MAUSAM, **75**, 3 (July 2024), 895-904

MAUSAM

Homepage: https://mausamjournal.imd.gov.in/index.php/MAUSAM

L E T T E R S

UDC No. 551.577.3 (540)
DOI : https://doi.org/10.54302/mausam.v75i3.6300

THE IMPACT OF DATA UNCERTAINTY ON IDENTIFYING PRECIPITATION TRENDS IN INDIA

1. The hydrological cycle is significantly affected by precipitation, which is a crucial component that displays spatial and temporal variations due to both climate change and human activities (IPCC, 2007). The analysis of precipitation's temporal and spatial variability is crucial for ecosystem functioning, agricultural productivity, and water resource management (Brunsell *et al*., 2010). The average temperature of the Earth has increased by 0.6 °C in the 20th century, according to the IPCC, 2007. Climate prediction models anticipate a significant temperature shift ranging from 1.4 to 5.4°C. Kumar & Jain, 2010 have stated that changes in rainfall distribution could affect the distribution of runoff, soil moisture, and groundwater reserves, while increasing the frequency of droughts and floods. Therefore, investigating the variations in rainfall patterns is crucial for water resource planning and management and designing appropriate strategies to deal with floods and droughts.

In the last two decades, trend analysis has been widely used as a popular approach to examine hydro-meteorological variables. Several techniques have been developed and applied to analyze trends in these variables, each with its own limitations and advantages. Two categories of trend analysis methods include parametric and non-parametric techniques. In recent years, non-parametric methods have become more popular among researchers for trend analysis due to their insensitivity to outliers and applicability to non-normally distributed series with missing values (Ahmadi *et al*., 2018). The Mann-Kendall (MK) test is a well-known non-parametric method proposed by Mann, 1945 and Kendall, 1975, and it has been recommended by the World Meteorological Organization (WMO) for analyzing hydro-meteorological time series trends (Kumar *et al*., 2009).

In India, the India Meteorological Department (IMD) uses a gridded rainfall product developed by Pai *et al*., 2014 that relies on rain gauges to monitor precipitation levels. Rain gauges are commonly used to estimate area-averaged precipitation over land surfaces by leveraging a network of uniformly distributed rain gauge stations. However, the optimal density and distribution of these stations have been the focus of research, as inadequate gauge density or poor station placement can lead to significant errors in estimating the areal rainfall of a region (Mishra *et al*., 2011). Moreover, studies have shown that the distance between rain gauge stations and the grid point can also affect the accuracy of rainfall estimates (Roy Bhowmik & Das, 2007).

The APHRODITE (Asian Precipitation-Highly-Resolved Observational Data Integration Toward Evaluation of Water Resources) project was launched with the objective of developing high-resolution, reliable and rain gauge-based precipitation products over the entire Asian region (Yatagai *et al*., 2005, 2009). These products were initially created to validate high-resolution climate model simulations and to make localized precipitation forecasts in response to future climate change caused by the anthropogenic greenhouse effect (Yatagai *et al*., 2005). The APHRODITE dataset is particularly useful for addressing the challenges of bias correction and downscaling of coarse-resolution climate simulation outputs, particularly in mountainous areas, because it uses an interpolation method that accounts for orographic effects and incorporates substantial rain gauge data. As a result, the APHRODITE dataset has become a valuable resource for researchers and has been widely used in many studies for trend estimation, dry/wet spells, and extreme event prediction (Kim *et al*., 2019; Kishore *et al*., 2016; Li *et al*., 2018; Malik *et al*., 2016; Soraisam *et al*., 2018).

The analysis of rainfall trends in the West Coast Plains of India was conducted by Saini *et al*., 2020, using the Rainfall gridded dataset developed by Pai *et al*., 2014. Various statistical techniques such as Modified Mann-Kendall's test, Innovative Trend Analysis (ITA), Linear Regression, Weibull's Recurrence Interval, Sen's Slope test, Pearson's Coefficient of Skewness, Consecutive Disparity Index, Kurtosis, *etc* were employed to examine the trends on monthly, seasonal, and decadal timescales. The results indicated significant decreasing rainfall trends in January and July months, whereas August and September months showed positive trends.

Singh, 2021, utilized ITA to examine the trends in seasonal and annual rainfall. The study revealed increasing trends in most sub-divisions of peninsular and

northwest India. However, winter rainfall displayed decreasing trends in most sub-divisions of the country. Rai & Dimri, 2020, studied rainfall seasonality using an individual seasonality index and linear regression for the period 1971-2015. Their findings indicated declining trends in the seasonality index, suggesting shorter dry periods resulting in more consistent rainfall throughout the year in some parts of India during the recent period. Praveen *et al*., 2020, analyzed the long-term Spatio-temporal changes in rainfall across India from 1901 to 2015 at the meteorological divisional level. The results indicated that most of the meteorological divisions exhibited significant negative trends in annual and seasonal scales.

These studies highlight the complex and varied changes in rainfall patterns across different regions of India. Although various investigations have conducted trend analyses, none of them have endeavoured to evaluate the impact of uncertainties, including rain gauge density and missing data, on trend analysis in India. APHRODITE products entail extensive rain gauge data and information on rain gauge density for each grid. Accordingly, the current study endeavours to examine the influence of uncertainty on trend estimation using this data.

