International Journal of Agricultural Sciences
Volume 13 | Issue 2 | June, 2017 | 208-214 ■ e ISSN-0976-5670
DOI:10.15740/HAS/IJAS/13.2/208-214
Visit us : *www.researchjournal.co.in*

RESEARCH PAPER

Modelling and forecasting of cultivated area and production of rice in India

PREMA BORKAR
Gokhale Institute of Politics and Economics, PUNE (M.S.) INDIA
Email: premaborkar@rediffmail.com

Abstract : In the present study, autoregressive integrated moving average (ARIMA) methodology has been applied for modeling and forecasting of yearly area and production of rice in India. Rice production data for the period of 1950-1951 to 2014-2015 of India were analyzed by time-series methods. Autocorrelation and partial autocorrelation functions have been estimated, which have led to the identification and construction of ARIMA models, suitable in explaining the time series and forecasting the future area and production.The diagnostic checking has shown that ARIMA (1, 0, 1) and ARIMA (0, 1, 1) is appropriate for rice area and production. The forecasts from 2015-2016 to 2024-2025 were calculated based on the selected model. The forecasting power of autoregressive integrated moving average model was used to forecast rice area and production for ten leading years. This projection is important as it helps to inform good policies with respect to relative production, price structure as well as consumption of rice in the country.

Key Words : ACF - autocorrelation function, ARIMA - Autoregressive integrated moving average, PACF - Partial autocorrelation function, Rice, Trends

View Point Article : Borkar, Prema (2017). Modelling and forecasting of cultivated area and production of rice in India. *Internat. J. agric. Sci.*, **13** (2) : 208-214, **DOI:10.15740/HAS/IJAS/13.2/208-214.**

Article History : Received : 11.01.2017**; Revised :** 06.04.2017**; Accepted :** 20.04.2017

INTRODUCTION

Paddy is the most important and extensively grown food crop in the world. It is the staple food of more than 60 per cent of the world population. It is of special importance for the nutrition of large reaches of the population in Asia. As a result, it plays a pivotal role for the food security of over half the world population. It is also a central component of the culture of a number of communities. For those reasons, rice is considered as a “stratregic” commodity in many countries, both developed and developing, and has consequently remained subject to a wide range of government controls and interventions. India has the largest area under paddy in the world and ranks second in the production after China. Country has also emerged as a major rice consumer.

Rice is cultivated in more than 100 countries in the world. In 2015-16, China produced some 145.5 million metric tons of milled rice, making the country the world’s largest producer of rice followed by India (103.5 million tons). Together these countries, accounted about half of world rice area and production. Indonesia (36.3 million tons), Bangladesh (34.6 million tons), Vietnam (28.2 million tons), Thailand (16.4 million tons), Burma (12.2 million tons), Philippines (11.5 million tons), Brazil (8 million tons), and Japan (7.9 million tons) are the other major rice producing countries. (*www.statista.com*). In case of productivity, Egypt ranks first with 9086 kg/ha followed by USA (7037 kg/ha), Japan (6702 kg/ha) and Korea Rep (6592 kg/ha). According to 4th advance estimate, India accounted for 439.49 lakh hectares area with production level of 106.54 million tons of rice. The productivity of rice in India is 2424 kg/ha which is less than other rice producing countries.

Forecasts have traditionally been made using structural econometric models. Concentration have been given on the univariate time series models known as auto regressing integrated moving average (ARIMA) models, which are primarily due to world of Box and Jenkins (1970). These models have been extensively used in practice for forecasting economic time series, inventory and sales modeling (Brown, 1959 and Holt *et al.,* 1960) and are generalization of the exponentially weighted moving average process. Several methods for identifying special cases of ARIMA models have been suggested by Box-Jenkins and others. Makridakis *et al.* (1982) and Meese and Geweke (1982) have discussed the methods of identifying univariate models. Among others Jenkins and Watts (1968); Yule (1926 and 1927); Bartlett (1964); Quenouille (1949); Ljune and Box (1978) and Prindycke and Rubinfeld (1981) have also emphasized the use of ARIMA models.

In this study, these models were applied to forecast the area and production of rice crop in India. This would enable to predict expected rice area and production for the years from 2015 onwards. Such an exercise would enable the policy makers to foresee ahead of time the future requirements for grain storage, import and/or export of rice thereby enabling them to take appropriate measures in this regard.

