



ISSN (E): 2277-7695

ISSN (P): 2349-8242

TPI 2025; 14(4): 49-55

© 2025 TPI

[www.thepharmajournal.com](http://www.thepharmajournal.com)

Received: 25-02-2025

Accepted: 30-03-2025

**Krishnaveni G**

Department of Agricultural  
Statistics, College of Agriculture,  
Indira Gandhi Krishi  
Viswavidyalaya, Raipur,  
Chhattisgarh, India

**Kuldeep Tandan**

Department of Agricultural  
Statistics, College of Agriculture,  
Indira Gandhi Krishi  
Viswavidyalaya, Raipur,  
Chhattisgarh, India

**Dr. ML Lakhera**

Department of Agricultural  
Statistics, College of Agriculture,  
Indira Gandhi Krishi  
Viswavidyalaya, Raipur,  
Chhattisgarh, India

**Dr. Sweta Ramole**

Department of Agricultural  
Statistics, College of Agriculture,  
Indira Gandhi Krishi  
Viswavidyalaya, Raipur,  
Chhattisgarh, India

## Forecasting of rice grain yield in long term fertilizer experiments (LTFE): An application of the Arima time series model

**Krishnaveni G, Kuldeep Tandan, ML Lakhera and Sweta Ramole**

### Abstract

Accurate forecasting of grain yield and straw yield of rice crop is essential for ensuring food security and sustainable agricultural planning. This study investigates the application of the Autoregressive Integrated Moving Average (ARIMA) model for predicting grain yield and straw yield of rice crop in long-term fertilizer experiments (LTFEs). The research utilizes historical yield data from LTFEs, incorporating different fertilizer treatments to analyze trends and patterns. The ARIMA model is employed to assess time-series data, identify optimal model parameters, and generate forecasts. Model performance is evaluated using statistical measures such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The analysis for rice grain yield and rice straw yield identified ARIMA (0,1,0) and ARIMA (2,2,0) as the optimal model specification respectively. By applying this model to the time series data, projections were made for the next five years with the objective of achieving the highest possible forecasting accuracy.

**Keywords:** ARIMA, rice yield forecasting, long-term fertilizer experiments, time series analysis

### 1. Introduction

Long-Term Fertilizer Experiments (LTFEs) are agricultural studies conducted over extended periods to evaluate the impact of continuous fertilizer application on soil health, crop productivity, and sustainability. These experiments help understand the long-term effects of different nutrient management practices, including chemical fertilizers, organic amendments, and integrated nutrient management. LTFEs are particularly important in staple crops like rice, wheat, and maize, as they provide valuable insights into nutrient-use efficiency and long-term soil health.

Reliable forecasting of crop production and yields supports agribusinesses and policymakers in supply chain planning and resource allocation. Statistical techniques help forecast these parameters, aiding decisions on food security, land use, and environmental concerns. Forecasting uses past data to predict future outcomes and plays a vital role in agriculture. Accurate forecasting is essential for long-term planning, especially for orchards and perennial crops, ensuring better resource use and maximizing profits.

This study aims to forecast the grain yield and straw yield of rice crop in long-term fertilizer experiments (LTFEs) for the next five years using the Autoregressive Integrated Moving Average (ARIMA) model. The ARIMA model, introduced by Box and Jenkins, is often considered superior for univariate time series with correlated lag variables, as it effectively captures the underlying data patterns and produces minimal forecast errors. However, ARIMA is most effective for short-term predictions, typically within a five-year horizon, relying solely on historical data. The ARIMA model was chosen for this study due to its ability to account for non-zero autocorrelation between consecutive values in the time series data, making it a reliable tool for trend prediction.

### 2. Review literature

Saeed *et al.* (2000) <sup>[11]</sup> used the Box-Jenkins ARIMA methodology to forecast wheat production in Pakistan, based on time series data from 1947-48 to 1988-89. Diagnostic checks indicated that the ARIMA (2,2,1) model was the most suitable. They forecasted wheat production for the period from 1999-2000 to 2012-13 with 95% confidence limits.

Sarda and Prajneshu (2002) <sup>[12]</sup> used the ARIMA time series method to forecast pesticide consumption in the country.

**Corresponding Author:****Krishnaveni G**

Department of Agricultural  
Statistics, College of Agriculture,  
Indira Gandhi Krishi  
Viswavidyalaya, Raipur,  
Chhattisgarh, India

They found that the ARIMA (2,1,0) model had the lowest goodness of fit criteria among all tested models, making it the best fit for the data. They also applied the Box-Ljung test to assess the independence of the model's errors.

