© Universiti Tun Hussein Onn Malaysia Publisher's Office

JST

Journal of Science and Technology

Journal homepage: http://penerbit.uthm.edu.my/ojs/index.php/jst

ISSN : 2229-8460 e-ISSN : 2600-7924

# On the Generalized Pareto Probability Density Function of the Mean Annual Rainfall of Katsina State

**Dauda Taofik O[1*], Adetayo Adewale O[1]**

[1]Institute of Agricultural Research and Training,
Obafemi Awolowo University, PMB 5029, Moor Plantation, Ibadan, NIGERIA

*Corresponding Author

DOI: https://doi.org/10.30880/jst.2021.13.02.002
Received 17 May 2021; Accepted 20 September 2021; Available online 2 December 2021

**Abstract:** This work was accomplished to assess the rainfall probability density function (pdf) of Katsina State and adopt the most parsimonious density function using 2003 – 2018 rainfall data from the Nigeria Institute of Meteorological Services. The data were subjected to cleaning, descriptive statistics and investigation of five distribution functions (General pareto, general extreme value, Gumbel maximum, Uniform and Normal distribution) through parameterizations. The results intimated that Daura experienced the highest mean annual rainfall ranging between 75.667 mm for 2003 and 82.417 mm for 2018. Both Dutsin-Ma and Safana local Government returned the least mean annual rainfall alternately and it ranged between 57.500 mm (2009 rainfall for Dutsin-Ma) and 67.167 mm for 2018 mean annual rainfall of Safana. The order of performance of the five (5) prominent pdf in each of the sites differs but the Generalized Pareto distribution (GPD) ranked highest across all the sites. The distribution characteristics for the GPD (the most parsimonious of the pdf) are, mean = 61.15098 (Duara), 58.93371 (Dutsinma) and 79.20851 (Safana) and variance of 9212.603 (Duara), 6709.908 (Dutsinma) and 6687.102 (Safana). The examination of the visual analysis of the residuals of the mean annual rainfall indicated that only the mean rainfall for Dutsin-Ma gave the most precise parameter estimates.

**Keywords:** Exceedance, immanent, intimate, parsimonious, fluctuation

## 1. Introduction

Different rainfall parameters (rainfall peak, ceasation and distribution) are important for agricultural planning and activities. It (rainfall) has been described as the best indicator of climate change [1] as well as an important environmental factor for agricultural success in Nigeria, particularly Katsina where agriculture depends largely on rainfall despite the irrigation practices. Meanwhile, rainfall prediction is said to be arduous due to climatic change and instability in the rainfall pattern [2]. The work of [3], [4] and [5] introduced the Generalized Pareto distribution. The Pareto distribution, a well-known statistical model found its applications in areas such as insurance, reliability analysis, finance, meteorology and environment, among others [6]. Its extension is the Generalized Pareto Distribution (GPD – [7]). In the work of Pkand [3] estimates are based on the k-top order statistics of the sample or the excesses over a high random threshold, $\mu$. Similarly, parametric statistical approach and estimation was adopted by [3] for modeling extreme values. Many works have been conducted on rainfall distribution in the northern region of Nigeria. These include but not limited to, an assessment of the rainfall and temperature variations in parts of Northern Nigeria [2], Best–fit Probability distribution model for peak daily rainfall of selected Cities in Nigeria [8] and Use of Probability Distribution in Rainfall Analysis [9]. Pikands [3] found out that the success of the transformed estimator will depend on how well the initial estimator is and in small samples for $\mathcal{E} \leq 0.5$.

This work derives its motivation from the work of [9] as well as [10]. About sixteen (16) probability distribution function were applied to their data to determine the best fit using the goodness of fit statistics (Kolmogorov-Smirnov,

*Corresponding author: todauda.iart@gmail.com

Anderson-Darling and Chi square statistics, These works notwithstanding, in view of the necessity for optimizing planning and preparations for extreme future rainfalls as articulated by [11] and recommendation of probability rainfall distribution as a better approach to rainfall distribution study [12], this work is therefore desirous. It was established by [2] that specific model building should be enabled upon in specific regions like Northern Nigeria for a better fitting of models to realities on the ground. Rainfall distribution is a veritable tool for profitable agriculture because it provides a guide for farming activities schedule as well as play a key role in planning defense for climate change effects. The goal of establishing rainfall probability distribution function has been to provide a robust and parsimonious estimation model thereby reducing the risk associated with crop failure due to wrong prediction. This work was therefore carried out to assess the rainfall probability density function of Katsina state and adopt the most parsimonious density function for the study area.

