

2024-01-30

Frequentist and Bayesian Approaches in Modeling and Prediction of Extreme Rainfall Series: A Case Study from Southern Highlands Region of Tanzania

Kyojo, Erick

Hindawi

<https://dspace.nm-aist.ac.tz/handle/20.500.12479/2538>

Provided with love from The Nelson Mandela African Institution of Science and Technology

Research Article

Frequentist and Bayesian Approaches in Modeling and Prediction of Extreme Rainfall Series: A Case Study from Southern Highlands Region of Tanzania

Erick A. Kyojo ^{1,2}, Silas S. Mirau ¹, Sarah E. Osima ³, and Verdiana G. Masanja ¹

¹Department of Mathematics, The Nelson Mandela African Institution of Science and Technology, P.O. Box 447, Arusha, Tanzania

²University of Dar es Salaam, Mkwawa University College of Education (MUCE), P.O. Box 2513, Iringa, Tanzania

³Tanzania Meteorological Authority (TMA), P.O. Box 3056, Dar es salaam, Tanzania

Correspondence should be addressed to Erick A. Kyojo; kyojoerick75@gmail.com

Received 28 July 2023; Revised 27 December 2023; Accepted 12 January 2024; Published 30 January 2024

Academic Editor: Marzuki Marzuki

Copyright © 2024 Erick A. Kyojo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study focuses on modeling and predicting extreme rainfall based on data from the Southern Highlands region, the critical for rain-fed agriculture in Tanzania. Analyzing 31 years of annual maximum rainfall data spanning from 1990 to 2020, the Generalized Extreme Value (GEV) model proved to be the best for modeling extreme rainfall in all stations. Three estimation methods— L -moments, maximum likelihood estimation (MLE), and Bayesian Markov chain Monte Carlo (MCMC)—were employed to estimate GEV parameters and future return levels. The Bayesian MCMC approach demonstrated superior performance by incorporating noninformative priors to ensure that the prior information had minimal influence on the analysis, allowing the observed data to play a dominant role in shaping the posterior distribution. Furthermore, return levels for various future periods were estimated, providing guidance for flood protection measures and infrastructure design. Trend analysis using p value, Kendall's tau, and Sen's slope indicated no statistically significant trends in rainfall patterns, although a weak positive trend in extreme rainfall events was observed, suggesting a gradual and modest increase over time. Overall, the study contributes valuable insights into extreme rainfall patterns and underscores the importance of L -moments in identifying the best fit distribution and Bayesian MCMC methodology for accurate parameter estimation and prediction, enabling effective measures and infrastructure planning in the region.

1. Introduction

Extreme rainfall events are no longer a distant possibility but a harsh reality that we must face [1–4]. These natural phenomena are becoming more frequent, more intense, and more devastating, affecting millions of people and ecosystems worldwide [3, 5]. Recent scientific evidence [1, 3–8] claims global warming to be the root cause of these extreme weather events. With rising temperatures and increasing levels of atmospheric carbon dioxide, the hydrological cycle is thrown off balance, leading to more frequent and severe floods and droughts across the globe [1, 4, 5, 9]. As per the findings of the Intergovernmental Panel on Climate Change (IPCC) Report [10], rural, underprivileged communities

within developing nations face the highest vulnerability to the impacts of climate change. The repercussions of climate change involve heightened variability in extreme rainfall patterns, leading to fluctuations in river flows and an increased occurrence of droughts and floods [11–14].

Like many other countries, Tanzania is vulnerable to these changes [15]. For example, according to the Tanzania Climate Status Statement of 2022 [16], on February 26, 2022, extreme rainfall and thunderstorms led to casualties and displacement in Sumbawanga, Nkasi, and Kalambo districts, Rukwa region. Additionally, in Ludewa district (Njombe region), extreme rainfall between February 22 and 23, 2022, caused the destruction of four bridges, disrupting road transport temporarily. Likewise, between April 26 and 28,

2022, extreme rainfall caused severe flooding in the Mbeya and Songwe regions of Southern Highlands Tanzania [16]. The flooding resulted in 5 fatalities, 21 injuries, 5 missing people, and the displacement of 3150 individuals. Numerous buildings, including houses, schools, and religious structures, were brutally damaged. Additionally, agricultural fields spanning 24,700 acres of farm fields, including paddy, maize, and banana farms were destroyed, and a significant number of livestock perished [16]. The flooding also led to the destruction of bridges and roads, impacting transportation in various areas [16]. Generally, extreme rainfall events have already taken a severe toll on the country's agricultural sector, causing significant losses in maize and rice production in the Southern Highlands [15, 17–19]. Tanzania and other countries that rely heavily on agriculture, the impact of extreme rainfall is already posing great challenges for their economies. Therefore, it is crucial for climatologists, meteorologists, economic planners, and other policymakers to have a better understanding of heavy rainfall patterns and their future behaviors for efficient and effective planning, decision-making, and mitigation purposes [17].

Extreme rainfall is characterized as the highest amount of rainfall experienced within a 24-hour period annually [20]. One way to analyze changes in extreme rainfall is to use statistical distribution models. These models are useful for studying extreme rainfall and can aid in flood mitigation and control [21, 22]. Numerous probability distributions can be employed to analyze extreme rainfall patterns. However, it is important to note that the data frequently exhibits non-normal distribution characteristics across multiple regions. While the Anderson-Darling and Kolmogorov-Smirnov tests have often been used for fitting probability distributions, they may not be sensitive enough to identify substantial deviations from an assumed distribution at a particular location [20]. To address this issue, researchers suggest utilizing *L*-moments as a reliable method to estimate higher statistical moments. The *L*-moments method allows for reasonable estimates to be obtained even with as small as 20 sample sizes and without making assumptions about the underlying distribution [20, 23]. This method also offers an excellent fit to the entire cumulative distribution function using only a limited number of parameters.

Several researchers [24–28] have used *L*-moments to fit probability distributions of extreme rainfall series, providing a powerful alternative for fitting probability distributions of non-normally distributed data, which is particularly important when dealing with extreme rainfall data where significant deviations from an assumed distribution can occur at a specific location.

This study aims at finding the most suitable probability distribution model for the extreme rainfall series utilizing data from Southern Highlands, Tanzania, and to estimate the extreme rainfall amounts that might occur once in 5, 10, 20, 50, and 100 years to come. The *L*-moments help guide the selection of candidate distributions and provide insights into the shape of the data. It is used to perform distribution

analysis to select the best distributions. The *L*-moments method is a popular approach for analyzing the statistical properties of hydrological variables, especially in regions with limited data [29]. This study applies five distinct probability distribution models, namely, the GEV, generalized pareto (GPD), generalized logistic (GLO), generalized normal (GNO), and the Pearson type three (PE3) distributions in order to evaluate the best fit model. These distributions are chosen due to their simplicity, superior performance, and widespread usage in frequency analysis of extreme events, as highlighted by [20, 23, 30]. It is widely acknowledged that numerous approaches have been extensively employed to estimate the parameters of potential probability distributions, such as the maximum likelihood method, method of moments, *L*-moments method, and least squares method [22, 31–33]. Nonetheless, the Bayesian method has not seen widespread utilization. In this research, we utilize the maximum likelihood estimation, *L*-moments, and Bayesian method to determine the parameters for the chosen probability distribution, and we conduct an extreme value analysis of annual maximum daily rainfall to compare the efficacy of the frequentist approach with the Bayesian approach. Bayesian modeling allows for more supplementary information via prior knowledge, which can improve statistical inference on extremes, considering that extreme data are rare by nature [8].

Overall, it is crucial to model extreme rainfall to understand its potential impacts and inform policy decisions [15, 17, 34, 35]. By using appropriate modeling techniques, we can better understand the behavior of extreme rainfall patterns and how often they occur in the Southern Highlands of Tanzania and help mitigate their effects on the economy, agriculture, and society [36]. The rest of this work is structured as follows: The data, study area, and methodologies for model fitting are described in Section 2 along with the best fitting techniques. Lastly, the statistical analysis results of the most suitable model are presented in Section 3, whereas Discussion and Conclusions are presented in Section 4 and Section 5, respectively.

2. Methods

2.1. Study Area. Figure 1 shows the study area that covers four stations in the Southern Highlands of Tanzania, which is the administrative zone comprised of the four regions: Iringa, Rukwa, Mbeya, and Ruvuma. The Highlands zone is at latitude $05^{\circ} - 11^{\circ}$ S and longitude $30^{\circ} - 39^{\circ}$ E. The rainfall amount in the area varies between 823 mm to 2,850 mm. The rainy season occurs from November to April. During dry season, higher elevation regions face some light rain and mists from May to August. Most rainfall is attributed to thunderstorms forming over the Lake Nyasa (Lake Malawi), and the areas facing towards the lake tend to receive more rain. The highlands have lower temperatures ranging from 13° C to 19° C than the surrounding lowlands. The highest elevated regions face nighttime frosts regularly from June to August.

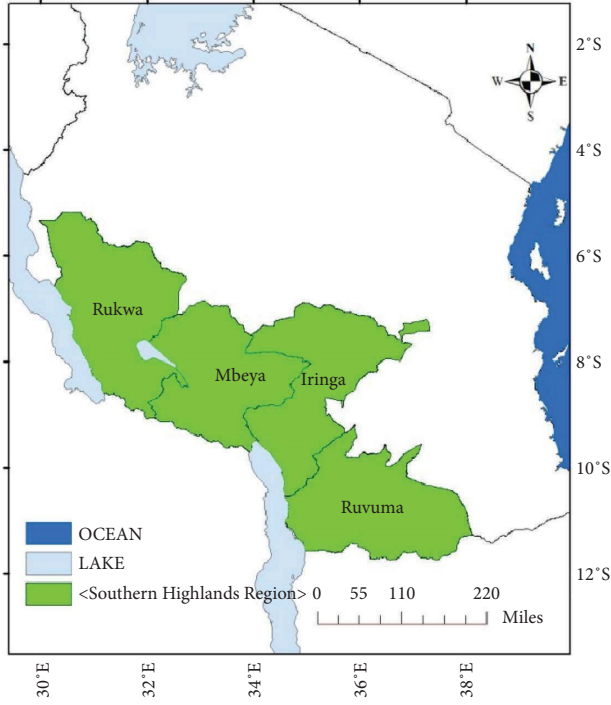


FIGURE 1: Tanzania map showing the Southern Highlands regions.

2.2. Data Description. Rainfall data (in millimeters) was collected from the Tanzania Meteorological Authority (TMA) from four different weather stations in the Southern Highlands over a period of 31 years from 1990 to 2020. Each station represents a different region in the highlands and has its own unique rainfall patterns and location. Because the data were analyzed station by station, a test for homogeneity was not relevant for this study.

