Journal of Pharmacognosy and Phytochemistry
Available online at www.phytojournal.com

**E-ISSN:** 2278-4136
**P-ISSN:** 2349-8234
JPP 2019; 8(1): 2158-2165
Received: 22-11-2018
Accepted: 24-12-2018

**Bhola Nath**
Department of Agricultural Economics and Agricultural Statistics, Institute of Agriculture, Visva-Bharati, Sriniketan, West Bengal, India

**DS Dhakre**
Department of Agricultural Economics and Agricultural Statistics, Institute of Agriculture, Visva-Bharati, Sriniketan, West Bengal, India

**Debasis Bhattacharya**
Department of Agricultural Economics and Agricultural Statistics, Institute of Agriculture, Visva-Bharati, Sriniketan, West Bengal, India

**Correspondence**
**Debasis Bhattacharya**
Department of Agricultural Economics and Agricultural Statistics, Institute of Agriculture, Visva-Bharati, Sriniketan, West Bengal, India

# Forecasting wheat production in India: An ARIMA modelling approach

**Bhola Nath, DS Dhakre and Debasis Bhattacharya**

**Abstract**
Box-Jenkins' ARIMA model: A time series modelling approach has been used to forecast wheat production for India. ARIMA (1,1,0) model was found to be the best ARIMA model for the present study. The efforts were made to forecast, the future wheat production for a period up to ten years as accurate as possible, by fitting ARIMA (1,1,0) model to our time series data. The forecast results have shown that the annual wheat production will grow in 2026-27. The wheat production will continuously grow with an average growth rate of approximately 4% year by year.

**Keywords:** ARIMA, forecast, modelling, production wheat, time series

## Introduction

In India, wheat is an important crop in terms of both production and consumption. It accounts for around 35% of total food production (98.38 million tonnes) and 21% of the total cultivated area (30.597 million hectares) in the country (2016-17). It is cultivated across the country with the exception of the southern and north-eastern states whose contributions to production are minimal. Nowadays, some of its products (biscuits, bread, noodles, etc.) have become very popular as a result of changes in lifestyle and the influence of western diets. Wheat provides both macro- (e.g. carbohydrates, fat and protein) and micro-nutrients (e.g. calcium and iron) and hence helps in building a healthy society (Tripathi and Mishra (2017)).

Wheat is mainly grown in the Rabi season (October–December to March–May) along with barley, lentils, peas, mustard and potatoes. The planting of winter wheat begins about 1$^{st}$ of October and runs through to the end of December. Wheat will usually begin to head in January, with the harvest following in March, April and May. Wheat acreage increased from 13% of the total cropped area (1990–1991) to about 15% (2009–2010). Wheat production is mainly confined to the Indo-Gangetic Plains Region, and three northern states, namely Uttar Pradesh (35.53 %), Punjab (18.96 %) and Haryana (13.39 %), and these states supply 72% of India's total wheat output. In addition, Rajasthan (8.31 %) and Madhya Pradesh (8.78 %) contribute a total output of 86 %. Wheat is one of the staple foods in India and it is a popular food item among both vegetarians and non-vegetarians. It provides nearly 50% of the calories and protein requirements for the vast majority of population. India is the second largest producer of wheat in the world, averaging an annual production of 66 million tonnes (which is increased by 98.98 million tonnes in 2016-17). On average, India consumes 65 million tonnes of wheat, ranking it as the second-largest consumer of wheat in the world. Although India has been self-sufficient in wheat, it also imports wheat. In recent years, India, on average, imports 1 million tonnes of wheat and, for various reasons, exports an average of 0.7 million tonnes. About 60% of India's total cropped area is still rain-fed and therefore dependent on the monsoon. India's food grain production, and especially wheat production, slumped in the early 2000s as a result of widespread drought in 2002–2003.

Thus, estimation of the future production of the wheat crop in context to growing population and need of food to them is an important issue. To meet this requirement, we are using the time series modelling approach known as Box-Jenkins ARIMA model to forecast the future production of the concerned crop. The ARIMA model utilizes the mix modelling approach of autoregressive and moving average methods. Also, ARIMA model has been frequently employed to forecast the future requirements in terms of internal consumption and export to adopt appropriate measures (Muhammad *et al*., 1992) [8].

The cultivation of wheat in India was started 5000 years ago. India produces 98.38 million tonnes of wheat with the second position in the world during the season 2017-18. Its cultivation area is 30.597 million hectares. As per the context wheat is a staple food in India, then it becomes our responsibility to take care of the wheat cultivation. Wheat is grown in almost all the states in India. The major wheat producing states are Uttar Pradesh, Punjab, Haryana, Rajasthan with the contribution of almost 80% to the total production in India. Only 13% area is rainfed under the wheat crop. Gujarat, Maharashtra, Madhya Pradesh, West Bengal and Karnataka have the major Rainfed areas under wheat cultivation in India. Total 33% of wheat area in India is covered by Central and Peninsular Zone. Only 33% area under irrigated wheat cultivation receive desired irrigations and remaining has limited irrigation. Table: 1 represents the 68 years' data on wheat production in India. Data is obtained from the online secondary source, *http://Indiastat.com*, for the period from 1949-50 to 2016-17.

In the present work, an effort is made to forecast wheat production for the ten leading years. The model developed for forecasting is an Autoregressive Integrated Moving Average (ARIMA) model. It was introduced first by Box and Jenkins (1960) and also known as Box-Jenkins Model which is used for forecasting a single variable. The ARIMA model adopts the non-zero autocorrelation in the time series data. This is the core reason for selecting the ARIMA model in the present study. The open source analytical software '*R*' (build 3.5.1) including packages such as 'tseries, 'forecast' and 'urca' etc. were used under this study.

