IJIE

INTERNATIONAL JOURNAL OF INTEGRATED ENGINEERING
ISSN: 2229-838X e-ISSN: 2600-7916

Vol. 18 No. 5 (2026) 153-165
https://publisher.uthm.edu.my/ojs/index.php/ijie

Multi-Test Homogeneity Assessment of Long-Term Rainfall Series in Peninsular Malaysia

Nur Zahidah Rosli[1], Jing Lin Ng[1]*, Norashikin Ahmad Kamal[1], Swee Pin Yeap[2], Nur Ilya Farhana Md Noh[1], Jin Chai Lee[3]

[1] *Faculty of Civil Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, MALAYSIA*
[2] *Department of Chemical & Petroleum Engineering, Faculty of Engineering, Technology & Built Environment, UCSI University, Kuala Lumpur, 56000, MALAYSIA*
[3] *Department of Civil Engineering, Faculty of Engineering, Technology & Built Environment, UCSI University, 56000 Kuala Lumpur, MALAYSIA*

*Corresponding Author: jinglin.ng@uitm.edu.my
DOI: https://doi.org/10.30880/ijie.2026.18.05.014

Article Info

Received: 3 October 2025
Accepted: 18 June 2026
Available online: 4 August 2026

Keywords

Homogeneity Test, Sequential Mann-Kendall (SQMK) Test, Von Neumann Ratio (VNR) Test, Pettitt Test, Buishand Range (BR) Test, Standard Normal Homogeneity Test (SNHT)

Abstract

The primary goal of this study is to analyse the homogeneity of rainfall in Peninsular Malaysia, crucial for understanding climate change impacts and water resource management. Despite the importance of rainfall data for hydrological and engineering applications, inconsistencies and inhomogeneities in long-term rainfall records can lead to misleading conclusions about climate trends and compromise infrastructure design. In Peninsular Malaysia, limited studies have systematically assessed rainfall data homogeneity, leaving uncertainty about the reliability of existing datasets. Given the rising implications of climate change, assessing the consistency and potential changes in rainfall patterns is essential for effective infrastructure and resource planning. Using rainfall data collected over a 21-year period (1998–2018) from 10 meteorological stations, the study employs various statistical methods, including the Von Neumann Ratio Test, Pettitt Test, Buishand Range Test, and Standard Normal Homogeneity Test, to evaluate temporal and spatial uniformity in rainfall patterns. Rainfall stations are categorised as 'useful', 'suspect', or 'doubtful' based on the results. 'Useful' stations show consistent and reliable data, while 'suspect' and 'doubtful' stations indicate potential inconsistencies. The findings in this study are Petaling Jaya in annual and Alor Setar in May were flagged as suspect, while Subang in May, Melaka in annual, Kuala Terengganu in August, and Bayan Lepas in January were considered doubtful. The remaining stations and time series exhibited consistent and reliable patterns. Additionally, the Sequential Mann-Kendall Test is used to detect trends and shifts in the data. The findings offer crucial insights into the homogeneity of rainfall distribution across Peninsular Malaysia, highlighting discrepancies that could impact water resource planning and climate change adaptation strategies. This research is valuable for engineers, researchers, and stakeholders involved in sustainable water resource management and infrastructure development, providing empirical evidence for hydrology and climate studies.

1. Introduction

A region with consistently similar rainfall over a long period is said to have homogeneous rainfall. Hydroclimatic studies in highly vulnerable locations require an extensive understanding of climate trends, variability and forecast [1]. However, before doing any research on climate change, agriculture, hydrology and water resources, data must be examined for consistency and inhomogeneities. To prevent unclear trend findings, the validity of experimentally observed data should be examined. Homogeneity tests on rainfall data have various advantages for engineers. Accurate rainfall data is essential for engineers building infrastructure, including stormwater management facilities, dams, and drainage systems. To ensure that designs sufficiently account for differences, homogeneity tests might help identify notable changes in rainfall patterns.

Various homogeneity tests have been developed to assess the consistency of rainfall data over time. The sequential Mann-Kendall (SQMK) test is an improvement over the classic Mann-Kendall (MK) test, intended to evaluate the homogeneity of trends in time-series data. Hanif et al. [2] conducted a study to examine the spatial variations in seasonal and annual rainfall patterns in Perak, Malaysia. The results revealed a clear upward trend in the monthly time series during the 1980s. The Von Neumann Ratio test (VNRT) is a nonparametric homogeneity test that is distinct from other tests [3]. Patakamuri et al. [4] conducted a study to assess the homogeneity, trend, and trend change points of rainfall data for the Anantapuramu district in the Andhra Pradesh State, India, using VNRT. The results showed that, of 27 stations, six experienced a clear increase in summer rainfall. In most locations, rainfall trends associated with the northeast and southwest monsoons were declining.

The Buishand Range (BR) Test is a parametric test used to determine breakpoints in a time series, assuming that the test results are independent and normally distributed. Khalil [5] conducted a BRT to determine the homogeneity of the rainfall series in the Mae Klong River Basin, Thailand. The results showed that 75% of the stations' monthly rainfall data could be considered 'useful'. The rainfall data from seven of the eight locations were considered "useful" on an annual basis, but one station was considered "suspect". The Pettitt test (PT) is a nonparametric test that is unaffected by the data's distribution. It is based on the Mann-Whitney test, which is rank-based and used to detect time-series shifts. Al-Bazaz et al. [6] conducted a study using PT to assess the homogeneity characteristics of monthly climatic data series from 1990 to 2020 in the Nineveh Governorate/Northern Iraq. The results showed that the homogeneity analysis of the highest monthly humidity was doubtful for several months.

The homogeneity tests employed in hydrological and climatological studies each provide distinct advantages and limitations. The SQMK test is effective in detecting step changes in time series, but can be biased due to its disregard for autocorrelation. The VNRT test provides a quick assessment of data continuity but does not indicate the strength or direction of the trend. The PT test is sensitive to mid-series breaks but requires large datasets for accuracy. The BR test assumes normality and independence but involves complex calculations. SNHT is useful for identifying anomalies in climate and hydrological data, but may produce errors when outliers or non-normal distributions are present.

In 2014, a massive flood in Peninsular Malaysia affected over 200,000 people and caused an estimated RM1 billion in economic losses [7]. Devastating floods occurred in Kota Tinggi in December 2006 and January 2007 because of cumulative rainfall exceeding 350mm over four days. Over 100,000 locals had to be evacuated due to this natural disaster, and it is estimated that USD 0.5 billion in economic losses have occurred. Understanding the effects and trends of climate change can be aided by identifying changes in rainfall patterns over time, such as variations in mean, variance, or distribution. Researchers usually identify regions lacking sufficient coverage or where contradictions arise by conducting a thorough literature review, examining previous research, and identifying gaps in the study. Once identified, these gaps become the primary focus of subsequent studies or inquiries, as they offer opportunities to add fresh perspectives, theories, methodologies, or empirical data. While the homogeneity test has been widely used worldwide, there has been little discussion in Peninsular Malaysia. In summary, this study aims to address the gap in the literature by conducting homogeneity testing of rainfall in Peninsular Malaysia.

