



OPEN ACCESS

EDITED BY

Sendhil R,
Pondicherry University, India

REVIEWED BY

Aditya K. S,
Division of Agricultural Economics, Indian
Agricultural Research Institute (ICAR), India
Achal Lama,
Indian Council of Agricultural Research, India
Zheng Pan,
Zhongkai University of Agriculture and
Engineering, China

*CORRESPONDENCE

S. Vishnu Shankar
✉ s.vishnushankar55@gmail.com

RECEIVED 28 April 2023

ACCEPTED 07 August 2023

PUBLISHED 31 August 2023

CITATION

Shankar SV, Chandel A, Gupta RK, Sharma S,
Chand H, Kumar R, Mishra N,
Ananthakrishnan S, Aravinthkumar A,
Kumaraperumal R and Gowsar SRN (2023)
Exploring the dynamics of arrivals and prices
volatility in onion (*Allium cepa*) using advanced
time series techniques.
Front. Sustain. Food Syst. 7:1208898.
doi: 10.3389/fsufs.2023.1208898

COPYRIGHT

© 2023 Shankar, Chandel, Gupta, Sharma,
Chand, Kumar, Mishra, Ananthakrishnan,
Aravinthkumar, Kumaraperumal and Gowsar.
This is an open-access article distributed under
the terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Exploring the dynamics of arrivals and prices volatility in onion (*Allium cepa*) using advanced time series techniques

S. Vishnu Shankar^{1*}, Ashu Chandel¹, Rakesh Kumar Gupta¹,
Subhash Sharma², Hukam Chand³, Rakesh Kumar¹, Neha Mishra¹,
S. Ananthakrishnan⁴, A. Aravinthkumar⁵, R. Kumaraperumal⁶ and
S. R. Naffees Gowsar⁷

¹Department of Basic Sciences, Dr. YS Parmar University of Horticulture and Forestry, Solan, India, ²Department of Social Sciences, Dr. YS Parmar University of Horticulture and Forestry, Solan, India, ³Department of Environmental Science, Dr. YS Parmar University of Horticulture and Forestry, Solan, India, ⁴Department of Soil Science and Water Management, Dr. YS Parmar University of Horticulture and Forestry, Solan, India, ⁵Division of Plant Pathology, Indian Agricultural Research Institute, New Delhi, India, ⁶Department of Remote Sensing and Geographic Information System, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India, ⁷Department of Agricultural Statistics, Uttar Banga Krishi Viswavidyalaya, Cooch Behar, West Bengal, India

Modeling the arrivals and prices of agricultural commodities is an essential requirement for farmers, consumers, and governmental organizations to make informed decisions. This is particularly important for perishable commodities such as vegetables, where spoilage can lead to significant losses for farmers and have a ripple effect on supply and demand dynamics. Volatility in the arrivals and prices of vegetables like onion is a serious issue affecting the common person in different ways. The study attempts to employ different time series models like the autoregressive integrated moving average (ARIMA), Artificial neural network (ANN), hybrid, and ensemble empirical mode decomposition (EEMD) techniques to analyze the pattern and trend of onions in Chandigarh and Delhi markets. From the results of the study, the amount of volatility in the data was found to range from medium to high among the markets. Decomposition techniques such as EEMD-ARIMA and EEMD-ANN performed better for the study data with the least mean absolute percentage error (MAPE) values, such as 17.74 and 6.78% for arrivals and 9.76 and 10.24% for prices at Chandigarh and Delhi markets, respectively. The EEMD techniques excelled in handling the non-linearity and non-stationarity by decomposing the data into different intrinsic modes and a residual, providing a better understanding of the fluctuation levels of data.

KEYWORDS

ensemble empirical mode decomposition, non-linearity, non-stationary, onion, time series forecasting, volatility

Introduction

The arrivals and prices of agricultural commodities are integral to farming communities and hold substantial interest for stakeholders. They exert a direct influence on the selling and purchasing power of the populace, particularly small and marginal farmers, by impacting their real income (Allen, 1994). The uncertainty of production, poor

infrastructure, perishable nature, seasonality of production, and hikes in fuel and transport costs are the major reasons behind the fluctuations in the arrivals and prices of agricultural commodities (Paul et al., 2015). In recent times, India has witnessed some serious fluctuations in the arrivals and prices of vegetables, particularly onions. The prices of onions are highly volatile due to their high demand and unstable supply, which indirectly affects their production across many places (Rakshit et al., 2021). Onions (*Allium cepa*) occupy a preeminent position as a fundamental ingredient in diverse cuisines and are among the most widely grown and globally consumed vegetables. India is the second-largest onion producer in the world after China, accounting for around 20% of the global onion production. The major onion-cultivating (Kumar et al., 2022) states in India are Maharashtra, Karnataka, Gujarat, Madhya Pradesh, and Bihar. Although the crop can be grown in different seasons, the onion supply chain is susceptible to a range of external factors, such as weather disturbances and policy regulations (Saxena et al., 2019). Furthermore, the perishable nature of onions and the limited availability of modern cold storage facilities pose additional challenges for the supply chain. As a result of these factors, the price of onions is prone to exhibiting high volatility. Thus, the volatility in onions can impose substantial issues for both farmers and consumers (Kumar et al., 2021), making it harder to stabilize prices and maintain an adequate supply to meet the rising demand for fresh produce. Therefore, timely forecasting of the arrivals and prices is essential to understand the pattern of fluctuations to give assured foresight on production and marketing in the future (Areef et al., 2020). All of these can support the farmers and other stakeholders to manage the risks and optimize their operations by ensuring a stable supply of fresh produce for consumers.

Forecasting the arrivals and prices of agricultural commodities is crucial and difficult as they are prompted by various factors such as an imbalanced demand and supply (Arjun, 2013), problems of hedging and speculation, and various services imposed on goods. In addition, many random factors like drought, flood, famine, the incidence of pests, and diseases also have an effect. The recent outbreak of COVID-19 has worsened the pattern of demand and supply. All these factors have resulted in high fluctuations in the arrivals and prices, imposing the characteristics (Paul et al., 2023) of non-linearity, non-stationarity, and noise on the data. In recent times, different time series models have been introduced which can handle the problems of volatility better than conventional models. Autoregressive integrated moving average (ARIMA) is a commonly used traditional time series method that can handle the linearity in the data (Pardhi et al., 2018). Artificial neural networks (ANN) are regarded as black box techniques of machine learning models (Singh, 2008) because of their complex internal structure, numerous interconnected nodes, lack of transparency in the training process, and ability to capture complex non-linear relationships, (Anjoy et al., 2017) which makes it challenging to understand the prediction process and interpret the results. Additionally, the absence of explicit rules further reinforces the notion of ANNs being black boxes. Generalized autoregressive conditional heteroskedasticity (GARCH) is a variance function model that can be used only when the error variance of the model is autocorrelated (Kumar and Thenmozhi, 2014). It can be used as a hybrid technique with ARIMA as the mean model. Even the ARIMA and ANN

models can be used in combination to capture linear and non-linear information (Ghani and Rahim, 2019). Thus, all these models have their own advantages and disadvantages. Empirical mode decomposition (EMD), a self-adaptive technique (Wang et al., 2014), is applied to time series data which can decompose the non-linear and non-stationary into a combination of simple orthogonal times series components. These smoothing techniques can overcome the flaws in previous models and increase the forecasting accuracy. A novel method for spectrum analysis called the ensemble empirical mode decomposition (EEMD) is considered to be an improved version of EMD which fills the gaps in conventional techniques.

