https://dx.doi.org/10.4314/dujopas.v10i2b.15
ISSN (Print): 2476-8316
ISSN (Online): 2635-3490

Dutse Journal of Pure and Applied Sciences (DUJOPAS), *Vol. 10 No. 2b June 2024*

Modelling Nigeria Rainfall Data Using Exponential Family of Distribution: A Comparative Study

Emmanuel S. Oguntade, Timileyin K. Babalola, Damilare M. Oladimeji *

Department of Statistics,
Faculty of Science,
University of Abuja,
Nigeria

Email: oladimejidamilare@yahoo.com

Abstract

In response to the growing concern about climate change, the study focuses on rainfall, a complex and challenging meteorological variable to predict compared to temperature. This research introduces an innovative approach to model rainfall data, leveraging probability distributions from the exponential family. The study utilizes secondary data encompassing monthly rainfall measurements in millimeters (mm) spanning from 1986 to 2020, obtained from the Nigeria Meteorological Agency (NiMET) – covering a span of thirty-five years. To identify the most suitable distribution, various model efficiency criteria, including log-likelihood, Kolmogorov-Smirnov test, Cramer-Von Mises, Anderson Darling, and Akaike Information Criterion (AIC), are employed. The analysis highlights the Weibull distribution as the most fitting model, exhibiting the highest likelihood value, along with the lowest values for the Kolmogorov-Smirnov test, Cramer-Von Mises, Anderson Darling, and AIC. For six States under consideration: Abuja, Lagos, Edo, Kebbi, Taraba, and Enugu, the study establishes the Weibull distribution as the superior choice. Notable log-likelihood and AIC values for each location reinforce its suitability for modeling rainfall data in Nigeria. Specifically, the log-likelihood and AIC values for Abuja are 1910.155 and 3824.309, respectively, and similar trends are observed across the other locations. Consequently, the study strongly recommends the adoption of the Weibull distribution as the preferred model for rainfall data modeling in Nigeria. Additionally, it encourages further research into rainfall data modeling employing the Weibull distribution within the context of time series models, both with and without seasonality, utilizing the well-established Box-Jenkins methodology.

Keywords: Climate, Rainfall, Probability, Exponential Distribution, Nigeria.

INTRODUCTION

The issue of climate change now poses a greater threat than ever to the long-term sustainability of any country's agricultural activities and the entirety of human existence (Igwenagu, 2015). As a result, the United Nations General Assembly's response in 1990 and the intergovernmental negotiating committee that developed, bargained, and ultimately accepted the United Nations Framework Convention on Climate Change (UNFCCC) on May 9th, 1992 were informed by this. Nigeria was one of the first 165 nations to sign the convention when it was made available for signature in June 1992. The agreement went into effect on March 21, 1994. The convention was ratified by Nigeria on August 29, 1994. In educational research, rainfall in Nigeria is a very important topic of study. Numerous academics have conducted excellent study on the patterns, advantages, impacts, and effects of rainfall in Nigeria. Their study work includes recommendations; however, the impact of such

Author for Correspondence

recommendations varies. In all of the studies, the most important statistical techniques for predicting and simulating rainfall in Nigeria were time series analysis and regression analysis. For instance, Adams & Bamanga (2020) use a SARIMA approach to model and forecast the monthly rainfall in Abuja, Nigeria in their work on "Modelling and Forecasting Seasonal Behaviour of Rainfall in Abuja, Nigeria". In their work, Adams & Bamanga (2020), flaws are exposed, including issues with the generalizability of their study the difficulty in obtaining the appropriate measure and model to represent the data used in the study and the study do not specify the specific distribution that can model Abuja rainfall data before carrying out their research.

In their study "Rainfall Trends and its Implication on Water Resources Management a Case Study of Ogbomosho City in Nigeria" published in 2019, Ogunbode, & Ifabiyi used Linear Regression Statistic (a parametric statistical test) and Mann Kendall (a non-parametric statistical test) to determine the rainfall trend. Without verifying the exact distribution that can model the rainfall data from Ogbomosho, Ogunbode, & Ifabiyi (2019) employ linear regression. One of the assumptions of linear regression analysis is that the set of data used in the study must follow a normal distribution.