2. Materials and Methods

2.1 Study Area

Peninsular Malaysia was used as the subject area for this case study, which is situated between latitudes 1° and 7°N and longitudes 100° and 104°E. The climate of Peninsular Malaysia is hot and muggy year-round due to its proximity to the equator. There is also consistent rainfall in the area. There are two main monsoon seasons that influence rainfall distribution in the study area: the Northeast Monsoon (NEM) and the Southwest Monsoon

(SWM). Additionally, there are the Inter-Monsoon 1 (INM1) and Inter-Monsoon 2 (INM2) seasons. The rainfall data used in this study were collected from 10 selected rainfall stations and obtained from the Malaysian Meteorological Department (MMD), covering a 20-year period from 1998 to 2018. The rainfall dataset was initially checked for quality and to ensure there were no missing rainfall values before proceeding with further in-depth research. If any rainfall values are missing, the gaps are filled using a multiple imputation approach. The list of stations and the distribution map are depicted in Table 1 and Fig. 1.

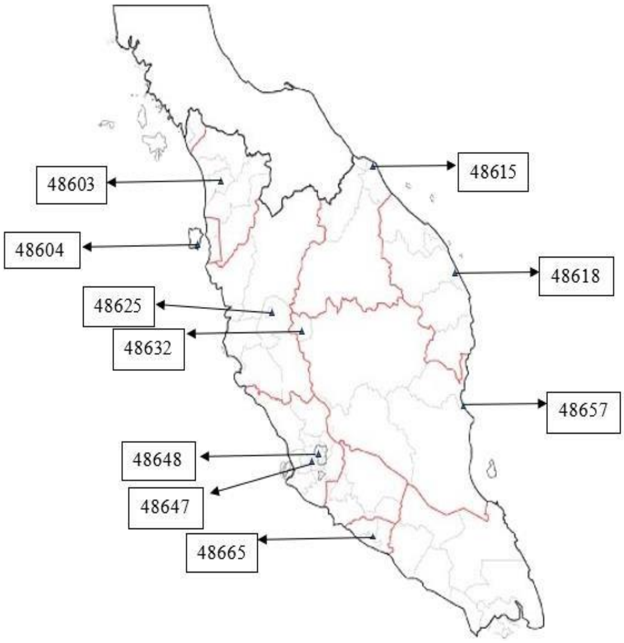


Fig. 1 *Rainfall stations distribution map*

Table 1 *Details of rainfall stations used in the study area*

No.	Station Code	Station Name	State	Record Period	Latitude	Longitude
1	48647	Subang	Selangor	1998 - 2018	3.11°N	101.55°E
2	48648	Petaling Jaya	Selangor	1998 - 2018	3.10°N	101.65°E
3	48615	Kota Bahru	Kelantan	1998 - 2018	6.17°N	102.28°N
4	48632	Cameron Highlands	Pahang	1998 - 2018	4.28°N	101.23°E
5	48657	Kuantan	Pahang	1998 - 2018	3.78°N	103.22°E
6	48665	Melaka	Melaka	1998 - 2018	2.27°N	102.25°N
7	48625	Ipoh	Perak	1998 - 2018	4.57°N	101.10°E
8	48618	Kuala Terengganu	Terengganu	1998 - 2018	5.32°N	103.13°E
9	48603	Alor Setar	Kedah	1998 - 2018	6.20°N	100.14°E
10	48604	Bayan Lepas	Penang	1998 - 2018	5.30°N	100.27°E

2.2 Missing Data Estimation

Data quality testing was conducted to ensure that the rainfall dataset was suitable for further analysis and to identify any missing rainfall values. The multiple imputation method was used to address gaps caused by missing data, where applicable. Homogeneity testing and missing data treatment were applied to the rainfall time series to ensure proper quality control of the dataset. The inverse distance weighted (IDW) interpolation method was used to repair missing data for a specific station identified during specific periods. This was done by utilising weighted-average data from any of the nine other stations adjacent to the station in question. IDW is a deterministic technique for multivariate interpolation based on a known, distributed collection of points. A weighted average of the values available at the known points is used to compute the assigned values to the unknown points. The formula is used to calculate the amount of precipitation at a place that is not included in the measurement ($\hat{P}_0$) [8].

$$\hat{P}_0 = \sum_i^n \frac{P_i}{D_{oi}^p} / \sum_i^n \frac{1}{D_{oi}^p} \quad (1)$$

where D_{0i} is the distance between the rain gauge and the estimated portion of the precipitation field, and P_i is the quantity of rainfall recorded at the rain gauge. The number of rain gauges at site (i, n) is used to estimate precipitation. The power exponent 0, $\hat{P}_0$ is in charge of giving each rain gauge a significant weight.

2.3 Extraction of Extreme Rainfall Data

Gathering and evaluating data on precipitation events that fall short of normal or anticipated thresholds in a particular area or time period is the process of extracting extreme rainfall data. Hydrology, climate research, engineering, and risk assessment are just a few of the fields that depend on this data. The extreme rainfall data were obtained as two series: monthly and annual. The monthly data were combined to become the total annual rainfall for all stations over time to discover any inhomogeneity in the rainfall data on an annual scale. Monthly rainfall data is the total quantity of precipitation in a specific location or region during a given month. The total daily rainfall measurements for each day of the month are used to construct this data. On the other hand, annual rainfall data show the total amount of precipitation received in a particular place or region over the course of a year. The monthly rainfall totals for each of the 12 months are added to compute the total.

2.4 Homogeneity Testing

In climatological investigations, homogeneity testing is essential to accurately represent the true fluctuations in weather and climate. Many factors, such as human error, changes in the instrument's surrounding areas, and technical errors, can cause inhomogeneity in climate data. Results indicating incorrect trends will be shown if homogeneity is not checked before trend analysis [4]. Five homogeneity tests are used to assess the homogeneity of the rainfall data: the Sequential Mann-Kendall (SQMK) Test, Von Neumann Ratio (VNR) Test, Pettitt Test, Buishand Range (BR) Test, and Standard Normal Homogeneity Test (SNHT).

The SQMK test was used to determine the sudden changes in rainfall indices. It is extremely helpful in determining the time series' consecutive step changes. The sequence statistic S_k for a time series xi where ($1<i\leq n$) must be calculated using Eqn. (2) as the first step in the SQMK test [2].

$$S_k = \sum_{i-1}^{k} r_i \; (k = 2,3,\ldots,n) \tag{2}$$

$$r_i = \{+1 \; if x_i > x_{j,} 0 if x_i \leq n_j\}(j = 1,2,\ldots,i) \tag{3}$$

where n indicates the length of the data series. x_i, x_j, on the other hand, are the data values at times i and j, respectively. r_i is the rank statistics for data pairings (x_i, $x_{j,}$), and S_k is the rank series. In this test, U(t) indicates the forward trend, calculated progressively from the beginning of the rainfall series, while U′(t) represents the backward trend, computed in reverse order from the end of the series.

