International Journal of Applied Mathematics
& Statistical Sciences (IJAMSS)
ISSN(P): 2319-3972; ISSN(E): 2319-3980
Vol. 6, Issue 4, Jun – Jul 2017; 67-72

International Academy of Science, Engineering and Technology
IASET
Connecting Researchers; Nurturing Innovations

ARIMA MODEL FOR FORECASTING OF RICE PRODUCTION IN INDIA BY USING SAS

G. RAMAKRISHNA & R. VIJAYA KUMARI

Department of Statistics & Mathematics, College of Agriculture,

Professor Jayashankar Telangana State Agricultural University, Rajendranagar, Hyderabad, Telangana, India

ABSTRACT

Autoregressive integrated moving average (ARIMA) approach has been applied for modeling and forecasting of rice production of India. Autocorrelation function (ACF) and partial autocorrelation function (PACF) functions were estimated, which led to the identification and construction of ARIMA models, for explaining the time series and forecasting the future production. A significant increasing linear trend in the total rice production in India has been found. The best identified model for the data under consideration was used for forecasting up to the year 2020 AD. The projected production by using ARIMA (011) would be increased to 112.90 million tonnes by 2020 AD.

KEYWORDS: ARIMA, ACF, PACF, AD, SAS

INTRODUCTION

India is in 2nd position in the world rice production. In India, rice is grown all over the country with an area of 44.10 million hectares, production of 105.52 million tones, and with the productivity of 2391kg/ha during 2014-15. (www.indian stat.com)

Meyler et al (1998) drew a framework for ARIMA time series models for forecasting Irish inflation. Prajneshu and venugopalan (1998) applied this model as well as other parametric statistical modeling techniques, like polynomial function fitting, and nonlinear mechanistic growth modeling for describing trends in marine fish production data of the country. Singh *et al.* (2007) applied statistical models for forecasting rice production in India. Elangbam Haridev Singh (2013) developed forecasts of international tourism demand in Bhutan, by selecting the appropriate model both ARIMA as well as exponential smoothing. Paul *et al.* (2014) applied models for livestock and dairy production in India under the time series framework.

Table 1: Rice Production in India (In Million Tonnes)

Year	Rice Production (In Million Tonnes)	Year	Rice Production (In Million Tonnes)	Year	Rice Production (In Million Tonnes)
1949-50	23.5	1972-73	39.2	1995-96	77.0
1950-51	20.6	1973-74	44.1	1996-97	81.7
1951-52	21.3	1974-75	39.6	1997-98	82.5
1952-53	22.9	1975-76	48.7	1998-99	86.1
1953-54	28.2	1976-77	41.9	1999-00	89.7
1954-55	25.2	1977-78	52.7	2000-01	85.0

Table 1: Contd.,					
Year	**Rice Production (In Million Tonnes)**	**Year**	**Rice Production (In Million Tonnes)**	**Year**	**Rice Production (In Million Tonnes)**
1955-56	27.6	1978-79	53.8	2001-02	93.3
1956-57	29.0	1979-80	42.3	2002-03	71.8
1957-58	34.6	1980-81	53.6	2003-04	88.5
1958-59	35.7	1981-82	53.2	2004-05	83.1
1959-60	31.7	1982-83	47.1	2005-06	91.8
1960-61	34.6	1983-84	60.1	2006-07	93.4
1961-62	35.7	1984-85	58.3	2007-08	96.7
1962-63	33.2	1985-86	63.8	2008-09	99.2
1963-64	37.0	1986-87	60.6	2009-10	89.1
1964-65	39.3	1987-88	56.4	2010-11	96.0
1965-66	30.6	1988-89	70.5	2011-12	105.3
1966-67	30.4	1989-90	73.6	2012-13	105.2
1967-68	37.6	1990-91	74.3	2013-14	106.6
1968-69	39.8	1991-92	74.7	2014-15	105.5
1969-70	40.4	1992-93	72.7	2015-16	104.4
1970-71	42.2	1993-94	80.3	2016-17-2nd Advance	108.9
1971-72	43.1	1994-95	81.8		

Source: www.indian stat.com

MATERIALS AND METHODS

Data Description

Rice production data, in million tonnes from the year 1949 to 2016 has been used for model development. The SAS 9.3 statistical software package has been used for data analysis.

Auto Regressive Integrated Moving Average (ARIMA) Model (Process): The ARIMA methodology is called as a Box-Jenkins methodology (Box and Jenkins 1976). The Box-Jenkins procedure is concerned with fitting a mixed ARIMA model to a given set of data. The main objective in the fitting ARIMA model is, to identify the stochastic process of the time series and predict the future values accurately. This process is an amalgamation to ARMA process when the time series {Yt} is Non-Stationary or "Integrated". It is obvious that, to develop the ARMA model in this situation, the series has to be "differenced" to make it stationary and this difference series, which is now stationary, has to be subjected to the fitting of ARMA model. This process is referred as ARIMA (p, d, q), where p and q refer to the number of AR and MA terms, and d refers to the order of differencing required for making the series a Stationary. ARIMA model mainly consist of model identification, parameter estimation and testing, modal validation and forecasting.

