Time Series Model to Forecast Area of Mangoes from Pakistan: An Application of Univariate Arima Model

Farooq Ahmed Jam

University of Malaya, Malaysia

Shafaqat Mehmood

Faculty of Commerce, University of Central Punjab, Pakistan

Zahid Ahmad (PhD)

Faculty of Commerce, University of Central Punjab, Pakistan

Abstract:
The present study is an attempt to find appropriate Univariate ARIMA model by using Box-Jenkins methodology to forecast the area of mangoes in Pakistan. Time series data for a period of 49 years from 1961 to 2009 were used. This study found ARIMA (0,1,0) as appropriate model to forecast the area. Further, this study concludes that in the year 2025 the area of mangoes in Pakistan would be 318.5 thousand hectares. This study also estimates increase in the area of mangoes in future. The finding of this study will be helpful for government, policy makers, exporters, producers, and importing countries.

Keywords: Forecasting, Arima Model, Area, Mangoes, Pakistan

Paper type: Research Paper

Introduction

Mango is known as "King of Fruits" for many years and it is also considered a major fruit crop of various countries such as Philippines, Mexico, India, Brazil, Pakistan, China and Thailand. A number of fruit varieties are produced and exported by Pakistan but among these fruits, Mangoes have central position due to its taste, color, area, production, and export.

Mangoes are grown all over Pakistan but provinces of Punjab and Sindh are main producers of mangoes. Chaunsa, Anwar Ratol, Fajri Kalan, Langra, Duseri, Samer Bahisht and Malda are popular mango varieties in Pakistan. Growth trend of area from 1960-61 to 2008-09 shows a nonstop increase. Area of mangoes in 1960-61 was 34.4 thousand hectares, while in 2008-09 it was 170.10 thousand hectares.

Problem statement

Various researchers, exporters, producers and farrmers are much interested to study the agriculture information including the areas of Pakistani mangoes. To provide an updated estimation about the forecasting of area of Mangoes in Pakistan over the next few years, a study is needed. Pakisdtan is a large producer of mangoes and many varities are produced in the country. Mangoes are also exported and export prices are much higher than the local prices. The area of growth of mangoes is dependent on the profit of the growers. It is needed to find out that if the existing trends continue what will be the cultivated area of mangoes in the future in Pakistan.

Objectives of study

The objectives of present research are mentioned below:

1. To examine the growth of area of Mangoes in Pakistan
2. To suggest appropriate ARIMA model for the generation of forecasting area of Mangoes in Pakistan
3. To generate forecasts of area of Manges in Pakistan by using appropriate ARIMA models

Data and Methodology

There have been extensive theoretical and empirical research to date that attempt to focus on the relationship between inflation and economic growth, both in the context of developed and developing countries.

Time series data for a period of 49-years from 1960-61 to 2008-09 were taken from FAOSTAT and Economic Survey of Pakistan. Units of area of Mangoes are mentioned in thousand hectares.

Ahmad & Mustafa (2006) mentioned that a number of methods are used for forecosting such as econometric models, judgmental approach, univaraite time series models, structure economic models and multivariate time series models. According to Celia, Ashish, Amar, & Les, (2003) Univariate and multivariate models are considered main modeling methods. Only one explanatory variable i.e. time is used in Univariate methods.

The current issue and full text archive of this journal is available at http://aocrj.org/archive/

Academy of Contemporary Research Journal
V II (I), 10-15, ISSN: 2305-865X

Received: November 2012
Revised: December 2012
Accepted: January 2013

Academy of Contemporary Research Journal
Volume 2, Issue 1, 2013, 10-15

The empirical studies of Mehmood (2012) used Univariate ARIMA model to forecast the Pakistan's exports to SAARC and concluded that ARIMA model is an appropriate model to forecast by using the time series data. Naz (2012) also used same model to forecast the dates exports in Pakistan.

