

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2024; SP-9(1): 22-29
© 2024 Stats & Maths
<https://www.mathsjournal.com>
Received: 24-11-2023
Accepted: 26-12-2023

Mohit Kumar

Ph.D. Scholar, Department of
Agricultural Economics, B. A.
College of Agriculture, Anand
Agricultural University, Anand,
Gujarat, India

Rohit Kumar Sharma

Ph.D. Scholar, International
Agribusiness Management
Institute, Anand Agricultural
University, Anand, Gujarat,
India

Upasana D Bhopala

Ph.D., Department of
Agricultural Economics, College
of Agriculture, Junagadh
Agricultural University,
Junagadh, Gujarat, India

Bhoomi Suthar

Assistant Professor, Department
of Agricultural Economics, ITM
University, Gwalior, Madhya
Pradesh, India

Corresponding Author:

Mohit Kumar

Ph.D. Scholar, Department of
Agricultural Economics, B. A.
College of Agriculture, Anand
Agricultural University, Anand,
Gujarat, India

Forecasting prices of onion in major wholesale markets of Gujarat

Mohit Kumar, Rohit Kumar Sharma, Upasana D Bhopala and Bhoomi Suthar

DOI: <https://doi.org/10.22271/math.2024.v9.i1Sa.1539>

Abstract

This research focuses into the complex dynamics of onion pricing in the Indian agricultural sector, with a specific emphasis on the volatility observed in onions, particularly within the wholesale markets of Gujarat. The research aims to provide a comprehensive understanding of price fluctuations and their impact on farmers, consumers, and the economy at large. The study employs time series analysis, specifically the ARIMA technique, to forecast future onion prices. The methodology involves unit root tests, model identification, estimation, and diagnostic checking. The selected models, ARIMA (3,1,2) for Mahuva, ARIMA (2,1,1) for Ahmedabad, and ARIMA (2,1,2) for Gondal, showcase the nuanced approach required for different markets. Post-sample period forecasts for 2021 reveal predicted onion prices. The forecasting accuracy is assessed using Mean Absolute Percentage Error (MAPE), with Mahuva exhibiting the lowest MAPE at 21.77 percent. The study emphasizes the market-specific nature of onion price dynamics, underscoring the importance of tailoring forecasting models to individual market characteristics.

Keywords: Onion, forecast, ARIMA, price volatility introduction

Introduction

Within the agricultural sector, pricing holds substantial significance in the context of the Indian economy. It serves as a crucial factor in the computation of farm revenues and exerts a direct influence on the welfare of farmers (Saxena and Chand, 2017) ^[1]. Onions are the most highly volatile crop among vegetables, exhibiting a notable tendency for unexpected price spikes and falls (Mulla *et al.*, 2020; Pradeep, 2015) ^[5, 6]. The pronounced price fluctuations in recent years have captured the attention of policymakers. Onions are traditionally cultivated in the northern regions during the winter (*rabi*) season. Meanwhile, in the western and southern states of Andhra Pradesh, Karnataka, Tamil Nadu, Gujarat, and Maharashtra, onions are grown during both the winter (*rabi*) and rainy (*kharif*) seasons.

Indian scenario

Maharashtra takes the lead among Indian states as the top onion producer, contributing significantly in terms of both area and production. In the 2018-19 period, Maharashtra's onion output accounted for a substantial 29.55 percent (authors calculation) of the total onion production in India. The Lasalgaon market in the Nasik district of Maharashtra holds a prominent position as the primary onion procurement market in India. It serves as a hub where onions are acquired from farmers and subsequently distributed across the country.

Gujarat scenario

Gujarat stands as the fifth-largest onion producer in India, contributing 3.96 percent (authors calculation) of the total cultivated area for onions in the country during 2018-19. Additionally, it holds a 5.9 percent share in the overall onion production in India.

Need of present study

The abrupt surge in onion prices in the market has a ripple effect, impacting both producers and consumers.

This spillover reaction extends to other onion markets, leading to elevated economic inflation. It's not just the producers who rely on price information; consumers and government agencies also depend on such data to formulate various policies. It's a crucial aspect that influences decision-making across different sectors of the economy. Being aware of long-term market fluctuations empowers producers to strategize their production decisions, determining what to produce and how much (Saxena *et al.*, 2019)^[7]. This not only shapes the nature of their enterprise but also has a lasting impact on individuals' pockets in the long run. It's like playing the long game in the dynamic field of agriculture.

The primary goal of time series analysis is to forecast future prices, providing insights into what is likely to occur (Pal, 2019)^[2]. It goes beyond just explaining why certain events will unfold, offering a forward-looking perspective that aids in making informed decisions in anticipation of future developments. The ultimate goal is to the present study is to provide suitable and appropriate price forecast so that proper policy options can be framed to minimize the recurrence of onion price shocks.

Materials and Methods

The present study is based on time series data of monthly duration, spanning from January 2004 to December 2020 focusing specifically on onion prices. Total 3 markets regional markets were selected from Gujarat state *viz.*, Mahuva, Ahmedabad, and Gondal. Market selection was based on the highest triennial ending average of onion arrivals from 2017 to 2020. Data analysis was conducted using R software to delve into the intricate patterns of onion prices and market dynamics.

