International Journal of Engineering & Technology, 7 (4.10) (2018) 1085-1088

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper



Predictive Modelling of Rainfall Data for Aurangabad Region by using ARIMA Method

Rupali D. Patil[1]*, O. S. Jadhav[2]

[1]Research Scholar
[2]Assistant Professor
**Corresponding author E-mail:Rupali.stat@gmail.com*

Abstract

In day to day life prediction of weather parameters play important role for planning in environment management fields. Predictive modelling of rainfall is very necessary criterion for water management resources. In time series analysis most effective autoregressive integrated moving average model (ARIMA) are used. In this paper, we set up various ARIMA models for prediction of rainfall from which ARIMA (1, 0, 0) (2, 0, 0)^{12} are best fitted models for future forecasting purpose. Forecasting the next three years has been described to decide water demand management priorities.

Keywords: *ARIMA; ARMA; Forecasting; Rainfall parameter; Time series.*

1. Introduction

Now a day's climate issue are becomes increasing seriously because of industrialization of world which cause great economic loss to mankind. Recently experts conducting research continuously in metrological area, so it is collected in large numbers. People get vast of meteorological forecasting knowledge through reviewing and brief experience from this document. The computer technology develops continuously in metrological area. These characteristics give to the informational structure of metrological sources. Mainly, data collected from two main stations viz. surface and aerial stations, there are number of weather stations, in which daily observed, acquired and proceed with an exponential manner [8, 9]. All the meteorological data primarily kept in a database file for examination to creation of weather forecasting. Weather prediction is one the most interesting and fascinating area and it plays significant role in metrology. There are several limitations in weather forecasting that implementation and used advanced techniques to get suitable as well as useful prediction due to some actively changes in atmosphere [2].
In present study, we tend to analyse the data of rainfall parameter and prediction for a particular period by using the strategy of Autoregressive Integrated Moving Average (ARIMA). In this study open source statistical R software used with their various time series packages such as 'tseries', 'UnitRoots', 'forecast' and 'TTR' etc for good prediction purpose of rainfall.

2. Review of Literature

It's most necessary and demanding working responsibilities accepted by metrological services all over the world.
V. K. Somvanshi et al. studied modelling tools and predicting the behavioural pattern of rainfall phenomena by using past observation. Forecasting the rainfall with two fundamental prediction techniques like ARIMA and artificial neural network (ANN) model and estimate the prediction efficiency for 104 years of mean annual rainfall observation at Hyderabad region (India) [3].
Md. Mahsin et al. construct the ARIMA model of thirty years rainfall data collected from Dhaka stations i.e. total 354 reading by using Box-Jenkins method. This study results forecasting next two year monthly rainfall to help the future water demand management [4].
Ette Harrison Etuk, Tariq Mahgoub Mohamed determined the seasonal ARIMA (0, 0, 0) (0,0,1)^{12} model of monthly rainfall at Gadaref, Sudan [5].
Metrine Chonge et al. fitted the ARIMA (0,0,0) (0,1,2)^{12} model of historical rainfall pattern of Uasin Gishu country, by using this model to predict two year expected average monthly rainfall. The data analysis and flow of data handle by using R software [6].

3. Research Methodology

3.1. Source of data

The present study collected month wise ten year rainfall data from WALMI (Water and Land Management Institute) for Aurangabad division, Maharashtra state. The data consists of 120 monthly observations from January 2006 to December 2015 at ten year. In this study, we used R software for the purpose of data analysis.

3.2 ARIMA model

In time-series analysis (Box-Jenkins models) ARIMA model is an important forecasting tool. An autoregressive model of order p is predictably classified as AR (p) and a moving average model with q as MA (q). A combined model that contains p autoregressive terms and q moving average terms is called ARMA (p, q) (Gujarati, 1995). The differenced d times to achieve stationary, the model is classified as ARIMA (p, d, q), where the symbol "I" stands for "integrated". ARIMA model is a arrangement of an autoregressive (AR) and a moving average (MA) process applied to a non-

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

stationary data series. The general non-seasonal ARIMA (p, d, q) model is as:
AR: p = parameter of the autoregressive,
I: d = difference and
MA: q = parameter of the moving average.
General equation of ARIMA (p, d, q) model is as follows:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - ..... - \theta_p e_{t-p} \quad (1)$$

Or in backshift notation

$$\left(1 - \phi_1 B - \phi_2 B^2 - .... - \phi_p B^p\right) Y_t = c + \left(1 - \theta_1 B - \theta_2 B^2 - .... - \theta_q B^q\right) e_t \quad (2)$$

Where,
c = constant,

ϕ_i =parameter of p[th] autoregressive,

θ_j = parameter q[th] moving average,

e_t = error term at time t and

B_k = K[th] order backward shift operator.

