Statistics and Applications {ISSN 2454-7395(Online)}
Volume 21, No. 1, 2023 (New Series), pp 99-112

Empirical Mode Decomposition Based Ensemble Hybrid Machine Learning Models for Agricultural Commodity Price Forecasting

Pankaj Das[1], Girish Kumar Jha[2] and Achal Lama[1]
[1]*ICAR- Indian Agricultural Statistics Research Institute, New Delhi, India*
[2]*ICAR- Indian Agricultural Research Institute, New Delhi, India*

Received: 12 April 2022; Revised: 20 May 2022; Accepted: 09 June 2022

Abstract

Agricultural commodity price is very volatile in nature due to its nonlinearity and nonstationary character. The volatility behaviour of the commodity price creates a lot of problems for both producer and consumer. The steady forecast of the price may reduce the problems and increase the profit for the stakeholders. In this study, an ensemble hybrid machine learning model based on empirical mode decomposition (EMD) has been proposed to forecast the commodity price. EMD decomposes the nonstationary and nonlinear price series into different stationary intrinsic mode functions (IMF) and a final residue. Then Machine learning techniques like Artificial neural network (ANN) and Support vector regression (SVR) were used to forecast each of the decomposed components. Finally, all the forecasted values of the decomposed components were aggregated to produce the final forecast. Two R modules were developed for the application of the proposed methodology. The proposed methodology has been applied to the monthly wholesale price index of vegetables. The results indicated that the ensemble hybrid machine learning model based on empirical mode decomposition has superior performance compared to generic models.

Key words: Agricultural commodity price; Machine learning; Empirical mode decomposition; Nonlinearity; Nonstationary; Artificial neural network; Support vector regression.

---

1. Introduction

The scientific and effective forecasting method is helpful to correctly guide producers and policy makers to match the supply and demand of the agricultural production and facilitate the decision-making process of the government. Agricultural price forecasting is not an easy task due to its dependency on many extraneous factors. Nonlinearity and nonstationary behaviour of data series are crucial problems in the agricultural price forecasting. Agricultural commodity prices are volatile in nature due to seasonality, inelastic demand, production uncertainty *etc*. Traditionally, time series forecasting has been dominated by linear methods like ARIMA (Box and Jenkins, 1970) and nonlinear models such as SETAR, STAR, *etc*. because they are well understood and effective in many situations. These traditional methods suffer from some limitations, such as linear models focusing on linear relationships, fixed temporal dependence *etc*. and nonlinear models

Corresponding Author: Pankaj Das
Email: pankaj.iasri@gmail.com

require the specific nonlinear relation of data generating process to be known a priori. Zhang *et al*. (1998) demonstrated that traditional methods requiring the time series data to be stable or stable after being differentiated. On the other hand, Machine learning (ML) models with their flexible functional designs and powerful self-learning capabilities have recently become a great alternative for time series data forecasting. Machine learning techniques like Artificial Neural Network (ANN) and Support Vector Machine (SVM) become popular to handle the nonlinearity problem in the dataset (Qin and Chiang, 2019). Darbellay and Slama (2000) highlighted that ANN which is non-linear, nonparametric and data driven self-adaptive method, is most suitable for forecasting agricultural price series which is inherently noisy and nonlinear in nature. Levis and Papageorgiou (2005) clarified the numerous advantages of SVR for nonlinear times series forecasting. Lu *et al*. (2009) demonstrated the generalization characteristics of SVR for finding a unique solution. An *et al.* (2012) reported that empirical mode decomposition (EMD) can reveal the hidden pattern and trends of time series which can effectively assist in designing forecasting models for various applications. Sugiyama and Kawanabe (2012) stated the inability of ML techniques to counter the nonstationarity behaviour of a dataset. Huang *et al.* (1998) pointed out how EMD is capable to deal with the problem of inherent nonstationarity or nonlinearity in a time series dataset. EMD has the power to isolate the high fluctuating data into respective smaller frequency components (Mumtaz *et al*., 2019). EMD also has the capability to reduce the influence of nonlinear characteristics of the stock series (Xuan *et al*., 2020)

It has been observed in the literature that a single model is not sufficient to deal with complex real-world systems such as agricultural price data which contains unknown mixed patterns. Besides, inherent non-stationarity and nonlinearity behaviour of price series create problems in robust forecasting (Taylor and Kingsman, 1978). ML algorithms are compatible and efficient to deal with nonlinear problems (Bishop, 1995). To handle non-stationarity features, the inputs for machine learning models need to be properly pre-processed (Wang *et al*., 2017). Therefore, some multi-resolution analysis techniques are widely used in many forecasting problems. In view of these factors, there is an ever-increasing need of using hybrid models to improve the accuracy of predictions. In order to improve forecast accuracy, hybridization is a good idea because it can capture various patterns in the data concurrently. Hybrid models, combining the benefits of different models, are suggested to achieve better prediction and the decomposition approaches such as EMD enhance the performance of hybrid models (Mo *et al*., 2020). These studies led to the development of novel hybrid models which are more robust as they often compliment the advantages of the individual technique involved and improve the forecasting accuracy. In this study, novel hybrid models have been proposed by combining EMD and ML algorithms like ANN and SVR to deal with nonlinearity and non-stationarity problems in a time series data. EMD counters with non-stationarity of a time series data by decomposing into several stationary components which is nonlinear in nature. Further ANN and SVR models have been used to forecast these nonlinear decomposed components. We have used monthly wholesale price index of vegetables for practical evaluation and compared the proposed model with other models, including hybrid and individual models.

