MAUSAM, **75**, 2 (April 2024), 433-442

MAUSAM

DOI : https://doi.org/10.54302/mausam.v75i2.5875
Homepage: https://mausamjournal.imd.gov.in/index.php/MAUSAM

UDC No. 551.577.37 (667.1)

Modeling rainfall extremes along the coastal and Northern parts of Ghana

SAMPSON TWUMASIANKRAH, WILHEMINA ADOMA PELS and SARALEES NADARAJAH*

Department of Statistics and Actuarial Science, Kwame Nkrumah Uni. of Science and Tech., Kumasi, PMB UPO, Ghana

**School of Mathematics, University of Manchester, UK*

(*Received 21 April 2022, Accepted 21 September 2023*)

e mails : sampson.ankrah@yahoo.com, wilheminapels2@gmail.com, saralees.nadarajah@manchester.ac.uk

सार – इस अध्ययन का मुख्य उद्देश्य घाना के तटीय और उत्तरी क्षेत्रों में अत्यधिक वर्षा के लिए उचित वितरण निर्धारित करना था। हितधारकों और नीति निर्माताओं के लिए बाढ़ और सूखे से होने वाले नुकसान को कम करने के लिए कम जोखिम वाले उचित नियम बनाने के लिएअत्यधिक वर्षा के बारे में उचित अनुमान लगाना आवश्यक है। इस अध्ययन मेंबहुचर और अविभाज्य चरम मूल्य डेटा विश्लेषण दृष्टिकोण दोनों का उपयोग किया गया है। इस अध्ययन में ब्लॉक मैक्सिमा दृष्टिकोण के साथ सामान्यीकृत चरम मूल्य (जीईवी) के साथ सामान्यीकृत पेरेटो डिस्ट्रीब्यूशन (जीपीडी) (यानी, सभी अतिरिक्त और डिक्लस्टर शिखर दृष्टिकोण) का चरम सीमा तक उपयोग किया गया। जलवायु अनुसंधान इकाई से 1970 से 2020 तक का ऐतिहासिक ग्रिडयुक्त मासिक अधिकतम वर्षा डेटा प्राप्त किया गया और इसे तटीय और उत्तरी स्टेशनों के रूप में समूहीकृत किया गया। मॉडल मापदंडों का अनुमान लगाने के लिए अधिकतम संभावना अनुमान पद्धति का उपयोग किया गया, और डेटा में प्रवृत्ति के परीक्षण के लिए यूनिट रूट परीक्षण और मान-केंडल परीक्षण दोनों का उपयोग किया गया। बहुचर चरम मॉडलिंग दृष्टिकोण के साथ, लॉजिस्टिक बाइवेरिएट GEV मॉडल को "सर्वश्रेष्ठ" मॉडल के रूप में चुना गया। हालाँकि, निर्भरता मान 0.965 था, इसलिए अत्यधिक वर्षा को अविभाज्य चरम मान दृष्टिकोण का उपयोग करके स्वतंत्र रूप से मॉडल किया जाना चाहिए। इसलिए, सूचना मानदंड और विचलन दृष्टिकोण के विश्लेषण के आधार पर, जीईवी वितरण को घाना के उत्तरी भाग के लिए अत्यधिक वर्षा डेटासेट के लिए "सर्वश्रेष्ठ" माना गया। इसके विपरीत, जीपीडी वितरण तटीय स्टेशन के लिए "सर्वोत्तम" पाया गया। तुलनात्मक रूप से, वर्ष 2020 में वर्षा की मात्रा के लिए, अगले दो वर्षों में घाना के तटीय क्षेत्र में अत्यधिक वर्षा अधिक होने का अनुमान है। साथ ही, 2 वर्षों में अत्यधिक वर्षा सितंबर 2020 में घाना के उत्तरी स्टेशन पर होने वाली अधिकतम वर्षा की घटना (279.267) से अधिक नहीं होगी।

ABSTRACT. The main objective of the study was to determine the appropriate distribution for extreme rainfall along the coastal and northern sectors of Ghana. For stakeholders and policymakers to make appropriate risk-mitigating measures to lessen the damage caused by flood and drought, it is necessary to make proper inferences about extreme rainfall. In this study, we used both the multivariate and univariate extreme value data analysis approaches. The Generalized Extreme Value (GEV) with the Block Maxima approach and Generalized Pareto Distribution (GPD) with the Peak over the threshold (that is all excesses and decluster peaks approaches) were used in this study. Historical gridded monthly maximum rainfall data from 1970 to 2020 were obtained from the Climatic Research Unit and were grouped as the coastal and northern stations. The Maximum Likelihood Estimation method was used to estimate the model parameters, and both the unit root test and the Mann-Kendall tests were used to test for trend in the data. With the multivariate extreme modelling approach, the logistic bivariate GEV model was chosen as the “best” model. However, the dependence value was 0.965, so the extreme rainfall should be modelled independently using the univariate extreme value approaches. Hence, based on the information criteria and analysis of deviance approaches, the GEV distribution was considered the “best” fit for the extreme rainfall dataset for the northern part of Ghana. In contrast, the GPD distribution was the “best” fit for the coastal station. Comparatively, for the volume of rainfall in the year 2020, the extreme rainfall is expected to be higher in the coastal station of Ghana in the next two years. Also, extreme rainfall in 2 years would not exceed the maximum occurrence of rainfall (279.267), which happened in September 2020 at the northern station of Ghana.