## 2. Methodology

**The Study area -** The study area is Katsina State of Nigeria $(12^0 59'26.95''N; 7^0 36'6.37''E)$ and located some 260 km from Sokoto State and 135 km from Kano State (Fig. 1). It has an estimated population of 5,801,584 people (by 2006 population census) and is majorly agrarian populace specializing in cotton, groundnut, guinea corn and millet cultivation. The main ethnic groups in Katsina state are Hausa and Fulani while other ethnic groups (Yoruba, Ebira, Igbo and Nupe) from other parts of Nigeria can also be found in minority. The specific locations whose data were collected for this study are Daura, Dutsin-Ma and Safana local government and represent 9% $(3/33)$ of the state. The areas were randomly selected and are representative of the state. The state has both savannah and arid zone climate with rainfall lasting up to 5 months and temperature ranging between 21-30$^0$C. The monthly rainfall data (2003 – 2018) for the selected local governments were obtained from the Nigerian Meteorological Agency (NIMET), Lagos, Nigeria. NIMET has the mandate for the measurement, control, and storage of rainfall data of the areas in Nigeria. The mean average of the rainfall were compared with rainfall of 2018.

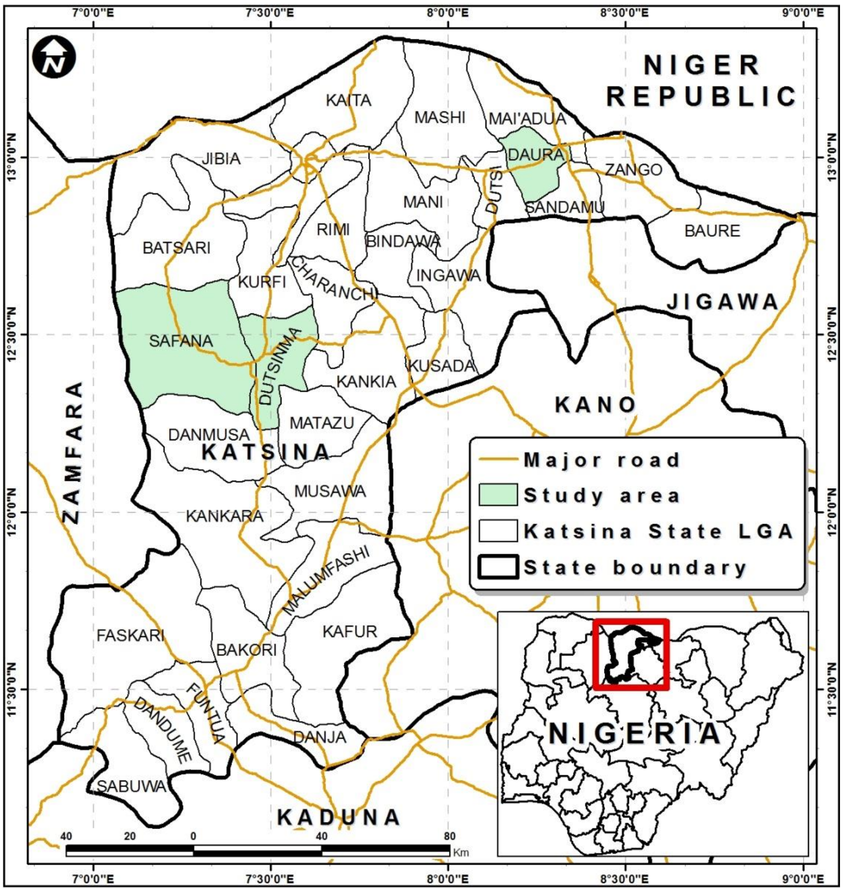


**Fig. 1 - Map of Katsina State Showing The Location Of The Study Area (Inset – Map of Nigeria)**

The data obtained from the NIMET were subjected to cleaning (by determination and elimination of outliers) and descriptive statistics including mean and variance of the rainfall for each of the sites. Outliers can be determined by subjecting the data to visual analysis (through graphing). The influence of outliere has been said to be insignificant importance in large sample size [13] hence the elimination of the outlier found. These variables (mean and variance) were computed across the years and months. Five distribution functions (General pareto, general extreme value, Gumbel maximum, Uniform distribution and Normal distribution) were investigated through parameterizations. Parameterizations have been adopted for deep convection using assumed pdf method [14]. The pdf parameterizations were eventually used to predict the pdf of subgrid variability of turbulence, clouds and hydrometeoris. Parametrization method has been adopted in our present study to establish the pdf form that produces more normal and less skewed distribution of the parameter values that occur using that form. This is in consonant with [15] and the distribution characteristics of the various pdf were computed. The most suitable pdf was determine based on three (3) different (Kolmogorov Smirnov, Anderson Darling and Chi-square) metric measures. The Kolmogorov Smirnov, Anderson Darling and Chi-square which measure the suitability of the distribution on the basis of maximum difference between an empirical and hypothetical cumulative distribution and have been recommended and adopted in the past studies [16, 17,18 and 19], The parameters of GPD distribution can be estimated by using maximum likelihood method which involves maximizing with respect to (wrt) parameters ξ, μ and σ.

