Journal of Civil Engineering, Science and Technology

Volume 13, Issue 1, April 2022, 33 – 48

COMPARISON OF L-MOMENT AND METHOD OF MOMENTS AS PARAMETER ESTIMATORS FOR IDENTIFICATION AND CHOICE OF THE MOST APPROPRIATE RAINFALL DISTRIBUTION MODELS FOR DESIGN OF HYDRAULIC STRUCTURES

Augustine Iyekeoretin Agbonaye[1*], Ebierin Akpoebidimiyen Otuaro[2] and Osadolor Christopher Izinyon[1]

[1]Department of Civil Engineering, University of Benin, Benin City. Nigeria

[2]Department of Civil Engineering, Maritime University, Okerenkoko, Warri, Nigeria

Date received: 17/11/2021 Date accepted: 10/03/2022

*Corresponding author's email: augustine.agbonaye@uniben.edu

DOI: 10.33736/jcest.4207.2022

Abstract — In rainfall frequency analysis, the choice of a suitable probability distribution and parameter estimation method is critical in forecasting design rainfall values for varying return periods at every location. Previously, some researchers in Nigeria used the method of moments (MoM) while others used the L-moment method (LMM) as parameter estimators. However, a more accurate result is obtainable if both estimators are used and their results are compared and ranked to obtain the most appropriate distribution models for each location This study compared the performance of two forms of parameter estimation, namely the method of moments (MoM) and the L-moment method (LMM). This was aimed at identifying and selecting the best fit probability distribution models among three distribution models for the design of hydraulic structures. These models are Generalized Pareto (GPA), Generalized Extreme Value (GEV), and Gumbel Extreme Value (EVI). Annual rainfall series of ten gauging stations with data from 33-50 years from ten southern States of Nigeria obtained from NIMET were used for Rainfall Frequency Analysis (RFA). At five locations, the best fit probability model was the GPA probability distribution model with L-Moment. EVI and GEV probability distribution models with the method of moments were the most appropriate probability models at two locations each. EVI probability distribution model with the L-moment was the most appropriate probability model at one place. The findings confirmed that no single distribution outperformed all others at all stations. Since no single model is regarded preferable for all practical purposes, the best-fit probability model with parameter estimator at any location is site-specific. Consequently, available models and parameter estimators are filtered based on the situation at hand and the type of data available. The identified best fit models with the most appropriate parameter estimator would be a tool to help decision-makers in sizing hydraulic structures in the area.

Keywords**:** Best-fit, goodness-of-fit, maximum rainfall, parameter estimator, probability distribution

1.0 INTRODUCTION

Flood of extreme events from higher rainfall is a major concern in Nigeria where almost all the cities in the Southern part are annually flooded. Most hydraulic structures originally designed to handle the volume of water from this extreme rainfall event have collapsed due to poor design caused by inaccurate estimations of rainfall frequency values and flood design values. On account of these, it is required that adequate rainfall frequency analysis be conducted to estimate the peak flood values of rainfall which is a tool to help decision-makers in sizing hydraulic structures such as culverts, dams, spillways, and bridges.

Estimation of rainfall peak values for accurate engineering design of water carrying hydraulic structures has remained one of the most challenging issues as sufficient hydrological data is limited. This is the paradigm in most gauging stations in Nigeria, where most of the sites are poorly gauged or not gauged at all.

Previously, some researchers in Nigeria used the method of moments (MoM) while others used the L-moment method (LMM) as parameter estimators. However, a more accurate result is obtainable if both estimators are used and their results are compared and ranked in order to obtain the most appropriate distribution models for each location.

This study was aimed to carry out an at-site frequency analysis of observed rainfall data at ten stations selected in Southern Nigeria using three probability distribution models where parameters were estimated by the Method of Moment and L-Moment approach.

The objectives of the study are to:

I. Derive annual maximum daily Rainfall series for observed rainfall data at selected stations in Southern Nigeria;

II. Estimate sample L-moments and L-Moment ratios namely: L-coefficient of variation (L-CV), L-coefficient of skewness(L-Skew), and L-coefficient of kurtosis(L-Kurtosis) and compute the descriptive statistics for MOM in favour of the selected stations;

III. Fit three (3) commonly utilized probability distribution models namely: Generalized Pareto (GPA), Generalized Extreme Value (GEV), and Gumbel Extreme Value (EVI) distribution models to the sample data at each station using Method of Moment and L-Moment's approach;

IV. Select the best fit probability distribution model at each station using statistical goodness of fit criteria or measures namely: root mean square error (RMSE), relative root mean square error (RRMSE), probability plot correlation coefficient (PPCC), maximum absolute error (MAE), mean absolute deviation index (MADI) and Diagnostic-Index(D-Index) tests and a scoring ranking scheme; and

V. Predict values of rainfall return level (RT) for return periods (T) of engineering importance (T = 2 years, 5 years, 10 years, 25 years, 50 years, 100 years, and 200 years) for each station using the best fit distribution model obtained.

For successful operation and cost-effective design of important hydraulic infrastructures such as dams, reservoirs, spillways, bridges, culverts, and urban drainage systems, realistic estimates of rainfall magnitude for a specific return time are required.

One of the most basic and well-known approaches for estimating parameters in statistical hydrology is the method of moments. Using this method, a sample is fitted with a probability distribution by equating the sample moments to the theoretical distribution moments and the parameters are calculated. Even though this technique is simple and the computations are straightforward, it has been discovered that sample instant numerical values may differ dramatically from those of the population from which the sample was drawn. This is especially true when the sample size is small and/or the sample skewness is significant [1]. Compared to traditional moments, L-moments can characterize a broader range of distributions. The L-moments can be utilized to determine the distribution even if any of the conventional moments are absent. [2].

Furthermore, L-moments are more resistant to data outliers than traditional moments [3] and allow for more trustworthy inferences about an original probability distribution from tiny data. The benefits of L-moments over conventional moments in hypothesis testing, boundedness of moment ratios, and distribution identification have all been thoroughly studied. [2,4,28,33]. Stedinger [5] outlined the links between the distribution parameters and the L-moments and explained the theoretical features of the various distributions widely utilized in hydrology.

At lower skewness, conventional moments are favoured, especially for smaller samples. [6], while L-moments are preferred at higher skewness, regardless of sample size. The best-fitted probability distributions were determined by [7] for at-site flood frequency analysis of the Ume River in Sweden. Their finding was that the generalized extreme value distribution with the L-moments estimation provided the most appropriate to the yearly maximum streamflow at two gauging sites. For the Prediction of Kanji Reservoir Inflows, Niger State, Nigeria, the Best-Fit Probability Distribution Models were examined by [8]. They selected the Gumbel (EVI) model as the best most appropriate distribution Model. EV1 method was used for flood frequency analysis of the Burhi Gandak river by [31]. A regional precipitation-frequency analysis was conducted by [9] using L-moment methods to estimate three characteristics (size, shape, and position) during the period 1965–2013. The GEV distribution was found to best suit the precipitation data in the R1 and R2 regions. GPD was also chosen for the R3 region.

For flood study and selection of parent distributions, [10] used L-moment-based regional frequency analysis. to fit extreme monthly precipitation data from 18 locations in Iran's Zayandehrood basin. As a result, the generalized extreme-value and Pearson type-III tests' distributions were selected and model parameters were estimated. [11], determined the most appropriate Probability Distribution for Monthly Rainfall Data from 1979 to 2013. They discovered that the monthly rainfall data for the three Bangladeshi sites were best distributed using a generalized extreme value distribution.