The remaining portion of the paper is organized as follows. Section 2 deals with the methodology. The data and results of the experiment are explained in the third section. The final section concludes the paper.

2. Materials and methods

2.1. Empirical mode decomposition (EMD)

The empirical mode decomposition method was introduced by Huang *et al*. in 1998. It assumes that the data have many coexisting oscillatory modes of significantly distinct frequencies and these modes superimpose on each other and form an observable time series. EMD decomposes original nonstationary and nonlinear data into a finite and small number of independent sub-series (including intrinsic mode functions and a final residue). Intrinsic Mode Function (IMF) is the finite additive oscillatory component decomposed by EMD. For example, let y_t be a time series (TS) dataset at time t consisting of high frequency part and low frequency part. After first decomposition, original time series data results into first IMF and the residue. The decomposed TS takes the following form:

$$y_t = d_t(1) + r_t(1) \tag{1}$$

where $d_t(1)$ = high frequency part *i.e.* IMF and $r_t(1)$ = low frequency part *i.e.* residue. EMD algorithm iterates over the slow oscillation component considered as a new signal. In the next iteration, the residue $r_t(1)$ will be treated as new signal for EMD decomposition.

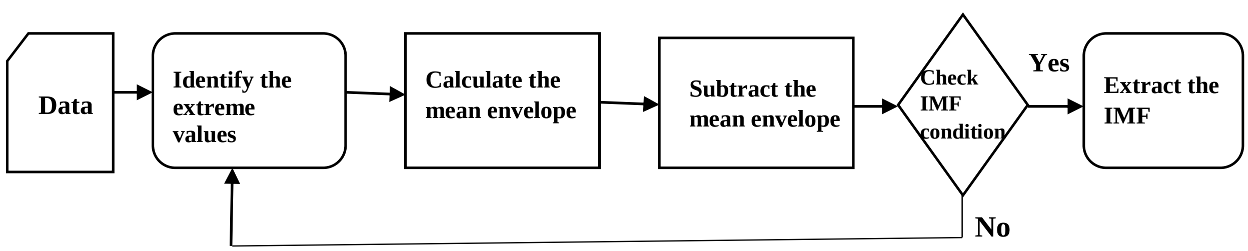
After second decomposition,

$$r_t(1) = d_t(2) + r_t(2)$$

Let us assume that y_t has been decomposed into k numbers of IMFs and final residues after the EMD decomposition process completes. Then, y_t can be expressed as follows:

$$y_t = \sum_{i=1}^{k} d_t(i) + R \tag{2}$$

Hence, original time series data = sum of IMFs + final residue (R). The stepwise EMD algorithm procedure is mentioned below:



All these steps come under the first iteration in the shifting process for y_t. The shifting process continues till we obtain an IMF. The point of termination of a sifting process is called stopping point k and the iteration is called k^{th} iteration. The stopping criteria of an iteration is the standard deviation (σ) between two IMF (d) values. The predefined threshold value ranges between 0.2 to 0.3 (Huang *et al*., 1998). When the σ value lies between 0.2 to 0.3, the iteration should stop. The threshold value is calculated using the given expression:

$$\text{Threshold value} = \sum_{l=0}^{T}\left[\frac{(d_t(l(k-1)-d_t(lk))^2}{d_t^2(l(k-1)}\right] \quad (3)$$

2.2. Artificial neural network (ANN) model

Artificial neural network is a non-linear, data driven self-adaptive approach as opposed to the traditional model based methods (Jha and Sinha, 2014). ANN can identify and learn correlated patterns between input data sets and corresponding target values. ANN imitates the learning process of the human brain and can process problems involving non-linear and complex data even if the data are imprecise and noisy. Thus, it is ideally suited for modelling of agricultural data which are known to be complex and often non-linear. Haykin (1999) stated mathematically that a neuron k can be defined by the following equations:

$$u_k = \sum_{j=1}^{m} w_{kj} x_j \quad (4)$$

$$y_k = \varphi(u_k + b_k) \quad (5)$$

Here bias (b_k), has the effect of increasing or lowering the net input of the activation function. $x_1, x_2, ..., x_m$ are the inputs; $w_{k1}, w_{k2}, \ldots, w_{km}$ are the weights of the neuron k; u_k is the linear combiner output due to input variables; $\varphi(.)$ is the activation function; y_k is the output of the neuron, w_{kj} the weight attached to the connection from j^{th} hidden node to the output node. The backpropagation algorithm can be implemented under the following components:

1. Data should contain input-output pair $(\vec{x}_i, \vec{y}_i)$, where $\vec{x}_i = \{y_{t-d}, y_{t-d+1}, ..., y_{t-1}\}$ the input, d is a user-defined parameter, which corresponds to the number of previous time-steps and $\vec{y}_i = y_t$ is the desired output. For T data of $X = \{(\vec{x}_1, \vec{y}_1), ..., (\vec{x}_T, \vec{y}_T)\}$.
2. Need a feedforward neural network. Let the parameters of the network be denoted by θ. The parameters of interest in backpropagation are the weights w_{ij}^k, node j in layer l_k and node i in layer l_{k-1} and bias b_i^k the bias for node i in layer l_k.

