Statistical Modeling and Forecasting of Gram Pulse Area and Production of Pakistan

Muhammad Wasim Amir*, Zeeshan Raza, Muhammad Amin

University of Sargodha, Sargodha, Pakistan

Abstract

The main aim of this study is to forecast the area and production of gram pulse in Pakistan using best selected time series model based on time series data i.e. 1947-1948 to 2016-2017. The area and production of gram were forecasted over 2017-2018 to 2029-2030. A variety of time series models were applied to forecasts the gram area and production of Pakistan. The Akaike Information Criteria (AIC) and some others showed that the ARIMA (1, 1, 1) and ARIMA (2, 1, 2) are the more adequate models for gram area and production, respectively. Some diagnostic tests of the selected time series model are also considered to evaluate the quality of the selected model. Moreover, the quality, of the selected models is measured based on the minimum value of the mean error, root mean square error, mean percentage error, and mean percentage absolute error. From ARIMA (1, 1, 1) and ARIMA (2, 1, 2) model, we noticed that gram area and production of Pakistan are forecasted to be 2515.08 thousand acers and 518.325 thousand tons in 2030 under the condition that there is no unusual event occurred.

Keywords: Gram Production, Gram Area, ARIMA, Time Series Models, Model Selection Criteria's.

Introduction

The agricultural sector plays an energetic role in the economic growth of Pakistan. Around one-third of Pakistani people are directly or indirectly linked with the agriculture sector. This sector contributes 19.8 percent to the Gross Domestic Product (GDP) which has an obvious influence over any other sector. Moreover, it is the major employing sector as well as it

* Corresponding Author: muhammadwasim113@gmail.com

contributes 42.3 percent of the country's labor force (Government of Pakistan, 2016). This sector provides many vegetables, crops, and fruits such as onion, garlic, mash, been, wheat, sugarcane, sesame, barley, rice, millet, banana, mango, apple, dates, and gram, etc. Some items such as gram, moong, masoor, contain a rich amount of protein which is a necessary part of the human body.

Pulses are commonly known as the poor man's meat. They are relatively inexpensive sources of protein in balancing human nutrition. Gram belongs to the family of legume crops that is rich in proteins and used as a source of vegetable proteins. For a human being, it is a low-cost item to get proteins as compared to animal meat. In Pakistan, pulses are cultivated on 5% of the total crops cultivated area which is minimum for the Pakistan population. The area under pulses production in Pakistan is 1.5 m hector (Pakistan Agriculture Research Council, 2019).

Forecasting of crops area and production is necessary for future policies and planning of a country to fulfill its population necessities. In the literature, researchers applied several statistical models for forecasting several crops area and production. Iqbal et al. (2005) applied the ARIMA model to forecast the production and area of wheat in Pakistan. For the desired purpose, they forecasted the production and are up to the year 2022 and the selected model showed that the wheat production would be 29774.8 thousand tons in 2022. Bilal and Shahbaz (2008) forecasted the Barley production of Punjab, Pakistan using ARIMA models based on 22 years' time series data i.e. 1981-2003. Khan et al. (2008) used the ARIMA and log-linear model for forecasting mango production. Awal and Siddique (2011) examined the growth pattern and found that the best model for rice production forecasting in Bangladesh. In their study, the ARIMA model was considered as the most efficient for rice production forecasting and they concluded that the selected model is useful for the researcher as well as for policymakers. Rani and Raza (2012) compared the double exponential smoothing and linear trend method to estimate the pulses prices in Pakistan. Naz (2012) designed a univariate analysis of exports dates in Pakistan. The purpose of his study was to forecast the export of dates for the next 15 years. For the desired purpose, time series data was taken from 1962 to 2008. Box-Jenkins methodology is considered for forecasting purposes. The results showed that the export of dates will be better in the future. Habib et al. (2013) used different time series models to investigate the future behavior of millet area and production in Pakistan. For this purpose, they used time series data during 1985-2012. The study disclosed that the quadratic model is the most appropriate model for forecasting the millet area and production of Pakistan. Amin et al. (2014) investigated the different time series models for forecasting wheat production in Pakistan. They concluded that the ARIMA (1, 2, 2) is the best time series model for long-term forecasting. Vishwajith et al. (2014) analyzed and forecasted the pulses production of major growing states in India. For the forecasting purpose, they utilized ARIMA and GARCH models. The comparative study disclosed that both models are not uniform for forecasting pulses production in India. Qureshi et al. (2014a) used the ARIM-X type model to forecasting the production of the citrus fruit of Pakistan. Qureshi et al. (2014b) forecasted the mango production of Pakistan by using the ARIMA-X model and also applied the various model diagnostics tools on residuals to ensure the adequacy of the best fitted model. Sahu et al. (2014) forecasted the yield, production, and area of wheat and rice in the South Asian Association for Regional Cooperation (SAARC) region. For the analysis purpose, they used descriptive statistics and Box-Jenkins methodology. Their study showed that wheat and rice production increases in upcoming years. Naheed et al. (2015) compared the various time series models to identify the suitable model for forecasting barely area and production in Punjab, Pakistan. In this study, they concluded that double exponential smoothing is the best model for forecasting the barely area and production. Rahman and Baten (2016) applied the ARIMA model to forecast the production and area of black gram pulse. They used the previous 47 years' time series data for forecasting the production of black gram pulse for the next 5 years. The study revealed that there is a downward trend in production and area of black gram pulse. Karadas et al. (2017) investigated the future trends of oil crops in Turkey and showed that the exponential smoothing models are more adequate for forecasting the oil crops in Turkey. Celik et al. (2017) forecasted the groundnut production of Turkey. The effort of their study is to build the finest model for groundnut production which indicates that the future trend in advance. ARIMA (0, 1, 1) was found to be the best model among all ARIMA models. Wali et al. (2017) designed a univariate analysis to forecast the production and area of cotton in India. For the forecasting purpose, the Box-Jenkins methodology is considered. The findings showed that there is an upward trend in both area and production in future. Masood et al. (2018) applied S-curve, exponential smoothing, double exponential smoothing, quadratic trend, and ARIMA model on wheat production of Pakistan. They found that the ARIMA (2, 1, 2) is an adequate model for forecasting wheat production in Pakistan. Ullah et al. (2018) forecasted the area and production of peach in Pakistan. Their study showed that the ARIMA (1, 1, 0) is the best model for forecasting the peach area and production of Pakistan. Khan et al. (2019) forecasted the guava area and production of Pakistan. For the said purpose the Box-Jenkins methodology is considered and found that the ARIMA (0, 0, 0) and ARIMA (1, 1, 0) models are most appropriate to forecast the guava area and production of Pakistan, respectively.

