JOURNAL OF AGRICULTURE & SOCIAL SCIENCES
1813–2235/2005/01–2–120–122
http://www.ijabjass.org

Use of the ARIMA Model for Forecasting Wheat Area and Production in Pakistan

NAJEEB IQBAL, KHUDA BAKHSH[1], ASIF MAQBOOL AND ABID SHOHAB AHMAD†
Department of Environmental and Resource Economics and Development Economics, University of Agriculture, Faisalabad–38040, Pakistan
†*Agricultural Analyst, National Commodity Exchange Limited, Karachi–Pakistan*
[1]Corresponding author's e-mail: kbmultan@hotmail.com

ABSTRACT

Wheat is the single most important crop, which plays vital role in the Pakistan's economy. It is imperative to assess scientifically the accurate future production potentials of this crop on the basis of past trends. Thus, the study has been made to forecast the area and production of wheat in Pakistan up to the year 2022. The ARIMA model showed that production of wheat would be 29774.8 thousand tons in 2002-22. The scope of higher area and production lies in adequate availability of inputs, educating and training the farming community, soil conservation and reclamation, and especially the supportive government policies regarding wheat cultivation in the country.

Key Words: ARIMA; Area; Forecasting; Production; Wheat; Pakistan

INTRODUCTION

Agriculture is the backbone of the Pakistan's economy because it contributes to the economic and social well being of the nation through its influence on the gross domestic product, employment and foreign exchange earnings. In food grain crops, wheat and rice are the most important in the agriculture sector, in that rice contributes 5.4% of value added in agriculture and 1.3% to GDP and wheat accounts for 13.8% in value added in agriculture and 3.4% in GDP (Government of Pakistan, 2004). Agriculture plays an important role in the betterment of the large proportion of the rural population in particular and overall economy in general. Agricultural development is desired in almost every part of the world today. The race between increasing population and food supply is a real grim. Wheat is the main staple food for the people of Pakistan. The unprecedented drought and water shortage conditions have severely affected the wheat crop during the last years.

Many studies have been conducted to forecast and determine constraints in the production of major crops such as wheat, cotton and rice in Pakistan. Despite these constraints, there are indeed good prospects for continued growth in the area and yield of wheat and other crops in Pakistan (Hamid *et al.,* 1987; Muhammad, 1989; Younis, 1995). Qureshi *et al.* (1992) analysed the relative contribution of area and yield to total production of wheat and maize in Pakistan and concluded that there was more than 100% increase in total wheat production that can be attributed to yield enhancement. Muhammad *et al.* (1992) conducted an empirical study of modeling and forecasting time series data of rice production in Pakistan. ARIMA model has been frequently employed to forecast the future requirements in terms of internal consumption and export to adopt appropriate measures (Muhammad *et al*., 1992; Shabur & Haque, 1993; Sohail, *et al*., 1994).

Factually other crops in general and wheat in particular provide linkages through which it can stimulate economic growth in other sectors. Wheat cultivation has been suffering from various problems, such as traditional methods of farming, low yields, shortage of key inputs and shortage of irrigation water. Pakistan has experienced ups and downs in wheat production. Prices of wheat and flour boosts up during low production seasons and falls drastically when there is a surplus wheat production, however, surplus wheat production occurred for few years and during such periods farming community suffered heavy losses due to inadequate marketing facilities in the country. On the other hand, farmers do not know future prospect of wheat production and prices while deciding to cultivate this and other crops. There is a dire need to forecast area, yield and production of wheat in Pakistan. Therefore, the objective of this paper is to determine future prospects of wheat in the four provinces of the country using past trends.

METHODOLOGY

Respective time series data for this study were collected from Government Publications such as Agricultural Statistics of Pakistan and Pakistan Economic Survey. Box and Jenkins (1976) linear time series model was applied. Auto Regressive Integrated Moving Average (ARIMA) is the most general class of models for forecasting a time series. Different series appearing in the forecasting equations are called "Auto-Regressive" process. Appearance of lags of the forecast errors in the model is called "moving average" process. The ARIMA model is denoted by ARIMA (p,d,q), where "p" stands for the order of the auto regressive process, 'd' is the order of the data stationary and 'q' is the order of the moving average process. The general form of the ARIMA (p,d,q) can be written as described by Judge, *et al*. (1988).