Analytical Techniques Used

In order to achieve the objectives of the present study, ARIMA technique was applied to forecast the prices on onion in selected markets.

Unit root test

A stochastic process is considered stationary if its mean and variance remain constant over time. Additionally, the covariance between two periods is said to be stationary if it depends solely on the time gap or lag between the two periods, irrespective of the actual time at which the covariance is computed. Such a stochastic process is called weak stationary or covariance stationary or second order stationary *etc.*

Mean: $E(Y_t) = \mu$

$\text{Var}(Y_t) = E(Y_t - \mu)^2 = \sigma^2$

Covariance (auto covariance) = $E[(Y_t - \mu)(Y_{t+k} - \mu)] = \gamma_k$

It is very important to test whether or not the time series is stationary because if a time series is not stationary, its behaviour can only be studied for the time period under consideration, it cannot be generalized to other periods & thus one cannot predict such a time series data. So in order to test the data is stationary or having unit root, the famous test known as Augmented Dickey-Fuller (ADF) test is used.

The presence of unit root (non-stationary) in the underlying series is tested by performing Augmented Dickey-Fuller test using the following regression.

$$\Delta Y_t = \alpha + \beta_i T + \delta_i Y_{it-1} + b_i \sum_{i=1}^p \Delta Y_{it-1} + e_t \quad (5)$$

Where,

Y_{it} = Price of a commodity in a given market 'i' at a time 't';
 $\Delta Y_{t-i} = (Y_{t-1} - Y_{t-2})$ (t-i - lagged prices & Δ is Differenced series);

e_t is pure white noise error-term,

α is the drift parameter,

T is the time trend effect,

β_i, δ_i & b_i is coefficients

p is the optimal lag value which is selected on the basis of Akaike Information Criterion (AIC)

The null hypothesis is that the coefficient of Y_{t-1} is zero.

The alternative hypothesis is: $\delta < 0$.

The possibility of acceptance & rejection of H_0 is based on the tau statistic or test (τ) & the estimation procedure of tau statistic (τ) is as follows.

1. Estimate the equation (1) by OLS method.
2. Divide the estimated coefficient of Y_{t-1} by its standard error & refer to the DF (Dickey-Fuller) table.
3. If the computed absolute value of the ($|\tau|$) exceeds the absolute DF or MacKinnon critical tau values, then the null hypothesis ($\delta = 0$) is rejected, in which the case of time series is stationary.

On the other hand if the computed absolute value of the ($|\tau|$) does not exceeds the absolute DF or MacKinnon critical tau values, then the null hypothesis ($\delta = 0$) is accepted, in which the case of time series is non stationary

ARIMA Model

An Autoregressive Integrated Moving Average (ARIMA) model is characterized by the notation ARIMA (p, d, q) where p, d & q denotes orders of auto-regression, integration (differencing) & moving average respectively. ARIMA is a parsimonious approach which can represent both stationary & non-stationary process. An ARMA (p, q) process is defined by the equation.

$$P_{1t} = \mu + \phi_1 P_{1t-1} + \phi_2 P_{1t-2} + \dots + \phi_p P_{1t-p} + \theta_0 \varepsilon_{1t} + \theta_1 \varepsilon_{1t-1} + \dots + \theta_q \varepsilon_{1t-q} \quad (11)$$

Where,

P_{1t} = price of series at time period

μ = constant term.

ϕ_i (i = 1, 2, ..., p) & θ_j (j=0, 1, 2, ..., q) are model parameters.

ε_{1t} = Random Error at time period t.

$\varepsilon_{1t} \sim \text{IID} (0, \sigma^2)$

However, in practical applications, residuals obtained after fitting of appropriate ARIMA model may have non-constant error variance. Engle (1982) proposed to model time-varying conditional variance with auto-regressive conditional heteroscedasticity (ARCH) process using lagged disturbance.

Diagnostic checking

The ARIMA (p, d, q) model is assumed to be efficient when the residuals estimated from it are of white noise which can be ensured only when the residuals of the fitted model are used for diagnostic checking (Jadav *et al.*, 2017)^[4]. In this way, the estimated model is checked to verify if it adequately represents the series. In this study, diagnostic checks was performed on the residuals to see if they are randomly &

normally distributed by using Jarque – Bera (JB) test for normality. In addition, the adequacy of the selected model was checked using Box-Ljung test.

The Jarque – Bera (JB) test of the following specification was used in the present study.

$$JB = \frac{n}{6} \left(s^2 + \frac{(K-3)^2}{4} \right) \sim \chi^2_{(2)}$$

Where in:

$$Skewness (s) = \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^3}{\left[\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \right]^{\frac{3}{2}}}$$

$$Kurtosis (k) = \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^4}{\left[\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \right]^2}$$

Where: n is the number of observations, \bar{y} is the mean of the underlying variable series under study & y_i refers to the individual values of the variable under study. The statistic JB has an asymptotic chi-square distribution with 2 degrees of freedom & can be used to test the hypothesis of skewness being zero & excess kurtosis being zero. If $JB > \chi^2_{(\alpha, 2)}$, then the null hypothesis was rejected & it was concluded that the data do not follow normal distribution.