2. *Study area and data used* - India, situated between approximately 8° N and 37° N latitude, is characterized by diverse climatic conditions due to its geographical location. The southern part of the country has tropical weather, while the northern region is dominated by the Himalayan range. The India Meteorological Department (IMD) classifies a calendar year into four seasons: winter (January and February) characterized by low temperatures, low humidity, and clear skies; pre-monsoon (March to May) with high temperatures and hot, dry winds; monsoon (June to September) with heavy rainfall across the country; and post-monsoon (October to December), which serves as a transition period from the monsoon to winter. This seasonal classification system is useful in understanding the climatic patterns in India, which are critical for various sectors such as agriculture, water management and disaster management.

The APHRODITE dataset provides daily precipitation data at a horizontal resolution of 0.25°, covering the Monsoon Asia region. The data is primarily sourced from a rain-gauge-observation network and is available for the period 1951 to 2015 (V1101: 1951- 2007, V1901: 2008-2015). The interpolation algorithm used to generate the dataset takes into account the local topography between the rain gauge and the interpolated point and the "number of observations" information is

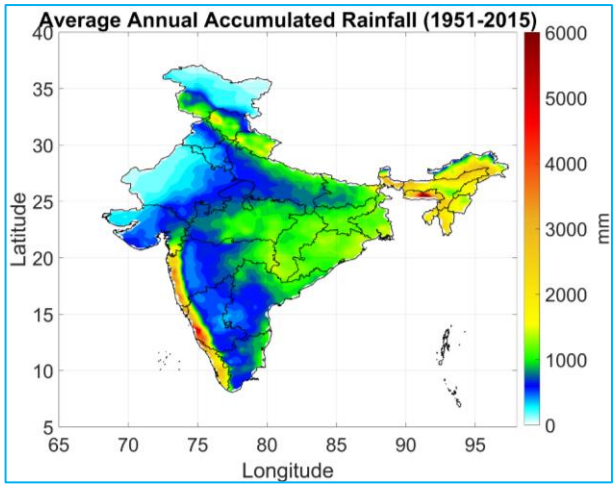

Fig. 1. Average Annual Rainfall (1951-2015)

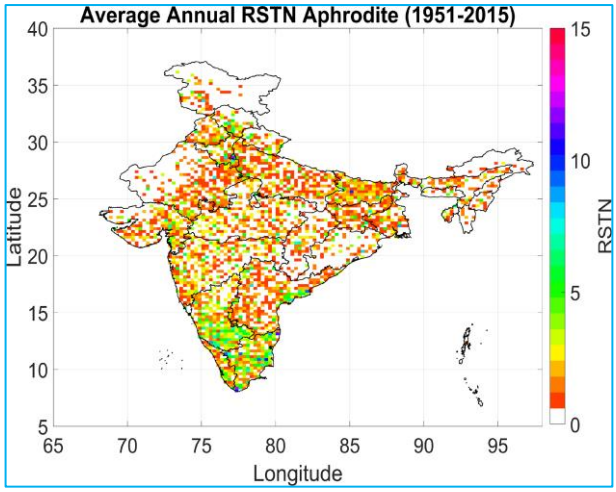

Fig. 2. Average RSTN for the period (1951-2015)

provided for each day on each grid. The count "ratio of station grids" (RSTN) is used to indicate the number of grids with at least one rain-gauge observation. The RSTN information is provided with the product on a daily basis and helps to determine whether each grid-box value is an interpolation or reflects an observed value within the grid box or nearby. When the 0.25 degree product is created, each grid box has 25 grids of 0.05 degree. For which, RSTN=4(%) means there is one grid with a station and RSTN=8(%) means there are two grids with a station.

Fig. 1 shows the average annual accumulated rainfall over India for the period (1951-2015) and Fig. 2 shows the average RSTN values for each grid. To prepare Fig. 1, for each year and each grid, accumulated annual rainfall was calculated and then for each grid, the accumulated annual rainfall was averaged temporally. However, for Fig. 2,