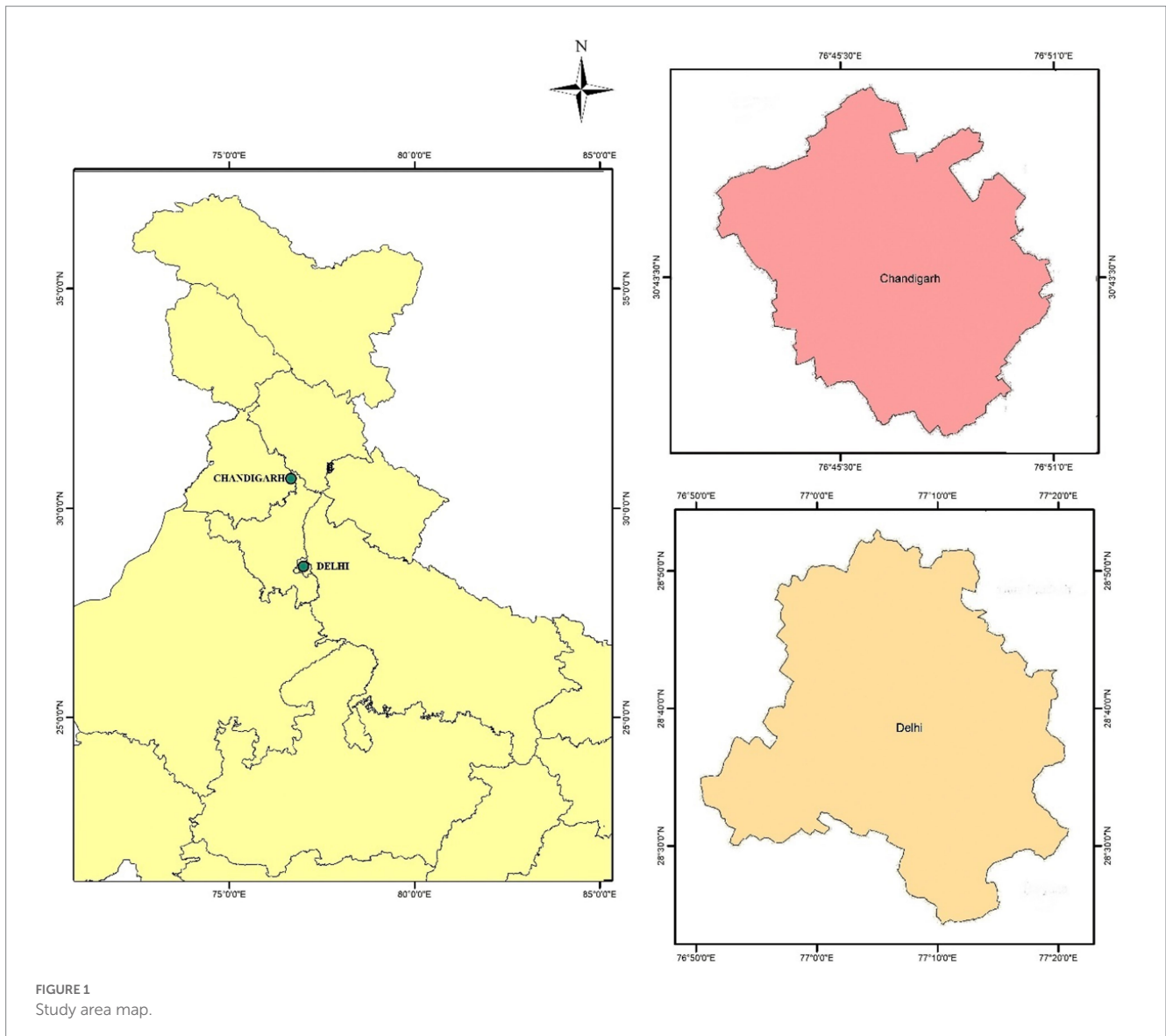
Thus, the study attempts to employ all possible time series techniques and examine the performance of the models on the basis of different error measurement criteria. Furthermore, the study emphasizes the effectiveness of the EEMD in handling complex time series data (Das et al., 2023). It is essential to recognize that there is no universally optimal model for all datasets. Instead, the data themselves determine the most suitable model based on its characteristics. Additionally, the study attempts to understand the pattern of volatility that existed in the arrivals and price of onions in Chandigarh and Delhi markets, where most previous research has dealt only with price series (Paul et al., 2023; Sinha, 2023). It has always been believed that the volatility in the prices is mainly due to fluctuations in the arrivals. To address this, the study endeavors to unravel the arrivals pattern, thereby providing insights into price trends. While prior research has predominantly concentrated on specific model types, such as ARIMA, ANN, hybrid, or GARCH (Kumar and Thenmozhi, 2014; Anjoy et al., 2017; Ghani and Rahim, 2019; Das et al., 2023; Paul et al., 2023), our study takes a holistic stance by comparing a wide range of models, aiming to underscore both their merits and drawbacks. The study's findings will contribute to a better understanding of the phenomenon of volatility in the onion market (Dahiya, 2022; Sujay et al., 2022), addressing the existing need to comprehend this aspect thoroughly.

Methodology

Different statistical techniques were employed to achieve the research objectives. Prior to analysis, it was ensured that all assumptions associated with the selected statistical techniques were adequately satisfied. The study was conducted in 2022 and involved the utilization of R software with different packages, in which the "Rlibeemd" package was used specifically for decomposing the data using ensemble empirical mode decomposition techniques.

Data and study area

The study makes use of the monthly arrivals (in metric tons, MT) and wholesale prices (in Rupees per quintal, Rs./Qtl) of onion, sourced from the National Horticultural Board (NHB) for the period January 2008 to December 2020. The total number of collected data points was 156 observations for each market. The last 6 months were reserved for model validation, while the remaining data points were used for model building. Figure 1 shows the map of the study area.



Analytical framework

Different univariate time series techniques were used to capture information from the past and forecast the future. The utilization of these techniques enabled a more efficient understanding of the complex pattern and challenges associated with the forecasting of agricultural commodities.

Autoregressive integrated moving average model

Autoregressive integrated moving average (ARIMA) is a conventional time series model used to estimate and forecast the time series data when they are linear and stationary (Box and Jenkins, 1976). There are three parameters for the ARIMA model: autoregression (p), integration (d), and the moving average (q). The ARIMA model can be expressed as follows:

$$\varnothing(B)(1-B)^d Y_t = \theta(B) \varepsilon_t$$

$$Y_t = \varnothing_1 Y_{t-1} + \varnothing_2 Y_{t-2} + \dots + \varnothing_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

where $\varnothing(B)$ is the autoregressive parameter, $\theta(B)$ is the moving average parameter, ε_t is the white noise or error term, d is the differencing term, and B is the backshift operator. The major steps involved in building the ARIMA model are:

Identification

The values of p and q can be found based on the number of lagged observations of the dependent variable and the number of lagged errors using the partial autocorrelation function (PACF) plot and autocorrelation function (ACF) plot. The integration (d) indicates the number of times the data is differenced to convert them into stationary

(Darekar and Reddy, 2017). Initially, the data must be checked for stationarity using the augmented Dickey–Fuller (ADF) test. If the data are non-stationary, then they must be differenced using appropriate lags to convert them to stationarity.

Parameter estimation

The parameters AR and MA of ARIMA models are estimated using maximum likelihood estimation (MLE). The model which gives the lowest Akaike information criterion (AIC) and Bayesian information criterion (BIC) values are fitted for the data. AIC measures the goodness-of-fit of a statistical model and it is given by

$$\text{AIC} = -2 \log(L) + 2k$$

where L is the likelihood function of the model and k is the number of parameters in the model. The AIC penalizes the models that have a large number of parameters relative to the amount of data. The BIC is similar to the AIC but includes a stronger penalty based on model complexity and it is given by

$$\text{BIC} = -2 \log(L) + k \log(n)$$

Diagnostics and forecasting

After fitting the model, the residuals are checked for the presence of white noise, i.e., the residuals should be uncorrelated. If the model has white noise, it is taken for forecasting, or else the model is refitted until it attains white noise. The presence of white noise is examined through a quantity Q known as the Box-Pierce statistic (a function of autocorrelations of residuals) whose approximate distribution is chi-square and is computed as follows:

$$Q = n \sum_{j=1}^k r_j^2$$

where n is the number of observations in the series and $r(j)$ is the estimated autocorrelation at lag j , extending the summation from 1 to k with k as the maximum positive integer lag. Q follows chi-square with $(k-m-1)$ degrees of freedom where $m-1$ is the number of parameters estimated in the model.

Artificial neural network model

An artificial neural network (ANN) is a computational model inspired by the structure and functionality of biological neural networks. They are a network of interconnected neurons mimicking the function of the human brain. A feedforward neural network (FFNN) is one of the basic neural networks that serve as a non-linear time series model for forecasting purposes (Jha and Sinha, 2013). They are made of input, hidden, and output nodes in which every unit in a particular layer is related to every unit in the previous layer. The data is given through the input node, and the result is obtained from the output node. The in-between hidden layer is the place where processing is done. Each layer consists of weights and biases. The number of input and hidden nodes is determined by experimentation,

as there is no theoretical base for finding these parameters. The mathematical representation of the ANN model is given as:

$$Y_t = \alpha_0 + \sum_{j=1}^q \alpha_j g \left(\beta_{0j} + \sum_{i=1}^p \beta_{ij} Y_{t-p} \right) + \varepsilon_t$$

where α_j ($j = 0, 1, 2, \dots, q$) and β_{ij} ($i = 0, 1, 2, \dots, p, j = 0, 1, 2, \dots, q$) are the model parameters, also called as the connection weights, p is the number of input nodes, q is the number of hidden nodes, and g is the activation function. The ANN model was applied only after confirming the non-linearity in the datasets. Two major steps are involved in the process of fitting a time series model using an ANN, i.e., choosing the right number of nodes for the input layer (p) and hidden layer (q). The choice of p is critical since it has a significant impact on the non-linear autocorrelation structure (Soumen and Debasis, 2018). The selection of q is also subject to faults that could lead to issues with underfitting and overfitting. Unfortunately, the selection of appropriate nodes for input and output layers does not have any prior procedure, so different combinations of networks are fitted until a suitable model is chosen.

