The Bioscan 9(3): 1063-1066, 2014
AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES
www.thebioscan.in

AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) APPROACH FOR PREDICTION OF RICE (*ORYZA SATIVA L.*) YIELD IN INDIA

PRITY KUMARI[1]*, G. C. MISHRA[1], ANIL KUMAR PANT[1], GARIMA SHUKLA[1] AND S. N. KUJUR[2]
[1]Department of Farm Engineering,
Institute of Agricultural Sciences, Banaras Hindu University,Varanasi - 221 005, INDIA
[2]Department of Genetics and Plant Breeding,
Institute of Agricultural Sciences, Banaras Hindu University,Varanasi - 221 005, INDIA
e-mail: psingh2506@gmail.com

KEYWORDS
Forecasting
Autoregressive Integrated Moving Average (ARIMA)
Rice and Yield

Received on :
05.10.2013

Accepted on :
06.07.2014

***Corresponding author**

ABSTRACT

In the present paper, different Autoregressive Integrated Moving Average (ARIMA) models were developed to forecast the rice yield by using time series data of sixty two years. The performance of these developed models were assessed with the help of different selection measure criteria and the model having minimum value of these criteria considered as the best forecasting model. Based on findings, it has been observed that out of eleven ARIMA models, ARIMA (1, 1, 1) is the best fitted model in predicting efficiently the rice yield as compared to others.

INTRODUCTION

Rice (*Oryza sativa L.*) production in India is an important part of the national economy as it is considered as staple food of most of the people of our country. Among the rice growing countries in the world, India has the largest area under rice crop and ranks second in production next to China [Directorate of Rice Development (www.drd.dacnet.nic.in); Rout and Tewari (2012); Balakrishna and Satyanarayana (2013); Singh *et. al.* (2013); Chowdhury *et al.* (2014)]. Around 65% of total population in India eats rice and it accounts for 40% of the nation's food production. Rice-based production systems provide the main source of income and employment for more than 50 million households. [International Rice Research Institute (IRRI), www.irri.org].

Heavy consumption of this food crop is become an important issue to increase the yield of rice of our country. Considering these facts in mind it is very important to have timely forecast of this crop in order to plan the strategies to increase the yield. To make the best forecast of rice yield in India, appropriate time series forecasting methods [Makridakis and Hibbon (1979); Pankratz (1983); Granger and Newbold. (1986); Makridakis *et. al.*(1998)] that can be able to describe the observed data successfully, are necessary. One of the most successful time series forecasting techniques is Box-Jenkin type ARIMA models [Box and Jenkins (1970)], which is the most commonly used as linear statistical model. It has been successfully applied in many time series forecasting and is a good tool to develop empirical model which is dependent on their own successive time series values [Iqbal *et. al.*(2005); Rahman (2010); Awal and Siddique (2011); Biswas and Bhattacharyya (2013); Jambhulkar (2013)].

The objective of the present study was to develop appropriate ARIMA model for forecasting the rice yield in India for the year 2012-13 with appropriate prediction interval.

MATERIALS AND METHODS

In the present study, time series secondary data on productivity (kg./ha.) of rice in India, were considered for the period 1950-51 to 2011-12 from Directorate of Economics and Statistics, Department of Agriculture and Cooperation, Ministry of Agriculture, Government of India. The time series secondary data were analyzed with the help of various ARIMA Models.

ARIMA is one of the most traditional methods of non-stationary time series analysis. In contrast to the regression models, the ARIMA model allows time series to be explained by its past, or lagged values and stochastic error terms. The models developed by this approach are usually called ARIMA models because they use a combination of autoregressive (AR), integration (I) - referring to the reverse process of differencing to produce the forecast and moving average (MA) operations

[Box and Jenkins (1994)]. An ARIMA model is usually stated as ARIMA (p, d, q). An autoregressive integrated moving average is expressed in the form:

If $w_t = \nabla^d r_t = (1-B)^d r_t$ then

$$w_t = \phi_1 w_{t-1} + \phi_2 w_{t-2} + ... + \phi_p \; w_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - ..... - \theta_q \varepsilon_{t-q}$$

Where r_t is time series data, B is Backshift operator, is autoregressive coefficient of order p, is moving average coefficient of order q and ε_t is white noise

Above equation is also written as:

Where $\phi(B)$ is a stationary autoregressive operator, is a stationary moving average operator, and is a constant. In the case of the pattern of seasonal time series, ARIMA model is written as follows:

Where $w_t = \nabla^d \nabla_s^D r_t$, $\nabla^d = (1-B)^d$ is number of regular differences, $\nabla_s^D = (1-B^s)^D$ is number of seasonal differences.