In addition, an overall check of the model adequacy was made using Box-Ljung test (Ljung & Box, 1978) [3]. The test statistics is given by.

$$Q = n(n + 2) \sum_{k=1}^m \frac{r_k^2}{n-k} \tag{15}$$

Where: n is the number of observations, r_k is the estimated autocorrelation of the series at lag $k = 1, 2, \dots, m$ & m is the number of lags being considered, $\chi^2_{(1-\alpha, h)}$ is the chi-square distribution table value with ‘h’ degrees of freedom & level of significance ‘ α ’ such that $P(\chi^2_{(h)} > \chi^2_{(1-\alpha, h)}) = 1 - \alpha$ & degrees of freedom, $h = (m-p-q)$; p & q are the numbers of AR and MA terms, respectively.

In the present study, a formal test regarding the overall fitness of the model was done using Box-Ljung test of the residuals in the following manner: (i) Null hypothesis (H_0): The errors are distributed randomly & (ii) Alternate hypothesis (H_1): The errors are non-random. Accordingly, the null hypothesis was rejected if $Q > \chi^2_{(1-\alpha, h)}$ & the errors are not considered to be independent. On the other hand, the null hypothesis was accepted *i.e.*, the errors are independent if $Q < \chi^2_{(1-\alpha, h)}$. Thereby, if the Q values happens to be significantly large than zero exceeding the table $\chi^2_{(1-\alpha, h)}$ value then it is to be concluded that the residuals of the estimated model are probably auto-correlated & the entire model was then has to be reformulated.

Forecasting

Once the three previous steps of ARIMA model are over, then we were able to obtain the forecasted values by estimating appropriate model. ARIMA models were used to forecast the corresponding variable. For that, the entire data was segregated into two parts: one for sample period forecasts & the other for post-sample period forecasts. The former was used to develop confidence in the model & the latter was used to generate genuine forecasts for use in future planning.

In this regard, the actual value of the left out period & the forecasted value of the left out period from the selected model are used for cross-validation. For this, the percentage error is calculated such as.

$$\% \text{ of Forecasting Error} = \left(\frac{Y - \hat{Y}}{Y} \right) \times 100 \tag{16}$$

Where, Y is the observed value of remaining twelve months (January, 2017 to December, 2020) & \hat{Y} is the forecast values of remaining the period under consideration.

Lower the value of forecasting error percentage, better is the prediction by the selected model. Besides, the accuracy of the forecasts for both ex-ante & ex-post was tested for the minimum values of Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) & maximum value of coefficient of determination (R^2). Further, Box-Jenkins ARIMA model was run to forecast the onion price under study for the ex-post facto period from January, 2021 to March, 2021.

Results and Discussion

The ARIMA model under study is a parsimonious model which tends to illustrate the data more accurately with few parameters. It involves both the Autoregressive (AR) & Moving Average (MA) components to explain the data. When it is compared with any other linear time series model, this particular model is preferred for eliminating the linear dependency in the data (Jadhav *et al.*, 2013) [8]. There are four steps involved in fitting the model *viz.*, identification, estimation, diagnostic checking and forecasting for the set of data under discussion (Xin and Can, 2016) [9]. Therefore, all the necessary steps were carried out for all the selected markets and outlined in the respective table.

Identification of parameter

The ARIMA model comprise of three parameter *viz.*, p, d and q where ‘p’ is the number of autoregressive (AR) terms while ‘q’ is the number of moving average (MA) terms. The number of times the sequence must be differenced in order for it to become stationary is denoted by ‘d.’ The parameter ‘d’ was earlier found by carrying out unit root test for the selected market and found integrated of order one, *i.e.* I (1) or ‘d = 1’ as mentioned in Table 1. Because the series is non-stationary, differencing was used to convert it to a stationary series. The Auto Correlation Function (ACF) helps in choosing relevant ordering values for moving average (MA) and Partial Autocorrelation Functions (PACF) for autoregressive (AR). The ACF and PACF coefficients were found to be under the standard error limits in all cases. ACF and PACF values were retrieved for 24 lags length which are illustrated in Table 2.

Estimation of parameter

The model having lowest values of MAPE & comparatively lower values of Akaike Information Criteria (AIC), Mean Absolute Error (MAE) & Root Mean Square Error (RMSE) were used as a criteria to determine optimum model for forecasting. ARIMA (3, 1, 2) model emerged as a best fit model out of several tried model. Least Squares Estimation for onion pricing is shown in Table 4. Almost all the coefficient of the ARIMA (3,1,2) model were found to be statistically significant at 5 percent level of significance except few parameter as illustrated in Table 4. Hence, the ARIMA (3, 1, 2) model can be utilized as the best model to forecast the prices of onion in Mahuva market.