As per information provided by the TMA in 2011 and 2019 [15, 17], rainfall that exceeds 50 mm within a 24-hour period is classified as intense (or extreme) and has the potential to result in flooding. Figure 2 presents the original daily rainfall data spanning from 1990-2020, depicting instances of extreme rainfall occurrences exceeding the threshold of 50 mm for each region as indicated by the values above the red line (i.e, rainfall > rbin 50 mm).

2.3. Probability Distributions. Extreme value theory has emerged as a crucial statistical discipline in the field of meteorology and hydrology, playing a significant role over the past six decades. This has led to its widespread adoption in numerous research studies. While many probability distributions can be applied to represent empirical rainfall extremes [22], this study focuses on five commonly used distributions: GPD, GLO, GNO, PE3, and GEV distributions. The aim is to identify the most suitable distribution that best fits the empirical annual rainfall extremes. Leveraging L -moments, we determine the best fit distribution and utilize it to analyze and predict extreme rainfall patterns in the Southern Highlands Region of Tanzania. By employing maximum likelihood, L -moments, and Bayesian

methods to estimate the parameters μ , σ , and ξ , we gained valuable insights into the behavior of extreme rainfall events. The best fit model facilitates the estimation of return levels for extreme rainfall, which is essential in designing robust infrastructure and effective flood protection measures.

2.3.1. Generalized Extreme Value (GEV). The GEV distribution arise from a combination of three models, namely, the Gumbel, Frechet, and Weibull whose cumulative distribution function as given by [9, 20, 22] is as follows:

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

defined on the set of values $\{x: 1 + (\xi(x - \mu)/\sigma) > 0\}$. The parameters of this family of models satisfy the conditions $-\infty < \mu < \infty, \sigma > 0$ and $-\infty < \xi < \infty$. The distribution in question was independently discovered by both [37, 38]. Its three parameters, namely, location μ , scale σ , and shape ξ are used to classify it into three types of extreme value distributions, with ξ being the key differentiating factor. When $\xi > 0$ and $\xi < 0$, the distribution falls under Frechet, and Weibull, respectively. When $\xi = 0$, it corresponds to the limit as $\xi \rightarrow 0$ of equation (1), leading to the Gumbel family of distributions.

2.3.2. Generalized Pareto (GPA). The cumulative distribution function for GPA is given as [20]:

$$F(x) = 1 - \exp(-y), \quad (2)$$

whereby

$$y = -\xi^{-1} \log \left(1 - \frac{\xi}{\sigma} (x - \mu) \right), \quad (3)$$

σ, μ , and ξ are scale, location, and shape parameters. The range of x is given by

$$\text{Range of } x = \begin{cases} \left(\mu + \frac{\sigma}{\xi} \right) \leq x < \infty, & \xi < 0, \\ \text{Exponential distribution,} & \xi = 0, \\ -\infty < x < \left(\mu + \frac{\sigma}{\xi} \right), & \xi > 0, \\ \text{Uniform distribution,} & \xi = 1. \end{cases} \quad (4)$$

2.3.3. Generalized Logistic (GLO). The cumulative distribution function for GLO given by [20, 39] is as follows:

$$F(x) = \frac{1}{1 - \exp(y)}, \quad (5)$$

whereby

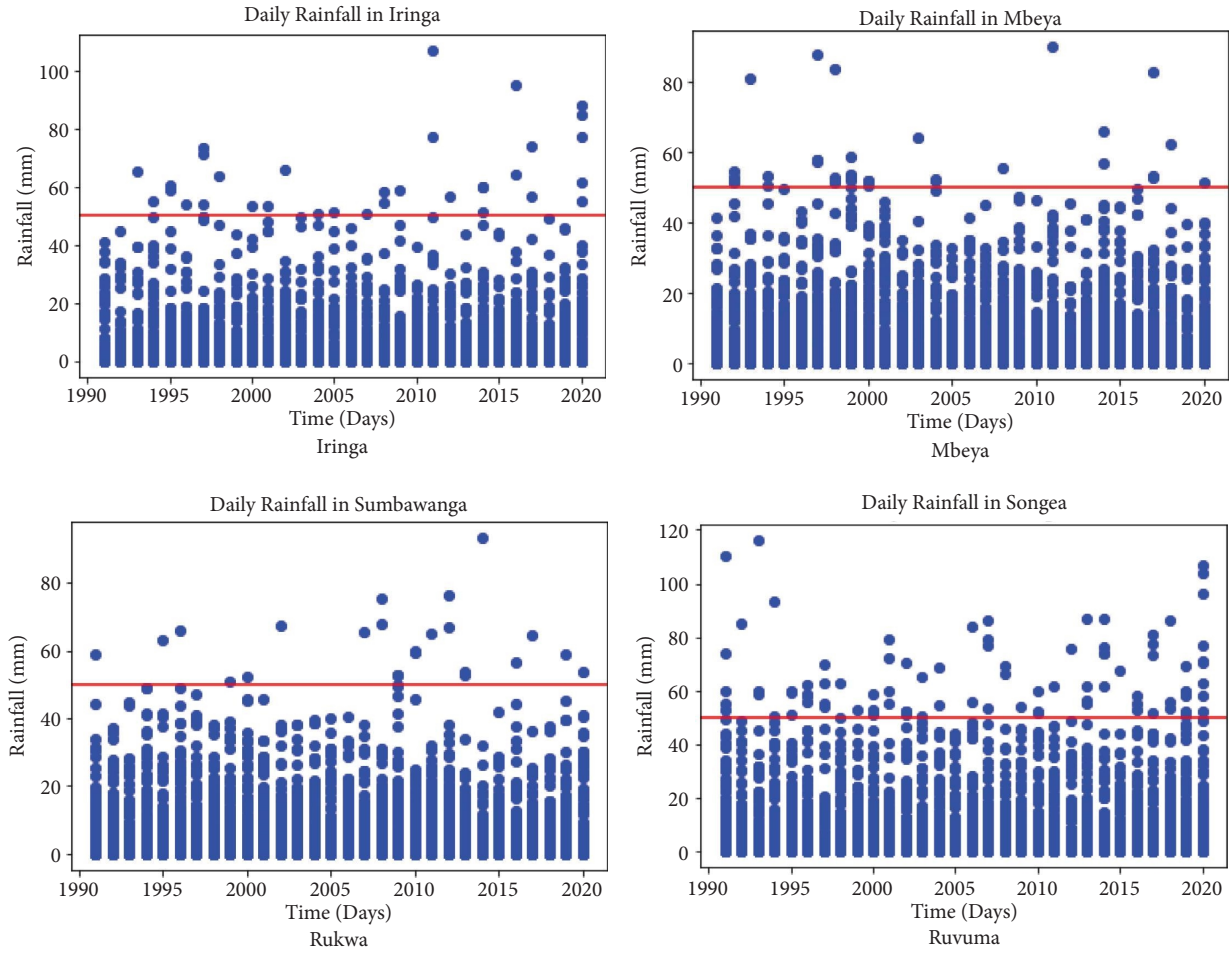


FIGURE 2: Daily rainfall distribution.

$$y = -\xi^{-1} \log \left(1 - \frac{\xi}{\sigma} (x - \mu) \right), \quad (6)$$

σ, μ , and ξ are scale, location, and shape parameters. The range of x is given by

$$\text{Range of } x = \begin{cases} \left(\mu + \frac{\sigma}{\xi} \right) \leq x < \infty, & \xi < 0, \\ \text{Logistic distribution,} & \xi = 0, \\ -\infty < x < \left(\mu + \frac{\sigma}{\xi} \right), & \xi > 0. \end{cases} \quad (7)$$

2.3.4. *Generalized Normal (GNO)*. The probability density function (PDF) and the cumulative probability distribution (CDF) for the GNO distribution are respectively given by

$$f(x) = \frac{\xi}{2\sigma\Gamma(1/\xi)} \exp \left[-\left(\frac{|x - \mu|}{\sigma} \right)^\xi \right], \quad (8)$$

$$F(x) = \frac{1}{2} + \frac{1}{2} \text{sign}(x - \mu) \left[1 - \exp \left(-\left(\frac{|x - \mu|}{\sigma} \right)^\xi \right) \right],$$

where σ, ξ , and μ are the scale, shape, and location parameters, respectively. The range of x is defined as follows:

$$\text{Range of } x = \begin{cases} -\infty < x < \infty, & \xi > 1, \\ \text{Semi - infinite interval,} & \xi = 1, \\ \text{Finite interval,} & \xi < 1. \end{cases} \quad (9)$$

2.3.5. *Pearson Type Three (PE3)*. PE3 has the following cumulative probability distribution [20]:

$$F(x) = \frac{1}{\sigma\Gamma(\xi)} \int_0^x \left(\frac{y - \mu}{\sigma} \right)^{\xi-1} \exp \left(-\frac{y - \mu}{\sigma} \right) dy, \quad (10)$$

μ, σ, ξ are the location, scale, and shape parameters and $\Gamma(\bullet)$ denotes the Gamma function given by

$$\Gamma(x) = \int_0^{\infty} (t)^{x-1} \exp(-t) dt. \quad (11)$$

2.4. L-Moments Method. *L*-moments are based on linear combinations of data or statistics that are sorted in ascending order [40, 41]. They are an improvement over normal moments for assessing probability distribution shape and estimating parameters, especially for small sample sizes. According to [42], *L*-moments can be computed linearly using Shifted Legendre Polynomials and the uniform distribution function as the foundation, providing a more robust estimate for a given amount of data. The first four *L*-moments of a probability distribution can be calculated using a sorted sample vector $(X_{i:n})$ of length n , ranked value i th, and expected value function E as follows:

$$\begin{aligned} \lambda_1 &= E(X_{1:1}), \\ \lambda_2 &= \frac{1}{2} E(X_{2:2} - X_{1:2}), \\ \lambda_3 &= \frac{1}{3} E(X_{3:3} - 2X_{2:3} + X_{1:3}), \\ \lambda_4 &= \frac{1}{4} E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}). \end{aligned} \quad (12)$$

Based on the information presented above, it is evident that the first *L*-moment represents the mean value of a data set by calculating the expected value of all individual values. The second *L*-moment, on the other hand, measures the expected difference between the largest and smallest values in the data set and is equal to half of this expected value.

2.4.1. L-Moments Ratios. The second *L*-moment is always normalized by the mean value, resulting in a *L*-moment coefficient of variability (*L*-Cv) calculated as the ratio of the second *L*-moment to the first *L*-moment:

$$L - Cv(\tau) = \frac{\lambda_2}{\lambda_1}. \quad (13)$$

Higher *L*-moment ratios are obtained by scaling the corresponding higher *L*-moments using the second *L*-moment as a reference point [29]:

$$\begin{aligned} L - Skewness(\tau_3) &= \frac{\lambda_3}{\lambda_2}, \\ L - Kurtosis(\tau_4) &= \frac{\lambda_4}{\lambda_2}. \end{aligned} \quad (14)$$

The *L*-Cv (τ) is similar to the standard CV and *L*-skewness (τ_3) measures the lack of symmetry in a distribution, while *L*-kurtosis (τ_4) measures the peakedness of the distribution. *L*-skewness and *L*-kurtosis are constrained to lie between $(-1,1)$ and *L*-kurtosis is bounded by τ_3 to be not

less than -0.25 . *L*-moments can estimate many probability distributions and are computed from probability-weighted moments. More information can be found in [43].

