Trans. Inst. Indian Geographers
ISSN 0970-9851
Indexed in Scopus

Systematic true trend detection based on homogeneity analysis: Application to rainfall time series for Thiruvallur district, Tamil Nadu, India

Venkatesan Madha Suresh, Tamil Nadu, India and Ibrahim Lawal Kane*, Nigeria

Abstract

This study investigates the monthly rainfall trends in Thiruvallur district, India through an analysis of historical rainfall data for every month for the period of 1971 to 2016. The individual test data is analysed using the Mann-Kendall trend test and the results revealed an upward trend in all the data sets excluding Vallur anicut and Thiruvottiyur which shows a decrease in rainfall over the years. However, anthropogenic factors such as rain gauge location, faulty equipment, and changes in instrument etc. affect the long-term climatic records, which may produce either inhomogeneity or bias in the series. These make most climate data unpredictable and may show spurious temporal climate variability. Pettitt's test is employed to avoid false trend detection and also change points in the time series. All the series change points were detected and were split according to the break dates; the homogeneity is confirmed using descriptive statistics and F-test. The results revealed homogeneity in all the series which confirmed the Mann Kendall trend test results. The outcomes show the inferences in water resource management help in future water demand in agricultural and basic water supply. The observed decrease in rainfall trend in Vallur anicut and Thiruvottiyur areas is alarming, therefore, in a need to implement water security policies to achieve sustainable water resource management.

Keywords: *Trend, homogeneity, Mann Kendall, Pettitt's test, t-test*

Introduction

The variability and the changes in the climate over the past few decades in the world is a serious concern. IPCC (2007) reports that an increase in global temperature, phenomenal changes in precipitation, and snow cover flight are some of the major factors attributed to climate change. The IPCC (2019) also reveals that climate change causes the degradation of land through an increase in rainfall intensity, drought, heat stress, flood and wave action, etc. modulated by land management. Climate change also affects the food security situation due to warming, changes in rainfall, etc. Ever since the reports were made public, trend detection in precipitation and temperature time series became one of the motivating research areas in climate change studies. Researchers working on climate change more specifically focus on the rainfall occurrence and its distribution and these are the most widely used approach for sustainable water resource management. India requires major research focus on the identification and quantification of climate change for sustainable development of agriculture (Meshram, 2017). The extreme weather patterns and events of a location make better decision-making and advance mitigation to understand the precipitation patterns (Bushra, 2020). It is well known that precipitation changes are not globally uniform; regional temporal variations exist across geographical regions. Consequently, the scientific community focuses on the analysis of rainfall trends on different temporal and spatial scale during the past century which has been a subject for discussion. Rainfall is the most important and essential element for a region which changes the climate and environmental conditions due to which the temporal and spatial patterns also undergo significant changes (Javari, 2016). Understanding the trend in rainfall variability is important in analysing the rainfall data (Dhakal*et al.* 2015). Pranuthi *et al.* (2014) made use of the Mann-Kendall and linear regression method to analyse long-term trends in the Dehradun, Almora, and Haridwar cities of Uttarakhand state of India. The change points were noticed and were detected using the cumulative deviation and the Worsley likelihood statistics. Haridwar district showed a significant increase in rainfall by the Mann-Kendall Test. A change point was observed with a distinct difference from 2009 onwards. Palaniswami and Muthiah (2018) worked on the changes caused due to rainfall and temperature associated with the northern parts of Tamilnadu using the Mann-Kendall, Sen's slope test, and the change point detection test. The study revealed a significantly rising trend in temperature for annual maximum, minimum, and mean temperatures. The results obtained from trends in northeast monsoon and southwest monsoon showed a higher magnitude in increasing rainfall too which is greater than the magnitude of change during summer and winter seasons and that the result was more or less uniform across all the rain gauge stations (Palaniswami and Muthiah, 2018). In order to get an idea about rainfall patterns and variability in different regions of Andhra Pradesh and Telangana district, Abhishek *et al.* (2018) applied some nonparametric statistical tests and change point detection methods to find increasing trends in annual rainfall series in both Andhra Pradesh and Telangana states. Nema *et al.* (2018) conducted a homogeneity test for the detection of change points, and a non-parametric trend test as well as Sen's slope estimator for the long-term rainfall record for 27 districts of Chhattisgarh state of India which revealed a decline in rainfall in 1961 as the change point year. The study conducted by Kocsis *et al.* (2019) investigated the rainfall in Keszthely Western Hungary, Europe that examined 150 years of past climate data for the analysis. Pettitt test and Mann-Kendall test were employed to detect change points and the time series trend detection respectively. The study showed that the homogeneity test was very important for analysis to avoid false trend detections. Patakamuri *et al.* (2020) studied the change points in rainfall in Ananthapuramu, Andhra Pradesh which revealed an increase in rainfall in summer for six out of 27 stations. Significantly however decreasing trend was observed in two stations for both the southwest and northeast monsoon. The rain gauge is an instrument for the direct measurement of precipitation and anthropogenic factors such as a change in instruments and rain gauge location affect long-term climatic data which may produce either inhomogeneity or bias in the series (Schroeer *et al.* 2018). These make most climate data unreliable and may show spurious temporal climate variability. "The large variability in time and space requires that, for accurate unbiased results, long-term climate analyses be based on homogeneous data" (Longobardi & Villani, 2010). Homogeneity infers that the series are almost similar and so the heterogeneous conditions are eliminated. Numerous homogeneity tests are available in the literatures. Among others, they include the Pettitt-Whitney-Mann test, autocorrelation test (ACF), von Neumann test, Normal Homogeneity test, and Buishand test, etc., which hypothesize that two or more time series are homogeneous.