Yaseen *et al.* (2005) <sup>[13]</sup> modeled and forecasted sugarcane yield in Pakistan using time series data from 1947 to 2002. They applied the ARIMA (2,1,2) model and predicted the yield up to 2008-09, with forecasted values for 1999-2000 to 2001-02 closely matching the actual yield.

Padhan (2012) <sup>[7]</sup> applied various ARIMA models to forecast the productivity of different crops in India. He fitted ARIMA models (1,0,1), (1,0,1), (2,1,2), (1,1,2), and (1,1,0) for crops such as gram, jowar, jute, pulses, and wheat, achieving adjusted  $R^2$  values of 0.17, 0.42, 0.78, 0.49, and 0.16, respectively.

Prabakaran *et al.* (2013) <sup>[8]</sup> applied ARIMA (1, 1, 1) and ARIMA (1, 1, 0) models to forecast wheat area and production in India from 1950-51 to 2011-12. The models predicted a wheat area of 31.46 thousand hectares for 2015, with a range of 34.25 to 31.46 thousand hectares, and a production of 97.73 thousand tonnes, with a range of 107.55 to 87.92 thousand tonnes.

Prabakaran and Sivapragasam (2014) <sup>[9]</sup> analyzed rice area and production data from 1950-51 to 2011-12 using time series methods, including ACF and PACF. They applied an ARIMA (1, 1, 1) model to forecast rice area and production for 2015, estimating the area at 44.75 thousand hectares and production at 104.37 thousand tonnes, with confidence intervals.

Jadhav *et al.* (2017) <sup>[4]</sup> demonstrated the effectiveness of farm price forecasting for crops like Paddy, Ragi, and Maize in Karnataka using time series data from 2002 to 2016. They applied univariate ARIMA models to predict prices and assessed accuracy using MSE, MAPE, and Theil's U coefficient. The results showed that the ARIMA model provided valid price forecasts with lower MSE, MAPE, and Theil's U values.

Hemavathi and Prabakaran (2018) <sup>[3]</sup> applied the ARIMA model to forecast the area, productivity, and growth trends of rice in Tamil Nadu's Thanjavur district using data from 1991 to 2015. Their predictions indicated a decline in rice cultivation area, productivity, and production over the next four years.

Bharati and Anil Kumar Singh (2019) <sup>[1]</sup> used the Box-Jenkins ARIMA model to predict India's rice production from 1950-51 to 2017-18. The ARIMA (0, 1, 1) model was found to be effective, with prediction errors below 3% for the years 2015-16 to 2017-18. The model's performance was evaluated using various statistical criteria, and forecasts for rice production were made up to 2099.

Praveen and Sharma (2019) <sup>[10]</sup> analyzed 50 years of data (1967-2016) across India, revealing that rising annual temperatures typically reduce land productivity for most crops, posing a food security risk. Their ARIMA forecast for the next 20 years suggests that higher temperatures and rainfall may boost yields of some crops like gram and sugarcane. However, climate-sensitive crops such as wheat, rice, and cotton could experience fluctuating production with increasing temperatures.

Delvadiya *et al.* (2023) <sup>[2]</sup> analyzed trends in groundnut area, production, and productivity in Gujarat from 1991-92 to 2019-20 using data from the Directorate of Agriculture. They found that a cubic model with a five-year moving average

best forecasted area, while a linear model was ideal for production and productivity. The ARIMA (2,1,0) model effectively forecasted area trends but was less suitable for production and productivity.

Mishra *et al.* (2024) <sup>[6]</sup> explored the challenges of estimating potato production for sustainable agricultural practices using historical data from agricultural markets in India, China, and the USA. They found that the ETS model outperformed ARIMA in predicting potato production. Their predictions indicated that China, India, and the USA would contribute 100,417, 61,882, and 18,229 thousand tonnes of potato production, respectively.

Ketali *et al.* (2024) <sup>[5]</sup> analyzed data from 1950-51 to 2017-18 for major pulse crops in India, including Bengal gram, Red gram, Green gram, and Black gram, using the Box-Jenkins ARIMA method to forecast production from 2018-19 to 2030-31. They found that production of all pulse crops showed an increasing trend. The best predictions were made with ARIMA models: (0,1,1) for Bengal gram (1.03 Mt), (2,1,3) for Red gram (0.20 Mt), (2,1,3) for Green gram (-1.47 Mt), and (2,1,2) for Black gram (-0.51 Mt).

