International Journal of Advanced Technology and Engineering Exploration, Vol 12(123)
ISSN (Print): 2394-5443 ISSN (Online): 2394-7454
http://dx.doi.org/10.19101/IJATEE.2024.111101018

Research Article

Identifying and validating optimal probability distributions for improved return period estimation of extreme events

Mohammad Amir Syahmi[1], Zahrahtul Amani Zakaria[1, 2*] and Nor Aida Mahiddin[1, 2]
Faculty of Computing and Informatics, Universiti Sultan Zainal Abidin, Kampus Besut, 22200 Besut, Terengganu, Malaysia[1]
East Coast Environmental Research Institute (ESERI), Universiti Sultan Zainal Abidin, Kampus Gong Badak, 21300, Kuala Terengganu, Terengganu, Malaysia[2]

Received: 13-June-2024; Revised: 12-February-2025; Accepted: 18-February-2025

Abstract
This engineering-focused study analyzes annual maximum daily rainfall data from the Department of Irrigation and Drainage (DID) Kemaman station in Terengganu, Malaysia, to enhance flood risk management and infrastructure resilience in the region. The study aims to identify the most effective probability distribution for modeling extreme rainfall and to estimate return periods for critical events. Using the robust L-moment method, various distributions were rigorously tested and initially selected through the L-moment ratio diagram (LMRD). The four-parameter Kappa distribution (K4D) emerged as the best fit, as determined by the mean absolute deviation index (MADI) and the mean squared deviation index (MSDI). The validated model enabled the estimation of return periods, indicating that a 2-year event corresponds to 188.66 mm of rainfall, while a 100-year event is expected to reach 475.48 mm. These quantitative insights are essential for designing durable, flood-resilient infrastructure, ensuring that regional development is both sustainable and adaptable to increasing weather variability.

Keywords
Extreme rainfall analysis, Probability distributions, Return period estimation, Flood risk management, L-moment method, Four-parameter kappa distribution.

1.Introduction
Global communities are increasingly susceptible to natural disasters, especially floods intensified by climate change and urbanization. These disasters, often triggered by significant rainfall events in poorly drained or flood-prone areas. This necessitates a deep understanding of rainfall variability and distribution. Historically, frequency analysis has played a crucial role in modeling these events, relying heavily on the accuracy of probability distributions to predict and mitigate potential impacts on public safety and economic stability [1].

The challenges in predicting extreme weather events lie in the variability and complexity of rainfall data. Traditional frequency analysis methods, although useful, often fall short when dealing with outliers and non-stationary data, which are prevalent in climate datasets [2].

This research addresses these limitations by utilizing the L-moment method, which offers a more robust and less outlier-sensitive approach to parameter estimation in hydrological studies.

L-moments are a powerful tool in hydrological data analysis, providing robust measures to summarize a data distribution's shape, scale, and location [3]. One of the primary benefits of employing the L-moment approach for parameter estimation in hydrological studies is its reduced sensitivity to outliers, leading to more stable and reliable parameter estimates [4]. The L-moment method has been successfully applied in flood frequency analysis in many countries, including Malaysia [5, 6], China [7], India [8], Norway [9], Iran [10, 11], Poland [12], and Turkey [13]. This widespread utilization highlights its critical role in hydrology. Implementing this methodology in regions with adequate data is essential for comprehensive analysis.

*Author for correspondence

The primary objective of this study is to apply the L-moment method to analyze annual maximum rainfall data from the Department of Irrigation and Drainage (DID) Kemaman station in Terengganu, Malaysia. This analysis aims to identify the most suitable probability distribution and estimate return periods for extreme precipitation events.

This paper contributes to the field by applying the L-moment technique in distribution parameter estimation and return period estimation, demonstrating its effectiveness in improving flood risk management and infrastructure resilience. It provides a comprehensive analysis of the best-fit probability distribution for extreme rainfall events, facilitating more accurate predictions and preparedness strategies.

The paper is structured as follows: Section 2 reviews related studies, Section 3 describes the methodology, Section 4 presents the results, Section 5 discusses the findings, and Section 6 concludes the paper.

2.Literature review

Understanding the variability and distribution of rainfall is critical for flood risk management. Numerous studies have focused on analyzing extreme rainfall events using different methodologies. This section discusses key studies in the field, highlighting their methodologies, results, advantages, and limitations.

2.1Background of L-moments

L-moments, powerful statistical tools for describing the shape of probability distributions, have been widely adopted in hydrological and meteorological studies for extreme event analysis. This review traces the background of L-moments, highlighting significant milestones and advancements that have enhanced their application in various fields.

L-moments were introduced by Hosking as an alternative to conventional moments for summarizing the characteristics of probability distributions. Traditional moments, such as mean, variance, skewness, and kurtosis, are sensitive to outliers and can be unstable for distributions with heavy tails or extreme values. L-moments, derived from linear combinations of order statistics, provide robust estimates that are less influenced by outliers and can describe the distribution shape more accurately [3]. The initial application of L-moments focused on hydrology, particularly for flood frequency analysis. Hosking and Wallis expanded the theory, providing a comprehensive framework for regional frequency analysis (RFA) using L-moments. This approach allowed for the effective pooling of data from multiple sites, improving the reliability of extreme rainfall and flood event estimates [4].

2.2L-moment ratio diagram (LMRD)

LMRDs have become an essential tool for hydrologists and statisticians. They are used to compare and select appropriate probability distributions for RFA and other applications. For example, LMRDs are employed to identify the best-fit distribution for rainfall, flood frequency, and other hydrological data.

Vogel and Fennessey demonstrated that LMRDs are superior to conventional moment ratio diagrams (MRDs) in hydrology. They found that LMRDs provided clearer insights into the distributional properties of daily streamflow data compared to traditional moment diagrams [14]. Peel et al. illustrated the utility of LMRDs for selecting regional probability distributions. They showed that LMRDs, combined with heterogeneity tests, effectively discriminate between distributions in both homogeneous and heterogeneous regional samples [15]. Haddad applied LMRDs to analyze temperature data in New South Wales, Australia, and found that LMRDs allowed for easy comparison of the fit of multiple distributions across several stations [16]. Similarly, Ouarda et al. used LMRDs to assess the fit of probability distributions for wind speed data in the United Arab Emirates, highlighting their effectiveness in various climatic contexts [17]. Hosking extended L-moments to trimmed L-moments for analyzing heavy-tailed distributions, proposing a trimmed LMRD as an enhancement for identifying distributions suited to extreme events. This approach further solidifies the utility of LMRDs in extreme value theory [18]. Bobée et al. emphasized the complementary nature of LMRDs with traditional MRDs, suggesting that integrating both can provide a more comprehensive understanding of distributional properties. Their study highlighted the adaptability of LMRDs in various hydrological applications [19].

LMRD have proven to be a robust and reliable tool for selecting and evaluating probability distributions, particularly in hydrology and environmental sciences. Their ability to provide nearly unbiased estimates and facilitate easy comparison of distributions makes them indispensable for statistical analysis of extreme events. Future research should explore improving

LMRDs' applicability to highly heterogeneous datasets and integrating them with advanced machine learning and statistical techniques.

2.3Integration of L-moments with other methods in hydrology

The linear combination of H-statistics (LH-moments), an extension of L-moments, is used to better characterize the upper part of distributions, which is crucial for predicting extreme events like floods. Murshed et al. formulated LH-moments for several distributions useful in hydrology, including the Generalized Gumbel and Kappa distributions, showing that LH-moments reduce the undue influences of small sample sizes on large return period event estimates [20]. Partial L-moments (PL-moments) are another extension designed to handle censored data, or to focus on specific data ranges. Junzhen and Songbai demonstrated that PL-moments provide better statistical properties for flood frequency analysis by improving the estimation accuracy of design floods, especially for large return periods [21]. Mudholkar and Hutson developed Linear combination of quantile functions (LQ)-moments, analogs of L-moments, which replace expectations with functionals inducing quick estimators like the median and trimean. LQ-moments offer simpler and more robust estimation procedures, especially useful in analyzing extreme events in hydrology [22].

Sung et al. proposed a distribution-free method to apply L-moments to nonstationary hydrological processes [23]. Their method adjusts traditional L-moment-based RFA for nonstationary conditions, providing more accurate return level estimates under changing climatic conditions [23]. Bahmani et al. integrated L-moments with the hydrologic engineering center – hydrologic modeling system (HEC-HMS) hydrological model to determine design flood hydrographs. Their approach used L-moments to estimate maximum precipitation for different return periods, which were then input into the hydrological model to simulate flood hydrographs, demonstrating the utility of combining statistical and physical modeling techniques [24]. Koutsoyiannis introduced K-moments, which combine the advantages of classical and L-moments to provide reliable estimates and describe high-order statistics. This integration helps in characterizing both marginal and joint distributions of stochastic processes, enhancing the robustness of hydrological modelling [25].