$$\Delta^d y_t = \delta + \theta_1 \Delta^d y_{t-1} + \theta_2 \Delta^d y_{t-2} + \text{-----} + \theta_p y_{t-p} + e_{t-1} \alpha\, e_{t-1} - \alpha_2\, e_{t-2}\, \alpha_q e_{t-2} \quad (1)$$

Where, Δd denotes differencing of order d,i.e., $\Delta y_t = y_t - y_{t-1}$, $\Delta^2 y_t = \Delta y_t - \Delta_{t-1}$ and so forth, Y_{t-1} ----- y_{t-p} are past observations (lags), δ, θ_1 -------- θ_p are parameters (constant and coefficient) to be estimated similar to regression coefficients of the Auto Regressive process (AR) of order "p" denoted by AR (p) and is written as

$$Y = \delta + \theta_1 y_{t-1} + \theta_2 y_{t-2} + \text{---------} + \theta_p y_{t-p} + e_t \quad (2)$$

Where, e_t is forecast error, assumed to be independently distributed across time with mean θ and variance $\theta_2 e$, e_{t-1}, e_{t-2} ------ e_{t-q} are past forecast errors, α_1, ---------- α_q are moving average (MA) coefficient that needs to be estimated.

While MA model of order q (i.e.) MA (q) can be written as

$$Y_t = e_t - \alpha_1 \alpha_{t-1} - \alpha_2 e_{t-2} \text{ ------------ } \alpha_q e_{t-q} \quad (3)$$

The major problem in ARIMA modeling technique is to choose the most appropriate values for the p, d, and q. This problem can be partially resolved by looking at the Auto correlation function (ACF) and partial Auto Correlation Functions (PACF) for the series (Pindyk & Rubinfeld, 1991). The degree of the homogeneity, (d) i.e. the number of time series to be differenced to yield a stationary series was determined on the basis where the ACF approached zero.

After determining "d" a stationary series Δd yt its auto correlation function and partial autocorrelation were examined to determined values of p and q, next step was to "estimate" the model. The model was estimated using computer package "Minitab".

Diagnostic checks were applied to the so obtained results. The first diagnostic check was to draw a time series plot of residuals. When the plot made a rectangular scatter around a zero horizontal level with no trend, the applied model was declared as proper. Identification of normality served as the second diagnostic check. For this purpose, normal scores were plotted against residuals and it was declared in case of a straight line. Secondly, a histogram of the residuals was plotted. Finding out the fitness of good served as the third check. Residuals were plotted against corresponding fitted values: Model was declared a good fit when the plot showed no pattern.

Using the results of ARIMA (p,q,d), forecasts from 2002 up to 2022 were made. These projections were based on the following assumptions.

- Absence of random shocks in the economy, internal or external.
- Agricultural price structure and polices will remain unchanged.
- Consumer preferences will remain the same.

RESULTS AND DISCUSSION

The maximum area of wheat was 8463.0 thousand hectares in 2000 and was minimum 5797.99 in 1972. For production, the maximum production of wheat in Pakistan was obtained in 1999-2000 year (21078.6 thousand tons) and minimum in 1971-72 year (6476.3 thousand tons). The ARIMA model was applied according to four steps namely model specification, model estimation, diagnostic checking and forecast. These four steps are explained in the text while discussing forecasts of wheat area, and production.

Forecasts of wheat area and production up to year 2021-2022. Last thirty years data of area and production of wheat was used for modeling purpose. The model specification involved the plots of the auto correlation function (ACF), partial auto correlation function (PACF) and the plot of the differenced series. Auto correlation function indicated the order of the autoregressive components 'q' of the model, while the partial correlation function gave an indication for the parameter p. First step was to check the stationarity of the data. The time series plot of area and production showed an increasing trend. Auto correlation function of both series showed non-stationary as auto correlation function did not fall as quickly as the log (k) increased. To make the series stationary for area and production, differenced series was used and first difference series of wheat area and production showed stationarity. To check the further stationarity, second difference of the original series for area and production was taken. The autocorrelation function of second difference series and correlogram showed some more stationarity than that of the first difference. Time series data of original series of wheat area and production was not much beneficial to get a stationary series, therefore, the value of parameter 'd' was decided to be equal to '1'. After this the values of autoregressive (AR) parameter "p" and moving average (MA) parameter 'q' was determined from correlograms of partial autocorrelation function and the auto correlation function, respectively. The correlogram of auto correlation function of first differenced series showed that the auto correlation function falls quickly after lag 1 for area and lag 2 for production, hence, the respective values of the parameter "q" was decided to be 1 and 2. Partial auto correlation function of the first differenced series of area and production was used to determine the parameter "p". It was observed that partial auto correlation function fell after lag 1 for area and lag 2 for production. Thus the value of "p" was decided equal to 1 and 2 for area and production respectively which gave good results consequently, the respective values of p,d,q were determined for ARIMA i.e. ARIMA (1,1,1) and ARIMA (2,1,2).