2.4.2. Probability Weighted Moments. According to [20, 39], probability-weighted moments (PWM) are calculated from sorted data values, where $x_{1:n} \leq \dots \leq x_{n:n}$ represents the sorted sample of size n . The PWM of order r is defined by the following equation:

$$\beta_r = E\{X[F(X)]^r\}, \quad (15)$$

where $E\{\bullet\}$ is the expectation and $F(X)$ is the cumulative distribution function of a random variable X . The estimators of PWM can be obtained using the equation as given by [20, 39]:

$$b_r = n^{-1} \sum_{j=r+1}^n \frac{(j-1)(j-2)\dots(j-r)}{(n-1)(n-2)\dots(n-r)} x_{j:n}. \quad (16)$$

The first four unbiased sample *L*-moments (estimators) are given by [20] as follows:

$$\begin{aligned} l_1 &= b_o \\ l_2 &= 2b_1 - b_o \\ l_3 &= 6b_2 - 6b_1 + b_o \\ l_4 &= 20b_3 - 30b_2 + 12b_1 - b_o. \end{aligned} \quad (17)$$

The sample *L*-moment ratios are given by

$$t_r = \frac{l_r}{l_2}, \quad (18)$$

where t and t_r are the natural estimators of τ and τ_r respectively, and the sample *L*-Cv is given by

$$t = \frac{l_2}{l_1}, 0 \leq t < 1. \quad (19)$$

2.4.3. L-Moment Ratio Diagram. The *L*-moment ratio diagram is a helpful tool for analyzing and testing the fit of probability distributions. This technique involves plotting *L*-Skewness (τ_3) versus *L*-Kurtosis (τ_4) and is commonly used to choose the most viable probability distribution for a given dataset. In the study by [20], two-parameter distributions were represented as points on the diagram, while three-parameter distributions were depicted as lines. The best fit distribution was determined by comparing the location of the data point (τ_3, τ_4) to the line representing the candidate distribution and by measuring the distance between (τ_3, τ_4) and (τ_3, τ_4^{Dist}) , where (τ_3) and (τ_4) represent the skewness and kurtosis values of the observed data, and (τ_4^{Dist}) is the kurtosis value of the candidate distribution [20].

2.4.4. Block Maxima Technique. The extreme values of a set of independent observations, X_1, X_2, \dots, X_n , were analyzed by splitting the data set into m blocks, each of length n , such that n is a large number. By taking the maximum value

within individual blocks, a new set $M_{n,1}, M_{n,2}, \dots, M_{n,m}$ is generated, that can be modeled using the appropriate extreme value probability distribution. It is important to choose an appropriate block size, since if n is too small, the limit model may not provide a good approximation, which can lead to biased parameter estimation and unreliable extrapolation. Conversely, when n is extra large, the number of blocked maximum will be insufficient, causing large variances in estimation. Thus, it is essential to find a compromise between bias and variance.

2.5. *L-Moments Estimation Method.* The *L*-moments method estimates distribution parameters by comparing the sample *L*-moments with the corresponding population *L*-moments [39]. *L*-moments offer quantifications of location, dispersion, skewness, kurtosis, and other aspects of probability distributions or datasets.

2.5.1. *L-Moments for Generalized Extreme Value Distribution.* The GEV parameter estimates distribution has three parameters: μ (location), α (scale), and ξ (shape). The *L*-moments parameter estimates for GEV were calculated using the following expressions when $\xi > -1$ [44]:

$$\begin{aligned}\lambda_1 &= \mu + \alpha \frac{\{1 - \Gamma(1 + \xi)\}}{\xi} \\ \lambda_2 &= \alpha \frac{(1 - 2^{-\xi})\Gamma(1 + \xi)}{\xi} \\ \tau_3 &= 2 \frac{(1 - 3^{-\xi})}{(1 - 2^{-\xi}) - 3} \\ \tau_4 &= \frac{\{5(1 - 4^{-\xi}) - 10(1 - 3^{-\xi}) + 6(1 - 2^{-\xi})\}}{(1 - 2^{-\xi})}\end{aligned}\quad (20)$$

The symbol $\Gamma\{\bullet\}$ represents the gamma function, which is defined as follows:

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt. \quad (21)$$

The parameter ξ is approximated using the expression $\xi \approx 7.8590c + 2.9554c^2$, where c is calculated as $c = (2/3 + \tau_3) - (\log 2/\log 3)$.

$$\begin{aligned}\alpha &= \frac{\lambda_2 \xi}{(1 - 2^{-\xi})\Gamma(1 + \xi)}, \\ \mu &= \lambda_1 - \alpha \frac{\{1 - \Gamma(1 + \xi)\}}{\xi}.\end{aligned}\quad (22)$$

For more information about *L*-moments of various probability distribution models, refer to [44].

2.5.2. *Maximum Likelihood Estimation Method.* MLE aims to identify model parameters that optimize the likelihood function with respect to those parameters. However, MLE is most effective when sample sizes are large (as $n \geq 30$). Despite this, some researchers still prefer to use MLE even with small sample sizes because it can provide accurate estimates for certain values of the shape parameter. [9] commented that MLE is reliable when the shape parameter (ξ) is greater than -0.5 , but its consistency decreases when the shape parameter is less than -0.5 .

2.5.3. *Computing Maximum Likelihood Parameter Estimates.* According to [9], if $\xi \neq 0$, the log-likelihood function of the GEV with three parameters can be expressed as follows:

$$l(\xi, \mu, \sigma) = -n \log \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^n \log \left[1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right] - \sum_{i=1}^n \left[1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right]^{(-1/\xi)}, \quad (23)$$

given that $[1 + \xi(x_i - \mu/\sigma)] > 0$ for $i = 1 \dots n$

In addition, assuming that $X = \{X_1, \dots, X_n\}$ is a random sample from the GPD with the extreme value $X_{(n)}$ above

a threshold u , the log-likelihood function of the GPD can be obtained using the following equation provided by [33, 45]:

$$l(X; \xi, \sigma) = \begin{cases} -n \log \sigma + \left(\frac{1}{\xi} - 1\right) \sum_{i=1}^n \log \left(1 - \frac{\xi X_i}{\sigma}\right), & \text{if } \xi \neq 0, \\ -n \log \sigma - \frac{1}{\sigma} \sum_{i=1}^n X_i, & \text{if } \xi = 0. \end{cases} \quad (24)$$

Similarly, MLE for other distributions were obtained but not presented in this work.

2.6. Bayesian Modeling Method. In the Bayesian approach, the model parameters may not be random, but they are treated as random by assigning probability distributions to describe the uncertainty of their values [21, 46]. According to [47], three critical steps are involved in the Bayesian approach to parameter estimation, which include specifying a probability model for the data, constructing the posterior distribution, and checking the model for its outputs before using them for inference.

2.6.1. Fundamental Principles. Let a dataset $\mathbf{x} = (x_1, \dots, x_n)$ representing realizations of a random variable with a density belonging to the parametric family $\mathbf{F} = f(\mathbf{x}; \theta)$: $\theta \in \Theta$ as detailed by [9]. Prior beliefs about the parameter θ can be expressed using the probability density function $\pi(\theta)$, independent of the observed data. According to [9], the likelihood for θ is given by:

$$\begin{aligned} L(\theta | \mathbf{x}) &= f(\mathbf{x} | \theta), \\ &= \prod_{i=1}^n f(x_i; \theta). \end{aligned} \quad (25)$$

The combination of prior knowledge and the probability distribution of the data is achieved through Bayes' theorem, leading to the posterior distribution of θ as given by [9]

$$\begin{aligned} \pi(\theta | \mathbf{x}) &= \frac{\pi(\theta)L(\theta | \mathbf{x})}{f(\mathbf{x})} \\ &= \frac{\pi(\theta)L(\theta | \mathbf{x})}{\int_{\Theta} \pi(\theta)L(\theta | \mathbf{x})d\theta} \end{aligned} \quad (26)$$

where $\pi(\theta)$ is the prior distribution for θ , $L(\theta | \mathbf{x})$ is the likelihood function, and $\pi(\theta | \mathbf{x})$ is the posterior distribution.

In extreme value analysis, the goal is often to predict the probability of extreme events in the future based on observed data. The predictive distribution, which describes the likelihood of different outcomes of a future experiment, can be obtained within the Bayesian framework. If y represents a forthcoming observation with a probability density function denoted as $f(y | \theta)$, and $\pi(\theta | \mathbf{x})$ corresponds to the posterior distribution of θ based on the observed data \mathbf{x} , the predictive distribution of y given \mathbf{x} can be described as follows [9]:

$$f(y | \mathbf{x}) = \int_{\Theta} f(y | \theta)\pi(\theta | \mathbf{x})d\theta. \quad (27)$$

Computing this integral may be difficult, but simulation-based techniques like Markov chain Monte Carlo (MCMC) can provide estimates of the posterior distribution through a simulated sample [9]. Bayesian procedures can be advantageous when a suitable prior distribution can be specified [9]. The main objective is to obtain a Markov chain that converges to the posterior distribution in Bayesian

statistics. To accomplish this, there are several algorithms available for implementing MCMC simulations, but this study uses the Metropolis-Hastings algorithm.

2.7. Return Level Estimates. After fitting a probability distribution model to the data, the next step is to estimate the return levels for rainfall. The T -year return level, denoted as x_T , is the level that is exceeded on average only once in T years. The cumulative probability of nonexceedance can be given by the following equation [9, 48]:

$$F(x) = \Pr(X \leq x_T) = 1 - \frac{1}{T}, \quad (28)$$

where T is the return period. For instance, in order to obtain x_T from the GEV model with return period T , we need to solve for x in equation (1) which gives the following [9, 48]:

$$x_T = \begin{cases} \mu - \frac{\sigma}{\xi} \left[1 - \left\{ -\ln\left(1 - \frac{1}{T}\right) \right\}^{-\xi} \right], & \text{if } \xi \neq 0, \\ \mu - \sigma \ln\left(-\ln\left(1 - \frac{1}{T}\right)\right), & \text{if } \xi = 0. \end{cases} \quad (29)$$

In addition, for the GPD model, the return level x_m is defined as the extreme level that is exceeded on average once every m observations [9]. This is given by

$$x_m = \begin{cases} \mu + \frac{\sigma}{\xi} \left[(m\zeta_u)^\xi - 1 \right], & \text{if } \xi \neq 0, \\ \mu + \sigma \log(m\zeta_u), & \text{if } \xi = 0, \end{cases} \quad (30)$$

where ζ_u represents the probability that the value X exceeds the threshold u , i.e., $\zeta_u = \Pr(X > u)$. The quantity x_m can also be defined as the m -observations return level [9].

