Applied Mathematics and Computational Intelligence
Volume 15, No. 1, 2026 [21-48]

A Comparison of Stationary and Non-Stationary Generalized Extreme Value Models with Climatic Covariates in Modeling Rainfall Extremes

Syafrina Abdul Halim[1*], Muhammad Hafiz Abdul Halim Toh Seng Onn[2]

[1]Institut Penyelidikan Matematik, Universiti Putra Malaysia, 43400 Serdang Selangor
[2]Department of Mathematics and Statistics, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

[*] Corresponding author: syafrina@upm.edu.my

Received: 20 March 2025
Revised: 29 August 2025
Accepted: 8 October 2025

ABSTRACT

Extreme rainfall in Peninsular Malaysia is closely linked to seasonality, primarily driven by the southwest monsoon (SWM) period, starting from May to September and northeast monsoon (NEM) period, starting from November to March. Extreme rainfall events may lead to secondary disasters such as floods, landslides and crop damage. The west part of Malaysia also known as Peninsular Malaysia is significantly influenced by large-scale global climate phenomena such as El Niño-Southern Oscillation (ENSO) that could affect the rainfall pattern across this region. Understanding the changing behavior of extreme rainfall and its relationship with ENSO will increase planners' ability to plan for, manage and respond to related flood events. This study investigates the trend and stationarity of extreme rainfall in Peninsular Malaysia by using the Mann-Kendall (MK) trend test and Augmented Dickey-Fuller (ADF) test. Rising trends in extreme rainfall were identified in most part of Peninsular Malaysia. This study also analyzes the suitability of stationary and non-stationary Generalized Extreme Value (GEV) models for modeling rainfall extremes. The non-stationary model integrates the Southern Oscillation Index (SOI) and a linear trend as covariates to capture potential climatic influences on extreme rainfall events. The model performance is assessed using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), enabling quantitative comparison. To assess the uncertainty of each model's parameter, bootstrap method will be applied in this study. Likelihood ratio tests are employed to evaluate the robustness and significance of the models, ensuring the best model is selected. Preliminary results suggest that incorporating SOI and trend improves the model's ability to explain variability in rainfall extremes, offering insights into climatic drivers of extreme events. Extreme rainfall return levels are used to quantify potential flooding risk. This research has practical implications for understanding and predicting rainfall extremes under changing climate conditions.

Keywords: generalized extreme value model, non-stationary modeling, rainfall extremes, southern oscillation index.

1 INTRODUCTION

Extreme rainfall events are a critical challenge in Peninsular Malaysia, where climatic and geographical factors heighten vulnerability. The region experiences distinct seasonal variations

driven by the northeast and southwest monsoons. The northeast monsoon, occurring from November to March, brings heavy and prolonged rainfall, often resulting in widespread flooding and landslides, especially in the east coast states such as Kelantan, Terengganu, and Pahang. Conversely, the southwest monsoon from May to September typically results in drier conditions. These seasonal dynamics underscore the need for a comprehensive analysis of rainfall patterns. This study investigates trends and stationarity in extreme rainfall using the Mann-Kendall test and Augmented Dickey-Fuller test, a robust method for detecting changes in hydrological data. By analyzing these trends and stationarity, the research aims to provide insights into how seasonal variations and long-term climatic changes affect extreme rainfall in Peninsular Malaysia. To understand the behavior of extreme rainfall more effectively, the study employs the Generalized Extreme Value (GEV) distribution, a statistical approach designed to model extreme and significant events. Both stationary and non-stationary GEV models are applied, with the latter incorporating the Southern Oscillation Index (SOI) as a covariate to account for the influence of the El Niño-Southern Oscillation (ENSO). The inclusion of SOI captures the global-scale climate variability that interacts with seasonal factors like the monsoons, shaping local rainfall extremes. By comparing the performance of stationary and non-stationary GEV models, the study evaluates how ENSO phases and seasonal monsoon patterns jointly influence extreme rainfall. These findings will contribute to improve climate resilience, disaster preparedness, and water resource management strategies in Peninsular Malaysia, addressing both seasonal and long-term climate variability.

Extreme Value Theory (EVT) is a subfield of probability and statistics focused on modeling and predicting rare and extreme events, such as hurricanes, floods, and droughts [1]. Developed by statisticians like Fisher and Gnedenko, EVT gained prominence for analyzing extremes in fields such as hydrology, finance, and environmental sciences [2]. EVT is widely used to model extreme rainfall such as the maximum daily, monthly or annual rainfall [3]. Modern advancements in EVT, including Bayesian inference and improved computational techniques, have enhanced its ability to model high-dimensional data and provide accurate forecasts [4]. Key EVT methods include block maxima, peaks over threshold (POT), and the GEV distribution [5]. These approaches are widely applied in environmental sciences to analyze extreme climate events, assess their intensity and frequency, and inform decision-making on climate adaptation and disaster preparedness [6]. Integrating EVT with spatial models and risk frameworks has further improved its utility in managing extreme events and evaluating their impacts on infrastructure, agriculture, and public health [7, 8, 9]. Extreme rainfall events, defined by periods of intense rainfall, are becoming increasingly frequent and severe worldwide due to climate change [10]. These events result in infrastructure damage, flooding, and erosion, with urban areas particularly vulnerable due to altered surface characteristics and inadequate drainage systems [11]. In Malaysia, extreme rainfall during monsoon seasons has led to significant flooding and landslides, with notable events attributed to global warming effects [12, 13]. Regional studies indicate that coastal areas, such as the east coast of Peninsular Malaysia, and highland areas, like Genting Highlands, experience distinct extreme rainfall patterns due to geographical diversity and orographic effects [14, 15, 16].

Recent advancements in climate modeling such as flood forecasting systems is critical for mitigating risks [17, 18]. Public awareness and community-based adaptation practices further enhance resilience to these increasingly severe events [19]. The ENSO is a periodic climate phenomenon characterized by fluctuations in sea surface temperatures (SSTs) and atmospheric pressure across the Pacific Ocean [20]. Southern Oscillation Index (SOI) is a key atmospheric indicator of ENSO, it measures the difference in air pressure between Tahiti, French Polynesia and Darwin, Australia, negative SOI value indicates El Niño while Positive SOI value indicates La Niña [21]. Advances in

satellite technologies, such as the Tropical Atmosphere Ocean (TAO) array, and predictive models have improved ENSO monitoring and forecasting [22, 23]. These tools help mitigate the socio-economic impacts of ENSO-related climate fluctuations, emphasizing the need for interdisciplinary approaches to enhance preparedness [24, 25]. Based on the Fisher-Tippett theorem, the GEV distribution encompasses three types: Gumbel (Type I), Fréchet (Type II), and Weibull (Type III), suitable for analyzing various tail behaviors [26]. It has been applied to assess extreme precipitation, floods, and storm frequencies, providing valuable insights for water resource planning and disaster risk reduction [27, 28]. Recent developments incorporate non-stationary models that account for covariates like ENSO, enhancing predictions under changing climate conditions [29]. A non-stationary GEV model is widely used to analyze and predict climate change in Malaysia particularly for seasonal Monsoon analysis [30]. The GEV distribution's ability to model return levels and quantify risks makes it essential for designing resilient infrastructure and developing climate adaptation strategies [31].

2 MATERIAL AND METHODS

2.1 Data Collection

Rainfall

The rainfall data used in this study were provided from NASA's POWER (Prediction of Worldwide Energy Resources) Data Access Viewer for 12 stations across Peninsular Malaysia from 1990 to 2022. The POWER database provides solar and meteorological datasets that are obtained from a high technology satellite. By focusing on capital cities for each station, the dataset captures the daily precipitation in (mm/day) unit, where the precipitation variable will be replaced with rainfall term in this study. Table 1 summarizes information for all station while Figure 1 illustrates the location of each station across Peninsular Malaysia.

Table 1: Study area information

Stations	Latitude	Longitude
Kangar, Perlis	6.44°N	100.19°E
Alor Setar, Kedah	6.12°N	100.37°E
Georgetown, P. Pinang	5.41°N	100.31°E
Ipoh, Perak	4.6°N	101.08°E
Kota Bharu, Kelantan	6.12°N	102.24°E
K. Terengganu, Terengganu	5.32°N	103.13°E
Kuantan, Pahang	3.82°N	103.34°E
Shah Alam, Selangor	3.07°N	101.52°E
Seremban, N. Sembilan	2.72°N	101.94°E
Malacca City, Melaka	2.2°N	102.25°E
Johor Bahru, Johor	1.46°N	103.76°E
Kuala Lumpur	3.15°N	101.7°E

Figure 1: Location of all 12 stations across Peninsular Malaysia.

ENSO

In this study, the Southern Oscillation Index (SOI) will be the indicator for ENSO. A standardized indicator known as the Southern Oscillation indicator (SOI) is derived from the observed changes in sea level pressure (SLP) between Tahiti (17.65°S, 149.43°W) and Darwin, Australia (46.18°S, 30.42°E). During El Niño and La Niña occurrences, there are significant variations in sea level pressure between the western and eastern tropical Pacific, which can be measured using the SOI. Events of El Niño are associated with a negative SOI, whereas events of La Niña are associated with a positive SOI. The SOI was obtained from NOAA's National Centers for Environmental Information website.

2.2 Data Preprocessing

Preprocessing data for extreme rainfall analysis is crucial to prepare raw rainfall data for statistical modeling, particularly when fitting distributions like the GEV distribution. Replacing missing data with the median value is a common method for imputation, especially for skewed distributions [32].

$$\text{Median}(X)=\begin{cases} X_{\frac{n+1}{2}} & \text{if } n \text{ is odd} \\ \frac{X_{\frac{n}{2}}+X_{\frac{n}{2}+1}}{2} & \text{if } n \text{ is even} \end{cases} \quad (1)$$

where X is the daily rainfall data and n is the number of non-missing values.

2.3 Mann-Kendall Trend Test

In this study, the Mann-Kendall (MK) method will be applied to analyze trends in extreme rainfall during seasonal (SWM and NEM) and nonseasonal period, setting a significance level at 5% under hypothesis test as shown in (2):

H_0: $There\ is\ insignificant\ trend\ in\ data\ over\ time$
H_1: $There\ is\ a\ significant\ trend\ in\ data\ over\ time$

$$Z = \begin{cases} \frac{S-1}{\sqrt{Var(S)}} & if\ S > 0 \\ 0 & if\ S = 0 \\ \frac{S+1}{\sqrt{Var(S)}} & if\ S < 0 \end{cases} \quad (2)$$

The magnitude of trend is predicted by the Sen's estimator. The slope T_i of all data pairs is computed as

$$T_i = \frac{x_j - x_k}{j-k} \quad \text{for } i = 1,2,\ldots,n \quad (3)$$

where x_j and x_k are considered as data values at time j and k ($j > k$) correspondingly.