### 3. Materials and Methods

The data for this study were obtained from the records of long-term fertilizer experiments conducted under the AICRP project by the Department of Soil Science, College of Agriculture, Indira Gandhi Krishi Vishwavidyalaya, Raipur. The dataset includes grain and straw yield data for rice crop over 20 years, from 2003-04 to 2022-23. They used the Mahamaya variety from 2003-04 to 2017-18 and the Rajeswari variety from 2019-20 to till date. The data was recorded on a treatment-wise and replication-wise basis after harvesting at maturity. Long term fertilizer experiments data, such as yield information, may not follow a normal distribution and could be skewed by outliers or extreme values. Since the input factors in our long-term fertilizer experiments have changed, this violates the standard rules for conducting long-term studies. However, it is understandable that older crop varieties may become obsolete or unavailable as they are replaced by newer ones. Changing the variety affects the assumption of normality in the data. Therefore, before applying the ARIMA model, the data is converted to the sustainable transformation. Transforming the data can help achieve a more normal distribution, facilitating the application of statistical tests and models that rely on this assumption. The transformed data was used in this study to analyze trends in the yield of the rice crop within the context of the long-term fertilizer experiments. Typically, long-term fertilizer experiments maintain consistent treatments and inputs for crop cultivation. In this study, the basic treatments remained unchanged throughout the period under consideration.

#### 3.1. ARIMA models for predicting yield trends

The Box and Jenkins model (1976) will be utilized for yield forecasting, with the Auto-Regressive Integrated Moving Average (ARIMA) serving as the fundamental category of models for time series forecasting. In these models, different series appearing in the forecasting equations represent the "Auto-Regressive" process. Meanwhile, the inclusion of lagged forecast errors in the model characterizes the "Moving Average" process. The ARIMA model is represented as ARIMA (p, d, q), where 'p' denotes the order of the auto-

regressive process, 'd' indicates the level of differencing required to achieve stationarity, and 'q' represents the order of the moving average process.

Auto Regressive process of order (p) is

$$Y_t = C + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_q Y_{t-q} + \epsilon_t$$

Where

$Y_t$  is the current observation,

C is a constant,

$\epsilon_t$  is the error at time t, and

$\phi_1$  to  $\phi_p$  are the moving average parameters

This is similar to multiple regression, but it uses lagged values of  $Y_t$  as predictors. This type of model is called an AR(p) model, meaning an autoregressive model of order p.

Moving Average process of order (q) is

$$Y_t = C + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q}$$

Where

$Y_t$  is the current observation,

C is a constant,

$\epsilon_t$  is the error at time t, and

$\theta_1$  to  $\theta_q$  are the moving average parameters.

The basic formulation of ARIMA (p,d,q) could be described as,

$$Y'_t = C + \phi_1 Y'_{t-1} + \phi_2 Y'_{t-2} + \dots + \phi_p Y'_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q}$$

Where

$Y'_t$  is the differenced and stationary time series at time t.

C is a constant or mean of the differenced series.

$\phi_1, \phi_2, \dots, \phi_p$  are autoregressive parameters representing the dependence on past values.

$\epsilon_t$  is the white noise error term at time t.

$\theta_1, \theta_2, \dots, \theta_q$  are MA parameters represents dependence on past forecast errors.

## 3.2. The Box-Jenkins modelling procedure

Instead of relying on traditional econometric methods, the Box-Jenkins approach is preferred for forecasting due to its mathematical robustness and reliability. This method follows a structured sequence of stages within the ARIMA modeling process to develop an effective model. The constructed models are then evaluated for accuracy using historical data. A well-fitted model is characterized by minimal residuals that contain little useful information and are randomly distributed. If the model does not perform satisfactorily, the entire process is repeated to refine the initial model using newly available data. This iterative procedure continues until the best-fitting model is identified. The key stages in developing an ARIMA forecasting model include:

1. Model specification
2. Model estimation
3. Diagnostic checking
4. Forecasting

### 3.2.1. Model specification

The primary objective of ARIMA modeling is to determine the most appropriate values for p, d, and q. This can be partially addressed by examining the Auto-Correlation

Function (ACF) and Partial Auto-Correlation Function (PACF) of the time series data (Pindyck & Rubinfeld, 1991). The ACF helps identify the order of the moving average component (q), while the PACF provides insight into the order of the autoregressive component (p). The first step is to assess whether the data is stationary. The degree of differencing required to achieve stationarity, denoted as d, is determined based on the point at which the ACF approaches zero. Once d is established, the stationary series is further analyzed using ACF and PACF to select appropriate values for p and q.