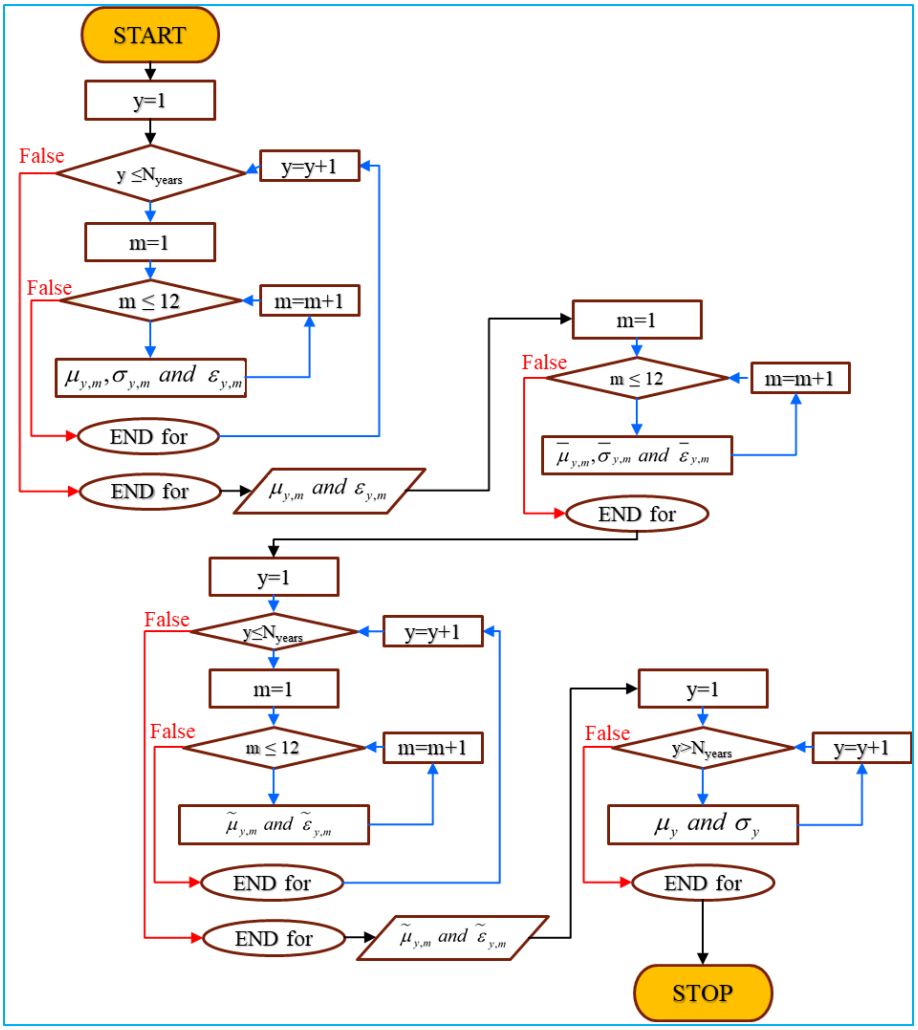

Fig. 3. Flow chart of methodology adopted for each grid

temporal average of RSTN values was taken for each grid. It can be seen that a large part of India is having average rainfall of less than 1000 mm whereas eastern regions received average rainfall of up to 1500 mm. There are some locations such as the western Ghats and East India, where the rainfall reaches up to 6000 mm in a year.

Apart from the high rainfall variability and missing observations (RSTN), rain gauge observations also have measurement errors. The major source of error is the systematic measuring error which results from evaporation out of the gauge and aerodynamic effects when droplets are drifted by the wind across the gauge funnel. Since the data set does not specify the magnitude of measurement errors, a 2% error in daily precipitation measurement was assumed in the present study.

3. *Mathematical formulation and Methodology* - In this study, the weighted linear regression (WLR) proposed by Soni *et al*., 2019, was used to enhance the accuracy of trend estimates for gridded rainfall products. The traditional approach of linear regression (LR)-based trend analysis assumes that every data point contributes equally precise information to the trend. However, for gridded rainfall products, the variability of information content changes each year due to missing values (rain gauge density, RSTN). As a result, this assumption made by LR does not hold for such products. Weighted linear regression addresses this issue by assigning weights to each data point.

To use the WLR method, weights must be assigned to each data point. However, assigning weights to each data point can be challenging when dealing with data that have measurement errors, large variability, strong seasonality and missing values, whether at random or not (Taylor, 1997). To address this issue, this study assumes that daily (*d*) precipitation data in a given month (*m*) and

year (y) are independent samples from a normal distribution with a mean, $\mu_{y,m}$ and standard deviation, $\sigma_{y,m}$, taking into account measurement errors that are also normally distributed with a known standard deviation, $\sigma_{y,m,d}$. This allows for appropriate weighting of data about measurement errors and missing values when estimating monthly and seasonal/annual rainfall anomalies. The process for determining these weighted anomalies and their standard errors is described in section 3.1, while section 3.2 outlines the WLR method for trend analysis using these weighted anomalies and their standard errors, as well as the LR method. The flow chart of the methodology adopted for each grid is shown in Fig. 3.

3.1. *Anomalies and their standard errors* - To eliminate the impact of seasonality, anomalies were utilized instead of actual values in the trend analysis.

3.1.1. *Monthly scale - Monthly Mean* - The initial stage involved the computation of monthly mean ($\mu_{y,m}$) and standard deviation ($\sigma_{y,m}$) using daily rainfall values ($x_{y,m,d}$) and their associated additive measurement errors. The measurement errors were assumed to conform to a normal distribution with zero mean and standard deviation, $\sigma_{y,m,d}$. As the daily observations were considered independent, the log-likelihood of the data (L) can be expressed using Equation *1*, wherein the unknown parameters are$\mu_{y,m}$ and $\sigma_{y,m}$.

$$L\left(\mu_{y,m},\sigma_{y,m}\right)=\sum_{\mathrm{d=1}}^{N_{\mathrm{days}}}\log\left[\frac{1}{\sqrt{\sigma_{y,m}^2+\sigma_{y,m,d}^2}}e^{\frac{-\left(x_{y,m,d}-\mu_{y,m}\right)^2}{2\left(\sigma_{y,m}^2+\sigma_{y,m,d}^2\right)}}\right] \tag{1}$$

To maximize the log-likelihood function, Nelder-Mead simplex direct search method was utilized to determine the optimal values for $\mu_{y,m}$ and $\sigma_{y,m}$. The summation was conducted over the number of daily measurements available in the month (N_{days}), with any missing values excluded. The standard error of the monthly mean ($\varepsilon_{y,m}$) was derived using

$$\varepsilon_{y,m}=\sqrt{\sum_{\mathrm{d=1}}^{N_{\mathrm{days}}}\left(\omega_{y,m,d}\right)^2\left(\sigma_{y,m}^2+\sigma_{y,m,d}^2\right)} \tag{2}$$

where

$$w_{y,m,d}=\frac{\frac{1}{\sigma_{y,m}^2+\sigma_{y,m,d}^2}}{\sum_{\mathrm{d=1}}^{N_{\mathrm{days}}}\left(\frac{1}{\sigma_{y,m}^2+\sigma_{y,m,d}^2}\right)} \tag{3}$$

