

Sequential wavelet-ANN with embedded ANN-PSO hybrid electricity price forecasting model for Indian energy exchange

Smitha Elsa Peter¹ · I. Jacob Raglend²

Received: 17 March 2015 / Accepted: 21 December 2015 / Published online: 21 January 2016
© The Natural Computing Applications Forum 2016

Abstract This research paper presents a novel sequential wavelet-artificial neural network (ANN) with embedded ANN-particle swarm optimization (PSO) for short-term day-ahead forecasting of market clearing price (MCP) in the Indian energy exchange. A precise price forecasting helps suppliers to set up bidding strategies, make investment decisions and be cautious against risks. Conversely, consumers can use price forecasting to exploit appropriate power purchasing strategies for maximum utility utilization. Here the most influential historical data, namely purchase bid and MCP, are considered for training the feed-forward back-propagation neural network. The proposed model involves three sequential phases. Initially, the raw historical data are smoothened by removing the high-frequency components using a wavelet transform method which may enable better training of neural network. Then, ANN is used to train historical patterns. More number of trials is carried out, and the final weights that give the least training error are stored. In the final phase, the stored weights that are obtained from various trials are used as the initial population for the embedded ANN-PSO model. Here the performance of the proposed forecasting model is carried out using three error indices, namely mean absolute percentage error, normalized mean square error and error variance.

Keywords Artificial neural networks · Particle swarm optimization · Wavelet transform · Mean absolute percentage error

1 Introduction

Electricity pricing and forecasting is one of the important functions in an electricity market since electricity has become an essential commodity in the modern society. The price signal is non-homogeneous, and its variations show a little cyclic property. The electricity price signal is volatile in nature due to volatility in fuel price, load uncertainty, transmission congestion, behaviour of market participant, market manipulations, etc. Because of the significant volatility, it is difficult to make an accurate forecast. Based on the literature, the accuracy of the price forecast varies around 10 % as compared to 3 % for load forecasting [24]. However, price forecasting accuracy is not as stringent as that of load forecasting. There are various electricity markets, around the world, using different types of time series models [11, 16, 18]. However, each market uses a suitable forecast model for its own method of functioning. As a result, it is necessary to develop an accurate forecasting model relevant to a particular electricity market. Here, we had considered an Indian electricity market, namely Indian Energy Exchange (IEX). In this market, market clearing price (MCP) is forecasted, which depends on the supply and demand of electricity in the power market. It should also be noted that the accuracy in forecasting MCP depends on the intrinsic and extrinsic factors that depend on each market. Therefore, MCP is a varying price signal that mostly depends on the dynamic behaviour of buyer and seller, analogous to the demand and supply in the market, respectively. When electricity MCP is

✉ Smitha Elsa Peter
smitha_peter@yahoo.com

I. Jacob Raglend
jacobraglend@rediffmail.com

¹ Department of ECE, PRIST University, Vallam, Thanjavur, Tamil Nadu 613 403, India

² Department of EEE, Noorul Islam University, Thuckalay, Kannyakumari, Tamil Nadu 629 180, India

determined, every supplier whose offering price is below or equal to the electricity MCP will be picked up by the Independent System Operator (ISO) to supply electricity at that hour. They will be paid at the same price, the electricity MCP and not at the price they offered. The reason for this is to keep fairness of the market and to avoid market manipulation [27].

Aggarwal et al. [2] investigated the state of price forecasting methodologies by reviewing 47 papers published during 1997 to November 2006. The review is based on the type of model used for forecasting, time horizon for prediction, input variables, output variables, analysis of results, data points used for analysis, pre-processing employed and the model architecture. They have observed that forecasting errors are still high from risk management perspective and the results obtained by different price forecasting methodologies are difficult to compare with each other. Catalao et al. [7] had proposed a novel approach combining wavelet transform, particle swarm optimization and adaptive-network-based fuzzy inference system to forecast short-term electricity prices forecasting of Spain market. Dev et al. [10] have investigated extensively on feed-forward back-propagation neural network and extreme value distributions to model electricity pool prices for the Australian National Electricity Market from 1998–2013. However, their work has shown that price spikes can be well-modelled using extreme value distributions, and this understanding based on their characteristics is an important component of understanding the extraordinary volatility of electricity spot prices. Gao et al. [12] used a three-layered feed-forward back-propagation artificial neural network method to forecast the Market Clearing Pricing and Market Clearing Quantity for the California day-ahead energy markets. The author discussed on one of critical issues in training neural networks which is over fitting. That is, it fitted the training set very well, but did not generalize well to new data outside the training set.