MATERIAL AND METHODS

Time series data were used for the study. The data were obtained from the website of Directorate of Economics and Statistics, Department of Agriculture and Cooperation, Ministry of Agriculture from 1950-51 to 2014-15. Box and Jenkin (1976) linear time series model was applied on the data. This model is commonly known as autoregressive integrated moving average model (ARIMA Model).

One of time series models which is popular and mostly used in ARIMA model. ARIMA (p, d, q) model is a mixture of autoregressive (AR) model which shows that there is a relation of a value in the present (Z) and value in the past (Z_{t-k}), added by random value and moving average (MA) model which shows that there is a relation between a value in the present (Z_t) and residuals in the past

$(Z_{t-k}\ k = 1,2...)$

with a non-stationary data pattern and d differencing order. The form of ARIMA (p,d,q) is:

$$c_p(B)\,(1- B)^d Z_t = {}^{-}_{q}\,(B)a_t$$

where, p is AR model order, q is MA model order, d is differencing order and:

$$c_p(B) = (1- c_1 B - c_2 B^2 - ...- c_p B^p)$$
$$c_p(B) = (1- c_1 B - c_2 B^2 - ...- c_q B^q)$$

Generalization of ARIMA model for a seasonal patter data, which is written as:

$$ARIM\,(p, d, q)(P, D, Q)^2 = c_p(B)\, c_p(B^2)\,(1-B)^d\,(1-B^s)\,(1-B^s)^D Z_t = {}^{-}(B)\, {}^{n}Q\,(B^s)at$$

where, s is seasonal period.

c p E_ sF= Z= ENJ c N_= J c 0_ 0= J= KKKJ c p_ pF
_ q E_ FZENJ _ N= _J _ 0= _ 0JKKKJ _ q_ qF

Model identification :

To determine whether the series isstationary or not we considered the graph of ACF. If a graph of ACF cuts of fairly quickly or dies down fairly quickly, then the time series value should be considered stationary. Model for non-seasonal series are called autoregressive integrated moving average model, denoted by ARIMA (p,d,q). Here p indicates the order of the autoregressive part, d indicates the amount average of difference and q indicates the order of the moving average part. If the original series is stationary, d = 0 and the ARIMA models reduce to the ARMA models.

The difference linear operator (Δ), denoted by:

$$aY_t = Y_t - Y_{t-1} = Y_t - BY_t = (1- B)Y_t$$

The stationary series:

$$W_t = a^d Y_t = (1-B)^d Y_t = \mu + {}^{-}_{q}(B)£_t$$
$$or\ c_p(B)W_t = \mu + {}^{-}_{q}(B)£_t$$

Model estimation and checking :

Estimate the parameters for a tentative model has been selected. The derived model must be checked for adequacy by considering the properties of the residuals whether the residuals from an ARIMA model is normal and randomly distribution. An overall check of the model adequacy is provided by Ljung-Box Q statistics. The test statistics Q is given in equation below:

$$Q_m = n(n+2)\sum_{k-1}^{n} \frac{r_k^2(e)}{n-k}\ \chi^2_{m-r}$$

where, $r_k(e)$ = The residual autocorrelation at lag K.
n = The number of residuals
m = The number of time lags includes in the test.

If the p-value associated with the Q statistics is small (p-value <“), the model is considered inadequate. The analysts should consider a new or modified model and continue the analysis until a satisfactory model has been determined.

Results and Discussion

The maximum cultivation Area of rice was 45.16 million hectares in 1999-00 and was minimum 29.83 million hectares in 1951-52. For production, the maximum production of rice in India was obtained in 2013-2014 year (106.65 million tons) and minimum in 1950-1951 year (20.58 million tons).

Last 65 years data of cultivation area and production of rice in India was used for modeling purpose. In model specification, we looked at the plots of auto-correlation function (ACF) for rice cultivation areas (Fig. 1) and production figures (Fig. 2). Also, partial auto correlation function (PACF) for rice cultivation area (Fig. 3) and production (Fig. 4). Auto correlation function indicated the order of the auto regression compounds “q” of the model while the partial correlation function gave an indication for the parameter p. The ACF and PACF of the residuals (Fig. 5 and 6) also indicate ‘good fit’ of the model.