Long-term monthly mean

The process of calculating the long-term monthly mean $\left(\bar{\mu}_m\right)$ and standard deviation $\left(\bar{\sigma}_m\right)$ for each month is analogous to that used for computing the monthly mean ($\mu_{y,m}$) and standard deviation ($\varepsilon_{y,m}$). In this case, the log-likelihood function was optimized using $\mu_{y,m}$ and $\varepsilon_{y,m}$ and as inputs.

$$L\left(\bar{\mu}_{y,m},\bar{\sigma}_{y,m}\right)=\sum_{\mathrm{y=1}}^{N_{\mathrm{years}}}\log\left[\frac{1}{\sqrt{\bar{\sigma}_m^2+\varepsilon_{y,m}^2}}e^{\frac{-\left(\mu_{y,m}-\bar{\mu}_m\right)^2}{2\left(\bar{\sigma}_m^2+\varepsilon_{y,m}^2\right)}}\right] \tag{4}$$

The calculation involves adding up all the years (N_{years}) with available monthly means, in order to obtain the summation. Equation 5 was then used to determine the standard error of the long-term monthly mean $\left(\bar{\varepsilon}_m\right)$.

$$\bar{\varepsilon}_{y,m}=\sqrt{\sum_{\mathrm{y=1}}^{N_{\mathrm{years}}}\left(\omega_{y,m}\right)^2\left(\bar{\sigma}_m^2+\varepsilon_{y,m}^2\right)} \tag{5}$$

where

$$w_{y,m}=\frac{\frac{1}{\bar{\sigma}_m^2+\varepsilon_{y,m}^2}}{\sum_{\mathrm{y=1}}^{N_{\mathrm{years}}}\left(\frac{1}{\bar{\sigma}_m^2+\varepsilon_{y,m}^2}\right)} \tag{6}$$

Monthly anomalies

The monthly anomalies $\left(\tilde{\mu}_{\mathrm{y,m}}\right)$ and their standard errors $\left(\tilde{\varepsilon}_{\mathrm{y,m}}\right)$ were estimated from $\bar{\mu}_{\mathrm{y,m}}$ and $\bar{\varepsilon}_{\mathrm{y,m}}$ using

$$\tilde{\mu}_{\mathrm{y,m}}=\mu_{y,m}-\bar{\mu}_{y,m} \tag{7}$$

$$\varepsilon_{y,m} = \sqrt{\varepsilon_{y,m}^2 + \bar{\varepsilon}_{y,m}^2} \quad (8)$$

respectively.

3.1.2. Seasonal (or annual) scale–From $\tilde{\mu}_{y,m}$ and $\tilde{\varepsilon}_{y,m}$ the seasonal (or annual) anomalies (μ_y) and their standard deviation (σ_y) were calculated for each year by maximizing log likelihood function given in

$$L(\mu_y, \sigma_y) = \sum_{m=1}^{N_{months}} \log\left[\frac{1}{\sqrt{(\sigma_y)^2 + (\tilde{\varepsilon}_{y,m})^2}} e^{\frac{-(\mu_y - \tilde{\mu}_{y,m})^2}{2\left[(\sigma_y)^2 + (\tilde{\varepsilon}_{y,m})^2\right]}}\right] \quad (9)$$

Here the summation is performed over all the months in the season (or year) for which monthly anomalies were present. The standard error (ε_y) of the seasonal (or annual) anomalies were calculated from

$$\varepsilon_y = \sqrt{\sum_{m=1}^{N_{months}} (\omega_m)^2 \left[(\sigma_y)^2 + (\tilde{\varepsilon}_{y,m})^2\right]} \quad (10)$$

where

$$w_m = \frac{\frac{1}{(\sigma_y)^2 + (\tilde{\varepsilon}_{y,m})^2}}{\sum_{m=1}^{N_{months}} \left(\frac{1}{(\sigma_y)^2 + (\tilde{\varepsilon}_{y,m})^2}\right)} \quad (11)$$

3.2. *Trend analysis*-The trends in Rainfall data were calculated by two different methods as described below.

3.2.1. *Linear regression (LR)* - The initial approach involved calculating annual and seasonal rainfall anomalies without factoring in data uncertainty, leading to the creation of simple anomalies (SA). Linear regression (LR) was then applied to determine trends in the annual and seasonal SSR anomalies.

3.2.2. *Weighted linear regression (WLR)* - The second approach involved using daily rainfall data to calculate the monthly means ($\mu_{y,m}$) and their standard error ($\varepsilon_{y,m}$), while taking into account the variability of the data, measurement uncertainties (assumed 2%), and missing values. On all the grids, RSTN=0 is replaced with RSTN=1, to avoid the zero division errors. This was done using:

