Physical Science International Journal

19(4): 1-13, 2018; Article no.PSIJ.44983
ISSN: 2348-0130

Probability Distribution of Daily Rainfall Pattern over Some Selected Stations in North Western Nigeria

I. Garba[1,2*], A. Akinbobola[1] and E. C. Okogbue[1]

[1]Department of Meteorology and Climate Science, Federal University of Technology, Akure, Nigeria.
[2]Nigeria Meteorological Agency, Abuja, Nigeria.

Authors' contributions

This work was carried out in collaboration between all authors. Author IG designed the study, performed the analysis, wrote the protocol and wrote the first draft of the manuscript. Author AA managed the analyses of the study. Author ECO managed the literature searches. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/PSIJ/2018/44983
Editor(s):
(1) Dr. Roberto Oscar Aquilano, School of Exact Science, National University of Rosario (UNR), Rosario, Physics Institute (IFIR)(CONICET-UNR), Argentina.
Reviewers:
(1) Olumide Adesina, Olabisi Onabanjo University, Nigeria.
(2) Ionac Nicoleta, University of Bucharest, Romania.
(3) Nitin Mishra, Graphic Era University, India.
Complete Peer review History: http://www.sciencedomain.org/review-history/27209

Original Research Article

Received 21 August 2018
Accepted 01 November 2018
Published 14 November 2018

ABSTRACT

Analysis of rainfall distribution is important in studying the impact of changing weather and climate on water resources planning and management. This study assessed the performance of three different probability distribution models, namely: Generalized Extreme Value Distribution (GEV), Lognormal Distribution (LNG) and Gumbel (EV1) Distribution to describe the rainfall distribution patterns in some selected stations (Gusau, Yelwa and Sokoto) in North-Western Nigeria. Thirty years of daily rainfall data for the period of (1985-2014) for the selected stations were obtained from the archives of the Nigerian Meteorological Agency (NIMET), Abuja. The aim of the study is to determine the distribution model that best describes the distribution of daily precipitation in North-Western Nigeria and also to identify the effect of plotting position on existing models. Root mean square error (RMSE) was used to determine the efficiency of the different plotting formulae on the existing model. The model performance was evaluated based on the statistical goodness of fit test, namely Kolmogorov-Smirnov (KS) at 95% (α=0.05) significant level. The method of moment (MOM), Probability weight-moment and maximum of likelihood were used for estimating the models

**Corresponding author: E-mail: ibrahimgarbakanwa@gmail.com;*

parameters. The result shows that the probability distribution model that best fit the data based on statistical of goodness of fit test is the EV1 followed by LNG then GEV. The EV1 model gave the smallest value of KS test for all stations except Gusau station where LNG model was the most suitable. The best plotting position formula with the entire distribution model was also observed to be the Weibull plotting position formula followed by Chegodayev, Gringoten and Hazen plotting position respectively. Hazen plotting position gave minimal errors with the EV1 and GEV probability distributions, while Weibull gave a minimal error with the LNG probability distribution for all the stations. The EV1 model was found to be the most suitable distribution for modelling the daily rainfall distribution in two out of the three stations investigated, while the LNG was observed to be only suitable for Gusau. The result of this work has provided information on rainfall probabilities as a vital tool for the design of water supply and supplemental irrigation schemes and the evaluation of alternative cropping and of soil water management plans.

Keywords: Daily rainfall; probability distribution; plotting position; generalized extreme value; Gumbel; Log-normal; goodness of fit; Sokoto; Gusau and Yelwa.

1. INTRODUCTION

Rainfall is the most important natural factor that determines the agricultural production in Nigeria, particularly in the North-Western part of Nigeria. The variability of rainfall and the pattern of extremely high or low precipitation are very important for agriculture as well as the economy of the state. It is well established that rainfall is changing on both the global and the regional scales due to global warming [1]. As the moves to encourage agriculture to ensure food security continues to gain ground and acceptability, information on rainfall probabilities is vital for the design of water supply and supplemental irrigation schemes and the evaluation of alternative cropping and of soil water management plans. Climatic variation in northern Nigeria is not altogether new because this part of Nigeria contains a significant portion of the Sudan-Sahel ecological zone of West Africa. However, since the early 1970s, climatic anomalies in the form of recurrent droughts, frightening dust storms and rampaging floods have overprinted their rhythms, creating short-duration climatic oscillations as against the normal cycles of larger amplitudes. Thus, the last 40years have witnessed four severe droughts, numerous dust storms and three killer floods. Indeed, the climate of the region has become highly unpredictable, making many people wonder what has happened to the climate [5]. Most of the droughts that occur in this region have been found to be associated with a late start of the rainy season and early cessation of the rains, resulting in drastic reductions of the length of the rainy season. For instance, a number of studies within the Sahel region have shown a significant trend towards false onset (false onset is a situation where the rainy season starts normally and then ceases abruptly, creating a dry period between the false onset and the true onset, late or delayed onset (late or delayed onset is a situation where the expected start of the rainy season is delayed) and early cessation (early cessation is a situation where the rainy season stops far ahead of the expected time) of the summer rains over a 30-year period from 1969 to 1998 [6,7,8]. The use of statistics has a wide range of important applications in climate research, as Climatology is usually the synthesis or the climatic elements for a resource the study of the statistics of length of time not less than 30 years.

The use of statistical distribution is applied to historical data, which is fit to the desired distribution. The parameters of the distribution were estimated, and from this information, the Cumulative Distribution Function (CDF) and Probability Density Function (PDF) are created. The distributions are also used to estimate the probability of future maximum occurrences, which is needed for design and planning with new ideas about more appropriate distributions emerging, studies must be done to ensure we are using the most accurate method available. Analysis of rainfall data strongly depends on its distribution pattern. It has long been a topic of interest in the fields of meteorology in establishing a probability distribution that provides a good fit to daily, monthly, seasonally and annual rainfall and of the most important problems in frequency analysis is the selection of an appropriate probability distribution for a given hydrological data set. In order to accurately design and manage flood control structures, including sewers, reservoirs, culvert and dams, and the appropriate way of estimating these extreme events must be determined. Engineers, Meteorologist and Hydrologist, as well as many other professional, have the responsibility of accurately assessing these risks and taking them into account during the design process.

