The Andhra Agric. J 70 (2): 246-253, 2023
10.61657/aaj.2023.37

Price Forecasting of Chilli in Warangal Market Using ARIMA Model

G Hareesh, Sk Nafeez Umar, B Ramanamurthy and B Aparna
Department of Statistics and Computer Applications, S. V. Agricultural College, Tirupati

ABSTRACT

This article presents a comprehensive study on the application of the Auto Regressive Integrated Moving Average (ARIMA) model for the purpose of price forecasting in the chilli market. The research employs historical price data of chilli spanning multiple years, collected from the Warangal market of Telangana State. Initially, the data is subjected to preprocessing (or) reprocessing techniques to ensure its quality and stationarity. The ARIMA model is applied to generate price forecasts for the upcoming periods, thereby enabling stakeholders to anticipate potential market trends and price fluctuations. The study considers performance metrics like Root Mean Squared Error (3.726), Mean Absolute Percentage Error (2.605) and Akaike Information Criterion (350.62) to evaluate the forecasting capabilities of each method. The results demonstrate that the ARIMA (2, 1, 2) model is appropriate for accurately predicting chilli prices in the Warangal market. The increasing volatility and uncertainty in agricultural commodity prices have necessitated the development of robust forecasting methods to aid farmers, traders and policymakers in making informed decisions.

Keywords*: *ARIMA, Chilli and Price forecasting.

Chilli, scientifically known as *Capsicum annuum*, is an essential spice crop that holds immense economic significance globally. Known for its pungent flavour and vibrant colours, chilli is a widely used ingredient in culinary practices around the world. Apart from its culinary applications, chilli also finds extensive use in the pharmaceutical, cosmetic, and food processing industries. Due to its versatile applications, the demand for chilli remains consistently high, making it a crucial commodity in the agricultural market.

In India, one of the largest producers and consumers of chilli, the market for this fiery spice is particularly dynamic. Among the various regional markets in India, the Warangal market plays a pivotal role in chilli trade. Warangal, situated in the state of Telangana, is known for its prominent agricultural activities and chilli cultivation forms a significant part of the region's agricultural landscape.

In 2019-20, Telangana ranked second in chilli area, production and productivity i.e., 0.85 lakh hectares (2.10 lakh acres), production 3.28 lakh tonnes and productivity 3859 Kg/ha (1561 Kg per acre) respectively.

The annual production of red chilli in Telangana in 2020-21 is 4.07 lakh tonnes. The state ranked second in production, the state accounts for 21% of red chilli produced in India. cultivation of chilli in this state is over 90,300 hectares with an yield of 4,510 kg / hectare. In India, second most chilli producing state is Telangana (3.04 lakh tonnes), bettered only by Andhra Pradesh (6.30 lakh tonnes) and is followed by Madhya Pradesh (2.18 lakh tonnes) and Karnataka (1.95 lakh tonnes) accounting for 33, 22, 14 and 10 per cent of all India production, respectively. (Source: APEDA agri exchange).

Currently in 2022-23, Telangana is ranked second in chilli production and productivity and fourth in cultivated area. Telangana accounts for 11.59% and 22% of all India area and production respectively. The major chilli growing districts are Khammam, Warangal (Rural), Mahabubabad, Gadwal and Surya pet. There is good demand for chilli hybrid varieties like 334 and Teja in the international market. Presently in 2022-23, Andhra Pradesh and Telangana, around 5.40 lakh tonnes chilli was in cold storages.

For farmers, traders and policymakers involved in the chilli market, accurate price forecasting is of paramount importance. Price fluctuations in the chilli market can have far-reaching consequences, impacting farmers income, trader's profits and consumers budgets. Moreover, policymakers rely on reliable price forecasts to implement effective market interventions and policies to stabilize prices and ensure food security.