[12] studied annual rainfall data from fourteen Sudanese rainfall locations from 1971 to 2010. For the yearly rainfall over the period, the normal and gamma distributions were chosen as the best-fit probability distributions. [13] determined the best-fit probability distributions for maximum monthly rainfall Statistical analysis and distribution types were used to analyze 30 years of data from 35 locations in Bangladesh. Generalized Extreme Value, Pearson type 3, and Log-Pearson type 3 were selected as the best-fitted probability distributions Also, the 10-year, 25-year, 50-year, and 100-year return periods of maximum monthly rainfall were calculated for all locations studied. For yearly maximum daily rainfall in India, the Lognormal type 2 distribution was the best-fitting probability distribution according to [14, 6].

[15] revealed that the log-Pearson type 3 distribution was the best-fit distribution by utilizing yearly maximum rainfall based on daily rainfall in the northern areas of Pakistan. The best-fitted probability distributions for at-site flood frequency analyses of the Ume River in Sweden were obtained by [7]. At the gauging sites Solberg and Stornorrfors Krv, the generalized extreme value distribution with L-moments estimation provided the greatest match to maximum annual streamflow. They also predicted the maximum flow of water for return times of five, ten, twenty-five, fifty, one hundred, two hundred, five hundred, and one thousand years using the best-fitted distribution for each measuring site.

[16] identified Log Pearson 3, Generalised Logistic, and Extreme Generalised Value Distribution Model as the most appropriate probability distribution models of extreme mean annual rainfall events in South Africa. [17] identified the most suitable probability distribution models for maximum, minimum, and mean streamflow for Tana River in Kenya. The lognormal and GEV distribution functions were the best-fit functions for the annual mean flows of the Tana River.

[18] conducted a flood frequency study on maximum monthly rainfall data for Patani, Niger Delta region of Nigeria. Rainfall data from 1981 to 2013 were obtained from NIMET and CBN to estimate the possibility of flooding and take required actions to mitigate it. They used five probability distributions: Normal, Lognormal, Log Pearson, Gumbel, and Foster's Type -1 with return periods of two, five, ten, twenty-five, fifty, one hundred, and two hundred years. Their result showed that the Gumbel Distribution best described the region's precipitation data and, as a result, can be used to forecast flooding in the area. [19] fitted three probability distribution models to yearly maximum series of discharge or flow data at three flow gauging stations in Nigeria during 32 years (1955–1986): Serav is on the Katsina-Ala River, Gasol is on the Taraba River, and Mayokam is on the Mayokam River. The most appropriate probability distribution models obtained for the different stations were, Log-Normal, and Log Pearson Type III for the stations at River Katsina-Ala at Serav, River Taraba at Garsol, and River Mayokam at Mayokam respectively. The most appropriate distribution model at each site was used to estimate return period floods for return periods of two, five, ten, twenty-five, fifty, one hundred, and two hundred years:

Since rainfall frequency analysis data must be independent and evenly distributed, possible distributions and parameter estimators must be chosen for the most appropriate to the data available at a specific place [20]. As a result, this research is required to screen and choose the most appropriate probability distribution models with parameter estimators for rainfall prediction in southern Nigeria.

2.0 MATERIALS AND METHODS

2.1 The Study Area

The purpose of this research is to look at the frequency of the yearly extreme series of diurnal rainfall totals in a few cities in Southern Nigeria. The cities that have been chosen are Ikeja, Akure, Ibadan, Benin City, Port Harcourt, Uyo, Calabar, Onitsha, Enugu, Owerri. Figure 1 is a map of Nigeria showing 10 selected cities and the ten (10) station coordinates are presented in Table 1.

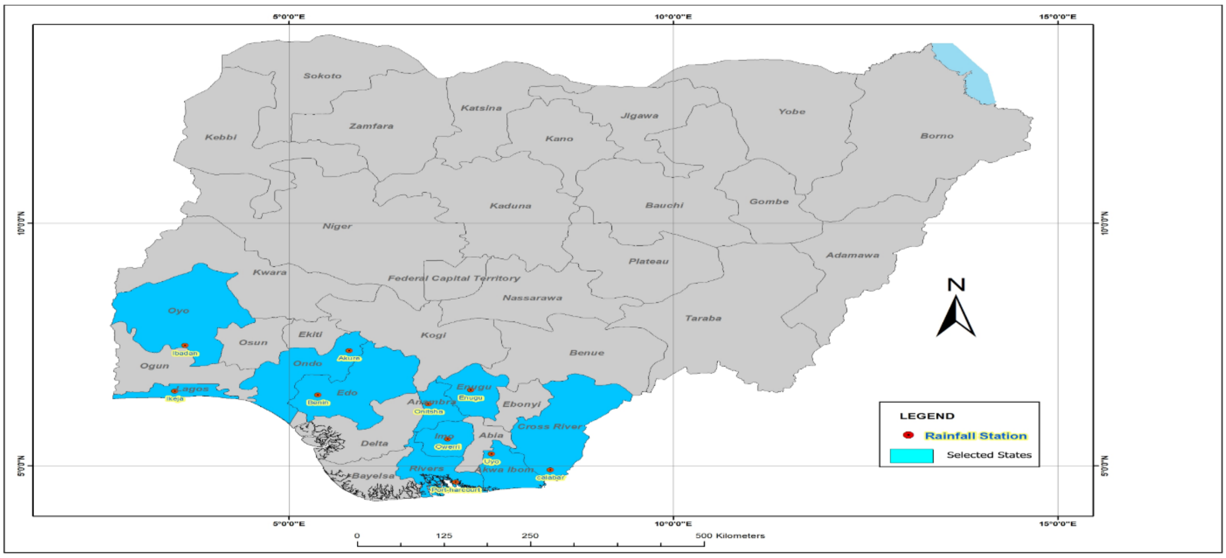

Figure 1 Nigerian Map with selected cities

Source: Adapted from Office of the Surveyor-General of the Federal Republic of Nigeria.

Table 1 Station coordination of selected cities in Nigeria

S/N	Station name	Geopolitical zone in Nigeria	State in Nigeria	Longitude	Latitude	Years of available data
1	Ikeja	South – West	Lagos	3.45°E	6.2°N	50
2	Akure	South – West	Ondo	5.5°E	7.15°N	33
3	Ibadan	South – West	Oyo	3.58°E	7.22°N	50
4	Benin City	South-South	Edo	5.31°E	6.20°N	48
5	Port Harcourt	South-South	Rivers	7.10°E	4.40°N	40
6	Uyo	South-South	Akwa Ibom	7.53°E	5.10°N	32
7	Calabar	South-South	Cross River	8.20°E	4.57°N	49
8	Onitsha	South – East	Anambra	6.42°E	6.60°N	32
9	Enugu	South – East	Enugu	7.30°E	6.30°N	49
10	Owerri	South – East	Imo	7.0°E	5.29°N	32

2.2 Data Used and Analysis method

The diurnal rainfall data for the chosen cities were acquired from the Nigerian Meteorological Agency (NIMET) in Oshodi, Lagos provided diurnal rainfall data for the chosen cities from 1965 to 2014 (50 years). In general, continuous data of 30 years is required to make a time series for the rainfall frequency analysis. For each of the stations, the yearly maximum series was generated by picking the wettest year for each of the years in question. As a result, the number of extreme value data was equal to the number of years of record. Rainbow Software was used to verify for homogeneity in the annual series data for each station, and HEC-SSP Software was used to look for outliers. Descriptive statistics (mean, standard deviation and skewness, and kurtosis) were computed using the Annual Maximum Rainfall (AMR) series data. Weibull's plotting position (WPP) formula was used to estimate observed rainfall magnitudes at different durations of T [35].