The time series plot (Fig. 7 and 8) of cultivation areas and production showed an increasing trend. ACF of both series showed non-stationary as ACF did not fall as quickly as the log K increased. To check the further stationary, second difference of the original series for cultivation area and production was taken. The auto

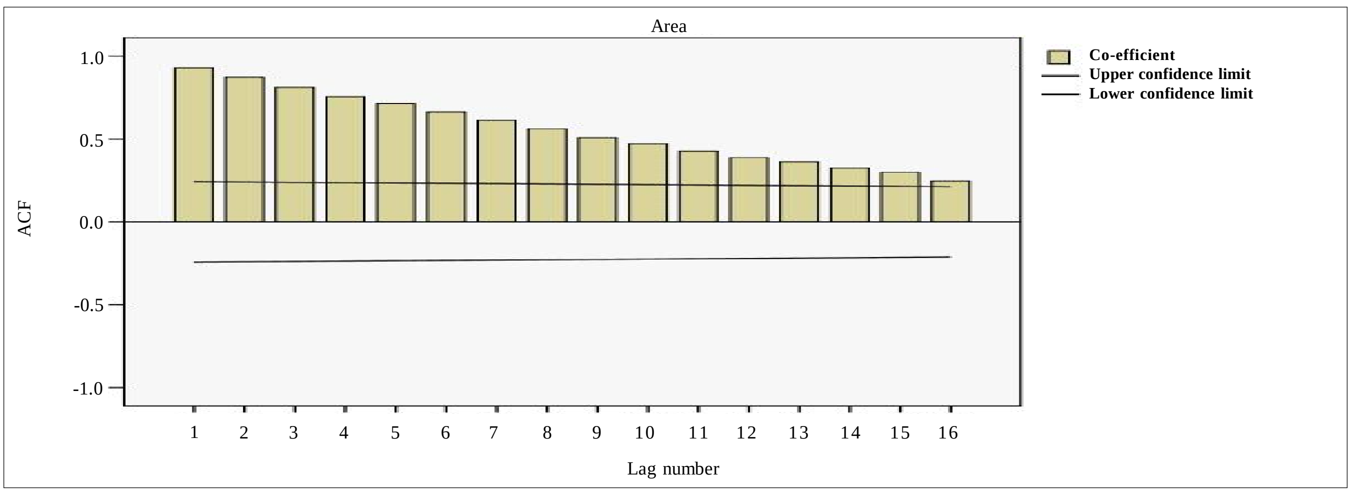

Fig. 1 : Autocorrelation function of rice cultivated area

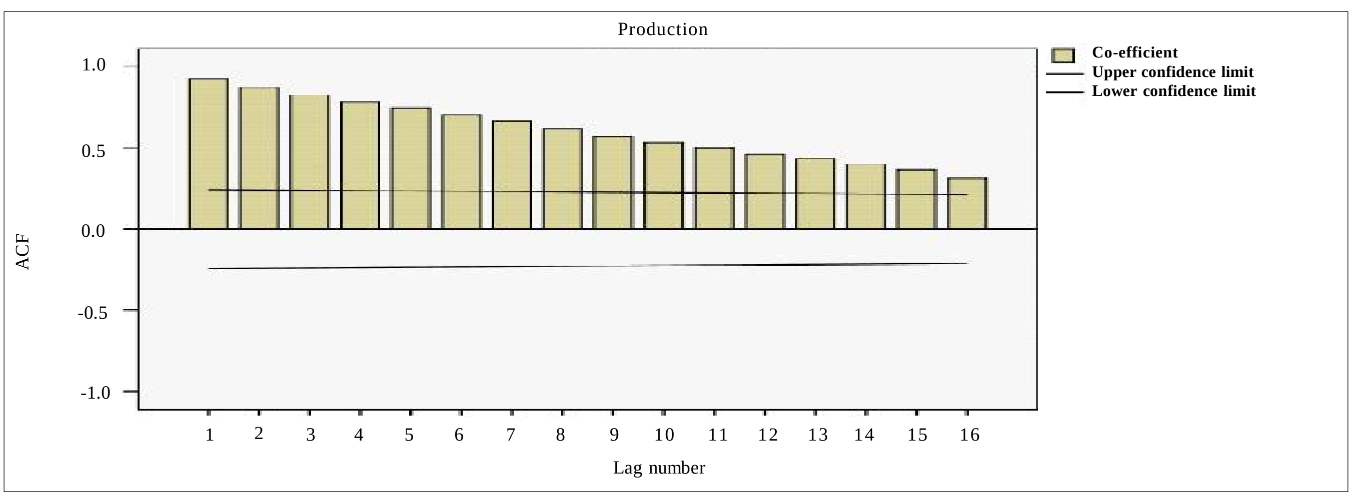

Fig. 2 : Autocorrelation function of rice production

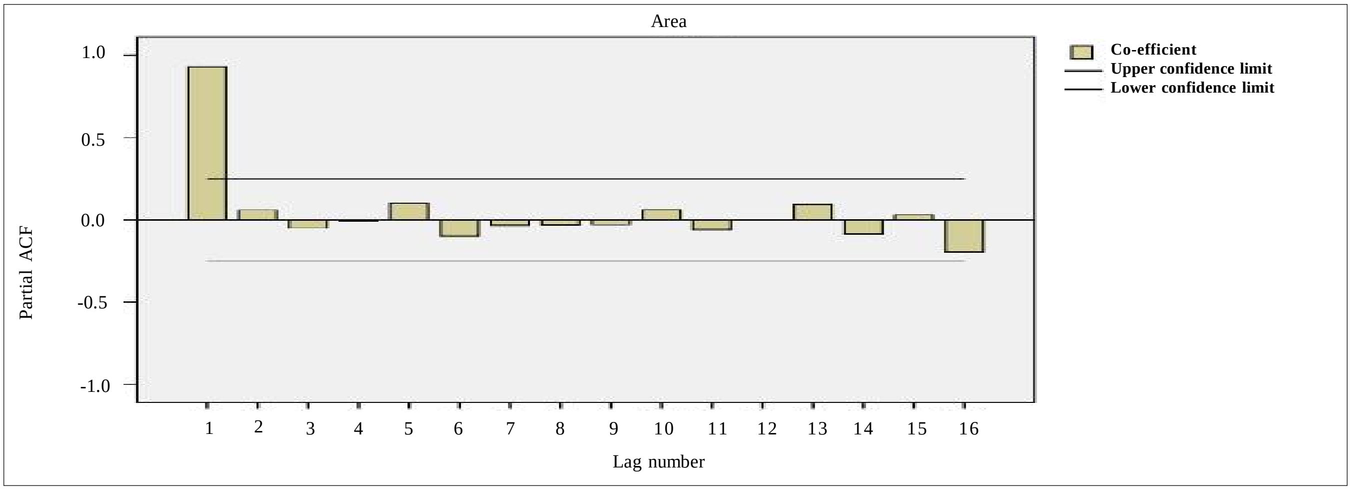

Fig. 3 : Partial autocorrelation function for rice cultivated area

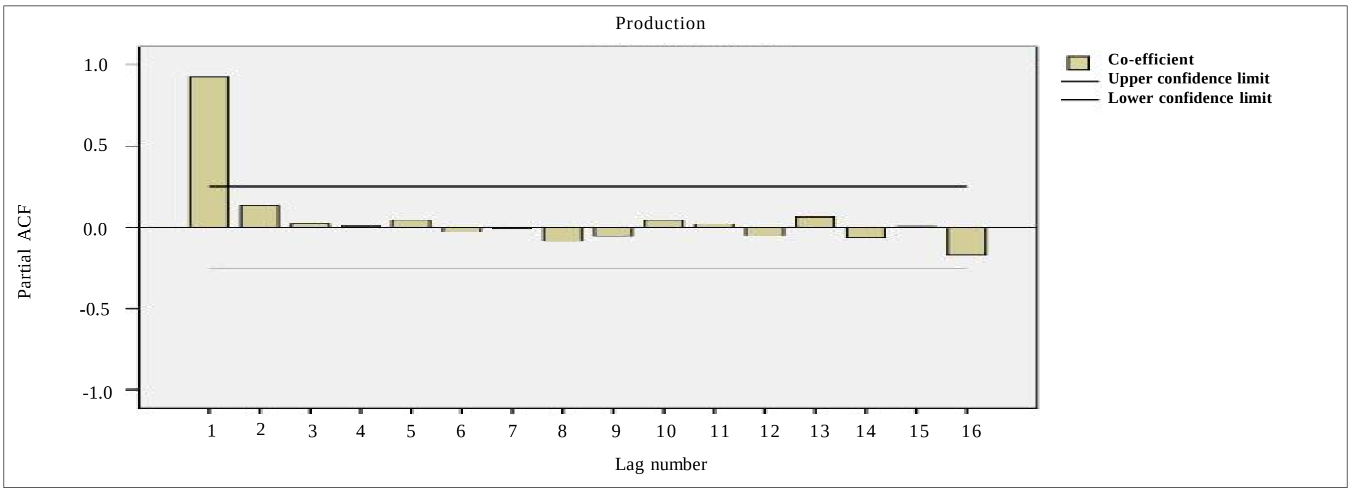

Fig. 4 : Partial autocorrelation function for rice production

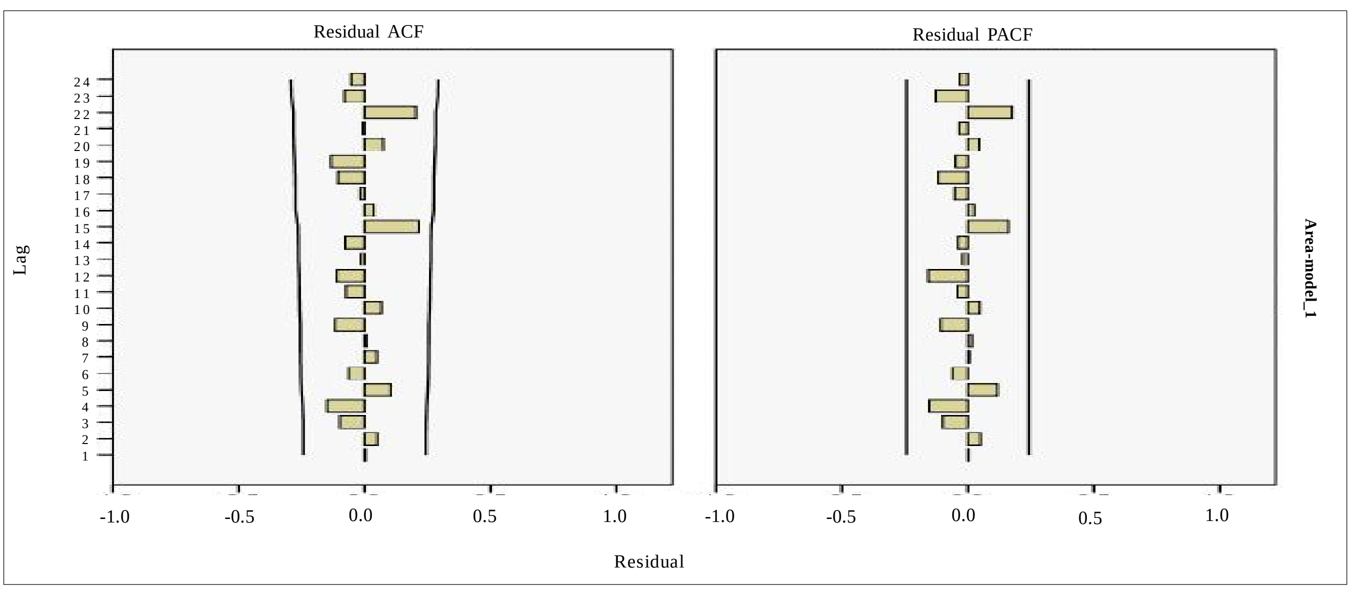

Fig. 5 : ACF of residuals of fitted ARIMA model

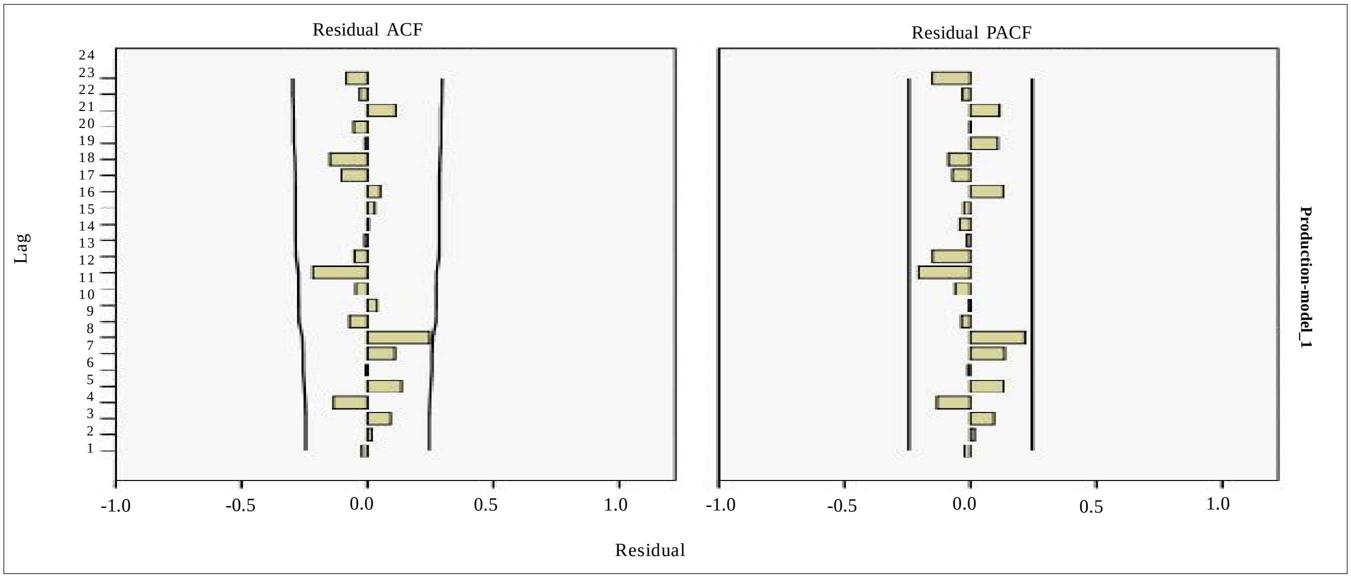

Fig. 6 : PACF of residuals of fitted ARIMA model

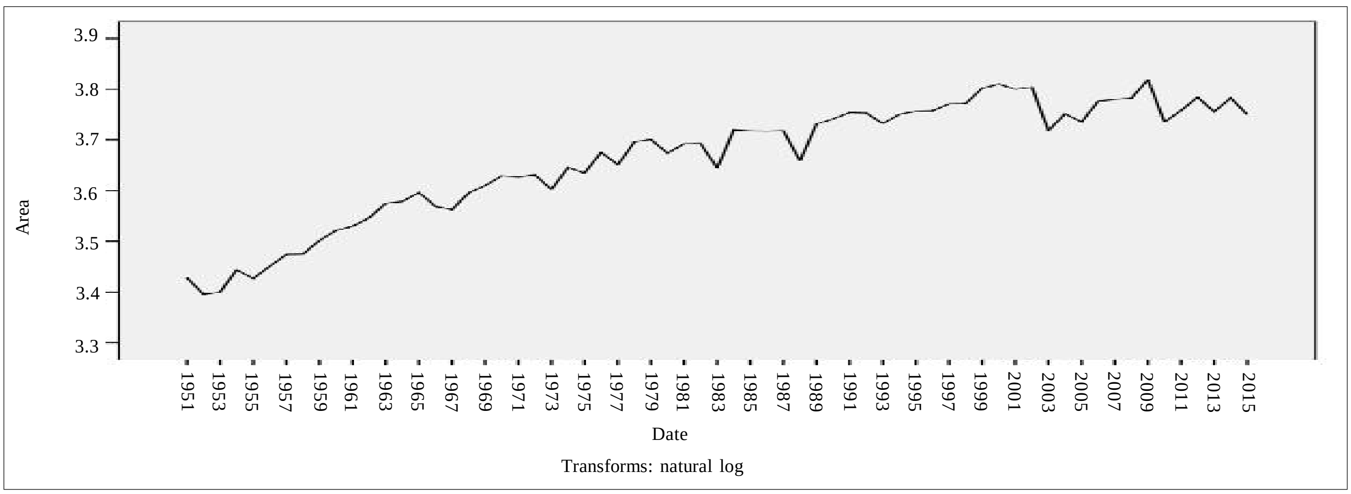