The literature designated that no research still has been done in Pakistan for forecasting gram area and production. This study fits the finest modeling for forecasting gram area and production. The main interest is to find the current behavior and future trends of the gram area and production of Pakistan.

Materials and Data Analysis Methods

As our interest is to forecast the gram area and production of Pakistan based on a suitable time series model. For this purpose, we consider the time series data of gram area and production from 1947-1948 to 2016-2017. The required time series data were collected from the agriculture marketing information service. For modeling time series data to forecast the output, some authors initially consider the time series plot and correlogram as a descriptive measure to see the behavior of time series data. Some also consider the autocorrelation (ACF) and partial autocorrelation function (PACF) plots of the original time series data as a descriptive tool to select the best time series model. Nowadays these descriptive tools are used jointly to show the time series plot as well as forecasted values on the same graph after model selection and also used as a diagnostic check. Because with the advancement of statistical software, best time series model can be achieved quickly. There are various time series models available for forecasting purposes but this study considered Box Jenkins (1976) methodology. The major use of this methodology is to estimate the parameters of the considered model and then forecasting based on the considered model. Autoregressive Integrated Moving Average (ARIMA) model is a popular Box-Jenkins (1976) time series model which is widely used for forecasting the time series data. The mathematical structure of the ARIMA (p, d, q) model is given as:

$$(1-\sum_{j=1}^{p}\theta_j L_j)(1-L)^d x_t = (1+\sum_{i=1}^{q}\varphi_i L_i)u_t, \qquad (1)$$

Where p, d, q are the order of Autoregressive (AR), Integrated, and Moving average (MA) respectively. The main problem in the ARIMA modeling is to select the best value for p, d, q.

Furthermore, the order of *p* and *q* for the ARIMA model are identified by Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF), respectively. Furthermore, the value of *d* was determined when the series becomes stationary. For more detail of mathematical forms of some time series models with different orders of p, d, q, the readers are referred to read (Chatfield, 1995; Amin et al. 2014; Amir et al. 2021).

Best Model selection criterion

The selection of an appropriate time series model to find out the most accurate forecasted value is a key issue. For selecting a suitable time series model, the AIC and SBC criterion is widely used. The AIC model selection criteria were proposed by Akaike's (1973) and mathematically it is defined as:

$$AIC = -2\log\left(maxmum likelihood\right) + 2k, \qquad (2)$$

Where $k = p + q + 1$(in case of inclusion of intercept in the model) otherwise $k = p + q$. The ARIMA model which displays the minimum value of AIC is preferred over other fitted ARIMA models (Tsay, 2005). Other criteria to find out the best time series model is the SBIC and is defined as:

$$SBIC = -2\log\left(maximum likelihood\right) + 2klog\left(n\right) \qquad (3)$$

It also concerns the minimum value for a model to be a good fit for the data (Tsay, 2005).

Selected model diagnostics

By using the model selection criterion, when the best model is obtained then it is the need of a time series model to full fill its assumptions for the accuracy of forecasting. The time series models require assumptions that the residuals are ought to be normally distributed, independent, and zero-order autocorrelation and no heteroscedasticity (Chatfield, 1995). The Periodogram is also used to test the normality of residuals. Run and white tests are used to identify the presence of autocorrelation and heteroscedasticity in the residuals (Gujarati, 2009; Amir et al., 2021).

Table 1

Forecasting Accuracy Measure Estimators

Accuracy measure tool	Estimator	References
ME	$ME = \frac{\sum_{t=1}^{n} u_t}{n}$	Makridakis et al. 2004
MAE	$MAE = \frac{\sum_{t=1}^{n} \lvert u_t \rvert}{n}$	Makridakis et al. 2004
MPE	$MPE = \frac{1}{n}\sum_{t=1}^{n} PE$	Makridakis et al. 2004
MAPE	$MPE = \frac{1}{n}\sum_{t=1}^{n} \lvert PE_t \rvert$	Makridakis et al. 2004
MSE	$MSE = \frac{\sum_{t=1}^{n} u_t^2}{n}$	Makridakis et al. 2004

Where $u_t = x_t - \hat{x}_t$, $PE_t = \left(\frac{x_t - \hat{x}_t}{x_t}\right) \times 100$ and $\hat{x}_t$ is the forecasted value over time *t*.