The flowchart for the adopted methodology in comparing different probability distribution models is presented in Figure 2.

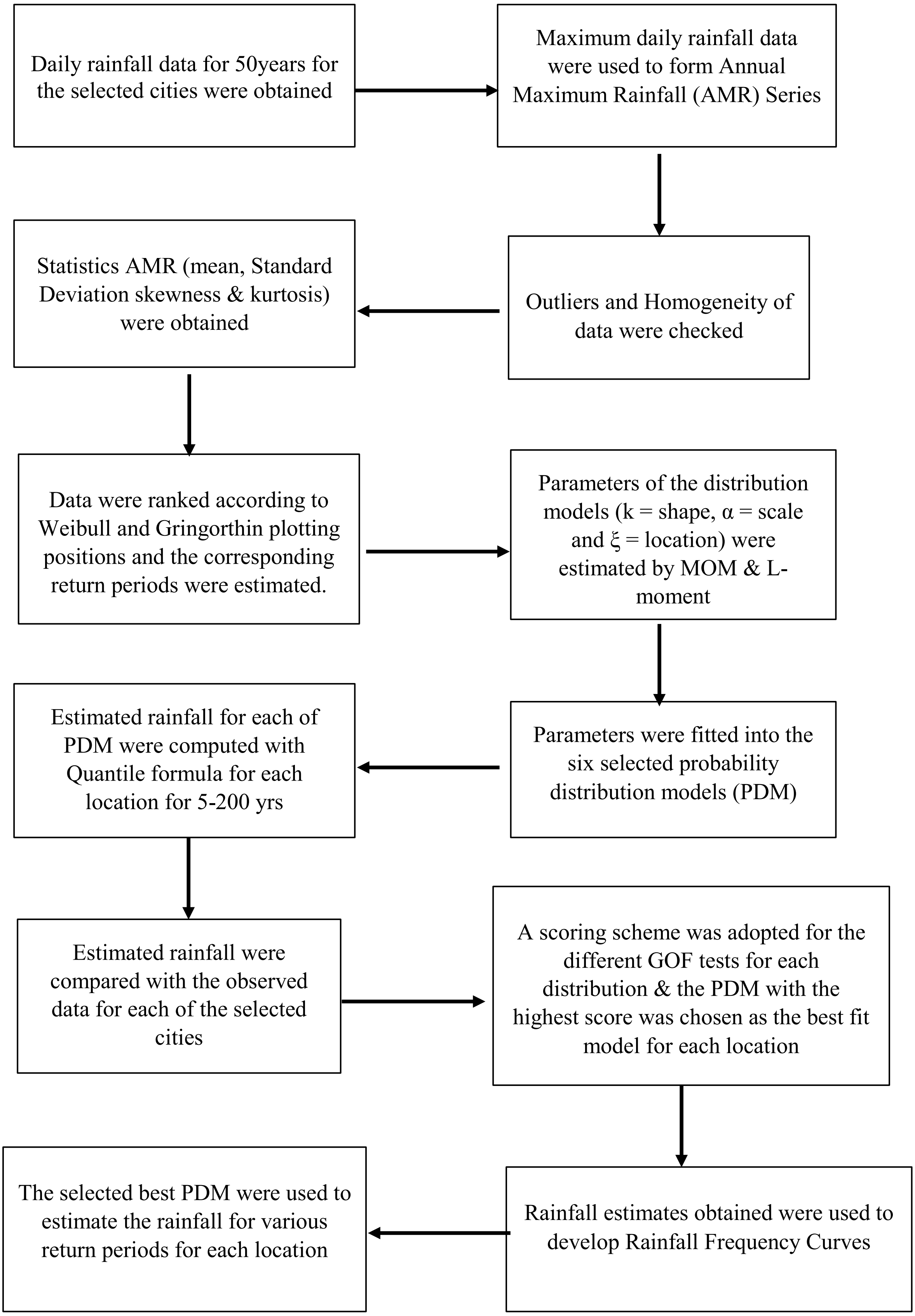

Figure 2 Flowchart for the adopted methodology in comparing different probability distributions

Three probability distribution models based on three factors (size, shape, and location) were studied and fitted to the yearly extreme series data in every location: Generalized Pareto (GPA), Generalized Extreme Value (GEV), and Gumbel Extreme Value type 1 (EVI).

According to [27], the cumulative distribution function (CDF) of R is represented in Table 2 by F(R) (or F), and P represents the chance of surpassing it α, ξ and k are the location scale and shape parameters, respectively. They were obtained from the average, standard deviation, and coefficient of skewness of the collected rainfall data. The estimated rainfall during a return time is given by the probability distribution; sign (k) is either plus or minus one, depending on the sign of k, sign (k) is either plus or negative one.

PWM theory is summarized in [23], which defines them as:

$$-r = EX \ldots F - s., x.. - r \tag{1}$$

Where, F-s., X. denotes x's cumulative distribution function (CDF), and -r. is the r-th. order of PWM.

The Quantile Function of Probability Distribution for Two Parameter Estimators is presented in Table 2.

Table 2 Quantile Function of Probability Distribution for Two Parameter Estimators [21, 22,27]

S/N	Dist.	Quantile function (R_T)	Parameter by MOM	Parameters by LMO
1	EV1	$R_T = \xi - \alpha \ln(-\ln F)$	$\xi = \bar{R} - 0.5772157\alpha$ $\alpha = \left(\sqrt{6/\pi}\right) S_R$	$\xi = l_1 - 0.5772157\ \alpha$; $\alpha = l_2 / \log 2$
2	GEV	$R_T = \xi + \frac{\alpha}{k}(-In F)^{-k} - 1$ Where $F = 1 - \frac{1}{T}$	$K = \frac{1}{3} - \frac{1}{0.31 + 0.91C_{SX} + \sqrt{(0.91C_{SX})^2 + 1.8}}$ $\alpha = C_1 S_X$ $C_1 = \frac{\lvert k \rvert}{\sqrt{\Gamma(1-2k) - \Gamma^2(1-k)}}$ $C_3 = \frac{(\Gamma(1-k)-1)}{k}$ $\xi = R_m - C_3$	$z = \left(2/(3+t_3) \right) - \left(In 2 / In 3 \right)$ $k = 7.8590z^2 + 2.9554z^2$ $\alpha = 1k / (1-2^{-k})\Gamma(1+k)$; $\xi = 1_1 + \left(\frac{\alpha(\Gamma(1+k)-1)}{k} \right)$
3	GPA	$R_T = \frac{\xi + \alpha(1-(1-F)^k)}{k}$	$\bar{R} = \xi + \alpha/(1+k)$; $S_R = \alpha^2/(1+2k)(1+k)^2$ $C_s = 2(1-k)(1+2k)^2/(1+3k)$	$\xi = l_1 + l_2(k+2)$ $k = \left(\frac{4}{(t_3+1)} \right)^{-3}$ $\alpha = (1+k)(2+k)l_2$

For any distribution, unbiased sample estimators of,-i. of the first four PWMs can be constructed as followed [24]:

$$\beta_o = \frac{1}{n} \sum_{j=1}^{n} X_{(j)} \tag{2a}$$

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

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

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

Where $X_{(j)}$ represents the ordered precipitation values with $X_{(1)}$ are largest precipitation value and as the tiniest.

