The Journal of Animal & Plant Sciences, 29(5): 2019, Page: 1396-1401
ISSN: 1018-7081

FORECASTING THE PRODUCTION OF SUGARCANE IN PAKISTAN FOR THE YEAR 2018-2030, USING BOX-JENKIN'S METHODOLOGY

Q. Mehmood[1], M. H. Sial[1], M. Riaz[2] and N. Shaheen[3]

[1]Government Post Graduate College Bahawalnagar, PhD Scholar at Department of Quantitative Methods, University of Management and Technology Lahore. [1]Department of Quantitative Methods, University of Management and Technology Lahore. [2]Department of Statistics Rahim Yar Khan Campus Islamia University, Bahawalpur. [3]Department of Statistics, Government Post Graduate Girls College Bahawalnagar.
Corresponding Author E-mail: qaisarm11@gmail.com

ABSTRACT

Agriculture is the mainstay of Pakistan's economy and contributes 24 percent of the *GDP. Considering its vital role*, planners and policy makers are always keen to have timely forecasts for the important crops such as wheat, cotton, rice and *sugarcane. Of these sugarcane* is a major cash crop and an important source of income for farmers in Pakistan. The present study is focused on developing and estimating time series models to forecast sugarcane production in Pakistan. Box-Jenkin (1976) methodology was employed to estimate production forecasting model using annual time series data as available from Pakistan Bureau of Statistics (PBS) and various issues of Pakistan Economic Survey for the years 1947-2017. An appropriate ARIMA (2, 1, 1) model was estimated to forecast the production of sugarcane crop in Pakistan for the years 2018-2029. Over this period the model predicts a significant increase (6.56%) in sugarcane output. These forecasts can be very useful for agricultural policy makers, sugar industry as well as farmers in making prudent resource allocation and production decisions for sugarcane in Pakistan.

Keywords: Autoregressive Integrated Moving Average, Model, Error, production, forecast.

INTRODUCTION

Sugarcane is an important cash crop of Pakistan. It is grown mostly for the production of sugar and gur. It is also used as raw material for industries like sugar, chip board, paper, fiber, chemicals, plastics, paints, synthetics, detergents and insecticides etc (GOP, 2012). Pakistan has the fourth position among the sugarcane producing countries: Brazil India, China, Pakistan, Thailand and Mexico (FAO. 2010). The production of sugarcane in our country has been increased much for the last thirty years. In order to increase the production of sugarcane crop, several initiatives were taken by the Government and the sugar mills association to get better yield production.

In South Asia, Pakistan is the largest consumer of sugar with 25.83 kg per capita consumption per year. In this connection, Pakistan produces about 99% of the sugar from sugarcane to meet the requirement and demand of sugar at domestic level (Azam and Khan, 2010). Presently Pakistan is 9th sugar exporting country in the world (USDA, 2019). The production of sugarcane for the years 2015-16 and 2016-17 was 65482 and 75482 tons, while in the year 2017-18 sugarcane production was 81102 tons, this increased amount of production was not fully utilized for sugar production but it created problems for farmers like lowest payments and incomplete purchasing of their crop by mill owners due to mismanagement. Mostly the farmers preferred to make gur, shakr and utilized their crop for animal feed instead of selling to the sugar mills. Sugarcane being a major cash crop and contributing significantly in the agricultural economy of the country, it is valuable to know about the production status of this crop in future. If past values of crop production are known, one can use past pattern of the data to forecast crop production by employing different forecasting models. The objective of this study was to propose an appropriate forecasting model for forecasting the production of sugarcane crop in Pakistan. Various models have been developed to forecast future values; however, in univariate time series analysis Box-Jenkins's, (1976) ARIMA model was extensively used. Forecasting of agricultural production and prices were proposed to be useful for the farmers, governments, and agribusiness industries (Allen, 1994). While production and yield of sugarcane in Pakistan was forecasted by applying ARIMA model and it was predicted that the forecasting values are very close to the actual values (Muhammad et al., 1992; Yaseen *et al.*, 2005).

