**DROUGHT ANALYSIS AND PREDICTION**

**INDEX**

|  |
| --- |
| 1. OBJECTIVE |
| 1. BACKGROUND  * What is Drought? * How do you calculate drought? * What are the different drought indices? * What is SPI? * What is SPAI? |
| 1. PREDICTION |
| 1. RESULT AND ANALYSIS  * SPI and SPAI series * Spatial Analysis * Temporal Analysis |
| 1. CONCLUSION |

**OBJECTIVE**

* Perform a drought characterization, analysis and prediction for two locations at your locality based on the available historical data.
* Prepare **SPI** and **SPAI** series along with Spatial Analysis and Temporal analysis
* To predict the drought condition (1971 onwards) by using different modelling approaches like ARIMA, Deep Learning etc.

**BACKGROUND**

**What is Drought?**

A **drought** is an event of prolonged shortages in the water supply, whether atmospheric (below-average [precipitation](https://en.wikipedia.org/wiki/Precipitation)), [surface water](https://en.wikipedia.org/wiki/Surface_water) or [ground water](https://en.wikipedia.org/wiki/Ground_water).

There are 3 types of drought:

1. **Meteorological** drought occurs when there is a prolonged time with less than average precipitation.
2. [**Agricultural**](https://en.wikipedia.org/wiki/Agriculture) droughts affect crop production or the ecology of the [range](https://en.wikipedia.org/wiki/Range_(biology)).
3. [**Hydrological**](https://en.wikipedia.org/wiki/Hydrology) drought is brought about when the water reserves available in sources such as [aquifers](https://en.wikipedia.org/wiki/Aquifer), [lakes](https://en.wikipedia.org/wiki/Lake) and [reservoirs](https://en.wikipedia.org/wiki/Reservoir) fall below a [locally significant](https://en.wikipedia.org/wiki/Descriptive_statistics) threshold.

Main Cause of Drought: **Precipitation Deficiency**



**How do you calculate Drought?**

Droughts are a normal part of the climate, and they can occur in any climate regime around the world, even deserts and rainforests. Droughts are one of the more costly natural hazards on a year-to-year basis; their impacts are significant and widespread, affecting many economic sectors and people at any one time.

**Drought indicators or indices** are often used to help track droughts, and these tools vary depending on the region and the season.

**Indicators** are variables or parameters used to describe drought conditions. Examples include precipitation, temperature, streamflow, groundwater and reservoir levels, soil moisture and snowpack

**Indices** are typically computed numerical representations of drought severity, assessed using climatic or hydrometeorological inputs including the indicators listed above. They aim to measure the qualitative state of droughts on the landscape for a given time period.

**What are the different Drought indices?**

There are lot of Drought indices which can be used to calculate the drought. Some of them are:

1. Aridity Anomaly Index (AAI)
2. Standardized Precipitation Index (SPI)
3. Weighted Anomaly Standardized Precipitation (WASP)
4. Rainfall Anomaly Index (RAI)
5. Standardized Anomaly Index (SAI)
6. Standardized Precipitation Evapotranspiration Index (SPEI)
7. Standardized Precipitation Anomaly Index (SPAI)

In our case study we have focused on **SPI** and **SPAI** to calculation and analysis of Drought.

**What is SPI?**

SPI is computed from an observed precipitation time series (at least 30–40 years) by fitting a gamma distribution to the raw precipitation data. The cumulative distribution function (CDF) of the gamma distribution is then transformed to standard normal variate (Z) to obtain SPI.

**What is SPAI?**

In the computation of SPAI, the precipitation anomalies are used instead of raw precipitation values. The anomalies of precipitation are given by subtracting long term mean precipitation with the actual precipitation value. Noteworthy is that the unit of the rainfall anomaly series is the same as that of the rainfall series. This needs to be standardized to convert to the scale of Z score. After obtaining the anomalies, a single probability distribution is fitted to the entire anomaly series. In our case empirical distribution is fitted. To obtain the empirical CDF of the rainfall anomaly series, the Weibull’s plotting position formula is found to be the best for plotting position and is expressed by:

p = m/(N+1)

where p = cumulative probability; m = rank of the dataset arranged in descending order; and N = sample size as explained before, i.e., the total number of time steps in the dataset.

After fitting the empirical distribution, the quantile values corresponding to each anomaly values are obtained. These quantile values, ranging from 0 to 1, may be designated as the reduced variates of the rainfall anomalies. Next, these reduced variates are transformed to standard normal variates (Z), i.e., the numbers on

the real line which would correspond to the values of reduced variates in a standard normal distribution are determined. The obtained standard normal variates (Z) are the required SPAI. Similar to the SPI, SPAI values also range between −∞ and +∞ where negative and positive values reflect drier and wetter conditions respectively.