2. METHODOLOGY

2.1 Description of the Study Areas

The study area is within north-western Nigeria, they are located in three states of Sokoto, Kebbi and Zamfara. The area is located between latitudes 10°N and 13°58’ N; and longitudes 4°8’E and 6°54’ E. It covers a land area of approximately 62,000km^2. It lies to the North-Western of Nigeria and shares its borders with the Niger Republic to the North, Katsina State to the East, Niger State to the South-east Kwara State to the South and Benin Republic to the west. The southern boundary is arbitrarily defined by the Sudan savanna. Like the rest of West Africa, the climate of the region is controlled largely by the two dominant air masses affecting the sub- region. These are the dry, dusty, tropical- continental (cT) air mass (which originates from the Sahara Desert), and the warm, tropical- maritime (mT) air mass (which originates from the Atlantic Ocean). The influence of both air masses on the region is determined largely by the movement of the Inter-Tropical discontinuity (ITD), a zone representing the surface demarcation between the two air masses. The interplay of these two air masses gives rise to two distinct seasons within the sub-region. The wet season is associated with the tropical maritime air mass, while the dry season is a product of the tropical continental air mass.

2.2 Method of Analysis

Thirty years (1985-2014) of daily rainfall data were obtained from Nigerian Meteorological Agency (NIMET), Abuja. The data cover the three synoptic meteorological stations located within the study areas. The Generalized Extreme Value (GEV) Probability Distribution, Gumbel (EVI) Probability Distribution and Log-Normal (LNG) Probability Distribution were used in this work because it has a wide variety of applications for estimating extreme values of given data sets. They are commonly used in hydrological applications. This explains how the data were fitted to the Probability Distribution, Parameters Estimation, Return period, Probability of Exceedance and testing of Goodness of fit. Coefficients of determination and root mean square errors (RMSE) between predicted and observed values were used to determine how well the predicted values were able to predict the observed values. The distribution functions are stated below:

(i) Generalized Extreme Value Distribution

The GEV distribution is a family of continuous probability distributions that combines the Gumbel (EVI), Fréchet and Weibull distributions. GEV makes use of 3 parameters: location, scale and shape. When shape parameter (k) = 0, this is the EV_1 distribution. When K>0, this is EV_2 (Fréchet), and when k<0 is the EV, (Weibull).

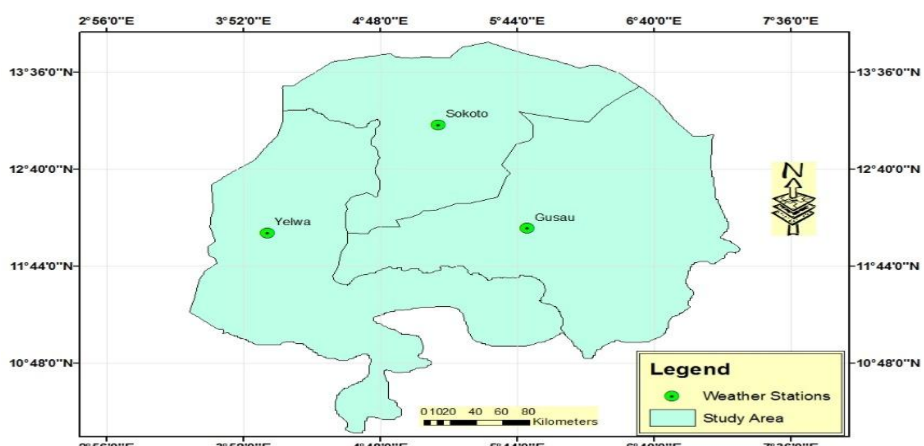

Fig. 1. Map of study the area (Sokoto, Yelwa and Guasu)

The Cumulative Distribution Function and Probability Density Function in [2] as:

$$F(X) = exp\{-(1 - k(x - \xi)/a)1 / k\} \quad (1)$$

$$f(x) = a^{-1}exp[-(1 - k)y - exp(-y)] \quad (2)$$

where $y = -k^{-1}log\left[1 - \frac{k(x-\xi)}{a}\right]$ $where\ k \neq 0$

Where, ξ is the location parameter, α is the scale parameter, and k is the shape parameter

(ii) Gumbel Distribution (EVI)

Gumbel distribution is a continuous probability distribution that used to model the distribution of the maximum (or the minimum) of a number of samples of various distributions [3,4]. It is useful in predicting extreme rainfall, flood or another natural disaster will occur.

The Cumulative Distribution Function and Probability Density Function as define in [9,10,11] are:

$$F(X) = \exp\left[-\exp\left(-\frac{(x-\xi)}{\alpha}\right)\right] \quad (3)$$

$$f(x) = \alpha^{-1}exp\left(-\frac{x-\xi}{\alpha}\right)exp\left[-\exp\left(-\frac{x-\xi}{\alpha}\right)\right] \quad (4)$$

Where, ξ is the location parameter, α is the scale parameter

(iii) Lognormal Distribution

The probability density function and the Cumulative Distribution Function for this distribution are:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}}exp\left(-\frac{(y-\mu_y)^2}{2\sigma_y^2}\right) \quad (5)$$

$$x > 0, y = logx$$

$$Fx(x, \mu, \sigma) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{1nx-\mu}{\sigma\sqrt{2}}\right)\right] \quad (6)$$

Where x $\geq$0 is the Rainfall data, $\mu > 0$ is the mean and $\sigma > 0$ is the standard deviation of the lognormal Distribution.

In determining the effect of the plotting position on the probability distribution for daily rainfall estimation for the study area, four plotting positions were plotted against the daily rainfall of the study area.

3. RESULTS AND DISCUSSION

3.1 Goodness of Fit

Table 1 shows the value of Kolmogrove-Sminorve test statistic for each of the distribution tested namely: Gumbel, Lognormal and Generalized extreme value (GEV) distributions for Gusau, Sokoto and Yelwa stations. Of the three (3) model tested the Gumbel has a value of 0.07204 as the lowest followed by 0.08614 and 0.15869 for Lognormal and GEV in Sokoto respectively. This shows that Gumbel is found to be the best fit distribution for Sokoto and followed closely by the Lognormal model that has similar value and GEV distribution being the last. It can be seen that Yelwa station for the Kolmogrove-Sminorve test Gumbel has smallest value compared to Lognormal and GEV distribution this implies that Gumbel is the best fit for Yelwa rainfall data and Lognormal been the second best and the GEV distribution is seen to be the worst fit. While reverse is the case in the Gusau stations the table shows that Lognormal is observed to be the best and followed by Gumbel while the GEV distribution has the worst.

3.2 The Probability Density Functions

The histogram shown in Fig. 2 are the frequency plot of the daily rainfall in the stations. From the figure, it can be observed that the data skewed to the right which implies the entire rainfall data assume the model being tested so that the three (3) distribution function can be used to model daily rainfall. They clearly showed that the Gumbel (EV1) and Lognormal exhibited similar probability densities which were different from that of the GEV distributions. This plot shows that all of the Probability Density Function is consistent with the result of skewness calculated for daily Rainfall. This plot shows that the distribution model fit the daily Rainfall data approximately Gumbel.

