Research Paper Computer Science

Rainfall Prediction using Seasonal Auto Regressive Integrated Moving Average model

S.Meenakshi Sundaram | Research Scholar, Faculty of Computing, Sathyabama University Tamilnadu, India

M.Lakshmi | Head, Faculty of Computing, Sathyabama University, Tamilnadu, India

ABSTRACT

Predicting rainfall is a very tough and needy problem. Researchers worldwide are trying to do it with different methods. Still no such prediction gives accurate result, because of unusual and unsure changes occurred. Here we tried to predict the monthly rainfall using Box-Jenkins Seasonal Auto Regressive Integrated Moving Average model, with 136 years of rainfall data.

KEYWORDS | ARIMA, seasonality, MAPE, prediction

1. Introduction

Monitoring the environmental changes especially climatic changes is the very essential part of building future. Rainfall prediction is the most concerned variable due to its inception and non uniformity. Many researchers contributed towards the prediction of rainfall in Tamilnadu. O.N.Dhar et.al. analyzed the average rainfall for the north east monsoon using standard methods. They analyzed trend, periodicities and variability for prediction annual rainfall in Tamilnadu. O.N.Dhar and Rekecha obtained the relation between southwest and northeast through a correlation analysis. The NEMR – STA association is exhibited by Balachandran et.al. In this research work, we have used the classical statistical seasonal ARIMA model for monthly rainfall prediction.

2. Study Area

Our study is within Tamilnadu region bounded on the north by Andhra Pradesh and Karnataka, on the east by the Bay of Bengal, on the south by the Indian Ocean and on the west by the state of Kerala. The state has a latitude between 8° 5′ N and 13 ° 35′ N and longitude between 78 ° 18′ E and 80 ° 20′ E. (see Figure 1). It receives rainfall by both southwest and northeast monsoons. The fate of the agricultural economy of the state is decided by the monsoons. Tamilnadu, located in southeast Peninsular India receives the major part of its annual rainfall during the northeast monsoon (the three-month period from October to December). While coastal Tamilnadu receives about 60% of its annual rainfall, interior Tamil Nadu receives about 40–50% of annual rainfall during Northeast monsoon season. Approximately 33% of annual rainfall in Tamilnadu is from the southwest monsoon and 48 % is from the northeast monsoon.

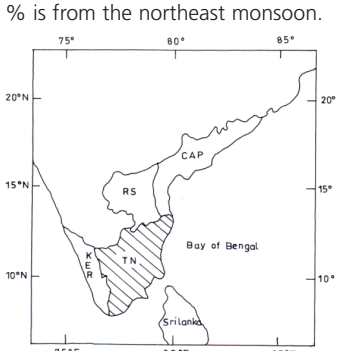


Fig1: Geographical location of area of study, Tamilnadu

3. Methodology

Data set:

The monthly rainfall for 136 years (1871 – 2006) had been obtained from Indian Institute of Tropical Meteorology (IITM), Pune, India.

ARIMA model:

ARIMA model is a major prediction tool in hydrology and widely used to predict surface water processes and stream flow events. It is a powerful statistical technique for modeling time series and rainfall prediction. The autoregressive models AR(p) base their predictions of the values of a variable x_t, on a number p of past values of the same variable number of autoregressive delays $x_{t-1}, x_{t-2}, \ldots x_{t-p}$ and include a random disturbance e_t. The moving average models MA(q) generate predictions of a variable x_t based on a number q of past disturbances of the same variable prediction errors of past values $e_{t-1}, e_{t-2}, \ldots, e_{t-q}$. In 1976, Box and Jenkins proposed the mathematical transformation of the non stationary time series into stationary time series by a difference process defined by an order of integration parameter d. This mathematical transformation transforms ARMA (p,q) models for non stationary transformed time series as the ARIMA (p,d,q) models, Autoregressive Integrated Moving Average models.

The ARIMA model building strategy includes iterative identification, estimation, diagnosis and forecasting stages. Once the model is selected it is used to forecast the monthly rainfall series. Time series analysis provides great opportunities for detecting, describing and modeling climatic variability and impacts. Ultimately, to understand the meteorological information and integrate it into planning and decision making process, it is important to study the temporal characteristic and predict lead times of the rainfall of a region. This can be done by identifying the best time series model using Box – Jenkins Seasonal ARIMA modeling techniques.

4. RESULTS AND DISCUSSIONS

A Time Series plot is a simple graph of the given observations(rainfall), shown in Fig. 2.