Traditional methods of price forecasting have often proven to be inadequate in capturing the intricate patterns and dynamics of the chilli market. In recent years, data-driven approaches and advanced statistical techniques have gained popularity for their ability to offer more accurate and insightful predictions. In this context, this article presents an innovative approach to price forecasting of chilli in the Warangal market using the Autoregressive Integrated Moving Average (ARIMA) model. ARIMA is a time series forecasting technique that combines autoregression, differencing and moving average components to model and predict sequential data (Darekar *et al*., 2016).

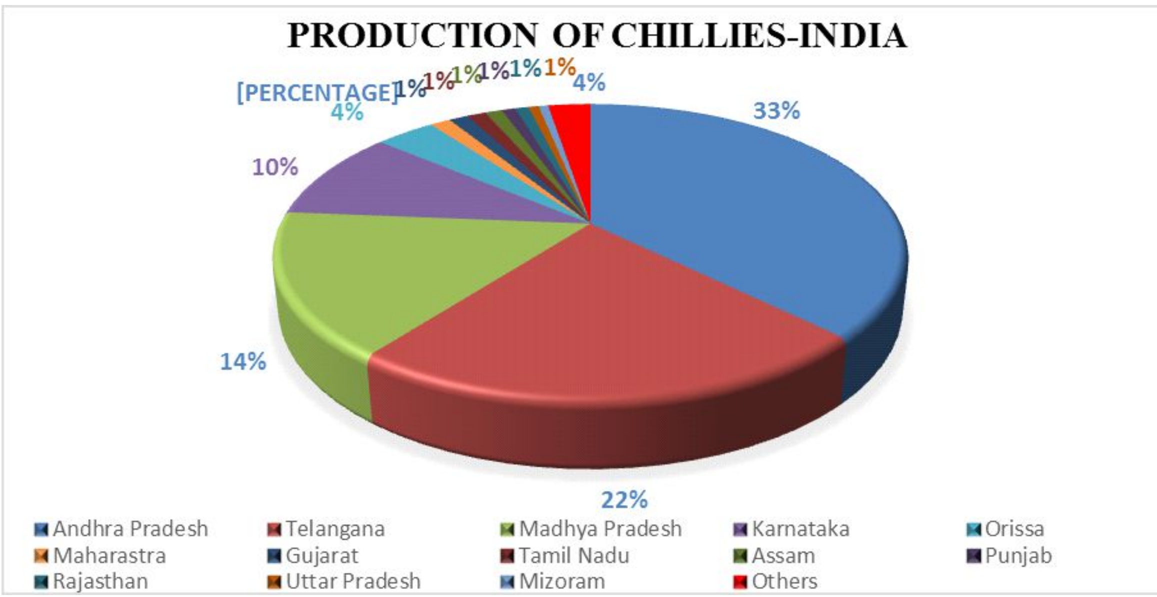

Figure 1: Percent share of states in Chilli production

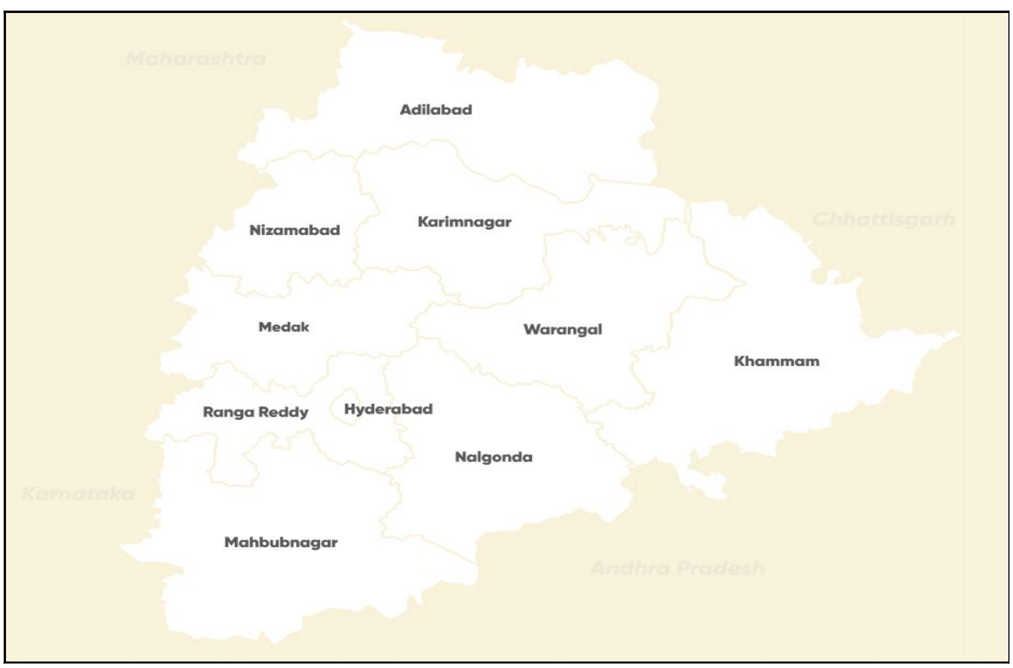

Figure 2: Selected Market in Districts of Telangana

The primary objective of this article is to demonstrate the effectiveness of the ARIMA model in forecasting chilli prices accurately and reliably. By analysing historical price data spanning several years, the ARIMA model aims to capture the underlying trends and seasonality patterns in the chilli market, providing stakeholders with valuable insights for informed decision-making.

The following sections will delve into the methodology employed, data and findings and model calibration. The model's forecasting performance will be rigorously evaluated using residual analysis and the results will be discussed in detail. Additionally, the implications of accurate price forecasts for farmers, traders and policymakers in the Warangal chilli market will be explored.

Ultimately, this study seeks to contribute to the growing body of research on agricultural price forecasting and establish the ARIMA model as a valuable tool for enhancing efficiency, sustainability and profitability in the chilli market. By empowering stakeholders with reliable forecasts, this approach can foster better planning, risk management and market stability in the Warangal chilli trade and potentially serve as a template for similar analysis in other agricultural markets worldwide.