**Table 1:** Wheat production in India (in '000 tonne)

| S. No. | Year | Production | S. No. | Year | Production | S. No. | Year | Production |
|---|---|---|---|---|---|---|---|---|
| 1 | 1949-50 | 6400 | 24 | 1972-73 | 24735 | 47 | 1995-96 | 62097 |
| 2 | 1950-51 | 6462 | 25 | 1973-74 | 21778 | 48 | 1996-97 | 69350 |
| 3 | 1951-52 | 6183 | 26 | 1974-75 | 24104 | 49 | 1997-98 | 66345 |
| 4 | 1952-53 | 7501 | 27 | 1975-76 | 28846 | 50 | 1998-99 | 71288 |
| 5 | 1953-54 | 8017 | 28 | 1976-77 | 29010 | 51 | 1999-00 | 76369 |
| 6 | 1954-55 | 9043 | 29 | 1977-78 | 31749 | 52 | 2000-01 | 69681 |
| 7 | 1955-56 | 8760 | 30 | 1978-79 | 35508 | 53 | 2001-02 | 72766 |
| 8 | 1956-57 | 9403 | 31 | 1979-80 | 31830 | 54 | 2002-03 | 65096 |
| 9 | 1957-58 | 7998 | 32 | 1980-81 | 36313 | 55 | 2003-04 | 72156 |
| 10 | 1958-59 | 9958 | 33 | 1981-82 | 37452 | 56 | 2004-05 | 68637 |
| 11 | 1959-60 | 10324 | 34 | 1982-83 | 42794 | 57 | 2005-06 | 69355 |
| 12 | 1960-61 | 10997 | 35 | 1983-84 | 45476 | 58 | 2005-06 | 75807 |
| 13 | 1961-62 | 12072 | 36 | 1984-85 | 44069 | 59 | 2007-08 | 78570 |
| 14 | 1962-63 | 10776 | 37 | 1985-86 | 47052 | 60 | 2008-09 | 80679 |
| 15 | 1963-64 | 9853 | 38 | 1986-87 | 44323 | 61 | 2009-10 | 80804 |
| 16 | 1964-65 | 12257 | 39 | 1987-88 | 45096 | 62 | 2010-11 | 86874 |
| 17 | 1965-66 | 10394 | 40 | 1988-89 | 54110 | 63 | 2011-12 | 94882 |
| 18 | 1966-67 | 11393 | 41 | 1989-90 | 49850 | 64 | 2012-13 | 93506 |
| 19 | 1967-68 | 16540 | 42 | 1990-91 | 55135 | 65 | 2013-14 | 95850 |
| 20 | 1968-69 | 18651 | 43 | 1991-92 | 55690 | 66 | 2014-15 | 86527 |
| 21 | 1969-70 | 20093 | 44 | 1992-93 | 57210 | 67 | 2015-16 | 92288 |
| 22 | 1970-71 | 23832 | 45 | 1993-94 | 59840 | 68 | 2016-17 | 98510 |
| 23 | 1971-72 | 26410 | 46 | 1994-95 | 65767 | | | |
| *Source: http://indiastat.com* | | | | | | | | |

## 2. Literature Review

Raymond Y.C. Tse, (1997) [9] suggested that the following two questions must be answered to identify the data series in a time series analysis: (1) whether the data are random; and (2) have any trends? This is followed by another three steps of model identification, parameter estimation and testing for model validity. If a time series is random, then the correlation between successive values for the time series will be close to zero. If the observations of a time series are dependent statistically on each other, then the ARIMA model is said to be appropriate for the analysis.

Meyler *et al* (1998) [7] illustrated a framework for ARIMA models to forecast the Irish inflation. In their study, they emphasized seriously on the optimizing forecast concert while concentrating more on minimizing the errors in the sample forecast rather than maximizing sample fitting i.e. goodness of fit.

Stergiou (1989) [10] used ARIMA modelling technique on a 17 years' (1964 to 1980 with 204 observations) time series data in his research which were of monthly catches of pilchard (*Sardina pilchardus*) for forecasting the values for the next 12 months in advance and forecasts were equated with actual data for the year 1981 which was not used in estimation of the concerned parameters. He found mean error as about 14% and suggested that ARIMA procedure could have the capability of forecasting the complex changing aspects of the Greek pilchard fishery, which, was difficult to predict because of the year by year changes in oceanographic and biological conditions.

Contreras *et al* (2003) [2] used ARIMA forecasting methodology, provided a method to foresee next-day electricity prices both for the spot markets and long-term contracts for central Spain and Californian markets. In fact, a plenty of research studies are available to justify that a vigilant and precise selection of ARIMA model could be fitted to the time series data on a single variable to forecast, the future values in the series with a great accuracy. This study was also an attempt to predict the future values for the Sugarcane production in India by fitting an ARIMA model for the time series data for 62 years' production of Sugarcane.

## 3. Box-Jenkins (ARIMA) Model: Introduction

A time series is defined as an arrangement of data observed over the specific time period. The ARIMA models are a class of models that have abilities to represent a stationary as well as non-stationary time series. ARIMA is used to produce accurate forecast based on an explanation of historical data on a single variable. Since it does not assume any specific pattern in the past data of the time series that is about to be forecasted, this model is different from other models which are used for forecasting the future values of the time series. The approach of Box-Jenkins procedure in order to construct an ARIMA model is based on the steps, viz., (1) Identification of Model, (2) Parameter Estimation and Selection, (3) Model

Validation or Diagnostic Checking and (4) Model's use.