2.4Recent L-moments methodology application and advancement

Agbonaye et al. compared the performance of the method of moments (MoM) and the L-moment method in identifying suitable probability distribution models for extreme rainfall data from Nigeria [26]. The study found that the L-moment method was more effective for the generalized pareto (GPA) distribution in five out of ten locations, highlighting the method's superior accuracy in parameter estimation for hydrology model design [26]. Vivekanandan conducted an analysis of rainfall estimation in Dahanu, India, using the L-moment method among other parameter estimation methods. The study concluded that the generalized extreme value (GEV) distribution fitted with the L-moment method provided the most reliable parameters for hydrological data in Dahanu [27]. In another study, Vivekanandan compared various extreme value distributions using the L-moment method for estimating annual 1-day maximum rainfall. The study reinforced the effectiveness of the L-moment method in providing robust parameter estimates for extreme rainfall events [28]. Sanusi et al. [29] applied the L-moment method to determine the type of distribution for rainfall data in South Sulawesi, Indonesia. The study used indicators such as root mean square error (RMSE) and mean absolute error (MAE) to select the best-fit distributions, demonstrated the method's practical utility in hydrological applications [29]. Guaykarernpanishk et al. explored the partial l-moments (PL-Moments) for the four-parameter kappa distribution (K4D), demonstrating its application for hydrological extremes from censored data. The study derived formulas for parameter estimation using PL-Moments, enhancing the accuracy of extreme event forecasts [30]. Anghel et al. conducted a comparative analysis of two-parameter probability distributions using L-moments and LH-moments methods. They found that the log-normal distribution (NOM) provided the best fit for the studied extreme events, with L-moments showing robust performance in approximating statistical indicators [31]. Chang et al. used L-moments for RFA of radar rainfall in Taiwan, clustering grids into homogeneous sub-regions. The study highlighted the effectiveness of L-moments in producing detailed regional rainfall estimates, crucial for large-scale hydrological applications [32].

The reviewed recent literature highlights the significant advancements and practical applications of the L-moments method in modeling extreme rainfall events. The L-moments method consistently

demonstrates superior performance in parameter estimation and model fitting across diverse geographical and hydrological scenarios. Its robust application in RFA further underscores its utility in designing resilient infrastructure and managing flood risks effectively. Future research should continue to integrate advanced statistical methods and real-time data to refine these models and enhance their predictive accuracy.

2.5Limitations and research gaps of L-moments

L-moments, since their introduction, have been widely used in hydrology and other fields for summarizing probability distributions. However, despite their robust performance and advantages over conventional moments, L-moments exhibit certain limitations and study gaps. This review discusses these limitations and identifies areas.

2.5.1Sensitivity to data quality

L-moments, like other statistical methods, are sensitive to the quality and quantity of available data. Although less sensitive to outliers compared to traditional moments, L-moments require high-quality data to produce accurate parameter estimates. Hinis and Geyikli pointed out that the accuracy of L-moment estimates can be compromised when dealing with sparse or noisy datasets [33].

2.5.2Limited flexibility with complex distributions

L-moments may struggle with very complex or highly skewed distributions. While effective for many common distributions, their performance can vary for distributions exhibiting extreme kurtosis or multimodal characteristics. Vivekanandan found that while L-moments provided reliable estimates for typical hydrological distributions, their performance was less consistent for more complex distributions [27].

2.5.3Computational intensity for large datasets

The computation of L-moments, particularly for large datasets or higher-order moments, can be computationally intensive. This poses challenges for real-time data analysis or scenarios requiring rapid processing. The need for advanced computational methods and optimization algorithms is critical to address this limitation and make L-moments more accessible for large-scale applications [34].

2.5.4Regionalization and homogeneity assumptions

A significant gap in L-moments research is the assumption of regional homogeneity in RFA. Determining homogeneous regions is crucial for accurate application, yet many studies rely on subjective methods. Guttman emphasized the need for objective and robust methods to define homogeneous regions, which remains an area for further research [35].

2.5.5Integration with machine learning and real-time data

There is a gap in integrating L-moments with modern machine learning techniques and real-time data analysis. Combining L-moments with machine learning algorithms could enhance predictive accuracy and computational efficiency, an area that remains largely unexplored [34].

2.5.6Evaluation of PL-moments

PL-moments extend L-moments to censored and incomplete data, but their full potential has not been extensively explored. More comprehensive studies are needed to evaluate PL-moments across different types of censored data and compare their effectiveness with other methods. Wang Junzhe demonstrated initial findings, but further validation is required [21].

2.5.7Assessment of methodological variations

Variations and modifications of L-moments, such as LH-moments and PL-moments, need more assessment across different contexts and applications. Comparative studies focusing on these variations can provide valuable insights into their strengths and limitations. Meshgi and Khalili suggested that while LH-moments offer advantages in certain scenarios, further validation and comparison are necessary [36].

L-moments have proven to be a robust tool for statistical analysis of extreme events, particularly in hydrology and meteorology. However, limitations such as sensitivity to data quality, computational intensity, and challenges with complex distributions need to be addressed. Additionally, significant study gaps in the integration of L-moments with machine learning, objective regionalization methods, and comprehensive evaluation of PL-moments require further research. Addressing these limitations and gaps will enhance the applicability and accuracy of L-moments, contributing to more reliable extreme event predictions and improved resource management strategies.

2.6Return period in hydrology

The concept of return period is fundamental in hydrology and risk management. It quantifies the expected frequency of occurrence for extreme hydrological events such as floods and droughts. This review covers the historical development, methodologies, applications, and limitations of return period analysis in hydrology.

2.6.1Definition and importance

The return period concept was first introduced by Fuller in 1914, gaining popularity due to its

simplicity and ease of understanding in hydrological and engineering applications. The return period (T) is traditionally defined as the inverse of the exceedance probability (P), representing the average time interval between events of a certain magnitude [37]. The return period is used to quantify the average interval between occurrences of extreme events [38]. It is critical for designing and assessing the reliability and safety of hydraulic structures. Traditional methods assume stationarity, where statistical properties of hydrological processes do not change over time [39].

2.6.2Mathematical formulation and methods

Fernández and Salas presented a comprehensive mathematical framework for estimating return periods and risks associated with hydrological events. Their approach addresses both simple and complex events, including dependent and independent annual flows and droughts. They emphasize the importance of considering dependencies among events, which can significantly influence the reliability of hydrological assessments [40].

2.6.3Nonstationarity and climate change

Climate change introduces nonstationarity in hydrological processes, complicating the estimation of return periods. Du et al. discuss how traditional stationary assumptions are no longer valid under changing climatic conditions. They propose methods incorporating meteorological covariates into nonstationary frequency analysis, providing more accurate estimates of return periods and associated risks under future climate scenarios [41].

2.6.4Multivariate analysis

Hydrological events often depend on multiple variables. Gräler et al. highlight the use of copulas in constructing multivariate return periods, allowing for the incorporation of dependencies between variables such as peak discharge, volume, and duration. This approach provides a more comprehensive understanding of extreme hydrological events [42].

2.6.5Data utilization and methodological advancements

Traditional methods often involve data decimation, discarding valuable information to fit models. Volpi et al. propose the complete time series analysis (CTA) method, which utilizes the entire dataset without decimation. This approach improves the reliability of return period estimates by preserving more information from the hydrological records [43].

2.6.6Case studies and applications

Practical applications of return period estimation are demonstrated in various studies. For instance, Shiau applied bivariate extreme value distributions to model flood events in Taiwan, showing good agreement between theoretical models and observed data. This study underscores the practical utility of advanced statistical methods in hydrological applications [44].

The concept of return period remains central to hydrological studies, but its application must evolve to address the challenges posed by nonstationary and complex dependencies among variables. Advances in statistical methods, such as the use of copulas and comprehensive data analysis techniques, provide more robust and accurate tools for estimating return periods. These developments are essential for effective water resource management and the design of resilient hydraulic structures.