Results and Discussion

The current study is based on the forecasting of area and production data of gram in Pakistan. To full fill the desired objective, various time series models are applied to the historical data (1948-2017) on the basis of statistical measures for reliable forecasts. The results of different time series models are reported in Table 3 and Table 5 for the reliability and selection criteria. The values of

Table 2

Descriptive statistics of gram area and production of Pakistan from 1948 to 2017

	Minimum	Mean	Standard Deviation	Kurtosis	Skewness	Maximum
Area	2027.78	2575.30	271.82	-0.21	0.19	3247.00
Production	284.40	545.62	134.28	-0.27	-0.05	868.20

AIC and SBIC indicating that the ARIMA (2, 1, 2) is the best model for forecasting the gram production and ARIMA (1, 1, 1) is the finest model for forecasting the gram cultivated area of gram in Pakistan. This is due to having the least values of AIC and SBIC and used these models for forecasting the area and production of a gram in Pakistan. To check the adequacy of the selected models, some diagnostic measures are considered. Four runs tests are used on the residuals are used for testing the randomness of the selected models and the results are given in Table 3 and Table 5. The runs tests of the selected model indicated that the selected model residuals are random which reveals that these selected models are best fitted and adequate for the data. Moreover, the results of different error measures are also provided in Tables 2 and 4 to show the predictability power of the selected models.

Table 3

Model validity and selection criteria's for Gram area

Model	AIC	HQC	SBIC	RMSE	MAE	MAPE	ME	MPE	RUNS	RUNM	AUTO	MEAN
(i)	11.05	11.05	11.05	250.58	184.69	7.31	3.20	-0.37	OK	OK	OK	OK
(ii)	11.09	11.10	11.12	252.39	184.37	7.31	0.00	-0.49	OK	OK	OK	OK
(ii)	11.24	11.25	11.27	271.82	217.65	8.58	0.00	-1.11	**	**	**	*
(iv)	11.20	11.22	11.26	262.46	205.74	8.20	0.00	-1.04	**	**	**	OK
(v)	11.24	11.28	11.34	264.38	205.75	8.20	0.00	-1.04	**	**	**	OK
(vi)	11.20	11.23	11.27	262.93	207.88	8.23	13.21	-0.53	**	**	**	OK
(vii)	11.28	11.31	11.35	273.66	217.21	8.50	14.03	-0.55	**	**	**	**
(viii)	10.99	11.01	11.02	240.31	180.40	7.18	-4.78	-0.76	*	OK	OK	OK
(ix)	10.99	11.00	11.02	239.69	183.00	7.25	0.30	-0.61	*	OK	OK	OK
(x)	11.16	11.17	11.19	260.96	202.76	7.99	4.88	-0.44	**	OK	OK	OK
(xi)	11.09	11.12	11.16	249.29	188.17	7.55	-43.44	-2.38	**	OK	OK	OK
(xii)	10.93	10.95	10.99	229.41	169.71	6.68	14.23	-0.20	OK	OK	OK	OK
(xiii)	10.95	10.98	11.04	228.21	166.60	6.56	16.49	-0.09	OK	OK	OK	OK
(xiv)	10.95	10.99	11.05	228.66	167.09	6.58	16.24	-0.11	OK	OK	OK	OK
(xv)	10.97	11.02	11.10	227.31	163.63	6.46	13.78	-0.18	OK	OK	OK	OK

(i) Random walk (ii) Random walk with drift = 3.19841 (iii) Constant mean = 2575.3 (iv) Linear trend = 2710.38 + -3.80501 t

(v) Quadratic trend = 2719.73 + -4.58416 t + 0.010974 t^2 (vi) Exponential trend = exp(7.89515 + -0.00132193 t)

(vii) S-curve trend = exp(7.85255 + -0.0626832 /t)

(viii) Simple moving average of 2 terms (ix) Simple exponential smoothing with alpha = 0.6279 (x) Brown's linear exp. smoothing with alpha = 0.2497 (xi) Holt's linear exp. smoothing with alpha = 0.4503 and beta = 0.0149 (xii) ARIMA(1,1,1)

(xiii) ARIMA(2,1,1) (xiv) ARIMA(1,1,2) (xv) ARIMA(2,1,2)

Table 4

Model Coefficient Summary ARIMA (1, 1, 1)

Parameter	Estimate	Stnd. Error	t-statistic	P-value
AR(1)	0.549091	0.100024	5.48959	0.000001
MA(1)	0.964562	0.024232	39.8054	0.000000

We have applied several time series models on the gram area data i.e. 1948-2017. On the evidence of Table 3 and Table 4, we found that the ARIMA (1, 1, 1) is the best model for gram area. The following model is given as:

$$\square \hat{x}_t = 0.549091(\hat{x}_{t-1} - \hat{x}_t) + 0.964562\hat{u}_{t-1} \quad (4)$$

where $\square \hat{x}_t = \hat{x}_t - \hat{x}_{t-1}$ is forecasted gram production for time t year

