Trend Analysis of Precipitation for Coastal Districts of Maharashtra

Amit Gangarde[1], Saha Dauji[2], and S. N. Londhe[3]
[1,3]Vishwakarma Institute of Information Technology
[2] Bhabha Atomic Research Centre & Homi Bhabha National Institute
Corresponding author email id: shreenivas.londhe@viit.ac.in

Abstract: The trend of annual rainfall for a region would be useful in planning of the irrigation patterns, reservoir operation, etc. and also to understand the climate change effect on the rainfall of a region. For the coastal belt of Maharashtra, the rainfall is very intense and this is different from the inland districts. The present study aims to analyse the trend of precipitation for the coastal districts of Maharashtra, namely, Sindhudurg, Ratnagiri, Raigad, Mumbai City, Mumbai Suburban, and Thane. The data used for analysis is monthly and annual precipitation totals from the year 1901-2000, obtained from Indian Meteorological Department. In order to examine the trend of rainfall or absence thereof, statistical tests are performed which include Mann-Kendall test and Spearman's *rho*; the trend value would be estimated by Sen's slope estimator.

Keywords: Coastal Precipitation, Trend Analysis, Mann-Kendall test, Spearman's *rho*; Sen's slope

1. Introduction

Rainfall is an important element of hydrological cycle on earth surface. Rainfall received in an area is important factor for meeting demands like agriculture, domestic water supply, hydroelectric power generation plants and industries. Climatic changes may affect the rainfall patterns that impacts availability of water along with the possibility of draughts and floods. Knowledge of meteorological parameters at a site plays a vital role in designing of engineering structures. So, it is essential to analyze rainfall on daily, monthly, seasonal and annual basis. Therefore, trend analysis of rainfall will help in construction of future climate scenarios.

Yue et al., (2002) determined the power of Mann Kendall and spearman's Rho method for detecting trend were similar and depends on pre-assigned significance level, magnitude of trend, sample size, and the amount of variation within a time series. In the study *(Machiwal et al., 2008)*, twenty-nine statistical tests for detecting time series characteristics were evaluated by applying them to analyse 46 years (1957-2002) of annual rainfall, 47 years (1956-2002) of 1-day maximum rainfall and consecutive 2-day, 3-day, 4-day, 5-day and 6-day maximum rainfalls at Kharagpur, West Bengal, India. In this study out of 12 trend detection tests, nine tests indicated no trends in rainfall. Paper concluded that decision about rejecting null hypothesis should be made by analysing critically the results of sufficient number of statistical tests (at least more than two tests). *(Rajeevan et al., 2008)* studied high resolution daily gridded data of 104 years (1901-2004) over India which demonstrated that there was decreasing trend in sea surface temperature (SST) and very high rainfall (VHR) in 1940-1960 and increasing trend in first four decades of 20th century. It suggests that in central India frequency of VHR and risk of high flood may increase due to present global warming scenario.

In another study (*Mondal et al., 2012)* modified Mann Kendall test and Sen's slope estimator method was used for detecting trend (1971-2010) of rainfall of river basin of Orissa near coastal region. In this study some months showed increasing trends while others showed decreasing. *Jagadeesh et al., (2014)* carried out trend analysis of Bharatpur river basin of Kerala (1976-

2008) using Mann Kendall test, Linear regression method and Sen's slope estimator which showed that stations in the north and east showed an increasing trend and stations in the south and west showed a decreasing trend. Nonparametric Mann-Kendall (MK) and Spearman's rho (SR) statistical tests were used to detect trends in monthly, seasonal, and annual precipitation in Swat river basin, Pakistan *(Ahmed et al.,2014)*. The result of test showed that the performance of Mann Kendall and Spearman's rho test was consistent at verified significance level. Trends in monthly and seasonal cumulative rainfall depth, number of rainy days and maximum daily rainfall, and in the monsoon, occurrence was analysed by *(Lacombe et al.,2014)* which confirm that paramount role of global warming in recent rainfall changes. In study of spatial and temporal analysis of rainfall and temperature *(Pingale et al.,2014)* for a period of (1971-2005) positive and negative trends were shown for both. Spatial variation in trends was determined by IDW method. The spatial maps were helpful for local stakeholders and water managers to understand the risks. In the trend analysis of rainfall (1901-2000) over ganga river basin out of 236 districts half of them showed a decreasing trend in annual rainfall in which 39 districts were statistically significant *(Bera, 2017)*.
Homogeneity *(Wijngaard et al.,2003)* of surface air temperature and precipitation from the European Climate Assessment dataset (1901-99) are statistically tested with respect to homogeneity and results of different tests are condensed into three classes: 'useful', 'doubtful' and 'suspect'.
Most of studies were carried out at various regions of India by Mann Kendall and Sen's slope estimator method. These methods are not enough to find the exact trend for rainfall. To interpret more accurate result, a more methods should be employed.