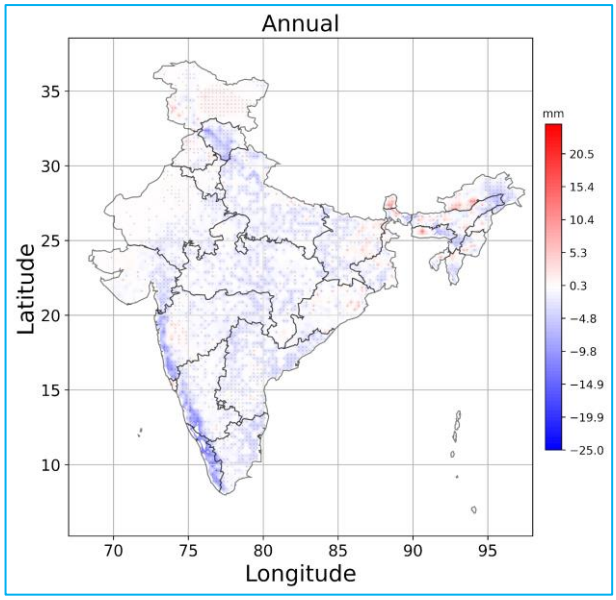

Fig. 4. Trends in annual precipitation using LR method for the period 1951-2015. The stippling's '.' and 'o' in the figure represent significance of trends at 95% and 99% CL, respectively

$$L(\mu_{y,m}, \sigma_{y,m}) = \sum_{d=1}^{N_{days}} \log\left[\frac{1}{\sqrt{\sigma_{y,m}^2 + \sigma_{y,m,d}^2}} e^{\frac{-(x_{y,m,d} - \mu_{y,m})^2}{2(\sigma_{y,m}^2 + \sigma_{y,m,d}^2)}}\right] \quad (12)$$

$$w_m = \frac{\frac{1}{(\sigma_y)^2 + (\tilde{\varepsilon}_{y,m})^2}}{\sum_{m=1}^{N_{months}} \left(\frac{1}{(\sigma_y)^2 + (\tilde{\varepsilon}_{y,m})^2}\right)} \quad (13)$$

The standard error of the monthly mean depended on the number of stations in the given month. The uncertainty due to missing values was also factored in, resulting in lower weights for the corresponding years in the *WLR* and propagating to the standard error of the annual and seasonal anomalies. The resulting anomalies were referred to as weighted anomalies (*WA*).

The trends in the annual and seasonal anomalies were then identified using the WLR, where weights were inversely proportional to the square of the standard error in the anomalies (refer to Appendix A). The general equation for weighted regression is provided in $\mu_y = At_y + B$. Equation 14, where μ_y represents the anomalies and t_y represents the year for the index *y*. The equations for

TABLE 1

Trend statistics over the entire India estimated using LR method for the period 1951-2015

Season	Positive Trends				Negative Trends			
	Total	95%	99%	NS	Total	95%	99%	NS
Winter	1385	117	48	1268	3254	1218	743	2036
Pre-monsoon	1813	436	239	1377	2826	1193	764	1633
Monsoon	1666	586	437	1080	2973	1603	1306	1370
Post monsoon	1217	105	61	1112	3422	1568	994	1854
Annual	1489	467	346	1022	3150	1805	1527	1345

calculating the slope (A) and the intercept (B), along with their uncertainties, are presented in Appendix A.

$$\mu_y = At_y + B \quad (14)$$

4. *Results and discussion* – 4.1. *Trends of Rainfall over India using Linear Regression (1951-2015)* - The spatial distribution of trends estimated using the LR method for the entire period of 1951-2015 is depicted in Fig. 4, where the stippling's '.' and 'o' indicate significant trends at 95% and 99% Confidence Level (CL), respectively. The analysis reveals that a large part of India, particularly over the western Ghats, experienced significant decreasing trends in precipitation, reaching up to -25 mm/year. However, some regions such as Jammu & Kashmir and parts of Northeast India, exhibited significant positive trends as well.

Further insights into the trends are presented in Table 1, which outlines the statistics of the trends observed during various seasons. The analysis shows that out of the total 4642 grids, 1489 grids exhibited positive trends, while 3150 grids showed negative trends on an annual scale. Notably, approximately 31% of positive trends and 57% of negative trends were significant at 95% CL. During the winter and premonsoon seasons, the number of significant negative trends dropped to around 37% and 42%, respectively. Similarly, the proportion of significant positive trends during these seasons was relatively low at approximately 8% and 24%, respectively.

4.2. *Recent trends of Rainfall over India using Weighted Linear Regression* - The spatial distribution of trends estimated using the WLR method for the entire period 1951-2015 is illustrated in Fig. 5. In comparison to the trends estimated using the LR method, we observe significant positive trends in many parts of India, including Rajasthan, Gujarat, Jharkhand, Orissa, West Bengal, Tamil Nadu, and Andhra Pradesh.

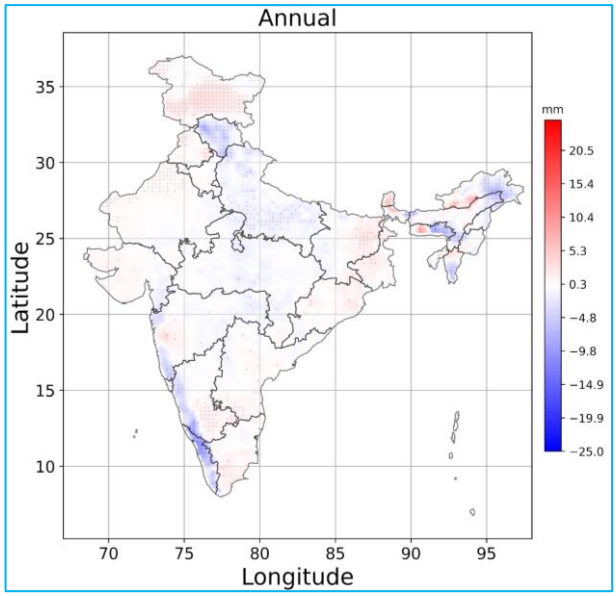

Fig. 5. Trends in annual precipitation using WLR method for the period 1951-2015. The stippling's '.' and 'o' in the figure represent significance of trends at 95% and 99% CL, respectively