3.Materials and methods

3.1L-moments

Hosking [3] introduced the concept of L-moments, which is grounded on the principles of order statistics. The primary four L-moments are expressed as follows in Equation 1:

$$\lambda_1 = \beta_0$$
$$\lambda_2 = 2\beta_1 - \beta_0$$
$$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0$$
$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \quad (1)$$

where λ_1 is the mean of the distribution; λ_2 is a measure of dispersion; λ_3 is a measure of skewness; and λ_4 is a measure of kurtosis. Here, β_r are probability weighted moments (PWMs), a concept previously defined by Greenwood et al. [45] and further devised by [46] as shown in Equation 2:

$$\beta_r = \int_0^1 x(F)F^r \, dF \quad (2)$$

F represents the nonexceedance probability. The ratios of these L-moments are then computed as per Equation 3:

$$\textit{Coefficient of L-variation}, \tau_2 = \frac{\lambda_1}{\lambda_2},$$
$$\textit{L-skewness}, \tau_3 = \frac{\lambda_3}{\lambda_2}, \quad (3)$$
$$\textit{L-kurtosis}, \tau_4 = \frac{\lambda_4}{\lambda_2}$$

For a sorted data sample, denoted as $x_1 \leq x_2 \leq \cdots \leq x_n$, Hosking [3] provides an unbiased formula for estimating sample PWMs as presented in Equation 4:

$$\widehat{\beta_r} = \frac{1}{n}\sum_{i=1}^{n} \frac{(i-1)(i-2)\ldots(i-r)}{(n-1)(n-2)\ldots(n-r)} x_i = b_r \quad (4)$$

Subsequently, unbiased estimators for L-moments can be articulated as per Equation 5:

$$l_1 = b_0,$$
$$l_2 = 2b_1 - b_0, \quad (5)$$
$$l_3 = 6b_2 - 6b_1 + b_0,$$
$$l_4 = 20b_3 - 30b_2 + 12b_1 - b_0$$

The sample estimates for τ are given as given in Equation 6:

$$t_r = \frac{l_r}{l_2} \text{ for } r = 3, 4 \text{ and } t_2 = \frac{l_2}{l_1} \quad (6)$$

Parameters corresponding to various statistical distributions investigated in this study and their quantile functions are summarized in *Table 1* [3, 4].

Table 1 Quantile functions of several probability distributions

Distributions	Quantile functions
Gumbel	$x(F) = \varepsilon - \alpha \ln(-\ln(F))$ where F, ε *and* α are the cumulative distribution function (CDF), location and scale parameter respectively
Generalized Logistic (GLO)	$x(F) = \epsilon + \frac{\alpha}{K}\left\{1 - \left[\frac{1-F}{F}\right]^K\right\}$ where F, ϵ, α and K are the CDF, location, scale and shape parameter respectively
Generalized Extreme Value (GEV)	$x(F) = \epsilon + \frac{\alpha}{K}\{1 - (-\ln F)^K\}$ where F, ϵ, α and K are the CDF, location, scale and shape parameter respectively
Generalized Pareto (GPA)	$x(F) = \epsilon + \frac{\alpha}{K}\{1 - [1-F]^K\}$ where F, ϵ, α and K are the CDF, location, scale and shape parameter respectively
Pearson Type III (Pe3)	The quantile function for the Pearson Type III distribution (Pe3) cannot be expressed in a simple closed-form equation.
Generalized Normal (GNO)	The quantile function for the Generalized Normal distribution (GNO) cannot be expressed in a simple closed-form equation.
Four-Parameters Kappa (K4D)	$x(F) = \epsilon + \frac{\alpha}{k}\left\{1 - \left(\frac{1-F^h}{h}\right)^k\right\}$ where F, ϵ, α are the CDF, location, scale parameter respectively and k, h are the two shape parameters

3.2Study area and dataset

The rainfall station of DID Kemaman is situated near Kemaman River in the southern part of Terengganu with latitude 4.2319° N and longitude 103.4222° E. Rainfall data for this study were extracted from DID Kemaman station. This station records daily rainfall (mm) and has been maintained by the DID, Ministry of Natural Resources and Environment, Malaysia, since 1971. Annual maximum rainfall data from 1971 to 2022 were used in this analysis. *Figure 1* presents the map location of the DID Kemaman Station.

3.2.1Data collection

The primary methodological approach involves extracting the annual maximum rainfall from the daily records for each year. This method focuses on identifying the single highest rainfall measurement for each year, which is then used for assessing the return periods of extreme rainfall events.

3.2.2Handling missing values in data

It was initially observed that the dataset contained some missing values, primarily attributed to equipment malfunctions or adverse environmental conditions. The missing data was handled by establishing a threshold where if the missing data within a single year exceeded 30% of the daily records, especially during the monsoon season, that year's data was excluded from the analysis. This threshold was set to ensure that missing data did not occur during critical periods of high rainfall, which could lead to underestimation of annual maximum values.

Figure 1 Map location of DID Kemaman station

3.3Goodness-of-fit criteria

3.3.1L-moments ratio diagram

The LMRD serves as a sophisticated graphical technique, predominantly employed in the disciplines of hydrology and statistics, for the comparative assessment and selection of appropriate probability distributions pertinent to a specific dataset. Within the LMRD, the third and fourth L-moment ratios commonly denoted as τ_3 and τ_4 are plotted on the x-axis and y-axis, respectively. These L-moment ratios function as analogous metrics to skewness and kurtosis in traditional moment statistics. The diagram facilitates a rigorous, visual comparison of the L-moment ratios calculated from empirical data against those derived from candidate probability distributions.

The framework for constructing the LMRD is delineated in seminal works by Hosking [47]. According to these relationships, certain distributions like two-parameter normal (NOM), Gumbel, logistic (LOG), and extreme value type I (EV1) are depicted as discrete points on the LMRD. In contrast, the three-parameter lognormal (LN3), GLO, GEV, GPA and Pe3 distributions manifest as curves or lines while K4D manifest as a plane bounded by the GLO curve above and overall lower bound below. The distribution that lies closest to the coordinate representing the average sample L-moment ratios (τ_3, τ_4) is considered the most appropriate for fitting the empirical data. Conversely, the distribution farthest from this coordinate is deemed the least suitable for accurately representing the dataset.

3.3.2Mean absolute deviation index (MADI) and mean squared deviation index (MSDI)

The MADI and MSDI serve as a sophisticated tool for assessing variability, particularly when examining data relative to theoretical benchmarks. The formula for MADI and MSDI can be written as follows in Equation 7 and 8 respectively:

$$\text{MADI} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i - z_i}{x_i}\right| \quad (7)$$

$$\text{MSDI} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - z_i}{x_i}\right)^2 \quad (8)$$

In this context, x_i represents the observed data points, while z_i denotes the predicted data points, corresponding to successive values calculated using Gringorten's plotting position formula [48, 49] for the empirical probability of exceedance. A smaller MADI or MSDI value for a particular distribution indicates a better fit to the actual data. Consequently, the distribution yielding the lowest MADI or MSDI value is considered the most suitable model, whereas

the one with the highest values is deemed the least appropriate for modeling the observed data.

3.3.3Return period

The return period is a statistical measure used to estimate the average time interval between occurrences of extreme events [50], like floods. It is commonly calculated using the formula of Equation 9:

$$T_x = \frac{1}{P_x} \quad (9)$$

where Tx represents the return period in years and Px is the annual exceedance probability as per Equation 10. In mathematical terms,

$$P_x = P(X \geq x_T) \quad (10)$$

Traditional methods often assume data stationarity, which is becoming less reliable due to factors like climate change and urbanization affecting the event patterns. Therefore, the methodology for calculating return periods is evolving to account for these non-stationary influences [39].

3.4Flow of enhanced return period estimation

Phase 0: **Initialization**

Step 0: Collect daily rainfall data from DID Kemaman Station for the period 1971 to 2022. Extract the annual maximum rainfall from the daily data for each year.

Phase 1: **L-moments calculation**

Step 1: Order the annual maximum data in ascending order for the calculation of L-moments.

Step 2: Calculate the L-moments. Refer section 3.1 Equation 4 and Equation 5.

Step 3: Calculate L-moments ratios. Refer section 3.1 Equation 6.

Phase 2: **Identify candidate distributions using LMRD.**

Step 4: Plot the τ_3 and τ_4 values on an LMRD to visually assess which theoretical probability distributions might provide the best fit.

Phase 3: **Parameter estimation using L-moments**

Step 5: Identify the appropriate probability distribution based on the LMRD analysis.

Step 6: Estimate the parameters of the chosen distribution using the relationships between L-moments and the distribution's parameters. This involves substituting the calculated L-moments into the parameter equations specific to the chosen distribution. The relationships between L-moments and the distribution's parameters can be found in Hosking [4].