Generalized Pareto distribution is defined as the random variable X, if its cumulative distribution is given by;

$$F(x;k,\sigma) = \begin{cases} 1-\left[1-\frac{kx}{\sigma}\right]^{1/k} k \neq 0, \sigma > 0 \\ 1-\ell^{-\frac{x}{\sigma}} k = 0, \sigma > 0 \end{cases} \quad (1)$$

where σ = scale parameter and k = position/shape parameter, $x > 0 \ \ for \ \ k \leq 0$. $x$ = form parameter.

$$x \in \left[0, \frac{\sigma}{k}\right] \ \ for \ \ k > 0$$

The random value x has the generalized pareto distribution. That is $X \sim GPD(\xi, \mu, \ \sigma)$. The sample observation of size n = 192 were available for our present study and the distribution characteristics (mean and variance –equation 2 and 3) were determined for each sites as follows;

The mean $E(X) = \mu + \frac{\sigma}{1-\varepsilon} \qquad r \wedge \ell, \qquad \varepsilon < 1$ (2)

$$Variance = \frac{\sigma^2}{(1-\varepsilon)^2(1-2\varepsilon)} \qquad r \wedge \ell, \qquad \varepsilon < 1/2 \quad (3)$$

where the Ɛ and σ are the mean and variance components. The residuals of the estimated values of the pdft for each of the sites were plotted to determine the randomness of the estimator Easy fit (version 5) and SAS (version 9) were used for the data analysis.

## 3. Results

### 3.1 The Descriptive Statistics of the rainfall

The results of the descriptive analysis indicated that Daura experienced the highest annual mean rainfall ranging between 75.667 mm for 2003 and 82.417 mm for 2018. Both Dutsin-Ma and Safana local Government returned the least mean annual rainfall alternately (Table 1). It ranged between 57.500 mm (2009 rainfall for Dutsin-Ma) and 67.167 mm for 2018 mean annual rainfall of Safana. The rainfall trend of each of the sites followed non-coherent but an increasing trend (Fig. 2) and the Safana rainfalls have the highest variability. It ranged between 6829.90 (for 2009) and 8499.76 for 2015 mean annual rainfall. The variability of the rainfall for those of Dustin-Ma was the least and the variability for the mean annual rainfall for all the sites were generally high (Table 1). The time plot of the variability of the mean annual rainfall showed that no two sites of the study areas returned similar trends though all the variability followed cyclical non-regular trends. Also, it is hard to discern which of the sites enjoined the highest variability because it (variance) fluctuates in each of the sites (Fig. 1).

Mean monthly rainfall for the 3 selected areas were markedly different from one another. The mean monthly rainfalls for Daura were the highest and ranged from 0.025 mm for February to 221.875 mm for August excluding both November and December which experienced 0mm mean monthly rainfall (Table 2). Dutsin-Ma enjoined the least (of the 3 sites) mean monthly rainfall which falls between 2.750 mm for April and 190.188 mm for September. Highest monthly rainfall variability was obtained for rainfall data for Dutsin-Ma and ranged between 4.60 (April) and 3847.90 (for June – Table 2). Daura experienced the least rainfall variability which falls between 0.010 for January and 163.183 for August.

## 3.2 The Probability Density Function (pdf)

The order of performance of the five prominent pdf in each of the sites differs but the Generalized Pareto distribution (GPD) ranked highest across all the sites (Table 3). The GPD for Daura has the function parameters of $k$ = 0.187 (position parameter), $\sigma$ = 61.779 (scale parameter) and $\mu$ = 14.811(form parameter) while Dutsin-Ma has $k$ = 0.05654, $\sigma$ = 72.782 and $\mu$ = -18.21. The GDP parameters for Safana are $k$ = -0.325, $\sigma$ = 139.18 and $\mu$ = -25.833 (Table 3). Other probability density functions for Daura include, generalized extreme value (GEV) with the parameters – $k$ = 0.35814 $\sigma$=38.027 and $\mu$=18.639, Gumbel Maximum with the parameters - $\sigma$=64.777 $\mu$= 23.771 as well as Uniform distribution having the parameters - $a$=-82.737 and $b$=205.06. The 5th probability density function for Daura is the Normal distribution with the parameters - $\sigma$=83.08 and $\mu$ = 61.161 (Table 3).

GEV for Dutsin-Ma has the parameters - $k$=0.27445 $\sigma$=41.284 $\mu$=19.919, Gumbel Maximum has the parameters- $\sigma$=62.815 $\mu$=22.676, Uniform distribution has the parameters - $a$=-80.606 and b=198.47 and Normal distribution has parameters - $\sigma$=80.563 and $\mu$=58.934. The parameter for GEV for Safana are - $k$=0.05082 $\sigma$=62.05 and $\mu$=40.117, for Gumbel Maximum they are both $\sigma$=64.023 and $\mu$=42.251 and for Uniform distribution, they are $a$=-63.018 $b$=221.43 (Table 3). Normal distribution for the mean annual rainfall for the Safana has parameters - $\sigma$=82.113 and $\mu$=79.206. From these results each of the parameters has different values for the sites and the distributions have different numbers of parameters. Some of them have just scale and shape parameters while others have form parameter in addition (Table 3).

