P-ISSN: 2618-0723
E-ISSN: 2618-0731

NAAS Rating: 5.04
www.extensionjournal.com

International Journal of Agriculture Extension and Social Development

Volume 7; SP-Issue 2; Feb 2024; Page No. 59-63

Received: 02-01-2024
Accepted: 06-02-2024

Indexed Journal
Peer Reviewed Journal

ARIMA model for forecasting of maize prices in Telangana state by using SAS

[1]M Maheshnath, [2]R Vijaya Kumari, [3]K Suhasini, [4]D Srinivasa Reddy and [5]A Meena

[1]Ph.D. Research Scholar, Department of Agricultural Economics, College of Agriculture, Professor Jayashankar Telangana State Agricultural University, Hyderabad, Telangana, India

[2]Deputy Director of Research, Professor Jayashankar Telangana State Agricultural University, Hyderabad, Telangana, India

[3]Sr. Professor & University Head, Department of Agricultural Economics, College of Agriculture, Rajendranagar, Hyderabad, Telangana, India

[4]Field Officer, Cost of Cultivation Scheme, Professor Jayashankar Telangana State Agricultural University, Hyderabad, Telangana, India

[5]Assistant Professor, Department of Statistics & Mathematics, College of Agriculture, Rajendranagar, Hyderabad, Telangana, India

DOI: https://doi.org/10.33545/26180723.2024.v7.i2Sa.375

Corresponding Author: M Maheshnath

Abstract

This study employs the Autoregressive Integrated Moving Average (ARIMA) approach to model and predict maize prices in Telangana State. The Autocorrelation (AC) and Partial Autocorrelation (PAC) functions are calculated to identify and construct suitable ARIMA models for explaining the time series and forecasting future production. Evaluation of forecasting performance is conducted using Akaike's Information Criterion (AIC) and Schwarz's Bayesian Information Criterion (BIC). The best-fitting model is then utilized for out-of-sample forecasting up to December 2023.

Keywords: Autoregressive integrated moving average (ARIMA), maize prices, forecasting

Introduction

Maize (*Zea mays* L.) referred to as the Queen of Cereals, is a vital crop in India, standing as the third cash crop after wheat and rice. With 15 million Indian farmers engaged in maize cultivation, states like Karnataka, Rajasthan, Madhya Pradesh, and Telangana contribute significantly to the country's maize production. To ensure remunerative prices for farmers, India must plan production by enhancing productivity and reorienting the value chain (IIMR, 2022-23).

In Telangana, maize ranks third among all crops, covering an extensive area of 12.74 lakh acres. The maize production in Telangana reached 28.65 lakh tonnes during 2022-23 (DES, 2022-23). Major maize-growing districts in Telangana include Warangal Rural, Khammam, Nirmal, Siddipet, Kamareddy, Mahabubabad, Nizamabad, Warangal Urban, Jagityal and Karimnagar. Over the last decade, both the area and production of maize have witnessed significant growth in the state (TS agriculture, 2022-23).

Survey of Literature

Stergiou (1989) [4] employed the ARIMA model technique to analyze a 17-year time series dataset of monthly pilchard catches from Greek waters. The aim was to forecast up to 12 months ahead, with forecasts compared against actual data from 1981, which was excluded from parameter estimation. Results indicated a mean error of 14%, suggesting the ARIMA methods capability in predicting the intricate dynamics of the Greek pilchard fishery, notorious for its unpredictable changes due to varying oceanographic and biological conditions.

Meyler *et al.* (1998) [2] established a framework for ARIMA time series models focusing on forecasting Irish inflation. Their research prioritized optimizing forecast performance by minimizing out-of-sample forecast errors rather than solely emphasizing in-sample 'goodness of fit'.

Prajneshu and Venugopalan (1998) [3] explored various parametric statistical modeling techniques, including the ARIMA model, polynomial function fitting, and nonlinear mechanistic growth modeling, to describe trends in marine fish production data in their country.

Contreras *et al.* (2003) [1] proposed an ARIMA methodology to predict next-day electricity prices for both spot markets and long-term contracts in mainland Spain and Californian markets.

Singh *et al.* (2007) [5] applied statistical models to forecast rice production in India, noting one of the advantages of the ARIMA approach is its ability to provide a comprehensive understanding of the system. This model has been the cornerstone of time series analysis for several decades.

