Trend Analysis of Rainfall in Prakasam District of Andhra Pradesh State in India

M. Satya Swarupa Rani[1*], R. Asha[2] and G. M. V. Prasadarao[3]

[1]Department of Agrometeorology, Krishivignan Kendra, Darsi, Prakasam, Andhra-523247, India.
[2]Department Agril. Economics, Acharya N G Ranga Agricultural University, Agricultural College, Bapatla, India.
[3]Department of Entomology, Krishivignan Kendra, Darsi, Prakasam, Andhra-523247, India.

Authors' contributions

This work was carried out in collaboration among all authors. Author MSSR designed the study, wrote the protocol and wrote the first draft of the manuscript. Author RA performed the statistical analysis and managed the analyses of the study. Author GMVP managed the literature searches. All authors read and approved the final manuscript.

Original Research Article

ABSTRACT

Globally, precipitation trend analysis in different space and time has great impact on crop-planning activities. To get accurate unbiased results a long-term climate analysis of a particular area required in large variability in both spatially, temporally. For sustainable crop production long term weather analysis act as vital role in alternation of existing cropping patterns. This study aimed at analysing the trend of rainfall events in Prakasam district of Andhra state of India the data consists of annual precipitation time series from 1991-2019. Initially study concerns with analysis of data base using descriptive statistics, later trend change was detected by using non parametric tests. The results indicate an increased trend in June and monsoon season, with a decreased trend in July and winter season at 5% level of significance.

_____

Corresponding author: E-mail: satyaswaruparani95@gmail.com;

1. INTRODUCTION

Rainfall is the most varying meteorological parameter from time to time. Changing precipitation pattern is one of the causes of climate change. It lead to farmers risk at field level, depending on the availability it is distributed to various sectors. Amount of water required for various sectors like agriculture, livestock, power generation determined by amount of rainfall received in specific area, under these situations [1]. Agriculture sector has a significant role on arrival of monsoon. India receives most of its total rainfall from June to September. In general rainfall is the key meteorological factor that directly influence crop production, the rainfall parameters include such as intensity, amount, distribution and no. of rainy days are major criteria to characterize the trend in rainfall in a district. Rainfall is the single most important factor in crop production particularly under dry lands areas. Among the climatic factors, rainfall is of greatest concern to the farmers in rainfed agriculture. The variation in seasonal and annual rainfall in space and time are well known and this inter-annual variability of seasonal rainfall has considerable impact on agricultural production. The agro climatic zone of India represent a wide range of rainfall distribution, temperature, humidity, topography, cropping and farming system. Modern agriculture need detailed information on rainfall to plan the cropping pattern. The objective of this work was to determine the homogeneity, trend, and trend change points in the rainfall data of Prakasam district.

1.1 Study Area

Prakasam district situated between 14 57° to 16 17° N latitude and 78 43° to 80 25° E longitude. District was bounded by bay of Bengal in the east, Guntur in the north, Kurnool in the west, Cuddapah and Nellore districts in the south. The central portion of district contains large tracts of low shrubs and forest area diversified with the rocky hills which is the peculiar feature of the district [2]. In the District the sea breeze renders the climate moderate both in winter and summer seasons in the coastal areas of the district. In the non-coastal areas of the district, the heat in the summer is severe especially in the tracts of upland areas and adjoining hills.