The principal aim of this study is to find potential trends, homogeneity, and the change points associated with rainfall time series for 9 stations of Thiruvallur district over the time interval 1971 – 2016. As far as we know, only a limited number of investigations into the analysis of climate trends have been conducted in this particular area. This study employs a unique methodology to examine the uniformity of rainfall patterns, relying on techniques involving identifying structural shifts, utilizing summary statistical measures, and applying the F-test. The findings of this research are expected to expressively benefit efforts directed at drought mitigation, resource management, water management, and socioeconomic development planning of the region.

Methods

Mann–Kendall Test

A trend estimate called the Mann-Kendall is used to demonstrate if data is noticeably declining or increasing over time. No distributional presumptions are made for this test, and the missing data or unevenly spaced measurement of the period is both acceptable. The Mann-Kendall trend test statistic has the following format:

$$S_{mk} = \sum_{i=1}^{n-1} \sum_{j=t+1}^{n} sgn(x_t - x_j)$$

where

$$sign(x_t - x_j) = \begin{cases} 1, & if\ (x_t - x_j) > 0 \\ 0, & if\ (x_t - x_j) = 0 \\ -1, & if\ (x_t - x_j) < 0 \end{cases}$$

This equals the difference between the positive and negative differences. There is neither an upward nor downward trend of the data if S = 0. If S is less than zero, it shows a decreasing trend, indicating the direction has been declining over time. If S is greater than 0, the concentration is growing over the interval, indicating an increasing trend.

Pettitt's test

There are many techniques for identifying change points in a time series, from parametric to nonparametric approaches. The non-parametric Pettitt's test can be used to determine whether rapid changes in the climate have occurred in the past (Gao *et al.*, 2011). Due to the sensitivity of time series breaks, it is highly helpful for change point detection (Dhorde and Zarenistanak, 2013). According to Jaiswal *et al.* (2015), the test can identify a significant shift in a time series' mean even if the specific timing of the change is unknown. According to Pettitt, the method was created in 1979. The non-parametric statistic is defined as;

$$U_t = \sum_{i=1}^{t} \sum_{j=t+1}^{n} sign(x_t - x_j)$$

$$sign(x_t - x_j) = \begin{bmatrix} 1, & if\ (x_t - x_j) > 0 \\ 0, & if\ (x_t - x_j) = 0 \\ -1, & if\ (x_t - x_j) < 0 \end{bmatrix}$$