2. Study area and methods

Study area

The study is carried out on Coastal districts of Maharashtra (Sindhudurg, Thane, Raigad,

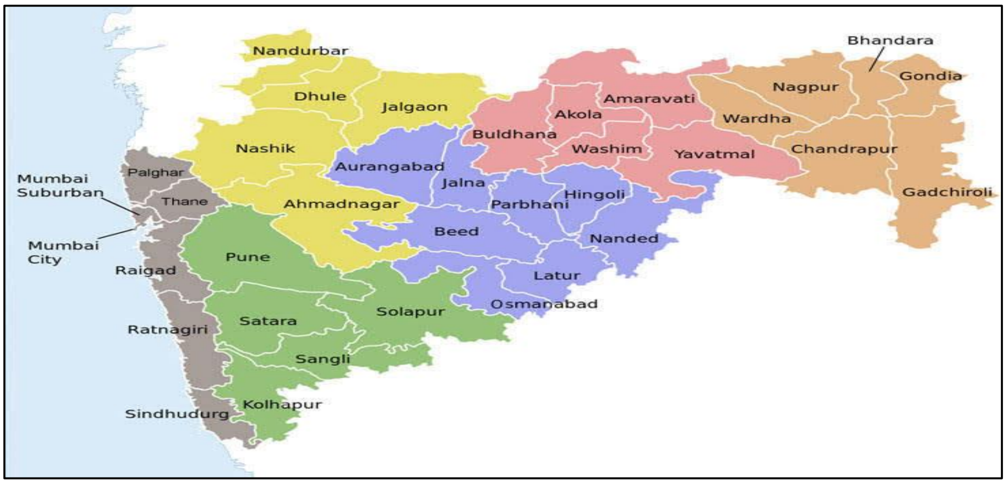

Fig.1 Map of Maharashtra with districts

Ratnagiri, Mumbai Suburban, Mumbai city). The study concentrates on the rainfall analysis from 1901-2000 for Coastal districts of Maharashtra. Annual rainfall data are obtained from the Indian Meteorological Department (IMD), Pune, India.

Methodology

In literature, most of the trend analysis reported detection of trend by Mann Kendall test and evaluated the slope of the trend line by Sen's Slope Estimation. However, it has also been highlighted in literature that inferences drawn from a single statistical test should be treated with caution. It has been recommended to examine the hypothesis using multiple tests and then conclude based on the combined results. This approach has been adopted in the present study. In the study, non-parametric methods, Mann Kendall test, Spearman's Rho test, Difference sign test, Run test and Run on successive differences are used to detect the trend in time series. Magnitude of trend is determined using Sen's slope estimator and linear regression method. Change point is detected using Pettitt test, Von Neumann test, Buishand test respectively. When different tests gave conflicting results, the concept from (Wijngaard et al., 2003) has been employed to arrive at the conclusions. This chapter gives the methodology and formulations of the various tests utilised in this study.