For any distribution, the first four L-moments $(\lambda_1, \lambda_2, \lambda_3 \ and \ \lambda_4)$ according to [24] gives the following results when stated as linear PWM combinations:

$$\lambda_1 = \beta_0 \tag{3a}$$

$$\lambda_2 = 2\beta_1 - \beta_0 \tag{3b}$$

$$\lambda_3 = 2\beta_2 - 6\beta_1 + \beta_0 \tag{3c}$$

$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \tag{3d}$$

The L-moments ratios that were utilized to express estimations of parameters are as followed:

$$\text{L- Variation coefficient (L-CV)} = \tau = \frac{\lambda_2}{\lambda_1} \tag{4}$$

$$\text{L- Skewness } (\tau_3) = \frac{\lambda_3}{\lambda_2} \tag{5}$$

$$\text{L- Kurtosis } (\tau_4) = \frac{\lambda_4}{\lambda_2} \tag{6}$$

The sample estimates of L-moments $(l_1, l_2, l_3 \ and \ l_4)$ are calculated by replacing $(\beta_0, \beta_1, \beta_2 \ and \ \beta_3)$ Equation (3) with $b_0, b_1, b_2 \ and \ b_3$ respectively [24]. The most relevant metrics for summarizing probability distribution are L-moments and (L-CV), as well as L-moment ratios.

For the yearly maximum rainfall data, Equations (2a), (2b), (2c), and (2d) were used to generate the associated probability weighted moments, $b-0, b-1, b-2 \ and \ b-3$. for the observed sample data. [24]. The sample estimates of L-moments (were computed by replacing the results in (3a), (3b), (3c), and (3d). Equations (4), (5), and (6) were used to calculate the L-moment ratios (i.e. (L-CV), L- Skewness, and L- Kurtosis) [28].

For each event, the cumulative probability of non-exceedance, F(P-i.), was calculated using the Weibull formula given by [25,29]:

$$F(P_i) = \frac{m}{n+1} \tag{7}$$

The i-th element of a sample of yearly maximum rainfall arranged in descending order of magnitude is P-i, where m is the rank and n is the number of annual maxima in the record or the sample size. Parameters obtained for both estimators were fitted into the three selected probability distribution models (PDM) to estimate rainfall for each PDM which were computed with a Quantile formula for each location for 5-200 years.

2.3 Goodness-of-fit test criteria

The goodness of fit indicates a comparison of the observed data with the data generated by a model.

According to [27] [32], the goodness of fit tests were used to determine how well a collection of rain data for each station matches a probability distribution model. The following tests are used: Relative root mean square error (RRMSE), Root mean square error (RMSE), Maximum absolute error (MAE), Mean absolute deviation index (MADI), Chi-square (X^2) test, and probability plot correlation coefficient (PPCC). Table 3 shows the Summary of Goodness-of-Fit Statistics.

Each of these goodness of fit tests has its limitations, especially with the sample size requirement. Some do a better job at fitting the tails of the distribution while others are only good at fitting the mid-range of the distribution. Consequently, accurate measures necessitated the use of many goodness of fit tests in this study. The major benefits are group comparison of the distribution models and the selection of the best fit models for each location by scoring/ranking.

Table 3 Summary of Goodness-of-Fit Statistics [27]

0	Test	Abbreviation	Mathematical Equations
1	Root mean square error	RMSE	$RMSE = \left(\frac{\sum (R_o - R_f)^2}{(n-m)} \right)^{\frac{1}{2}}$
2	Relative Root Mean Square Error	RRMSE	$RRMSE = \left(\frac{\sum \left(\frac{R_{oi} - R_{fi}}{R_o} \right)^2}{(n-m)} \right)^{\frac{1}{2}}$
3	Mean Absolute Deviation Index	MADI	$MADI = \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{R_o - R_f}{R_o} \right\|$
4	Maximum Absolute Error	MAE	$MAE = \max(\|R_o - R_f\|)$
5	Probability Plot Correlation Coefficient	PPCC	$PPCC = \frac{\sum \left\{ (R_{0i} - R_m)(R_{fi} - \overline{R_{fm}}) \right\}}{\left[\sum (R_{0i} - \overline{R}_m)^2 \sum (R_{fi} - \overline{R_{fm}})^2 \right]^{\frac{1}{2}}}$
6	Diagnostic Test	D-Index	$D - Index = \left(\frac{1}{R} \sum_{i-1}^{6} \|R_1 - R_i\| \right)$

Except for the PPCC criterion, where the closer the numerical value is to 1, the better the distribution when examining the efficacy of a probability distribution model at a certain place, the lower the value of the Goodness-of-fit test scores, the better the distribution. The mean values of the observed and predicted values are represented by R_m and R_{mf}, respectively.

2.4 Scoring and ranking scheme

Based on the results of the goodness of fit tests, the best performing distribution in terms of the test criterion obtained a score of six (6), the next best test received a score of five, and the worst test received a score of one. The best distribution model for the station was found through a ranking mechanism that used the Goodness of Fit metric to choose the best distribution.

2.5 Model validation

The best-fit probability models were validated using Probability Plot Correlation Coefficient (PPCC) Goodness of fit Statistic. If the PPCC score is close to 1, it indicates that prediction with the model is an exact match with the observed rainfall.

The plotting position correlation coefficient (PPCC) is a metric for the relationship between ordered data and fitted values as determined by a plotting position equation. PPCC was calculated using the relationship [26].

In Table 3, the mean values of the observed and predicted values are represented by Rm and R_{mf}, respectively.

The observed data may have come from the fitted distribution at a specific location if the PPCC score is close to 1.

The hypothesis is as stated:

Null Hypotheses (H0) Probability Distribution Model fits the data

Alternative Hypotheses (Ha): The probability Distribution Model does not fit the data

Accept H_0: if PPCC test statistics r is greater than PPCC Critical values r* at 5% significance level

Or reject H_0): if PPCC test statistics r is less than PPCC Critical values r* at 5% significance level

E. Forecasting rainfall return levels at the station

The rainfall return levels at a station were estimated by using the applicable best distribution's Quantile function in Table 2 and plugging in the relevant return period.

3.0 RESULTS AND DISCUSSION

The summary of descriptive statistics which describe characteristics of a data set of extreme yearly rainfall series are presented in Table 4. These were used to compute the parameters of location (α), scale (ε), and shape (k) of the selected probability distribution for the method of moment estimation of quantile values

Table 4 Summary of descriptive statistics of extreme yearly rainfall.

S/N	Station Location	Mean	Standard Deviation	Skewness	Kurtosis
1	Ikeja	107.67	44.56	1.26	1.38
2	Akure	86.03	24.40	1.35	1.36
3	Ibadan	69.71	27.35	-1.16	1.72
4	Benin city	103.53	48.79	-0.25	0.78
5	Port Harcourt	99.14	42.87	-0.64	0.82
6	Uyo	98.54	25.66	0.64	0.03
7	Calabar	111.73	49.47	-0.58	0.72
8	Onitsha	97.77	38.28	-0.70	1.21
9	Enugu	85.35	39.02	-0.33	0.70
10	Owerri	111.02	36.48	0.16	0.18

The values of the computed sample probability-weighted moments (PMWs) were obtained by applying Equation (2a) – (2d) to the observed data at the different stations are given in Table 5.

