Journal of Experimental Agriculture International

Volume 46, Issue 5, Page 146-154, 2024; Article no.JEAI.114606
ISSN: 2457-0591
(Past name: American Journal of Experimental Agriculture, Past ISSN: 2231-0606)

A Comparative Analysis of Time Series Models for Onion Price Forecasting: Insights for Agricultural Economics

Vinay H T [a++*], Pavithra V [a++], Jagadeesh M S [b], G. Avinash [c] and Harish Nayak G H [c]

[a] Department of Agricultural Statistics, Uttar Banga Krishi Viswavidyalaya, Cooch Behar-736165, West Bengal, India.
[b] Division of Agricultural Economics, ICAR-Indian Agricultural Research Institute, New Delhi-110012, India.
[c] ICAR-Indian Agricultural Statistics Research Institute, New Delhi-110012, India.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JEAI/2024/v46i52365

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/114606

Original Research Article

Received: 06/01/2024
Accepted: 11/03/2024
Published: 14/03/2024

ABSTRACT

Onion price forecasting plays a critical role in agricultural planning, market stability, and consumer welfare. By predicting onion prices, stakeholders can make informed decisions regarding planting, harvesting, trading, and consumption, mitigating risks and ensuring sustainable supply chains. The present study aims to forecast monthly wholesale onion prices in Bangalore market by using various statistical techniques like Exponential Smoothing, ARIMA, SARIMA, BATS and TBATS models. The time series price data from January 2010 to December 2023 was utilized for the study. Models were trained on 80% of the data and validated on the remaining 20%. The performance of each model was compared based on the two metrics like Root Mean Square Error (RMSE) and

[++] *Ph.D. Scholar;*
**Corresponding author: E-mail: vinaygkvkstat@gmail.com;*

Mean Absolute Percentage Error (MAPE). Results revealed that the TBATS model was performing better as compared to other six models with less RMSE of 281.30 and MAPE of 0.1544 Additionally, the residuals of the TBATS model exhibited a random distribution. Therefore, the onion prices for January 2024 to December 2024 were forecasted using the TBATS model.

Keywords: Forecasting; ARIMA; SARIMA; BATS; TBATS.

1. INTRODUCTION

Onion is one of the most important vegetable crops for household consumption and also for foreign exchange earner among the fruits and vegetables. It is considered as a most sensitive commodity due to sudden price fluctuations [1]. Karnataka, with an area of 230.40 hectares, is the second-largest cultivator of onions after Maharashtra, producing a yield of 11,548 quintals (Agricultural statistics at a glance) [2]. Price forecasting of agricultural commodities, such as onions, holds significant importance in agricultural economics, particularly in regions like Bangalore, Karnataka, where the onion market plays a crucial role in the local economy.

Price forecasting is a critical component in the economic management of a nation, particularly in the agricultural sector where prices can fluctuate across various markets within a country. These fluctuations are influenced by factors such as supply and demand dynamics, weather conditions, transportation costs, government policies, and market infrastructure. Therefore, accurate forecasting techniques, including statistical models, econometric models, and machine learning algorithms, are essential for predicting future price movements. Understanding regional differences and incorporating local factors into forecasting models is also crucial for precise predictions. By effectively forecasting agricultural commodity prices, policymakers and stakeholders can make informed decisions to mitigate risks, optimize resource allocation, and ensure stability in agricultural markets, thus contributing to the overall economic stability of the nation. Price drops directly impact the farmers when the price is lower than the cost of cultivation and on the other hand price spikes can disrupt consumer budgets. A forehand estimate can help to adapt to probable changes depending on the trend, seasonality or cyclical pattern. Forecasting using several statistical techniques empowers the analysts with the ability to make forehand preparations in times of prospective predicted changes.

This study aims to utilize statistical models for forecasting and to identify the model that best captures variations in the data. Modeling serves as a fundamental analytical tool in modern statistical analysis. Exponential smoothing models are forecasting techniques that assign weights to observed time series data based on the timing of observations. Recent observations are given higher weights, while past observations receive less weight. In this analysis, various exponential smoothing methods such as Simple Exponential Smoothing, Double Exponential Smoothing, and Triple Exponential Smoothing are employed. Additionally, the classical time series Autoregressive Integrated Moving Average (ARIMA) model and Seasonal Autoregressive Integrated Moving Average (SARIMA) model is utilized, where present observations are regressed on past observations and error terms. Furthermore, BATS and TBATS models helped improve forecasting accuracy and capture complex seasonal behaviour in time series data. Different forecasting methods are employed for forecasting the commodity prices in literature. Shahini [3] forecasted potato price using Holt-Winters and ARIMA models. Kozuck et al. [4] compared the neural networks with classical approaches like ARIMA, BATS and TBATS in case of timber price forecasting. Various researchers like Areef [5], Suresha et al. [6], Beniwal [7], and Kumar and Kumar [8] employed different statistical and machine learning techniques to forecast the onion prices in different markets of India.

