RESEARCH
www.commprac.com
ISSN 1462 2815

NON-PARAMETRIC, PARAMETRIC TEST AND LINEAR REGRESSION TO ANALYSE TRENDS OF RAINFALL

Sheetal Shende [1], Shweta Dour [2], Rahul Mehta [3] and Kishori Shekokar [4]

[1] Assistant Professor, Data Science Department, Navrachana University, Vadodara, India.
[2] Assistant Professor, Electrical and Electronics Engineering Department, Navrachana University, Vadodara, India.
[3] Assistant Professor, Applied Sciences and Humanities Department, SVIT, Vasad, India.
[4] Assistant Professor, Computer Science and Engineering Department, Parul University, Vadodara, India.

DOI: 10.5281/zenodo.13292505

Abstract

Water availability is mostly determined by the measure of precipitation obtained in a given area. Variations in precipitation patterns have a considerable impact on people's livelihoods, particularly in the region where rainfed cultivation predominates. The present study looked at the temporal variance of rainfall in Panchmahal district of Gujarat, India. Mann–Kendall trend analysis was applied to examine monthly, seasonal, and yearly rainfall series from 1981 to 2017. The trends were also checked using a parametric linear regression test. Using the Wetness Index and standardised rainfall difference, a deficiency in annual total rainfall was detected. To inspect the distribution of rainfall on monthly basis, the Precipitation Concentration Index (PCI) value has been computed. The paper presents, both parametric and non-parametric tests forecast an increased inclination in rainwater from the months of March to September, whereas they predict a decrease trend for all other months. This is reflected in an increase in the study area's annual total rainfall. The average coefficient of variation (CV) for rain water data was 59.25959, indicating that there was a lot of variance within and between years. Rainfall nonuniformity was also highlighted by PCI readings.

INTRODUCTION

The earth system's climate is one of its extremely crucial elements. Weather and climate are made of various elements like rainfall, temperature, humidity, and air pressure Climate refers to the long-term average of the weather for a given place over an extended period of time. It is the overall average and variations of climatic variables over time scales ranging from a few months to many years, to place it in a different way (Shende S. V., 2019).

In this field, (New M, (2001)) reported that climate change has been significantly impacted by global warming everywhere in the world. (Giri RK, 2015), (Nnaji CC, 2016), (Belayneh A, 2016), (Huntington TG, 2006) noticed changes in rainfall pattern in several geographical region throughout the planet. (Shende S. V., August 2021) analysed rainfall and temperature pattern trends. (Sinha Ray K. C., 2000) have studied a trend of declining rainfall over the majority of the nation, with the exception of northwest India and a few northern Indian stations. Many experts have checked out the yearly variability of monsoon rainfall (Rupa Kumar K., 1992) (Parthasarathy B., 1993) (Pant G. B., 1997) (Kripalani R. H., 2001) (Shende S. V., 2019). (Bhatla R, 2014).

Variations in the patterns of rainfall have a major impact on the lives of people and means of living, especially in agricultural nations like India. With quick climate change, it is more important than ever to review rainfall trends, particularly in areas of the country where agriculture is more reliant. The amount of rain that falls between June and November has a significant impact on agricultural production. In this case,

ISSN 1462 2815

calculating monthly predictable rainfall would be useful in estimating the crops' irrigation needs

Study Area and Data

Panchmahal, commonly known as Panch Mahals, is a district in Gujarat State, western India. The Panchmahals District is located in Gujarat's northern region. The district is defined by latitudes of 20.30 to 23.30 north and longitudes of 73.15 to 74.03 south. It has an area of 5210 sq.kms.