Diagnostic checking of the fitted ARIMA model

The autocorrelation of various lags of the residuals of ARIMA (3, 1, 2) model were estimated up to 24 lags. The graphical representation of residuals of the ARIMA (3, 1, 2) is shown in Figure 1. Except lag 8, 10 and lag 22, none of the other lag residuals found outside the confidence limit, thus explaining absence of autocorrelation. Therefore, to confirm

the absence of autocorrelation among residuals Ljung-Box test statistics was carried out and it was found that residuals were non-significant at the 5 percent level of significance, highlighting absence of autocorrelation among error terms of the fitted model as mentioned in Table 5. This shows that the selected ARIMA (3, 1, 2) model was found appropriate & suitable model for forecasting the prices of onion.

Table 1: ACF & PACF values at level & first differenced price series for Mahuva market

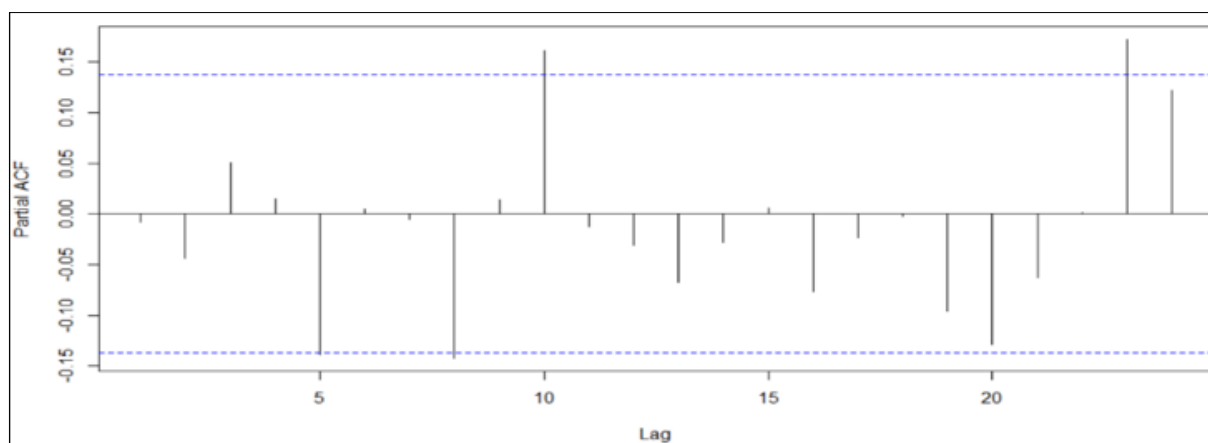
| Lag | ACF | | PACF | |
|-----|----------|---------------------|----------|---------------------|
| | At Level | At First Difference | At Level | At First Difference |
| 1 | 0.79 | 0.02 | 0.79 | 0.02 |
| 2 | 0.58 | -0.08 | -0.14 | -0.08 |
| 3 | 0.39 | -0.13 | -0.07 | -0.13 |
| 4 | 0.39 | -0.02 | 0.04 | -0.02 |
| 5 | 0.26 | -0.24 | -0.07 | -0.27 |
| 6 | 0.15 | 0.01 | 0.18 | -0.09 |
| 7 | 0.14 | -0.09 | -0.05 | -0.16 |
| 8 | 0.14 | -0.10 | 0.09 | -0.20 |
| 9 | 0.25 | -0.03 | 0.13 | -0.06 |
| 10 | 0.20 | 0.25 | 0.03 | 0.10 |
| 11 | 0.16 | -0.02 | -0.16 | -0.10 |
| 12 | 0.16 | -0.01 | 0.02 | -0.07 |
| 13 | 0.09 | 0.01 | -0.03 | -0.04 |
| 14 | 0.05 | 0.03 | 0.02 | -0.05 |
| 15 | 0.02 | -0.01 | -0.01 | 0.02 |
| 16 | -0.06 | -0.01 | -0.08 | -0.09 |
| 17 | -0.02 | -0.01 | 0.04 | 0.09 |
| 18 | -0.03 | -0.03 | -0.07 | -0.01 |
| 19 | -0.03 | -0.0 | -0.06 | -0.15 |
| 20 | -0.01 | -0.10 | 0.08 | -0.20 |
| 21 | 0.07 | -0.06 | 0.17 | -0.17 |
| 22 | 0.18 | 0.012 | 0.16 | -0.12 |
| 23 | 0.29 | 0.16 | 0.12 | 0.01 |
| 24 | 0.32 | -0.01 | -0.05 | 0.03 |

Table 2: Summary of the ARIMA model for onion price for Mahuva market

| Model | AIC | MAPE | MAE | RMSE |
|---------------|--------|-------|--------|--------|
| ARIMA (3,1,2) | 3044.5 | 21.76 | 217.87 | 421.40 |

Table 3: Parameter estimates for fitted ARIMA (3, 1, 2) model for onion prices for Mahuva market

| Variables | Estimates | Standard Error | t Ratio | Probability |
|-----------|-----------|----------------|---------|-------------|
| AR1 | -0.01 | 0.09 | 0.14 | 0.88 |
| AR2 | 0.68 | 0.06 | 9.95 | < 0.01 |
| AR3 | -0.25 | 0.07 | 3.46 | 0.05 |
| MA1 | -0.05 | 0.07 | 0.77 | 0.43 |
| MA2 | -0.87 | 0.08 | 12.24 | < 0.01 |