Phase 4: **Validate with goodness-of-fit test (MADI and MSDI)**

Step 7: Use the quantile function (*Table 1*) of the chosen probability distribution with the estimated parameters to calculate predicted data points (z_i) for each empirical probability.

Step 8: Employ the Gringorten plotting position formula to transform the observed and predicted data points before calculating MADI and MSDI.

Step 9: Compute the MADI and MSDI to assess the goodness of fit. Refer section 3.3.2 Equation 7 and 8.

Step 10: Determine the best-fit distribution. Lower MADI and MSDI means better fit.

Phase 5: **Return period calculation**

Step 11: Insert the estimated parameters into the CDF of the chosen probability distribution. Refer section 4.5 Equation 10.

Step 12: Calculate the return periods by evaluating the CDF for different rainfall intensities. Refer section 4.5 Equation 11.

3.5Software's and hardware

Hardware: All analyses were performed on an AMD Ryzen 5 7500F 6-Core Processor with 32GB RAM, running Windows 11. The analyses hardware requirements heavily depend on the size of data.

Software: Statistical analyses were conducted using Python 3.9. Key packages included lmoments3 for L-moments estimation, pandas for data manipulation, NumPy for numerical calculations, and matplotlib for data visualization. R statistical software was also used in conducting LMRD analysis.

4.Results and discussion

4.1Time series plot

The time series plot depicted in *Figure 2* illustrates the annual maximum daily precipitation levels recorded at the DID Kemaman station for the period spanning from year 1971 to year 2022. An examination of the plot reveals that the lowest recorded annual maximum daily precipitation was 76.5 mm in the year 2010. Conversely, the highest observed value for annual maximum daily rainfall at

this station was registered at 440.3 mm in the year 2022. The year 1996 record was excluded from analysis for having too many missing values throughout the year.

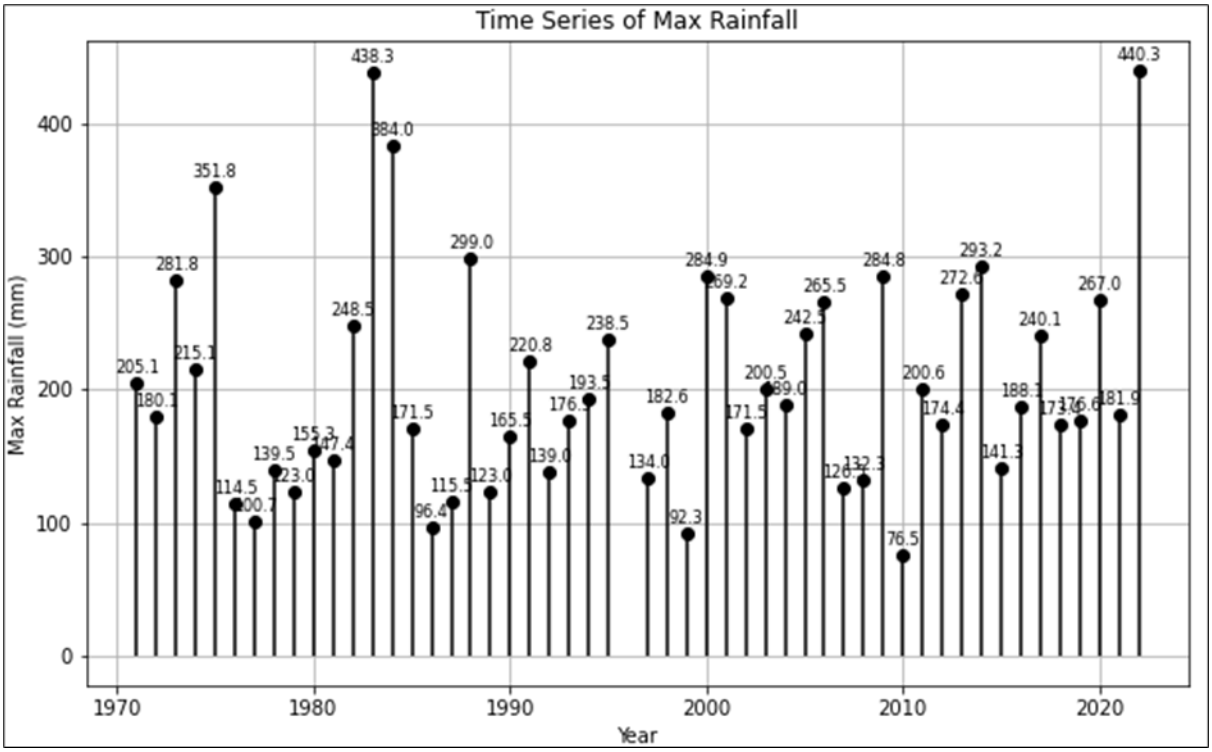

Figure 2 Time series plot of annual maximum rainfall of DID Kemaman

4.2L-moment ratio

The precipitation data was subsequently examined to derive the L-metrics in *Table 2*, specifically l_1, l_2, t_3 and t_4, which correspond to L-mean, L-scale, L-skewness, and L-kurtosis respectively.

Table 2 L-metrics for DID Kemaman station annual maximum rainfall data

Station name (ID)	n	l_1	l_2	t_3	t_4
DID. Kemaman at Terengganu (4234109)	51	204.4	45.7	0.19	0.15

The values for t_3 and t_4 were then plotted on an LMRD in *Figure 3* to provide a graphical representation of the dataset's shape characteristics, including its skewness and kurtosis. This diagram helps in visualizing how closely the rainfall data aligns with the theoretical distributions. The LMRD's analysis indicated that the values of t_3 and t_4 were in closest alignment with the GEV, GNO, Pe3, K4D, and Gumbel distributions. The K4D distribution, while not explicitly marked in the diagram's legend, is represented by the area enclosed by the GLO and the lower boundary curve labeled ALL.LB. Such closeness between the observed L-moment ratios and these distributions implies that the DID Kemaman rainfall data set is likely to conform to these distributions. The distributions were selected for parameter estimation phase based on their proximity to the observed L-moment ratios in the LMRD which serves as an initial screening tool to identify potential distributions that may provide a good fit for the data. Subsequent validation of the distribution fit will be conducted using MADI and MSDI to ensure the selected distributions accurately represent the underlying statistical characteristics of the DID Kemaman rainfall dataset. The distributions curves present are pre-determined. The dependent variable here is the t_3 (L-Skewness) and t_4 (L-Kurtosis) of DID Kemaman data, marked by the red dot in *Figure 3*.

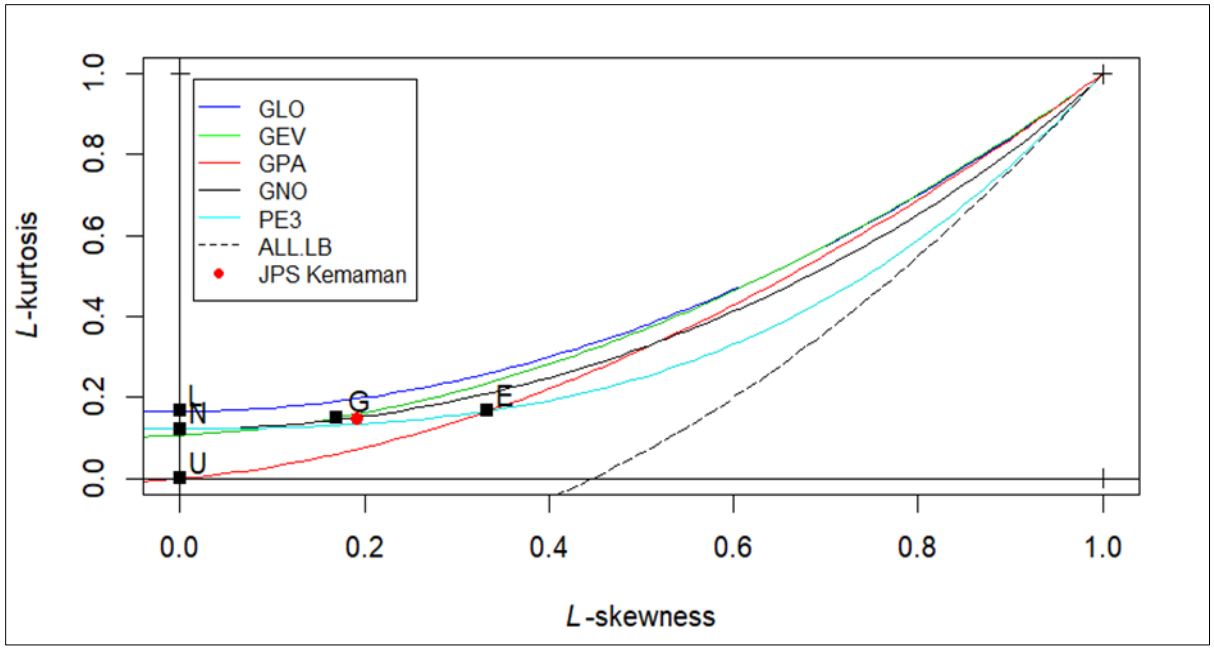

Figure 3 L-moment ratio diagram

4.3L-moment parameter estimation

The dataset containing the annual maximum daily rainfall measurements for DID Kemaman was subsequently subjected to computational analysis. The parameters for the distributions were estimated through the L-moment method. A comprehensive summary of the estimated parameters for each distribution is presented in *Table 3*. The selection of the distributions is based on the closeness of the t_3 and t_4 of DID Kemaman with the distributions' curves in the LMRD depicted in *Figure 3.*