In this study, the ARIMA model was utilized to model and forecast rice production in India. The trend of production over the past three decades was examined.

Materials and Methods

Data for maize prices were sourced from official records of the Directorate of Economics and Statistics, Government of Telangana and relevant websites. The ARIMA model, an extension of ARMA models, involves three stages: identification, estimation and diagnostic checking. The methodology includes differencing to make the series stationary and parameters are estimated using nonlinear least squares. Model adequacy is assessed through Akaike's Information Criterion (AIC) and Schwarz's Bayesian Information Criterion (BIC).

The Box-Jenkins methodology guides the identification, fitting, and checking of ARIMA models. The process involves estimating parameters for Autoregressive (AR), Moving Average (MA), and Integrated (I) components. Stationarity is achieved through differencing, and the series is prepared for model selection.

A fundamental illustration of a non-stationary process that transforms into a stationary one through differencing is the Random Walk. A series {yt} is considered to adhere to an Integrated ARMA model, abbreviated as ARIMA (p, d, q).

The ARIMA methodology progresses through three key stages: identification, estimation, and diagnostic checking. Initially, parameters of the tentatively chosen ARIMA model are identified, followed by estimation of these parameters in the subsequent stage. The adequacy of the selected model is then tested in the diagnostic checking stage. If the model is deemed inadequate, the process iterates until a satisfactory ARIMA model is selected for the time-series in question. Box *et al.* (2007) [6] provide a comprehensive discussion of various aspects of this approach. Additionally, most standard software packages such as SAS, SPSS, and E views include programs designed for fitting ARIMA models.

The Box-Jenkins methodology encompasses a series of procedures aimed at identifying, fitting, and verifying ARIMA models using time series data. Forecasts are derived directly from the structure of the fitted model.

A pth-order autoregressive model: AR (p), which has the general form.

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + \varepsilon_t$$

Where

Y_t= Response (dependent) variable at time t.

$Y_{t-1}, Y_{t-2}, \cdots, Y_{t-p}$ = Response variable at time lags t - 1, t - 2, ., t -p, respectively.

$\varphi_0, \varphi_1, \varphi_2, \cdots, \varphi_p$ = Coefficients to be estimated.

ε_t = Error term at time t.

A qth-order moving average model: MA (q), which has the general form.

$$Y_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \cdots - \theta_q\varepsilon_{t-q}$$

Where

Y_t = Response (dependent) variable at time t.

μ = Constant mean of the process.

$\theta_1, \theta_2, \cdots, \theta_q$ = Coefficients to be estimated.

ε_t = Error term at time t.

$\varepsilon_{t-1}, \varepsilon_{t-2}, \cdots, \varepsilon_{t-q}$ = Errors from preceding time periods that are integrated into the response.

Y_t

Autoregressive Moving Average Model: ARMA (p, q), which has the general form.

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \cdots - \theta_q\varepsilon_{t-q}$$

The autoregressive integrated moving average model, represented as ARIMA (p, d, q), where p denotes the autoregressive component's order, d signifies the degree of differencing, and q represents the moving average component's order. When the original series is stationary, d equals 0, and the ARIMA models simplify to ARMA models.

The difference linear operator (Δ), defined by

$$Y_t = Y_t - Y_{t-1} = Y_t - BY_t = (1 - B)\, Y_t$$

The stationary series W_t obtained as the dth difference (Δ^d) of Y_t,

$$W_t = {}^dY_t = (1 - B)^d Y_t$$

ARIMA (p, d, q) has the general form

$\varphi_p(B)\,(1 - B)^d Y_t = \mu + \theta_q(B)\varepsilon_t$ or

$\varphi_p(B)\, W_t = \mu + \theta_q(B)\varepsilon_t$

After achieving stationarity in a series, proceed to determine the model's structure.

Estimation of Parameters

Parameter estimation for the ARIMA model is carried out using the Nonlinear Least Squares method. Software packages like SAS are employed for fitting ARIMA models. AIC and BIC values aid in selecting the model with the best fit, and the model with the smallest AIC or BIC values is preferred.

Akaike's information criterion (**AIC**)

$$AIC = \log \hat{\sigma}^2 + 2\frac{p+q}{n}$$

Where $\hat{\sigma}^2$ is the estimated variance of e_t.