Mandal et al. [17] had presented a hybrid intelligent algorithm utilizing a data filtering technique. It is based on wavelet transform (WT), an optimization technique based on firefly (FF) algorithm and a soft computing model based on fuzzy ARTMAP (FA) network. Here, it forecasts day-ahead electricity prices in the Ontario market and was evaluated against PJM market data. Aggarwal et al. [1] and Pindoriya et al. [21] had used WT with neural network (WNN) for short-term price forecasting for the day-ahead prediction of the market clearing price (MCP) in the Ontario, Spain and PJM electricity markets. The performance of WNN was found to be encouraging, since the data had been pre-processed by WT. The benefit of WT for pre-processing is also found in [26] when it is used with Extreme Learning Machine for electricity price forecasting. Pousinho et al. [22] had proposed a hybrid PSO–

ANFIS (adaptive-network-based fuzzy inference system) approach for short-term price forecasting. It is a combination of particle swarm optimization and adaptive-network-based fuzzy inference system on the electricity market of mainland Spain. The author demonstrated its effectiveness regarding forecasting accuracy and reduced computation time. Amjady [3] developed a fuzzy neural network model for day-ahead electricity price forecasting (EPF) for Spanish electricity market and claimed the model to be satisfactory.

Sharma and Srinivasan [25] had proposed a hybrid intelligent model based on the recurrent neural network (RNN) and excitable dynamics for price prediction in the deregulated electricity market. The developed model consists of three components: a Fitz-Hugh Nagumo model (FHN) model to mimic spiking behaviour, RNN unit to regulate FHN and feed-forward neural network to model the residue of RNN-FHN. Success of this synergistic combination of RNN and coupled system of equations presents exciting opportunities for future work in day-ahead prediction in this time series system using multi-scale system. Short-term electricity price forecasting (i.e. day-ahead hourly electricity price forecasting) in case of organized power exchanges in developing nations (like India where power sector is getting deregulated due to reforms such as Electricity Act 2003) is one of the directions for future research. The Central Electricity Regulatory Commission issued a discussion paper for setting up a common platform for trading of electricity on February 6, 2007. After much debate and discussion, the plan for setting up power exchanges within the country was formulated. Applications from two exchanges, namely Power Exchange of India and Indian Energy Exchange, were submitted and approved by June and September 2008, respectively [23]. Both these exchanges have active participation from various utilities, and both provide electronic platforms for trading electricity on a day-ahead basis. Indian electricity market is one of the primitive electricity markets in the world which has much resistance to completely deregulate the electricity market based on the market equilibrium. There are very limited studies related to short-term electricity price forecasting for developing countries, particularly India, where electricity markets are getting deregulated. With practical relevance and stakes for implications of price forecasting for a Generator/Indian Power Producer (IPP)/firm being high (i.e. next-day price forecasting is a crucial need for producers, consumers and energy service companies), accurate forecasting tools and techniques is valued by power market participants, especially in an emerging economy like India. One of the directions for future research is developing a Time Series Econometric forecasting model. The model has to accurately forecast day-ahead hourly electricity

prices in Indian Electricity market. Therefore, the proposed research work mainly focuses on developing ANN-based forecasting tools for the Indian Energy Exchange Ltd. (IEX) launched in June 2008.

In the literature, though similar-day approach is carried out for electricity price forecasting [19], most of the neural network-based research papers are using day-ahead electricity pricing approach [4, 6, 8, 15]. Here, the electricity price data of the previous day (n th day) are mapped with the next day ($n + 1$)th day while training neural networks. The reason may be that the strength of the correlation of both linear and nonlinear parameters between n th and ($n + 1$)th day is stronger. For example, the MCP profile on Monday of the previous day is mapped to Tuesday of the next day. So when a test input of the n th day is fed into the forecast model, the MCP of the ($n + 1$)th MCP is forecasted. However, it should be noted the problem formulation is short-term electricity price forecasting in which the electricity price or MCP is forecasted for a day or a week.

2 Proposed work

The proposed research work focuses to develop a hybrid ANN forecast engine solely to Indian Energy Exchange (IEX), since only meagre work is carried out in this market. Hybrid ANN models, which combine heuristic search algorithms, such as genetic algorithms, particle swarm optimization, artificial bee colony algorithms for updating the weights, show some better performance. Among these optimization algorithms, particle swarm optimization [13, 14] is regarded as a promising method in several engineering applications as is evident in the literature. Therefore, the proposed research work develops a novel ANN-based training scheme using PSO for forecasting the MCP. Even though similar-day and day-ahead approaches are available in the literature, day-ahead approach is found to be more appropriate. Therefore, this

research work investigates day-ahead training approach for the electricity price forecasting. In day-ahead training, the correlation is between n th and ($n + 1$)th day. It is also understood that the performance of forecasting accuracy is improved by pre-processing of data before training. Among the data processing techniques, WT method can be pointed out as the prevailing approach [21] due to its easy implementation and adaptive ability of time frequency analysis. Therefore, WT is used for pre-processing in the proposed work.

3 Indian energy exchange and its historical database

IEX is India's premier power trading platform. By providing a transparent, neutral, demutualized and automated platform for physical delivery of electricity, IEX enables efficient price discovery and price risk management for participants in the electricity market, including industries eligible for open access.