2. METHODOLOGY

2.1 Data Description

Data on onion prices spanning from January 2010 to December 2023 were gathered from the AGMARKNET website (http://agmarknet.gov.in), with a particular emphasis on the Bangalore market in Karnataka state. These prices reflect the modal prices in a given month, which are deemed superior to monthly average prices as they represent the predominant proportion of the commodity marketed during that month.

2.2 Methods

The prices of onion were modelled using Exponential Smoothing methods, ARIMA model, SARIMA model, BATS model and TBATS model.

2.2.1 Single Exponential Smoothing (SES)

In 1963, Brown [9] introduced a method to estimate future values using a single weight or parameter. This technique assigns more weightage to recent observations and less weightage to distant observations.

$$F_{t+1} = \alpha Y_t + (1-\alpha)F_t$$

Where, α is a smoothing parameter taking values in the interval (0, 1).

F_t = the forecasted price at time t.
Y_t =actual price at time t.

2.2.2 Double Exponential Smoothing (DES)

It is also known as Holt's linear method. Simple exponential smoothing (SES) is not effective in predicting time series with a local linear trend to address this limitation, Holt [10] proposed an extension of SES called Holt's method. This method includes an additional updating equation for the slope (trend), resulting in improved forecasts for time series with a local linear trend.

$$L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + b_{t-1})$$
$$b_t = \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1}$$
$$F_{t+m} = L_t + mb_t$$

Where,

L_t= Level at time t
b_t= Trend at time t
F_{t+m}= Forecast value for m period ahead
α, β = Smoothing parameters ranging from 0 to 1. The combination of these parameters is selected based on minimum RMSE value.

2.2.3 Triple Exponential Smoothing (TES)

It is also known as Holt-Winter's Exponential Smoothing (H-WES) method. The Holt- Winter's method is based on three smoothing equations, one for level, one for trend, and one for seasonality. It is similar to Holt's method, with an additional equation to deal with seasonality. Holt-Winter's Exponential Smoothing (H-WES) methods are widely used when the data shows trend and seasonality. In this study the multiplicative model is used as the seasonal variation over time is observed. The four equations for the model are given as follows:

$$L_t = \alpha \frac{Y_t}{S_t} + (1-\alpha)(L_{t-1} + b_{t-1})$$
$$b_t = \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1}$$
$$S_t = \gamma \frac{Y_t}{L_t - Y_t} + (1-\gamma)S_{t-s}$$
$$F_{t+m} = L_t + mb_t + S_{t-s+m}$$

Where,

s= length of seasonality
L_t= Level at time t
b_t= Trend at time t
S_t= Seasonal component at time t
F_{t+m}= Forecast value for m period ahead
α, β and γ are level, trend and seasonal smoothing constants or the weights respectively, which lies between 0 and 1. The combination of these parameters is selected based on minimum RMSE and MAPE value.

2.2.4 Autoregressive Integrated Moving Average (ARIMA) model

In an autoregressive integrated moving average model, the future value of a variable is assumed to be a linear function of several past observations and random errors (Box and Jenkins) [11]. Theoretically ARIMA model includes three components: Auto-Regressive (AR), Moving-Average (MA), and Integrated (I) terms. The first two components are expressed in equation.

$$\nabla^d y_t = \underbrace{\emptyset_1 \nabla^d y_{t-1} + \cdots + \emptyset_p \nabla^d y_{t-p}}_{\text{AR terms}} + \varepsilon_t - \underbrace{\theta_1 \varepsilon_{t-1} - \cdots - \theta_p \varepsilon_{t-p}}_{\text{MA terms}}$$

where ϕ is a number strictly between −1 and +1, and θ are the weights, and *p* is the order of the AR model, and *q* is the order of the MA model. Here, ε_t 's are independently and normally distributed with zero mean and constant variance σ^2 $\forall$ *t* =1,2, …, *n*.