Figure 1: Map of India and Gujarat (Wikipedia)

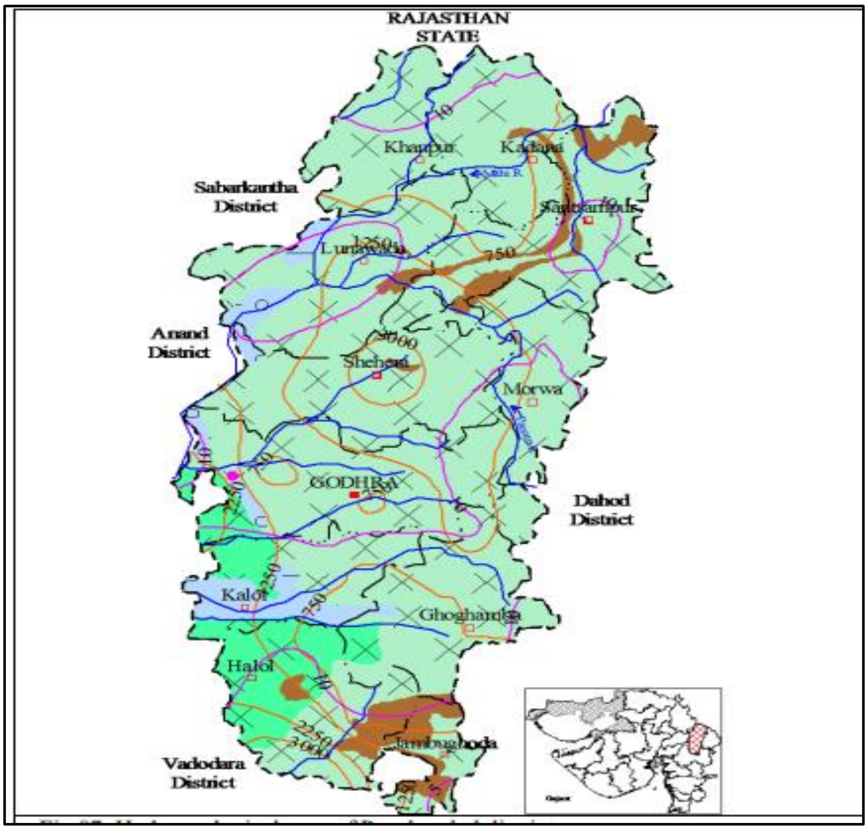


Figure 2: Hydrological Map of Panchmahal District

Government of India Ministry of Water Resources Central Ground Water Board

The states of Madhya Pradesh, Rajasthan, and Godhra district encircle it. The residents of this district mostly rely on rainfall for agricultural purposes because the irrigation potential that has been developed thus far is extremely limited. Panchmahal is having subtropical climates. Summers are hot, while winters are freezing. Throughout the year, it receives 421mm to 1323mm of rain. Since agriculture relies on rainwater, droughts occur every other year. As a result, the crop suffers greatly, and people migrate from one location to another in search of job.

Methodology and Data set

The Mann-Kendall test is applied in order to examine the pattern of annual monthly, seasonal and annual rainfall. (Mann, 1945) (Kendall, 1975) (Mann 1945; Kendall 1975) The most important benefit of non-parametric test like M-K test is that they are independent of nonlinearity.

Table 1: Range of Rainfall and other Rainfall Indices

	Maximum	***Minimum***	**Average**
January	18.17	0	2.567027
February	19.74	0	2.176757
March	40.34	0	2.827568
April	68.99	0	4.908108
May	88.87	0	11.11649
June	544.57	4.96	116.637
July	525.09	53.99	312.7397
August	471.59	37.22	245.5589
September	349.57	3.96	121.5332
October	172.71	0.06	23.74568
November	55.11	0	5.466486
December	10.7	0	1.857838
Annual	1323.77	421.76	851.1349
JF (January-February)	22.34	0	4.743784
MAM (March to May)	158.25	0	18.85216
JJAS (June to September)	1300.54	315.94	796.4689
OND (October to December)	172.83	1.01	31.07
Coefficient of Variation (CV)	68.42635	44.07856	59.25959
Precipitation Concentration Index (PCI)	47.64989	24.64865	30.95121
Standardized rainfall anomaly (Z)	3.936462	-3.57616	0
Index of Wetness (IOW)	155.53	49.55266	100
% of Monsoon	99.76624	69.08968	92.91617

Rainfall data has been gathered through POWER Project funded by the NASA Earth Science/Applied Science Program. The data was gathered from 1981-2017 for district of Panchmahal.