Error function $E(X,\theta) = \frac{1}{2T}\sum_{t=1}^{T}(\hat{y}_t - y_t)^2$; where $\hat{y}_t$ are the computed output of the network on input $\vec{x}_t$ and y_t is the target value for input-output pair $(\vec{x}_i, \vec{y}_i)$.

In the present study, Logistic function was used as an activation function and resilient backpropagation algorithm was used to adjust the weights in the multi-layered feed-forward network.

2.3. Support vector regression (SVR) model

Vapnik (1998) introduced support vector regression model by incorporating a loss function. SVR fits linear regression in the outer space through mapping input vectors into a high dimensional space. In the present study, a modified SVR model named least squares support vector regression (LS-SVR) proposed by Suykens *et al.* (2002) has been used. LS-SVR model focuses on set of linear equations instead of a quadratic programming problem in SVR model.

LS-SVR model is represented as:

$$y = w^T \varphi(x) + b \tag{6}$$

with $x \in R^T$ and φ, mapping function $R^n \to R^{n_t}$ to high dimensional feature space, bias b and error e. For a given training set $\{x_t, y_t\}_{t=1}^{T}$, optimization problem becomes as follows:

$$\min_{\{w,e,b\}} J(w,e) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{t=1}^{T} e_t^2 \tag{7}$$

where $e_t = y_t - w^t \varphi(x_t) + b$ is the fitting errors, subject to equality constraints

$$y_t = w^T \varphi(x_t) + b + e_t\,;\ t = 1, 2, \ldots, T. \tag{8}$$

This is a form of ridge regression. Now incorporating Lagrange multiplier α_t

$$L(w,b,e;\alpha) = J(w,e) - \sum_{t=1}^{T} \alpha_t \{w^T \varphi(x_t) + b + e_t - y_t\} \tag{9}$$

with following conditions of optimality

$$\begin{cases} \dfrac{\partial L}{\partial w} = 0 \to w = \displaystyle\sum_{t=1}^{T} \alpha_t \varphi(x_t) \\ \dfrac{\partial L}{\partial b} = 0 \to \displaystyle\sum_{t=1}^{T} \alpha_t = 0 \\ \dfrac{\partial L}{\partial e_t} = 0 \to \alpha_t = \gamma e_t, \qquad t = 1, 2, \ldots, T \\ \dfrac{\partial L}{\partial \alpha_t} = 0 \to w^T \varphi(x_t) + b + e_t - y_t = 0, \quad t = 1, 2, \ldots, T \end{cases} \tag{10}$$

$$\text{Solution will be } \begin{bmatrix} 0 & \vec{1}^T \\ 1 & \Omega + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{11}$$

with $y = [y_1, \ldots, y_T], \vec{1} = [1, \ldots, 1],\ \alpha = [\alpha_1, \ldots, \alpha_T]$. By applying Mercer's condition

$$\Omega_{tu} = \Omega(\alpha_t, \alpha_u) = \sum_{t=1}^{T} d_t(\alpha_t - \alpha_t^{'}) - e\sum_{t=1}^{T}(\alpha_t - \alpha_t^{'}) - \frac{1}{2}\sum_{t=1}^{T}\sum_{u=1}^{T}(\alpha_t - \alpha_t^{'})(\alpha_u - \alpha_u^{'})K(x_t, x_u) \tag{12}$$

where d = scalar vector of x, $K(x_t, x_u)$ is the kernel function.

The final LS-SVR model can be written as:

$$y(x) = \sum_{t=1}^{T} \alpha_t K(x_t, x) + b \tag{13}$$

In the present study, least squares SVR model with Radial basis function (RBF) kernel was used for nonlinear mapping of dataset.

2.4. Proposed ensemble hybrid model

Ensemble method is a machine learning approach which combines multiple base models to produce an optimal predictive model. The proposed EMD-SVR/ANN consists of three steps depicted in Figure 1. In the first step, original nonlinear and nonstationary dataset is decomposed into a finite and often small numbers of independent sub-series by EMD technique. This sub-series contain k intrinsic mode functions (IMFs) and a final residue. Secondly, these IMFs and residue is modelled and predicted through ANN or SVR. Then, all the forecasted values of the IMFs and residue are summed up to produce ensemble forecast for the original series. The prediction model of ANN (Equation 5) and SVR (Equation 13) were used with decomposed components as inputs. The input lags were selected based on Akaike information criterion (AIC) and Bayesian information criterion (BIC).

