Research Paper

Volume : 2 | Issue : 8 | August 2013 • ISSN No 2277 - 8179

Modeling of Rice Production in Punjab using ARIMA Model

Agriculture

KEYWORDS : Rice production, ARIMA, Punjab

N. N. Jambhulkar Central Rice Research Institute, Cuttack – 753006

ABSTRACT

Rice is the chief food grain of Indian population. Punjab ranks in the top five rice producing states in the country. Due to increase in demand of rice over the years, the modeling and forecasting of rice production over the years is very important. An Auto Regressive Integrated Moving Average (ARIMA) methodology has been successful in describing and forecasting the rice production in the past studies. In the present study, ARIMA stochastic modeling is used for describing rice production in Punjab. The yearly rice production data of Punjab from 1960-61 to 1999-2000 has been taken for model building and the data from 2000-01 to 2009-10 has been used for validation of the model. The best model has been selected based on the minimum Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values. It has been found that ARIMA (1, 1, 2) model described the rice production data in Punjab.

1. INTRODUCTION

Rice is one of the chief food grains of India and it is one of the principal food crop of the country. India is one of the world's largest producer of rice. India ranks second in the world rice production only by China. India's rice production reaches to a record of 104.32 million tones in the year 2011-12. In India, Punjab ranks fourth with the production of 11.24 million tones in the rice production of the country. Due to increase in population, its accurate forecast is important for efficient planning of rice production. Forecasting of rice production is a formidable challenge. In view of globalization, it is important to study the trend of rice production by employing sound statistical modeling techniques that, in turn will be beneficial to the planners in formulating suitable policies to face the challenges ahead. Several attempts have been made in the past to develop yield forecast models for various commodities. Rahman (2010) studied the forecasting of boro rice production in Bangladesh. Badmus and Ariyo (2011) have studied the forecasting of Maize production. Awal and Siddique (2011) have studied the rice production in Bangladesh. Zakari and Ying (2012) have studied the forecasting of Niger grain production. Time series analysis of rice production has been an important tool for rice production management. Among all states of India, Punjab ranks fourth in the rice production. Thus, modeling and forecasting of rice production of Punjab over the years is important. To this end, autoregressive integrated moving average (ARIMA) methodology has been successful in describing and forecasting the rice production in the past. For the prediction purpose one or both of two types of models, usually known as structural regression models and time series models are often used in practice. The use of structural regression models requires information about the factors affecting the time series. On the other hand, time series analysis, especially Box-Jenkin type ARIMA models, let the data speak for themselves i.e. the future movements of a time series are determined using its own present and past values. (Box and Jenkins, 1976). Among the stochastic time series models ARIMA types are very powerful and popular as they can successfully describe the observed data and can make forecast with minimum forecast error. In Box-Jenkins model, the response variable at any time '*t*' is assumed to be expressive as a linear function of its values at past epochs *t*-1, *t*-2, ... Thus, the role of various predictor variables enter into the model 'implicitly' through response variable observations at past epochs. In the present study, ARIMA stochastic modeling is used on the rice production of Punjab for forecasting purpose.

2. MATERIAL AND METHODS

The time series data of rice production in Punjab has been collected from the website of Directorate of Economics and Statistics, Department of Agriculture and Cooperation, Ministry of Agriculture from 1960-61 to 2009-2010. ARIMA stochastic modeling is used on the rice production of Punjab for forecasting purpose.