Mann-Kendall's Test

The Mann-Kendall test (Mann, 1945) (Kendall, 1975), which is non-parametric, is frequently used to identify monotonic trends in environmental, hydrological, or climate data series. Benefits of the Mann-Kendall test are as follows: It is not necessary for the data to be regularly distributed since the test does not presume that the data will be delivered in accordance with any specific set of guidelines. Other than the fact that there are fewer sample points, missing data has no effect on it; hence, it could negatively impact statistical significance. Uneven spacing between the time points of measurement has no effect on it. The time series' duration has no bearing on it.

As the data does not possess any trends and the data originates with independent realizations the null hypothesis or H_0, has been considered. According to the alternative hypothesis, or H_A, the trend of the data is monotonic. Computation of the Mann-Kendall test is obtained by using the following formula:

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

$$where, \quad sgn\ (x) = \begin{cases} 1, & if\ x > 0 \\ 0, & if\ x = 0 \\ -1, & if\ x < 0 \end{cases}$$

here n is the sample length, x_k and x_j are from $k = 1, 2, \ldots, n-1$ and $j = k+1, \ldots, n$. If n is bigger than 8, then S is follow normal distribution approximately. The mean of $S\ is\ 0$ and the variance of S can be calculated using following formula:

$$var(S) = \frac{n(n-1)(2n+5) - \sum_{j=1}^{p} t_j(t_j - 1)(2t_j + 5)}{18}$$

where p is the total of tied groups in the data set and t_j is the number of data points in the j^{th} tied group. Test statistic Z is denoted by:

$$Z_s = \begin{cases} \frac{S-1}{\sqrt{var(S)}}, & if\ S > 0 \\ 0, & if\ S = 0 \\ \frac{S+1}{\sqrt{var(S)}}, & if\ S < 0 \end{cases}$$

A very extreme positive value of S indicates a rising trend, whereas a very low negative value shows a declining trend. The presence of a statistically considerable trend is evaluated using Z_s value. If $Z_s > 0$, it indicates an increasing trend, and vice versa. This test has been computed using R software. (Shende S. V., August 2021)

Sen's Slope

Sen's method is a reliable nonparametric technique for estimating a time series' trend slope. It is employed to calculate the actual inclination of a existing trend, like the annual change. Sen's nonparametric approach is applied, and R software is used to conduct the test. Sen's slope indicates an upward or increasing trend in the time series with a positive value and a downward or decreasing trend with a negative value.

Parametric Test

The trends in rainfall were also analyzed using trend lines from linear regression. Both the entire time series and the decadal data were fitted with the regression trend lines. In addition to trend analysis, the rainfall is described using the following indices.

Normal Annual Rainfall

The average of the 30-year rainfall series that followed was used to determine the normal yearly rainfall. The study under consideration examined rainfall series on a monthly, seasonal, and annual basis. Since the average annual rainfall is revised

ISSN 1462 2815

every ten years, the average annual rainfall was computed, and its trend examined after every ten years. Based on Subramanya (2008), one can estimate the average annual rainfall as

$$N = \frac{\sum_{i=0}^{30} P_i}{30},$$

where P_i is the precipitation in i^{th} year.

Index Of Precipitation Concentration

The consistency or irregularity of rainfall over a given period is determined by the Precipitation Concentration Index (PCI). The degree of rainfall nonuniformity increases with PCI value. According to Oliver (1980) and Asfaw et al. (2018), PCI values in range 11 and 15 indicate moderate rainfall concentration, 16 < PCI < 20 indicates high rainfall concentration, and PCI > 21 indicates excessive concentration. PCI < 10 indicates uniform distribution of rainfall. Yearly PCI was calculated for all the entire data set. PCI is represented as

$$PCI = \frac{\sum_{i=0}^{12} p_i^2}{\left(\sum_{i=0}^{12} p_i\right)^2} \; x\; 100$$

Where p_i is the rainfall in i^{th} month of the year.