Similarly in India sugarcane yield was forecasted through ARIMA model considering weather variable as a regressor (Suresh and Priya, 2011). Taking time series data, forecasting the production and yield of sugarcane and cotton ARIMA model was appropriate for Pakistan (Sajid *et al.,*2015). In Turkey production and area of strawberry was estimated by using ARIMA model, three exponential smoothing, holt's exponential smoothing and Brown's smoothing model were appropriate in forecasting (Akin and Eyduran, 2017). For forecasting the production of major food crops in Khyber Pakhtunkhwa, Pakistan, Box-Jenkin's (ARIMA) forecasting technique was found to be adequate (Shah *et al.,*2017).

MATERIALS AND METHODS

Secondary data of Agriculture Statistics "50 years of Pakistan: volume-iii (1947-1997)" published by Pakistan Bureau of Statistics (PBS) and various issues of Economic Survey of Pakistan (GOP, various issues) were used in this study. The data set covers the history of Pakistan and it ranges from 1947 to 2018. Time series modeling and forecasting has fundamental importance in various useful practical domains. A lot of the forecasting models have been planned in the literature for improving the efficiency and accuracy of forecasting the time series data**.** Due to comprehensiveness of ARIMA model technique it is extensively employed in literature to forecast the specific area as well as production related to different major crops (Gujrati, 2003). Autoregressive Integrated Moving Average *(*ARIMA) model is a generalized form of ARIMA model introduced by Box and Jenkins (1976) which includes both autoregressive as well as moving average parameters, and also includes the differencing in the formulation of this model. ARIMA model is summarized as ARIMA (p, d, q). In ARIMA (p, d, q) model where p, d and q are the non-negative integers referred to as the order of the autoregressive integrated moving average process. It is an important part of Box Jenkins approach to time series modeling. It can be written as;

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \ldots + \varphi_p Y_{t-p} + u_t + \theta_1 \mu_{t-1} + \theta_2 \mu_{t-2} + \ldots + \theta_q \mu_{t-q}$$

Where Y_t is production of sugarcane crop, while Y_{t-1}, Y_{t-2}, …..Y_{t-p} is lag values of the production. φ_1……..φ_p are the parameter of autoregressive model, θ_1,…………θ_q are the parameter of moving average model and u_t, u_{t-1}……….u_{t-q} are the residual term and its lag values.

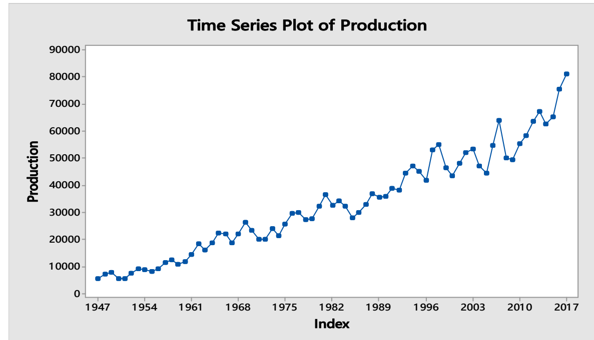

Fig.1. Time Plot of the Sugarcane Production.

Time plot graph for the production of the sugarcane crop shows upward trend, which indicates the observed series was non stationary, to achieve the stationary of the series graphs of Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF), and Augmented Dickey-Fuller (Dickey and Fuller, 1979) of unit root test is constructed.

The stationary of the data is tested by both graphs and Augmented Dickey-Fuller (ADF) unit root Test. From the graphs of Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF), and Augmented Dickey-Fuller (ADF) of unit root test First order Integration (Difference) is suggested to make the stationary of the above series.

In the graph of autocorrelation function there is only one negative spike, we use first order of moving average (MA) model, and there is one positive spike in partial autocorrelation function graph we consider first order Autoregressive (AR) model. Tentative autoregressive integrated moving average (ARIMA) model is ARIMA (1,1,1). Starting from this model we chose the best forecast ARIMA (p,d,q) model by comparing all possible fitted model diagnostically.

