ST. XAVIER'S COLLEGE (AUTONOMOUS), KOLKATA

DISSERTATION REPORT

TIME SERIES ANALYSIS OF RAINFALL IN INDIA

NAME: MEGHNA MONDAL

DEPARTMENT: STATISTICS (STSA)

ROLL NO.: 0451

REGISTRATION NO.: A01-2112-0803-18

SEMESTER: 6

NAME OF THE SUPERVISOR: PROF. PALLABI GHOSH

STUDENT'S DECLARATION:

I affirm that I have identified all my sources and that no part of my dissertation paper uses unacknowledged materials.

SIGNATURE: Meghna Mondal

DATE: **10.05.2021**

CONTENTS

TOPICS	PAGE NO.
INTRODUCTION	3-7
STUDYAREA AND DATA USED	7-9
Methodology	10-24
Results and Discussion	25-63
Conclusion	64-65
Acknowledgement	66
Bibliography	66
Appendix	67-74

INTRODUCTION

Time series analysis accounts for the fact that data points taken over time may have a meaningful reasoning (such as autocorrelation, trend or seasonal variation) that should be accounted for. Time series analysis provides a body of techniques to better understand a dataset. In mathematics, a **time series** is a series of data points indexed (or listed or graphed) in time order. Most commonly, a time series is a sequence taken at successive equally spaced points in time. Thus it is a sequence of discrete-time data. **Time series analysis** comprises methods for analysing time series data in order to extract meaningful statistics and other characteristics of the data.

Here, we are interested for Time Series Analysis of Rainfall Pattern in India for different parts of our country individually. The study of precipitation trends is critically important for a country like India whose food security and economy are dependent on the timely availability of water. In this work, monthly, seasonal and annual trends of rainfall have been studied using monthly data series of 146 years (1871–2016) for 5 sub-divisions (sub-regions) in India. Although the subject area of climate change is vast, the changing pattern of rainfall is a topic within this field that deserves urgent and systematic attention, since it affects both the availability of freshwater and food production.

Example: Based on an experiment, at New Delhi, India in 2007, it has been reported that a 1^0C rise in temperature throughout the growing period will reduce wheat production by 5 million tonnes.

The global average precipitation is projected to increase, but both increases and decreases are expected at the regional and continental scales. Higher or lower rainfall, or changes in its spatial and seasonal distribution would influence the spatial and temporal distribution of runoff, soil moisture and groundwater reserves, and would affect the frequency of droughts and floods. Further, temporal change in precipitation distribution will affect cropping patterns and productivity.

According to the Intergovernmental Panel on Climate Change (IPCC, 2007), future climate change is likely to affect agriculture, increase the risk of hunger and water scarcity, and lead to more rapid melting of glaciers. Freshwater availability in many river basins in India is likely to decrease due to climate change. This decrease, along with population growth and rising living standards, could adversely affect many people in India by the 2050s. Accelerated glacier melt is likely to cause an increase in the number and severity of glacier melt related floods, slope destabilization and a decrease in river flows as glaciers recede (IPCC, 2007). Lal (2001) discussed the implications of climate change on Indian water resources quantifying the impact of climate change on the water resources of Indian river systems. It is found that the yield of wheat, mustard, barley and chickpea show signs of stagnation or decrease following a rise in temperature in four northern states of India.

Under the conditions of skewed water availability and its mismatch with demand, large storage reservoirs may be needed to redistribute the natural flow of streams in accordance with the requirements of a specific region. The general practice of designing a reservoir is based on the assumption that climate is stationary. Changes in rainfall due to global warming will influence the hydrological cycle and the pattern of stream flows. This will call for a review of reservoir design and management practices in India.

The Indian climate is dominated by the southwest monsoon. About 80% of the rainfall in India occurs during the four monsoon months (June–September) with large spatial and temporal variations over the country. Such a heavy concentration of rainfall results in a scarcity of water in many parts of the country during the non-monsoon period. Therefore, for India, where agriculture has a significant influence on both the economy and livelihood, the availability of adequate water for irrigation under changed climatic scenarios is very important. The agricultural output is primarily governed by timely availability of water. In future, population growth along with a higher demand for water for irrigation and industries will put more pressure on water resources.

With the growing recognition of the possibility of adverse impacts of global climate change on water resources, an assessment of future water availability at various spatial and temporal scales is needed. It is expected that the response of hydrological systems, erosion processes and sedimentation could significantly alter due to climate change. An understanding of the hydrological response of a river basin under changed climatic conditions would help solve problems associated with floods, droughts and allocation of water for agriculture, industry, hydropower generation, domestic and industrial use.

Scenarios of changes in runoff and its distribution depend on the future climate scenarios.

In India, attempts have been made in the past to determine trends in the rainfall at national and regional scales. Most of the rainfall studies were confined to the analysis of annual and seasonal series for individual or groups of stations. In the present study, a much wider view has been taken, and changes in rainfall have been studied on seasonal and annual scales for five main regions. Further, the time series of rainfall data used in this study spans more than 100 years. Thus, the present analysis is a significant improvement over the studies carried out for less span of time.

SIGNATURES OF CLIMATE CHANGE OVER INDIA

Studies carried out by several investigators have shown that the trend and magnitude of warming over India/the Indian sub-continent over the last century is broadly consistent with the global trend and magnitude analysed the seasonal and annual air temperatures from 1881–1997 and have shown that there has been an increasing trend of mean annual temperature, at the rate of $0 \cdot 57^0$C per 100 years. Studies have found a warming trend in seven of the nine river basins in northwest and central India.

Some past studies relating to changes in rainfall over India have concluded that there is no clear trend of increase or decrease in average annual rainfall over the country. Though no trend in the monsoon rainfall in India is found over a long period of time, particularly on the all-India scale, pockets of significant long-term rainfall changes have been identified.

Recent studies show that, in general, the frequency of more intense rainfall events in many parts of Asia has increased, while the number of rainy days and total annual amount of precipitation has decreased. Daily rainfall data has been used to show the significant rising trends in the frequency and magnitude of extreme rain events, and a significant decreasing trend in the frequency of moderate events over central India during the monsoon seasons from 1951 to 2000. The frequency of heavy rainfall events during the monsoon season was found to be increasing over the Andaman and Nicobar Islands, Lakshadweep, the west coast and some pockets in central and northwest India, whereas it was found to be decreasing in winter, pre-monsoon and post-monsoon seasons over most parts of India. It has been inferred that there has been a westward shift in rainfall activity over the Indo-Gangetic Plain region. An increase in intense rainfall events leads to more severe floods and landslides. The number of cyclones originating from the Bay of Bengal and the Arabian Sea has decreased since 1970, but their intensity has increased. Moreover, the damage caused by intense cyclones has risen significantly in India. In three consecutive years since 2002, there were large floods in the north-eastern states of India, on 26–27 July 2005, a record 944 mm of rain fell in Mumbai, but the seasons of 2006 and 2007 saw deficient rainfall. Severe floods were observed in many parts of Gujarat and Rajasthan during the monsoon seasons of 2006 and 2007 (India Meteorological Department, 2006, 2007).

STUDYAREA AND DATA USED

Based on meteorological considerations, India has been divided into 5 meteorological sub-divisions. The subdivisions on the mainland are shown in Fig. 1. In this study, rainfall over the whole of India (except the Hilly region and islands) was considered. The geographical area of the sub-divisions considered in this study is 2.88x 10^6 km^2 (3.29 x10^6 km^2, being the whole of India) excluding islands. Sub -divisional monthly rainfall data of India prepared by the Indian Institute of Tropical Meteorology (IITM) were used in this study. A network of 306 stations (one representative station per district) over 5 meteorological subdivisions was used to prepare the sub-divisional data. The monthly (January–December) area weighted rainfall series for each meteorological sub-division were prepared by assigning the district area as the weight for each rain gauge station in that sub-division. The station rainfall data were obtained from the India Meteorological Department (IMD). Before releasing the data, the IMD carries out quality checks to ensure that error-free data are used in analysis and design. Thus, the quality of this data set is very good and it is one of the most reliable long series of data. The monthly data were available for 146 years (1871–2016) for 5 sub-divisions. Rainfall data of six meteorological sub-divisions, namely Jammu and Kashmir, Uttaranchal, Himachal Pradesh, Arunachal Pradesh, Lakshadweep and Andaman & Nicobar Islands, were not available. The area-weighted monthly rainfall (by assigning the sub-division area as the weight) series of these five regions, were available at htpp://www.tropmet. res.in and were also analysed separately. For the trend analysis, monthly rainfall series were used to form annual Rainfall analysis was carried out for the whole year series of these variables.

Region name	**Sub-divisions forming the region**	**Area (km2)**	**% of all regions**	**No. of rainfall stations used**
North East India	Assam & Meghalaya, Nagaland, Manipur, Mizoram & Tripura, Sub-Himalayan West Bengal & Sikkim, Gangetic West Bengal	267 444	9.29	30
Central North East India	Odisha, Jharkhand, Bihar, East Uttar Pradesh, West Uttar Pradesh	573 006	19.89	75
North West India	Haryana, Punjab, West Rajasthan, East Rajasthan, Gujrat, Saurashtra, Kutch & Diu	634272	22.02	66
West Central India	West Madhya Pradesh, East Madhya Pradesh, Konkan & Goa, Madhya Maharashtra, Marathwada, Vidarbha, Chhattisgarh, Telangana, North Interior Karnataka	962 694	33.42	85
Peninsular India	Coastal Andhra Pradesh, Rayalaseema, Tamil Nadu & Pondicherry, Coastal Karnataka, South Interior Karnataka, Kerala	442 908	15.38	50
Whole study area		2 880 324	100	306


North West
Central
North East
North
East
West Central
Peninsular

Figure 1

Methodology

Assuming a multiplicative model, we analyse each data graphically and study the nature of the series followed by the identification of several components. Clearly, this data contains seasonal fluctuations. We then check for the cyclical component and obtain the residual series. Basically, we isolate and measure each component and analyse them separately. Thus, we get an overall idea about the behaviour of rainfall pattern in India. This analysis will help us understand the change in rainfall pattern due to the climate change over the span of 146 years and help us taking preventive measures.

Now, we will discuss the procedure and various steps of the decomposition of a Time Series Data.

Decomposition of Time Series:

A plot of a time series gives us an overall impression of haphazard movement. A critical study of the series, however, reveals that the movements are not completely haphazard, and, at least a part of it can be accounted for. The part which can be accounted for is called systematic part and the other part is called unsystematic or irregular.

The systematic movement of a time series may be attributed to three broad factors:

1. Trend (or Secular trend)
2. Seasonal Variation.
3. Cyclical Fluctuation.

Trend (or Secular trend)

It is a smooth, regular and long-term movement of a time series. Some series may exhibit an upward or a downward trend. Increase in demand of a commodity may exhibit an upward trend, while, decrease in demand of a product due to unavailability of raw materials, or, a substitute taking its place may lead to a decreasing or downward trend. Also, some series may exhibit a constant trend. For instance, per capita consumption of salt may exhibit such a trend. Again, some series, after a period of growth, may reverse their course and enter a period of decline. But sudden or frequent changes are incompatible with the idea of trend.

In time series analysis, the trend value at any time point is taken as the "normal value" or "mean value" at that time point.

A study of secular trend is useful for the following purposes:

(a) For describing the underlying pattern of behaviour which has characterized the series in the past, we can study the trend of growth or decline – as the case may be.

(b) For studying the influence of seasonal, cyclical and irregular forces of change, the secular trend is removed from the observed values. The combined force of change due to seasonal, cyclical and irregular components is called ***residual forces of change***.

(c) For extrapolation on the assumption that the past behaviour will continue in the future.

Seasonal Variation

It is an oscillatory movement in a time series where the period of oscillation is **not** more than a year. A periodic movement in a time series is one which recurs or repeats at regular intervals of time. This type of movement is generally governed by climatic changes and the customs and habits that people follow regularly at different points of time.

A study of seasonal variation of the time series is useful for the following purposes:

(a) To analyse the seasonal pattern in observed values in a short-term time series.

(b) Once the seasonal factor is known, it can be used for separating the cyclical and irregular forces of change from the residual forces of change.

(c) The seasonal factor can be used for adjustment in the value forecasted on the basis of projection of trend. Short term forecasts are always based on seasonal factor.

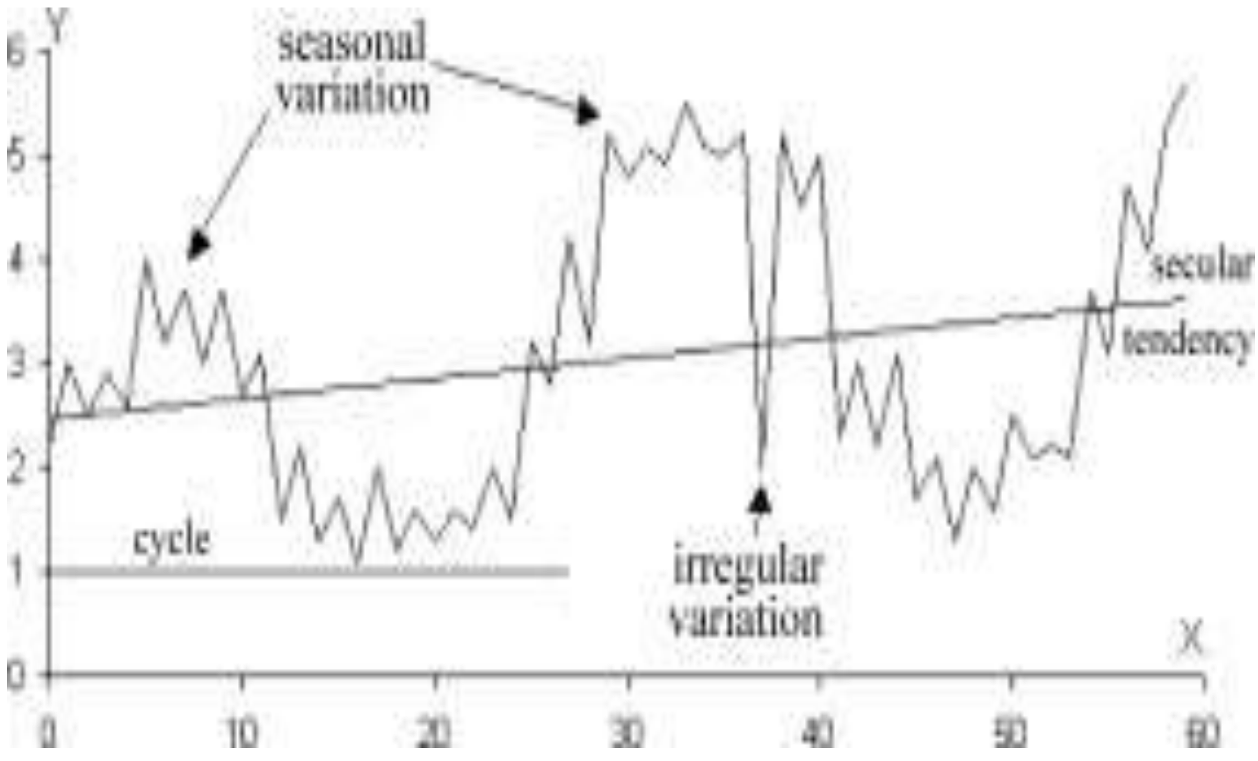
Cyclical Fluctuation

It is also an oscillatory movement in a time series where the period of oscillation is more than a year. This type of periodic movement is generally observed in economic and business cycles which undergo four phases – boom (or prosperity), decline, depression and recovery. The time lag between two consecutive booms or two consecutive depressions is known as the period of a cycle.

The knowledge of cyclical fluctuation is very useful in predicting the turning point in business activities. In fact, for forecasting purposes the seasonal factor is to be considered with reference to the basic forces of cyclical variation.

Irregular Component

It is that part of a time series which encompasses all types of movements that are not accounted for by trend, seasonal and cyclical components. It is caused by unforeseen events like wars, floods, strikes, pandemics, etc.



By decomposition of a timeseries we mean breaking down a time series in its various components. The analysis of time series consists two major steps – (a) Identifying the various forces of influence whose interaction produces variations in the time series and (b) isolating, analysing and measuring the effect of these forces separately and independently, by holding the other things as constant.

There are two broadly used models for decomposition of a time series into its various components.

These are

1. **Additive Model:** According to this model, decomposition of time series is done on the assumption that the effects of various components are additive in nature. In other words,

$Xt = Tt + St + Ct + It$

where Xt is the value of the time series at time point t while Tt, St, Ct and It are respectively the values of trend, seasonal component, cyclical component and Irregular variation at time point t.

In this model St, Ct and It are absolute quantities and can have positive or negative values.

The model assumes that the four components of the time series are independent of each other and none has any effect on the remaining three components. In actual practice, very often, this assumption does not hold good.

2. **Multiplicative Model:** In this model the decomposition is done on the assumption that the effects of the four components are multiplicative and they are not necessarily independent of each other. The model is given by

$Xt = Tt.St.Ct.It$

In this model St, Ct and It are not absolute amounts as in the case of the Additive model. They are relative variations and are expressed as rate or indices fluctuating above and below unity. It is presumed that the geometric mean of all St, Ct and It (for varying values of t) would be unity.