Box and Jenkins (1976) methodology of Univariate ARIMA model is considered the most flexible method and is used by a numbers of researchers for forecasting time series data. This study employs Univariate ARIMA model by considering best among the other methods.

According to the Box and Jenkins (1976) a non-seasonal ARIMA model is denoted by ARIMA (p,d,q). This model is combination of auto regressive (AR) and moving average (MA) with an order of integration or differencing (d). Where p is order of autocorrelation and q is order of moving average. ARIMA in general form is written as:

$$\Delta^d Z_t = C + (\varnothing_1 \Delta^d Z_{t-1} + \ldots + \varnothing_p \Delta^d Z_{t-p}) - (\Phi_1 a_{t-1} + \ldots.. + \Phi_p a_{t-p}) + a_t$$

Where 'C' is a constant and D denotes "difference operator"

$$\Delta Z_t = Z_t - Z_{t-1}$$
$$\Delta^2 Z_{t-1} = \Delta Z_t - \Delta Z_{t-1}$$

$Z_{t-1} \ldots Z_{t-p}$ are values (lags) of past series values and $\varnothing$is the coefficient to be estimated by auto-regressive model. The auto-regressive model of order 'p' denoted by AR (P) is:

$$Z_t = C + \varnothing_1 Z_{t-1} + \varnothing_2 Z_{t-2} + \ldots + \varnothing_p \Delta^d Z_{t-p} + a_t$$

Where a random variable with zero mean and constant variance is denoted by at. In the moving average (MA) model, F is denoting the coefficient. The Moving Average Model is of order 'q' or MA (q) which can be written as:

$$Z_t = a_t - \Phi_1 a_{t-1} - \Phi_2 a_{t-2} - \ldots - \Phi_p a_{t-p}$$

This model is used to analyze the quantitative relationship of data and to forecast area of Mangoes in Pakistan. Box and Jenkins (1976) ARIMA methodology consists of four steps i.e. Identification, Estimation, Diagnostics and Forecast, which are applied in this research.

Data Analysis, Results, and Discussion

In the time series research, the selection of appropriate models is an art. Different researches have mentioned different criteria for selection of model because there is no hard and fast rule for the selection of appropriate model in time series studies. A model with minimum value of BIC, non-significance B-J Q statistic, and with high correlation of coefficient is considered an appropriate model for the purpose of forecasting. In Figure 1 summery indicates that it is Leptokurtic having value 4.2565 and positively skewed with a value of 0.781764.

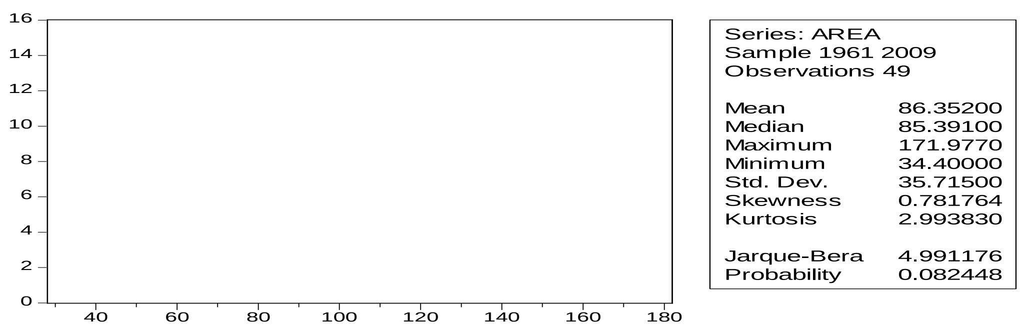

Figure 1. Summery Statistics of Area of Mangoes

Model Identification

For the data stationarity requirement Augmented Dickey-Fuller test, ACF, PACF, Sequence Plots and correlogram are used, which give same results indicating that time series data of area of mangoes in Pakistan become stationary at first difference. The study of Mehmood (2012) used ADF test to check the stationarity of time series in case of investigating relationship between exports and economic growth of Pakistan. Mehmood (2012) also conduct a study on effect of factors on gross domestic product and concluded that ADF is best test to investigate the unit root within the time series data.

