FORECASTING THE PRODUCTION OF SUGARCANE CROP Of 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 Raheem Yar Khan Campus Islamia University, Bahawalpur. [3]Department of Statistics, Government Post Graduate Girls College Bahawalnagar.
Corresponding Author E-mail: qaisarm11@gmail.com

ABSTRACT

Sugarcane is a major cash crop of the Pakistan. This study was present to forecast the production of sugarcane crop in Pakistan. Historical data for the production sugarcane from Pakistan Bureau of Statistics (PBS) and various issues of Economic Survey of Pakistan for the year 1947-2017, were used to predict the production of sugarcane for the years 2018-2030, using Box-Jenkin's (1976) methodology. A suitable ARIMA (2, 1, 1) model was proposed to forecast the production of Sugarcane crop from the year 2019-2030, which show a significant increase, from 75394 tons to 86792 tons. These forecast values are useful for Government, Sugar mills, researchers and business men for information and planning their resources as well as farmers decisions regarding the production of Sugarcane crop in Pakistan.

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

INTRODUCTION

Sugarcane is an important cash crop of the Pakistan. It is mostly grown for the production of sugar and gur. Sugarcane crop is an important source of income and employment for the farmers of the country. It also used as raw material for industries like sugar, chip board, paper, chemicals, plastics, fiber , paints, synthetics, detergents and insecticides etc (GOP, 2012). The production of Sugarcane in our country has increased much for the last thirty years. Pakistan ranks an important position in sugarcane producing countries in the world. Pakistan has the fourth position of sugarcane producing countries followed by Brazil India, China, Pakistan, Thailand and Mexico (FAO. 2010). 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. Several researches have been done to increase the production of Sugarcane which includes chemical composition as well as agricultural reforms to increase the recovery and more %age yield.

In South Asia, Pakistan is the largest consumers 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). Being an agricultural country, presently Pakistan is not self sufficient in sugarcane as well as sugar production, even for domestic requirement in the previous years. The production of sugarcane for the years 2016-17 and 2017-18 was 75482 ton and 81102 ton, while in the year 2015-16 sugarcane production was 65482 ton, this increased amount of production was not fully utilizes for sugar production but it create problems for farmers like lowest payments and incomplete purchasing of their crop by mills owners due to mismanagement and stay order for some mills by court in the country. Mostly the farmers prefer to make gur, shakr and utilize their crop for their animal feed instead of selling to the sugar mills. However, Government of Pakistan some time take initiatives to promote sugarcane production through market support prices. 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 purpose an appropriate forecasting model used 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 sugarcane in Pakistan by using an appropriate measure as an ARIMA model was helpful and appropriate for policy making (Muhammad et al., 1992). Forecasting of agricultural production and prices were proposed to be useful for the farmers, governments, and agribusiness industries (Allen, 1994). The study of wheat yield forecasting a case study in Canada, they were used different forecasting techniques but found that quadratic smoothing technique was best (Boken, 2000). Forecasting the yield of sugarcane in Pakistan, they were used the ARIMA model and predict that the forecasting values are very close to the actual

values (Yaseen et al., 2005). Significance of sugarcane in the province of KPK Pakistan, was discussed and conclude that the results shows that the sum of the elasticities was higher than the unity, it means that the agricultural sector was producing in the stage of increasing return to scale, which means that the allocation of inputs in this sector is not optimal (Azam and Khan, 2010). In the conducting a study of pre harvest of sugarcane yield forecasting by using climatic variables in India, they were developing a forecast model by using weather variable as a regressor and conclude that forecast model able to explain 87% variation in the sugarcane yield before two months harvest (Krishna and Priya, 2010). Forecasting the maize area and production in Pakistan, they use trend methods. They found best forecasting method was quadratic trend model and conclude that the predicted values are very close to actual values and have a positive increasing trend (Ayesha and Nusrat, 2013). Climate and other environmental changes in the developing world and the African continent has become a major threat to their agricultural economy, therefore; the results of this study indicate that the ARMA model is preferable over other time series models considered (Askar and James, 2014). Forecasting the yield of major crops sugarcane and cotton in Pakistan, ARIMA model for forecasting was used and found that ARMA model was suitable for forecasting the sugarcane yield and ARIMA model for cotton yield after using some diagnostic test on fitted model that was verified (Sajid et al., 2015). Forecasting harvest area and production of strawberry in turkey, non stationary data was used after making stationary at first difference for the methodology of ARIMA and three exponential smoothing (Holt, Brown and Damped), holt's exponential smoothing and Brown smoothing model were most appropriate and the results help the policy makers (Akin and Eyduran, 2017). Conducting a study forecasting the major food crops production in Khyber Pakhtunkhwa, the secondary data was used by applying ARIMA forecasting technique they found that the results of the ARIMA model were 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) was used in this study. The data set covers the history of Pakistan and it ranges from 1947-2018. 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.