The multiplicative model may also be expressed in terms of logarithms as given below: $logXt = logTt + logSt + logCt + logIt$

We assume our data to follow multiplicative model as the various factors affecting such time series are not independent of each other.

The multiplicative models can be used to obtain the values of various components of a time series by division. Now, for each dataset we will perform the following methods to analyse each component.

Estimation of Trend

We are given monthly data, that’s why method of moving average of period 12 is the best method for estimation of trend.

Method of Moving Averages

A simple moving average of period ‘k’ is a series of simple Arithmetic means each of k successive observations. We start with first k observations. At the next step, we leave the first and include the $(k + 1)^{th}$ observation. This process is continued until we arrive at the last k observations.

If k, the period of moving average is odd, each moving average value corresponds to the tabulated time value. However, if k is even (we have k=12), each moving average value falls mid-way between two tabulated time values. In this case we calculate a subsequent twoitem moving average to make the resultant Moving Average values correspond to the tabulated time values. This is known as centring.

Let $y_1, y_2, \cdots, y_n$ be the values of the time series corresponding to the time points $t_1, t_2, \cdots, t_n$. We know, in a time series, there are four components – trend, seasonal variation, cyclical fluctuation and Irregular variation. Let Tt, St, Ct and It be respectively the values of these components at the time point t.

for a multiplicative model, we have, $yt = Zt*Tt$, then in the Z-series there will be seasonal, cyclical and irregular components only.

Here, we get the values of Tt by the Method of Moving Average. Now, we detrend the data by dividing the yt's by Tt.

Measurement of Seasonal Variation

Seasonal patterns are exhibited by most of the business and economic time series and the study of this phenomenon is essential because of the following reasons:

(i) To isolate seasonal variation, i.e., to determine the effect of the season on the value of the variable under consideration.

(ii) To eliminate the seasonal variation, i.e., to study what would be the value of the variable had there been no seasonal effect.

The determination of seasonal effects is of paramount importance in planning business efficiencies, production programming, etc. Moreover, the isolation and elimination of seasonal factor from the data is necessary to study the effect of cycles.

As we have used the Method of Moving Average for estimating Trend Ratio to Moving Average Method will be best applicable for computation of Seasonal Indices.

Ratio to Moving Average Method

We know that moving average eliminated the periodic movement if the period of the moving average equals the period of oscillatory movement to be eliminated. Thus, for monthly data, a 12-month moving average should completely eliminate the seasonal movement if they are of constant pattern and intensity.

The method of obtaining seasonal indices by this method involves the following steps:

(i) We have calculated the centred 12-month moving average of the data. These moving average values will then give estimates of the combined effect of trend and cyclical variation.

(ii) Express the original data except for six months in the beginning and six months at the end as percentages of the moving average values. As we use the multiplicative model, these percentages would then represent the combined effect of seasonal and irregular components.

(iii) The preliminary seasonal indices are obtained by eliminating the irregular or random component by averaging the percentages for each month.

(iv) The sum of these seasonal indices (=S, say), will not, in general, be 1200. These seasonal indices are called "unadjusted seasonal indices". Finally, an adjustment is to be done by multiplying the indices throughout by the factor $1200/S$ to make the sum of these indices 1200. These new indices are known as "adjusted seasonal indices".

(v) Thus, we get the seasonal indices (St) for the 12 months. We repeat the set of the Adjusted Seasonal Indices for each year.

for a multiplicative model, we have, Zt =St*u_t, then in the u-series there will be cyclical and irregular components only. To eliminate the seasonal component, we divide the Z-series by St.

Measurement of Cyclical Fluctuations

At the final stage it is necessary to remove It from $u_t = Ct \times It$ by some process of smoothing. A satisfactory method of determining the cyclical component is the method of ***Periodogram analysis***.

Periodogram Analysis

As we have already eliminated trend and seasonal effects from our time series.

Let u_t(t = 1, 2, …, n) represent this residual series. Here our objective is to find out whether u_t contains a harmonic term with period μ.

Now, consider the quantities:

$$A = \frac{2}{n}\sum_{t=1}^{n} u_t \cos\frac{2\pi t}{\mu} \ldots\ldots\ldots\ldots..(1)$$

$$B = \frac{2}{n}\sum_{t=1}^{n} u_t \sin\frac{2\pi t}{\mu} \ldots\ldots\ldots\ldots..(2)$$

Let us further write $R_\mu^2 = A^2 + B^2$ … … … … (3)

$R_\mu{}^2$ is called the **intensity** corresponding to the trial period μ.

Let us consider a simple model, according to which u_t is composed of two components – one periodic with period λ and amplitude a, and the other, an irregular component b_t. Thus, we have,

$$u_t = a\,sin\frac{2\pi t}{\lambda} + b_t \dots\dots\dots\dots(4)$$

We now take a number of trial period μ round about the true period λ, which may be guessed by plotting the data, and calculate R_μ^2 in each case. Finally, we draw a graph plotting R_μ^2 against μ. The diagram, called a **Periodogram**, is a simple device for finding the true cyclical period λ in a time series by equating it to that value of μ for which R_μ^2 attains a maximum.

Harmonic Analysis

Having obtained the true period λ through Periodogram analysis, we may now try to fit a sine-cosine curve trough the residual series u_t. This is known as harmonic analysis. Let the curve to be fitted be

$$u_t = A_0 + Acos\frac{2\pi t}{\lambda} + Bsin\frac{2\pi t}{\lambda}, t = 1,2,\dots,n \dots\dots\dots\dots(1)$$

The constants A_0, A and B may be obtained by the method of least squares by minimizing

$$S = \sum_{t=1}^{n}(u_t - A_0 - Acos\frac{2\pi t}{\lambda} - Bsin\frac{2\pi t}{\lambda})^2.$$

The normal equations are then given by

$$\sum_{t=1}^{n} u_t = nA_0 \Rightarrow \hat{A}_0 = \frac{1}{n}\sum_{t=1}^{n} u_t = \bar{u}$$

$$\sum_{t=1}^{n} u_t cos\frac{2\pi t}{\lambda} = A \times \frac{n}{2} \Rightarrow \hat{A} = \frac{2}{n}\sum_{t=1}^{n} u_t cos\frac{2\pi t}{\lambda}$$

$$\sum_{t=1}^{n} u_t sin\frac{2\pi t}{\lambda} = B \times \frac{n}{2} \Rightarrow \hat{B} = \frac{2}{n}\sum_{t=1}^{n} u_t sin\frac{2\pi t}{\lambda}$$

Now using the estimates $\hat{A}_0,\ \hat{A},\ \hat{B}$ in equation (1) we obtain the periodic component.

Let, $C_t = \hat{A}_0 + \hat{A}\,cos\frac{2\pi t}{\lambda} + \hat{B}\,sin\frac{2\pi t}{\lambda}$ for t=1,2, …, n

Now if we divide the u_t's by C_t we are only left with the Irregular component It. Thus, we decompose the whole time series in its constituent component. Now, we can analyse them separately.

As this is a seasonal data not expected to show any major trend, we would like to check whether the data actually have any trend lying underneath. For that we will perform the Mann-Kendall Non parametric Test for Monotonic Trend using annual data for each dataset.

Mann-Kendall Test for Monotonic Trend

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward (downward) trend means that the variable consistently increases (decreases) through time, but the trend may or may not be linear. The MK test can be used in place of a parametric linear regression analysis, which can be used to test if the slope of the estimated linear regression line is different from zero. The regression analysis requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the MK test, that is, the MK test is a non-parametric (distribution-free) test.

Assumptions

The following assumptions underlie the MK test:

- When no trend is present, the measurements (observations or data) obtained over time are independent and identically distributed. The assumption of independence means that the observations are not serially correlated over time.
- The observations obtained over time are representative of the true conditions at sampling times.
- The sample collection, handling, and measurement methods provide unbiased and representative observations of the underlying populations over time.

There is no requirement that the measurements be normally distributed or that the trend, if present, is linear. The assumption of independence requires that the time between samples be sufficiently large so that there is no correlation between measurements collected at different times.

Calculations

The MK test tests whether to reject the null hypothesis (Ho) and accept the alternative hypothesis (Ha),

Ho: No monotonic trend

Ha: Monotonic trend is present

The initial assumption of the MK test is that the Ho is true and that the data must be convincing beyond a reasonable doubt before Ho is rejected and Ha is accepted.

The MK test is conducted as follows:

1. List the data in the order in which they were collected over time, x_1, x_2, ..., x_n, which denote the measurements obtained at times 1, 2, ..., n, respectively.

2. Determine the sign of all n(n−1)/2 possible differences $x_j - x_k$, where j>k. These differences are

$$x_2 - x_1,\ x_3 - x_1,\ ...,\ x_n - x_1,\ x_3 - x_2,\ x_4 - x_2,\ ...,\ x_n - x_{n-2},\ x_n - x_{n-1}$$

3.Let $\text{sgn}(x_j - x_k)$ be an indicator function that takes on the values 1, 0, or -1 according to the sign of $(x_j - x_k)$, that is,

$$\text{sgn}(x_j - x_k) \quad = 1 \text{ if } (x_j - x_k) > 0$$

$$= 0 \text{ if } (x_j - x_k) = 0,$$

$$= -1 \text{ if } (x_j - x_k) < 0$$

For example, if $(x_j - x_k) > 0$, that means that the observation at time j, denoted by x_j, is greater than the observation at time k, denoted by x_k.

4. Compute

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

which is the number of positive differences minus the number of negative differences. If S is a positive number, observations obtained later in time tend to be larger than observations made earlier. If S is a negative number, then observations made later in time tend to be smaller than observations made earlier.

5. Compute the variance of S as follows:

$$\text{VAR(S)}=\frac{1}{18}\left[n(n-1)(2n+5)-\sum_{p-1}^{g} t_p(t_p-1)(2t_p+5)\right]$$

Where g is the number of tied groups and t_p is the number of observations in the p^{th} group.

6. Compute the MK test statistic, z_{Mk}, as follows:

$$z_{Mk} = \begin{cases} \frac{S-1}{\sqrt{Var(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \frac{S+1}{\sqrt{Var(S)}} & \text{if } S < 0 \end{cases}$$

A positive (negative) value of z_{Mk} indicates that the data tend to increase (decrease) with time.

7. Suppose we want to test the null hypothesis

H_0: No monotonic trend

versus the alternative hypothesis

H_A: Upward monotonic trend

at the Type I error rate α, where $0 < α < 0.05$. (Note that α is the tolerable probability that the MK test will falsely reject the null hypothesis.) Then H_0 is rejected and H_A is accepted if $z_{Mk} \geq \tau_{1-\alpha}$, where $\tau_{1-\alpha}$ is the $100(1-a)^{th}$ percentile of the standard normal distribution.

8. To test H_0 above versus

H_A: Downward monotonic trend

at the Type I error rate α, H_0 is rejected and Ha is accepted if $z_{Mk} \leq -\tau_{1-\alpha}$.

9. To test the H_0 above versus

H_A: Upward or downward monotonic trend

at the Type I error rate α, H_0 is rejected and H_A is accepted if $| z_{Mk} | \geq \tau_{1-\alpha/2}$, where the vertical bars denote absolute value."

10. Otherwise we can compute the p-value,

at the Type I error rate α, H_0 is rejected and H_a is accepted if p-value< α

where, α is the level of significance.

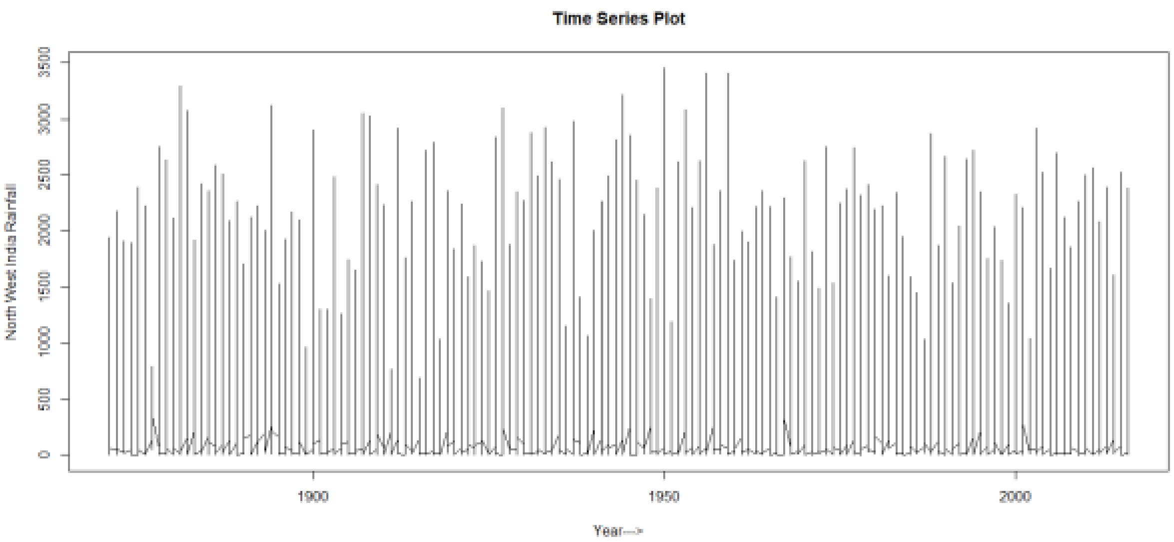
Results and Discussion

1.North West India-

The North West India consists of Haryana, Punjab, West Rajasthan, East Rajasthan, Gujrat, Saurashtra, Kutch & Diu with 634272 sq.km contributing 22.02 % of the whole study area. This area receives on an average 400-700 mm of rainfall annually which is much less than the other areas concerned. This area includes the dessert area of Rajasthan which receives very less amount of rainfall. North west India receives monsoon rainfall from July to September through the monsoon winds from Arabian sea and Bay of Bengal.

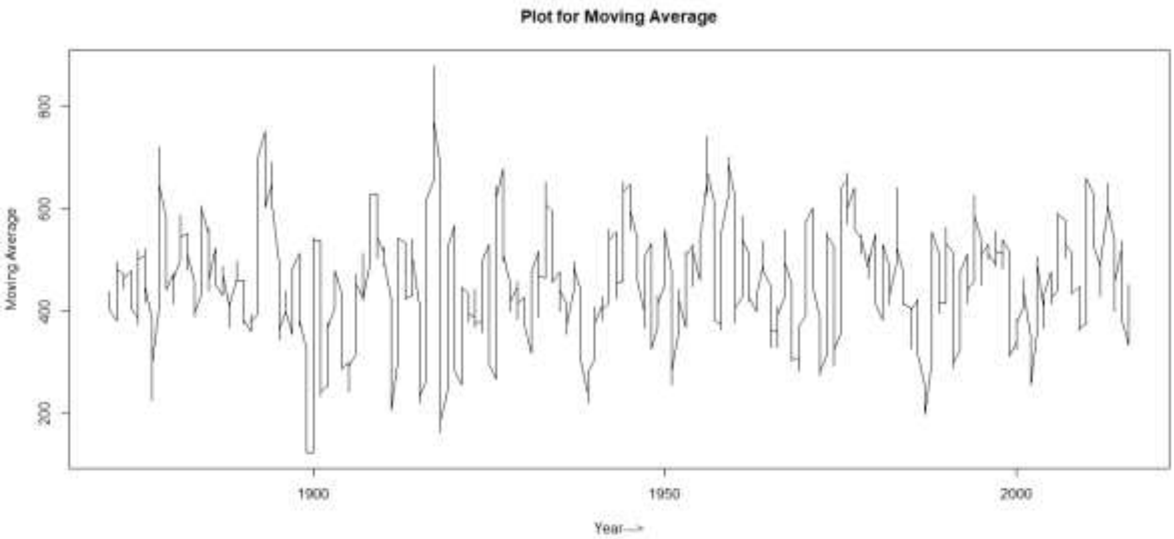
During winter, low pressure areas forming over Mediterranean Sea moves east and precipitates over Pakistan and North west India. Such low-pressure areas are called as Western disturbances. These precipitates rain to the plains and snow to the western Himalayas. It is good for wheat crop and other rabi crops in Punjab and Haryana.

Time Series Plot:



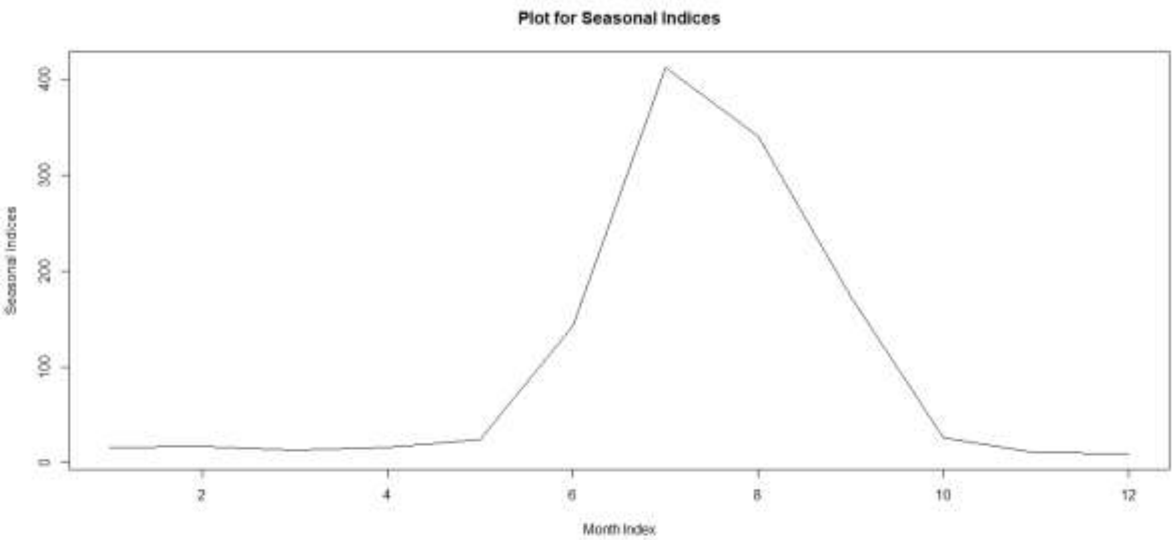
The above Time Series Plot doesn't indicate any clear trend but seems to have seasonal and cyclical fluctuations. So, we are further interested in decomposing the time series data.