2.2.5 Seasonal Autoregressive Integrated Moving Average (SARIMA) model

When time series data have seasonal component, SARIMA model is employed. SARIMA model is characterized by SARIMA (*p*, *d, q*) $(P, D, Q)_s$ Here, *p* and *q* are orders of non-seasonal autoregressive and moving average parameters respectively, whereas *P* and *Q* are the seasonal autoregressive and moving average parameters respectively. Also 'd' and '*D*' denote non-seasonal and seasonal differences respectively Makridakis *et al.,* [12]) and it is given by

$$(1-\varphi_p B)(1-\Phi_P B^s)(1-\mathrm{B})(1-B^s)y_t = (1-\theta_q B)(1-\Theta_Q B^s)\varepsilon_t$$

where,

B=backshift operator,
s =seasonal lag,
ε_t= sequence of error ~ N (0, σ^2),
Φ's and *φ's* = the seasonal and non-seasonal autoregressive parameters respectively
Θ's and *θ's* = the seasonal and non-seasonal moving average parameters.

2.2.6 BATS (B: Box-Cox transformation A: ARIMA errors T: Trend S: Seasonal components) model

The Box-Cox transformation is a method used to stabilize the variance and mean of a time series by applying a power transformation. It helps in making the series stationary. The BATS model, an extension of the Double-Seasonal Holt-Winter's (DSHW) method, incorporates the Box-Cox transformation, ARMA errors, trend, and multiple seasonal patterns to improve forecasting accuracy and capture complex seasonal behaviour in time series data. De Livera, [13].

$$y_t^{(\omega)} = \begin{cases} \frac{y_t^{\omega}-1}{\omega};\ \omega \neq 0 \\ \log y_t \quad \omega = 0 \end{cases}$$
$$y_t^{(\omega)} = l_{t-1} + \emptyset b_{t-1} + \sum_{i=1}^{T} s_{t-m_i}^{(i)} + d_t$$
$$l_t = l_{t-1} + \emptyset b_{t-1} + \alpha d_t$$
$$b_t = (1-\emptyset)b + \emptyset b_{t-1} + \beta d_t$$
$$s_t^{(i)} = s_{t-m_i}^{(i)} + \gamma_i d_t$$
$$d_t = \sum_{i=1}^{p} \varphi_i\, d_{t-i} + \sum_{i=1}^{q} \theta_i\, \varepsilon_{t-i} + \varepsilon_t$$

where, $y_t^{(\omega)}$ represents Box–Cox transformed observations with a parameter ω at time *t*, $m_{1,\dots,}$ m_T denote the seasonal periods, l_t is the local level at time *t, b* is the long-run trend and b_t is the short-run trend at time *t*, $s_t^{(i)}$ indicates the *i*th seasonal component at time *t*, d_t represents an ARMA *(p, q)* process, ε_t is a Gaussian white-noise process with zero mean and constant variance σ^2, and the smoothing parameters are given by *α*, *β*, and γ_i for *i*= 1,…, *T.* The model was represented by BATS (*ω*,(*p*, *q*), ∅, m_1, m_2 ,…, m_T), where, *ω* is the Box–Cox transformed value, (*p*, *q*) is ARMA components, ∅ dampening parameter, m_i represents *i*th seasons.

2.2.7 TBATS (T: Trigonometric B: Box-Cox transformation A: ARIMA errors T: Trend S: Seasonal components) model

To address issues with high frequency and non-integer seasonality, the TBATS (Trigonometric seasonality, Box-Cox transformation, ARMA errors, Trend, and Seasonal components) model was introduced as an extension of the BATS (Box-Cox transformation, ARMA errors, Trend, and Seasonal components) model. It adapts the following equations: De Livera *et al.,* [14].

$$s_t^{(i)} = \sum_{j=1}^{k_i} s_{j,t}^{(i)}$$
$$s_{j,t}^{(i)} = s_{j,t-1}^{(i)} \cos\lambda_j^{(i)} + s_{j,t-1}^{*(i)} \sin\lambda_j^{(i)} + \gamma_1^{(i)} d_t$$
$$s_{j,t}^{*(i)} = -s_{j,t-1} \sin\lambda_j^{(i)} + s_{j,t-1}^{*(i)} \cos\lambda_j^{(i)} + \gamma_2^{(i)} d_t$$

where, $\gamma_1^{(i)}$ and $\gamma_2^{(i)}$ are the smoothing parameters, $\lambda_j^{(i)} = \frac{2\pi j}{m_i}$, $s_{j,t}^{(i)}$ describe the stochastic level of the *i*th seasonal component, $s_{j,t}^{*(i)}$ describe the stochastic growth of the *i*th seasonal component, k_i is the number of harmonics required for the *i*th seasonal component, $k_i = \frac{m_i}{2}$ for even values of m_i, and $k_i = \frac{(m_i-1)}{2}$ for odd values of m_i.

