Journal of Pharmacognosy and Phytochemistry

Available online at www.phytojournal.com

E-ISSN: 2278-4136
P-ISSN: 2349-8234
JPP 2017; 6(5): 1418-1421
Received: 01-07-2017
Accepted: 02-08-2017

Anosh Graham
Department of Environmental Sciences and NRM, College of Forestry, Sam Higginbottom University of Agriculture, Technology & Sciences (formerly Allahabad Agriculture Institute) Allahabad, Uttar Pradesh, India

Ekta Pathak Mishra
Department of Environmental Sciences and NRM, College of Forestry, Sam Higginbottom University of Agriculture, Technology & Sciences (formerly Allahabad Agriculture Institute) Allahabad, Uttar Pradesh, India

Correspondence
Anosh Graham
Department of Environmental Sciences and NRM, College of Forestry, Sam Higginbottom University of Agriculture, Technology & Sciences (formerly Allahabad Agriculture Institute) Allahabad, Uttar Pradesh, India

Time series analysis model to forecast rainfall for Allahabad region

Anosh Graham and Ekta Pathak Mishra

Abstract
The prediction of Rainfall on monthly and seasonal time scales is not only scientifically Challenging but is also important for planning and devising agricultural strategies. Various research groups attempted to predict rainfall on a seasonal time scales using different techniques. This paper describes the Box-Jenkins time series seasonal ARIMA (Auto Regression Integrated Moving Average) approach for prediction of rainfall on monthly scales. Seasonal ARIMA model (0, 0, 0) (0, 1, 0) for rainfall was identified the best model to forecast rainfall for next 5 year's with confidence level of 95 percent by analyzing last 31 year's data (1985-20015). Previous years data is used to formulate the seasonal ARIMA model and in determination of model parameters. The performance evaluations of the adopted models are carried out on the basis of correlation coefficient (R^2) and root mean square error (RMSE).The study conducted at Allahabad, Uttar Pradesh (India). The results indicate that the seasonal ARIMA model provide consistent and satisfactory predictions for rainfall parameters on monthly scale.

Keywords: rainfall, ARIMA, correlation coefficient (R^2), root mean square error (RMSE).

Introduction
The agricultural practices and crop yields of India are heavily dependent on the climatic factors like rainfall. Out of 142 million ha cultivated land in India, 92 million ha (i.e. about 65%) are under the influence of rain fed agriculture. Unlike irrigated agriculture, rain fed farming is usually diverse and risk prone. The monsoon season is the principal rain bearing season and in fact a substantial part of the annual rainfall over a large part of the country occurs in this season. Small variations in the timing and the quantity of monsoon rainfall have the potential to impact on agricultural output. Rainfall is the most important climatic element that influences agriculture. Monthly rainfall forecasting plays an important role in the planning and management of agricultural scheme and water resources systems. The main objective of the present study is to develop a valid stochastic model to simulate monthly rainfall in Allahabad region. Rainfall is a seasonal phenomenon with twelve months period. Seasonal time series are often modeled by SARIMA techniques. Recently, a few researchers modeled monthly rainfall using SARIMA methods. Nimarla and Sundaram fitted a SARIMA (0, 1, 1) x (0, 1, 1)12 model to monthly rainfall in Tamil Nadu, India. Etuk and Mohamed fitted a SARIMA (0, 0, 0)x(0, 1, 1)12 model to monthly rainfall in Gadaref, Sudan. The prediction of rainfall and temperature has been made using auto regressive integrated moving average method and is examined using data for the period of 1985-2015. The modeling and prediction of rainfall is done through the statistical method based on autoregressive integrated moving average (ARIMA). In this paper, modeling and forecasting of rainfall is made through the conventional method called box-jerkins seasonal ARIMA model.