Atedhor, (2016) investigated the rainfall trends in "Growing Season Rainfall Trends, Alterations and Drought Intensities in the Guinea Savanna Belt of Nigeria: Implications on Agriculture" using a simple linear regression model using a second order polynomial investigating which model is suitable for the rainfall data in the Guinea Savanna Belt of Nigeria.

Kumar *et al*., (2010) in their study examined long-term rainfall trends in India. Sen's estimate (Sen, 1968) was used in the study to quantify the size of the trend in the time series. To name a few, a lot of research has been done regarding rainfall in and outside of Nigeria using the study of the preceding scholars. All of these scholars didn’t specifically identify the actual probability distribution the rainfall data they used in their respective studies followed. Thus, all of the aforementioned researchers used time series models such as SARIMA, Sen's estimate and Linear Regression Model (LRM) to predict rainfall within and outside Nigeria. None of these scholars modelled rainfall data from probability process and probability distribution standpoint. Also, the data used in those researches only focused on a particular state of a country no study has ever modelled rainfall data in Nigeria by geopolitical zones.
Therefore, this research will model rainfall data in the six geopolitical zones of Nigeria using probability distribution (exponential family of distributions) approach.

METHODOLOGY

Study Design

This study was a retroactive study that centered on Modelling Nigeria Rainfall data using Exponential Distribution Family.

Data Source

The data used in this study was secondary data on monthly rainfall in millimeters (mm), sourced from the Nigeria Meteorological Agency (NiMET), Nigeria. The data covers a period of 35 years, from 1986 to 2020. The rainfall data for this research was obtained from NiMET, Abuja, Nigeria, and includes rainfall information categorized by state. Specifically, data from six states, representing six geopolitical zones in Nigeria, were collected and utilized for this research. These states are Lagos, Abuja, Kebbi, Taraba, Enugu, and Edo.

Exponential Family of Distributions

A parametric distribution family of univariate continuous distributions is said to be an exponential family if and only if the probability density function of any member of the family can be written as $f(x) = h(x)\exp[\eta(\Theta)^{T}T(x) - A(\Theta)]$. (1)

Where; h is real number of functions that depends only on x, ϴ is a K * 1 vector of parameters, η : is a vector-valued function of the vector of parameters ϴ, *T*: is a real number vector- valued function of x, $\eta(\Theta)^{T}\,T(x)$ is also the dot product between η and *T* and *A*: is a function of ϴ.

Normal Distribution with parameter μ and σ

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}} e - \frac{1}{2\sigma^2}(x-\mu)^2 & 0<x<\infty \\ 0, & \end{cases} \quad (2)$$

Normal Distribution Mean and Variance

The random variable x is said to have a normal distribution if the density function X is given

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}} e - \frac{1}{2\sigma^2}(x-\mu)^2 & \infty<x<\infty \\ 0, & \end{cases} \quad (3)$$

$$F(X)=\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e - \frac{1}{2\sigma^2}(x-\mu)^2 dx \quad (4)$$

$$k = \frac{(x-\mu)}{\delta}$$

$$E(X)=\mu \quad (5)$$

Variance of A Normal distribution

$$Var(X)=E(x-\mu)^2$$

$$Var(X)= \sigma^2 \quad (6)$$

Exponential Distributions with parameter ϴ

$$f(x) = \{\lambda\, e^{-\lambda x} \qquad X\ \varepsilon\ (0, e\ \Theta\ \theta), \quad (7)$$

$$E(x) = \frac{1}{\lambda} \quad (8)$$

$$Var(X)= \frac{1}{\lambda^2} \quad (9)$$

Gamma Distribution with Parameter λ **and** α

$$f(x) = \begin{cases} \frac{\lambda^{\alpha} x^{\alpha-1} e^{-\lambda x}}{r(\alpha\}} & 0<x<\infty \\ 0 & \end{cases} \quad (10)$$