Table 3 Estimated parameters of probability distributions for DID Kemaman station

Distribution	Estimated parameters
Gumbel	$\hat{\epsilon} = 166.3341$, $\hat{\alpha} = 65.9984$
GLO	$\hat{\epsilon} = 190.2906$, $\hat{\alpha} = 43.0416$, $\hat{K} = -0.1913$
GEV	$\hat{\epsilon} = 165.3589$, $\hat{\alpha} = 63.9563$, $\hat{K} = -0.0330$
GPA	$\hat{\epsilon} = 96.5760$, $\hat{\alpha} = 146.4247$, $\hat{K} = -0.3576$
Pe3	$\hat{\epsilon} = 204.4294$, $\hat{\alpha} = 84.5408$, $\hat{K} = 1.1585$
GNO	$\hat{\epsilon} = 188.8249$, $\hat{\alpha} = 75.9759$, $\hat{K} = -0.3950$
K4D	$\hat{\epsilon} = 158.9837$, $\hat{\alpha} = 70.6798$, $\hat{k} = 0.0118$, $\hat{h} = 0.1582$

4.3.1Parameter implications to rainfall behavior

Location parameter (ε)

The location parameter ε represents the central or median value of a distribution. Accurately estimating ε is crucial because it establishes the baseline for what is considered normal or typical rainfall within any model. A higher ε value indicates that extreme rainfall events occur more frequently within the region's climate pattern and are not merely outliers. Precise estimation of this parameter is essential, as minor errors can significantly misrepresent the central tendency, impacting decisions related to water resource management and agricultural planning.

Scale parameter (α)

The scale parameter α determines the spread or variability of rainfall data. Accurately capturing α is vital for understanding the potential range of rainfall events. A larger α implies greater variability, suggesting that the region may experience a broad range of rainfall intensities. Precise estimation of α is critical because return periods are sensitive to changes in variability. Underestimating α could understate the risk of flooding, while overestimating might lead to economically inefficient infrastructure planning due to excessive caution.

Shape parameters (k & h)

These parameters significantly influence the distribution's tail behavior and asymmetry. A positive k value indicates a heavier tail, which suggests a higher likelihood of severe events. The parameter h adjusts the asymmetry, impacting how skewed the data might appear toward higher extremes. Accurately estimating k and h is crucial as they determine the upper extremes of rainfall predictions, key for disaster preparedness and emergency management strategies. Errors in these estimates can lead to inadequate preparedness for extreme events.

The accurate and precise estimation of ε, α, k, and h is essential not only as a statistical concern but also as a practical necessity. Given how sensitive return period calculations are to these parameters, small deviations can lead to significantly different outcomes in model predictions. This highlights the need for robust statistical methods and high-quality data to ensure that models reliably reflect the true risks and characteristics of extreme rainfall events.

4.4Goodness-of-fit

4.4.1Goodness-of-fit (MADI and MSDI)

Following the initial evaluation via the LMRD, the analytical methodology was expanded to include other goodness-of-fit metrics specifically, MADI and MSDI. The collective outcomes from these metrics indicate that the K4D yields the lowest value of MADI and MSDI of all the distributions, hence serves as the most fitting model for the dataset under study. The outcomes for the metrics are summarised in *Table 4.*

Table 4 MADI and MSDI goodness-of-fit tests

Distributions	MADI	MSDI
Gumbel	0.0340	0.0018
GLO	0.0410	0.0029
GEV	0.0315	0.0015
GPA	0.0476	0.0042
Pe3	0.0298	0.0014
GNO	0.0301	0.0013
K4D	0.0294	0.0013

Based on the results presented in *Table 4*, which summarizes the outcomes of the goodness-of-fit tests using MADI and MSDI, it is concluded that the K4D distribution provides the most suitable fit for the dataset among the distributions evaluated. This conclusion is drawn from the observation that the K4D distribution yields the lowest values for both MADI and MSDI, with scores of 0.0294 and 0.0013, respectively. These metrics are critical in assessing the goodness of fit, where lower values indicate a closer match between the observed data and the model's predictions.

4.4.2 Goodness-of-fit

Following the quantitative assessment of the goodness-of-fit through MADI and MSDI metrics, graphical visualization offers an intuitive means to further validate the model's fit. The visualization through the transformed Q-Q plot depicted in *Figure 4* typically involves plotting the observed data against the model predictions to visually assess how well the model captures the underlying distribution of the data. The blue line represents theoretical quantiles, while the dashed black line represents actual data quantiles.

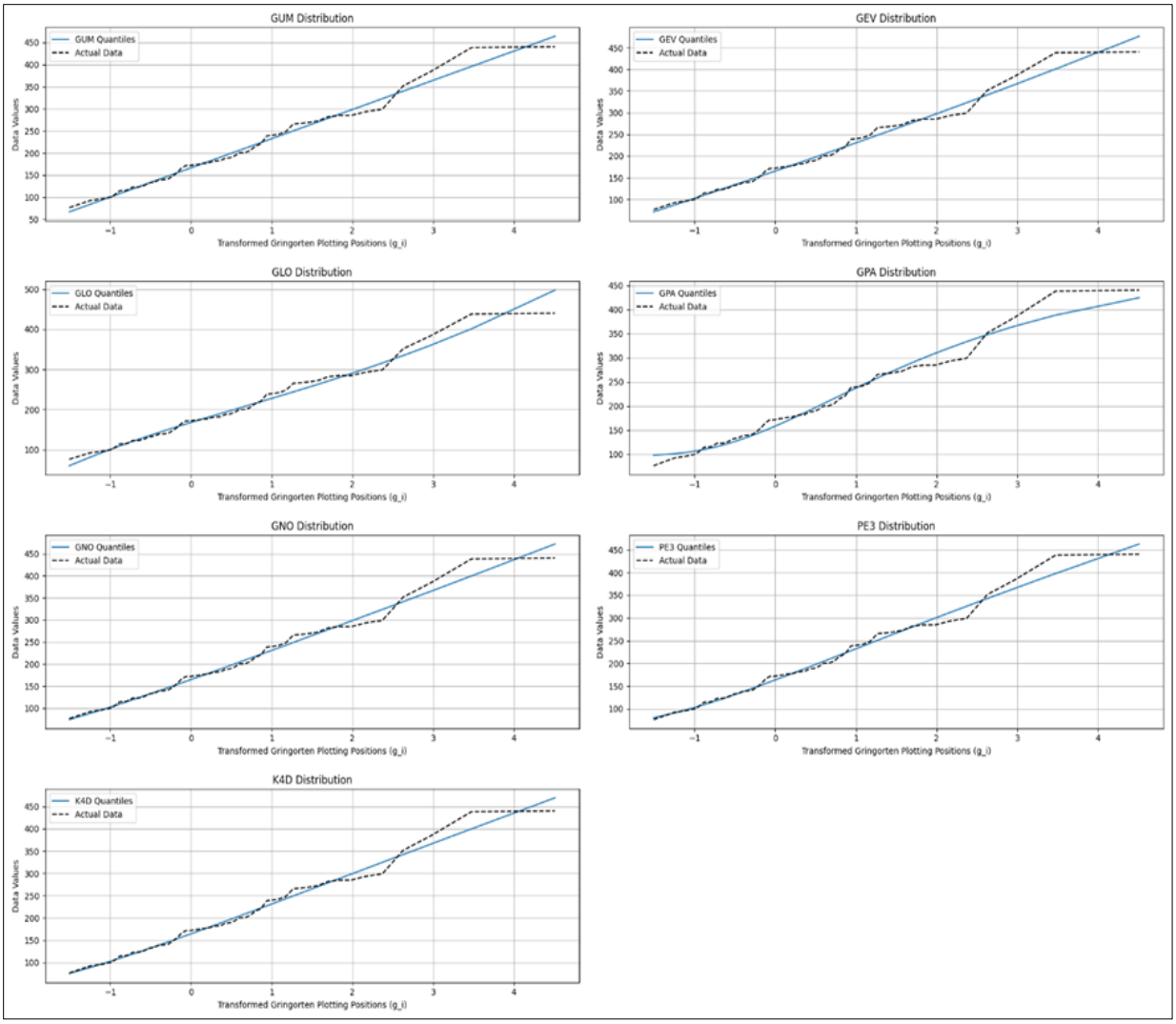

Figure 4 Q-Q plot comparing the quantiles of distributions against the actual data, using transformed Gringorten plotting positions