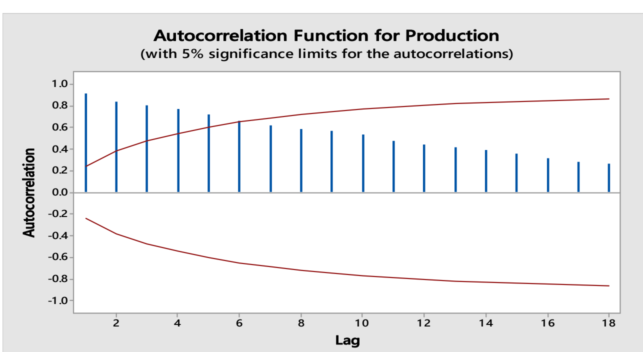

Fig.2. ACF plot for Sugarcane Production

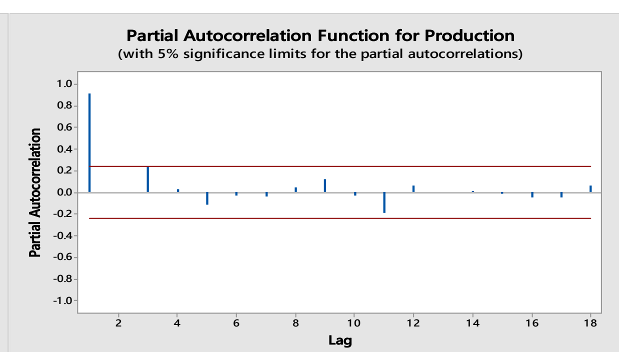

Fig.3. PACF Plot of the Production of Sugarcane

Table 1. Augmented Dickey Fuller Unit Root Test.

Variables	Observed ADF Test	Lag Level	1^{st}. Difference ADF Test	Lag Lavel	Order of Integration
Production	-1.7487	2	-11.6471	1	I (1)

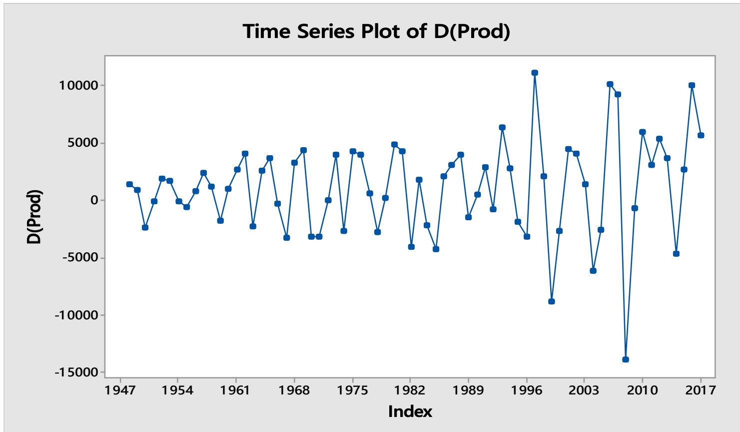

Fig.4 Time Graph of the First difference of the Production

Graph of first difference of the original Production of Sugarcane crop shows the stationary of the data.

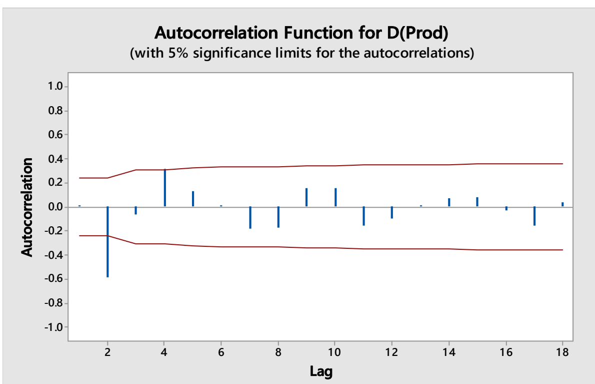

Fig.5. ACF Plot of the First difference Production

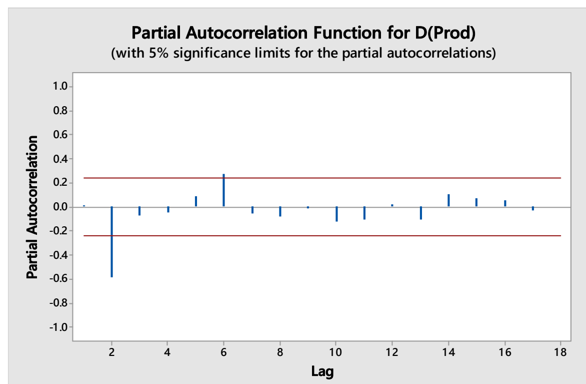

Fig.6. PACF Plot of the First difference Production

Table: 2 Diagnostic Statistics of all possible ARIMA Models.