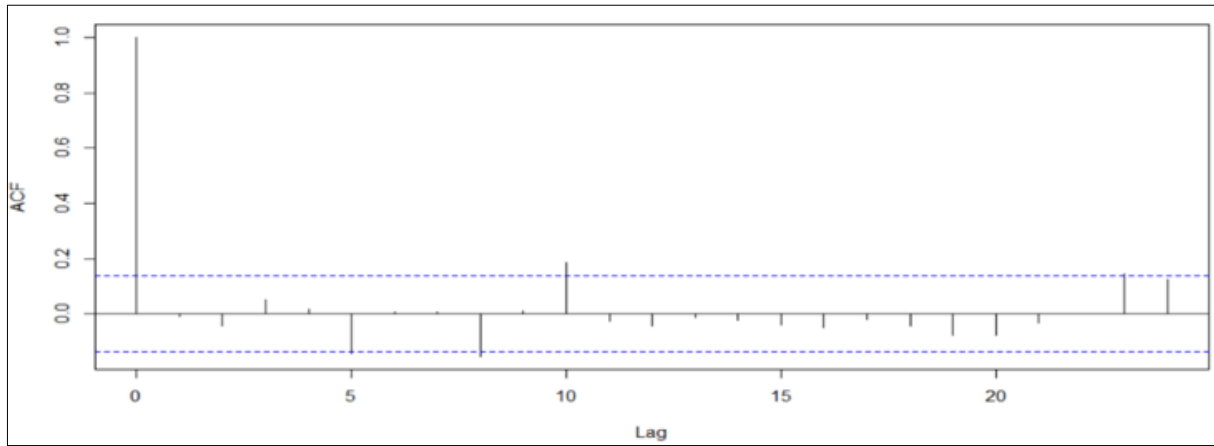


Fig 1: Residual correlogram of ACF & PACF for fitted ARIMA (3, 1, 2) model

Table 4: AIC & Ljung-Box test statistic value for the fitted ARIMA (3, 1, 2) model for onion prices for Mahuva market

| | | |
|---------------------|--|---------------------------|
| Fitted model | (3,1,2) | |
| AIC | 3044.5 | |
| Ljung-Box | 31.26 (test statistics) (Test Statisticis) | 0.14 (p-value) (p- value) |

Similar steps were carried out for Ahmedabad and Gondal market as well. Details are illustared in the table Ahmedabad

Table 5: ACF & PACF values at level & first differenced price series for Ahmedabad market

| Lag | ACF | | PACF | |
|-----|----------|---------------------|----------|---------------------|
| | At Level | At First Difference | At Level | At First Difference |
| 1 | 0.82 | 0.82 | 0.82 | 0.16 |
| 2 | 0.60 | -0.25 | -0.25 | -0.09 |
| 3 | 0.38 | -0.09 | -0.09 | -0.05 |
| 4 | 0.21 | -0.03 | -0.03 | -0.21 |
| 5 | 0.12 | 0.11 | 0.11 | -0.07 |
| 6 | 0.08 | 0 | 0 | -0.13 |
| 7 | 0.06 | 0.03 | 0.03 | -0.21 |
| 8 | 0.10 | 0.15 | 0.15 | -0.09 |
| 9 | 0.17 | 0.08 | 0.08 | -0.01 |
| 10 | 0.20 | -0.06 | -0.06 | 0.01 |
| 11 | 0.19 | -0.07 | -0.07 | -0.14 |
| 12 | 0.15 | 0.04 | 0.04 | -0.06 |
| 13 | 0.09 | -0.04 | -0.04 | -0.03 |
| 14 | 0.05 | 0.03 | 0.03 | -0.01 |
| 15 | 0.07 | -0.07 | -0.07 | -0.06 |
| 16 | -0.03 | -0.06 | -0.006 | -0.04 |
| 17 | -0.05 | 0.01 | 0.01 | 0.02 |
| 18 | -0.06 | -0.05 | -0.05 | -0.09 |
| 19 | -0.05 | 0.02 | 0.02 | -0.16 |
| 20 | -0.05 | 0.11 | 0.11 | -0.16 |
| 21 | 0.08 | 0.17 | 0.17 | -0.15 |
| 22 | 0.21 | 0.16 | 0.16 | -0.11 |
| 23 | 0.33 | 0.10 | 0.10 | 0.06 |
| 24 | 0.38 | -0.06 | -0.06 | 0.10 |

Table 6: Summary of the ARIMA model for onion price for Ahmedabad market

| Model | AIC | MAPE | MAE | RMSE |
|---------------|---------|-------|--------|--------|
| ARIMA (2,1,1) | 3092.02 | 22.99 | 273.93 | 478.88 |