Hybrid time series model

Time series data are complex data structures composed of information in linear or non-linear forms. The hybrid time series model is the combination of different time series techniques which adds the strengths of different techniques to improve the accuracy of forecasting (Rathod et al., 2017). Some of the hybrid techniques employed in the study are given as follows:

ARMA-GARCH

ARMA-GARCH is the combination of an autoregressive moving average (ARMA) with a generalized autoregressive conditional heteroskedasticity (GARCH) model which helps to capture the linear information along with volatile information present in the residuals of the mean model. ARMA is the mean model which helps in modeling the linearity of the time series, while GARCH helps in modeling the volatility or the variance (Bollerslev, 1986). ARMA-GARCH models are specified by four parameters: the orders of the autoregressive process (p) and moving average process (q) for the mean component, and the orders of the autoregressive process (r) and the moving average process (s) for the volatility component.

$$\text{Mean equation: } Y_t = \varphi_1 * Y_{t-1} + \dots + \varphi_p * Y_{t-p} + \theta_1 * e_{t-1} + \dots + \theta_q * e_{t-q}$$

$$\text{Variance equation: } \sigma_t^2 = \alpha_0 + \alpha_1 * e_{t-1}^2 + \beta_1 * \sigma_{t-1}^2$$

where φ_1, φ_p and θ_1, θ_q are the coefficients of the AR and MA terms, e_t is the error term at time t , and σ_t^2 is the conditional variance of the error term at a time t with non-negative α and β terms. Checking for heteroskedasticity in the squared residuals (Lama et al., 2016) of the mean model is the first and most important step before proceeding with fitting the of volatility model. ARCH-Lagrange Multiplier (LM)

test is based on the regressing the squared residuals from a given model onto their lagged values. The null hypothesis of the ARCH-LM test is that there is no ARCH effect in the residuals. The formula for the ARCH-LM test statistic is:

$$LM \text{ test} = nR^2$$

where n is the sample size and R^2 is the R-squared value from the auxiliary regression of the squared residuals onto their p -lagged values. The GARCH model is fitted only if heteroskedasticity is found in the residuals. AIC and BIC values are used to choose the optimal model.

ARIMA-ANN model

The hybrid time series model with the combination of ARIMA and ANN combines the strengths of both linear and non-linear techniques to improve the accuracy of time series forecasting (Zhang et al., 2015). The general form of hybrid methodology is given by

$$y_t = L_t + N_t$$

where L_t and N_t are the linear and non-linear components present in the time series y_t , respectively. First, the ARIMA model is fitted to the data to capture the linear dependencies with the prediction series denoted as \hat{L}_t . Second, the residuals ($e_t = y_t - \hat{L}_t$) of the ARIMA model are used as inputs to the ANN model. The ANN is trained to learn the non-linear patterns in the residuals that are not captured by the ARIMA model. Finally, the outputs of the ARIMA and ANN models are combined to produce the final forecast values.

$$\hat{y}_t = \hat{L}_t + \hat{N}_t$$

where \hat{L}_t and \hat{N}_t are the forecasted linear and non-linear components. The ANN model is applied to the residuals of ARIMA only after confirming the presence of non-linearity using the Brock–Dechert–Scheinkman (BDS) test (Naveena and Subedar, 2017). This approach offers a robust and effective means of analyzing time series data, combining the abilities of both linear

and non-linear models. The working flow chart for the hybrid model is given in Figure 2.

Empirical mode decomposition

Empirical mode decomposition (EMD) is a type of self-adaptive time series decomposition approach used to handle non-stationary and non-linear time series data, especially the arrivals and prices of agricultural commodities. The fundamental idea of EMD is to decompose a signal into a finite number of oscillatory components called intrinsic mode functions (IMFs) with a residual term (Huang et al., 1998). It is a data-driven technique that adaptively extracts the oscillatory modes from the signal based on the local extrema. Each IMF has a unique amplitude and frequency modulation for each set of data. There are two requirements that an IMF must meet: (i) the total number of zero crossings and extreme values in the data series must be equal or differ by not more than one, and (ii) the mean value of the envelope determined by the local maxima and minima must be always zero. Generally, the data are present in the form where fast oscillation signals (Das et al., 2020) are likely superimposed over the slow oscillations, i.e.,

$$x(t) = d_1(t) + r_1(t)$$

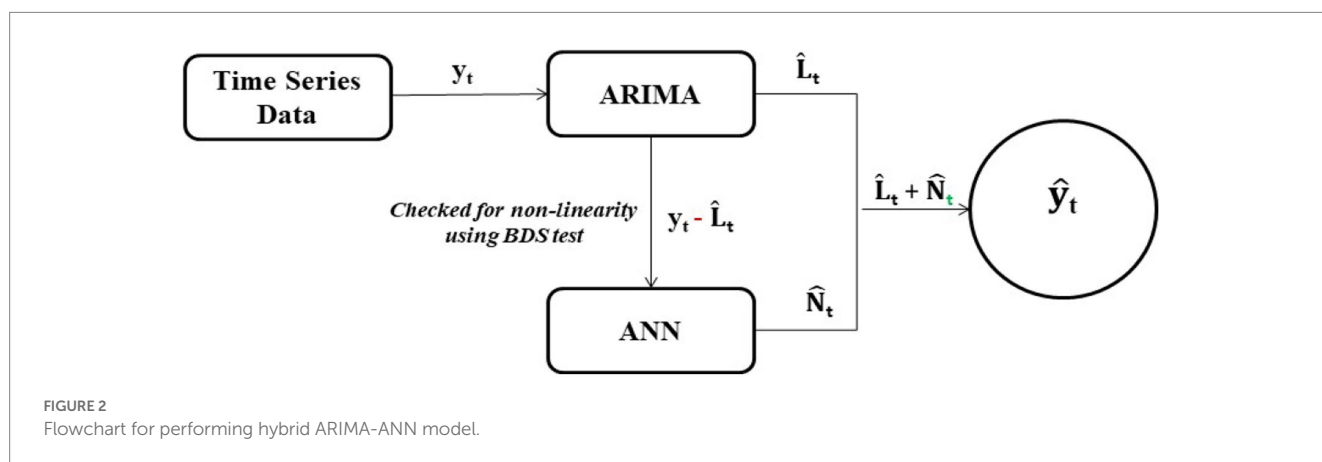
where $d(t)$ = high frequency part, i.e., IMF, and $r(t)$ = low frequency part, i.e., residuals. The EMD algorithm iterates on the slow oscillation component by considering it as a new signal.

$$r_1(t) = d_2(t) + r_2(t)$$

After full decomposition, the data are decomposed into many IMFs and residuals.

$$\text{Data} = \text{sum of IMFs} + \text{final residue.}$$

The stepwise procedure of the EMD algorithm for the price series x_t is mentioned below:



Step 1: Identify all extrema of $x(t)$.

Step 2: Interpolate the local maxima to form an upper envelope $u(x)$.

Step 3: Interpolate the local minima to form a lower envelope $l(x)$.

Step 4: Calculate the mean envelope: $m(t) = \frac{u(x) + l(x)}{2}$

Step 5: Extract the mean from the signal: $h(t) = x(t) - m(t)$

Step 6: Check whether $h(t)$ satisfies the IMF condition.

- YES: $h(t)$ is an IMF, stop shifting.
- NO: let $x(t) = h(t)$, keep shifting.

However, EMD encounters limitations such as mode mixing and boundary effects, which compromise the accuracy of decomposition (Choudhary et al., 2019). Ensemble empirical mode decomposition (EEMD), an enhanced version of EMD, was developed to address the limitations of EMD. EEMD is an ensemble-based variation of EMD that adds noise to the signal to improve decomposition quality and robustness. This helps to alleviate mode mixing issues by averaging the decompositions and enhances the signal-to-noise ratio that results in providing better decomposition results, especially in the presence of noise. The procedure for performing EEMD is:

- Generate and add a quantity of Gaussian white noises $n_i(j)$ into data series $x_i, n_i(j) \sim N(0, \sigma^2)$

$$x_t(j) = x_t + n_t(j)$$

- Apply EMD to newly formed price series $x_t(j)$ and decompose it into a set of IMFs $c_i(ij)$ and residual $r_i(j)$ where $c_i(ij)$ is j^{th} ($i = 1, 2, \dots, n$) IMF decomposed by EMD after adding the $n_i(j)$ for j^{th} ($j = 1, 2, \dots, w$) time.
- Repeat the above two steps for w times to obtain all IMFs and residuals.

The process of EEMD begins with the decomposition of data into different intrinsic mode functions and a residue. After the decomposition, each IMFs and residue are considered as separate datasets, and the desired time series model is applied like the usual procedure. The fitted and forecasted values of each set are bagged together at the end to get the final fitted and forecasted values. The graphical flowchart of EEMD which employs the ARIMA, ANN, and GARCH models is given in Figure 3.

Performance metrics

Mean absolute percentage error, root mean square percentage error (RMSPE), and mean absolute error (MAE) were employed to find the best-performing time series model for the data.

$$\text{MAPE} = \left(\sqrt{\sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right|^2} / n \right) * 100$$

$$\text{RMSPE} = \left(\sqrt{\sum_{t=1}^n \left\{ \frac{Y_t - \hat{Y}_t}{Y_t} \right\}^2} / n \right) * 100$$

$$\text{MAE} = \frac{1}{n} \sum_{t=1}^n |Y_t - \hat{Y}_t|$$

where Y_t is the original observation for the period t , \hat{Y}_t is the corresponding forecasted value from the fitted model, and n is the number of observations.