Model Identification

Model identification involves, finding out the order of the AR, MA parameters and differentiation value of the model, first we saw if the series was stationary or not, if the series was non stationary we went for differences in the series. To find if the series is stationary or not, we found out, by graph or by Augmented Dickey-Fuller (ADF) Test.

Table 2: Augmented Dickey-Fuller (ADF) Test for Original Series Data

Augmented Dickey-Fuller Unit Root Tests							
Type	**Lags**	**Rho**	**Pr < Rho**	**Tau**	**Pr < Tau**	**F**	**Pr > F**
Zero Mean	6	1.1570	0.9303	3.51	0.9998		
	8	1.1107	0.9240	2.25	0.9936		
Single Mean	6	0.4372	0.9731	0.90	0.9948	7.77	0.0010
	8	0.5277	0.9759	0.86	0.9943	3.10	0.2917
Trend	6	-6.2848	0.7049	-1.56	0.7974	1.90	0.7982
	8	-16.1655	0.1144	-2.14	0.5118	3.08	0.5

Table 3: Augmented Dickey-Fuller (ADF) Test for Differenced Series Data

Augmented Dickey-Fuller Unit Root Tests							
Type	**Lags**	**Rho**	**Pr < Rho**	**Tau**	**Pr < Tau**	**F**	**Pr > F**
Zero Mean	6	-123.534	0.0001	-4.55	<.0001		
	8	-71.7947	<.0001	-3.60	0.0004		
Single Mean	6	-126.595	0.0001	-4.54	0.0003	10.32	0.0010
	8	-73.6621	0.0013	-3.59	0.0070	6.46	0.0023
Trend	6	-127.748	0.0001	-4.55	0.0017	10.39	0.0010
	8	-73.7795	0.0005	-3.60	0.0332	6.56	0.0440

From Table 1 we observed that, probability of original series data at zero mean, single mean and trend less than 0.05, it indicated that series is non stationary, but we need a stationary data, hence the next step is, we go for differences in the series data from table 2, we observed that the probability for original series data at zero mean, single mean and trend greater than 0.05, it indicated that the series is stationary. From figure 1 and 2, also showing same results, this difference series can be used for model identification, parameter estimation and selection, modal validation and forecasting

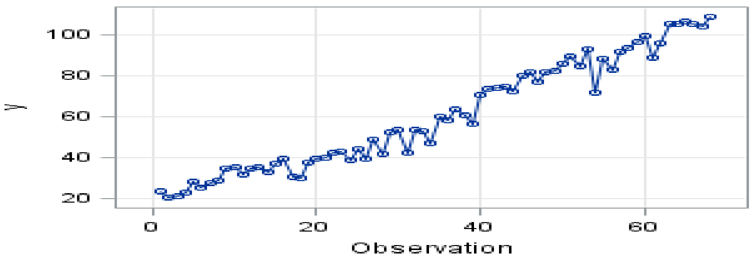

Figure 1: Line Plot of the Original Series Rice Production Data

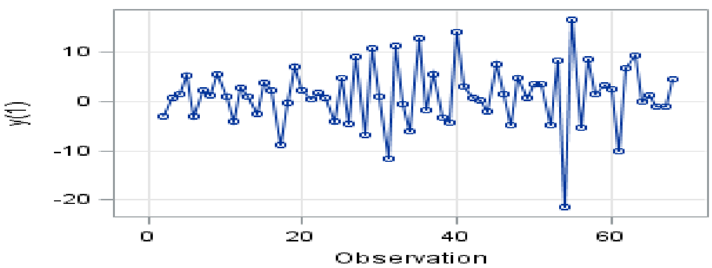

Figure 2: Line Plot of the First Order Differenced Rice Production Data

Estimation and Testing

After identifying the suitable ARIMA (p, d, q)'s structure, subsequent steps of parameter estimation and testing were performed. Estimation stage, consists of using the data to estimate and make inferences about the parameters of tentatively identified model. The parameters are estimated such that an overall measure of residuals is minimized. The last stage of model building, is the testing or diagnostic checking of model adequacy. After identifying the tentative model, the process was again followed by the stage of parameter estimation and model verification. Diagnostic information may help to suggest alternative model(s). Now, the series is stationary and several models were selected based on their ability of reliability prediction. On examining its autocorrelation functions (ACF), partial autocorrelation functions (PACF), on the basis of minimum AIC, BIC, significance of AR and MA parameters, the ARIMA (*011*) model is selected. Parameter estimates along with corresponding standard errors of fitted ARIMA (*011*) model were reported in Table 4.