Key words – Information criteria, Generalized pareto distribution, Generalized extreme value, Block maxima, decluster peaks.

1. Introduction

Generally, extreme low or high rainfall affects the socioeconomic activities of people. Extreme maxima or minima of rain are natural events that result in floods or droughts. Severe floods cause critical harm to farms, ecology, infrastructure (both building and transport), disruption of human activities and energy supply, injuries, loss of lives, disruption of communication service, and others (Ansah *et al*., 2020). According to Ofori-Sarpong (1986), in Ghana's modern history, the most severe drought began in 1981. This poses a significant threat to national development, which manifests in increasing desertification levels in the northern savannah, undermining the agricultural potential and economic viability of the north of the ecological zone.

Recently, major cities in Ghana, like Accra, the capital town, get flooded with the slightest rains causing huge losses. Hence, appropriate models that can predict extreme events like rainfall are a significant concern for researchers and policymakers. Extreme value theory can capture and quantify the stochastic behavior of processes that are rarely large or small levels (Hasan *et al*., 2012).

Several studies detail the relevance of modeling rainfall using extreme value theory in various parts of the world like Europe (Buishand *et al*., 2008); France (Carreau *et al*., 2013); Taiwan (Chu *et al*., 2013); China (Ender and Ma, 2004); Australia (Li *et al*., 2005); Mozambique (Maposa *et al*., 2014); Bangladesh (Shahid, 2011), Colombo (Varathan *et al*., 2010), and South Korea (Nadarajah and Choi, 2007). A common objective of these studies is searching for appropriate distribution for extreme rainfall in the respective locations.

However, there is limited research done on modeling rainfall using extreme value theory in Ghana. To the best of our knowledge, the only publications entirely related to our work that used extreme value theory are Ibn Musah *et al*. (2018) and Nkrumah (2017). Ibn Musah *et al*., (2018) used Structural Equation Modeling (SEM) regression analysis to evaluate the potential impacts of weather extremes on Ghana's agriculture production. They did not focus much on the extreme value analysis of rainfall. They assumed that the GEV approach was appropriate for modeling rainfall extremes in Ghana. Nkrumah (2017) fitted GEV and GPD models for temperature and rainfall for some selected regions in Ghana. He did not compare these distributions to know which one is appropriate for rainfall data in Ghana. In 2007, Kwaku and Duke conducted a study on Accra's annual rainfall. They analyzed the maximum rainfall for one day as well as the maximum rainfall for two to five consecutive days. The study used the normal, lognormal, and gamma distributions to determine which probability distribution best fitted the data. Their approach has a disadvantage due to the utilization of distributions that are not extreme value distributions.

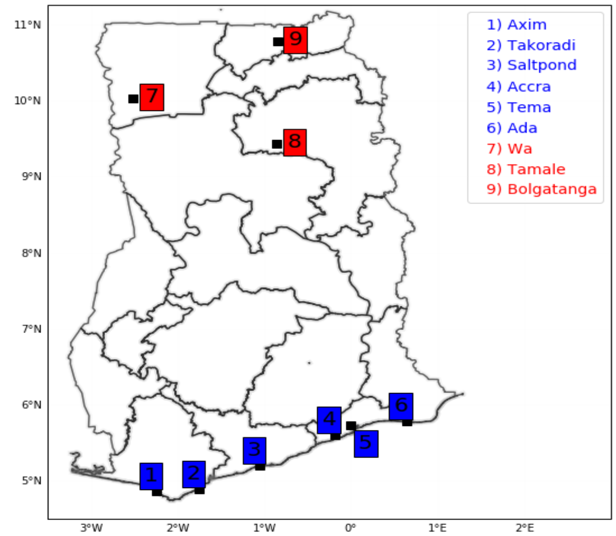

Fig. 1. The map of Ghana showing the selected stations or towns

Hence, this study seeks to model the behavior of extreme rainfall events in two significant locations in Ghana, that is, considering flood-prone areas (*i.e*., coastal part of Ghana) and areas that are vulnerable to climate change (*i.e*., the northern part of Ghana) using the GEV and GPD approaches. Our initial focus was to determine whether a bivariate extreme value approach was appropriate for the extreme rainfall data in these sites.

2. Methods

2.1. *Dataset and study area*

Monthly maximum rainfall datasets for nine cities in Ghana spanning from 1970-2020 were obtained from the climate monitor unit hosted by the Climatic Research Unit, University of East Anglia. Therefore, the total sample size is 612. The average of the six sites along the coastal part of Ghana (*i.e*., Accra, Ada, Axim, Saltpond, Tema, and Takoradi) was taken to form an aggregated dataset for the coastal station. Again, we obtained an aggregated dataset for the northern station by taking the average of three northern sites (*i.e*., Bolgatanga, Tamale, and Wa) of Ghana (see Fig. 1).