Schwarz's Bayesian Information criterion (SC, BIC, or **SBC**)

$$BIC = \log \hat{\sigma}^2 + 2\frac{p+q}{n}\log(n)$$

ACF and PACF plots of the residuals

After selecting an ARIMA model, a residual analysis is conducted to ensure the residuals resemble white noise. Significance checks of residual autocorrelations are performed, comparing them with two standard error bounds ($\pm 2/\sqrt{n}$). The adequacy of the chosen model is crucial for accurate forecasting.

Data Preparation and Model Selection
To apply the ARIMA model, the input time series must be stationary. The identification phase determines the necessary differencing to achieve stationarity and identifies non-seasonal AR and MA orders. Differencing and transformation are applied to remove trends and stabilize variance. The selected model is then used for prediction, considering both non-seasonal and seasonal components.

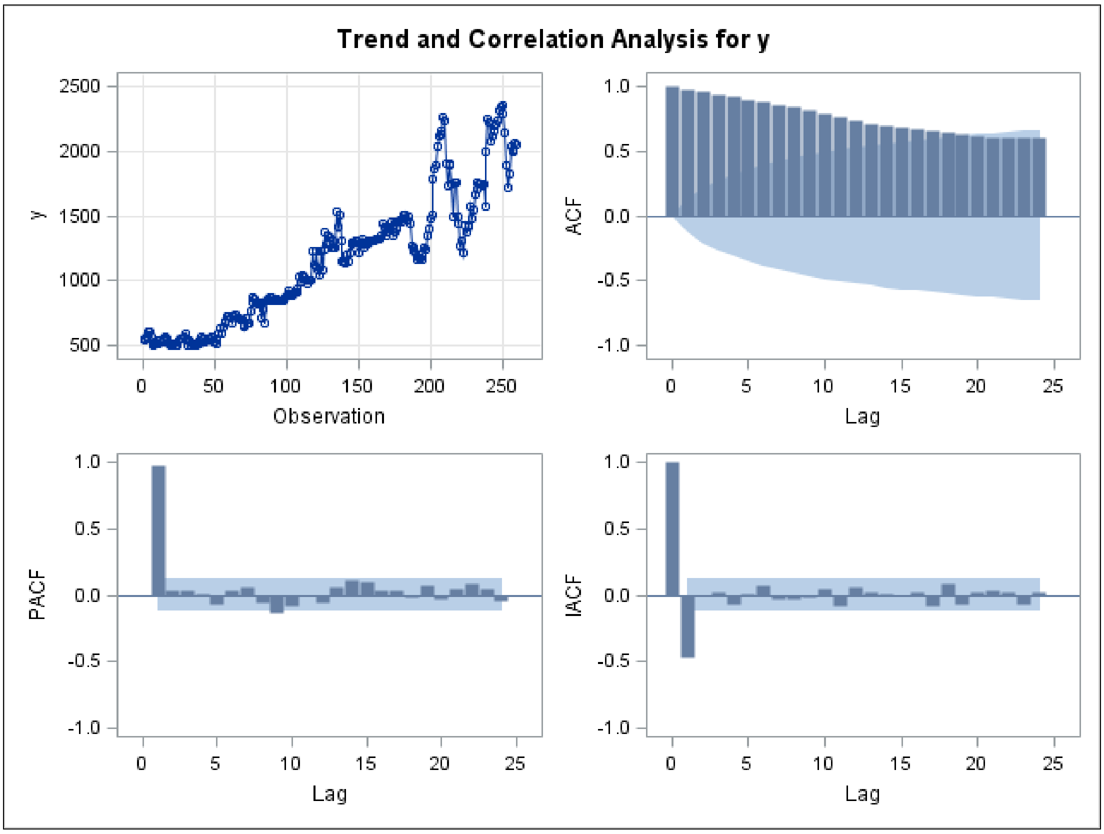

Fig 1: Line plot of the original series maize price data

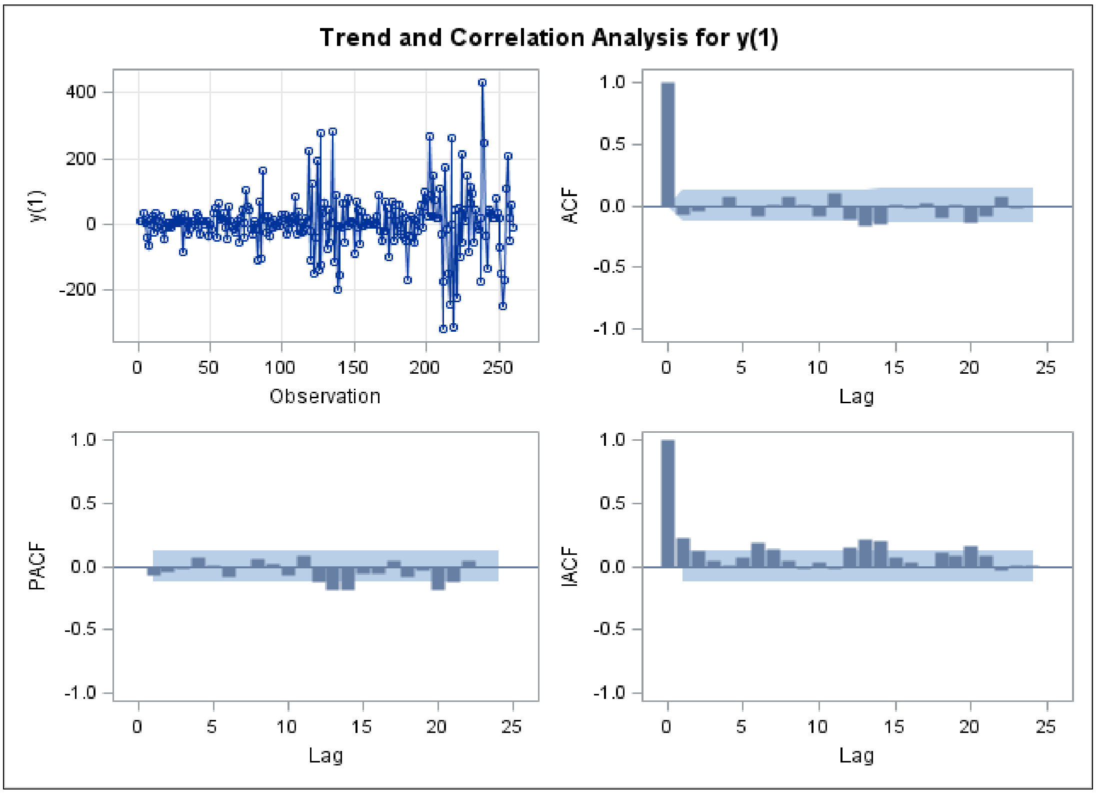

Fig 2: Line plot of the first order differenced maize price data

Estimation and Verification

Once the appropriate ARIMA (p, d, q) structure is identified, the subsequent steps involve parameter estimation and testing. During the estimation stage, data is utilized to estimate and draw inferences about the parameters of the tentatively identified model. These parameters are estimated to minimize the overall measure of residuals. The final stage of model development is the testing or diagnostic check of model