Table 5 Computed sample Probability weighted moments (PWMs) at stations

	Sample PWMs			
Station	b_0	b_1	b_2	b_3
Ibadan	71.19	48.53	37.17	30.57
Port Harcourt	99.33	56.81	40.72	32.04
Enugu	92.80	53.92	39.07	31.02

According to [29], the values of sample L- Moments values and L-Moment ratios for the observed data are obtained by applying Equations (3a) – (3c), (4), (5), and (6). The Computed sample L-moments and L-moment ratios at stations for the observed data are presented in Table (6).

Table 6 Computed sample L-moments and L-moment ratios at stations

Station	l_1	l_2	l_3	l_4	l_{-CV}	l_{-CS}	l_{-CK}
Ibadan	71.19	25.85	3.06	7.48	0.36	0.12	0.29
Port Harcourt	99.33	14.29	2.78	1.75	0.14	0.19	0.12
Enugu	92.80	15.04	3.72	2.50	0.16	0.25	0.17

These were used to compute the parameters of location (**α**), scale (ε **),** and shape (**k)** of the selected probability distribution for the L-moment approach of estimation of quantile values.

The parameters of location (**α**), scale (ε **),** and shape (**k)** of the selected probability distribution estimated by the relevant equation in Table 2 using the method of moments are presented in Table 7A.

Also, the parameters of location (**α**), scale (ε **),** and shape (k**)** of the selected probability distribution were estimated by the relevant equation in Table 2. Table7B displays the results of applying the L-moment approach.

Table 7A Estimated parameters using Method of Moments

Location	EVI		GEV			GPA		
	α	ε	K	α	ε	k	α	ε
Benin	38.06	81.55	-0.36	50.8	88.06	0.74	9.14	92.53
Port Harcourt	33.40	79.85	-0.51	52.80	82.22	0.46	10.39	92.00
Uyo	20.00	89.99	-0.09	22.07	87.55	0.45	10.10	91.56
Calabar	38.57	89.82	-0.48	52.74	98.64	0.49	14.47	101.80
Ikeja	34.50	88.82	0.23	22.00	93.51	0.17	9.80	99.94
Ibadan	20.09	59.5	-0.72	42.26	66.56	0.02	5.23	65.95
Akure	19.02	75.05	0.04	10.95	79.95	0.15	6.39	80.44
Enugu	30.42	67.79	-0.39	41.13	73.46	0.68	16.05	75.77
Onitsha	29.84	80.54	-0.53	41.49	88.97	0.42	11.92	89.38
Owerri	28.44	96.68	-0.22	35.18	99.20	0.82	17.88	103.28

Table 7B Estimated parameters using L- Moments

Location	EVI		GEV			GPA		
	α	ε	K	α	ε	k	α	ε
Benin	162.84	-6.56	0.3174	0.09	87.28	0.60	204.07	-40.06
Port Harcourt	47.47	74.93	0.0094	-0.03	99.32	0.44	45 .59	65.5
Uyo	60.23	66.28	0.444	-0.21	100.67	0.05	39.22	63.81
Calabar	61.72	71.93	0.0021	0.01	124.34	0.44	65.5	79.45
Ikeja	77.97	64.44	0.445	-0.18	109.09	0.10	54.27	60.13
Ibadan	85.88	21.62	0.178	0.08	71.13	0.58	105.87	4.58
Akure	46.28	52.45	0.178	-0.14	79.03	0.16	34.93	49.07
Enugu	45.96	62.96	0.13	-0.12	92.71	0.21	40.08	59.6
Onitsha	53.21	67.56	-0.09	-0.08	98.33	0.28	46.68	61.77
Owerri	66.7	75.27	0.241	0.13	113,65	0.67	89.29	60.21

(k = shape, α = scale, ε = location)

3.1. Selection of the most appropriate Probability Distribution with Goodness of Fit Test Results

Six goodness of fit tests and diagnostic tests. D – Index, were used to determine the most appropriate probability distribution of each location. Table 8 gives the values obtained for these tests by applying the equation in Table 3.

Table 8 The distributions' Goodness-of-Fit test results at Ikeja

S/N	Distribution	RMSE	RRMSE	MADI	MAE	PPCC	D-INDEX
1	EVI	7.44	0.075	0.0076	25.76	0.9865	0.595
2	GEV	12.34	0.139	0.0696	35.78	0.9890	0.908
3	GPA	14.23	0.760	0.540	98.62	0.7100	2.61
4	EVI/ L- Moment	8.5325	0.0792	0.0663	23.78	0.9835	2.895
5	GEV/ L-Moment	44.376	0.4088	0.3206	124.8	0.9915	2.835
6	GPA/ L-Moment	5.4735	0.04725	0.0353	13.43	0.9935	0.267

3.2. The Goodness of Fit Test Values are used to evaluate probability distribution models

The total score acquired from all of the tests were used to evaluate the probability distribution models. Each distribution model was given a score from one to six (1-6) based on the criteria that the distribution model with the highest score was chosen as the optimal distribution model for the data of a specific city. The best-supported distribution receives a score of six (6), the second-best receives a score of five (5), and so on in descending order. Table 9 was constructed utilizing the above-mentioned grading scheme. The results of the overall ranking are shown in Table 9.

Table 9 Scoring and ranking scheme for distribution at Ikeja

DIST	RMSE	RRMSE	MADI	MAE	PPCC	D - INDEX	Total	Rank
EVI/MOM	5	5	6	4	3	5	28	2
GEV/MOM	3	3	3	3	4	4	20	3
GPA/MOM	2	1	1	2	1	3	10	5
EVI/ L- Moment	4	4	4	5	2	1	20	3
GEV/ L-Moment	1	2	2	1	5	2	13	4
GPA/ L-Moment	6	6	5	6	6	6	35	1st

3.3. Summary of the Presentation of The Six Probability Models

Table 10 displays a summary of the six probability models presentation from their overall scores. Table 10 listed, at each site as the most appropriate model and the second most appropriate model as well as their scores. For Ikeja, the total score of GPA/ L-Moment for the different Goodness-of-Fit tests was 35 points. This being the highest score, was ranked the first GPA estimated by L-Moment was considered the best-fit distribution model. This method was adopted for other locations.

Table 10 The Selected Model for the Peak Rainfall and the Goodness of Fit Test.

S/N	Location	Most Appropriate Model	Total Max Score	Second Most Appropriate Model	Total Max Score
1	Ikeja	GPA/ L- Moment	35	EVI/MOM	28
2	Akure	EVI/MOM	33	EVI/L-M	26
3	Ibadan	GEV/MOM	34	EVI/MOM	26
4	Benin city	GEV/MOM	31	GEV/L-M	28
5	Port Harcourt	GPA/ L- Moment	33	EVI/ L- Moment	30
6	Uyo	EVI/MOM	35	GEV/MOM	31
7	Calabar	EVI/ L- Moment	32	GPA/ L- Moment	31
8	Onitsha	GPA/ L- Moment	32	EVI/ L- Moment	28
9	Enugu	GPA/ L- Moment	32	EVI/ L- Moment	29
10	Owerri	GPA/ L- Moment	32	GEV/MOM	31

In summary, the most appropriate estimators were L-Moments in six locations and Methods of Moments in four locations.. This indicates that L-Moments are the preferred estimator in six locations over the conventional method of moments because L-moments, being linear functions of the data, suffer less from the effects of sampling variability [2]. However, according to [6], for lower skewness, and the absence of outliers, the conventional method of moments is favored, especially for smaller samples. Hence, Methods of Moments is the preferred estimator in four locations.

