 = b_0
\]

\[
\lambda_2 = L_2 = (2b_1 - b_0)
\]

\[
\lambda_3 = L_3 = (6b_2 - 6b_1 + b_0)
\]

\[
\lambda_4 = L_4 = (20b_3 - 30b_2 + 12b_1 - b_0)
\]

The parameters \( \lambda_1 = L_1 ; \lambda_2 = L_2 ; \lambda_3 = L_3 ; \lambda_4 = L_4 \) can easily be determined by using the developed Microsoft excel algorithm that requires forty year’s annual maximum monthly rainfall or discharge data.

**Step Three: Computation of L-Moment Ratio**

L-Moment ratio used for expressing the parameter estimates are as follows (Hosking and Wallis, 1997).

\[
L - CV \quad (\text{Coefficient of variability}) = (\tau_2)
\]

\[
L - \text{Skewness} = (\tau_3)
\]
$$L - Kurtosis = (\tau_4)$$

$L$-Cv is a dimensionless measure of variability. For a distribution or sample data that only has positive values, $L$-Cv is normally in the range of $0 < |L-Cv| < 1$. Negative values of $L$-Cv are only possible if the at-site mean has a negative value (Sanjib, 2016; Herlina, 2015). The descriptions of the relative magnitude of variability is presented in Table 1

**Table 1: Magnitude of $L$-Cv**

| Range of $L$-Cv | Descriptions of Relative Magnitude of $L$-Cv |
|----------------|---------------------------------------------|
| $0.000 < |L-Cv| < .025$ | minimal variability |
| $.025 < |L-Cv| < .075 | minor variability |
| $.075 < |L-Cv| < .150 | moderate variability |
| $.150 < |L-Cv| < .400 | large variability |
| $.400 < |L-Cv| | very large variability |

$L$-Skewness is a dimensionless measure of asymmetry, which may take on positive or negative values. For a distribution or sample data, $L$-skewness is in the range $0 < |L$-Skewness$| < 1$ (CEH, 2001). The descriptions of the relative magnitude of asymmetry is presented in Table 2

**Table 2: Relative magnitude of Asymmetry**

| Magnitude of $L$-skewness | Descriptions of Relative Magnitude of $L$-Skewness |
|---------------------------|---------------------------------------------------|
| $L$-skewness = 0.0 | symmetrical distribution |
| $0.000 < |L - skewness| \leq 0.050$ | minor skewness |
| $0.050 < |L - Skewness| \leq 0.150$ | moderate skewness |
| $0.150 < |L - Skewness| \leq 0.300$ | large skewness |
| $0.300 < |L - Skewness| | very large skewness |

$L$-kurtosis refers to any measure of the "peakedness" of the probability distribution of a real-valued random variable. The parameters ($\tau_2$, $\tau_3$, and $\tau_4$) are computed using the formula below (Hosking & Wallis, 1997; Gubareva & Gartsman, 2010).

$$\tau_2 = \frac{\lambda_2}{\lambda_1} = \frac{L_2}{L_1}$$  \hspace{1cm} (12)

$$\tau_3 = \frac{\lambda_3}{\lambda_2} = \frac{L_3}{L_2}$$  \hspace{1cm} (13)

$$\tau_4 = \frac{\lambda_4}{\lambda_3} = \frac{L_4}{L_3}$$  \hspace{1cm} (14)

The parameters ($\tau_2$, $\tau_3$, and $\tau_4$) can easily be determined by using the developed Microsoft excel algorithm that requires forty year’s annual maximum monthly rainfall or discharge data.

### 2.1 Advantages of L-Moment

The main advantage of L-moment over conventional moments is that L-moments, being linear functions of the data, suffer less from the effects of sampling variability and are more robust compared to conventional moments in handling outliers. In addition (Hosking & Wallis, 1997). Some of the underlying simplicity of L-Moments are:

i. L-moment is based on linear combination of data that have been arranged in ascending order of magnitude.

   It provides an advantage as it is easier to work with, and more reliable since it is less sensitive to outliers.

ii. The method of L-moment calculates more accurate parameter than method of moment (MOM) technique especially for smaller sample size.

iii. MOM techniques only apply to limited range of parameters, whereas L-moment can be more widely used, and are also nearly unbiased
iv. L-Moment allow for the generation of ratio diagrams which are helpful in identifying the distribution properties of highly skewed data

v. L-Moment is a dramatic improvement over conventional product moment statistics for characterizing the shape of a probability distribution and estimating the distribution parameters, particularly for environmental data where sample sizes are commonly very small

vi. In practice, an in-depth knowledge and accurate estimation of the L-moment ratios of L-Cv ($T_2$), L-Skewness ($T_3$) and L-kurtosis ($T_4$) is a key determinant in quantifying the success of the regional frequency analysis in computing quantile estimates for selected stations (Hosking & Wallis, 1997).