In this exchange, there are about 4000 participants across utilities from 29 States, 5 Union Territories (UTs), 1000 + private generators (both commercial and renewable energy). More than 3000 open access consumers are leveraging the exchange platform to manage their power portfolio in the most competitive and reliable way. A typical market snapshot of IEX can be obtained from <http://www.iexindia.com/marketdata/marketsnapshot.aspx>. The most influential historical data [20], namely hourly Purchase Bid (MW) and Market Clearing Price (INR), are taken from the website for the proposed work. The historical data considered for performance analysis are taken for a time period from 4 January 2014 to 1 November 2014. The smoothing feature of the Daubechies [9] wavelet of order 4 (db4) is used to remove higher-frequency component. The PB and MCP for the same time frame are shown in Figs. 1 and 2 without and with pre-processing using WT technique.

Fig. 1 Market clearing price

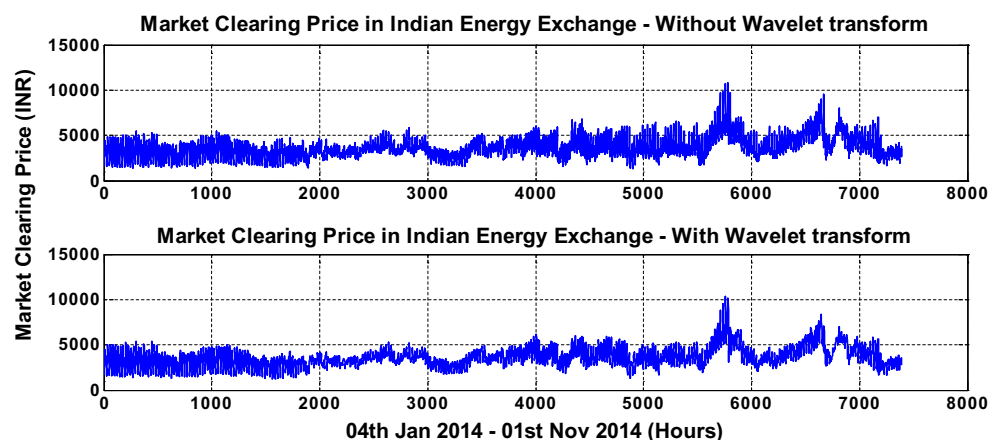
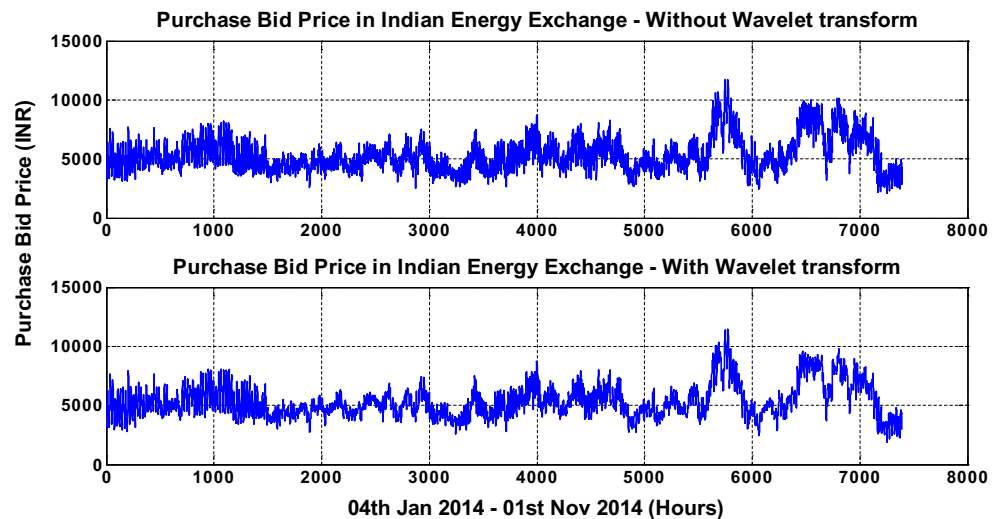


Fig. 2 Purchase bid price

The historical dataset is usually not used directly in process modelling of ANNs due to the difference in magnitude of the process variables. Therefore, the data need to be scaled to a fixed range to prevent unnecessary domination of certain variables, and to prevent data with larger magnitude from overriding the smaller and impede the premature learning process. The choice of range depends on the transfer function of the output nodes in ANN. Typically, $[0, 1]$ for sigmoid function and $[-1, 1]$ for hyperbolic tangent function. However, due to nonlinear transfer function having an asymptotic limit, the range of dataset is always set slightly less than the lower and upper limits. In this work, since the sigmoid function is adopted, the data are normalized in the range of $[0.1, 0.9]$. That is, if x_1 and x_2 are the maximum and minimum values of the training set, then the normalized data are given by $N(x)$ as in (1).

$$N(x) = \left(\frac{(x - x_1) \times (0.1 - 0.9)}{(x_2 - x_1)} \right) + 0.9 \quad (1)$$

Here, the total number of input nodes for the feed-forward back-propagation neural network is equal to 48. The first 24 inputs represent the hourly PB in the day, and the next 24 inputs represent the hourly MCP. The total number of output nodes is 24 since MCP has to be forecasted using day-ahead forecasting. According to literature, generally one hidden layer is sufficient for most neural network applications. The number of hidden neurons in the hidden layer is fixed either by trial and error or by statistical evaluation.