Detection of trend

Mann-Kendall test

The nonparametric Mann–Kendall test *(Bera 2017),* which is commonly, used for identifying trends in hydrologic time series data. One advantage of this test is that the data need not follow any particular distribution. The Mann-Kendall test statistic *S* is determined using the following formula

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} sgn(x_j - x_k) \tag{1}$$

where x_j and x_k are the annual values in years *j* and *k*, *j* > *k*, respectively, and *n* is the number of data points. The value of sgn ($x_j - x_k$) is computed as follows:

$$\text{sgn}(x_j - x_k) = \begin{cases} 1 \text{ if } x_j - x_k > 0 \\ 0 \text{ if } x_j - x_k = 0 \\ -1 \text{ if } x_j - x_k < 0 \end{cases} \tag{2}$$

The variance of *S* is computed by the following equation

$$\text{VAR(S)} = \frac{1}{18} * \ (n(n-1)(2n+5) \tag{3}$$

where, *n* is the number of data points. *Z* statistics or Normal Approximation test generally used when the sample size is greater than 10. The Normal approximation test (*Z* test) is developed by both the values of *S* and VAR(*S*).

$$z = \begin{cases} \frac{S-1}{\sqrt{VAR(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \frac{S+1}{\sqrt{VAR(S)}} & \text{if } S < 0 \end{cases} \tag{4}$$

The positive and negative values of *Z* denote increasing and decreasing trend respectively. The statistic *Z* has a normal distribution. To test for either an increasing or decreasing monotone trend (a two-tailed test) at α level of significance, H_0 is rejected if the absolute value of Z is greater than $Z_{1-\alpha/2}$, where $Z_{1-\alpha/2}$ is acquired from the standard normal cumulative distribution tables *(Ang and Tang 2007).*

Spearman's Rho Test

Spearman's rho test *(Ahmed et al., 2014)* is another rank-based nonparametric method used for trend analysis. In this test, which assumes that time series data are independent and identically distributed, the null hypothesis (H_0) again indicates no trend over time; the alternate hypothesis (H_1) is that a trend exists and that data increase or decrease with i. The test statistics R_{SP} (equation 4.1.2.1) and standardized statistics Z_{sp} (4.1.2.2) are defined as

$$R_{SP} = 1 - \frac{6\sum_{i=1}^{n}(Di-i)^2}{1-R_{SP}^2} \tag{5}$$

$$Z_{sp} = R_{sp}\sqrt{\frac{n-2}{1-R_{sp}^2}} \tag{6}$$

In these equations, Di is the rank of ith observation, i is the chronological order number, n is the total length of the time series data, and Z_{sp} is Student's t-distribution with $(n-2)$ degree of freedom. The positive values of Z_{sp} represent an increasing trend across the hydrologic time series; negative values represent the decreasing trends. The critical value of t at a 0.05 significance level of Student's t-distribution table *(Ang and Tang 2007)* is defined as $(n-2, 1-\alpha/2)$. If $|Z_{sp}| > (n-2, 1-\alpha/2)$, ($H_0$) is rejected and a significant trend exists in the hydrologic time series.

Run test

In this method values above the median considered as positive and values below the median considered as negative. A run *(Machiwal et al., 2008)* is defined as a series of consecutive positive (or negative) values. The runs test is defined as (equation 4.1.3.1):

H_0 : the sequence was produced in random manner

H_1 : the sequence was not produced in a random manner

Test statistic:

$$z = \frac{R - \bar{R}}{s_R} \tag{7}$$

Where R is the observed number of runs $\bar{R}$, is the expected number of runs, and s_R is the standard deviation of number of runs. The values of $\bar{R}$ (4.1.3.2) and s_R (4.1.3.3) are computed as follows:

$$\bar{R} = \frac{2n_1 n_2}{n_1 + n_2} + 1 \tag{8}$$

$$s_R = \frac{2n_1 n_2(2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)} + 1 \tag{9}$$

With n_1 and n_2 denoting the number of positive and negative values in the series. H_0 is rejected if the absolute value of *Z* is greater than $Z_{1-\alpha/2}$, where $Z_{1-\alpha/2}$ is acquired from the standard normal cumulative distribution tables *(Ang and Tang 2007).*

Run test on successive differences

In this test *(Machiwal et al., 2008)* successive difference of data is taken and then next procedure is similar to run test.

