Growth, Variability and Forecasting of Wheat and Sugarcane Production in Khyber Pakhtunkhwa, Pakistan

Syed Asghar Ali Shah¹, Sajad Ali² and Arshad Ullah Jadoon*³

¹Agricultural Research Institute Tarnab, Pakistan
²Govt College of Management Sciences, Higher Education Department, Pakistan
³School of Economics Northeast Normal University, China

Submission: September 22, 2018, Published: October 15, 2018

*Corresponding author: Arshad Ullah Jadoon, PhD Scholar, School of Economics Northeast Normal University, International students dormitory Room.509, Changchun, China, Tel: 008613009140277; Email: arshadjadoon22@gmail.com

Abstract

The present study was undertaken to investigate forecasting of major food crops production KPK Pakistan. Time series data were used from 1984-85 to 2015-16, for empirical analysis. The data were analyzed by using statistical packages Gretl.1.9.4 and SPSS. Whereas, the growth models have been employed to fit the best growth model and Cuddy Della Vella Index was applied to find variability in major food crops production [1]. Moreover, auto regressive integrated moving average (ARIMA) model has been employed to fit the best time series model. Based on the results of analyzed data, it was found that in for both wheat and sugarcane crops Production, the cubic growth model was found suitable based on the R² criteria and fitted trend line. After selecting best fitted model for each crop, the growth rate was calculated by using the selected fitted models which were found to be 10.97% and 45.31% respectively. Moreover, the variability for each food crops production was found to be 1.53% and 0.44% respectively. Also, Different ARIMA models were examined for Wheat and Sugarcane. Among the studied models the ARIMA (0, 2, 1) for both food crops were found to be the best on the basis accuracy measures criteria i.e. Akaike Information criteria (AIC) and Schwartz Information criteria (SIC). From the results of analyzed data, time series models were found adequate for forecasting major food crops production in Khyber Pakhtunkhwa.

Keywords: Diagnostics measures; Auto regressive integrated; Accuracy measures; Moving average; Parameter estimates; ADF test; Cuddy della vella index

Abbreviations: ARIMA: Auto Regressive Integrated Moving Average; AIC: Akaike Information Criteria; SIC: Schwartz Information Criteria; ACF: Auto Correlation Function; PACF: Partial Auto Correlation Function

Introduction

Agriculture sector plays significant role in the economy of each and every country. Agriculture occupies central position in Pakistan and contributes nearly 20.9% to the GDP and about 43.5% of the labor force is engaged in agriculture (Pakistan Economic Survey, 2014-2015). Agriculture sector contributes significantly in the economy of Pakistan and also considered as the largest sector as well as the hub of economic activities. With the increase in the growth rate of population, it was found the growth rate of agriculture had slowed down i.e. over 3% in 1980’s, while, during the year 2012-13, it was reported to be 2% which was still comparatively high. Moreover, it was reported that in the scenario of current rate of increase in the population growth, it can be expected that Pakistan will attain fifth position from current running status of sixth most populous country in the world ranking by the year 2050 [2].

The main food crops of Khyber Pakhtunkhwa province are sugarcane, wheat, rice and maize. Diverse varieties of these food grains are developed in different regions of the province as compared to barley, bajra and jowar which are not mature comprehensively. The most important wheat varieties developed in Khyber Pakhtunkhwa are Fakhre- Sarhad, Auqab-200, Bakhtawar-92, Khyber-87, Haidar-2002, Saleem-200 and Tatara. The major sugarcane varieties grown in Khyber Pakhtunkhwa are CP77-440, CP51-52, CP65-357, SPF 238, HSF 240 ans SPF 242. Food-grain crops generally wheat, rice, maize and sugarcane are dissimilar in stipulations of price and production on the identical dimension of land. According to Hazell [3], who observed that due to their vast scope and adoption of modern prevailing technology, the variability in production of world cereal as well as Indian food grain were also been increased.