Standardized Rainfall Anomaly

The severity of the drought is indicated by the standardized rainfall anomaly (Z). Drought severity can be categorized as severe (- 1.28 > Z > - 1.65), moderate (-1.28< Z < - 0.84), no drought (Z > - 0.84), and extreme (Z < - 1.65) based on Z value (Agnew and Chappel, 1999). Z may be calculated as

$$z = \frac{(P_i - \overline{P_i})}{s}$$

where P_i is the rainfall in the i^{th} year; s = standard deviation of rainfall.

Index Of Wetness

It is possible to compare the annual rainfall to the mean annual rainfall using the Index of Wetness. A water deficit equal to the difference from 100 is indicated by an Index of Wetness value less than 100.

$$index\ of\ Wetness = \frac{Annual\ Rainfall}{Normal\ Annual\ Rainfall}\; x\; 100$$

Coefficient Of Variation

Rainfall changeability is classified as low (CV < 20), moderate (20 < CV < 30), and high (CV > 30) based on the coefficient of variation (Hare 2003; Asfaw et al. 2018). Calculating the coefficient of variation can be done as

$$Coefficient\ of\ Variation\ (CV) = \frac{Averae\ Rainfall}{Standard\ Deviation\ of\ Rainfall}\; x\; 100$$

Dependable Rainfall

The time series needs to be arranged and ranked in descending order to compute reliable rainfall. The plotting position formula was then used to determine the probability of occurrence.

$$p(\%) = \frac{m}{N+1} \times 100$$

Where p is the occurrence percentage, m is the rank and N is the total number of observations.

RESULT AND DISCUSSION

Rainfall data from the Panchmahal district was analysed for the period 1981–2017 in order to characterise the region's rainfall and identify its pattern. Table 1 shows monthly rainfall parameters such as extreme, lowest, and average rainfall, as well as various other rainfall attributes faced by the region from 1981 to 2017. The intra-annual variation in rainfall is well depicted in Table 1. It was discovered that the maximum recorded precipitation for the calendar month of January (18 mm), February (19 mm), and December (10.7 mm) was less than the minimum observed rainfall for the months of July (53.99 mm) and August (37 mm). The driest months of the year—January, February, March, April, November, and December—see very little precipitation. To calculate intra-annual variability, PCI was computed for the complete data set spanning from 1981 to 2017.

The PCI values, which ranged from 24.07 to 68.42, indicate extremely high rainfall concentrations, necessitating the determination of the season in which these concentrations occur. Because of the high concentration of precipitation shown in Table 1 during the wet season months, the percentage of monsoon rainfall that contributed to total rainfall was calculated, and it ranged from 69% to 99%. Approximately 92% of the total rainfall was, on average, attributed to the monsoon season. The large value of PCI also explains these results. Within a month a major variation in rainfall is observed. Maximum variation in rainfall was observed in month of July (265.55) and minimum variation in rainfall was observed in March (37.36) while average coefficient of variation was calculated as (59.25).

Significant intra- and interannual rainfall changes are indicated by high value of CV and PCI values. The % of monsoon rainfall and PCI values were utilized to address the intra-annual variability. The standardized rainfall anomaly (z) was calculated on yearly basis in order to analyze the inter-annual variability of rainfall. Standardized rainfall anomaly was computed using the annual rainfall average. Standardized rainfall was computed using the average annual rainfall. Normal annual rainfall was calculated as 787 mm (1981-2000), 894 mm (1991-2010), 841.511 (1981-2010). Annual rainfall more than the average annual rainfall is represented by $Z > 0$ values, while annual rainfall less than the normal annual rainfall is represented by $Z < 0$ values. Drought is a rare occurrence in this region, according to the computed Z values. The region has experienced a severe drought in year 1987 ($Z = -2.6798$) and an extreme drought in year 1988 ($Z = -3.5761$). Whereas the year 1994 was observed as extremely wet year with Z value 3.9364. Data analysis also revealed that 1987 was the driest year with 421.78 mm of observed rainfall, while 1994 was the wettest year with total rainfall of 1323 mm. Both years 1994 and 1987 observed maximum (155.53) and minimum (49.55) Index of Wetness respectively.

