ISSN No. 2395-0900

IJESTA International Journal of Engineering Studies and Technical Approach

Development of a Monthly Rainfall Prediction Model Using Arima Techniques in Krishnanagar Sub-Division, Nadia District, West Bengal

Alivia Chowdhury[*] and Amit Biswas[**]

[*]*Assistant Professor, Department of Soil and Water Engineering, Faculty of Agricultural Engineering Bidhan Chandra Krishi Viswavidyalaya (BCKV), Mohanpur, West Bengal, India*

[**]*Research Scholar, Agricultural & Food Engineering Department, IIT Kharagpur, West Bengal, India*

ABSTRACT

Hydrological parameters are complicated phenomena, which require modern computer modelling and simulation techniques for accurate forecasting. Thorough understanding of the rainfall process is critical for the proper planning of integrated water resources management projects which relates agriculture, climate change, and natural hazards such as floods and droughts. In the present study, an autoregressive integrated moving average (ARIMA) model was developed in MATLAB environment for simulating and forecasting the rainfall data for the study area using Box-Jenkins methodology. Rainfall data from 1971-2000 (30 years) were used for the development of the model while data from 2001 to 2010 was used to verify the developed model. Four basic steps i.e. model identification; estimation of model parameters; diagnostic checking for the identified model appropriateness for modeling and application of the model (i.e. forecasting) as suggested by Box and Jenkins were followed in the present study to develop a suitable rainfall prediction model for the study area. The ARIMA model $(1,0,0)(1,1,1)^{12}$ was found to be the most suitable for forecasting monthly rainfall over the study area. The selected model can also be used for forecasting the monthly rainfall data for the up-coming years.

Keywords: Rainfall, time series, ARIMA, Uni-variate autoregressive parameters

INTRODUCTION

The agricultural practices and crop yields of India were heavily dependent on the climatic factors like rainfall. Prior knowledge of monsoon behaviour helped Indian farmers and also policy makers, to take advantage of good monsoons and also to minimize crop damages and human hardship during adverse monsoons. Rainfall is also a major factor for planning and management of irrigation projects and agricultural production such as reservoir operation, irrigation area, and irrigation water requirement. As a result, forecasting of monsoon on time scales of daily, weekly and monthly has become a major scientific issue in the field of monsoon meteorology. However various statistical methods are often useful to predict rainfall (Bisgaard and Kulahci, 2011). Among which the most effective approaches for analyzing time series data is the model introduced by Box and Jenkins (1976) and modified by Box, et al. (1994), also known as ARIMA (Autoregressive Integrated Moving Average). ARIMA has widely been exercised over the years to predict the rainfall trend (Delleur and Kavvas 1978, Thapaliyal 1981, Weesakul and Lowanichchai 2005, Somvanshi et al. 2006, Kaushik and Singh 2008, Momani et al.2009, Shamsnia et al. 2011, Mahsin et al. 2012,

ISSN No. 2395-0900
IJESTA International Journal of Engineering Studies and Technical Approach

Janhabi and Jha (2013)), The method has some interesting features that made it more desirable for researchers. It eases the forecasting process allowing researchers to use only single variable time data series while also allow multiple for more complex cases.Time-series analysis became a major tool in this context. Rainfall prediction modelling involved a combination of computer models, observation and knowledge of trends and patterns. Many methods and approaches for formulating forecasting models were available in the literature dealing with time series forecasting, the most relevant is Box-Jenkins methodology, in particular, the Auto Regressive Integrated Moving Average (ARIMA) method. It allowed the manager who had only data on past years' quantities, rainfall as an example, to forecast future ones without having to search for other related time series data, for example temperature. In forecasting and analysis of time series data, it is well demonstrated that ARIMA was very effective in handling practical applications. With this background, the present study aims at developing a rainfall prediction and forecasting model using ARIMA techniques using 40 (1971-2010) years' daily rainfall data for Krishnanagar rain gauge station.