Table 4: AIC and BIC Values for Different ARIMA Models

Model	AIC	BIC
ARIMA(011)	406.59	411.23
ARIMA(110)	416.06	420.54
ARIMA(111)	408.13	414.75
ARIMA(210)	412.94	419.56
ARIMA(012)	408.54	418.92
ARIMA(211)	410.10	418.36
ARIMA(112)	408.21	414.84
ARIMA(212)	411.92	422.94

Table 5: Parameter Estimates along with Corresponding Standard Errors

Conditional Least Squares Estimation					
Parameter	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**	**Lag**
MU	1.28268	0.17469	7.34	<.0001	0
MA1,1	0.72038	0.08613	8.36	<.0001	1

Diagnostics Checking

The models that are estimated are acceptable only when, the residuals are random. For this purpose, several alternative models that may be appropriate were to be fitted. The ACF and PACF of the residuals of these models are then estimated. If the plot of these ACF and PACF exhibit a non-significant pattern, then the corresponding model is considered as valid and can be considered for forecasting. Figure 3 shows ACF and PACF of the residuals of fitted ARIMA (011) model, exhibit a non-significant pattern, this model is considered as valid and can be considered for forecasting

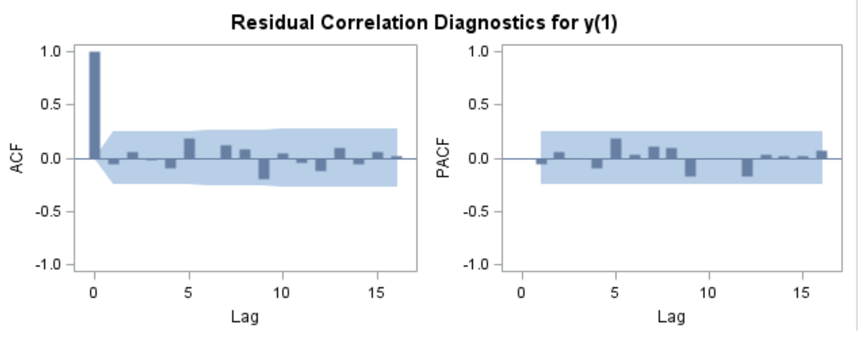

Figure 3: ACF and PACF of the Residuals of Fitted ARIMA Model

CONCLUSIONS

Forecasting: The model that satisfies all the diagnostic checks was considered for forecasting. For forecasting, our objective was to predict the 4 future values of time series. Table 4 shows yearly forecasted results with confidence limits for time series. It has been found that, there is a significantly increasing trend in the total rice production in India, the projected production would be increased to 112.90 million tonnes by 2020 AD.

Table 6: Forecasts of Rice Production (In Million Tonnes)

Year	Forecast	Std Error	95% Confidence Limits	
2017-18	109.06	4.95	99.35	118.77
2018-19	110.34	5.14	100.26	120.42
2019-20	111.62	5.32	101.18	122.07
2020-21	112.90	5.50	102.11	123.69

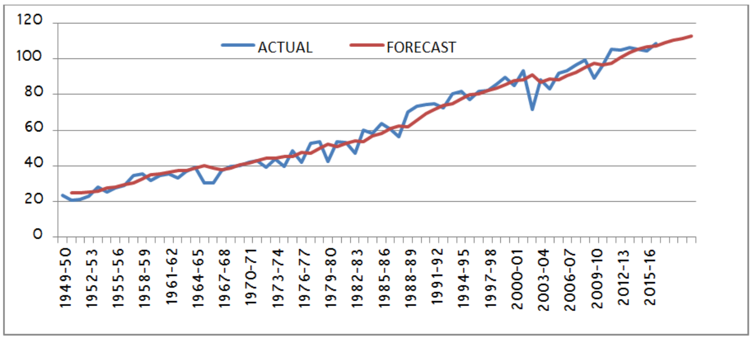

Figure 4: Rice Production Actual and Forecasted Data in Million Tonnes for the Year 1949 to 2020

REFERENCES

1. Box GEP and Jenkin GM 1976. Time series of analysis. Forecasting and Control, Sam Franscico, Holden Day, California, USA.

2. Meyler, Aidan; Kenny, Geoff and Quinn, Terry (1998), Forecasting Irish inflation using ARIMA models, Central Bank and Financial Services Authority of Ireland Technical Paper Series, Vol. 1998, No. 3/RT/98 (December 1998), pp. 1-48.

3. Prajneshu and Venugopalan R. 1998. On nonlinear procedure for obtaining length – weight relationship. *Indian Journal of Animal Sciences* 68 (1), 452-456.

4. R K Paul, Wasi Alam and A K Paul 2014. Prospects of livestock and dairy production in India under time series framework. *Indian Journal of Animal Sciences 84 (4): 462–466.*

5. Singh S, RamasubramanianV and Mehta S C. 2007. Statistical Models for forecasting rice production of India. *Journal of Indian Society of Agricultural Statistics* 61 (2), 2007: 80- 83.

6. Haridev Singh. E 2013 Forecasting Tourist Inflow in Bhutan using Seasonal ARIMA. *International Journal of Science and Research* 9 (2), 242-245