Table 2: Monthly Normal Rainfall for Different Periods

	1981-2000	1991-2010	1981-2010
JAN	3.107	2.3985	2.301
FEB	2.077	2.176	2.341667
MAR	1.2595	3.189	2.620667
APR	6.115	1.76	4.871
MAY	14.4665	9.0765	12.86367
JUN	101.5875	144.314	119.493
JUL	301.1555	328.6145	301.237
AUG	218.802	246.6665	250.342
SEP	100.863	128.425	111.3093
OCT	30.911	19.375	25.725
NOV	4.655	6.7095	6.641333
DEC	2.0075	1.373	1.766
ANN	787.0075	894.0755	841.511
JJAS	722.408	848.02	782.3813

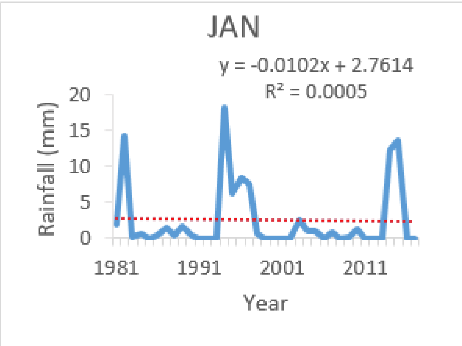
Table 3: Mann-Kendal Test Data for Different Duration

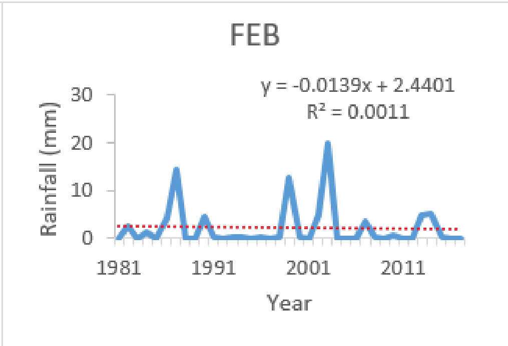
Rainfall/Duration	1981-2017	1981-1999	2000-2017
Annual	1.7395	0.69971	0.53029
JJAS	1.8703	0.83965	0.45453

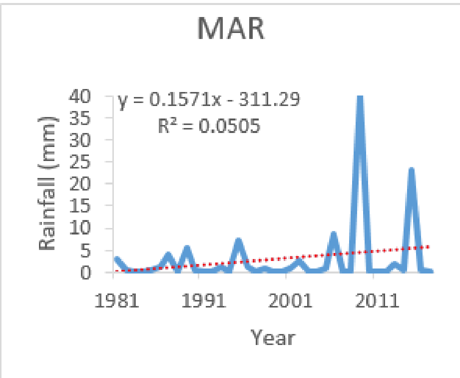
Table 4: Sen's Slope

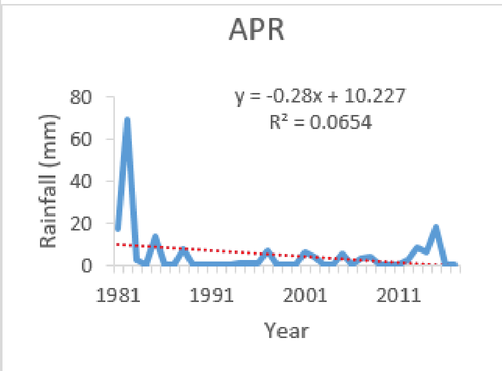
Rainfall/Duration	1981-2017	1981-1999	2000-2017
Annual	134	21	15
JJAS	144	25	13