2.4 Augmented Dickey-Fuller Test

The Augmented Dickey-Fuller (ADF) test will be conducted on the same dataset via t-statistics as shown in (4) with the hypothesis test:

H_0: $The\ series\ is\ nonstationary\ (\gamma = 0)$
H_1: $The\ series\ is\ stationary\ (\gamma \neq 0)$

$$t_\gamma = \frac{\hat{\gamma} - \gamma}{se(\hat{\gamma})} \quad (4)$$

where $\hat{\gamma}$ represents the estimate of γ and $se(\hat{\gamma})$ is the coefficient standard error. Once t_γ statistics is calculated, it can be compared with the p-value at significant level of (0.05). If the value of t_γ is less than 0.05, then null hypothesis will be rejected to conclude that there is no unit root, and the time series of extreme rainfall is stationary.

2.5 Extreme Value Analysis

Block maxima method

In the block maxima method, a block size is defined, and the maximum events are selected for each block. The block size, time interval in this study, is separated by 2 blocks which are Southeast Monsoon (SWM) and Northeast Monsoon (NEM) and the sample is defined by selecting the maximum rainfall value for each block by year as indicated in (5),

$$M_i = \max\left(X_{t_1}, X_{t_2}, \dots, X_{t_n}\right) \tag{5}$$

where M_i is the maximum rainfall in a defined block. In this study, all these 3 blocks will be blocked for all 12 stations of daily rainfall datasets across Peninsular Malaysia from 1990 to 2022.

Stationary GEV model

Extreme rainfall data for SWM, NEM and nonseasonal blocks were fitted by using GEV distribution. The cumulative distribution function of the GEV distribution can be expressed as follow

$$\mathrm{G}(x;\mu,\sigma,\xi) = \begin{cases} \exp\left\{-\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right)\right\}, & \text{for } \xi \neq 0 \\ \exp\left\{-\exp\left(-\frac{x-\mu}{\sigma}\right)\right\}, & \text{for } \xi = 0 \end{cases} \tag{6}$$

where μ, σ, ξ denote the location, scale, and shape parameters, respectively. x denotes extreme rainfall whether during SWM, NEM and nonseasonal throughout 1990 until 2022. The GEV has three distribution types: Frechet, Weibull and Gumbel. In the stationary hypothesis, the location, scale, and shape parameters are assumed to be fixed over time.

Non-stationary GEV model (NSGEV)

The NSGEV model is a statistical model used to analyze extreme events over time and its parameters model are allowed to vary over time, capturing the extreme event behaviour. The objective of this study is to explore the impacts of ENSO on extreme precipitation by developing the non-stationary GEV distribution models. The NSGEV cumulative distribution function is expressed as follows,

$$G(x;\mu(t),\sigma(t),\xi)=\begin{cases}\exp\left\{-\left(1+\xi\left(\frac{x-\mu(t)}{\sigma(t)}\right)^{-\frac{1}{\xi}}\right)\right\}, & \text{for } \xi\neq 0\\ \exp\left\{-\exp\left(-\frac{x-\mu(t)}{\sigma(t)}\right)\right\} & \text{for } \xi=0.\end{cases} \quad (7)$$

In the analysis of non-stationarity, parameters are expressed as functions of covariates over time and the SOI (as an indicator for ENSO) will be the covariate in the location and scale parameter's function. However, the shape parameter remains constant. For each station, a total of ten models will be derived, including one GEV and nine NSGEV models during SWM, NEM and nonseasonal period. Four model types are briefly outlined as follows:

Type I: Stationary GEV (GEV0) model, with all parameters constant over time. X_t, the extreme rainfall in year t, might be suitable as stated in (8):

$$X_t \sim GEV(\mu,\sigma,\xi), \quad (8)$$

Type II: Non-stationary GEV (NSGEV) model, with a non-stationarity in location parameter only,

$$X_t \sim GEV(\mu(t),\sigma,\xi), \quad (9)$$

where

$$\mu(t)=\beta_0+\beta_1 SOI(t), \quad (10)$$

where $SOI(t)$ denotes the standardized value of the SOI in year t. However, a possible trend in extreme rainfall through time, suggesting

$$\mu(t)=\beta_0+\beta_1(t), \quad (11)$$

where $t=1,2,\ldots$. Equation (10) and (11) can be combined to allow for a dependence on time and SOI by letting

$$\mu(t)=\beta_0+\beta_1 SOI(t)+\beta_2(t); \quad (12)$$

Type III: Non-stationary GEV (NSGEV) model, with a non-stationary in scale parameter only.

$$X_t \sim GEV(\mu,\sigma(t),\xi), \quad (13)$$

where

$$ln\sigma(t) = \beta_2 + \beta_3 SOI(t), \quad (14)$$

$$ln\sigma(t) = \beta_2 + \beta_3(t), \quad (15)$$

$$ln\sigma(t) = \beta_3 + \beta_4 SOI(t) + \beta_5(t), \quad (16)$$

Type IV: Non-stationary GEV (NSGEV) model, with a non-stationary in location and scale parameters.

$$X_t \sim GEV(\mu(t), \sigma(t), \xi), \quad (17)$$

where

$$\mu(t) + ln\sigma(t) = \beta_0 + \beta_1 SOI(t) + \beta_2 + \beta_3 SOI(t), \quad (18)$$

$$\mu(t) + ln\sigma(t) = \beta_0 + \beta_1(t) + \beta_2 + \beta_3(t), \quad (19)$$

$$\mu(t) + ln\sigma(t) = \beta_0 + \beta_1 SOI(t) + \beta_2(t) + \beta_3 + \beta_4 SOI(t) + \beta_5(t) \quad (20)$$

A total of ten models will be developed for each station, as summarized in **Table 2**. Among these, one model will be stationary, with all parameters assumed constant over time, while the remaining nine will be non-stationary models, where only the shape parameter remains constant over time. Both seasonal (Southwest Monsoon, SWM, and Northeast Monsoon, NEM) and non-seasonal maximum rainfall data were fitted using the Generalized Extreme Value (GEV) distribution and its non-stationary counterpart (NSGEV) for all stations. The R programming language was utilized for fitting GEV models, allowing both stationary and non-stationary GEV modeling.

Table 2: Developed GEV and NSGEV models for each station.

Model ID	Parameters of model	Remarks
GEV0	μ(constant), σ(constant), ξ(constant)	Stationary model
NSGEV1	$\mu(t) = \beta_0 + \beta_1 SOI(t)$	Nonstationary model with SOI for location parameter
NSGEV2	$\mu(t) = \beta_0 + \beta_1(t)$	Nonstationary model with time dependent location parameter
NSGEV3	$\mu(t) = \beta_0 + \beta_1 SOI(t) + \beta_2(t)$	Nonstationary model with SOI and time for location parameter
NSGEV4	$ln\sigma(t) = \beta_2 + \beta_3 SOI(t)$	Nonstationary model with SOI for scale parameter
NSGEV5	$ln\sigma(t) = \beta_2 + \beta_3(t)$	Nonstationary model with time dependent scale parameter
NSGEV6	$ln\sigma(t) = \beta_3 + \beta_4 SOI(t) + \beta_5(t)$	Nonstationary model with SOI and time for scale parameter
NSGEV7	$\mu(t) + ln\sigma(t) = \beta_0 + \beta_1 SOI(t) + \beta_2 + \beta_3 SOI(t)$	Nonstationary model with SOI for location and scale parameter
NSGEV8	$\mu(t) + ln\sigma(t) = \beta_0 + \beta_1(t) + \beta_2 + \beta_3(t)$	Nonstationary model with time dependent location and scale parameter
NSGEV9	$\mu(t) + ln\sigma(t) = \beta_0 + \beta_1 SOI(t) + \beta_2(t) + \beta_3 + \beta_4 SOI(t) + \beta_5(t)$	Nonstationary model with SOI and time for location and scale parameter

Maximum likelihood estimator

The stationary and non-stationary GEV model parameters are estimated using the maximum likelihood method as follows:

$$L(\mu, \sigma, \xi; \boldsymbol{x}) = -m \log \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{m} log \left[1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right] - \sum_{i=1}^{m} \left[1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}; \quad \xi \neq 0 \qquad (21)$$

$$L(\mu, \sigma, \xi; \boldsymbol{x}) = -m \log \sigma - \sum_{i=1}^{m} log \left(\frac{x_i - \mu}{\sigma}\right) - \sum_{i=1}^{m} \exp\left\{-\left(\frac{x_i - \mu}{\sigma}\right)\right\}; \quad \xi = 0 \qquad (22)$$

where m is the number of block maxima $x_1, x_2, \ldots, x_m$.

Goodness-of-fit tests

The significance of the covariate is tested using the goodness-of-fit tests; the Akaike information criterion (AIC) and the Bayesian Information Criterion (BIC) as follows,

$$AIC = 2K - 2\,ln(L) \tag{23}$$

$$BIC = Kln\,(n) - 2\,ln\,(L) \tag{24}$$

Here, L is the maximum value of the likelihood function for a fitted model, K represents the estimated parameters for the model and n is the number of samples. The model with the lowest AIC and BIC explains most of the variations in the data in the simplest manner and thus is preferable to the other models. Meanwhile, the best fitting model is then distinguished by comparing the deviance statistic D between the negative log-likelihood $(nllh)$ of two models M_0 and M_1 with $C_\propto$. The deviance statistic D, for the LRT is used to test model the following hypothesis, $H_0 = M_0$ model is stationary versus $H_1 = M_1$ model is non-stationary with $D = 2 \cdot [nllh(M_0) - nllh(M_1)] > C_\propto$ where $C_\propto$ is the quantile of the χ_k^2 distribution. Here, k is the difference between the number of estimated parameters in M_0 and M_1. If $D > \chi_k^2$, null hypothesis is rejected at the $\alpha = 0.05$ significance level which concludes that M_1 model is suitable for the monthly maximum rainfall.