$$E(X) = \beta\alpha \text{ (which is equivalent to } \frac{\alpha}{\lambda}) \quad (11)$$

$$Var(X)= \alpha\beta^2 \text{ or } \frac{\alpha}{\lambda^2}) \quad (12)$$

Weibull Distribution with Parameter ϴ **and β**

$$f(x/\Theta,\beta) = \frac{\Theta}{\beta} x^{\alpha-1} e^{-x^{\frac{\Theta}{\beta}}} \quad (13)$$

$$0 \le x \le \infty, \Theta > 0, \beta > 0$$

$$E(X)= \beta^{\frac{1}{\Theta}}\, \Gamma\,(1+\frac{1}{\Theta}) \quad (14)$$

$$Var(X) = \beta^{\frac{2}{\Theta}}[\Gamma\left(1+\frac{2}{\Theta}\right) - \Gamma(1+\frac{1}{\Theta}]$$

RESULTS AND DISCUSSION

Table 1 presents the descriptive statistics of the locations chosen in the six geopolitical zones in Nigeria rainfall data starting from the first location f Abuja to the sixth location Enugu. This table summarizes the mean and standard deviation of monthly rainfall data for selected locations in six geopolitical zones of Nigeria. In North Central, Abuja has a mean rainfall of 160.5943 mm with a standard deviation of 117.7307 mm. Lagos State in the South West region has a mean rainfall of 134.9171 mm and a standard deviation of 110.9603 mm. Edo State in the South-South region has the highest mean rainfall of 198.0864 mm and a standard deviation of 143.2249 mm. Kebbi in the North West region has a mean rainfall of 138.0905 mm with a standard deviation of 101.5811 mm. In the North East, Taraba has a mean rainfall of 131.818 mm and a standard deviation of 88.91717 mm. Lastly, Enugu in the South East region has a mean rainfall of 177.911 mm and a standard deviation of 121.909 mm.

Table 1. Descriptive Statistics for the Rainfall Data Based on Geopolitical Zones in Nigeria.

Geopolitical zone	Location	Mean (mm)	Median (mm)	min (mm)	Max (mm)	standard Dev	skewness	Kurtosis
North Central	Abuja	160.5943	154	0.4	554.9	117.7307	0.6002	2.9684
North West	Lagos State	134.9171	109.2	0.3	619.5	110.9603	1.2370	4.7046
South South	Edo State	198.0864	109.2	0.2	722.5	143.2249	0.6690	2.9741
North West	Kebbi	138.0905	126.4	0.4	609.9	101.5811	0.8579	4.1561
North East	Taraba	131.818	126.6	0.3	453.2	88.91717	0.4570	2.8659
South East	Enugu	177.911	188.7	0.5	508.3	121.909	0.1348	2.0069

Table 2 to 12 show the model efficient statistics and goodness of fit criteria for the best distributions for modelling rainfall data which is the distribution with the highest value of log likelihood, lowest value of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The analysis of rainfall data from multiple states, including Abuja, Lagos, Edo, Kebbi, Taraba, and Enugu, reveals a consistent pattern in terms of the preferred probability distribution for modeling the data. Across all states, the Weibull distribution consistently emerges as the optimal choice based on various goodness-of-fit criteria (see Table 2 to Table 12). In Abuja, Lagos, Edo, Kebbi, Taraba, and Enugu states, the Weibull distribution consistently exhibits the highest log-likelihood values and lowest values for both AIC and BIC among the distributions considered. For instance, in Abuja, the Weibull distribution has a log-likelihood of -1910.155, AIC of 3824.309, and BIC of 3831.814 (see Table 2). Similarly, in Lagos, the Weibull distribution shows a log-likelihood of -2267.408, AIC of 4538.816, and BIC of 4546.723 (see Table 4). This consistent performance across diverse geographical regions underscores the reliability and versatility of the Weibull distribution in modeling rainfall patterns. The Weibull distribution's superiority in fitting the data suggests its robustness in capturing the variability of rainfall across different states.

Table 2: Model Efficient Statistics and Goodness of Fit Criteria of the Distributions Fitted for Abuja Rainfall Data.