$\hat{x}_{t-1}$ is the forecasted gram area forecasted of previous one year

$\hat{u}_{t-1}$ is residual of one year behind

Table 5

Model validity and selection criteria's for gram production

Model	*AIC*	*HQC*	*SBIC*	*RMSE*	*MAE*	*MAPE*	*ME*	*MPE*	*RUNS*	*RUNM*	*AUTO*	*MEAN*
(i)	10.17	10.17	10.17	161.91	119.03	23.09	-1.96	-5.14	OK	OK	***	OK
(ii)	10.22	10.23	10.25	163.09	119.00	23.04	0.00	-4.75	OK	OK	***	OK
(iii)	9.83	9.84	9.86	134.28	105.44	22.31	0.00	-7.18	*	*	OK	OK
(iv)	9.84	9.87	9.91	133.20	101.26	21.26	0.00	-6.86	*	*	OK	OK
(v)	9.88	9.92	9.98	134.16	101.12	21.26	0.00	-6.87	*	*	OK	OK
(vi)	9.86	9.88	9.92	134.45	101.47	20.55	16.73	-3.48	*	*	OK	OK
(vii)	9.89	9.91	9.95	136.20	107.49	21.96	17.49	-3.71	*	*	OK	OK
(viii)	9.87	9.88	9.90	136.90	104.17	21.58	-7.55	-6.61	OK	OK	OK	OK
(ix)	9.84	9.85	9.87	135.17	108.38	23.10	-10.33	-8.71	OK	OK	OK	OK
(x)	9.88	9.89	9.91	137.92	111.63	23.76	-10.09	-8.66	OK	OK	OK	OK
(xi)	9.91	9.94	9.98	138.01	110.60	23.52	-13.32	-9.11	OK	OK	OK	OK
(xii)	9.91	9.92	9.94	139.60	113.57	24.08	-7.05	-8.09	OK	OK	OK	OK
(xiii)	9.80	9.85	9.92	126.53	99.20	20.55	1.37	-5.55	OK	OK	OK	OK
(xiv)	9.83	9.87	9.92	130.41	103.34	21.62	-0.22	-6.40	*	OK	OK	OK
(xv)	9.84	9.87	9.90	133.17	102.29	21.29	-1.11	-6.86	OK	OK	OK	OK

(i) Random walk (ii) Random walk with drift = -1.96087 (iii) Constant mean = 545.617 (iv) Linear trend = 586.414 + -1.14922 t

(v) Quadratic trend = 592.705 + -1.67343 t + 0.0073833 t^2 (vi) Exponential trend = exp(6.36939 + -0.00282349 t) (vii) S-curve trend = exp(6.25981 + 0.135338 /t) (viii) Simple moving average of 2 terms

(ix) Simple exponential smoothing with alpha = 0.2005 (x) Brown's linear exp. smoothing with alpha = 0.1049

(xi) Holt's linear exp. smoothing with alpha = 0.2483 and beta = 0.0301 (xii) Brown's quadratic exp. smoothing with alpha = 0.0715 (xiii) ARIMA(2,1,2) (xiv) ARIMA(2,1,1) (xv) ARIMA(1,1,1)

Table 6

ARIMA (2, 1, 2) Model Coefficients Summary

Parameter	Estimate	Stnd. Error	t-statistic	P-value
AR(1)	-0.235099	0.266442	-0.882365	0.380831
AR(2)	0.452412	0.11887	3.80595	0.000315
MA(1)	0.456855	0.297369	1.53632	0.129314
MA(2)	0.493142	0.281889	1.74942	0.084939

We have applied several time series models on the gram production data i.e. 1948-2017. On the evidence of Table 5 and Table 6, we observed that the ARIMA (2, 1, 2) is the best model for forecasting gram production. The estimated model is given as:

$$\Delta \hat{x}_t = -0.235099(\hat{x}_{t-1} - \hat{x}_t) + 0.452412(\hat{x}_{t-2} - \hat{x}_{t-1}) + 0.456855\hat{u}_{t-1} + 0.493142\hat{u}_{t-2}, \quad (5)$$

where $\Delta \hat{x}_t = \hat{x}_t - \hat{x}_{t-1}$ is forecasted gram production for time t year

$\hat{x}_{t-1}$ is the forecasted gram production for previous one year

$\hat{x}_{t-2}$ is the forecasted gram production of previous two years

$\hat{u}_{t-1}$ is residual of one year behind

$\hat{u}_{t-2}$ is residual of two years behind

Testing Selected Model Assumptions

The forecasted value of the selected model is accurate and reliable if the selected model satisfies its assumptions like normality, autocorrelation, and heteroscedasticity. It is noticed from Table 7 that the residuals of ARIMA

(2, 1, 2) and ARIMA (1, 1, 1) models are uncorrelated and as well as independent. From Figure 1 and Figure 5, we found that the residuals of ARIMA (2, 1, 2) and ARIMA (1, 1, 1) models are normally distributed. The normality of residuals is also tested by Periodogram as exposed in Figure 4 and Figure 8 respectively, which indicated normality of selected model residuals.