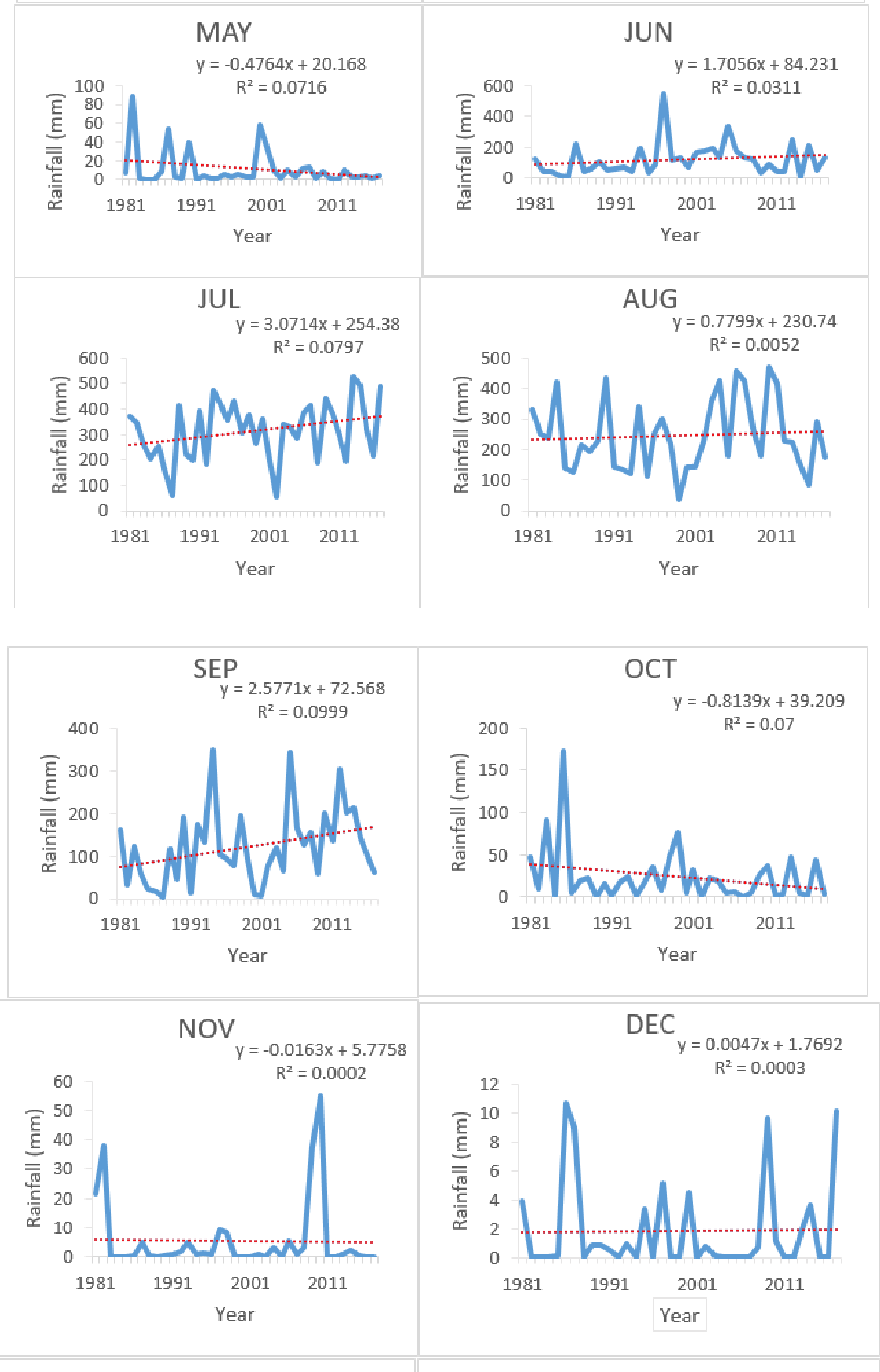
Table 3 reveals that significant positive trends in rainfall exist over the last three decades from 1981-2017. This conclusion can be verified with Sen's slop (Table 4) which gives us positive value over the same period.