Model identification involves defining the orders (*p*, *d*, and *q*) of the AR and MA components of the time series model. Basically, it tries to find the answers to the question i.e. whether the time series data is stationary or non-stationary? What will be the order of differentiation (*d*), that makes the time series stationary?

## 4. Building up the ARIMA models

The given set of data (Table: 1) is used to develop the forecasting model. To identify the pattern in the given data set the line plot of wheat production in India for the period of 1949-50 to 2016-17 is shown in the Figure: 1.

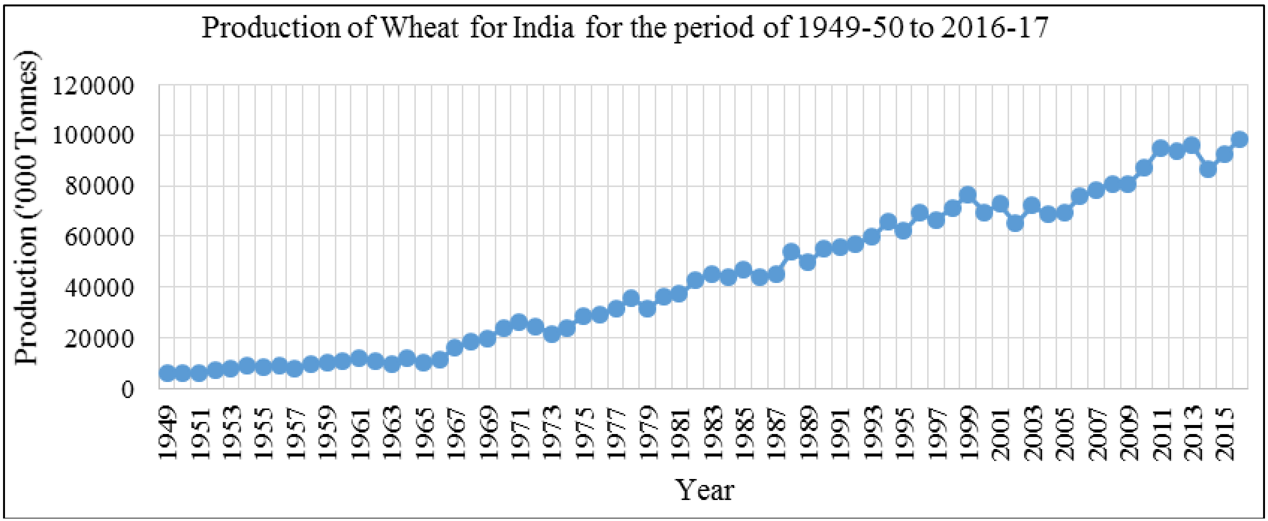


**Fig 1:** Wheat Production ('000 Tonnes) in India from 1949-50 to 2016-17

Since we have already discussed that to construct an ARIMA model for forecast of the particular variable requires following steps: 1) Identification of Model, (2) Parameter Estimation & Selection, and (3) Model Validation; before it can be (4) used for forecasting. Therefore, an attempt is to be made to identify the model for fitness.

## 5. Identification of Model

The first stage of ARIMA modelling is to identify that the variable, which is about to be forecasted, is a stationary time series or not. Stationary means the values of variable over time fluctuates around a constant mean and variance. The plot of the wheat production data in Figure 1 above undoubtedly shows that the data is not stationary (it is showing an increasing trend in the time series). The ARIMA model cannot be made until we make the series stationary. First, we have to take the difference of the time series 'd' times to obtain a stationary series to obtain an ARIMA (*p*, *d*, *q*) model, where 'd' is the order of differencing. There should a Caution is to be taken in differencing the time series as over-differencing will lead to increase the standard deviation, rather than the reduction. The best idea is to start the differencing with the lowest order (first order, $d = 1$) and then test the unit root problems for the data. Using the above approach, the obtained time series of differencing first order will be checked for the unit root problem. The line plot of the time series ($d = 1$) of wheat production data is plotted in figure: 2.