2.1 Development of ARIMA models

The development of ARIMA models was based on the methodology described in the classic work of Box and Jenkins. Univariate ARIMA models use only the information contained in the series itself. Thus, models are constructed as linear functions of past values of the series and/or previous random shocks (or errors). Forecasts were generated under the assumption that the past history could be translated into predictions for the future. ARIMA modeling was developed following the standard three steps procedures. (i) Identification of the model; (ii) Parameter estimation and (iii) Diagnostic and verification of the model. The identification step determines (i) whether the process is stationary and the possible transformations to obtain stationarity and (ii) whether the form of the process is autoregressive (AR), moving average (MA) or both (ARMA), and its orders. Three parameters are used in summarizing an ARIMA model are the AR parameter *p*, integration parameter *d* and MA parameter *q*. Parameters *p* and *q* denote the order of AR and MA, while *d* denotes the degree of differencing the series to obtain stationarity. The autocorrelation function (ACF) and partial autocorrelation functions (PACF) of a series together are the most powerful too, usually applied to reveal the correct values of the parameters. The ACF gives the autocorrelations calculated at lags 1, 2 and so on, while PACF gives the corresponding partial autocorrelations, controlling the autocorrelations at intervening lags. Parameter estimation of tentative models was determined using maximum-likelihood methods. The final results included the parameter estimates, standard errors, estimates of residual variance, standard error of the estimate, natural log likelihood, Akaike's Information Criterion (AIC), and Schwartz's Bayesian criterion (SCB) or Bayesian Information Criterion (BIC). Model selection was based on the minimization of AIC and BIC. These criteria are descriptors of the model's parsimony as they simultaneously account for the model's fit onto the observed series alongside number of parameters used in the fit. The ability to forecast using ARIMA models was tested by applying the ARIMA methodology to available data (1960-61 to 1999-2000), excluding the data from 2000-01 to 2009-10 which were used for testing the forecasting power of the established models.

The autoregressive moving average (ARMA) model, denoted by ARMA (*p*,*q*), is given by

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \ldots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \quad (1)$$

or equivalently by

$$\varphi(B) y_t = \theta(B) \varepsilon_t \quad (2)$$

where $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^p$ and

$\theta(B) = 1 - \theta_1 B - \theta_2 B^q$

B is the backshift operator defined by $By_t = y_{t-1}$.

A generalization of ARMA models which incorporates a wide class of nonstationary time-series is obtained by introducing 'differencing' into the model. The simplest example of a nonstationary process which reduces to a stationary one after differencing is 'Random Walk'. A process $\{y_t\}$ is said to follow autoregressive integrated moving average (ARIMA), denoted by $ARIMA(p,d,q)$, if $\nabla^d y_t = (1-B)^d \varepsilon_t$ is $ARMA(p,q)$. The model is written as

$$\varphi(\boldsymbol{B})(1-\boldsymbol{B})^d \boldsymbol{y}_t = \theta(\boldsymbol{B})\varepsilon_t \quad (3)$$

where ε_t are identically and independently distributed as $N(0,\sigma^2)$ The integration parameter *d* is a nonnegative integer. When $d = 0$, the ARIMA(*p*, *d*, *q*) model reduces to ARMA(*p*, *q*) model.

2.2 Estimation of Parameter

Estimation of parameters for ARIMA model is generally done through Nonlinear least squares method. Several software packages are available for fitting of ARIMA models. The Akaike information criterion (AIC) and Bayesian information criterion (BIC) values are used for parameter estimation.

3. RESULTS AND DISCYSSION

The graph of the rice production of Punjab is plotted in figure 1. Perusal of the graph indicates that the rice production of Punjab fluctuates from 500 thousand tones and increases upto 9000 thousand tones. The fluctuation of the production also indicates that the dataset is non stationary. It can also be visualized from the plot of ACF and PACF (Figure 2 (a)) of the series. After first differentiation of the original series the decay rate becomes high (Figure 2(b)) resulting the identification of the order of the model very easy. To this end, Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests were used for the test of stationarity. On the basis of minimum AIC and BIC values and considering the ACF and PACF of the rice production series, ARIMA (1, 1, 2) model is selected.