Fig. 7 : Time series plot for cultivated area

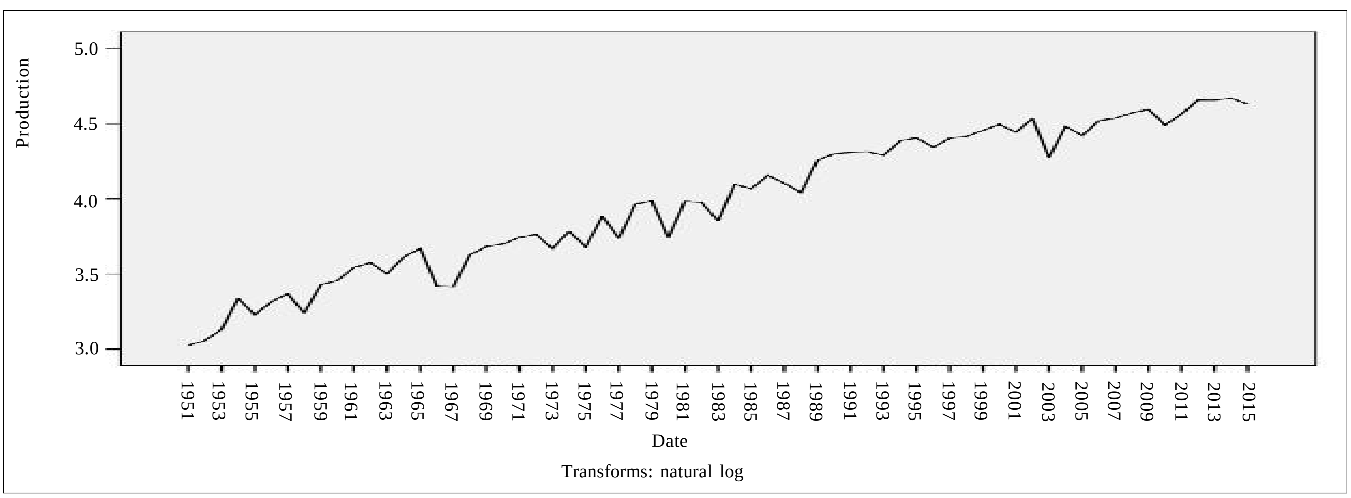

Fig. 8 : Time series plot of production area

correlation formulation of second series and correlogram shows some more stationary than that of the first different. The corellogram of the auto correlation function of first difference series showed that the auto correlation function falls finally after lag 1 for area and production, hence the respective values of the parameter "q" decided to be 1.

PAC function of the first differenced series of the cultivation area and production was used to determine parameter "p". Thus, we chose "p" to be 1 and 0 for cultivation area and production respectively which gave good results consequently, the respective value of p,d,q were determined for ARIMA, that is ARIMA (1,0,1) and ARIMA (0,1,1).

Model estimation :