Plot of Moving Average:



The plot of Moving Averages shows clear cyclical pattern. That's why proceed for Harmonic Analysis.

Plot of Seasonal Indices:

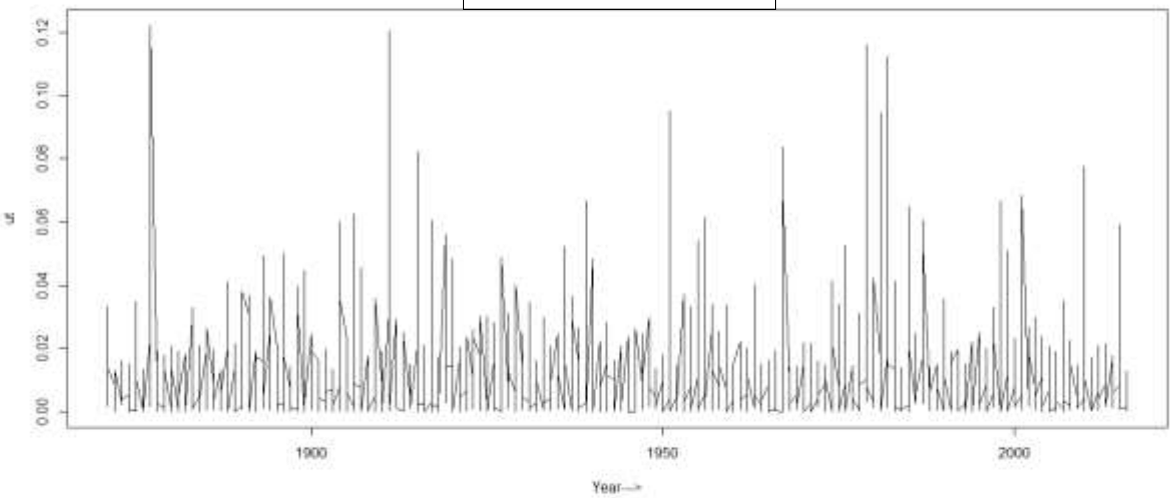


The plot for the Adjusted Seasonal Indices shows it maximum in the month of July. Moreover, it implies very small amount of rainfall in the area from Jan to May and then there is a steep increase in June and maximises in July and slow decrease after Aug. The seasonal indices for June -Sept shows a comparative higher value than others. Hence, monsoon in North West India lasts for these four months.

Periodogram Analysis:

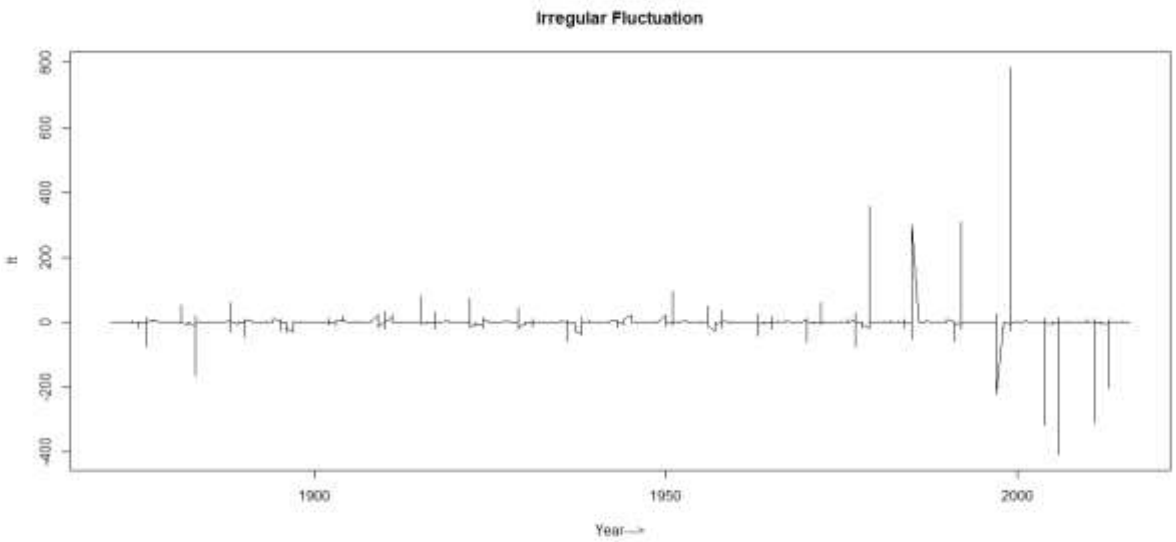
Let u_t be the residual series after eliminating the trend and seasonal component.

Plot of u_t



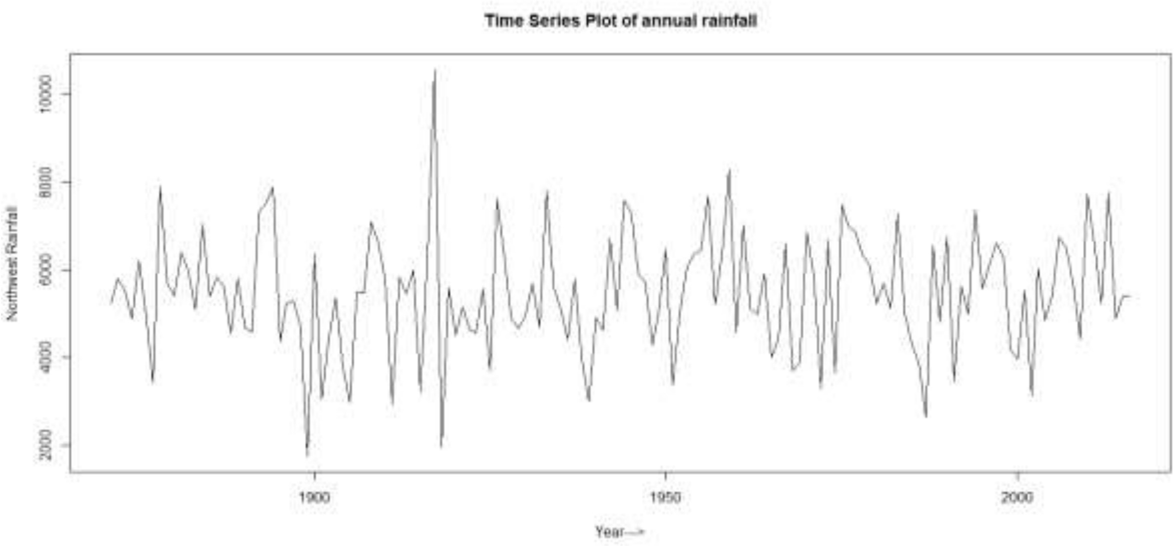
Eliminating the seasonality, we have only the cyclical and irregular component in the Periodogram. The diagram helps us in getting an idea of true period in the Periodic Analysis of the time series data. Here, in the analysis we get the true period 82 months i.e., it takes almost 7 years to complete a cycle. Now, we would eliminate the cyclical components to get the irregular components which remains unexplained.

Plot of Irregular Component:



In the plot of irregular components, it shows much less irregularity till 1970's. After that the irregularity in rainfall has increased. We observe that, the irregular fluctuations increase with time implying more unusual behaviour in the data. More irregularity in rainfall may be looked upon as an alarming sign of climate change in that area.

Mann-Kendall Test:



This is the plot of Annual Rainfall of North West India over the years concerned. As we can't determine the direction of trend from graph, we take the hypothesis of Mann- Kendall Test as-

H_0: No monotonic trend is present in the Time Series data

$$vs$$

H_1: Upward or Downward monotonic trend is present in the Time Series data

Let, x_1, x_2, ..., x_n, denote the annual rainfall in North West India obtained at times 1, 2, ..., n, respectively.

Let us define, S= $\sum_{k-1}^{n-1}\sum_{j-k+1}^{n} \text{sgn}(x_j - x_k)$

And variance of S comes out to be,

$$\text{VAR(S)}=\frac{1}{18}\left[n(n-1)(2\text{n}+5) - \sum_{p-1}^{g} t_p(t_p-1)(2t_p+5)\right]$$

Where g is the number of tied groups and t_p is the number of observations in the p^{th} group.

Test statistic-

The MK test statistic is defined as-

$$z_{Mk} \quad = \frac{S-1}{\sqrt{Var(S)}} \quad \text{if S} > 0$$

$$= 0 \quad \text{if S} = 0$$

$$= \frac{S+1}{\sqrt{Var(S)}} \quad \text{if S} < 0$$

Critical region-

We reject Ho in favour of H_1 , if $|z_{Mk}| \geq \tau_{1-\alpha/2}$, where, α is the level of significance.

Calculation-

n= 146

α= 0.05

Using the “trend” package in the statistical software R we get,

S= 400

z_{Mk}= 0.6751

$\tau_{1-\alpha/2}$= $\tau_{0.975}$= 1.959964

Decision-

Since, $|z_{Mk}|$= 0.6751 < $\tau_{0.975}$= 1.959964,

Hence, here Ho is accepted 0.05 level of significance.

Conclusion-

In the light of the given data, it seems that there is no monotonic trend present in the Annual Rainfall of North West India.

2.North East India-

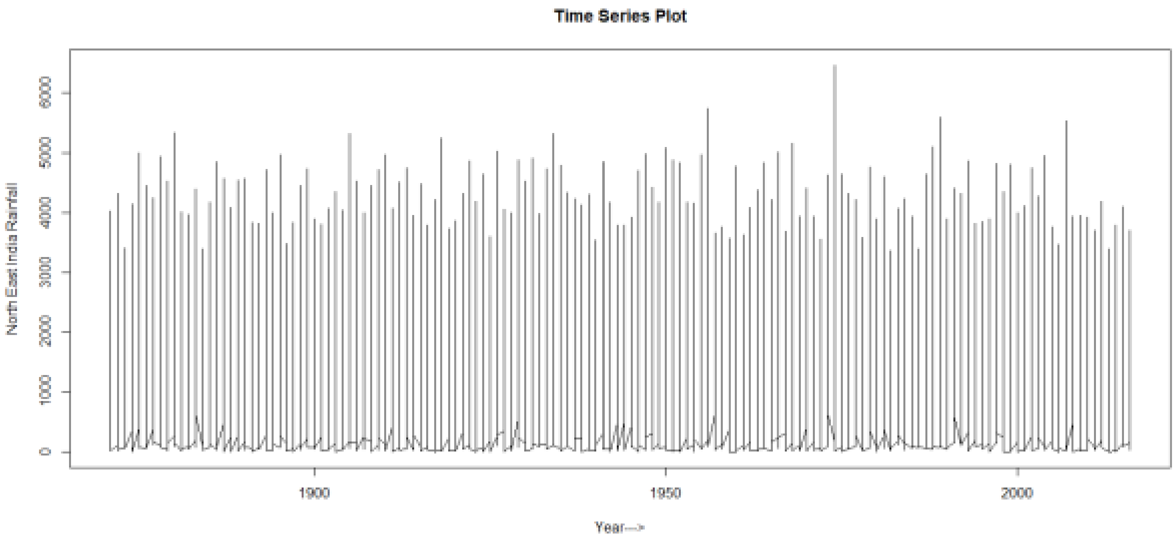
The North East India consists of Assam & Meghalaya, Nagaland, Manipur, Mizoram & Tripura, Sub-Himalayan West Bengal & Sikkim, Gangetic West Bengal with 267444sq.km contributing only 9.29 % of the whole study area. This area receives on an average 1800-2000 mm of rainfall annually which is much higher than the other areas concerned.

Lying very close to the Tropics, North-East India displays, to a large extent, the character of tropical climate, especially in the valleys. The region has a monsoon climate with heavy to very heavy rains, confined within four summer months from June to September. The southwest monsoon is the main source of rain, and July is the rainiest month. There are three seasons in the area, winter, summer and rainy season, though rainy season, as in the rest of India, coincides with summer months. There is a climatic contrast between the valleys and the mountainous region. While the mean January temperature in the valley region of Assam is around 16 °C, the temperatures in the mountainous region of Arunachal Pradesh and Nagaland hover around a maximum of 14 °C and a sub-zero minimum temperature. The summer temperatures in the plains vary between 30 and 33 °C, while the hills have a mean summer temperature of around 20 °C with a mean minimum of 15 °C. Nowhere in the region, there is heavy snow except in the higher parts of Arunachal Pradesh. No part of North-East India receives rainfall below 1,000 mm. Shillong-Plateau with its southern limit marked by a 1,200-m-high scarp overlooking the Bangladesh plain receives very heavy rains. Mawsynram, situated on the top of the scarp, receives a mean annual rainfall of 11,465 mm. The average rainfall of Brahmaputra valley is around 2,000 mm with local variations. Guwahati, being in the rain shadow of the Meghalaya plateau, receives only 1,717 mm of rain. About 90 % of the rain is received during the southwest summer monsoon, and June is by far the rainiest month. The hilly areas of the region receive 2,000–3,000 mm of rain, though places like Kohima in

Nagaland and Imphal in Manipur, because of their being in the shadow of the mountains, receive less than 2,000 mm of rains.

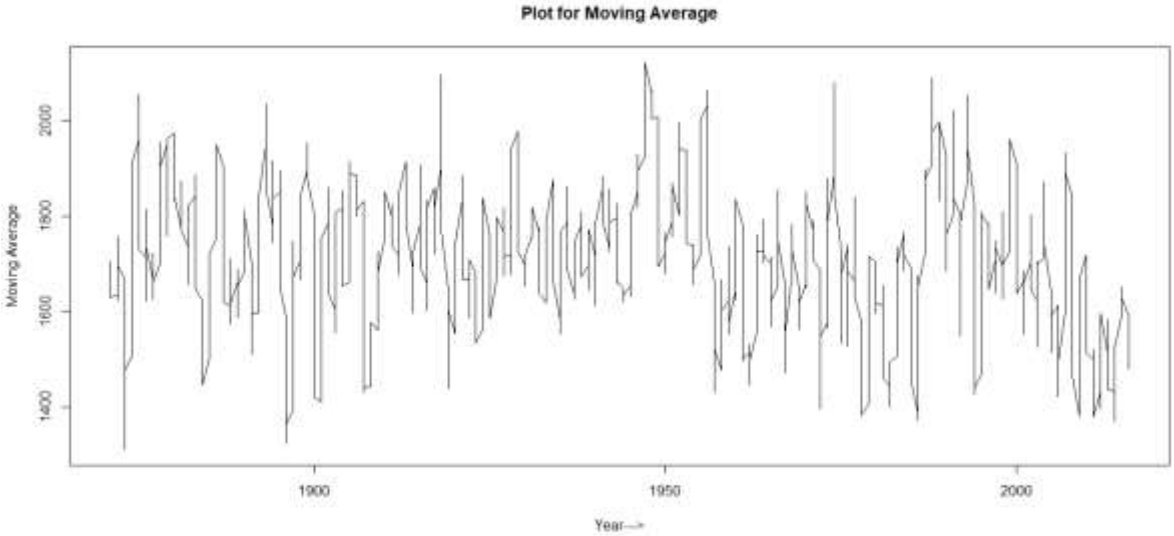
These heavy amount of rainfall in these regions resulted dense rainforest as we know the northeast states of the country poses higher percentages of forest cover. Timely monsoon and adequate amount of rainfall has ensured West Bengal to be enriched with higher production of various agricultural products.