adequacy. This step aims to ascertain whether residuals exhibit independence, homoscedasticity, and a normal distribution. Diagnostic statistics and residual plots are employed to assess the goodness of fit. Following the tentative model identification, the process is reiterated with the stage of parameter estimation and model verification. Diagnostic information may suggest alternative model (s). The series is now stationary, and various models are chosen based on their reliability in prediction. After evaluating autocorrelation functions (ACF) and partial autocorrelation functions (PACF), and considering minimum AIC and BIC values, as well as the significance of AR and MA parameters, the ARIMA (212) model is selected. Parameter estimates, along with corresponding standard errors, for the fitted ARIMA (212) model are presented in Table 2.

Table 1: AIC and BIC values for different ARIMA models

Model	AIC	BIC
ARIMA (000)	3975.55	3979.11
ARIMA (111)	3047.66	3058.33
ARIMA (011)	3045.95	3058.34
ARIMA (110)	3046.10	3053.21
ARIMA (210)	3047.47	3058.14
ARIMA (012)	3047.52	3058.19
ARIMA (211)	3049.45	3063.68
ARIMA (112)	3049.52	3063.74
ARIMA (212)	3040.56	3053.06

Table 2: Parameter estimates along with corresponding standard errors

Conditional Least Squares Estimation					
Parameter	**Estimate**	**Standard Error**	**t Value**	**Approx Pr > |t|**	**Lag**
MU	17.62148	29.48285	0.60	0.5506	0
MA1,1	-0.44722	0.05386	-8.30	<.0001	1
MA1,2	-0.92103	0.05296	-17.39	<.0001	2
AR1,1	-0.47983	0.08036	-5.97	<.0001	1
AR1,2	-0.80511	0.07994	-10.07	<.0001	2