DATA COLLECTION AND STUDY AREA DESCRIPTION

Daily rainfall data of 40 years (1971-2010) has been collected for Krishnanagar weather station. The rainfall data were obtained from the records of the Indian Meteorological Department (IMD), Alipore, Kolkata. The weather station is located in Krishnagar Sub-dividion, Nadia District, West Bengal at Latitude 23^0 $24^{'}$ N and Longitude 88^0 $33^{'}$ E (Fig. 1). This Subdivision is the administrative headquarter of Nadia district and situated on the southern banks of the Jalangi river. Krishnanagar lies between the Tropic of cancer and the Tropic of Capricorn. The climate of the sub-division is characterized by oppressively hot summer, high humidity and high rainfall during the monsoon. Winter starts from the middle of November which continues up to the end of February. During the monsoon from June to September about 71% of annual rainfall occurs. The rainiest month is August. The mean annual rainfall is about 1400 mm. Average elevation of the area is 14 m from mean sea level.

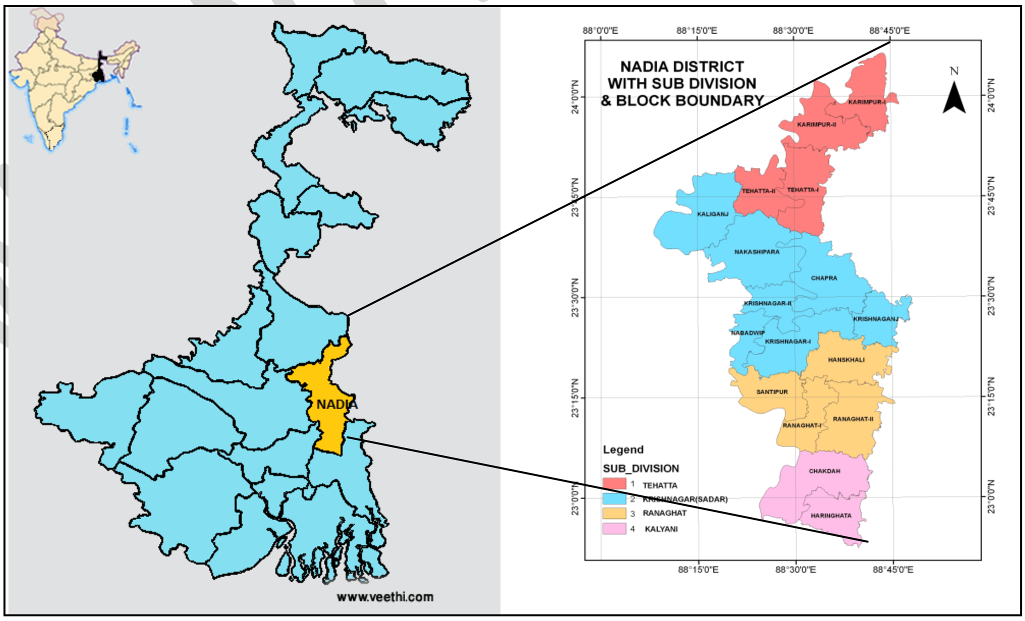


Fig-1. Study area map (Source: www.veethi.com, NRDMS)

International Journal of Engineering Studies and Technical Approach
ISSN No. 2395-0900

DEVELOPMENT OF ARIMA MODEL

The present study used Autoregressive Integrated Moving Average (ARIMA) model, proposed by Box and Jenkins (1976), for model building and forecasting the rainfall data. The Box and Jenkins methodology is a powerful approach to the solution of many forecasting problems (Johnson and Montgomery, 1976) and it can provide extremely accurate forecasts of time series and offers a formal structured approach to model building and analysis. To identify a perfect ARIMA model for a particular time series data, Box and Jenkins (1976) proposed a methodology that consists of four phases: i) Model identification; ii) Estimation of model parameters; iii) Diagnostic checking for the identified model appropriateness for modeling and iv) Application of the model (i.e. forecasting). ARIMA or Box-Jenkins method is usually presented as ARIMA(p, d, q) × (P, D, Q)S, with p = non-seasonal Auto Regressive (AR) order, d = non-seasonal differencing, q = non-seasonal Moving Average (MA) order, P = seasonal AR order, D = seasonal differencing, Q = seasonal MA order, and S = time span of repeating seasonal pattern (Box et al. 1994). In the present study MATLAB (2012a) software was used for the development of the ARIMA model. A general flowchart representing the application of Box-Jenkins methodology is given in figure below (Fig. 2).