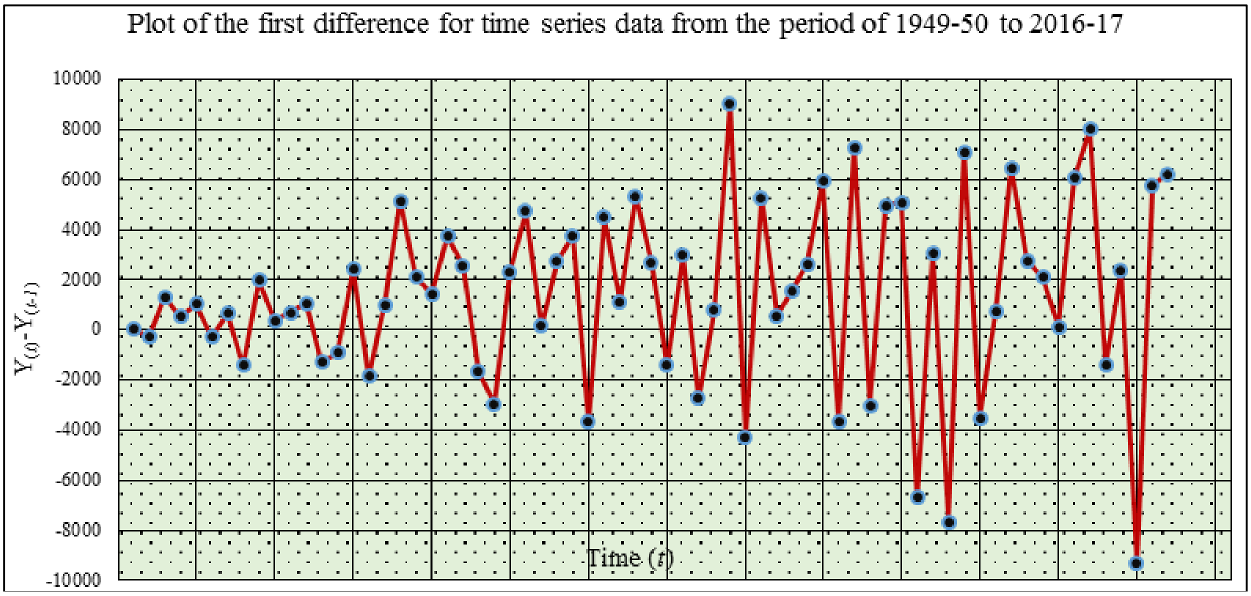


**Fig 2:** Line plot of the wheat production data of first order ($d$=1)

It can be seen from the graph (Figure 2) that the time series looks to be stationary with respect to both in the mean and variance. Before moving further, we will have to test the differenced time series for stationarity (unit root problem) by Augmented Dickey-Fuller test (ADF).

**6. Augmented Dickey-Fuller (ADF) Test: A test for stationarity**

In case of stationarity the hypothesis will become:

**$H_0$**: The time series is non-stationary against

**$H_1$**: The time series is stationary

Then this hypothesis is tested by carrying out appropriate difference ($d = 1$) of the data and applying the ADF test to the differenced time series. First order difference ($d = 1$) means we have generated a differenced time series data of current ($Y_t$) and immediate previous one values [$Y = Y_{(t)} - Y_{(t-1)}$]. The

ADF test result, is given below:

Dickey-Fuller = -4.86, Lag order = 1, $p$-value = 0.01

Therefore, $H_0$ cannot be accepted. Hence, alternative hypothesis is true and it can be concluded that the time series ($d$ =1) is stationary in respect to mean and variance. Thus, there is no need of further differencing the time series and then the adopted difference order is $d = 1$ for the ARIMA ($p$, $d$, $q$) model.

This test allows to go further in the steps for ARIMA model development which are to find out suitable values of $p$ in Auto Regressive (AR) and $q$ in Moving Average (MA) in our model. In order to determine the values of $p$ and $q$, there is need to examine the auto correlation function (ACF) and partial auto correlation function (PACF) of the stationary time series ($d = 1$).

**7. ACF and PACF analysis**

The plot of correlogram ACF (Figure: 3) of lags 1 to 20 for the first order differenced time series data of the wheat production in India.

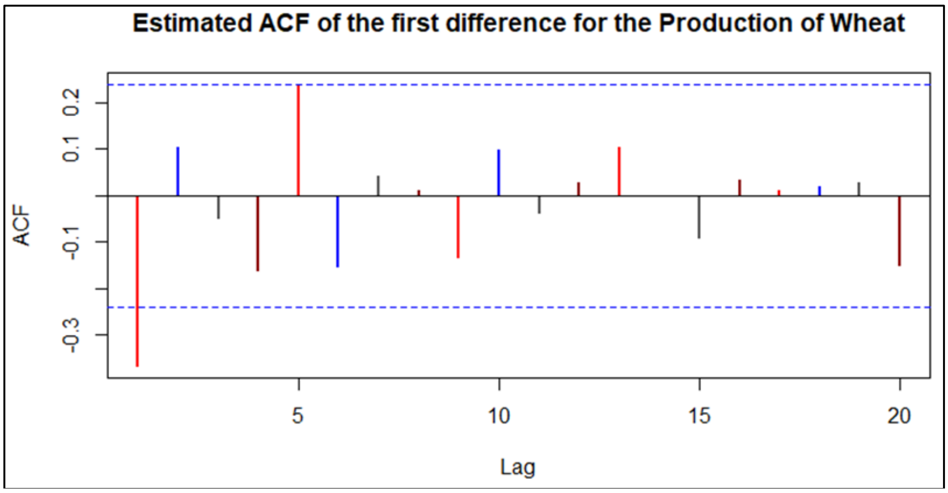


**Fig 3:** Autocorrelation Function (ACF) of first differenced series by lag

From the above correlogram it can be inferred that the auto-correlation at lag 1 and 2 are definitely exceeding the significance limits and it becomes zero after lag 2. All the Rest coefficients between lag 3 and lag 20 are well within the limits.

The Figure 4 below represents the partial correlogram of the lags 1 to 20 for the differenced time series.