2.2.8 Model evaluation criteria

The model performance was evaluated on important criteria like Root Mean Square Error (RMSE).

$$\text{RMSE}= \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2}$$

where, *n* is number of observations, Y_t and $\hat{Y}_t$ are the actual and forecasted price at time *t*. lower value for the metrics denotes a more accurate prediction from the model.

3. RESULTS AND DISCUSSION

3.1 Descriptive Statistics

The descriptive statistics is depicted in Table 1. Which reveals that the average price of onion in

Bangalore market is 1447 Rs/quintal and the price ranges between 461 to 8568 during the study period and high CV of 69.88% indicates so much variations in the price data. Time plot of the average monthly onion price for the original series is presented in Fig. 1.

Table 1. Descriptive statistics of Onion price in Bangalore market of Karnataka

Statistics	Price
Observations	168
Mean (Rs/quintal)	1447
Minimum	461
Median	1130
Maximum	8568
SD	78.02
CV	69.88
Skewness	3.01
Kurtosis	15.14

3.2 Modelling of Price Data

Initially, Exponential Smoothing techniques were employed to model the price data. The estimated smoothing parameters for all three exponential smoothing models are presented in Table 2. Before fitting an ARIMA model, the Augmented Dickey-Fuller (ADF) test was employed to test the stationarity of the data. The data was found to be non-stationary; hence, first-order differencing was used to make the data stationary. The results of the ADF test are depicted in Table 3.

By using the auto. arima function in R software, we found that ARIMA (1,1,2) is the best model with a lower AIC value of 2279.316. The parameters estimated for this model are given in Table 4. To check the seasonality of the data, we used the Kruskal-Walli's test. The test suggests that there is seasonality; hence, we utilized SARIMA models to fit the data. Based on the lower AIC value, we suggest that ARIMA (1,1,2) (1,0,0) [12] is the best fit model, and the parameter estimates of the model are presented in Table 5.

Further, we employed BATS and TBATS models to deal with high frequency and non-integer seasonality present in the data. The parameter estimates of both the BATS and TBATS models are depicted in Table 6.

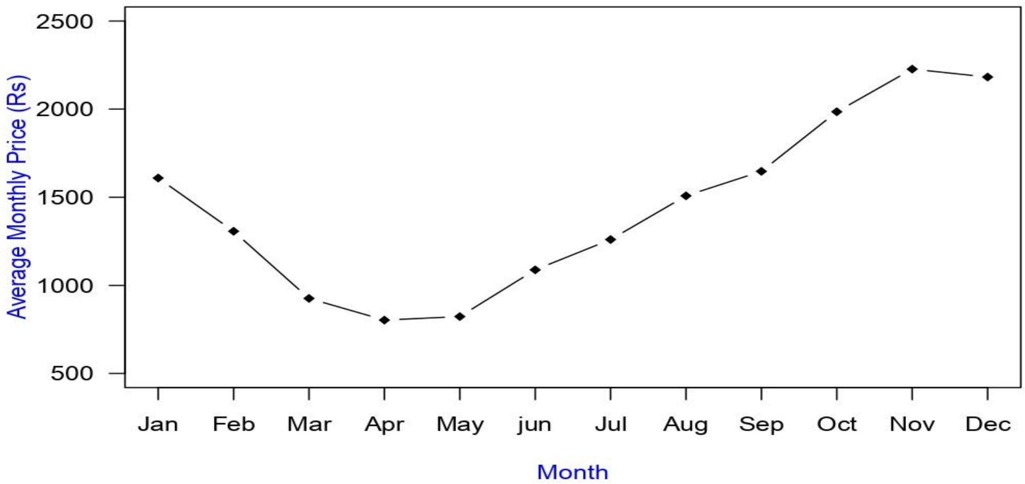

Fig. 1. Monthly average wholesale price of Onion in Bangalore market, Karnataka
3.2 Modelling of Price data