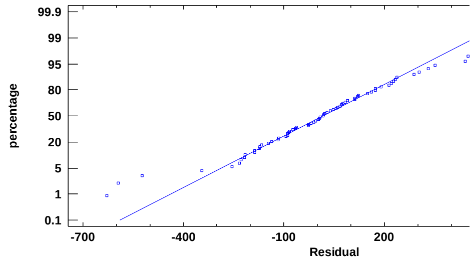

Figure 1.

Normal Probability Plot of Gram Area residuals based on ARIMA (1, 1, 1)

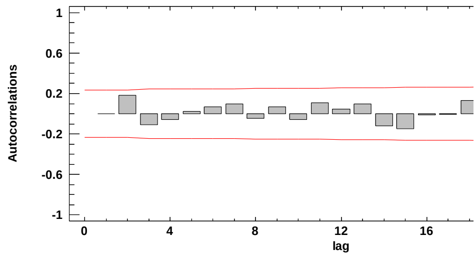

Figure 2.

Autocorrelation Plot of Gram Area residuals based on ARIMA (1, 1, 1)

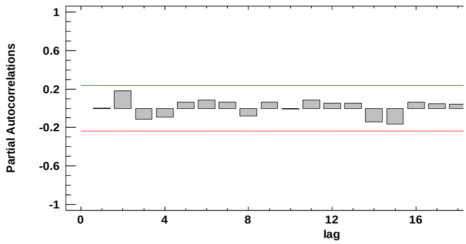

Figure 3.

Partial Autocorrelation Plot of Gram Area residuals based on ARIMA (1, 1, 1)

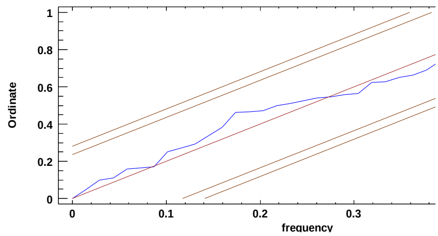

Figure 4.

Residuals Periodogram of Gram Area based on ARIMA (1, 1, 1)

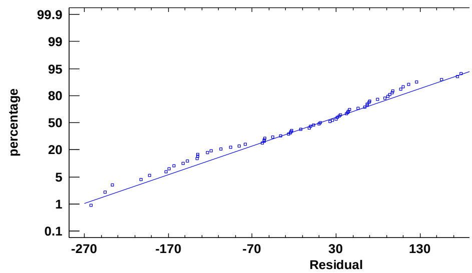

Figure 5

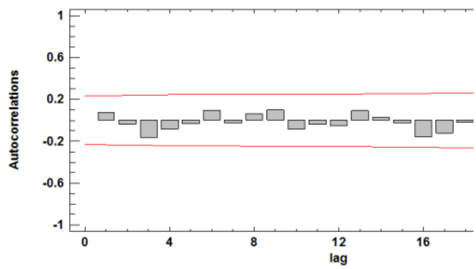

Figure 6

Normal Probability Plot of Gram Production residuals based on ARIMA (2, 1, 2)

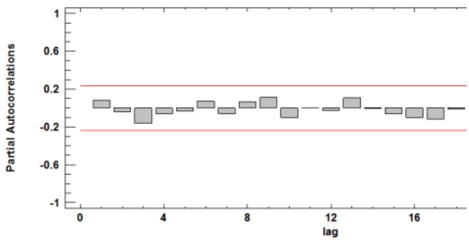

Figure 7.

Partial Autocorrelation Plot of Gram Production residuals based on ARIMA (2, 1, 2)

Autocorrelation Plot of Gram Production residuals based on ARIMA (2, 1, 2)

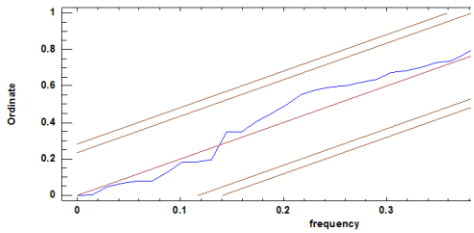

Figure 8.

Residuals Periodogram of Gram Production based on ARIMA (2, 1, 2)

Table 7

Tests for Autocorrelation and Randomness

Test	Test Statistic	P-value	Expected Number of Runs
Gram Area			
Runs overhead and underneath median	1.5883	0.1122	35
Runs up and down	0.3375	0.7356	44
Box-Pierce Test	11.9471	0.941	--
Gram Production			
Runs overhead and underneath median	0.6109	0.54	35
Runs up and down	0.6269	0.53	45
Box-Pierce Test	9.22	0.96	--

To check the randomness sequences of the residuals of selected models for area and production of Gram another three test runs. Usually, another name of the randomness sequence is white noise. In the first test the, randomness scenario is identified by counting the number of times below or above the median. If there is randomness between the residuals then the number of such runs equals to the expected number. The P-value of the 1st test is greater than 0.05 with a 95% confidence interval which shows there is randomness in the residuals. The second test checks the randomness of residuals by counting the number of times rose or fell. The sequence is random if the expected number is compared with runs. The Pr-value of the

2nd test is also insignificant which reveals that the residuals of selected models are random. The 3rd test identifies the random sequences by using sum squares of the first 24 autocorrelation functions. The statistic of the white test is reported for gram area and production respectively in Table 8 and Table 9 for testing the heteroscedasticity of the selected model residuals. These results showed that the selected model residuals are homoscedastic for forecasting gram area and production of Pakistan.