H_0 : the sequence was produced in random manner

H_1 : the sequence was not produced in a random manner

The test statistic is (equation (4.1.4.1):

$$z = \frac{R - \bar{R}}{s_R} \tag{10}$$

Where R is the observed number of runs $\bar{R}$, is the expected number of runs, and s_R is the standard deviation of number of Runs. The values of $\bar{R}$ (equation 4.1.4.2) and s_R (4.1.4.3) are computed as follows:

$$\bar{R} = \frac{2n_1 n_2}{n_1 + n_2} + 1 \tag{11}$$

$$s_R = \frac{2n_1 n_2(2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)} + 1 \tag{12}$$

With n_1 and n_2 denoting the number of positive and negative values in the series. H_0 is rejected if the absolute value of *Z* is greater than $Z_{1-\alpha/2}$, where $Z_{1-\alpha/2}$ is acquired from the standard normal cumulative distribution tables *(Ang and Tang 2007).*

2.2 Estimation of slope

Sen's slope estimator

Sen's nonparametric method *(Bera 2017*) was used to estimate the magnitude of trends in the time series data (equation 4.2.1.1)

$$Q_i = \frac{x_j - x_k}{j - k} \tag{13}$$

where x_j and x_k are the data scores at times *j* and *k* (*j* > *k*) respectively. If there are *n* values x_j in the time series, we get as many as $N = n_{(n-1)/2}$ slope estimates of Q_i. The Sen's estimator of slope is the median of these *N* values of Q_i. The *N* values of Q_i are ranked from the smallest to the largest and the Sen's estimator is median of this *Q* values.

Linear regression method

In linear regression test *(Jaiswal et al., 2015)*, a straight line is fitted to the data and the slope of the line may be significantly different from zero or not. For a series of observations x_i, i=1, 2, 3…*n*, a straight line in the form of *y=a+bx* is fitted to the data and slope of line is calculated.

2.3 Test for change point detection

Pettitt's test

$$U_t = \sum_{i=1}^{t} \sum_{j=t+1}^{n} sign(x_t - x_j) \tag{14}$$

$$sgn(x_t - x_j) = \begin{cases} 1 \; if \;\; x_i - x_j > 0 \\ 0 \; if \;\; x_i - x_j = 0 \\ -1 \; if \;\; x_i - x_j < 0 \end{cases} \tag{15}$$

The test statistic *K* (equation 4.3.1.3) and the associated confidence level (*ρ*) (4.3.1.4) for the sample length (*n*) may be described as *(Jaiswal et al., 2015)*:

$$K = Max|U_t| \tag{16}$$

$$\rho = \exp\left(\frac{-k}{n^2 + n^3}\right) \tag{17}$$

The test statistic *K* can also be compared with standard values *(Jaiswal et al., 2015)* at different confidence level for detection of change point in a series.

Von Neumann Ratio Test

The test statistics for change point detection *(Jaiswal et al., 2015)* in a series of observations x_1, x_2, x_3…x_n can be described as (eqution 4.3.2.1):

$$N = \frac{\sum_{i=1}^{n-1}(x_i - x_{i-1})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{18}$$

The critical values of *N (Jaiswal et al., 2015)* at 1 and 5 % confidence levels can be used for identification of non-homogeneous series with change point.

Buishand's Range Test

The adjusted partial sum (S_k), that is the cumulative deviation from mean for kth observation of a series x_1, x_2, x_3 …. x_k…. x_n with mean ($\bar{x}$)can be computed using following equation *(Jaiswal et al., 2015)*:

A series may be homogeneous without any change point if $S_k \cong 0$, because in random series, the deviation from mean will be distributed on both sides of the mean of the series. The significance of shift can be evaluated by computing rescales adjusted range (R) using the following equation (4.3.3.1):

$$R = \frac{Max(S_k) - Min(S_k)}{\bar{x}} \qquad (19)$$

The computed value of $R/\sqrt{n}$ is compared with critical values *(Jaiswal et al., 2015).*