Model	Exponential	Normal	Gamma	Weibull
Loglikelihood	-1914.848	-1948.511	-1913.778	-1910.155
Kolmogorov Smirnov	0.1407389	0.08646505	0.1300674	0.1114714
Cramer-Von Mises	1.477027	0.395646	1.238623	0.954773
Anderson-Darling	7.231102	3.291786	6.443831	5.641013
AIC	3831.695	3901.022	3831.555	3824.309
BIC	3835.448	3839.06	3824.309	3831.814

Table 3: Two Parameter Weibull Value for the Shape and the Scale Parameters

Parameters	Weibull	Std Error
Shape	1.162166	0.05523941
Scale	167.960409	8.49394436

$$f(x) = \frac{1.162166}{167.960409}\frac{x}{167.960409}^{1.162166-1} e^{-(\frac{x}{167.960409})^{1.162166}} \quad \text{where; } x \geq 0$$

Table 4: Model Efficient Statistics and Goodness of Fit Criteria of the Distribution Fitted to Lagos State Rainfall Data.

Model	Exponential	Normal	Gamma	Weibull
Loglikelihood	-2273.295	-2358.822	-2269.532	-2267.408
Kolmogorov Smirnov	0.07198007	0.1122246	0.04502291	0.03996904
Cramer-Von Mises	0.61833905	1.4167711	0.20575817	0.10543959
Anderson-Darling	3.13555391	9.1108642	1.19024184	0.73897545
AIC	4548.59	4721.644	4543.064	4538.816
BIC	4552.544	4729.551	4550.971	4546.723

Table 5: Two Parameter Weibull Value for the Shape and the Scale Parameters

Parameters	Weibull	Std Error
Shape	1.155808	0.04716511
Scale	141.638939	6.55614563

$f(x) = (\frac{1.155808}{141.638939})(\frac{x}{167.960409})^{1.155808-1}e^{-(\frac{x}{167.960409})^{1.155808}}$ where; x ≥ 0

Table 6: Model Efficient Statistics and Goodness of Fit Criteria of the Distribution Fitted to Edo State Rainfall Data.

Model	Exponential	Normal	Gamma	Weibull
Loglikelihood	-2421.15	-2457.091	-2415.629	-2409.45
Kolmogorov Smirnov	0.1174292	0.08326448	0.08908047	0.0657639
Cramer-Von Mises	1.5307960	0.7000325	0.77760802	0.4619229
Anderson-Darling	7.8269495	4.69618930	4.55091277	3.2716735
AIC	4844.3	4918.182	4835.259	4822.9
BIC	4848.253	4926.088	4843.165	4830.806

Table 7: Two Parameter Weibull Value for the Shape and the Scale Parameters

	Parameters	Std Error
Shape	1.238865	0.05234087
Scale	210.639356	9.05357199

$f(x) = (\frac{1.238865}{210.639356})(\frac{x}{210.639356})^{1.238865-1}e^{-(\frac{x}{210.639356})^{1.238865}}$ where; x ≥ 0

Table 8: Model Efficient Statistics and Goodness Fit Criteria of the Distribution Fitted to Kebbi State Rainfall Data Set

Model	Exponential	Normal	Gamma	Weibull
Loglikelihood	-1564.968	-1594.005	-1563.788	-1560.275
Kolmogorov Smirnov	0.1157613	0.08722421	0.1049327	0.09000855
Cramer-Von Mises	1.1591785	0.38128401	0.8406151	0.53431175
Anderson-Darling	6.1206083	2.82175453	4.9115413	3.73925644
AIC	3131.936	3192.011	3131.575	3124.550
BIC	3135.512	3199.162	3138.727	3131.702

Table 9: Two Parameter Weibull Value for the Shape and the Scale Parameters

	Parameters	Std Error
Shape	1.176983	0.06046458
Scale	144.805658	7.88852230

$f(x) = (\frac{1.176983}{144.805658})(\frac{x}{144.805658})^{1.176983-1}e^{-(\frac{x}{144.805658})^{1.176983}}$ where; x ≥ 0

Table 10: Model Efficient Statistics and Goodness Fit Criteria of the Distribution fitted to Taraba State Rainfall Data.