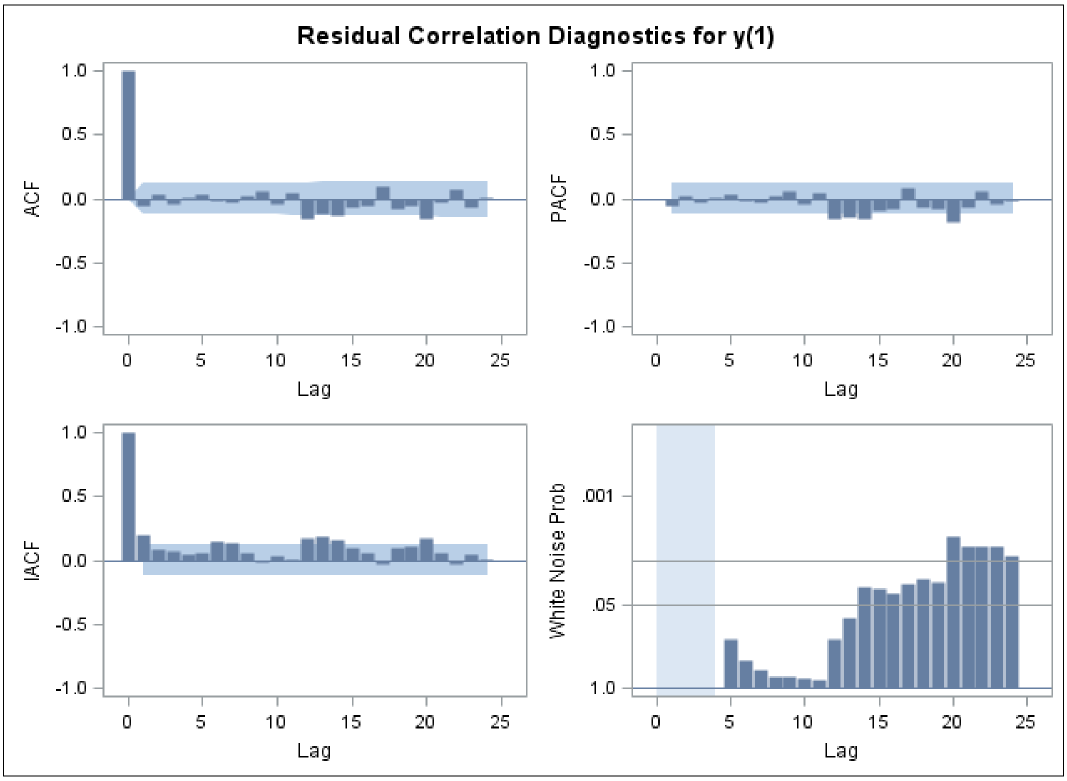

Fig 3: ACF and PACF of the residuals of fitted ARIMA model

Forecasting

Design on the selected models, we further refine the model by examining residuals to ensure they exhibit white noise characteristics. If the model's residuals demonstrate white noise, it can be used for forecasting. The residual analysis of the ARIMA (212) model indicates white noise (refer to Fig 3). With this confirmation, we shift our focus to the application, specifically forecasting. The objective is to predict future values of the time series. Table 3 displays monthly forecasted results along with confidence limits for the time series.

Table 3: Forecasts of maize prices in Telangana State up to December 2023.

Months	Forecast	Std Error	95% Confidence Limits	
Jan-23	2182.17	83.35	2018.81	2345.52
Feb-23	2186.00	109.57	1971.24	2400.76
Mar-23	2189.84	130.27	1934.53	2445.16
Apr-23	2202.11	120.16	1954.32	2399.02
May-23	2234.74	80.47	2077.03	2392.46
Jun-23	2279.72	106.05	2071.87	2487.57
Jul-23	2297.14	127.03	2048.15	2546.12
Aug-23	2282.47	150.35	1987.79	2577.16
Sep-23	2008.90	84.89	1844.12	2176.87
Oct-23	2067.01	119.01	1806.87	2273.38
Nov-23	2054.37	144.07	1809.94	2374.67
Dec-23	2067.91	166.25	1742.06	2393.75