ARIMA Model	MSE	R Square	MAPE	MAE	BIC	Box-Pierce Test	DF	P.Value
(1, 1, 1)	15410509	0.954	10.140	0.002988	16.821	35.664	16	0.003
(1, 1, 2)	11900927	0.965	9.225	.002612	16.625	21.100	15	0.134
(1, 1, 3)	15303171	0.967	8.847	0.002506	16.650	14.811	14	0.391
(2, 1,1*)	11348544*	0.970*	8.917	0.002559	16.489*	12.975*	15	0.604*
(2, 1, 2)	11509620	0.970	8.922	0.002565	16.564	13.143	14	0.515

After comparing all these possible models we propose ARIMA (2,1,1) model is the best model for forecasting, as its Value of R-Square "0.970" is maximum among all possible models, while Mean Absolute Error, Mean Absolute Percentage Error and Bayesian Information Criterion is minimum for the ARIMA (2,1,1) model among the all possible ARIMA models. All the parameters of ARIMA (2,1,1) model are significant and also the Box-Pierce (Ljung-Box) Chi-Square statistic is satisfy for the proposed model.

RESULTS AND DISCUSSION

Table 3 shows all the estimates of the parameters of ARIMA (2,1,1) model are significant which shows that this model is suitable for forecasting the sugarcane production. Graphically diagnostic checks for ARIMA (2,1,1) Model in the figure 7 and 8 are the residual plot for the autocorrelations and partial autocorrelations functions. From the Plot all the autocorrelations and partial autocorrelations lie between the 95% confidence interval. Thus, the model is correctly specified. Normal probability plot for residuals in Fig. 9 show that the residuals lie along a straight line and the residuals are symmetric showing in the figure 10 for histogram of residuals tells that the model is correctly specified.

After fitting the adequate model it is utilized for forecasting. Our objective in forecasting is to predict the future values of the Production of Sugarcane crop for the year 2018-19 to 2029-30. Figure 11 shows the Time Plot of the Forecast values of Sugarcane Production and table 4 provide the Forecast values for the twelve year at 95% confidence Limits.

Table 3. Estimates of the ARIMA (2,1,1) Models.

ARIMA Model With (2,1,1)				
Type	**Coefficient**	**S.E(Coef)**	**T-test**	**P-Value**
AR(1)	0.1494	0.1507	0.99	0.325
AR(2)	-0.6491	0.1031	-6.30	0.000
MA(1)	0.2595	0.1912	1.36	0.179
Constant	1485.6	299.0	4.97	0.000