Model	Exponential	Normal	Gamma	Weibull
Loglikelihood	-1676.205	-1682.893	-1673.046	-1666.5
Kolmogorov Smirnov	0.1120907	0.06920897	0.1298363	0.09000855
Cramer-Von Mises	0.9535903	0.22834201	1.4507011	0.53431175
Anderson-Darling	6.1626200	1.93863537	7.8938970	3.73925644
AIC	3354.410	3369.785	3350.092	3337.000
BIC	3358.063	3377.090	3357.397	3344.305

Table 11: Two Parameter Weibull Value for the Shape and the Scale Parameters

	Parameters	Std Error
Shape	1.262366	0.05234087
Scale	140.126882	9.05357199

$$f(x) = \left(\frac{1.262366}{140.126882}\right)\left(\frac{x}{140.126882}\right)^{1.262366\ -1} e^{-\left(\frac{x}{140.126882}\right)^{1.262366}} \quad \text{where; } x \geq 0$$

Table 12: Model Efficient Statistics and Goodness Fit Criteria of the Distribution fitted to Enugu State Rainfall Data.

Model	Exponential	Normal	Gamma	Weibull
Loglikelihood	-2163.45	-2177.274	-2163.346	-2159.649
Kolmogorov Smirnov	0.168038	0.09312194	0.1662062	0.153488
Cramer-Von Mises	2.657073	0.59669658	2.5749149	2.140743
Anderson-Darling	13.932793	4.50276104	13.7709310	13.130721
AIC	4328.9	4358.548	4330.692	4323.299
BIC	4332.758	4338.408	4338.408	4331.015

Table 13: Two Parameter Weibull Value for the Shape and the Scale Parameters

	Parameters	Std Error
Shape	1.14118	0.05324341
Scale	184.64157	8.98153889

$$f(x) = \left(\frac{1.14118}{184.64157}\right)\left(\frac{x}{184.64157}\right)^{1.14118\ -1} e^{-\left(\frac{x}{184.64157}\right)^{1.14118}} \quad \text{where; } x \geq 0$$

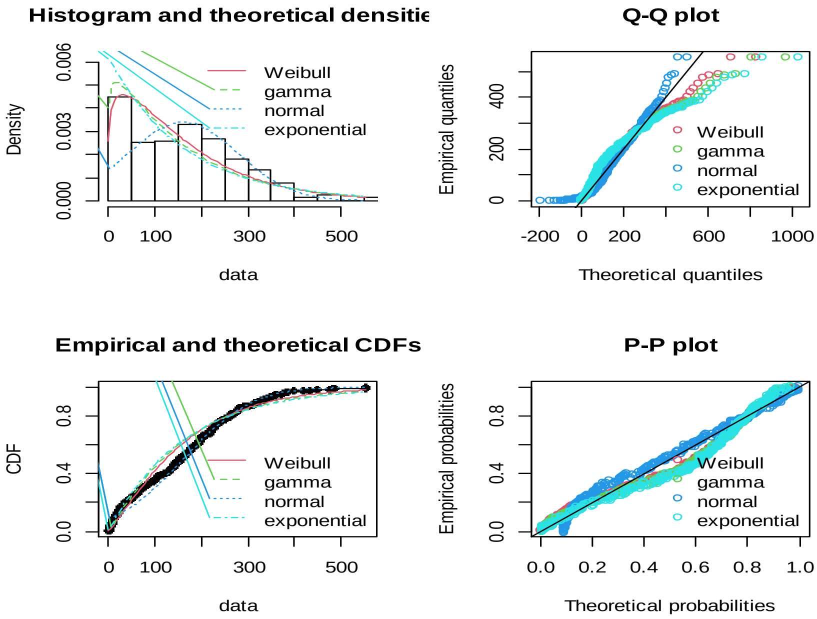

Figure 1: Four Goodness-of-Fit Plots for the Four Distributions fitted to Abuja Rainfall Data.