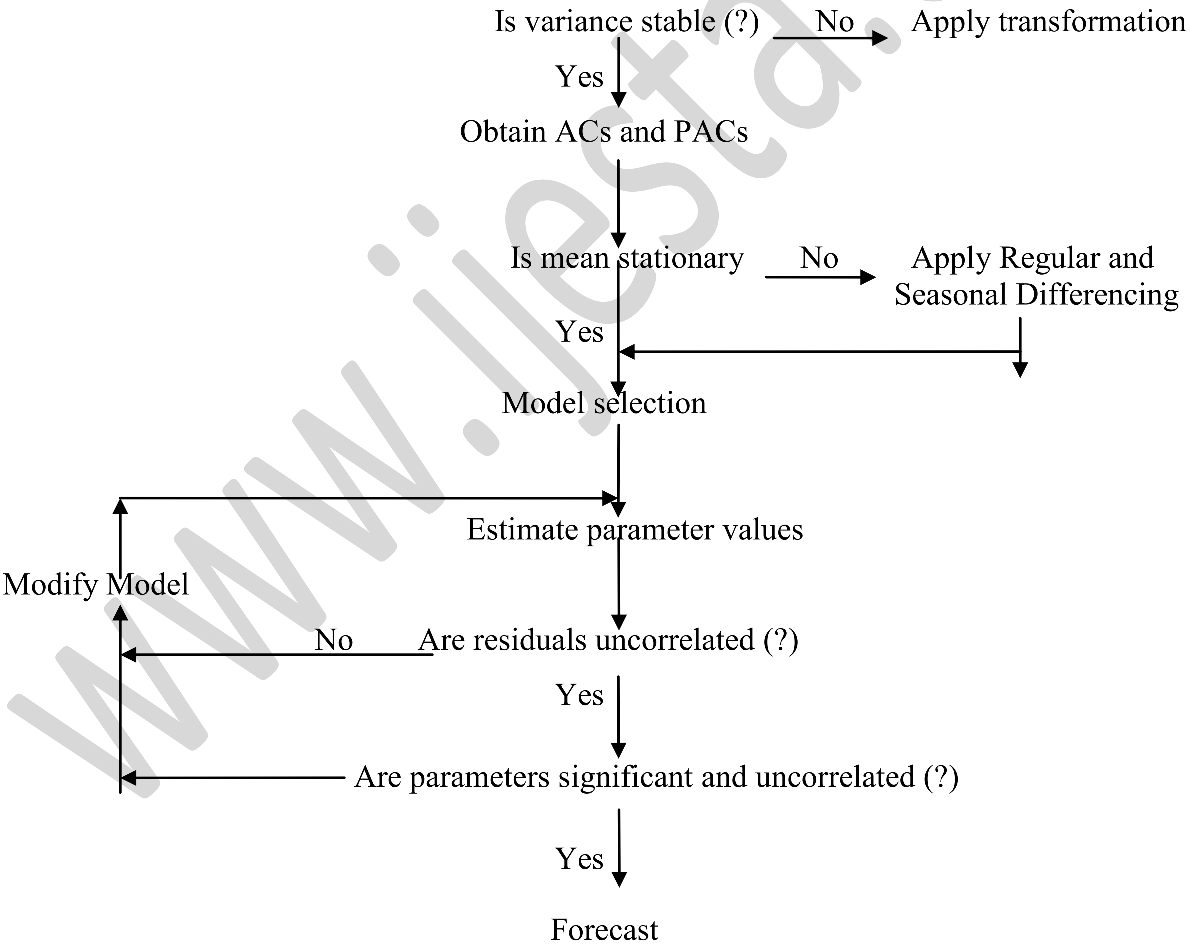


Fig-2: Box and Jenkins Methodology

RESULTS AND DISCUSSION

Time Series Analysis with ARIMA Model

(i) Model Identification

The foremost step in the process of modelling is to check for the stationarity of the series as the estimation procedures are only available for stationary series. In this study, a total of 480 monthly rainfall data point obtained during the period of 1971-2010 were used for the analysis. The plot of monthly rainfall (Fig. 3) shows that there is a seasonal cycle of the series with periodicity of one year (twelve month) in the data. A cursory look at the data structure of autocorrelation and partial autocorrelation coefficients may provide clues for the stationarity in the data series. The ACF and PACF plots of the original data prior to any transformation and differencing (Fig. 4 (a and b) confirmed the presence of stationarity in the data series.