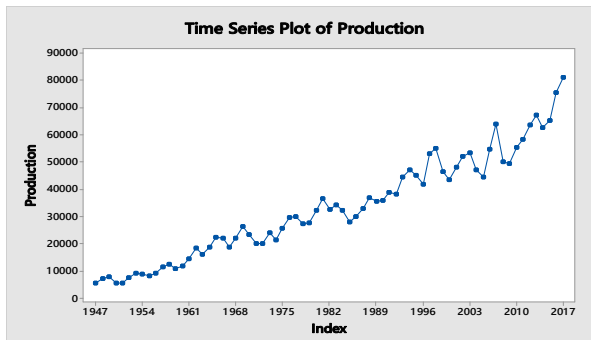

Fig.1. Time Plot of the Sugarcane Production.

Time plot graph for the production of the Sugarcane crop shows upward trend, which indicate 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 was constructed.

The stationary of the data was 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) was suggested to make the Stationary of the above series.

From the graph of autocorrelation function there was only one negative spike, we use first order of moving average (MA) model, and there was one positive spike in partial autocorrelation function graph we consider first order Autoregressive (AR) model. Tentative autoregressive integrated moving average (ARIMA) model was 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.

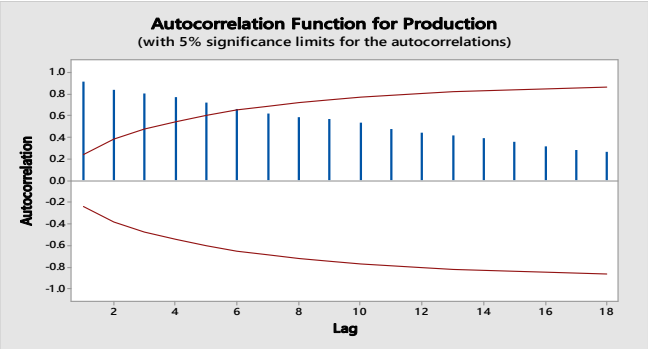

Fig.2. ACF plot for Sugarcane Production

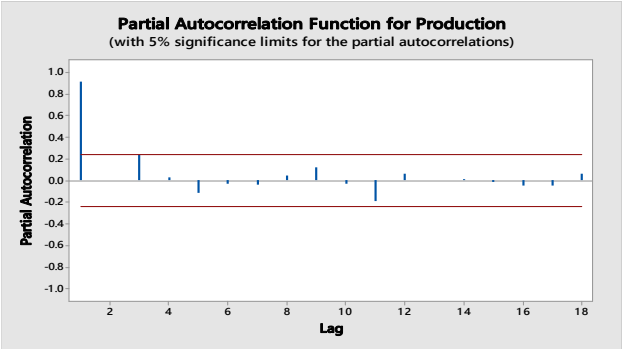

Fig.3. PACF Plot of the Production of Sugarcane

Table 1. Augmented Dickey Fuller Unit Root Test.