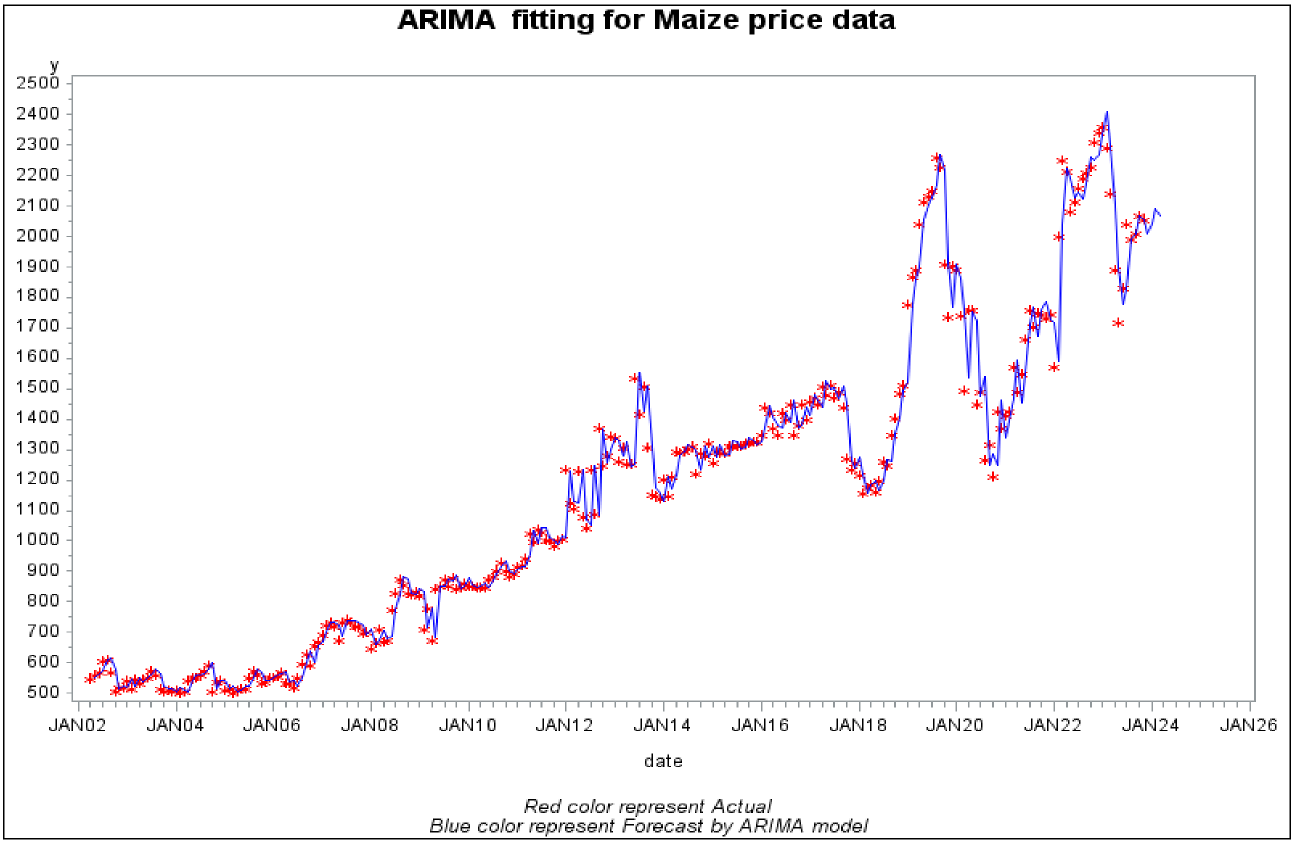

Fig 4: Maize prices in Telangana state actual and forecasted data from the month and year April 2002 to December 2023

References

1. 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.
2. Aidan M, Geoff K, Terry Q. Forecasting Irish inflation using ARIMA models. Central Bank and Financial Services Authority of Ireland Technical Paper Series, 1998, 3(1-48).
3. Prajneshu, Venugopalan R. On nonlinear procedure for obtaining length - weight relationship. Indian Journal of Animal Sciences. 1998;68(1):452-456.
4. 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.
5. Singh S, Ramasubramanian V, Mehta SC. Statistical Models for forecasting rice production of India. Journal of Indian Society of Agricultural Statistics. 2007;61(2):80-83.
6. Box SJ, Mumby PJ. Effect of macroalgal competition on growth and survival of juvenile Caribbean corals. Marine Ecology Progress Series. 2007 Jul 24;342:139-149.