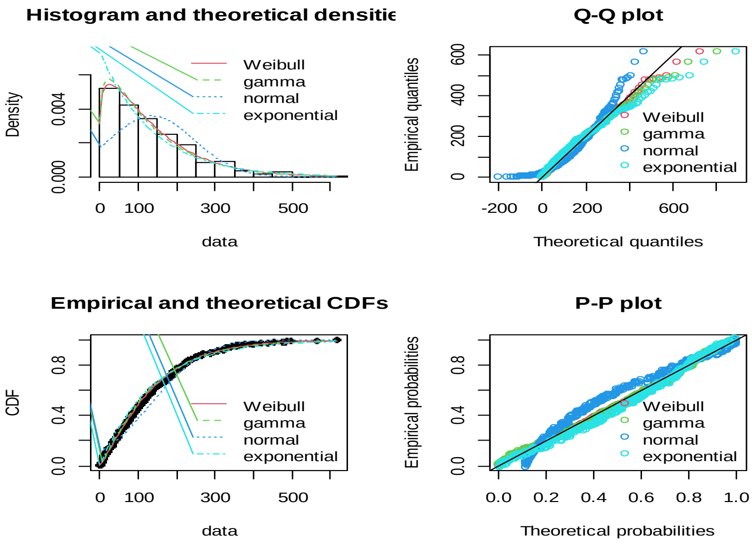

Figure 2: Four Goodness-of-Fit Plots for the Four Distributions Fitted to Lagos State Rainfall Data.

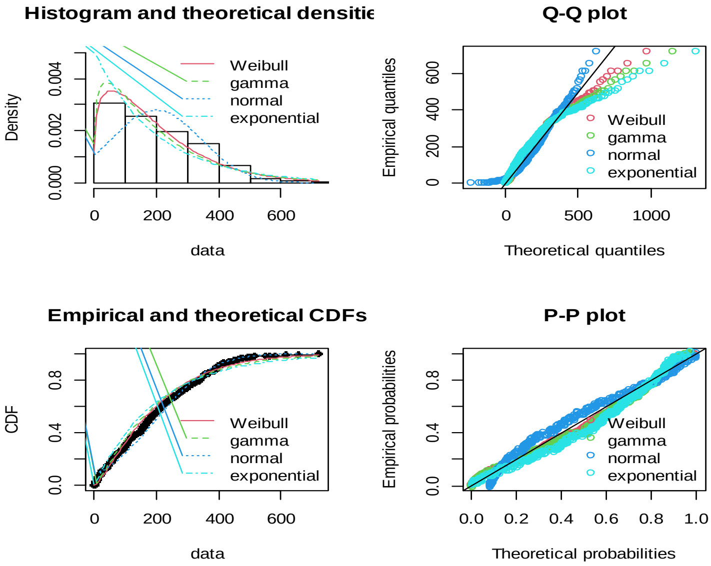

Figure 3: Four Goodness-of-Fit Plots for the Four Distributions Fitted to Edo State Rainfall Data.

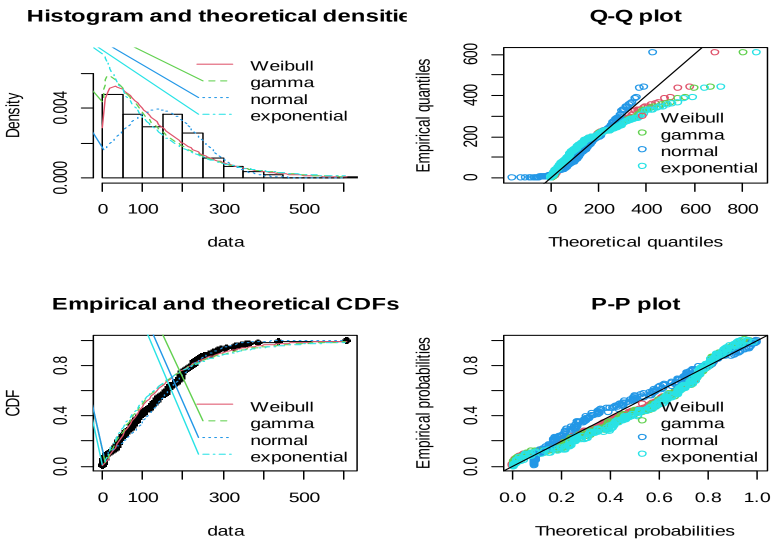

Figure 4: Four Goodness-Of-Fit Plots for the Four Distributions fitted to Kebbi State Rainfall Data.