TABLE 2

Trend statistics over the entire India estimated using WLR method for the period 1951-2015

Season	Positive Trends				Negative Trends			
	Total	95%	99%	NS	Total	95%	99%	NS
Winter	1727	148	51	1579	2912	664	238	2248
Pre-monsoon	3186	700	337	2486	1453	147	65	1306
Monsoon	2327	630	413	1697	2312	581	341	1731
Post monsoon	1364	59	21	1305	3275	592	275	2683
Annual	2274	673	399	1601	2365	613	353	1752

TABLE 3

Effect of incorporating uncertainty in Trend analysis for the period 1951-2015

LR	WLR	Positive				Negative			
		TOTAL	95%	99%	NS	TOTAL	95%	99%	NS
Positive	TOTAL	**86.23%**	36.20%	22.10%	50.03%	13.77%	0.54%	0.07%	13.23%
	95%	95.50%	**68.74%**	49.68%	26.77%	4.50%	0.43%	0.00%	4.07%
	99%	94.80%	71.39%	**53.76%**	23.41%	5.20%	0.29%	0.00%	4.91%
	NS	82.00%	21.33%	9.49%	**60.67%**	18.00%	0.59%	0.10%	17.42%
Negative	TOTAL	31.43%	4.25%	2.22%	0.00%	**68.57%**	19.21%	11.17%	0.00%
	95%	25.21%	2.99%	1.44%	22.22%	74.79%	**28.48%**	17.95%	46.32%
	99%	21.81%	2.42%	1.11%	19.38%	78.19%	30.71%	**19.97%**	47.48%
	NS	39.78%	5.95%	3.27%	33.83%	60.22%	6.77%	2.08%	**53.46%**

5. Table 2 presents the statistics of these trends during different seasons. On an annual scale, 2274 grids exhibited positive trends, while 2365 grids showed negative trends. However, only about 25-30% of the grids showed significant trends at an annual timescale.

5.1. Effect of adding uncertainty in Trend analysis - To estimate the impact of data uncertainty on trend analysis, we conducted a grid-level analysis, which revealed interesting findings. As presented in Table 3, when using the WLR method instead of the LR method, about 86% of positive trends and 68% of negative trends remained unchanged. However, we observed that 23-26% of significant positive trends and 46% of significant negative trends were converted to non-significant trends. Additionally, approximately 0.5% of positive and 2-3% of negative significant trends reversed to significant negative and positive trends, respectively. We also noted that about 4-6% of non-significant negative trends converted to significant positive trends and approximately 30% of significant negative trends changed their significance level from 99% to 95% CL. As per the Table 1, total number of grids with positive trends are lesser than the grids with negative trends in all the seasons. However, a reversal is seen in the Table 2 during pre-monsoon and monsoon season. The major reason for the same is the higher rainfall mean and standard deviation during these two seasons compared to other seasons. In the WLR method, standard error and thus the weight for the regression is proportional to both RSTN and rainfall variability. Hence higher rainfall values as well as high variability during these seasons result in larger differences between trend estimated using LR and WLR methods. These results emphasize the importance of considering data uncertainty in trend analysis and highlight the need for more robust methods to account for such uncertainties.

6. *Summary and Conclusions* - In this study, we investigated the impacts of uncertainties in gridded rainfall data arising from rain gauge density and seasonal variations. The daily precipitation data with a horizontal resolution of 0.25° were obtained from the APHRODITE (Asian Precipitation - Highly-Resolved Observational Data Integration Towards Evaluation) dataset developed by the Research Institute for Humanity and Nature (RIHN) and the Meteorological Research Institute of Japan Meteorological Agency (MRI/JMA). These datasets were primarily derived from the rain gauge observations network, such as IMD for India. Additionally, the datasets provide information on the number of observations for each day on each grid, represented as the "ratio of station grids (RSTN)." This approach enabled us to account for the effects of the spatial and temporal variability of rain gauge density, which are crucial for improving the accuracy of precipitation estimates.

Through the application of simple linear regression, the study estimates the trends over the entire India, revealing that a considerable portion of the country has experienced decreasing trends in precipitation, reaching up to -25 mm/year, with only small patches of positive trends, particularly over the western Ghats. The analysis indicates that out of the total grids, 1489 (31% significant at 95% CL) exhibited positive trends on an annual scale, while 3150 (57% significant at 95% CL) showed negative trends.

Interestingly, the comparison of the LR method with the WLR method reveals a significant difference in the trend estimates for certain regions. While the LR method

estimates negative trends over large parts of India, including Western (Rajasthan & Gujrat), Eastern (States of Jharkhand, Orissa & West-Bengal) and south India (States of Tamil Nadu and Andhra Pradesh), the WLR method gives significant positive trends. Specifically, a total of 2274 grids showed positive trends, whereas 2365 were negative, with approximately 25-30% of the grids exhibiting significant trends at an annual timescale.