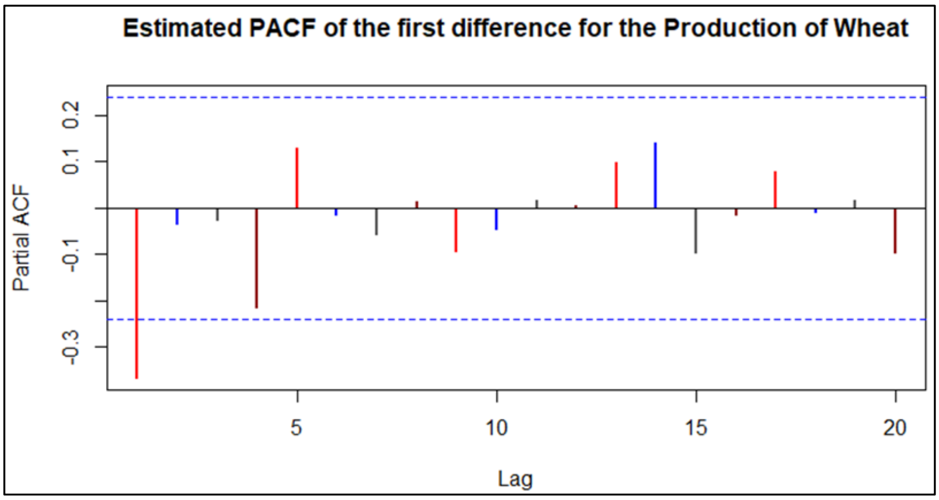


**Fig 4:** Partial Autocorrelations Function of first differenced time series by lag

The PACF (Figure: 4) also infers that the coefficients do not exceed the significance limits at lag 2 and after lag 2 to lag 20 it tails off to zero. Although there is one outlier at lag 1, which we can suppose that it is an error and happened by chance alone because all the other PACFs from lag 2 to 20 are inside the significance limits.

The ACF and PACF coefficients for lag 1 to 20 of the first order differenced series are given (Table: 2).

**Table 2:** ACF and PACF Coefficients for lag 1 to 20

| Lag | ACF | PACF | Lag | ACF | PACF |
|---|---|---|---|---|---|
| 1 | -0.368 | -0.368 | 11 | -0.039 | 0.016 |
| 2 | 0.105 | -0.035 | 12 | 0.027 | 0.006 |
| 3 | -0.050 | -0.026 | 13 | 0.103 | 0.099 |
| 4 | -0.162 | -0.215 | 14 | -0.001 | 0.141 |
| 5 | 0.240 | 0.130 | 15 | -0.090 | -0.097 |
| 6 | -0.152 | -0.015 | 16 | 0.034 | -0.016 |
| 7 | 0.042 | -0.056 | 17 | 0.011 | 0.078 |
| 8 | 0.012 | 0.013 | 18 | 0.020 | -0.009 |
| 9 | -0.133 | -0.093 | 19 | 0.027 | 0.016 |
| 10 | 0.097 | -0.046 | 20 | -0.151 | -0.096 |

Since, the ACF tailing off to zero after lag 2 (omitting the outliers 1 and 2) and the PACF is about to zero after lag 1, then the following possible Auto Regressive Moving Average (ARMA) models can define for the first differenced time series data on wheat production in India:

1. An ARMA (1, 0) model i.e. autoregressive model of order $p = 1$ since the partial autocorrelation is zero after lag 1 and the autocorrelation is also zero.
2. An ARMA (0, 2) model i.e. moving average model of the order $q = 2$ since the ACF is zero after lag 2 and the PACF are zero.
3. An ARMA ($p$, $q$) model which is a mix model with the orders $p$ and $q$ both greater than 0 because ACF and PACF both tail off to the value of zero.

## 8. Selecting the candidate model for forecasting

ARMA (1, 0) has 1 parameter, ARMA (0, 2) has 2 parameters and ARMA ($p$, $q$) has at least 2 parameters. Therefore, from the help of principle of parsimony, the models ARMA (1,0) and ARMA ($p$, $q$) would be the best models for the purpose of study. Further, there is a need to obtain the best ARIMA model using the ARMA (1,0), ARMA (p, q) mixed model (with p & q both greater than 0), and differencing order (d = 1). Based upon the conditions, we can have the following three expected ARIMA ($p$, $d$, $q$) models:

ARIMA ($p$, $d$, $q$): ARIMA (1, 1, 0), ARIMA (1, 1, 1) and ARIMA (1, 1, 2)

To select the best suitable model for forecasting out of above three, the model will be chosen, the one with lowest Akaike Information Criterion (AIC). The summary of each of the fitted ARIMA model in the time series for wheat production data can be seen for further reference (Table: 3).

**Table 3:** Summary of the fitted three ARIMA models with AIC

| Model | Coefficients | | $\sigma^2$ (Estimated) | Log-Likelihood | AIC |
|---|---|---|---|---|---|
| (1,1,0) | AR (1) | -0.3725 (0.1139) | 11865410 | -639.81 | 1285.63 |
| | Drift | 1360.342 (303.244) | | | |
| (1,1,1) | AR (1) | -0.4537 (0.2764) | 14459179 | -647.41 | 1300.81 |
| | MA (1) | 0.2461 (0.2879) | | | |
| (1,1,2) | AR (1) | -0.0938 (0.7728) | 14046093 | -646.47 | 1300.93 |
| | MA (1) | -0.0947 (0.7931) | | | |
| | MA (2) | 0.2027 (0.1274) | | | |

**Note:** AR1 (Autoregressive process of order $p = 1$), MA1 (Moving Average of order $q = 1$), MA2 (Moving Average of order $q = 2$) and figures in the parentheses are the standard error for the concerned parameter.

We can clearly observe in the Table 3 that the lowest AIC value is for the ARIMA (1,1,0) with drift model ($p = 1$, $d = 1$ and $q = 0$). Hence, the ARIMA (1,1,0) model can be the best model to forecast the future values of the given time series data. The model summary is given below for the ARIMA (1,1,0) with drift:

## 9. Forecast using selected ARIMA (1, 1, 0) model

The above selected model ARIMA (1, 1,0), which we are fitting to our time series data, means that we are fitting ARMA (1,0) model of first order difference to our time series. Also, ARMA (1,0) model, which has two parameters in it, can be rewritten an AR model of order 1, or AR (1) model, since q is zero in MA. Therefore, this model can be expressed as:

$X_t = \mu + (\beta_1 * (Z_{t-1} - \mu)) + (\beta_2 * (Z_{t-2} - \mu)) + \varepsilon_t$, Where $X_t$ is the stationary time series under study, μ is the mean of the time series $X_t$, $\beta_1$ and $\beta_2$ are parameters to be estimated (the AR1 and Drift terms in the fitted ARIMA (1,1,0) model have the values as above in Table 3, i.e. AR1 = -0.3725 and Drift = 1360.342) and $\varepsilon_t$ is white noise error term following mean zero and constant variance. Now the chosen ARIMA (1,1,0) will be fitted and used to forecast the future values of the time series. The forecast for the next 10 years with 80% and 95% (low and high) prediction intervals is done (Table: 4).