Table 7: Parameter estimates for fitted ARIMA (2, 1, 1) model for onion price in Ahmedabad market

| Variables | Estimates | Standard Error | t Ratio | Probability |
|-----------|-----------|----------------|---------|-------------|
| AR1 | 1.03 | 0.06 | 15.14 | <0.01 |
| AR2 | -0.29 | 0.08 | 4.32 | <0.01 |
| MA1 | -0.96 | 0.02 | 46.74 | < 0.01 |

Table 8: AIC & Ljung-Box test statistic value for the fitted ARIMA (2, 1, 1) model for Ahmedabad market

| | | |
|---------------------|---|--------------------------------------|
| Fitted model | (2, 1, 1) | |
| AIC | 3092.0 | |
| Ljung-Box | 31.38 (test statistics) (p- value) (Test Statistic) | 0.14 (p-value) (p- value) (P- value) |

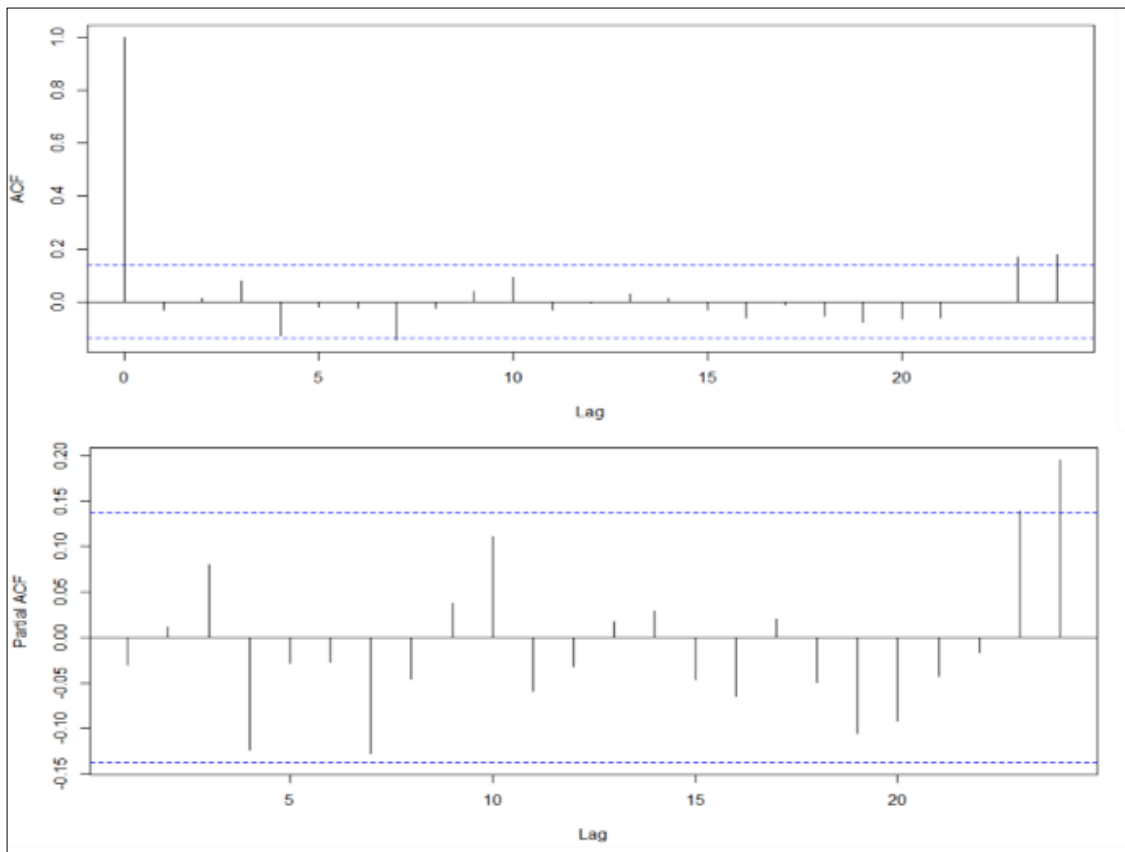


Fig 2: Residual correlogram of ACF & PACF of fitted ARIMA (2, 1, 1) model

Table 9: ACF & PACF values at level & first differenced of onion price series for Gondal market