3.0: Development and Execution of the Algorithm

The Microsoft Excel algorithm was designed to capture three probability distributions, namely: Generalized Extreme Value (GEV), Generalized Logistics (GLO) and Generalized Pareto (GPA). The GEV distribution is a flexible distribution which has been found to fit flood and rainfall extremes in a variety of environments. The underlying equations including the quantile function ($x_p$) corresponding to the non-exceedance probability ($p$) and the return period ($T$) corresponding to the non-exceedance probability ($p$) have been incorporated into the algorithm. The probability density functions and or cumulative distribution functions, range and moments for the selected distributions captured by the algorithm are presented in Table 3 and 4 respectively while the goodness of fit statistics captured by the algorithm is presented in Table 5.

Table 3: Probability distribution model captured by the algorithm

| Distribution                  | Parameters/Range                                      | Probability density function ($f(x)$)                                  | Cumulative distribution function ($F(x)$)                                      | Quantile function ($x_p$) |
|------------------------------|-------------------------------------------------------|------------------------------------------------------------------------|----------------------------------------------------------------------------------|---------------------------|
| Generalized Extreme Value (GEV) | Parameters: Location ($\xi$), Scale ($\alpha$), Shape ($\kappa$) | $f(x) = \frac{1}{\alpha} \left[ 1 - \left( 1 - \frac{x - \xi}{\kappa} \right)^{\alpha} \right]^{\alpha-1}$ | $F(x) = \left( 1 - \frac{1 - \left( 1 - \frac{x - \xi}{\kappa} \right)^{-\alpha}}{\alpha} \right)^{\frac{1}{\alpha}}$ | $x_p = \frac{\kappa}{\alpha} \left[ \left( 1 - \frac{1}{p} \right)^{\alpha} - 1 \right]$ |
| Generalized Pareto (GPA)      | Parameters: Location ($\xi$), Scale ($\alpha$), Shape ($\kappa$) | $f(x) = \frac{1}{\alpha} \left( 1 - \frac{x - \xi}{\kappa} \right)^{\alpha-1}$ | $F(x) = 1 - \left[ 1 - \frac{x - \xi}{\kappa} \right]^\gamma$ | $x_p = \frac{\kappa}{\alpha} \left[ \left( 1 - \frac{1}{p} \right)^{\alpha} - 1 \right]$ |
| Generalized Logistics (GLO)   | Parameters: Location ($\xi$), Scale ($\alpha$), Shape ($\kappa$) | $f(x) = \frac{1}{\alpha} \left[ 1 - \left( 1 - \frac{x - \xi}{\kappa} \right)^{\alpha} \right]^{\alpha-1}$ | $F(x) = 1 - \left( 1 - \frac{x - \xi}{\kappa} \right)^{\alpha}$ | $x_p = \frac{\kappa}{\alpha} \left[ \left( 1 - \frac{1}{p} \right)^{\alpha} - 1 \right]$ |
Table 4: L-Moment parameter estimates for probability distributions captured by the algorithm

| Distribution                  | Quantile function \( x_p \)                                                                 | L-Moment Parameter Estimates Equations |
|-------------------------------|--------------------------------------------------------------------------------------------|---------------------------------------|
| Generalized Extreme Value \(\text{GEV})\) | \[ x_p = \xi + \frac{\alpha}{k} \left(1 - \left[-\ln(F)\right]^k\right) \] | \[ \alpha = \frac{l_2 K}{\Gamma(1 + K)\Gamma(1 - 2^{-K})} \]  
\[ \xi = l_1 + \frac{\alpha(\Gamma(1 + K) - 1)}{K} \]  
\[ K = 7.8590C + 2.9554C^2 \]  
\[ C = \frac{2}{3 + \tau_3} = \frac{\ln 2}{\ln 3} \] |
| Generalized Pareto \(\text{GPA})\) | \[ x_p = \xi + \frac{\alpha}{k} \left[1 - \left(1 - F\right)^{k}\right] \] | \[ \alpha = l_2\left[(K + l)(K + 2)\right] \]  
\[ \xi = l_1 = l_2(K + 2) \]  
\[ K = \left(1 - 3\tau_3\right) \] |
| Generalized Logistics \(\text{GLO})\) | \[ x_p = \xi + \frac{\alpha}{k} \left[1 - \left(\frac{1 - F}{F}\right)^{k}\right] \] | \[ \alpha = \frac{l_2}{\Gamma(1 + K)\Gamma(1 - K)} \]  
\[ \xi = l_1 + \frac{(l_2 - \alpha)}{K} \]  
\[ K = -\tau_3 \] |