2.6 Bootstrap Method

The bootstrap method is used to estimate parameters and their uncertainty in both stationary and non-stationary GEV best-fitted models. Bootstrap resampling involves generating 1000 multiple extreme rainfall datasets by resampling the observed extreme values with replacement, then refitting the GEV model to each resampled extreme rainfall dataset, and extracting the parameter estimates (location μ, scale σ, and shape ξ) for stationary GEV model whilst at least one of the parameters (μ, σ, and ξ) varies with covariates (SOI and time) for non-stationary GEV model. This process provides a bootstrapping of parameter estimates by calculating the mean of the parameter estimates across all iterations (R=1000). The bootstrap parameter can be used to calculate the 95 % confidence interval to assess parameter significance, then the original parameter that has been estimated by using MLE will be compared with the bootstrap parameter to determine the uncertainty of the original parameter. Further examination, if the original parameter falls within the bootstrap 95 % confidence interval, it can be concluded that the original parameter is reliable. This method is offering a flexible and robust approach for parameter estimates from the best-fitted model under both stationary and non-stationary GEV model. Nevertheless, only nonseasonal best-fitted model will be conducted in this study.

2.7 Return Level

Return levels for stationary GEV model

Extreme rainfall return levels are used to quantify potential flooding risk and generally are adopted in engineering design. The equation can be expressed as follows:

$$\mathrm{RL}_T = \begin{cases} \mu - \frac{\sigma}{\xi}\left[1 - \left\{-log\left(1 - \frac{1}{T}\right)\right\}^{-\xi}\right] & , \ \xi \neq 0 \\ \mu - \sigma log\left\{-log\left(1 - \frac{1}{T}\right)\right\} & , \qquad \xi = 0 \end{cases} \tag{25}$$

Return levels for non-stationary GEV model

In the non-stationary state, the location and scale parameters vary over time with inclusion of SOI. Time-varying location and scale parameters will clearly have time-varying estimates of extreme rainfall return levels. These parameters were considered simultaneously when calculating return levels. Thus, the return level calculation can be performed as follows:

$$\mathrm{RL}_T = \begin{cases} \mu(t) - \frac{\sigma(t)}{\xi}\left[1 - \left\{-\log\left(1 - \frac{1}{T}\right)\right\}^{-\xi}\right], & \xi \neq 0 \\ \mu(t) - \sigma(t)\log\left\{-\log\left(1 - \frac{1}{T}\right)\right\}, & \xi = 0 \end{cases} \tag{26}$$

T in (25) and (26) is the return period, whereas RL is the return level. The return level will assess by establishing a return period of 10, 20, 50 and 100 years. Thus, RL_{10}, RL_{20}, RL_{50} and RL_{100} are used to display the potential effects of a warming climate on extreme rainfall during seasonal (SWM and NEM) and nonseasonal period across Peninsular Malaysia.

3 RESULTS AND DISCUSSION

Table 3 illustrates the data preprocessing step for all twelve stations of daily rainfall datasets, there are no missing values for all datasets thus imputation method for replacing missing values will not be conducted in this study.

Table 3: Missing values detection for all 12 stations.

Stations	Missing Values	Stations	Missing Values
Kangar	0	Kota Bharu	0
Kuala Terengganu	0	Kuantan	0
Kuala Lumpur	0	Shah Alam	0
Alor Setar	0	Seremban	0
Georgetown	0	Malacca City	0
Ipoh	0	Johor Bahru	0

3.1 Data Visualization

The boxplot of the Southwest Monsoon (SWM), Northeast Monsoon (NEM) and non-seasonal are illustrated in Figure 2. Each boxplot represents rainfall (precipitation) in mm/day for Kangar, Kuala Terengganu and Kuala Lumpur station. Most of the stations exhibit a narrow Interquartile Range (IQR), suggesting that most daily rainfall values are clustered closely together and less dispersed. All stations have prominent outliers, indicating occurrence of extreme rainfall event from 1990 until 2022. For NEM and non-seasonal boxplot, Kuala Terengganu station exhibits greater outliers and variability in daily rainfall compared to Kangar and Kuala Lumpur station. The range for rainfall in mm/day for NEM and non-seasonal is much wider compared to SWM, suggesting that heavy rainfall frequently occurring throughout the wet season while less rainfall during the dry season. This further strengthens the evidence from MET Malaysia that most states on east coast of peninsular Malaysia (Kuala Terengganu) have maximum rainfall throughout November, December and January while June and July are the dry month in most states.

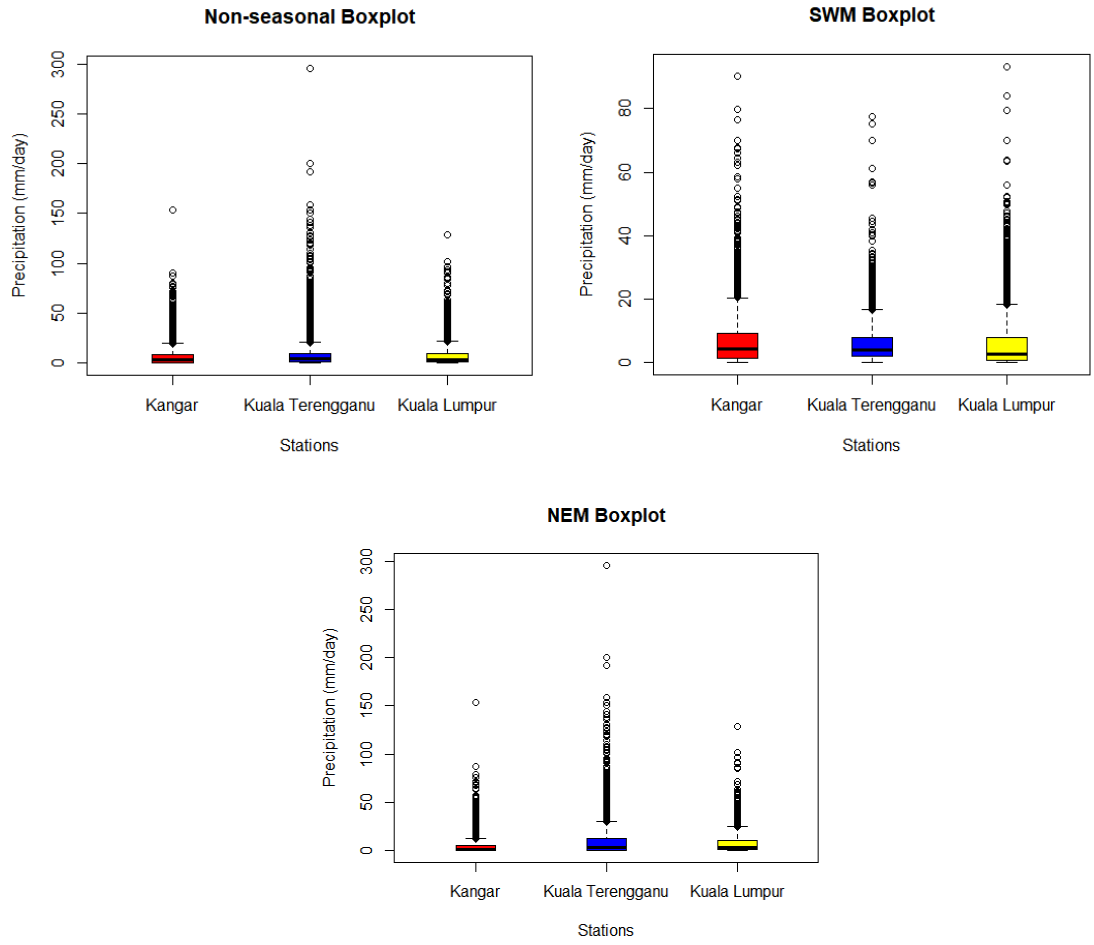
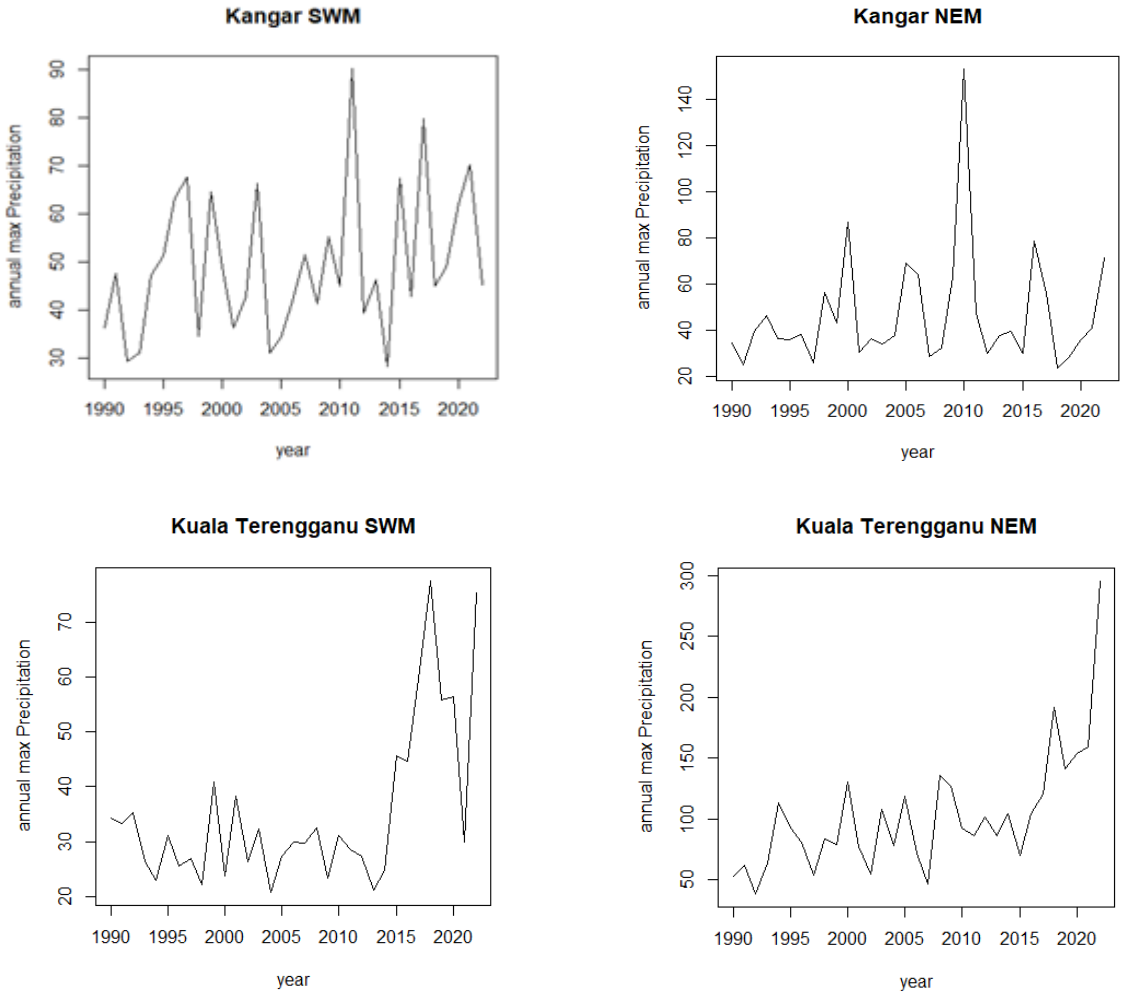


Figure 2: Southeast Monsoon (SWM), Northeast Monsoon (NEM) and non-seasonal boxplot.