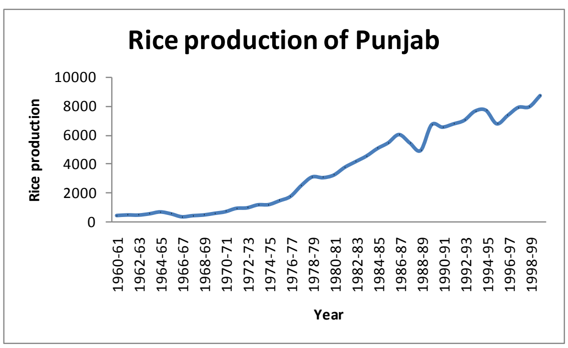

Figure 1: Rice production (in '000 tones) of Punjab

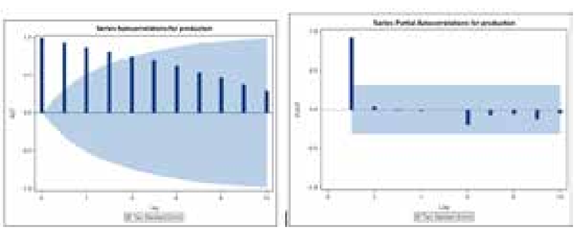

Figure 2 (a): ACF and PACF of the rice production data series

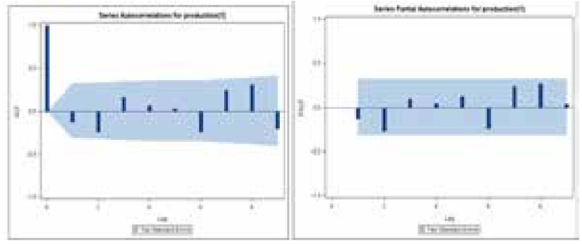

Figure 2 (b): ACF and PACF of the first differentiated rice production data series

The estimate of the parameters with corresponding standard error for ARIMA (1, 1, 2) is given in Table 1.

Table 1: Parameter estimate of the ARIMA (1, 1, 2) model

Parameters	Estimate	SE
Constant	211.045	50.544
AR1	-0.402	0.405
MA1	0.319	0.170
MA2	-0.502	0.412

The forecast and actual values of yearly rice production (in thousand tones) of Punjab for the period 2000-01 to 2009-10 is given in Table 2.

Table 2. Forecast of rice production of Punjab

Year	Actual	Forecast by ARIMA(1, 1, 2)	UCL	LCL
2000-01	9154	8570.80	4749.90	12391.69
2001-02	8816	8781.84	4912.69	12650.99
2002-03	8880	8992.88	5076.07	12909.70
2003-04	9656	9203.93	5240.02	13167.83
2004-05	10437	9414.97	5404.53	13425.42
2005-06	10193	9626.02	5569.57	13682.47
2006-07	10138	9837.06	5735.13	13939.00
2007-08	10489	10048.11	5901.18	14195.03
2008-09	11000	10259.15	6067.72	14450.58
2009-10	11236	10470.20	6234.73	14705.67

4. CONCLUSION

ARIMA model being stochastic in nature, it could be successfully used for modeling as well as forecasting the rice production of Punjab. The model demonstrated a good performance in terns of explaining variability and predicting power. The forecasting of rice production can help to both farmers as well as the planner for future planning.

REFERENCE

Awal, M. A. and Siddique, M. A. B. (2011). Rice production in Bangladesh employing by ARIMA model. Bangladesh Journal of Agricultural Research, 36(1), 51-62. | Badmas, M. A. and Ariyo, O. S. (2011). Forecasting cultivated area and production of Maize in Nigerian using ARIMA Model. Asian Journal of Agricultural Sciences, 3(3), 171-176. | Box, G.E.P. and G.M. Jenkin, 1976. Time Series of analysis, forecasting and control, Sam Franscico, Holden-Day, California. USA. | Rahman, N. M. F. (2010). Forecasting of boro rice production in Bangladesh: An ARIMA approach. Journal of Bangladesh Agricultural University, 8(1), 103-112. | Zakari S. and Ying L. (2012). Forecasting of Niger production and harvested area. Asian Journal of Agricultural Sciences, 4(4), 308-313. | http://eands.dacnet.nic.in/