In general, the application of the WLR method resulted in comparable positive and negative trends as the LR method, with approximately 86% positive and 68% negative trends remaining unchanged. However, a significant proportion of the trends changed when using the WLR method. Specifically, approximately 23-26% of significant positive trends and 46% of significant negative trends observed using the LR method were converted to non-significant trends with the WLR method. Additionally, around 0.5% of positive significant trends and 2-3% of negative significant trends were reversed, resulting in significant negative and positive trends, respectively. Notably, the WLR method also converted about 4-6% of non-significant negative trends to significant positive trends. These results underscore the importance of carefully selecting appropriate statistical methods for detecting precipitation trends to avoid erroneous conclusions.

The conclusions of the study can be enumerated as:

(*i*) The LR method revealed that at an annual scale, about 31% of grids (1489) showed significant positive trends at a 95% confidence level (CL), while 57% (3150) showed significant negative trends at a 95% CL. However, the WLR method showed that a total of 25-30% of grids had significant trends at an annual timescale, with 2274 showing positive trends and 2365 showing negative trends.

(*ii*) Precipitation trends over India ranged from about -25 to +25 mm/year at an annual timescale.

(*iii*) The WLR method preserved 86% of positive and 68% of negative trends obtained by the LR method.

(*iv*) The incorporation of uncertainty into the WLR method led to the conversion of 23-26% of significant positive trends and about 46% of significant negative trends obtained using the LR method into non-significant trends.

(*v*) Some significant trends were reversed, with 0.5% of positive trends becoming significant negative trends and 2-3% of negative trends becoming significant positive trends. Additionally, 4-6% of non-significant negative trends were converted into significant positive trends.

This study emphasizes the significance of incorporating missing records and data variability over time for precise trend analysis. Neglecting these uncertainties in trend estimation could lead to erroneous trends, emphasizing the need to consider them in trend analysis. The disparate trends observed across the country underline the necessity of utilizing multiple methods to estimate precipitation trends to enhance comprehension of the spatial and temporal variations in India's precipitation patterns. Additionally, persistent monitoring and analysis of precipitation trends can assist in informed decision-making for water resource management and adaptation strategies amidst the changing climate conditions.

Conflict of Interests : The authors declare no conflict of interest.

References

Ahmadi, F., Nazeri Tahroudi, M., Mirabbasi, R., Khalili, K. and Jhajharia, D., 2018, "Spatiotemporal trend and abrupt change analysis of temperature in Iran", *Meteorological Applications*, **25**, 2, 314-321.

Brunsell, N. A., Jones, A. R., Jackson, T.and Feddema, J., 2010, "Seasonal trends in air temperature and precipitation in IPCC AR4 GCM output for Kansas, USA: evaluation and implications", *International Journal of Climatology*, **30**, 8, 1178-1193.

IPCC, 2007, Climate change 2007: The physical science basis. Geneva.

Kendall, M., 1975, Rank Correlation Measures; Charles Griffin Book Series. Oxford University Press, London, UK.

Kim, I.-W., Oh, J., Woo, S. and Kripalani, R., 2019, "Evaluation of precipitation extremes over the Asian domain: Observation and modelling studies", *Climate Dynamics*, **52**, 1317-1342.

Kishore, P., Jyothi, S., Basha, G., Rao, S., Rajeevan, M., Velicogna, I. and Sutterley, T. C., 2016, "Precipitation climatology over India: Validation with observations and reanalysis datasets and spatial trends", *Climate Dynamics*, **46**, 541-556.

Kumar, S., Merwade, V., Kam, J.and Thurner, K., 2009, "Streamflow trends in Indiana: Effects of long term persistence, precipitation and subsurface drains", *Journal of Hydrology*, **374**, 1-2, 171-183.

Kumar, V. and Jain, S. K., 2010,"Rainfall trend in Ganga-Brahmputra-Meghna river basins of India (1951-2004)", *Hydrol. J.*, **33** (Special Issue), 59-66.

Li, H., Haugen, J. E. and Xu, C.-Y., 2018,"Precipitation pattern in the Western Himalayas revealed by four datasets", *Hydrology and Earth System Sciences*, **22**, 10, 5097-5110.

Malik, N., Bookhagen, B. and Mucha, P. J., 2016, "Spatiotemporal patterns and trends of Indian monsoonal rainfall extremes", *Geophysical Research Letters*, **43**, 4, 1710-1717.

Mann, H. B., 1945, "Nonparametric tests against trend", *Econometrica: Journal of the Econometric Society*, 245-259.

Mishra, A. K., Gairola, R., Varma, A. and Agarwal, V. K., 2011, "Improved rainfall estimation over the Indian region using

satellite infrared technique", *Advances in Space Research*, **48**, 1, 49-55.

Pai, D. S., Sridhar, L., Rajeevan, M., Sreejith, O. P., Satbhai, N. S., Mukhopadhyay, B. and Mukhopadyay, B., 2014,"Development of a new high spatial resolution (0. 25° × 0. 25°) Long Period (1901-2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region", *MAUSAM*, **65**, 1, 1-18.

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

Rai, P. and Dimri, A., 2020, "Changes in rainfall seasonality pattern over India", *Meteorological Applications*, **27**, 1, e1823.

Roy Bhowmik, S.and Das, A. K., 2007, "Rainfall analysis for Indian monsoon region using the merged rain gauge observations and satellite estimates: Evaluation of monsoon rainfall features", *Journal of Earth System Science*, **116**, 187-198.

Saini, A., Sahu, N., Kumar, P., Nayak, S., Duan, W., Avtar, R. and Behera, S., 2020, "Advanced rainfall trend analysis of 117 years over West Coast Plain and Hill Agro-climatic region of India", *Atmosphere*, **11**, 11, 1225.