Table 1. Goodness of fit test for Gumbel, lognormal and GEV distribution

Distribution	Kolmogrove-Sminorve test		
	Sokoto	Yelwa	Gusau
Gumbel	0.07204	0.05431	0.06795
Lognormal	0.08614	0.05814	0.08090
GEV	0.15869	0.19992	0.16228

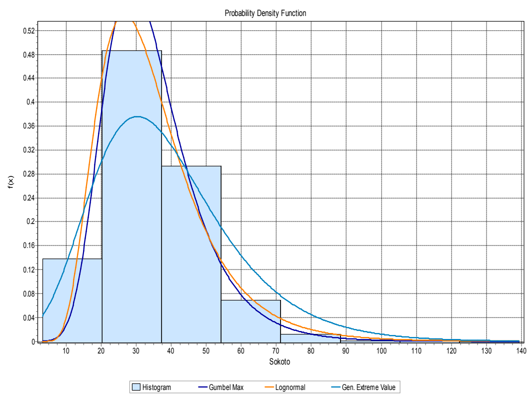

Fig. 2.1. The probability density function of daily rainfall of Sokoto

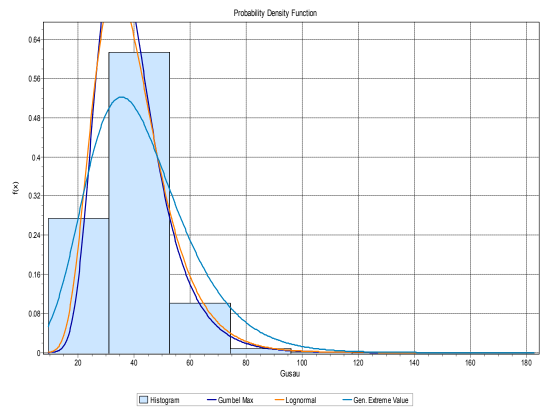

Fig. 2.2. The probability density function of data daily rainfall data of Gusau

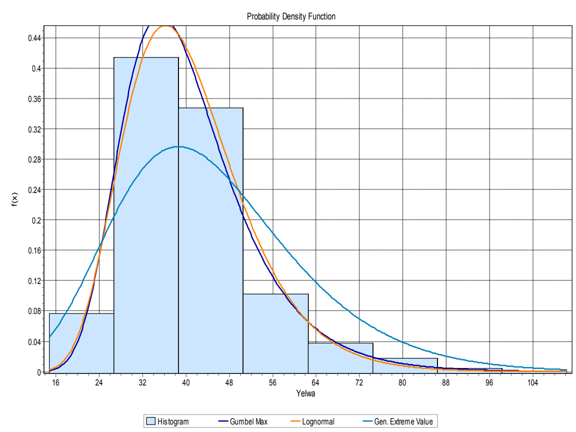

Fig. 2.3. The probability density function of daily rainfall data of Yelwa

3.3 Coefficient of Determination and Root Mean Square Error (RMSE)

From the distributions and plotting positions charts were constructed for each of the distributions and plotting positions used in the study, Fig. 3(a) – Fig. 11(d), it displays the charts for all stations. The estimated data is the one being estimated by the models or the distributions used which gives the coefficient of determination $R\text{^}2$ to determine how well each of the plotting position matches with the three distributions used, and others values of $R\text{^}2$ were tabulated in Tables 2-4.

Table 2. Coefficient of determination and root mean square error of Gusau station

Station	Plotting position	Probability distributions		
Gusau		**Gumbel**	**Lognormal**	**GEV**
Coefficient of Determination	Hazen	0.9341	0.8824	0.9429
	Gringoten	0.937	0.8849	0.9452
	Chegodayev	0.9426	0.8901	0.9496
	Weibull	0.9508	0.8984	0.9564
Root Mean Square Error				
	Hazen	-0.0019	-0.1610	-0.0091
	Gringoten	-0.0016	-0.1613	-0.0090
	Chegodayev	-0.0011	-0.1618	-0.0087
	Weibull	-0.0004	-0.1628	-0.0084