Generally, during ANN training, the weights are updated based on conventional gradient descent method. It should be noted that the weights can be updated either in incremental or in batch modes. In incremental mode of training, weights are updated every time an individual

Table 1 Batch mode versus incremental mode

Hidden neurons	Batch mode		Incremental mode	
	AMSE	Time taken (s)	AMSE	Time taken (s)
1	0.0013	337.51	0.0015	2.11E+03
3	0.0012	430.58	0.0012	2.13E+04
5	0.0012	503.19	0.0014	4.81E+03
Average	0.0012	423.76	0.0014	9.39E+03

pattern in a training set is sent into the network. While in batch mode of training, the weights are updated simultaneously for the entire training set. In order to choose between either modes of training, performance analysis between incremental and batch modes is carried out using historical data of IEX from 12 October 2014 to 01 November 2014. The number of hidden nodes is varied from 1 to 5. The final result is tabulated in Table 1. The result from Table 1 infers that the average mean square error (AMSE) and the time taken are found to be better for batch mode of training. Therefore, in the proposed model, batch mode training is used.

4 Proposed methodology

The proposed methodology involves three phases (Phase A, Phase B and Phase C) of operation. In Phase A, the raw data are pre-processed by removing the high-frequency components as mentioned in Sect. 3. In Phase B, the training of feed-forward back-propagation neural network is carried out in the batch mode using the conventional gradient-based neural network. Before the start of the training, the weights are initialized randomly and the