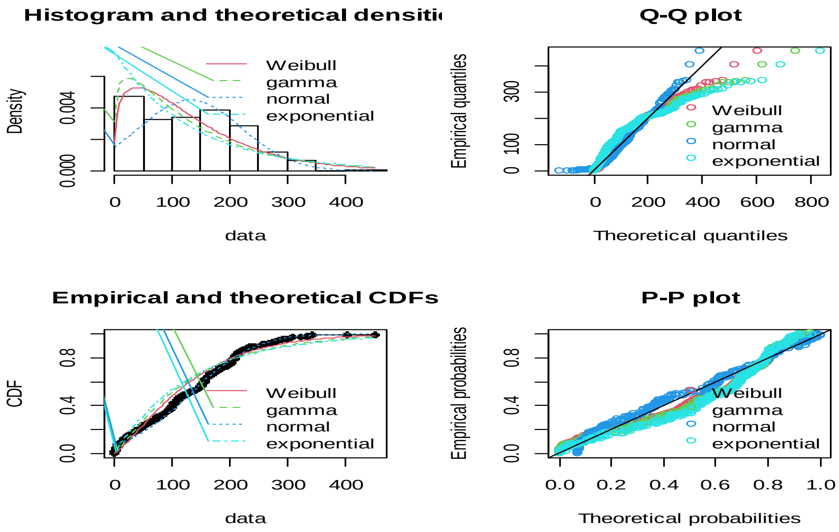

Figure 5: Four Goodness-Of-Fit Plots for the Four Distributions fitted to Taraba State Rainfall Data.

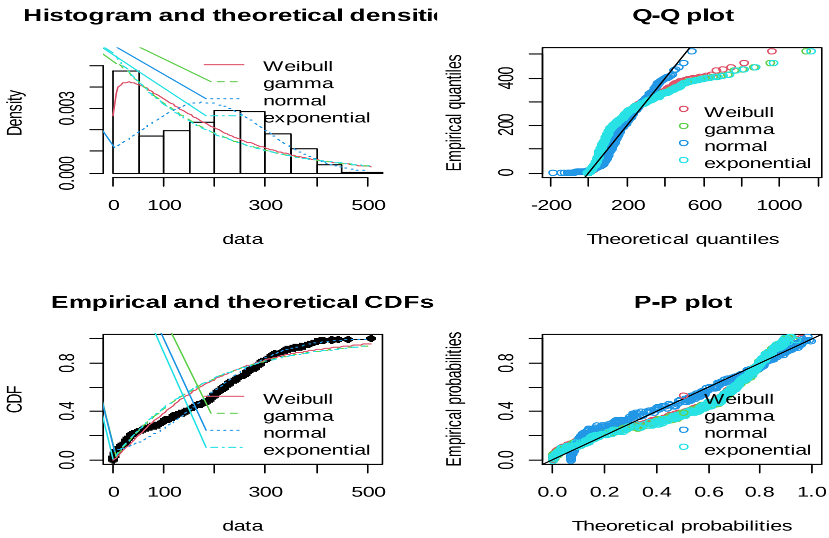

Figure 6: Four Goodness-of-Fit Plots for the Four Distributions fitted to Enugu State Rainfall Data.