MATERIAL AND METHODS

The study was conducted on a major spice *i.e.,* Chilli as this spice has more demand in our daily life due to its culinary and flavour, health benefits, medicinal and therapeutic applications, agricultural, environmental, cultural and social significance.

The time series data on monthly prices (Rs/qtl) on highly consumed spice namely chilli for the period of 5 years 3 Months *i.e*., from January 2018 to March 2023 at selected market *i.e*., Warangal market in Warangal district of Telangana state was collected for the study, which resulted in 63 data points respectively.

The AutoRegressive Integrated Moving Average (ARIMA) methodology developed by Box-Jenkins is the most widely used model for analysing time series data. The basic purpose of fitting the ARIMA model is to accurately characterise and forecast the time series stochastic process (Box and Jenkins, 1970). Initially, George Box and Gwilym Jenkins conducted substantial research on ARIMA models and their names were frequently associated with the broad ARIMA method used in time series analysis, forecasting and control. The ARIMA model can only be used with stationary data.

a) Auto Regressive (AR) Model (p, 0, 0)

In this model, the current value of the process is expressed as a finite, linear aggregate of previous values of the process and a shock ε_t. Let us denote the values of a process at equally spaced time epochs t, t-1, t-2……by y_t, y_{t-1}, y_{t-2}, …., then y_t can be described by the following expression

The autoregressive process of first order AR (1) is given by Devika and Mohideen (2022)

$$y_t=\emptyset_1 y_{t-1}+\varepsilon_t \quad \text{-------------------------------- (i)}$$

The autoregressive process of second order AR (2) is given by

$$y_t=\emptyset_1 y_{t-1}+\emptyset_2 y_{t-2}+\varepsilon_t \quad \text{---------------------- (ii)}$$

And similarly, the autoregressive process of p- order AR (p) is given by

$$y_t=\emptyset_1 y_{t-1}+\emptyset_2 y_{t-2}+\dots\dots+\emptyset_p y_{t-p}+\varepsilon_t \quad \text{----------- (iii)}$$