**SPI and SPAI classification**

|  |  |
| --- | --- |
| SPI/SPAI values | Condition |
| 2.0 and more | Extremely wet |
| 1.5 to 2.0 | Very wet |
| 1.0 to 1.5 | Moderately wet |
| -1.0 to 1.0 | Near Normal |
| -1.5 to -1.0 | Moderately Dry |
| -2.0 to -1.5 | Severely Dry |
| -2.0 and less | Extremely Dry |

**PREDICTION**

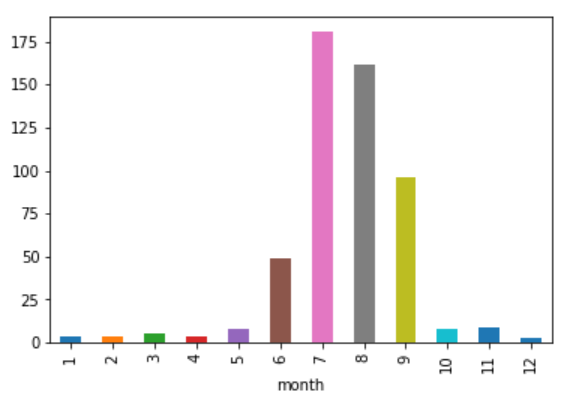
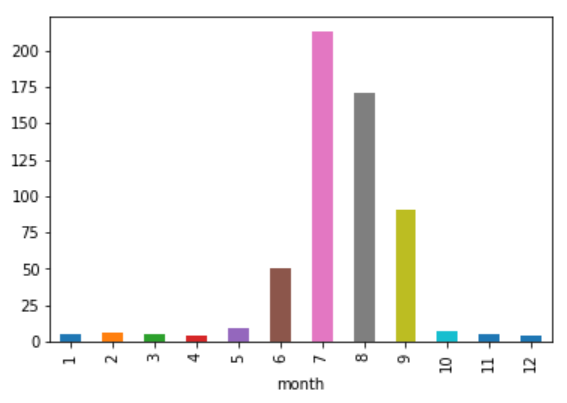
Precipitation data(monthly) form year **1901** to **1970** of two major cities of Rajasthan (Western India) i.e., **Jaipur** and **Ajmer** are used in our Analysis.

**What is the process used?**

With help of monthly Precipitation data from **Jan** **1901** to **Dec** **1970**, we have forecasted the rainfall from **Jan** **1971** to **Dec** **2000** using different Time Series Analysis Model. From the Rainfall data SPI and SPAI series is prepared which are then used for Spatial and Temporal Analysis.

\*\*Python Programming Language has been used for the project\*\*

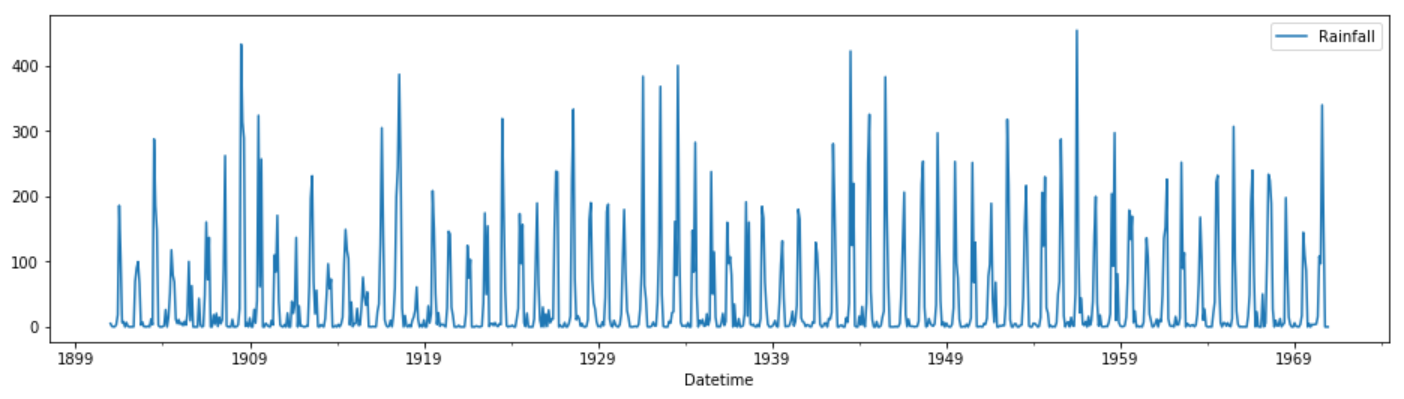
1. Plot cumulative monthly rainfall.