4.5Return period

Within the framework of the current research, the analytical scope is extended to include the calculation of return periods, utilizing metrics such as MADI and MSDI. These quantitative measures collectively identify the K4D distribution as the most appropriate statistical model for the examined rainfall dataset. This methodological approach allows for the deployment of computational techniques, offering a more detailed and accurate evaluation compared to graphical methods. Therefore, return period analysis was conducted under the assumption that the K4D distribution yields the most reliable fit. The selection of the K4D model aligns well with the overarching aim of this research—to provide a comprehensive and rigorous assessment of extreme precipitation events—establishing a solid foundation for future risk assessment and decision-making. Employing previously estimated parameters for the K4D, return values were computed for the dataset under consideration. Specifically, the K4D quantile function was utilized to derive a set of projected return values pertinent to the rainfall data. *Table 5* presents these calculated return values for DID Kemaman, assuming the dataset conforms to a K4D distribution.

Table 5 Estimated return values of annual maximum rainfall for several return periods of DID Kemaman

Return period (year)	Return value (mm)
2	188.66
5	265.29
10	316.51
20	365.55
30	393.64
40	413.38
50	428.61
60	440.99
70	451.43
80	460.45
90	468.39
100	475.48

The return period for rainfall events that either surpass or match the peak observed value of 440.3mm is calculated by inserting the parameters previously ascertained for the K4D distribution into its CDF namely, $\hat{\epsilon}$ = 158.9837, $\hat{\alpha}$ = 70.6799, $\hat{k}$ = 0.0118, $\hat{h}$ = 0.1582 and x = 440.3. The formula for the CDF of the K4D distribution is presented in Equation 11:

$$F(x) = \left[1 - h\left(1 - \frac{k(x-\epsilon)}{\alpha}\right)^{1/k}\right]^{1/h} \quad (11)$$

In this example, when considering a specific rainfall magnitude of interest, x = 440.3 for example, the CDF which denoted with F(x) yields 0.983. This value represents the annual probability that rainfall will not exceed 440.3 mm. To find the probability that this amount will be exceeded in any given year, denoted as P_x, calculation of 1 − F(x) is used. For x = 440.3, this calculation results in P_x = 0.017, indicating a 1.7% chance that annual maximum rainfall will exceed 440.3 mm in any given year. A probability value of Px was calculated to be 0.017 as shown in Equation 12:

P(annual maximum >= 440.3) = 1 - F(440.3) = 0.017

Estimated return period $(T_x) = (0.017)^{-1} = 59.39 \approx 59$ years (12)

4.6Sensitivity analysis of return period estimates to K4D parameters

The sensitivity analysis conducted on the K4D model examined the impact of perturbations in the scale (α), shape (k), second shape (h), and location (ϵ) parameters on return period estimates. Perturbations of ±5% were applied to each parameter, and their effects on the predicted return values for periods ranging from 2 to 100 years were analyzed. Key findings from the analysis include both α and ϵ parameters demonstrated significant sensitivity, with changes in these parameters resulting in proportional and substantial shifts in return values, respectively. These parameters are crucial for accurately predicting the magnitude and baseline of extreme events. Modifications to k and h also impacted the return values but to a lesser extent compared to α and ϵ parameters. These parameters primarily influence the distribution's tail behavior and peakedness, affecting the likelihood and characterization of extreme events. *Table 6* to *Table 9* displays the sensitivity analysis for the K4D parameters.

Table 6 Sensitivity analysis of α parameter for K4D

Return Period (years)	Baseline Value (mm)	0.95x Value (mm)	0.95x % Change	1.05x Value (mm)	1.05x % Change
2	188.66	180.71	-4.21	196.60	4.21
5	265.29	257.34	-3.00	273.23	3.00
10	316.51	308.56	-2.51	324.46	2.51
20	365.55	357.60	-2.17	373.49	2.18
50	428.61	420.66	-1.85	436.56	1.85
100	475.48	467.53	-1.67	483.43	1.67

Table 7 Sensitivity analysis of ϵ parameter for K4D

Return Period (years)	Baseline Value (mm)	0.95x Value (mm)	0.95x % Change	1.05x Value (mm)	1.05x % Change
2	188.66	187.17	-0.79	190.14	0.79

Return Period (years)	Baseline Value (mm)	0.95x Value (mm)	0.95x % Change	1.05x Value (mm)	1.05x % Change
5	265.29	259.97	-2.00	270.60	2.00
10	316.51	308.64	-2.49	324.39	2.49
20	365.55	355.22	-2.83	375.87	2.83
50	428.61	415.13	-3.15	442.09	3.15
100	475.48	459.66	-3.33	491.31	3.33

Table 8 Sensitivity analysis of k parameter for K4D

Return Period (years)	Baseline Value (mm)	0.95x Value (mm)	0.95x % Change	1.05x Value (mm)	1.05x % Change
2	188.66	188.66	0.00	188.65	-0.00
5	265.29	265.33	0.02	265.24	-0.02
10	316.51	316.62	0.03	316.41	-0.03
20	365.55	365.73	0.05	365.36	-0.05
50	428.61	428.92	0.07	428.30	-0.07
100	475.48	475.91	0.09	475.05	-0.09

Table 9 Sensitivity analysis of h parameter for K4D

Return Period (years)	Baseline Value (mm)	0.95x Value (mm)	0.95x % Change	1.05x Value (mm)	1.05x % Change
2	188.66	188.47	-0.10	188.84	0.10
5	265.29	265.22	-0.02	265.35	0.02
10	316.51	316.48	-0.01	316.54	0.01
20	365.55	365.53	-0.00	365.56	0.00
50	428.61	428.60	-0.00	428.61	0.00
100	475.48	475.48	-0.00	475.48	0.00

The results highlight the importance of precise parameter estimation, particularly for the scale and location parameters, due to their pronounced impact on model outputs. Accurate parameter calibration is essential for enhancing the predictive accuracy and reliability of hydrological models. Employing L-moments for parameter estimation provides a significant advantage, as it predicts parameters more accurately compared to traditional methods. This increased accuracy is crucial in sensitivity analyses, ensuring that the models are both robust and reliable.

5.Discussions

5.1Summary of key findings

Through L-moment parameter estimation, various probability distributions were analyzed, with the K4D distribution emerging as the most suitable model for the dataset. This conclusion was supported by the goodness-of-fit tests using MADI and MSDI, where the K4D distribution yielded the lowest deviation values, indicating the best fit shown in *Table 4*. The study effectively utilized the K4D to estimate the return periods of extreme rainfall events at the DID Kemaman station.

A critical finding from this analysis is the calculation of return period for significant rainfall events, particularly those that match or exceed the peak observed value of 440.3 mm. By inserting the estimated parameters (ϵ = 158.9837, α = 70.6799, k = 0.0118, h = 0.1582) into the CDF of the K4D, it was determined that the probability of annual rainfall exceeding 440.3 mm is 1.7% (P_x = 0.017). This corresponds to an estimated return period of approximately 59 years.

5.2Implications of findings

The methodology applied in this study extends beyond extreme rainfall events. For example, if an analyst wants to determine the return period of a 150 mm rainfall event, they can do so by adjusting the input value of x in the CDF function. By substituting x = 150 into the K4D CDF function, the probability P_x of exceeding this threshold can be determined, followed by calculating its return period using the inverse probability formula. This model's flexibility enables stakeholders, including hydrologists and urban planners, to assess rainfall frequency for specific needs, such as agricultural water resource planning or infrastructure development.