Table 5: Goodness of fit statistics captured by the algorithm

| Statistics                              | Equation                                      |
|-----------------------------------------|-----------------------------------------------|
| Root Mean Square Error (RMSE)           | \[ \text{RMSE} = \left(\frac{\sum (x_i - y_i)^2}{n - m}\right)^{\frac{1}{2}} \] |
| Relative Root Means Square Error (RRMSE) | \[ \text{RRMSE} = \left(\frac{\sum \left\{\frac{x_j - y_j}{x_i}\right\}^2}{n - m}\right)^{\frac{1}{2}} \] |
| Mean Absolute Deviation Index           | \[ \text{MADI} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_j - y_j}{x_i}\right| \] |
| Maximum Absolute Error                  | \[ \text{MAE} = \max \left\{|x_i - y_i|\right\} \] |
| Probability Plot Correlation Coefficient | \[ \text{PPCC} = \left[\frac{\sum (x_j - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}\right]^{\frac{1}{2}} \] |

Root mean square error (RMSE), relative root means square error (RRMSE) and maximum absolute deviation index (MADI) were selected since they can adequately assess the fitted distribution at a site. They possess an added advantage of being able to summarize the deviation between observed precipitation and predicted precipitation. In addition, RRMSE can also provide a better picture of the overall fit of a distribution as it calculates each error in proportion to the size of observation thus helping to eliminate or reduce the effects of bias commonly associated with hydrological data (Tao et al, 2008).
To develop the excel algorithm, the underlying mathematics of L-moment were employed. The complex equations were first digested and rewritten in a more simplified format that is usable by Microsoft Excel. Two algorithms were developed, namely; **algorithm one** and **algorithm two**. Algorithm one uses the ranked discharge or precipitation data (Xi) (40 years annual maximum data) as the input data to calculates the following parameters:

i. Variance  
ii. Standard deviation  
iii. Skewness  
iv. Kurtosis  
v. Coefficient of variation (CV)  
vi. Probability weighed moment’s parameters (b0, b1, b2 and b3)  
vii. L-Moment values (ƛ1, ƛ2, ƛ3 and ƛ4)  
viii. L-Moment ratio values (Ʈ2, Ʈ3 and Ʈ4)  
ix. Scale parameter (α) of GEV, GPA and GLO probability distribution  

Algorithm two uses the 40 years precipitation or discharge data including the calculated shape parameter (k), scale parameter (α) and location parameter (ξ) from algorithm one as input data to calculates the following parameters:

i. The predicted rainfall or discharge values (yi)  
ii. The goodness of fit statistics (RMSE, RRMSE, MADI, MAE and PPCC)  
iii. Quantile estimates at various return periods  
iv. Exceedance and non-Exceedance probabilities

### 3.1: Computation of L-moment parameters Using Algorithm One

The following steps are involved in using algorithm one to compute the parameters of GEV, GLO and GPA probability distributions.

i. The algorithm is available on request. Table 6a and 6b shows a section of the algorithm  
ii. The algorithm requires 40 years annual maximum data (precipitation or discharge) which must be ranked in ascending order of magnitude. The data must be prepared in excel, copy and paste on cell B2 – B41 of Algorithm one  
iii. Descriptive statistics of the data presented to the algorithm will be gotten immediately from cell D4 – D10  
iv. The probability weighted moment statistics (b0, b1, b2 and b3) will be gotten immediately from cell I10 - I13  
v. The L-moment values (ƛ1 = L1 ; ƛ2 = L2 ; ƛ3 = L3 ; ƛ4 = L4) will be gotten immediately from cell K10 - K13  
vi. The L-moment ratio values (τ2, τ3 and τ4) will be gotten immediately from cell M10 – M12  

### 3.2: Predicted values and Quantile Estimates based on selected Return Period Using Algorithm Two

The following steps are involved in using algorithm two to determine the quantile estimates for selected return periods based on the probability distribution.