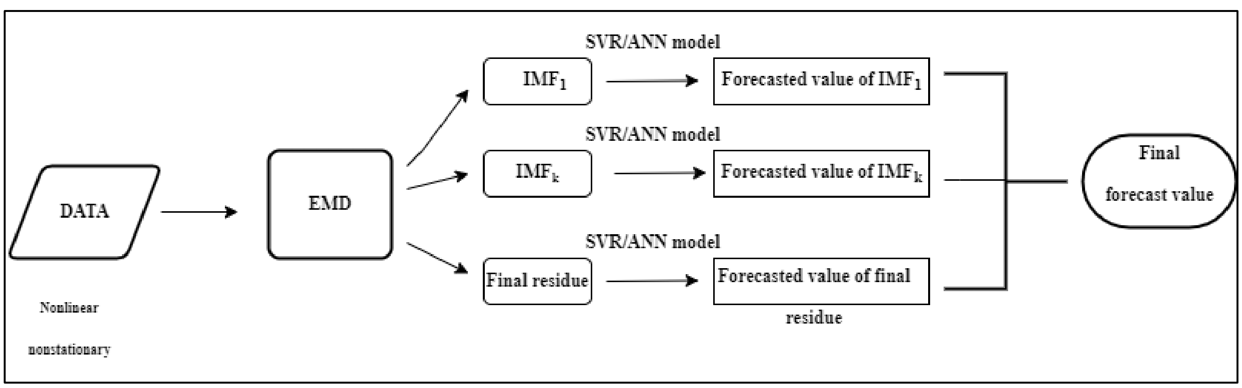


Figure 1: EMD based ensemble hybrid machine learning model for a dataset

2.5. Assessment of the fitted models

The fitted models were assessed using the performance measures like root mean squared error (RMSE), mean absolute deviation (MAD), mean absolute percentage error (MAPE) and maximum error (ME). These performance measures can be expressed as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{T}(y_t - \hat{y}_t)^2}{T}} \qquad \text{MAD} = \frac{\sum_{t=1}^{T}\left|y_t - \hat{y}_t\right|}{T}$$

$$\text{MAPE} = \frac{1}{T}\sum_{t=1}^{T}\frac{\left|y_t - \hat{y}_t\right|}{y_t} \qquad \text{ME} = \max\sum_{t=1}^{T}\left|y_t - \hat{y}_t\right|$$

where y_t and $\hat{y}_t$ are the actual value and predicted value of response variable and T is the number of data points. For comparing the forecasting performance, the DM (Diebold and Mariano, 1995) test was used.

3. Results and discussion

The complete analysis of the present study was done using RStudio. For analysis, two R packages *i.e. EMDANNhybrid* (Das *et al*., 2021) and *EMDSVRhybrid* (Das *et al*., 2021) have been developed and are available in CRAN repository. Considering the time dependency of observations, the data was divided into training and testing set for all models.

3.1. Data source

In the present study, monthly price index of vegetables was used to evaluate the performance of the proposed EMD-ANN and EMD-SVR models. The dataset was obtained from the Office of the Economic Advisor, Ministry of Commerce, Government of India (*https://eaindustry.nic.in/*). Figure 2 illustrates the monthly data of wholesale price index (WPI) of vegetables (January 2005 to November 2020) containing 191 data points. The descriptive statistics, stationarity test and normality of sample data were presented in Table 1. The statistics obtained through Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) test were insignificant *i.e*. null hypothesis of the unit root test cannot be refused. It indicated that the given dataset was nonstationary. Jarque-Bera test (Table 1) indicated the nonnormality of data.

Table 1: The descriptive statistics, stationarity and normality tests of data

Descriptive statistics		Stationarity test		Normality test
Numbers of observations	191	Augmented Dickey-Fuller test (*p* value)	Phillips-Perron test (*p* value)	Jarque-Bera test (*p* value)
Maximum	284.90	0.23	0.30	<0.01
Minimum	43.70			
Mean	121.97			
Standard deviation	52.17			
Skewness	0.64			
Kurtosis	-0.12			

Brock-Dechert-Scheinkman (Brock *et al*., 1996) test was used in the dataset for checking the nonlinearity of data. The results of the BDS test (Table 2) indicated that the test statistics were far bigger than the critical values. It provided a piece of evidence to reject the null hypothesis that the price series is linearly dependent. The results obtained from various tests revealed that the monthly vegetables WPI dataset is nonlinear and nonstationary in nature. These characteristics of the dataset enabled us to implement and evaluate the performance of the proposed EMD-ANN/SVR models with existing individual models.

