Available online at
https://mijuitm.com.my

e-ISSN: 2735-0703

Mathematical Sciences and Informatics Journal 6(2) 2025, 97 – 113.

Mathematical Sciences and Informatics Journal

L-Moment Methodology for Flood Frequency Analysis: A Case Study of Triang River

Nur Amira Abdul Aziz[1], Basri Badyalina[2*], Ani Shabri[3], Fatin Farazh Ya'acob[4] & Ira Syaqira Sukimin[5]

[1]*Faculty of Computer and Mathematical Science, Universiti Teknologi MARA (UiTM)Cawangan Negeri Sembilan, Kampus Seremban, 70300 Seremban, Negeri Sembilan, Malaysia.*
[2*] *Faculty of Computer and Mathematical Science, Universiti Teknologi MARA Cawangan Johor, Kampus Segamat, 85000 Johor, Malaysia*
[3]*Department of Mathematical Sciences, Faculty of Science,Universiti Teknologi Malaysia, 81310, Johor, Malaysia.*
[4]*Faculty of Business and Management, Universiti Teknologi MARA,Cawangan Johor, Kampus Segamat, 85000 Johor, Malaysia*
[5]*Academy of Language Studies, Universiti Teknologi MARA,Cawangan Johor, Kampus Segamat, 85000 Johor, Malaysia*

ARTICLE INFO

Article history:
Received 12 August 2025
Revised 30 August 2025
Accepted 20 September 2025
Online first
Published 31 October 2025

Keywords:
L-Moment
Flood Frequency Analysis
Extreme Value Distribution
Hydrology in Peninsular Malaysia

DOI:
10.24191/mij.v6i2.4654

ABSTRACT

This study investigates flood frequency analysis in Peninsular Malaysia using extreme value distributions derived from the L-Moment method. This study evaluates the relevance of three probability distributions: Generalized Extreme Value (GEV), Generalized Pareto (GPA), and Generalized Logistic (GLO), for modelling flood magnitudes. The research focuses on parameter estimation, performance evaluation, and quantile estimation to assess flood risks. Utilizing hydrological data from 1961 to 2023, the research employs the goodness-of-fit tests alongside performance metrics such as Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), Root Mean Square Percentage Error (RMSPE), and Coefficient of Determination (R^2).The results reveal that the GEV distribution consistently outperforms GLO and GPA distributions, achieving superior accuracy across metrics and reliable quantile estimates for different return periods. The findings contribute to the field of hydrology by offering robust models for flood risk analysis and practical strategies for mitigation. Policymakers and urban planners could leverage these insights to improve disaster management and infrastructure planning in flood-prone areas of Peninsular Malaysia, addressing the challenges posed by persistent monsoonal floods.

[2*] Corresponding author. *E-mail address*: basribdy@uitm.edu.my
https://doi.org/10.24191/mij.v6i2.4654

1. INTRODUCTION

Flooding is a significant issue for Malaysia, particularly in Peninsular Malaysia, due to its tropical climate, monsoon seasons, and geographical features. The Northeast Monsoon, occurring from November to March, often causes heavy and prolonged rainfall, resulting in frequent floods in areas like Kelantan, Terengganu, and Johor (Rezali et al.,2025). These floods disrupted lives, damaged infrastructure, and incurred substantial economic losses. Globally, studies indicated that once-in-a-century floods could impact approximately 1.81 billion people or 23% of the population (Chancel et al., 2023). Effective flood management is critical, and flood frequency analysis (FFA), employing statistical tools like L-moments and extreme value distributions, offered essential insights for disaster mitigation and infrastructure planning (Hamed & Rao, 2019).

This study focuses on advancing flood risk analysis in Peninsular Malaysia by employing Flood Frequency Analysis (FFA) with L-moment methods to estimate the parameters of extreme value distributions, including the Generalized Pareto, Generalized Extreme Value, and Generalized Logistic distributions. By analysing historical streamflow data from 1961 to 2023, the research seeks to identify the best-fitted distribution for accurately modelling flood frequency. The study addressed the escalating threats of urbanization, deforestation, and climate change (Abid et al., 2021), offering critical insights for policymakers, engineers, and planners. These findings aim to enhance flood preparedness, safeguard communities, and promote sustainable development, despite challenges related to data acquisition and regional variations in flood behaviour (Mohd Baki et al., 2014). By analysing long-term streamflow data, the research evaluates the performance of the Generalized Extreme Value (GEV), Generalized Pareto (GPA), and Generalized Logistic (GLO) distributions to determine the best fit for regional flood characteristics. The findings are expected to address challenges such as regional variability and data limitations, contributing to enhanced flood risk assessment, infrastructure design, and climate adaptation strategies.

2. LITERATURE REVIEW

2.1 Flood Frequency Analysis (FFA)

FFA has seen significant advancements over the years. In the 1990s, methods were developed to incorporate censored or individual peak flow data and outlier-detection tests were introduced. By the 2000s, historical data became instrumental in estimating flood quantiles. Today, FFA remains a core method for assessing and managing flood risks, with continuous efforts to improve accuracy and confidence in estimates (Ali and Rahman, 2022; Dalrymple, 1960; Stedinger, 1993). The growing body of FFA literature has prompted many authors to synthesize and review its development (Hamed & Rao, 2019). FFA plays a vital role in flood-prone regions by supporting flood forecasting through streamflow data analysis (Hamzah et al., 2020). In Malaysia, FFA has been widely applied in areas such as river basin management, climate change adaptation, and flood system management, contributing to mitigation planning, flood prediction, and improved drainage infrastructure (Bakri, 2022; Ahmad et al., 2023; Che Ilias et al., 2021). Predicting the maximum flood magnitude for specific return periods is a crucial requirement for designing hydraulic structures and managing infrastructure (Badyalina et al., 2015; Krishna & Veerendra, 2015). FFA uses probability distributions to analyse high runoff events and their recurrence intervals, making it indispensable for understanding and managing flood risks (Hamed & Rao, 2019).