Results and discussion

The data analysis conducted in this study revealed notable disparities between the two markets in terms of the average arrivals and prices, as depicted in Table 1. The data exhibited a positive skewness, with mean arrivals of 2,468.25 MT and 24,208.81 MT in the Chandigarh and Delhi markets, respectively. Correspondingly, the mean wholesale prices were recorded as 1,419.76 (Rs./Qtl) and 1,514.92 (Rs./Qtl) for the respective markets. The Cuddy Della Valle Index (CDVI) was employed to assess the volatility around the trend of time series data (Dudhat et al., 2017), while the coefficient of variation (CV) was utilized to capture deviations from the mean. The CDVI indicated a substantial level of instability in both arrivals and prices, with the exception of the Delhi market, which exhibited a moderate instability index (Sinha et al., 2018). These findings were further supported by the CV results. The arrivals in Chandigarh and Delhi markets showed a positive compound growth rate (CGR) of 10.95 and 3.82%, respectively. Similarly, the growth rates for prices were observed as 8.34 and 7.79%, respectively (Agarwal et al., 2018). Figure 4 displays the time series graphs for the data, illustrating the non-linear and non-stationary nature of the arrivals and prices in both markets, with several breakpoints observed over the years. The typical behavior of agricultural commodities was further confirmed in onion data by conducting the BDS test, autocorrelation (AC) test, and ADF test (Table 2). The BDS test conducted at different embedding dimensions confirmed the non-linearity in data with statistical significance (Saha et al., 2020). The data were found to be autocorrelated with a value of p of 0.00. The ADF test showed that the data were non-stationary at levels and were converted to stationary after the first differencing. These preliminary tests provided the necessary foundation for the application of subsequent time series techniques to the dataset.

ARIMA model

Table 3 presents the outcomes of the linear models that best fit the data. For the arrivals of Chandigarh and Delhi markets, the best-fitted models were found to be ARIMA (0,1,1) and ARIMA (1,1,1), respectively, based on their low values of AIC and BIC. Likewise, the prices in both markets were modeled using ARIMA (1,1,2) and ARIMA (2,1,1) (Mishra et al., 2013). All the

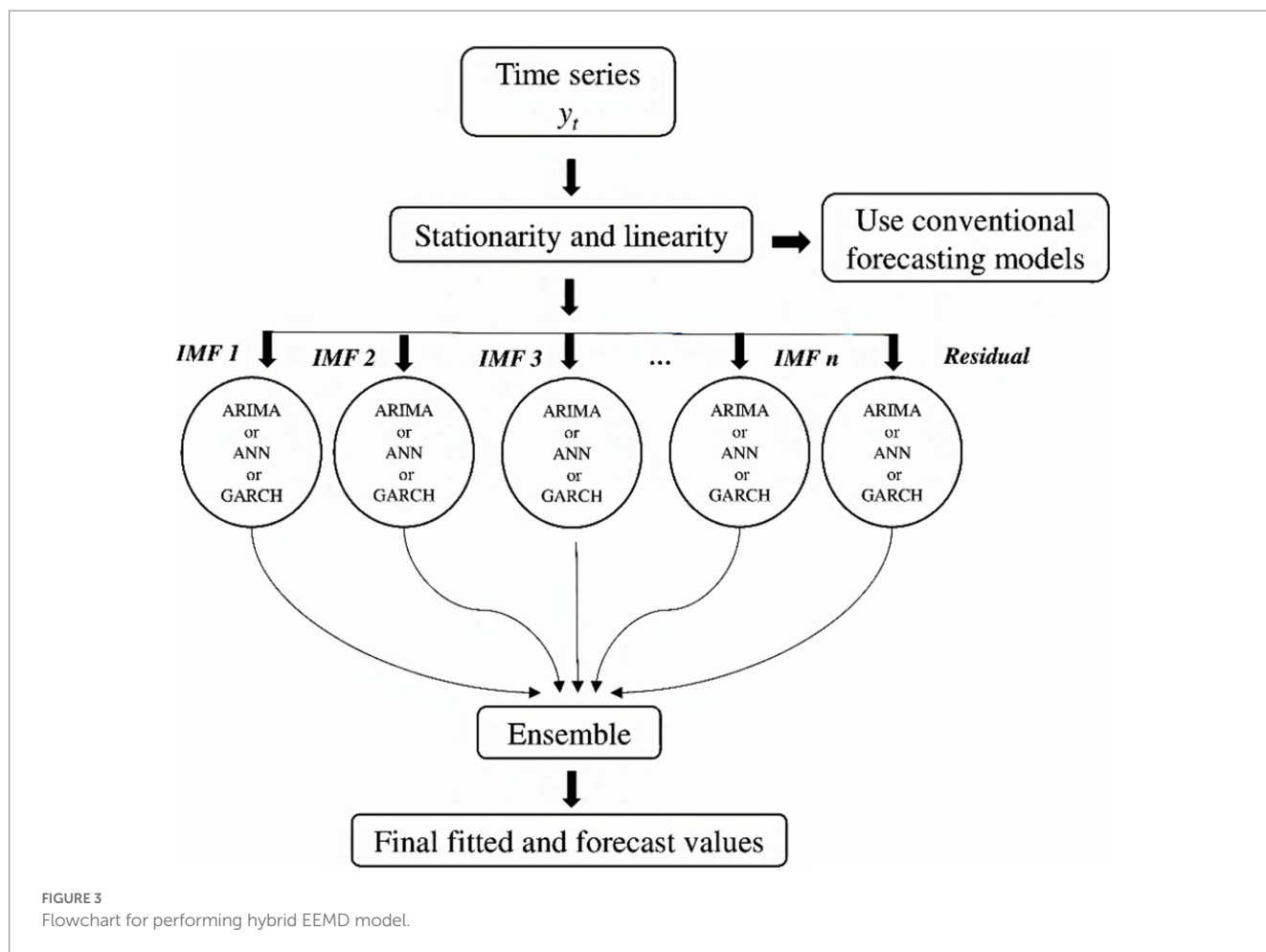


TABLE 1 Descriptive statistics on the arrivals and prices of onion.

Statistics	Onion			
	Arrivals (MT)		Prices (Rs./qtl)	
	Chandigarh	Delhi	Chandigarh	Delhi
Mean	2,468.25	24,208.81	1,419.76	1,514.92
SD	1,390.62	5,854.91	924.27	1,049.25
Skewness	1.55	0.90	2.47	1.88
Kurtosis	2.04	3.73	8.86	4.62
Range	7,441	42,925	6,491	6,568
CV (%)	56.34	24.19	65.10	69.26
CDVI	36.94	19.65	54.08	60.38
	High	Medium	High	High
CGR	10.95	3.82	8.34	7.79

parameters of the ARIMA model were found to be statistically significant with a value of p of 0.01. Results of the Box-Pierce (BP) test indicated that all the fitted models were white noise. However, the Arch-LM test revealed the presence of heteroskedasticity (Jalilkatti and Patil, 2015) only in the residuals of the ARIMA models of the price series. Therefore, to model the price series, a GARCH model was employed while the ARIMA model was used solely for comparative purposes, despite the non-linear nature of the data.

ARMA-GARCH model

The outcomes of the ARMA-GARCH model are presented in Table 4. The models ARMA (1,2)+GARCH (2,1) and ARMA (2,1)+GARCH (2,1) were found to be the best fit based on their AIC and BIC values. Low AIC and BIC values were found of 14.66 and 14.82 for onion prices in Chandigarh and 14.68 and 14.85 for prices in Delhi, respectively (Bawa et al., 2021). Both the series were fitted with the GARCH (2,1) model, i.e., α_1 , α_2 , and β_1 . The lags mean