MAY
y = -0.4764x + 20.168
R² = 0.0716
Rainfall (mm)
Year
JUN
y = 1.7056x + 84.231
R² = 0.0311
Rainfall (mm)
Year
JUL
y = 3.0714x + 254.38
R² = 0.0797
Rainfall (mm)
Year
AUG
y = 0.7799x + 230.74
R² = 0.0052
Rainfall (mm)
Year
SEP
y = 2.5771x + 72.568
R² = 0.0999
Rainfall (mm)
Year
OCT
y = -0.8139x + 39.209
R² = 0.07
Rainfall (mm)
Year
NOV
y = -0.0163x + 5.7758
R² = 0.0002
Rainfall (mm)
Year
DEC
y = 0.0047x + 1.7692
R² = 0.0003
Rainfall (mm)
Year

ISSN 1462 2815

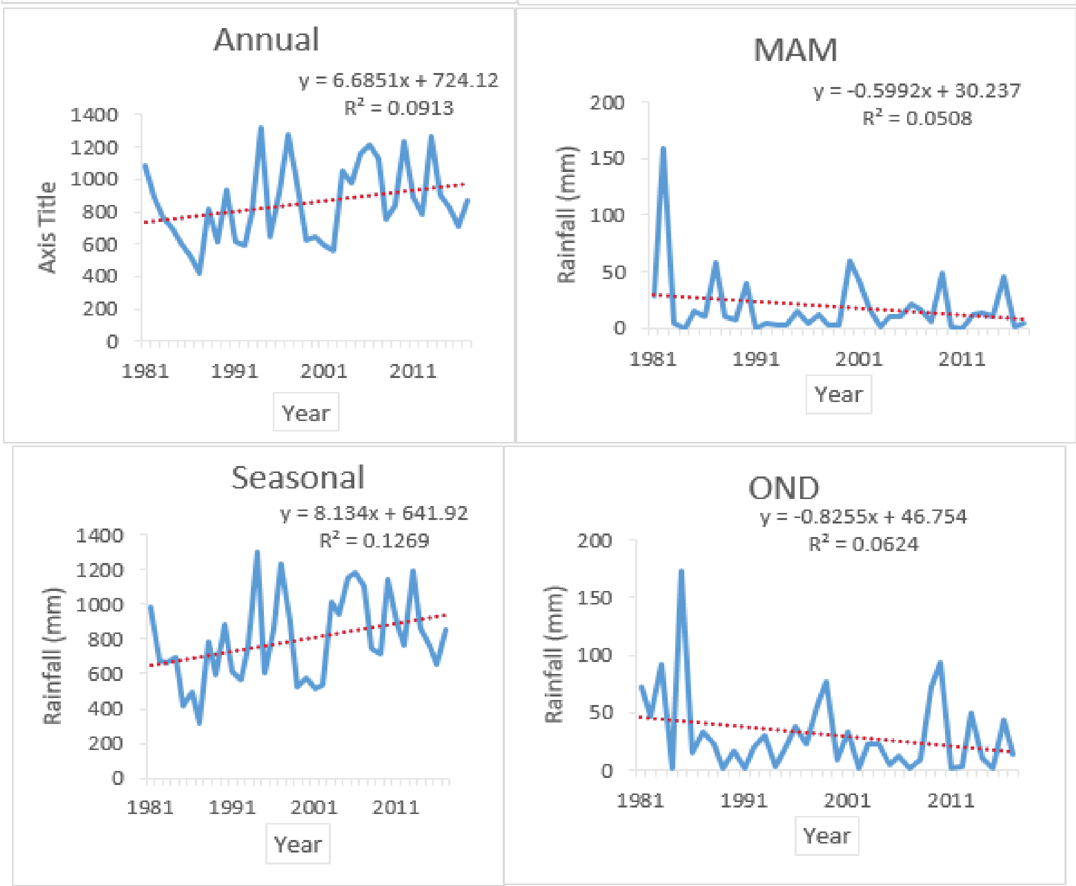


Figure 3: Monthly, Annual and Seasonal Rainfall Trends

CONCLUSION

The rain amount that falls in a particular area has a substantial impact on several aspects of water resource management, including planning, use, distribution, and monitoring of water quality. The sequence of variations in rainfall in Gujarat, India's Panchmahal district, has been studied.

Sen's slope, parametric linear regression testing, and the Mann-Kendall test were utilized to analyze trends in seasonal and annual rainfall series for the years 1981–2017. To determine whether a given year had insufficient rainfall, the Index of Wetness and the Standardized Rainfall Anomaly were computed. Precipitation Concentration Index(PCI) was also calculated to verify the month wise distribution of rainfall.

Very high diversity of rainfall represented by PCI value. Monsoon months from June to September marks high concentration (94%) of rainfall, which is shown by analysis of data. Linear regression revels decline in rainfall in the months January, February, April, May, October which are non- monsoon months.

A meaningful positive tendency in total precipitation was observed over the period from 1981 to 2017. If implemented, a large-scale rainwater gathering initiative in Panchmahal district might minimise flood risk while also securing freshwater and groundwater supplies.

References

1) Belayneh A, A. J. (2016). Short-term SPI drought forecasting in the Awash River Basin in Ethiopia using wavelet transforms and machine learning methods. *Sustain Water Res Manag 2(1):*, 87–101.