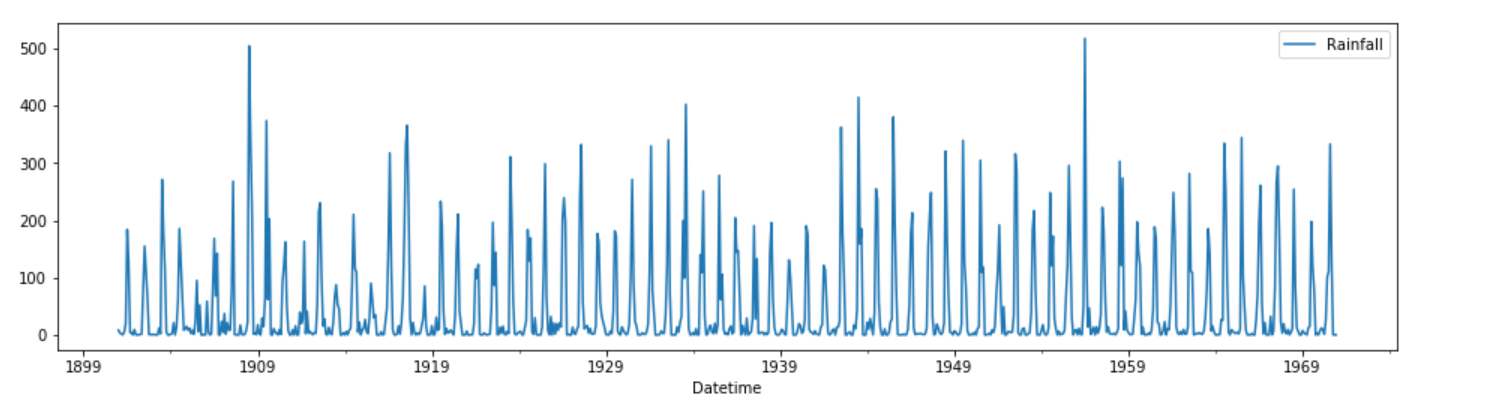
 

AJMER JAIPUR

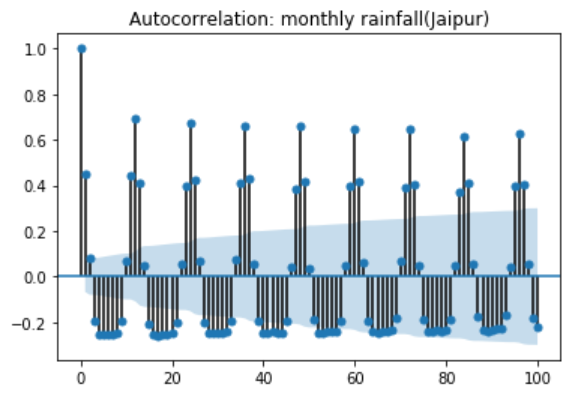
It can be observed that highest rainfall is observed during the month of JULY and AUGUST. These are the rainy seasons.

1. Visualizations

 AJMER

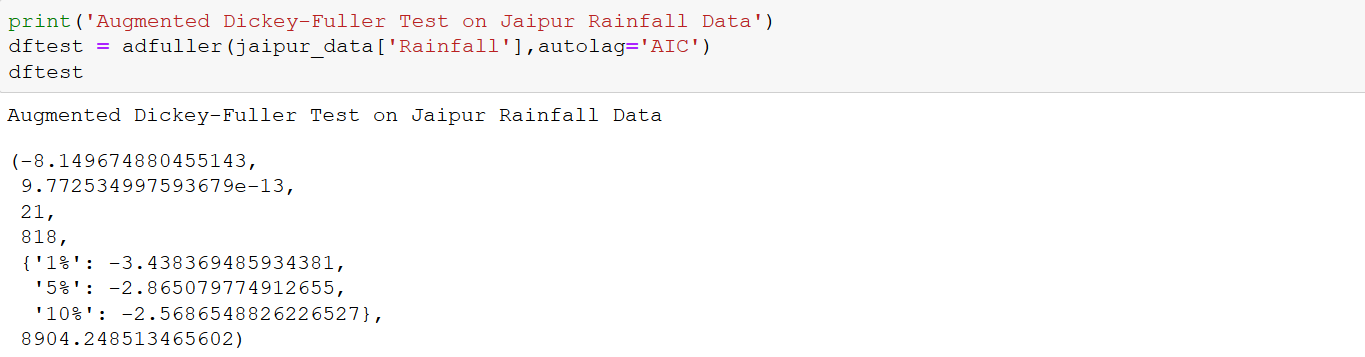
JAIPUR

1. Plotting Auto-Correlation function to check the seasonality



It can easily be observed that our rainfall data is seasonal.

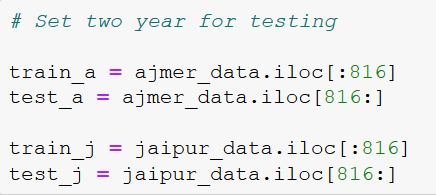
1. Checking Stationarity of data using Augmented Dickey-Fuller Test



For augmented dickey-fuller test we assumed the null hypothesis as non-stationary data and calculate p value which comes out to be 9.77e-13.

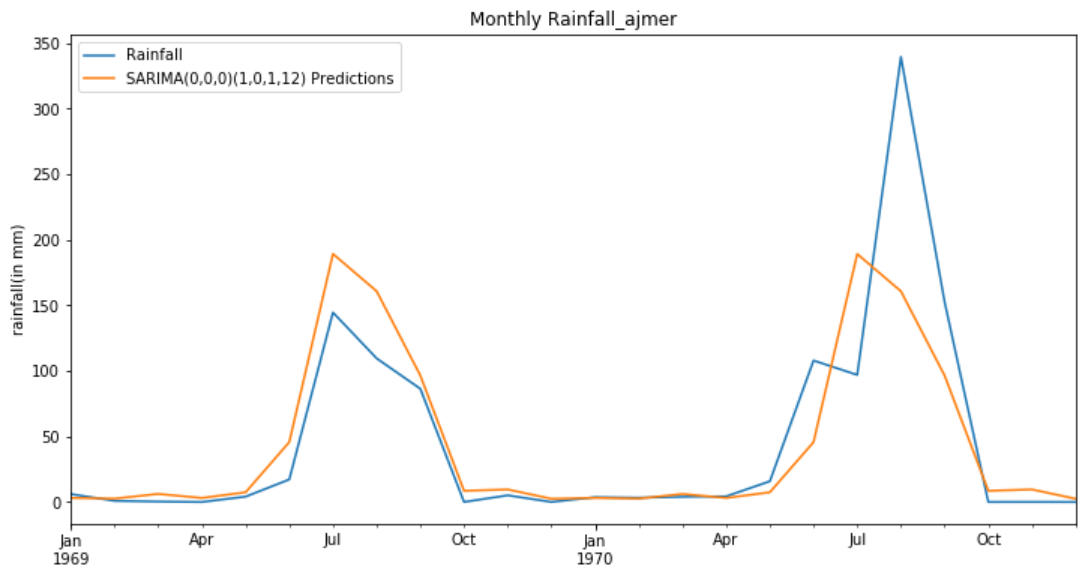
Since p < 0.05 therefore we reject the null hypothesis and can say that our data is stationary.