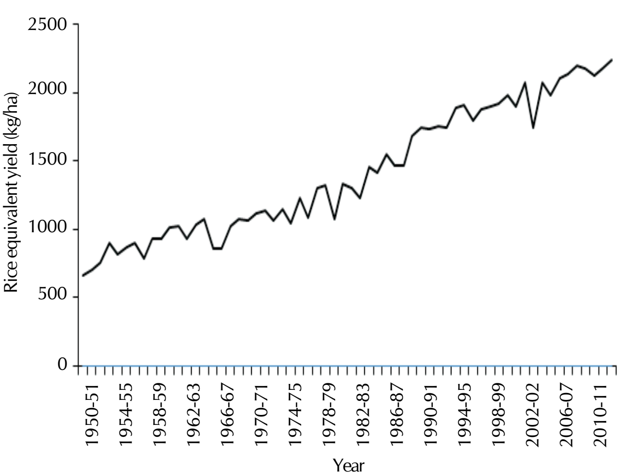

Figure 1: Yearly Trend of Rice Yield in India during 1950-51 to 2011-12

Seasonal ARIMA model is to denoted by (p, d, q) (P, D, Q), where P denotes the number of seasonal autoregressive components, Q denotes the number of seasonal moving average terms and D denotes the number of seasonal differences required to induce stationarity (Box and Jenkins, 1994).

The steps which are followed in order to define an ARIMA model as stated by Box & Jenkins:

Identifying a model

Estimating the parameters of the model

Diagnostic checking

In the present paper, time series yearly data on productivity (kg./ha.) of rice of India were considered so there is no seasonal variation in the data which means non seasonal ARIMA (p, d, q) models are applicable only. Comparison among family of different parametric combination of ARIMA (p, d, q) was done on the basis of minimum value of selection criteria which are Root Mean Squared Error (RMSE), Mean absolute percentage error (MAPE), Mean absolute error (MAE), Maximum absolute percentage error (MaxAPE), Maximum absolute error (MaxAE) and Bayesian information criteria.

RESULTS AND DISCUSSION

Fig. 1 represents the time (yearly) plot of rice yield in India from 1950-51 to 2011-12. After seeing the graph it can be inferred that the series is not stationary because the mean of the time series is increasing with the increase in time i.e. reflection of increasing trend.

Figure 2 show the autocorrelation function (ACF) of the rice yield in India and of the yield after one differencing. From the first figure (ACF of Yield_Rice) of Fig. 2 it is clear that Auto correlation coefficients do not cut off to statistical insignificance fairly quickly. All the first 16 autocorrelations are significantly different from zero at about the 5% level which indicates that the series is non stationary. To make the series stationary differencing is necessary. The second figure (ACF of Yield_Rice_1) of Fig. 2 clearly depicts that the series has now become mean stationary because autocorrelations decay to statistical insignificance rather quickly.

After making the series stationary, different parametric combinations of ARIMA (p, 1, q) model were tried to analyze the sixty two years data (1950-51 to 2011-12) of rice yield