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

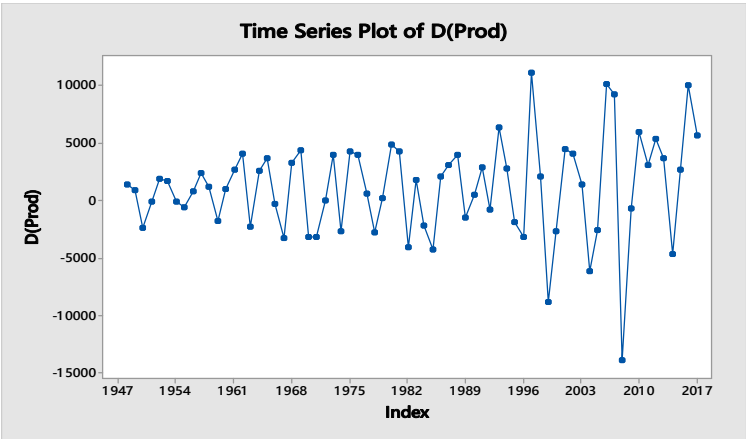

Fig.4 Time Graph of the First difference of the Production

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

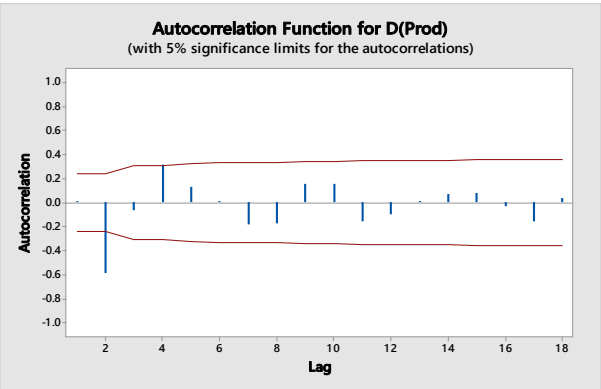

Fig.5. ACF Plot of the First difference Production

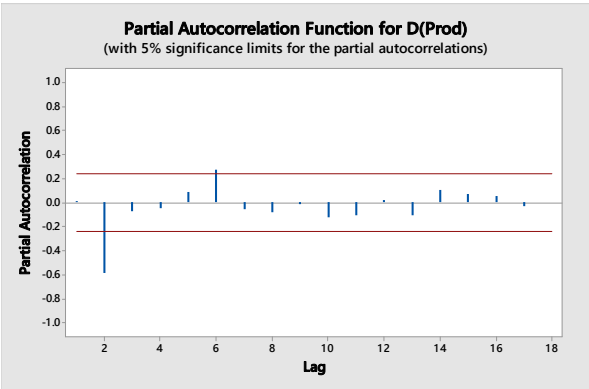

Fig.6. PACF Plot of the First difference Production

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

RESULTS AND DISCUSSION

Table 3 shows all the estimates of the parameters of ARIMA (2,1,1) model were significant which shows that this model was suitable for forecasting the sugarcane production. Graphically diagnostic checks for ARIMA (2,1,1) Model in the figure 7 and 8 were 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 was correctly specified. Normal probability plot for residuals in Fig. 9 show that the residuals lie along a straight line and the residuals were symmetric showing in the figure 10 for histogram of residuals tells that the model was correctly specified.

After fitting the adequate model it is utilizes 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