Time Series Plot:



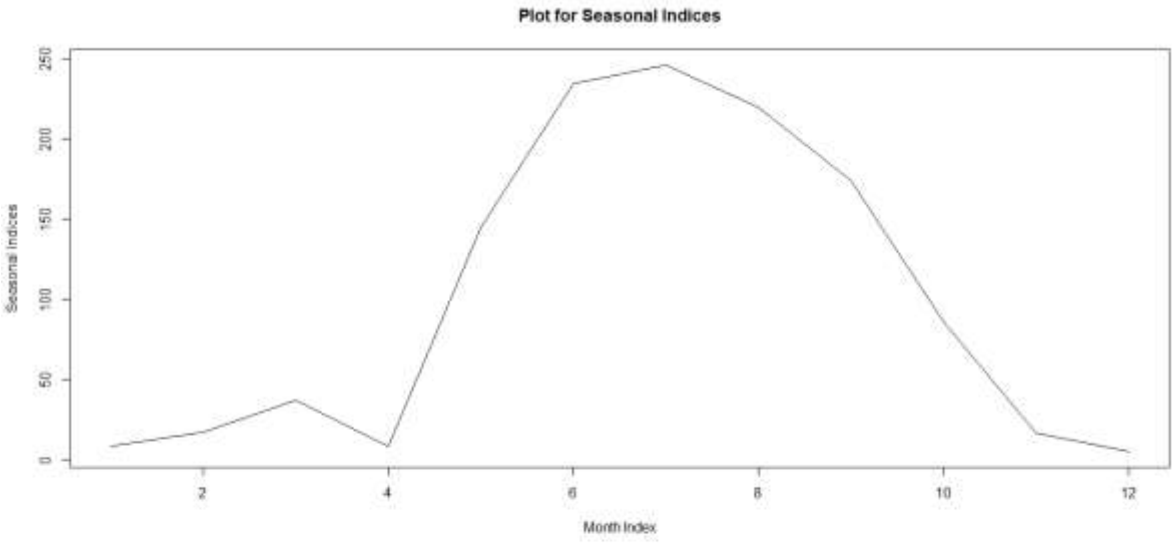
The above Time Series Plot doesn’t indicate any clear trend but seems to have seasonal and cyclical fluctuations. So, we are further interested in decomposing the time series data.

Plot of Moving Average:



The plot of Moving Averages shows clear cyclical pattern. That's why proceed for Harmonic Analysis.

Plot of Seasonal Indices:

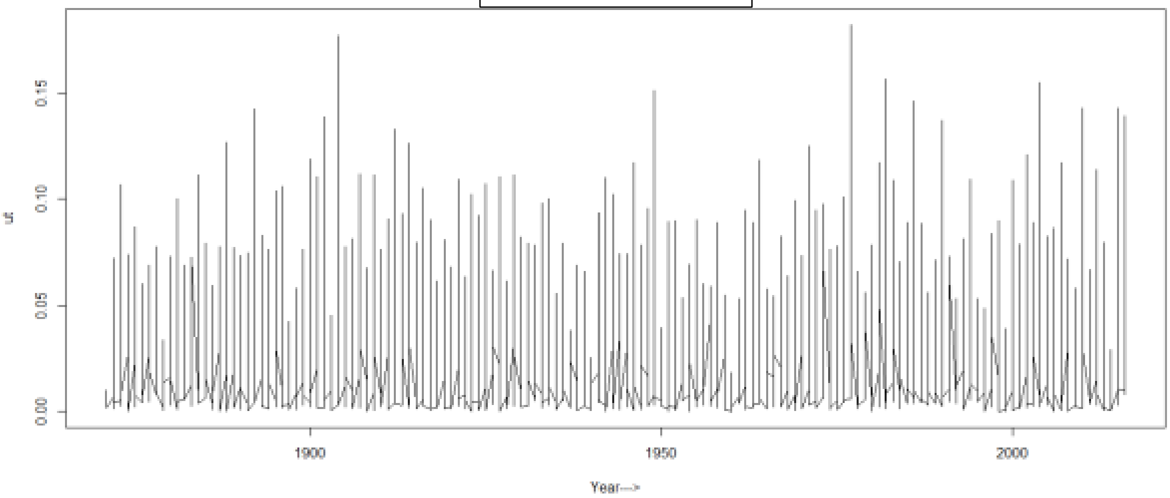


The plot for the Adjusted Seasonal Indices shows it maximum in the month of July. Moreover, it implies comparatively less amount of rainfall in the winter from Nov to Feb and increase in March then decrease in April and then there is a steep increase in May and maximises in July and decrease in Oct. The seasonal indices for May -Oct shows a comparative higher value than others i.e., North East India experiences higher rainfall for these long six months.

Periodogram Analysis:

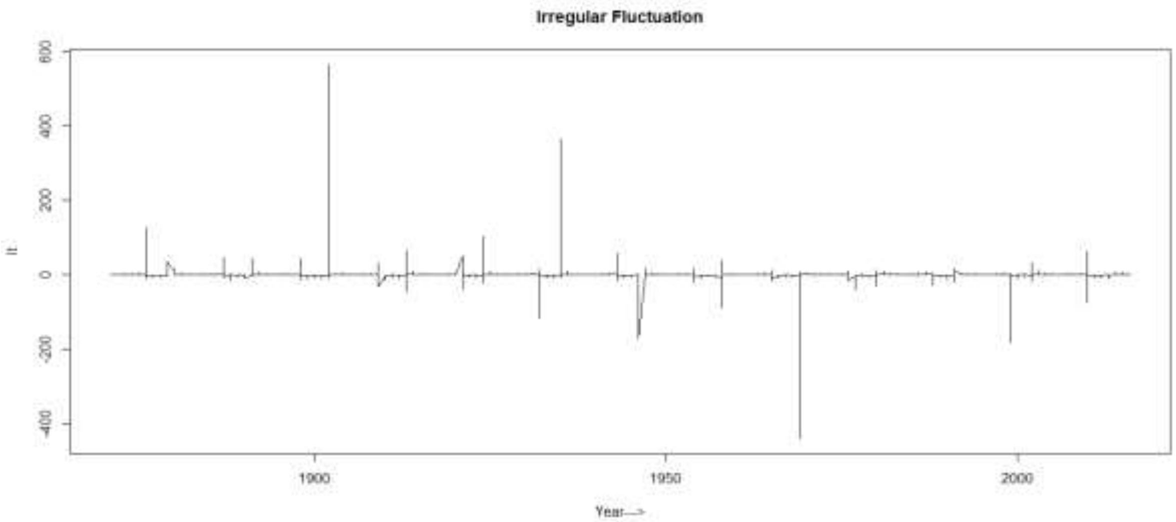
Let u_t be the residual series after eliminating the trend and seasonal component.

Plot of u_t



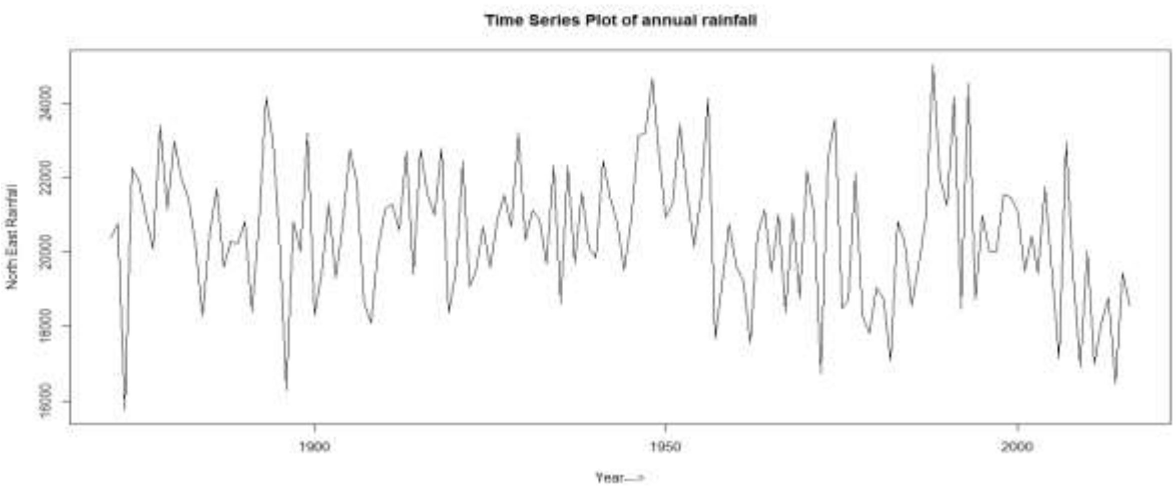
Eliminating the seasonality, we have only the cyclical and irregular component in the Periodogram. The diagram helps us in getting an idea of true period in the Periodic Analysis of the time series data. Here, in the analysis we get the true period 134 months i.e., it takes almost 11 years to complete a cycle. Now, we would eliminate the cyclical components to get the irregular components which remains unexplained.

Plot of Irregular Component:



In the plot we observe the irregular components to be more or less randomly scattered. That implies we have been able to explain the variability of the time series data by decomposition of it. Moreover, it is not affected much by external effects (for e.g., climate change, global warming). These North East India has also been reported less amount of pollution than the rest of the country. In fact, Sikkim has been declared as the only organic state of India.

Mann-Kendall Test:



This is the plot of Annual Rainfall of North East India over the years concerned. In the last portion of graph (from 1980's), we suspect downward monotonic trend in the data, so we take the hypothesis of Mann- Kendall Test as-

H_0: No monotonic trend is present in the Time Series data

$$vs$$

H_1: Downward monotonic trend is present in the Time Series data

Let, x_1, x_2, ..., x_n, denote the annual rainfall in North East India obtained at times 1, 2, ..., n, respectively.

Let us define, S= $\sum_{k-1}^{n-1}\sum_{j-k+1}^{n} \text{sgn}(x_j - x_k)$

And variance of S comes out to be,

$$\text{VAR(S)}=\frac{1}{18}[n(n-1)(2n+5) - \sum_{p-1}^{g} t_p(t_p - 1)(2t_p + 5)]$$

Where g is the number of tied groups and t_p is the number of observations in the p^{th} group.

Test statistic-

The MK test statistic is defined as-

$$z_{Mk} = \frac{S-1}{\sqrt{Var(S)}} \quad \text{if } S> 0$$

$$= 0 \quad \text{if } S = 0$$

$$= \frac{S+1}{\sqrt{Var(S)}} \quad \text{if } S< 0$$

Critical region-

We reject Ho in favour of H_1 , if z_{Mk}≤ -$\tau_{1-\alpha}$, where, α is the level of significance.

Calculation-

n= 146

α= 0.05

Using the "trend" package in the statistical software R we get,

S= -1302

z_{Mk}= -2.2013

$\tau_{1-\alpha}$= $\tau_{0.95}$= 1.644854

Decision-

Since, z_{Mk}= -2.2013 < - $\tau_{0.95}$= -1.644854,

Hence, here Ho is rejected at 0.05 level of significance.

Conclusion-

In the light of the given data, it seems that monotonic decreasing trend is present in the Annual Rainfall of North East India.

Here, the value of S comes out to be negative which implies the observations made later in time tend to be smaller than observations made earlier. That is the data seems to have a decreasing trend.

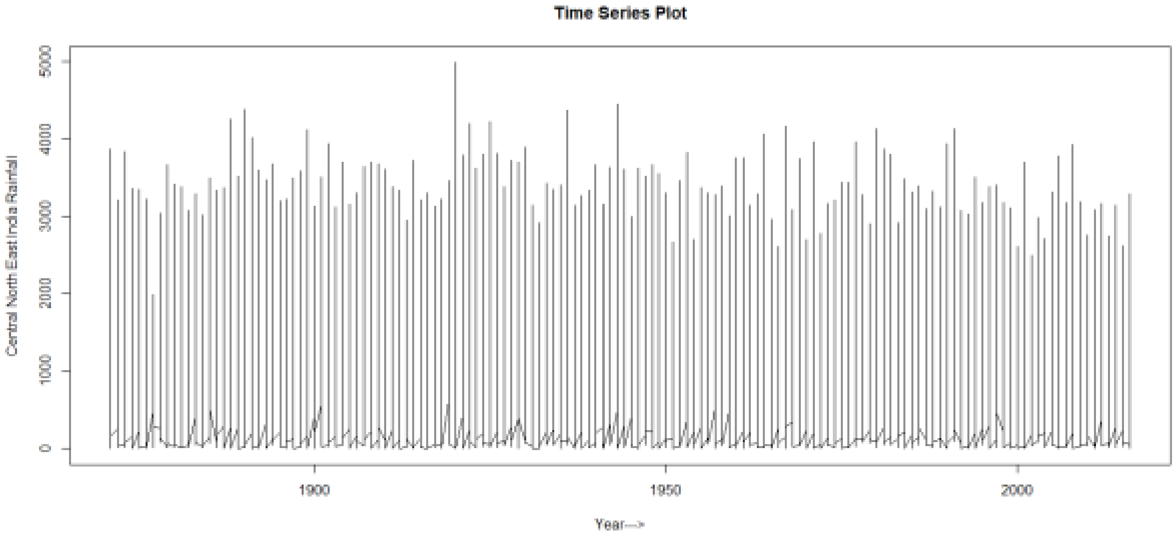
3.Central North East India-

The Central North East India consists of Odisha, Jharkhand, Bihar, East Uttar Pradesh, West Uttar Pradesh with 573006 sq.km contributing 19.89% of the whole study area. This area receives on an average 1000-1500 mm of rainfall annually.

The climate of this region is primarily defined as humid subtropical with dry winter. Alternatively, some authors refer to it as *tropical monsoon*. Variations do exist in different parts of the large area, however the uniformity of the vast Indo-Gangetic Plain forming bulk of the area gives a predominantly single climatic pattern with minor regional variations. With temperatures fluctuating anywhere from 0 °C to 50 °C in several parts of the region and cyclical droughts and floods due to unpredictable rains, the summers are extremely hot, winters cold and the rainy season can be either very wet or very dry.

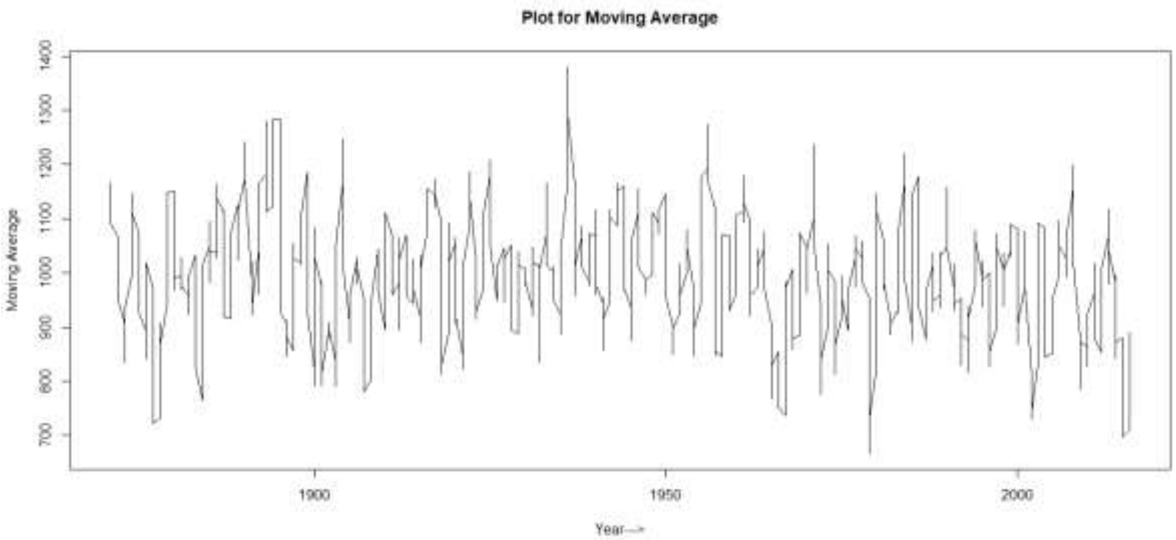
Primarily a summer phenomenon, the Bay of Bengal branch of the Indian Monsoon is the major bearer of rain in most parts of central north east India. It is the South-West Monsoon which brings most of the rain here, although rain due to the *western disturbances* and North-East Monsoon also contribute small quantities towards the overall precipitation. Given the concentration of most of this rainfall in the 4 months of Monsoon period, excess rain can lead to floods and shortage to droughts. As such these two phenomena of floods and droughts are a common recurrence in the region.

Time Series Plot:



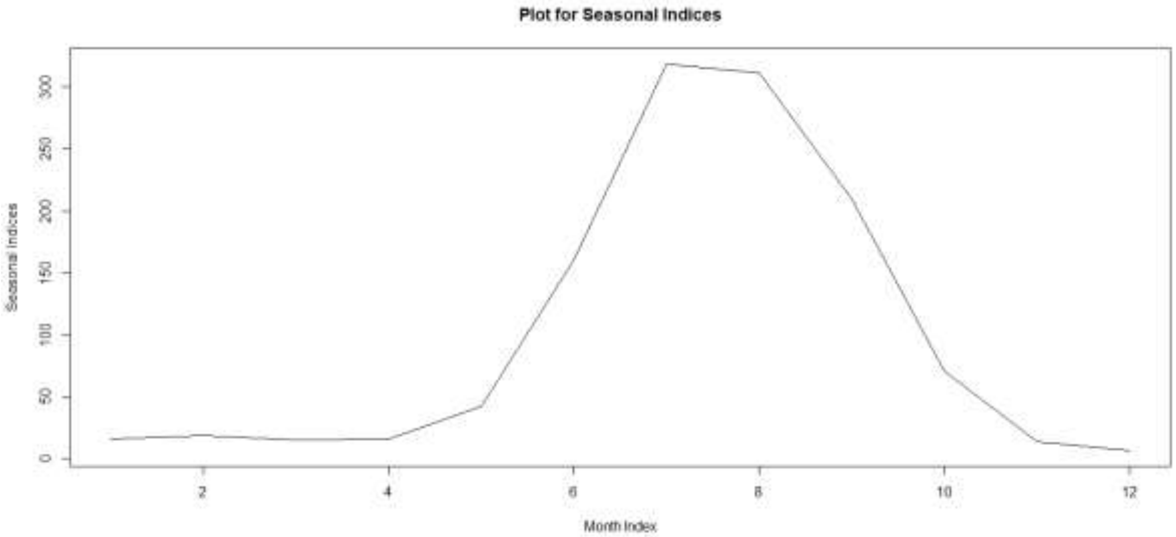
The above Time Series Plot doesn't indicate any clear trend but seems to have seasonal and cyclical fluctuations. So, we are further interested in decomposing the time series data.