ARIMA (1,0,1) and (0,1,1) model were estimated using SPSS and estimation of the models for the rice area and production are given in Table 1 and 2. Goodness of fit of the model given in as the diagnostic check of the estimated model.

Residual analysis :

The time series plot of the residual cultivated area and production data showed scattered trend, therefore, models were fitted properly by residual analysis.

For normality test, Shapiro-wilk test was used. The test was significant and assumption of normality was accepted. Since the series fitted shows normality, the model is a good fit.

Forecast of area and production :

ARIMA (1,0,1) and ARIMA(0,1,1) were taken for 10 years ahead and forecasts for rice area and production are given in Table 3 at 95 per cent confidence interval values.

For 2015-16, a forecast of rice area was 42.82 million ha with lower and upper limit of 40.27 and 45.37 million ha, respectively. A rice area forecast for the year 2025 was 42.45 million ha with lower and upper limit of 36.82 and 48.08 million ha, respectively. Forecast of rice production showed an increasing trend. For 2015-16, a rice production was about 111.23 million tons with lower and upper limit 89.4 and 133.06 million tons. The rice production forecast for the year 2025 is about 139.25 million tons.

Conclusion :

In our study, the developed model for rice area and production was found to be ARIMA (1,0,1) and (0,1,1).

Table 1: Estimate of area parameters

Type	Co-efficients	S.D.	t-ratio
Constant	3.601	0.185	19.452
AR1	0.993	0.015	66.602
MA1	0.331	0.127	2.605

Table 2: Estimate of production parameters

Type	Co-efficients	S.D.	t-ratio
Constant	0.025	0.003	7.657
AR1	0.998	0.001	9.654
MA1	0.732	0.092	7.969

Table 3 : Forecasts for rice area and production (2015-16 to 2024-2025)

Years	Area (Million hectares)	Production (Million tonnes)
2015-2016	42.82±2.55	111.23±21.83
2016-2017	42.78±3.06	114.04±23.23
2017-2018	42.74±3.50	116.92±24.66
2018-2019	42.69±3.89	119.88±26.12
2019-2020	42.65±4.24	122.91±27.62
2020-2021	42.61±4.56	126.02±29.16
2021-2022	42.57±4.85	129.20±30.74
2022-2023	42.53±5.13	132.47±32.36
2023-2024	42.49±5.38	135.82±34.03
2024-2025	42.45±5.63	139.25±35.75

The forecasts of rice area and production, lower control limits (LCL) and upper control limits (UCL) are presented in Table 3. The validity of the forecasted values can be checked when the data for the lead periods become available. ARIMA model being stochastic in nature, it could be successfully used for modeling as well as forecasting the rice area and production of India. The model demonstrated a good performance in terms of explaining variability and predicting power. The supply projection of an agricultural commodity especially rice plays a vital role in the adjustment of supply to demand. These projections help the government to make policies with regards to relative price structure, production and consumption patterns and also, to establish relationship with other countries of the world.