Model validation results obtained using PPCC Goodness of fit Statistic is presented in Table 11

Table 11 Model validation result with PPCC Goodness of Fit Statistic

Location	Most Appropriate Model	Test Performed	Calculated Values PPCC test= r	Numbers of Observations	Critical values at 5% significance Level = r*	Decision: Reject Ho : when PPCC r < r*
Ikeja	GPA/L- Moment	PPCC	0.9935	50	0.977	Accept H_O
Akure	EVI/MOM	PPCC	0.974	33	0.967	Accept H_O
Ibadan	GEV/MOM	PPCC	0.978	50	0.977	Accept H_O
Benin city	GEV/MOM	PPCC	0.979	48	0.976	Accept H_O
Port Harcourt	GPA/L- Moment	PPCC	0.973	40	0.972	Accept H_O
Uyo	EVI/MOM	PPCC	0.9916	32	0.966	Accept H_O
Calabar	EVI/ L- Moment	PPCC	0.988	49	0.976	Accept H_O
Onitsha	GPA/L- Moment	PPCC	0.979	32	0.966	Accept H_O
Enugu	GPA/L- Moment	PPCC	0.9913	49	0.976	Accept H_O
Owerri	GPA/L- Moment	PPCC	0.992	32	0.966	Accept H_O

The PPCC test statistics in Table 11 were calculated using Equation (6) and the Critical values of r* at 5% significance level were interpolated in PPCC Table in Standard textbooks. The decision column of the table implies that we have to accept the null hypotheses (H_o): that the data is well-fitting with the Probability Distribution Model. Th**is** has validated our models outlined in Table 5 and therefore it can be reliably used for future rainfall prediction as was carried out in Table 7.

3.4 Forecasted Rainfall (mm) for Varying Return Periods at Each Location

The Quantile values in Table 12 were calculated using the best-fit probability distribution provided in Table 10. The results of several analyses led to Table 12 which shows the rainfall return levels (mm) for given return times ranging from five to two hundred years, as well as the most appropriate probability distribution model for each station.

Table 12 Model validation result with PPCC Goodness of Fit Statistic

S/N	Location	Most Appropriate Distribution	Return Period (Years)					
			5	10	25	50	100	200
1	Ikeja	GPA/ L-Moment	140.77	171.67	209.33	236.63	231.76	256.90
2	Akure	EVI/MOM	103.58	117.85	134.39	149.27	162.54	175.78
3	Ibadan	GEV/MOM	105.7	112.5	119.65	121.96	123.32	124.18
4	Benin city	GEV/MOM	146.9	166.34	184.45	194.39	202.07	208.006
5	Port Harcourt	GPA/ L- Moment	121.65	137.39	152.99	161.79	168.64	173.97
6	Uyo	EVI/MOM	116.99	131.98	149.39	165.03	179.99	192.92
7	Onitsha	GPA/ L- Moment	122.30	141.09	160.94	172.93	182.81	190.96
8	Enugu	GPA/ L- Moment	114.45	132.98	153.74	167.01	178.51	188.47
9	Owerri	GPA/ L- Moment	148.30	165.24	178.41	184.19	187.83	190.96
10	Calabar	EVI/ L- Moment	151.78	171.89	197.31	216.16	234.88	253.53

3.5 Rainfall Frequency Curves (RFCs)

The Rainfall Frequency Curves was developed using the yearly extreme rainfall. Table 7 shows the estimates for the probability distribution models depicted in Figure 3,4, and 5.