| Lag | ACF | | PACF | |
|-----|----------|---------------------|----------|---------------------|
| | At Level | At First Difference | At Level | At First Difference |
| 1 | 0.78 | 0.08 | 0.78 | 0.08 |
| 2 | 0.52 | -0.06 | -0.24 | -0.06 |
| 3 | 0.29 | -0.14 | -0.08 | -0.13 |
| 4 | 0.13 | -0.19 | 0.01 | -0.17 |
| 5 | 0.06 | -0.09 | 0.05 | -0.09 |
| 6 | 0.03 | -0.05 | -0.01 | -0.09 |
| 7 | 0.02 | -0.09 | 0.02 | -0.16 |
| 8 | 0.05 | -0.08 | 0.09 | -0.16 |
| 9 | 0.11 | -0.08 | 0.10 | -0.19 |
| 10 | 0.21 | 0.12 | 0.14 | 0.05 |
| 11 | 0.25 | 0.2 | -0.05 | 0.06 |
| 12 | 0.20 | -0.01 | -0.12 | -0.15 |
| 13 | 0.12 | 0.02 | 0.02 | -0.03 |
| 14 | 0.06 | 0.02 | 0.05 | 0.02 |
| 15 | 0 | -0.03 | -0.09 | -0.04 |
| 16 | -0.04 | -0.06 | -0.02 | -0.10 |
| 17 | -0.05 | 0.01 | 0.04 | 0.01 |
| 18 | -0.07 | -0.05 | -0.06 | -0.04 |
| 19 | -0.06 | -0.10 | 0.02 | -0.11 |
| 20 | -0.01 | -0.11 | 0.06 | -0.16 |
| 21 | 0.09 | -0.07 | 0.13 | -0.21 |
| 22 | 0.22 | 0.06 | 0.18 | -0.08 |
| 23 | 0.33 | 0.20 | 0.10 | 0.07 |
| 24 | 0.34 | 0.26 | -0.07 | 0.01 |

Table 10: Summary of the ARIMA model for onion price for Gondal market

| Model | AIC | MAPE | MAE | RMSE |
|--------------|------|-------|--------|--------|
| ARIMA(2,1,2) | 3090 | 27.51 | 241.99 | 473.79 |

Table 11: Parameter estimates for fitted ARIMA (2, 1, 2) model for Gondal market

| Variables | Estimates | Standard Error | t ratio | Probability |
|-----------|-----------|----------------|---------|-------------|
| AR1 | 1.37 | 0.16 | 8.44 | <0.01 |
| AR2 | -0.58 | 0.12 | 4.84 | <0.01 |
| MA1 | -1.43 | 0.18 | 7.71 | <0.01 |
| MA2 | 0.45 | 0.18 | 2.52 | 0.01 |

Table 12: AIC & Ljung-Box test statistic value for the selected ARIMA (2, 1, 2) model for Gondal market

| Fitted model | (2,1,2) | |
|--------------|-------------------------|-----------------|
| AIC | 3092.0 | |
| Ljung-Box | 28.16 (test statistics) | 0.25 (p- value) |

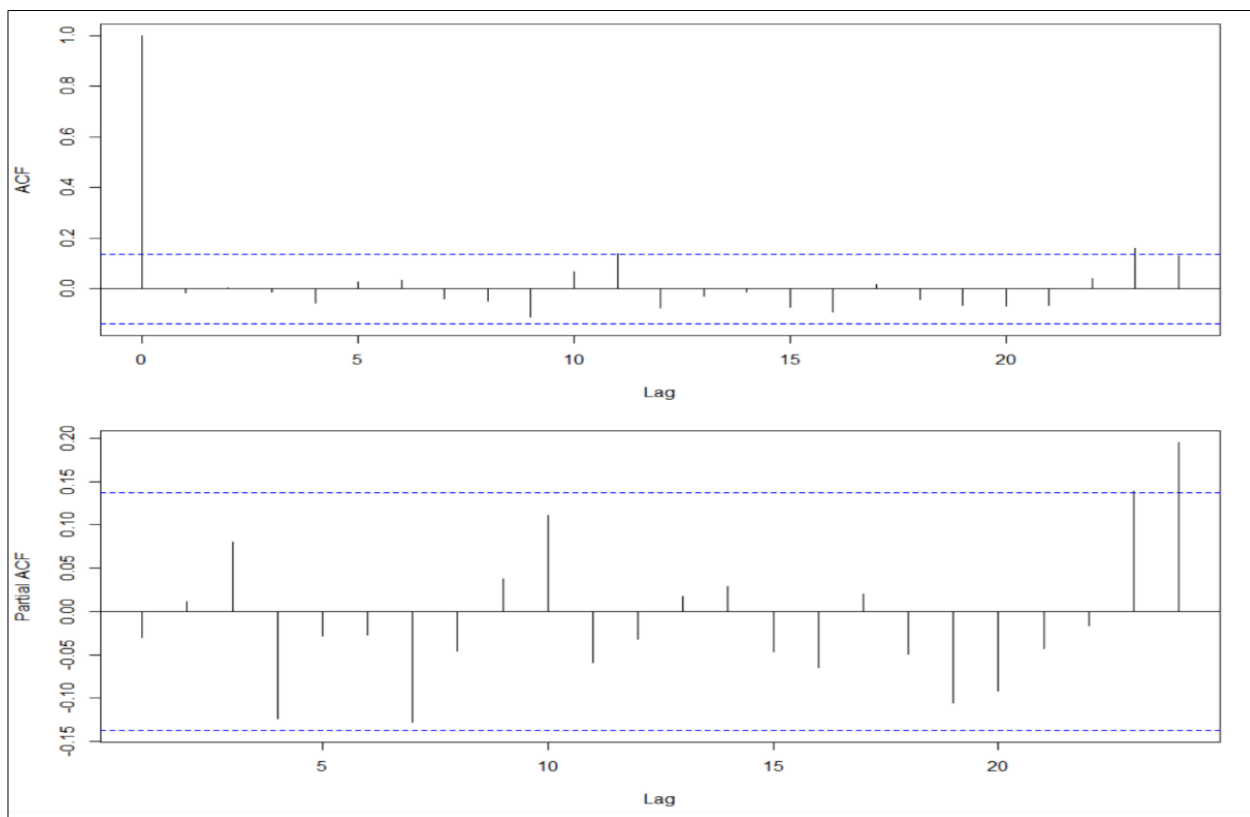


Fig 3: Residual correlogram of ACF & PACF of fitted ARIMA (2, 1, 2) model

Post sample period forecast

Table 13: Forecasting of onion prices by for the year 2021 using best fitted ARIMA model