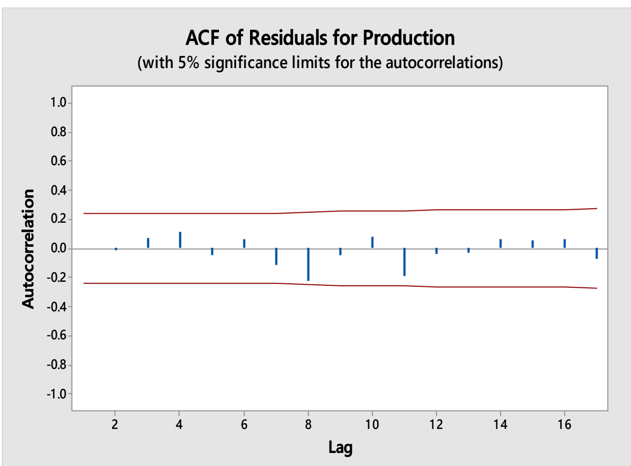

Fig.7. Plot of ACF of Residual for Production

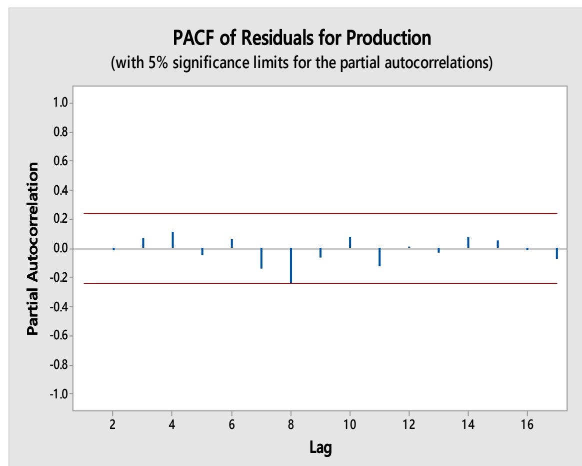

Fig.8. Plot of PACF of Residual for Production

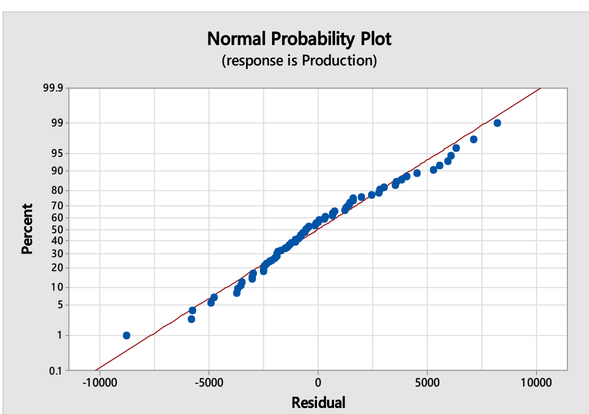

Fig.9. Residuals Normal Probability Plot

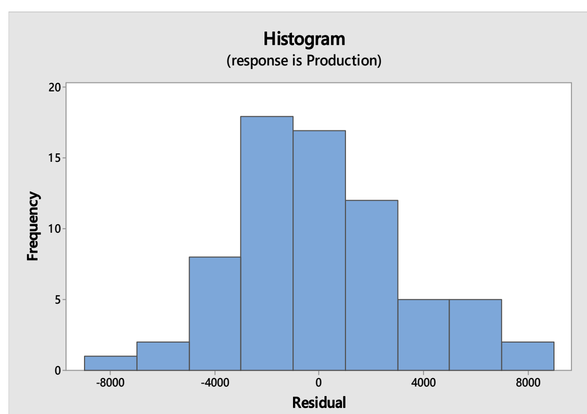

Fig.10. Histogram (Response is Production

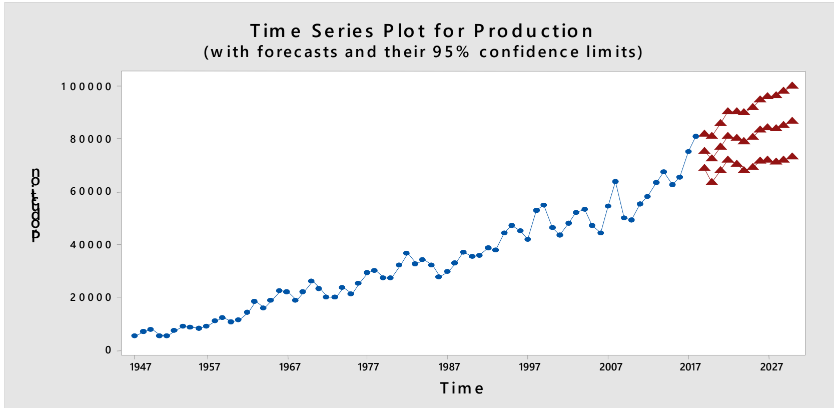

Fig.11. Time Plot of the Forecast Sugarcane Production for the year 2018-2030

Table: 4 Forecast from year 2018-19 to year 2029-30 at 95% Limits.