2.2. *Trend analysis*

Two statistical tests, namely the Mann-Kendall trend test and the unit root test, were used to check for a trend in our dataset. The Mann-Kendall trend test is a statistical test that determines whether there is a trend in a time series dataset. It is a non-parametric test and is based on the following hypothesis.

H_0 : There is no existence of a trend in the data series

H_1 : The existence of a trend in the data series

In this study, the four-unit root tests used are Augmented Dickey-Fuller (ADF), Phillips & Perron (PP), the Kwiatkowski Phillips Schmidt and Shin (KPSS) and the Kilic nonlinearity tests (Kilic, 2011). The ADF and the PP have the same null hypothesis; the time series has a unit root (that is, it is not stationary), and the alternative hypothesis is that the time series does not have a unit root (that is, it is stationary). However, the KPSS has its null hypothesis as the series is stationary with alternative as the series is not stationary. Also, the Kilic nonlinearity has the null hypothesis, existence of linear unit root against the alternative of stationary exponential smooth transition autoregressive model

$$\Delta y_t = \sum_{t=1}^{p} \delta_i \Delta y_{t-i} + \phi y_{t-1}\left[1 - exp\left(-yz_t^2\right)\right] + u_t \quad (1)$$

where $z_t = \Delta y_{t-d}$ is strictly stationary.

2.3. *Model Selection Criteria*

In this study, the two commonly used information criteria are utilized: The Akaike information criterion (AIC) and the Bayesian information criterion (BIC). AIC is formulated as:

$$\text{AIC} = -2\,ln\,L\left(\hat{\theta}_k\right) + 2p \quad (2)$$

Where $L(\hat{\theta}_k)$ is the likelihood of the fitted model, and p is the number of unknown parameters. The formulation of the BIC is given as:

$$\text{BIC} = -2\,ln\,L\left(\hat{\theta}_k\right) + p\,ln(n) \quad (3)$$

where n is the sample size.

2.4. *Modeling approaches*

In this study, we considered two modeling approaches to extreme events, namely multivariate extreme modeling and univariate extreme modeling [Generalized Extreme Value (GEV) and Generalized Pareto Distribution (GPD)]. For GEV, the block maxima approach was considered as annual maxima (*i.e.*, maximum temperature value each year over the entire 51-year period). For the GPD, all excess above the threshold and the decluster approach were considered.

For the Block Maxima approach, the Generalized Extreme Value Distribution with three parameters was used, and it is given as:

$$G_{\mu\sigma\xi}(x) = exp\left\{-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{\frac{-1}{\xi}}\right\} \quad (4)$$

where μ is the location, σ is the scale, and ξ is the shape. The distributions of GEV can be divided into three corresponding distributions by the values of the shape parameter (ξ):

(*i*) $\xi = 0$ signifiesa Gumbel distribution

(*ii*) $\xi > 0$ signifies a Fréchet distribution

(*iii*) $\xi < 0$ signifies a Weibull distribution

And for the Peaks over Threshold method, Generalized Pareto Distribution (GPD) which is a limiting distribution for excesses over a sufficiently high threshold u for a random variable X is defined as:

$$F_u(x) = P\left(X - u \le x \middle| X > u\right) \quad (5)$$

for $0 \le x < x_F - u$ where z_F is the right endpoint of F.

2.5. *Modelling the dependence structure*

There are many families of distributions within the class of bivariate extreme value models that can be utilized to model the dependence structure of the dataset. These distributions are

(*i*) Logistic Model-symmetric introduced by [17].

$$F(x,y) = \exp\left\{-\left(x^{-\frac{1}{a}} + y^{-\frac{1}{a}}\right)^{a}\right\} \quad (6)$$

$$x > 0,\, y > 0, \text{for } a \in (0,1)$$

(*ii*) Asymmetric Logistic Model introduced by Tawn, 1988.

$$G(z_1, z_2) = \exp\left\{-(1-t_1)_{y_1} - (1-t_2)y_2 - \left[(t_1 y_1)^{\frac{1}{r}} + (t_2 y_2)^{\frac{1}{r}}\right]^{r}\right\} \quad (7)$$