Mehmood [4], used uni-variate model to forecast the exports of Pakistan to SAARC and argued that the ARIMA model was found suitable for forecasting time series data. Similarly, ARIMA model technique has also been used widely in literature for forecasting...
area and production. Furthermore, efforts have also been made to estimate production and productivity of sugarcane using ARIMA models [5,6]. Also, regarding ARIMA models including forecasting of sugarcane production in Pakistan [7], forecasting of area, production and productivity of different crops for Tamil Nadu State [8], forecasting of wheat production in Canada and Pakistan [9,10], forecasting agronomic production at state level [11], estimating sugarcane area and yield for Pakistan [12], estimating and modeling of wheat yield in Pakistan [13], sugarcane yield estimation for Tamil Nadu, and predicting productivity in India [14].

Materials and Methods

We used annual time series data which are from 1984 to 2014 i.e. time series data of 30 years, for to forecast the increase in production for onward ten years which include major food crop including wheat, maize, sugarcane and rice by considering all related factors as constant in Khyber Pakhtunkhwa, Pakistan. The data were collected from different secondary sources such that various issues of Crop Statistics, Crop Reporting Service of Khyber Pakhtunkhwa and were analyzed by using Statistical Packages SPSS version 21 and Gred 1.9.4.

The study has been carried out by using different growth models including Linear growth model, Quadratic growth model, Cubic growth model, Power growth model, S-curve model, Logistic growth model, Exponential growth model, Compound growth model, Growth curve, Inverse curve, Logistic Curve to find the best fitted model for production of major food crops of Khyber Pakhtunkhwa. By selecting the best fitted growth model based on the criteria of R² and fitted trend line to forecast the food crops production for onward ten years.

Autoregressive integrated moving average (ARIMA)

ARIMA model is a generalized form of Auto regressive moving average (ARMA) model introduced by Box & Jenkins [15], 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 as the order of the autoregressive integrated moving average process. It is an important part of Box & Jenkins [15] approach to time series modeling. It can be written as:

\[ \Delta^d Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \ldots + \alpha_p Y_{t-p} + \epsilon_t + \beta_1 \epsilon_{t-1} + \beta_2 \epsilon_{t-2} + \ldots + \beta_q \epsilon_{t-q} \]  

Where, \( \Delta^d \) represents differencing of order d i.e. \( \Delta^d Y_t = Y_t - Y_{t-d} \), \( \Delta^d Y_t = \Delta^d Y_{t-1} \), and so forth, \( Y_t, \ldots Y_{t-p} \) shows lags of the variables. Also, \( \alpha_0, \alpha_1, \ldots, \alpha_p \) and \( \epsilon_t \) are the parameter estimates and error term of the model.

Generally, ARIMA model technique has extensively been employed in literature to forecast the specific area as well as production related to different major crops [16,17].

Stationary test

The first step in Box- Jenkins methodology is to find whether data follows stationarity or not. There are a number of tests which can be used to decide about the stationarity of the variables. Augmented – Dickey Fuller [18], abbreviated as ADF, is the more popular test in literature due to its simplicity and powerfulness.

Mathematically

\[ \Delta^d Y_t = a_0 + \lambda t + \sum_{i=1}^{p} \beta_i Y_{t-i} + \epsilon_t \]  

There are three options in this equation:

a. \( a_0 \) is the intercept or drift parameter of the time series.

b. \( t \) is the time trend in time series. There may be downward or upward linear trend in the data.

c. It is also possible that both drift and time trend exist in the data.

Diagnostic measures for selection of best forecasting model

There are few diagnostic checks which each estimated model has to fulfill and are as follows:

a. Residuals are normally distributed (Jarque- Bera test).

b. Residuals of the projected model are not serially correlated (Q-Statistics).

c. Model selection criteria

Generally, the model selection criteria statistics are used to compare the fits of different forecasting and smoothing method and also contributes a great deal of information by comparing the fits obtained through different methods [19]. For selecting best suitable models the Accuracy measures including Akaike information criteria (AIC) and Schwartz information criteria (SIC) were employed. Smaller values of these accuracy measures indicate a good fitted model with minimum forecasting error.