Where, y_t is the value of variable for forecasting at time 't' (Arrivals and Prices in the present case)

$\phi_1, \phi_2, \dots\dots \phi_p$: Regression Coefficients and

ε_t : Random error

b) Moving Average (MA) Model (0, 0, q)

The moving average process of q^{th} order MA (q) model is defined as

$$y_t=\varepsilon_t-\theta_1\varepsilon_{t-1}-\theta_2\varepsilon_{t-2}-\dots\dots-\theta_q\varepsilon_{t-q} \quad \text{--------- (iv)}$$

The moving average process of first order MA (1) is given by

$$y_t=\varepsilon_t-\theta_1\varepsilon_{t-1} \quad \text{-------------------- (v)}$$

The moving average process of second order MA (2) is given by

$$y_t=\varepsilon_t-\theta_1\varepsilon_{t-1}-\theta_2\varepsilon_{t-2} \quad \text{------------------- (vi)}$$

Where, y_t is the value of the variable for forecasting at time 't' (Arrivals and Prices in the present case)

ε_t : Random error and

θ_p to θ_q : Partial regression coefficients

c) Auto Regressive Moving Average (ARMA) Model (p, 0, q)

To achieve greater flexibility in fitting of actual time-series data, it is sometimes advantages to include both autoregressive and moving average process. This leads to the mixed autoregressive-moving average model (Kumar *et al*., 2018)

$$Y_t=\emptyset_1 y_{t-1}+\emptyset_2 y_{t-2}+\ldots\ldots+\emptyset_p y_{t-p}+\varepsilon_t-\theta_1\varepsilon_{t-1}-\theta_2\varepsilon_{t-2}-\ldots\ldots-\theta_q\varepsilon_{t-q} \quad \text{--------------- (vii)}$$

This is written as ARMA (p, q) model. In practice, it is frequently true that adequate representation of actually occurring stationary time-series can be obtained with autoregressive, moving average or mixed models.

d) Auto Regressive Integrated Moving Average (ARIMA) Model (p, d, q)

The ARIMA model essentially requires identification of three constants p, d, q *i.e.,* the order of AR terms (p), order of differencing (d) and the order of MA terms (q). The ARIMA methodology is also called as Box-Jenkins methodology. The first step in developing ARIMA model is to examine data for stationarity. This can be identified through Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) of actual data. The characteristics p and q are determined on the basis of PACF and ACF of the stationary data.

The ARIMA (p, d, q) process is given by (Mallikarjun *et al*., 2019)

$$y_t=\theta_0+\emptyset_1 y_{t-1}+\emptyset_2 y_{t-2}+\ldots\ldots+\emptyset_p y_{t-p}+\varepsilon_t-\theta_1\varepsilon_{t-1}-\theta_2\varepsilon_{t-2}-\ldots\ldots-\theta_q\varepsilon_{t-q} \quad \text{------------- (viii)}$$

Where, y_t and ε_t are the actual value and random error at time period t, respectively.

$\emptyset_p$ (i=1, 2, ..., p) and θ_j (j=1, 2, ..., q) are model parameters. p and q are integers and often referred to as orders of the model. Random errors are assumed to be independently and identically distributed with a mean of zero and a constant variance of the ARIMA model incorporates non-seasonal factors in a multiplicative model and is denoted as: ARIMA (p, d, q)

Where, p = order of non- seasonal Auto Regressive (AR)

d = order of non- seasonal difference

q = order of non- seasonal Moving Average (MA)

Box and Jenkins propose a practical four stage procedure for finding a good model. The four-stage univariate Box Jenkins procedure is summarized schematically in figure 3.