i. The algorithm is available on request. Table 7a and 7b shows a section of the algorithm  
ii. The same 40 years annual data (precipitation or discharge) used in Algorithm one is also needed to run Algorithm two. Go to cell B2 – B41 of Algorithm one where you have your 40 years annual maximum data (precipitation or discharge) copy and paste on cell E2 – E41 of GEV, GLO and GPA in Algorithm two  
iii. Copy the computed parameters (Shape parameter (k), Scale parameter (α) and Location parameter (ξ)) of each distribution from algorithm one. Replicate them to 40 years data and paste in cell F2 – F41, G2 – G41 and H2 – H41 of GEV, GLO and GPA in Algorithm two  
iv. The predicted discharge or precipitation values of generalized extreme value (GEV) will be gotten from
cell N2 – N41 of Algorithm two; that of generalized logistics (GLO) will be gotten from cell O2 – O41 of Algorithm two and that of generalized pareto (GPA) will be gotten from cell N2 – N41 of algorithm two.

v. The goodness of fit statistics for selecting the probability distribution that best fit your data, namely; root mean square error (RMSE), relative root mean square error (RRMSE) and maximum absolute deviation index (MADI)) will be gotten from cell R3, V3 and X 3 of algorithm two for generalized extreme value distribution (GEV), cell S3, W3 and Y 3 of algorithm two for generalized logistics distribution (GLO) and cell R3, W3 and Y3 of algorithm two for generalized pareto distribution (GPA)

vi. Cell AE2 to AE11 of Algorithm two gives the computed quantile estimates for the generalized extreme value distribution (GEV), generalized logistics distribution (GLO) and generalized Pareto distribution (GPA).

vii. The cumulative probability of non-exceedance is also computed in addition to the graphical visualization of your data with respect to GEV, GLO and GPA probability distribution.

Table 6a: Sectional view of algorithm one

| Rank | BB1 | BB2 | BB3 | BB4 | BB5 | BB6 | BB7 | BB8 | BB9 | BB10 |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| 1    | 222.3 | 222.3 | 222.3 | 222.3 | 222.3 | 222.3 | 222.3 | 222.3 | 222.3 | 222.3 |
| 2    | 333.4 | 333.4 | 333.4 | 333.4 | 333.4 | 333.4 | 333.4 | 333.4 | 333.4 | 333.4 |
| 3    | 444.2 | 444.2 | 444.2 | 444.2 | 444.2 | 444.2 | 444.2 | 444.2 | 444.2 | 444.2 |
| 4    | 555.6 | 555.6 | 555.6 | 555.6 | 555.6 | 555.6 | 555.6 | 555.6 | 555.6 | 555.6 |
| 5    | 666.2 | 666.2 | 666.2 | 666.2 | 666.2 | 666.2 | 666.2 | 666.2 | 666.2 | 666.2 |
| 6    | 777.8 | 777.8 | 777.8 | 777.8 | 777.8 | 777.8 | 777.8 | 777.8 | 777.8 | 777.8 |
| 7    | 888.4 | 888.4 | 888.4 | 888.4 | 888.4 | 888.4 | 888.4 | 888.4 | 888.4 | 888.4 |
| 8    | 999.2 | 999.2 | 999.2 | 999.2 | 999.2 | 999.2 | 999.2 | 999.2 | 999.2 | 999.2 |
| 9    | 1010.2 | 1010.2 | 1010.2 | 1010.2 | 1010.2 | 1010.2 | 1010.2 | 1010.2 | 1010.2 | 1010.2 |
| 10   | 1111.8 | 1111.8 | 1111.8 | 1111.8 | 1111.8 | 1111.8 | 1111.8 | 1111.8 | 1111.8 | 1111.8 |

Table 6b: Sectional view of algorithm one

| L-Moment Equation | L-Moment Ratio | Parameters Estimation |
|-------------------|----------------|----------------------|
| (3) GENERALIZED EXTREME VALUE (GEV) | | |
| K = 7.88506 + 2.95845 L | | |
| C = [0.05+0.01L] | | |
| α = [0.05+0.01L] | | |
| λ = [0.05+0.01L] | | |
| β = [0.05+0.01L] | | |
| (1) = [0.05+0.01L] | | |
| (2) = [0.05+0.01L] | | |
| (3) = [0.05+0.01L] | | |

3.3 Table 6b: Sectional view of algorithm one
4.0: Application of the Algorithms for modeling Rainfall Data

4.1: Description of study area

Benin City, the capital of Edo State, Nigeria, the study area, is located in the rain forest zone, with geographical coordinates of latitudes 6° 17' N, 6° 26' N and longitudes 5° 35' E, with an annual mean temperature of 27.5°C (Ikhuoria, 1987). It has two main seasons, wet and dry; from April to November and November to April respectively, with an annual mean rainfall of about 2095mm. In matters of hydrogeology, Benin City lies on the Benin Formation, with an aquifer of mean depth of 136m. Schools, hospitals, markets and cemeteries are among the social services provided in the City. Figure 1 shows the digitized map of the study area.
4.2: Data collection
The data used for this study was collected from the Nigerian Meteorological Agency, Oshodi; Lagos State, Nigeria.
The data includes monthly precipitation data for 40 years spanning between; 1974 to 2013. The data were then sorted to obtain the annual maximum precipitation records.