The BR test can detect the shift in the intermediate. It can be more vulnerable to time series breakpoints in the middle. The definition of the adjusted partial sums is as follows [9]:

$$s_o^* = o \; and \; s_k^* = \sum_{i=1}^{k}(P_i - \overline{P}) \; where \; k = 1,2,3,\ldots,n. \tag{4}$$

where P is the average, and P_i is the time series that needs to be evaluated. If the series is homogeneous, there are no variations found in the P_i values from the average. If the series is homogeneous, s_k^* values will fluctuate about zero. The following expression for the value of R can be used to determine the significance of the change in the average:

$$R = \frac{(\max_{0\leq k\leq n} S_k^* - \min_{0\leq k\leq n} S_k^*)}{s} \tag{5}$$

The VNR test assesses the randomness of the time series. It offers a quick and easy method to assess whether the data is broken or non-homogeneous. The test is sensitive to the point at which a given time series' homogeneity breaks down. The test calculated the variance by dividing the average of the squared consecutive (year-to-year) differences. Below is a definition of the statistic [9]:

$$N = \frac{\sum_{i=1}^{n-1}(P_i - P_{i+1})^2}{\sum_{i=1}^{n}(P_i - \overline{P})^2} \tag{6}$$

where P is the average, and P_i is the time series that needs to be evaluated.

The homogeneity of a time scale according to an element's ranks is assessed using the non-parametric Pettitt Test. The following is a definition for the statistic [9]:

Int. Journal of Integrated Engineering Vol. 18 No. 5 (2026) p. 153-165

$$x_k = 2\sum_{i=1}^{k} r_i - k(n+1) \text{ where } k = 1,2,3,\ldots,n \quad (7)$$

The test's variates are assumed to be regularly distributed by the SNHT. Inhomogeneity near the beginning and end of a given time scale can be detected using the SNHT. The following is an expression for the statistic [9]:

$$\overline{z_1} = \frac{1}{k}\sum_{i=1}^{k}\frac{P_i - \overline{P}}{s} \text{ and } \overline{z_2} = \frac{1}{n-k}\sum_{i=k+1}^{n}\frac{Y_i - \overline{Y}}{s} \quad (8)$$

where the arithmetic averages of the z_i sequences before and after the shift are represented by the numbers $\overline{z_1}$ and $\overline{z_2}$. The time series that need to be examined are Y_i and P_i. The standard deviation is represented by the letter s, while the averages are P and Y.

3. Result and Discussion

As shown in Fig. 2, the trends in monthly rainfall in Subang, Petaling Jaya, Kota Bharu, and Cameron Highlands from 1998 to 2018, using the SQMK test, were displayed in February. In the graph, the blue line represents the prograde trend, U(t), while the orange line indicates the retrograde trend, U'(t). Progradеd trends show a long-term upward trend in the data. This indicates that, when analysed sequentially, the variable (rainfall, temperature, etc.) shows a general upward trend. In Subang, the rainfall trend in February shows noticeable fluctuations. In the u(t) and u'(t) graph, the curves intersect at three points, specifically in the years 2000, 2014 and 2016. The rainfall trend series increased from 2000 to 2003. These findings are consistent with those of Jamaluddin et al. [10], which summarised that, because the Titiwangsa Mountains act as a rain shadow, causing the west coast to receive comparatively less rainfall than the east coast, Subang Jaya experiences fluctuations in rainfall trends during the Northeast Monsoon. Furthermore, the p-values for the PT (0.221), SNHT (0.204), and BR (0.165) tests in Subang were all above 0.05. This indicates that no statistically significant inhomogeneity is detected.

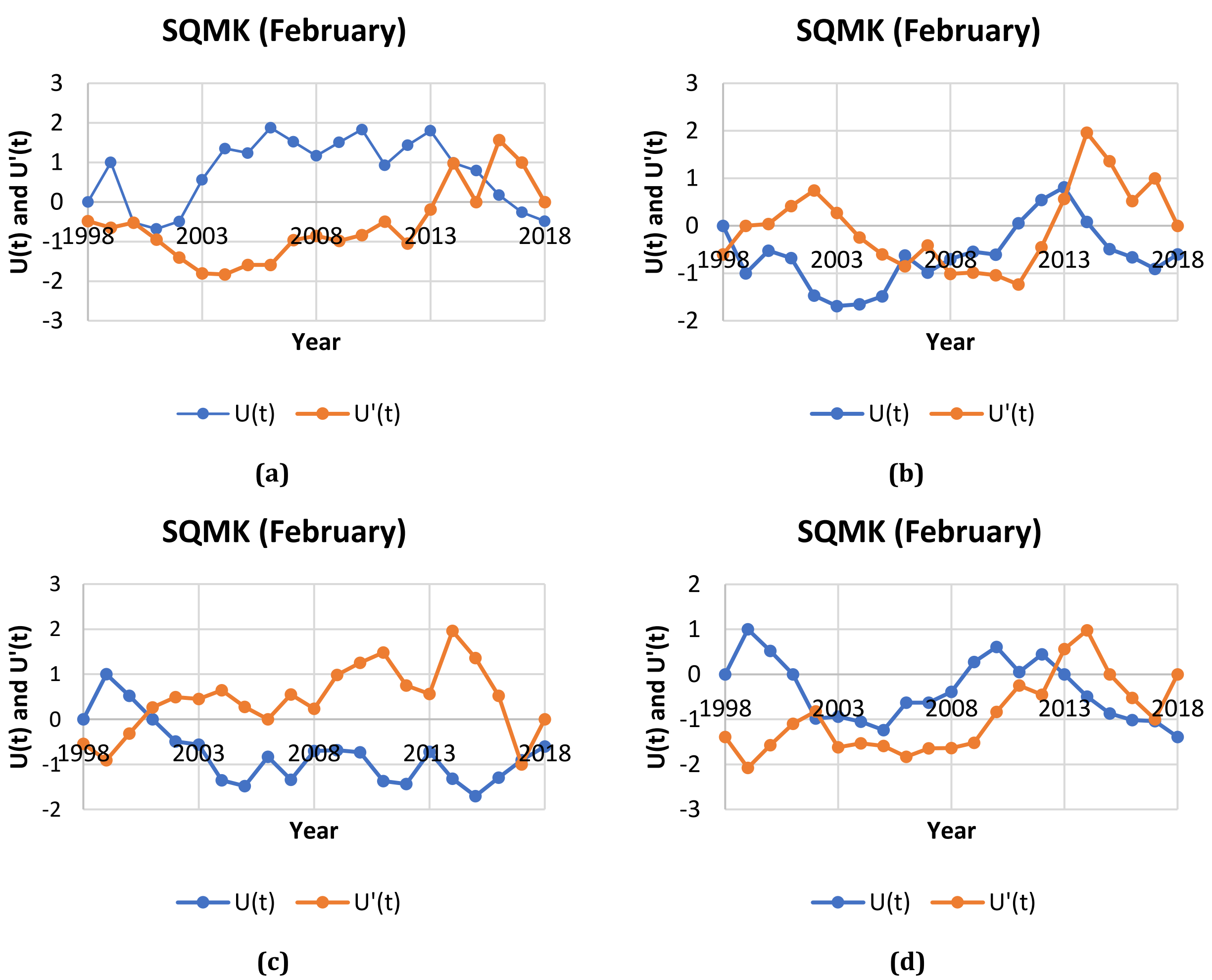


Fig. 2 *Trend of monthly rainfall (February) in (a) Subang; (b) Petaling Jaya; (c) Kota Bharu; and (d) Cameron Highlands. U(t) indicates the forward trend, and U′(t) represents the backward trend*

As shown in Fig. 3, the trends in monthly rainfall in Subang, Petaling Jaya, Kota Bharu, and Cameron Highlands from 1998 to 2018, using the SQMK test, were displayed in May. In Cameron Highlands, there was a fluctuation in the rainfall trend in May. The graphs of u(t) and u'(t) had a notable trend break point of rainfall in 2013 and continued increasing until 2018. These findings are consistent with those of Rasdi et al. [11], who found that rainfall in the Cameron Highlands has been increasing in some months. In particular, March to May are the months when the area receives the most rainfall. This pattern is consistent with earlier research showing increased rainfall in the area, which raises the risk of landslides and flooding. The p-values in Cameron Highlands for the PT (0.427), SNHT (0.341), and BR (0.183) tests were all above 0.05.