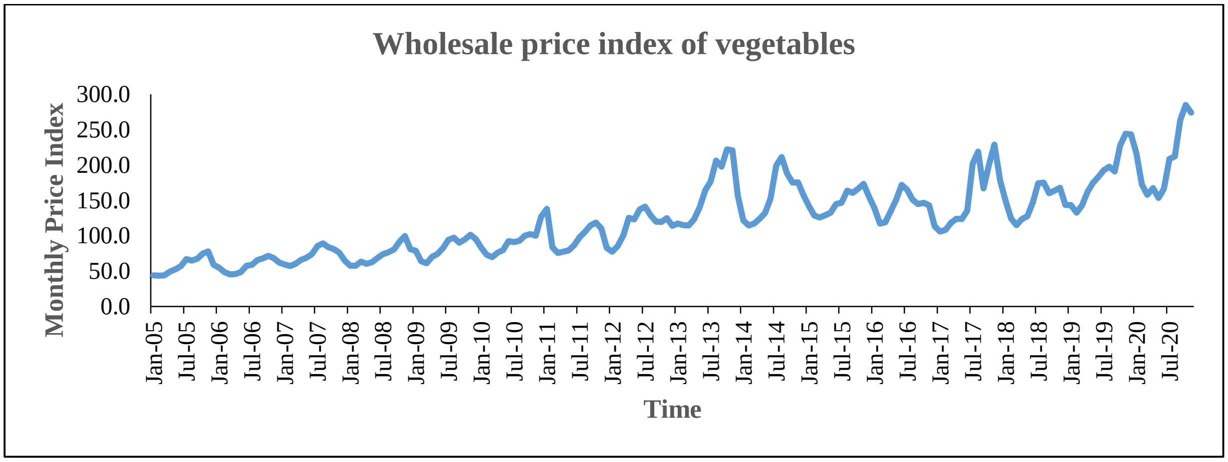


Figure 2: Time plot of vegetables monthly price index (2004-05=100)

Table 2: Results of Brock- Dechert-Scheinkman (BDS) test

Embedding dimension				Conclusion
2		3		
Statistics	Probability	Statistics	Probability	
102.75	< 0.001	172.21	< 0.001	Nonlinear
54.88	< 0.001	68.12	< 0.001	
40.38	< 0.001	42.96	< 0.001	
37.68	< 0.001	37.17	< 0.001	

3.2. EMD decomposition

The EMD, as an adaptive decomposition technique is quite effective in extracting characteristic information from nonlinear and nonstationary time series. EMD methodology has been employed to decompose the series. Firstly, EMD algorithm finds the local extreme values (*i.e*. maxima and minima) from the dataset. Local extrema are the points where slope sign is changed. The envelopes are the curve that passing through the local extrema. The curves passing through the local minima and local maxima are known as upper envelope and lower envelope respectively. Mean envelope are the curves that is passing through mean values of local maxima and minima. The whole process of EMD decomposition has been visualized in Figure A.1. The figure gives an idea of how the EMD algorithm finds the envelopes and residue from a dataset. In the given figure red line and blue line indicates the upper and lower envelope respectively. The mean envelope and the residue are denoted using black and green lines in Figure A.1. The original series has been decomposed into four IMFs and one final residue using EMD (Figure A.2). It has been observed that the frequencies and amplitudes of IMFs were different and independent from each other. Thus, the different hidden oscillatory modes in the original datasets were separated by EMD. Each decomposed IMF contains certain characteristics of the dataset which needs to be modelled and forecasted using appropriate model.

After decomposition, it is pertinent to check the stationarity of IMFs and residue. The results (Table 3) of the test indicated that all IMFs were stationary, but the final residue was nonstationary because it contains the remaining portion of the data that cannot be decomposed by the EMD algorithm. As stationarity is one of the important assumptions for forecasting, hence, the nonstationary residue as such cannot be used in forecasting. The residue was transformed into stationary by first differencing. The stationary decomposed parts *i.e.* IMFs and the differenced residue were used for forecasting.

Table 3: Unit root test of decomposed components of vegetables WPI dataset

Components	Augmented Dickey-Fuller test (*p*-value)	Augmented Dickey-Fuller test (p-value)	Remarks
IMF_1	<0.01	<0.01	Stationary
IMF_2	<0.01	<0.01	Stationary
IMF_3	<0.01	<0.01	Stationary
IMF_4	<0.01	<0.01	Stationary
Residue	0.14	0.14	Nonstationary

3.3. ANN training and forecasting

ANN model was employed to different IMFs and final residue for forecasting purpose as it is capable of handling nonlinear and complex data. For this purpose, the *EMDANNhybrid* R-Package has been developed and used for analysis. Backpropagation training algorithm was used for ANN fitting. In practice, ANN with a small number of parameters namely input lags and hidden nodes often perform better for out of sample forecasting. This may be because over-fitting is a common problem in case of neural network models with a large number of parameters. In this study, we varied input lags and hidden nodes from one to five. ANN model with three input lags and four hidden nodes was found to be the most suitable model for the given dataset in terms of accuracy criterion. The other parameters like the maximum number of iterations for the neural network was fixed at 200. We averaged the results of 26 neural networks for getting the final output. The number of neural networks to be averaged was selected based on the minimum error criterion. We tried averaging 10 to 50 neural networks and obtained the best result on 26 number of neural networks. In our study, 80% of data as training set and the remaining 20% as testing set were used.

3.4. SVR training and forecasting

Similarly, the SVR model was also fitted to different IMFs and final residue. The each component (IMFs and residue) was modelled and forecasted using the *EMDSVRhybrid* R-package, developed under this study. The SVR model was preferred over other machine learning algorithms due to its capability to handle nonlinear systems as well as its suitability for a small sample size.