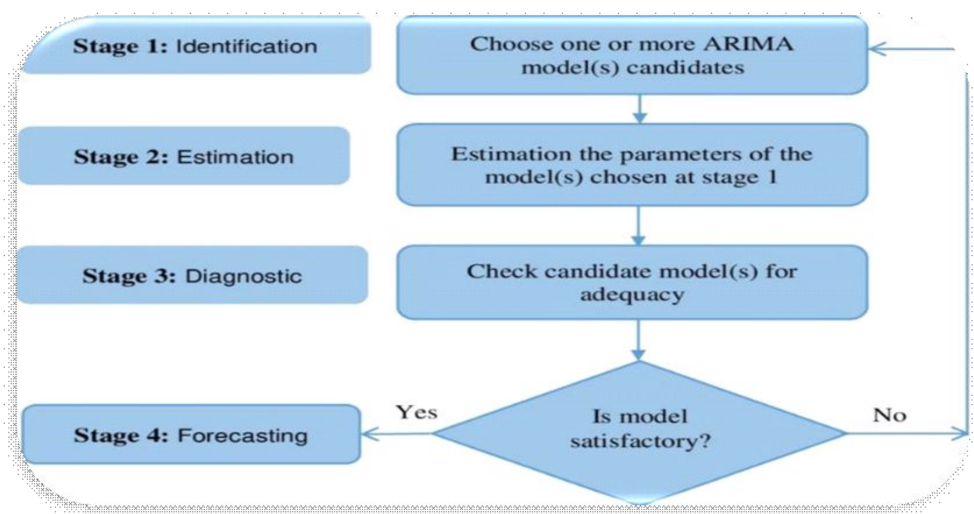

Figure 3: Flow chart of Box-Jenkins Methodology

The main stages in setting up a Box-Jenkins forecasting model are as follows:

Stage 1: Identification

There are two kinds of stationarity, *viz*., stationarity in 'mean' and stationarity in 'variance'. Identification of the model was concerned with deciding the appropriate values of (p, d, q). It was done by the auto.arima function in R which uses a variation of the Hyndman-Khandakar algorithm, performs an iterative procedure to select the model that minimizes the value of each criterion (RMSE & MAPE).

Stage 2: Estimation of parameters

We attempt to obtain precise estimates of parameters of the model by least square as advocated by Box and Jenkins. Standard computer software like R programming is available for finding the estimates of relevant parameters using iterative procedures.

Stage 3: Diagnostic checking

Here, if it shows random residuals, then it indicates that the tentatively identified model was adequate. When an inadequacy is detected, the checks should give an indication of how the model needs to be modified, after which further fitting and checking takes place. Some of the diagnostic checks are mentioned below under Performance models.

Stage 4: Forecasting

The forecast of the variable in the t^{th} year (y_t) in ARIMA model is based on its past values. After the identification of the model and its adequacy check, it is used to forecast the future values. The forecasts can be obtained from the estimated ARIMA model.

RESULTS AND DISCUSSION

Secondary data of monthly chilli prices from Warangal market between January-2018 and March-2023 was taken, has a total number of 63 observations which are used to estimate and forecast the model. It must be noted that the data is divided into two parts: first part is the Training data which covers the period from January-2018 to September-2022 that includes 57 observations and is used to estimate the model, second part is testing data which covers the period from October-2022 to March-2023 and is used for forecasting. The data used in the study is taken from http://agrimarketing.telangana.gov.in/indexnew.jsp. Summary statistics reveal that, least price was noticed in November-2018 (Rs. 7941/qtl) and peak price was recorded in March-2023 (Rs. 18283/qtl).

FORECASTING MODEL

The secondary data on monthly prices of chilli from Warangal market at Warangal district of Telangana was subjected to statistical analysis using the methodologies given earlier highlights of the outcomes obtained in accordance with the methodology involved were presented below.

a) Identification of the model

The price series of chilli follows Normality, that can be done by using the Shapiro-Wilk test on the price data (p-value is 0.059(>0.05)). The price series data of chilli in Warangal clearly exhibited non-stationarity, which was converted to stationary, by using Augmented Dickey-Fuller Test, by the first difference. Several tentative models were identified by examining the Autocorrelation Function (ACF) (figure 4) and Partial Auto Correlation Function (PACF) (figure 5) As there is no evidence of seasonality, no seasonal models were verified. The possible models are initially found (Table 1.) shows the values for all of the models that were tentatively recognised. Based on least RMSE (3.726), MAPE (2.605) and AIC (350.62) values, the model ARIMA (2, 1, 2) was identified for prices of chilli in Warangal market.