| Month & Year | Forecast (Rs./q) |
|----------------------------|------------------|
| Mahuva (3, 1, 2) | |
| Jan-21 | 1216 |
| Feb-21 | 1468 |
| Mar-21 | 1206 |
| Ahmedabad (2, 1, 1) | |
| Jan-21 | 2713 |
| Feb-21 | 2562 |
| Mar-21 | 1076 |
| Gondal (2, 1, 2) | |
| Jan-21 | 2567 |
| Feb-21 | 2268 |
| Mar-21 | 1304 |

Forecasting accuracy

The accuracy of forecast models was evaluated using error measures, such as Mean Absolute Percentage Error. The findings are organised under the Table 15.

Table 14: Accuracy of the different forecast model

| Market | MAPE (%) |
|-----------|----------|
| | ARIMA |
| Mahuva | 21.77 |
| Ahmedabad | 22.99 |
| Gondal | 27.51 |

Among all the selected markets, forecasted onion prices were reportedly found better with less MAPE value. The results were found contrary to the findings of Shruthi (2015) [10]. Thus explaining the need to use more accurate & elaborative model which can capture the volatility in onion prices.

Summary and Conclusion

For the Mahuva market, the ARIMA model (3, 1, 2) emerged as the most effective, demonstrating its ability to capture and predict the intricate patterns in onion price fluctuations in that specific region. On the other hand, the Ahmedabad market exhibited optimal forecasting results with the implementation of the ARIMA model (2, 1, 1). This model's performance highlights its suitability for capturing the unique dynamics and trends influencing onion prices in Ahmedabad, showcasing the importance of tailoring modeling choices to specific market characteristics. In the case of the Gondal market, the ARIMA model (2, 1, 2) stood out as the most proficient in forecasting onion prices. Its success underscores the significance of considering different model specifications for accurate predictions, as market-specific factors play a crucial role in determining the effectiveness of forecasting models. Ultimately, the study emphasizes the need for a nuanced approach to time series forecasting, acknowledging the diversity of market behaviors and dynamics. By tailoring ARIMA model parameters to specific markets, stakeholders can enhance the precision of onion price predictions and make more informed decisions in the volatile agricultural market.

References

1. Saxena R, Chand R. Understanding the Recurring Onion Price Shocks: Revelations from Production-Trade-Price Linkages. Policy Paper 33, ICAR-National Institute of Agricultural Economics & Policy Research (NIAP), New Delhi; c2017.
2. Pal V. Price forecasting of Brinjal & Chilli (Green) - A Statistical Evaluation. M.Sc. (Unpublished Thesis), Anand Agricultural University, Anand; c2019.
3. Ljung GM, Box GE. On a measure of lack of fit in time series models. *Biometrika*. 1978;65(2):297-303.
4. Jadav CM, Reddy BVC, Gaddi GM. Application of ARIMA model for forecasting agricultural prices. *Journal of Agricultural Science & Technology*. 2017;19(4):981-992.
5. Mulla A, Seelam R, Nalamaru V, Naidu GM. Price behaviour & forecasting of onion prices in Kurnool market, Andhra Pradesh state. *Agricultural Economics & Social Science Research Association (AESSRA)*. 2020;65(1):43-50.
6. Pradeep M. Statistical model for forecasting arrival & price behaviour of potato in major regular markets of Karnataka (Master's thesis, University of Agricultural Sciences, Bengaluru); c2015.
7. Saxena R, Paul RK, Pavithra S, Singh NP, Kumar R. Market intelligence in India: price linkages & forecasts. Policy paper 34, ICAR-National Institute of Agricultural Economics & Policy Research (NIAP), New Delhi; c2019.
8. Jadhav V, Reddy BVC, Sakamma S. Forecasting Monthly Prices of Arecanut and Coconut Crops in Karnataka. *International Journal of Agricultural and Statistical Sciences*. 2013;9(2):597-606.
9. Xin W, Can W. Empirical study on agricultural products price forecasting based on internet-based timely price information. *International Journal of Advanced Science & Technology*. 2016;87:31-36.
10. Nethravathi PC, Shruthi GS, Suresh D, Nagabhushana H, Sharma SC. *Garcinia xanthochymus* mediated green synthesis of ZnO nanoparticles: Photoluminescence, photocatalytic and antioxidant activity studies. *Ceramics International*. 2015 Aug 1;41(7):8680-7.