The sample (n), the statistics K, and the corresponding confidence (ρ) are expressed as given below:

$$K = Max|U_t|$$

$$\rho = exp\left(\frac{-K}{n^2 + n^3}\right)$$

The hypothesis is disproved when it is less than a particular confidence level. The definition of an approximate probability (p) for the change point is as follows:

$$p = 1 - \rho$$

It is clear that a substantial transition point occurs; the series is divided into two sub-series at that specific place. For the purpose of identifying a transition point in a series, the statistic K can also be contrasted with the standard values at various confidence levels.

Descriptive statistics

Mean: For the calculation of the mean the observed results are totalled and divided by the number of observations used for the study. The following symbols are a useful way to represent this well-known procedure:

$$\overline{\boldsymbol{x}} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Where n - total number of rainfall values and x_i - observed rainfall value at time i.

Standard deviation: This is an overall indicator of how far each observation deviates from the mean. The square of values obtained by subtracting from the mean is added. Then the Standard deviation is calculated by dividing the number of samples obtained and the square root gives the Standard Deviation.

$$\sigma = \sqrt{\frac{\sum(x - \overline{\boldsymbol{x}})^2}{n}}$$

Coefficient of variation

The variability between the values can be observed using the statistical measure of coefficient of variation (CV) which is expressed as;

$$CV = \frac{\mu}{\sigma}$$

where σ - standard deviation and μ - mean

F-Test

The F test divides two variances, s_1 and s_2. The variances are always positive as it is a square term; the outcome is always a positive number. The formula for F-test is given by:

$$F = s^2_1 / s^2_2$$

s^2_1 and s^2_2 - sample variances. The higher the evidence for unequal population variances is when it divides more from the ratio value 1. Therefore, the assumption that the variances are equal will always be the null hypothesis; H0: $\sigma_1^2 = \sigma_2^2$

Results and discussion

Rainfall trend analysis

The trends of the monthly rainfall for the whole considered datasets are shown in Figure 1a and b. For the period 1971-2016, visual inspection of the figures could not reveal evidence of significantly increasing or decreasing trends in all locations, excluding the Vallur Anicut and it is clear that all the series seem to fluctuate over time with little or no deviation. The Mann-Kendall trend test statistic depicted in Table 1 gave clear trend results where increasing or positive trends

Table 1: Mann Kendall linear and seasonal trend results

Test	Mann Kendall		Seasonal Mann Kendall	
Statistic	*S*	*P-Value*	*S*	*P-Value*
Thiruvallur	0.78869	0.4303	1.2039	0.2286
Sholingur	2.171	0.02993	2.6252	0.00866
Ponneri	0.56463	0.5723	0.75749	0.4488
Poondi	0.54603	0.585	0.68506	0.4933
Vallur_Anicut	-1.631	0.1029	-2.4105	0.01593
Red_Hills	0.06040	0.9518	0.10988	0.9125
Thiruvottiyur	-1.0096	0.3127	-1.7218	0.0851
Thamaraipakkam	1.2084	0.2269	1.9083	0.05636
Korattur_Dam	0.46297	0.6434	0.69931	0.4844

Table 2: Pettitt's test results for single change-point detection

Parameter	U*	p-value	Probable change point at time k	Break date
Thiruvallur	6184	0.4897	264	5/1993
Sholingur	10958	0.06411	193	6/1987
Ponneri	4884	0.8275	432	5/2007
Poondi	3935	1.131	264	5/1993
Vallur_Anicut	8855	0.1135	355	12/2000
Red_Hills	3743	1.194	487	12/2011
Thiruvottiyur	7461	0.2464	294	11/1995
Thamaraipakkam	5710	0.6026	145	6/1983
Korattur_Dam	5548	0.6444	266	7/1993

are evident in all locations excluding Vallur Anicut and Thiruvottiyur which experienced decreasing or negative rainfall trends in both linear and seasonal constituents.