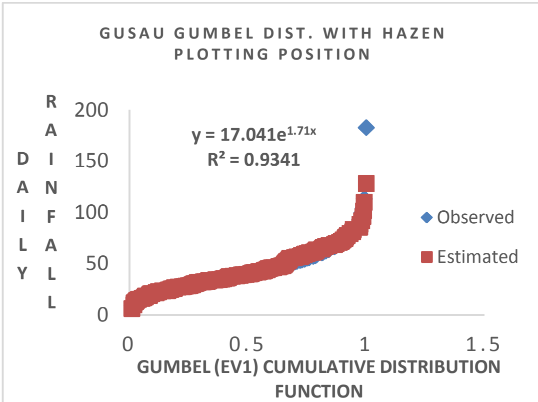

Fig. 3a. Gumbel distribution with Hazen plotting Position for Gusau station

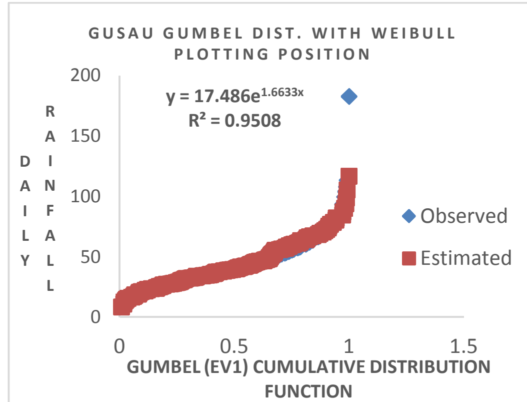

Fig. 3b. Gumbel distribution with Weibull Plotting Position for Gusau station

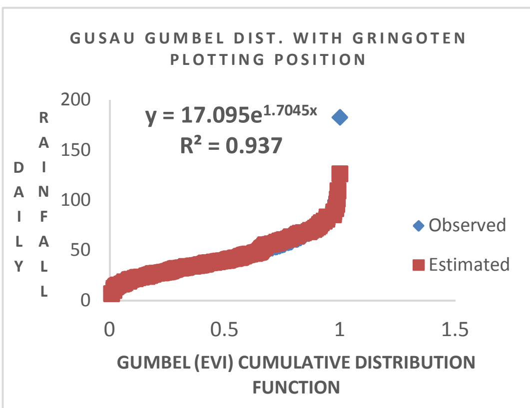

Fig. 3c. Gumbel distribution with Gringoten plotting position for Gusau station

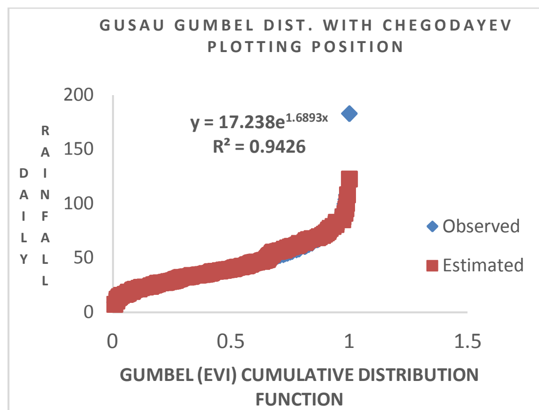

Fig. 3d. Gumbel distribution with Chegodayev plotting position for Gusau station

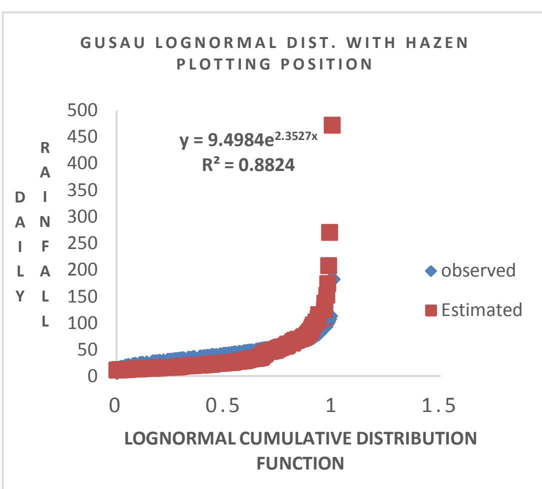

Fig. 4a. Lognormal distribution with Hazen Plotting Position for Gusau station

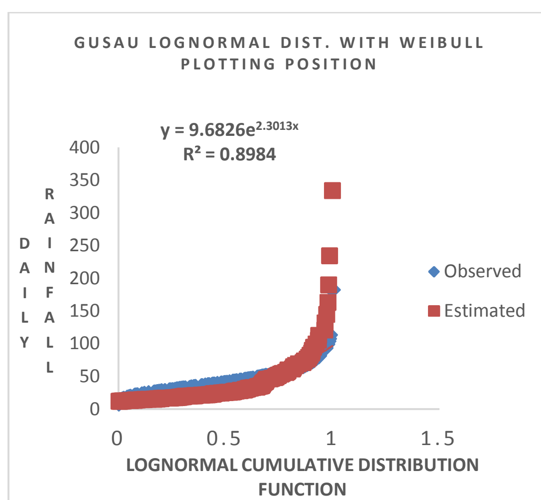

Fig. 4b. Lognormal distribution with Weibull Plotting Position for Gusau station

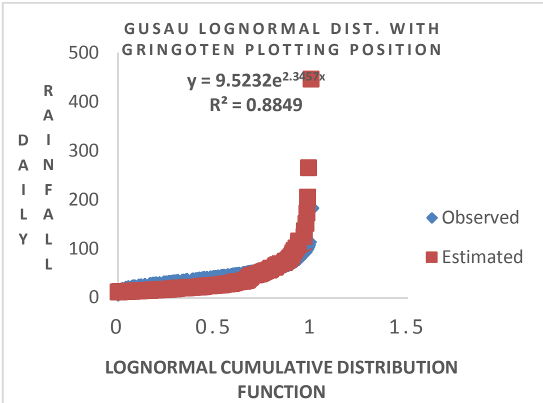

Fig 4c. Lognormal distribution with Gringoten plotting position for Gusau station

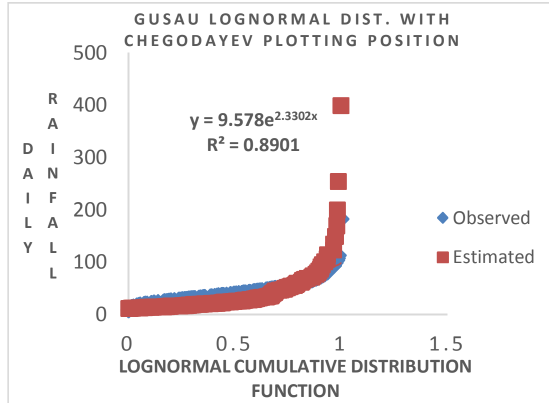

Fig. 4d. Lognormal distribution with Chegodayev plotting position for Gusau station

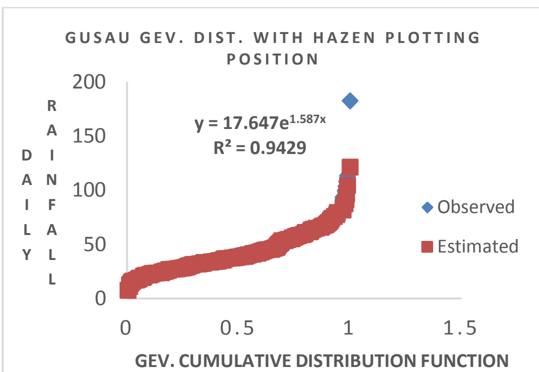

Fig. 5a. GEV distribution with Hazen plotting position position for Gusau station

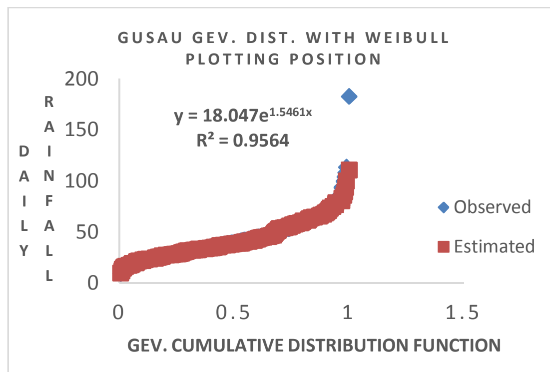

Fig. 5b. GEV distribution with Weibull plotting for Gusau station

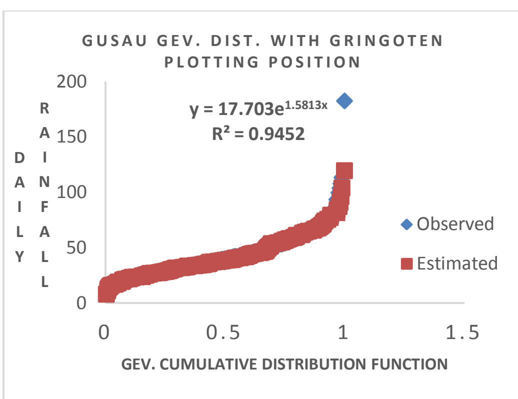

Fig. 5c. GEV distribution with Gringoten plotting position position for Gusau station