1. Split the dataset into train and test set



1. Applying the models
2. **ARIMA**
3. Ajmer

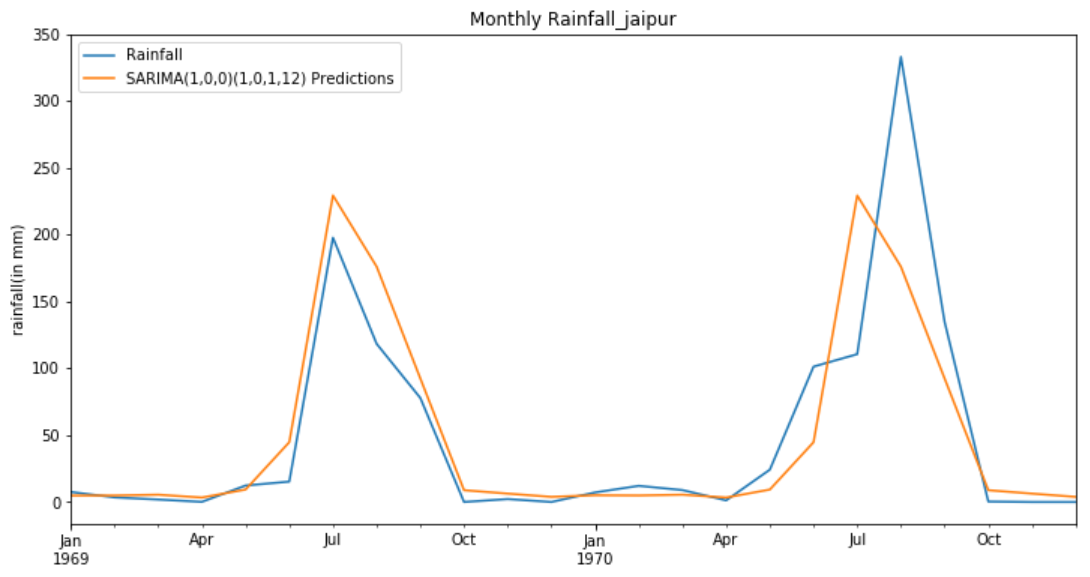
Model used Seasonal-ARIMA model = SARIMAX(0, 0, 0)x(1, 0, 1, 12)



RMSE value = 47.18

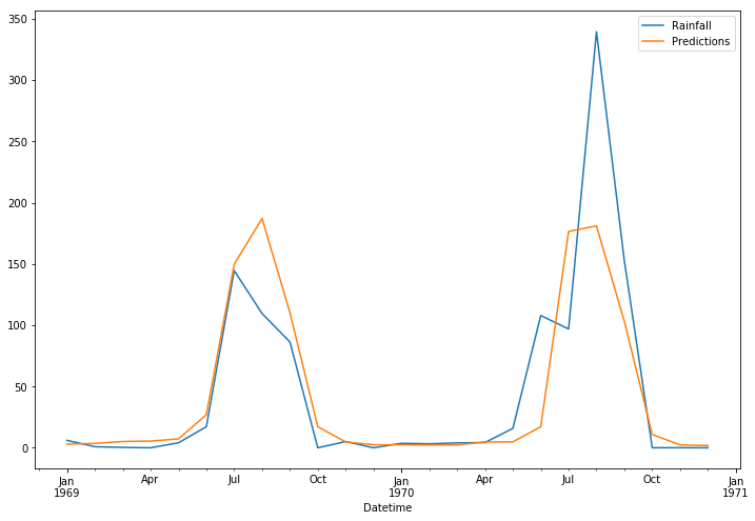
1. Jaipur

Model used Seasonal-ARIMA model = SARIMAX(1, 0, 0)x(1, 0, 1, 12)