Materials and Methods
Study area
The district of Allahabad is located between 24° 47´ N and 25° 43´ N latitudes and between 81° 31´ E to 82° 21´ E longitudes. The climate of Allahabad district is typical humid subtropical. It has three seasons: hot dry summer, warm humid monsoon and cool dry winter. The winter usually extends from mid-November to February and is followed by the summer which continues till about the middle of June. Allahabad experiences severe fog in January resulting in massive traffic and travel delays. The summer season is long and hot with the maximum temperatures ranging from 40 °C (104 °F) to 45 °C (113 °F) accompanied by hot local winds called as "loo". The monsoon season starts from mid of June to September. About 88 percent of the annual rainfall is received during the monsoon season July and August being the months of maximum rainfall.

The normal rainfall in the district is 975.4 mm. (38.40 inches) but the variation from year to year is appreciable.

Data Collection

Daily rainfall data for the past 31 years from 1985 to 2009 was collected from IMD, while data from 2010 to 2015 was collected from the agro meteorological observatory of College of Forestry, SHUATS, Allahabad, for forecasting.

Software used: SPSS

SEASONAL Auto Regressive Integrated Moving Average (ARIMA) models were selected using **SPSS** software to find the best fit of a time series to past values of this time series in order to make forecasts.

Methodology

A time series is defined as a set of observations arranged chronologically i.e. a sequence of observations usually ordered in time. The principal aim of a time series analysis is to describe the history of movements in time of some variable at a particular site. The objective is to generate data having properties of the observed historical record. To compute properties of a historical record, the historical record or time series is broken into separate components and analyzed individually to understand the casual mechanism of different components. Once properties of these components are understood, these can be generated with similar properties and combined together to give a generated future time series. Analysis of a continuously recorded rainfall and temperature data time series is performed by transforming the continuous series into a discrete time series of finite time interval.

Mathematical modeling of rainfall data is a stochastic process. Several mathematical models based on the probability concept are available. These models help in knowing the probable weekly, monthly or annually rainfall. Over the past decade or so, a number of models have been developed to generate rainfall and runoff. Monthly rainfall and temperatures were analyzed using time series analysis. Time series models have been extensively studied by Box and Jenkins (1976) [1] and as there names have frequently been used with synonymously with general ARIMA process applied to time series analysis and forecasting.

Box and Jenkins (1976) [1] have effectively put together in a comprehensive manner, the relevant information required to understand and use time series ARIMA models. A detailed strategy for the construction of linear stochastic equation describing the behavior of time series was examined. Consider the function Z_t represents forecasted rainfall and temperature at time t month. Y_t is series of observed data of rainfall and temperature at time t. If series is stationary then a ARIMA process can be represented as

$$\nabla pZ_t = \nabla qY_t \;....(1)$$

Where ∇ is a back shift operator. If series Y is not stationary then it can be reduced to a stationary series by differencing a finite number of times.

$$\nabla pZ_t = \nabla q\,(1\text{-}B)\,d\,Y_t \;....(2)$$

Where d is a positive integer, and B is back shift operator on the index of time series so that

$BY_t= Y_t\text{-}1$; $B2Y_t = Y_t\text{-}2$ and so on. Thus further equation (2) can be simplified into following equation.

$(1\text{-}\Phi1B\text{-}\Phi2B2\text{-}.......\text{-}\Phi pBp)\ Zt = \theta0+ (1\text{-}\theta1B\text{-}\theta2B2\text{-}\ ....\text{-}\theta qBq)\ at$ ..... (3)

Where a_t's a sequence of identically distributed uncorrelated deviates, referred to

as "white noise". Combining equations (2) and (3) yields the basic Box-Jenkins models for non stationary time series

$(1\text{-}\Phi1B\text{-}\Phi2B2\text{-}.......\text{-}\Phi pBp)\ (1\text{-}B)d\ \ Yt =\theta0+ (1\text{-}\theta1B\text{-}\theta2B2\text{-}.........\text{-}\theta qBq)\ at$ ....(4)

Equation (4) represents an ARIMA process of order (p,d,q). Seasonal ARIMA model represented as follows for a stationary series i.e. differencing parameters (d &$d_s = 0$) equal to Zero, used for forecasting rainfall and temperature.

$$\nabla_{ps} \nabla^{p} Z_t = \nabla_{qs} \nabla^{q} Y_t \;....(5)$$

Where p_s and q_s are the seasonal parameters corresponding to AR and MA process. Model of type of equation (5) was fitted to given set of data using an approach consists of mainly three steps (a) identification (b) estimation (c) application (forecasting) or diagnostic checking. At the identification stage tentative values of p,d,q and p_s,d_s,q_s were chosen. Coefficients of variables used in model were estimated. Finally diagnostic checks were made to determine, whether the model fitted adequately describes the given time series. Any inadequacies discovered might suggest an alternative form of the model, and whole iterative cycle of identification, estimation and application was repeated until a satisfactory model was obtained.