Table 2. Estimated parameters of the Exponential Smoothing models

Model	Parameter	Estimate	AIC
SES	Alpha (Level)	0.9999	2628.066
DES	Alpha (Level)	0.9999	2632.183
	Beta (Trend)	0.0001	
TES	Alpha (Level)	0.9991	2462.176
	Beta (Trend)	0.0082	
	Gamma (Seasonal)	0.0005	

Table 3. Augmented dicky fuller test results

Data	ADF test	Lag order	p-value
Original	-3.19	12	0.093
First Differenced	-4.94	12	0.01

Table 4. Estimated parameters of ARIMA (1,1,2) model

Parameters	Estimate	S.E.	p-value
AR1	0.53	0.10	0.000***
MA1	-0.53	0.11	0.000***
MA2	-0.43	0.10	0.000***

****: Significant at 0.1%*

Table 5. Parameter estimates of SARIMA (1,1,2) (1,0,0) $_{[12]}$ model

Parameters	Estimate	S.E.	p-value
AR1	0.54	0.10	0.000***
MA1	-0.54	0.11	0.000***
MA2	-0.42	0.10	0.000***
SAR1	-0.02	0.09	0.80

****: Significant at 0.1%*

3.3 Model Accuracy Evaluation

Model performance was evaluated using Root Mean Square Error (RMSE) and Mean Absolute Percent Error (MAPE) metrics. The fitted model along with its RMSE and MAPE is given in Table 7. TBATS model outperformed the other six models with less RMSE of 281.30 and MAPE of 0.1544 for the testing data set. The randomness of residuals of TBATS model was tested using Ljung- Box test and results presented in Table 8 indicates that the errors are random in nature hence onion prices can be forecasted using TBATS model.

3.4 Forecasting of Monthly Onion Prices

TBATS model was employed to forecast the monthly onion prices in Bangalore market from January 2024 to December 2024. The forecasted prices are depicted in Table 8 and the onion prices are expected to be high in November and December. The graph showing observed and fitted values is depicted in Fig. 2. These results will assist farmers and stakeholders in decision-making regarding planting, harvesting, trading, and consumption. This will help mitigate risks and ensure sustainable supply chains.

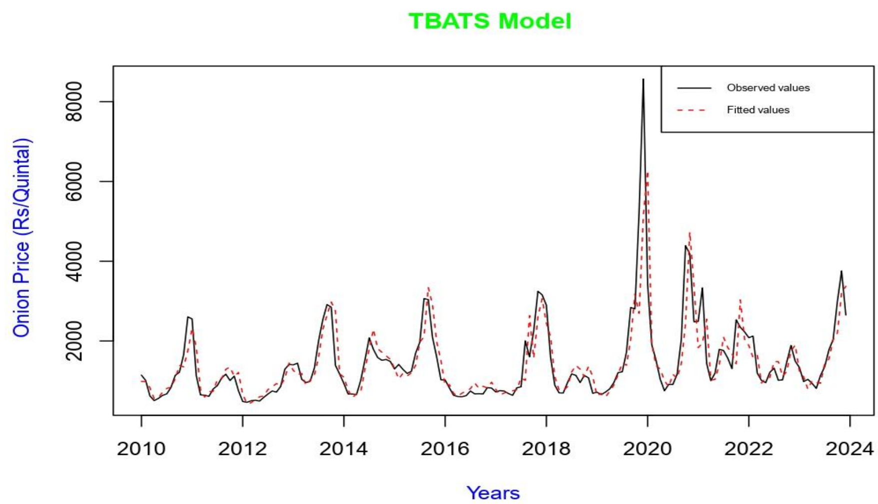

Fig. 2. Plot showing observed vs fitted values by TBATS model

Table 6. Parameter estimates of BATS and TBATS model

Model	***Box-Cox transformation (Omega)**	**Smoothing parameter**			**Phi**	**ARMA coefficients**		**Prediction error**	
		Alpha	**Beta**	**Gamma**		**AR coefficients**	**MA coefficients**	**Sigma**	**AIC**
BATS (0.002, {1,2}, -, {12})	0.002	-0.002	-	-0.108	-	0.865	0.247 -0.122	0.228	2368.70
TBATS (0, {0,0}, 0.8, {<12,2>})	0	1.396	-0.355	Gamma 1 -0.0004 Gamma 2 0.001	0.8	-	-	0.240	2368.95