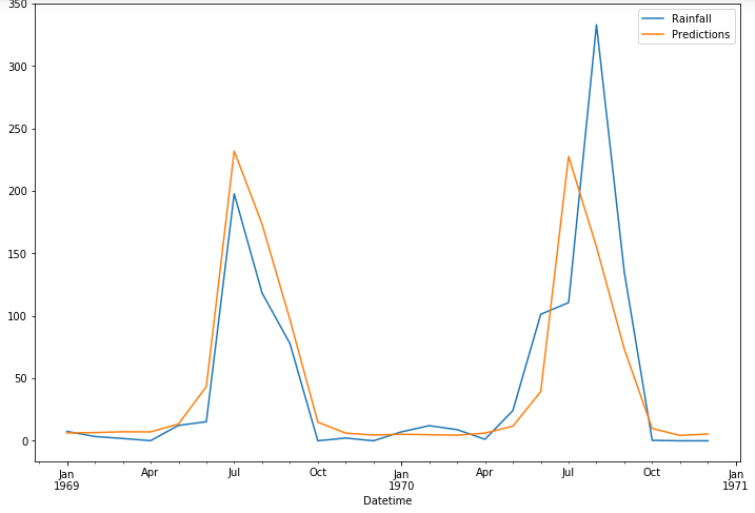
RMSE value = 45.51

1. **Deep Learning (LSTM)**
2. Ajmer



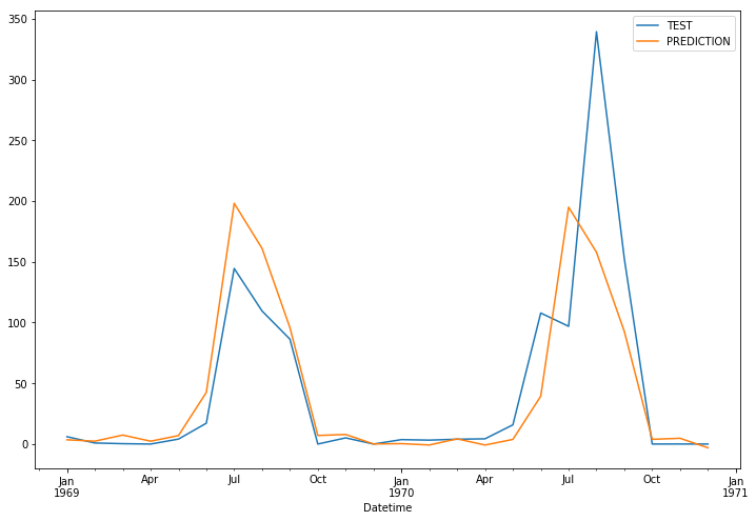
RMSE value = 45.40

1. Jaipur



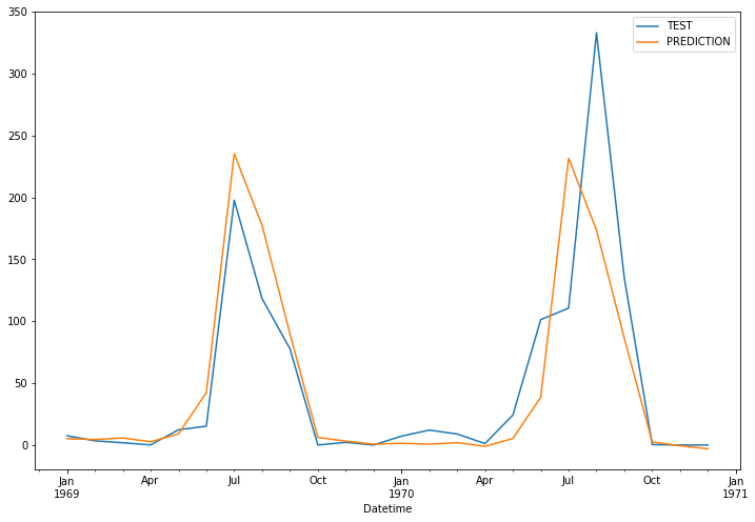
RMSE value = 49.57

1. **Holt’s Winter Method (Triple Exponential Smoothing)**
2. Ajmer



RMSE value = 48.95

1. Jaipur



RMSE value = 46.88

1. Results



1. Choosing the best model and refit it to forecast future rainfall

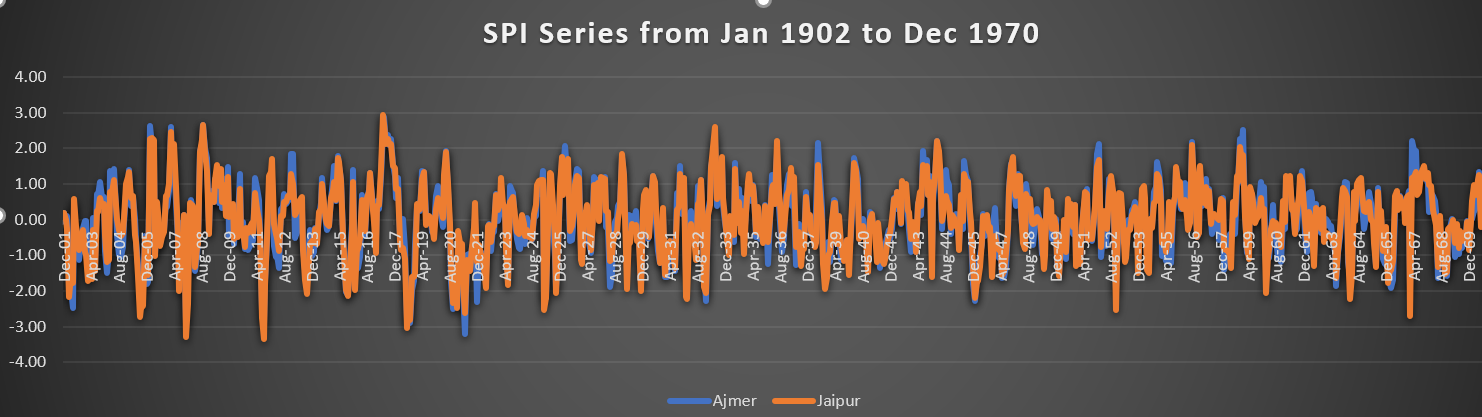
It can be observed that lowest RMSE value comes out from LSTM model on Ajmer Rainfall data and ARIMA model on Jaipur Rainfall data, therefore we would refit the above best-chosen model on the whole data and forecast the future rainfall and save it.

Now we have the monthly Rainfall data from 1901 to 2000.