**Table 1 - Descriptive statistics of annual rainfall of the Katsina state**

| **Years** | **Daura** | | **Dutsin-Ma** | | **Safana** | |
|---|---|---|---|---|---|---|
| | Mean ± SE | Variance | Mean ± SE | Variance | Mean ± SE | Variance |
| **2003** | 75.667± 23.752 | 6770.06 | 59.333± 24.066 | 6950.06 | 59.417 ± 24.234 | 7047.17 |
| **2004** | 76.833± 23.688 | 6733.24 | 60.583 ±25.067 | 7540.27 | 57.750 ± 24.248 | 7055.66 |
| **2005** | 76.708± 23.549 | 6654.74 | 58.083 ± 23.716 | 6749.36 | 59.000 ± 24.867 | 7420.18 |
| **2006** | 79.167± 24.635 | 7282.33 | 58.4167 ± 24.001 | 6912.63 | 59.9167± 24.678 | 7307.90 |
| **2007** | 80.667± 25.153 | 7592.24 | 57.500 ± 24.011 | 6918.45 | 59.833 ± 24.563 | 7239.97 |
| **2008** | 79.167± 24.744 | 7347.24 | 60.500 ± 24.696 | 7318.45 | 59.333 ± 24.514 | 7211.15 |
| **2009** | 78.767± 24.327 | 7101.71 | 60.917 ± 24.816 | 7389.90 | 58.917 ± 23.857 | 6829.90 |
| **2010** | 81.317± 24.968 | 7480.78 | 58.083 ± 24.419 | 7155.36 | 61.500 ± 25.396 | 7739.18 |
| **2011** | 79.667± 24.239 | 7050.24 | 58.750 ± 24.083 | 6960.02 | 60.833 ± 25.110 | 7566.15 |
| **2012** | 80.617 ±24.907 | 7444.60 | 58.583 ± 24.321 | 7097.90 | 60.917 ± 25.462 | 7779.72 |
| **2013** | 81.167 ±24.957 | 7474.15 | 59.333 ± 24.066 | 6950.06 | 59.167 ± 24.279 | 7073.79 |
| **2014** | 82.083± 25.531 | 7822.27 | 60.083 ± 24.113 | 6977.17 | 60.333 ± 24.629 | 7279.15 |
| **2015** | 75.750 ±23.825 | 6811.48 | 59.667 ± 24.629 | 7278.97 | 65.167 ± 26.614 | 8499.79 |
| **2016** | 77.600± 24.930 | 7457.85 | 61.250 ± 24.352 | 7116.39 | 64.167 ± 25.181 | 7609.06 |
| **2017** | 79.708 ±25.450 | 7772.38 | 64.727 ±25.692 | 7261.02 | 65.167 ± 25.546 | 7830.88 |
| **2018** | 82.417± 26.145 | 8202.99 | 60.167 ± 25.154 | 7592.88 | 67.167 ± 26.196 | 8235.06 |

**Table 2 - Descriptive statistics of monthly rainfall of the Katsina state**

| Months | Daura | | Dutsin-Ma | | Safana | |
|---|---|---|---|---|---|---|
| | Mean ± SE | Variance | Mean ± SE | Variance | Mean ± SE | Variance |
| **January** | 0.125 ± 0.066 | 0.070 | 0 | 0 | 0 | 0 |
| **February** | 0.025 ± 0.025 | 0.010 | 0 | 0 | 0 | 0 |
| **March** | 0.325 ± 0.161 | 0.417 | 0 | 0 | 1.125 ± 0.272 | 1.183 |
| **April** | 57.250 ± 1.512 | 36.60 | 2.750 ± 0.536 | 4.60 | 9.688 ± 0.546 | 4.763 |
| **May** | 117.438 ± 2.604 | 108.529 | 12.563 ± 1.720 | 47.329 | 51.500 ± 2.351 | 88.400 |
| **June** | 143.000 ± 1.420 | 32.267 | 157.813 ± 15.508 | 3847.90 | 115.500 ± 3.699 | 218.933 |
| **July** | 178.250 ± 1.010 | 16.333 | 146.875 ± 14.047 | 3157.18 | 189.875 ± 2.498 | 99.850 |
| **August** | 221.875 ± 3.194 | 163.183 | 134.563 ± 11.742 | 2206.00 | 251.500± 1.906 | 58.133 |
| **September** | 186.438 ± 1.049 | 17.596 | 190.188 ± 1.860 | 55.363 | 106.500 ± 1.557 | 38.800 |
| **October** | 45.750 ± 1.815 | 52.733 | 26.188 ± 14.495 | 3361.63 | 8.250 ± 0.479 | 3.667 |
| **November** | 0 | 0 | 18.563 ± 12.681 | 2573.06 | 0 | 0 |
| **December** | 0 | 0 | 23.438 ± 16.016 | 4104.26 | 0 | 0 |