Period	Forecast	Lower	Upper
2018-19	75394	68790	81998
2019-20	72379	63538	81220
2020-21	77119	68155	86083
2021-22	81270	72212	90327
2022-23	80299	70347	90250
2023-24	78945	67955	89936
2024-25	80859	69489	92228
2025-26	83509	71909	95108
2026-27	84148	72076	96220
2027-28	84009	71324	96695
2028-29	85059	71950	98168
2029-30	86792	73376	100207

Conclusions and Recommendations: From table. 2 results suggest through Box-Jenkin's Methodology the ARIMA (2,1,1) model is best forecasting model for the production of Sugarcane crop for the years 2018-19 to 2029-30. These forecasting results for the production of sugarcane crop are, presented in table. 4 show a significant increase, from 75394 to 86792 tons for the years 2018-19 to 2029-30. Actual production for the year 2017-2018 of sugarcane was 81102 tons highest in the historical data which will be decreased to 72379 tons as forecast production for the year 2019-2020, then it will be increased for the upcoming years gradually. The prediction of this study may provide help for the policymakers in their macro-level policies for food security, reasonable support price and better planning for the production of sugarcane in Pakistan. It also provides help in micro level policies to sugar mill owners, use of high yield varieties, fertilizer, pest control measures, adequate water management for irrigation and other inputs for sugarcane crop.

The proposed ARIMA model for forecasting is recommended for researchers and business men for information and planning their resources as well as farmers decisions regarding the production of Sugarcane crop in Pakistan.

REFERENCES

Akin, M., and S. P. Eyduran (2017). Forecasting harvest area and production of strawberry using time series analyses. Journal of Agricultural Faculty of Gaziosmanpasa University. 34 (3): 18-26.

Allen, P. G. (1994). Economic forecasting in agriculture. Intl. J. Forecasting. 10(1): 81-135.

Azam, M., and M. Khan (2010). Significance of the sugarcane crops with special and reference to NWFP. Sarhad J. Agric. 26: 289-295.

Box, G. E. P., and G. M. Jenkins (1976). Time series analysis: Forecasting and Control. Revised Ed. Holden Day.

Dickey, A., and W. A. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit-root. J. American Statistical Association. 74: 427-431.

FAO. (2010). Crop production. Food and Agriculture Organization of the United Nations. Retrieved. 17.

GOP. (Various issues). Pakistan Economic Survey. Finance Division, Economic Adviser's Wing. Islamabad.

GOP. (2012). Pakistan Economic Survey. Finance Division, Economic Advisor's Wing. Islamabad.

Gujarati, D.N. (2003). Basic Econometrics. 4th ed. McGraw-Hill Book Company Inc. New York. 465-7.

Muhammad, F., M. S. Javed, and M. Bashir (1992). Forecasting sugarcane production in Pakistan using ARIMA Models. Pakistan J. Agri. Sci., 9(1): 31-36.

PBS. 50 years of Pakistan: volume-iii (1947-1997). Finance division, Economic Advisors Wing, Islamabad.

Sajid, A., N. Badar, and H. Fatima (2015). Forecasting production and yield of sugarcane and cotton crops of Pakistan for 2013-2030. Sarhad J. Agric. 31(1): 1-10.

Shah, S. A. A., N. Zeb, and Alamgir (2017). Forecasting major food crops production in Khyber Pakhtunkhwa, Pakistan. J. App. Ad. Res. 2(1): 21–30.

Suresh, K. K., and S. R. K. Priya (2011). Forecasting sugarcane yield of Tamilnadu using ARIMA model. Sugar Technology. 13 (1): 23-26.

USDA(2019). United States Department of Agriculture Retrieved fromhttps://usdasearch.usda.gov/search?utf8=%E2%9C%93&affiliate=usda&query=list+of+sugar+exporting+countries&commit=Search

Yaseen, M., M. Zakriya, I. D. Shahzad, M. I. Khan, and M. A. Javed (2005). Modeling and forecasting the sugarcane yield of Pakistan. Int. J. Agric. Biol. 7: 180-183.