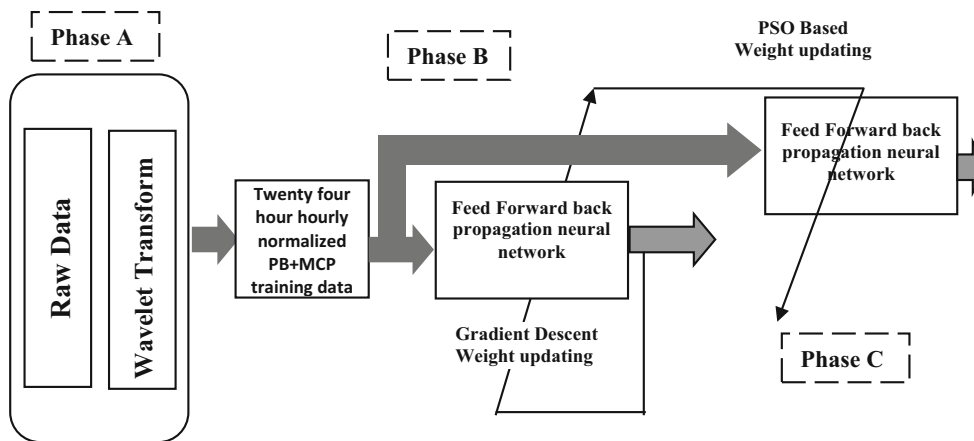


Fig. 3 Block diagram of the proposed wavelet-based ANN-ANN-PSO model

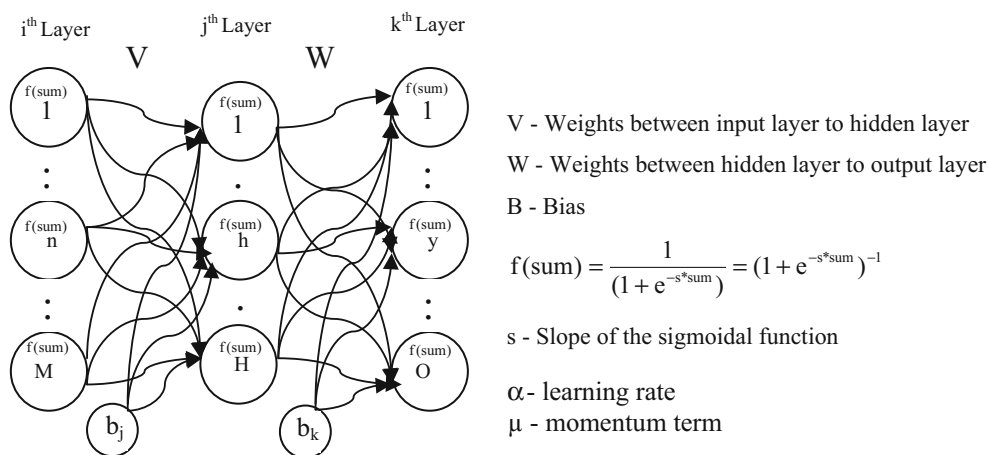


Fig. 4 Architecture of feed-forward back-propagation neural network

training is continued till the training error gets minimized and no further improvement is possible. Then the final weights obtained for the network are recorded. This phase is repeated for a sufficient number of trials, and the final weights obtained in each of the trials are recorded. In the Phase C, the final weights obtained from the trials will be the initial population for the ANN-based training using PSO. The optimal number of population is fixed from the number of trials carried out in the first phase of operation. Since, it involves four techniques used in three phases of operation; the model can be referred to as wavelet-based ANN-ANN-PSO. The PSO algorithm will improve upon the training of weights from the point where the conventional gradient-based training gets stagnated. The proposed wavelet-based ANN-ANN-PSO model improves upon the PSO-based training in the right direction. The block diagram shown in Fig. 3 gives an outline picture on the proposed model.

4.1 Batch mode training algorithm for FFBPNN

The generalized architecture of FFBPNN (Phase B) is given in Fig. 4.

Step 1 Set epoch $ep = 1$.

Step 2 Generate the weights $(V_{n \times h}, W_{h \times y})$ randomly to small random values between 0 and 1 to ensure that the network is not saturated by large values of weights. Let I and T be the normalized input and target training vector from set of P number of training patterns.

Step 3 For the training data, present one set of inputs and outputs. Present the complete input matrix $[I]$ to the input layer.

- $[I]_{P \times M}$ input training set
- $I_p = (i_1, \dots, i_n, \dots, i_M)$ p th input training vector
- $[T]_{P \times O}$ output target set
- $T_p = (t_1, \dots, t_y, \dots, t_O)$ p th target training vector

Step 4 Compute the inputs to the hidden layer by multiplying corresponding weights as in (2)

$$[\text{sum}]_{H \times P} = B_{H \times P} + [V]_{M \times H}^T \times [I]_{M \times P} \quad (2)$$

Step 5 Evaluate the hidden layer units' output using the sigmoidal function in (3)

$$f(\text{sum}_{H \times P}) = \frac{1}{1 + \exp(-\text{sum}_{H \times P})}. \quad (3)$$

Step 6 Compute the inputs to the output layer by multiplying corresponding weights of synapses as in (4)

$$[\text{sum}]_{O \times P} = B_{O \times P} + [W]_{H \times O}^T \times f(\text{sum}_{H \times P}) \quad (4)$$

Step 7 Let the output layer units evaluate the output using sigmoidal function as in (5)

$$[K]_{O \times P} = f(\text{sum}_{O \times P}) = \frac{1}{1 + \exp(-\text{sum}_{O \times P})} \quad (5)$$

Step 8 Calculate the squared error $[\varepsilon]_{O \times P}$, the difference between the network output $[K]_{O \times P}$ and the desired target $[T]_{O \times P}$, for all the training pairs as in (6) and then the average mean squared error (AMSE) as in (7), which is calculated for every epoch. Update $\text{ep} = \text{ep} + 1$.

$$[\varepsilon]_{O \times P} = [[T]_{O \times P} - [K]_{O \times P}]^2 \quad (6)$$

$$\text{AMSE} = \frac{\sum_{y=1}^O [\varepsilon]_{O \times P}}{P \times O} \quad (7)$$

Step 9 Calculate the updation of weights ($W_h \times y$) between hidden layer and output layer.

$$[d]_{O \times P} = 2 \times s \times [[T]_{O \times P} - [K]_{O \times P}] \times [K]_{O \times P} \times [[T]_{O \times P} - [K]_{O \times P}] \quad (8)$$

$$[Y]_{H \times O} = f(\text{sum}_{H \times P}) \times [d]_{O \times P}^T \quad (9)$$

$$[YB]_{O \times P} = [d]_{O \times P} \times [1]_{P \times P} \quad (10)$$

$$[\Delta W]_{H \times O}^{ep+1} = \mu \times [Y]_{H \times O} + \alpha \times [Y]_{H \times O} \quad (11)$$

$$[\Delta BW]_{O \times P}^{ep+1} = \mu \times [YB]_{O \times P} + \alpha \times [YB]_{O \times P} \quad (12)$$

$$[W]_{H \times O}^{ep+1} = [W]_{H \times O} + [\Delta W]_{H \times O}^{ep+1} \quad (13)$$

$$[BW]_{O \times P}^{ep+1} = [BW]_{O \times P} + [\Delta BW]_{O \times P}^{ep+1} \quad (14)$$

Step 10 Calculate the updation of weights ($V_n \times h$) between input and hidden layer.