2. MATERIALS AND METHODS

2.1 Data Analysis

To determine the precipitation trend in this study descriptive statistics like mean, standard deviation, coefficient variation were used for data analysis. Annual average values calculated to understand the inter annual variation; lower standard deviation value indicates less rainfall variation it doesn't not mean in terms of quantity, lower CV shows consistent rainfall and higher CV specify irregularity in the rainfall distribution. Daily data was analysed for monthly, seasonal, annual precipitation in a time period.
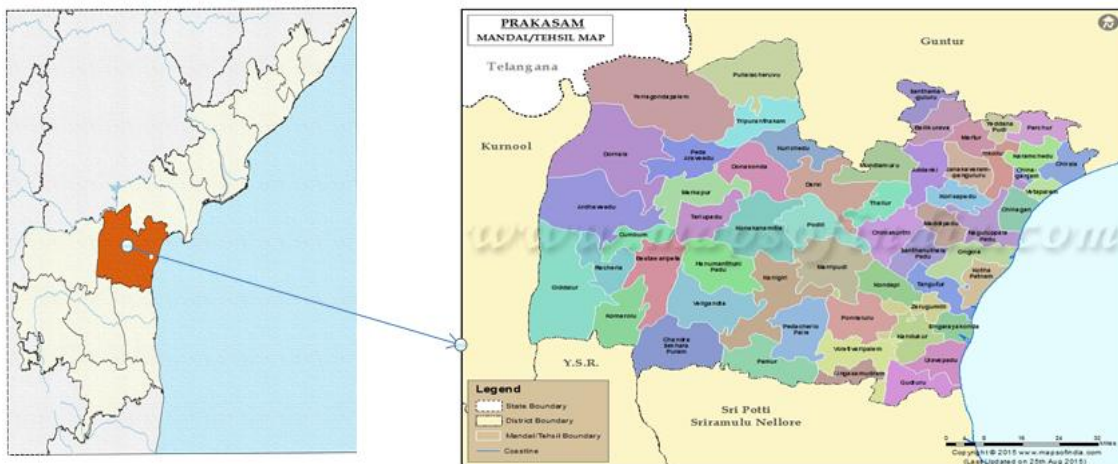

Fig. 1. Study area: Prakasam district

Various statistical tests were performed to analyse the homogeneity and trend in the time series data, after thorough quality checks and preprocessing of the data trend analysis were done. For homogeneity tests pettitt's test, was carried out and then data was tested to see the auto correlation problem prevailed or not. Mann-kendal test conducted to determine trend analysis.

To test the rainfall data follow a normal distribution, the skewness and kurtosis were computed. Skewness is a measure of symmetry, or more precisely, the lack of symmetry. The data set is said to be symmetric if it looks the same to the left and right from the center point. The skewness for a normal distribution is zero.

Kurtosis is a measure of data peakness or flatness relative to a normal distribution. That is, data sets with a high kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. The standard normal distribution has a kurtosis of zero. Positive kurtosis indicates a peaked distribution and negative kurtosis indicates a flat distribution [3].

2.1.1 Pettitt's test

Pettitt test is a non-parametric rank test that has been used in a number of time series data studies to detect abrupt changes in the mean of the distribution of the variable of interest. The test is based on the Mann-Whitney two-sample test (rank-based test), and allows to detection of a single shift at an unknown point in time. This test is often used to detect shifts in extremes because of lack of distributional assumptions

$$Y_k = 2 \sum_{i=1}^{k} ri - k(n-1) \text{ where } k = 1, 2, \dots, n \quad (1)$$

Where: Ranks r_1, r_2,...r_n of the sample series $X_1, X_2, \dots, X_n$.

When a break occurs at year K, the statistic is maximum or minimum at year k = K,

$$YK = \max |Y_k| \quad (2)$$

1<k<n

The value of Y_K is compared with critical values given by [4] to test for statistical significance.

2.1.2 Test for serial correlation

To detect the trend in a time series, the statistical tests assume the subsequent data in the series to be independent. The power of trend tests is highly influenced by the presence of serial correlation in the data [5,6]. A positive serial correlation leads to wrongful rejection of the null hypothesis of no trend when it is true (type I error). Similarly, a negative serial correlation leads to acceptance of the null-hypothesis of no trend when it is false (type II error). To test for serial correlation in the data, lag-1 serial correlation coefficients are calculated. In several of the trend studies, the time series was tested for serial correlation [7,8] by calculating the lag-1serial correlation coefficient ρ_1. For anytime series X_i = x_1, x_2, ...,x_n, lag-1serial correlation coefficient (ρ_1) [9,10] D is calculated as

$$\rho_1 = \frac{\frac{1}{n-1}\sum_{i=1}^{n-1}(x_i - E(x_i)).(x_{i+1} - E(x_i))}{\frac{1}{n}\sum_{i=1}^{n}(x_i - E(x_i))^2} \quad (3)$$

where $E(x_i)$ is the mean of the sample and n is the sample size

$$E(x_i) = \frac{1}{n}\sum_{i=1}^{n} x_i \quad (4)$$

The probability limits for ρ_1 on the correlogram of an independent series is given by Anderson [11].