The time series plot for non-seasonal and seasonal (SWM and NEM) extreme rainfall at Kangar, Kuala Terengganu, and Kuala Lumpur stations were illustrated in Figure 3 and Figure 4, respectively. All stations from NEM and non-seasonal exhibit the same trend, indicating that both the NEM season (November to March) and non-seasonal which the entire year (January to December) are experiencing similar changes over time, such as an increase or decrease in extreme rainfall event. Most of the stations show an increasing trend from 1990 to 2022, observing that rainfall is increasing from time to time. Kuala Terengganu station displayed a prominent increasing trend.



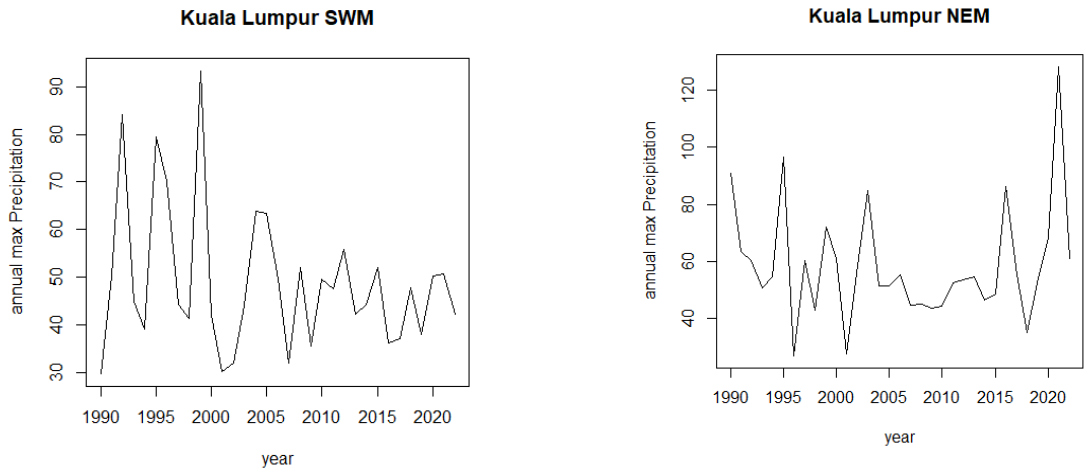


Figure 3: Time series plot during seasonal period.

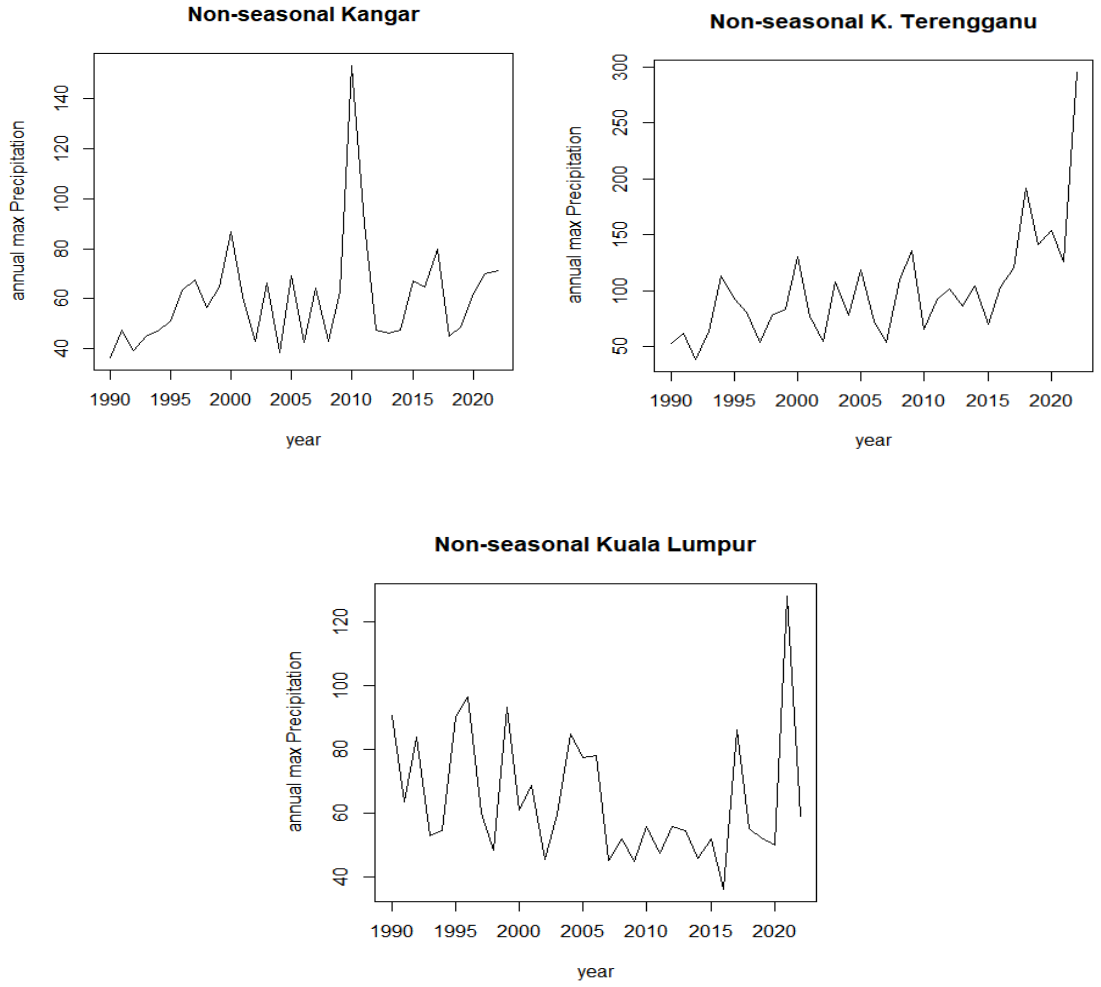


Figure 4: Time series plot during nonseasonal period.

3.2 MK Trend Test and ADF Test

The results of the Mann-Kendall (MK) trend test for non-seasonal and seasonal (NEM and SWM) extreme rainfall are summarized in Table 4, while Table 5 provides an overview of the number of stations exhibiting statistically significant or insignificant trends. For non-seasonal trends, all stations demonstrate a significant positive trend in extreme rainfall. During the NEM period, two stations show a significant positive trend, eight stations exhibit an insignificant positive trend, and two stations display an insignificant negative trend. Meanwhile, during the SWM period, one station shows a significant positive trend, six stations exhibit an insignificant positive trend, and five stations display an insignificant negative trend. Importantly, no stations demonstrate a significant decrease in extreme rainfall trends across both seasonal and non-seasonal analyses. The Sen's slope has been calculated and used for the determination of trend where positive and negative indicate the magnitude of trend.

Table 4: MK trend test for extreme rainfall

Stations		p-value	Sen's slope	Trend
Kangar	SWM	0.1250	0.4312	Insignificant increase
	NEM	0.5664	0.1624	Insignificant increase
	Non-seasonal	1.3e-11	3.2e-05⁺	Significant increase
Georgetown	SWM	0.2209	0.3565	Insignificant increase
	NEM	0.0039	0.7700⁺	Significant increase
	Non-seasonal	2.2e-16	0.0000⁺	Significant increase
Alor Setar	SWM	0.1176	-0.3951	Insignificant decrease
	NEM	0.3771	0.2881	Insignificant increase
	Non-seasonal	2.2e-16	7.0e-05⁺	Significant increase
Ipoh	SWM	0.5664	0.1153	Insignificant increase
	NEM	0.5052	0.2108	Insignificant increase
	Non-seasonal	2.2e-16	0.0001⁺	Significant increase
Kota Bharu	SWM	0.7216	0.0692	Insignificant increase
	NEM	0.0528	1.2776	Insignificant increase
	Non-seasonal	2.1e-06	3.1e-05⁺	Significant increase
K. Terengganu	SWM	0.0491	0.4429⁺	Significant increase
	NEM	6.8e-05	2.7462⁺	Significant increase
	Non-seasonal	0.0057	2.0e-05⁺	Significant increase
Kuantan	SWM	0.5457	-0.1820	Insignificant decrease
	NEM	0.1328	1.3919	Insignificant increase
	Non-seasonal	0.0015	2.2e-05⁺	Significant increase
Shah Alam	SWM	0.9382	0.0216	Insignificant increase
	NEM	0.9382	-0.0216	Insignificant decrease
	Non-seasonal	7.2e-10	5.3e-05⁺	Significant increase
Seremban	SWM	0.5151	0.1724	Insignificant increase
	NEM	0.1328	1.3919	Insignificant increase
	Non-seasonal	4.7e-07	3.5e-05⁺	Significant increase
Malacca City	SWM	0.3036	-0.2944	Insignificant decrease
	NEM	0.9876	0.0090	Insignificant increase
	Non-seasonal	2.2e-16	7.8e-05⁺	Significant increase

Johor Bahru	SWM	0.0748	-0.6808	Insignificant decrease
	NEM	0.1328	1.3919	Insignificant increase
	Non-seasonal	2.2e-16	0.0002^{+}	Significant increase
Kuala Lumpur	SWM	0.6310	-0.1274	Insignificant decrease
	NEM	0.7332	-0.1529	Insignificant decrease
	Non-seasonal	2.2e-16	6.0e-05^{+}	Significant increase

$^{+}$: significant increase and $^{-}$: significant decrease

Table 5. The number of stations showing statistically significant or insignificant trends based on the MK test.