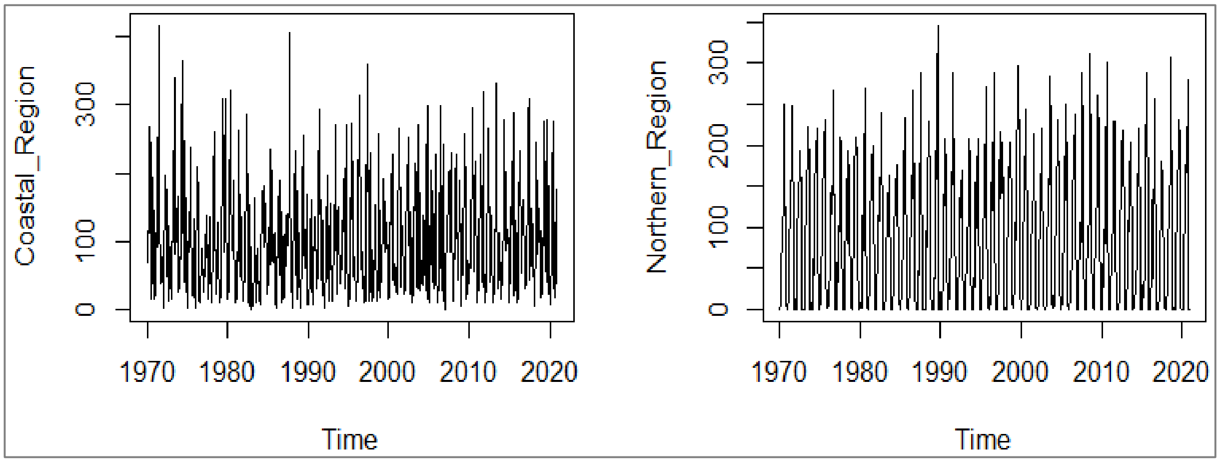

Fig. 2. Time series plot of monthly rainfall for coastal and northern station of Ghana

TABLE 1

Descriptive statistics of monthly rainfall in the coastal and northern stations of Ghana

Stations	N	Min	Max	Mean	S. Dev	Skewness	Shapiro-Wilk Test
Coastal	612	0.15	416.30	100.35	78.14	1.07	0.912(<0.01)
Northern	612	0.00	345.50	86.07	82.78	0.66	0.890(<0.01)

where $0 < r \leq 1$ and $0 \leq t_1, t_2 \leq 1$. When $t_1 = t_2 = 1$ the asymmetric logistic model is equivalent to the logistic model. Independence is obtained when either $r = 1$, $t_1 = 0$ or $t_2 = 0$. Complete dependence is obtained in the limit when $t_1 = t_2 = 1$ and r approaches 0.

(*iii*) Negative Logistic Model introduced by [18].

$$F(x,y) = \exp\left\{\frac{1}{x} + \frac{1}{y} - \left(x^a + y^a\right)^{\frac{1}{a}}\right\} \tag{8}$$
$$x > 0, y > 0, \text{for } a > 0$$

(*iv*) Asymmetric Negative Logistic Model introduced by [19].

$$G(z_1, z_2) = \exp\left\{-y_1 - y_2 + \left[(t_1 y_1)^{-r} + (t_2 y_2)^{-r}\right]^{-\frac{1}{r}}\right\} \tag{9}$$

where r > 0 and $0 < t_1, t_2 \leq 1$.

(*v*) Husler-Reiss model introduced by [20]. Its distribution function is

$$G(z_1, z_2) = \exp\left(-y_1\phi\left\{r^{-1} + \frac{r\left[\log\left(\frac{y_1}{y_2}\right)\right]}{2}\right\} - y_2\phi\left\{r^{-1} + \frac{r\left[\log\left(\frac{y_2}{y_1}\right)\right]}{2}\right\}\right) \tag{10}$$

where $\phi()$ is the standard normal distribution function and $r > 0$.

(*vi*) Asymmetric Mixed Model introduced by[21].

This has a dependence function:

$$A(w) = \phi w^3 + \theta w^2 - (\theta + \phi)w + 1 \tag{11}$$
$$(\theta \geq 0,\ \theta + \phi \leq 1, \theta + 2\phi \leq 1, \theta + 3\phi \geq 0)$$
$$\theta \in [0, 1.5] \text{ and } \phi \in [-0.5, 0.5]$$

TABLE 2

Unit root Tests on for monthly Rainfall in Ghana

Stations	P-value for ADF test	P-value for PP test	P-value for KPSS test	Conclusion
Coastal	0.01	0.01	0.1	The series is stationary. No trend in the series
Northern	0.01	0.01	0.1	The series is stationary. No trend in the series
The Kilic Nonlinear Unit Root Tests				
Stations		Estimate (Standard Error)	P-value	Conclusion
Coastal	Z. lag.1	-6.015e-06(4.247e-07)	<0.001 ***	The series is stationary. No trend in the series
	Z. diff. lag	1.292e-01 (3.983e-02)	0.001 **	
Northern	Z. lag.1	-5.940e-06(4.187e-07)	<0.001 ***	The series is stationary
	Z. diff. lag	2.998e-01 (3.765e-02)	0.001 ***	No trend in the series

TABLE 3

Mann-Kendall test of trend for monthly rainfall

Stations	Test Statistic	2 sided P-value	Decision
Coastal	0.0227	0.40162	Do not reject the Null hypothesis. No trend in the series
Northern	0.0134	0.61998	Do not reject the Null hypothesis. No trend in the series

3. Results and discussion

3.1. *Descriptive measures*

Table 1 displays the main descriptive statistics for the monthly rainfall at the coastal and northern stations in Ghana. The observations indicate that the monthly rainfall at both stations follows a distribution that is skewed towards the right. After conducting the Shapiro-Wilk test to test the null hypothesis of a normal distribution, it was found that the *p*-value was less than the significance level of 5%, which makes it suitable for conducting extreme value analysis.