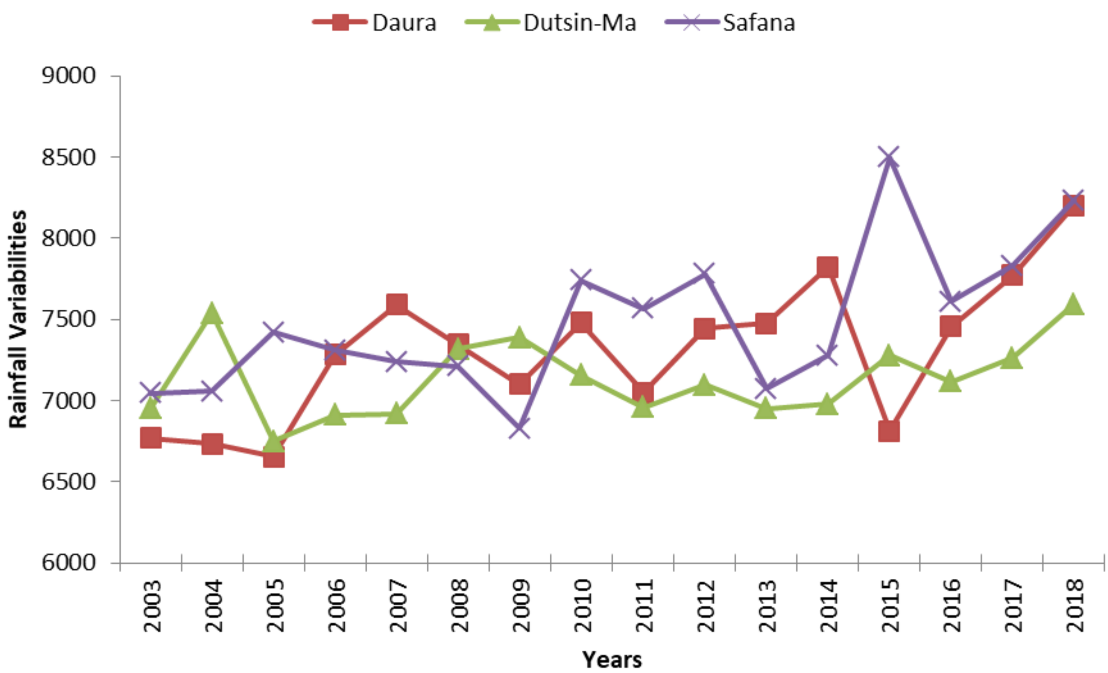


**Fig. 2 - Time plots of annual rainfall variability of the 3 selected LG of Katsina state**

**Table 3 - Five (5) best rainfall probability density functions' parameters**

| s/n | Pdf | Daura | Dutsin-Ma | Safana |
|---|---|---|---|---|
| **1** | General Pareto | $k$=0.18683 $\sigma$=61.779 $\mu$= -14.811 | $k$=0.05654 $\sigma$=72.782 $\mu$=-18.21 | $k$=-0.32501 $\sigma$=139.18 $\mu$= -25.833 |
| **2** | General Extreme Value | $k$=0.35814 $\sigma$=38.027 $\mu$=18.639 | $k$=0.27445 $\sigma$=41.284 $\mu$=19.919 | $k$=0.05082 $\sigma$=62.05 $\mu$=40.117 |
| **3** | Gumbel Maximum | $\sigma$=64.777 $\mu$=23.771 | $\sigma$=62.815 $\mu$=22.676 | $\sigma$=64.023 $\mu$=42.251 |
| **4** | Uniform Distribution | $a$=-82.737 $b$=205.06 | $a$=-80.606 $b$=198.47 | $a$=-63.018 $b$=221.43 |
| **5** | Normal Distribution | $\sigma$=83.08 $\mu$=61.161 | $\sigma$=80.563 $\mu$=58.934 | $\sigma$=82.113 $\mu$=79.206 |

Note: $\sigma$ and $k$ are the scale and shape parameters while $\mu$ is form parameter.

The ranking of the distribution by Kolmogorov Smirnov, Anderson Darling and Chi square statistics were used to adjudged the best pdf. GDP ranked highest for both Daura and Dutsin-Ma with Kolmogorov Smirnov statistics of both 0.237 and 0.282 (Table 4). For Safana however, the Kolmogorov Smirnov statistics ranked third while it ranked first using Anderson Darling statistics of 10.837. The Kolmogorov Smirnov statistics for other pdf obtained for mean annual rainfall of Daura are 0.27619 (GEV), 0.27671 (Gumbel Maximum), 0.28748 (Uniform distribution and 0.30111(Normal distribution – Table 4). For other pdf obtained for Dutsin-Ma, their Kolmogorov Smirnov statistics are 0.32295(GEV), 0.31508 (Gumbel Maximum), 0.29727 (Uniform distribution and 0.34536 (Normal distribution). The pdf for the mean annual rainfall for Safana has Kolmogorov Smirnov statistics of 0.26201 (for GEV), 0.26334 (for Gumbel Maximum), 0.22154 (for Uniform distribution and 0. 243312 (for Normal distribution – Table 4). Judging by the performance statistics of the pdf, it can be noticed that the 3 test statistics ranked the distribution differently.