Plot of Moving Average:



The plot of Moving Averages shows clear cyclical pattern. That’s why proceed for Harmonic Analysis.

Plot of Seasonal Indices:

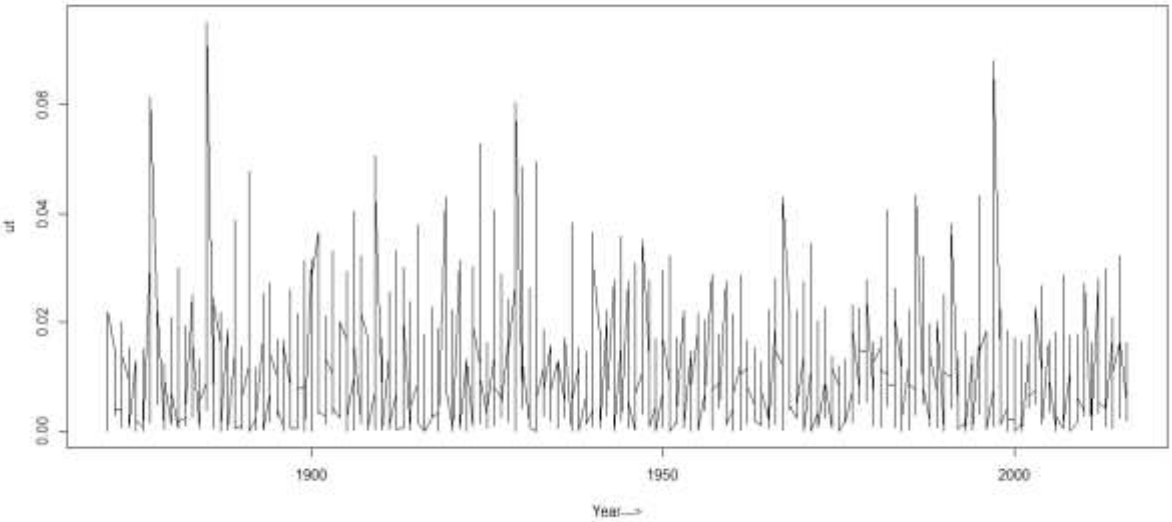


The plot for the Adjusted Seasonal Indices shows it maximum in the month of July. Moreover, it implies very small amount of rainfall in the area from Jan to May and then there is a steep increase in June and maximises in July and slow decrease after Aug. The seasonal indices for June -Sept shows a comparative higher value than others. Hence, monsoon in Central North East India lasts for these four months.

Periodogram Analysis:

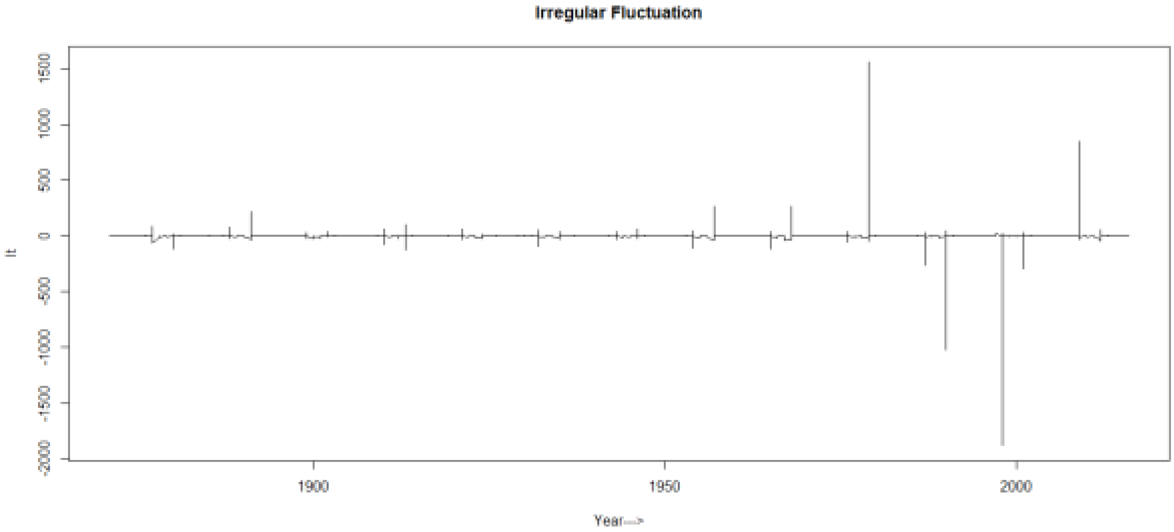
Let u_t be the residual series after eliminating the trend and seasonal component.

Plot of u_t



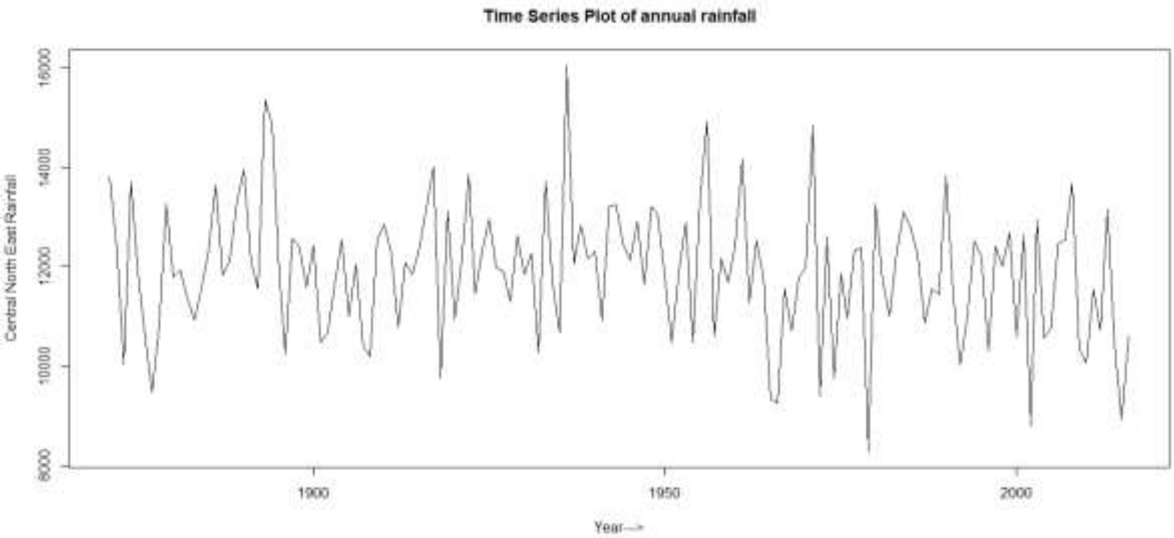
Eliminating the seasonality, we have only the cyclical and irregular component in the Periodogram. The diagram helps us in getting an idea of true period in the Periodic Analysis of the time series data. Here, in the analysis we get the true period 132 months i.e., it takes almost 11 years to complete a cycle. Now, we would eliminate the cyclical components to get the irregular components which remains unexplained.

Plot of Irregular Component:



In the plot of irregular components, it shows much less irregularity till 1970's. After that the irregularity in rainfall has increased. We observe that, the irregular fluctuations increase with time implying more unusual behaviour in the data. More irregularity in rainfall may be looked upon as an alarming sign of climate change in that area.

Mann-Kendall Test:



This is the plot of Annual Rainfall of Central North East India over the years concerned. As we can't determine the direction of trend from graph, we take the hypothesis of Mann- Kendall Test as-

H_0: No monotonic trend is present in the Time Series data

$$vs$$

H_1: Upward or Downward monotonic trend is present in the Time Series data

Let, x_1, x_2, ..., x_n, denote the annual rainfall in Central North East India obtained at times 1, 2, ..., n, respectively.

Let us define, S= $\sum_{k-1}^{n-1}\sum_{j-k+1}^{n} \text{sgn}(x_j - x_k)$

And variance of S comes out to be,

$$\text{VAR(S)}=\frac{1}{18}[n(n-1)(2\text{n}+5) - \sum_{p-1}^{g} t_p(t_p-1)(2t_p+5)]$$

Where g is the number of tied groups and t_p is the number of observations in the p^{th} group.

Test statistic-

The MK test statistic is defined as-

$$z_{Mk} = \frac{S-1}{\sqrt{Var(S)}} \quad \text{if } S> 0$$

$$= 0 \quad \text{if } S = 0$$

$$= \frac{S+1}{\sqrt{Var(S)}} \quad \text{if } S< 0$$

Critical region-

We reject Ho in favour of H_1, if $|z_{Mk}| \geq \tau_{1-\alpha/2}$, where, α is the level of significance.

Calculation-

n= 146

α= 0.05

Using the “trend” package in the statistical software R we get,

S= -907

z_{Mk}= -1.5329

$\tau_{1-\alpha/2}$= $\tau_{0.975}$= 1.959964

Decision-

Since, $|z_{Mk}|$= 1.5329 < $\tau_{0.975}$= 1.959964,

Hence, here Ho is accepted 0.05 level of significance.

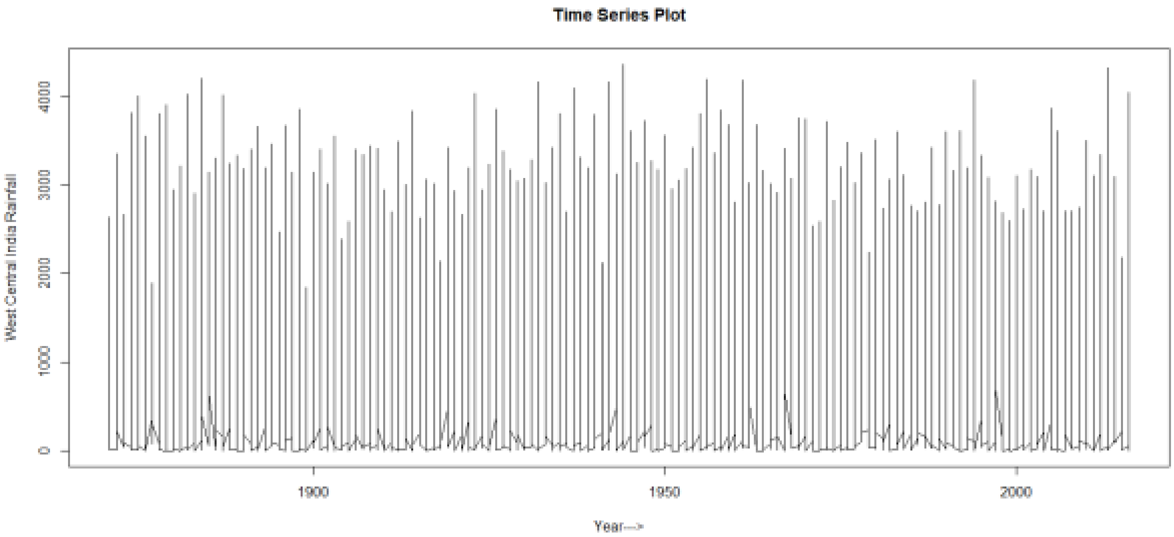
Conclusion-

In the light of the given data, it seems that there is no monotonic trend present in the Annual Rainfall of Central North East India.

4.West Central India-

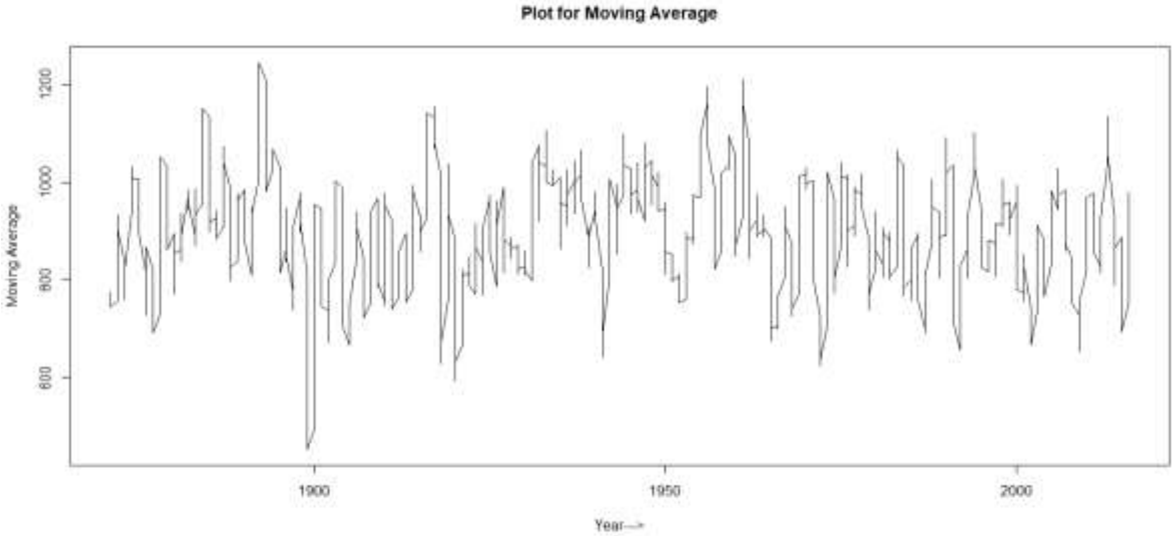
The West Central India consists of West Madhya Pradesh, East Madhya Pradesh, Konkan & Goa, Madhya Maharashtra, Marathwada, Vidarbha, Chhattisgarh, Telangana, North Interior Karnataka with 962694 sq.km contributing 33.42 % of the whole study area. This area receives on an average 900-1300 mm of rainfall annually. The area is drought-prone, as it tends to have less reliable rainfall due to sporadic lateness or failure of the southwest monsoon.

Time Series Plot:



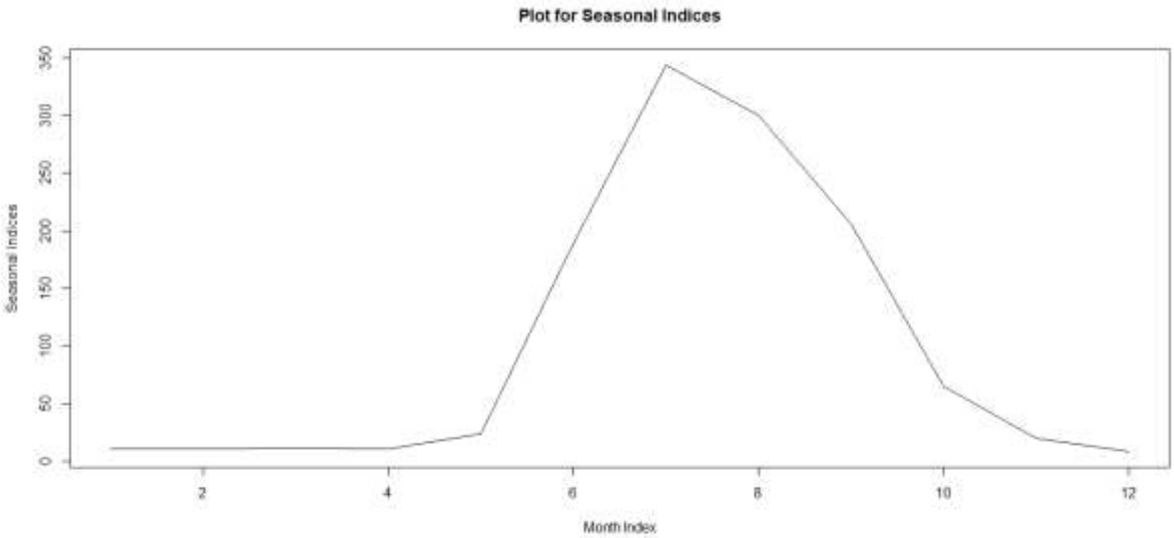
The above Time Series Plot doesn't indicate any clear trend but seems to have seasonal and cyclical fluctuations. So, we are further interested in decomposing the time series data.

Plot of Moving Average:



The plot of Moving Averages shows clear cyclical pattern. That's why proceed for Harmonic Analysis.