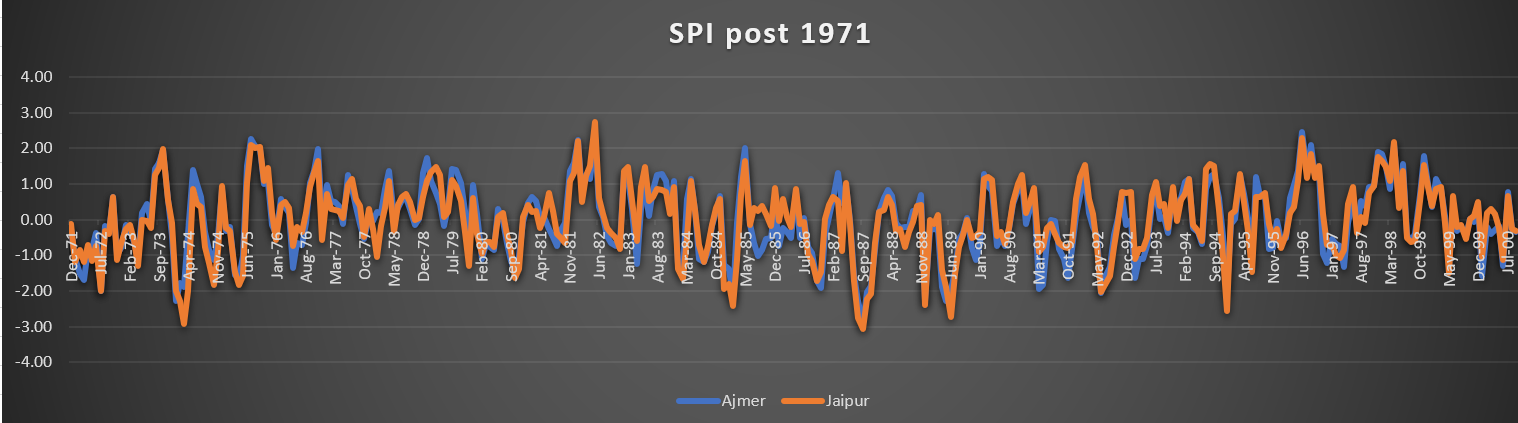
**RESULT AND ANALYSIS**

**SPI and SPAI Series**

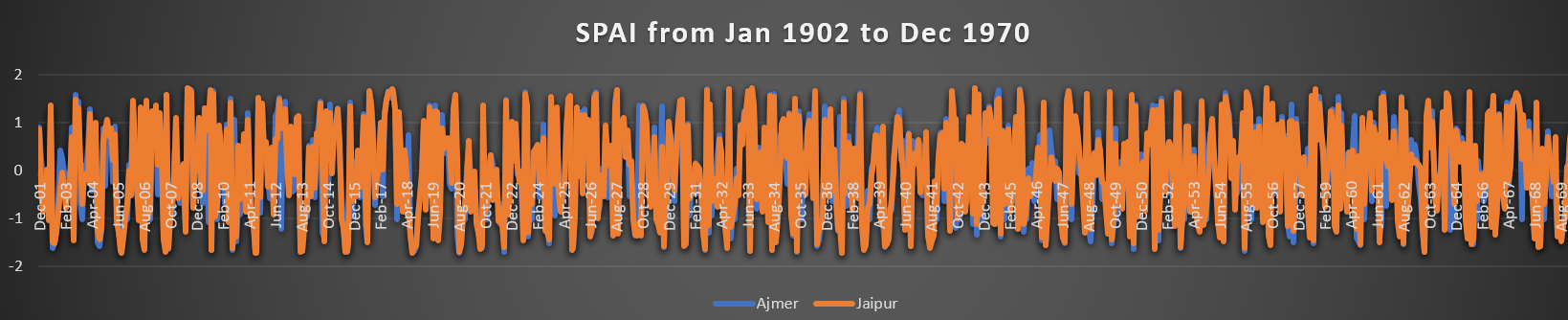
1. **SPI series from 1901 to 1970**



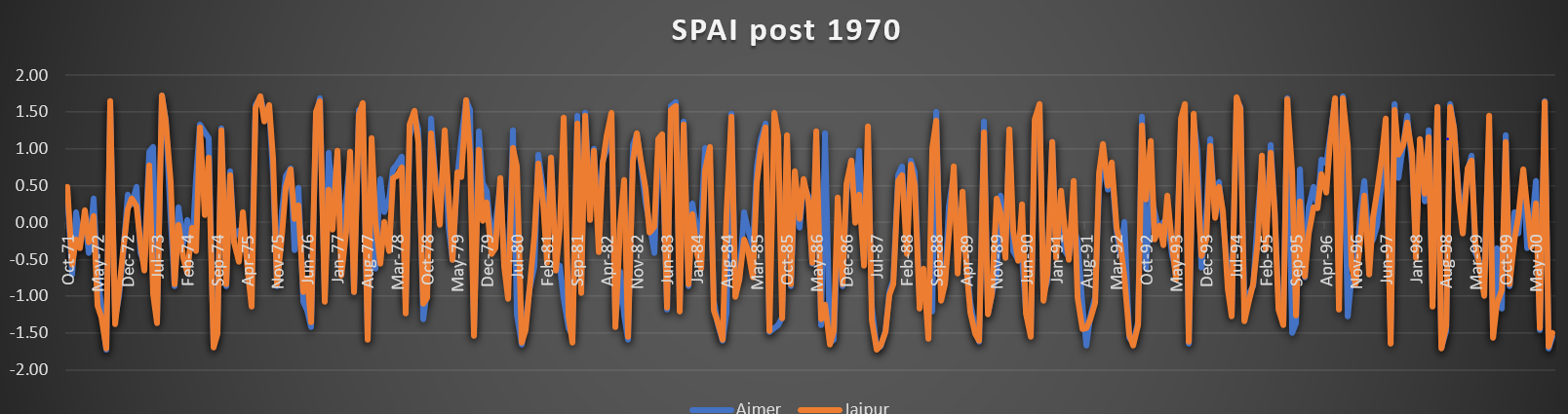
1. **SPI series post 1970**



1. **SPAI series from 1901 to 1970**



1. **SPAI series post 1970**



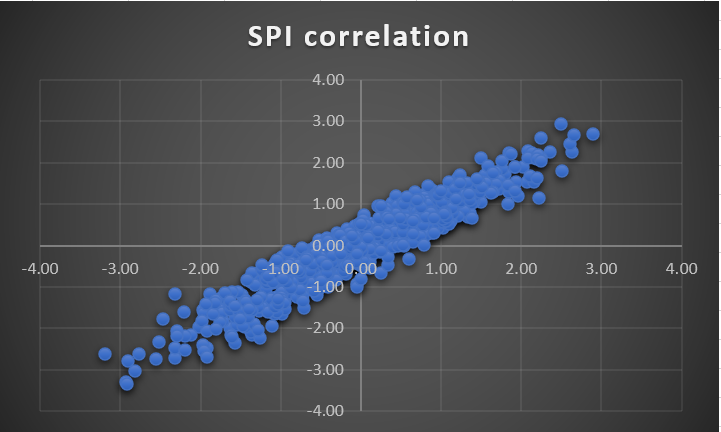
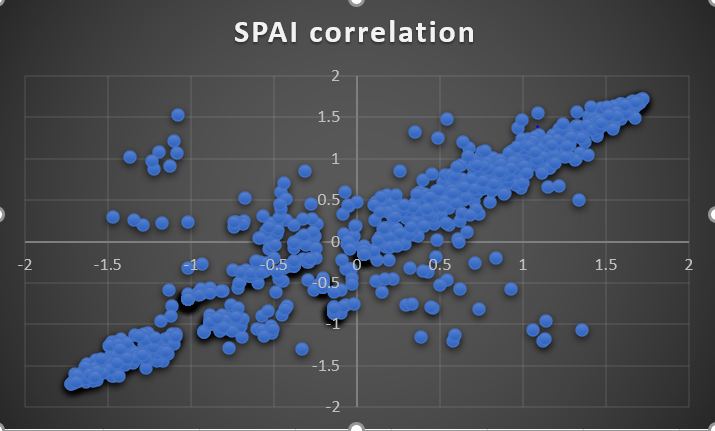
\*\* Note: It has been observed from experiment and research that SPAI series gave mush better result than SPI series. Therefore, we will use SPAI drought characterization in our Analysis

**SPATIAL ANALYSIS**

For spatial analysis we have to compare drought characteristics at the two stations (i.e., Jaipur and Ajmer in our case)