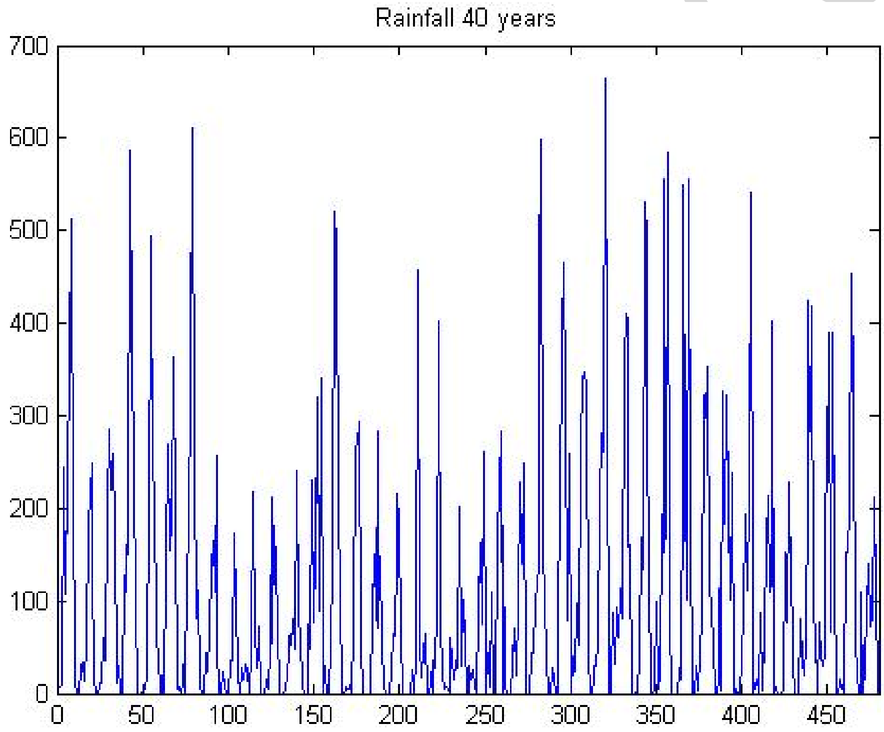


Fig-3. Monthly rainfall data for Krishnagar Weather Station for the period 1971-2010

In order to fit an ARIMA model stationary data in both variance and mean are needed. In the present study Box-Cox power transformation and 1st order seasonal differencing of the original monthly rainfall data upto lag 20, respectively was done. Thereafter, ACF and PACF for the differenced series were tested to check stationarity. The ACF and PACF (Fig. 5) show that one order seasonal differencing is adequate. Although further differencing shows a similar result but first seasonal differencing has a minimum standard deviation. Therefore, one order difference is enough for the data series. From this, a preliminary ARIMA (p, 0, q) × (P, 1, Q)12 was selected.

ISSN No. 2395-0900
IJESTA
International Journal of
Engineering Studies and Technical Approach

(a)

(b)