Results and Discussion

The model that seems to represent the behaviour of the series is searched, by the means of autocorrelation function (ACF) and partial auto correlation function (PACF), for further investigation and parameter estimation. The behaviour of ACF and PACF is to see whether the series is stationary or not. For modelling by ACF and PACF methods, examination of values relative to auto regression and moving average were made. An appropriate model for estimation of monthly rainfall for Allahabad station was finally found. Many models for Allahabad station, according to the ACF and PACF of the data, were examined to determine the best model. The model that gives the minimum Bayer's Information Criterion (BIC) is selected as best fit model, as shown in Table 1. Obviously, model SARIMA (0, 0, 0) (0, 1, 0) has the smallest values of BIC. Observed and predicted values of next fives years are determined and plotted as shown in figure: 4.

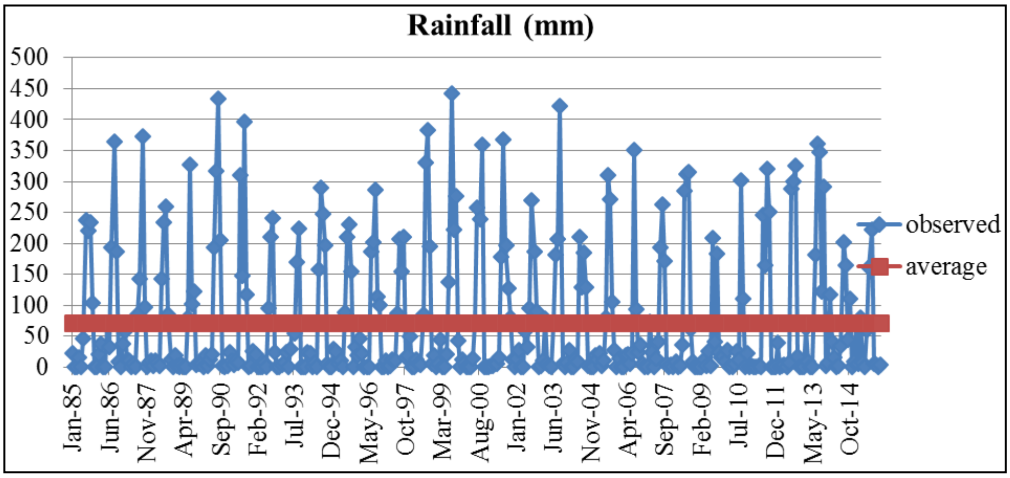

Fig 1: Observed rainfall in Allahabad district (Jan. 1985-Dec. 15)