	Increase		**Decrease**	
	Significant increase	**Insignificant increase**	**Significant decrease**	**Insignificant decrease**
SWM	1	6	0	5
NEM	2	8	0	2
Non-seasonal	12	0	0	0

Next, the Augmented Dickey-Fuller (ADF) test was conducted on the same dataset to assess the stationarity of the time series. As shown in Table 6, the non-seasonal extreme rainfall data for all stations are stationary at the 5% significance level, indicating the absence of any prominent trend in the time series. In contrast, the seasonal extreme rainfall series is non-stationary for all stations except Alor Setar station during NEM, suggesting the presence of underlying trends or variability over time.

Table 6: Stationarity for each station based on ADF test.

Stations		**p-value**	**Conclusion**
Kangar	SWM	0.2300	Nonstationary
	NEM	0.1141	Nonstationary
	Non-seasonal	0.0100*	Stationary
Alor Setar	SWM	0.4283	Nonstationary
	NEM	0.0172*	Stationary
	Non-seasonal	0.0100*	Stationary
Georgetown	SWM	0.5461	Nonstationary
	NEM	0.1288	Nonstationary
	Non-seasonal	0.0100*	Stationary
Ipoh	SWM	0.7585	Nonstationary
	NEM	0.4922	Nonstationary
	Non-seasonal	0.0100*	Stationary
Kota Bharu	SWM	0.7638	Nonstationary
	NEM	0.1913	Nonstationary
	Non-seasonal	0.0100*	Stationary
K. Terengganu	SWM	0.6536	Nonstationary
	NEM	0.9900	Nonstationary

	Non-seasonal	0.0100*	Stationary
Kuantan	SWM	0.5481	Nonstationary
	NEM	0.1939	Nonstationary
	Non-seasonal	0.0100*	Stationary
Shah Alam	SWM	0.6160	Nonstationary
	NEM	0.6263	Nonstationary
	Non-seasonal	0.0100*	Stationary
Seremban	SWM	0.3176	Nonstationary
	NEM	0.5077	Nonstationary
	Non-seasonal	0.0100*	Stationary
Malacca City	SWM	0.1025	Nonstationary
	NEM	0.3391	Nonstationary
	Non-seasonal	0.0100*	Stationary
Johor Bahru	SWM	0.0847	Nonstationary
	NEM	0.1939	Nonstationary
	Non-seasonal	0.0100*	Stationary
Kuala Lumpur	SWM	0.1502	Nonstationary
	NEM	0.6045	Nonstationary
	Non-seasonal	0.0100*	Stationary

*: stationary time series

3.3 GEV and NSGEV Model Selection

The GEV and NSGEV models were developed for three selected stations: Kangar, Kuala Terengganu, and Kuala Lumpur. For Kangar station, Table 7 indicates no significant improvement of NSGEV over GEV during the Southwest Monsoon (SWM). In other words, at the 5% significance level, the inclusion of time and SOI variables did not improve the GEV0 model. However, during the Northeast Monsoon (NEM), NSGEV1 showed significant improvement over GEV0. Further comparison of NSGEV1 with higher-order models (NSGEV2 and beyond) revealed NSGEV1 as the best-fitting model. Meanwhile during nonseasonal, NSGEV2 exhibited the best-fitted model, indicating addition of time variable improved the model. These findings align with the AIC and BIC values, where GEV0, NSGEV1 and NSGEV2 yielded the lowest AIC and BIC values during both seasonal and nonseasonal, respectively. Notice that there was a convergence issue during NEM of Kangar station where NSGEV7 and NSGEV9 showed difficulty to fit when considering SOI and time variable. Kuala Lumpur station also exhibited the same issue for NSGEV9 during nonseasonal period. Typically, convergence equal to zero means that the optimization algorithm successfully converged, while any nonzero value suggests that the optimization process failed to converge. For Kuala Terengganu station, based on the D values, NSGEV5 was identified as the most suitable model for extreme rainfall during SWM, while NSGEV8 was the best model during NEM and nonseasonal. This is consistent with the AIC and BIC values, which were minimized for these models. Nevertheless, only during nonseasonal the AIC value is not the lowest but supported by lowest BIC and significant D. Similarly, for Kuala Lumpur station, NSGEV5 was determined to be the appropriate models for extreme rainfall during SWM while NSGEV6 during NEM and nonseasonal, supported by the lowest AIC and BIC values and significant improvement of D values. Table 8 summarized the best fitted GEV models for each station, showing

3 models were associated with SOI variable to indicate that extreme rainfall in Kangar and Kuala Lumpur has influenced on ENSO during wet season.

Table 7: Seasonal (SWM and NEM) and Nonseasonal AIC, BIC and D value.

Station	Model	SWM			NEM			Non-seasonal		
		AIC	BIC	D	AIC	BIC	D	AIC	BIC	D
Kangar	GEV0	272.65*	277.14*	-	279.15	283.64	-	281.43	285.92	-
	NSGEV1	274.19	280.17	0.48	276.53*	282.52*	4.62*	282.99	288.98	0.45
	NSGEV2	272.34	278.32	2.32	281.13	287.11	0.02	279.31*	285.29*	4.13*
	NSGEV3	274.29	281.77	2.36	276.39	283.88	2.14	281.28	288.76	0.02
	NSGEV4	274.64	280.63	0.00	281.05	287.03	0.10	280.99	286.98	2.44
	NSGEV5	274.57	280.55	0.08	280.80	286.79	0.35	282.99	288.97	0.45
	NSGEV6	276.54	284.02	0.11	280.97	288.45	2.18	282.75	290.24	2.68
	NSGEV7	276.19	283.67	0.46	-	-	-	282.07	289.56	3.36
	NSGEV8	273.75	281.23	2.90	282.34	289.82	0.81	280.59	288.07	0.72
	NSGEV9	277.60	288.07	3.05	-	-	-	282.01	292.48	3.30
K. Terengganu	GEV0	251.26	255.75	-	341.87	346.35	-	338.42	342.91	-
	NSGEV1	253.12	259.11	0.14	342.85	348.83	1.01	335.50	341.48	4.92
	NSGEV2	253.17	259.15	0.09	330.81	336.79	13.06	327.45	333.44	12.97
	NSGEV3	255.05	262.53	0.20	332.81	340.29	0.00	327.97	335.46	1.48
	NSGEV4	253.23	259.22	0.02	342.02	348.01	0.64	340.31	346.30	0.10
	NSGEV5	248.64*	254.63*	4.61*	341.85	347.84	2.01	339.54	345.53	0.88
	NSGEV6	250.56	258.05	0.08	343.84	351.33	2.02	341.24	348.73	1.18
	NSGEV7	255.08	262.56	0.18	340.46	347.94	5.41	333.95	341.43	8.47
	NSGEV8	249.41	256.89	1.24	327.38*	334.86*	5.43*	325.26	332.74*	4.19*
	NSGEV9	253.40	263.87	1.25	327.65	338.13	3.73	325.10*	335.58	4.16
	GEV0	266.83	271.32	-	288.98	293.47	-	285.81	290.30	-
	NSGEV1	268.30	274.29	0.53	290.95	296.94	0.02	287.76	293.75	0.05
	NSGEV2	267.76	273.75	1.07	290.77	296.76	0.20	283.01	288.99	4.80

	NSGEV 3	269.62	277.10	1.21	292.67	300.16	0.30	283.25	290.73	1.76
	NSGEV 4	268.56	274.54	0.27	288.00	293.99	2.98	283.56	289.55	4.25
Kuala Lumpur	NSGEV 5	260.70 *	266.69 *	8.13 *	290.42	296.41	0.55	287.65	293.64	0.15
	NSGEV 6	262.62	270.11	0.08	283.38 *	290.86 *	9.59*	214.88 *	222.36 *	70.13 *
	NSGEV 7	270.83	278.31	0.57	289.91	297.40	3.06	285.36	292.84	4.45
	NSGEV 8	262.65	270.13	0.05	292.21	299.69	0.76	285.00	292.48	4.81
	NSGEV 9	266.40	276.87	0.31	287.24	297.72	0.14	-	-	-

***:** lowest AIC, BIC and significant improvement of D

Table 8: Best Fitted Models.

Stations	SWM	NEM	Non-seasonal
Kangar	GEV0	NSGEV1*	NSGEV2
Kuala Terengganu	NSGEV5	NSGEV8	NSGEV8
Kuala Lumpur	NSGEV5	NSGEV6*	NSGEV6*

*: significant inclusion of SOI

3.4 Parameter Estimation

GEV distribution is parameterized with location parameters μ, scale parameter σ, and shape parameter ξ. Nonstationary GEV reveals more than one parameter as function of linear trend with covariate SOI and time. μ and σ are expressed as linear trend functions while ξ is constant over time. As in a standard simple linear regression, $\beta_1, \beta_2, \beta_3, \beta_4$ and β_5 represent the slope while β_0 illustrates as the intercept. All the best-fitted model parameters during both seasonal (SWM and NEM) and non-seasonal are listed in Table 9 with their 95% confidence interval. In this study, the nonstationary models consist of unfixed location and scale parameters with the shape parameter will be fixed. Based on the 95% confidence interval, most of the shape parameters include 0 within the 95% confidence interval for both stationary and nonstationary model, suggesting that the extreme rainfall data does not provide enough evidence to conclude whether the shape parameter is significantly different from 0 at the 5% significance level except Kangar station during NEM. In other words, the GEV distribution in this study might be the Gumbel distribution where the shape parameter is equal to zero ($\xi = 0$) rather than a heavy-tailed (Fréchet, $\xi > 0$) or bounded-tailed (Weibull, $\xi < 0$) distribution. These findings aligned with Ragulina and Reitan [33] who examined the extreme rainfall for GEV shape parameter and discusses the shape parameter is expected to belong in a narrow range, approximately from 0.127 to 0.150 with 95% confidence, suggesting shape parameter is often close to 0.