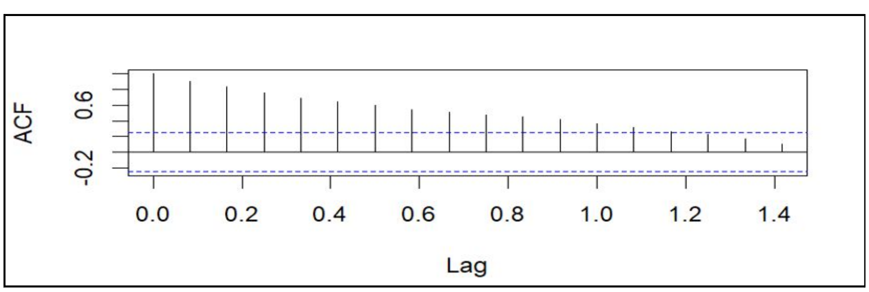

Figure 4: ACF chart for chilli prices of Warangal market

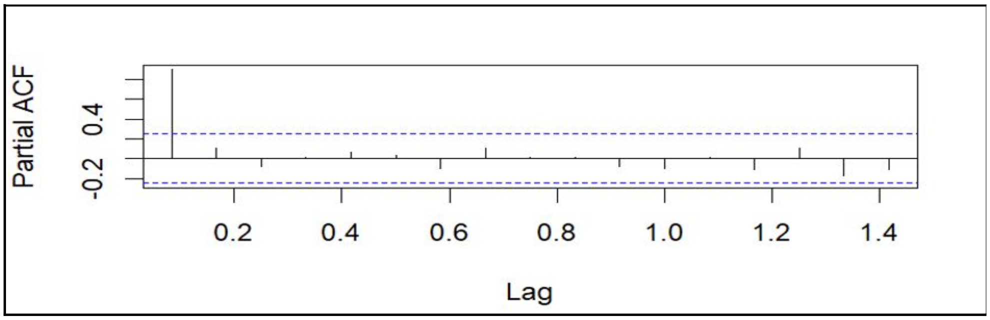

Figure 5: PACF chart for chilli prices of Warangal market

Table 1. The tentative models for chilli prices in Warangal market

ARIMA model	RMSE	MAPE	AIC
ARIMA (1, 1, 0)	3.953	2.716	351.44
ARIMA (1, 1, 1)	3.931	2.677	352.78
ARIMA (1, 1, 2)	3.921	2.695	354.46
ARIMA (2, 1, 0)	3.928	2.686	352.67
ARIMA (2, 1, 1)	3.927	2.683	354.65
ARIMA (2, 1, 2)	**3.726**	**2.605**	**350.62**
ARIMA (3, 1, 0)	3.926	2.678	354.63

Table 2. ARIMA (2, 1, 2) model parameter table

Model	**Coefficients**		**Estimate**	**SE**
ARIMA (2, 1, 2)	**AR**	Lag 1	-0.8509	0.0767
		Lag 2	-0.8948	0.0694
	MA	Lag 1	0.7618	0.0553
		Lag 2	-0.9998	0.0719