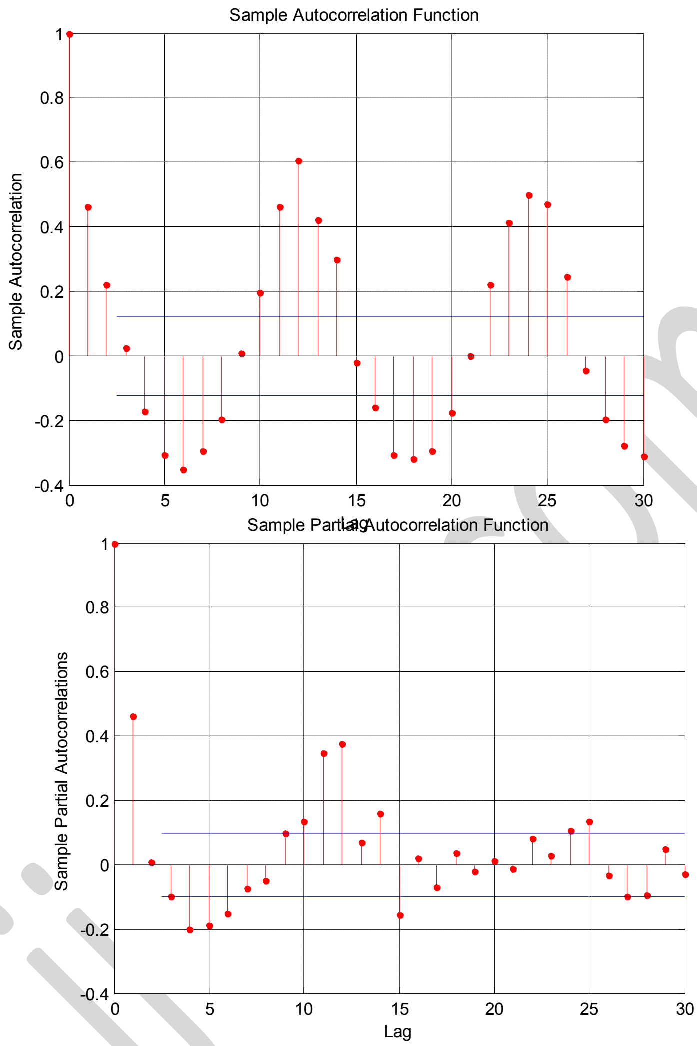
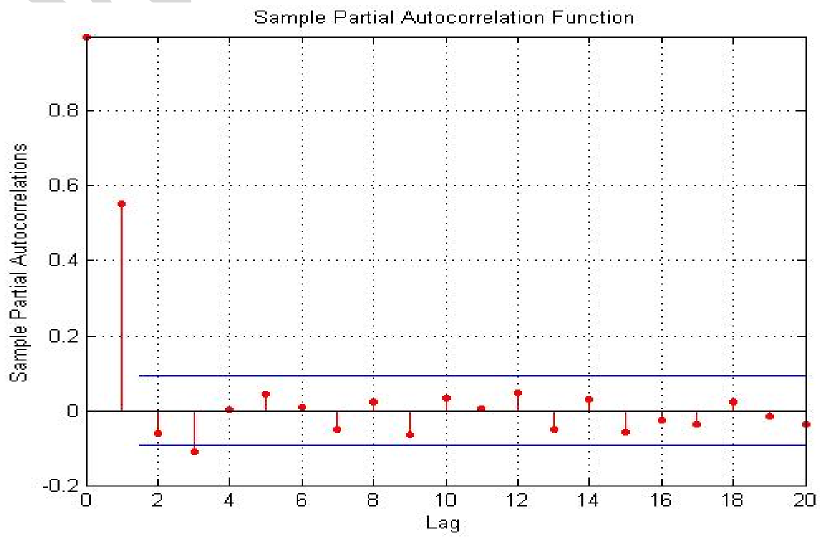


Fig-4 (a) Plot of ACF (b) Plot of PACF for observed rainfall data

(a)



(b)

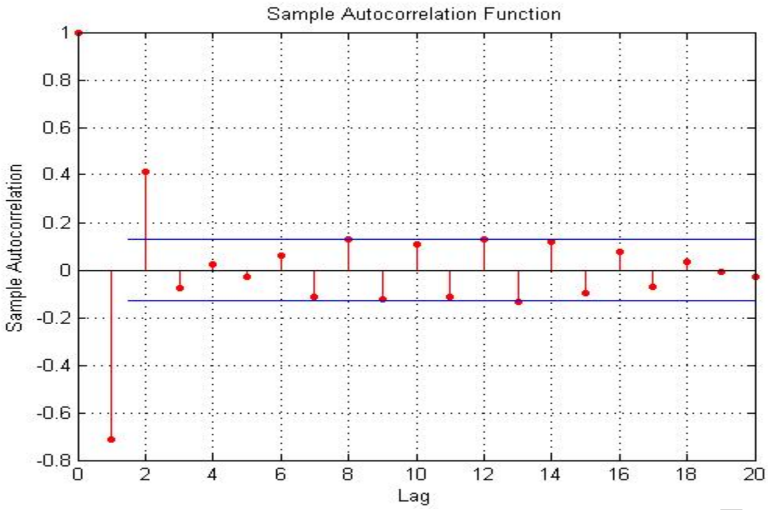


Fig- 5 (a) Plot of ACF (b) Plot of PACF for first order seasonal differencing and de-seasonalized original rainfall data