**Table 4:** 10 - Year Forecast of Wheat Production in India

| Season | Forecast | Low 80% | High 80% | Low 95% | High 95% |
|---|---|---|---|---|---|
| 2017-18 | 98059.21 | 93644.75 | 102473.70 | 91307.88 | 104810.50 |
| 2018-19 | 100094.26 | 94882.74 | 105305.80 | 92123.93 | 108064.60 |
| 2019-20 | 101203.25 | 94990.22 | 107416.30 | 91701.25 | 110705.30 |
| 2020-21 | 102657.23 | 95689.33 | 109625.10 | 92000.75 | 113313.70 |
| 2021-22 | 103982.69 | 96298.61 | 111666.80 | 92230.91 | 115734.50 |
| 2022-23 | 105356.02 | 97029.30 | 113682.70 | 92621.40 | 118090.60 |
| 2023-24 | 106711.52 | 97784.07 | 115639.00 | 93058.17 | 120364.90 |
| 2024-25 | 108073.67 | 98584.93 | 117562.40 | 93561.89 | 122585.40 |
| 2025-26 | 109433.34 | 99414.18 | 119452.50 | 94110.36 | 124756.30 |
| 2026-27 | 110793.93 | 100271.24 | 121316.60 | 94700.87 | 126887.00 |

The graph of the forecasted values of the production of wheat in India using ARIMA (1,1,0) with drift is plotted (Figure: 5).

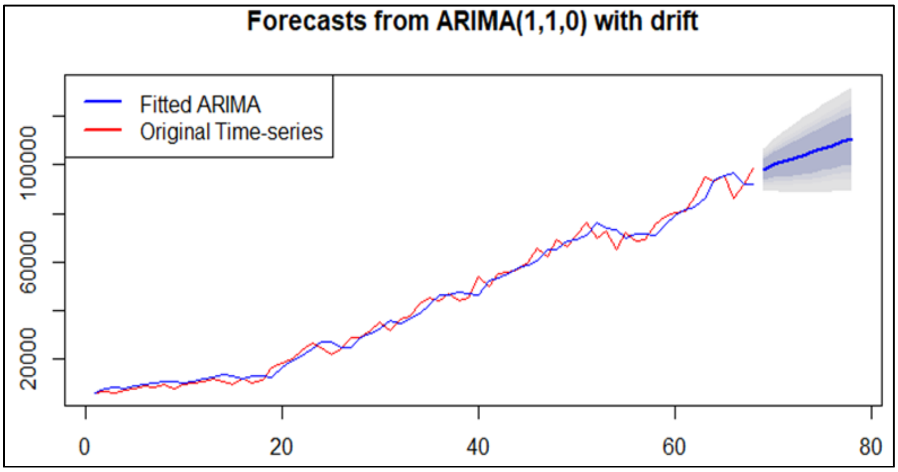


**Fig 5:** Plot of forecast from ARIMA (1,1,0)

In the Figure: 5, there are two shaded zones of forecast representing the 80% and 95% confidence intervals (with lower and upper side) projection of forecast intervals.

Now, there is a need to check the distribution of the forecast errors of the ARIMA (1,1,0) model, whether these are normally distributed by means of mean zero and constant variance. The diagnostic of the auto correlation in the residuals (forecast errors).

To identify the distribution of residuals, the errors (standard residuals) will be plotted. There are three plots viz., Figure: 6(a), 6(b) and 6(c) below show plot of residuals, histogram of the residuals and Normal Q-Q plot of residuals of the fitted ARIMA (1,1,0) model respectively.