2.2 Generalized Extreme Value Distribution (GEV)

The GEV distribution plays a critical role in FFA by modelling the distribution of extreme values, such as annual maximum flood levels, using three parameters: location (ξ), scale (α), and shape (k). These parameters respectively described central tendency, spread, and tail behaviour, with the GEV unifying the

https://doi.org/10.24191/mij.v6i2.4654

Gumbel, Fréchet, and Weibull distributions, based on the shape parameter (Hossain et al., 2021). Gumbel models light-tailed data ($k = 0$), Fréchet is suited for heavy-tailed data ($k > 0$), and Weibull applied to bounded upper-tailed data ($k < 0$) (Legrand, 2022). Studies, such as those by Ahmad et al. (2023), Yusoff et al. (2022), and Hamzah et al. (2020), highlighted the effectiveness of GEV in FFA across various locations in Malaysia, often outperforming other distributions like the Generalized Pareto (GPA) and Generalized Logistic (GLO). Bakri (2022) further affirmed GEV's suitability in Regional Flood Frequency Analysis in Peninsular Malaysia, demonstrating its robustness and versatility for extreme value modelling.

2.3 Generalized Pareto Distribution (GPA)

The GPA distribution is widely utilised in hydrological studies to model peak exceedance above a defined threshold, aiding in flood risk assessment, management, and hydraulic structure design. Its parameters include the location (ξ), setting the threshold; the scale (α), defining data dispersion; and the shape (k), influencing tail behaviour and flood event severity. A positive k indicates a heavy upper tail with a higher probability of extreme events, a negative k resulted in a finite upper tail, and $k = 0$ corresponded to an exponential distribution with a light tail (Campos-Aranda, 2016). GPA's reliability and robustness, particularly with L-moment parameter estimation, make it a valuable tool in FFA. Hassim et al. (2022) found GPA to outperform GEV in Kelantan River Basin studies, while Hamzah (2020) applied GPA to analyse high tides at Port Klang. Badyalina et al. (2022) utilized GPA for flood return period estimation in Labis, Johor, suggesting that GEV and GLO could also be explored due to annual variations in peak flow datasets.

2.4 Generalized Logistic Distribution (GLO)

The GLO distribution is a versatile tool in flood frequency analysis, offering insights into the probability and magnitude of extreme flood events through its ability to model diverse tail behaviours. Defined by three parameters that are location (ξ), scale(α), and shape (k)—it effectively represented tail behaviour, with positive k indicating a heavy upper tail, negative k implying a finite upper limit, and $k = 0$ suggesting symmetric tails (Hamed & Rao, 2019). GLO is widely applied in studies across Peninsular Malaysia. Che Ilias et al. (2021) identified GLO as the most robust distribution for Region III, covering the southern area, while Mohd Baki et al. (2014) found it more accurate than GEV for Region 5. Similarly, Badyalina et al. (2021) employed GLO in Segamat River analysis, and Bakri (2022) observed that GLO better fit station data compared to GEV and GPA distributions, further underscoring its suitability for regional flood frequency studies.

2.5 L-Moment (LMO)

LMO are widely used for parameter estimation of probability distributions in flood frequency analysis, aiding in the design of flood control infrastructure and hazard management. Introduced by Hosking (1990) as a refinement of Probability Weighted Moments (PWM), LMO provide robust, nearly unbiased estimates, even for skewed or heavy-tailed data, and performed well with small sample sizes (Othman et al., 2025; Marsani et al., 2022; Jan et al., 2016; Vivekanandan, 2015; Hosking, 1990). Studies demonstrated the utility of LMO in hydrological applications: Badyalina et al. (2021) used it for flood frequency analysis in Segamat River, Johor, while Mohd Baki et al. (2014) applied it with GEV and GLO distributions for regional flow analysis in Peninsular Malaysia, identifying GLO as more suitable for estimating design runoff. Hamzah et al. (2020) employed LMO in Pelabuhan Klang to assess GEV and GPA models for high tide data, and Bakri (2022) applied LMO for extreme rainfall analysis across 28 rain gauge stations, concluding that GEV was the best fit for the data.

Performance measurement is essential for evaluating flood frequency models and was categorized into accuracy performance metrics, the L-Moment Ratio Diagram (LMRD), and goodness-of-fit (GOF) tests. Accuracy metrics, such as MAE, MAPE, RMSE, RMSPE, and R^2, assess the alignment between

https://doi.org/10.24191/mij.v6i2.4654

observed and estimated flood data, with lower error values and higher R^2 indicate better model performance. Studies, such as those by Badyalina et al. (2021) and Hassim et al. (2022), applied these metrics to evaluate distributions like GEV, GPA, and GLO, often ranking GPA highest for certain datasets. The LMRD method examined L-moment ratios to determine the most suitable distribution, with studies like Prahadchai et al. (2024) concluding that GEV was optimal in specific regions. GOF tests, including the Anderson-Darling and Kolmogorov-Smirnov tests, were used to validate model suitability, with research by Acharya and Joshi (2020) and Yusoff et al. (2022) consistently finding GEV as the best fit for datasets. Collectively, these methods, combined with the robust application of L-moments, confirm the suitability of GEV, GPA, and GLO distributions for flood frequency analysis in Malaysia, supporting effective flood control and risk mitigation. This study aims to build on advancements in FFA by identifying the most suitable probability distribution for modelling flood events in Peninsular Malaysia, utilizing L-moments for parameter estimation.

3. METHODOLOGY

3.1 Study Area

This study focuses on the station Triang River (2920432) as shown in Fig. 1, located in Jelebu, Malaysia, which was strategically chosen for its susceptibility to recurrent flooding due to monsoon seasons impacting both rural and urban populations (Syukri, 2021). Streamflow data spanning 63 years (1961–2023), obtained from the Department of Irrigation and Drainage Malaysia, provide a robust basis for flood frequency analysis. The station's long-term peak flow records enable accurate parameter estimation, validating the use of extreme value distributions for modelling flood magnitudes. The region's diverse hydrological characteristics and extensive river systems further enhance the study's relevance, with findings expected to have significant implications for flood risk management, infrastructure planning, and disaster preparedness.