+6
8

$$\rho 1 = \frac{\frac{-1+1.645\sqrt{n-2}}{n-1}}{\frac{-1\pm1.645\sqrt{n-2}}{n-1}} \quad (5)$$

Significance of serial correlation was evaluated by comparing the ρ_1 value with the critical values of Student's t-distribution values.

2.1.3 Mann–Kendall test

The Mann–Kendall test is perhaps [12,13] the most widely used nonparametric test for detecting trends in hydro-meteorological and environmental data. It is a non-parametric test for monotonic trend detection. It does not assume the data to be normally distributed and is flexible to outliers in the data. The test assumes a null hypothesis, H_0, of no trend and alternate hypothesis, HA, of increasing or decreasing monotonic trend.

For a time series X_i = x_1, x_2, ... , x_n, the Mann–Kendall test statistic S is calculated using Equation (6)

$$S = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} sign(x_j - x_i) \quad (6)$$

where n is the number of data points, x_i and x_j are the data values in time series i and j (j > i), respectively, and sign function as sign $(x_j − x_i)$is the sign function as

$$\text{Sign } (x_i\text{- } x_j) = \begin{cases} -1 \; if \; (x_j - x_i) < 0 \\ 0 \; if \; (x_j - x_i) = 0 \\ 1 \; if \; (x_j - x_i) > 0 \end{cases} \quad (7)$$

Statistics S is normally distributed with parameters E(S) and variance V(S) as given below:

E(S) = 0

$$V(s) = \frac{n(n-1)(2n+5) - \sum_{k=1}^{m} t_k(k)(k-1)(2k+5)}{18} \quad (8)$$

where n is the number of data points, m is the number of tied groups, and t_k denotes the number of ties of extent k. Standardized test statistic Z is calculated using the formula below.

$$Z = \begin{cases} \frac{s-1}{\sqrt{var\,(s)}} \; if \; S > 0 \\ 0 \; if \; S = 0 \\ \frac{s+1}{\sqrt{var\,(s)}} \; if \; S < 0 \end{cases} \quad (9)$$

To test for a monotonic trend at an∝ significant level, the alternate hypothesis of trend is accepted if the absolute value of standardized test statistic Z is greater than the $Z_{1-\alpha/2}$ value obtained from the standard normal cumulative distribution tables. A positive sign of the test statistic indicates an increasing trend and a negative sign indicates a decreasing trend.

2.1.4 Sen's slope estimator

To estimate the true slope of an existing trend (as change per year) the Sen's nonparametric method is used. The Sen's method can be used in cases where the trend can be assumed to be linear.

$Q_i = Qt + B'$ (10)

Where Q is the slope, B is a constant and t is time. To get the slope estimate Q, the slopes of all data value pairs isfirst calculated using the equation:

$Q_i = \frac{x_j - x_k}{j-k}$ (11)

Where x_j an d x_k are data values at time j and k (j>k) respectively. If there are n values x_j in the time series there will be as many as N = n(n-1)/2 slope estimates 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

$Q = Q_{\frac{(N+2)}{2}}$, if N is odd   (12)

Or

$Q = \frac{1}{2}\left(Q_{\left[\frac{n}{2}\right]} + Q_{\frac{(N+2)}{2}}\right)$, if N is even   (13)

To obtain an estimate of B in Equation f(t) the n values of differences $x_i − Qt_i$ are calculated. The median of these values gives an estimate of B

In order to reduce the impact of serial correlation on the MKtest, pre-whitening by [14] was used to remove serial correlation from time series as shown in Equation (14)

$X'_t = X_t − r_1 X_{t-1}$ (14)

where r_1 is the lag 1 serial correlation coefficient of sample data. The lag 1 serial correlation coefficient r_1 can only be estimated from sample data by an autocorrelation function as given in the work of [15]. Equation has been used to remove an AR(1) process from sample data in the trend-detection studies of [16].