3. Results

3.1. Statistical Analysis. The statistical analysis presented in this section reveals some interesting insights about the yearly maximum daily rainfall in four stations. According to the results shown in Table 1, the Ruvuma region receives the highest amount of rainfall, indicating that this region may be more susceptible to flooding and other related hazards. In contrast, Rukwa, the station with the lowest rainfall amount, may experience some challenges related to water scarcity and drought particularly during the dry seasons when water sources dry up. These findings underscore the significance of considering local weather patterns and climatic conditions when planning for development and resource management in the region.

Using the block maxima method, Figure 3 displays the annual maxima daily rainfall data collected over a time period of 31 years from four stations in the Highlands Region. Upon closer inspection, the results indicate that there is no clear trend in the data, suggesting that these stations have experienced a relatively stable pattern of extreme rainfall events over the past few decades.

TABLE 1: Descriptive statistical analysis of annual maximums.

Station	Period	Minimum	1 st quartile	Median	Mean	3 rd quartile	Maximum
Iringa	1990–2020	39.50	46.85	53.95	58.20	62.85	106.90
Mbeya	1990–2020	37.10	42.80	53.25	54.47	64.08	93.20
Rukwa	1990–2020	32.60	45.05	50.45	54.39	61.40	89.90
Ruvuma	1990–2020	43.80	61.95	69.60	74.31	85.92	116.00

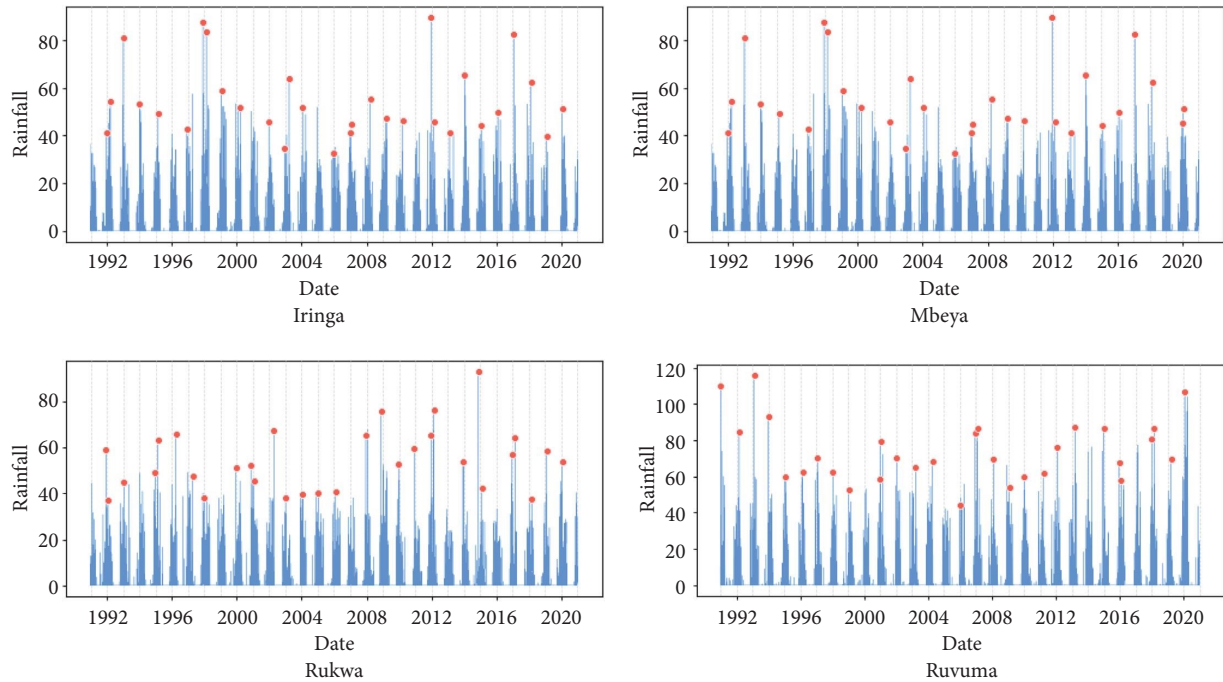


FIGURE 3: Scatter plots for annual maxima rainfall.

3.2. Data Assumptions and Trend Detection Analysis. Before engaging in extreme value modeling of rainfall data, a thorough examination of the data's underlying assumptions was conducted. Potential trends in the series, capable of influencing the modeling process, were carefully assessed for normality, homogeneity, and stationarity. Tests such as Shapiro [49, 50] and augmented Dickey-Fuller (ADF) [51, 52] were employed for this purpose. Given the diverse distribution of stations with varying rainfall patterns and geographical locations, a station-by-station data analysis was conducted, rendering the homogeneity test inapplicable for this study. For more detailed procedures, refer to [11, 19, 53, 54] and [55]. Table 2 displays the results, indicating that the data fulfill the mentioned assumptions and can be used for further analysis. The study also analyzed the Mann-Kendall trend test and Sen's slope estimator for annual maxima rainfall in all stations. Kendall's tau, a measure of the strength and direction of the trend in the data, indicates a weak positive correlation between time and rainfall. All p values are greater than 0.05, suggesting insufficient evidence to reject the null hypothesis of no trend. Sen's slope values are relatively small, indicating a gradual increase in rainfall over time. The trend analysis results, illustrated in

Figure 4, suggest a weak positive trend in rainfall, aligning with the return levels obtained in Table 3. The findings are consistent with the study conducted by [56].

3.3. L -Moments Ratios. L -moments results, as shown in Table 4, provide valuable insights into the characteristics of extreme rainfall in these regions. The λ_1 values represent the mean rainfall depth, with Ruvuma having the highest average rainfall depth of 74.3066. The λ_2 values indicate the variability in rainfall depths, with Iringa having the lowest spread (scale) of 8.34. Kurtosis (τ_4) and Skewness (τ_3) values reveal the heaviness of the tails, and the departure from symmetry, respectively. Iringa shows a slightly right-skewed distribution (0.3177) and a relatively high kurtosis (0.2267). The λ_5 values reflect the departure from symmetry in the upper tail, with Iringa having a thicker upper tail (0.9665). The τ_r values represent the overall shape, and all stations exhibit similar values, indicating moderate peakedness. These findings enable the construction of L -moment ratio diagrams to identify the best-fit distribution for extreme rainfall in each region, facilitating a comprehensive understanding of the rainfall characteristics and aiding in effective hydrological modeling.

TABLE 2: Assumptions on extreme rainfall series using different tests.

Station	Shapiro-Wilk	ADF test	Kendall's tau	Sen's slope
Iringa	0.8967 (0.218)	-14.399 (0.00008)	0.053 (0.694)	0.169
Mbeya	0.8941 (0.341)	-3.7967 (0.0029)	0.0368 (0.789)	0.18
Rukwa	0.8639 (0.089)	-4.2256 (0.000596)	0.0621 (0.643)	0.100
Ruvuma	0.9013 (0.175)	-3.566 (0.000643)	0.0322 (0.817)	0.125

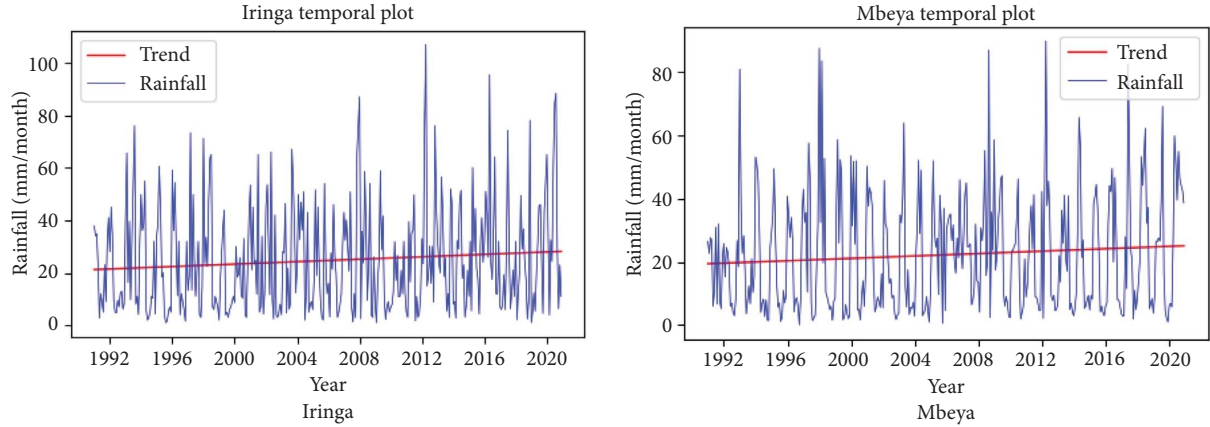


FIGURE 4: Temporal plots for monthly maxima rainfall.

TABLE 3: Return level estimates using MLE, *L*-moments, and Bayesian MCMC.

Return period	MLE method	<i>L</i> -moments method	Bayesian MCMC method
5-years	66.81 (58.20–75.41)	59.47 (66.70–77.26)	60.20 (72.51–92.87)
10-years	77.69 (62.87–9.50)	65.31 (77.53–93.11)	67.86 (87.84–133.75)
20-years	90.22 (64.92–115.51)	71.26 (89.51–111.64)	75.72 (107.68–201.07)
50-years	110.16 (62.22–158.10)	77.55 (108.11–151.27)	85.11 (145.42–364.48)
100-years	128.45 (54.56–202.344)	81.20 (124.74–194.38)	91.83 (188.40–584.76)

TABLE 4: *L*-moments results.

Parameter	Iringa	Ruvuma	Mbeya	Rukwa
λ_1	58.1	74.3066	55.6933	54.4667
λ_2	8.34	9.7536	8.7222	7.6584
λ_3	2.649	1.4908	0.0297	1.0784
λ_4	1.857	1.2427	1.1082	0.6319
λ_5	0.9665	0.1976	-0.4153	0.9003
τ_3 -skewness	0.3117	0.1528	0.2373	0.1408
τ_4 -kurtosis	0.2267	0.1274	0.1271	0.0825
τ_r	0.1433	0.1313	0.1566	0.1406

3.4. *L*-Moments Ratio Diagrams. The study examined the characteristics of extreme rainfall in the research area by analyzing two key parameters: the coefficient of variation denoted as τ and the mean represented by λ_1 . To ensure the reliability of the results, the *L*-moments ratio diagram was employed to assess the relative distribution of extreme rainfall. This diagram played a crucial role in selecting the most appropriate distribution among GPA, GNO, GEV, GLO, and PE3 across the four stations. Based on the *L*-moments ratio diagram results of Figure 5, it was evident that the GEV distribution exhibited the best fit for extreme

rainfall in all four stations. Consequently, parameter estimation for the GEV distribution was carried out using MLE, *L*-moments, and Bayesian MCMC.