In present study seasonal ARIMA model perform on rainfall data. The parameters are: seasonal autoregressive (P), seasonal differencing (D) and seasonal moving average (Q). The general form of the Seasonal ARIMA (p, d, q) (P, D, Q) ^{S} model using backshift notation is given by,

$$\phi_{AR}(B)\phi_{SAR}(B^S)(1-B)^d\left(1-B^S\right)^D Y_t = \theta_{MA}(B)\theta_{SMA}(B^S)e_t \quad (3)$$

Where, s = Seasonal period term,

ϕ_{AR} = Parameter of non-seasonal autoregressive,

θ_{MA} = Parameter of non-seasonal moving average,

θ_{SMA} = Parameter of seasonal moving average [4].

4. Result and Discussion

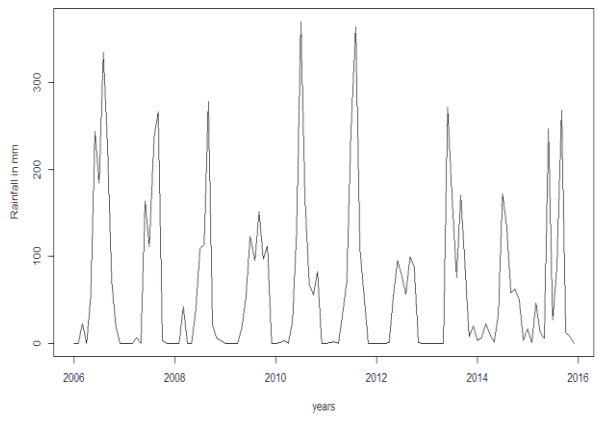


Fig.1: Line plot of monthly rainfall in mm against the year

The Figure 1 shows the line plot of monthly rainfall data fluctuates against the years, depicting no trend. This implies that the series is stationary. After again we confirm that rainfall series data are stationary or not by using the Augmented Dickey Fuller (ADF) test. We suppose null hypothesis (H_0) the series is non-stationary while alternative hypothesis (H_1) the series is stationary. The ADF test result, as obtained:
Augmented Dickey-Fuller = -5.37, Lag order = 4, p-value = 0.01

We conclude that, the reject H_0 at 5% level i.e. the series is stationary in its value of mean and variance. Thus, there is no need of differencing the time series plot and we adopt D = 0 for our ARIMA model.
From the Figure 2 observe that autocorrelation function (ACF) and partial autocorrelation function (PACF) of original rainfall data in which the presence of seasonality behaviour as well as stationary. Also examine the strong seasonal interval (S=12) and difference (d=0), of the original data in order to have stationary series. An ARIMA model was identified above, the p, q, P and Q parameters need to identify for our model.

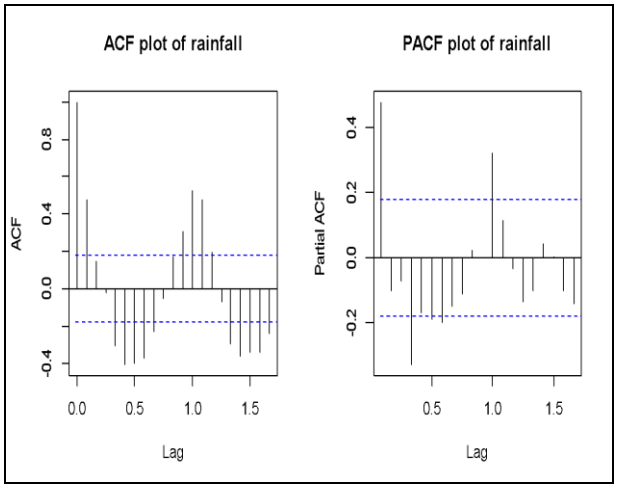


Fig.2: Plot of ACF and PACF for original rainfall data

The Table 1 gives parameter estimation and their standard errors for the $ARIMA(p,d,q)(P,D,Q)^{12}$ model.

Table 1: Parameter Estimation for $ARIMA(0,0,1)(2,0,0)^{12}$

Coefficients	Estimation	S.E
θ_{1MA}	0.2939	0.1087
ϕ_{1SAR}	0.4425	0.1087
ϕ_{2SAR}	0.1959	0.1033

σ_e^2 = 4575: log likelihood = -676.62, AIC=1363.23, AICc=1363.76, BIC=1377.17
Figure 3, show the plot of the diagnostics of standardized residuals, the ACF of the residuals; calculated model good fits the data well.