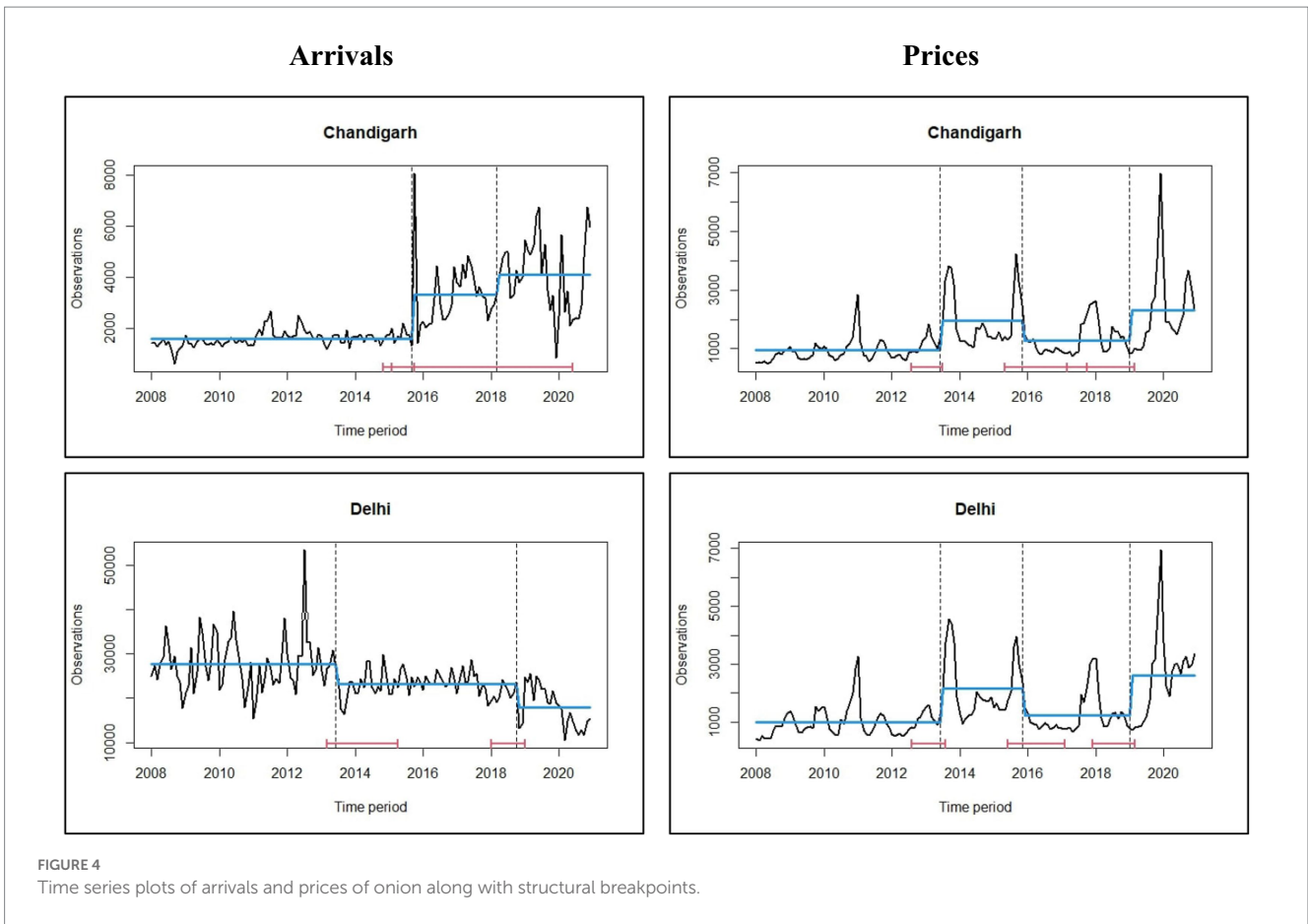


FIGURE 4 Time series plots of arrivals and prices of onion along with structural breakpoints.

TABLE 2 Prerequisites test for time series data.

Statistical test		Chandigarh	Delhi	Chandigarh	Delhi
BDS test (m = 2)	Parameter	695.31	2,927.45	462.13	524.63
	Statistics	18.73	7.67	20.10	21.55
	Probability	<0.001	<0.001	<0.001	<0.001
AC test	Chi-squared	75.54	59.35	102.10	109.48
	Probability	<0.001	<0.001	<0.001	<0.001
ADF test	At levels	-2.05	-3.49	-3.11	-2.77
	Probability	0.55	0.04	0.11	0.26
	First differencing	-4.22	-3.84	-4.28	-3.93
	Probability	<0.001	<0.001	<0.001	<0.001

model were considered from the ARIMA model. The mean value of both the models along with omega and alpha1 of Delhi markets were statistically significant with a value of p of 0.01, whereas the rest of the mean and volatility parameters were non-significant (Ghosh et al., 2020).

ANN model

Considering the confirmation of non-linearity in the arrivals and prices through the BDS test from Table 2, the application of ANN was

more justifiable. The FFNN was employed to build the ANN network for the data with a sigmoid activation function (Nelson et al., 1999). Various combinations of input and output nodes were tested for the data with a minimum of 25 iterations to identify the optimal network configuration. Table 5 shows the results of the ANN model with the final network selected. For the arrival series, 3–2–1 and 3–11–1 networks were fitted while 2–2–1 and 2–2–1 networks were fitted for the price series of Chandigarh and Delhi markets, respectively. Notably, the onion arrivals at the Delhi market required more lags in the hidden node (Jha and Sinha, 2014). The results of the BP test confirmed the adequacy of the fitted ANN model.

TABLE 3 Estimates of parameters of the ARIMA model.

Arrivals				Delhi			
Chandigarh				Delhi			
ARIMA (0,1,1)				ARIMA (1,1,1)			
Parameter	Estimate	Std. error	Value of <i>p</i>	Parameter	Estimate	Std. error	Value of <i>p</i>
MA1	-0.70	0.07	<0.001	AR1	0.43	0.08	<0.001
				MA1	-0.94	0.03	<0.001
AIC	2,444.21			AIC	2,936.39		
BIC	2,450.22			BIC	2,945.4		
BP test	0.25 (-0.61)	LM test	2.27 (-0.99)	BP test	0.05 (-0.82)	LM test	4.50 (-0.97)
Prices				Delhi			
Chandigarh				Delhi			
ARIMA (1,1,2)				ARIMA (2,1,1)			
	Estimate	Std. error	Value of <i>p</i>	AR1	Estimate	Std. error	Value of <i>p</i>
AR1	0.62	0.08	<0.001	AR2	1.13	0.08	<0.001
MA1	-0.49	0.10	<0.001	AR3	-0.38	0.08	<0.001
MA2	-0.46	0.09	<0.001	MA2	-0.97	0.03	<0.001
AIC	2,282.41			AIC	2,295.69		
BIC	2,294.43			BIC	2,307.71		
BP test	0.03 (-0.85)	LM test	57.55 (<0.001)	BP test	0.02 (-0.87)	LM test	36.67 (<0.001)

Probability values are given in brackets.

TABLE 4 Parameter estimation of the ARIMA-GARCH model.

Chandigarh				Delhi			
ARMA (1,2) + GARCH (2,1)				ARMA (2,1) + GARCH (2,1)			
Parameter	Estimate	Std. error	Value of <i>p</i>	Parameter	Estimate	Std. error	Value of <i>p</i>
mu	432.90	150.50	<0.001	mu	325.60	58.16	<0.001
ar1	0.54	0.11	<0.001	ar1	1.00	0.10	<0.001
ma1	0.20	0.12	0.09	ar2	-0.31	0.06	<0.001
ma2	0.16	0.11	0.15	ma1	0.03	0.12	0.82
Omega	52,850.00	31,940.00	0.10	Omega	40,420.00	6,481.00	<0.001
alpha1	1.00	0.67	0.13	alpha1	0.92	0.26	<0.001
alpha2	0.00	0.80	1.00	alpha2	0.24	0.13	0.08
beta1	0.00	0.79	1.00	beta1	0.00	0.09	0.01
AIC	14.66			AIC	14.68		
BIC	14.82			BIC	14.85		

ARIMA-ANN

The initial step involved performing the BDS test on the residuals derived from the applied ARIMA model, aiming to detect the presence of non-linearity. The results of the BDS test indicated that all residuals from the fitted ARIMA model exhibited unexplained non-linear characteristics, as evidenced by a statistically significant value of *p* of 0.01 (Table 6). In order to address this issue, the ANN model was employed to capture the non-linear patterns present in the residuals. The ANN model was configured with appropriate lag lengths for both the input and hidden nodes (Alam et al., 2018). The findings revealed that 6-11-1 and 10-1-1 were the best networks fitted for the arrivals,

while the 7-4-1 and 8-4-1 networks were fitted to the price series of Chandigarh and Delhi markets, respectively (Zhang, 2003). Furthermore, the adequacy of the fitted models was confirmed by the value of *ps* obtained from the BP test. Table 7 represents the weights and parameters utilized in fitting the networks.

Ensemble empirical mode decomposition-based time series models

To ensure the suitability of the data for analysis, the application of EEMD techniques was contingent upon verifying the presence of

TABLE 5 Parameter estimation of the ANN model.

Markets		Network	Weights	parameters	BP test
Arrivals	Chandigarh	3-2-1	11	11	1.08 (0.30)
	Delhi	3-11-1	56	56	0.62 (0.43)
Prices	Chandigarh	2-2-1	9	9	0.54 (0.46)
	Delhi	2-2-1	9	9	0.01 (0.94)

Probability values are given in brackets.

TABLE 6 Results of BDS test on residuals of ARIMA model.

Markets		BDS test		
		Parameter	Statistics	Probability
Arrivals	Chandigarh	434.27	5.94	<0.001
	Delhi	2,233.44	0.19	<0.001
Prices	Chandigarh	246.67	5.37	<0.001
	Delhi	257.95	5.93	<0.001

TABLE 7 Results of ANN model on residuals of ARIMA model.