3.5. *GEV Parameter Estimates*. To estimate the GEV distribution, we extracted the blocked maximum of daily annual maximum rainfall data from all four stations, using blocks of $n = 365$ days to ensure that our samples were sufficiently large. We then applied maximum likelihood estimation to fit the model to the $N = 31$ annual maximum.

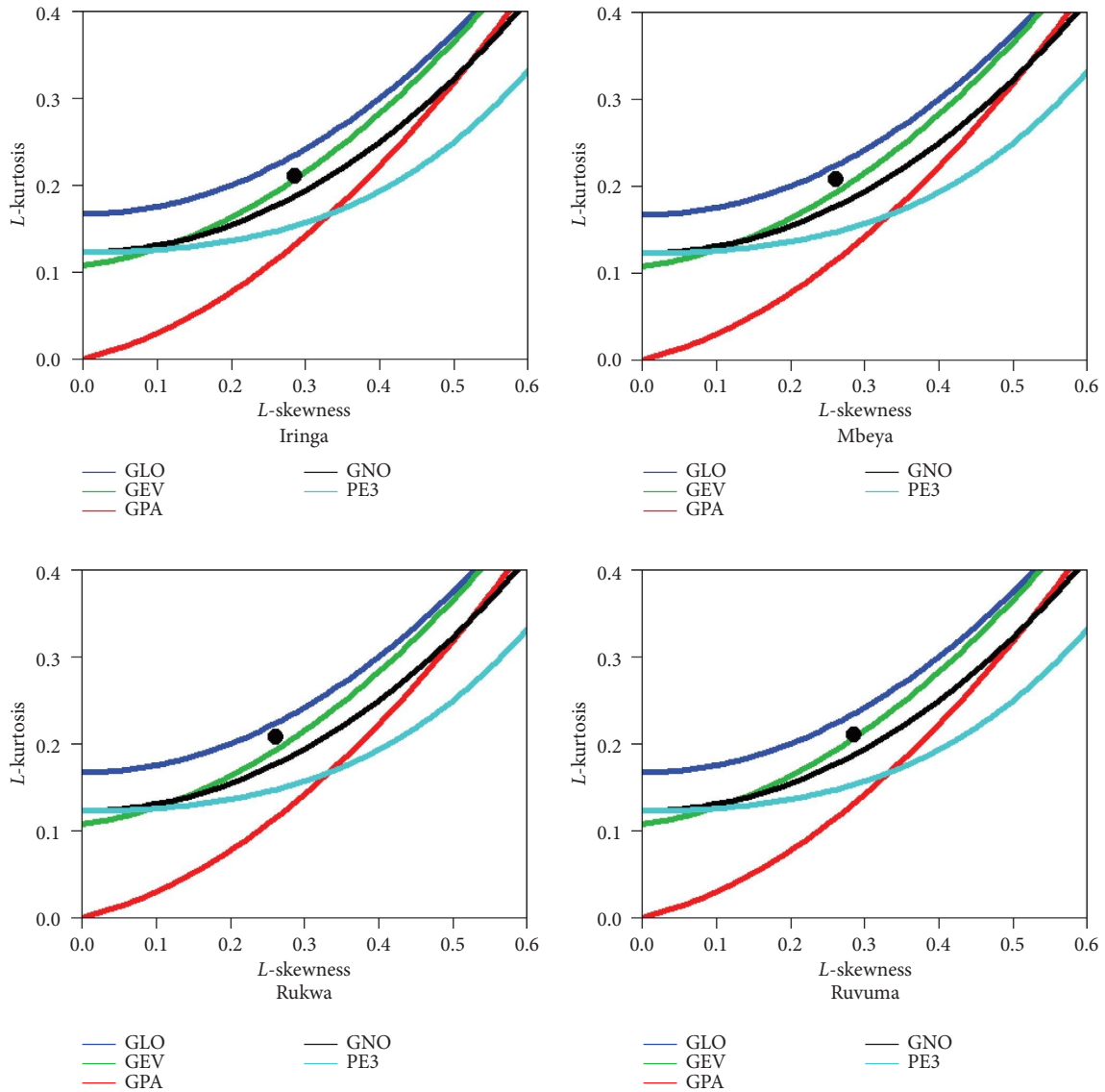


FIGURE 5: *L*-moments ratio diagrams.

Also, the Bayesian MCMC was used to provide the Bayesian parameters of the annual maxima rainfall using non-informative priors. When selecting noninformative priors for GEV modeling, we minimized prior influence to allow the data to play a key role in shaping the posterior distribution [57]. For the location parameter μ , we deployed a Normal distribution that is broad and symmetric distribution centered at 0. The scale parameter σ , was modeled with a wide normal distribution to encompass the expected range of extreme values. Similarly, a normal distribution that permits a wide range of shapes, indicating minimal prior information about the tail behavior was used for the shape parameter ξ . Following the methodology outlined by [2, 57, 58], the joint density of our parameters α, μ , and ξ was represented as $f(\alpha, \mu, \xi) = f_{\alpha}(\alpha)f_{\mu}(\mu)f_{\xi}(\xi)$.

Thus for the noninformative priors we employed the following: $f_{\alpha}(\alpha) \sim N(0, 10000)$, $f_{\mu}(\mu) \sim N(0, 10000)$, and $f_{\xi}(\xi) \sim N(0, 10000)$. Here, $N(0, 10000)$ denotes a normal distribution with a mean of 0 and a variance of 10,000, ensuring a high variance to affirm the absence of external information.

Table 5 shows the outcomes for MLE, *L*-moments, and Bayesian posterior parameter estimations for the GEV over all four stations. We plotted the diagnostic plots for the GEV model for all three methods as shown in Figures 6–8. The model diagnostic plots for rainfall at Iringa exhibit a satisfactory fit for the GEV distribution across all methods. Although we conducted similar diagnostic plots for the other stations, we have not included them in this study. Nevertheless, all important diagnostic plots including quantile

plot, the probability plot, density plot, and the return level plot present compelling evidence that GEV distribution is an appropriate model that fits the block maxima data in all stations.

From Table 5, the MLE shows that the shape parameter ξ , was estimated to be 0.249 for Iringa, 0.089 for Mbeya, -0.06 for Rukwa, and 0.0199 for Ruvuma. The positive values for Iringa and Mbeya indicate that these two regions have an upper bound, while the negative value for Rukwa indicates a lower bound. The value for Ruvuma is close to zero, indicating that the distribution for this region is approximately unbounded. Notably, the confidence interval for ξ contains 0, indicating that a Gumbel distribution can also provide a better fit for the data. This finding is supported by the profile confidence interval for ξ , which clearly shows that the value $\xi = 0$ falls well within the interval for all stations. The values have small difference when L -moments and Bayesian MCMC were used. The MCMC estimation results revealed the existence of an upper finite endpoint to the distribution due to positive shape parameter. This estimation is in agreement with the MLE approach when used for Iringa station. The diagnostic plots for GEV estimation indicated in Figures 6–8 shows in all methods the empirical quantiles are linearly related with model quantiles. In both methods simulated data are right-distributed and affected by extreme values.

3.6. Trace Plots and Posterior Densities Analysis. Trace plots and posterior densities obtained from Bayesian MCMC analysis for the GEV distribution provide valuable insights into the parameters of extreme rainfall distribution in the regions. According to Figure 9, the trace plots show the convergence and mixing of the MCMC chains, indicating reliable parameter estimation. The absence of a trace plot for the shape parameter (ξ), indeed shows that the estimated shape parameter is close to zero. This indicates that the data may also be well-fitted by the Gumbel distribution, which is a special case of the GEV distribution with a shape parameter of zero. The reason behind this behavior is that the shape parameter determines the tail behavior of the distribution. When the estimated shape parameter is close to zero, it suggests that the tails of the data are not significantly heavy or light compared to the Gumbel distribution, which has a standard exponential tail. Therefore, the model is effectively converging towards the Gumbel distribution, and the trace plot for the shape parameter may not be displayed.

The posterior densities of the location parameter (μ), and scale parameter (σ) provide information about the uncertainty associated with these parameters. Based on the trace plots and well-defined posterior densities, we can conclude that the MCMC algorithm has successfully explored the parameter space and converged to stable estimates for μ , σ , and ξ . This indicates that the Bayesian MCMC method is effective in estimating the parameters of the GEV distribution for extreme rainfall in Iringa. Similarly, fairly

good trace plots and posterior densities for the GEV distribution parameters were obtained for other stations but not presented in this work.

3.7. Rainfall Return Levels for Different Return Periods. The return level estimates obtained using different methods provide important insights into the potential intensity of extreme rainfall events in Iringa. The 100-year return level estimated using maximum likelihood estimation (MLE), is projected to be 128.45 mm of annual daily maximum rainfall, with a confidence interval ranging from 54.56 mm to 202.344 mm. This implies that, without intervention to mitigate climate change, Iringa could experience intense rainfall events of such magnitudes once every hundred years. The Bayesian MCMC method, which incorporates uncertainty quantification, provides a slightly lower estimate for the 100-year return level at 91.83 mm, with a wider confidence interval ranging from 188.40 mm to 584.76 mm. This suggests that the Bayesian approach acknowledges the inherent uncertainties in the estimation process and presents a broader range of possible outcomes. Comparatively, the L -moments method estimates the 100-year return level at 81.20 mm, with a narrower confidence interval ranging from 124.74 mm to 194.38 mm. These variations in estimates across methods highlight the importance of considering different approaches and their associated uncertainties in assessing extreme rainfall patterns. As the return period increases, both the estimated return levels and the width of the confidence intervals tend to increase, indicating a higher likelihood of more intense rainfall events occurring over longer periods. Table 3 presents the estimations of return levels for different return periods in Iringa. Similar return levels for different return periods were made for other stations but were not presented in this study. These findings provide valuable information for understanding the potential risks associated with extreme rainfall in the regions and can aid in decision-making for infrastructure planning and climate adaptation strategies.

3.8. Performance of MLE, L -Moments, and Bayesian Methods. Table 6 presents the performance comparison of Bayesian MCMC, MLE, and L -moments, in estimating the parameters of the GEV distribution. The evaluation is based on two metrics, MAE (mean absolute error) and RMSE (root mean square error). The results show that the Bayesian MCMC method achieved the lowest MAE (0.312) and RMSE (0.243) values compared to the other two methods. This indicates that the Bayesian MCMC method provides estimates that are closer to the true values of the GEV distribution parameters, resulting in smaller prediction errors on average. The MLE method obtained slightly higher MAE (0.327) and RMSE (0.268) values compared to the Bayesian MCMC method, indicating slightly larger prediction errors. On the other hand, the L -moments method yielded the highest MAE (0.342) but had a lower RMSE (0.253) compared to the MLE

TABLE 5: MLE, Bayesian MCMC, and L -moments parameter estimates.