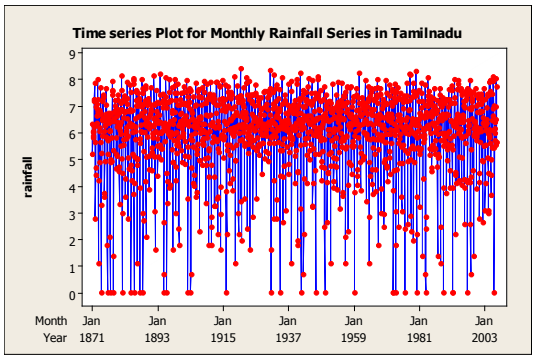


Fig.2: Time series plot of Monthly Rainfall in Tamilnadu

The descriptive statistics are given in the following table

Mean	Std Deviation	Coeffi of Variation	Median
5.9562	1.6065	26.97	8.4109

Table 1: Descriptive Statistics of log transformed monthly rainfall series

An important guide to the properties of a time series is provided by a series of quantities called sample autocorrelation coefficents or serial correlation coefficent, which measure the correlation between observations at different distances apart.

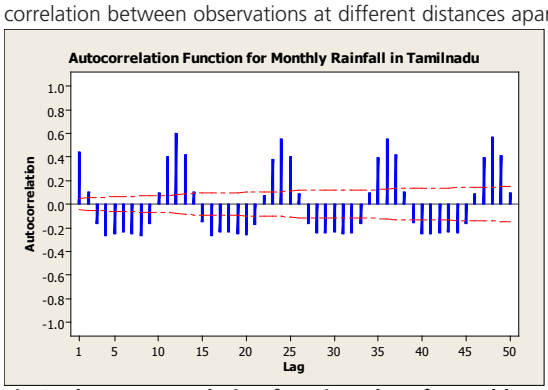


Fig.3: The Autocorrelation function plot of Monthly Rainfall in Tamilnadu

These coefficents often provide insight into the probability model which generated the data. The sample autocorrelation coefficent is similar to the ordinary correlation coefficent between two variables (x) and (y) except that it is applied to a single time series to see if successive observations are correlated.

The autocorrelation coefficient is a valuable tool for investigating properties of an empirical time series. Autocorrelations can be used to determine whether there is any pattern like AR, MA, ARMA and ARIMA in the time series. The absence of such patterns shows that the time series is random. The autocorrelation plot shows how values of the time series are correlated with the past values of the series.

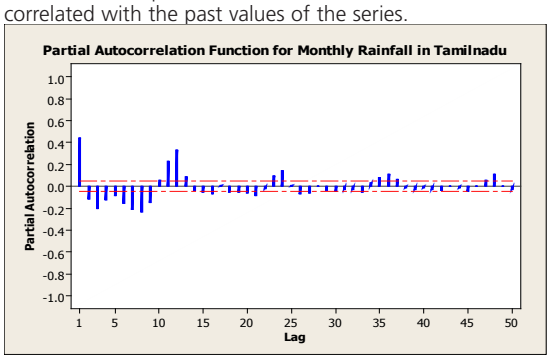


Fig.4: The Partial Autocorrelation function plot of Monthly Rainfall in Tamilnadu

The prediction methods need the assumption that the time series is stationary and in the way we get meaningful statistics such as means, variances and correlations. Seasonal patterns of time series can be examined through correlograms. Seasonality can be found by identifying those autocorrelation coefficients of more than two or three time lags that are significantly different from zero. The autocorrelation function coefficient and the partial autocorrelation coefficient for the first twelve lags are given in table 2. Since the decay of seasonal ACF is very gradual, the series remains seasonally nonstationary and is in need of seasonal differencing.

LAG	ACF	PACF
1	0.376	O.376
2	0.104	-0.044
3	-0.139	-0.190
4	-0.268	-0.177
5	-0.270	-0.114
6	-0.241	-0.133
7	-0.273	-0.248
8	-0.266	-0.265
9	-0.133	-0.155
10	0.120	0.023
11	0.373	0.159
12	0.571	0.334

Table 2: Autocorrelation and partial autocorrelation functions coefficients for first twelve lags for the monthly rainfall in Tamilnadu

The Box – Jenkins SARIMA model:

A powerful model for describing stationary and non - stationary seasonal time series is Seasonal Autoregressive Integrated Moving Average process (SARIMA) of order (p, d, q) X (P,D,Q) s. Model identification can be made through the autocorrelation function and partial autocorrelation function plots (fig. 3 & 4). The tentative SARIMA model for the monthly rainfall series of Tamilnadu is identified as $(0,1,1)X(0,1,1)_{12}$. .