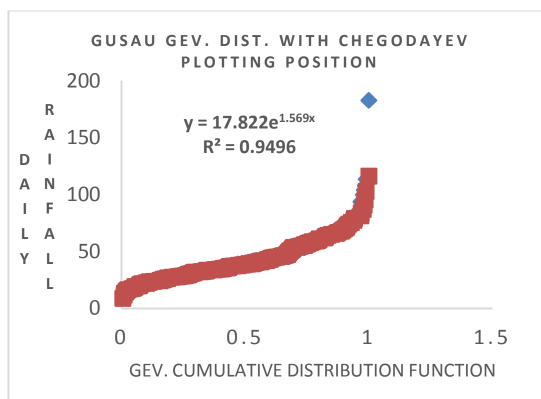

Fig. 5d. GEV distribution with Chegodayev Plotting for Gusau station

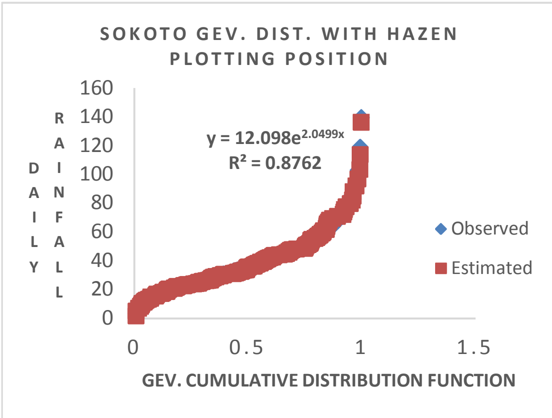

Fig. 6a. GEV distribution with Hazen plotting position position for Sokoto station

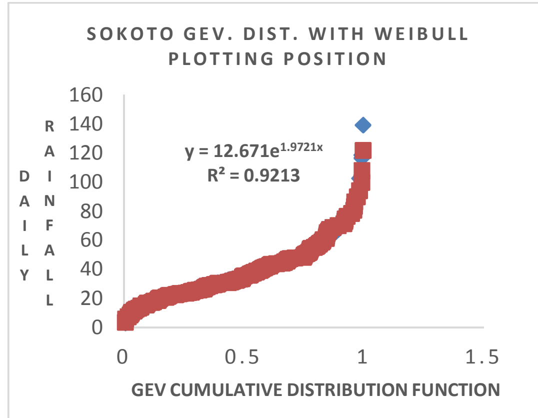

Fig. 6b. GEV distribution with Weibull plotting for Sokoto station

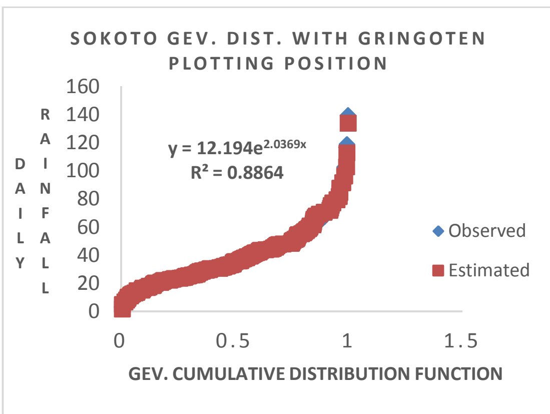

Fig. 6c. GEV distribution with Gringoten plotting position position for Sokoto station

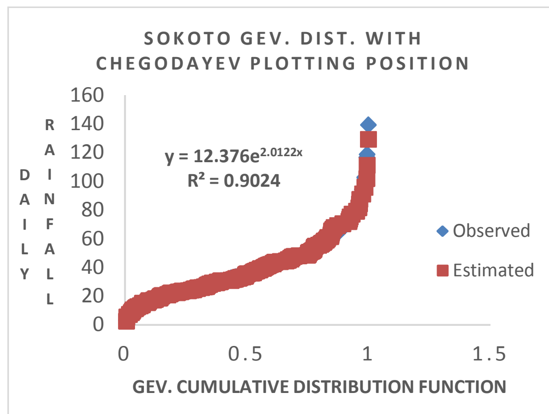

Fig. 6d. GEV distribution with Chegodayev plotting position for Sokoto station

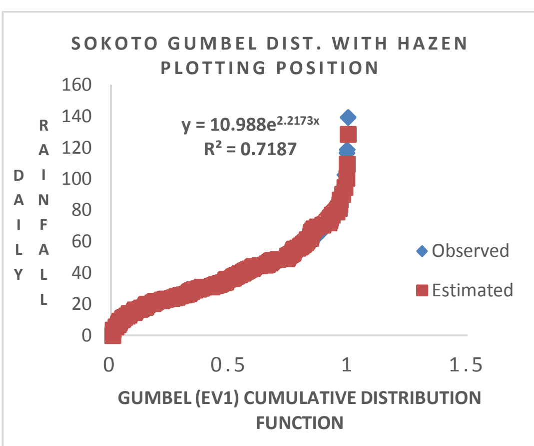

Fig. 7a. Gumbel distribution with Hazen plotting position Sokoto station

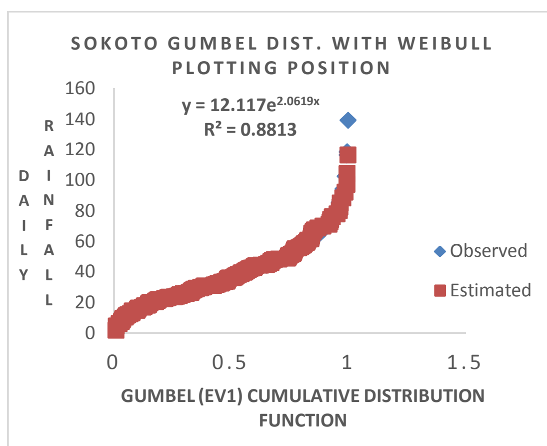

Fig. 7b. Gumbel distribution with Weibull plotting for position for Sokoto station