3.2. *Data visualization*

Time Series Plot:Based on Fig. 2, there has been no significant pattern in monthly rainfall for both the northern and coastal stations of Ghana over the 51-year period from January 1970 to December 2020. Therefore, both series are showing stationarity.

3.3. *Test of stationarity/trend analysis*

Table 2 shows three unit root tests commonly used to determine if a series is stationary: Augmented Dickey-Fuller, Phillips-Perron, and Kwiatkowski Phillips Schmidt and Shin. With a significance level of 5%, these tests provide enough evidence to reject the hypothesis that the monthly rainfall in both the coastal region and northern station of Ghana are non-stationary.

To confirm the presence of a trend, the Kilic nonlinear unit root test was applied, assuming the null hypothesis of a unit root (*i.e.*, a nonlinear trend). The results in Table 2 confirmed the stationarity of both series. Therefore, both linear and nonlinear trend analysis concluded that there is no trend in the monthly rainfall in Ghana's coastal and northern regions.

A Mann-Kendall (MK) test was conducted to determine if there was a trend in the monthly rainfall data collected from the coastal and northern stations in Ghana. The null hypothesis assumed no trend, and the results are available in Table 3. It was found that none of the series showed a trend, which confirmed the findings from Fig. 1 and Table 2.

3.4. *Threshold selection*

To determine the optimal threshold for GPD modeling, we utilized the threshr package in R, as

TABLE 4

"Best" threshold for various rainfall stations

Station	Best threshold	Best threshold. quantile	Values Above threshold
Coastal	226.119	91	55
Northern	188.4833	85	92

TABLE 5

Dependence models for Annual maximum rainfall in Ghana

Model	GEV		GPD	
	AIC	SIC	AIC	SIC
Logistic	**1091.319**	**1096.826**	2301.8	2305.736
Asymmetric Logistic	1093.94	1101.022	2305.84	2311.35
Husler-Reiss	1091.55	1097.058	2299.99	2303.919
Negative Logistic	1092.46	1097.963	**2299.98**	**2303.917**
Asymmetric Negative Logistic	1095.43	1102.514	2304.33	2309.832
Asymmetric Mixed	1093.64	1099.934	2315.19	2319.913

introduced by [16]. We fitted a quantile sequence between 0.1 and 0.95 and evaluated their predictive performance to identify the most effective threshold value and quantile. The "best" threshold for each station is presented in Table 4.

3.5. *Multivariate modelling of extreme rainfall in ghana*

In order to assess the correlation between maximum rainfall in coastal and northern Ghana, six different multivariate extreme value models were used. These included the logistic and asymmetric logistic models, Husler-Reiss, negative logistic models, asymmetric negative logistic models, and asymmetric mixed models.

3.6. *Selection of "best" dependence models for both GEV and GPD*

When analyzing annual extreme rainfall data in Ghana, we considered two approaches for parameter estimation: the annual blocking method (which looks at the maximum month within a year) and the peak over threshold method. In Table 5, we present the results for bivariate GEV and GPD models using data from coastal and northern stations. According to the minimum information criteria, the Logistic model was the best GEV model for the annual blocking approach, while the Negative Logistic model was chosen as the best GPD model for the peak over threshold approach.

3.7. *Comparative analysis of GPD and GEV models on rainfall stations in Ghana*

We used the analysis of deviance and minimal information criteria approaches to determine the best dependence model. The logistic model was chosen as the best option with an AIC of 1091.319, along with a test statistic and *p*-value from the analysis of deviance as 1212.7 and <0.001, respectively. Using the logistic bivariate GEV model, we estimated the three GEV parameters (location, scale and shape) using Maximum Likelihood Estimation (MLE). It was observed that the strength of dependence approaches 1, indicating independence between maximum rainfall for the coastal station and that of the northern station. Therefore, we recommend that the maximum rainfall at these stations should be modeled separately.

3.8. *Univariate modeling of extremes*

3.8.1. *GEV modeling*

Table 7 displays the output of the GEV modeling conducted on the annual extreme rainfall data of the coastal and northern stations in Ghana. The Block

TABLE 6

Estimated GEV parameters for Annual maximum rainfall in Ghana

<table>
<tr><th>Stations</th><th>Location (SE)</th><th>Scale (SE)</th><th>Shape (SE)</th><th>Dep</th><th>AIC</th></tr>
<tr><td>Coastal</td><td>243.373* (8.480)
[226.752, 259.993]</td><td>55.189* (5.774)
[43.873, 66.505]</td><td>-0.1996* (0.083)
[-0.363, -0.037]</td><td rowspan="2">0.965* (0.106)
[0.757, 1.172]</td><td rowspan="2">1091.319</td></tr>
<tr><td>Northern</td><td>224.129* (5.878)
[212.609, 235.649]</td><td>37.236* (4.186)
[29.032, 45.440]</td><td>-0.172 (0.109)
[-0.386, 0.041]</td></tr>
</table>