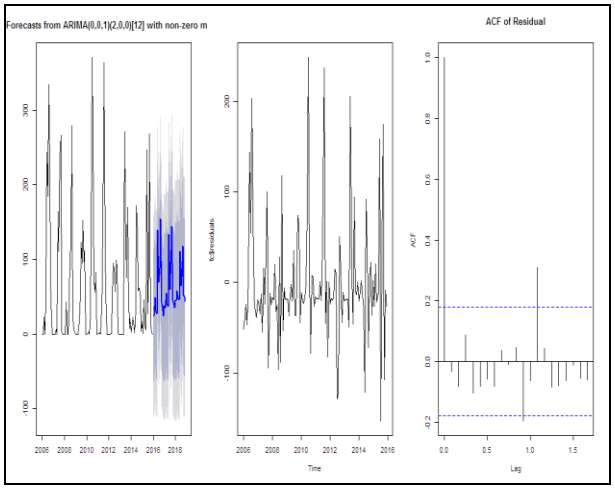


Fig.3: Diagnostics for the $ARIMA(0,0,1)(2,0,0)^{12}$ fit on the rainfall data

After that to test residual of fitted model come from normally distributed series or not, by using the Shapro-wilks test. Shapro-

wilks test also confirms this fact from the P-value 8.876e-10 while it is small than 0.05. Then determine the residual from the fitted model do not come from a normal distribution. Also from the Figure 3 observed that the AFC some apparent and pick few lag out of significant limit. This therefore, estimated $ARIMA(0,0,1)(2,0,0)^{12}$ model not good for prediction, at that moment some transformation used for data set of rainfall series to again fulfil the criterion of normality.

At that time we use the Box-Cox Transformation for rainfall time series, with optimum value $\lambda = 0.5$. Hence we use maximum likelihood estimation method for transformed total monthly rainfall to estimate the parameter.

Table.2: Parameter Estimation for $ARIMA(1,0,0)(2,0,0)^{12}$ with Box-Cox Transformation

Coefficients	Estimation	S.E
ϕ_{1AR}	0.2097	0.1005
ϕ_{1SAR}	0.5210	0.1009
ϕ_{2SAR}	0.2457	0.1030

σ_e^2 =51.57, log likelihood=-409.48, AIC=828.96 AICc=829.49 BIC=842.9

The above Table 2 presents the variance estimated value, standard error, Bayesian information criterion (BIC) and Akaike information criterion (AIC) which are necessary to the model. The AIC, BIC, variance of the estimate σ_e^2 are used to decide best fitted selected models of monthly rainfall observations. The value of information criteria is small then the determined model best fit the future value of rainfall series. Since ARIMA $(1,0,0)(2,0,0)^{12}$ model has lower AIC, BIC and estimate variance than other fitted models. The selected model is best fit and accuracy in forecasting.

If the model fits the data well, the residual of the fitted model are random. The diagnostic statistics and plots of the residuals can be used to examine the goodness of fit of the selected model to the data.

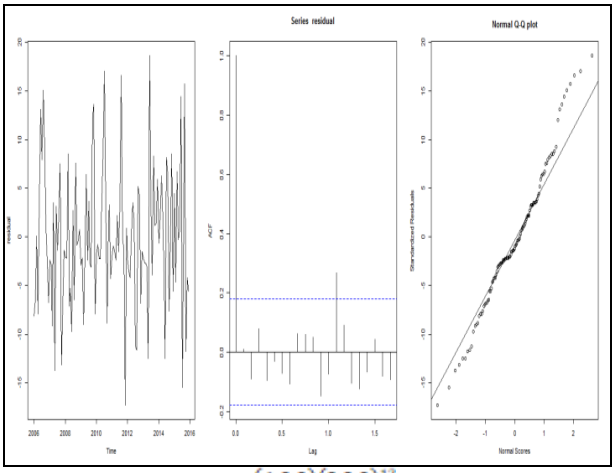


Fig.4: Diagnosis of ARIMA $(1,0,0)(2,0,0)^{12}$ with Box-Cox Transformation

From the Figure 4, shown that no any particular trend in the residual autocorrelation plot for the fitted models i.e. the autocorrelation coefficient within the significant limit of series. Also the Shapiro-Wilk test result (p-value = 0.358) maintain the normality of transformed series.