Table 1: Performaces of Different ARIMA (p,d,q) Models of Rice Yield in India

Models	Model Selection Criteria					
	RMSE	MAPE	MAE	MaxAPE	MaxAE	BIC
ARIMA(0,1,0)	120.927	6.796	87.746	27.949	360.230	9.658
ARIMA(1,1,0)	99.932	6.088	76.456	26.627	282.047	9.344
ARIMA(2,1,0)	94.851	5.622	71.637	23.179	280.105	9.307
ARIMA(3,1,0)	95.332	5.561	70.822	23.685	279.273	9.384
ARIMA(1,1,1)	93.603	5.353	68.101	25.201	276.910	9.280
ARIMA(1,1,2)	94.465	5.398	69.517	25.334	281.552	9.366
ARIMA(2,1,1)	94.375	5.365	68.432	24.966	278.142	9.364
ARIMA(2,1,2)	95.667	5.476	69.307	24.928	269.478	9.459
ARIMA(0,1,1)	94.255	5.317	67.858	25.926	293.909	9.227
ARIMA(0,1,2)	93.603	5.370	68.580	25.062	279.362	9.280
ARIMA(0,1,3)	94.386	5.358	68.229	25.043	277.253	9.364

Table 2: ARIMA (1,1,1) Model parameter values for predicting rice yield in India

ARIMA Model Parameters							
	Numerator			Estimate	SE	t	Sig.
Yield_Rice-Model_1	Yield_Rice	Constant		24.725	4.097	6.035	.000
		AR	Lag 1	-.230	.177	-1.298	.199
		Difference		1			
		MA	Lag 1	.591	.149	3.951	.000

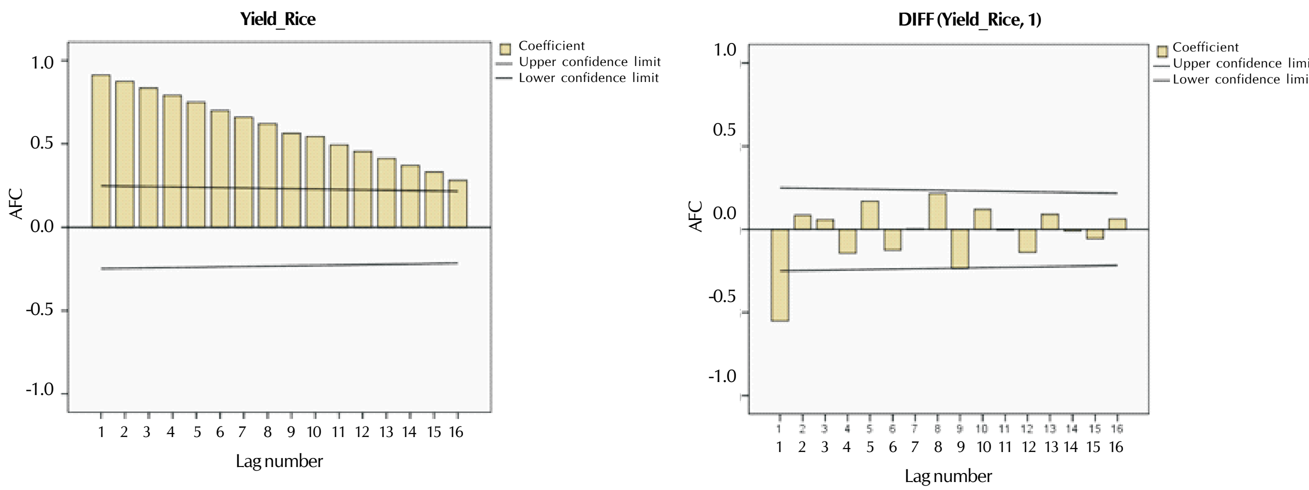

Figure 2: Autocorrelation Function of actual series (Yield_Rice) and differenced series (Yield_Rice_1)

with help of Statistical software SPSS 19.0 and the best fitted model is accepted on the basis of minimum value of all selection criteria as mentioned above in methodology. The results of performance of developed ARIMA (p, 1, q) model is presented in Table 1:

Table 1 exposed the performance of eleven ARIMA Models out of which ARIMA (1, 1, 1) was best out of all. ARIMA (0, 1, 2) and ARIMA (0,1,1) ranks second and third respectively while remaining ARIMA Models are not as good as these three. Therefore it was concluded that the appropriate model for forecasting the yield of rice in India during 2012-2013 was ARIMA (1, 1, 1) having minimum value of all selection criteria as compared to remaining ten models.