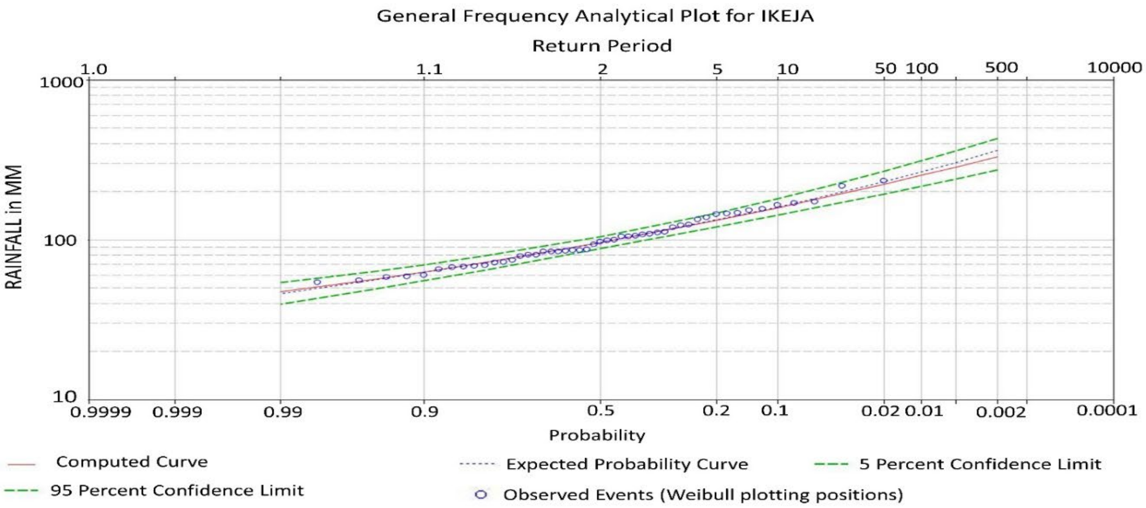

Figure 3 Rainfall frequency curve (RFC) for Ikeja

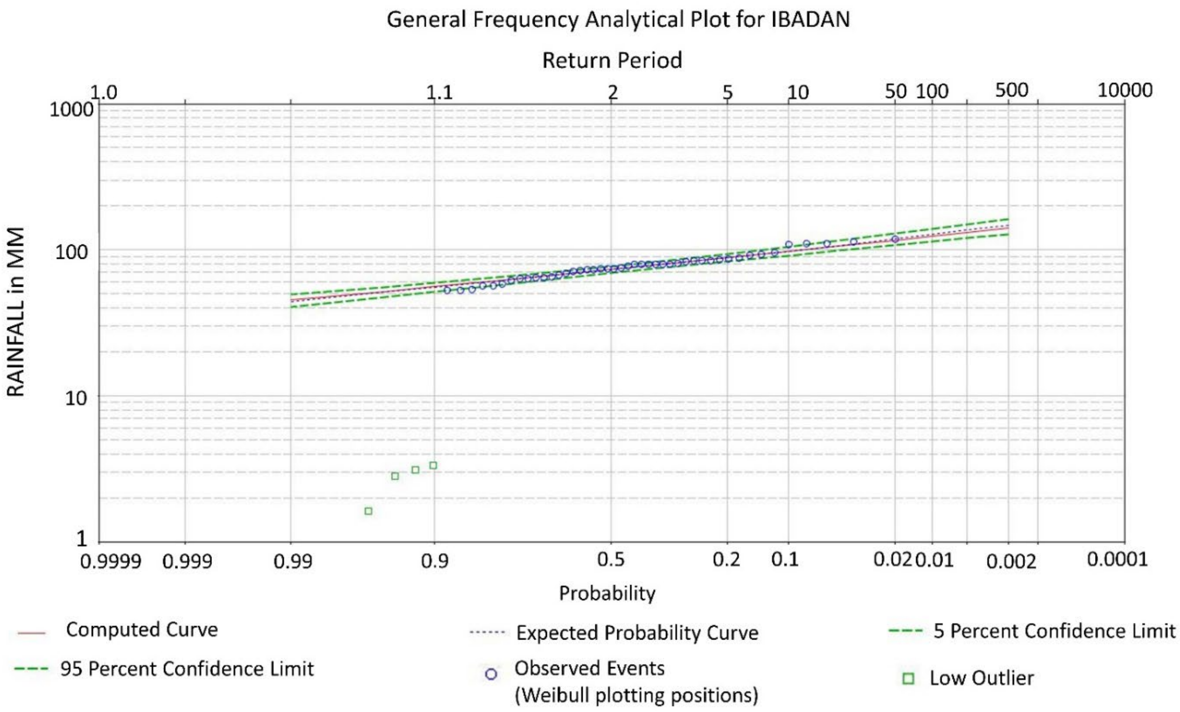

Figure 4 Ibadan's rainfall frequency curve (RFC)

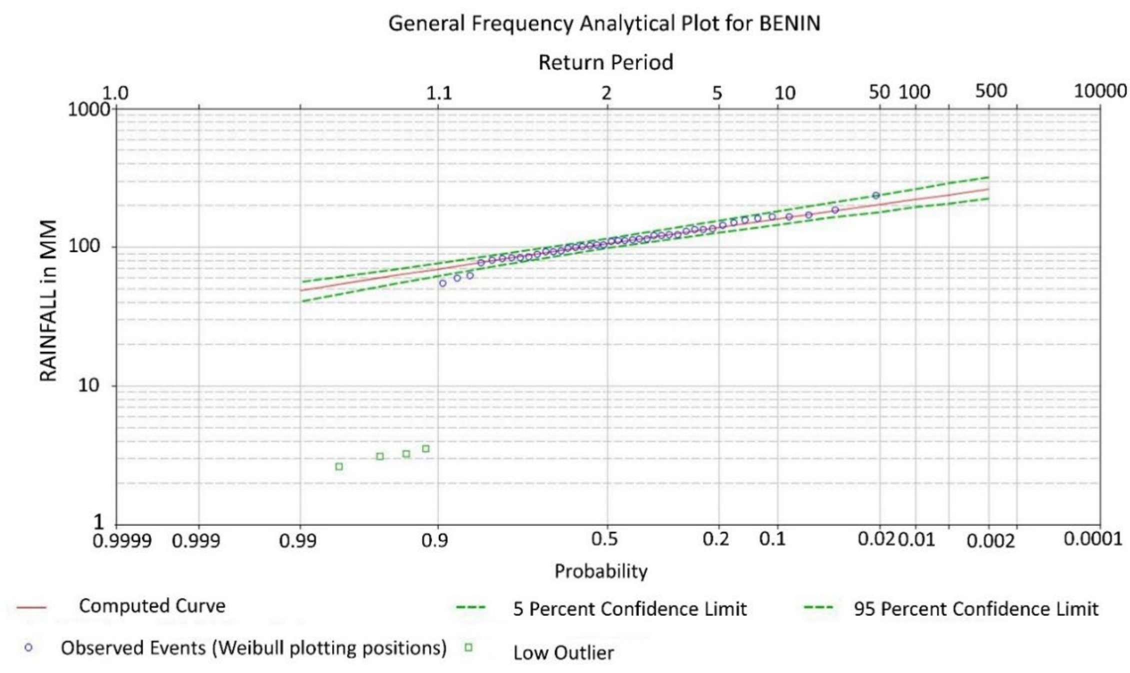

Figure 5 Rainfall frequency curve (RFC) for Benin City

These Rainfall Frequency Curves (RFC), which represent ten curves in the study area are useful and effective tools for forecasting rainfall frequency distributions and determining T-year rainfall. They can also be used to plan, build, and at a variety of sites, manage hydraulic systems for flood reduction and flood damage avoidance.

4.0 CONCLUSION

The most appropriate probability distribution model for each city was found by applying two-parameter estimators to three probability distributions for the selected cities in Southern Nigeria. Each site's rainfall return levels of engineering importance were forecasted using the most appropriate probability distribution models. The findings lead to the following conclusions:

1. At five sites, the GPA probability distribution model with the L-Moment parameter estimator was the best-fit probability model. At two sites, the GEV and EVI probability distribution models with the Method of Moments parameter estimator were the best-fit probability models. At one site, the EVI probability distribution model with L-Moment parameter estimator was the best-fit probability model.

2. The best-fit probability model with parameter estimator is site-specific at any location. particular as a result, no single model is thought to be preferable for all practical purposes; instead, models are chosen based on the circumstances at hand and the nature of the data presented.

3. Quantile estimations or expected rainfall levels for various return durations for appropriate design criteria needed for the preparation and the construction of flood-relieving hydraulic infrastructure and flood prevention in diverse areas have been provided.

4. Curves of rainfall frequency for Ikeja, Ibadan, and Benin City have been supplied. They were created using predicted rainfall values generated with the most appropriate distribution model. These are useful for water resources design guidelines.

5. In this period of frequent bridge, home, dam, and drainage collapse as a result of improper design, engineers must incorporate all relevant criteria. For effective works, statistical inputs during the design process are required. This research has supplied useful engineering design parameters for improved hydrological design and planning which are required for effective Hydraulic Structure Design for Flood Control and Mitigation, and Prevention in a range of settings. Engineers that apply these features will see a significant reduction in or complete elimination of design failures.

Conflict of Interests

There are no potential conflicts of interest among the authors.

Acknowledgement

There was no sponsorship for this research work.

References

[1] Sankarasubramanian, A. & Srinivasan, K. (1999). Investigation and comparison of sampling properties of L-moments and conventional moments. Journal of Hydrology, 218 (1–2), 13–34. https://doi.org/10.1016/S0022-1694(99)00018-9

[2] Hosking, J. R. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. Journal of the Royal Statistical Society: Series B (Methodological), 52(1), 105-124.https://doi.org/10.1111/j.2517-6161.1990.tb01775.x

[3] Vogel, R. M., & Fennessey, N. M. (1993). L moment diagrams should replace product moment diagrams. Water resources research, 29(6), 1745-1752. https://doi.org/10.1029/93WR00341

[4] Hosking, J. R. M. (1986). The theory of probability weighted moments, Res. Report RC12210, IBM Research, Yorktown Heights.

[5] Stedinger, J. R., Vogel, R. M., & Georgiou, E.F. (1993). Frequency analysis of extreme events Chap. 18. In: Maidment, D.J. (Ed.). Handbook of Hydrology, McGraw Hill, New York.

[6] Singh, R. K. (2001). Probability analysis for prediction of annual maximum rainfall of Eastern Himalaya (Sikkimmid hills). Indian Journal of Soil Conservation, 29, 263–265.