It can be observed from above graphs (SPI and SPAI series) that rainfall variations are almost same in both the cities.

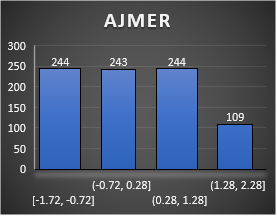
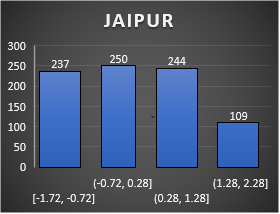
**Correlation between two cities**

It can be observed that correlation between two cities is approximately equal to 1 in case of SPI series which is practically difficult to believe as there are always some months which created good amount of difference in intensity of rainfall which is clearly shown using SPAI correlation which is slightly less than 1.

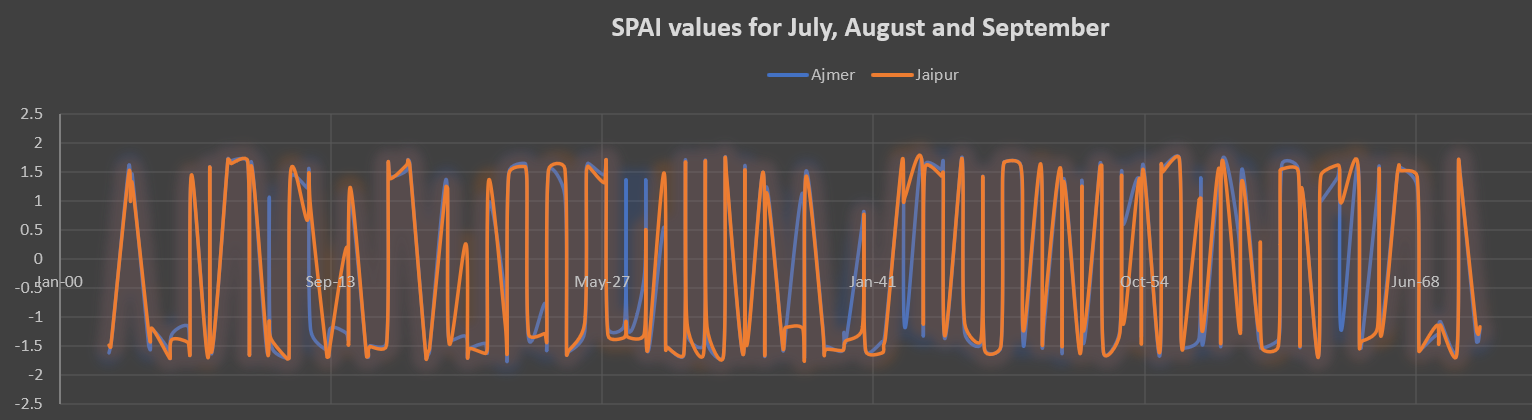
Since both the cities are from the same state of Rajasthan and approximately 200 km apart therefore correlation between them are nearly 1.