3. RESULTS AND DISCUSSION

The distribution of rainfall is very much erratic and uneven so drought is occurring frequently in different regions of the district. Thus, the agricultural production is highly unstable. Even during monsoon season, the district suffers from simultaneous problems of disposal of continuous dry spells and water deficit due to lack of adequate rainfall in other parts. In this study meteorological variable, Daily precipitation data is collected for the district of Prakasam, Andhra Pradesh, India for a period of 29 years and is aggregated to monthly, annual and seasonal totals. It was examined from 1991 to 2019 in Prakasam district to identify the variability, trend from past historical data by computing the data of annual and seasonal rainfall, we concluded that mean annual rainfall was maximum during 2 seasons i.e., post monsoon, monsoon period with 419.65 mm, 404.14 mm respectively with minor variation and the minimum mean rainfall recorded during winter season as 15.94 mm. The descriptive statistics of data is represented in

Table 1. Another statistic variable standard deviation shows that, an amount of 221.33 mm of annual rainfall was deviated from the annual mean value and highest deviation was observed during post monsoon and lowest in winter period, in terms of coefficient of variation highest variability was noticed in winter season and lowest in monsoon and post monsoon period.

The coefficient of skewness of pre monsoon season and annual was near zero which indicates that the data was normally distributed. For monsoon, post monsoon and winter were positively skewed it indicates that data was not normally distributed. All seasonal data points

except post monsoon are near to zero and less than three which indicates nearly all are normally distribution.

Time series data of monthly rainfall not followed normal distribution. To know if there is any abrupt changes in the mean of the distribution of the variable of interest of rainfall data pettitt's test were done using in a two-tailed test in which the null hypothesis was that there is no shift in the mean in the dataset and the alternative hypothesis was that there is a certain data at which a change point can be detected and the mean of the dataset shifts at this break point.

Table 1. Descriptive statistics of rainfall

|	Mean	Median	Standard error	Standard deviation	Sample variance	Kurtosis	Skewness	Coefficient of variance
January	9.37	1.16	3.27	17.63	310.66	5.84	2.50	188.09
February	6.58	0.26	3.02	16.28	265.02	15.83	3.75	247.37
March	6.30	1.77	2.86	15.42	237.87	16.17	3.86	244.75
April	13.86	8.75	3.11	16.75	280.46	4.71	1.94	120.80
May	40.87	19.70	9.49	51.12	2613.20	3.47	2.04	125.07
June	72.07	41.39	12.53	67.47	4552.28	3.88	2.10	93.62
July	95.57	94.22	7.06	38.03	1446.33	0.37	0.62	39.79
August	120.49	113.48	8.10	43.62	1903.13	0.47	0.25	36.21
September	116.01	109.15	8.55	46.02	2117.82	0.05	0.17	39.67
October	195.61	202.28	20.38	109.75	12045.88	-0.42	0.28	56.11
November	178.35	119.95	34.93	188.10	35379.81	13.13	3.17	105.46
December	45.68	41.07	7.91	42.60	1814.54	1.50	1.15	93.25
Pre monsoon	61.04	43.43	9.35	50.35	2535.00	1.76	1.50	82.49
Monsoon	404.14	380.17	21.19	114.13	13024.97	-0.87	0.55	28.24
Post monsoon	419.65	386.29	39.10	210.57	44340.59	5.82	1.81	50.18
Winter	15.95	5.74	4.00	21.54	464.16	2.87	1.84	135.06
Annual	900.78	872.80	41.10	221.33	48986.79	2.24	0.67	24.57
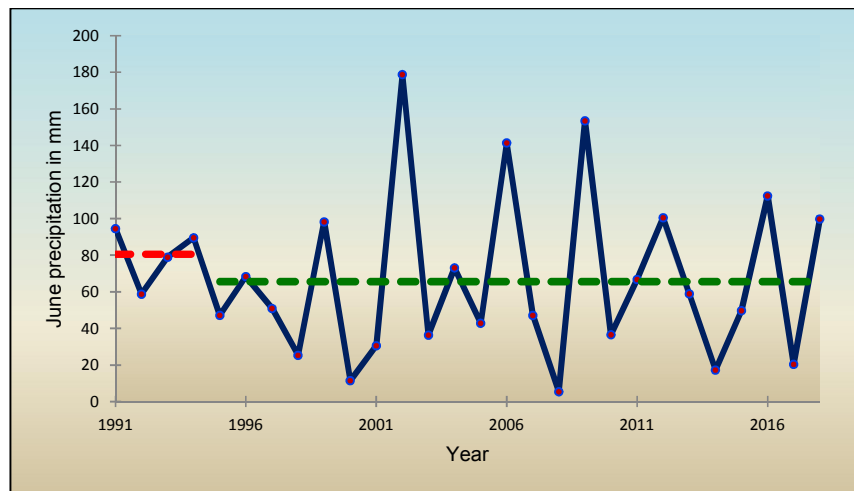
Fig. 2. Significant change points and downward shifts in the mean in the precipitation amounts for June (1991–2019)