(ii) Model Selection

It is required to determine the model that best fits the data based on observing the ACF and the PACF of the differenced data. After carefully examining ACF and PACF, following nine models were identified for test. These models are: ARIMA $(1,0,1)(1,1,1)^{12}$, ARIMA $(1,0,1)(0,1,1)^{12}$, ARIMA $(1,0,1)(1,1,0)^{12}$, ARIMA $(1,0,0)(1,1,1)^{12}$, ARIMA $(1,0,0)(0,1,1)^{12}$ ARIMA $(1,0,0)(1,1,0)^{12}$, ARIMA$(0,0,1)(1,1,1)^{12}$, ARIMA $(0,0,1)(0,1,1)^{12}$, ARIMA$(0,0,1)(1,1,0)^{12}$.

(iii)Selection of Best ARIMA Model:

To choose the best model among the nine models selected above, the Akaike Information Criteria (AIC) was used. The model which had the minimum AIC value was considered as the best model for Krishnanagar rainfall analysis. Table 1 showed the best three ARIMA models for the study area with the values of AIC. It was clear that the ARIMA model $(1,0,0)(1,1,1)^{12}$ had the minimum value of AIC. Therefore, the ARIMA model $(1,0,0)(1,1,1)^{12}$ was the ultimate best model for the study area.

Table 1: ARIMA models with different numbers of parameters

Sl. No.	ARIMA model	AIC value
1	$(1,0,1)(1,1,1)^{12}$	6728.20
2	$(1,0,1)(0,1,1)^{12}$	6754.80
3	$(1,0,1)(1,1,0)^{12}$	6727.20
4	$(1,0,0)(1,1,1)^{12}$	6726
5	$(1,0,0)(0,1,1)^{12}$	6772.60
6	$(1,0,0)(1,1,0)^{12}$	6820.20
7	$(0,0,1)(1,1,1)^{12}$	6752.80

ISSN No. 2395-0900
International Journal of
Engineering Studies and Technical Approach

8	$(0,0,1)(0,1,1)^{12}$	6752.80
9	$(0,0,1)(1,1,0)^{12}$	6772.60

(iv) Goodness of Fit Test

Time series plots and normal probability plots of residuals resulting from the ARIMA model $(1,0,0)(1,1,1)^{12}$ were essential to find the existence of any correlation between the residuals.