Table 8

Test of heteroscedasticity for fitted ARIMA (1, 1, 1)

F-statistics	2.522	Prob. F(3,64)	0.0656
Obs*R-squared	7.1901	Prob. Chi-Square(3)	0.0661
Scaled explained SS	10.909	Prob. Chi-Square(3)	0.0122

Table 9

Test of heteroscedasticity for fitted ARIMA (2, 1, 2)

F-statistics	1.7717	Prob. F(3,64)	0.1616
Obs*R-squared	5.2129	Prob. Chi-Square(3)	0.1569
Scaled explained SS	7.0253	Prob. Chi-Square(3)	0.0711

Table 10

Comparison of Actual and forecasted values of Gram area and production based on ARIMA (1, 1, 1) and ARIMA (2, 1, 2) model using 1948-2017 data

	Gram Area			**Gram Production**		
Period	**Data**	**Forecast**	**Residual**	**Data**	**Forecast**	**Residual**
1948	2179			465		
1949	2797	2295.98	501.02	754	533.187	220.813
1950	2398	2653.07	-255.073	599	554.73	44.2699
1951	2756	2424.95	331.053	744	637.071	106.929
1952	2110	2633.25	-523.253	422	569.104	-147.104
1953	2032	2260	-227.997	316	577.776	-261.776

	Gram Area			Gram Production		
Period	**Data**	**Forecast**	**Residual**	**Data**	**Forecast**	**Residual**
1954	2561	2209.09	351.912	562	387.38	174.62
1955	3046	2512.03	533.971	594	505.527	88.4734
1956	3247	2797.26	449.739	688	571.238	116.762
1957	3162	2923.57	238.434	681	583.405	97.5952
1958	2998	2885.34	112.657	653	623.006	29.9944
1959	3013	2799.28	213.716	568	594.585	-26.5846
1960	2821	2815.09	5.90592	598	572.67	25.3303
1961	2732	2709.88	22.1222	600	554.03	45.9703
1962	2951	2661.79	289.207	613	579.609	33.391
1963	3013	2792.29	220.707	667	572.924	94.0762
1964	2751	2834.16	-83.1576	600	600.74	-0.74031
1965	2991	2687.35	303.651	661	594.127	66.8728
1966	2643	2829.89	-186.891	531	586.161	-55.1612
1967	3074	2632.18	441.815	625	581.383	43.6171
1968	2769	2884.5	-115.5	473	551.363	-78.3627
1969	2368	2712.93	-344.934	520	565.553	-45.5528
1970	2293	2480.52	-187.525	503	499.639	3.36138
1971	2259.08	2432.7	-173.617	493.8	549.188	-55.3884
1972	2383.13	2407.92	-24.7896	510.3	511.919	-1.61873
1973	2513.85	2475.16	38.6941	553.1	530.313	22.7875
1974	2737.98	2548.3	189.676	610.1	540.89	69.2097
1975	2462.2	2678.09	-215.894	550.2	573.206	-23.0063
1976	2640.12	2519.01	121.105	601.4	566.45	34.9497
1977	2704.62	2621	83.6193	649.4	557.642	91.7581
1978	2715.99	2659.38	56.6096	613.5	602.123	11.3766
1979	2531.39	2667.63	-136.24	537.8	593.209	-55.4086
1980	2788.64	2561.44	227.201	313.4	559.059	-245.659
1981	2082.89	2710.74	-627.855	336.9	471.463	-134.563
1982	2227.94	2300.97	-73.0336	293.7	412.474	-118.774
1983	2206.45	2378.03	-171.581	491	435.11	55.8905
1984	2272.42	2360.15	-87.7307	521.9	458.11	63.7905
1985	2504.95	2393.27	111.685	523.7	547.191	-23.4915
1986	2553.39	2524.9	28.4867	586.2	516.531	69.6692
1987	2673.98	2552.51	121.469	583.3	552.077	31.2234

	Gram Area			Gram Production		
Period	**Data**	**Forecast**	**Residual**	**Data**	**Forecast**	**Residual**
1988	2027.78	2623.03	-595.25	371.5	563.636	-192.136
1989	2420.2	2247.11	173.087	456	492.363	-36.3628
1990	2558.58	2468.72	89.8586	561.9	451.676	110.224
1991	2697.21	2547.89	149.321	531	542.808	-11.8075
1992	2463.44	2629.3	-165.861	512.8	537.213	-24.4133
1993	2489.88	2495.06	-5.18232	347.3	520.075	-172.775
1994	2582.3	2509.4	72.9034	410.7	468.948	-58.2475
1995	2630.49	2562.73	67.7628	558.5	432.734	125.766
1996	2764.91	2591.59	173.321	679.6	523.703	155.897
1997	2718.7	2671.54	47.1597	594.4	584.753	9.64691
1998	2723.89	2647.84	76.052	767.1	587.931	179.169
1999	2661.13	2653.38	7.74711	697.9	601.341	96.5587
2000	2401.41	2619.2	-217.786	564.5	659.831	-95.3313
2001	2236.35	2468.87	-232.519	397	560.491	-163.491
2002	2307.76	2370	-62.2356	362.1	497.731	-135.631
2003	2379.67	2407	-27.3307	675.2	437.114	238.086
2004	2427.36	2445.52	-18.1573	611.1	543.916	67.1845
2005	2703.14	2471.06	232.08	868.2	619.716	248.484
2006	2542.51	2630.71	-88.2029	479.5	632.104	-152.604
2007	2600.33	2539.39	60.9434	837.8	634.378	203.422
2008	2735.01	2573.29	161.715	474.6	560.032	-85.4324
2009	2670.27	2652.98	17.2928	740.5	660.802	79.6982
2010	2636.42	2618.04	18.3781	561.5	519.391	42.1092
2011	2604.05	2600.11	3.94358	496	665.339	-169.339
2012	2489.63	2582.47	-92.8421	284.4	487.015	-202.615
2013	2451.33	2516.35	-65.0249	751.3	480.588	270.712
2014	2346.56	2493.02	-146.46	399	522.043	-123.043
2015	2313.44	2430.3	-116.862	379.2	615.77	-236.57
2016	2323.58	2407.97	-84.3946	286.2	393.226	-107.026
2017	2399.69	2410.55	-10.8616	329.7	464.664	-134.964