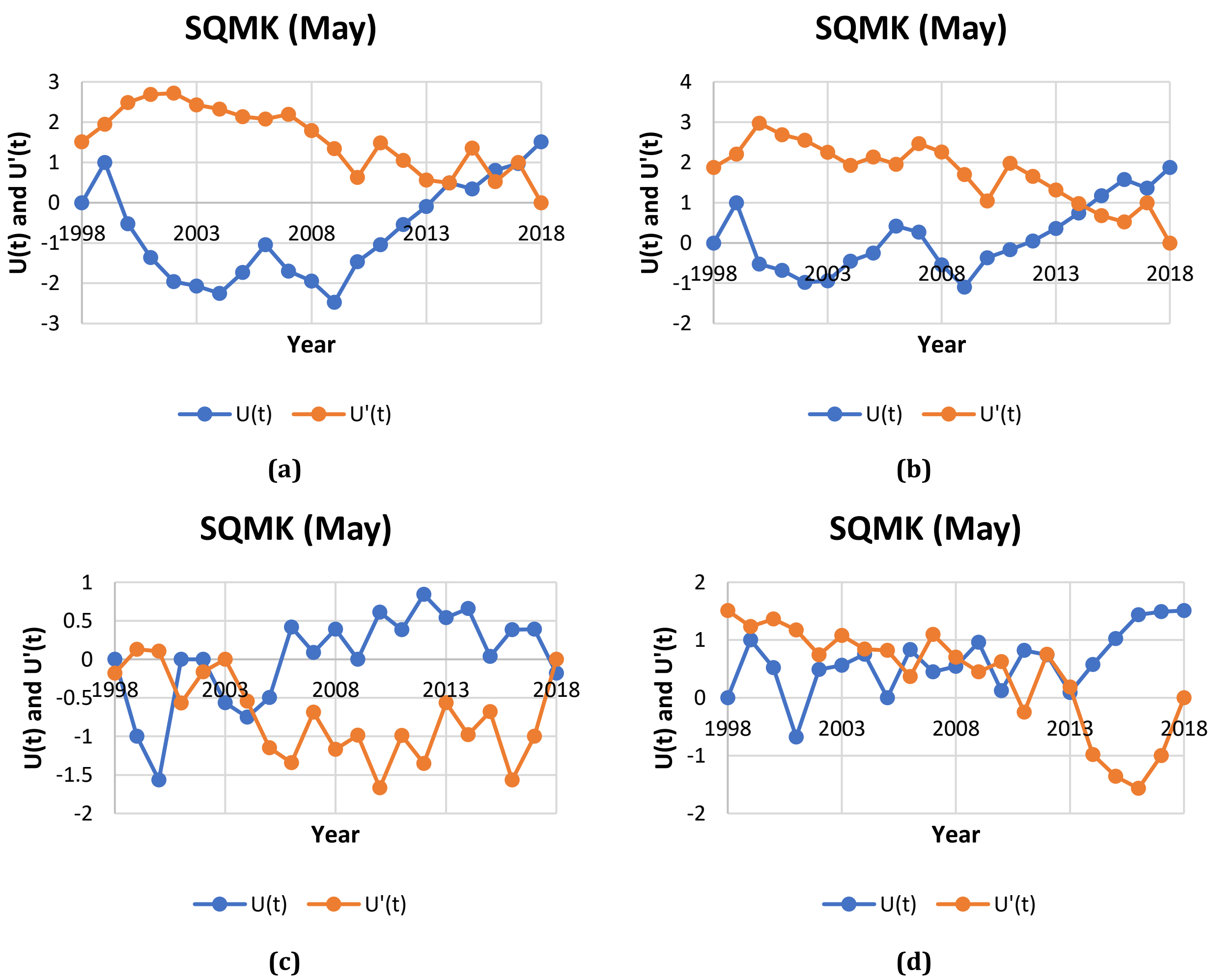


Fig. 3 *Trend of monthly rainfall (May) in (a) Subang; (b) Petaling Jaya; (c) Kota Bharu; and (d) Cameron Highlands. U(t) indicates the forward trend, and U′(t) represents the backward trend*

On the other hand, as shown in Fig. 4, the trends in monthly rainfall in Subang, Petaling Jaya, Kota Bharu, and Cameron Highlands from 1998 to 2018, using the SQMK test, were displayed in August. Kota Bahru showed a clear break trend of rainfall in August. The graph's u(t) and u'(t) curves displayed a clear intersection in 2006, and the trend continued to exhibit an increasing trend of rainfall until 2015. These results are in line with the Malaysian Meteorological Department (2024), which stated that the general seasonal rainfall pattern in Kota Bharu and the surrounding area includes heavy rainfall in August. August falls under the wetter season because there are two rainfall maxima: one from October to November and another from April to May. Besides that, Fig. 5 illustrates the trends in monthly rainfall in Subang, Petaling Jaya, Kota Bharu, and Cameron Highlands from 1998 to 2018, as determined by the SQMK test, which was displayed in September. It was found that rainfall fluctuated in Subang in September. The curves in the u(t) and u'(t) graphs intersect clearly in 2002, which showed a break point of the rainfall trend. Then the graph continued to show a decreasing trend from 2002 until 2017. Similar to Petaling Jaya, there was a clear cross visible in 2011 on the curves in the u(t) and u'(t) graphs. From 2011 to 2018, rainfall showed a decreasing trend. These findings are similar to those of Saimi et al. [12], who reported a statistically significant drop in rainfall in September, which falls within the Southwest Monsoon season. In Subang, the p-values from the Pettitt (0.988), SNHT (0.336), and Buishand Range (0.566) tests for September were all greater than the 0.05 significance level, indicating that no statistically significant breakpoints or inhomogeneities were detected in the rainfall series. Similarly, in Petaling Jaya, the p-values for the Pettitt (0.651), SNHT (0.791), and Buishand Range (0.793) tests in September also exceeded 0.05.