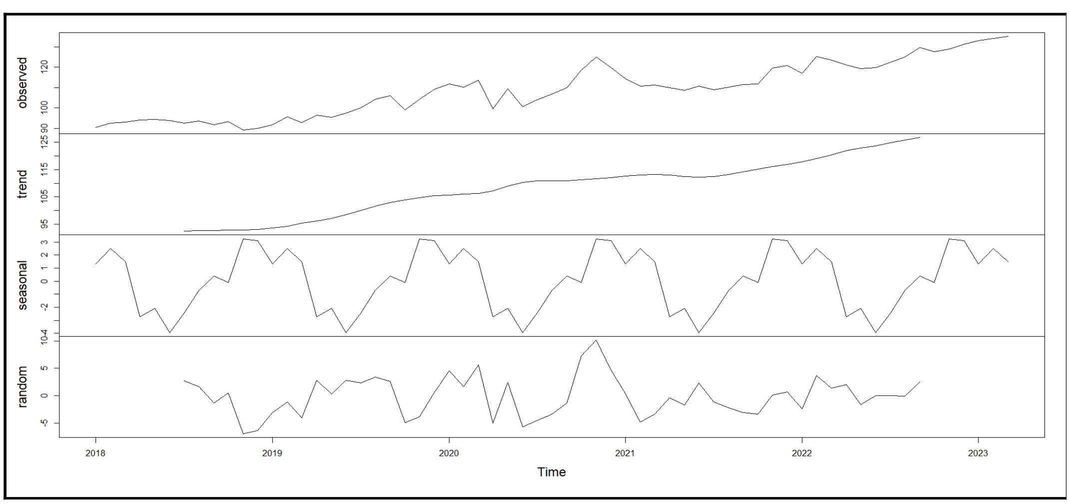

Figure 6: Decomposition plot for chilli price data of Warangal market

The Decomposition plots for the chilli price data of Warangal market, which are graphical representations used to analyse time series data by breaking it down into its individual components (Figure 6). These components typically include trend, seasonality and residuals.

Residual analysis can be used to determine the model's suitability. The residuals should be independent according to one of the basic assumptions of ARIMA model (Figure 3.4.) The adequacy of the fitted model can be found out by residual analysis. The p-value of the Ljung-box test for chilli prices was 0.283 (>0.05), indicating residual independence.

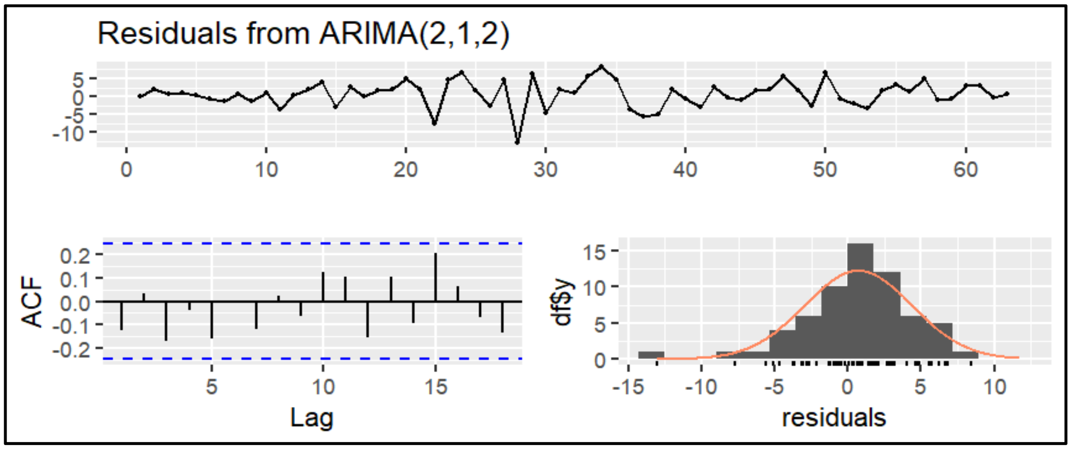

Figure 7: Residuals of chilli prices in Warangal market

Forecasted values of prices on chilli in Warangal market using ARIMA

The forecasted values of chilli prices in Warangal market using the ARIMA (2, 1, 2) model (Table 3.)