The distribution characteristics for the GPD (the most parsimonious of the pdf) are, mean = 61.15098 (Duara), 58.93371 (Dutsinma) and 79.20851 (Safana – Table 5) and variance of 9212.603 (Duara), 6709.908 (Dutsinma) and 6687.102 (Safana). The examination of the visual analysis of the residuals of the mean annual rainfall (Figure 3) indicated that only the mean rainfall for Dutsin-Ma gave the most efficient parameter estimates (for Mean) with the sum of residual of 13.036. The sum of square of the residual annual mean rainfall of other 2 areas (Daura and Safana) were 288.886 and -288.751 for both Daura and Safana (Fig. 3). The results imply that while the mean annual rainfalls of Daura were exaggerated, the mean annual rainfalls of Safana were understated. The difference between the GPD of the three study sites were based on the value of the shape parameter. The key feature of these GPD was that increasing shape parameter (k) for both Daura and Dustin-Ma bestow increasing values on the GPD while the declining shape parameter of the Safana bestows increasing GPD values.

## 4. Discussion

The rainfall trends were similar to that of locations in the Sahara region [20] but differ from both the declining trends obtained for Northern Nigeria [2] and the zig-zag trends (a very wide range of flunctuation) obtained for the entire Nigeria [21]. The disparity may however be due to the scope of the research in each of the study. The latter [21] covered the whole Country (Nigeria) while the present study covers just a state of the Federation. Similarly, the rainfall pattern of Abuja, a site within the same region as this study site (Katsina) unlike the present study gave a declining trend over the 31 years [22]. The high variability for the mean annual rainfall established in our study have been the peculiarity of the Nigerian rainfall since the early 19th century [23]. This variability according to [23] were still evident when the rainfall periods were partitioned into rainfall retreat and rainy season. Similarly, high rainfall variability was established for both time and space by [19]. The high variability of the mean annual rainfall established in our present study may be attached to the aggregate monthly variability of the rainfall. This conforms with [21] which established disparity in the variability of rainfall of different decades. The differences in the GPD of the different sites as bestowed by the shape parameter (k) established in our present study confirm with properties of Generalized Extreme value [10]. By this, the GPD can be adjudged as a parsimonious model for rainfall distribution in the face of rainfall instability in the area. The generalized pareto distribution established in the present study did not feature among the suitable pdf in the rainfall peak prediction of selected cities in Nigeria [8]. Similarly, none of the favoured pdf in [8] featured in our current work.

**Table 4 - Fitting statistics for the best Pdf**

| Sites | Pdf | Kolmogorov Smirnov | | Anderson Darling | | Chi-Squared | |
|---|---|---|---|---|---|---|---|
| | | Statistic | Rank | Statistic | Rank | Statistic | Rank |
| **Daura** | General Pareto | 0.23696 | 1 | 14.381 | 2 | 52.428 | 2 |
| | General Extreme Value | 0.27619 | 2 | 18.216 | 4 | 80.008 | 6 |
| | Gumbel Maximum | 0.27671 | 3 | 16.498 | 3 | 97.065 | 10 |
| | Uniform Distribution | 0.28748 | 4 | 79.495 | 20 | NA | |
| | Normal Distribution | 0.30111 | 5 | 20.109 | 5 | 67.174 | 4 |
| **Dutsin-Ma** | General Pareto | 0.28192 | 1 | 52.877 | 6 | 129.4 | 2 |
| | General Extreme Value | 0.32295 | 5 | 66.155 | 10 | 190.84 | 3 |
| | Gumbel Maximum | 0.31508 | 4 | 54.783 | 7 | 349.37 | 7 |
| | Uniform Distribution | 0.29727 | 2 | 68.43 | 12 | NA | |
| | Normal Distribution | 0.34536 | 7 | 59.294 | 9 | 355.99 | 8 |
| **Safana** | General Pareto | 0.22987 | 3 | 10.837 | 1 | 51.376 | 2 |
| | General Extreme Value | 0.26201 | 7 | 14.421 | 3 | 70.192 | 4 |
| | Gumbel Maximum | 0.26334 | 9 | 14.549 | 4 | 72.832 | 6 |
| | Uniform Distribution | 0.22154 | 2 | 41.12 | 13 | 0.22154 | 2 |
| | Normal Distribution | 0.24312 | 4 | 13.214 | 2 | 70.476 | 5 |

**Table 5 - Distribution characteristics**

| Parameter | Daura | Dutsin-Ma | Safana |
|---|---|---|---|
| ***μ*** | -14.811 | -18.21 | -25.833 |
| ***Σ*** | 61.77 | 72.782 | 139.18 |
| ***K*** | 0.18683 | 0.05654 | -0.325 |
| $\bar{x}$ | **61.15098** | **58.93371** | **79.20851** |
| | | | |
| ***Σ*** | 61.77 | 72.782 | 139.18 |
| ***K*** | 0.18683 | 0.05654 | -0.325 |
| ***Variance*** | **9212.603** | **6709.908** | **6687.102** |