As shown in Fig. 6, the trends in monthly rainfall in Subang, Petaling Jaya, Kota Bharu, and Cameron Highlands from 1998 to 2018, using the SQMK test, were displayed in October. It was found that Petaling Jaya showed a clear break in rainfall in 2006, and the trend continued to increase until 2008. These findings are consistent with those of Muhammad et al. [7], who concluded that September and October are important months for rainfall patterns in a study of Peninsular Malaysia's rainfall features. It is specifically mentioned that these months are a part of the intermonsoon season, which is characterised by afternoon convective rain events. While in Kota Bharu, it showed a clear break in the rainfall pattern and the u(t) and u'(t) curves on the graph showed a distinct intersection in 2003. The rainfall trend continued to rise until 2005. These findings are in line with those of Muhammad et al. [7], which summarised that the Northeast Monsoon season is the main cause of the sharp rise in rainfall Kota Bharu experiences throughout October. The region's vulnerability to significant rainfall during the monsoon season is highlighted by this trend, which is consistent with observations from other stations in the area. In Petaling Jaya, the p-values obtained in October for the Pettitt (0.514), SNHT (0.346), and Buishand Range (0.185) tests were all greater than 0.05, indicating that no statistically significant breakpoints were detected in the rainfall series. Similarly, in Kota Bharu, the October p-values for the Pettitt (0.665), SNHT (0.298), and Buishand Range (0.377) tests also exceeded 0.05.

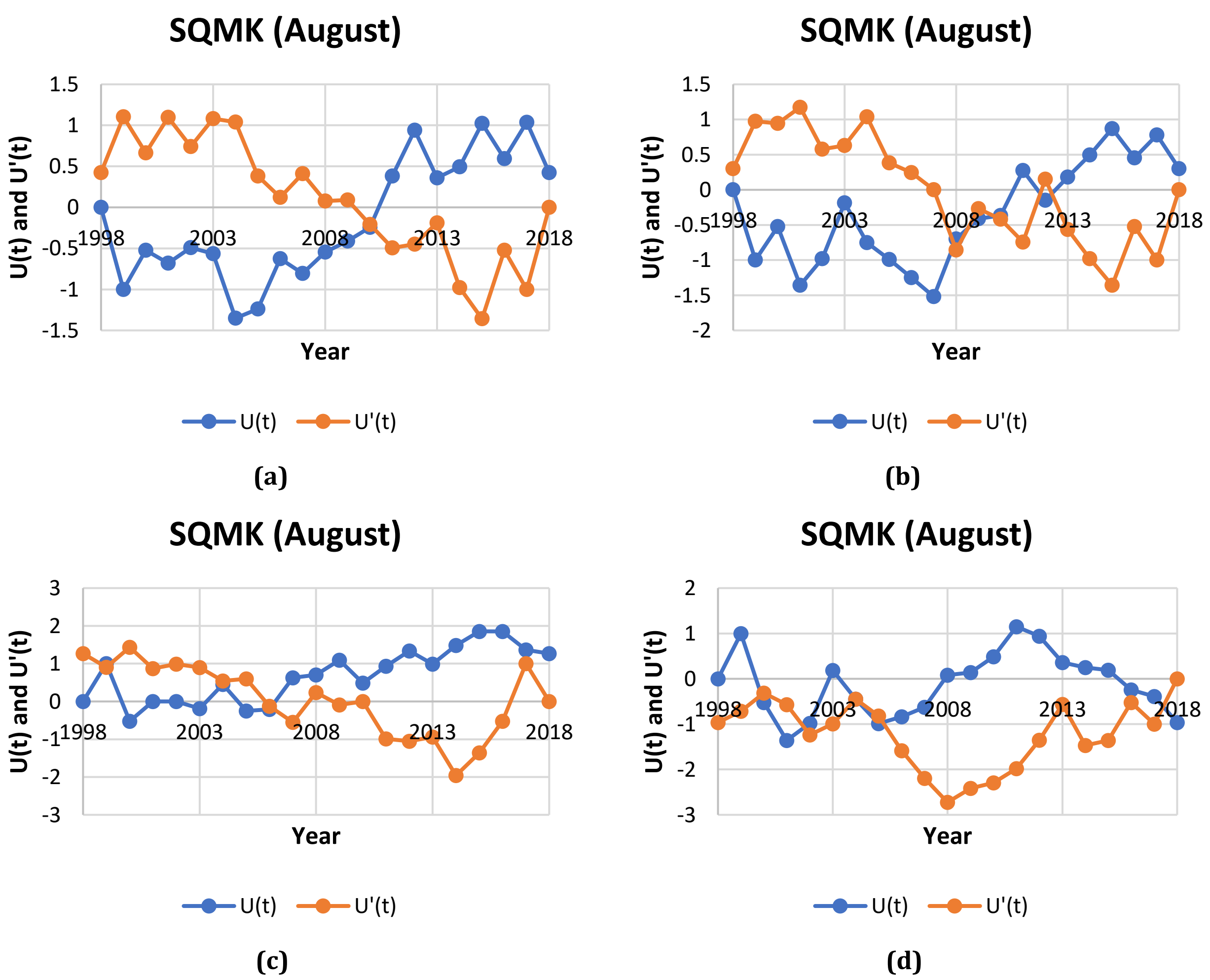


Fig. 4 *Trend of monthly rainfall (August) in (a) Subang; (b) Petaling Jaya; (c) Kota Bharu; and (d) Cameron Highlands. U(t) indicates the forward trend, and U′(t) represents the backward trend*