5.0: Results and Discussion
The time series plot of the data presented in Figure 2 shows the presence of seasonal variability since rainfall depth varies within the period understudy as some years experienced extreme precipitation compared to others.

On whether the rainfall data used in this study are from the same population distribution, homogeneity test was performed using hydrological software (RAINBOW). Result of the test presented in Figure 3 shows that the
The data used is homogeneous.

Figure 3: Homogeneity test of data

For a homogeneous record, the rainfall data points normally fluctuate around the zero-center point in the residual mass curve as observed in Figure 3. The descriptive statistics of the rainfall data and the computed probability weighted moment statistics (b0, b1, b2 and b3), the L-moment values (λ1, λ2, λ3 and λ4) and the L-moment ratio values (Τ2, Τ3 and Τ4) were obtained from algorithm one and presented in Table 8

Table 8: Descriptive and L-moment statistics obtained from algorithm one

| Probability Weighted Moment Values | L-Moment Values | L-Moment Ratios | Basic Statistics |
|-----------------------------------|----------------|----------------|-----------------|
| b0 = 461.565                      | λ1 = 461.565   | L-CV = 0.128   | n = 40           |
| b1 = 260.3925                     | λ2 = 59.220    | L-Skewness = 0.0347 | Mean = 461.565  |
| b2 = 183.807                      | λ3 = 2.0527    | L-Kurtosis = 0.2151 | Variance = 11280.1 |
| b3 = 143.190                      | λ4 = 12.739    |                | Standard Dev. = 106.208 |

Table 9: Probability parameter estimate based on L-moment using algorithm one

| Probability Distribution Model | Shape Parameter | Scale Parameter | Location Parameter |
|-------------------------------|-----------------|-----------------|-------------------|
| GEV                           | 0.223348438     | 74.54719112     | 540.6930545       |
| GLO                           | -0.03466283     | 57.98293611     | 425.8765149       |
| GPA                           | 0.865993717     | 316.7041936     | 291.8408521       |

The estimated parameters for the different distributions were applied to the relevant quantile function given in Table 3. The predicted annual maximum precipitation records based on L-moment using the three-probability distribution model, namely; GEV, GLO and GPA was obtained from algorithm two and presented in Table 10
Table 10: Observed and predicted annual maximum precipitation based on GEV, GLO and GPA using L-moment obtained from algorithm two