Fig. 1. Location of study in Peninsular Malaysia

3.2 L-Moment (LMO)

LMO, derived from PWM, are widely used in hydrological studies due to their robustness and ability to determine the scale and shape of probability distributions (Hamed & Rao, 2019). LMO offer advantages over traditional moments, including reduced sensitivity to outliers and increased effectiveness with small sample sizes. It is particularly valuable in estimating flood recurrence intervals, drought severity, and precipitation trends. By providing reliable estimates of statistical characteristics, LMO enhance the accuracy of fitting probability distributions to hydrological data, improving risk assessment and water resource management. The expression for PWM to estimate the r^{th} LMO was introduced by Hosking (1990) described as follows:

https://doi.org/10.24191/mij.v6i2.4654

$$b_r = \frac{1}{n}\binom{n-1}{r}^{-1} \sum_{i=r+1}^{n} \binom{i-1}{r} x_{i:n} \tag{1}$$

where $x_{i:n}$ was the ordered reading of streamflow, was the PWM, n was the sample size. Thus, the first four moments of an unbiased sample estimator were given as followed:

$$b_0 = \frac{1}{n}\sum_{i=1}^{n} x_{i:n} \tag{2}$$

$$b_1 = \frac{1}{n}\sum_{i=2}^{n} \frac{(i-1)}{(n-1)} x_{i:n} \tag{3}$$

$$b_2 = \frac{1}{n}\sum_{1=3}^{n} \frac{(i-1)(i-2)}{(n-1)(n-2)} x_{i:n} \tag{4}$$

$$b_3 = \frac{1}{n}\sum_{i=4}^{n} \frac{(i-1)(i-2)(i-3)}{(n-1)(n-2)(n-3)} x_{i:n} \tag{5}$$

The first four sample estimates for LMO are the mean of distribution (l_1), the measure of scale (l_2), the measure of skewness (l_3), and the measure of kurtosis (l_4) respectively, which are referred to as:

$$l_1 = b_0 \tag{6}$$

$$l_2 = 2b_1 - b_0 \tag{7}$$

$$l_3 = 6b_2 - 6b_1 + b_0 \tag{8}$$

$$l_4 = 20b_3 - 30b_2 + 12b_1 - b_0 \tag{9}$$

LMO can also be used to derive the LMO ratios, which were analogous to conventional moment ratios (Hosking, 1990). Hence, the samples of the LMO ratio are given as:

$$t_2 = \frac{l_2}{l_1} \tag{10}$$

$$t_3 = \frac{l_3}{l_2} \tag{11}$$

$$t_4 = \frac{l_4}{l_1} \tag{12}$$

3.2.1 Generalized Extreme Value (GEV) Distribution

The GEV distribution has been widely used for the analysis of extreme events. The GEV distribution is widely applied in the frequency analysis of both flood and drought phenomena, (Zeng et al., 2015). The probability density function of the GEV distribution is as follows (Hamed & Rao, 2019):

$$f(x) = \frac{1}{\alpha}\left[1 - k\left(\frac{x-\xi}{\alpha}\right)\right]^{-\frac{1}{k}-1} \exp\left\{-\left[1 - k\left(\frac{x-\xi}{\alpha}\right)\right]^{-1/k}\right\} \tag{13}$$

where ξ is the location parameter, $\alpha > 0$ is the scale parameter, k is the shape parameter.

https://doi.org/10.24191/mij.v6i2.4654

The cumulative density function of the GEV is denoted as $F(x)$.

$$F(x) = \exp\left\{-\left[1 - k\left(\frac{x-\xi}{\alpha}\right)\right]^{-1/k}\right\} \tag{14}$$

Consequently, the parameters ξ, α and k associated with GEV distribution can be estimated as outlined in Hosking et al. (1985).

$$\hat{k} = 7.8590 + 2.9554C^2 \tag{15}$$

$$\hat{\alpha} = \frac{l_2\hat{k}}{\Gamma(1+\hat{k})(1-2^{-k})} \tag{16}$$

$$\hat{\xi} = l_1 + \frac{\hat{\alpha}}{\hat{k}}\left[\Gamma(1+\hat{k}) - 1\right] \tag{17}$$

where $C = \frac{2}{3+t_3} - \frac{\log 2}{\log 3}$.

3.2.2 Generalized Pareto Distribution (GPA)

The GPA has been used in hydrology for modelling the probability of an extreme flood occurrence. This distribution is especially effective for fitting high-accuracy tail data in flood distributions. The probability density function of the three-parameters GPA distribution is given by:

$$f(x) = \frac{1}{\alpha}\left[1 - \frac{k}{\alpha}(x-\xi)\right]^{\frac{1}{k}-1} \tag{18}$$

where ξ represented the location parameter, α the scale parameter, and k the shape parameter (Hamed & Rao, 2019).

The equation below depicts the cumulative distribution function of GPA given by:

$$F(x) = 1 - \left[1 - \frac{k}{\alpha}(x-\xi)\right]^{\frac{1}{k}} \tag{19}$$

where ξ is the location parameter, $\alpha > 0$ is the scale parameter, k is the shape parameter.

Thus, the three parameters of ξ, α and in the GPA distribution can be estimated as in Hamed and Rao (2019).

$$\hat{k} = \frac{1-3t_3}{1+t_3} \tag{20}$$

$$\hat{\alpha} = l_2(1+\hat{k})(2+\hat{k}) \tag{21}$$

$$\hat{\xi} = l_1 - l_2(2+\hat{k}) \tag{22}$$

3.2.3 Generalized Logistic Distribution (GLO)

The GLO distribution is a statistical distribution used to estimate the highest annual flood peak values. It is mainly adopted to estimate the parameters of data sets that exhibit a distribution pattern similar to the logistic distribution. This method is quite useful in predicting the frequency and severity of extreme flooding events. Hence, the probability density function of this distribution can be expressed as follows:

https://doi.org/10.24191/mij.v6i2.4654

$$f(x) = \frac{1}{\alpha}\left[1 - \frac{k}{\alpha}(x-\xi)\right]^{\frac{1}{k}-1} \tag{23}$$

where x is a random variable, ξ is the location parameter, α is the scale parameter, and k is shape parameter.