[7] Kousar, S., Khan, A. R., Ul Hassan, M., Noreen, Z., & Bhatti, S. H. (2020). Some best-fit probability distributions for at-site flood frequency analysis of the Ume River. Journal of Flood Risk Management, 13(3), e12640. https://doi.org/10.1111/jfr3.12640

[8] Mamman, M. J., Martins, O. Y., Ibrahim, J., & Shaba, M. (2017). Evaluation of Best-Fit Probability Distribution Models for the Prediction of Inflows of Kainji Reservoir, Niger State, Nigeria Air Soil and Water Research vol 10, 1-7 https://doi.org/10.1177/1178622117691034

[9] Gocic, M., Velimirovic, L., Stankovic, M., & Trajkovic, S. (2021). Determining the best fitting distribution of annual precipitation data in Serbia using L-moments method. Earth Science Informatics, 14(2), 633-644. https://doi.org/10.1007/s12145-020-00543-9
[10] Eslamian, S. S., & Feizi, H. (2007). Maximum monthly rainfall analysis using L-moments for an arid region in Isfahan Province, Iran. Journal of Applied Meteorology and Climatology, 46(4), 494-503. https://doi.org/10.1175/JAM2465.1
[11] Ghosh, S., Roy, M. K., & Biswas, S. C. (2016). Determination of the best fit probability distribution for monthly rainfall data in Bangladesh. American Journal of Mathematics and Statistics, 6(4), 170-174.
[12] Mohamed, T. M., & Ibrahim A. A. A. (2016). Fitting Probability Distributions of Annual Rainfall in Sudan. SUST Journal of Engineering and Computer Sciences (JECS), Vol. 17, No.2
[13] Alam, M. A., Emura, K., Farnham, C., & Yuan, J. (2018). Best-fit probability distributions and return periods for maximum monthly rainfall in Bangladesh. Climate, 6(1), 9. https://doi.org/10.3390/cli601000
[14] Anil, K. (2000). Prediction of annual maximum daily rainfall of Ranichauri (Tehri Garhwal) based on probability analysis. Indian Journal of Soil Conservation, 28(2), 178-180.
[15] Amin, M. T., Rizwan, M., & Alazba, A. A. (2016). A best-fit probability distribution for the estimation of rainfall in northern regions of Pakistan. Open Life Sciences, 11(1), 432-440. https://doi.org/10.1515/biol-2016-0057
[16] Masereka, E. M., Otieno, F. A. O., Ochieng, G. M., & Snyman, J. (2015). Best fit and selection of probability distribution models for frequency analysis of extreme mean annual rainfall events. International Journal of Engineering Research and Development, 11(4), 34-53.
[17] Langat, P. K., Kumar, L., & Koech, R. (2019). Identification of the most suitable probability distribution models for maximum, minimum, and mean streamflow. Water, 11(4), 734. https://doi.org/10.3390/w11040734
[18] Okeke, O. B., & Ehiorobo, J. O. (2017). Frequency analysis of rainfall for flood control in Patani, Delta State of Nigeria. Nigerian Journal of Technology, 36(1), 282-289. https://doi.org/10.4314/njt.v36i1.34
[19] Izinyon, O. C., & Ajumuka, H. N. (2013). Probability distribution models for flood prediction in Upper Benue River Basin-Part II. Civil and Environmental Research, 3(2), 62-74.
[20] Ologhadien, I. (2021). Selection of Probabilistic Model of Extreme Floods in Benue River Basin, Nigeria. European Journal of Engineering and Technology Research, 6(1), 7-18. https://doi.org/10.24018/ejers.2021.6.1.2300
[21] Anandan, V. (2014). Comparison of probability distributions for frequency analysis of annual maximum rainfall. International Journal of Research and Innovative Technology, IJRIT, 1(3), 50-55.
[22] Vivekanandan, N. (2015). Flood frequency analysis using method of moments and L-moments of probability distributions. Cogent engineering, 2(1), 1018704., https://doi.org/10.1080/23311916.2015.1018704
[23] Greenwood, J. A., Landwehr, J. M., Matalas, N. C., & Wallis, J. R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. Water resources research, 15(5), 1049-1054. https://doi.org/10.1029/WR015i005p01049
[24] Chadwick, A., Morfett, J., & Borthwick, M. (2021). Hydraulics in civil and environmental engineering. Crc Press. Taylor & Francis Group,5TH Edition, 340-342. https://doi.org/10.1201/b14556
[25] Ojha, C., Berndtsson, R., & Bhunya, P. (2008). Engineering Hydrology, Oxford University Press, New Delhi, India Ch 7, 248-289
[26 Filliben, J. J. (1975). The probability plot correlation coefficient test for normality. Technometrics, 17(1), 111-117. https://doi.org/10.1080/00401706.1975.10489279
[27 Agbonaye A. I. and Izinyon O. C. (2017). Best-Fit Probability Distribution Model for Rainfall Frequency Analysis of Three Cities in South-Eastern Nigeria, Nigerian Journal of Environmental Sciences and Technology, 1(1), 34-42. https://doi.org/10.36263/nijest.2017.01.0024
[28] Hao, W., Hao, Z., Yuan, F., Ju, Q., & Hao, J. (2019). Regional frequency analysis of precipitation extremes and its spatio-temporal patterns in the Hanjiang River Basin, China. Atmosphere, 10(3), 130. https://doi.org/10.3390/atmos10030130
[29] Izinyon, O. C., & Ehiorobo, J. O. (2015). L-moments method for flood frequency analysis of river Owan at Owan in Benin Owena River basin in Nigeria. Current Advances in Civil Engineering, 3(1), 1-10. https://doi.org/10.4314/njt.v33i1.2
[30] Kumar, R. (2019). Flood frequency analysis of the Rapti river basin using log pearson type-III and Gumbel Extreme Value-1 methods. Journal of the Geological Society of India, 94(5), 480-484. https://doi.org/10.1007/s12594-019-1344-0
[31] National Water Development Authority (2019). Detailed Project Report (DPR) Burhi Gandak-Noon-Baya-Ganga intra-state link of Bihar (Chapter III Hydrology). http://nwda.gov.in/upload/upload files/files/ DPR_BG_N_BG_Ch_3.pdf. Accessed 28 March 2019.
[32] Ul Hassan, M., Hayat, O., & Noreen, Z. (2019). Selecting the best probability distribution for at-site flood frequency analysis; a study of Torne River. SN Applied Sciences, 1(12), 1-10. https://doi.org/10.1007/s42452-019-1584-z
[33] Drissia, T. K., Jothiprakash, V., & Anitha, A. B. (2019). Flood frequency analysis using L moments: a comparison between at-site and regional approach. Water Resources Management, 33(3), 1013-1037. https://doi.org/10.1007/s11269-018-2162-7
[34] Ul Hassan, M., Noreen, Z., & Ahmed, R. (2021). Regional frequency analysis of annual daily rainfall maxima in Skåne, Sweden. International Journal of Climatology, 41(8), 4307-4320. https://doi.org/10.1002/joc.7074.
[35] Bajirao, T. S. (2021). Comparative performance of different probability distribution functions for maximum rainfall estimation at different time scales. Arabian Journal of Geosciences, 14(20), 1-15. https://doi.org/10.1007/s12517-021-08580-4