Statistically significant change points were detected in June and January. In the time series of June and January precipitation amounts, the significant shift of the mean is downward, and the break point occurs in 1995 and 1996 respectively. There is no significant monotonic shift present in monthly and seasonal precipitation data. The mean monthly (June) amount of precipitation was before 106.22 mm in the period before the date of the break point. Due to the shift in the average, the figure for the mean amount of precipitation in spring was 63.16 mm in the period after the break point (represented in Fig. 2). The mean monthly (January) amount of precipitation was before 12.07 mm in the period before the date of the break point. Due to the shift in the average, the figure for the mean amount of precipitation in spring was 7.66 mm in the period after the break point (represented in Fig. 3).

To detect the autocorrelation problem in time series data correlation coefficient was calculated and then serial auto correlation value calculated. In this data positive serial autocorrelation problem were detected. If we analysed raw data for trend analysis the results may non significant due to serial auto correlation. To avoid this
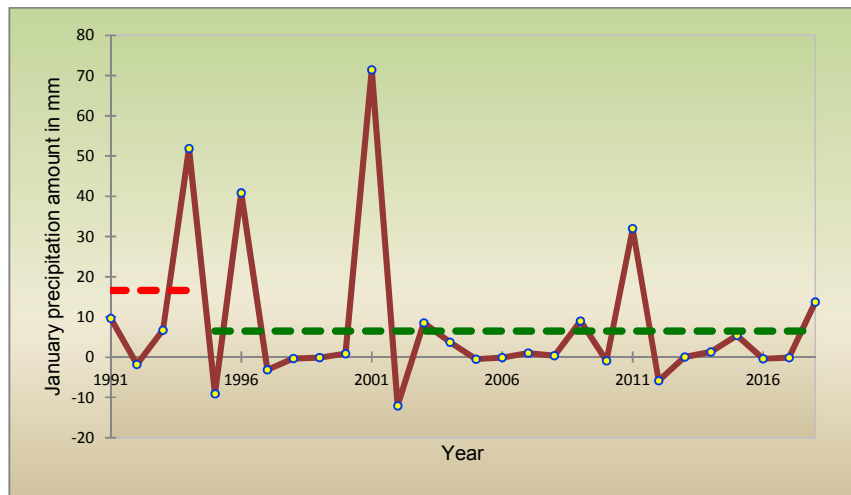

Fig. 3. Significant change points and downward shifts in the mean in the precipitation amounts for January (1991–2019)

Table 2. Monthly, seasonal and annual rainfall for the Mann–Kendall and Sen's Slope test (1991–2019)

	Kendall's tau	S	Var(S)	p-Value	Sen's slope
January	0.003	14	2561	0.070	0.005
February	0.005	2	2562	0.984	0.000
March	0.111	42	2562	0.418	0.042
April	0.011	4	2562	0.953	0.019
May	0.069	26	2562	0.621	0.014
June	0.302	114	2562	0.026*	1.507
July	-0.037	-7	2562	0.018*	-0.411
August	0.074	28	2562	0.594	0.401
September	0.164	62	2562	0.228	1.609
October	-0.143	-54	2562	0.295	-3.313
November	-0.180	-68	2562	0.186	-2.515
December	-0.101	-38	2562	0.465	-0.633
Pre monsoon	0.079	30	2562	0.567	0.380
monsoon	0.085	12	2562	0.040*	2.747
Post monsoon	-0.190	-72	2562	0.161	-6.334
Winter	-0.026	-3	2562	0.049*	-0.084
Annual	-0.085	-32	2562	0.540	-4.316

*Note: * indicated significant at 5% level*

problem the data should be prewhitening before the analysis. Annual trends of precipitation and their magnitude (in mm/year) obtained by the Mann–Kendall test and Sen's slope estimator are given in Table 2. Both positive and negative trends were identified by the statistical tests in data. However, most of the maximum trends were non-significant at 95% confidence levels. Among the positive trends, four significant trends were observed at the 95% confidence levels.