Table 3. Forecasted values of prices on chilli in Warangal market using ARIMA

	Months	**Forecasted values using ARIMA (2, 1, 2) model**
Prices (Rs/qtl)	April-2023	17812.84
	May-2023	18202.56
	June-2023	18295.01
	July-2023	17864.43
	August-2023	18147.77
	September-2023	18295.11

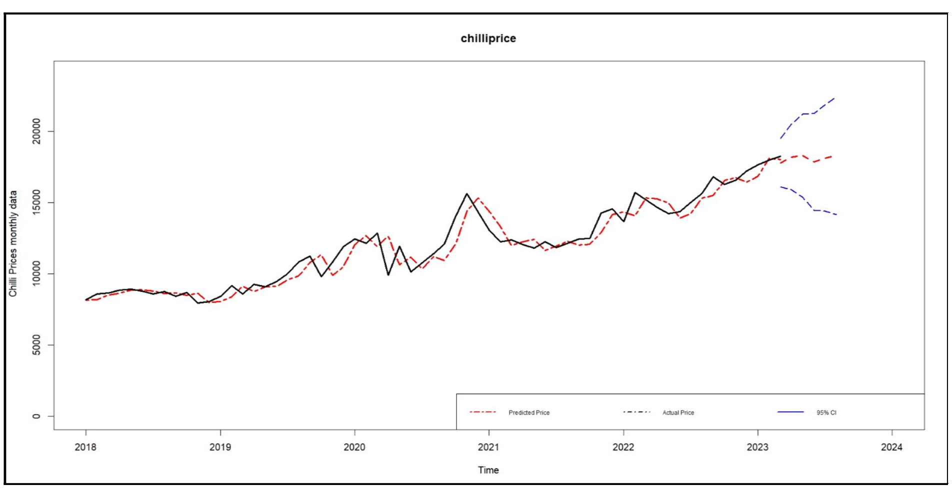

Figure 8: Ex-post forecast of chilli prices in Warangal market

The study aims to investigate the application of Auto Regressive Integrated Moving Averages (ARIMA) for forecasting the Chilli prices in Warangal market from January-2018 to March-2023.

Based on our study, the forecasting of Chilli prices identified by ARIMA (2, 1, 2) model is found fruitful and is essential for several reasons, as it provides valuable insights and benefits to various stakeholders involved in the chilli industry and the broader economy. Some of the key reasons for the need for chilli price forecasting include Agricultural planning, Marketing analysis, Risk management, Supply chain optimization, Consumer impact, export and import decisions, Policy formulation, Investment decisions and economic stability. It benefits various stakeholders, from farmers to consumers, by providing valuable information for planning and strategizing in an unpredictable market environment.

LITERATURE CITED

Baser U, Bozoglu M, Eroglu N A and Topuz B K 2018. Forecasting chestnut production and export of Turkey using ARIMA model. *Turkish Journal of Forecasting*. 2(2): 27-33.

Darekar A S, Pokharkar V G and Datarkar S B 2016. Onion price forecasting in Kolhapur market of Western Maharashtra using ARIMA technique. *International Journal of Information Research and Review*. *3*(12): 3364-3368.

Devaki C and Mohideen A K 2022 A sugarcane production in tamilnadu using a comparison of arima, state space and linear mixed models. *International Journal of Mechanical Engineering,* pp.970-974.

Kumar K, Bhat A R S, Pavithra N L and Megha J 2018. Comparative performance of ANN and ARIMA models in redgram price forecasting-Kalaburagi market. *Journal of Pharmacognosy and Phytochemistry*. 7(5): 1630-1632.

Mallikarjuna H B, Paul A, Paul A, Noel A S and Sudheendra M 2019. Forecasting of black pepper price in Karnataka state: An application of ARIMA and ARCH models. *International Journal of Current Microbiology Agricultural Science*. 8(1): 1486-1496.

Mashadihasanli T 2022. Stock Market Price Forecasting Using the Arima Model: an Application to Istanbul, Turkiye. *Journal of Economic Policy Researches*, *9*(2), pp.439-454.

Received on 13.04.2023 and Accepted on 11.06.2023