### 3.2.2. Model estimation

The next step involves estimating the model using a statistical software package. The objective is to obtain parameter estimates for the initially proposed ARIMA model based on the selected values of p and q. The ARIMA coefficients ( $\phi$  and  $\theta$ ) are determined using a nonlinear least squares approach. One of the most widely used techniques for estimating ARIMA models is known as "Marquardt's compromise."

### 3.2.3. Diagnostic checking

Diagnostic checks are essential for evaluating the accuracy of a time series model. The first diagnostic check involves residual analysis, where a time series plot of residuals is examined. If the residuals exhibit a random scatter around the zero line without any noticeable trend or pattern, the model is considered appropriate. The second diagnostic check assesses the normality of residuals. The first normality test involves plotting normal scores against residuals; if the points align along a straight line, the model is deemed a good fit. Additionally, a histogram of residuals is analyzed to confirm their normal distribution. The third diagnostic check evaluates the model's goodness of fit by plotting residuals against the corresponding fitted values. If no discernible pattern emerges in this plot, the model is considered well-fitted to the time series data.

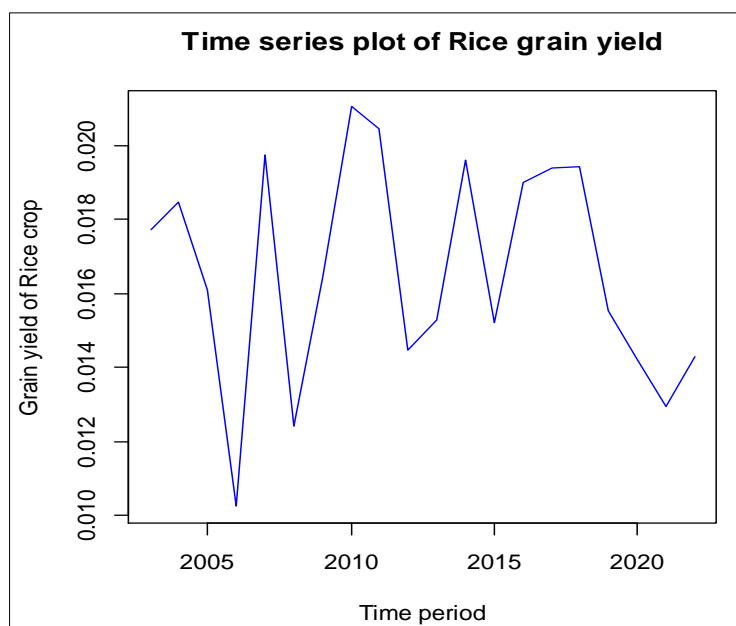
### 3.2.4. Forecasting

After evaluating the predictive performance of the fitted ARIMA model, along with the 95% confidence interval estimates, a forecast is conducted for a period of up to five years. A longer forecast horizon is avoided, as prediction errors tend to increase significantly when projecting too far into the future.

The statistical analysis for this study was conducted using "R: The R Project for Statistical Computing". R is a widely utilized open-source programming language, renowned for its capabilities in statistical analysis, data manipulation, and machine learning. The open-source statistical software 'R' (version 4.2.2) was employed, along with various statistical and time series packages, including 'tseries' (version 0.10-54) and 'forecast' to facilitate model development and forecasting.

## 4. Results and discussion

To understand the data's behavior, time series graphs were generated. A time series analysis was conducted on rice grain yield from 2003 to 2022. Trends and variations over time were examined using statistical visualization. This approach helps identify patterns, fluctuations, and potential influencing factors. The Figure 1 below represents the line plot of rice grain yield recorded between 2003 and 2022.



**Fig 1:** Rice grain yield recorded between 2003 and 2022

To build ARIMA model, the preliminary steps followed were Unit root test, identification of the parameter, then estimation of parameters and at last diagnostic checking of the model was done. Then the best fitted ARIMA model is used for forecasting of data for next 5 years.

#### 4.1. Unit root test (stationarity test)

The stationarity of the time series for rice grain yield in LTFE

was evaluated using the Augmented Dickey-Fuller (ADF) test, with the results presented in Table 1. The test indicated that the original series was non-stationary, as the p-value exceeded 0.05. To achieve stationarity, the series was first differenced ( $d=1$ ), resulting in a significant ADF test outcome with a p-value below 0.05. Consequently, this differenced series was utilized for developing the ARIMA model to analyze rice grain yield.