TABLE 7

Estimated GEV parameters for Annual maximum rainfall in Ghana

Stations	Location (SE)	Scale (SE)	Shape (SE)	AIC	BIC
Coastal	243.338* (8.460) [226.761, 259.951]	55.096* (5.772) [43.816, 66.522]	-0.206* (0.0794) [-0.362, -0.050]	563.314	565.674
Northern	224.155* (5.878) [212.670, 235.712]	37.343* (4.218) [29.078, 45.596]	-0.1811 (0.104) [-0.386, 0.023]	526.118	528.478

TABLE 8

GPD parameter estimates for annual maximum rainfall in Ghana (All Excesses Approach)

Station	Scale (SE)	Shape(SE)	AIC	SIC
Coastal	59.184 * (10.823) [37.971, 80.397]	-0.184 (0.126) [-0.431, 0.064]	542.8273	549.9743
Northern	51.822 * (6.969) [38.163, 65.481]	-0.256 (0.089) [-0.430, -0.082]	867.1302	874.2772

TABLE 9

GPD parameter estimates for annual maximum rainfall in Ghana (Cluster Peaks)

Station	Scale(SE)	Shape(SE)	AIC	SIC
Coastal	68.80* (12.98) [43.352, 94.240]	-0.260* (0.127) [-0.509, 0.011]	481.243	484.985
Northern	73.519* (11.745) [50.499, 96.539]	-0.427* (0.102) [-0.626, -0.227]	530.057	534.035

Maxima approach was used. The results indicated that the appropriate distribution for annual maximum rainfall of both stations is the Weibull distribution (ξ<0), based on the shape parameter.

3.8.2. *Modeling with GPD approach*

When modeling with GPD, two methods were evaluated: including all excesses (all data points above the threshold) and using a decluster approach, as was done in this study.

3.8.3. *Model estimates for all excesses*

Table 8 displays the GPD model's estimates of the scale and shape parameters for all stations using all excess data points above the threshold. The shape parameter for both stations indicates a light-tailed distribution with $\xi < 0$.

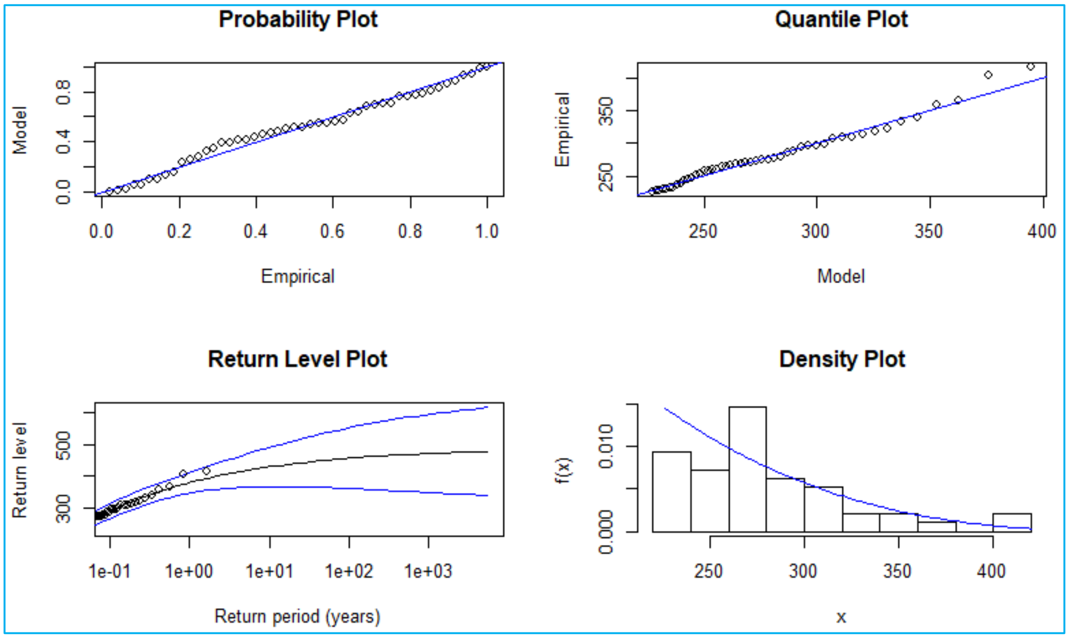

Fig. 3(a). Model diagnostic for annual maximum rainfall at the coastal station of Ghana