Figure 2 shows that at first difference the time series data of area of mangoes becomes stationary. In Figure 3 auto-correlation function shows that at first difference approximately all values of coefficients lie within the upper and lower confidence limits. Further, figure-4 provides same results for Partial Autocorrelation Function at first difference.

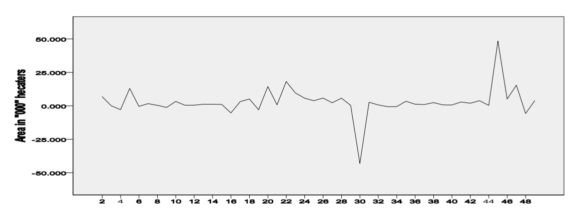

Figure 2. Sequence Plots of Area of Mangoes at first difference

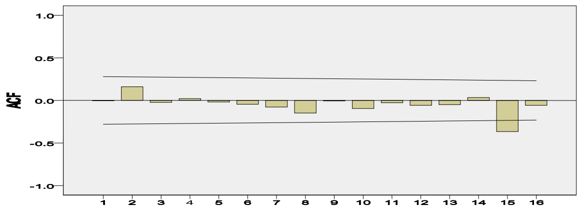

Figure 3. ACF of Area of Mangoes at first difference

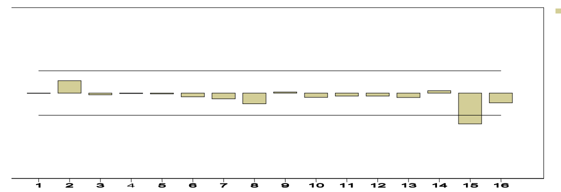

Figure 4. PACF of Area of Mangoes at first difference

Table 1. ADF Test of Area of Mangoes at first difference

		t-Statistic	Probability
Augmented Dickey-Fuller test statistic		-6.744500	0.0000
Test critical values:	1% level	-3.577723	
	2% level	-2.925169	
	3% level	-2.600658	

Hypotheses related to time series data in Augmented Dickey-Fuller test are as follows:
H0: Time series data are non stationary ($p > 0.05$)
H1: Time series data are stationary ($p \leq 0.05$)

Figure 5. Correlogram of Area of Mangoes at first difference

Autocorrelation	Partial Correlation	Lags	AC	PAC	Q-Stat	Prob
. |. |	. |. |	1	-0.004	-0.004	0.0008	0.978
. |*. |	. |*. |	2	0.161	0.161	1.3497	0.509
. |. |	. |. |	3	-0.022	-0.022	1.3758	0.711
. |. |	. |. |	4	0.020	-0.006	1.3987	0.844
. |. |	. |. |	5	-0.018	-0.011	1.4167	0.922
. |. |	. |. |	6	-0.044	-0.048	1.5263	0.958
.*|. |	.*|. |	7	-0.077	-0.074	1.8757	0.966
.*|. |	.*|. |	8	-0.148	-0.139	3.1928	0.922
. |. |	. |. |	9	-0.006	0.014	3.1954	0.956
.*|. |	. |. |	10	-0.095	-0.056	3.7699	0.957
. |. |	. |. |	11	-0.027	-0.037	3.8185	0.975
. |. |	. |. |	12	-0.057	-0.037	4.0386	0.983
. |. |	. |. |	13	-0.049	-0.056	4.2046	0.989
. |. |	. |. |	14	0.033	0.031	4.2822	0.994
***|. |	***|. |	15	-0.366	-0.400	14.041	0.522
. |. |	.*|. |	16	-0.058	-0.128	14.292	0.577