Region	MLE			Bayesian MCMC			L -moments		
	μ	σ	ξ	μ	σ	ξ	μ	σ	ξ
Iringa	50.31	9.072	0.249 (0.174) (-0.091, 0.5899)	48.92	9.446	0.298 (0.135)	50.212	9.333	0.218 (0.152)
Mbeya	48.15	11.19	0.089 (0.156) (-0.216, 0.3948)	51.27	12.45	0.172 (0.112)	50.41	10.74	0.137 (0.168)
Rukwa	66.81	14.24	-0.06 (0.142) (-0.338, 0.218)	50.72	11.1	0.258 (0.143)	50.61	9.9	0.184 (0.212)
Ruvuma	48.15	10.50	0.0199 (0.194) (-0.359, 0.3993)	52.99	9.94	0.255 (-0.079)	52.52	8.83	0.174 (0.188)

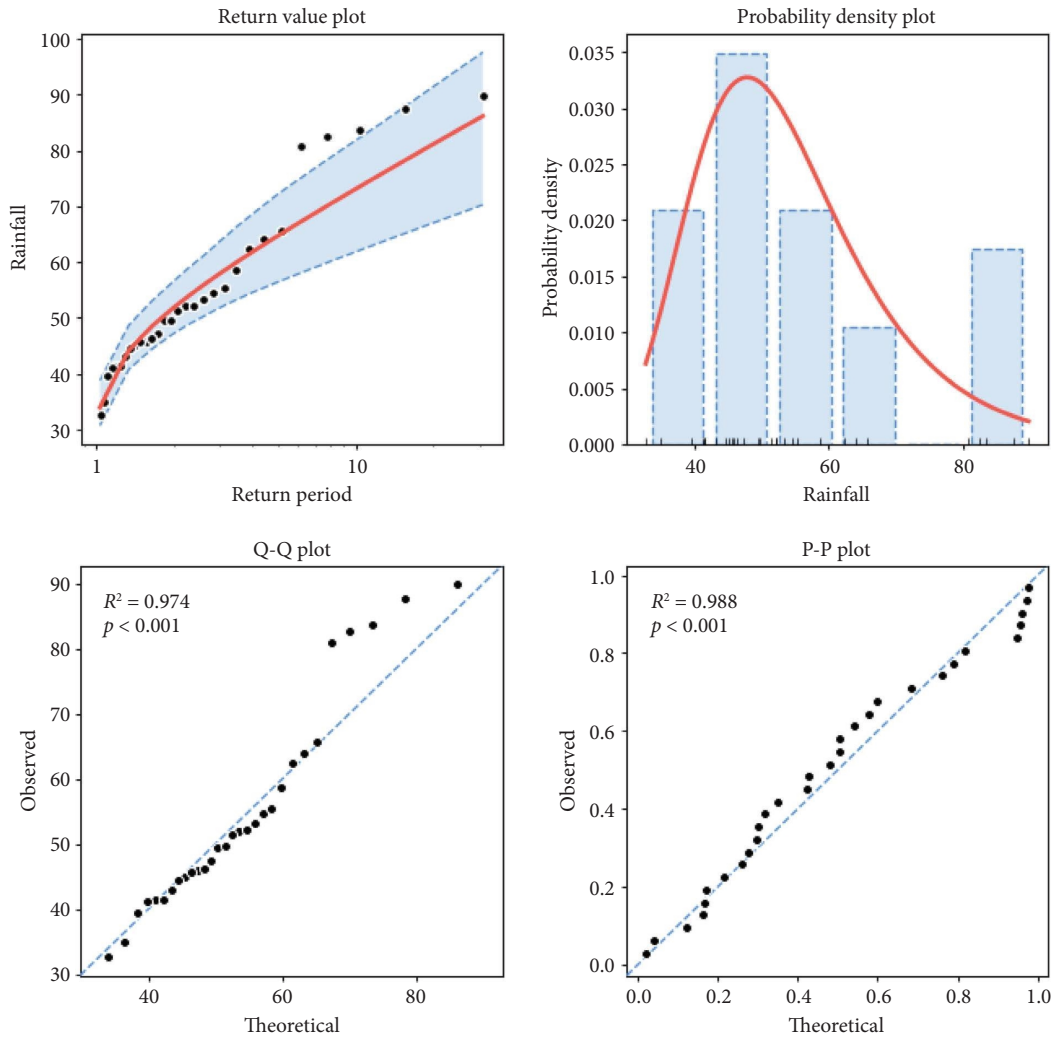


FIGURE 6: Diagnostic plots of the MLE method in Iringa station.

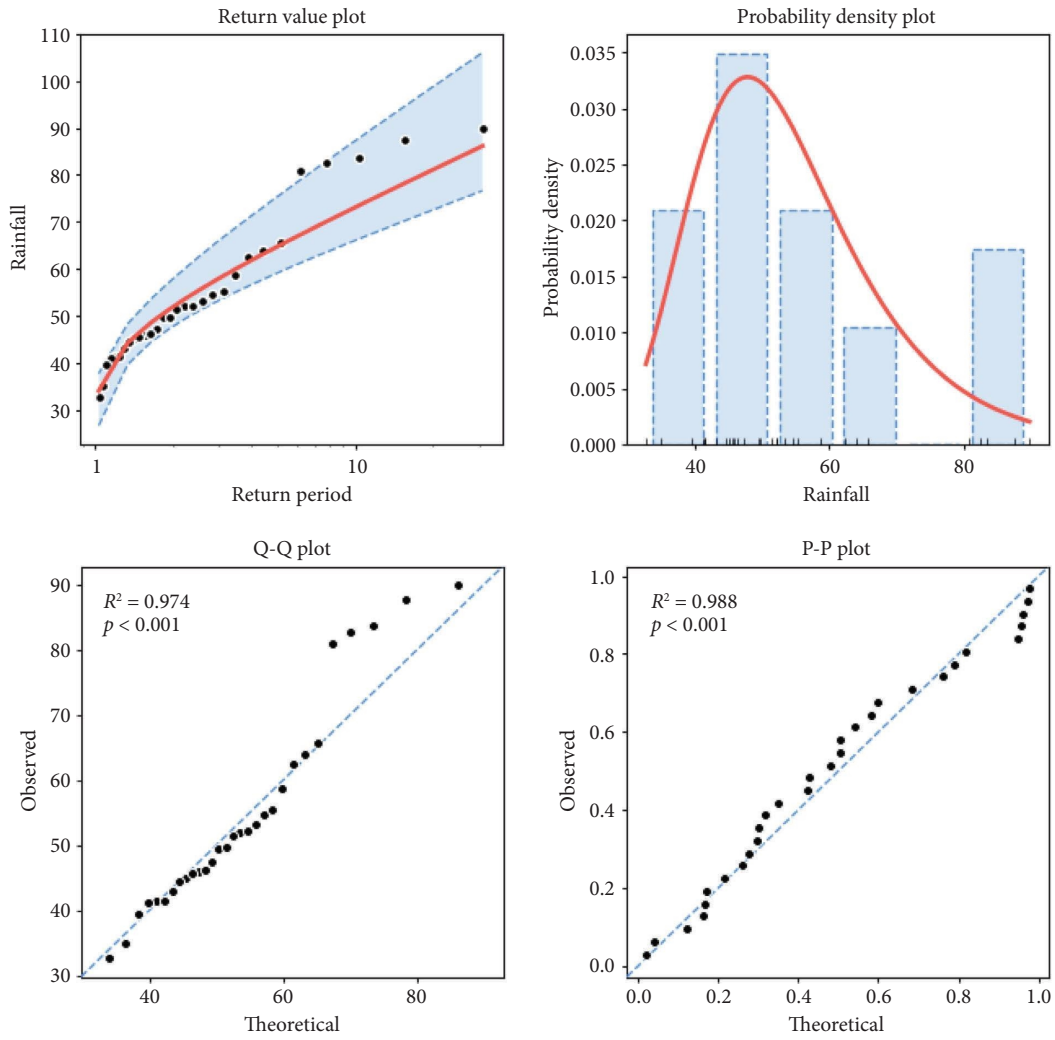


FIGURE 7: Diagnostic plots of Bayesian method in Iringa station.

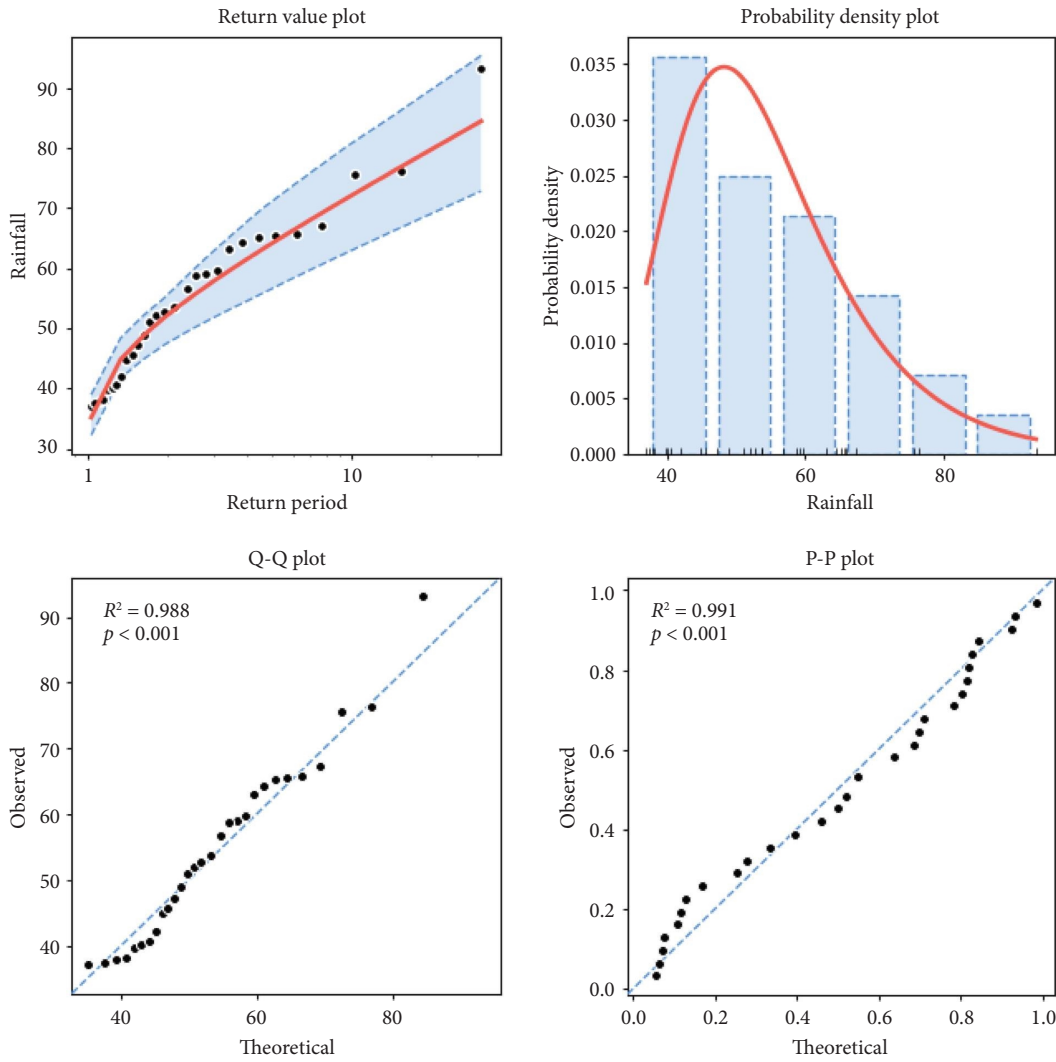


FIGURE 8: Diagnostic plots of L -moments method in Iringa station.