**Frequency distribution bar graph**

It can be observed from the above graph that both Jaipur and Ajmer cities has suffered the scarcity of good amount of rainfall as out of all the months from 1901 to 1970 only 109 months had SPAI values greater than 1.28 in both Ajmer and Jaipur.

For comparison of Drought at two locations SPAI values of July, August and September are compared as these are rainy seasons which decide whether the year is a drought year or not.



When monthly SPAI values are compared for drought characterization at the two location it has been observed that the driest year is 1911 and the wettest year is 1956 for both the cities.

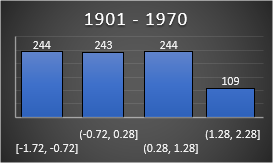
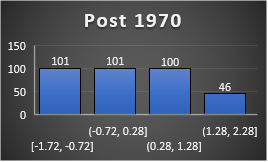
Year with severe drought are 1905, 1911, 1913, 1915, 1918, 1920, 1930, 1938, 1968 in Ajmer while 1905, 1911, 1913, 1915, 1918, 1920, 1921, 1925, 1930, 1938, 1939, 1968 in Jaipur where SPAI values are less than -1.5.

It is observed that Jaipur has faced more scarcity of rainfall as compared to Ajmer in long term.

**TEMPORAL ANALYSIS**

For Temporal Analysis we have to compare drought characterization between two time periods (in our case we have compared time period of 1951-1970 and 1971-1990) at both the locations.

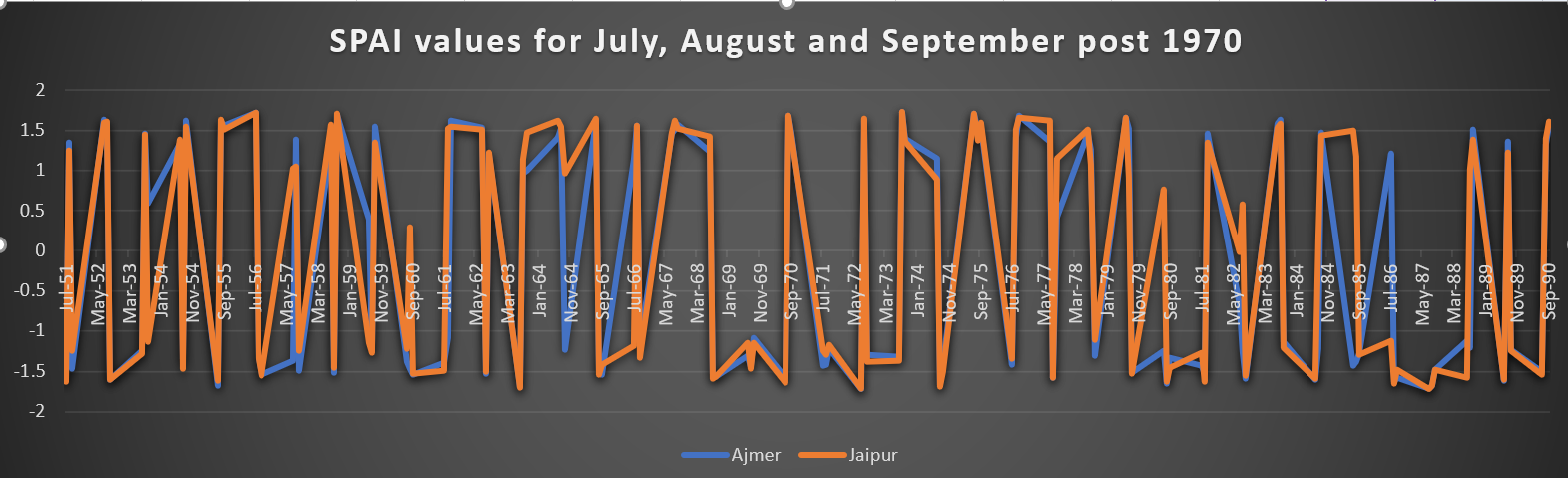
**Combined Frequency distribution bar graph of both the location**

It is clearly seen that Rainfall pattern remains the same and both the cities of Rajasthan suffered from Rainfall before and after 1970.

Since we have already described the drought characterization before 1970 at both the locations. Now we will look at the situation after 1970

**SPAI values from 1950 - 1990**



It has been observed that the driest year is 1972 in Ajmer and 1987 in Jaipur and the wettest year is 1973 for both the cities.

Year with severe drought are 1972, 1986 and 1987 in Ajmer while 1974, 1987 and 1989 in Jaipur where SPAI values are less than -1.5.

It is observed that Drought conditions are nearly remains the same in both the time period as shown in above graph.

**CONCLUSION**

From our case study we concluded that Eastern part of Rajasthan (Ajmer and Jaipur) had suffered a prolonged shortage of rainfall from many years.

There is lot of variations in drought conditions in both the cities and ultimate drought curve remains constant over years in both the cities.