Plot of Seasonal Indices:

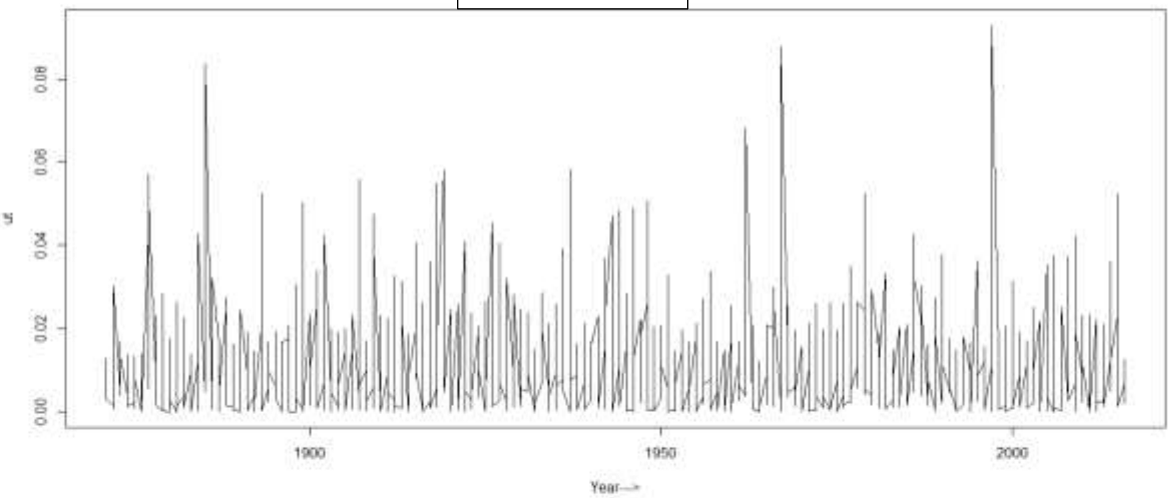


The plot for the Adjusted Seasonal Indices shows it maximum in the month of July. Moreover, it implies very small amount of rainfall in the area from Jan to May and then there is a steep increase in June and maximises in July and slow decrease after Aug. The seasonal indices for June -Sept shows a comparative higher value than others. Hence, monsoon in West Central India lasts for these four months.

Periodogram Analysis:

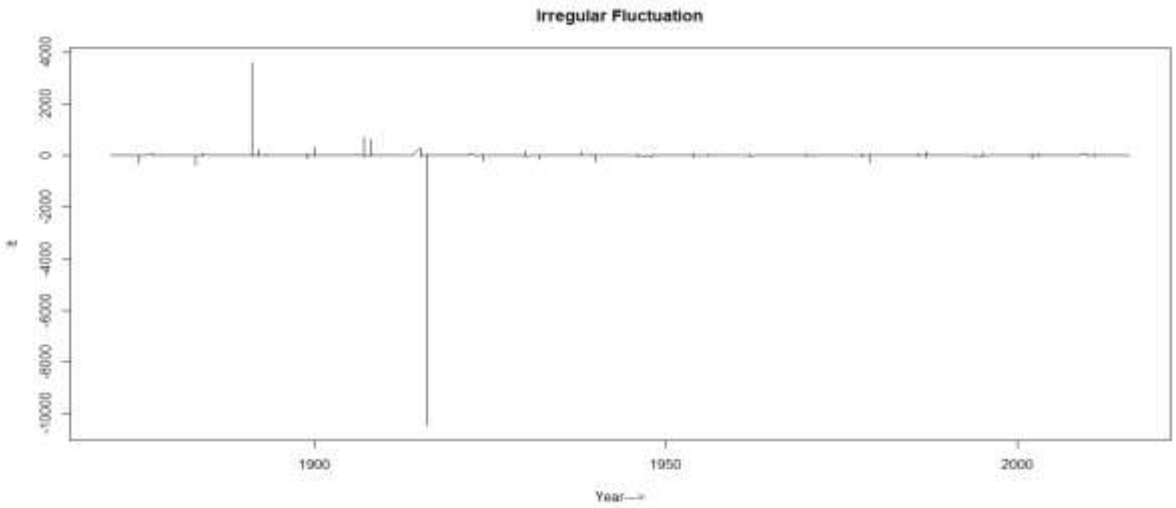
Let u_t be the residual series after eliminating the trend and seasonal component.

Plot of u_t



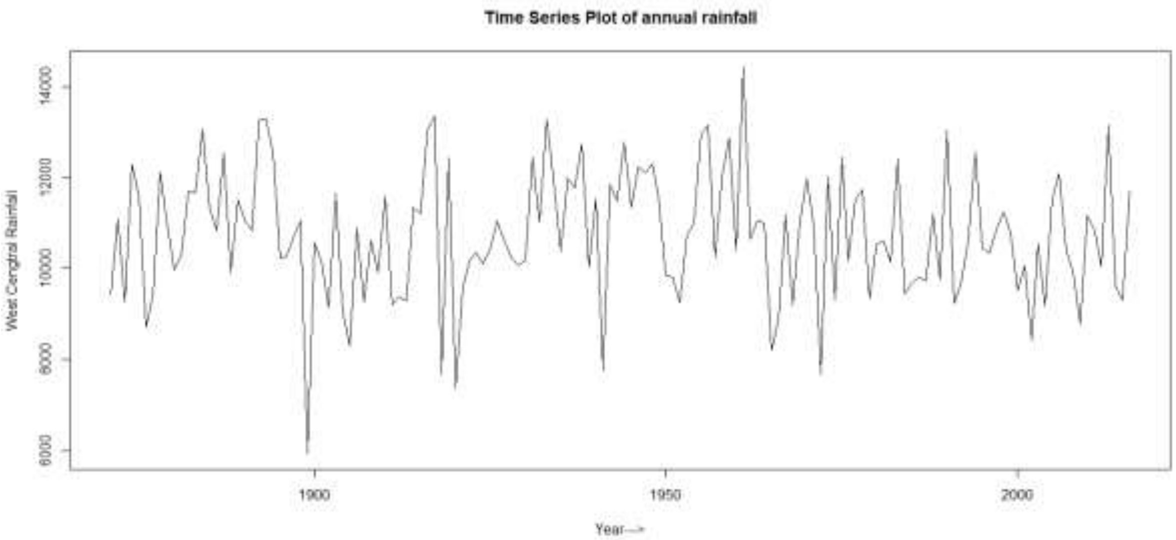
Eliminating the seasonality, we have only the cyclical and irregular component in the Periodogram. The diagram helps us in getting an idea of true period in the Periodic Analysis of the time series data. Here, in the analysis we get the true period 95 months i.e., it takes almost 8 years to complete a cycle. Now, we would eliminate the cyclical components to get the irregular components which remains unexplained.

Plot of Irregular Component:



In the plot we observe the irregular components to be more or less randomly scattered except two extreme values. Probably these two time points have shown some unusual amount of rainfall. Otherwise, we have been able to explain the variability of the time series data by decomposition of it. Moreover, it is not affected much by external effects (for e.g., climate change, global warming).

Mann-Kendall Test:



This is the plot of Annual Rainfall of West Central India over the years concerned. As we can't determine the direction of trend from graph, we take the hypothesis of Mann- Kendall Test as-

H_0: No monotonic trend is present in the Time Series data

$$vs$$

H_1: Upward or Downward monotonic trend is present in the Time Series data

Let, x_1, x_2, ..., x_n, denote the annual rainfall in West Central India obtained at times 1, 2, ..., n, respectively.

Let us define, S= $\sum_{k-1}^{n-1}\sum_{j-k+1}^{n} \operatorname{sgn}(x_j - x_k)$

And variance of S comes out to be,

$$\text{VAR(S)}=\frac{1}{18}[n(n-1)(2n+5) - \sum_{p-1}^{g} t_p(t_p-1)(2t_p+5)]$$

Where g is the number of tied groups and t_p is the number of observations in the p^{th} group.

Test statistic-

The MK test statistic is defined as-

$$z_{Mk} \quad = \frac{S-1}{\sqrt{Var(S)}} \quad \text{if } S>0$$

$$= 0 \quad \text{if } S=0$$

$$= \frac{S+1}{\sqrt{Var(S)}} \quad \text{if } S<0$$

Critical region-

We reject Ho in favour of H_1 , if $|z_{Mk}|\geq\tau_{1-\alpha/2}$, where, α is the level of significance.

Calculation-

n= 146

α= 0.05

Using the "trend" package in the statistical software R we get,

S= -504

z_{Mk}= -0.85107

$\tau_{1-\alpha/2}$= $\tau_{0.975}$= 1.959964

Decision-

Since, $|z_{Mk}|$= 0.85107 < $\tau_{0.975}$= 1.959964,

Hence, here Ho is accepted 0.05 level of significance.

Conclusion-

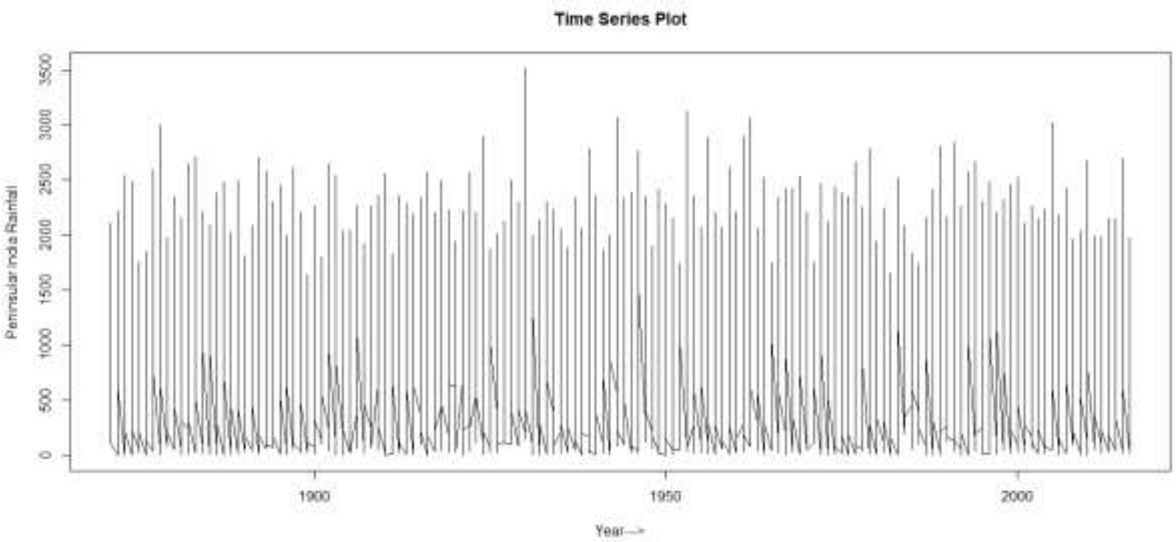
In the light of the given data, it seems that there is no monotonic trend present in the Annual Rainfall of West Central India.

5.Peninsular India-

The Peninsular India consists of Coastal Andhra Pradesh, Rayalaseema, Tamil Nadu & Pondicherry, Coastal Karnataka, South Interior Karnataka, Kerala with 442908 sq.km contributing 15.38% of the whole study area. This area receives on an average 1000-1400 mm of rainfall annually.

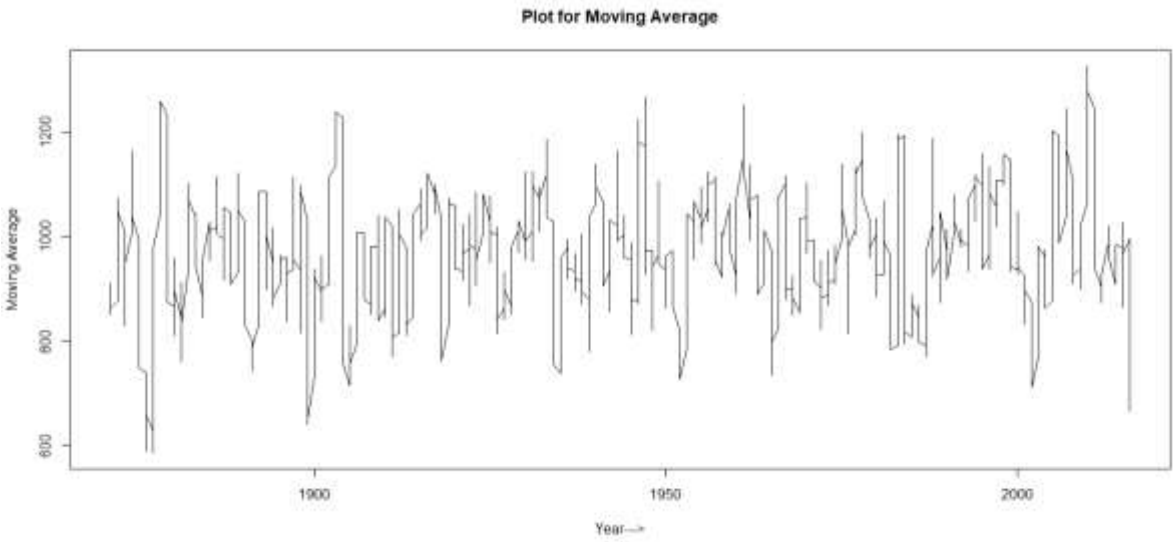
Peninsular India's climate is affected by two seasonal winds - the northeast monsoon and the southwest monsoon. The north-east monsoon, commonly known as winter monsoon blows from land to sea, whereas south-west monsoon, known as summer monsoon blows from sea to land after crossing the Indian Ocean, the Arabian Sea, and the Bay of Bengal. The south-west monsoon brings most of the rainfall during a year in the country.

Time Series Plot:



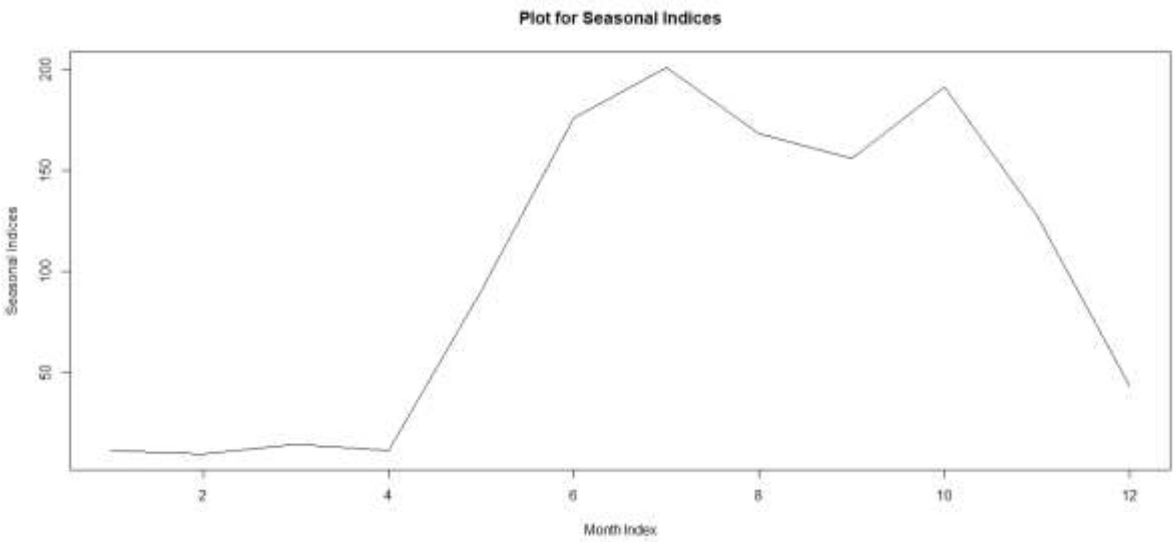
The above Time Series Plot doesn't indicate any clear trend but seems to have seasonal and cyclical fluctuations. So, we are further interested in decomposing the time series data.

Plot of Moving Average:



The plot of Moving Averages shows clear cyclical pattern. That's why proceed for Harmonic Analysis.

Plot of Seasonal Indices:

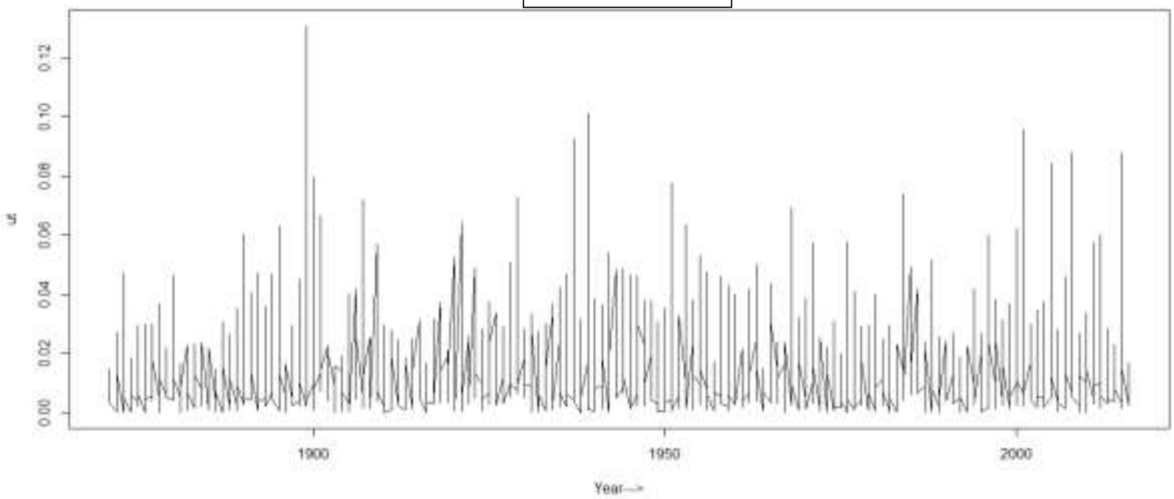


The plot for the Adjusted Seasonal Indices shows it maximum in the month of July. Moreover, it implies very small amount of rainfall in the area from Jan to Apr and then there is a steep increase in May and maximises in July and slow decrease in Aug-Sept and again increase in Oct and decrease afterwards. The seasonal indices for May -Dec shows a comparative higher value than others. Moreover, it has two peak values one in July and the other in Oct. This is how the rainfall pattern in Peninsular India is different from the other parts of country as this area experiences double monsoon.