REFERENCES

Bartlett, M.S. (1964). On the theoretical specification of sampling properties of autocorrelated time series. *J. Roy. Stat. Soc.,* **8** : 27–41.

Box, G.E.P. and Jenkins, G. M. (1970). *Time series analysis: Forecasting and control*, Holden Day, San Francisco, CA.

Box, G.E.P. and Jenkins, G.M. (1976).*Time series analysis: forecasting and control*. Rev. Holden-Day, Ed. San Francisco.

Brockwell, P.J. and Davis, R.A. (1996). *Introduction to time series and forecasting,* Springer.

Brown, R.G. (1959). *Statistical forecasting for inventory control*. McGraw-Hill, NEW YORK, U.S.A.

Holt, C.C., Modigliani, F., Muth, J.F. and Simon, H.A. (1960). Planning, production, inventores and work force. Prentice Hall,Englewood Cliffs, NJ, USA.

Iqbal, N., Bakhsh, K., Maqbool, A. and Ahmad, A.S. (2005). Use of the ARIMA Model for forecasting wheat area and production in Pakistan. *J. Agric. & Soc.Sci.*, **2**:120-122.

Jambhulkar, N. N. (2013). Modeling of rice production in Punjab using ARIMA Model. *Internat. J. Scientif. Res.*,**2**(8):1-2.

Jenkins, G. M. and Watts, D.G. (1968).*Spectral analysis and its application*, Day, San Francisco, California, USA.

Kendall, M.G. and Stuart, A. (1966). *The advanced theory of statistics*. Vol. **3**. Design and Analysis and Time-Series. Charles Griffin & Co. Ltd., LONDON, UNITED KINGDOM.

Ljunge, G.M. and Box, G.E.P. (1978). On a measure of lack of fit in time series models. *Biometrika*, **65**: 67–72.

Makridakis, S., Anderson, A., Filds, R., Hibon, M., lewandowski, R., Newton, J., Parzen, E. and Winkler, R. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition, *J. Forecasting Competition. J. Forecasting*, **1**: 111–53.

Meese, R. and Geweke, J.(1982). A comparison of autoregressive univariate forecasting procedures for macroeconomic time series. Manuscript, University of California,Berkeley, CA, USA.

Muhammad, F., Javed, M. S. and Bashir, M. (1992). Forecasting sugarcane production in Pakistan using ARIMA Models, *Pakistan. J. Agric. Sci.,* **9** (1) : 31-36.

Padhan, P.C. (2012). Application of ARIMA model for forecasting agricultural productivity in India. *J. Agric. & Soc. Sci.,* **8** : 50-56.

Paul, R.K., Panwar, S., Sarkar, S.K., Kumar, A. Singh, K.N. and Farooqi, S. (2013). Modelling and forecasting of meat exports from India. *Agric. Econ. Res. Rev.,* **26** (2): 249-255.

Prabakaran, K., Sivapragasam, C., Jeevapriya, C. and Narmatha, A. (2013). Forecasting cultivated areas and production of wheat in India using ARIMA Model. *Golden Research Thoughts*, **3**(3):1-7.

Prindycke, R.S. and Rubinfeld, D.L. (1981). Econometric models and economic forecasts, 2nd Ed. McGraw-Hill, NEW YORK, U.S.A.

Quenouille, M.H. (1949). Approximate tests of correlation in time- series. *J. Roy. Stat. Soc.,* **11**: 68–84.

Saeed, N., Saeed, A., Zakria, M. and Bajwa, T.M. (2000). Forecasting of wheat production in Pakistan using ARIMA models. *Internat. J. Agric. & Biol.,* **4** : 352-353.

Sarika, Iquebal and Chattopadhyay, M. A. (2011). Modelling and forecasting of pigeonpea (*Cajanus cajan*) production using autoregressive integrated moving average methodology. *Indian J. Agric.Sci.,***81**(6): 520-523.

Yule, G.U. (1926). Why do we sometimes get nonsence-corrleations between times series.A study in sampling and the nature of series. *J. Roy. Stat. Soc.,* **89** : 1–69.

Yule, G.U. (1927). On a method of investigation periodicities in disturbed series, with specia; Reference ToWolfer's Sunspot Number. *Phill. Trans.,* **98** (A): 226: 267.