2.4 Qualitative interpretation of Results

Next step is overall evaluation of all tests.
Class 1: ‘useful’ — one or zero tests reject the null hypothesis at the 5% level.
Class 2: ‘doubtful’ — two tests reject the null hypothesis at the 5% level.
Class 3: ‘suspect’ — three or four tests reject the null hypothesis at the 5% level.
The qualitative interpretation *(Wijngaard et al., 2003)* of the categories is as follows:
Class 1: ‘useful’. No clear signal of an inhomogeneity in the series is apparent. Hence, inhomogeneities that may be present in the series are sufficiently small with respect to the inter-annual standard deviation of the testing variable series that they will largely escape detection. The series seems to be sufficiently homogeneous for trend analysis and variability analysis.
Class 2: ‘doubtful’. Indications are present of an inhomogeneity of a magnitude that exceeds the level expressed by the inter-annual standard deviation of the testing variable series. The results of trend analysis and variability analysis should be regarded very critically from the perspective of the existence of possible inhomogeneities.
Class 3: ‘suspect’. It is likely that an inhomogeneity is present that exceeds the level expressed by the inter-annual standard deviation of the testing variable series. Marginal results of trend and variability analysis should be regarded as spurious. Only very large trends may be related to a climatic signal.

4. Results and Discussion

Table 1 Detection of trend

Table 1 shows that Positive trend is observed for Sindhudurg site and Mumbai city. Trend is observed to be ‘Suspect’ *(Wijngaard et al., 2003)* for Thane, Raigarh, Mumbai Suburban site. ‘Doubtful’ trend is observed only for Ratnagiri site.

Table 2 Estimation of slope

Site name	Sen’s slope	Intercept	Linear regression	Intercept
Sindhudurg	6.93	-10535	6.35	-9401
Thane	0.97	435.25	1.17	84.31
Raigad	3.45	-3580.69	4.35	-5311.32
Ratnagiri	2.52	-1462.79	2.42	-1286.96
Mumbai Suburban	4.39	-6444.49	5.84	-9133.028
Mumbai city	6.25	-10407	6.97	-11590.196

Table 2 and Fig.2, Fig.3, Fig.4, Fig.4, Fig.5 and Fig.6 shows that Slope observed by Sen’s slope estimator and linear regression are almost similar. Higher slopes are observed for Sindhudurg, Mumbai Suburban, Mumbai city. Slopes are very smaller for Thane district.

Site Name	Mann-Kendall test	Spearman’s Rho Test	Run test	Run test on successive differences
Sindhudurg	4.02	4.22	3.81	2.12
Thane	0.58	0.53	-0.10	2.23
Raigad	1.73	1.83	-0.60	4.14
Ratnagiri	1.05	1.07	2.41	2.72
Mumbai Suburban	1.50	1.23	1.47	4.39
Mumbai city	3.65	3.78	0	4.06
Critical value	1.96	1.96	1.96	1.96

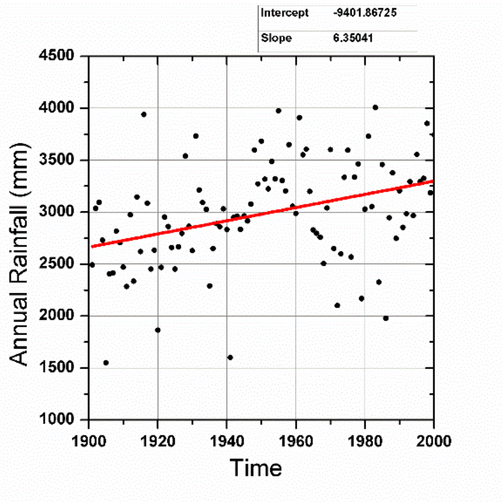

Fig.2 Linear regression (Sindhudurg site)

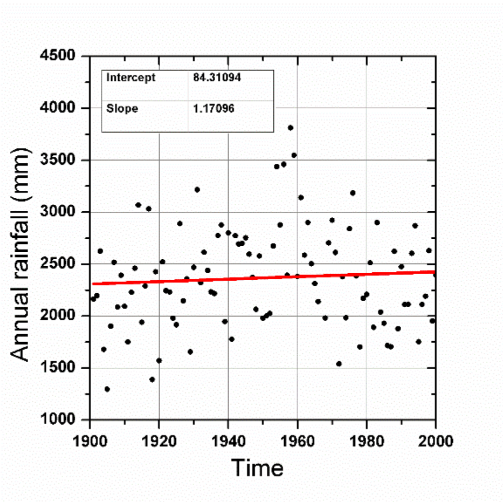

Fig.3 Linear regression (Thane site)