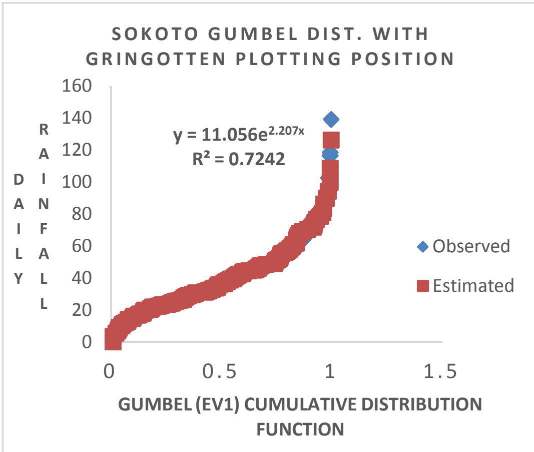

Fig. 7c. Gumbel distribution with Gringoten plotting position for Sokoto station

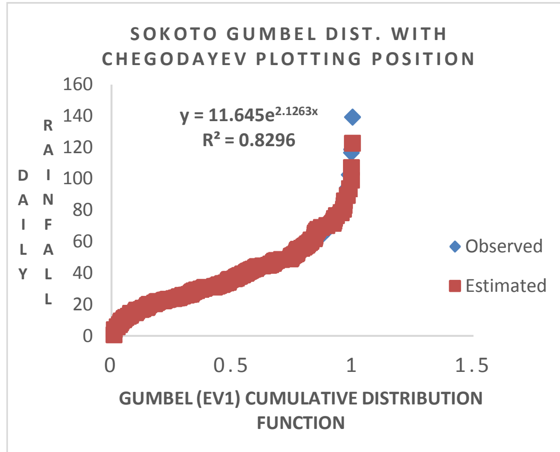

Fig. 7d. Gumbel distribution with Chegodayev plotting position for Sokoto station

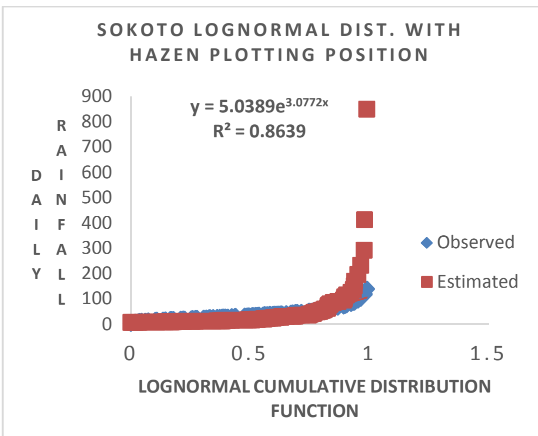

Fig. 8a. Lognormal distribution with Hazen plotting position for Sokoto station

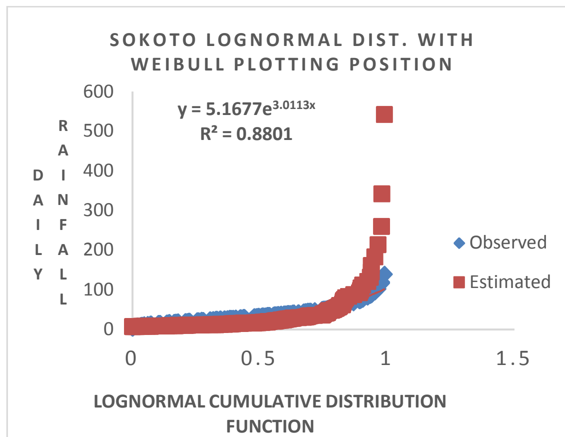

Fig. 8b. Lognormal distribution with Weibull plotting position for Sokoto station

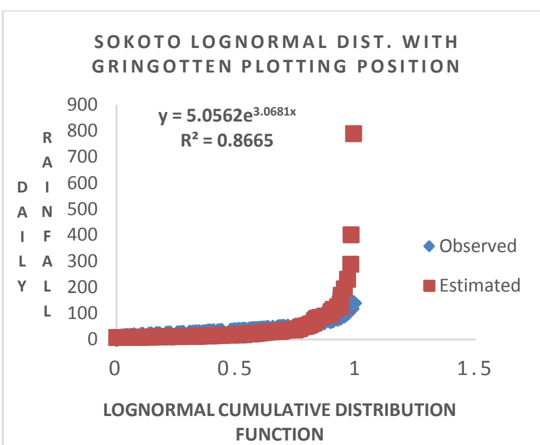

Fig. 8c. Lognormal distribution with Gringoten plotting position for Sokoto station

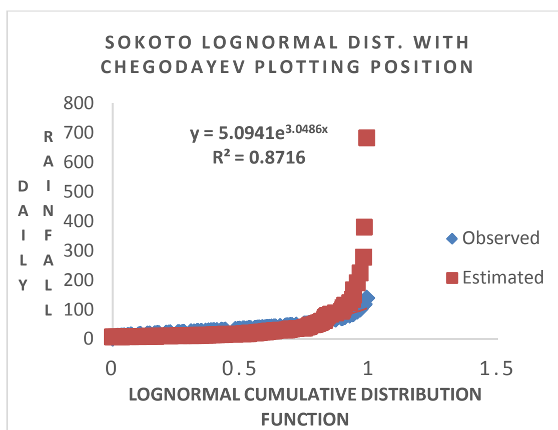

Fig. 8d. Lognormal distribution with Chegodayev plotting position for Sokoto station

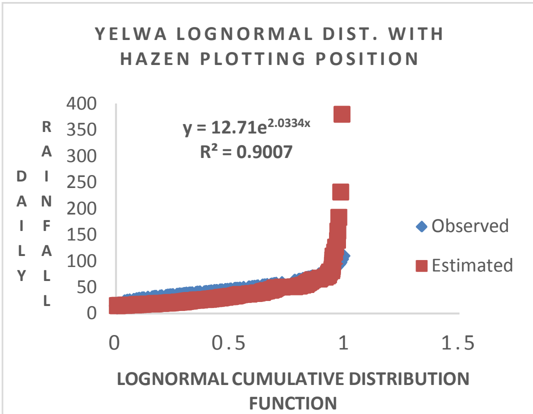

Fig. 9a. Lognormal distribution with Hazen plotting position for Yelwa station

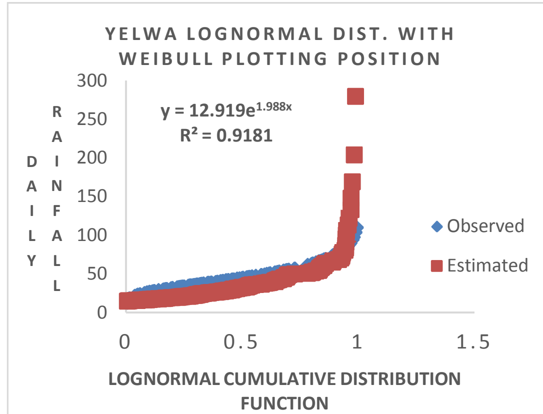

Fig. 9b. Lognormal distribution with Weibull plotting position for Yelwa station

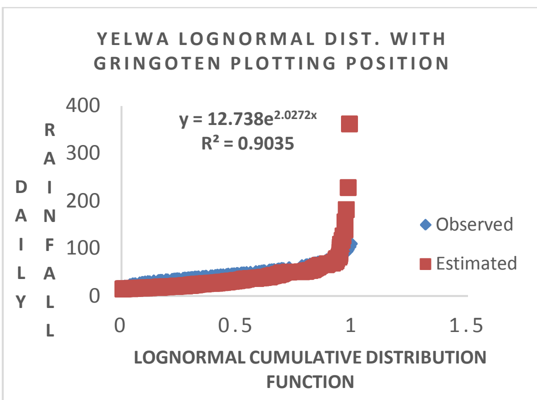

Fig. 9c. Lognormal distribution with Gringoten plotting position for Yelwa station