| Rank | Observed annual maximum precipitation (xi) | Predicted annual maximum precipitation (yi) based on GEV | Predicted annual maximum precipitation (yi) based on GLO | Predicted annual maximum precipitation (yi) based on GPA |
|------|------------------------------------------|--------------------------------------------------------|--------------------------------------------------------|--------------------------------------------------------|
| 1    | 222.3                                    | 427.0489297                                          | 225.0946321                                          | 299.5780861                                           |
| 2    | 235.2                                    | 447.2249508                                          | 262.213348                                           | 307.3412874                                           |
| 3    | 305.2                                    | 460.7610137                                          | 284.951664                                           | 315.1312125                                           |
| 4    | 338.4                                    | 471.3912158                                          | 301.7342325                                          | 322.9486605                                           |
| 5    | 342.9                                    | 480.3639873                                          | 315.2418773                                          | 330.7944768                                           |
| 6    | 351.2                                    | 488.26164                                            | 326.681309                                           | 338.6695567                                           |
| 7    | 357.1                                    | 495.4068129                                          | 336.7022738                                          | 346.5748505                                           |
| 8    | 391                                      | 501.9987395                                          | 345.6961962                                          | 354.5113683                                           |
| 9    | 393.3                                    | 508.1701477                                          | 353.9181094                                          | 362.4801854                                           |
| 10   | 394.8                                    | 514.0146504                                          | 361.5442227                                          | 370.4824493                                           |
| 11   | 403                                      | 519.6013347                                          | 368.7020615                                          | 378.5193866                                           |
| 12   | 413.6                                    | 524.9831494                                          | 375.4875291                                          | 386.5923113                                           |
| 13   | 425.2                                    | 530.2020282                                          | 381.9751792                                          | 394.7026345                                           |
| 14   | 427.8                                    | 535.29218                                            | 388.2247225                                          | 402.8518751                                           |
| 15   | 433.9                                    | 540.2822948                                          | 394.2853316                                          | 411.0416725                                           |
| 16   | 445.4                                    | 545.1970825                                          | 400.1986036                                          | 419.2738008                                           |
| 17   | 453.1                                    | 550.0583886                                          | 406.0006771                                          | 427.5501866                                           |
| 18   | 454                                      | 554.8660348                                          | 411.7238054                                          | 435.872928                                            |
| 19   | 454.5                                    | 559.6984815                                          | 417.3975769                                          | 444.4443186                                           |
| 20   | 455.4                                    | 564.5133759                                          | 423.0499088                                          | 452.6667855                                           |
| 21   | 458.8                                    | 569.3480325                                          | 428.7079054                                          | 461.1433725                                           |
| 22   | 461.4                                    | 574.2198818                                          | 434.3986499                                          | 469.6768808                                           |
| 23   | 462.4                                    | 579.1469178                                          | 440.1499871                                          | 478.2708183                                           |
| 24   | 462.5                                    | 584.1481753                                          | 445.9913571                                          | 486.9290101                                           |
| 25   | 467.8                                    | 589.2442708                                          | 451.9547409                                          | 495.657641                                           |
| 26   | 472.5                                    | 594.4580483                                          | 458.0757995                                          | 504.4559663                                           |
| 27   | 476                                      | 599.8153864                                          | 464.3953135                                          | 513.3352023                                           |
| 28   | 490.4                                    | 605.3462472                                          | 470.9610814                                          | 522.2999154                                           |
| 29   | 491.8                                    | 611.0860883                                          | 477.8305154                                          | 531.3576148                                           |
| 30   | 506.1                                    | 617.0778262                                          | 485.0743203                                          | 540.5171568                                           |
| 31   | 515.9                                    | 623.3746618                                          | 492.7818904                                          | 549.7891329                                           |
| 32   | 550.8                                    | 630.0442946                                          | 501.0695375                                          | 559.1864205                                           |
| 33   | 564                                      | 637.1754714                                          | 510.0935822                                          | 568.7249902                                           |
| 34   | 572.7                                    | 644.8886268                                          | 520.0722686                                          | 578.425385                                           |
| 35   | 580.8                                    | 653.3545277                                          | 531.3248014                                          | 588.313466                                            |
| 36   | 614.5                                    | 662.8283195                                          | 544.3465967                                          | 598.4262568                                           |
| 37   | 615.1                                    | 673.7206674                                          | 559.9703872                                          | 608.8157625                                           |
| 38   | 623.1                                    | 686.7649251                                          | 579.7661189                                          | 619.5633949                                           |
| 39   | 656.2                                    | 703.5073961                                          | 607.2890118                                          | 630.8123007                                           |
| 40   | 722.5                                    | 728.4379637                                          | 654.0454185                                          | 642.8811053                                           |

The best from among candidate distributions fitted to the observed data at the station was selected by subjecting their respective predicted precipitation values to five statistical goodness-of-fit tests. The computed goodness of fit statistics based on the three probability distributions was obtained from algorithm two and presented in Table 11

Table 11: Goodness of fit statistics based on L-moment obtained from algorithm two

| GoF Statistics | RMSE   | RRMSER | MADI    | MAE     | PPCC   |
|----------------|--------|--------|---------|---------|--------|
| GEV            | 115.6429728 | 0.326199623 | 0.269482909 | 0.45673 | 0.9675 |
| GLO            | 37.79785402 | 0.079475417 | 0.001362853 | 0.00354 | 0.9793 |
| GPA            | 30.20182487 | 0.089637382 | 0.006644185 | 0.0674 | 0.9171 |

The overall goodness of fit of each distribution was judged using a ranking scheme by comparing the three categories of test criteria based on the relative magnitude of the statistical test results. The distribution with the
lowest RMSE, lowest RRMSE, lowest MADI, lowest MAE and highest PPCC was assigned a score of 3, the next 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 and the distribution with the highest total point score was selected as the best fit distribution model. The scoring scheme and the overall ranking of the distributions models at the stations based on the goodness of fit tests is presented in Table 12