Next, the cumulative density function for GLO is given by:

$$F(x) = 1 - \left[1 - \frac{k}{\alpha}(x-\xi)\right]^{\frac{1}{k}} \tag{24}$$

where x is a random variable, ξ is the location parameter, α is the scale parameter, and k is shape parameter.

Therefore, the three parameters of ξ, α and k in GLO distribution can be estimated as follows (Hamed & Rao, 2019):

$$\hat{k} = -t_3 \tag{25}$$

$$\hat{\alpha} = \frac{l_2}{\Gamma(\hat{k})\left[\Gamma\left(1-\hat{k}\right) - \Gamma(2-\hat{k})\right]} \tag{26}$$

$$\hat{\xi} = l_1 - \frac{\hat{\alpha}}{\hat{k}} + \hat{\alpha}\hat{\Gamma}(\hat{k})\Gamma(1-\hat{k}) \tag{27}$$

3.2.4 Gringorten Plotting Position

The Gringorten plotting position formula, introduced by Gringorten (1963), is a method for effectively plotting ordered observations, particularly useful for small sample sizes (less than 20). The formula, $P_i = \frac{i-0.44}{n+0.12}$ calculates the plotting position P_i for the i^{th} ranked value in a dataset of size n. This method is favoured for its accuracy and effectiveness in data visualization and probability plotting.

3.2.5 Quantile Estimate using L-Moment

Quantile estimates are computed after the estimation of parameters related to different return periods. The probability of non-exceedance is defined as $\mathrm{F} = 1 - \frac{1}{\mathrm{T}}$ (Acharya & Joshi, 2020), where T is the return period, such as T = 2, 10, 50 and 100 years. Thus, the quantile estimates for the GEV using L-moments are given as shown below:

$$x_T = \xi + \frac{\alpha}{k}\{1 - [-\log(F)]\}^k \tag{28}$$

where ξ is the location parameter, k is the shape parameter, and F is the probability of non-exceedance.

Next, the quantile estimates for the GPA using L-moments are given as:

$$x_T = \xi + \frac{\alpha}{k}[1 - (1-F)]^k \tag{29}$$

where ξ is the location parameter, α is the scale parameter, $k > 0$ is the shape parameter, and T is the return period of formula of probability of non-exceedance $\mathrm{F} = 1 - 1/\mathrm{T}$.

Lastly, the quantile estimates for the GLO using L-moments is:

$$x_T = \xi + \frac{\alpha}{k}[1 - \{(1-F)/F\}^k] \tag{30}$$

where x is random variable, ξ is the location parameter, α is the scale parameter, and k is the shape parameter, and F is the probability of non-exceedance.

https://doi.org/10.24191/mij.v6i2.4654

3.3 Performance Measurement

3.3.1 Accuracy Measure Performance

Five accuracy performance measures used in this research are the mean absolute error (MAE), the mean absolute percentage error (MAPE), the root mean square error (RMSE), the root mean square percentage error (RMSPE), the coefficient of determination (R^2) and Euclidean distance (d).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|F(y_i) - F(\hat{y}_i)| \tag{31}$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{F(y_i) - F(\hat{y}_i)}{F(y_i)}\right| \tag{32}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(F(y_i) - F(\hat{y}_i))^2}{n}} \tag{33}$$

$$RMSPE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{F(y_i) - F(\hat{y}_i)}{F(y_i)}\right)^2} \times 100\% \tag{34}$$

$$R^2 = \frac{\sum_{i=1}^{n}F(y_i) - F(\bar{y}_i)^2}{\sum_{i=1}^{n}F(\hat{y}_i) - F(\bar{y}_i)^2 + \sum_{i=1}^{n}F(y_i) - F(\hat{y}_i)^2} \tag{35}$$

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{36}$$

where n is the number of observations, $\mathrm{F}(y_i)$ represented the actual values, $\mathrm{F}(\hat{y}_i)$ represented the predicted values for the i^{th} observation, and $\mathrm{F}(\bar{y}_i)$ represented the mean of the actual values.

MAE measures the average size of prediction errors by taking the absolute difference between predicted and actual values. It is straightforward, treats all errors equally, and is less affected by extreme values, making it useful for general error assessment. However, it doesn't emphasize larger errors, which may limit its use in some cases. MAPE shows the average error as a percentage of actual values, making it easy to interpret across different data scales. It's common in forecasting but can be unreliable when actual values are near zero, as this causes large or undefined percentage errors. RMSE averages the squared differences between predicted and actual values, then takes the square root. It highlights larger errors due to squaring, which is helpful when big mistakes are costly. However, it's sensitive to outliers, which can skew results. RMSPE is like RMSE but expresses errors as percentages of the actual values. It is useful for comparing performance across datasets with different scales but shares MAPE's issue with near-zero values, making it less reliable in those cases. R^2 indicates how much of the variation in the data is explained by the model, with values closer to 1 showing a better fit. It is widely used in regression to gauge model quality but does not reveal the size of prediction errors and can mislead in complex models. Euclidean distance in L-moment ratio diagrams quantifies the separation between two points in a two-dimensional space, typically defined by L-moment ratios like L-skewness (t_3) and L-kurtosis (t_4). In this application, it measures how closely a sample's L-moment ratios match those of theoretical distributions.