Model estimation. ARIMA (1,1,1) and (2,1,2) models were estimated using MINITAB computer package and estimation of the models for the wheat area and production data are given in Tables I and II.

Diagnostic checking. For diagnostic checking of the estimated models, different diagnostic checks were applied for whether these were properly fitted or not. Goodness of fit of the models is given as under:

Residual analysis. One of the indicators of the properly fitted model is that of scattered residuals in a rectangular shape around the zero at horizontal level. The time series plot of residuals of area and production data showed scatter trend, therefore, models were fitted properly by residual analysis.

For first normality test, plot of normal scores and residuals for area and production data of wheat showed an approximate straight line showing normality, which is necessary condition for normality. The second normality test was to plot the histogram of residuals. If the histogram shows normality, the model is a good fit. The histogram of residuals of wheat area and production series showed normality. Residuals versus fitted values of wheat area and production showed that there were no regular patterns found indicating goodness of fit.

Forecasts of area and production. ARIMA (1,1,1) and ARIMA (2,1,2) were taken for 20 year ahead and forecasts for wheat area and production which are given in Table III along with 95% confidence interval values. For 2000-2001 forecasts of wheat area was about 8451.5 thousand hectares with lower and upper limits of 8095.6 and 8807.3 thousand hectares, respectively. A wheat area forecast for the year 2022 was 8475.1 thousands hectares with lower and upper limits of 3246.2 and 13704.2 thousand hectares, respectively. Forecasts of wheat production showed an increasing trend. For 2000-2001, a forecast of wheat production was about 20670.8 thousands tons with lower and upper limits of 19006.8 and 22334.0 thousand tons respectively. Wheat production forecast for the year 2022 came to be about 29774.8 thousand tons with lower and upper limits of 27542.1 and 32007.6 thousand tons, respectively.

SUGGESTIONS

Selection of high yielding varieties, massive education of farmers through a net work of agriculture officers to make improvements in cultural practices, adequate supply of inputs and full scale use of latest technology are important to increase the yield and production in future.

Total cropped area will increase in future, if reclamation and conservation measures are adopted. Gypsum should be provided to farmers on free of cost to control salanity in their lands. In barani areas, artificial irrigation methods should be used to increase the area under cultivation.

The supply projections of an agricultural commodity, especially wheat plays a vital role in the adjustments of supply and demand in the future. These projections help the government to make policies with regard to relative price structure, production and consumption and also to establish relations with other countries of the world.

Table I. Final estimates of area parameters

Type	Coefficients	St. Dev.	t-ratio
AR_1	-1.6351	0.2492	-6.56
AR_2	-1.4860	0.2973	-5.0
AR_3	-0.7226	0.2101	-3.44
MA_1	-0.0581	0.3101	-0.19

Differencing: 2 regular differences, No. of observations: Original Series 31, after differencing 29, Residuals SS = 823707 (back forecast excluded), MS = 32948, DF = 25

Modified Box-Pierce (Chi. Square-statistics)

Lag	12	24	36	48
Chi-square	13.5(DF=8)	20.4(DF=26)	DF**	DF = 0

Table II. Final Estimates of Production Parameters

Type	Co-efficient	St. Dev.	t. ratio
AR_1	-0.9283	0.1681	-5.52
AR_2	-0.5696	0.1807	-3.15
MA_1	1.0595	0.1639	6.46

Differencing=2 regular differences, Number of observations= Original series 31, after differencing 29; Residuals SS = 256859 (back forecast excluded), MS = 9879 DF: 26