Pettitt's change point detection test

Pettitt's test evaluates the significance of the detected shifts through a hypothesis test. If the test rejects the null hypothesis, it indicates strong evidence that there is a change-point of data, suggesting a significant difference in the underlying distribution or average value. Pettitt's test allows for detecting the time for which the changes happen. The results in Table 2 show ρ values are greater than 0.05 confidence level; therefore, the null hypothesis that a significant change point exists is accepted. The break dates are depicted in table 2.

Rainfall data homogeneity

The homogeneity in this regard is based on a systematic approach by utilizing the Pettitt test result and descriptive statistics. The series are segmented according to their changing points to allow true detection of homogeneity

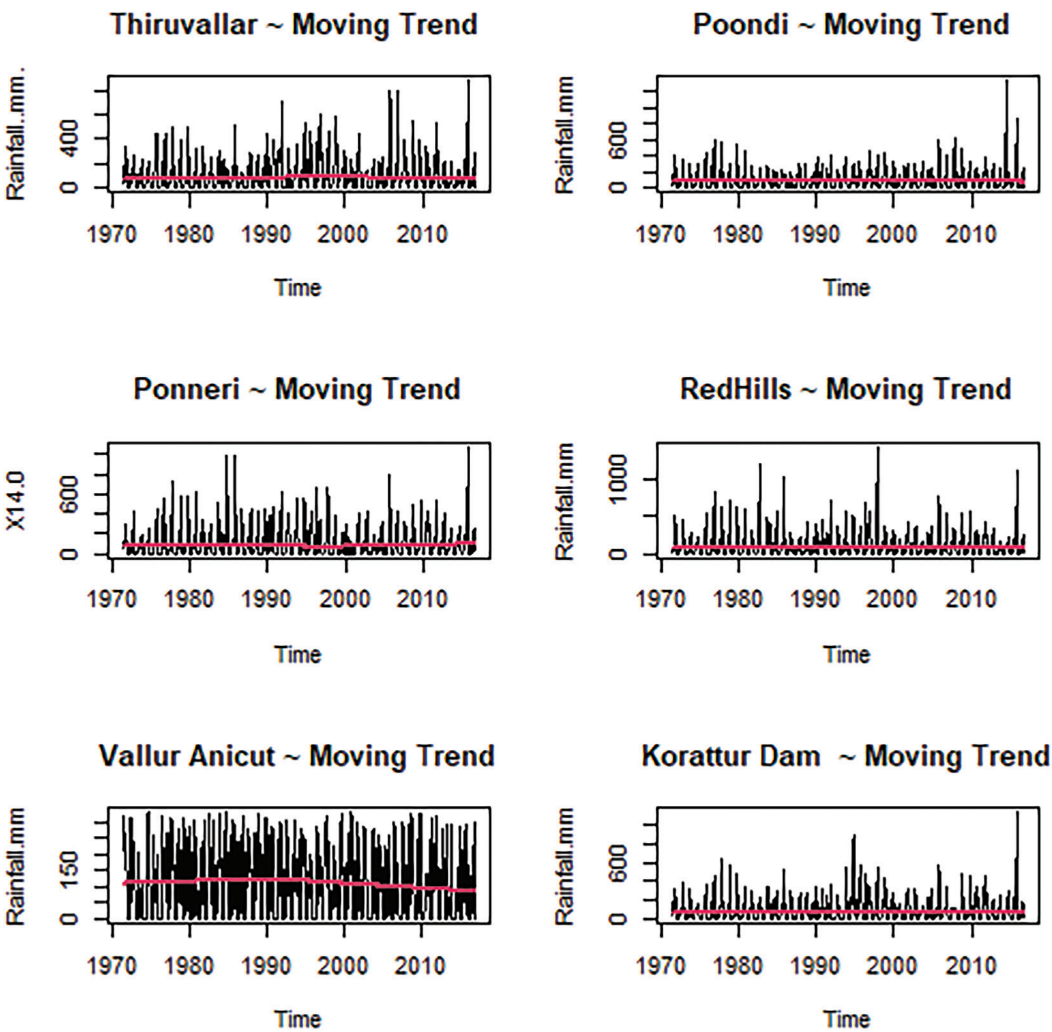

Fig. 1a: Moving trend of monthly rainfall time series

conditions in all the data sets. It can be observed from Table 3 that the descriptive statistics for all the data sets behave similarly with mean values lower than the standard deviation values for all the original series and the series before and after breakpoints. The statistical indicator to show data points in a data points are distributed around mean values. According to Chandramouli (2013), series with a CV less than 1 are regarded as having low variability, whereas those with a CV greater than 1 are regarded as having high variability. According to Table 3, all of the data sets' CV values are greater than 1. These show that the amount of rainfall varies dramatically throughout time. These therefore