3.3.2 L-Moment Ratio Diagram (LMRD)

The LMRD, proposed by Hosking and Wallis (1997), is a useful tool for determining an appropriate distribution that accurately depicted the catchment's streamflow series. To create a ratio diagram, one must have a straightforward explicit expression for t_4 in terms of t_3 for the chosen probability distributions. The following types of polynomial approximations are used:

https://doi.org/10.24191/mij.v6i2.4654

$$t_4 = A_0 + A_1 t_3 + A_2(t_3)^2 + A_3(t_3)^3 + A_4(t_3)^4 + A_5(t_3)^5 + A_6(t_3)^6 + A_7(t_3)^7 + A_8(t_3)^8 \quad (37)$$

Table 1. Polynomial approximations of t_4 as a function of t_3 based on L-moment method

	GPA	GEV	GLO
A_0	0	0.10701	0.16667
A_1	0.20196	0.11090	0
A_2	0.95924	0.84838	0.83333
A_3	-0.20096	-0.06669	0
A_4	0.04061	0.000567	0
A_5	0	-0.04208	0
A_6	0	0.03763	0
A_7	0	0	0
A_8	0	0	0

Table 1 provided the coefficients of Ak for the GEV, GPA, and GLO distributions, respectively, based on L-moment, where $k = 0,1,2,\ldots,8$. By substituted t_3 over the range $-0.5 \leq t_3 \leq 0.9$, yield values of t_4 for each distribution respectively. The range selected because mostly from fall between the range -0.5 and 0.9. The constructed LMRD for GEV, GPA, and GLO distributions respectively will be illustrated in Fig. 2.

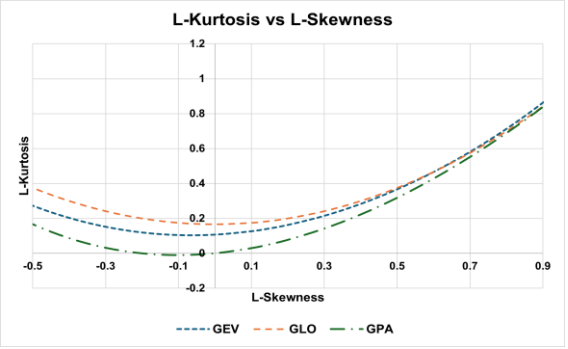

Fig. 2. L-Moment Ratio Diagram

3.3.3 Goodness of Fit Test

Anderson Darling Test

The Anderson-Darling (AD) test is a statistical tool designed to evaluate the goodness-of-fit between a dataset and a specified distribution, making it particularly valuable for flood frequency analysis. Its sensitivity to deviations in the distribution's tails enhanced its effectiveness in analysing extreme events like flooding.

H_0: The data follow the specified distribution.
H_1: The data do not follow the specified distribution.

P-value must be less than α to reject the null hypothesis. Hence, AD test statistic is as follows:

$$AD = -n - \frac{1}{n}\sum_{i=1}^{n}\left[(2i-1)ln\left(F(X_{(i)})\right) + \ln(1 - F\left(X_{(n-i+1)}\right))\right] \quad (38)$$

where n is the sample size and $F(X(i))$ is the empirical cumulative distribution function of the ordered sample $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$.

https://doi.org/10.24191/mij.v6i2.4654

Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov (K-S) test is a nonparametric statistical method widely used to compare a dataset with a specified distribution (Hamed & Rao, 2019).

H_0: The data follow the specified distribution.
H_1: The data do not follow the specified distribution.

P-value must be less than α to reject null hypothesis. Hence, the test statistic is as follows:

$$D_{Nj} = \max |F_N(x) - F_0(x)| \tag{39}$$

where N_j is the cumulative number of sample events at class limit j, and the values of $F_N(x)$ are estimated; $1/k, 2/k, \ldots$, etc. are the values of $F_0(x)$, where k is the number of class intervals.

4. RESULTS AND DISCUSSION

4.1 LMO

Table 2 presents the initial four elements involved in the computation of L-moments. These components are mean (l_1), scale (l_2), skewness (l_3), and kurtosis (l_4), provide insights into the statistical characteristics of the flood data at this station.

Table 2. First Four L-Moment Component

l_1 (Mean)	l_2 (Scale)	l_3 (Skewness)	l_4 (Kurtosis)
4.3715365	1.6051485	0.2264257	0.2397717

The first component l_1, (4.3715), represents the central tendency, reflecting the average magnitude of flood events. The second component l_2, (1.6051), indicates the variability around the mean, demonstrating moderate variation in flood magnitudes. The third component l_3, (0.2264) signifies a minor positive skewness in the data distribution, suggesting that high-magnitude floods are infrequent but possible. Finally, the l_4 (0.2398), reflects light tails in the distribution, implying that extreme flood events are rare.

4.2 Parameter Estimation

Table 3 shows the estimated parameters for GEV, GPA and GLO distributions using the LMO method. From the table, GPA distribution exhibits the highest $\hat{\alpha}$ value (6.0547) among the distributions. The GLO distribution has the highest value $\hat{\xi}$ (4.0027). This value signifies that the flood magnitude modelled by the GLO distribution is the highest. The GLO distribution also shows a negative $\hat{k}$ value that specifies a bounded tail distribution. This value suggests that the model has more frequent events with fewer extremes, which may require consistent but moderate flood management strategies (Hamed & Rao, 2019).

Table 3. Estimated Distribution Parameters using LMO Methods

Distribution	$\hat{\alpha}$	$\hat{\xi}$	$\hat{k}$
GEV	2.41087887	3.08395543	0.04542615
GPA	6.0547059	0.34982765	0.5055058
GLO	1.5531227	4.0027214	-0.1410622

4.2 Gringorten Plotting Position

Fig. 3 presents the Gringorten Plotting Position graph, comparing observed peak flow measurements with predicted values from the GEV, GPA, and GLO distributions. The x-axis represents the Gringorten

https://doi.org/10.24191/mij.v6i2.4654

plotting position probability, while the y-axis shows peak flow discharge in m^3/s. The graph indicates that GPA and GLO distributions closely align with observed values, demonstrating a strong fit to the data (Samat & Othman, 2023). To assess the accuracy of the plotting positions relative to the fitted distributions, the Euclidean distance metric is used to provide a quantitative measure of goodness-of-fit by evaluating deviations between observed data and theoretical quantiles from the GEV, GPA, and GLO distributions.