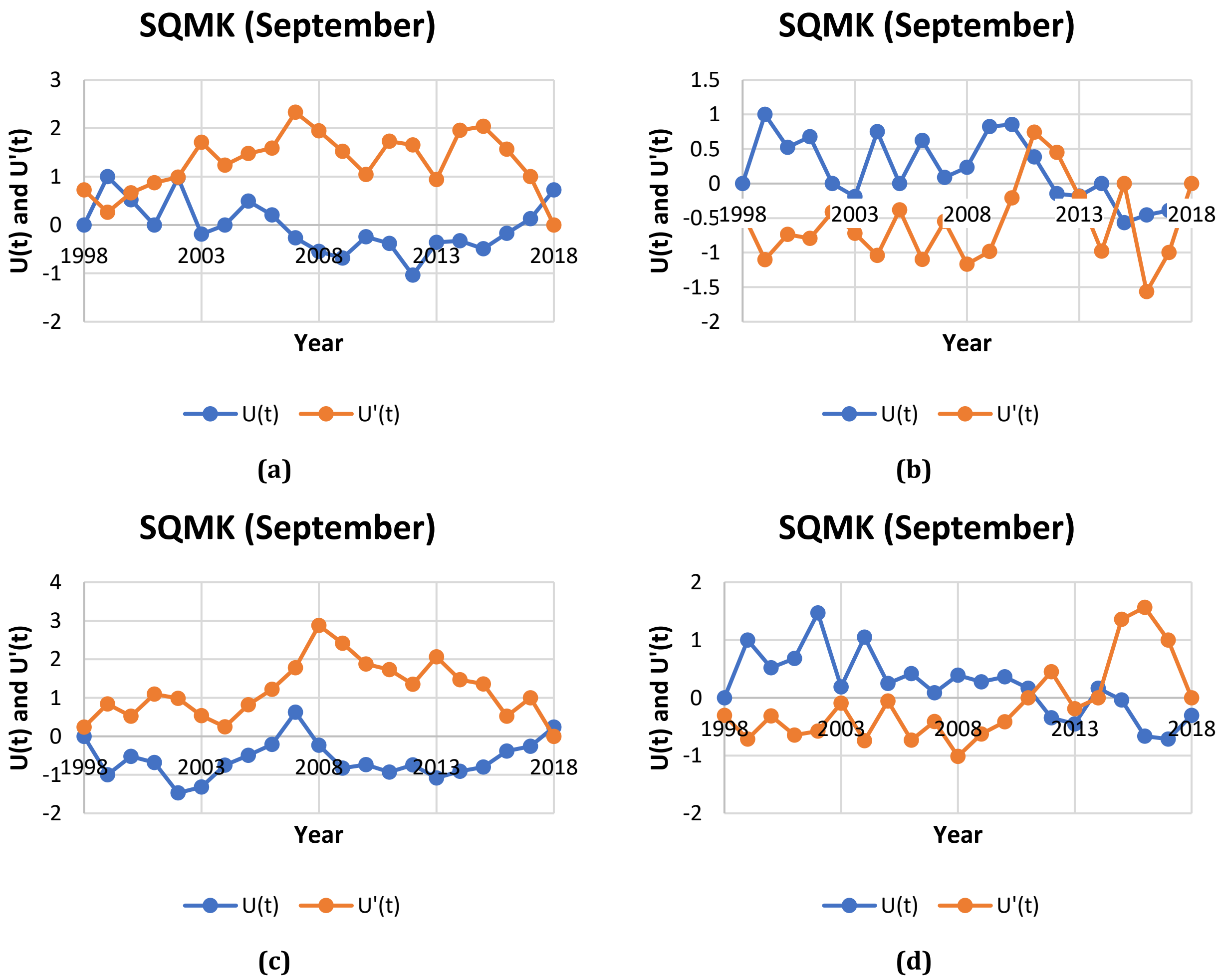
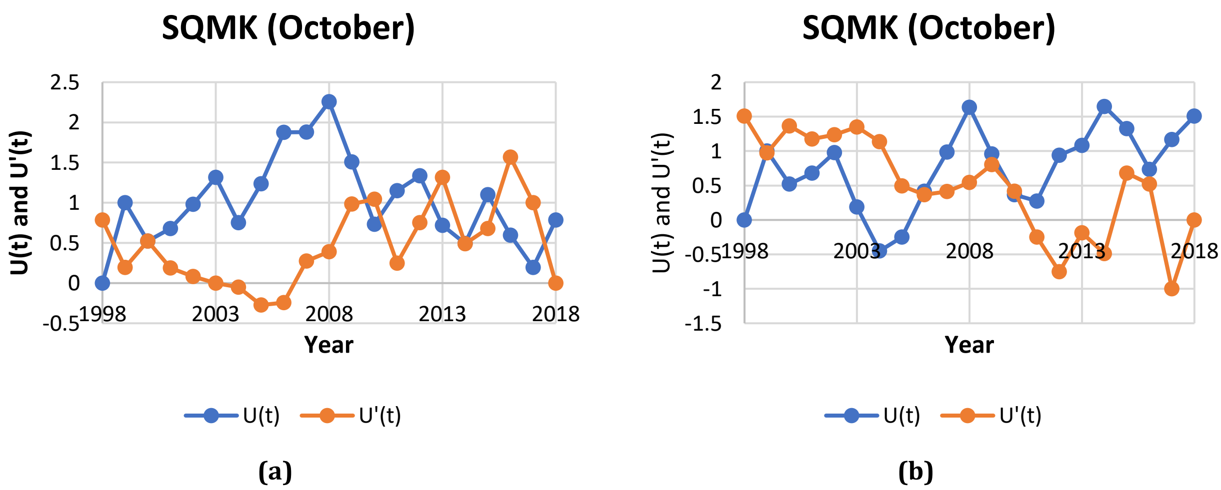


Fig. 5 *Trend of monthly rainfall (September) in (a) Subang; (b) Petaling Jaya; (c) Kota Bharu; and (d) Cameron Highlands. U(t) indicates the forward trend, and U′(t) represents the backward trend*



Int. Journal of Integrated Engineering Vol. 18 No. 5 (2026) p. 153-165

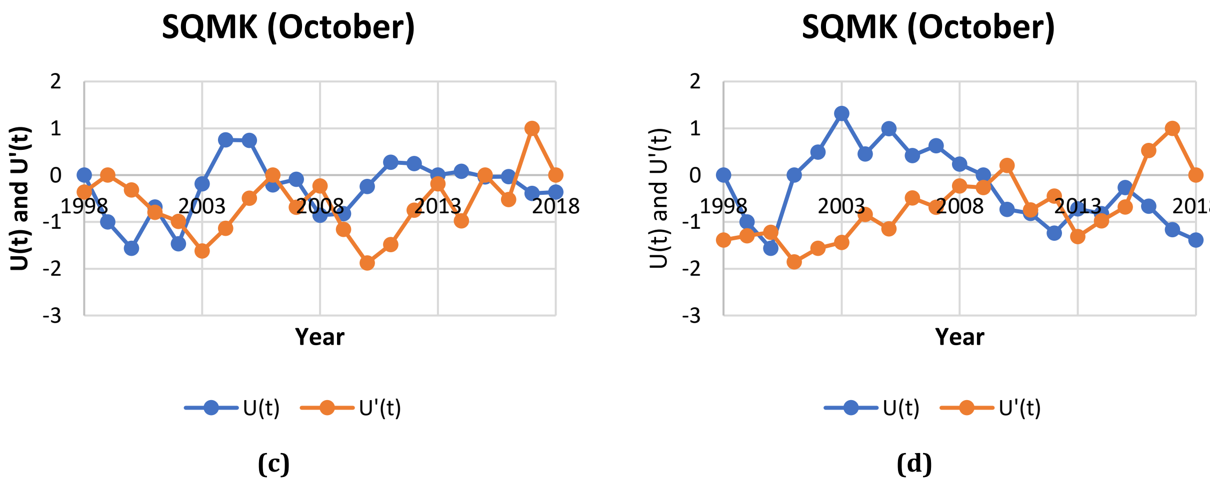


Fig. 6 *Trend of monthly rainfall (October) in (a) Subang; (b) Petaling Jaya; (c) Kota Bharu; and (d) Cameron Highlands. U(t) indicates the forward trend, and U′(t) represents the backward trend*