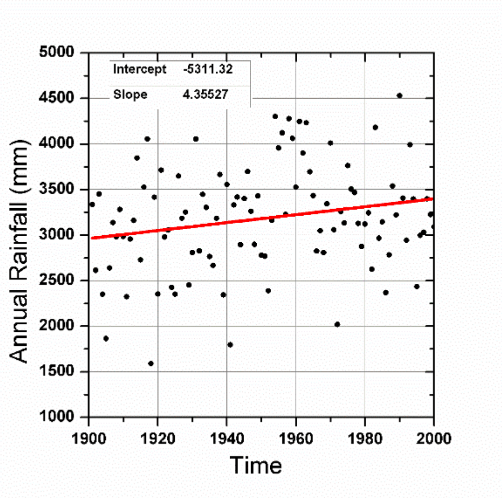

Fig.4 Linear regression (Raigad site)

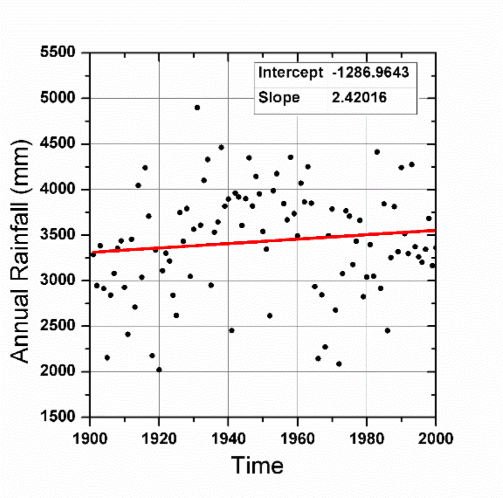

Fig.5 Linear regression (Ratnagiri site)

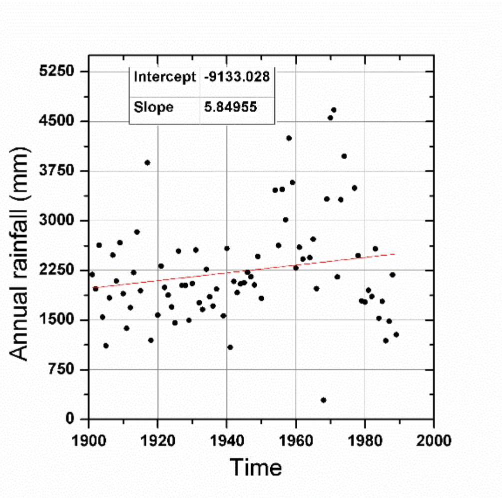

Fig.6 Linear regression

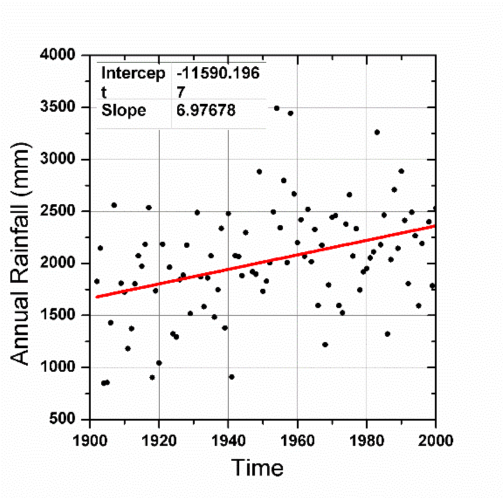

Fig.7 Linear regression

(Mumbai subarban) (Mumbai city)

Table 3 Detection of Change point

Site Name	Pettitt test	Buishand test	Von Neumann test
Sindhudurg	1308	0.48	1.79
Thane	632	0.52	1.63
Raigad	615	0.50	1.79
Ratnagiri	923	0.48	1.50
Mumbai Suburban	556	0.61	1.23
Mumbai city	1039	0.51	1.68
Critical value	677	1.62	1.67

Table shows that ‘Doubtful change point is observed for Sindhudurg, Mumbai city. Change point is ‘Suspect’ for all other districts.

5. Summary and Conclusion

The trend analysis of rainfall helps in forming more robust decisions regarding the future scenarios. When performed for rainfall, it helps to judge the future increase or decrease in precipitation for the Coastal districts of Maharashtra, which forms an important consideration for the design margin in fixing the design basis flood level and ascertaining the design parameters for the various hydraulic structures such as storm water drains and culverts for the facility. In the climate change scenario, the extremes are in general expected to increase in magnitude and frequency.