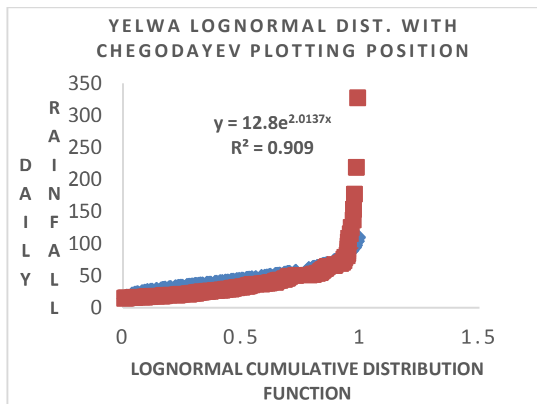

Fig. 9d. Lognormal distribution with Chegodayev plotting position for Yelwa station

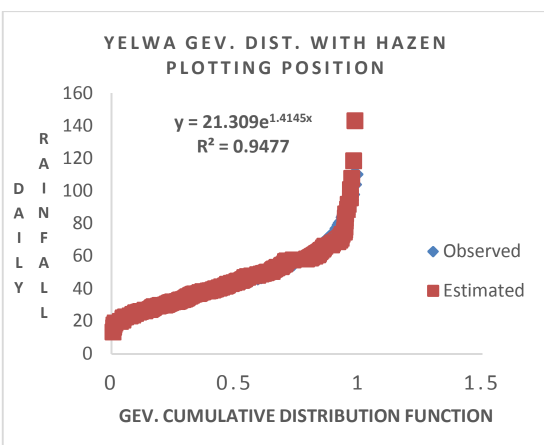

Fig. 10a. GEV distribution with Hazen plotting position for Yelwa station

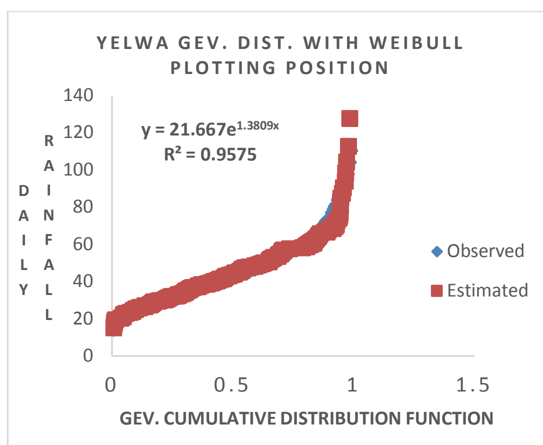

Fig. 10b. GEV distribution with Weibull plotting position for Yelwa station

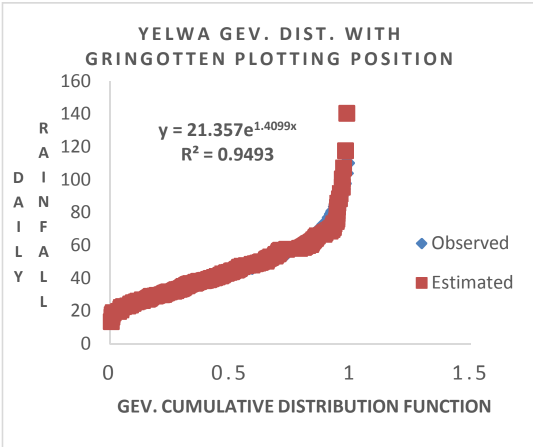

Fig. 10c. GEV distribution with Gringoten plotting position for Yelwa station

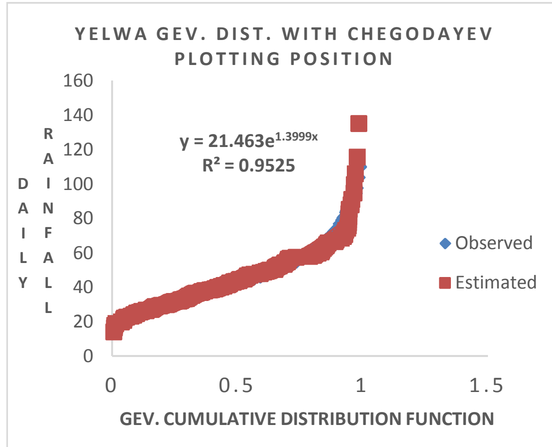

Fig. 10d. GEV distribution with Chegodayev plotting position for Yelwa station

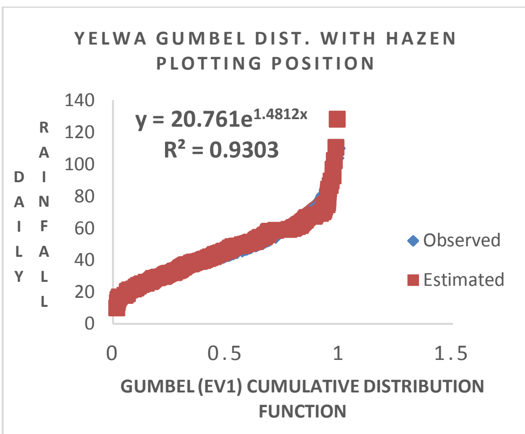

Fig. 11a. Gumbel distribution with Hazen plotting position for Yelwa Station

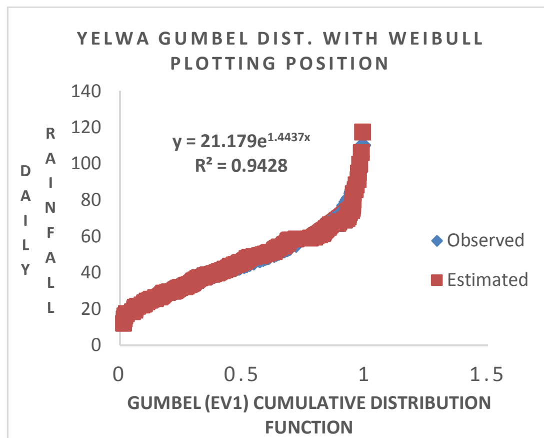

Fig. 11b. Gumbel distribution with Weibull plotting position for Yelwa station

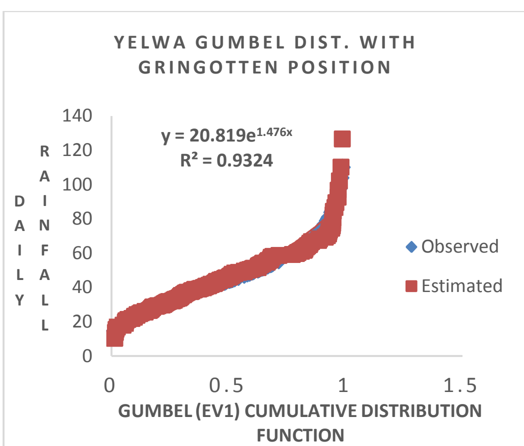

Fig. 11c. Gumbel distribution with Gringoten plotting position for Yelwa station

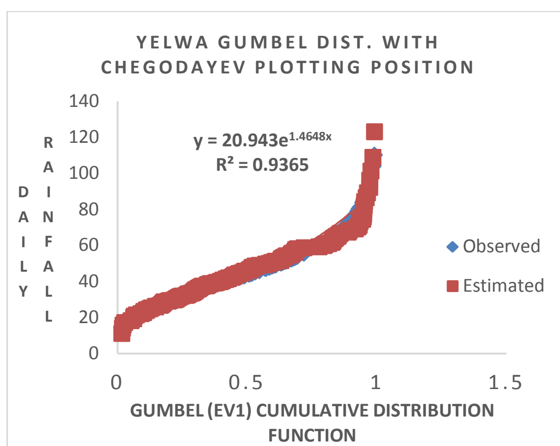

Fig. 11d. Gumbel distribution with Chegodayev plotting position for Yelwa station

Table 3. Coefficient of determination and root mean square error of Sokoto station

Station	Plotting position	Probability distributions		
Sokoto		**Gumbel**	**Lognormal**	**GEV**
Coefficient of Determination	Hazen	0.7187	0.8639	0.8762
	Gringoten	0.7242	0.8665	0.8864
	Chegodayev	0.8296	0.8716	0.9024
	Weibull	0.8813	0.8801	0.9213
Root Mean Square Error				
	Hazen	-0.0634	-0.2399	-0.0022
	Gringoten	-0.0580	-0.2401	-0.0011
	Chegodayev	-0.0097	-0.2405	0.0004
	Weibull	-0.0018	-0.2413	0.0020

Table 4 Coefficient of determination and root mean square error of Yelwa station

Station	Plotting position	Probability Distribution		
Yelwa		**Gumbel**	**Lognormal**	**GEV**
Coefficient of Determination	Hazen	0.9303	0.9007	0.9477
	Gringoten	0.9324	0.9035	0.9493
	Chegodayev	0.9365	0.909	0.9525
	Weibull	0.9428	0.9181	0.9575
Root Mean Square Error				
	Hazen	-0.0034	-0.1360	-0.00167
	Gringoten	-0.0033	-0.1362	-0.00166
	Chegodayev	-0.0029	-0.1367	-0.00164
	Weibull	-0.0024	-0.1376	-0.00163

4. DISCUSSION

The result show that the probability distribution models that best fit the data based on the statistical goodness of fit test is the EV1 followed by LNG and GEV. The Gumbel model gave the smallest value of Kolmogrove-Sminorve test for all stations except Gusau station where LNG model was the most suitable. The best plotting position formula with the entire distribution model was also observed to be Weibull plotting position formula followed by Chegodayev, Gringoten and Hazen plotting position respectively. Hazen plotting position gave minimal errors with the GEV and EVI probability distribution for the three station, while Weibull plotting position gave minimal error with the LNG probability distribution for the three station respectively. The EVI model was found to be the most suitable distribution for modelling the daily rainfall distribution in two out of the three stations investigated, while LNG was observed to be only suitable for Gusau.