Periodogram Analysis:

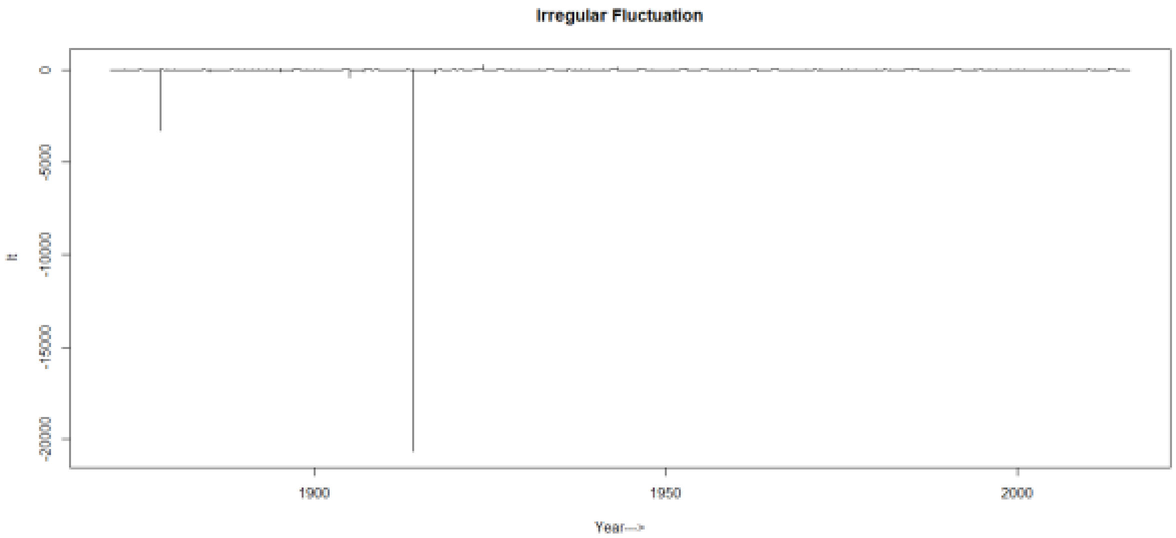
Let u_t be the residual series after eliminating the trend and seasonal component.

Plot of u_t



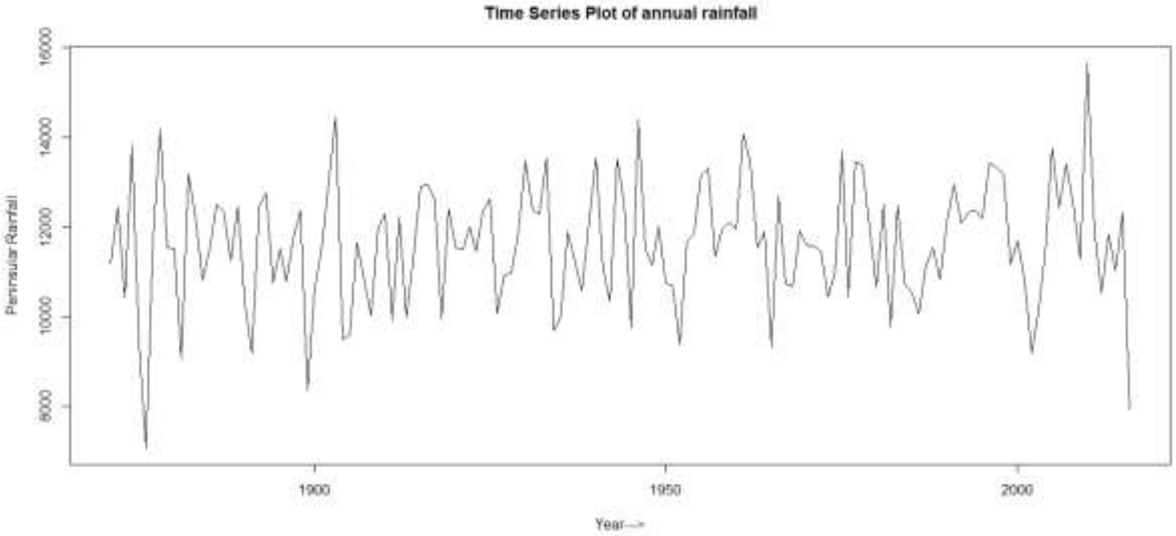
Eliminating the seasonality, we have only the cyclical and irregular component in the Periodogram. The diagram helps us in getting an idea of true period in the Periodic Analysis of the time series data. Here, in the analysis we get the true period 116 months i.e., it takes almost 10 years to complete a cycle. Now, we would eliminate the cyclical components to get the irregular components which remains unexplained.

Plot of Irregular Component:



In the plot we observe the irregular components to be more or less randomly scattered except two extreme values. Probably these two time points have shown some unusual amount of rainfall. Otherwise, we have been able to explain the variability of the time series data by decomposition of it. Moreover, it is not affected much by external effects (for e.g., climate change, global warming).

Mann-Kendall Test:



This is the plot of Annual Rainfall of Peninsular India over the years concerned. As we can't determine the direction of trend from graph, we take the hypothesis of Mann- Kendall Test as-

H_0: No monotonic trend is present in the Time Series data

$$vs$$

H_1: Upward or Downward monotonic trend is present in the Time Series data

Let, x_1, x_2, ..., x_n, denote the annual rainfall in Peninsular India obtained at times 1, 2, ..., n, respectively.

Let us define, S= $\sum_{k-1}^{n-1} \sum_{j-k+1}^{n} \text{sgn}(x_j - x_k)$

And variance of S comes out to be, VAR(S)= $\frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p-1}^{g} t_p(t_p-1)(2t_p+5)\right]$

Where g is the number of tied groups and t_p is the number of observations in the p^{th} group.

Test statistic-

The MK test statistic is defined as-

$$z_{Mk} = \frac{S-1}{\sqrt{Var(S)}} \quad \text{if } S> 0$$

$$= 0 \quad \text{if } S = 0$$

$$= \frac{S+1}{\sqrt{Var(S)}} \quad \text{if } S< 0$$

Critical region-

We reject Ho in favour of H_1 , if $|z_{Mk}| \geq \tau_{1-\alpha/2}$, where, α is the level of significance.

Calculation-

n= 146

α= 0.05

Using the “trend” package in the statistical software R we get,

S= 599

z_{Mk}= 1.0118

$\tau_{1-\alpha/2} = \tau_{0.975}$= 1.959964

Decision-

Since, $|z_{Mk}|$= 1.0118 < $\tau_{0.975}$= 1.959964,

Hence, here Ho is accepted 0.05 level of significance.

Conclusion-

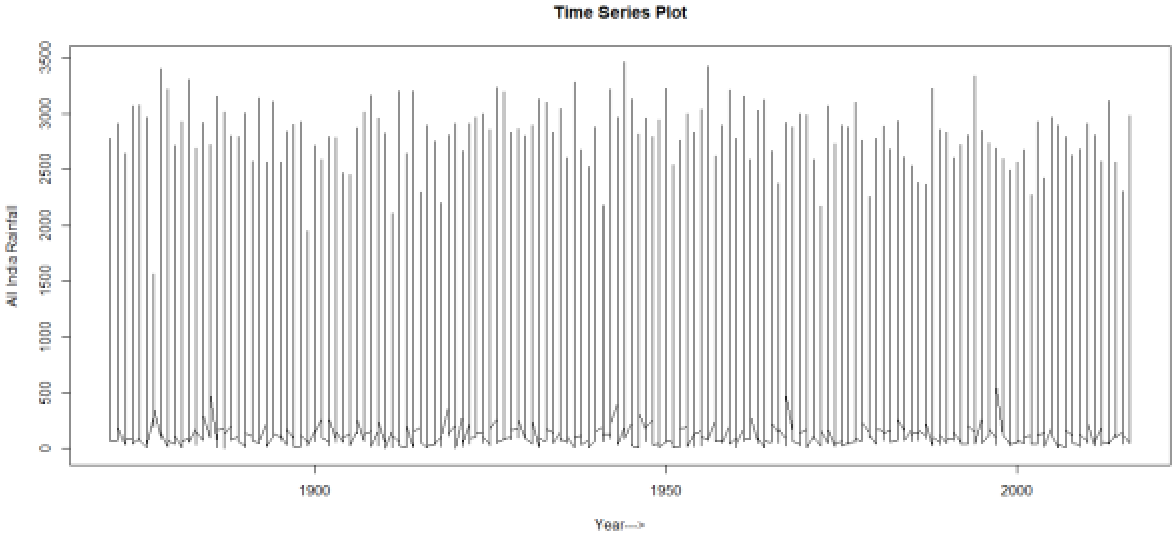
In the light of the given data, it seems that there is no monotonic trend present in the Annual Rainfall of Peninsular India.

6. All India-

All India rainfall provides us the collective rainfall of all the five regions with the area of 2880324 sq.km expect the six meteorological sub-regions, namely Jammu and Kashmir, Uttaranchal, Himachal Pradesh, Arunachal Pradesh, Lakshadweep and Andaman & Nicobar Islands which data were not available. This area receives on an average 900-1200 mm of rainfall annually. India is a country with huge landmass and diverse climates. The analysis of this data will give us an overall idea of rainfall in India.

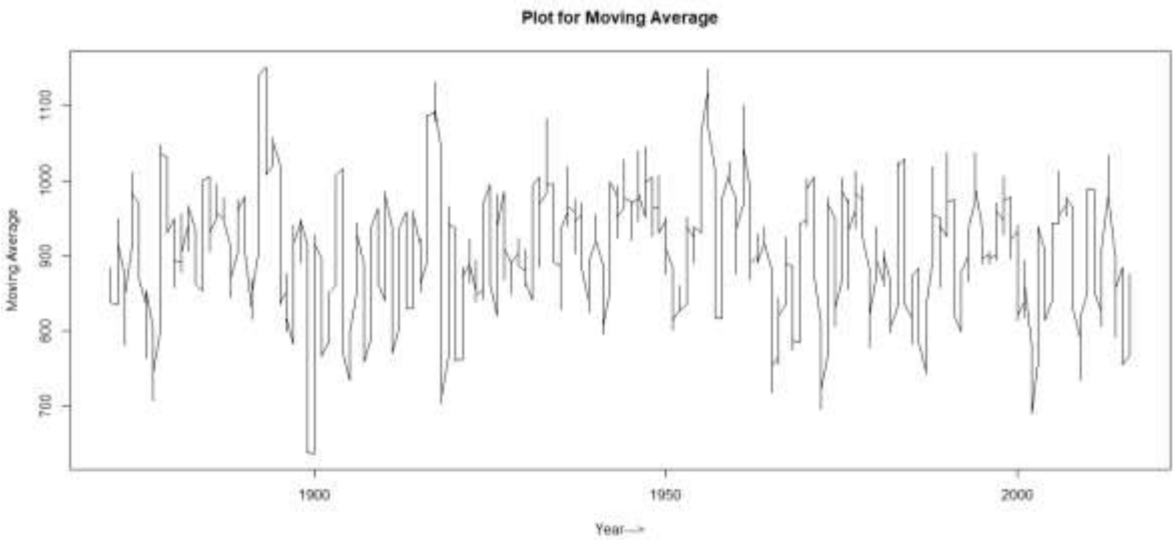
Overall, the climate in India is regarded as “Tropical”, as the Tropic of Cancer runs centrally over sub-continent. The IMD – India Meteorological Department has differentiated the nation’s tropical climate into 4 seasons – Summer, Monsoon, Post Monsoon and Winter. The summer is experienced during the months from March to May, Monsoon or the Rainy season from June to September, Post Monsoon or Retreating Monsoon from October to November and the Winters are seen from December to February.

Time Series Plot:



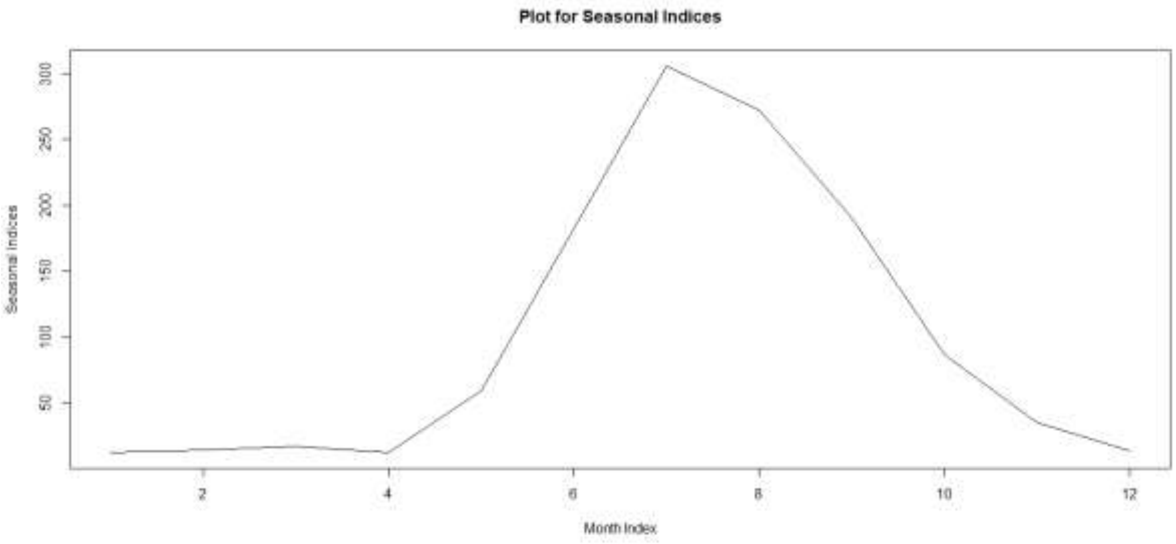
The above Time Series Plot doesn't indicate any clear trend but seems to have seasonal and cyclical fluctuations. So, we are further interested in decomposing the time series data.

Plot of Moving Average:



The plot of Moving Averages shows clear cyclical pattern. That’s why proceed for Harmonic Analysis.

Plot of Seasonal Indices:

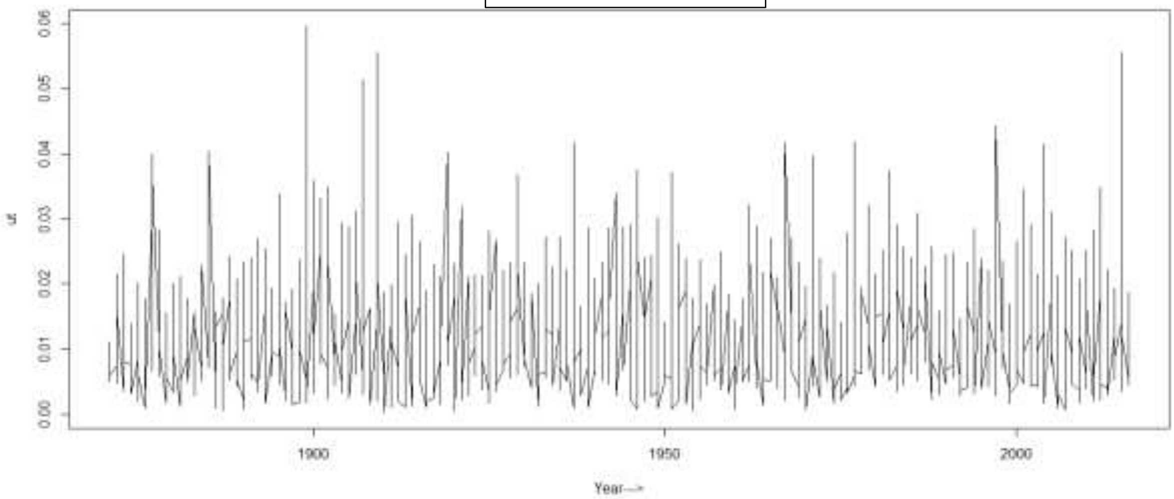


The plot for the Adjusted Seasonal Indices shows it maximum in the month of July. Moreover, it implies very small amount of rainfall in the area from Jan to May and then there is a steep increase in June and maximises in July and slow decrease after Aug. The seasonal indices for June -Sept shows a comparative higher value than others. Hence, monsoon in India lasts for these four months.

Periodogram Analysis:

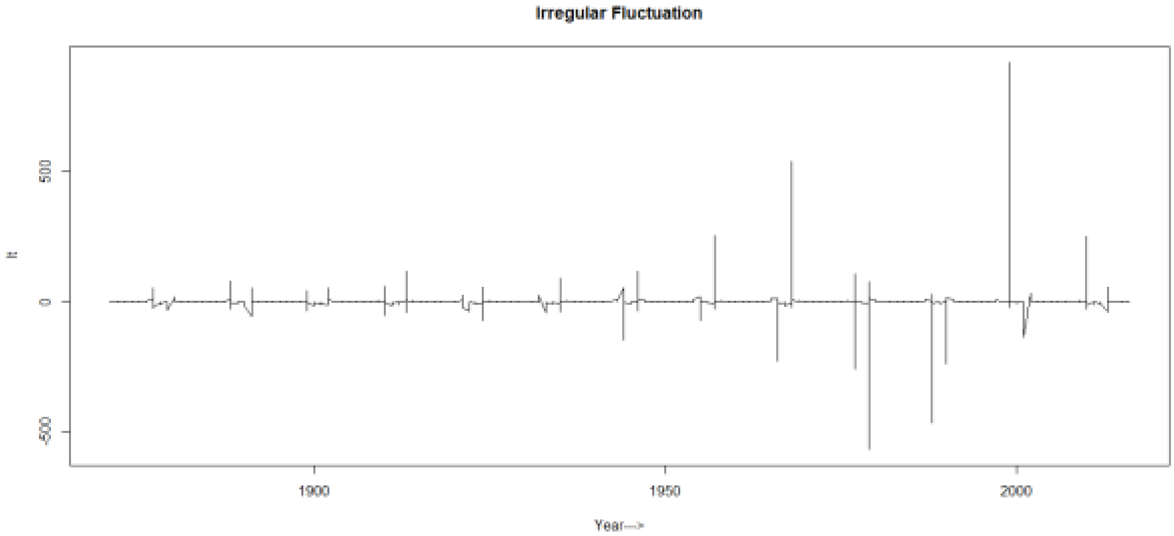
Let u_t be the residual series after eliminating the trend and seasonal component.

