INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
1560–8530/2000/02–4–352–353

Forecasting of Wheat Production in Pakistan using Arima Models

NADEEM SAEED, ASIF SAEED†, MUHAMMAD ZAKRIA, TARIQ MAHMOOD BAJWA‡
Departments of Mathematics & Statistics, and †Plant Breeding & Genetics, University of Agriculture, Faisalabad–38040, Pakistan; ‡College of Veterinary Science, Lahore, Pakistan

ABSTRACT

The paper describes an empirical study of modeling and forecasting time series data of Wheat Production in Pakistan. The Box Jenkins ARIMA methodology has been used for forecasting. The diagnostic checking has shown that ARIMA (2, 2, 1) is appropriate. The forecasts from 1998-99 to 2012-13 are calculated based on the selected model. These forecasts would be helpful for the policy makers to foresee ahead of time the future requirements of grain storage import and/or export and adopt appropriate measures in this regard.

Key Words: Forecasting; Wheat production; Arima models

INTRODUCTION

Wheat being the staple food grain plays a remarkable role in meeting the diversified food requirements of both the Urban and Rural population of the country. Area under wheat has remained around 7.2 and 7.7 million hectares between 1991-92 and 1998-99. Efforts have moreover been made to increase wheat yield. During this period, wheat production increased from 11.30 million tons to 14.41 million tons by making optimum use of improved farm inputs (Anonymous, 1998-99).

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; 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, 1927), Bartlett (1964), Quenouille (1949), Ljune and Bos (1978) and Pindyck and Tubinfeld (1981) have also emphasized the use of ARIMA models.

In this study, these models were applied to forecast the production of wheat crop in Pakistan. This would enable to predict expected wheat production for the years from 1998 onward. Such an exercise would enable the policy makers to foresee ahead of time the future requirements for grain storage, import and/or export of wheat thereby enabling them to take appropriate measures in this regard. The forecasts would thus help save much of the precious resources of our country which otherwise would have been wasted.

MATERIALS AND METHODS

Our aim is to model the annual production of wheat for the years 1947-48 to 1988-89 and to forecast the production of future years. A choice is often to be made as to which type of the model should be developed. This choice may be difficult. A choice was made from a class of liner time series models introduced by Box and Jenkins (1970), which are now widely used and accepted. According to them, the ARIMA model is denoted by ARIMA (p, d, q) where 'p' is the order of the autoregressive process; 'd' is the order of homogeneity i.e. the number of difference to make the series stationary; and 'q' the order of the moving average process. The general from of the ARIMA (p, d, q) is

$$Z_t = C + (F_1 Z_{t-1} + ... + F_p Z_{t-p}) - (\theta_1 a_{t-1} + ... + \theta_q a_{t-q}) + a_t$$

Here C is a constant Z_{t-1}, ...,Z_{t-p} are pasts series values (1ags), the Φ_i's are the coefficients, similar to regression coefficients, to be estimated of the autoregressive model where autoregressive (AR) model of order p, denoted by AR (p) is

$$Zt = C + F_1 Z_{t-1} + F_2 Z_t\text{-}2 + \theta_p a_t\text{-}p + a_t$$

a_t is a random variable with zeros mean and constant variance. Θ's are coefficients in the moving average (MA) model, where moving average model of order q or MA (q) is

$$Z_t = a_t + \theta_1 a_{t-1} - \theta_2 a_{t-2} ... - \theta_q a_{t-q}$$

RESULTS AND DISCUSSION

As stated above, the annual production of wheat crop have been used for modeling purposes. The used data associated with this crop is for the years from 1947-48 to 1998-99 (Anonymous, 1998–99). The modeling of the time series involved the steps of model specification, model estimation, diagnostic checking and forecasts. The plot of the second differenced series showed that the parameter d is 2' the correllogram of autocorrelation function of the second differenced series falls of quickly after lag 1, so q=1, the correllogram for the partial autocorrelation function falls off quickly after lag 2 , so the p=2, the suggested that ARIMA (2,2,1) is appropriate for this data.