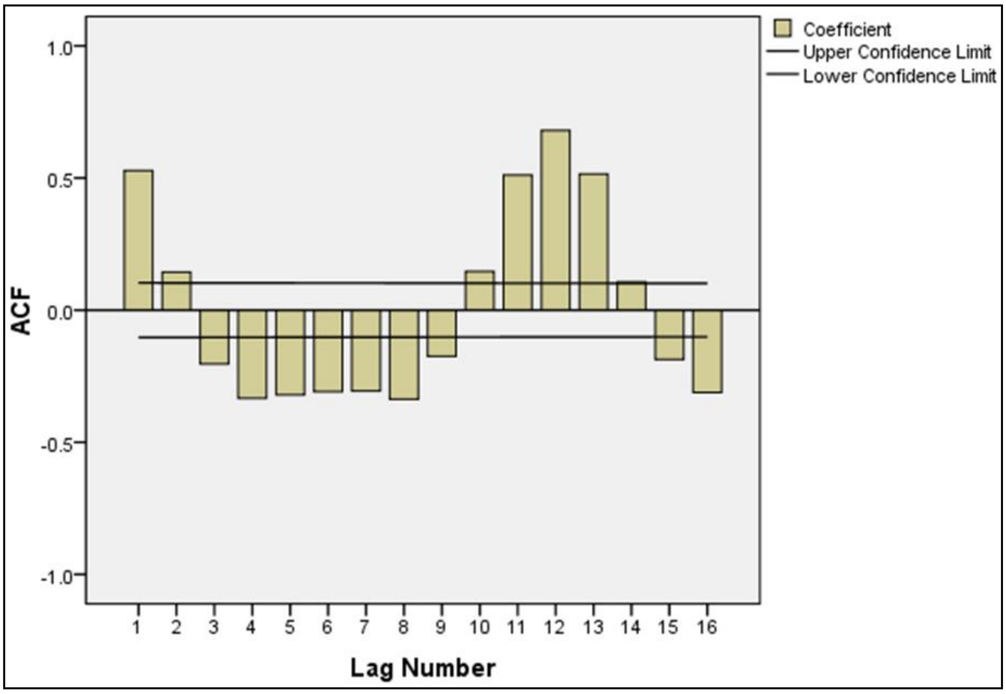

Fig 2: Autocorrelation function of rainfall.

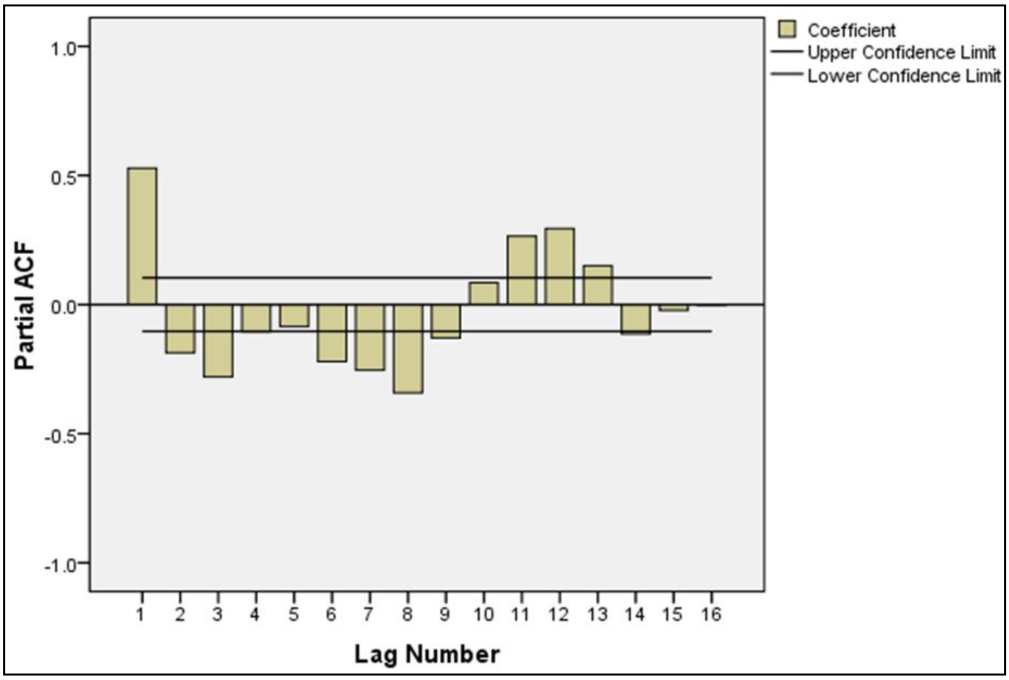

Fig 3: Autocorrelation function of rainfall.

Table 1

Parameters	Rainfall model Sarima (0,0,0) (0,1,0)
R^2	0.384706
RMSE	83.132

In the analysis of observed data there are many missing values which results into less accuracy of predicted model in case of rainfall.