For employing the SVR model, we divided the dataset into training and testing sets. The training set is used for model building purposes whereas, the testing set allows us to understand the generalization ability of the developed model. In this study, we have used 80% of the data as a training set and the remaining 20% as a testing set. The developed SVR model for each decomposed component (IMFs and residue) was used to forecast the respective components. Then all the forecasted values of IMFs and residue were summed up to get an ensemble forecast of the data. Radial Basis Function (RBF), polynomial, linear and sigmoid kernel functions were implemented in SVR model fitting. The best result was obtained using the RBF kernel function. To overcome the problem of overfitting, 10-fold cross-validation was also done.

3.5. Performance comparison of fitted models

The performance (in-sample and out-sample) of the EMD-ANN and EMD-SVR model was compared with the individual ANN and SVR model (Table 4 and 5) with forecasting horizon of six months. Both the in-sample and out-sample performance of EMD-SVR was relatively superior as compared to other competing models. EMD based ANN and SVR models outperformed the individual models like ANN and SVR. The reason behind the poor performance of singular ANN and SVR models can be attributed mainly to the fact that these models could not handle the nonstationary behaviour of the given dataset. On the other hand, the hybrid models EMD-ANN and EMD-SVR performed better due to the ability to capture both nonlinearity and nonstationarity patterns of the dataset.

Table 4: In-sample performance of fitted models

	ANN	SVR	EMD-ANN	EMD-SVR
RMSE	10.68	28.25	5.13	3.15
MAPE	5.22	0.21	0.04	0.03
MAD	6.67	20.84	4.36	2.71
ME	0.87	98.19	19.63	15.76

Table 5: Out-sample performance of fitted models

	ANN	SVR	EMD-ANN	EMD-SVR
RMSE	44.39	54.39	25.36	23.69
MAPE	0.26	0.24	0.15	0.10
MAD	41.33	44.80	23.12	17.12
ME	69.81	115.83	67.23	68.74

Further DM test was employed to assess the accuracy of the EMD-SVR model compared to the EMD-ANN model. The null hypothesis of the DM test was that both models have the same accuracy. Results of Table 6 clearly indicated that the EMD-SVR model was superior to the EMD-ANN model in terms of all the criteria. The test also indicated that the forecasting performance of both EMD-ANN and EMD-SVR gave better results compared to the individual ANN and SVR models. The novelty of the proposed ensemble approach is that it can handle nonlinear and nonstationary data which is difficult for the existing time series methods. Our empirical findings suggest that the proposed EMD-SVR and EMD-ANN models can be considered as an alternative tool for volatile agricultural price series forecasting.

Table 6: Results of Diebold-Mariano (DM) test

Hypothesis	DM value	*p* value	Remarks
H_0: The accuracy of both EMD-SVR and EMD-ANN is same. H_1: The accuracy of EMD-SVR is superior to EMD-ANN.	5.48	<0.01	The accuracy of EMD-SVR is superior to EMD-ANN.
H_0: The accuracy of both EMD-SVR and SVR is same. H_1: The accuracy of EMD-SVR is superior to SVR.	3.02	<0.01	The accuracy of EMD-SVR is superior to SVR.
H_0: The accuracy of both EMD-ANN and ANN is same. H_1: The accuracy of EMD-ANN is superior ANN.	6.23	<0.01	The accuracy of EMD-ANN is superior ANN.

4. Conclusion

In this study, EMD based hybrid machine learning models for forecasting have been proposed to deal with inherent nonlinearity and non-stationarity behaviour in a time series dataset. Single models fail to capture both aforementioned characteristics in a dataset. To deal with these problems, one decomposition technique namely EMD and two machine learning algorithms *i.e*. ANN and SVR were combined to formulate two hybrid models. The hybrid models, EMD-ANN and EMD-SVR are capable to deal with the nonlinearity and non-stationarity in a dataset. The performance of hybrid models was also evaluated with a real dataset. The empirical results clearly demonstrated the superior forecast accuracy of the proposed hybrid models (EMD-ANN and EMD-SVR) as compared to the individual ANN and SVR model.

Acknowledgements

The authors are thankful to the editor and the anonymous reviewer for their valuable suggestions which helped in the improvement of the paper. The authors are also thankful to the Director, ICAR-Indian Agricultural Statistics Research Institute, New Delhi for providing necessary facilities for carrying out the present research work.

References

An, X., Jiang, D., Zhao, M. and Liu, C. (2012). Short time prediction of wind power using EMD and chaotic theory. *Communication in Nonlinear Science and Numerical Simulation*, **17**, 1036-1042.

Bishop, M. C. (1995). *Neural Networks for Pattern Recognition*. Oxford University Press, New York.

Brock, W. A., Scheinkman. J. A., Dechert, W. D. and LeBaron, B. (1996). A test for independence based on the correlation dimension. *Econometric Reviews*, **15**, 197-235.

Darbellay, G. A. and Slama, M. (2000). Forecasting the short-term demand for electricity: Do neural networks stand a better chance? *International Journal of Forecasting*, **16**, 71-83.