5. CONCLUSION

All the three distributions had the highest coefficient of determination using Weibull's plotting position this implies Weibull matched with the existing models more than other plotting positions, Weibull plotting position matched with EV1, LNG and GEV distribution in this order. Hazen plotting position had minimal root mean square error with GEV and EV1 distribution, while Weibull plotting position had minimal root mean square error with LNG. This indicate that using Hazen to predict statistically with GEV and EV1, while Weibull will be predicted statistically with LNG in the entire station the error will be minimal.

6. RECOMMENDATION

The need for more studies of application of Gumbel, Lognormal and GEV on other stations in the region at large is recommended to ensure its countrywide applicability. It also be recommended that for understanding of impact of extreme event more meteorological and hydrological stations should be increased across the region.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Hulme M, Osborn TJ, Johns TC. Precipitation sensitivity to global warming: Comparison of observations with HADCM2 simulations. Geophysical Research Letters. 1998;25:3379-3382.
2. Hosking JRM, Wallis JR. Regional frequency analysis. Cambridge University Press: Cambridge, UK; 1997.
3. Gumbel EJ. The return period of flood flows. Ann. Math. Statist. 1941;12(2):163-190.
4. Gumbel EJ. Statistics of extremes. Columbia University Press. New York; 1958.
5. Ekpoh IJ. The effects of climatic variability on agriculture in northern Nigeria. PhD Dissertation, University of Birmingham, Birmingham B 15 2TT, United Kingdom; 2011.
6. Ekpoh IJ. Rainfall and peasant agriculture in northern Nigeria. Global Journal of Pure and Applied Sciences. 1999a;5(1):123-128.
7. Houndenou C, Hernandez K. Climate of the Sahel 1961-1990. Benin, Secheresse. 1998;9:23-34.
8. Camberlin P, Diop M. Application of daily rainfall principal component analysis to the assessment of the rainy season characteristics in Senagal. Climate Research. 2003;23:159-169.

   DOI: 10.3354/cr023159

   Available:http://dx.doi.org/10.3354/cr023159
9. Hosking JRM. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. J. Roy. Statist. Soc. Ser B. 1990;52:105–124.
10. Hosking JRM, Wallis JR. Some statistics useful in regional frequency analysis. Water Resouces. Res. 1993;29(3):271–281.
11. Hosking JRM, Wallis JR. Regional frequency analysis: An approach based on L-moments. Cambridge University Press, Cambridge, UK; 1997.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history/27209