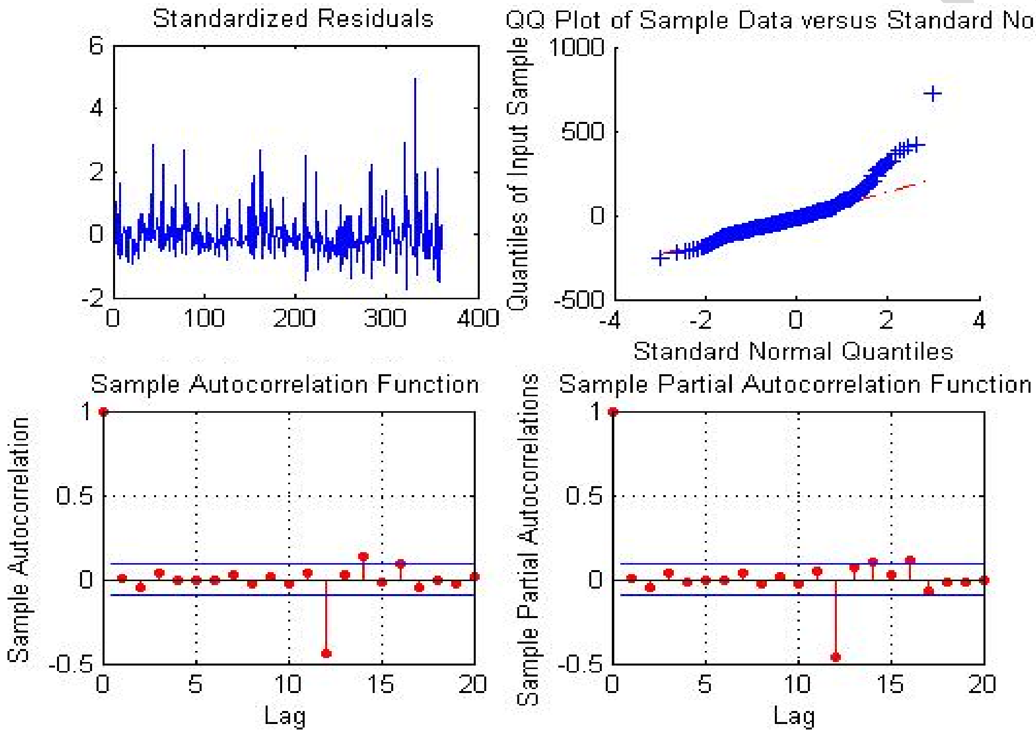


Fig-6: Time series plot of residuals for checking goodness of fit

Fig. 6 showed the time-series plots and normal probability plots of residuals resulting from the ARIMA model $(1,0,0)(1,1,1)^{12}$. Time series plots of the residuals showed a few outliers and the normal probability plots of the residuals a few departures from the expected normal value at the tail due to the outliers. The ACF and the PACF of the standardized residuals showed a few departures from the model assumptions.

Forecasting using ARIMA Model

ARIMA (1, 0, 0) $(1, 1, 1)^{12}$ was used to predict the monthly rainfall data from 2001 to 2010. Figure 8 represents the graphical plot of the observed rainfall values and the predicted rainfall values for the 120 observations obtained for the years 2001-2010. From this plot, it could be concluded that the chosen model is a good model and could be used to forecast the upcoming rainfall in Krishnagar sub-division.