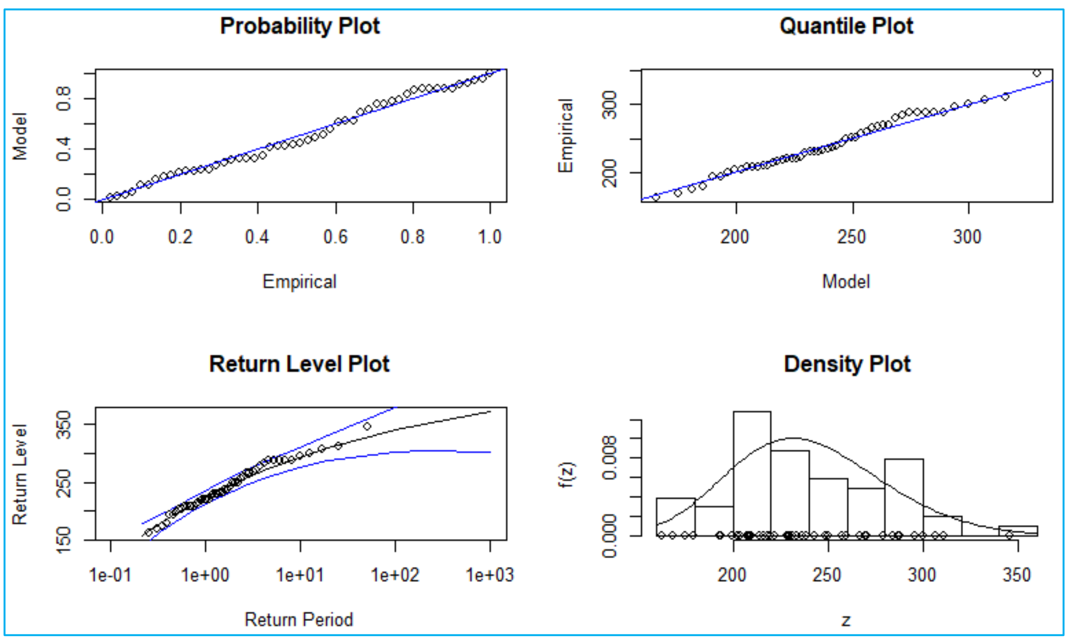

Fig. 3(b). Model diagnostic for annual maximum rainfall at the northern station of Ghana

3.8.4. *Model estimates for cluster peaks*

Table 9 shows the estimates of the scale and shape parameters fitted by the GPD model on two stations using clustered peaks. The results indicate that both stations have a light-tailed shape.

3.8.5. *GPD model selection*

Based on the information criterion (AIC) measure, the cluster peaks approach seems more appropriate than "all excesses" at all stations, with its point estimates significantly different from zero. Hence, the GPD fitted to

TABLE 10

Rainfall stations with their respective extreme value distribution

Station	Lowest AIC	Distribution
Coastal	481.243	GPD
Northern	528.478	GEV

TABLE 11

Return Levels for period units in years for maximum monthly rainfall in Ghana

Station	2-year	5-year	10-year	20-year	30-year	50-year	100-year
Coastal	398.386 [332.04, 464.74]	417.945 [332.64, 503.26]	429.932 [327.66, 532.20]	439.939 [318.46, 561.42]	445.01 [311.16, 578.85]	450.678 [299.90, 601.46]	457.258 [280.68, 633.83]
Northern	237.431 [225.25, 249.61]	273.228 [259.01, 287.45]	293.190 [276.44, 309.9]	309.948 [288.68, 331.22]	318.64 [293.84, 343.44]	328.635 [298.64, 358.63]	340.709 [302.72, 378.70]

the cluster peak approach is more appropriate than "all excesses" approach.

3.8.6. *Comparative analysis of GPD and GEV models on rainfall stations in Ghana*

Based on the minimum AIC, it is determined that the GPD extreme value model with the decluster approach is the best fit for the coastal station rainfall data in Table 10. However, for the northern station, the GEV model provides a better fit.

3.9. *Model Diagnostics and Return Levels*

Fig. 3(a) illustrates the diagnostic plots for the annual maximum rainfall at the coastal station in Ghana, using the "best" GPD model (*i.e*., the cluster peak model). Meanwhile, Fig. 3(b) shows the diagnostic plots for the annual maximum rainfall at the northern station using the GEV model. The QQ-plot for all stations indicates that all points are linear along the unit diagonal. Therefore, the GPD model is a good fit for the maximum rainfall along the coastal sectors, and the GEV model is suitable for the maximum rainfall at the northern station in Ghana. Both figures (*i.e*., Figs. 3(a&b) suggest that there is no real cause for concern about the quality of the fitted model.

In Table 11, the estimated return periods for the cluster peak approach, along with their 95% confidence intervals are presented for the coastal station. Additionally, the GEV return periods for the northern station in Ghana are also given. Based on the findings, it is predicted that the highest amount of rainfall in the next 2 years will not surpass the maximum rainfall that occurred in September 2020 (279.267 mm) at the northern station in Ghana. However, for the coastal station, the results suggest an increase in rainfall volume over the next 2 years compared to the rainfall that occurred in 2020.

4. Conclusion

It is becoming more common for parts of Ghana to experience flooding due to climate change. This can have serious consequences for people, animals, and property. To better understand the extreme rainfall patterns in Ghana, we used multivariate and univariate modeling techniques to analyze rainfall data.

By examining both the coastal and northern regions, we observed that extreme rainfall events were best modeled independently using a logistic bivariate GEV approach. The decluster peak approach was also found to be more effective than the GPD method for analyzing extreme rainfall patterns.