Table 12: Scoring and ranking scheme for selected probability distribution models

| Test Criteria | Distribution Scoring |
|---------------|----------------------|
|               | GEV                  | GLO                  | GPA                  |
| RMSE          | 1                    | 2                    | 3                    |
| RRMSE         | 1                    | 3                    | 2                    |
| MADI          | 1                    | 3                    | 2                    |
| MAE           | 1                    | 3                    | 2                    |
| PPCC          | 2                    | 3                    | 1                    |

Total Score/ Rank 6 (3rd) 14 (1st) 10 (2nd)

Based on the result of Table 12, generalized logistics probability distribution (GLO) with the highest total score of 14 was selected as the best probability distribution model for analyzing annual maximum rainfall series in Benin City followed by GPA and then GEV. The quantile estimates (Qt) based on 2, 5, 10, 20, 50, 100, 200 and 500 years was also obtained based on L-moment procedure using algorithm two and results obtained is presented in Tables 13

Table 13: Computed quantile estimates based on selected return periods for GLO using algorithm two

The graphical visualization of the observed and predicted annual maximum precipitation for GLO was obtained from algorithm two and presented in Figure 4

Figure 4: Observed versus predicted precipitation based on GLO distribution

The computed coefficient of determination (r²) between the observed and predicted precipitation was observed to be 0.9793 for generalized logistics distribution (GLO). Based on the computed coefficient of determination, it was concluded that generalized logistics distribution had a better fit of the annual maximum
precipitation data.

6.0 Conclusion
This paper gave a detailed description of the current method of statistical parameter estimation of selected probability distribution model, namely; Generalized Extreme Value (GEV), Generalized Logistics (GLO) and Generalized Pareto (GPA) probability distributions – the L-moments method. A simple to use Microsoft Excel Algorithm have been developed for estimating basic descriptive statistics such as the sample mean, variance, standard deviation, skewness, kurtosis, and coefficient of variation. Other exciting features include; the computation of Probability weighted moment parameters (b0, b1, b2 and b3), L-Moment values (ƛ1, ƛ2, ƛ3 and ƛ4), L-Moment ratio values (Τ2, Τ3 and Τ4), and goodness of fit statistics (RRMSE, RMSE and MADI). The algorithm will not only find use in the practice of engineering hydrological computation, it will also help design engineers in estimating the magnitude of peak rainfall/discharge for various return periods.

References
Ahmad, U. N., Shabri, A. and Zakaria, Z.A., (2011); Flood frequency Analysis of Annual Maximum Flows Using L-Moments and TL-Moments Approach, Applied Mathematics Sciences, Vol. 5 (5), 243-253
Centre for Ecology and Hydrology (CEH), FLOODS version 1.1, Regional flood frequency analysis software manual, Water Resources section, CEH Wallingford, U.K.2001
ECOST (European Corporation in Science and Technology); Review of statistical Methods for Flood frequency Analysis in Europe, Centre for Ecology and Hydrology, Wallingford, United Kingdom, 2012
Ehiorobo, J. O and Izinyon, O. C., (2013); Flood frequency analysis of Oshun River in Asejire Dam Site, Nigeria, Journal of Earth Science and Engineering, vol. 3, 292-300
Eslamian, S.S and Feizi, H (2007); Maximum Monthly Rainfall Analysis Using L-Moments for an Arid Region in Isfahan Province, Iran, Journal of Applied Meteorology and Climatology, Vol.46, 494-503.
Eregno, F.E (2014), Regional flood frequency analysis using L-moment in the tributaries of Upper Blue Nile River, South Western Ethiopia, Merit Journal of Engineering, Pure and Applied Sciences, vol. 2(2), 012 – 021
Gubareva, T. S and Gartsman, B. I., (2010); Estimating Distribution Parameters of Extreme Hydrometeorological Characteristics by L-Moment Method; Water Resources, Vol. 37(4), 437–445
Herlina, H (2015); Modeling Extreme Rainfall with Gamma-Pareto Distribution, Applied Mathematical Sciences, Vol. 9 (121), 6029 – 6039
Hosking, J.R.M. and Wallis, J.R., (1997); Regional Frequency Analysis: An Approach Based On L-Moments, Cambridge: Cambridge University Press.
Hosking, J.R.M., (1990); L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics, Journal of the Royal Statistical Society, Series B, vol. 52, 105-124
Izinyon, O.C and Ehiorobo, J.O., (2015); L-moments Method for Flood Frequency Analysis of River Owana at Owana in Benin Owena River Basin in Nigeria, Current Advances in Civil Engineering, Volume 3, Issue 1, 1-10
Landwehr, J.M., Matalas, N.C., and Wallis, J.R., (1979); Probability Weighted Moments Compared with Some Traditional Techniques in Estimating Gumbel Parameters and Quantiles, Water Resources Research, vol. 15 (5), 1055-1064
Maleki-Nezhad, H (2006); Regional flood frequency analysis using L-moments approach, 7th International Congress on Civil Engineering, Tehran Iran, May 8-10, 1-7
Sanjib, G.; Manindra, K. R. and Soma, C. B, (2016), Determination of the Best Fit Probability Distribution for Monthly Rainfall Data in Bangladesh, American Journal of Mathematics and Statistics, vol. 6(4), 170-174
Sharma, M.A and Singh, J. B (2010); Use of probability distribution in rainfall analysis, New York Science Journal, vol. 3 (9), 40-49
Vivekanandandan, N., (2015); Flood frequency analysis using method of moments and L-moments of probability distribution, Vivekanandandan Cogent Engineering, vol. 2, 1-10