$$[e]_{H \times P} = [W]_{H \times O} \times [d]_{O \times P} \quad (15)$$

$$[dx]_{H \times P} = [e]_{H \times P} \times f(\text{sum}_{H \times P}) \times [1 - f(\text{sum}_{H \times P})] \quad (16)$$

$$[X]_{M \times H} = [I]_{M \times P} \times [dx]_{H \times P}^T \quad (17)$$

$$[XB]_{H \times P} = [dx]_{H \times P} + [1]_{P \times P} \quad (18)$$

$$[\Delta V]_{M \times H}^{ep+1} = \mu \times [X]_{M \times H} + \alpha \times [X]_{M \times H} \quad (19)$$

$$[\Delta BV]_{H \times P}^{ep+1} = \mu \times [XB]_{H \times P} + \alpha \times [XB]_{H \times P} \quad (20)$$

$$[V]_{M \times H}^{ep+1} = [V]_{M \times H} + [\Delta V]_{M \times H}^{ep+1} \quad (21)$$

$$[BV]_{H \times P}^{ep+1} = [BV]_{H \times P} + [\Delta BV]_{H \times P}^{ep+1} \quad (22)$$

$$[V]^{t+1} = [V]^t + [\Delta V]^{t+1} \quad (23)$$

$$[W]^{t+1} = [W]^t + [\Delta W]^{t+1} \quad (24)$$

Step 11 Repeat steps 3 to 10, if $\text{ep} < \text{TE}$ (total number of epochs) or if AMSE has reached a desired minimum value or if the validation error is increasing such that the number validation checks is greater than the validation count (VC), then stop the training.

4.2 Step-by-step algorithm of ANN-ANN-PSO architecture

The iterative approach of PSO for ANN training (Phase C) can be described by the following:

Step 1 Initialize the population size which is also equal to the number of trials performed in Phase B. Get the previously stored recorded weights and biases obtained in Phase B. Also, initialize the positions and velocities of agents

Step 2 Set the current best fitness achieved by particle p as $pbest$. Set the $pbest$ with best value as $gbest$, and store the value

Step 3 Evaluate the desired optimization fitness function f_p for each particle as the AMSE over a given data set

Step 4 Compare the evaluated fitness value f_p of each particle with its $pbest$ value. If $f_p < pbest$, then $pbest = f_p$ and $pbest_{xp} = f_p$, x_p is the current coordinates of particle p , and $pbest_{xp}$ is the coordinates corresponding to particle p as the best fitness so far

Step 5 Calculate objective function value for new positions of each particle. If a better position is achieved by an agent, $pbest$ value is replaced by the current value. As in step 2, $gbest$ value is selected among $pbest$ values. If the new $gbest$ value is better than the previous $gbest$ value, the $gbest$ value is replaced by the current $gbest$ value and this value is stored. If $f_p < gbest$, then $gbest = p$, where $gbest$ is the particle having the overall best fitness over all particles in the swarm

Step 6 Change the velocity and location of the particle according to Eq. (25) and (26), respectively [13, 14]

$$V_i = wV_{i-1} + acc \times rand() \times (best_{xp} - x_p) + acc \times rand() \times (best_{xgbest} - x_p) \tag{25}$$

where *acc* is the acceleration constant that controls how far particles *p* move from one another, and *rand()* returns a uniform random number between 0 and 1, *V_i* is the current velocity, *V_{i-1}* is the previous velocity,

$$x_p = x_{pp} + V_i \tag{26}$$

x_p is the present location of the particle, *x_{pp}* is the previous location of the particle, and *i* is the particle index. Here, the coordinates *best_{xp}* and *best_{xgbest}* are used to pull the particles towards the global minimum

Step 7 Fly each particle *p* according to Eq. (26)

Step 8 If the maximum number of predetermined iterations is exceeded, then stop; otherwise, go to step 3 until convergence

The flow chart for the proposed wavelet-based ANN-ANN-PSO model is given in Fig. 5.

5 Performance evaluation

The experimental results in this case study are evaluated based on three error indices. They are: mean absolute percentage error (MAPE), normalized mean square error (NMSE) and error variance (EV). The accuracy of the forecasted results is evaluated by mean absolute percentage error (MAPE) which is defined by the following Eq. (27).

$$MAPE = \frac{1}{NH} \sum_{i=1}^{NH} \left| \frac{P_{For(t)} - A_{Ac(t)}}{A_{Ac(t)}} \right| \tag{27}$$

where *P_{For(t)}* and *A_{Ac(t)}* are the forecasted and actual data at time *t*, respectively, and *NH* is the total number of predictions.