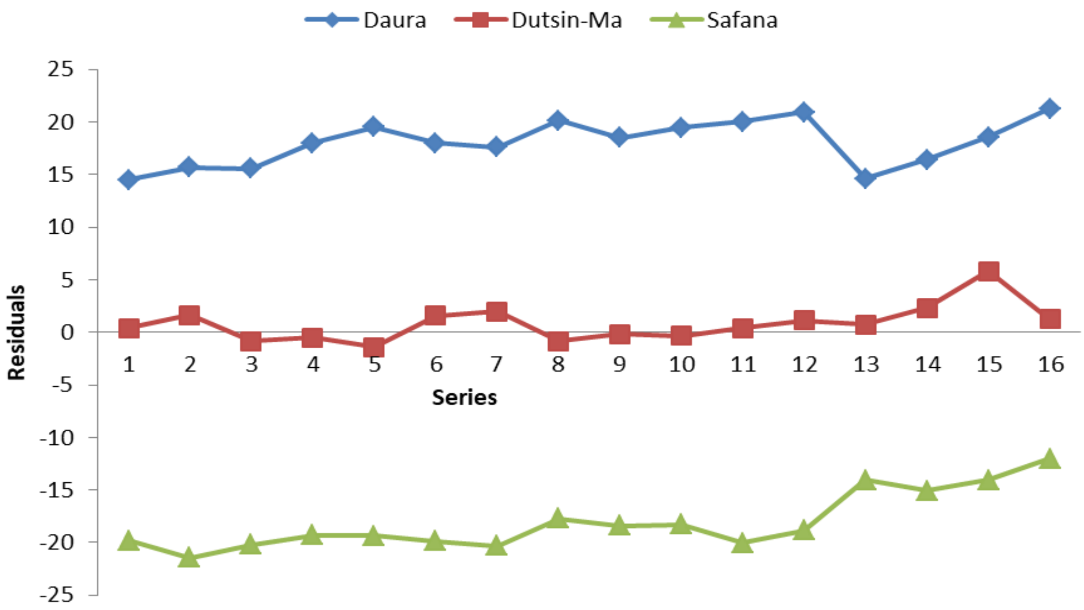


**Fig. 3 - Sum of the residual of the mean annual rainfall for Daura, Dutsinma and Safana**

This can be hinged on the dissimilarity of the target prediction variable in both cases. The present work predicts rainfall distribution while [8] predicted rainfall peak. Also, unlike in the present study, the best performance probability distribution function for modelling visibility for free-space optical link in Nigeria was lognormal distribution [24]. This might be due to differences in the variables being predicted. The suitability of the generalized pareto distribution as the best pdf for the 3 sites of Katsina states contrast with [8]. No single best pdf was arrived at for all the sites and this could be due to spread of the study site which is over Nigeria (including South and North). The implication of this is that the probability density function would differ according to the coveted variables and also according to the aspect of such coveted variable.

The exaggeration and underestimation of the GPD as established in our present study coincides with exceedance characteristics of the GPD [25].

Lastly, the GPD for all the sites satisfy 2 of the properties of the pdf which include;

i) $f(x) \geq 0$ for any value of *x* (that is the probability for any value is greater than zero or exist) and

ii) $\int_{-\infty}^{\infty} f(t)dt = 1$

Based on these properties, it could conveniently be established that the GPD of the rainfall for all the sites exist.

## 5. Conclusion

From this study it can be concluded that variability is believed to be an immanent character of climate due to the fluctuation of the climatic variables especially rainfall. The present work has utilized a probability density function (Generalized Pareto Distribution - GPD) to establish a more parsimonious model for rainfall distribution for the State. The practical application of this work is that it injects specificity into the study of the rainfall distribution thereby reducing the risk of prediction failure for agricultural activities. This study recommends that the scope of the study can be extended to cover more states of the Northern origin and that the model can be employed in risk assessment for optimum agricultural practices.

## Acknowledgement

The authors wish to appreciate the Department of Forestry and Wildlife Management, Federal University, Dutsin-Ma, Katsina and Institute of Agricultural Research and Training, Obafemi Awolowo University, Moor Plantation, Ibadan, Nigeria.