Table 7. Model accuracy evaluation

Model	RMSE		MAPE	
	Training	Testing	Training	Testing
SES	749.56	407.61	0.2282	0.2510
DES	749.87	411.39	0.2311	0.2539
TES	592.86	425.27	0.2370	0.2676
ARIMA	673.35	328.83	0.2304	0.1975
SARIMA	673.17	329.03	0.2302	0.1981
BATS	559.33	294.99	0.1894	0.1750
TBATS	589.36	281.30	0.1713	0.1544

Table 8. Residual diagnostics test

Models	Ljung-Box test	
	Test statistic	p-value
TBATS	0.096	0.757

Table 9. Monthly wholesale prices of onion in Bangalore market forecasted using TBATS model

Month	Price (Rs/quintal)
Jan-24	1971
Feb-24	1473
Mar-24	1131
Apr-24	994
May-24	1032
Jun-24	1195
Jul-24	1414
Aug-24	1622
Sep-24	1810
Oct-24	1993
Nov-24	2112
Dec-24	2032

4. CONCLUSION

The study employed various statistical techniques, including Single Exponential Smoothing (SES), Double Exponential Smoothing (DES), Triple Exponential Smoothing (TES), Autoregressive Integrated Moving Average (ARIMA), Seasonal Autoregressive Integrated Moving Average (SARIMA), BATS, and TBATS, to model the monthly time series data of onion price. Among the fitted models, TBATS outperformed the other six models with a lower RMSE in the testing dataset. The residuals of the TBATS model were found to be random in nature, indicating its suitability for forecasting. Consequently, it was selected to forecast the monthly wholesale prices of onions in the Bangalore market from January 2024 to December 2024. The study predicts that onion prices are expected to be high in November and December month of 2024. These results will assist farmers in planning production and in assessing and managing price risks.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Sinha K, Panwar S, Alam W, Singh KN, Gurung B, Paul RK, Mukherjee A. Price volatility spillover of Indian onion markets: A comparative study. Indian Journal of Agricultural Sciences. 88(1):114-20.
2. Statistics A. Statistics at a Glance. Statistics Division, Directorate of Economics & Statistics, Ministry of Agriculture, Government of India; 2023.
3. Sahinli MA. Potato price forecasting with Holt-Winters and ARIMA methods: A case

study. American Journal of Potato Research. 2020;97(4):336-46.

4. Kozuch A, Cywicka D, Adamowicz K. A Comparison of Artificial Neural Network and Time Series Models for Timber Price Forecasting. Forests. 2023;14(2):177.
5. Areef M, Rajeswari S, Vani N, Naidu GM. Forecasting of onion prices in Bangalore market: An application of time series models. Indian Journal of Agricultural Economics. 2020;75(2):217-27.
6. Suresha HP, Mutturaj U, Krishna KT. Indian Commodity Market Price Comparative Study of Forecasting Methods-A Case Study on Onion, Potato and Tomato. Asian J. Res. Comput. Sci. 2021;12:147-59.
7. Beniwal A, Poolsingh D, Shastry SS. Trends and Price Behaviour Analysis of Onion in India. Indian Journal of Agricultural Economics. 2022;77(4):632-42.
8. Kumar S, Kumar M. Forecasting of onion price in Patna district through Arima model. Environment and Ecology 2023; 41(3): 1299-1308.
9. Brown RG. Smoothing, forecasting and prediction of discrete time series. Courier Corporation; 2004.
10. Holt CC. Forecasting seasonals and trends by exponentially weighted moving averages. International journal of forecasting. 2004;20(1):5-10.
11. Box GE, Jenkins GM, Reinsel GC, Ljung GM. Time series analysis: forecasting and control. John Wiley & Sons; 2015.
12. Makridakis S, Wheelwright SC, Hyndman RJ. Forecasting methods and applications. John wiley & sons; 2008.
13. De Livera AM. Automatic forecasting with a modified exponential smoothing state space framework. Monash Econometrics and Business Statistics Working Papers. 2010;10(10):6.
14. De Livera AM, Hyndman RJ, Snyder RD. Forecasting time series with complex seasonal patterns using exponential smoothing. Journal of the American statistical association. 2011;106(496): 1513-27.

© Copyright (2024): Author(s). The licensee is the journal publisher. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/114606