Table 2 shows the estimates of parameter of ARIMA (1, 1, 1). It is clear from the table that only MA (1) and constant term is statistically highly significant having standard error 0.149 and 4.097 respectively while AR (1) was non-significant. Since all selection criteria measures of this model is least and R square of this model is 0.96 which means 96% of variation in the data series, is explained by this model. It is, therefore, considered to be the best fitted model amongst all. The forecasted value of rice yield in India, for the year 2012-13, as obtained from ARIMA (1, 1, 1) was obtained as 2262.34 Kg./ha. with Upper and Lower confidence of 2454.52 Kg./ha. and 2070.16 Kg./ha.

CONCLUSIONS

This paper aimed to forecast the rice yield during 2012-13 in India, by Autoregressive Integrated Moving Average (ARIMA) Approach. On basis of results obtained it is concluded that ARIMA (1,1,1) Model having minimum value of all measures of selection criteria was found to be the appropriate model amongst all for predicting the rice yield of India. The model showed a good performance in case of explaining variability in the data series and also its predicting ability. The forecasting of rice yield can help farmers as well as the policy makers for future planning.

REFERENCES

Annual Report 2011-12. Directorate of Economics and Statistics, Department of Agriculture and Co-operation, Ministry of Agriculture, Govt. of India. http://www.agricoop.nic.in.

Awal, M. A. and Siddique, M. A. B. 2011. Rice Production in Bangladesh Employing by Arima Model, Bangladesh, *J. Agril. Res.* **36(1):** 51-62.

Balakrishna, B. and Satyanarayana, P. V. 2013. Genetics of Brown Planthopper (Nilaparvata Lugens Stal.) Resistance in elite donors of rice (*Oryza sativa* L.). *The Bioscan*. **8(4):** 1413-1416.

Biswas, R. and Bhattacharyya, B. 2013. ARIMA modeling to forecast area and production of rice in West Bengal. *J. Crop and Weed*. **9(2):** 26-31.

Chowdhury, Md. Riton, K., Vinod Sattar, A. and Brahmachari, K. 2014. Studies on the water use efficiency and nutrient uptake by rice under system of intensification. *The Bioscan*. **9(1):** 85-88

Box, G. E. P. and Jenkins, G. M. 1970. Time series analysis: forecasting and control. *San Francisco: Holden-Day.*

Granger, C. W. T. and Newbold, P. 1986. Forecasting Economics Time Series, Second Edition, Academic Press, Inc., *Orlando, Florida.*

International Rice Research Institute (IRRI). http://www.irri.org.

Iqbal, N., Bakhsh, K., Maqbool, A. and Ahmad Abid, S. 2005. Use of the ARIMA model for forecasting wheat area and production in Pakistan. *J. Agriculture and Social Sciences*. **1(2):** 120-122.

Jambhulkar, N. N. 2013. Modeling of rice production in Punjab using ARIMA model. *International J. scientific Research*. **2(8):**

Makridakis, S. and Hibbon, M. 1979. Accuracy of forecasting: An empirical investigation. *J. Roy. Statist. Soc. A*. **41(2):** 97-145.

Makridakis, S., Wheelwright, S. C. and Hyndman, R. J. 1998. Forecasting: methods and applications, *J. Wiley and Sons:* New York.

Pankratz, A. 1983. Forecasting with univariate box - Jenk ins Models - Concepts and Cases. *J. Willey and Sons,* New York.

Rahman, N. M. F. 2010. Forecasting of boro rice production in Bangladesh: An ARIMA approach. *J. Bangladesh Agril. Univ*. **8(1):** 103-112.

Rice In India During Tenth Plan. Directorate of Rice Development, Patna. www.drd.dacnet.nic.in.

Rout, S. and Tewari, S. N. 2012. Amalab-E, a formulated botanical product potential against rice blast incitant *Pyricularia grisea*. *The Bioscan*. **7(3):** 547-552.

Singh Chandra, M., Babu, G. S., Kumar, B. and Mehandi, S. 2013. Analysis Of quantitative variation and selection criteria for yield improvement In exotic germplasm of upland rice (*Oryza sativa* L.). *The Bioscan*. **8(2):** 485-492.