Result and Discussion

Growth model for wheat and sugarcane crops production in KP

| Equation      | Model Summary | Parameter Estimates |
|---------------|---------------|---------------------|
|               | R Square      | F  | df1 | df2 | Sig. | Constant | b1  | b2  | b3  |
| Linear        | 0.192         | 6.67 | 1   | 28  | 0.015 | 989.972  | 6.876 | 0   | 0   |
| Logarithmic   | 0.251         | 9.369 | 1   | 28  | 0.005 | 894.331  | 81.259 | 0   | 0   |
| Inverse       | 0.235         | 8.592 | 1   | 28  | 0.007 | 1142.697 | -346.515 | 0   | 0   |

How to cite this article: Syed AA S, Sajad A, Arshad U J. Growth, Variability and Forecasting of Wheat and Sugarcane Production in Khyber Pakhtunkhwa, Pakistan. Agri Res& Tech. Open Access J. 2018; 18(2): 556056. DOI: 10.19080/ARTOAJ.2018.18.556056.
The growth analysis technique is employed to check best fitted trend model for wheat and sugarcane production in Khyber Pakhtunkhwa province of Pakistan. Among the growth models including linear growth model, Quadratic growth model, cubic growth model, compound growth model, exponential growth model, power growth model, S-Curve model, inverse trend model, growth model, logarithmic growth model, logistic growth model etc., the best fitted model on the basis of $R^2$ criteria and fitted trend line the cubic growth curve and exponential growth curve were found to be best fitted curve for each crop i.e. wheat and sugarcane crops, presented in Table 1 & 2 along with their respective Figure 1 & 2 respectively [20].

The parameter estimates of all selected models presented in Table 1 is found to be significant, except Quadratic growth model having $p \leq 0.05$. Also, cubic trend model is found to be the best suitable one among the studied trend models, on the basis of fitted trend curve and accuracy measures.

Figure 1 shows that the fitted line provides a good fit as it passes through majority of the data points. Moreover, based on the
parameter estimates of best fitted growth model, the growth rate for wheat production is calculated to be 10.97% having significant contribution in wheat production over last 30-years. The growth analysis of sugarcane production is made after selecting best fitted model among the studied models; it has been found that exponential growth model is the best fitted model on the basis of $R^2$ criteria and fitted trend line. Trend curve and also its parameter estimates are presented in Table 2 and Figure 2.

From Table 2, it is showed that parameter estimates of all the selected models are found to be significant at 5% level of significance. Also, the cubic growth model is calculated to be best suitable model among the studied trend models, on the basis of accuracy measures and best fitted trend line.

Figure 2 shows that the fitted trend line provides good fit as compared to other selected models and it passes through majority of the data points. Furthermore, growth of sugarcane crop production is found to be 45.31% which shows higher growth rate as compared to other major food crops under study and it is encouraging factor for boosting economy of Pakistan.

Variability of Major Food Crops Production in Khyber Pakhtunkhwa

The variability of major food crops production has pivotal role for sustainable production as well as for the economy of the country. Therefore, an effort has been made to estimate the relative variability in major fruits production presented in Table 3. The variability in production of wheat and sugarcane crops of Khyber Pakhtunkhwa is estimated with effect from period 1984-85 to 2013-14. It is observed that wheat and sugarcane crops production possess variability (1.53%) whereas, the sugarcane crop production having the variability index of (0.44%).

| Food Crop  | CV (%) | $AdR^2$ | Variability Index ($I = CV = \sqrt{1 - AdR^2}$) |
|-----------|--------|---------|---------------------------------------------|
| Wheat     | 12.586 | 0.985383| 1.53                                        |
| Sugarcane | 9.2278 | 0.997722| 0.44                                        |

ARIMA modeling for wheat crop production
At first, it is very essential to find out the stationarity in wheat crop production. It is important to consider that which order difference of time series sequence of wheat crop production satisfies the stationarity conditions. The Dickey fuller unit root test is performed and plot of time series of wheat crop production and is shown in Figure 3 & 4 respectively. From the results of Dickey fuller test, it showed that the original series is non-stationary due to unit root in the data (Table 4).