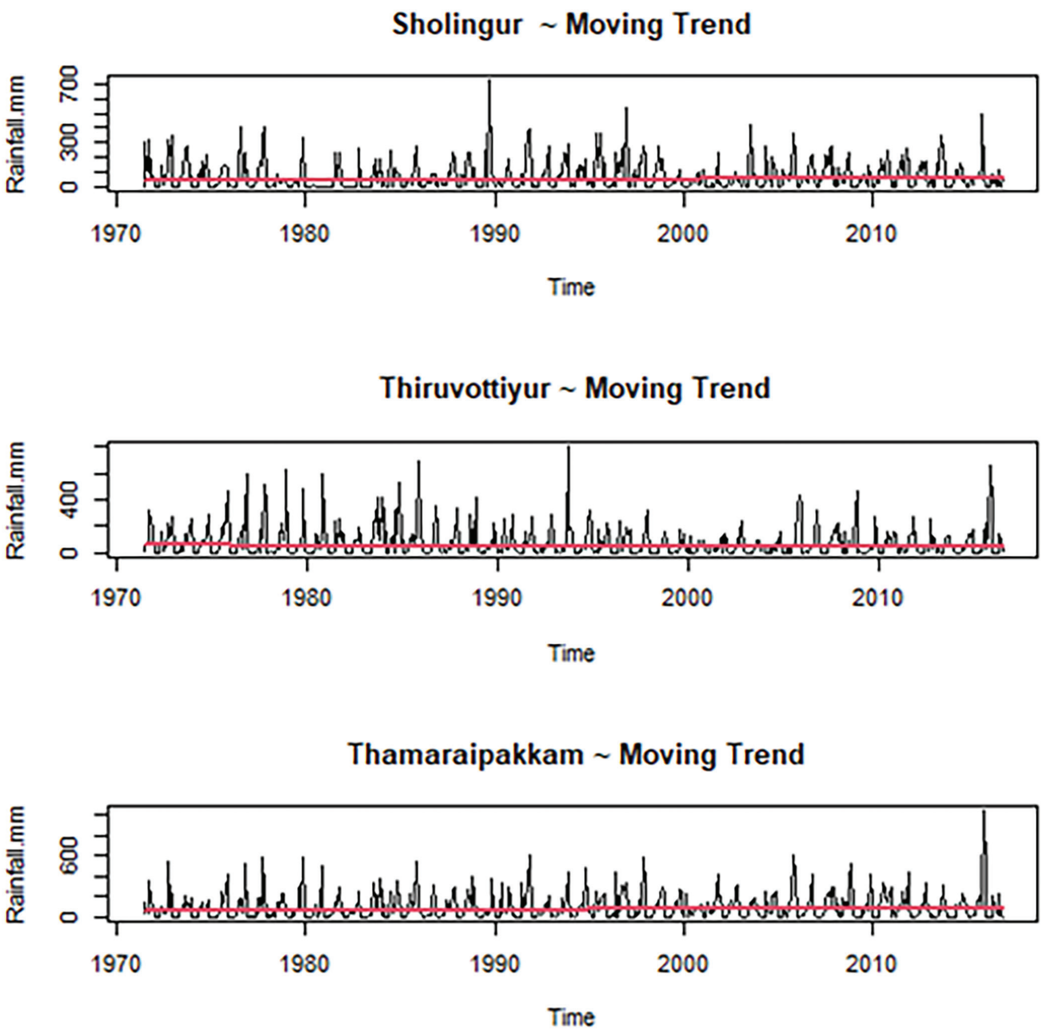

Fig. 1b: Moving trend of monthly rainfall time series

demonstrate that the same statistical patterns are followed by all series.

Conclusion

In the results temporal trends and homogeneity of monthly rainfall series are analysed. The analysis covers the Thiruvallur district of Tamil Nadu state of India. The areas covered are Thiruvallur, Sholingur, Ponneri, Poondi, Vallur Anicut, Red Hills, Thiruvottiyur, Thamaraipakkam, and Korattur Dam. Most of these areas practice rain-fed agriculture; therefore, knowledge of the temporal dynamic of rainfalls in the areas is paramount. The Mann-Kendall trend test applied to the data sets revealed an upward

Table 3: Homogeneity test results

Station	Parameter	Original Series	Series Before Break	Series after Break	F-Statistic	p-value
Thiruvallur	Mean	108.74	95.530	121.11	0.66889	0.0005431
	SD	139.99	121.74	154.34		
	CV	1.29	1.2744	1.2743		
Sholingur	Mean	73.30	61.733	79.611	0.84397	0.09484
	SD	94.92	89.423	97.338		
	CV	1.28	1.4486	1.2227		
Ponneri	Mean	114.63	112.52	121.86	1.13561	0.209808
	SD	160.33	162.35	152.35		
	CV	1.39	1.4429	1.2502		
Poondi	Mean	101.11	92.975	108.73	0.61181	2.99E-05
	SD	133.30	115.59	147.78		
	CV	1.32	1.2432	1.3591		
Vallur_Anicut	Mean	113.64	113.25	90.785	1.4995	0.00097
	SD	110.19	158.13	129.13		
	CV	0.97	1.3963	1.4224		
Red_Hills	Mean	117.61	121.46	85.827	1.17711	0.22374
	SD	174.87	176.10	162.31		
	CV	1.49	1.4498	1.8912		
Thiruvottiyur	Mean	77.85	77.847	106.17	0.62409	0.00042
	SD	115.49	115.49	146.19		
	CV	1.48	1.4835	1.3769		
Thamaraipakkam	Mean	99.55	85.717	103.96	0.94335	0.34940
	SD	128.83	121.78	125.38		
	CV	1.29	1.4207	1.2061		
Korattur_Dam	Mean	102.24	89.042	114.77	0.537	1.99E-07
	SD	138.23	114.69	156.55		
	CV	0.97	1.2881	1.3640		