The findings of this study have several implications for flood risk management in Terengganu and similar regions. The robust modeling of extreme rainfall

events using the K4D distribution enables more accurate predictions of future events, which is critical for designing resilient infrastructure. The detailed return period analysis provides essential data for planning and decision-making, ensuring flood defenses and emergency response strategies rely on reliable statistical models.

Furthermore, L-moments and the K4D distribution apply to other regions with similar climatic and hydrological conditions, providing a valuable tool for comparative studies and hydrological risk assessment.

5.3Limitations of study

While the ability to adjust input magnitudes for calculating return periods offers significant advantages, it also introduces complexities that must be managed carefully. The accuracy of these calculations depends heavily on the model's reliability and the assumptions underpinning it, such as the stationarity of climate patterns and the quality of historical data. Changes in climatic conditions, particularly those driven by climate change, could alter precipitation patterns and challenge the assumption of stationarity on which many models rely. Furthermore, the sensitivity of return periods to specific model parameters highlights the need for precision in data collection and parameter estimation.

5.4Recommendations for future research

To enhance the robustness and practical applicability of these findings, future research should focus on Integrating non-stationary models to account for climate change impacts. Additionally, the incorporation of real-time data and advanced computational techniques, such as machine learning, could improve predictive accuracy and computational efficiency. Expanding the dataset to include more rainfall stations and longer observation periods would also provide a more detailed and reliable analysis. Comparative studies across different climatic regions using the L-moments method and the K4D distribution could further validate the effectiveness of these models.

A complete list of abbreviations is listed in *Appendix I.*

6.Conclusion

This research embarked on a detailed examination of the annual maximum daily rainfall data at the DID Kemaman station, with the dual objectives of identifying the most fitting probability distribution and estimating the return periods of extreme rainfall events. Employing the L-moment method for parameter estimation, a variety of distributions were initially considered, guided by the insights provided by the LMRD. The subsequent employment of goodness-of-fit criteria, specifically through the MADI and MSDI, facilitated a meticulous evaluation of these distributions' suitability in modeling the observed rainfall data. The K4D distribution was identified as the most accurate model, exhibiting the lowest values in both MADI and MSDI assessments. The results further reinforced this finding, confirming the superior accuracy of the K4D distribution. Moreover, the application of the K4D distribution for further analysis yielded significant insights into the return periods of extreme rainfall events. Notably, the analysis found that a 2-year return period corresponds to a rainfall event of 188.66 mm, while a 100-year return period predicts a significant rainfall event of 475.48 mm.

Acknowledgment

This research was funded by a grant from Ministry of Higher Education of Malaysia (MoHE) (FRGS Grant RR458). Rainfall data of DID Kemaman was retrieved from the Department of Irrigation and Drainage Malaysia (DIDM).

Conflicts of interest

The authors have no conflicts of interest to declare.

Data availability

The data used in this study are sourced from the DID, Malaysia. Due to restrictions related to confidentiality and data protection, the raw data are not publicly available. However, data are available from the authors upon reasonable request and with permission of DID Malaysia.

Author's contribution statement

Mohammad Amir Syahmi: He formulated the research problem, conducted the main data analysis focusing on rainfall extremes, and interpreted the results. Supervised the manuscript preparation. **Zahrahtul Amani Zakaria**: Provided expertise in statistical modeling, specifically in the application of L-Moments. Assisted in the theoretical framework development, reviewed the manuscript for statistical accuracy, and contributed to the interpretation of the results. **Nor Aida Mahiddin**: Focused on providing the necessary data for the study, ensuring the data was comprehensive and adequately prepared for analysis. Assisted in manuscript preparation, specifically ensuring the data presentation was clear and methodologically sound.

References

[1] Mahiddin NA, Sarkar NI. Improving the performance of MANET gateway selection scheme for disaster

recovery. In 18th international conference on high performance computing and communications; 14th international conference on smart city; 2nd international conference on data science and systems (HPCC/SmartCity/DSS) 2016 (pp. 907-12). IEEE.
[2] Khaliq MN, Ouarda TB, Ondo JC, Gachon P, Bobée B. Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: a review. Journal of Hydrology. 2006; 329(3-4):534-52.
[3] Hosking JR. L-moments: analysis and estimation of distributions using linear combinations of order statistics. Journal of the Royal Statistical Society Series B: Statistical Methodology. 1990; 52(1):105-24.
[4] Tasker G. Regional frequency analysis: an approach based on L-moments. Journal of the American Statistical Association. 1998; 93(443):1233.
[5] Shabri A, Ariff NA. Frequency analysis of maximum daily rainfalls via l-moment approach. Sains Malaysiana. 2009; 38(2):149-58.
[6] Hamzah FM, Yusoff SH, Jaafar O. L-moment-based frequency analysis of high-flow at Sungai Langat, Kajang, Selangor, Malaysia. Sains Malaysiana. 2019; 48(7):1357-66.
[7] Hassan BG, Ping F. Regional rainfall frequency analysis for the Luanhe Basin–by using L-moments and cluster techniques. APCBEE Procedia. 2012; 1:126-35.
[8] Krishna GS, Veerendra G. Flood frequency analysis of Prakasam barrage reservoir Krishna district, Andhra Pradesh using Weibull, Gringorten and L-moments formula. International Journal of Civil, Structural, Environmental and Infrastructure Engineering Research and Development. 2015; 5(2):57-62.
[9] Hailegeorgis TT, Alfredsen K. Regional flood frequency analysis and prediction in ungauged basins including estimation of major uncertainties for mid-Norway. Journal of Hydrology: Regional Studies. 2017; 9:104-26.
[10] Mosaffaie J. Comparison of two methods of regional flood frequency analysis by using L-moments. Water Resources. 2015; 42:313-21.
[11] Malekinezhad H, Nachtnebel HP, Klik A. Comparing the index-flood and multiple-regression methods using L-moments. Physics and Chemistry of the Earth, Parts A/B/C. 2011; 36(1-4):54-60.
[12] Rutkowska A, Żelazny M, Kohnová S, Łyp M, Banasik K. Regional L-moment-based flood frequency analysis in the upper Vistula River basin, Poland. Geoinformatics and Atmospheric Science. 2018:243-63.
[13] Seckin N, Haktanir T, Yurtal R. Flood frequency analysis of Turkey using L-moments method. Hydrological Processes. 2011; 25(22):3499-505.
[14] Vogel RM, Fennessey NM. L moment diagrams should replace product moment diagrams. Water Resources Research. 1993; 29(6):1745-52.
[15] Peel MC, Wang QJ, Vogel RM, Mcmahon TA. The utility of L-moment ratio diagrams for selecting a regional probability distribution. Hydrological Sciences Journal. 2001; 46(1):147-55.
[16] Haddad K. Selection of the best fit probability distributions for temperature data and the use of L-moment ratio diagram method: a case study for NSW in Australia. Theoretical and Applied Climatology. 2021; 143(3):1261-84.
[17] Ouarda TB, Charron C, Chebana F. Review of criteria for the selection of probability distributions for wind speed data and introduction of the moment and L-moment ratio diagram methods, with a case study. Energy Conversion and Management. 2016; 124:247-65.
[18] Hosking JR. Some theory and practical uses of trimmed L-moments. Journal of Statistical Planning and Inference. 2007; 137(9):3024-39.
[19] Bobee B, Cavadias G, Ashkar F, Bernier J, Rasmussen P. Towards a systematic approach to comparing distributions used in flood frequency analysis. Journal of Hydrology. 1993; 142(1-4):121-36.
[20] Murshed MS, Park BJ, Jeong BY, Park JS. LH-moments of some distributions useful in hydrology. Communications for Statistical Applications and Methods. 2009; 16(4):647-58.
[21] Junzhen WA, Songbai SO. Study on application of partial L-moments to flood frequency analysis. Journal of Hydroelectric Engineering. 2015; 12(1):1-10.
[22] Mudholkar GS, Hutson AD. LQ-moments: analogs of L-moments. Journal of Statistical Planning and Inference. 1998; 71(1-2):191-208.
[23] Sung JH, Kim YO, Jeon JJ. Application of distribution-free nonstationary regional frequency analysis based on L-moments. Theoretical and Applied Climatology. 2018; 133:1219-33.
[24] Bahmani R, Eslamian S, Khorsandi M, Hosseinipour EZ. Combination of L-moments method and hydrological model for design flood hydrograph determination. In world environmental and water resources congress 2013: showcasing the future 2013 (pp. 3236-46).
[25] Koutsoyiannis D. Knowable moments for high-order stochastic characterization and modelling of hydrological processes. Hydrological Sciences Journal. 2019; 64(1):19-33.
[26] Agbonaye AI, Otuaro EA, Christopher O. Comparison of L-moment and method of moments as parameter estimators for identification and choice of the most appropriate rainfall distribution models for design of hydraulic structures. Journal of Civil Engineering. 2022; 13(1):33-48.
[27] Vivekanandan N. Intercomparison of probability distributions for selecting a best fit for estimation of rainfall. i-Manager's Journal on Civil Engineering. 2022; 12(3):42-7.
[28] Vivekanandan N. Intercomparison of estimators of extreme value family of distributions for rainfall frequency analysis. Mausam. 2022; 73(1):59-70.
[29] Sanusi W, Chaerunnisa S, Annas S, Side S, Abdy M. Estimated parameters of rain flow distribution using L-moment method in South Sulawesi, Indonesia.