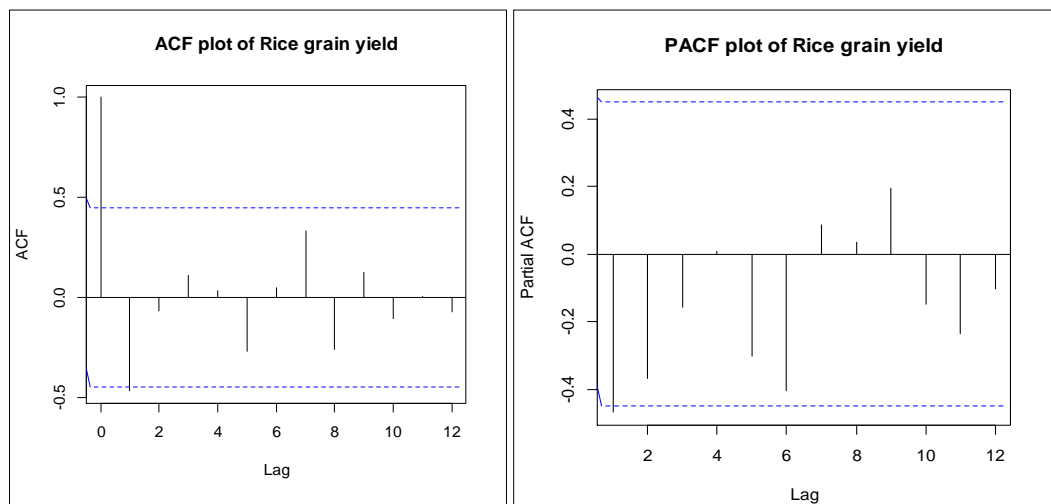
**Table 1:** Unit root test for assessing stationarity of rice grain yield in LTFE

	Original values		1 <sup>st</sup> Differencing	
	ADF value	P value	ADF value	P value
Grain yield	-1.9228	0.6018	-4.1505	0.01801

#### 4.2. Parameter identification

The differencing parameter 'd' was predetermined, and for rice grain yield in LTFE, it was set to  $d=1$ . The autoregressive (p) and moving average (q) parameters were identified using the PACF and ACF of the first-differenced stationary series respectively. The ACF and PACF plots for rice grain yield, presented in Figure 2, were used to determine the appropriate

values of 'q' and 'p'. The ACF revealed a significant lag exceeding the standard error limit, while the PACF indicated that a specific number of lags also surpassed this threshold. Based on these observations, the values of 'p' and 'q' for rice grain yield in LTFE were determined, facilitating the selection of the optimal ARIMA model using the insights gained from the first-differenced series.



**Fig 2:** ACF and PACF plots of first differenced series



#### 4.3. Estimation of parameter

The Augmented Dickey-Fuller (ADF) test confirmed that the series is auto-correlated (Table 1). After first differencing the series, it became stationary, and the autocorrelation was removed. All the models were estimated based on the identified values of p, d, and q for rice grain yield and the results are presented in Table 2. The best-fitting ARIMA model for the rice grain yield was ARIMA (0,1,0) as indicated by the lowest AIC and BIC values (Table 3). Eighty percent of the data was used for model identification, after which the parameters were estimated using the maximum likelihood method. The Summary of the best fitted model for rice grain yield is presented in table 3 which includes log-likelihood, AIC, BIC and AICc values. The model's performance was evaluated using rest of twenty percent of holdout data, with error metrics such as MAPE, MAE, RMSE, and MASE (Table 4). Therefore, the ARIMA (0, 1, 0) model appears to be the best model for forecasting the future values of rice grain yield in LTFE. However, the model still needs to be validated through diagnostic checks of the residuals.

**Table 2:** Preliminary Models

MODEL	AIC
ARIMA(2,1,2) with drift	Inf
ARIMA(0,1,0) with drift	-149.504
ARIMA(1,1,0) with drift	-151.179
ARIMA(0,1,1) with drift	Inf
ARIMA(0,1,0)	-151.983
ARIMA(1,1,1) with drift	Inf