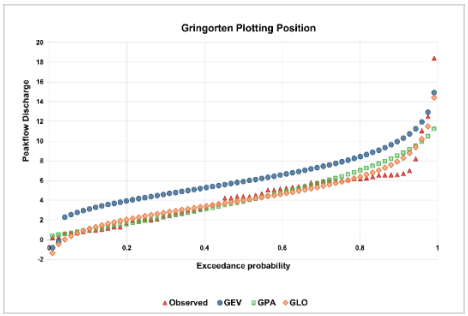

Fig. 3. Gringorten Plotting Position

Table 4. Euclidean Distance of candidate distribution

Distribution	Euclidean Distance
GEV	7.050357
GPA	9.196464
GLO	6.224768

Table 4 summarizes the Euclidean distances between observed and predicted values for the GEV, GPA, and GLO distributions using Gringorten plotting positions. The results show that the GLO distribution, with the smallest Euclidean distance of 6.2248, is the most suitable model for flood frequency analysis at the analysed station. The GEV distribution, with a Euclidean distance of 7.0504, demonstrates a better fit than the GPA distribution but is less accurate than the GLO distribution. The GPA distribution, having the highest Euclidean distance of 9.1965, is the least suitable model among the three distributions.

4.3 Performance Measurement

4.3.1 Accuracy Performance Measures

Table 5. Test Performance Measurement for Distributions

Distribution	GEV	GPA	GLO
MAPE	0.2275573	0.1270529	0.3044755
MAE	0.536894	0.5826525	0.5396244
RMSE	0.8882615	1.158646	0.784247
RMSPE	65.05872	19.93939	99.92112
R^2	0.9180464	0.858907	0.9366164

Regarding performance metrics in Table 5, the GPA distribution shows the lowest MAPE (0.1271), indicating that it has the smallest average percentage error, thus providing the most accurate predictions on average among the distributions. For MAE, the GEV (0.5369) performs better than the other distributions, indicating more accurate absolute estimates. The RMSE values highlight that the GLO distribution (0.7842) has the lowest value, indicating it performs best in minimising large prediction errors. For the metric

https://doi.org/10.24191/mij.v6i2.4654

RMSPE, the GPA distribution (19.9394) has the lowest value, which suggests it performs better than GEV (65.05872) and GLO (99.9211). Finally, the R^2 values indicate that the GLO distribution (0.9366) explains the highest proportion of variance in the data, followed by GEV (0.9180) and the GPA distribution (0.8589).

4.3.2 L-Moment Ratio Diagram

Table 6 presents L-moment Ratios (LMR) for a river station in Peninsular Malaysia, capturing key hydrological attributes using three LMO statistics: t_2 (L-CV), t_3 (L-skewness), and t_4 (L-kurtosis), which describe variability, asymmetry, and tail heaviness, respectively. The t_2 value of 0.3672 indicates moderate variability, balancing consistency and dispersion. The $t_3$value of 0.1411 reflects moderate right skewness, suggesting a slightly longer tail on the right side. The t_4 value of 0.1494 indicates moderately heavy tails, implying the potential presence of extreme values without being overly pronounced. These LMR statistics provide a comprehensive understanding of the hydrological data, enhancing flood prediction models. The LMRD for the three distributions is depicted in Fig. 4.

Table 6. LMR

L-CV (t_2)	L-Skewness (t_3)	L-Kurtosis (t_4)
0.3672	0.1411	0.1494

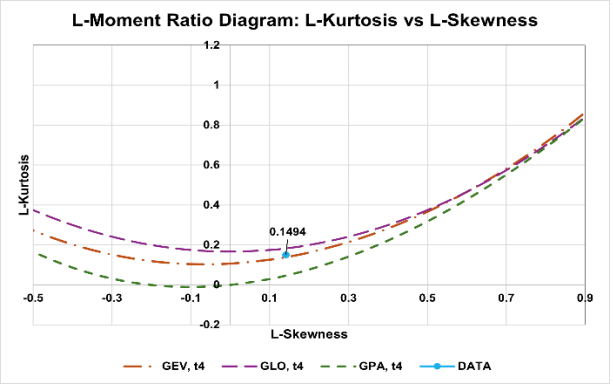

Fig. 4. LMRD

Fig. 4 shows the LMRD depicting the relationship between L-skewness (along the x-axis) and L-kurtosis (along the y-axis) for the observed data along with three theoretical extreme value distributions: GEV, GPA and GLO distributions. Theoretical curves for each distribution are shown in the diagram, and the observed data point (0.1494) is plotted at a position determined by its respective L-Skewness and L-Kurtosis values.

The observed data point (0.1494) is closest to the curve of the GEV distribution, which implies that the GEV model is the most suitable for the data. The observed data show more scatter from the GLO and GPA curves. This indicates that both distributions are less accurate in modelling the observed data. The major outcome from the LMRD reveals that the GEV distribution fits best for this dataset, and is effective in flood frequency analysis, promising in correctly representing the statistical properties of observed data.

4.3.3 Goodness of Fit Test

Table 7 presents the p-value from the GOF tests for the GEV, GPA and GLO distributions. In terms of GOF, the AD test p-value for GEV (0.3979) indicates the best fit, while GPA has the smallest AD p-

https://doi.org/10.24191/mij.v6i2.4654

value (0.000009524), suggesting a poor fit to the data. Similarly, the K-S test p-value further confirms that GLO (0.5923) fits the data well, while the GPA distribution (0.1567) performs worse than GEV and GLO.

Table 7. P-value of GOF Test for Each Distributions

Distribution	GEV	GPA	GLO
AD	0.3979	0.000009524	0.3684
K-S	0.4244	0.1567	0.5923

4.4 Measurement Ranking

Table 8 puts the Generalized Extreme Value (GEV) distribution at the forefront as the most suitable model for flood frequency model for this study. This can be supported by its outstanding performance in the GOF tests, specifically the AD and K-S tests, combined with its good results in other accuracy measures. Besides, GEV obtained the highest cumulative rank score in Table 8, further proving its appropriateness for this station. By contrast, the GLO and GPA distributions are the least favourable, with their total ranks scores being the lowest among the three. The present study concludes that the GEV distribution is the best and most appropriate model for the purpose of flood frequency analysis in the river system of Peninsular Malaysia.