The present study on Coastal districts of Maharashtra from 1901-2000 using multiple tests to have clear understanding about rainfall trends. In trend analysis, rainfall shows increasing trend for Sindhudurg and Mumbai city. No clear trend is observed for Thane, Raigad, Ratnagiri, Mumbai Suburban districts. Increase in rainfall for Sindhudurg district is about 6.93 mm/year (Sen’s slope), 6.35 mm/year (Linear regression) and for Mumbai city 6.25 mm/year (Sen’s slope), 6.97 mm/year (Linear regression). Slope value obtained for Thane district are smaller. Doubtful change point is observed for Sindhudurg district and Mumbai city. No significant change point is found for other districts.

Overall, it is concluded that the application of several tests for the same objective in a rainfall analysis increases the chance of accurate results. Therefore, the vital decision about the rejection of the null hypothesis should be made by analysing critically the results of a sufficient number of statistical tests (at least more than two tests).

Acknowledgements

The author would like to thank the Indian Meteorological Department (IMD), Pune, India for providing the rainfall data.

References

Ahmed, I., Tang, D., Wang, T., Wang, M., and Wagan, B. (2014). Precipitation Trends over Time Using Mann-Kendall and Spearman's rho Tests in Swat River Basin, Pakistan. *Advances in Meteorology Volume 2015*, 15, doi.org/10.1155/2015/431860

Ang, A. and Tang, W. (2007). Probability Concepts in Engineering (2nd Edition) –Emphasis on Applications in Civil and Environmental Engineering, John Wileyand Sons. Inc, United States of America.

Bera, S. (2017).Trend Analysis of Rainfall in Ganga Basin, India during 1901-2000." *American Journal of Climate Change,* 6, 116-131.

Jagadeesh, P., Anupama, and C. (2014). Statistical and trend analyses of rainfall: a case study of Bharathapuzha river basin, Kerala, India. *ISH Journal of Hydraulic Engineering,* 20:2, 119-132, DOI:10.1080/09715010.2013.843280.

Jaiswal, R., Lohani, A., Tiwari, H. (2015). Statistical analysis for change detection and trend assessment in climatological parameters. *Springer International Publishing*, 2, 729-749, DOI 10.1007/s40710-015-0105-3

Lacombe, G., and McCartney, M., (2014). Uncovering consistencies in Indian rainfall trends observed over the last half century, *Climatic Change*, doi: 10.1007/s10584-013-1036-5

Machiwal, D., and Jha, M. (2008). Comparative evaluation of statistical tests for time series analysis: application to hydrological time series. *Hydrological Sciences Journal*, 53(2), 353-366, DOI: 10.1623/hysj.53.2.353

Mondal, A., Kundu, S., and Mukhopadhyay, A. (2012). Rainfall trend analysis by Mann Kendall test: a case study of north-eastern part of Cuttack district, Orissa. *International Journal of Geology, Earth and Environmental Sciences, ISSN: 2277-2081*

Pingale, S., and Khare, D. (2014). Spatial and temporal trends of mean and extreme rainfall and temperature for the 33 urban centers of the arid and semi-arid state of Rajasthan, India. *Atmospheric Research*, 138, 73–90.

Rajeevan, M., Bhate, J., and Jaswal, A. (2008). Analysis of variability and trends of extreme rainfall events over India using 104 years of gridded daily rainfall data." *Geophysical research letters*, 35, doi:10.1029/2008GL035143.

Wijngaard, J., Klein Tank, A., and Konnen, G. (2003). Homogeneity of 20th century European daily temperature and precipitation series. *International journal of climatology*, 23, 679-692, doi: 10.1002/joc.906

Yue, S., Pilon, P., and Cavadias, G. (2002). Power of Mann -Kendall test and spearman's Rho test for detecting monotonic trends in hydrological series. *Journal of Hydrology*, 259, 254-271p of Maharashtra with districts-https://www.maharashtra.gov.in/1128/Districts last access-23/09/2019,13:50