The normalized mean square error (NMSE) is an estimator of the overall deviations between predicted and measured values. Here, in the NMSE the deviations (absolute values) are summed instead of the differences. For this reason, the NMSE generally shows the most striking differences among models. If a model has a very low NMSE, then it is well performing both in space and in time. On the other hand, high NMSE values do not necessarily mean that a model is completely wrong. NMSE [5] is defined by Eq. (28) given below

$$NMSE = \left[\frac{1}{\Delta^2 NH} \sum_{t=1}^{NH} (P_{For(t)} - A_{Ac(t)})^2 \right] \tag{28}$$

where

$$\Delta = \frac{1}{NH - 1} \sum_{t=1}^{NH} (A_{Ac(t)} - A_{Ave})^2 \tag{29}$$

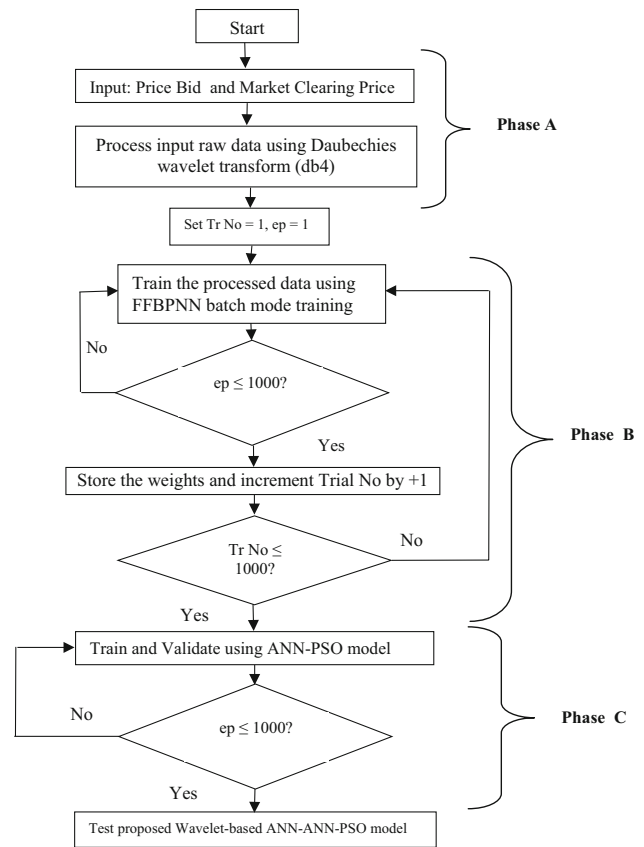


Fig. 5 Flow chart for the proposed wavelet-based ANN-ANN-PSO model

where *A_{Ave}* is average of actual data.

An index of the uncertainty of a model is the variability of what is still unexplained after fitting the model, which can be measured through the estimation of the variance of the error term. The smaller the variance, the more precise the prediction of prices [8]. EV [5] is defined by the following Eq. (30)

$$\sigma^2 = \frac{1}{NH} \sum_{i=1}^{NH} \left(\left| \frac{P_{For(t)} - A_{Ac(t)}}{A_{Ac(t)}} \right| - MAPE \right)^2 \tag{30}$$

6 Optimal ANN architecture and training

In order to obtain an accurate forecasting for training of neural networks, the selection of a suitable ANN architecture with an optimal size of historical data becomes important. During training, weights are updated simultaneously only when an epoch is completed. The performance valuation can be accomplished by forecasting the MCP for a particular week using day-ahead approach. The

Table 2 Statistical performance measures for 1-week training data

No. of trials = 25, no. of epochs = 1000, without wavelet transform				
Training data		Source range	18–24 October 2014	
		Target range	19–25 October 2014	
Test data		Input range	25–31 October 2014	
		Output range	26 October 2014–01 November 2014	
No. of hidden nodes	Avg. MAEP (for 25 trials)	Min. mean absolute error percentage (MAEP)	Normalized mean square error (NMSE)	Error variance (EV)
<i>1</i>	<i>15.0914</i>	<i>15.0135</i>	<i>5.71E–06</i>	<i>220.8800</i>
3	16.8084	16.4807	6.20E–06	266.2247
5	17.3327	17.1081	6.82E–06	286.8803
7	18.5415	16.4600	5.81E–06	265.5519
9	21.9267	16.1963	6.54E–06	257.1161
11	24.9352	16.6153	7.03E–06	270.5952
13	25.9612	16.7656	8.09E–06	275.5155
15	26.4648	19.0757	8.73E–06	356.6609
20	26.1575	17.0041	6.55E–06	283.3991

Table 3 Statistical performance measures for 2-week training data

No. of trials = 25, no. of epochs = 1000, without wavelet transform				
Training data		Source range	11–24 October 2014	
		Target range	12–25 October 2014	
Test data		Input range	25–31 October 2014	
		Output range	26 October–01 November 2014	
No. of hidden nodes	Avg. MAEP (for 25 trials)	Min. mean absolute error percentage (MAEP)	Normalized mean square error (NMSE)	Error variance (EV)
<i>1</i>	<i>25.7586</i>	<i>24.4491</i>	<i>1.07E–05</i>	<i>585.8902</i>
<i>3</i>	<i>18.9983</i>	<i>17.7463</i>	<i>6.27E–06</i>	<i>308.6750</i>
5	22.4375	14.3346	4.54E–06	201.4031
7	26.9552	14.0058	5.44E–06	192.2738
9	28.4593	15.4736	5.93E–06	234.6831
11	28.2916	14.5007	6.91E–06	206.1069
13	27.7325	15.1655	6.34E–06	225.4317
15	29.6471	16.2006	6.80E–06	257.2529
20	29.3321	15.4846	6.01E–06	235.0181