By taking order first difference, it was found that stationarity condition is satisfied with the p-value =0.000044, which strongly suggested that there is no unit root (Figure 5). Moreover, it is clear from Auto correlation function (ACF) & Partial auto correlation function (PACF) of the differenced series that a tentative selected ARIMA model among the subset models, to forecast the wheat crop production in Khyber Pakhtunkhwa is ARIMA (0, 2, 1).

Based on the Accuracy measures, it is obvious that ARIMA (0, 2, 1) is an adequate model for wheat crop production at 5% level of significance. Table 5 shown below is the model parameter estimates along with their significance. In literature similar studies had been carried out by Yaseen et al. [3], Karim R, Awala A, Akhter M [18] & Bajpai and Venugopaln [4], had used the same criteria for forecasting wheat production.

### Table 4: Dickey Fuller Test for Wheat crop Production.

| Production Series | Null Hypothesis | P-Value   | Remarks     |
|-------------------|-----------------|-----------|-------------|
| Original Series   | a=1             | 0.1512    | Non-Stationary |
| 1st Order Difference | a=1            | 0.00004374 | Stationary   |

### Table 5: ARIMA (0, 2, 1) Model using observations 1986-2013 (t = 28).

| Dependent Variable: Second Order Difference for Wheat Crop Production | Coefficient | Std. error | Z     | P-value |
|-----------------------------------------------------------------------|-------------|------------|-------|---------|
| const                                                                 | 0.311069    | 2.62655    | 0.1184| 0.9057  |
| q1                                                                    | -1          | 0.105667   | -0.464| 2.97e-021 *** |

Akaike criterion | 355.0832 | Schwarz criterion | 359.0798

---

00114 How to cite this article: Syed A A S, Sajad A, Arshad U J. Growth, Variability and Forecasting of Wheat and Sugarcane Production in Khyber Pakhtunkhwa, Pakistan. Agri Res& Tech: Open Access J. 2018; 18(2): 556056. DOI: 10.19080/ARTOAJ.2018.18.556056.
Model diagnostics
To check the auto correlation assumption, the "Ljung-Box" test is used. Test statistic gives $Q' = 15.55$, with $p$-value $= 0.556$, which suggested that we may accept the assumption that there is no autocorrelation among the residuals of the fitted ARIMA model at 5% level of significance. To check the normality assumption, the Jarque Bera test is used. The test statistic value $= 0.690693$, with $p$-value $0.707975$, suggesting normality assumption of the residuals is valid. Graphical residuals diagnostics are shown in Figure 6 & 7. The Q-Q plots and histogram of residuals also show normality. So, it can be concluded from the graphical and formal tests that the selected model ARIMA $(0, 2, 1)$ is an adequate model to forecast wheat crop production in Khyber Pakhtunkhwa.

![Figure 6: Q-Q plot of residual.](image)

![Figure 7: Histogram of residual.](image)

Forecast for Wheat Crop Production

| Year | Predicted Production | Std. Error | 95% Interval |
|------|----------------------|------------|--------------|
| 2014 | 1381.7               | 116.11     | 1154.1 - 1609.3 |
| 2015 | 1400.6               | 164.21     | 1078.7 - 1722.4 |
| 2016 | 1419.8               | 201.11     | 1025.6 - 1814.0 |
| 2017 | 1439.3               | 232.22     | 984.2 - 1894.5 |
| 2018 | 1459.2               | 259.64     | 950.3 - 1968.0 |
| 2019 | 1479.3               | 284.42     | 921.9 - 2036.7 |
The selected model is used for forecasting the wheat crop production. In Table 3, the predicted values, standard errors, lower and upper confidence limits for ten years onward values are given, for Khyber Pakhtunkhwa based on the sample data Table 6. By comparing the original and forecasted series it is obvious that the original series of wheat crop production shows increasing tendency with the passage of time and following the same pattern in the next stages and finally showing increasing pattern. Similarly, the forecasted series shows the same pattern. In the forecast plot the in sample and out sample forecasting part is shown in Figure 8.