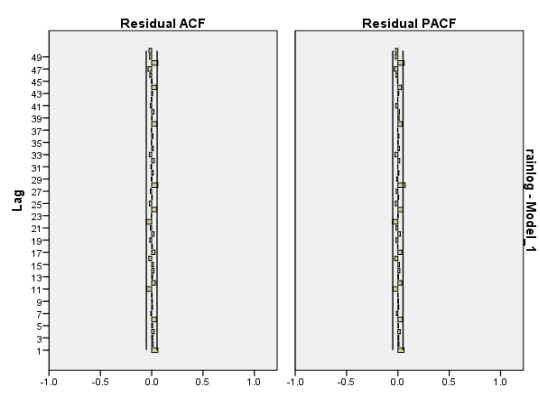


Fig 5: The residual plots of ACF and PACF of Monthly Rainfall Series

Performance Evaluation Criteria:

The error measure used in this research work is the Mean Absolute Percentage Error (MAPE) measure, which is given by

$$MAPE = 100 \times \frac{\sum_{t=1}^{N} \frac{|E_t|}{Y_t}}{N} \qquad \text{------- (1)}$$

where Y_t and E_t represent desired inputs and corresponding errors at t = 1,2,...,N respectively. For this model SARIMA $(0,1,1)$ X $(0,1,1)_{12}$, the error measure, Mean Absolute Percentage Error is 17.545. Also, the residual plots (fig. 5) show that the autocorrelation and partial autocorrelation coefficients all lie inside the confidence limits, which shows that the fit is a good fit. The graph (fig. 6) shows the observed and fitted val-

ues obtained through SARIMA Model for the monthly rainfall in Tamilnadu.

5. Conclusion

Study reveals that the time series model SARIMA is a very useful tool for prediction of rainfall. This is evidenced from the following figure 6 and also seeing the MAPE value which is very less.

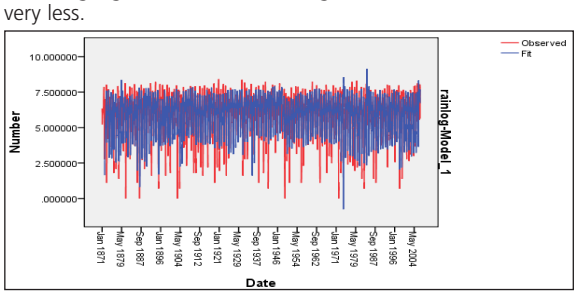


Fig.6: Graph showing observed and fitted values through the model SARIMA (0,1,1) X (0,1,1) $_{12}$

REFERENCES

Balachandran S., Asokan R. and Sridharan S (2006). Global surface temperature in relation to northeast monsoon rainfall over Tamilnadu, Journal of Earth System Science, vol. 115, no.3, pp. 349-362. | [2] Fallah G-A , Mousavi M. Habiby M, (2008), Using Artificial neural network (ANN) for seasonal rainfall forecasting based on tele-connection patterns, Geophysical Research Abstracts, Vol. 10, EGU 2008 -A- 11664. | [3] Gadgil, S., Rajeevan, M. & Nanjundiah, R., (2005) Monsoon prediction – Why yet another failure? Current Science 88: 1389-1499. | [4] George E. P. Box, Gwilym M. Jenkins and Gregory C. Reinsel (1994). Time Series Analysis – Forecasting and Control, 3rd Edition, Pearson Education, | [5] Hung N Q, Babel M S, Thailand, Weesakul S, and Tripathi N K, An artificial neural network model for rainfall forecasting in Bangkok, Hydrology and Earth System Sciences discussions, vol. 5, pp.183–218, 2008. | [6] Khaing Win Mar and Thinn Thu Naing (2008), Optimum Neural Network architecture for Precipitation Prediction of Myanmar, Proceedings of World Academy of Science, Engineering and Technology, vol.36, December. | [7] Meaza Demissie (2003), Characterizing trends and Identification of rainfall predicting model: An Application to Melksa observatory Nazreth rainfall series, Proceedings of Biometric society, pp. 132 – 134. | [8]. Moller A F(1993), A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning, Neural Networks, Vol.6, pp.525 – 533. | [9]. O. N. Dhar, P. R. Rakhecha, A. K. Kulkarni (1982). Fluctuations in northeast monsoon rainfall of Tamilnadu, International Journal of Climatology Volume 2, Issue 4, Pages339 – 345. | [10]. O. N. Dhar, P. R. Rakhecha, A. K. Kulkarni (1982). Trends and fluctuations of seasonal and annual rainfall of Tamilnadu, Proceedings of Indian Academy of Sciences (Earth planetary Sciences), vol. 91, No.2, pp. 97 –104. | | [11]. O. N. Dhar and P. R. Rakhecha (1983). Foreshadowing Northeast Monsoon Rainfall over Tamilnadu, Monthly Weather Review, vol.3, issue 1,pp. |