trend in Thiruvallur, Sholingur, Ponneri, Poondi, Red Hills, Thamaraipakkam, and Korattur Dam locations whereas the results show a decrease in rainfall over the years in Vallur Anicut and Thiruvottiyur. An upward rainfall trend may lead to extreme events which may lead to floods and a downward trend is an alarming signal of drought and shortage of water. Due to the magnitude of trend studies and in order to avoid false classification, the change point detection method was utilized and the probable changes were detected and identified in all the series and splitting the series based on their change-point dates. The series was tested for homogeneity using descriptive statistics and F-test. The results revealed homogeneity in all the series which confirmed the Mann Kendall trend test results. An increasing trend in rainfall can lead to higher water availability. This is beneficial for agriculture as it provides sufficient moisture for crop growth and reduces the need for irrigation. It can also replenish water reservoirs, lakes, and groundwater, improving water supply for human consumption. Higher rainfall can raise the risk of flooding though. If the intensity of rainfall exceeds the capacity of drainage systems or the natural absorption capacity of the soil, excess water can lead to flash floods, river overflows, and urban flooding. This poses threats to infrastructure, property, and human lives. Likewise decreasing rainfall can result in water scarcity, particularly in regions heavily reliant on precipitation for freshwater supplies. This can lead to drought conditions, reduced water availability for agriculture, increased competition for limited water resources, and potential conflicts over water usage. Insufficient rainfall negatively impacts agricultural productivity. Reduced water availability can hinder crop growth, decrease yields, and potentially lead to crop failures. This study identifies the areas to be closely monitored within the study areas for better water management policies.