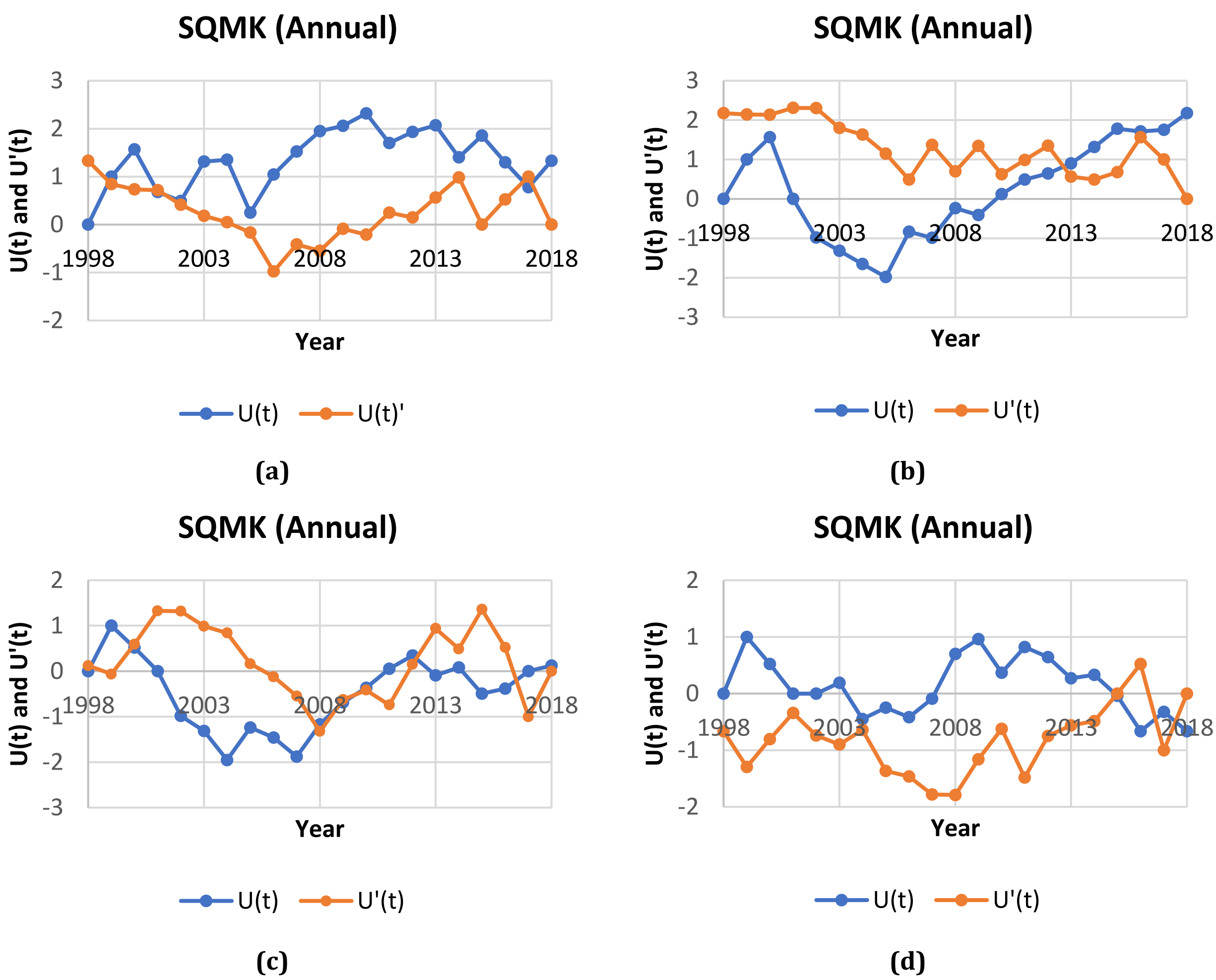


Fig. 7 *Sequential Mann-Kendall plots for (a) Subang, annual; (b) Petaling Jaya, annual; (c) Kota Bharu, annual; (d) Cameron Highlands, annual. U(t) indicates the forward trend, and U′(t) represents the backward trend*

Fig. 7 illustrates the annual rainfall trend fluctuations in Subang, Petaling Jaya, Kota Bharu, and Cameron Highlands, as determined using the SQMK test. In Subang, the variation in the rainfall trend showed a trend break point in the year 2002, and the trend showed continuous increasing of rainfall until 2015 in the curves of the u(t) and u'(t) graphs. Comparable to Petaling Jaya, which showed the increasing trend from 2005 until 2018, and there was a point of crossover in 2013. These results were consistent with the findings of Ramli et al. [13], who used statistical techniques to analyse the considerable changes in rainfall patterns observed in the Klang Valley. The study concluded that climate change was likely responsible for the differences in rainfall patterns, which have consequences for flood frequency in the area.

Table 2 *P-Values of Subang*

Month/Annual	Pettitt	SNHT	Buishand Range	Von Neumann Ratio
January	0.200	0.350	0.814	0.127
February	0.221	0.204	0.165	0.386
March	0.516	0.921	0.756	0.295
April	0.897	0.187	0.531	0.272
May	**0.021**	0.061	**0.018**	0.058
June	0.310	0.174	0.268	**0.019**
July	0.884	0.705	0.562	0.417
August	0.072	0.193	0.963	0.845
September	0.988	0.336	0.566	0.754
October	0.197	0.508	0.983	0.707
November	0.762	0.553	0.677	0.897
December	0.900	0.631	0.376	0.395
Annual	0.251	0.147	0.077	0.734

Table 3 *P-Values of Petaling Jaya*

Month/Annual	Pettitt	SNHT	Buishand Range	Von Neumann Ratio
January	0.878	0.609	0.547	0.567
February	0.672	0.286	0.296	0.053
March	0.770	0.185	0.471	0.499
April	0.504	0.160	0.184	0.180
May	0.096	0.072	**0.034**	0.098
June	0.896	0.606	0.641	**0.045**
July	0.492	0.689	0.425	0.394
August	0.874	**0.047**	0.848	0.511
September	0.651	0.791	0.793	0.927
October	0.514	0.346	0.185	0.285
November	0.200	0.897	0.939	0.982
December	0.180	0.185	0.090	0.169
Annual	**0.020**	**0.007**	**0.006**	0.331

Table 4 *P-Values of Kota Bharu*

Month/Annual	Pettitt	SNHT	Buishand Range	Von Neumann Ratio
January	0.589	0.127	0.097	0.217
February	0.372	0.217	0.364	0.402
March	0.586	0.585	0.470	0.281
April	0.067	0.463	0.713	0.487
May	**0.031**	0.717	0.847	0.675
June	0.503	0.235	0.220	0.661
July	0.765	0.971	0.980	0.574
August	0.419	0.561	0.303	0.745
September	0.745	0.771	0.631	0.419
October	0.388	0.955	0.824	0.261
November	0.765	0.981	0.884	0.717
December	0.628	0.616	0.847	0.937
Annual	0.854	0.676	0.492	0.202

Based on the results of the other homogeneity tests, as shown in Table 2 to Table 5, the rainfall data time series can be further classified into three classes: Class I (useful), Class II (doubtful) and Class III (suspect). The classification of rainfall data series was based on the results of the homogeneity tests at the 5% significance level [14]. A series was considered useful if all tests accepted the null hypothesis of homogeneity (p-value > 0.05), or if at most one test failed. It was deemed doubtful when two tests rejected homogeneity and suspect when three or more tests indicated inhomogeneity. Table 6 provided an overview of the homogeneity test outputs, which were classified. All rainfall data for Class I were uniform across all time scales except the monthly and annual scales, as shown in the table, making them suitable for further study. Next, rainfall data for Subang in May, Melaka in Annual, Kuala Terengganu in August and Bayan Lepas in January were categorised as Class II (doubtful) due to observed inhomogeneity, which may be attributed to factors such as rapid urban development, station relocation, or changes in measurement instruments. If they were to be considered for further analysis, they needed to be handled cautiously. Since the rainfall data for class III were not uniform across all time scales, they could not be used for additional study.