Table 9: Parameter values for each station

Stations	Seasonal	Best Model	Parameters	Standard error	95% C.I.
Kangar	SWM	GEV0	μ = 42.4160	2.3369	(37.84, 47.00)
			σ = 11.4726	1.7622	(8.02, 14.93)
			ξ = 0.0300	0.7622	(-0.29, 0.35)
	NEM	NSGEV1	β_0= 34.2993	1.8904	(30.59, 38.00)
			β_1= 3.1927	1.5303	(0.19, 6.19)
			σ = 9.4604	1.7189	(6.09, 12.83)
			ξ = 0.4239	0.1654	(0.1, 0.75)
	Non-seasonal	NSGEV2	β_0= 44.2865	3.4390	(37.55, 51.03)
			β_1= 0.3451	0.1629	(0.03, 0.66)
			σ = 10.5703	1.9754	(6.70, 14.44)
			ξ = 0.3136	0.2212	(-0.12, 0.75)
Kuala Terengganu	SWM	NSGEV5	β_2= 27.4764	1.2988	(24.93, 30.02)
			β_3= 3.1122	1.3200	(0.53, 5.70)
			μ = 0.2411	0.1232	(-0.0003, 0.48)
			ξ = 0.2167	0.238	(-0.25, 0.68)
	NEM	NSGEV8	β_0= 48.2008	6.6543	(35.16, 61.24)
			β_1= 2.2626	0.4889	(1.30, 3.22)
			β_2= 11.7695	4.8841	(2.20, 21.34)
			β_3= 0.9059	0.3735	(0.17, 1.64)
			ξ = -0.0280	0.1272	(-0.28, 0.22)
	Non-seasonal	NSGEV8	β_0= 48.4428	6.5452	(35.61, 61.28)
			β_1= 2.0974	0.4584	(1.20, 3.00)
			β_2= 12.1289	4.8598	(2.61, 21.67)
			β_3= 0.7561	0.3635	(0.04, 1.47)
			ξ = 0.0422	0.1446	(-0.24, 0.33)
Kuala Lumpur	SWM	NSGEV5	β_2= 42.3831	1.5962	(39.25, 45.51)
			β_3= 17.5733	4.3960	(8.96, 26.19)
			μ = -0.4247	0.1579	(-0.73, -0.12)
			ξ = -0.0120	0.2074	(-0.42, 0.39)
	NEM	NSGEV6	β_3= 49.4171	1.4526	(46.57, 52.26)
			β_4= 20.5841	3.3512	(14.02, 27.15)
			β_5= 6.8546	1.8034	(3.32, 10.39)
			μ = -0.4754	0.1222	(-0.71, -0.24)
			ξ = 0.0158	0.1201	(-0.22, 0.25)
	Non-seasonal	NSGEV6	β_3= 52.0500	2.004e-06	(52.05, 52.05)
			β_4= 16.1093	2.000e-06	(16.11, 16.11)
			β_5= 6.0470	2.000e-06	(6.05, 6.05)
			μ = -0.2940	2.000e-06	(-0.29, -0.29)
			ξ = 0.2110	1.106e-01	(-0.01, 0.43)

3.5 Bootstrapping

Bootstrap method was applied for the nonseasonal best-fitted model only. The uncertainty of these parameters will involve resampling the data with replacement and re-fitting the best-fitted model to each dataset, then calculating the mean of the parameter estimates across all iterations (R=1000) with 95% confidence intervals. Table 10 provides the precision and uncertainty of each parameter estimates. All stations exhibit the parameters in the GEV distribution estimated from the original extreme rainfall dataset are approximately the same as the parameters estimated by using a bootstrap method, indicating that the parameter estimates are stable and robust to resampling variability. This is a sign that the nonseasonal extreme rainfall dataset suitably represent the behaviour for extreme value analysis. Other than that, this similarity between the original and bootstrap parameter estimates, suggesting there is no bias in the parameter estimation method by using MLE. Bootstrapping also provides the 95% confidence interval where all the original parameter estimates fall within bootstrap range, supporting that the original parameters are consistent and reliable. Since all the GEV parameters estimate also align closely with the bootstrap parameters, revealing that all best-fitted model is well-suited to the extreme rainfall data and to validate the uncertainty of the original parameters are consistent and robust. Hence, there is no doubt that the best-fitted model during seasonal (SWM and NEM) period also exhibited the same reliable of parameter estimation. This bootstrapping method in this study is corresponding with Diop and Deme [34] who examined a GEV model and applying parametric bootstrap methods to estimate the confidence intervals of parameters.

Table 10: Bootstrapping of the parameter estimates.

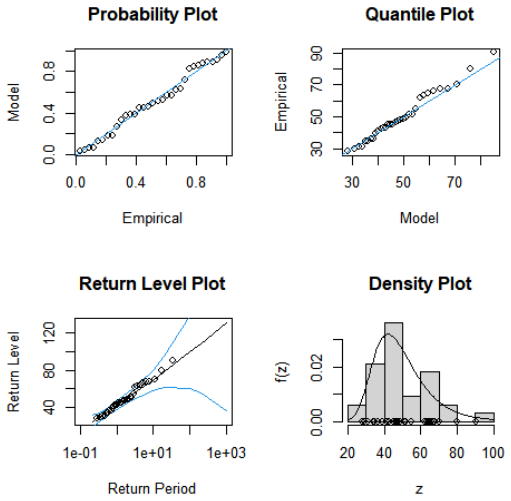
Stations	Best-fitted Model (nonseasonal)	Original Parameters	Bootstrap Parameters	Bootstrap 95% C.I.
Kangar	NSGEV2	β_0= 44.2865	β_0= 44.4134	(37.95, 52.92)
		β_1= 0.3451	β_1= 0.3683	(0.02, 0.74)
		σ = 10.5703	σ = 10.4002	(5.37, 14.43)
		ξ = 0.3136	ξ = 0.3501	(-0.52, 1.14)
Kuala Terengganu	NSGEV8	β_0= 48.4428	β_0= 49.5093	(31.87, 67.13)
		β_1= 2.0974	β_1= 2.0814	(0.96, 3.27)
		β_2= 12.1289	β_2= 11.2312	(-3.93, 32.58)
		β_3= 0.7561	β_3= 0.8390	(-0.36, 2.36)
		ξ = 0.0422	ξ = 0.0625	(-0.48, 0.63)
Kuala Lumpur	NSGEV6	β_3= 52.0500	β_3= 53.0422	(48.48, 61.44)
		β_4= 16.1093	β_4= 17.1818	(3.05, 36.56)
		β_5= 6.0470	β_5= 4.8070	(-4.30, 10.19)
		μ = -0.2940	μ = -0.2744	(-0.95, 0.58)
		ξ = 0.2110	ξ = 0.1947	(-0.79, 0.72)

3.6 Model Diagnostics

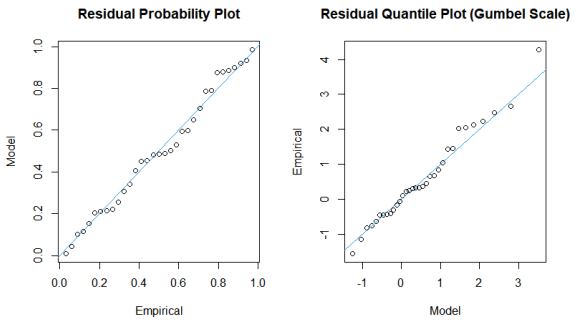
Figure 5 exhibits the model diagnostics plot from the best-fitted model during seasonal period (SWM and NEM) whilst Figure 6 shows for nonseasonal period. For stationary model (GEV0), only Kangar station during SWM contains four plots which are probability plot, quantile plot, return level plot and density plot. The probability plot is to compare the theoretical cumulative distribution function (CDF)

of the best-fitted stationary model against the empirical CDF derived from the extreme rainfall. Next, the quantile plot is to compare the theoretical quantiles of the fitted GEV0 to the empirical quantiles of the extreme rainfall data. The return level plot against the return periods is to evaluate the tail behavior of the fitted GEV0 model. Moreover, the density plot is to compare the histogram of the observed data with the theoretical probability density function (PDF) of the fitted GEV0 model. During SWM, only Kangar station exhibit GEV0 as the best-fitted model, indicating that the stationary model (GEV0) is well-fitted with the extreme rainfall data. The residual probability plot and residual quantile plot on the Gumbel scale are to evaluate the goodness-of-fit of a non-stationary GEV model (NSGEV). During both seasonal and nonseasonal period, all stations exhibit almost the same pattern for NSGEV diagnostic plot where all points align well with the diagonal line, suggesting the NSGEV models are well-fitted with the extreme rainfall data with inclusion of covariates (SOI and time). These diagnostic plots reveal that all the best-fitted model whether stationary or non-stationary are well-fitted for all stations.