The model AIMA (2,2,1) is estimated using the Minitab statistical computer package. The brief output is given below.

Final Estimates of Parameters

Type	Estimate	St. Deviation	t-ratio
AR 1	-0.6337	0.1816	-3.49
AR 2	-0.0424	0.1810	-0.23
MA 1	0.8583	0.1310	6.45

Differencing: 2 regular Differences

No. of observations: Original series 42, after differencing 40 Residuals SS=22379630 (backforecasts excluded)

MS = 508628 DF = 44

Modified Box-Pierce chisquare statistic

Lag	12	24	36	48
Chisquare	18.9(7)	33.3(19)	46.5(44)	54.3(43)

The modified Box and Pierce (1970) statistic for wheat production, calculated above, for lag 12 is 18.9 which has the observed significance level 0.7197. It indicates that it is non significant at SL = 0.05.

Table I. Forecasts for Wheat Production (million tons)

Period	Forecast	Lower	Upper
1998-1999	1485.3	17087.2	19883.5
1999-2000	19641.7	18073.6	21209.8
2000-2001	19749.9	18180.4	21319.4
2001-2002	20491.5	18770.7	22212.4
2002-2003	20876.2	19075.6	22676.7
2003-2004	21460.2	19.483.2	23437.2
2004-2005	21933.0	19783.2	24082.8
2005-2006	22467.9	20089.7	24846.0
2006-2007	22968.1	20344.2	25592.1
2007-2008	23487.7	20583.1	26392.3
2008-2009	23996.5	20790.4	27202.5
2009-2010	24511.2	20978.8	28043.7
2010-2011	25022.7	21144.3	28901.0
2011-2012	25536.0	21292.0	29780.0
2012-2013	26048.3	21421.4	30676.2

Hence, the fit is good. Also for the normal score correlation test (test of correlation between residuals and its normal scores), the critical value at SL =0.01 for n = 40 is 0.960 and for SL = 0.05 is 0.972. The current observed correlation is 0.978, so the normality of residuals is assured.

Using ARIMA (2,2,1) the 15 years ahead forecasts and their 95% confidence interval are calculated and are given in Table I.

REFERENCES

Anonymous, 1998–99. *Economic Survey of Pakistan,* Government of Pakistan.

Bartlett, M.S., 1964. "On The Theoretical Specification of Sampling Properties of Autocorrelated Time Series". *J. Roy. Stat. Soc.,* B 8: 27–41.

Box, G.E.P. and D.A. Pierce, 1970. "Distribution of Residual Autocorrelations in Atuoreressive-Integrated Moving Average Models." *J. American Stat. Assoc.,* 65: 1509–26.

Box, G.E.P. and G.M. Jenkins, 1976. *Time Series Analysis: Forecasting and Control*. Rev. Ed. San Francisco. Holden-Day.

Brown, R.G., 1959. *Statistical Forecasting For Inventory Control*. New York, McGraw-Hill,

Holt, C.C., F. Modigliani, J.F. Muth and H.A. Simon, 1960. *Planning, Production, Inventores, and Work Force*. Prentice Hall, Englewood Cliffs, NJ, USA.

Jenkins, G. M. and D.G. Watts, 1968. *Spectral Analysis and its Application,* Day, San Francisco, California, USA.

Ljunge, G.M. and G.E.P. Box, 1978. "On a Measure of Lack of Fit in Time Series Models". *Biometrika,* 65: 67–72.

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

Meese, R. and J. Geweke, 1982. *A Comparison of Autoregressive Univariate Forecasting Procedures for Macroeconomic Time Series*. Unpublished Manuscript, University of California, Berkeley, CA, USA.

Prindycke, R.S. and D.L. Rubinfeld, 1981. *Econometric Models and Economic Forecasts,* 2nd Ed. New York, McGraw-Hill.

Quenouille, M.H., 1949. "Approximate Tests of Correlation in Time-Series. *J. Roy. Stat. Soc.,* B11: 68–84.

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 To Wolfer's Sunspot Number". *Phill. Trans.,* A 226: 267–98.

(Received 20 July 2000; Accepted 15 September 2000)