Table 2. Initial estimate of parameters of different ARIMA models

Model	Parameters	Estimate	S.E.	t-ratio	P-value
ARIMA (0, 1, 0)	C	0.033	0.016	2.123	0.039
ARIMA (0,1,1)	C	2.827	1.538	1.838	0.073
	MA1	0.003	0.147	0.018	0.986
ARIMA (1,1,0)	C	2.827	1.536	1.840	0.072
	AR1	-0.004	0.147	-0.026	0.979
ARIMA (1,1,1)	C	2.811	1.503	1.870	0.068
	AR1	-0.757	1.087	-0.696	0.490
	MA1	-0.698	1.164	-0.600	0.551

Diagnostic Checking

This study observed ARIMA (0, 1, 0) Random Walk as an optimal and appropriate model to achieve one part of present study first objective that to forecast the Area of Mangoes in Pakistan. The selected model also approximately fulfills the basic criteria of model selection with minimum value of BIC, non-significance B-J Q statistic and with high correlation of coefficient.

Table 3. Comparative Results from various ARIMA models for Area of mangoes in Pakistan

Model	S.R	R^2	RMSE	MAPE	MAE	MaxAE	MaxAE	BIC
ARIMA (0,1,0)	0.03	.91	10.61	6.20	5.24	58.20	48.13	4.82
ARIMA (0,1,1)	1.12E-5	.91	10.68	6.18	5.15	55.55	45.93	4.89
ARIMA (1,1,0)	1.51E-5	.90	10.68	6.18	5.15	55.55	45.94	4.89
ARIMA (1,1,1)	0.00	.90	10.76	6.15	5.10	56.03	46.35	4.99

Table 4. Optimal model for forecasting Area of mangoes in Pakistan

Model	Parameters	Estimate	S.E.	t-ratio	P-value
ARIMA (0, 1, 0)	C	0.033	0.016	2.123	0.039

Forecasting with optimal Model

Further to Forecast the Area of mangoes in Pakistan with 95% confidence interval, ARIMA (0,1,0) model was used. The forecast values of Area of Mangoes in Pakistan for the period of 16 years, from 2009-010 to 2024-25 are mentioned in Table 5.