Dr. Idowu Rudolph Ilaboya
Dr. Idowu Rudolph Ilaboya holds a National Diploma (ND) and a Higher National Diploma (HND) in Polymer Technology from Federal Polytechnic Auchi, a Post Graduate Diploma (PGD) in Chemical Engineering from Federal University of Technology Minna, Master’s degree (M.Eng) and a Doctorate degree (PhD) in Water Resources and Environmental Engineering from University of Benin, Benin City. He is a recipient of both the Alumni Merit Award for best graduating National Diploma (ND) student and Rector’s Merit Award for best graduating Higher National Diploma (HND) student in Polymer Technology. He currently holds the prestigious University of Benin Annual Research Day (UBARD) Award for best presenter from the Faculty of Engineering for both the second and third edition. He is a member of the NEPAD-AFRICAN NETWORK of center of excellence on water science and technology University of Benin, member University of Benin Taskforce
Committee on Utilization and Training on Equipment, member Department of Civil Engineering Standards and Accreditation Committee, member Faculty of Engineering Research/Conference Committee, member Nigerian Institution of Professional Engineers and Scientist, member Nigerian Association of Technologists in Engineering, member Material Society of Nigeria and the Polymer Institute of Nigeria. Dr. Ilaboya teaches hydraulics/hydrology, pollution transport systems and solid waste disposal in the Department of Civil Engineering, Faculty of Engineering, University of Benin, Benin City and has published in local, national and foreign journals.

Dr. Ebierin Akpoebidimiyen Otuaro
Engr. Dr. Ebierin Akpoebidimiyen Otuaro obtained a Diploma in Physics Education from the prestigious Ahmadu Bello University, Zaria, Kaduna State in September 1997 and a Bachelor of Agriculture Engineering (B. Eng.) from the same University in December, 2001. She bagged a Master’s Degree in Agriculture Engineering (M. Eng.) from Federal University of Technology, Akure, Ondo State in August, 2006. She has successfully defended her PhD Thesis from the University of Benin, Benin City in 2018. She started her working career as a Feed Mill Operator in Bayuf Farm Ketu, Lagos State in 1999 and Natural Centre for Agricultural Mechanization, Ilorin, Kwara State as a Trainee Engineer in 2000. She taught at Cherubim and Seraphim College, Sanya, Surulere, Lagos State as a Mathematics Teacher in 2001. Engr. Ebierin Otuaro worked as a Trainee Engineer with Ministry of Agriculture and Natural Resources, Abakaliki, Ebonyi State as a youth corps member in September 2002-August, 2003. Engr. Ebierin Otuaro started lecturing at Delta State Polytechnic, Ozoro from December, 2009 to November 2017. She was a lecturer in the Department of Agronomy, Delta State University, Asaba Campus from November 2017 - October 2019. Presently, she lectures in the Civil Engineering Department, Faculty of Engineering, Maritime University, Okenrenko, in Warri South-West Local Government Area. Her philanthropic gesture has earned her lots of awards at local, national and international levels which include Top Ten Pastoral Care Fellowship Leader’s Award (PCF) December, 2013; Nigeria Society of Engineer, Warri Branch Award, 2018; Association of Professional Women Engineers in Nigeria (APWEN) Award of Appreciation, 2018; and Nigeria Society of Engineers Merit Award, June, 2018. She is a distinguished member of Council for the Regulation of Engineering in Nigeria (COREN) and also a member of Nigeria Society of Engineers (NSE).