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 Journal of Extension Education*, **56**, 7-12.

Das, P., Lama, A. and Jha, G. K. (2021). R Package *EMDSVRhybrid*. (https://CRAN.R-project.org/package=EMDSVRhybrid).

Das, P., Lama, A and Jha, G. K. (2021). Package *EMDANNhybrid*. https://CRAN.R-project.org/package=EMDANNhybrid.

DES (Directorate of Economics and Statistics) *Cost of Cultivation of Principal Crops in India* (*various issues*), Government of India, New Delhi.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, **84**, 427–431.

Diebold, F. X. and Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, **13**, 253-265.

Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*. Prentice Hall, Ontario.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. L., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C. and Liu, H. H. (1998). The empirical mode decomposition and Hilbert spectrum for nonlinear and nonstationary time series analysis. *Proceeding of the Royal Society London A*, **454**, 909-995.

Jarque, C. M. and Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, **55**, 163–172.

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 Computing and Applications*, **3-4**, 563-571.

Levis, A. and Papageorgiou, L. (2005). Customer demand forecasting via support vector regression analysis. *Chemical Engineering Research and Design*, **83**, 1009-1018.

Lu, C. J., Lee, T. S. and Chiu, C. C. (2009). Financial time series forecasting using independent component analysis and support vector machine. *Decision Support Systems*, **47**, 115-125.

Mo, Y., Li, Q., Karimian, H., Fang, S., Tang, B. and Chen, G. (2020). A novel framework for daily forecasting of ozone mass concentrations based on cycle reservoir with regular jumps neural networks. *Atmospheric Environment*, **220**, 117072. https://doi.org/10.1016/j.atmosenv.2019.117072.

Mumtaz, A., Prasad, R., Xiang, R., Zaher, Y. and Yaseen, M. (2020). Complete ensemble empirical mode decomposition hybridized with random forest and kernel ridge regression model for monthly rainfall forecasts. *Journal of Hydrology*, **584**, 124647. https://doi.org/10.1016/j.jhydrol.2020.124647.

Qin, S. and Chiang, L. (2019). Advances and opportunities in machine learning for process data analytics. *Computers and Chemical Engineering*, **126**, 465-473.

Sugiyama, M. and Kawanabe, M. (2012). *Machine Learning in Non-Stationary Environments-Introduction to Covariate Shift Adaptation*. 2nd Edition, MIT Press, Cambridge, Massachusetts, London, England.

Taylor, S. J. and Kingsman, B. G. (1978). Non-stationarity in sugar prices. *Journal of the Operational Research Society*, **29**, 971–980.

Vapnik, N. V. (1998). *Statistical Learning Theory*. 1st Edition, Wiley-Interscience, New York.

Wang, D., Wei, S., Luo, H., Yue, C. and Grunder, O. (2017). A novel hybrid model for air quality index forecasting based on two phase decomposition technique and modified extreme learning machine. *Science of 6e Total Environment*, **580**, 719–733.

Xuan, Y., Yu, Y. and Wu, K. (2020). Prediction of short-term stock prices based on EMD-LSTM-CSI neural network method. *5th IEEE International Conference on Big Data Analytics (ICBDA)*, (pp. 135-139). Xiamen, China.

Zhang, G. P., Zhang, B. E. and Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, **14**, 35-62.