Markets		Network	Weights	parameters	BP test
Arrivals	Chandigarh	6-11-1	89	89	0.02 (0.90)
	Delhi	10-1-1	13	13	0.67 (0.41)
Prices	Chandigarh	7-4-1	37	37	0.09 (0.76)
	Delhi	8-4-1	41	41	0.45 (0.50)

Probability values are given in brackets.

non-linearity and non-stationarity. Following confirmation through the BDS test and ADF test, the arrivals and price data of onions at the Chandigarh and Delhi markets were decomposed using EEMD. Through the decomposition process, the raw datasets were effectively separated into six intrinsic mode functions (IMFs) and one residual component. This decomposition was visually depicted in Figure 5. It was observed that the decomposed IMFs exhibited a pattern characterized by decreasing fluctuations and increasing amplitude (Guo et al., 2012). Subsequently, time series models such as ARIMA, ANN, and GARCH were individually fitted to each IMF and the residual component using the usual procedures (Wu and Huang, 2009). For every IMF and the residual component, fitted and forecasted values were calculated. Finally, all the fitted and forecasted values obtained from the IMFs and the residual were combined, resulting in a unified output. This process involved aggregating the individual values generated for each IMF and the residual, allowing for a comprehensive and consolidated analysis of the data.

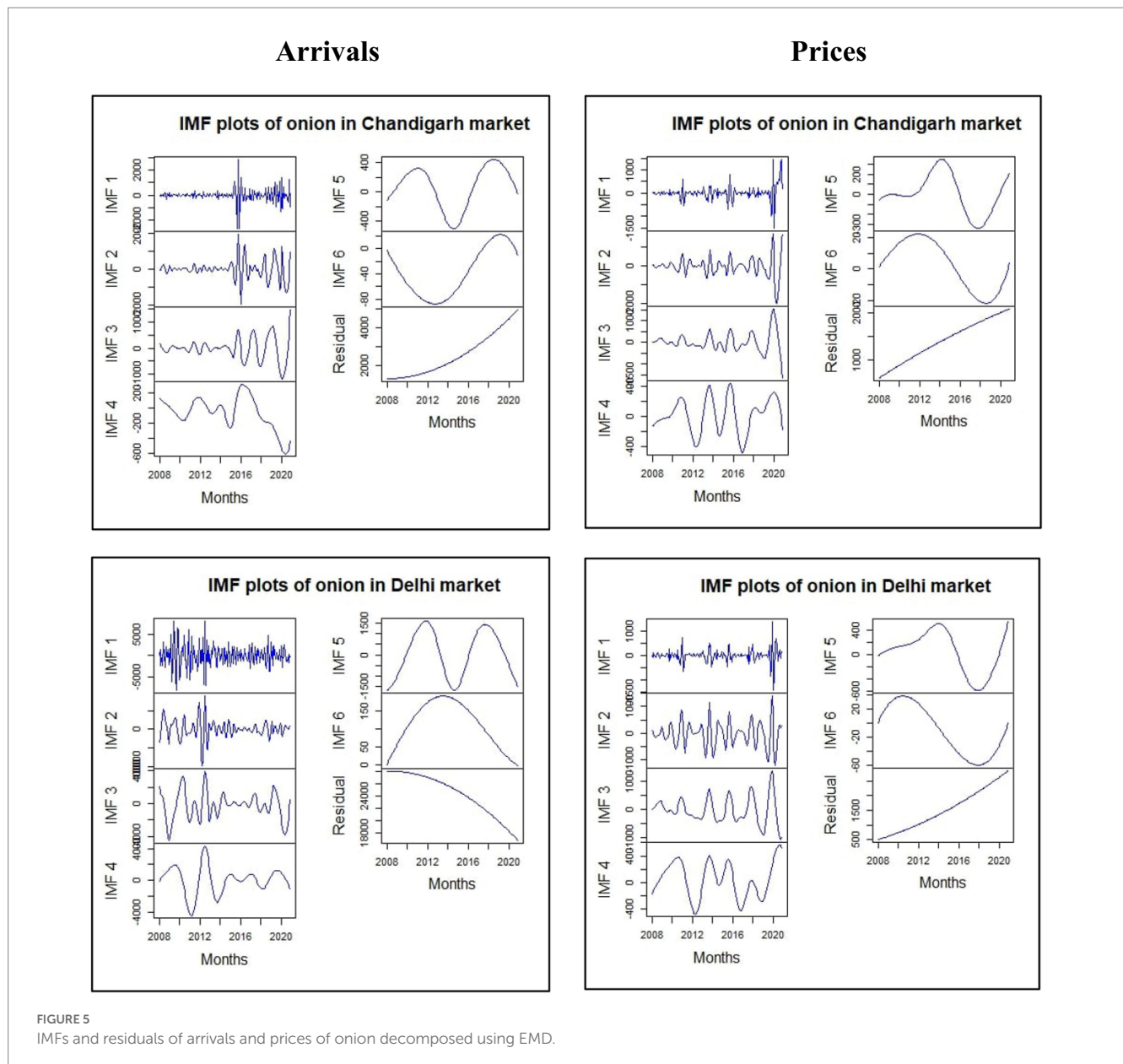
Selection of best-fitting model

In order to evaluate the accuracy of the fitted models, a range of error metrics, including mean absolute percentage error (MAPE), root mean percentage squared error (RMPSE), and mean absolute error (MAE), were employed. Table 8 presents the results of these metrics for the different models. It was observed that the EEMD-ARIMA model exhibited the best fit for the arrival series of

Chandigarh, while the EEMD-ANN model performed well for the remaining data (Choudhury et al., 2019). The MAPE values for the best-fitted models were found to be 17.74 and 6.78% for the arrival series and 9.76 and 10.24% for the price series of Chandigarh and Delhi, respectively. Overall, the findings suggest that the EEMD-ANN and EEMD-ARIMA models (Fang et al., 2020), followed by the ANN model, were the most effective in forecasting both the arrival and price series. Interestingly, the GARCH model, which is typically known for its efficiency in handling volatile data, did not perform well for the dataset under study. Consequently, the arrival and price series were forecasted based on the best-fitted models. Figure 6 illustrates the forecasted values (Wang et al., 2017) of the entire dataset using the respective best-fitted models.

General discussion

Despite being major urban centers in Northern India, cities like Chandigarh and Delhi face challenges in achieving self-sustainability in agricultural production. These cities rely heavily on neighboring regions for the supply of commodities, including onions which can be prominently observed from the previous studies (Mishra et al., 2013; Jalikatti and Patil, 2015; Sinha et al., 2018; Dahiya, 2022; Garai et al., 2023). As a result, they depend on transportation from hub areas, which adds additional costs to the procurement of vegetables. The time series plot showed many breakpoints which indicates the irregular variation in the arrivals and prices pattern of onions. The more the fluctuation in the trends,



the more the volatility which results in an unstable supply and prices (Saha et al., 2020). These variations may be induced by natural calamities like drought, famine, and flood. The implication of policy related to agriculture would also cause such an impact on the pattern of onions. Different kinds of machine learning models are employed to figure out the volatility in onions (Babu and Mallikharjuna, 2022). The present study attempts to employ all these techniques additionally with ensemble techniques. The ensemble techniques combined with the machine learning models were found to perform better for the data. The outperformance of decomposition techniques over the other models is eventual evidence of the presence of complexity in the data which was further supported by the results of Table 2. The forecasting graphs also showed the incidence of volatility especially in prices. Thus, the volatile nature of onion was further confirmed by the results of the study. The government should provide proper guidance and support to onion cultivators regarding the fluctuation phenomena. Proper

forecasting should be done every year so that the farmers can give considerable attention to onion cultivation and marketing. It could also help the government to optimally manage fund allocations for the agricultural sector. As there is no minimum support price (MSP) for vegetables, the government should find possible options to avail MSP for volatile crops, especially onions. This study would help in framing the optimum MSP and other beneficial policies for onion as suggested by previous studies (Babu and Mallikharjuna, 2022). The government should take strict actions against the middlemen who are involved in hedging and hoarding practices that artificially create fluctuations in the markets.

Conclusion

The study evaluates the performance of different forecasting models like ARIMA and ANN, hybrid models like ARMA-GARCH

TABLE 8 Performance metric of fitted models.

Markets		Models	MAPE (%)	RMSPE	MAE
Arrivals	Chandigarh	ARIMA	59.07	37.62	1,662.15
		ANN	51.80	34.91	1,536.27
		ARIMA-ANN	80.55	43.67	1,839.30
		EEMD-ARIMA	17.74	32.00	812.88
		EEMD-ANN	22.17	33.16	901.76
	Delhi	ARIMA	30.65	47.60	5,832.01
		ANN	11.64	15.76	1,694.67
		ARIMA-ANN	27.61	43.44	5,075.23
		EEMD-ARIMA	9.24	13.55	1,293.09
		EEMD-ANN	6.78	10.67	946.57
Prices	Chandigarh	ARIMA	66.34	39.31	1,090.94
		ANN	51.21	33.9	926.98
		ARMA-GARCH	163.04	59.66	1,658.15
		ARIMA-ANN	39.15	29.37	756.93
		EEMD-ARIMA	10.78	10.64	311.42
		EEMD-ANN	9.76	9.83	283.43
		EEMD-GARCH	87.53	67.75	1,098.32
	Delhi	ARIMA	38.32	28.64	812.36
		ANN	14.94	13.87	699.05
		ARMA-GARCH	135.72	55.38	1,604.66
		ARIMA-ANN	46.69	32.51	928.5
		EEMD-ARIMA	11.1	18.63	527.79
		EEMD-ANN	10.24	12.09	434.91
		EEMD-GARCH	48.45	67.14	1,219.28

Bold values indicate the error measures of the best fitted model.