#### Best model: ARIMA (0,1,0)

**Table 3:** Summary of the best fitted model for rice grain yield

Model	Log-likelihood	AIC	BIC	AICc
ARIMA(0,1,0)	77.11	-152.22	-151.27	-151.98

**Table 4:** The values of MAPE, MAE, MASE and RMSE of best fitted model for rice grain yield

MODEL	MAPE	MAE	MASE	RMSE
ARIMA(0,1,0)	20.32029	0.003132886	0.9502687	0.004074691

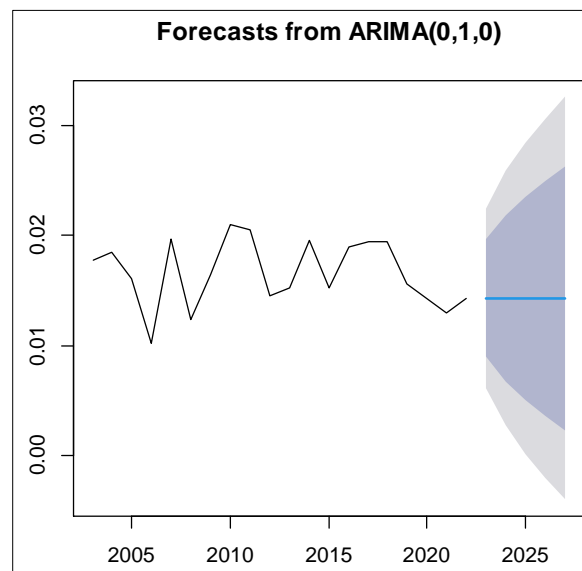
#### 4.4. Forecasting

The table 5 presents forecasted values from an ARIMA model for the years 2023 to 2027, along with 80% and 95% confidence intervals. The forecasted values for the years 2023 to 2027 show a stable trend, maintaining a constant forecast value of 0.01428 kg/ha across all years. Despite this stability, the confidence intervals widen progressively over time, indicating increasing uncertainty in long-term projections. The 80% confidence intervals reveal a gradual expansion, starting from 0.00892 kg/ha to 0.01964 kg/ha in 2023 and widening to 0.0023 kg/ha to 0.02626 kg/ha by 2027. This increase in range suggests that while the forecast remains unchanged, the variability in potential outcomes grows over time. Similarly, the 95% confidence intervals also demonstrate increasing uncertainty. In 2023, the lower bound is 0.00609 kg/ha and the upper bound is 0.02247 kg/ha, whereas by 2027, the range extends from 0.00404 kg/ha to 0.03260 kg/ha. This broadening range underscores the potential variability in the forecasted parameter and highlights the need for cautious interpretation of long-term predictions. Figure 3 shows the five-year forecasted data of rice grain yield, generated by applying the ARIMA (0,1,0) model to the

time series data.

**Table 5:** Prediction of rice grain yield for next 5 years from ARIMA (0, 1, 0) model (kg/ha)

Year	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2023	0.01428	0.008922	0.019638	6.09E-02	0.022474
2024	0.01428	0.006703	0.021857	2.69E-02	0.025868
2025	0.01428	0.005	0.02356	8.81E-02	0.028472
2026	0.01428	0.003565	0.024995	2.11E-02	0.030667
2027	0.01428	0.0023	0.02626	4.04E-02	0.032602



**Fig 3:** Forecast graph from ARIMA (0,1,0)

Table 6 illustrates the predicted rice grain yield over a five-year period, based on the ARIMA (0,1,0) model. This table shows yield data after converting the transformed forecasted data which shown in table 5 into original yield data. The forecasted values for the years 2023 to 2027 remain constant at 1204.95 kg/ha, suggesting a stable trend over the forecast period. This indicates that the projected variable is expected to experience minimal fluctuations, maintaining a steady level across the years. The 80% confidence intervals show a relatively narrow range, indicating high confidence in the stability of the forecast. In 2023, the 80% confidence interval spans from 1177.74 kg/ha to 1232.17 kg/ha, and by 2027, it slightly widens to a range of 1144.10 kg/ha to 1265.81 kg/ha. The gradual expansion of the interval suggests a small increase in uncertainty over time, but overall, the values remain within a controlled range. The 95% confidence intervals, which provide a broader range of potential outcomes, also indicate a stable forecast with minor variations. The lower bound starts at 1163.33 kg/ha in 2023 and slightly decreases to 1111.89 kg/ha in 2027, while the upper bound increases from 1246.58 kg/ha to 1298.02 kg/ha over the same period.