Appendix A

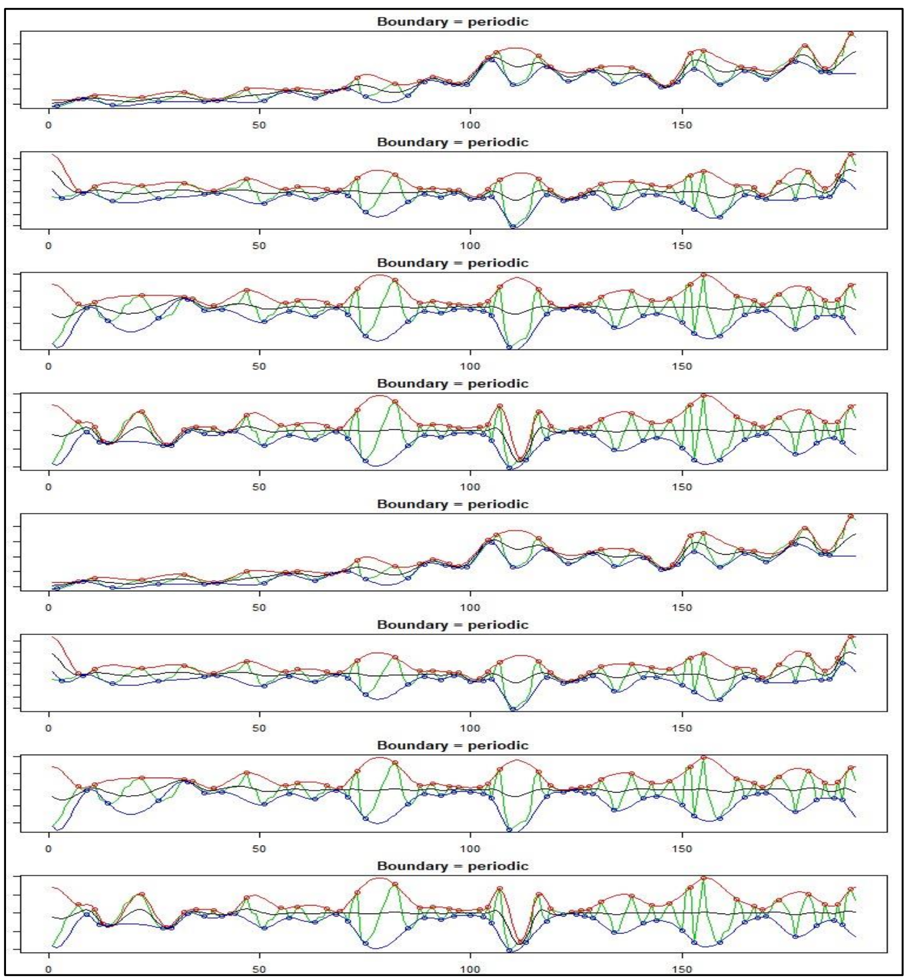


Figure A.1: EMD process in the monthly vegetables WPI dataset

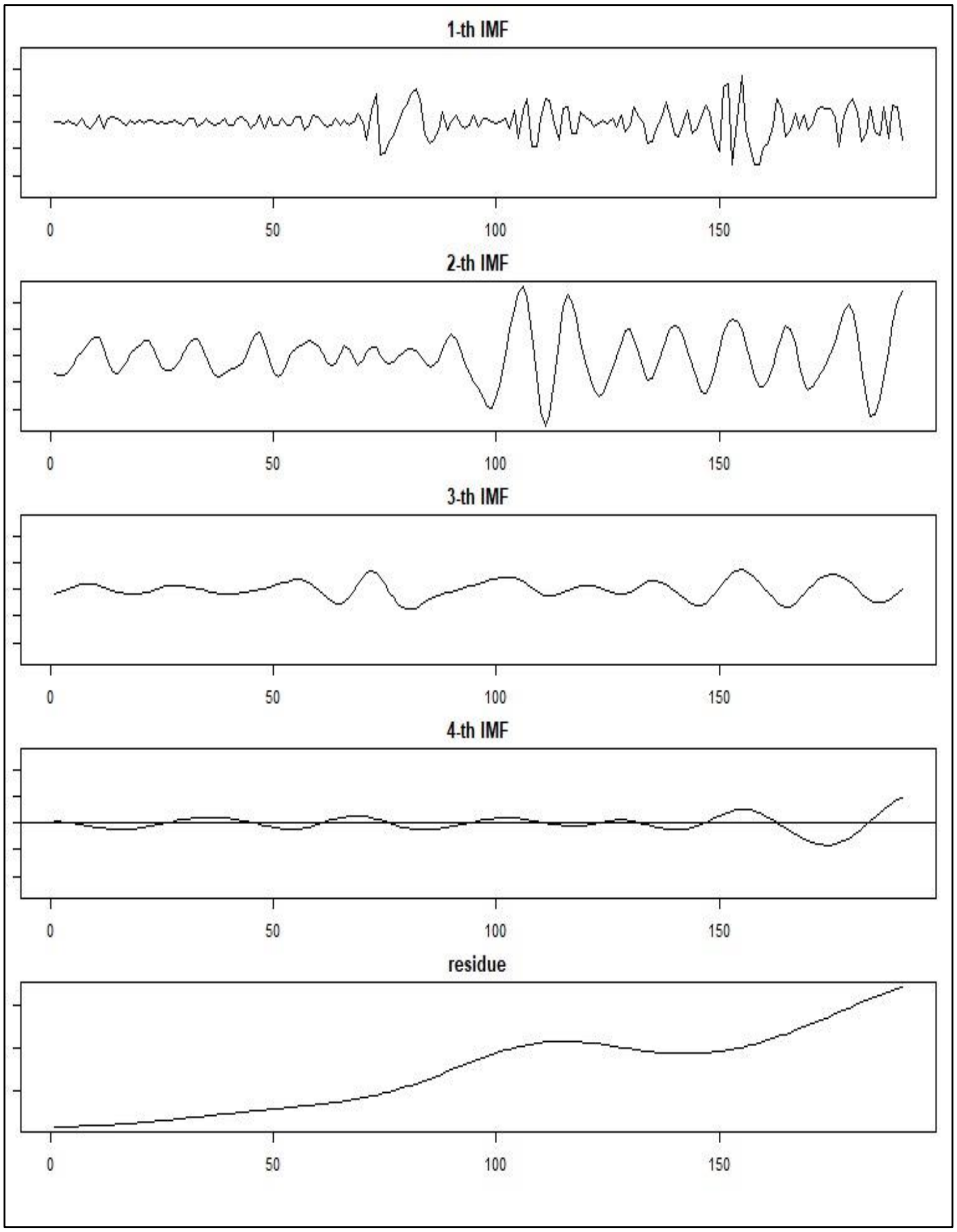


Figure A.2: Decomposed components of monthly vegetables WPI dataset