## References


[1] Adakayi, P.E., Oche, C.Y. & Ishaya, S. (2016) Assesment of Pattern of Rainfall in Northern Nigeria. Ethiopian Journal of Environmental Studies and Management. 9(5):554-566

[2] Adakayi, P. E., 2012. An assessment of the rainfall and temperature variations in parts of northern Nigeria. Doctoral thesis, University of Jos, Plateau State, Nigeria available at http://irepos.unijos.edu.ng/, 378pp

[3] Pickands, James (1975). "Statistical inference using extreme order statistics". *Annals of Statistics*. **3**: 119–131

[4] Smith, R.L. (1984), "Threshold Methods for Sample Extremes," In Statistical Extremes and Applications, ed. J. Tiago de Oliveira, Dordrecht, Reidel, pp. 621 – 638

[5] Castillo, E. & Hadi, A.S.(1997), "Fitting the Generalized Pareto Distribution to Data," Journal of the American Statistical Association 92: 440 (1997). 1609-1620

[6] Bermudez, P. D.Z. & Kotz, S. (2010). Parameter estimation of the generalized Pareto distribution-Part I. *Journal of Statistical Planning and Inference*, *140*, 1353C1373: 21-29

[7] Andrade, T.A.N., Fernandez, L.M.Z., Silva, F.S.G. & Cordeiro, G.M. (2017). The Gamma Generalized Pareto Distribution with Applications in Survival Analysis. International Journal of Statistics and Probability; Vol. 6 (3): 140 – 156

[8] Olofintoye, O.O, Sule, B.F & Salami, A.W. (2009). Best–fit Probability distribution model for peak daily rainfall of selected Cities in Nigeria. New York Science Journal, 2(3): 1 – 12

[9] Sharma, M.A. and Singh, J.B. (2010). Use of Probability Distribution in Rainfall Analysis. *New York Science Journal* 3(9):40 – 49

[10] Zhanling L., Zhanjie L., Wei Z. and YuehuaW. (2015). Probability Modeling of Precipitation Extremes over Two River Basins in Northwest of China. *Advances in Meteorology*, Article ID 374127, 13 pp

[11] Andersson, E. & Nilsson, E. (2018). An Extreme Value Approach to Modeling Risk of Extreme Rainfall in Bangladesh. Bachelor of Science thesis at the Lund Univeristy, Faculty of Science, Mathematical Statistics, 39pp

[12] Odekunle, T.O. (2004). Rainfall and the Length of the Growing Season in Nigeria. International Journal of Climatology, 24: 467–479

[13] Cousineau D. and Chartier, S. (2010). Outliers detection and treatment: a review. *International Journal of Psychological Research*, 3 (1), 59-68

[14] Storer, R. L., Griffin , B. M., Höft , J., Weber, J. K., Raut, E., Larson, V. E., Wang, M. and Rasch P. J. (2015). Parameterizing deep convection using the assumed probability density function method. Development and technical paper, Geosci. Model Dev., 8, 1–19, https://doi.org/10.5194/gmd

[15] Marsaglia G, Tsang WW, Wang J (2003). "Evaluating Kolmogorov's Distribution". *Journal of Statistical Software*. **8** (18): 1–4

[16] Kikawa, C.R., Shatalov, M.Y., Kloppers, P.H. Mkolesia, A. (2016). Parameter estimation for a mixture of two univariate gaussian distributions: A comparative analysis of the proposed and maximum likelihood methods. Journal of Advances in Mathematics and Computer Science, 1-8

[17] Marozzi, Marco (2013). "Nonparametric Simultaneous Tests for Location and Scale Testing: a Comparison of Several Methods". *Communications in Statistics – Simulation and Computation*. **42** (6): 1298–1317

[18] Dauda, Taofik O., Asiribo, O.E and Akintoye, N.A. (2012) On the Estimation of a Scaled Weibull Distribution of Rainfall Data of South West Nigeria. *International Journal of Modern Mathematical Sciences*, 6(3): 1- 11

[19] Dimitrova DS, Kaishev VK, Tan S (2020). "Computing the Kolmogorov–Smirnov Distribution when the Underlying cdf is Purely Discrete, Mixed or Continuous". *Journal of Statistical Software*. **95** (10): 1–42

[20] Oyewole,J.A, Thompson,A.M, Akinpelu,J.A & Jegede O.O (2014). Variation of Rainfall and Humidity in Nigeria. Journal of Environment and Earth Science, 4(2):29 – 37

[21] Adenodi, R.A. (2018). A Centurial Analysis of Rainfall Variability in Nigeria. *Nigerian Journal of Technology (NIJOTECH), 37(2): 543-547*

[22] Itiowe, T., Hassan, S. M. & Agidi, V. A. (2019). Analysis of Rainfall Trends and Patterns in Abuja, Nigeria. *Current Journal of Applied Science and Technology, 34(4):1-7*

[23] Olaniran, O.J. & Sumner, G.N. (1989). A Study of Climatic Variability in Nigeria Based on the Onset, Retreat, and Length of the Rainy Season. *International Journal of Climatology*, 9:253-269

[24] Ojo, J.S., Ojo, O.L. and Ibitola, G.A 2018. Performance probability distribution function for modeling visibility for free space optical link in Nigeria. *World Scient i f ic News 109 (2018) 211-234*

[25] Kotz, S., & S. Nadarajah. (2000). Extreme Value Distributions: Theory and Applications. London: Imperial College Press, 356pp