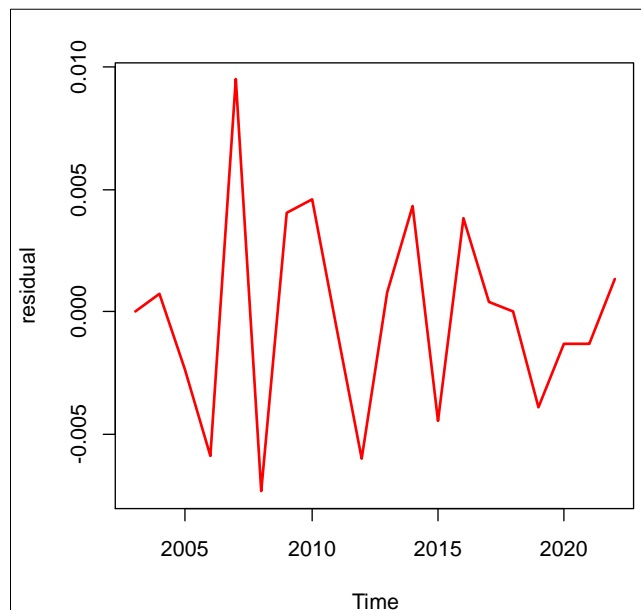
**Table 6:** Five-Year Forecast of Rice Grain Yield Using ARIMA (0,1,0) Model

Year	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2023	1204.954	1177.739	1232.168	1163.332	1246.575
2024	1204.954	1166.466	1243.441	1146.092	1263.815
2025	1204.954	1157.816	1252.091	1132.863	1277.044
2026	1204.954	1150.524	1259.383	1121.71	1288.197
2027	1204.954	1144.099	1265.808	1111.885	1298.022

#### 4.5. Diagnostic check of residuals

The figure 4 shown is a residual plot for the ARIMA (0,1,0) model, which represents the differences between the observed and predicted values of grain yield over time. The residuals appear to fluctuate randomly around zero, indicating that the ARIMA (0,1,0) model effectively captures the underlying pattern in the data without systematic bias. However, the

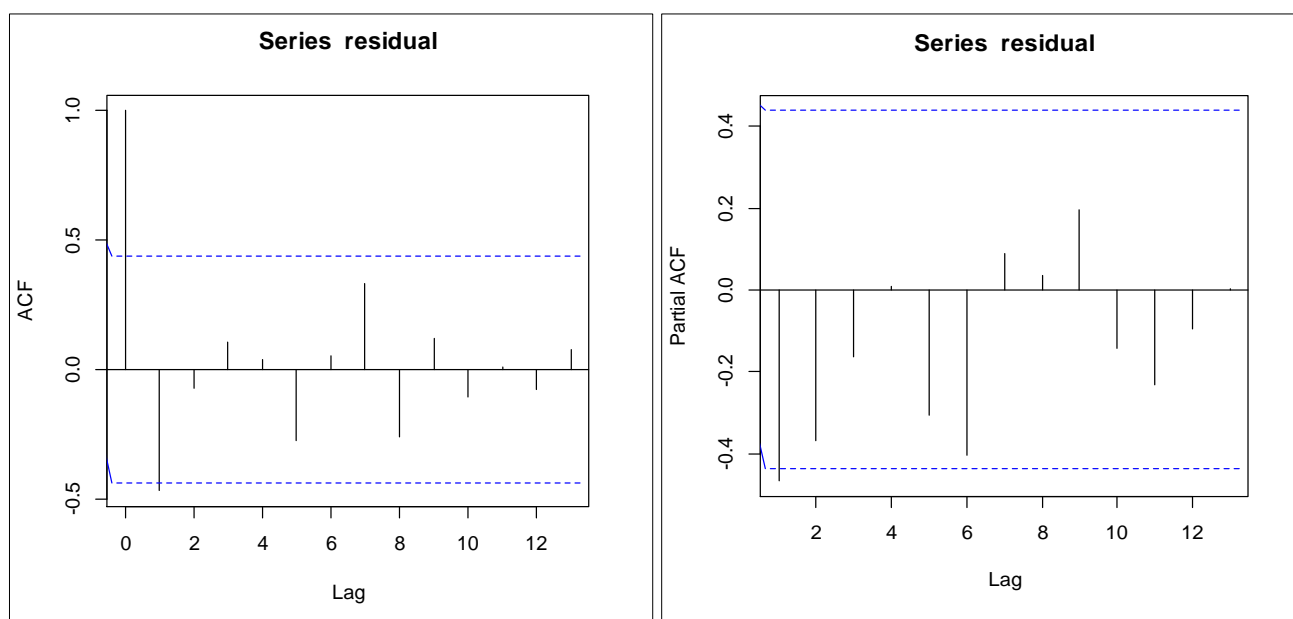
presence of some large spikes implies that certain unexpected factors or outliers may have influenced the yield, which the model could not fully account for. Overall, the residual plot supports the validity of the ARIMA (0,1,0) model for forecasting rice grain yield, with random and normally distributed errors indicating a well-fitted model.