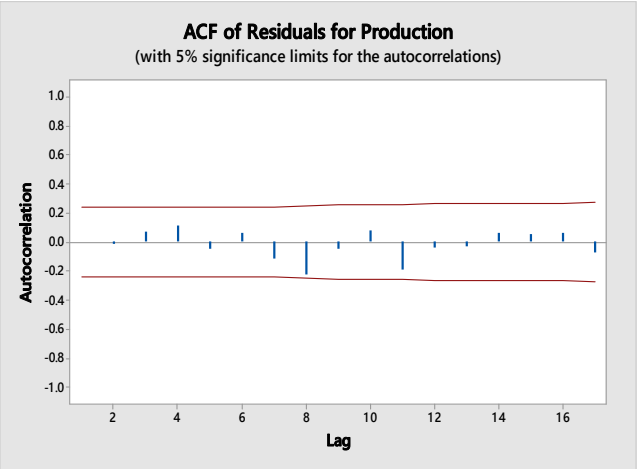

Fig.7. Plot of ACF of Residual for Production

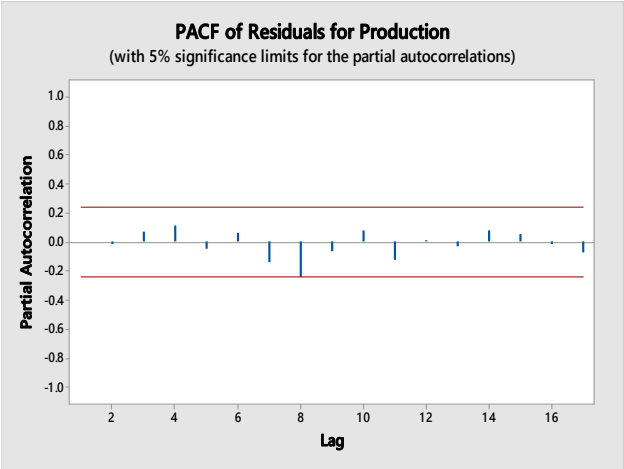

Fig.8. Plot of PACF of Residual for Production

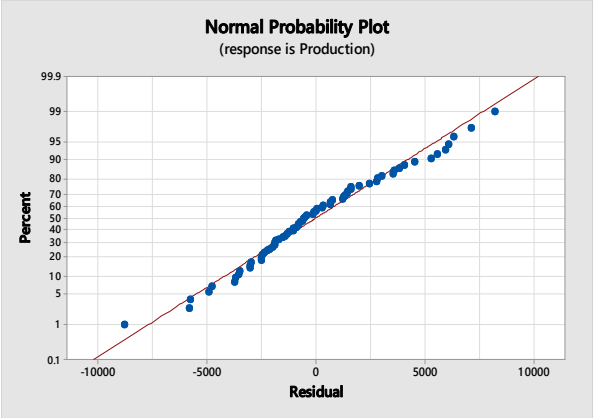

Fig.9. Residuals Normal Probability Plot

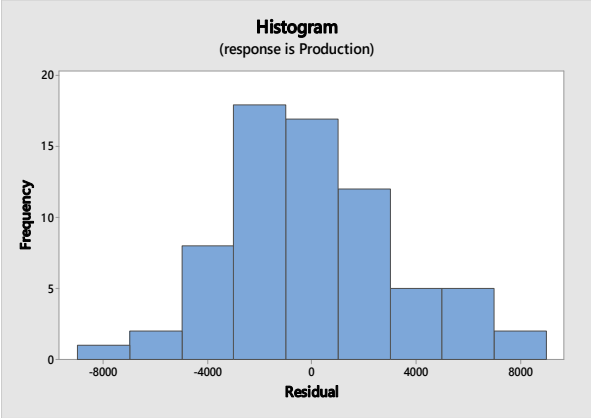

Fig.10. Histogram (Response is Production

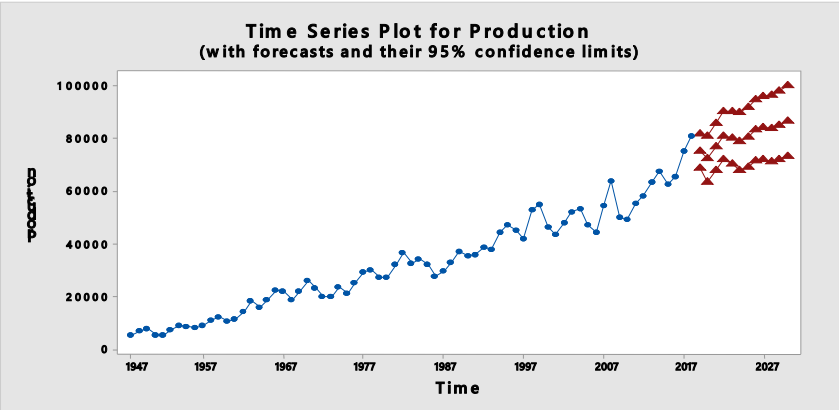

Fig.11. Time Plot of the Forecast Sugarcane Production for the year 2018-2030

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 was best forecasting model for the production of Sugarcane crop for the year 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 ton to 86792 ton for the year 2018-19 to 2029-30. Actual production for the year 2017-2018 of sugarcane was 81102 ton highest in the historical data which will be decreased to 72379 ton as forecast production for the year 2019-2020, then rapidly increasing for the upcoming years. The prediction of this study may provide help policy makers in their macro-level policies for food security, reasonable support price and for better planning for the production of sugarcane in Pakistan. It also provide help micro level policies to sugar

mills 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. International Journal of Forecasting. 10(1): 81-135.

Askar, C., and J. James (2014). Crop yield prediction using time series model. Journal of Economic and Economic Education Research. 15: 53-67.

Ayesha, T., and H. Nusrat (2013). Forecasting of maize area and production in Pakistan. ESci J. Crop Prod. 2 (2): 44-48.

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

Boken, V. K. (2000). Forecasting spring wheat yield using time series analysis. Agronomy Journal. 92(6): 1047-1053.

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. Journal of the 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 Journal of Agricultural Sciences. 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.

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.