**ARIMA modeling for sugarcane crop production**

The time series plot for sugarcane production is presented in Figure 9 & 10, indicates trend in data and does not show constant mean and variance over time i.e. up-ward trend initially and then down-ward trend in the end which indicates non-stationarity. The first order difference of the data makes the series stationary.
The results of Dickey fuller unit root test showed stationarity at first order difference as its p-value = 0.013288, which suggest that there is no unit root at first order difference and have constant variance as shown in Table 7.

The tentative model based on ACF and PACF plot are given in Figure 11 is ARIMA (0, 2, 1). Also, various subset ARIMA models were fitted among which the ARIMA (0, 2, and 1) having lowest accuracy measures i.e. AIC and SIC’s and was suitable as compared to other subset models Table 8. Similar studies had been carried out by Ullah et al. [13], for estimating sugarcane in Pakistan and for sugarcane productivity in India.
of residuals is valid. The Q-Q plot and histogram in Figure 12 & 13 show picture of approximate normality for residuals. Thus, our fitted model ARIMA (0, 2, 1) is an adequate model for forecasting Sugarcane production in Khyber Pakhtunkhwa.

Figure 12: Q-Q plot of residual.

Figure 13: Histogram of residual.

Forecast for Sugarcane Production

Figure 14: Forecasting plot for sugarcane production.
Table 9: Forecast for Sugarcane Production with effect from 2014-2023.

| Year | Predicted | Std. Error | Confidence Interval |
|------|-----------|------------|---------------------|
| 2014 | 5402.3    | 272.48     | 4868.3 - 5936.4     |
| 2015 | 5443.5    | 385.34     | 4688.3 - 6198.6     |
| 2016 | 5484.9    | 471.94     | 4559.9 - 6409.9     |
| 2017 | 5526.5    | 544.95     | 4458.4 - 6594.6     |
| 2018 | 5568.4    | 609.27     | 4374.2 - 6762.5     |
| 2019 | 5610.4    | 667.43     | 4302.3 - 6918.6     |
| 2020 | 5652.7    | 720.9      | 4239.8 - 7065.7     |
| 2021 | 5695.3    | 770.68     | 4184.8 - 7205.8     |
| 2022 | 5738      | 817.43     | 4135.9 - 7340.2     |
| 2023 | 5781      | 861.64     | 4092.2 - 7469.8     |

The selected model is used to forecast the sugarcane production for onward period of ten years i.e. from 2014 to 2023. The predicted values, standard errors, lower and upper confidence limits are presented in Table 9. Moreover, it is possible to compare the original and predictions for sugarcane production are graphically given in Figure 14. By comparing the original and forecasted sugarcane production, it is quite dear that sugarcane production shows upward production tendency and the forecasted series also shows the same upward production tendency. The in-sample and out-sample forecasting shows same pattern for forecasting sugarcane production in Khyber Pakhtunkhwa [21,22].

Conclusion and Recommendations

From the analyzed data of each selected food crop i.e. wheat and sugarcane, the best model for each was calculated on the basis of best fitted trend line and accuracy measures. The best fitted model for each food crops were cubic growth model respectively. Based on the parameter estimate of best selected model for both food crops, the growth rates were calculated and found to be 10.97% and 45.31% respectively. Moreover, the variability measure for wheat crop production was recorded higher i.e. 1.53%, as compared to sugarcane which possess least variability i.e. 0.44%. It may be due to fluctuating crop acreage, area and productivity. From the time series modeling of both wheat and sugarcane crops production it can be concluded that for each food crop i.e. wheat and sugarcane, the forecasting model ARIMA (0, 2, 1) were found adequate for forecasting purpose based on forecast evaluation criteria. Also, it can be recommended that these selected models could be used by researchers, business men, policy makers and food crop producers for information, planning their resources as well as decision making regarding food crops production in Khyber Pakhtunkhwa. Also, at the same time Box-Jenkins ARIMA model give good representation of short time forecasting.