2) Bhatla R, T. A. (2014). The study of rainfall and temperature variability over Varanasi. *Int J Earth Atmos Sci 1(2):*, 90–94.

3) Giri RK, P. D. (2015). Rainfall comparison of automatic weather stations and manual observations over Bihar region. . *Int J Phys Math Sci 5(2):*, 1–22.

4) Huntington TG. (2006). Evidence for intensification of the global water cycle: review and synthesis. *J Hydrol 319(1):*, 83–95.

5) Kendall, M. G. (1975). Rank Correlation Methods . *U.K.:Charles Griffin*.

6) Kripalani R. H., K. A. (2001). “Monsoon rainfall variations and tile connections over South and East Asia. *International Journal of Climatology, 21,* , 603-616.

7) Mann, H. B. (1945). . Non-parametric tests against trend,. *Econometrica, 13, 245-259.*, 245-259.

8) New M, T. M. ((2001)). *Precipitation measurements and trends in the 20th century.* Int J Climatol 21(15):1889–1922.

9) Nnaji CC, M. C. (2016). Hierarchical analysis of rainfall variability across Nigeria. *Theor Appl Climatol 123(1–2):* , 171–184.

10) Pant G. B., R. K. (1997). Climates of South Asia, . *John Wiley & Sons: Chichester, UK,* , 320 pp.

11) Parthasarathy B., R. K. (1993). , “Homogeneous Indian monsoon rainfall: variability and prediction . *Proceding of Indian Academy of Science Earth and Planetary Science, 102, 121-155.*, 121-155.

12) Rupa Kumar K., P. G. (1992). Spatial and sub-seasonal patterns of the long-term trends of Indian summer monsoon rainfall. *International Journal of Climatology, 12,* , 257-268. .

13) Shende, S. V. (2019). Artificial Neural Network in Analysis of. *Journal of Emerging Technologies and Innovative Research (JETIR), 6*(2).

14) Shende, S. V. (August 2021). Rainfall and temperature pattern trend analysis of Dahod district of Gujarat, India. *Turkish Online Journal of Qualitative Inquiry (TOJQI), 12*, 255-266.

15) Sinha Ray K. C., S. A. (2000). Is there any change in extreme events like drought and heavy rainfall?”. *Current Science, 79,*, 155-158.