Singh, R. M., 2021, Discussion of "Multi objective Optimization of Sensor Placement for Precipitation Station Monitoring Network Design" by Ke Wang, Jia Yang, Yuling Peng, Qianqian Wu, and Chuli Hu. *Journal of Hydrologic Engineering*, **26**, 7, 07021005.

Soni, P., Srivastava, R.and Tripathi, S., 2019, "Implication of data uncertainty in the detection of surface radiation trends and observational evidence of renewed solar dimming over India", *Theoretical and Applied Climatology*, **137**, 3, 2663-2680.

Soraisam, B., Karumuri, A. and Pai, D., 2018,"Uncertainties in observations and climate projections for the North East India", *Global and Planetary Change*, **160**, 96-108.

Taylor, J. R., 1997, "An introduction to error analysis", p327. Univ. Sci. Books, Sausalito, Calif.

Yatagai, A., Arakawa, O., Kamiguchi, K., Kawamoto, H., Nodzu, M. I.and Hamada, A., 2009,"A 44-year daily gridded precipitation dataset for Asia based on a dense network of rain gauges",*Sola*, **5**, 137-140.

Yatagai, A., Xie, P. andKitoh, A., 2005,"Utilization of a new gauge-based daily precipitation dataset over monsoon Asia for validation of the daily precipitation climatology simulated by the MRI/JMA 20-km-mesh AGCM", *Sola*, **1**, 193-196.

PRAMOD SONI

Department of Civil Engineering, Indian Institute of Technology (BHU) Varanasi
(*Received 27 June 2023, Accepted 8 April 2024*)
e mail :pramodsoni41@gmail.com

APPENDIX A

The seasonal (or annual) SSR anomalies (μ_y) and corresponding standard errors (ε_y) were employed in weighted linear regression (WLR) to find the trends. The equation of the trend line fitted by the WLR method is given by the Equation 13 in which A and B represent the slope and the intercept, respectively.

$$\mu_y = At_y + B \qquad \text{Equation 13}$$

A and B were found using Equation 14 and Equation 15.

$$A = \frac{\sum_{y=1}^{N_{years}} w_y t_y^2 \sum_{y=1}^{N_{years}} w_y t_y - \sum_{y=1}^{N_{years}} w_y t_y \sum_{y=1}^{N_{years}} w_y t_y \mu_y}{\Delta} \qquad \text{Equation 14}$$

$$B = \frac{\sum_{y=1}^{N_{years}} w_y \sum_{y=1}^{N_{years}} w_y t_y \mu_y - \sum_{y=1}^{N_{years}} w_y t_y \sum_{y=1}^{N_{years}} w_y \mu_y}{\Delta} \qquad \text{Equation 15}$$

where,

$$\Delta = \sum_{y=1}^{N_{years}} w_y \sum_{y=1}^{N_{years}} w_y t_y^2 - \left(\sum_{y=1}^{N_{years}} w_y t_y \right)^2 \qquad \text{Equation 16}$$

In the equations above, t_y and μ_y represent year and seasonal (or annual) SSR anomaly, respectively, for they thdata point, and N_{years} denotes the number of years for which data are available. The weights (w_y) are inversely proportional to the square of the standard error (ε_y) of SSR anomaly μ_y. The model error (σ_M) is given by Equation 17 and the uncertainties in slope ΔA and intercept (ΔB) are given by Equation 18 and Equation 19, respectively

$$\sigma_M = \sqrt{\frac{1}{N-2} \times \frac{N}{\sum_{\forall y} w_y} \sum_{\forall y} w_y \left(\mu_y - At_y - B\right)^2} \qquad \text{Equation 17}$$

$$\Delta A = \frac{1}{\Delta}\sqrt{\left(T_1 + T_2 - T_3\right)} \qquad \text{Equation 18}$$

$$\Delta B = \frac{1}{\Delta}\sqrt{\left(T_5 + T_6 - T_7\right)} \qquad \text{Equation 19}$$

where T_i's are given by,

$$T_1 = \left(\sum_{y=1}^{N_{\text{years}}} w_y t_y^2\right)^2 \sum_{y=1}^{N_{\text{years}}} T_{4,y} w_y^2$$

$$T_2 = \left(\sum_{y=1}^{N_{\text{years}}} w_y t_y\right)^2 \sum_{y=1}^{N_{\text{years}}} T_{4,y} w_y^2 t_y^2$$

$$T_3 = 2\left(\sum_{y=1}^{N_{\text{years}}} w_y t_y^2\right)\left(\sum_{y=1}^{N_{\text{years}}} T_{4,y} w_y^2 t_y\right)\left(\sum_{y=1}^{N_{\text{years}}} w_y t_y\right)$$

$$T_4 = \left(\frac{1}{w_y} + \sigma_M^2\right)$$

$$T_5 = \left(\sum_{y=1}^{N_{\text{years}}} w_y\right)^2 \sum_{y=1}^{N_{\text{years}}} T_{4,y} w_y^2 t_y^2$$

$$T_6 = \left(\sum_{y=1}^{N_{\text{years}}} w_y t_y\right)^2 \sum_{y=1}^{N_{\text{years}}} T_{4,y} w_y^2$$

$$T_7 = 2\left(\sum_{y=1}^{N_{\text{years}}} w_y\right) \sum_{y=1}^{N_{\text{years}}} T_{4,y} w_y^2 t_y \times \left(\sum_{y=1}^{N_{\text{years}}} w_y t_y\right)$$