**Fig 4:** Residual plot

ACF and PACF plots of residuals are shown in figure 5. The ACF plot of the residuals from the ARIMA model indicates that the model provides a good fit to the data. In the ACF plot, most of the autocorrelation values are within the 95% confidence limits, suggesting that the residuals are free from significant autocorrelation. The PACF plot of the residuals from the ARIMA model indicates that the model effectively captures the underlying structure of the data. In the plot, all partial autocorrelation values fall within the 95% confidence

limits, suggesting that there are no significant lagged dependencies left unexplained by the model. This implies that the residuals behave like white noise, indicating that the model has successfully captured the temporal dependencies without overfitting or underfitting. Therefore, the ARIMA model can be considered well-specified and suitable for forecasting. This implies that the ARIMA model has effectively captured the underlying structure of the time series data.



**Fig 5:** ACF and PACF of residual plots

## 5. Summary and conclusion

The ARIMA (0,1,0) model was chosen as the best model for producing forecasts for up to five years for the rice grain yield in LTFE utilizing a 20-year time series data in this study. ARIMA was chosen because of its ability to anticipate utilizing time series data with autocorrelations between successive values in the time series and with any type of pattern. The study also found that the consecutive residuals in the fitted ARIMA time series were not associated, and that the residuals were not normally distributed. As a result, we may infer that the ARIMA (0,1,0) model chosen appears to be an appropriate forecasting model for the rice grain yield in LTFE. Despite the fact that ARIMA, like any other predictive model in forecasting, has limits in terms of prediction accuracy, it is nevertheless commonly employed for projecting future consecutive values in time series

## Acknowledgement

We would like to express our sincere gratitude to the developers of the open-source R programming libraries which were instrumental in implementing and evaluating the machine learning models. Special thanks to our colleagues for their valuable feedback and suggestions during the preparation of this manuscript. Finally, we acknowledge the support of our institution for providing the necessary resources and infrastructure to carry out this research.

## References

1. Bharati RC, Singh AK. Predicting rice production using autoregressive integrated moving average model. *Journal of AgriSearch*. 2019;6(4):205-210.
2. Delvadiya JB, Chaudhar GK, Marviya PB. Evaluation of statistical models for forecasting area, production and productivity of groundnut in Gujarat. *Journal of Oilseeds Research*. 2023;40(3):143-150.
3. Hemavathi M, Prabakaran K. ARIMA model for forecasting of area, production and productivity of rice and its growth status in Thanjavur District of Tamil Nadu, India. *International Journal of Current Microbiology and Applied Sciences*. 2018;7(2):149-156.
4. Jadhav V, Reddy BVC, Gaddi GM. Application of ARIMA model for forecasting agricultural prices. *Journal of Agricultural Science and Technology*. 2017;19:981-992.
5. Ketali SS, Swaminathan B, Aiswarya S. Towards Atmanirbharta (self-reliance) in the production of pulse crops in India: A situational analysis of future demand and supply. *International Journal of Humanities Social Science and Management (IJHSSM)*. 2024;4(1):241-254.
6. Mishra P, Alhussan AA, Khafaga DS, Lal P, Ray S, Abotaleb M, *et al.* Forecasting production of potato for a sustainable future: Global market analysis. *Potato Research*. 2024;1-20.
7. Padhan PC. Applications of ARIMA model for forecasting agricultural productivity in India. *Journal of Agriculture and Social Sciences*. 2012;8:50-56.
8. Prabakaran K, Sivapragasam C, Jeevapriya C, Narmatha A. Forecasting cultivated areas and production of wheat in India using ARIMA model. *Golden Research Thoughts*. 2013;3(3):1-7.
9. Prabakaran K, Sivapragasam C. Forecasting areas and production of rice in India using ARIMA model. *International Journal of Farm Sciences*. 2014;4(1):99-106.
10. Praveen B, Sharma P. Climate variability and its impacts on agriculture production and future prediction using autoregressive integrated moving average method (ARIMA). *Journal of Public Affairs*. 2019;e2016. <https://doi.org/10.1002/pa.2016>
11. Saeed N, Saeed A, Zakria M, Bajwa TM. Forecasting of wheat production in Pakistan using ARIMA models. *International Journal of Agriculture and Biology*. 2000;2(4):352-353.
12. Sarda C, Prajneshu. Modelling and forecasting country's pesticide consumption data using ARIMA time-series approach. *Annals of Agricultural Research New Series*. 2002;23(4):719-722.
13. Yaseen M, Zakria M, Shahzad IUD, Khan MI, Javed MA. Modelling and forecasting the sugarcane yield of Pakistan. *International Journal of Agriculture and Biology*. 2005;7(2):180-183.