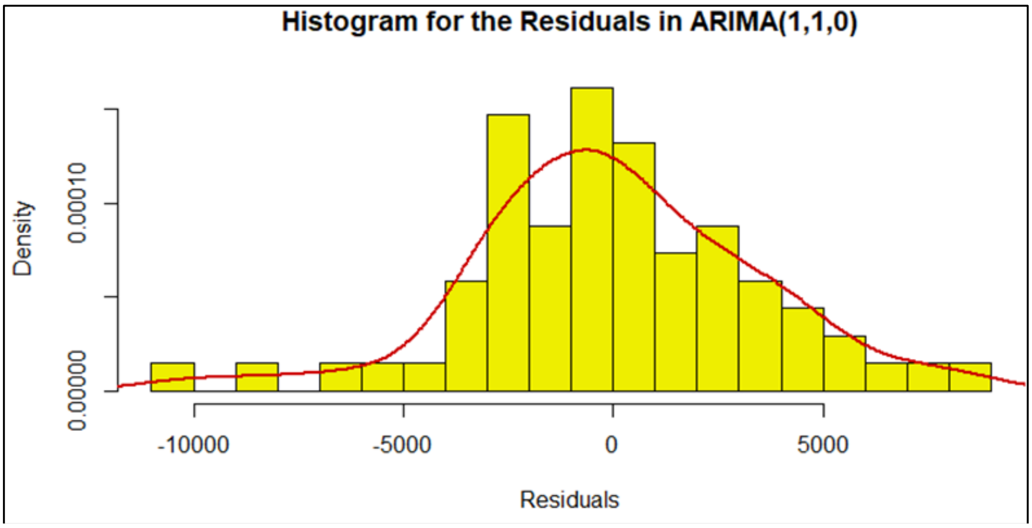


**Fig 6 (a):** Plot of Residual of fitted ARIMA (1,1,0)

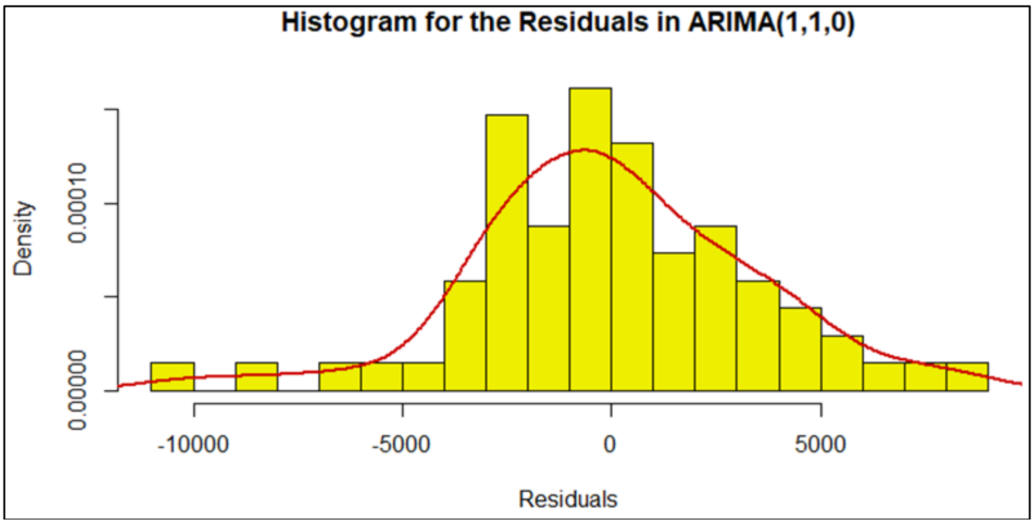


**Fig 6 (b):** Histogram of Forecast Errors (Residuals) – ARIMA (1,1,0)

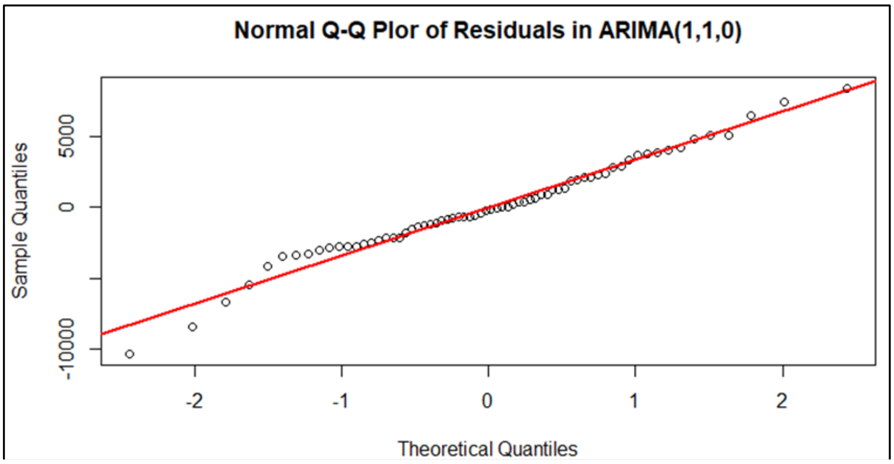


**Fig 6 (c):** Normal Q-Q Plot of Residuals (Forecast Errors) – ARIMA (1,1,0)

From these plots (Figure 6(a) and 6(b)), line and histogram It can be clearly seen that residuals are normally distributed. Q-Q plot of standard residuals in the fitted ARIMA (1,1,0) model, it can be concluded that standard errors are roughly constant with respect to mean and variance overtime.

To check whether there are any auto correlations in forecast errors, there is a need to plot correlograms; ACF and PACF of the forecast errors.

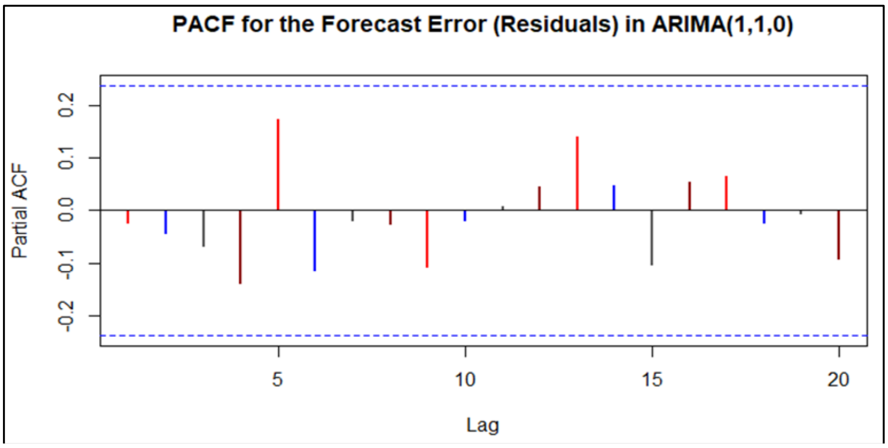


**Fig 7(a):** Estimated ACF of Residuals for the fitted ARIMA (1, 1,0)

It can be clearly concluded from the ACF plot that none of the ACF coefficients are breaking the significance limit. it means all the values of ACF are inside the significance bounds.