Table 8. Rank Score for Distributions

Distribution	GEV	GPA	GLO
AD	3	1	2
K-S	2	1	3
MAPE	2	3	1
MAE	3	1	2
RMSE	2	1	3
RMSPE	2	3	1
R^2	2	1	3
Total Score	16	11	15

4.5 Quantile Estimate

Table 9 demonstrates the GEV distribution's effectiveness in modeling flood magnitudes, showing a consistent increase in discharge with longer return periods. For a 10-year period (p = 0.90), the estimated flood discharge is 8.2412 m^3/s, rising to 11.7045 m^3/s, for 50 years (p = 0.98) and 13.0923 m^3/s for 100 years (p = 0.99). Compared to other distributions, GEV provides slightly lower estimates than GLO for shorter periods but higher than GPA, while offering more conservative predictions for longer periods, such as 100 years. This makes GEV particularly suitable for applications requiring conservative flood risk assessments, ensuring realistic and consistent predictions. Its reliability across return periods underscores its value in long-term flood risk management and infrastructure design capable of withstanding extreme hydrological events.

Table 9. Quantile Estimate Based on Return Periods

Return Period (Years)	Probability (p)	Estimated Flood Discharge (m^3/s)		
		GEV	GPA	GLO
10	0.90	8.241226	8.228045	8.382849
50	0.98	11.704503	12.111101	10.652548
100	0.99	13.092262	14.074898	11.153635

https://doi.org/10.24191/mij.v6i2.4654

5. CONCLUSION

This research presents a comprehensive framework for understanding and managing flood risks in Peninsular Malaysia, a region highly vulnerable to monsoon-induced flooding. By employing advanced statistical methods, including L-moments and extreme value distributions (GEV, GPA, and GLO), the study presents a robust model for predicting flood magnitude and frequency. The findings highlight the efficiency of these techniques in analysing long-term records, estimating return periods, and identifying suitable probability distributions, with the Generalized Extreme Value (GEV) distribution emerging as the most reliable model. The GEV distribution consistently outperforms others in goodness-of-fit tests and accurately predicts flood magnitudes, providing valuable quantile estimates for 10-, 50-, and 100-year return periods at 8.24 m^3/s, 11.70 m^3/s, and 13.09 m^3/s, respectively. The research emphasizes the importance of integrating these findings into regional planning and flood management strategies to address challenges from urbanization and climate change, ultimately improving infrastructure resilience and community preparedness.

The study offers several recommendations to enhance flood management and preparedness. Policymakers and engineers are advised to prioritize the use of the GEV distribution for designing flood control structures, as it is identified as the most effective model for extreme flood events. Utilizing quantile estimates in high-risk areas can aid in targeted strategies and efficient resource allocation. Collaboration between government bodies and academic institutions is crucial for improving real-time streamflow data accuracy, which will enhance the precision of flood frequency models and future flood management efforts. Future research should apply the L-Moment method to regions with varying hydrological conditions or limited data to validate its robustness and adaptability. Expanding the analysis to other river systems and regions in Malaysia is also recommended, particularly to assess the impacts of climate change on flood magnitudes and return periods, thereby providing critical insights for scenario planning and improving preparedness strategies.

6. ACKNOWLEDGEMENTS/FUNDING

The authors acknowledge the financial support from the Ministry of Higher Education Malaysia through Fundamental Research Grant Scheme FRGS-EC/1/2024/STG06/UITM/02/12.

7. CONFLICT OF INTEREST STATEMENT

The authors confirm that this research was conducted in the absence of any self-benefits, commercial or financial conflicts, and they declare the absence of conflicting interests with the funders.

8. AUTHORS' CONTRIBUTIONS

Nur Amira Abdul Aziz conducted the main research activities, performed data analysis, and contributed significantly to the writing and revising the manuscript. Basri Badyalina conceptualized the research framework, supervised the overall project, contributed to the methodology design, and finalized the manuscript for submission. Ani Shabri provided expertise in mathematical modelling and statistical validation and contributed to reviewing and refining the analytical methods. Fatin Farazh Ya'acob assisted in the literature review, data interpretation, and contributed to manuscript editing. Ira Syaqira Sukimin contributed to the linguistic editing of the manuscript, improving language clarity, and assisted in preparing the manuscript for submission.

https://doi.org/10.24191/mij.v6i2.4654

REFERENCES

Abid, S., Sulaiman, N., Wei, C., & Nazir, U. (2021). Flood vulnerability and resilience: Exploring the factors that influence flooding in Sarawak. IOP conference series: earth and environmental science. https://doi.org/10.1088/1755-1315/802/1/012059

Acharya, B., & Joshi, B. (2020). Flood frequency analysis for an ungauged Himalayan River basin using different methods: a case study of Modi Khola, Parbat, Nepal. *Meteorology Hydrology and Water Management. Research and Operational Applications*, *8*(2), 46-51.

Ahmad, A. M., Romali, N. S., & Sulong, S. (2023). Flood frequency analysis of annual maximum stream flows for Kuantan River Basin. *AIP Conference Proceedings, 2688*, 040002. https://doi.org/10.1063/5.0111746

Ali, S., & Rahman, A. (2022). Development of a kriging-based regional flood frequency analysis technique for South-East Australia. *Natural Hazards*, *114*(3), 2739-2765. https://doi.org/10.1007/s11069-022-05488-4

Badyalina, B., Mokhtar, N. A., Jan, N. A. M., Hassim, N. H., & Yusop, H. (2021). Flood frequency analysis using L-moment for Segamat river. *Matematika*, 47-62. https://doi.org/10.11113/matematika.v37.n2.1332

Badyalina, B., Mokhtar, N. A., Ya'acob, F. F., Ramli, M. F., Majid, M., & Jan, N. A. M. (2022). Flood Frequency Analysis Using L-Moments at Labis in Bekok River Station. *Applied Mathematical Sciences*, *16*(11), 529-536. https://doi.org/10.12988/ams.2022.917192

Badyalina, B., & Shabri, A. (2015). Flood frequency analysis at ungauged site using group method of data handling and canonical correlation analysis. *Modern Applied Science*, *9*(6), 48.