Plot of u_t



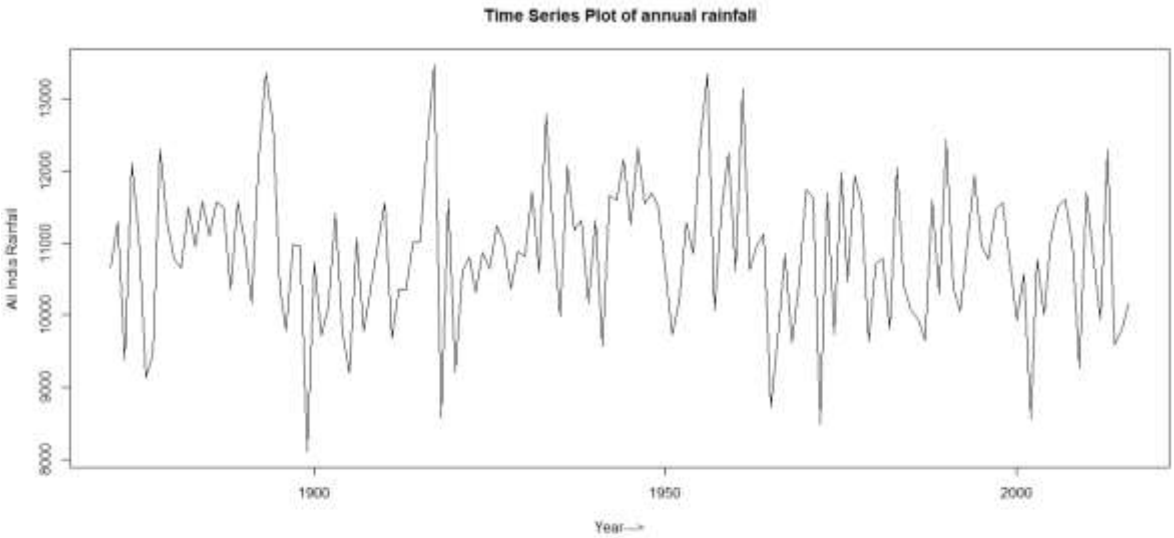
Eliminating the seasonality, we have only the cyclical and irregular component in the Periodogram. The diagram helps us in getting an idea of true period in the Periodic Analysis of the time series data. Here, in the analysis we get the true period 133 months i.e., it takes almost 11 years to complete a cycle. Now, we would eliminate the cyclical components to get the irregular components which remains unexplained.

Plot of Irregular Component:



In the plot of irregular components, it shows much less irregularity till 1970's. After that the irregularity in rainfall has increased. We observe that, the irregular fluctuations increase with time implying more unusual behaviour in the data. More irregularity in rainfall may be looked upon as an alarming sign of climate change or global warming.

Mann-Kendall Test:



This is the plot of Annual Rainfall of All India over the years concerned. As we can't determine the direction of trend from graph, we take the hypothesis of Mann- Kendall Test as-

H_0: No monotonic trend is present in the Time Series data

$$vs$$

H_1: Upward or Downward monotonic trend is present in the Time Series data

Let, x_1, x_2, ..., x_n, denote the annual rainfall in All India obtained at times 1, 2, ..., n, respectively.

Let us define, S= $\sum_{k-1}^{n-1}\sum_{j-k+1}^{n} \text{sgn}\left(x_j - x_k\right)$

And variance of S comes out to be,

$$\text{VAR(S)}=\frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p-1}^{g} t_p(t_p-1)(2t_p+5)\right]$$

Where g is the number of tied groups and t_p is the number of observations in the p^{th} group.

Test statistic-

The MK test statistic is defined as-

$$z_{Mk} \quad = \frac{S-1}{\sqrt{Var(S)}} \quad \text{if } S>0$$

$$= 0 \quad \text{if } S=0$$

$$= \frac{S+1}{\sqrt{Var(S)}} \quad \text{if } S<0$$

Critical region-

We reject Ho in favour of H_1 , if $|z_{Mk}| \geq \tau_{1-\alpha/2}$, where, α is the level of significance.

Calculation-

n= 146

α= 0.05

Using the “trend” package in the statistical software R we get,

S= -341

z_{Mk}= -0.57528

$\tau_{1-\alpha/2}$= $\tau_{0.975}$= 1.959964

Decision-

Since, $|z_{Mk}|$= 0.57528 < $\tau_{0.975}$= 1.959964,

Hence, here Ho is accepted 0.05 level of significance.

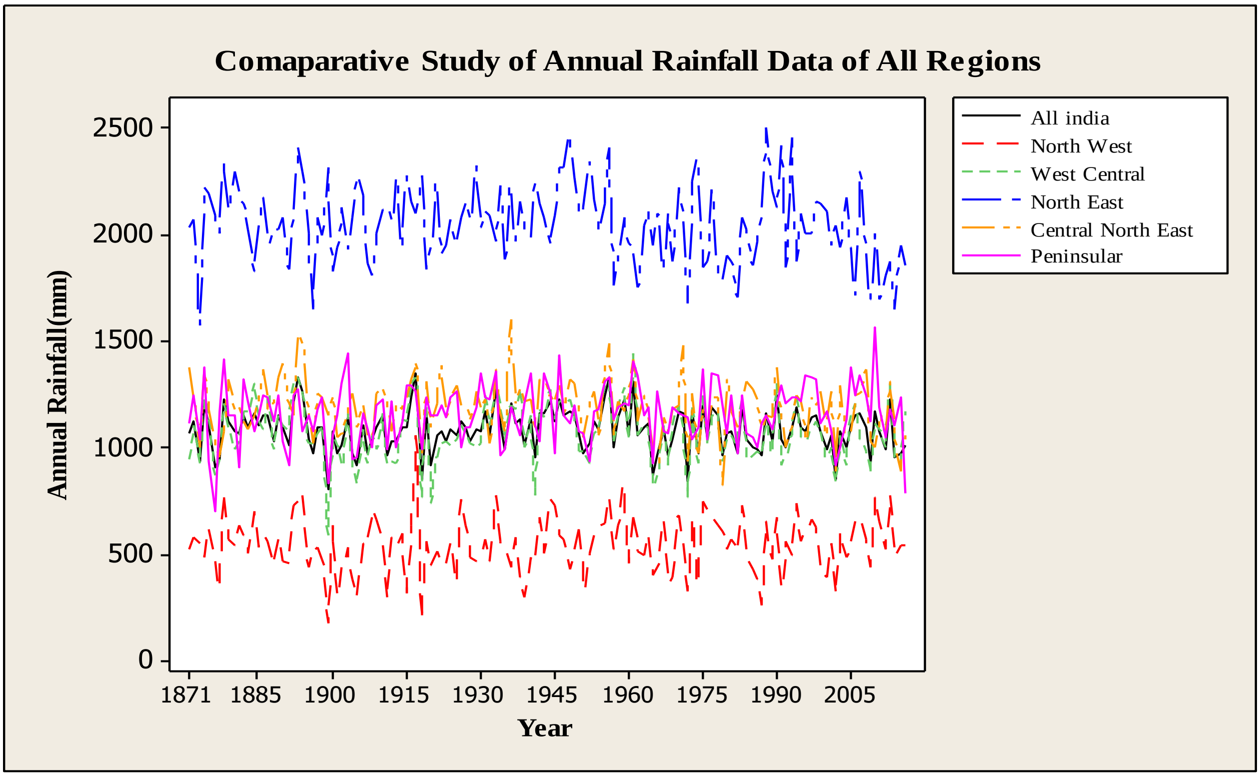
Conclusion-

In the light of the given data, it seems that there is no monotonic trend present in the Annual Rainfall of All India.

Conclusion

- From the time series analysis, we get an idea of the rainfall pattern in India and some of its sub-regions. The seasonal indices show maximum value on the month of July in all the sub-divisions and comparatively higher value for four months (June- Sept) in most part of the country. But the Peninsular India is an exception which experience rainfall also in Oct-Dec.

- From comparative study we observe that the average rainfall in North East India shows higher and the North West India shows lower than the average rainfall for all over India. The rainfall of other three provinces namely, Central North East, West Central, Peninsular India show values around the average rainfall for all India.



- Also, we get an idea of the period of cyclical movement of different provinces of India. It is 82 months for North West India, 134 months for North East India, 132 months for Central North East India, 95 months for West Central India, 116 months for Peninsular India, 133 months for All India dataset. That is, we get the period of cycles lie between 7-12 years.

- We also have performed the Mann-Kendall test for Monotonic Trend for annual data for each sub-division. All the sub-divisions except the North East India seemed to have no monotonic trend present in their annual rainfall. But the North East India's annual rainfall seems to have a decreasing trend. This can be an alarming sign of climate change or global warming in that area and this should be considered as a matter of concern. Some of the areas here reported drastic fall in amount of rainfall may be caused by rapid deforestation in those areas.

Acknowledgements

I wish to offer this endeavour to Professor Pallabi Ghosh, my supervisor, whose invaluable expertise, mentorship and guidance made it possible for me to work on this topic. I shall forever remain indebted to her for constant support throughout the project, in every possible way.

I further extend my gratitude to the other teachers of my department for their counsel, the students of my department for their perennial co-operation, my parents for their persistent encouragement and my friends for their support in times when I was woebegone and distressed regarding my dissertation.

Bibliography

1. The Analysis of Time Series- Chris Chatfield

2. Mathematical Statistics- Parimal Mukhopadhyay

3. Hydrological Sciences Journal- Vijay Kumar, Sharad K. Jain & Yatveer Singh (2010) Analysis of long-term rainfall trends in India, Hydrological Sciences Journal – Journal des Sciences Hydrologiques

4. Mann-Kendall Test- https://vsp.pnnl.gov/help/vsample/design_trend_mann_kendall.htm

5. Indian Institute of Tropical Meteorology- https://www.tropmet.res.in

6. Wikipedia- https://en.wikipedia.org/wiki/Climate_of_India

Appendix

We have used the statistical software R to perform the Time Series Analysis of each dataset. Here, I have attached the R code used for time series analysis of the All-India Data. Changing the file location for each dataset we implement the same R code for each dataset.

R-code used for All-India Monthly Data

```
data1=read.csv (file="C:/Users/MONDAL/Desktop/SEM 6/Project/Final/all india.csv") #read
the file
data1
mat plot(data1[,1:2], data1[,3], type="l", x lab="Year--->", y lab="All India Rainfall",
main="Time Series Plot")
mat plot(data1[,1:2], data1[,4], type="l", x lab="Year--->", y lab="Moving Average",
main="Plot for Moving Average")
Org= data1$ All. India. Rainfall                                  #Amount of Rainfall
MA=data1$Moving.Average                                               #Moving Average
Values
Ratio=(Org/MA)
data=data. frame (data1, Ratio) #attaching the Percentages in data frame
data
USI=array(dim=1)
Per Jan=data [data$ Month == "Jan",5]                 #Extracting Percentages for Jan
M1=Per Jan [-1]                                        #discarding the NA from array
USI [1] =mean(M1)
```

```
Per Feb=data [data$ Month == "Feb",5]                    #Repeating for Feb
M2=Per Feb [-1]
USI [ 2] =mean(M2)

Per Mar=data [data$ Month == "Mar",5]                    #Repeating for March
M3=Per Mar [-1]
USI [ 3] =mean(M3)

Per Apr=data [data$ Month == "Apr",5]                    #Repeating for Apr
M4=Per Jan [-1]
USI [ 4] =mean(M4)

Per May=data [data$ Month == "May",5]                    #Repeating for May
M5=Per May [-1]
USI [ 5] =mean(M5)

Per Jun=data [data$ Month == "Jun",5]                    #Repeating for June
M6=Per Jun [-1]
USI [ 6] =mean(M6)

Per Jul=data [data$ Month == "Jul",5]                    #Repeating for July
M7=Per Jul [-146]
USI [7] =mean(M7)

Per Aug=data [data$ Month == "Aug",5]                    #Repeating for Aug
```

```
M8=Per Aug [-146]
USI [ 8] =mean(M8)

Per Sept=data [data$ Month == "Sept",5]                    #Repeating for Sept
M9=Per Sept [-146]
USI [ 9] =mean(M9)

Per Oct=data [data$ Month == "Oct",5]                      #Repeating for Oct
M10=Per Oct [-146]
USI [ 10] =mean(M10)

Per Nov=data [data$ Month == "Nov",5]                      #Repeating for Nov
M11=Per Nov [-146]
USI [ 11] =mean(M11)

Per Dec=data [data$ Month == "Dec",5]                      #Repeating for Dec
M12=Per Dec [-146]
USI [ 12] =mean(M12)

USI                      #Unadjusted Seasonal Indices
A f=1200/sum (USI) #Adjustment Factor
ASI=USI*A f
ASI
Month=1:12
```

```
plot (Month, ASI, type="l", x lab="Month Index", y lab="Seasonal Indices", main="Plot for
Seasonal Indices")
Seasonal Indices=rep (ASI, each=1, times=146)        #Adjusted Seasonal Indices
u t=Ratio/Seasonal Indices                    #Eliminating Seasonality
mat plot(data1[,1:2], u t, type="l", x lab="Year--->", y lab="u t", main="Periodogram")
#Periodogram
u=u t [-(1:6)]
v t=u [-(1741:1746)] #Eliminating NA Values
k=as. array (v t)
k

sum1=sum2=0
a=array(dim=1)
b=array(dim=1)
A=array(dim=1)
B=array(dim=1)
R mu=array(dim=1)

for (i in 1:135)
{
  for (t in 1:1740)
  {
     a[t]=k[t]*cos(2*3.147*t/i)
     sum1=sum1+a[t]
     b[t]=k[t]*sin(2*3.147*t/i)
```

```
    sum2=sum2+b[t]
  }
  A[i]= ( 2/ 1740)* sum1
  B[i] = ( 2/ 1740)*sum2
  R mu[i]=(A[i]^2) + (B[i]^2)
}
R mu
x=as. vector (R mu)
max=max (R mu)
lambda=which (x==max, T) #true period
lambda
A0=mean(k)
A0
A1=A[lambda]
A1
B1=B[lambda]
Ut=array(dim=1)
for (t in -5:1746)
{
  Ut[t]=A0+(A1*cos(2*3.14*t/lambda)) + (B1*sin(2*3.14*t/lambda))
}
Ut
It=u t /Ut
It
mat plot (data1[,1:2], It, type="l", x lab="Year--->", y lab="It", main="Irregular Fluctuation")
```

We store the annual data for each region in an EXCEL file and use the following R-code for performing the Mann-Kendall Test.

R-code for Mann- Kendall Test

```
data=read. csv (file="C:/Users/MONDAL/Desktop/SEM 6/Project/Final/Annual Data.csv")
#read the file

data

All India=data$ All. India

All India

Northwest=data$ North. West

Northwest

West central=data$ West. Central

West central

Northeast=data$ North. East

Northeast

Central northeast= data$ Central North. East
```

```
Central northeast

Peninsular=data$ Peninsular

Peninsular

library (trend) #call the library

mk. test (All India) #Mann-Kendall Test

mk. test (Northwest)

mk. test (West central)

mk. test (Northeast, alternative="less")

mk. test (Central northeast)

mk. test (Peninsular)

crit=q norm (0.975, lower. tail=TRUE)

crit

crit northeast=q norm (0.95, lower. tail=TRUE)

crit northeast
```

mat plot (data [,1], data [,2:7], type="l", x lab="Year--->", y lab="Rainfall", main="Time Series Plot")

mat plot (data [,1], data [,2], type="l", x lab="Year--->", y lab="All India Rainfall", main="Time Series Plot of annual rainfall")

mat plot (data [,1], data [,3], type="l", x lab="Year--->", y lab="Northwest Rainfall", main="Time Series Plot of annual rainfall")

mat plot (data [,1], data [,4], type="l", x lab="Year--->", y lab="West Central Rainfall", main="Time Series Plot of annual rainfall")

mat plot (data [,1], data [,5], type="l", x lab="Year--->", y lab="North East Rainfall", main="Time Series Plot of annual rainfall")

mat plot (data [,1], data [,6], type="l", x lab="Year--->", y lab="Central North East Rainfall", main="Time Series Plot of annual rainfall")

mat plot (data [,1], data [,7], type="l", x lab="Year--->", y lab="Peninsular Rainfall", main="Time Series Plot of annual rainfall")