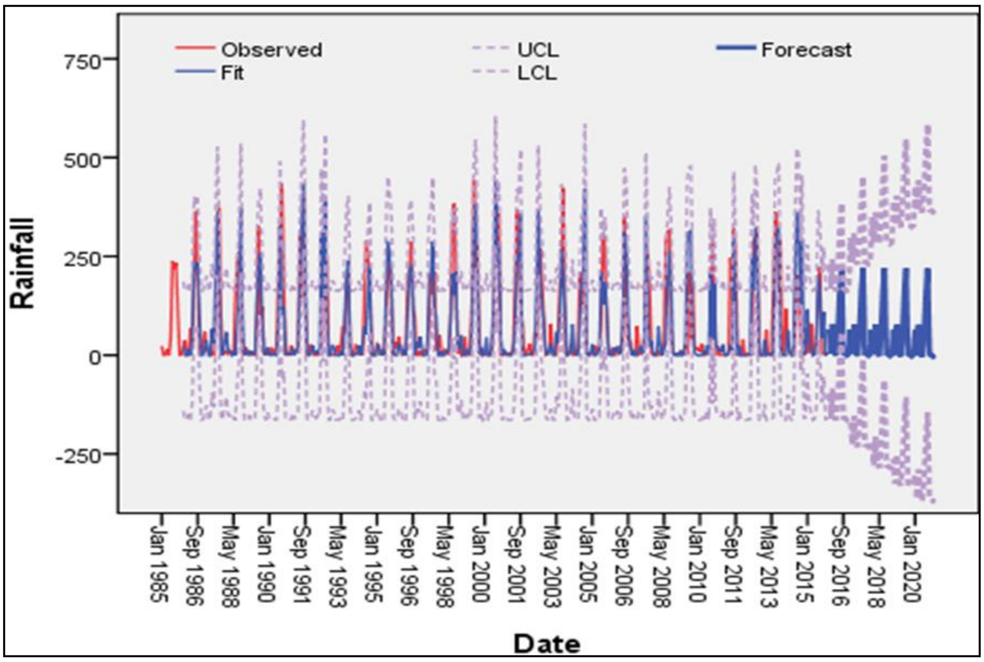

Fig 4: Observed and fitted values of rainfall series.

Conclusion

The Box-Jenkins ARIMA methodology was used to develop monthly rainfall of Allahabad. The monthly rainfall is panning over the period of 1985-2015 at Allahabad. The performance of resulting ARIMA model was evaluated by using the data from the year 1985-2015 through graphical comparison between the forecasted and observed data. In ARIMA model the forecasted and observed data of rainfall showed good results. The study reveals that Box-Jenkins methodology can be used as an appropriate tool to forecast rainfall in Allahabad for upcoming years. The accuracy of predictions made for rainfall by seasonal ARIMA model is less because data is abrupt and contains many missing values, which increases white noise in the system; because no value is assigned to missing values and are assumed zero. Accuracy of these predictions can be increased in future by using these predicted values for missing values.

Reference

1. Box GEP, Jenkins GM. Time series analysis: forecasting and control, Prentice Hall, Inc, 1976, 575.
2. Nirmala M, Sundaram SM. A Seasonal Arima Model for Forecasting monthly rainfall in Tamil Nadu. National. Journal on Advances in Building Sciences and Mechanics. 2010; 1(2):43-47.
3. Etuk EH, Mohamed TM. Time Series Analysis of Monthly Rainfall data for the Gadaref rainfall station, Sudan, by SARIMA Methods. International Journal of Scientific Research in Knowledge. 2014; 2(7):320-327.
4. District Statistical Bulletin, Allahabad. Department of Statistics and Economics, Directorate of Statistics and Planning, Uttar Pradesh, Vikas Bhawan, Allahabad, 2011, 46-47.
5. Kaushik I, Singh SM. Seasonal ARIMA model for forecasting of monthly rainfall and temperature. International Journal of Agriculture Sciences. 2008; 5(8):112-25.
6. Mohamed TM, Ibrahim A. Time Series Analysis of Nyala Rainfall Using ARIMA Method Journal of Engineering and Computer Science (JECS). 2016; 17(1):5-11.