DISCUSSION OF FINDINGS

This study was designed to model the rainfall data in the six geopolitical zones in Nigeria using the suitable probability distribution that is very good for modelling the rainfall data. The data used in the study is a secondary data obtained from Nigeria Meteorological Agency (NiMET). The data covers 35 years spanning from 1986 to 2020. The data has been analyzed using the Fit Distribution in which four distributions were fitted to know which distribution is best for modelling rainfall data in Nigeria in the location chosen in the six geopolitical zones in Nigeria. The four-distribution fitted to the data were the Exponential distribution, the Normal distribution, the Gamma distribution, and the Weibull distribution. The four distributions used in the study all belong to the exponential family of distributions with light tailed (Oladimeji et al., 2023). The essence was to fit an appropriate distribution to model the rainfall data in the location chosen in the six geopolitical zones in Nigeria.

The study revealed that Weibull distribution is the best distribution that can model the rainfall data in the six geopolitical zones in Nigeria with the maximum value of log likelihood, the Kolmogorov Smirnov test, Cramer-Von Mises, and Anderson Darling which depicts the data follows the Weibull distribution for all the locations. This current result agrees that Weibull distribution as the best distribution for modelling as reported in Jalgaon and Coimbatore divisions in India ((Muralidharan & Lathika, 2022; Ilori et al., 2021).

Also, the result based on Akaike Information Criterion reported that Weibull distribution is the distribution with the minimum value out of the four-distribution fitted. Hence, Weibull distribution is the best for modelling the rainfall data in the locations. The result based on Akaike Information Criterion is in line with Adams and Bamanga (2020) the model efficiency of the four selected distributions were judged on the basis of their information criterion statistics and the AIC reported that Weibull distribution is the distribution with the minimum value out of the four-distribution fitted. Hence, Weibull distribution is the best for modelling the rainfall data in the locations.

The results of this current study based on the five-model efficiency used revealed that Weibull distribution is the best and appropriate distribution that is reliable to be used in modelling the rainfall data in the six geopolitical zones in Nigeria.

CONCLUSION

Having applied the fit distribution using the exponential family of distributions in modelling rainfall data in the six geo-political zones in Nigeria and with the four-distribution selected which are Exponential distribution, Normal distribution, Gamma distribution, and Weibull distribution. The analysis reported with the result derived from all the model efficiency statistics used that Weibull distribution is the best and reliable distribution that should be used in modelling the rainfall data in Nigeria. Therefore, Weibull distribution should be used in modelling rainfall data in Nigeria.

REFERENCE

Adams, S. O., & Bamanga, M. A. (2020). Modelling and Forecasting Seasonal Behavior of Rainfall in Abuja, Nigeria; A SARIMA Approach. *American Journal of Mathematics and Statistics*, 10(1), 10-19. DOI: 10.5923/j.ajms.20201001.02

Atedhor, G. O. (2016). Growing Season Rainfall Trends, Alterations and Drought Intensities in the Guinea Savanna Belt of Nigeria: Implications on Agriculture. *Journal of Environment and Earth Science*, 6(3)

Kumar, V., Jain, S. K., & Singh, Y. (2010). Analysis of long-term rainfall trends in India. *Hydrological Sciences Journal – Journal des Sciences Hydrologiques*, 55(4), 484-496. DOI: 10.1080/02626667.2010.481373.

Igwenagu, C. M. (2015). Trend Analysis of Rainfall Pattern in Enugu State, Nigeria. *International Journal of Scientific & Engineering Research*, 5(10), 1317.

Ilori, A. K., Okafor, F. M., Matthew, O. D., & Kayode, D. V. (2021). Mathematical Study of Weibull – New Weighted Exponential Distribution, 1, 164–168.

Muralidharan, K., & Lathika, P. (2022). Statistical modelling of rainfall data using modified Weibull distribution. Mausam, 56(4), 765–770. https://doi.org/10.54302/mausam.v56i4.1026

Ogunbode, T. O., & Ifabiyi, P. I. (2019). Rainfall trends and its implications on water resources management: A case study of Ogbomosho city in Nigeria. *International Journal of Hydrology*. DOI: 10.15406/ijh.2019.03.00182

Oladimeji, D. M., Oguntade, E. S., & Olarenwaju, S. O. (2023). Heavy Tailed Distribution of Binary Classification Model. *Journal of Sciences*, 34(3), 239– 246. https://doi.org/10.22059/jsciences.2023.363551.1007821