Academy of Contemporary Research Journal
Volume 2, Issue 1, 2013, 10-15

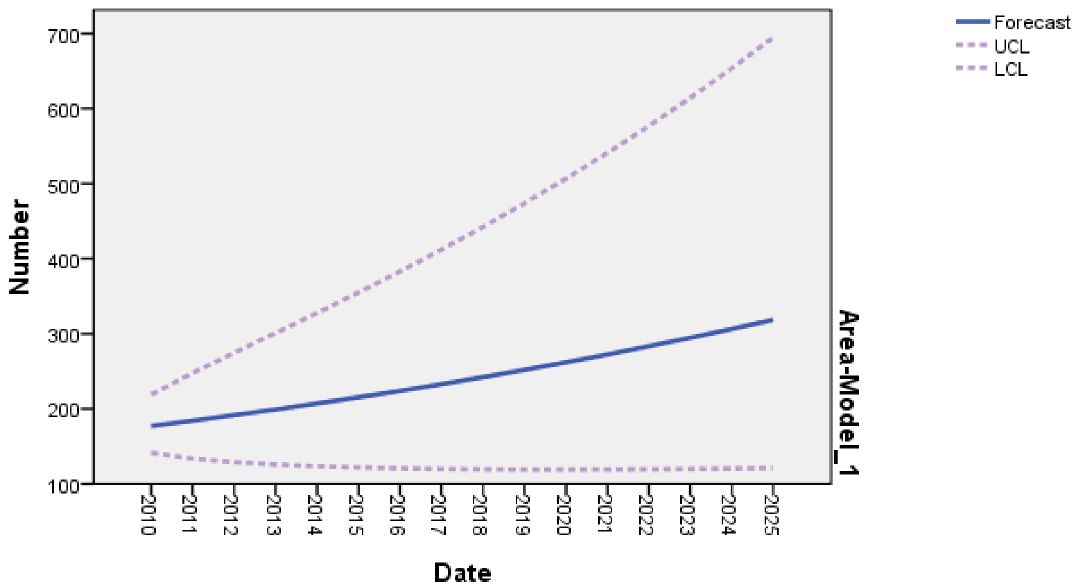

Figure 6. Trend of Forecast values of Area of mangoes

Table 5. Forecast for the Area of mangoes in Pakistan

S. No.	Year	Forecasts in "000" hectares	95 % LCL	95% UCL
1	2009-10	176.901	141.323	218.835
2	2010-11	183.974	133.458	247.691
3	2011-12	191.330	128.714	274.504
4	2012-13	198.979	125.501	300.921
5	2013-14	206.935	123.223	327.589
6	2014-15	215.209	121.587	354.858
7	2015-16	223.814	120.424	382.959
8	2016-17	232.763	119.627	412.058
9	2017-18	242.069	119.124	442.294
10	2018-19	251.748	118.865	473.785
11	2019-20	261.814	118.812	506.638
12	2020-21	272.282	118.937	540.957
13	2021-22	283.168	119.219	576.839
14	2022-23	294.489	119.642	614.384
15	2023-24	306.264	120.192	653.688
16	2024-25	318.509	120.859	694.8529

Conclusions

The following are the conclusions of this study:

- ARIMA is an appropriate model to forecast the area of mangoes in Pakistan.
- ARIMA (0,1,0) is the most appropriate model to forecast area of mangoes in Pakistan.
- The forecast indicates that in the year 2025 the cultivated area of mangoes in Pakistan will be 318,509 hectares, which is 148,409 hectares more than the area today.
- The increase in area of mangoes in Pakistan in the year 2025 will be 87% as compared to today, which is more than 6% per year.

Academy of Contemporary Research Journal
Volume 2, Issue 1, 2013, 10-15

- A comparison with the export of mangoes can be studied to use the finding meaningfully in Pakistan.

References

Ahmad, B., and K. Mustafa, 2006, Forecasting Kinnow Production in Pakistan: An Econometrics Analysis, International Journal of Agriculture & Biology, 455-458.

Box, G.E. and Jenkins, G.M., 1976, Time Series Analysis, Forecasting and Control, Holden Day,San Francisco

Celia, F., G. Ashish, G. Amar and S. Les, 2003, Forecasting Women's Apparel usingMathematical Modeling, International Journal of Clothing, Science and Technology, 15(2), 107-125.

FAO, 2011, FAOSTAT: Food and Agriculture Organization of the United Nations, Retrieved from www.fao.org.com:http://faostat.fao.org/site/570/DesktopDefault.aspx?PageID=570#ancor

FAO, 2010, Food and Agricultue Organization of United Nation.

Mehmood, S. (2012). Dynamics of Exports and Economic Growth at Regional Level: A Study on Pakistan's Exports to SAARC. *Journal of Contemporary Issues in Business Research , 1* (1), 11-19.

Mehmood, S., & Ahmad, Z. (2012). Forecasting Pakistan's Exports to SAARC: An Application of Univiriate ARIMA Model. *Journal of Contemporary Issues in Business Research , 1* (3), 41-54.

Mehmood, S. (2012). Effect of Different Factors on Gross Domestic Product: A Comparative Study of Pakistan and Bangladesh. *Academy of Contemporary Research Journal , 1* (1), 18-35.

Naz, F. (2012). A UNIVARIATE TIME SERIES MODELLING OF DATES EXPORTS IN PAKISTAN. *Journal of Contemporary Issues in Business Research , 1* (2), 37-48.

Pakistan, G.,2009, Agriculture Statistics of Pakistan 2008-09, Ministry of Food and Agriculture, Islamabad: Economic Wing.

Pakistan, G., 2002, Pakistan Economy Survey 2001-02, Finance Division, Islamabad.