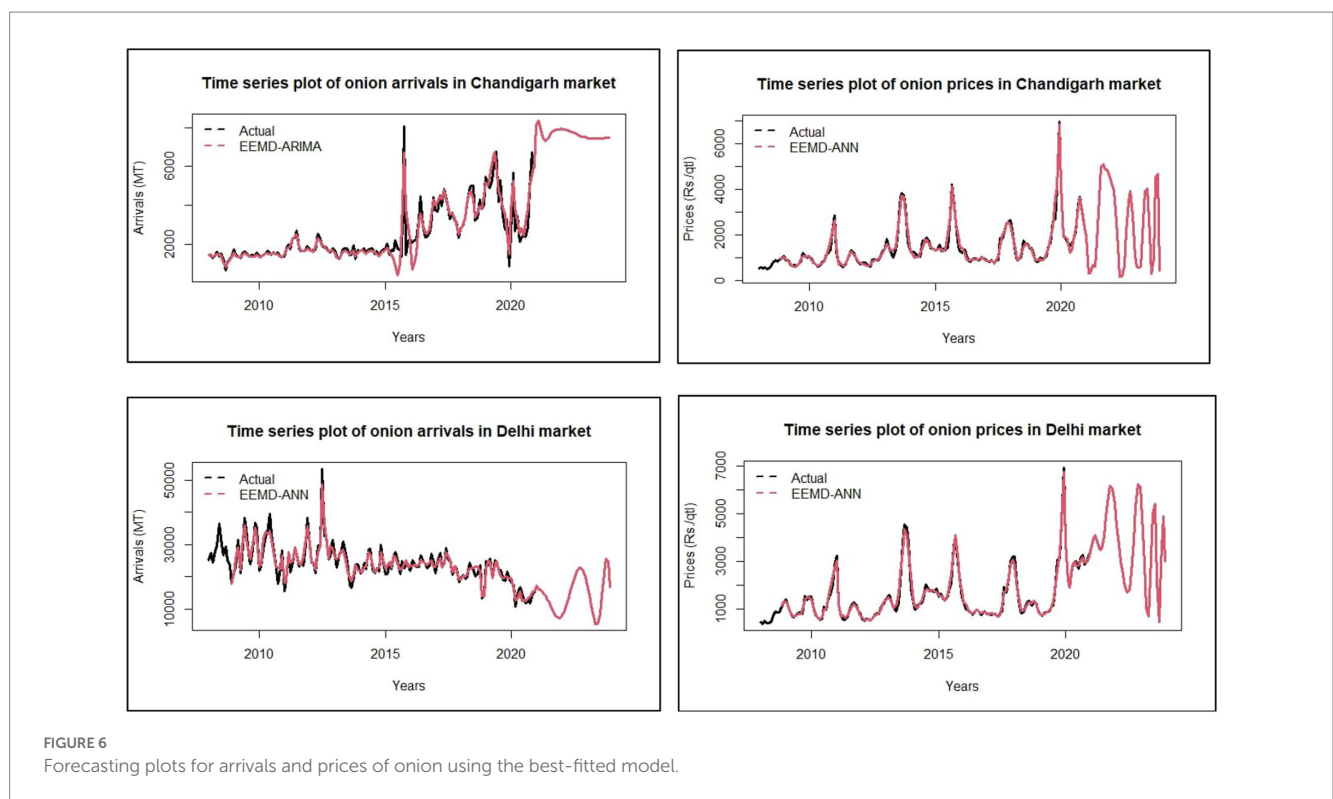


FIGURE 6 Forecasting plots for arrivals and prices of onion using the best-fitted model.

and ARIMA-ANN, and decomposition models like EEMD-ARIMA and EEMD-ANN over multifaceted data. The results of the summary statistics on the arrivals and prices of onion at Chandigarh and Delhi markets showed that the mean arrivals and mean prices were high at Delhi markets, viz., 24,208.81 MT and 1,514.93 Rs./Qtl. The time series plot showed many structural breaks in the data over the years which were due to a high instability index. The arrivals of onions at Delhi markets were observed with a medium instability index. The preliminary test confirmed that data used for the study were autocorrelated, non-linear, and non-stationary at levels with positive growth rates over the years. Therefore, forecasting these unpredictable agricultural arrivals and prices would be helpful for the farmers, stakeholders, and governmental bodies for better planning in the future. From the results of the study, it was found that the ANN model does not consider the stationarity condition but is good at handling non-linearity in data, whereas the ARIMA model can be applied only when the data are stationary and linear. These limitations are overcome by applying EEMD techniques which can analyze the data with both the issues and explain the different levels of fluctuations using the intrinsic mode function. The overall results conclude that EEMD-ANN and EEMD-ARIMA were the best performing models for the study data and had less error rate which were observed using the error metrics like MAPE, RMSPE and MAE. The data were forecasted for the next 3 years using decomposition techniques. All the extracted information from this study could help in better decision-making regarding the production, marketing, and policymaking on the arrivals and prices of onion.

Thus, the attempts to highlight the problems of instability in the arrivals and prices of onion at Chandigarh and Delhi markets through different statistical measures were shown in this study. Additionally, the performance of different time series models was examined, and the best-performing model was chosen to forecast the future pattern of arrivals and price of onions in study markets. In the future, the accuracy of forecasting can be further improved by accounting for weather

parameters, as they are one of the prime factors influencing the patterns of agricultural commodities.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: <https://www.nhb.gov.in/OnlineClient/MonthlyPriceAndArrivalReport.aspx>.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer AK declared a shared affiliation with the author AA to the handling editor at the time of review.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Agarwal, P., Singh, R., and Singh, O. P. (2018). Dynamics of prices and arrivals of major vegetables: a case of Haldwani and Dehradun markets, Uttarakhand. *J. Agric. Dev. Policy* 28:1.
- Alam, W., Sinha, K. A., Kumar, R. R., Ray, M. R., Rathod, S. A., Singh, K. N., et al. (2018). Hybrid linear time series approach for long term forecasting of crop yield. *Indian J. Agric. Sci.* 88, 1275–1279. doi: 10.56093/ijas.v88i8.82573
- Allen, P. G. (1994). Economic forecasting in agriculture. *Int. J. Forecast.* 10, 81–135. doi: 10.1016/0169-2070(94)90052-3
- Anjoy, P., Paul, R. K., Sinha, K., Paul, A. K., and Ray, M. (2017). A hybrid wavelet based neural networks model for predicting monthly WPI of pulses in India. *Indian J. Agric. Sci.* 87, 834–839. doi: 10.56093/ijas.v87i6.71022
- Areef, M., Rajeswari, S., Vani, N., and Naidu, G. M. (2020). Price behaviour and forecasting of onion prices in Kurnool market, Andhra Pradesh State. *Econ. Aff.* 65, 43–50. doi: 10.30954/0424-2513.1.2020.6
- Arjun, K. M. (2013). Indian agriculture-status, importance and role in Indian economy. *Int. J. Agric. Food Sci. Technol.* 4, 343–346.
- Babu, K. S., and Mallikharjuna, R. K. (2022). *Onion Price prediction using machine learning approaches*. In Proceedings of International Conference on Computational Intelligence and Data Engineering: ICCIDE 2021. Singapore: Springer Nature Singapore, pp. 175–189
- Bawa, M. U., Dikko, H. G., Shabri, A., Garba, J., and Sadiku, S. (2021). Forecasting performance of hybrid ARIMA-FIGARCH model and hybrid of ARIMA-GARCH model: a comparative study. *J. Math. Prob. Equ. Stat.* 2, 48–58.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econ.* 31, 307–327. doi: 10.1016/0304-4076(86)90063-1
- Box, G. E., and Jenkins, G. M. (1976). *Time series analysis, control, and forecasting*. San Francisco, CA: Holden Day, p. 10.
- Choudhary, K. A., Jha, G. K., Kumar, R. R., and Mishra, D. C. (2019). Agricultural commodity price analysis using ensemble empirical mode decomposition: a case study of daily potato price series. *Indian J. Agric. Sci.* 89, 882–886. doi: 10.56093/ijas.v89i5.89682
- Choudhury, K., Jha, G. K., Das, P., and Chaturvedi, K. K. (2019). Forecasting potato price using ensemble artificial neural networks. *Indian J. Ext. Educ.* 55, 73–77.
- Dahiya, P. (2022). Modeling and forecasting of all India monthly average wholesale Price volatility of onion: an application of GARCH and EGARCH techniques. *Curr. J. Appl. Sci. Technol.* 41, 91–100.
- Darekar, A., and Reddy, A. (2017). Predicting market price of soybean in major India studies through ARIMA model. *J. Food Legum.* 30, 73–76. doi: 10.2139/ssrn.3089035
- Das, P., Jha, G. K., and Lama, A. (2023). *Empirical mode decomposition based ensemble hybrid machine learning models for agricultural commodity Price forecasting*, No. 21. pp. 99–112.
- Das, P., Jha, G. K., Lama, A., Parsad, R., and Mishra, D. (2020). Empirical mode decomposition based support vector regression for agricultural price forecasting. *Indian J. Ext. Educ.* 56, 7–12.
- Dudhat, A. S., Yadav, P., and Venujayakanth, B. (2017). A statistical analysis on instability and seasonal component in the price series of major domestic groundnut markets in India. *Int. J. Curr. Microbiol. Appl. Sci.* 6, 815–823. doi: 10.20546/ijemas.2017.611.096
- Fang, Y., Guan, B., Wu, S., and Heravi, S. (2020). Optimal forecast combination based on ensemble empirical mode decomposition for agricultural commodity futures prices. *J. Forecast.* 39, 877–886. doi: 10.1002/for.2665