Overall, the GEV distribution was found to be the best fit for the northern station's rainfall data, while the GPD distribution was best suited for the coastal station's data. We also observed an increasing trend in the return levels for all stations, indicating that extreme rainfall is likely to become more frequent in the coming years.

Specifically, the coastal station can expect even higher levels of extreme rainfall in the next two years than was experienced in 2020. However, the northern station's maximum occurrence of rainfall in September 2020

(279.267 mm) is unlikely to be exceeded in the next two years.

Author Declarations

Funding : No funding was received for this study.

Competing interests : The authors declare no competing interests.

Availability of data and material (data transparency).

The datasets analyzed during the current study are available in the GitHub repository [https://github.com/stankrah2017/RAINFALLData].

Code availability (software application or custom code) : Available upon request from the corresponding author.

Authors' contributions (optional : please review the submission guidelines from the journal whether statements are mandatory).

STA, SN conceived and designed the study, STA and WAP conducted the analysis and wrote the first draft, SN reviewed the first draft. All authors approved the final copy.

Ethics approval (include appropriate approvals or waivers): Not applicable.

Consent to participate (include appropriate statements) : Not applicable.

Disclaimer : The contents and views expressed in this research paper/article are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

References

Ansah, S. O., Ahiataku, M. A., Yorke, C. K., Otu-Larbi, F., Yahaya, B., Lamptey, P. N. L and Tanu, M., 2020, "Meteorological Analysis of Floods in Ghana", *Advances in Meteorology*.

Buishand, T. A., De Haan, L. and Zhou, C., 2008, "On spatial extremes: With application to a rainfall problem", *Annals of Applied Statistics*, **2**, 2, 624-642.

Carreau, J., Neppel L., Arnaud, P. and Cantet, P., 2013, "Extreme Rainfall Analysis at Ungauged Sites in the South of France", *Comparison of Three Approaches*, **154**, 2, 119-138.

Chu, L. F., Mcaleer, M. and Chang, C. C., 2013, "Statistical Modelling of Extreme Rainfall in Taiwan", Tinbergen Institute, Tinbergen Institute Discussion Papers, 13-006/III; 6536.

Ender, M. and Ma T., 2004, "Scientific and Innovative Mathematical Research", **2**, 1,23-36.

Gumbel, E. J., 1960, "Distributions des valeurs extremes enplusiers dimensions", Publ. Inst. Statist. Univ. Paris, **9**, 171-3.

Hasan, H. B., Ahmad, N. F., Suraiya, B. K., 2012, "Modeling of extreme temperature using generalized extreme value (GEV)distribution: A case study of Penang", In: Ao, Si., Gelman L., Hukins, D. W. L., Hunter A., Korsunsky, A. M., editors Proceedings of the World Congress on Engineering. 1 London, UK. Hong Kong: News wood Limited; 1-6.

Ibn, Musah, A. A., Du, J., Bilaliib, U. T. and Abubakari, S.M., 2018, "Maxima of normal random vectors: between independence and complete dependence. and the future risk in Ghana", *Climate*, **6**, 4,86.

Kılıç, R., 2011, "Testing for a unit root in a stationary ESTAR process", *Econometric Reviews*, **30**, 3, 274-302.

Kwaku, X. S. and Duke, O.,2007, "Characterization and frequency analysis of one day annual maximum and two to five consecutive days maximum rainfall of Accra, Ghana", *ARPN J. Eng. Appl. Sci.*,**2**, 5, 27-31.

Li, Y., Cai, W. and Campbell, E., 2005, "Statistical modeling of extreme rainfall in southwest Western Australia", *Journal of Climate*, **18**, 1999, 852-863.

Maposa, D., Cochran, J. J., Lesaoana, M. and Sigauke, C., 2014, "Estimating high quantiles of extreme flood heights in the lower Limpopo River basin of Mozambique using model-based Bayesian approach", *Natural Hazards and Earth System Sciences Discussions*, **2**, 8, 5401-5425.

Nadarajah, S. and Choi, D., 2007, "Maximum daily rainfall in South Korea", *Journal of Earth System Science*, **116**, 4, 311-320.

Nkrumah, Stephen, 2017, "Extreme Value Analysis of Temperature and Rainfall: Case Study of Some Selected Regions in Ghana", University of Ghana, mphil thesis.

Ofori-Sarpong, E., 1986, "The 1981-1983 Drought In Ghana", *Singapore Journal of Tropical Geography*, **7**, 2.

Shahid, S., 2011, "Trends in extreme rainfall events of Bangladesh", *Theoretical and Applied Climatology*, **104**, 489-499.

Tawn, J. A., 1988, "Bivariate extreme value theory: models and estimation", *Biometrika*, **75**, 3, 397-415.

Varathan, N., Perera, K., Wikramanayake, N., 2010, "Statistical Modelling of Daily Extreme Rainfall in Colombo", *International Conference on Sustainable Built Environments*, 13-14.