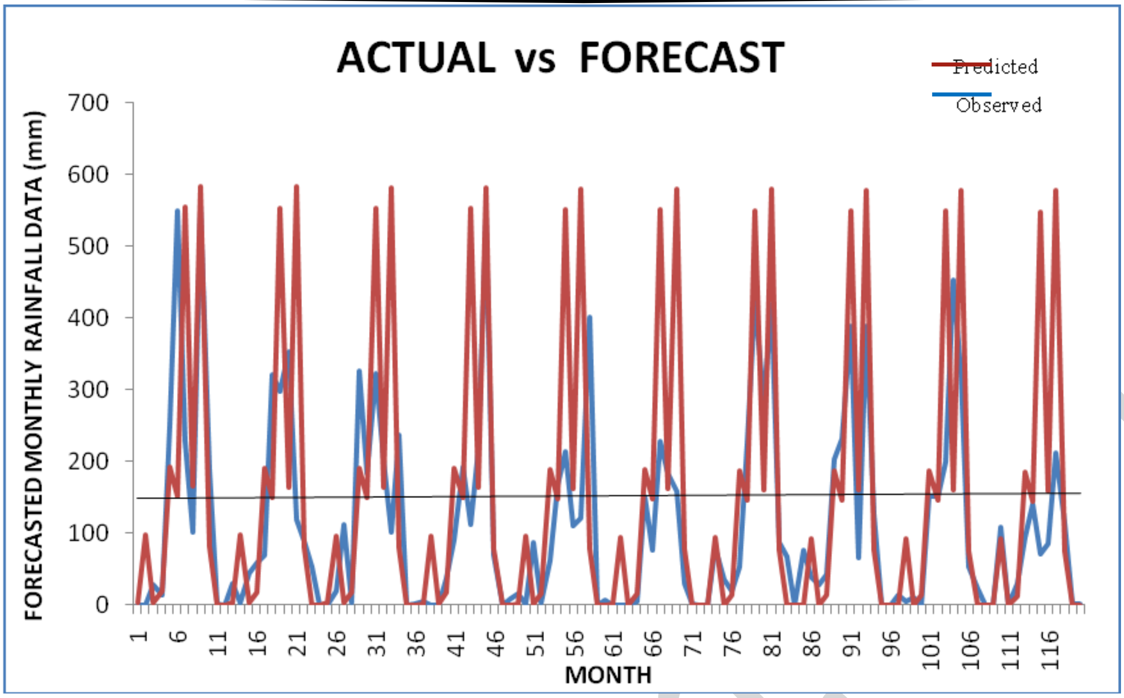


Fig-7: Forecasting of monthly rainfall using developed ARIMA model (1,0,0)(1,1,1)^{12}

CONCLUSIONS

The present study used Autoregressive Integrated Moving Average (ARIMA) model, proposed by Box and Jenkins (1976), for model building and forecasting the rainfall data. For the development of ARIMA model, the monthly rainfall data of 30 years (1971-2000) were used in this study. Nine models were identified for this study and AIC values for all models were estimated. The model with minimum AIC value was selected as the best model for the study area. In this study, the ARIMA model (1,0,0)(1,1,1)^{12} performed the best. The selected ARIMA model (1, 0, 0) (1, 1, 1)^{12} was used to predict the monthly rainfall data from 2001 to 2010. The comparison of the predicted and actual rainfall data series shows a good accuracy in prediction. Therefore, the selected model could be fairly used to forecast the upcoming rainfall in Krishnanagar sub-division, Nadia district, West Bengal.

REFERENCES

i. Bisgaard, S. and Kulahci, M. (2011). Time series analysis and forecasting by example. John Wiley & sons, Aug 24, mathematics - 400 pages

ii. Delleur, J.W. And Kavvas, M.L., (1978). Stochastic models for monthly rainfall forecasting and synthetic generation. J. Appl Meteorology., 17:1528-1536.

iii. Mahsin, M.D., Yesmin, A. and Monira, B., (2012). Modeling rainfall in Dhaka division of Bangladesh using time series analysis. Journal of mathematical modeling and application, vol.1, no.5, 67-73 pp.

iv. Meher J, and Jha, R. (2013). Time series analysis of monthly rainfall data for the Mahanadi river basin, India. Sciences in cold and arid regions, 5(1): 0073-0084

v. Mohammadi, K., Eslami, H.R. and Dardashts. D. I. (2005). Comparison of regression arima and ann models for reservoir inflow forecasting using snow melt equivalent (a case study of karaj), Journal of agriculture science technology, 7, pp. 17-30.

vi. Momani, P.E..N..M., (2009). Time series analysis model for rainfall data in jordan: case study for using time series analysis. Am. J. Environ. Science., 5: 599-604. Doi:10.3844/AJESSP.2009.599.604

vii. Inderjeet. K. and Singh. S. M. (2008). Seasonal arima model for forecasting of monthly rainfall and temperature. Journal of environmental research and development vol. 3 no. 2, October-December

viii. Shamsnia, S.A., Shahidi, N., Liaghat, A., Sarraf, A. Vahdat, S.F (2011). Modeling of weather parameters using stochastic methods (arima model) (case study: abadeh region, Iran), International Conference on Environment and Industrial Innovation, ipcbee vol.12, IACSIT press, Singapore.

ix. Somvanshi, V.K., Pandey, O.P., Agarwal, P.K., Kalankerl, N.V., Prakash, M.R. and Chand, R. (2006). Modeling and prediction of rainfall using artificial neural network and arima techniques. J. Ind. Geophys. Union (april) vol. 10, no.2, 141-151.

x. Thapliyal, V., and Kulshrestha, s. M., (1992). Rrecent models for long range forecasting of southwest monsoon rainfall over India. S Mausam, 43, 239–248.

xi. Weesakul, U. and Lowanichchai, S. (2005). Rainfall forecast for agricultural water allocation planning in Thailand. Thammasat int. Journal of science and technology. Vol.10, no. 3,

xii. Yurekli K. and Kurunc, A. (2004). Simulation of drought periods using stochastic models. Turkish journal of engineering and environmental sciences, 28: 181-190.