- Garai, S., Paul, R. K., Rakshit, D., Yeasin, M., Emam, W., Tashkandy, Y., et al. (2023). Wavelets in combination with stochastic and machine learning models to predict agricultural prices. *Mathematics* 11:2896. doi: 10.3390/math11132896
- Ghani, I. M., and Rahim, H. A. (2019). *Modeling and forecasting of volatility using ARMA-GARCH: Case study on Malaysia natural rubber prices*. In IOP Conference Series: Materials Science and Engineering, Vol. 548. p. 012023.
- Ghosh, S., Singh, K. N., Thangasamy, A., Datta, D., and Lama, A. (2020). Forecasting of onion (*Allium cepa*) price and volatility movements using ARIMAX-GARCH and DCC models. *Indian J. Agric. Sci.* 90, 1009–1013. doi: 10.56093/ijas.v90i5.104384
- Guo, Z., Zhao, W., Lu, H., and Wang, J. (2012). Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model. *Renew. Energy* 37, 241–249. doi: 10.1016/j.renene.2011.06.023
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., et al. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. London A Math. Phys. Eng. Sci.* 454, 903–995. doi: 10.1098/rspa.1998.0193
- Jalikatti, V. N., and Patil, B. L. (2015). Onion price forecasting in Hubli market of northern Karnataka using ARIMA technique. *Karnataka J. Agric. Sci.* 28, 228–231.
- Jha, G. K., and Sinha, K. (2013). Agricultural price forecasting using neural network model: an innovative information delivery system. *Agric. Econ. Res. Rev.* 26, 229–239.
- Jha, G. K., and Sinha, K. (2014). Time-delay neural networks for time series prediction: an application to the monthly wholesale price of oilseeds in India. *Neural Comput. Appl.* 24, 563–571. doi: 10.1007/s00521-012-1264-z
- Kumar, P., Badal, P. S., Paul, R. K., Jha, G. K., Venkatesh, P., Anbukani, P., et al. (2021). Forecasting onion price for Varanasi market of Uttar Pradesh, India. *Agric. Res.* 91, 249–253. doi: 10.56093/ijas.v91i2.111603
- Kumar, M., Shaikh, A. S., and Sharma, R. K. (2022). Market integration and Price transmission analysis of onion in wholesale Markets of India. *Asian J. Agric. Exten. Econ. Sociol.* 40, 164–171. doi: 10.9734/ajaees/2022/v40i121778
- Kumar, M., and Thenmozhi, M. (2014). Forecasting stock index returns using ARIMA-SVM, ARIMA-ANN, and ARIMA-random forest hybrid models. *Int. J. Bank. Account. Financ.* 5, 284–308. doi: 10.1504/IJBAAF.2014.064307
- Lama, A., Jha, G. K., Gurung, B., Paul, R. K., Bharadwaj, A., and Parsad, R. (2016). A comparative study on time-delay neural network and GARCH models for forecasting agricultural commodity price volatility. *J. Indian Soc. Agric. Stat.* 70, 7–18.
- Mishra, P., Sarkar, C., Vishwajith, K. P., Dhekale, B. S., and Sahu, P. K. (2013). Instability and forecasting using ARIMA model in area, production and productivity of onion in India. *J. Crop Weed* 9, 96–101.
- Naveena, K., and Subedar, S. (2017). Hybrid time series modelling for forecasting the price of washed coffee (Arabica plantation coffee) in India. *Int. J. Agric. Sci.* 9, 975–3710.
- Nelson, M., Hill, T., Remus, W., and O'Connor, M. (1999). Time series forecasting using neural networks: should the data be deseasonalized first? *J. Forecast.* 18, 359–367. doi: 10.1002/(SICI)1099-131X(199909)18:5<359::AID-FOR746>3.0.CO;2-P
- Pardhi, R., Singh, R., and Paul, R. K. (2018). Price forecasting of mango in Lucknow market of Uttar Pradesh. *Int. J. Agric. Environ. Biotechnol.* 11, 357–363.
- Paul, R. K., Bhardwaj, S. P., Singh, D. R., Kumar, A., Arya, P., and Singh, K. N. (2015). Price volatility in food Commodities in India-an Empirical Investigation. *Int. J. Agric. Stat. Sci.* 11, 395–401.
- Paul, R. K., Das, T., and Yeasin, M. (2023). Ensemble of time series and machine learning model for forecasting volatility in agricultural prices. *Natl. Acad. Sci. Lett.* 46, 185–188. doi: 10.1007/s40009-023-01218-x
- Rakshit, D., Paul, R. K., and Panwar, S. (2021). Asymmetric Price volatility of onion in India. *Indian J. Agric. Econ.* 76, 245–260.
- Rathod, S., Mishra, G. C., and Singh, K. N. (2017). Hybrid time series models for forecasting banana production in Karnataka state, India. *J. Indian Soc. Agric. Stat.* 71, 193–200.
- Saha, N., Kar, A., Jha, G. K., Kumar, P., Venkatesh, P., and Kumar, R. R. (2020). Forecasting of onion Price in Lasalgaon market and potato Price in Agra market. *J. Commun. Mob. Sustain. Dev.* 15, 460–466. doi: 10.5958/2231-6736.2020.00029
- Saha, A., Singh, K. N., Ray, M., and Rathod, S. (2020). A hybrid spatio-temporal modelling: an application to space-time rainfall forecasting. *Theor. Appl. Climatol.* 142, 1271–1282. doi: 10.1007/s00704-020-03374-2
- Saxena, R., Singh, N. P., Paul, R. K., and Kumar, R. (2019). Market linkages for the major onion markets in India. *Indian J. Hortic.* 76, 133–140. doi: 10.5958/0974-0112.2019.00019.7
- Singh, R. K. (2008). Artificial neural network methodology for modelling and forecasting maize crop yield. *Agric. Econ. Res. Rev.* 21, 5–10.
- Sinha, K. (2023). “The Price of onions” in *The future of India's rural markets: A transformational opportunity*. ed. B. S. Sergi (Bingley, England: Emerald Publishing Limited), 49–51.
- Sinha, K., Panwar, S., Alam, W., Singh, K. N., Gurung, B., Paul, R. K., et al. (2018). Price volatility spillover of Indian onion markets: a comparative study. *Indian J. Agric. Sci.* 88, 114–120.
- Soumen, P., and Debasis, M. (2018). Forecasting monthly rainfall using artificial neural network. *RASHI* 3, 65–73.
- Sujay, R., Parvati, V. K., Biradar, S. R., Kerur, V. V., and Huilgol, S. S. (2022). Forecasting of onion prices. In *Contemporary issues in communication, cloud and big data analytics: Proceedings of CCB 2020 2022*. Springer: Singapore, pp. 339–343
- Wang, H., Liu, L., Qian, Z., Wei, H., and Dong, S. (2014). Empirical mode decomposition-autoregressive integrated moving average: hybrid short-term traffic speed prediction model. *Transp. Res. Rec.* 2460, 66–76. doi: 10.3141/2460-08
- Wang, D., Yue, C., Wei, S., and Lv, J. (2017). Performance analysis of four decomposition-ensemble models for one-day-ahead agricultural commodity futures price forecasting. *Algorithms* 10:108. doi: 10.3390/a10030108
- Wu, Z., and Huang, N. E. (2009). Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Adv. Adapt. Data Anal.* 01, 1–41. doi: 10.1142/S1793536909000047
- Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* 50, 159–175. doi: 10.1016/S0925-2312(01)00702-0
- Zhang, J. L., Zhang, Y. J., and Zhang, L. (2015). A novel hybrid method for crude oil price forecasting. *Energy Econ.* 49, 649–659. doi: 10.1016/j.eneco.2015.02.018