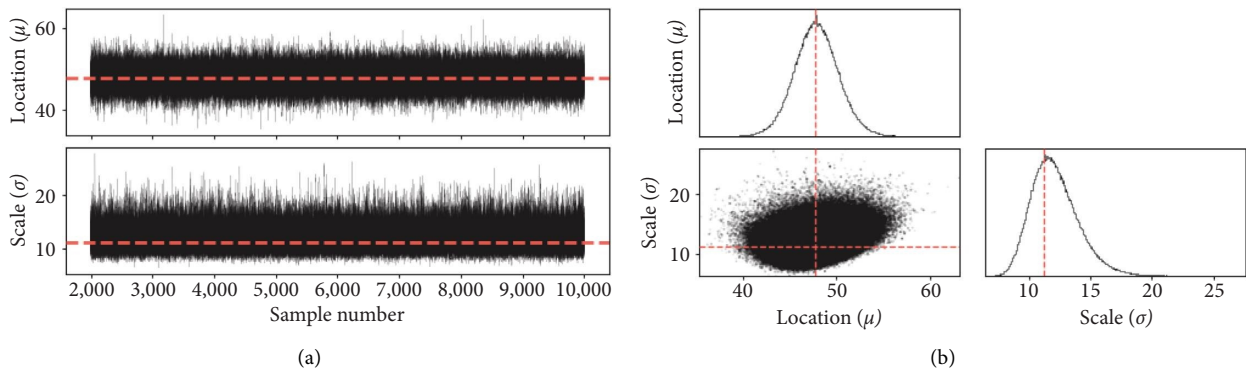


FIGURE 9: Trace plots and posterior densities of location μ , and scale σ parameters for Iringa station.

method. This suggests that the L -moments method has a higher average prediction error but exhibits less variability in the predictions. Based on these findings, we can conclude that the Bayesian MCMC method outperforms the MLE and

L -moments methods in estimating the parameters of the GEV distribution for the given rainfall data. It provides more accurate and precise parameter estimates, resulting in better predictions of extreme rainfall events.

TABLE 6: Performance comparison of MLE, L -moments, and Bayesian MCMC.

Methods	MAE	RMSE
Bayesian MCMC	0.312	0.243
MLE	0.327	0.268
L -moments	0.342	0.253

4. Discussion and Conclusion

The statistical analysis of yearly maximum daily rainfall at four stations revealed valuable insights into the rainfall patterns and characteristics in the Highlands Region. The results demonstrated significant differences in rainfall amounts among the stations, with Ruvuma receiving the highest amount and Rukwa experiencing the lowest. This suggests that the Ruvuma region may be more susceptible to flooding and related hazards, while Rukwa may face challenges related to water scarcity and drought during the dry seasons. The analysis of the annual maxima rainfall data using the Block Maxima method indicated a relatively stable pattern of extreme rainfall events over the past few decades, without a clear trend. These findings imply that the risk of flooding or other related hazards in these regions may remain constant over time. However, it is crucial to continue monitoring weather patterns and extreme events to ensure the implementation of appropriate measures for protecting people and infrastructure from the impacts of climate change.

The L -moments ratios provided further insights into the characteristics of extreme rainfall. The values of λ_1 (mean rainfall depth) showed that Ruvuma had the highest average rainfall depth, while Rukwa had the lowest spread (scale) represented by λ_2 . The kurtosis (τ_4) and skewness (τ_3) values indicated the heaviness of the tails and the departure from symmetry, respectively. Iringa exhibited a slightly right-skewed distribution and relatively high kurtosis. The λ_5 values reflected the departure from symmetry in the upper tail, with Iringa having a thicker upper tail. Overall, the L -moments ratios suggested that the extreme rainfall in these regions had moderate peakedness and varying degrees of tail heaviness and asymmetry.

The L -moments ratio diagrams played a crucial role in identifying the best-fit distribution for annual extreme rainfall in each region. The analysis showed that the GEV distribution provided the best fit for extreme rainfall in all four stations, as indicated by Figure 5. The finding was further supported by the parameter estimation using maximum likelihood estimation (MLE), L -moments, and Bayesian Markov Chain Monte Carlo (MCMC) methods. This finding corroborates previous research by [17, 59] highlighting the effectiveness of the GEV distribution in characterizing extreme events.

The GEV parameter estimates revealed important characteristics of the rainfall distributions in the study area. The shape parameter (ξ) values indicated whether the distributions had upper or lower bounds or were unbounded. Iringa and Mbeya exhibited positive shape parameters, suggesting an upper bound, while Rukwa had a negative shape parameter, indicating a lower bound.

Ruvuma had a shape parameter close to zero, implying an approximately unbounded distribution. The confidence intervals for ξ contained the value of 0, indicating that a Gumbel distribution could also provide a reasonable fit for the data. This finding is supported by the absence of a trace plot representing the shape parameter (ξ), indicating that the estimated shape parameter is near zero. This suggests that the data could be well-suited for fitting the Gumbel distribution, which is a special case of the GEV distribution with a shape parameter of zero. This behavior stems from the fact that the shape parameter influences the tail characteristics of the distribution. When the estimated shape parameter is close to zero, it implies that the data's tails are not significantly heavier or lighter compared to the Gumbel distribution, which has a standard exponential tail.

Additionally, the estimation of return levels based on the GEV distribution provided valuable information on the magnitude and frequency of extreme rainfall events in the Highlands Region. The results of this study provide valuable insights into extreme rainfall patterns and their impacts in the Southern Highlands Region of Tanzania. However, it is crucial to acknowledge that each region may possess unique characteristics that influence the occurrence and consequences of extreme rainfall. As such, caution should be exercised when generalizing these findings to other locations. Nonetheless, the methodology and analysis employed in this study can be universally applied as a framework for deducing the best-fit distribution that characterizes extreme rainfall behaviors. By utilizing a rigorous statistical approach and considering relevant climatic factors, this study has contributed to understanding the underlying patterns of extreme rainfall events in Southern Highlands Region of Tanzania. The findings can serve as a basis for similar studies in different regions, adapting the methodology to the specific local context.

Furthermore, the identification of potential links between extreme rainfall and climate change in Tanzania highlights the need for further research and analysis. To achieve a more comprehensive understanding of the universal aspects of extreme rainfall patterns and their implications for climate change adaptation worldwide, it is essential to undertake broader studies encompassing multiple regions and considering a longer time span. Such studies could explore the influence of global climate change phenomena on regional extreme rainfall events, facilitating the development of more effective adaptation and mitigation strategies.

Data Availability

The corresponding author will provide data used in this research upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The authors acknowledge the Tanzania Meteorological Authority (TMA) for providing the data used in this study. This research was financially facilitated by Mkwawa University College of Education (MUCE) through Higher Education for Economic Transformation (HEET) Project.