Mann kendall test was used to identify the trend pattern in terms of increase or decrease trend, results showed that during July, October, November, December a negative trend was observed and remaining months found to be positive trend. Post monsoon, winter and annual seasons were showing negative trend and pre monsoon and monsoon periods are having positive trend. Magnitude of trend was verified by Sen's slope method which showed increased trend in premonsoon, monsoon period and decreasing trend during winter, post monsoon periods. The significance at 5% level trend were noted positively in June, monsoon season and negatively in July and winter season. The positive, increasing trend in June is due to south west monsoon.

4. CONCLUSION

From this study we could conclude that, by analysing the past 29 years of historical data (1991 -2019) of Prakasam district *in Andhra* is showing positive trend *during* pre monsoon, monsoon period and negative trend for post monsoon and winter. The coefficient of skewness of pre monsoon season and annual was near zero it indicates that the data was normally distributed. For monsoon, post monsoon and winter were positively skewed. *In case of kurtosis* all seasonal data points except post monsoon are near to zero and less than three which indicates nearly all are normally distribution Significant change points were detected in June and January. It was obtained that precipitation is the most varying meteorological factor which affected the both livelihood and agricultural sector. Hence the need for crop planning activities due to the change in climate.

ACKNOWLEDGEMENT

COMPETING INTERESTS

REFERENCES

1. Nyatuame M, Owusu-Gyimah V, Ampiaw F. Statistical analysis of rainfall trend for Volta region in Ghana. International Journal of Atmospheric Sciences; 2014. Article ID: 203245.
2. District Hand Book of Prakasam, Chief Planning Office, Prakasam District, Andhra Pradesh; 2013.
3. Joanes DN, Gill CA. Comparing measures of sample Skewness and Kurtosis. Journal of Royal Statistical Society. 1998;47(1):183-189.
4. Pettitt AN. A non-parametric approach to the change-point problem. Appl Stat. 1979;28(2):126–135.
5. Yue S, Pilon P, Phinney B, Cavadias G. The influence of autocorrelation on the ability to detect trend in hydrological series. Hydrol. Process. 2002;16:1807–1829.
6. Onyutha C. Statistical uncertainty in hydrometeorological trend analyses. Adv. Meteorol. 2016;8701617.
7. Kendall MG, Stuart A. The Advanced Theory of Statistics; Charles Griffin & Company Ltd.: London, UK. 1968;2. ISBN: 0-85264-069-2.
8. Anderson RL. Distribution of the serial correlation coefficient. Ann. Math. Stat. 1942;1(13):1–13.
9. Chattopadhyay S, Edwards D. Long-term trend analysis of precipitation and air temperature for Kentucky, United States. Climate. 2016;4:10.
10. Basistha A, Arya DS, Goel NK. Analysis of historical changes in rainfall in the Indian Himalayas. Int. J. Climatol. 2009;29:555–572.
11. Anderson RL. Distribution of the serial correlation coefficient. Ann. Math. Stat. 1942;1(13):1–13.
12. Mann HB. Nonparametric tests against trend. Econometrica. 1945;13:245–259.
13. Kendall MG. Rank correlation methods. Charles Griffin: London, UK; 1955.
14. Von Storch H. Misuses of statistical analysis in climate research. In Analysis of Climate Variability: Applications of

Statistical Techniques, Edited by H. V. Storch and A. Navarra. Springer-Verlag, New York. 1995;11–26.

15. Salas JD, Delleur JW, Yevjevich VM, Lane WL. Applied modeling of hydrologic time series. Water Resources Publication:

Littleton, CO, USA; 1980. ISBN: 0-918334-37-3.

16. Douglas EM, Vogel RM, Knoll CN. Trends in flood and low flows in the United States: Impact of spatial correlation. J. Hydrol. 2000;240:90–105.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle4.com/review-history/58856