Modified Box – Pierce chi-square statistics

Lag	12	24	36	48
Chi-square	7.1(DF=9)	17.5(DF=21)	DF=**	DF**

Table III. Forecast of wheat area up to year 2022

Years	Area (000 hectares)	Production (000 tons)
2000-01	8451.5 ± 355.9	20670.8 ± 1664.0
2001-02	8291.0 ± 386.4	20882.1 ± 1670.3
2002-03	8359.4 ± 510.7	20763.6 ± 1828.7
2003-04	8296.5 ± 686.3	21912.9 ± 1923.5
2004-05	8451.3 ± 853.4	21942.4 ± 1932.4
2005-06	8279.9 ± 982.3	22486.6 ± 2007.4
2006-07	8413.3 ± 1226.6	23055.0 ± 2023.2
2007-08	8375.7 ± 1412.8	23349.0 ± 2037.8
2008-09	8400.5 ± 1613.2	23908.3 ± 2055.0
2009-10	8357.1 ± 1846.7	24333.1 ± 2058.5
2010-11	8456.1 ± 2094.2	24764.8 ± 2062.4
2011-12	8378.5 ± 2314.0	25254.8 ± 2063.3
2012-13	8427.3 ± 2582.9	25682.6 ± 2063.3
2013-14	8429.0 ± 2847.2	26144.9 ± 2063.6
2014-15	8447.5 ± 3110.3	26602.6 ± 2066.1
2015-16	8417.2 ± 3389.5	27048.2 ± 2071.4
2016-17	8475.7 ± 3687.2	27508.3 ± 2080.8
2017-18	8449.4 ± 3972.0	27959.6 ± 2096.2
2018-19	8465.1 ± 4277.7	28412.7 ± 2117.8
2019-20	8474.0 ± 4592.0	28868.2 ± 2147.2
2020-21	8492.8 ± 4906.8	29320.4 ± 2185.4
2021-22	8475.1 ± 5228.9	29774.8 ± 2232.7

REFERENCES

Box, G.E.P. and G.M. Jenkins, 1976. *Time Series Analysis, Forecasting and Control*. San Francisco, Holden–Day, California, USA.

Government of Pakistan, 2004. *Economic Survey 2003–2004. Finance Division,* Economic Adviser's Wing, Islamabad.

Hamid, N., P. Thomas, V. Alberto and G. Suzanne, 1987. The wheat economy of Pakistan setting and prospects. *International Food Policy Research Institute, Ministry of Food and Agriculture*, Government of Pakistan, Islamabad, Pakistan

Judge, G.G., R. Carter Hill, E.G. William and I. Helmut, 1988. *Introduction to the Theory and Practice of Econometrics.* 2nd ed., John Wiley and Son, INC. New York, Toronto, Singapore

Muhammad, F., M. Siddique, M. Bashir and S. Ahmad, 1992. Forecasting rice production in Pakistan–using ARIMA Models. *J. of Animal Plant Sci.*, 2: 27–31

Muhammad, K., 1989. *Description of the Historical Background of Wheat Improvement in Baluchistan, Pakistan.* Agriculture Research Institute, Sariab, Quetta, Baluchistan, Pakistan

Pindyck, R., S. Daniel and L. Rubinfed, 1991. *Economic Models and Economic Forecasts,* 3rd ed. McGraw Hill International editions (Economic Survey) New York, USA

Qureshi, K., A.B. Akhtar, M. Aslam, A. Ullah and A. Hussain, 1992. An analysis of the relative contribution of area and yield to total production of wheat and maize in Pakistan. Directorate of crop production and Management, University of Agriculture, Faisalabad, *Pakistan. J. Agric. Sci.,* 29: 166–9

Shabur, S.A. and M.E. Haque, 1993. An analysis of rice price in Mymensing town market pattern and forecasting. *Bangladesh J. of Agric. Econ.*, 16: 120–3

Sohail, A., A. Sarwar and M. Kamran, 1994. Forecasting total food grains in Pakistan. Department of Math and Stat, University of Agriculture, Faisalabad. *J. Eng. Appl. Sci.*, 13: 140–6

Younis, M., 1995. *Present Status and Future Prospects of Wheat in Pakistan–An Economic Review.* Agricultural Economic Section, Ayub Agricultural Research Institute, Faisalabad. Research Study No. 30.

(Received 10 January 2005; Accepted 20 February 2005)