References

- [1] S. Solomon, D. Qin, M. Manning et al., "Ipcc fourth assessment report (ar4)," *Climate Change*, vol. 374, 2007.
- [2] T. Ahmad, I. Ahmad, I. A. Arshad, and N. Bianco, "A comprehensive study on the bayesian modelling of extreme rainfall: a case study from Pakistan," *International Journal of Climatology*, vol. 42, no. 1, pp. 208–224, 2022.
- [3] Y. Xu, V. Ramanathan, and D. G. Victor, *Global Warming Will Happen Faster than We Think*, Springer Nature, Berlin, Germany, 2018.
- [4] H.-O. Pörtner, D. C. Roberts, V. Masson-Delmotte et al., "Technical summary. IPCC special report on the ocean and cryosphere in a changing climate," 2019, https://www.ipcc.ch/site/assets/uploads/sites/3/2019/11/04_SROCC_TS_FINAL.pdf.
- [5] N. Singh and U. Mina, "An assessment of methane emission from the cng cylinder testing stations in Delhi and its implication for global warming," *Environmental Monitoring and Assessment*, vol. 195, no. 3, p. 369, 2023.
- [6] S. Leng, A. Huete, J. Cleverly et al., "Assessing the impact of extreme droughts on dryland vegetation by multi-satellite solar-induced chlorophyll fluorescence," *Remote Sensing*, vol. 14, no. 7, p. 1581, 2022.
- [7] W. Z. W. Zin, A. Aziz Jemain, K. Ibrahim, and S. Jamaludin, "Bayesian changepoint analysis of the extreme rainfall events," *Journal of Mathematics and Statistics*, vol. 8, no. 1, pp. 85–91, 2012.
- [8] I. Ahmad, T. Ahmad, and I. M. Almanjahie, "Modelling of extreme rainfall in Punjab: Pakistan using bayesian and frequentist approach," *Applied Ecology and Environmental Research*, vol. 17, no. 6, pp. 13729–13748, 2019.
- [9] E. Smith, *Bayesian modelling of extreme rainfall data*, Ph.D. thesis, University of Newcastle upon Tyne, England, UK, 2005.
- [10] T. F. Stocker, D. Qin, G.-K. Plattner et al., "Climate change 2013: the physical science basis," *contribution of working group i to the fifth assessment report of ipcc the intergovernmental panel on climate change*, Cambridge University Press, Cambridge, UK, 2014.
- [11] S. Twisa and M. F. Buchroithner, "Seasonal and annual rainfall variability and their impact on rural water supply services in the wami river basin, Tanzania," *Water*, vol. 11, no. 10, p. 2055, 2019.
- [12] R. K. Pachauri and A. Reisinger, *Synthesis Report Fifth Assessment Report Of The Intergovernmental Panel On Climate Change*, IPCC, Geneva, Switzerland, 2007.
- [13] A. Georgakakos, P. Fleming, M. Dettinger et al., "Chapter 3 water resources," in *Climate Change Impacts in the United States: The Third National Climate Assessment*, J. M. Melillo, T. C. Richmond, and G. W. Yohe, Eds., pp. 69–112, US Global Change Research Program, Washington, DC, USA, 2014.
- [14] A. E. Vrochidou, I. K. Tsanis, M. G. Grillakis, and A. G. Koutroulis, "The impact of climate change on hydro-meteorological droughts at a basin scale," *Journal of Hydrology*, vol. 476, pp. 290–301, 2013.
- [15] T. M. A. Book, *Statement on the Status of tanzania Climate in 2019*, Tanzania Meteorological Authority, Dar es Salaam, Tanzania, 2019.
- [16] T. M. A. Book, *Statement on the Status of tanzania Climate in 2022*, Tanzania Meteorological Authority, Dar es Salaam, Tanzania, 2022.
- [17] T. J. Ngailo, J. Reuder, E. Rutalebwa, S. Nyimvua, and M. Mesquita, "Modelling of extreme maximum rainfall using extreme value theory for Tanzania," *International Journal of Scientific and Innovative Mathematical Research*, vol. 4, no. 3, pp. 34–45, 2016.
- [18] L. B. Chang'a, A. L. Kijazi, K. B. Mafuru et al., "Assessment of the evolution and socio-economic impacts of extreme rainfall events in october 2019 over the east africa," *Atmospheric and Climate Sciences*, vol. 10, no. 3, pp. 319–338, 2020.
- [19] A. A. Lukali, S. E. Osima, Y. Lou, and K. H. Kai, "Assessing the impacts of climate change and variability on maize (zea mays) yield over Tanzania," *Atmospheric and Climate Sciences*, vol. 11, no. 3, pp. 569–588, 2021.
- [20] W. Sanusi, S. Side, and K. A. Muhammad, "Probability distribution modeling of extremes rainfall series in makassar city using the l-moments method," *Asian Journal of Applied Sciences*, vol. 3, no. 5, 2015.
- [21] O. I. Martins, B. O. Sam, and S. N. David, "Classical and bayesian Markov chain Monte Carlo (mcmc) modeling of extreme rainfall (1979-2014) in makurdi, Nigeria," *International Journal of Water Resources and Environmental Engineering*, vol. 7, no. 9, pp. 123–131, 2015.
- [22] Z. Li, Z. Li, W. Zhao, and Y. Wang, "Probability modeling of precipitation extremes over two river basins in northwest of China," *Advances in Meteorology*, vol. 2015, Article ID 374127, 13 pages, 2015.
- [23] Y. Li, W. Cai, and E. P. Campbell, "Statistical modeling of extreme rainfall in southwest western Australia," *Journal of Climate*, vol. 18, no. 6, pp. 852–863, 2005.
- [24] S. S. Eslamian and H. Feizi, "Maximum monthly rainfall analysis using l-moments for an arid region in isfahan province, Iran," *Journal of Applied Meteorology and Climatology*, vol. 46, no. 4, pp. 494–503, 2007.
- [25] K. K. Kar, S.-K. Yang, J.-H. Lee, and F. K. Khadim, "Regional frequency analysis for consecutive hour rainfall using l-moments approach in jeju island, korea," *Geo-environmental Disasters*, vol. 4, no. 1, p. 18, 2017.
- [26] L. Marx and J. L. Kinter, *Estimating the Representation of Extreme Precipitation Events in Atmospheric General Circulation Models Using L-Moments*, Center for Ocean-Land-Atmosphere Studies, Virginia, VA, USA, 2007.
- [27] B. Lin, G. M. Bonnin, D. L. Martin, T. Parzybok, M. Yekta, and D. Riley, "Regional frequency studies of annual extreme precipitation in the United States based on regional l-moments analysis," in *World Environmental and Water Resource Congress 2006: Examining the Confluence of Environmental and Water Concerns*, pp. 1–11, ASCE, Reston, VA, USA, 2006.
- [28] C. S. Ngongondo, C.-Y. Xu, L. M. Tallaksen, B. Alemaw, and T. Chirwa, "Regional frequency analysis of rainfall extremes in southern Malawi using the index rainfall and l-moments

- approaches,” *Stochastic Environmental Research and Risk Assessment*, vol. 25, no. 7, pp. 939–955, 2011.
- [29] J. R. Hosking, “L-moments: analysis and estimation of distributions using linear combinations of order statistics,” *Journal of the Royal Statistical Society: Series B*, vol. 52, no. 1, pp. 105–124, 1990.
- [30] S. Feng, S. Nadarajah, and Q. Hu, “Modeling annual extreme precipitation in China using the generalized extreme value distribution,” *Journal of the Meteorological Society of Japan*, vol. 85, no. 5, pp. 599–613, 2007.
- [31] S. G. Coles and M. J. Dixon, “Likelihood-based inference for extreme value models,” *Extremes*, vol. 2, no. 1, pp. 5–23, 2000.
- [32] R. Deidda and M. Puliga, “Performances of some parameter estimators of the generalized pareto distribution over rounded-off samples,” *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 34, no. 10–12, pp. 626–634, 2009.
- [33] S. D. Grimshaw, “Computing maximum likelihood estimates for the generalized pareto distribution,” *Technometrics*, vol. 35, no. 2, pp. 185–191, 1993.
- [34] B. C. Bates and E. P. Campbell, “A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling,” *Water Resources Research*, vol. 37, no. 4, pp. 937–947, 2001.
- [35] J. He, P. Feng, B. Wang et al., “Centennial annual rainfall pattern changes show an increasing trend with higher variation over northern Australia,” *Journal of Hydrometeorology*, vol. 23, no. 8, pp. 1333–1349, 2022.
- [36] S. M. S. Maliondo, E. J. Mpeti, and J. Olson, *Climate Change and Food Security in Tanzania: An Analysis of Current Knowledge and Research Gaps and Recommendations for a Research Agenda*, Ohio State University and Sokoine University of Agriculture, Columbus, OH, USA, 2012.
- [37] R. von Mises, “On the mathematical theory of the storage of grains,” *Journal of Applied Mathematics and Mechanics*, vol. 34, pp. 249–256, 1954.
- [38] A. F. Jenkinson, “The frequency distribution of the annual maximum (or minimum) values of meteorological elements,” *Quarterly Journal of the Royal Meteorological Society*, vol. 81, no. 348, pp. 158–171, 1955.
- [39] W. Sanusi, U. Mulbar, H. Jaya, P. Purnamawati, and S. Side, “Modeling of rainfall characteristics for monitoring of the extreme rainfall event in makassar city,” *American Journal of Applied Sciences*, vol. 14, no. 4, pp. 456–461, 2017.
- [40] N. Millington, S. Das, and S. P. Simonovic, “The comparison of gev, log-pearson type 3 and gumbel distributions in the upper thames river watershed under global climate models,” 2011, <https://ir.lib.uwo.ca/wrrrr/40/>.
- [41] R. Modarres, “Regional dry spells frequency analysis by l-moment and multivariate analysis,” *Water Resources Management*, vol. 24, no. 10, pp. 2365–2380, 2010.
- [42] E. A. H. Elamir and A. H. Seheult, “Trimmed l-moments,” *Computational Statistics and Data Analysis*, vol. 43, no. 3, pp. 299–314, 2003.
- [43] R. D. Reiss and M. Thomas, *Statistical Analysis of Extreme Values*, Birkhäuser Verlag, Basel, Switzerland, 2nd edition, 2001.
- [44] J. R. M. Hosking and J. R. Wallis, *Regional Frequency Analysis*, Cambridge University Press, Cambridge, UK, 1997.
- [45] E. Singirankabo and E. Iyamuremye, “Modelling extreme rainfall events in kigali city using generalized pareto distribution,” *Meteorological Applications*, vol. 29, no. 4, 2022.
- [46] A. Gelman and D. Nolan, *Teaching Statistics: A Bag of Tricks*, Oxford University Press, Oxford, UK, 2002.
- [47] A. Gelman, J. Carlin, H. Stern, D. Dunson, V. Aki, and D. Rubin, *Bayesian Data Analysis*, Chapman and Hall/CRC, Boca Raton, FL, USA, 3rd edition, 2013.
- [48] M. O. Isikwue, S. B. Onoja, and D. S. Naakaa, “Classical and bayesian Markov chain Monte Carlo (mcmc) modelling of extreme rainfall in makurdi, Nigeria,” *International Journal of Water Resources and Environmental Engineering*, vol. 7, no. 9, pp. 123–131, 2015.
- [49] B. W. Yap and C. H. Sim, “Comparisons of various types of normality tests,” *Journal of Statistical Computation and Simulation*, vol. 81, no. 12, pp. 2141–2155, 2011.
- [50] B. Yazici and S. Yolacan, “A comparison of various tests of normality,” *Journal of Statistical Computation and Simulation*, vol. 77, no. 2, pp. 175–183, 2007.
- [51] J. H. Lopez, “The power of the adf test,” *Economics Letters*, vol. 57, no. 1, pp. 5–10, 1997.
- [52] R. Mushtaq, *Augmented dickey Fuller Test*, SSRN, New York, NY, USA, 2011.
- [53] A. H. Abdalla, K. H. Kai, S. A. Khamis et al., “The influence of climate change and variability on spatio-temporal rainfall and temperature distribution in zanzibar,” *Atmospheric and Climate Sciences*, vol. 13, no. 02, pp. 282–313, 2023.
- [54] M. Naghettini, *Fundamentals of Statistical Hydrology*, Springer, Berlin, Germany, 2017.
- [55] P. K. Sen, “Estimates of the regression coefficient based on kendall’s tau,” *Journal of the American Statistical Association*, vol. 63, no. 324, pp. 1379–1389, 1968.
- [56] H. Zuberi, Y. Lou, and A. Ojara Moses, “Spatial-temporal analysis of drought characteristics in Tanzania from 1978 to 2018,” *North American Academic Research*, vol. 5, no. 1, 2022.
- [57] T. Ahmad, I. Ahmad, I. A. Arshad, and I. M. Almanjahie, “An efficient bayesian modelling of extreme winds in the favour of energy generation in Pakistan,” *Energy Reports*, vol. 9, pp. 2980–2992, 2023.
- [58] S. Coles and J. Tawn, “Bayesian modelling of extreme surges on the UK east coast,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 363, pp. 1387–1406, 2005.
- [59] E. Iyamuremye, S. Wanyonyi, and D. Mbete, “Analysis and modelling of extreme rainfall: a case study for dodoma, Tanzania,” *Asian Journal of Mathematics and Statistics*, vol. 3, 2019.