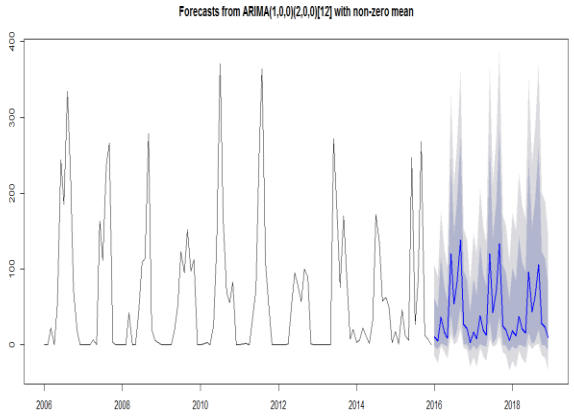


Fig.5: Plot of Forecast ARIMA $(1,0,0)(2,0,0)^{12}$ with Box-Cox Transformation

Figure 5, shows prediction of three years rainfall from the past rainfall data and developed the seasonal ARIMA model for the period 2006 to 2015, while its error will be relatively minor. The accuracy of forecast model is regularly measured using performance measurement such as root mean square error (RMSE), mean absolute error (MAE), Mean absolute percentage error (MAPE).

5. Conclusion

Box and Jenkins proposed the autoregressive integrated moving average used in this paper to examine the structure and forecasting the rainfall series. At the present study several model find out for particular data set but best model may be selected on the basis of smallest value of AIC as well as BIC. In this study, ARIMA (1, 0, 0) (2, 0, 0)^{12} model create to predict the three year monthly rainfall data. The selected model can be prediction of upcoming rainfall in Aurangabad region at Maharashtra state.

Acknowledgement

I acknowledged my honest and deep appreciation towards my guide, Dr. Omprakash S. Jadhav, for his valuable guidance, enthusiastic concentration and support of this work. And I also recognize my sincere gratitude to authorities of Dr. Babasaheb Ambedkar Marathwada University, Aurangabad and other teaching staff of Statistics department for their valuable support and guidance. I am pleased to Water and Land Management Institute, Aurangabad for providing me weather data and also thankful to my friends for their cooperation.

References

[1] Jain, G., & Mallick, B., "A Study of Time Series Models ARIMA and ETS", *I.J. Modern Education and Computer Science*, Vol.4, (2017), pp. 57-63.

[2] Basvanth Reddy,B.A Patil, "Weather Prediction Based on Big Data Using Hadoop Map Reduce Technique", *International Journal of Advanced Research in Computer and communication Engineering*, Vol.5, No. 6, (2016), pp. 2278-1021.

[3] Somvanshi, V. K., Pandey, O. P., Agrawal, P. K., Kalanker, N. V., Prakash, M. R., & Chand, R, Modeling and prediction of rainfall using artificial neural network and ARIMA techniques. *J. Ind. Geophys. Union* , Vol.10,No.2, (2006), pp. 141-151.

[4] Mahsin, M., "Modeling rainfall in Dhaka division of Bangladesh using time series analysis", *Journal of Mathematical Modelling and Application*, Vol. 1, No.5, (2011), pp. 67-73.

[5] Etuk, E. H., & Mohamed, T. M., "Time series analysis of monthly rainfall data for the Gadaref rainfall station, Sudan, by SARIMA

methods", *International Journal of Scientific Research in* Knowledge, Vol.2, No.7, (2014), pp. 320-327.
[6] Metrine, C., Kennedy, N., Omukoba, M., Lucy, M., & Frankline, T., "A time series model of rainfall pattern of Uasin Gishu County", *IOSR Journal of Mathematics (IOSR)*, Vol.11,No. 5, (2015), pp. 77-84.
[7] Tularam, G. A., & Ilahee, M., "Time series analysis of rainfall and temperature interactions in coastal catchments", *Journal of Mathematics and Statistics*, Vol. 6, No.3, (2010), pp. 372-380.
[8] Y.W.Dou, L. Lu, X. Liu and Daiping Zhang, "Metrological data storage and management system", *Computer Systems and Application* , Vol.20, No. 7, (2011), pp. 116-120.
[9] Li, L., Ma, Z., Liu, L., & Fan, Y., "Hadoop-based Arima algorithm and its application in weather forecast",. *International Journal of Database Theory and Application*, Vol. 6, No.5, (2013), pp. 119-132.
[10] https://en.wikipedia.org/wiki/Ljung%E2%80%93Box_test