Table 5 *P-Values of Cameron Highlands*

Month/Annual	Pettitt	SNHT	Buishand Range	Von Neumann Ratio
January	0.098	**0.043**	0.061	0.052
February	0.264	0.443	0.344	0.346
March	0.898	0.556	0.396	0.224
April	0.208	0.699	0.932	0.591
May	0.427	0.341	0.183	0.829
June	0.777	0.556	0.354	0.256
July	0.901	0.817	0.610	0.257
August	0.596	0.460	0.457	0.070
September	**0.013**	0.990	0.978	0.979
October	0.502	0.300	0.144	0.679
November	0.497	0.461	0.581	0.623
December	0.293	0.998	0.976	0.971
Annual	0.526	0.844	0.715	0.552

Table 6 *Summary of Homogeneity Results*

Station Name	Class II (Doubtful)	Class III (Suspect)
Subang	May	-
Petaling Jaya	-	Annual
Kota Bharu	-	-
Cameron Highlands	-	-
Kuantan	-	-
Melaka	Annual	-
Ipoh	-	-
Kuala Terengganu	August	-
Alor Setar	-	May
Bayan Lepas	January	-

4. Conclusion

In this study, homogeneity tests of monthly and annual rainfall were investigated in Peninsular Malaysia using data from 10 stations, consisting of Subang, Petaling Jaya, Kota Bharu, Cameron Highlands, Kuantan, Melaka, Ipoh, Kuala Terengganu, Alor Setar and Bayan Lepas, over a 21-year period. The Buishand Range (BR) test, Standard Normal Homogeneity (SNHT), Von Neumann Ratio (VNR) test, Pettitt test (PT), and Sequential Mann-Kendall test (SQMK) were used to assess the homogeneity of the dataset collected for this study. All stations showed 'useful' results except for January, May, August, and Annual for Bayan Lepas, Subang, Kuala Terengganu, and Melaka, respectively, which showed 'doubtful' according to the homogeneity test results. In addition, there were two rainfall stations classified as 'suspect': Annual for Petaling Jaya and May for Alor Setar.

The successful completion of a thorough investigation into the homogeneity of Peninsular Malaysia's rainfall series has provided important insights into the dependability and consistency of the region's precipitation data. A variety of disciplines, including meteorology, engineering and agriculture, depend on homogeneity testing of rainfall data. In engineering practice, reliance on homogeneous data can lead to the underestimation of design storms, resulting in undersized drainage systems, flood-prone developments or other hydraulic structures. Homogeneity testing ensures that rainfall data used in stormwater design, watershed modelling, and flood risk mapping reflect real climatic trends rather than artificial inconsistencies. It also improves the accuracy of hydrological forecasts, informs irrigation planning, and supports agricultural resilience. The study contributes to SDGs 6 and 11 by supporting data-driven planning for water infrastructure and flood resilience in urban areas, particularly amid increasing rainfall variability driven by climate change (SDG 13). To improve future homogeneity assessments and ensure data suitability for hydrological applications, future studies should extend the temporal span of rainfall data to over 50 years to capture long-term trends and inhomogeneities more accurately. Moreover, developing and applying statistical correction methods, such as quantile matching, regression-based adjustments, or machine learning techniques, can improve the reliability of historical rainfall data. Integrating climate indices such as ENSO, the Indian Ocean Dipole (IOD), and monsoon indices is also recommended to explore the physical linkages between climate oscillations and rainfall variability. Furthermore, future studies are encouraged to incorporate spatio-temporal synthesis techniques, such as spatial trend mapping, cluster analysis, or Principal Component Analysis (PCA), to more effectively capture and interpret the spatial variability and underlying patterns in rainfall homogeneity across different regions.

Acknowledgement

The authors would like to express gratitude to the Faculty of Civil Engineering, Universiti Teknologi MARA (UiTM), for the financial support. Special appreciation goes to the Malaysia Meteorological Department (MMD) for providing the weather data.

Conflict of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

Author Contribution

The authors confirm contribution to the paper as follows: ***Study conception and design:*** *Jing Lin Ng, Norashikin Ahmad Kamal, Swee Pin Yeap;* ***Data collection:*** *Nur Zahidah Rosli;* ***Analysis and interpretation of results:*** *Nur Zahidah Rosli, Jing Lin Ng, Jin Chai Lee;* ***Draft manuscript preparation:*** *Nur Zahidah Rosli, Jin Chai Lee, Nur Ilya Farhana Md Noh.* ***All authors*** *reviewed the results and approved the final version of the manuscript.*

References

[1] Daba, M. H., Ayele, G. T., & You, S. (2020). Long-term homogeneity and trends of hydroclimatic variables in Upper Awash River Basin, Ethiopia. Advances in Meteorology, 2020, 1–21,

[2] Hanif, M. F., Mustafa, M. R. U., Liaqat, M. U., Hashim, A. M., & Yusof, K. W. (2022). Evaluation of long-term trends of rainfall in Perak, Malaysia. Climate, 10, 44.

[3] Ng, C. K., Ng, J. L., Huang, Y. F., Tan, Y. X., & Mirzaei, M. (2020). Tropical rainfall trend and stationarity analysis. Water Supply, 20, 2471–2483.

[4] Patakamuri, S. K., Muthiah, K., & Sridhar, V. (2020). Long-term homogeneity, trend, and change-point analysis of rainfall in the arid district of Ananthapuramu, Andhra Pradesh State, India. Water, 12, 211.

[5] Khalil, A. (2021). Inhomogeneity detection in the rainfall series for the Mae Klong River Basin, Thailand. Applied Water Science, 11, 147.

[6] Al-Bazaz, H. J. (2023). A study of the homogeneity of climatic data for rain, temperature and humidity for Nineveh Governorate. Al-Rafidain Engineering Journal, 28, 173–185.

[7] Muhammad, N. S., Abdullah, J., & Julien, P. Y. (2020). Characteristics of rainfall in Peninsular Malaysia. Journal of Physics: Conference Series, 1529, 052014.

[8] Gilewski, P. (2021). Impact of the grid resolution and deterministic interpolation of precipitation on rainfall-runoff modeling in a sparsely gauged mountainous catchment. Water, 13, 230.

[9] Tan, X. Y. (2021). Trend analysis of seasonal rainfall patterns in Peninsular Malaysia. IOP Conference Series: Materials Science and Engineering, 1101, 012017.

[10] Jamaluddin, A. F., Kamal, M. I. M., Abdullah, M. H., & Marodzi, A. N. (2022). Comparison between satellite-derived rainfall and rain gauge observation over Peninsular Malaysia. Sains Malaysiana, 51, 67–81.

[11] Rasdi, R. M., & Rahman, M. H. A. (2022). Vulnerability of agriculture to climate change events in the upper basin system: A review. Malaysian Journal of Tropical Geography, 48, 32–40.

[12] Saimi, F. M., Hamzah, F. M., Toriman, M. E., Jaafar, O., & Tajudin, H. (2020). Trend and linearity analysis of meteorological parameters in Peninsular Malaysia. Sustainability, 12, 9533.

[13] Ramli, M. K., & Masbah, D. I. (2023). Statistical analysis of temporal variations in patterns of rainfall and temperature in the Klang Valley. Construction, 3, 316–325.

[14] Ng, J. L., Aik, H. C., Jin, J. C. L., Noh, N. I. F. M., Abdulkareem, M., Syamsunur, D., & Siaw, Y. T. (2025). Machine learning-based approach for filling gaps in streamflow data. Semarak International Journal of Machine Learning, 5, 46–63.