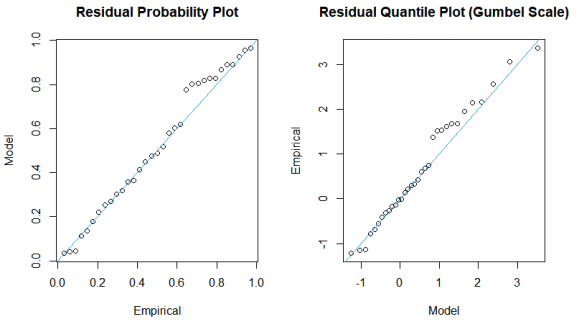
Kangar SWM



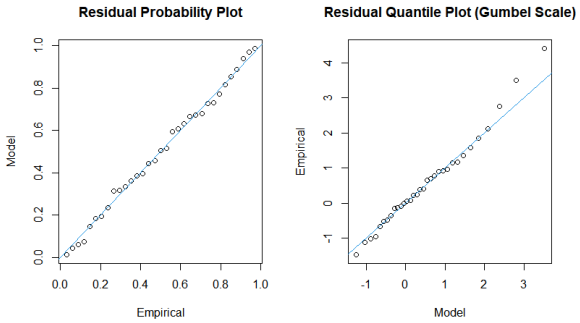
Kangar NEM



Kuala Terengganu SWM



Kuala Terengganu NEM



Kuala Lumpur SWM

Kuala Lumpur NEM

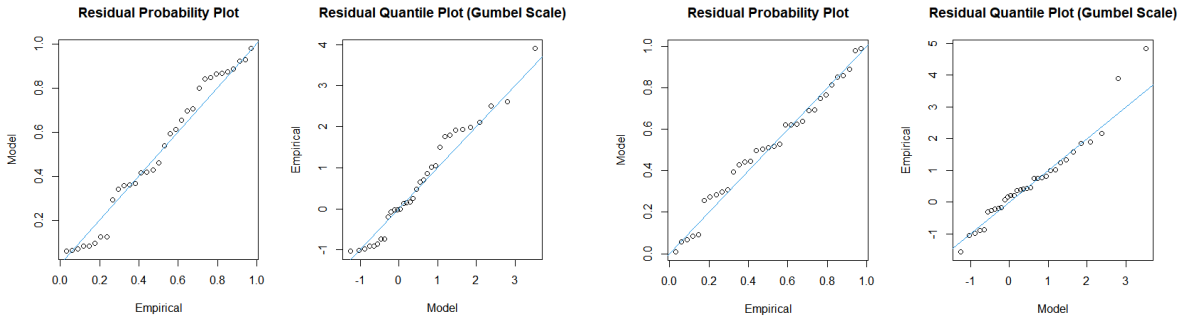
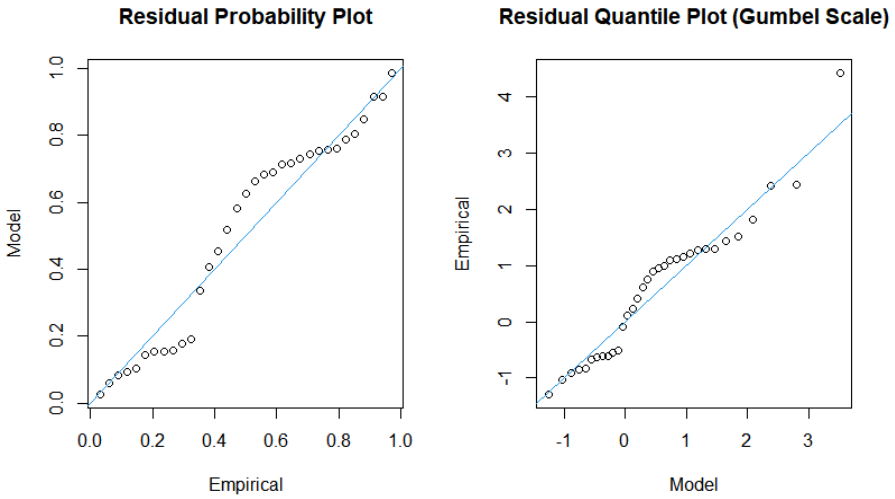
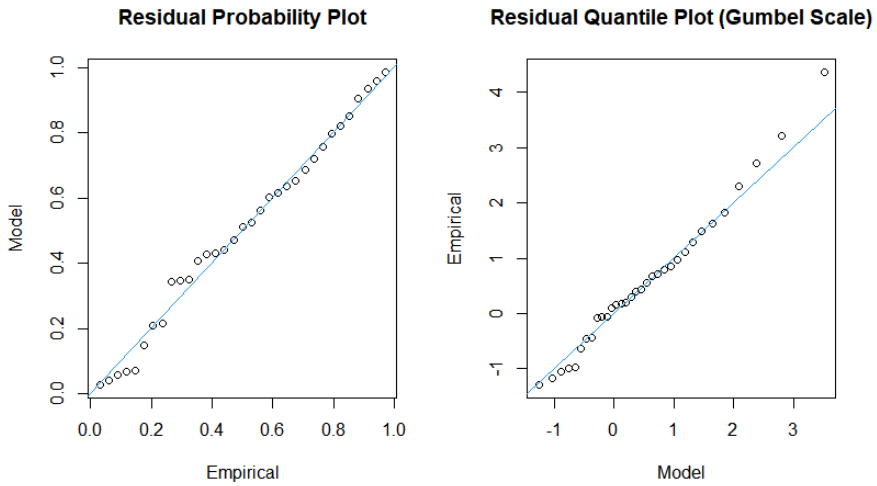


Figure 5: Diagnostic plots for seasonal model.

Kangar



Kuala Terengganu



Kuala Lumpur

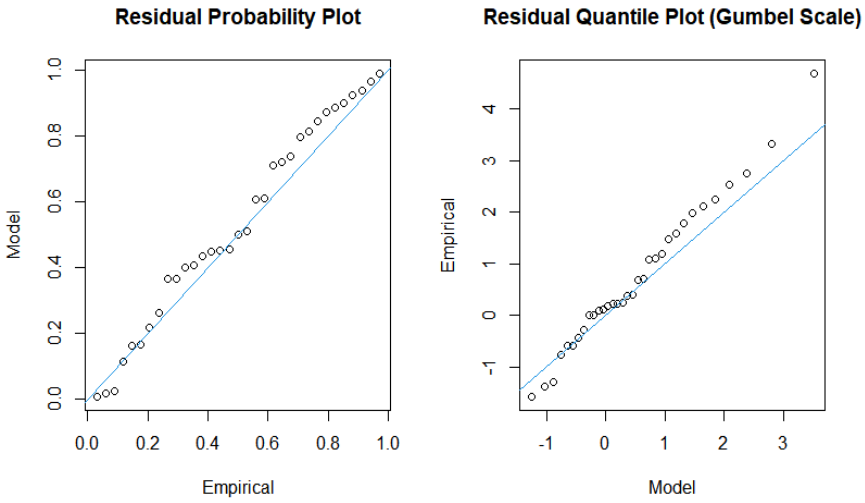


Figure 6: Diagnostic plots for nonseasonal model

3.7 Return Level Estimate

Table 11 summarizes the return level estimate of extreme rainfall for both seasonal and nonseasonal best-fitted model for all stations. The return level in this study provides the value of an extreme rainfall event expected to occur, once in each period (10 years, 20 years, 50 years, and 100 years). From this table, it can be observed that as the return period increase, the return level estimates increase as well. For stationary model (GEV0), the return level RL_T for a given return period T is calculated using the inverse of the GEV cumulative distribution function in (25). All the parameters (location μ, scale σ and shape ξ) are constant over time because all the parameters do not change, which is why the return levels for Kangar station during SWM are single value without any variables (SOI and time). Meanwhile for non-stationary model (NSGEV), the return level RL_T is calculated by using the time-dependent parameters in (26). The location parameter (μ) and scale parameter (σ) are varying over time with covariates (SOI and time), hence the return level changes with time or SOI. Understanding how the return levels develop over time is crucial especially for climate risk management and planning [35].

Table 11: Return level estimate of extreme rainfall for T=10, 20, 50, 100.

Stations	Seasonal	Model	T=10	T=20	T=50	T=100
Kangar	SWM	GEV0	69.92	78.06	89.92	99.02
	NEM	NSGEV1	69.91+ 3.19SOI(t)	90.59+ 3.19SOI(t)	128.65+ 3.19SOI(t)	168.84+ 3.19SOI(t)
	Non-seasonal	NSGEV2	121.26+0.35t	138.55+0.35t	167.58+0.35t	195.63+0.35t
K. Terengganu	SWM	NSGEV5	0.24+2.9 $e^{27.48+3.11t}$	0.24+4.17 $e^{27.48+3.11t}$	0.24+6.13 $e^{27.48+3.11t}$	0.24+7.89 $e^{27.48+3.11t}$
	NEM	NSGEV8	(48.2+2.26t)+ 2.18$e^{27.48+3.11t}$	(48.2+2.26t)+ 2.85$e^{27.48+3.11t}$	(48.2+2.26t)+ 3.7$e^{27.48+3.11t}$	(48.2+2.26t)+ 4.32 $e^{27.48+3.11t}$
	Non-seasonal	NSGEV8	(48.44+2.1t)+ 2.36$e^{12.14+0.76t}$	(48.44+2.1t)+ 3.16$e^{12.14+0.76t}$	(48.44+2.1t)+ 4.24$e^{12.14+0.76t}$	(48.44+2.1t)+ 5.08$e^{12.14+0.76t}$

Kuala Lumpur	SWM	NSGEV5	-0.42+2.22 $e^{42.38+17.57t}$	-0.42+2.92 $e^{42.38+17.57t}$	-0.42+3.81 $e^{42.38+17.57t}$	-0.42+4.48 $e^{42.38+17.57t}$
	NEM	NSGEV6	0.48+2.29 $e^{49.42+20.58SOI(t)+6.85t}$	0.48+3.04 $e^{49.42+20.58SOI(t)+6.85t}$	0.48+4.02 $e^{49.42+20.58SOI(t)+6.85t}$	0.48+4.77 $e^{49.42+20.58SOI(t)+6.85t}$
	Non-seasonal	NSGEV6	-0.29+2.88 $e^{52.05+16.11SOI(t)+6.05t}$	-0.29+4.13 $e^{52.05+16.11SOI(t)+6.05t}$	-0.29+6.06 $e^{52.05+16.11SOI(t)+6.05t}$	-0.29+7.77 $e^{52.05+16.11SOI(t)+6.05t}$

4 CONCLUSION

In conclusion, this study investigated the trends in both seasonal (SWM and NEM) and nonseasonal extreme rainfall obtained from 12 stations across Peninsular Malaysia from 1990 to 2022 by using the MK test. A total of 100% of stations in the Peninsular Malaysia displayed significant positive extreme rainfall trends during nonseasonal period while only Kuala Terengganu station and Georgetown station exhibited a significant increasing trend during seasonal period. These two stations experienced higher frequency and intensity of rainfall, making them the most affected areas during the NEM compared to other stations. Only Kuala Terengganu station showed significant positive during both seasonal and nonseasonal due to geographical factor as Terengganu's coastal region exposes it to heavy rainfall influenced by the strong NEM winds. The ADF test further confirms the presence of trends in the extreme rainfall series with all stations are stationary during nonseasonal whilst only Kedah station is stationary during NEM. Only 3 stations were selected which are Kangar, Kuala Terengganu and Kuala Lumpur station were modelled by using the GEV distribution to detect whether extreme rainfall was associated with ENSO. When considering separately location and scale as time varying parameters with inclusion of SOI, Kangar and Kuala Lumpur showed associations with ENSO where two models during NEM and one model during nonseasonal, revealing a significant relationship between extreme rainfall and ENSO. To assess the uncertainty of the parameter estimates by using the MLE, bootstrap method was applied by generating 1000 bootstrap samples to provide robustness and reliability of GEV parameter estimates nevertheless only nonseasonal models were applied. The model diagnostics were conducted to examine the best-fitted model for all stations by inspecting their plots. Then, the return level estimation of the best-fitted GEV model has been assessed in the period of 10 years, 20 years, 50 years and 100 years.