On the basis of selected time series models i.e. ARIMA (1, 1, 1) and ARIMA (2, 1, 2), the comparison of the actual and forecasted gram area and production for the years 1948 to 2017 is given in Table 10. While gram area and production forecasted values using ARIMA (1, 1, 1) and ARIMA (2, 1,

2) model for the years 2018 to 2030 are given in Table 11. From Table 11, it is shown that the area and production of a gram in Pakistan would become 2505.82 thousand acers and 502.597 thousand tons with minimum/maximum gram area and production may be expected to be 1935.9/3083.74 thousand acers and 209.27/795.923 thousand tons in 2022. While gram area and production of Pakistan for the year 2030 is forecasted to be 2515.56 thousand acers and 518.325 thousand tons.

Table 11

Gram area and production forecasts

	Gram area			Gram production		
Year	**Forecast**	**Lower**	**Upper**	**Forecast**	**Lower**	**Upper**
2018	2451.96	1990.75	2913.17	391.837	138.219	645.454
2019	2480.66	1946.44	3014.88	463.465	198.087	728.843
2020	2496.42	1937.48	3055.36	474.737	187.825	761.649
2021	2505.07	1936.06	3074.08	504.492	216.709	792.276
2022	2509.82	1935.9	3083.74	502.597	209.27	795.923
2023	2512.43	1935.64	3089.22	516.504	223.024	809.984
2024	2513.86	1935.09	3092.64	512.377	216.681	808.072
2025	2514.65	1934.31	3094.99	519.639	223.822	815.456
2026	2515.08	1933.38	3096.78	516.064	219.014	813.114
2027	2515.32	1932.36	3098.27	520.19	222.967	817.414
2028	2515.45	1931.3	3099.6	517.603	219.527	815.679
2029	2515.52	1930.21	3100.83	520.078	221.761	818.394
2030	2515.56	1929.11	3102.01	518.325	219.335	817.316

Conclusion

The Population of Pakistan is increasing over time, so it is necessary to design a strategy that fulfills the nation's gram requirements. For this purpose, forecasting is a strategic tool to alarm the nation's requirements in advance. For this purpose, time series modeling is employed to forecast gram area and production. We applied multiple time series models on gram area and production data. For the best model selection, we used the model selection criteria's SBC and AIC, etc. We selected the best model for gram area and production of Pakistan for forecasting purpose time series model selection criterion. The best models were found to be ARIMA (1, 1, 1) and ARIMA (2, 1, 2) for gram area and production respectively. The selected model founds that gram area and production of Pakistan are forecasted to be 2525.56 thousand acers and 1854.33 thousand tons in 2030 under the condition that there is no irregular event happened in a country. The authors suggested the government gives the awareness of the former about the proper use of the latest technology. The government provided different research projects on gram which may play a significant role to increase gram production and reduce their price.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Pet rov & F. Caski (Eds.), *Proceedings of the Second International Symposium on Information Theory*. Akademiai Kiado.

Amin, M., Amanullah, M., & Akbar, A. (2014). Time series modeling for forecasting wheat production of Pakistan. *JAPS: Journal of Animal & Plant Sciences*, *24*(5), 1444-1451.

Awal, M. A., & Siddique, M. A. B. (2011). Rice production in Bangladesh employing by ARIMA model. *Bangladesh Journal of Agricultural Research*, *36*(1), 51-62. https://doi.org/10.3329/bjar.v36i1.9229

Amir, M. W., Amin, M., & Nazir, H. Z. (2021). Statistical modeling and forecasting for onion production of Pakistan. *Journal of Agricultural Research*, *59*(2), 171-176.

Agriculture Marketing Information Service. (2019). *Directorate of agriculture (Economics and Marketing) Punjab, Lahore*. http://www.amis.pk/Aboutus.aspx

Bilal, M., & Shahbaz, M. Q. (2008). Modeling barley production in Punjab. *Pakistan Journal of Statistics and Operation Research*, *4*(1), 25-33. https://doi.org/10.18187/pjsor.v4i1.58

Box, G. E. P., & Jenkins, G. M. (1976), *Time series analysis forecasting and control* (revised ed.). San Francisco

Chatfield, C. (1995). *The Analysis of time series an introduction* (5th ed.). Chapman & Halljcrc.