Acknowledgement

The authors sincerely acknowledge the ICSSR, New Delhi for the financial support to accomplish the research paper.

Competing interest

The authors declare that they have no conflict of interest.

References

Dhakal, N., Jain, S., Gray, A., Dandy, M., & Stancioff, E. (2015). Nonstationarity in seasonality of extreme precipitation: A nonparametric circular statistical approach and its application. *Water Resources Research, 51*(6), 4499-4515.

Javari, M. (2016). Trend and homogeneity analysis of precipitation in Iran. *Climate, 4*(3), 44.

Kocsis, T., Kovács-Székely, I., & Anda, A. (2020). Homogeneity tests and non-parametric analyses of tendencies in precipitation time series in Keszthely, Western Hungary. *Theoretical and Applied Climatology, 139*(3-4), 849-859. https://doi.org/10.1007/s00704-019-03014-4

Longobardi, A., & Villani, P. (2010). Trend analysis of annual and seasonal rainfall time series in the Mediterranean area. *International journal of Climatology, 30*(10), 1538-1546.

Meshram, S. G., Singh, V. P., & Meshram, C. (2017). Long-term trend and variability of precipitation in Chhattisgarh State, India. *Theoretical and Applied Climatology, 129*, 729-744.

Nema, M. K., Khare, D., Adamowski, J., & Chandniha, S. K. (2018). Spatio-temporal analysis of rainfall trends in Chhattisgarh State, Central India over the last 115 years. *Journal of Water and Land Development*, (36).

Palaniswami, S., & Muthiah, K. (2018). Change Point Detection and Trend Analysis of Rainfall and Temperature Series over the Vellar River Basin. *Polish Journal of Environmental Studies, 27*(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*(1), 211. www.mdpi.com/journal/water.

Pranuthi, G., Dubey, S. K., Tripathi, S. K., & Chandniha, S. K. (2014). Trend and change point detection of precipitation in urbanizing Districts of Uttarakhand in India. *Indian Journal of Science and Technology, 7*(10), 1573-1582.

Praveen, B., Talukdar, S., Shahfahad, Mahato, S., Mondal, J., Sharma, P., Islam, T. A. R. M. & Rahman, A. (2020). Analyzing trend and forecasting of rainfall changes in India using non-parametrical and machine learning approaches. *Scientific reports, 10*(1), 10342. https://doi.org/10.1038/s41598-020-67228-7

Schroeer, K., Kirchengast, G., & Sungmin, O. (2018). Strong dependence of extreme convective precipitation intensities on gauge network density. *Geophysical Research Letters, 45*(16), 8253-8263. https://doi.org/10.1029/ 2018GL077994.

Waghaye, A. M., Rajwade, Y. A., Randhe, R. D. & Nidhi Kumari (2018). Trend analysis and change point detection of rainfall of Andhra Pradesh and Telangana, India. *Journal of Agrometeorology*, (20) 2, Short Comminucation, 160-163.

Venkatesan Madha Suresh
ICCSR – Senior Fellow, Centre for Environmental Sciences,
School of Earth and Atmospheric Sciences,
University of Madras, Chennai, 600025, India

Ibrahim Lawal Kane, Nigeria*
Department of Mathematics and Statistics,
Umaru Musa Yar'adua University,
2218, Katsina – Nigeria.

*Author for Correspondence
E-mail: Ibrahim.lawal@umyu.edu.ng