This study provides valuable insight of the influence ENSO on extreme rainfall, offering evidence of the interaction between global climatic phenomena and the local weather patterns. By identifying areas most affected with the extreme rainfall, such as Kuala Terengganu and Georgetown, this study contributes to better disaster preparedness and mitigation plans. Hence, the developed non-stationary GEV modelling in this study demonstrates in understanding how climate variability like ENSO affects extreme rainfall, providing a more dynamic perspective compared to stationary models. Limitations of this study is that only SOI considered as a covariate. There are more potential climate drivers such as local and global temperature, sea-surface temperature, Indian Ocean Dipole (IOD) or local topography to were not included in this study to examine the local extreme rainfall. The analysis was limited at twelve stations across capital cities in Peninsular Malaysia which may not fully represent the spatial heterogeneity of extreme rainfall. Increasing the number of stations could capture more valuable insight of the extreme rainfall across Peninsular Malaysia. For recommendation, this study should be extended by incorporating future climate projections to assess

how climate change influences extreme rainfall trends in Peninsular Malaysia. The climate models and downscaled data can be conducted to predict future extreme events. This study also recommends investigating additional covariates such as sea surface temperature, IOD or other global climate factors to provide better understanding the drivers of local extreme rainfall. Other than that, this study suggests increasing the number of stations should be accounted to carry out the future research instead of only considering the capital cities of Peninsular Malaysia for research purpose. On the other hand, this study emphasizes the significant of incorporating non-stationary GEV analysis to better understand the nature of extreme rainfall influenced by climate variability, particularly ENSO. By addressing the limitations and recommendations, researchers and policymakers can better prepare for the challenges posed by extreme weather across Peninsular Malaysia.

ACKNOWLEDGEMENT

This work was supported by Ministry of Higher Education under Fundamental Research Grant Scheme (FRGS/1/2024/STG06/UPM/02/6).

REFERENCES

[1] D. Rypkema and S. Tuljapurkar, "Modeling extreme climatic events using the generalized extreme value (GEV) distribution," in *Handbook of Statistics*, vol. 44, pp. 39–71, Elsevier, 2021.

[2] S. Tsuzuki, "A hydrodynamic approach to reproduce multiple spinning vortices in horizontally rotating three-dimensional liquid helium-4," *arXiv preprint*, arXiv:2405.03980, 2024.

[3] K. Makubyane and D. Maposa, "Forecasting short- and long-term wind speed in Limpopo Province using machine learning and extreme value theory," *Forecasting*, vol. 6, no. 4, pp. 885–907, 2024.

[4] H. Li and H. Chen, "Hierarchical mortality forecasting with EVT tails: An application to solvency capital requirement," *Int. J. Forecast.*, 2022.

[5] B. Safari, "Modelling extreme rainfall with block maxima and peak-over threshold methods in Rwanda," unpublished.

[6] S. Sippel, D. Mitchell, M. T. Black, A. J. Dittus, L. Harrington, N. Schaller, and F. E. Otto, "Combining large model ensembles with extreme value statistics to improve attribution statements of rare events," *Weather Clim. Extremes*, vol. 9, pp. 25–35, 2015.

[7] F. Zwiers, "Long period return level estimates of extreme precipitation," in *Int. Workshop on Flood-Resistant Buildings: Book of Abstracts*, Ottawa, Canada, Feb. 26–27, 2020, Nat. Res. Council Canada.

[8] T. S. Kiessé, B. Lemercier, M. S. Corson, Y. Ellili-Bargaoui, and C. Walter, "Analysis of extreme values of soil ecosystem services predicted from associated soil properties and weather conditions," *Eur. J. Soil Sci.*, vol. 74, no. 1, p. e13342, 2023.

[9] D. Malka, K. Janot, M. Pasi, J. P. Desilles, G. Marnat, I. Sibon, and G. Boulouis, "Effects of weather conditions on endovascular treatment case volume for patients with ischemic stroke," *J. Neuroradiol.*, vol. 50, no. 6, pp. 593–599, 2023.

[10] J. E. Walsh, T. J. Ballinger, E. S. Euskirchen, E. Hanna, J. Mård, J. E. Overland, and T. Vihma, "Extreme weather and climate events in northern areas: A review," *Earth-Sci. Rev.*, vol. 209, p. 103324, 2020.

[11] V. R. Golroudbary, Y. Zeng, C. M. Mannaerts, and Z. B. Su, "Detecting the effect of urban land use on extreme precipitation in the Netherlands," *Weather Clim. Extremes*, vol. 17, pp. 36–46, 2017.

[12] Y. L. Tew, M. L. Tan, L. Juneng, K. P. Chun, M. H. B. Hassan, S. B. Osman, and M. H. Kabir, "Rapid extreme tropical precipitation and flood inundation mapping framework (RETRACE): Initial testing for the 2021–2022 Malaysia flood," *ISPRS Int. J. Geo-Inf.*, vol. 11, no. 7, p. 378, 2022.

[13] M. Z. M. Amin, A. J. Shaaban, A. Ercan, K. Ishida, M. L. Kavvas, Z. Q. Chen, and S. Jang, "Future climate change impact assessment of watershed scale hydrologic processes in Peninsular Malaysia by a regional climate model coupled with a physically-based hydrology model," *Sci. Total Environ.*, vol. 575, pp. 12–22, 2017.

[14] O. O. Mayowa, S. H. Pour, S. Shahid, M. Mohsenipour, S. B. Harun, A. Heryansyah, and T. Ismail, "Trends in rainfall and rainfall-related extremes in the east coast of Peninsular Malaysia," *J. Earth Syst. Sci.*, vol. 124, pp. 1609–1622, 2015.

[15] N. A. Majid, M. Taha, and S. Selamat, "Historical landslide events in Malaysia 1993–2019," *Indian J. Sci. Technol.*, vol. 13, no. 33, pp. 3387–3399, 2020.

[16] N. M. Nasidi, A. Wayayok, A. Abdullah, and M. S. Mohd Kassim, "Dynamics of potential precipitation under climate change scenarios at Cameron Highlands, Malaysia," *SN Appl. Sci.*, vol. 3, pp. 1–17, 2021.

[17] R. A. Adewoyin, P. Dueben, P. Watson, Y. He, and R. Dutta, "TRU-NET: A deep learning approach to high resolution prediction of rainfall," *Mach. Learn.*, vol. 110, pp. 2035–2062, 2021.

[18] L. Cea and P. Costabile, "Flood risk in urban areas: Modelling, management and adaptation to climate change—A review," *Hydrology*, vol. 9, no. 3, p. 50, 2022.

[19] H. W. Fischer, "Decentralization and the governance of climate adaptation: Situating community-based planning within broader trajectories of political transformation," *World Dev.*, vol. 140, p. 105335, 2021.

[20] C. Wang and P. Fiedler, "ENSO variability and the eastern tropical Pacific: A review," *Prog. Oceanogr.*, vol. 69, pp. 239–266, 2006.

[21] T. Q. Lap and N. L. Bang, "Teleconnections between sea surface temperature anomalies and the Southern Oscillation Index in the Pacific and Indian Oceans on drought events in the Mekong Delta," *Water Resour.*, vol. 51, no. 5, pp. 764–779, 2024.

[22] D. Miller and Z. Wang, "Assessing seasonal predictability sources and windows of high

predictability in the Climate Forecast System, Version 2," *J. Climate*, 2019.

[23] E. Hackert, S. Akella, L. Ren, K. Nakada, J. Carton, and A. Molod, "Impact of the TAO/TRITON array on reanalyses and predictions of the 2015 El Niño," *J. Geophys. Res.: Oceans*, 2023.

[24] A. Capotondi, A. Wittenberg, M. Newman, E. Lorenzo, J. Yu, P. Braconnot, J. Cole, B. Dewitte, B. Giese, E. Guilyardi, F. Jin, K. Karnauskas, B. Kirtman, T. Lee, N. Schneider, Y. Xue, and S. Yeh, "Understanding ENSO diversity," *Bull. Am. Meteorol. Soc.*, vol. 96, pp. 921–938, 2015.

[25] W. Fang, Y. Sha, and V. S. Sheng, "Survey on the application of artificial intelligence in ENSO forecasting," *Mathematics*, vol. 10, no. 20, p. 3793, 2022.

[26] F. N. Ogana, J. S. A. Osho, and J. J. Gorgoso-Varela, "Application of extreme value distribution for assigning optimum fractions to distributions with boundary parameters: An eucalyptus plantations case study," unpublished, 2018.

[27] M. Šraj, A. Viglione, J. Parajka, and G. Blöschl, "The influence of non-stationarity in extreme hydrological events on flood frequency estimation," *J. Hydrol. Hydromech.*, vol. 64, pp. 426–437, 2016.

[28] F. C. Onwuegbuche, S. B. Affognon, E. P. Enoc, and M. O. Akinade, "Application of extreme value theory in predicting climate change induced extreme rainfall in Kenya," *Int. J. Stat. Probab.*, vol. 8, no. 4, pp. 85–94, 2019.

[29] H. Hammami, J. Carreau, L. Neppel, S. Elasmi, and H. Feki, "Smooth spatial modeling of extreme Mediterranean precipitation," *Water*, vol. 14, no. 22, p. 3782, 2022.

[30] J. L. Ng, Y. F. Huang, S. K. Tan, J. C. Lee, N. I. F. Md Noh, and S. Y. Thian, "Comparative evaluation of various parameter estimation methods for extreme rainfall in Kelantan River Basin," *Theor. Appl. Climatol.*, vol. 155, no. 3, pp. 1759–1775, 2024.

[31] M. Lakatos, B. C. Kovácsné Izsák, O. Szentes, L. Hoffmann, A. Bíróné Kircsi, and Z. Konkolyné Bihari, "Return values of 60-minute extreme rainfall for Hungary," *Idojaras*, vol. 124, no. 2, pp. 143–156, 2020.

[32] E. Acuna and C. Rodriguez, "The treatment of missing values and its effect on classifier accuracy," in *Classification, Clustering, and Data Mining Applications: Proc. Meeting Int. Fed. Classification Societies (IFCS)*, Illinois Inst. Technol., Chicago, Jul. 15–18, 2004, pp. 639–647.

[33] G. Ragulina and T. Reitan, "Generalized extreme value shape parameter and its nature for extreme precipitation using long time series and the Bayesian approach," *Hydrol. Sci. J.*, vol. 62, no. 6, pp. 863–879, 2017.

[34] A. Diop and E. H. Deme, "Parametric bootstrapping in a generalized extreme value regression model for binary response," *arXiv preprint*, arXiv:2105.00489, 2021.

[35] T. Milojevic, J. Blanchet, and M. Lehning, "Determining return levels of extreme daily precipitation, reservoir inflow, and dry spells," *Front. Water*, 2023.