Bakri, S. N. F. A. (2022). At-Site and Regional Frequency Analysis of Extreme Rainfall Modelling in Peninsular Malaysia. *Applied Mathematics and Computational Intelligence (AMCI)*, *11*(2), 534-549.

Campos-Aranda, D. F. (2016). Fitting of the GEV, GLO and GPA distributions with trimmed L moments (1, 1) in 21 annual flood records of the hydrological region no. 10 (Sinaloa), Mexico. *Agrociencia*, *50*(1), 15-31.

Chancel, L., Bothe, P., & Voituriez, T. (2023). *Climate Inequality Report 2023*. World Inequality Lab.

Che Ilias, I. S., Wan Zin, W. Z., & Jemain, A. A. (2021). Regional frequency analysis of extreme precipitation in Peninsular Malaysia.

Dalrymple, T. (1960). *Flood-frequency analyses*. US Government Printing Office.

Gringorten, I. I. (1963). A plotting rule for extreme probability paper. *Journal of Geophysical Research, 68*(3), 813–814. https://doi.org/10.1029/JZ068i003p00813

Hamed, K., & Rao, A. R. (2019). *Flood frequency analysis*. CRC press.

Hamzah, F. M., Tajudin, H., Abdullah, S. M. S., Toriman, E., & Syed Abdullah, S. M. (2020). Tidal Flooding Frequency Analysis using Partial Duration Series. *International Journal of Advanced Trends in Computer Science and Engineering, 9(1.2)*, 21–27. https://doi.org/10.30534/ijatcse/2020/0591.22020

https://doi.org/10.24191/mij.v6i2.4654

Hassim, N. H., Badyalina, B., Mokhtar, N. A., Abd Jalal, M. Z. H., Zamani, N. D., Lee, C. K., Zainoddin, A. I., Ahmad Fadzil, A. S., & Shaari, N. F. (2022). In situ flood frequency analysis used for water resource management in Kelantan River Basin. *International Journal of Statistics and Probability, 11*(5), 1–8. Retrieved from https://ccsenet.org/journal/index.php/ijsp/article/view/0/47687

Hosking, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society: Series B*, 52, 105–124. https://doi.org/10.1111/j.2517-6161.1990.tb01775.x

Hosking, J. R. M., Wallis, J. R., & Wood, E. F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, *27*(3), 251-261.

Hossain, I., Imteaz, M., & Khastagir, A. (2021). Study of various techniques for estimating the generalised extreme value distribution parameters. IOP Conference Series: Materials Science and Engineering,

Jan, N. A. M., Shabri, A., Ismail, S., Badyalina, B., Abadan, S. S., & Yusof, N. (2016). Three-parameter lognormal distribution: Parametric estimation using L-moment and TL-moment approach. *Jurnal Teknologi (Sciences & Engineering)*, *78*(6-11). https://doi.org/10.11113/jt.v78.9202

Krishna, G. S., & Veerendra, G. (2015). Flood frequency analysis of Prakasam barrage reservoir Krishna district, Andhra Pradesh using Weibull, Gringorten and L-moments formula. *Int. J. Civil. Struct. Env. Infr. Eng. Res. Dev*, *5*.

Legrand, J. (2022). *Simulation and assessment of multivariate extreme models for environmental data* Université Paris-Saclay].

Marsani, M. F., Shabri, A., Badyalina, B., Jan, N. A. M., & Kasihmuddin, M. S. M. (2022). Efficient market hypothesis for Malaysian extreme stock return: Peaks over a threshold method. *Matematika*, 141-155. https://doi.org/10.11113/matematika.v38.n2.1396

Mohd Baki, A., Mat Yusof, D. A., Atan, I., & Mohd Halim, N. F. (2014). Regional flow frequency analysis on Peninsular Malaysia using lmoments. *Jurnal Intelek*, *9*(1), 63-68.

Othman, N. H., Badyalina, B., Izan, S. H. S., Shabri, A., Marsani, M. F., & Ya'acob, F. F. (2025). Hydrological Risk Assessment of the Johor River: Flood Frequency Analysis with L-Moments. *Journal of Computing Research and Innovation*, *10*(2), 144-155.

Prahadchai, T., Busababodhin, P., & Park, J.-S. (2024). Regional flood frequency analysis of extreme rainfall in Thailand, based on L-moments. *Communications for Statistical Applications and Methods*, *31*(1), 37-53. https://doi.org/10.29220/CSAM.2024.31.1.037

Rezali, M. S. N., Ramli, M. W. A., & Bakar, M. A. A. (2025). Analysis of Precipitation Patterns in the East Coast States of Peninsular Malaysia from 1981 to 2019. *Journal of Asian Geography*, *4*(1), 71-81. https://doi.org/10.36777/jag2025.4.1.9

Samat, S., & Othman, N. (2023). Plotting Position for Low Flow Frequency Analysis at Jempol River Streamflow Station. IOP Conference Series: Earth and Environmental Science,

Stedinger, J. R. (1993). Frequency analysis of extreme events. *Handbook of hydrology*.

Syukri, N. (2021). Jelebu Kawasan Terkering Namun Alami Banjir Terburuk di Negeri Sembilan. *BERNAMA*.

https://doi.org/10.24191/mij.v6i2.4654

Vivekanandan, N. (2015). Flood frequency analysis using method of moments and L-moments of probability distributions. *Cogent engineering*, *2*(1), 1018704. https://doi.org/10.1080/23311916.2015.1018704

Yusoff, S., Hamzah, F., & Jaafar, O. (2022). Multiparameter probability distributions of at-site Lmoment-based frequency analysis in Malaysia. *International Journal of Mechanical Engineering*, *7*(4), 4

Zeng, X., Wang, D., & Wu, J. (2015). Evaluating the three methods of goodness of fit test for frequency analysis. *Journal of Risk Analysis and Crisis Response*, *5*(3), 178-187. https://doi.org/10.2991/jrarc.2015.5.3.5

https://doi.org/10.24191/mij.v6i2.4654