Celik, S., Karadas, K., & Eyduran, E. (2017). Forecasting the production of groundnut in Turkey using Arima model. *The Journal of Plant and Animal Science*, *27*(3), 920-928.

Karadas, K., Celik, S., Eyduran, E., & Hopoglu, S. (2017). Forecasting production of some oil seed crops in Turkey using exponential smoothing methods. *The Journal of Animal and Plant Science*, *27*(5), 1719-1729.

Gujarati, D. N. (2009). *Basic econometrics*. Tata McGraw-Hill Education

Government of Pakistan. (2016). *Agricultural statistics of Pakistan*. Ministry of food and agriculture, Islamabad. http://www.mnfsr.gov.pk/pubDetails.aspx

Habib, N., Anuar, M. Z., & Saeed, I. (2013). Forecasting of millet area and production in Pakistan. *Journal of Social Walefare and Human Rights*, *1*(1), 147-52. https://doi.org/10.9734/AJEA/2016/27930

Iqbal, N., Bakhsh, K., Maqbool, A., & Ahmad, A. S. (2005). Use of the ARIMA model for forecasting wheat area and production in Pakistan. *Journal of Agriculture and Social Sciences*, *1*(2), 120-122.

Khan, M., Mustafa, K., Shah, M., Khan, N., & Khan, J. Z. (2008). Forecasting mango production in Pakistan an econometric model approach. *Sarhad Journal of Agriculture, 24*(2), 363-370.

Khan, D., Ullah, A., Bibi, Z., Ullah, I., Zulfiqar, M., & Khan, Z. U. (2020). Forecasting area and production of guava in Pakistan: An econometric analysis. *Sarhad Journal of Agriculture*, *36*(1), 272-281. http://dx.doi.org/10.17582/journal.sja/2017/33

Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (2003). *Forecasting: methods and applications* (3rd ed.). John Wiley and Sons.

Masood, M. A., Raza, I., & Abid, S. (2018). Forecasting wheat production using time series models in Pakistan. *Asian Journal of Agriculture and Rural Development*, *8*(2), 172-177. https://doi.org/10.18488/journal.1005/2018.8.2/1005.2.172.177

Naheed, S., Raza, I., Anwar, M. Z., Habib, N., Zahra, N., & Siddiqui, S. (2015). Forecasting area and production of barley in Punjab, Pakistan. *Pakistan Journal of Agricultural Research*, *28*(3), 304-309.

Naz, F. (2012). A univariate time series modelling of dates exports in Pakistan. *Journal of Contemporary Issues in Business Research*, *1*(2), 57-68.

Naheed, S., Raza, I., Anwar, M. Z., Habib, N., Zahra, N., & Siddiqui, S. (2015). Forecasting area and production of barley in Punjab, Pakistan. *Pakistan Journal of Agricultural Research*, *28*(3), 304-309.

Pakistan Agriculture Research Council. (2019). *Pakistan agriculture research council: Ministry of national food security & research*. PARC. http://www.parc.gov.pk/

Qureshi, M. N., Bilal, M., Ali, M. K., & Ayyub, R. M. (2014a). Modeling of citrusfruits production in Pakistan. *Science International (Lahore)*, *26*(2), 733-737.

Qureshi, M. N., Bilal, M., Ayyub, R. M., & Ayyub, S. (2014). Modelling on Mango Production in Pakistan. *Science International (Lahore)*, *26*(3), 1227-1231.

Rani, S., & Raza, I. (2012). Comparison of trend analysis and double exponential smoothing methods for price estimation of major pulses in Pakistan. *Pakistan Journal of Agricultural Research*, *25*(3), 233-239.

Rahman, N. M., & Baten, M. (2016). Forecasting area and production of black gram pulse in Bangladesh using ARIMA models. *Pakistan Journal of Agricultural Sciences*, *53*(4), 759-765. https://doi.org/10.21162/PAKJAS/16.1892

Sahu, P. K., Mishra, P., Dhekale, B. S., Vishwajith, K. P., & Padmanaban, K. (2014). Modelling and forecasting of area, production, yield and total seeds of rice and wheat in SAARC countries and the world towards food security. *American Journal of Applied Mathematics and Statistics*, *3*(1), 34-48. https://doi.org/10.12691/ajams-3-1-7

Tsay, R. S. (2005). *Analysis of financial time series* (Vol. 543). John Wiley & Sons.

Ullah, A., Khan, D., & Zheng, S. (2018). Forecasting of peach area and production wise econometric analysis. *Journal of Animal and Plant Sciences*, *28*(4), 1121-1127.

Vishwajith, K. P., Dhekale, B. S., Sahu, P. K., Mishra, P., & Noman, M. D. (2014). Time series modeling and forecasting of pulses production in India. *Journal of Crop and Weed*, *10*(2), 147-154.

Wali, V. B., & Lokesh, D. B. H. (2017). Forecasting of Area and Production of Cotton in India and Karnataka Using ARIMA Model. *Indian Journal of Economics and Development*, *13*(4), 723-728. https://doi.org/10.5958/2322-0430.2017.00236.0