References

1. Cuddy JDA, Della VPA (1978) Measuring of Instability of Time Series Data. Oxford Bulletin of Economics and Statistics 40(1): 79-85.
2. (2014-2015) Pakistan Economic Survey. Economic Advisory Wing, Ministry of Finance, Islamabad. p. 18.
3. Hazell PBR (1982) Instability in Indian Food grain Production. International Food Policy Research Institute. Washington, USA. p. 60.
4. Mehmed S, Carter D (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.
5. Yaseen M, Rashid M (2014) Effectiveness of calcium carbide formulations in emergence stimulation and growth of wheat (Triticum aestivum L.) Seedlings.
6. Bajpai PK, Venugopalan R (1996) Forecasting sugarcane production by time series modeling. Indian Journal of Sugarcane Technology 11(1): 61–65.
7. Muhammad S, Hussain I, Khan IA, Rab A, Jan I, et al. (2013) Influence of Organic Mulches on Growth and Yield Components of Pea’s Cultivars. Greener Journal of Agricultural Sciences 3(8): 652-657.
8. Balanagamml D, Ranganathan CR, Sundaresan K (2000) Forecasting of agricultural scenario in Tamilnadu: A time series analysis. jou Ind Soc Ag Statistcs 53(3): 273-286.
9. Boken VK (2000) Forecasting spring wheat yield using time series analysis: A case study for the Canadian prairies. Agricultural Journal 92(6): 1047-1053.
10. Saeed M, Guo M, Ulhab I, Tabbasam N, Zafar Y, et al. (2000) QTL mapping for physiology yield and plant architecture traits in cotton (Gossypium hirsutum L.) grown under well-watered versus water-stress conditions. Electronic Journal of Biotechnology 1(43).
11. Indira R, Datta A (2003) Univariate forecasting of state-level agricultural production. Economic and Political Weekly 38(18): 1800-1803.
12. Masood MA, Javed MA (2004) Dermatoprotective effects of some plant extracts (genus Ficus) against experimentally induced toxicological insults in rabbits. Pak J Agri Sci 41(1-2): 1-6.
13. Ulhab AK, Gilani AH (2010) Antiseccretory and analgesic activities of Terminalia bellirica. African Journal of Biotechnology 9(18): 2717-2719.
14. Padhan (2012) Congestion Management in Deregulated Power System Based on Optimal Rescheduling of Generators, Proceedings of National Conference on Recent Advances in Modern Power Systems (RAMPS-2012), pp. 197-203.
15. Box GEP, Jenkins GM (1976) Time Series Analysis: Forecasting and Control. Holden-Day San Francisco. USA, pp. 575.
16. Munir M, Siddique M, Ali Q (2008) Comparative efficacy of standard AGID and precipitinogen inhibition test with monoclonal antibodies based competitive ELISA for the serology of Peste des Petits Ruminants in sheep and goats. Trop Anim Health Prod 41(3): 413-420.
17. Gujarati DN (2003) Basic Econometrics, (4th edn.). McGraw-Hill Companies Inc., New York, USA, pp. 465-467.
18. Dickey DA, Fuller WA (1981) Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. Econometrica 49(4): 1057-1072.
19. Karim R, Awala A, Akhter M (2010) Forecasting of wheat production in Bangladesh. Bangladesh J Agric Res 35(1): 17-28.
20. Chandran KP, Pranjeshu (2005) Non-parametric regression with jump point’s methodology for describing country’ soil seed yield data. Journal of Indian Society of Agricultural Statistics 59(2): 126-130.
21. Government of Khyber Pakhtunkhwa (2013) Crops Statistics of Khyber Pakhtunkhwa, Crops Reporting Services, Agriculture, Livestock Cooperative Department, Peshawar, Pakistan, pp. 121.
22. Government of Khyber Pakhtunkhwa (2009) Crops Statistics of Khyber Pakhtunkhwa, Crops Reporting Services, Agriculture, Livestock Cooperative Department, Peshawar, Pakistan.