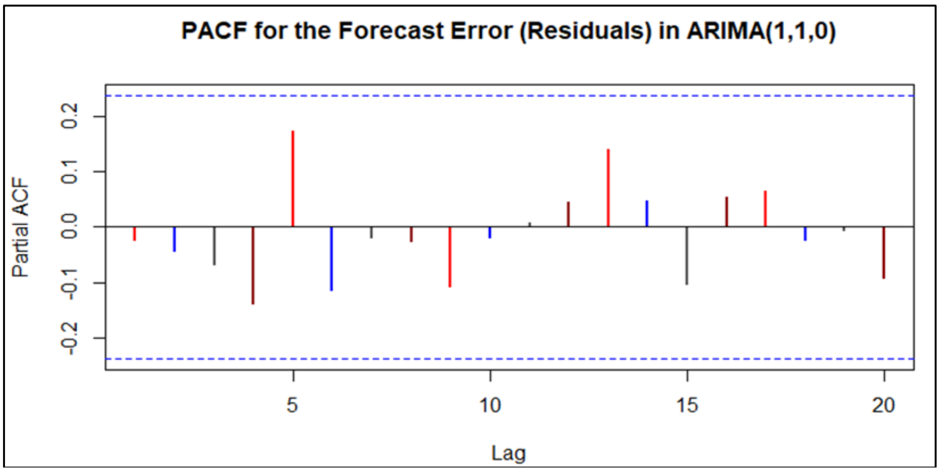


**Fig 7(b):** Estimated PACF of Residuals for the fitted ARIMA (1, 1,0)

All the ACFs and PACFs of the residuals for the fitted ARIMA (1,1,0) from lag 1 to lag 20 are within the significance limits. Then, from the ACF and PACF it can be concluded that there are not non-zero autocorrelations in the errors (or standard residuals) at between the lag 1 and 20 in the fitted ARIMA (1, 1,0) model.

**Test for the autocorrelation in ACF and PACF of the Residuals**

The Box-Ljung and Box-Pierce test results are conducted to test the autocorrelation in the lags of ACF and PACF (Table: 5).

**Table 5:** Box-Ljung and Box-Pierce test statistics

| Test | $X^2$ | *df* | *p*-value |
|---|---|---|---|
| Box-Ljung | 12.198 | 19 | 0.877 |
| Box-Pierce | 10.034 | 19 | 0.952 |

The large *p*-values in both of the tests are suggesting to accept the null hypothesis that all of the autocorrelation functions between the lags (1 to 20) are zero. Thus, it can be concluded that there is no any evidence (or almost nil) for non-zero autocorrelations in the forecast errors between lags (1 to 20) in the fitted ARIMA (1,1,0) model.

## 10. Conclusions

In the present study, the ARIMA (1,1,0) with drift was the best fitted model, then used for prediction upto 10 years of the production of wheat in India using the 68 years' time series data. ARIMA (1,1,0) was used because the reason of its capability to make prediction using the time series data with any kind of patterns and with auto correlated successive values of the time series. The study was also validated and statistically tested that the successive residuals in the fitted ARIMA (1,1,0) were not correlated, and the residuals appear to be normally distributed with the mean zero and constant variance. Hence, it can be concluded that the selected ARIMA (1,1,0) seems to provide a satisfactory predictive model for the wheat production in India for the period of 2017-18 to 2026-27.

The ARIMA (1,1,0) model projected an increment in the production for the duration of 10 years from 2017-18 to 2026-27. The prediction for 2026-27 is resulted approximately 110793 ’000 tonnes. Like any other predictive models for forecasting, ARIMA model has also limitations on accuracy of the predictions yet it is widely used for forecasting the future values for the time series.

## 11. References


1. Coghlan A. A Little Book of R for Time Series, Readthedocs.org, 2010. Available online at: *http://a-little-book-of-r-for-time-series.readthedocs. org/en/latest/src/timeseries.html*
2. Contreras J, Espinola R, Nogales FJ, Conejo AJ. ARIMA Models to Predict Next- day Electricity Prices, IEEE Transactions on Power Systems. 2003; 18(3):1014-1020.
3. Hannan E. The Estimation of the Order of ARMA Process. 1980; 8:1071-1081.
4. Jeffrey EJ. Evaluating Methods for Forecasting Earnings Per Share, Managerial Finance. 1990; 16 (3):30–35.
5. Kumar M, Anand M. An application of time series ARIMA forecasting model for predicting sugarcane production in India, Studies in Business and Economics. 2014; 9(1):81-94.
6. Ljung GM, Box GEP. On a measure of lack of fit in time series models, Biometrika. 1978; 67:297-303.
7. Meyler Aidan, Kenny Geoff, Quinn Terry. Forecasting Irish inflation using ARIMA models, Central Bank and Financial Services Authority of Ireland Technical Paper Series. 1998, 3/RT/.98-48.
8. Muhammad F, M Siddique, M Bashir, S Ahmad. Forecasting rice production in Pakistan–using ARIMA Models, J of Animal Plant Sci. 1992; 2:27-31.
9. Raymond YC Tse. An application of the ARIMA model to real-estate prices in Hong Kong, Journal of Property Finance. 1997; 8(2):152-163.
10. Stergiou KI. Modeling and forecasting the fishery for pilchard (Sardina pilchardus) in Greek waters using ARIMA time-series models, ICES Journal of Marine Science. 1989; 46(1):16-23.
11. Tripathi A, Mishra AK. The Wheat Sector in India: Production, Policies and Food Security, the Eurasian Wheat Belt and Food Security. 2017, 275-296.