Journal of Applied Mathematics and Computation. 2022; 6(1):30-40.
[30] Guayjarernpanishk P, Bussababodhin P, Chiangpradit M. The partial L-moment of the four kappa distribution. Emerging Science Journal. 2023; 7(4):1116-25.
[31] Anghel CG, Stanca SC, Ilinca C. Two-parameter probability distributions: methods, techniques and comparative analysis. Water. 2023; 15(19):1-35.
[32] Chang CH, Rahmad R, Wu SJ, Hsu CT. Spatial frequency analysis by adopting regional analysis with radar rainfall in Taiwan. Water. 2022; 14(17):1-27.
[33] Hinis MA, Geyikli MS. Accuracy evaluation of standardized precipitation index (SPI) estimation under conventional assumption in Yeşilırmak, Kızılırmak, and Konya closed Basins, Turkey. Advances in Meteorology. 2023; 2023(1):1-13.
[34] Cao S, Lu H, Peng Y, Ren F. A novel fourth-order L-moment reliability method for L-correlated variables. Applied Mathematical Modelling. 2021; 95:806-23.
[35] Guttman NB. The use of L-moments in the determination of regional precipitation climates. Journal of Climate. 1993; 6(12):2309-25.
[36] Meshgi A, Khalili D. Comprehensive evaluation of regional flood frequency analysis by L-and LH-moments. I.A re-visit to regional homogeneity. Stochastic Environmental Research and Risk Assessment. 2009; 23:119-35.
[37] Fuller WE. Closure to flood flows. Transactions of the American Society of Civil Engineers. 1914; 77(1):676-94.
[38] Volpi E. On return period and probability of failure in hydrology. Wiley Interdisciplinary Reviews: Water. 2019; 6(3):e1340.
[39] Volpi E, Fiori A, Grimaldi S, Lombardo F, Koutsoyiannis D. One hundred years of return period: strengths and limitations. Water Resources Research. 2015; 51(10):8570-85.
[40] Fernández B, Salas JD. Return period and risk of hydrologic events. I: mathematical formulation. Journal of Hydrologic Engineering. 1999; 4(4):297-307.
[41] Du T, Xiong L, Xu CY, Gippel CJ, Guo S, Liu P. Return period and risk analysis of nonstationary low-flow series under climate change. Journal of Hydrology. 2015; 527:234-50.
[42] Gräler B, Van DBMJ, Vandenberghe S, Petroselli A, Grimaldi S, De BB, et al. Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation. Hydrology and Earth System Sciences. 2013; 17(4):1281-96.
[43] Volpi E, Fiori A, Grimaldi S, Lombardo F, Koutsoyiannis D. Save hydrological observations! return period estimation without data decimation. Journal of Hydrology. 2019; 571:782-92.
[44] Shiau JT. Return period of bivariate distributed extreme hydrological events. Stochastic Environmental Research and Risk Assessment. 2003; 17:42-57.
[45] Greenwood JA, Landwehr JM, Matalas NC, Wallis JR. Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. Water Resources Research. 1979; 15(5):1049-54.
[46] https://dominoweb.draco.res.ibm.com/reports/RC12210.pdf. Accessed 30 January 2025.
[47] Hosking JR. Approximations for use in constructing L-moment ratio diagrams. IBM Research Division, TJ Watson Research Center; 1991.
[48] Guo SL. A discussion on unbiased plotting positions for the general extreme value distribution. Journal of Hydrology. 1990; 121(1-4):33-44.
[49] Cook NJ, Harris RI. The Gringorten estimator revisited. Wind & Structures. 2013; 16(4):355-72.
[50] Benjamin JR, Cornell CA. Probability, statistics, and decision for civil engineers. Courier Corporation; 2014.

Mohammad Amir Syahmi was born in Kuala Terengganu, Malaysia in 1995 and raised in Malaysia. He initiated his higher education journey at the University of Nottingham in the UK, where he completed an Undergraduate Diploma in Software Engineering. Following this, he earned an undergraduate degree in Financial Mathematics from Universiti Malaysia Terengganu, Malaysia. He is currently pursuing a Master's Degree in Mathematical Sciences at Universiti Sultan Zainal Abidin, Malaysia. Mohammad is a Certified Data Science Analyst, accredited by Fusionex, and a Certified Big Data Analyst with expertise in Hadoop from NIIT. Presently, he serves as a Graduate Research Assistant, where he is actively involved in a grant-funded research project focusing on rainfall extreme events. His work contributes to the understanding of tropical rainfall patterns and their implications on regional and global climate models. His academic and professional pursuits reveal a deep commitment to leveraging Data Science and Mathematical expertise to address pressing environmental challenges, particularly those related to extreme weather conditions.
Email: amirsyahmi.jamsari@gmail.com

Dr. Zahrahtul Amani Zakaria was born in Terengganu, Malaysia. She received her B.Sc. in Industrial Mathematics and her Ph.D. in Mathematics from the University of Technology Malaysia. She is currently an Associate Professor at Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin (UniSZA), Malaysia. Her academic expertise is deeply rooted in Mathematics and Statistics, specifically focusing on areas such as Probability and Distribution Theory, Statistical Modelling, Industrial Statistics, and Applied Mathematics. Her work involves significant contributions to research and education in the

field of applied probabilities specifically involving L-Moments, providing a substantial foundation in both theoretical and applied aspects of her discipline. This aligns with her role in developing and nurturing future professionals in these fields through comprehensive academic programs and research initiatives.
Email: zahrahtulamani@unisza.edu.my

Dr. Nor Aida Mahiddin received her B.S. Degree in Information Technology from the National University of Malaysia, her Master's Degree in Computer Science with a major in Distributed Computing, and her Ph.D. in Computer and Information Sciences from Auckland University of Technology, New Zealand. She currently holds a position as a senior lecturer at the Faculty of Informatics and Computing at the University Sultan Zainal Abidin, Malaysia, and is affiliated with the East Coast Environmental Research Institute (ESERI). Dr. Nor Aida is the author of numerous papers published in peer-reviewed journals and conference proceedings. She is also an active member of professional organizations, including the Institute of Electrical and Electronics Engineers (IEEE), the Internet Society, and The Society of Digital Information and Wireless Communications (SDIWC). Her research interests span across a diverse range of areas, with a primary focus on Disaster Recovery Planning and Infrastructure. Additionally, she is deeply engaged in exploring emergency communication frameworks and their optimization, as well as Ad Hoc Network Design and Traffic Flow Frameworks.
Email: aidamahiddin@unisza.edu.my

Appendix I

S. No.	Abbreviation	Description
1	CDF	Cumulative Distribution Function
2	CTA	Complete Time Series Analysis
3	DID	Department of Irrigation and Drainage
4	EV1	Extreme Value Type I
5	GEV	Generalized Extreme Value
6	GLO	Generalized Logistic
7	GNO	Generalized Normal Distribution
8	GPA	Generalized Pareto
9	HEC-HMS	Hydrologic Engineering Center – Hydrologic Modeling System
10	K4D	Four-Parameter Kappa Distribution
11	LH-moments	Linear Combination of H-Statistics
12	L-moments	Linear Combination of Probability Weighted Moment
13	LMRD	L-Moment Ratio Diagram
14	LN3	Three-Parameter Lognormal
15	LQ-moments	Linear combination of Quantile functions
16	MADI	Mean Absolute Deviation Index
17	MAE	Mean Absolute Error
18	MoM	Method of Moments
19	MRD	Moment Ratio Diagram
20	MSDI	Mean Squared Deviation Index
21	NOM	Normal Distribution
22	Pe3	Pearson Type III Distribution
23	PL-moments	Partial L-Moments
24	PWM	Probability Weighted Moment
25	RFA	Regional Frequency Analysis
26	RMSE	Root Mean Square Error