Table 4 Statistical performance measures for 3-week training data

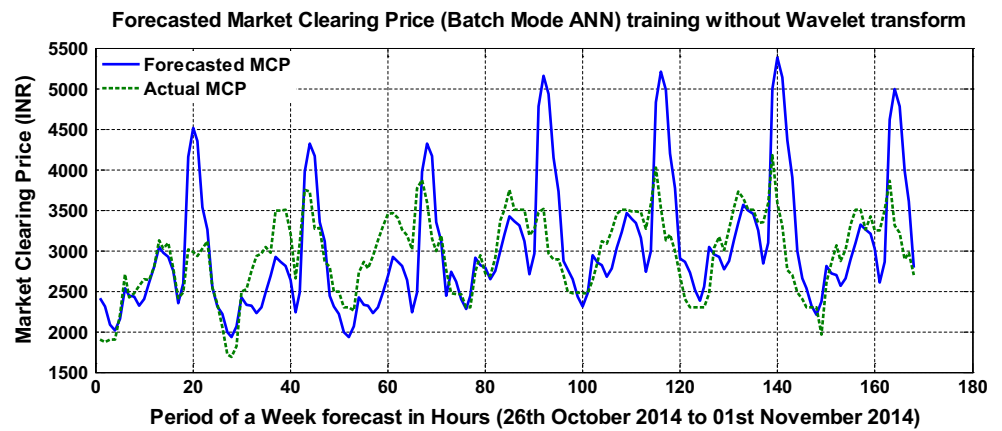
No. of trials = 25, no. of epochs = 1000, without wavelet transform				
Training data		Source range	27 September–24 October 2014	
		Target range	28 September–25 October 2014	
Test data		Input range	25–31 October 2014	
		Output range	26 October–01 November 2014	
No. of hidden nodes	Avg. MAEP (for 25 trials)	Min. mean absolute error percentage (MAEP)	Normalized mean square error (NMSE)	Error variance (EV)
<i>1</i>	<i>37.4548</i>	<i>17.1723</i>	<i>5.26E–06</i>	<i>289.0378</i>
3	75.5396	25.4230	1.32E–05	633.4905
5	73.2888	18.0471	5.88E–06	319.2302
7	70.1434	17.2489	5.36E–06	291.6219
9	67.3691	18.8415	6.52E–06	347.9541
11	57.7053	17.0209	5.22E–06	283.9588
13	64.133	17.4066	5.50E–06	296.9765
15	68.203	17.5463	5.62E–06	301.7608
20	71.6385	17.7191	5.76E–06	307.7312

same week is forecasted with various combinations of hidden neurons and training size. Tables 2, 3 and 4 shown below furnish statistical evaluation of the performance for selecting the best ANN architecture and training size. The number of trials and number of epochs performed for a typical ANN architecture are fixed to 25 and 1000, respectively, in order to have a reference platform for better comparison. Having considered 1 week of training data, Fig. 6 presents the best forecasted MCP for a week. The italicized values in Tables 2, 3 and 4 provide the

optimal architecture of the neural network with 1, 3 and 1 hidden neurons, respectively and are compared with respect to average MAEP obtained from all the 25 trials.

If the value of average MAEP is more, then the performance is not found to be satisfactory. Therefore, the best architecture is selected only if the average MAEP is minimal. It is inferred from Tables 2, 3 and 4 that as the number of weeks is increased, i.e. the size of the training data, the performance is not improving. Therefore, it is best to consider the training data solely for 1 week.

Fig. 6 Forecasted market clearing price



6.1 Implementation

Once the optimal architecture for ANN is finalized, the performance analysis part is carried out in two parts. In the first part, the raw data are taken as such, whereas in the second part, the raw data are filtered for high-frequency components using WT technique. The resolution of the signal, which is a measure of the amount of detailed information, is determined by filtering operations. Discrete WT decomposes a time domain signal into approximations and details by successive low-pass and high-pass filtering. Here, Daubechies wavelet transform is used. This wavelet offers an appropriate trade-off between wavelength and smoothness, resulting in an appropriate behaviour for MCP forecasting. A non-decimated wavelet function of type Daubechies of order 4 (abbreviated as db4) and decomposition level 4 is used as the mother wavelet.

Embedded ANN-PSO simulation in Phase C is carried out using the weights obtained from all the trials from Phase B. These weights are the initial population for ANN-PSO training. The architecture of the ANN-PSO is the same as that of Phase B ANN architecture. Table 5 presents the controlling parameters for the PSO algorithm. The parameter setting for the PSO algorithm is fixed based on trial and error approach. Here, the convergence of the error plot is not as smooth as that of ANN Phase B training. The reason is the stochastic nature of the particle search in finding the optimal weights for minimizing the MAPE while training and also satisfying the validation check. A comparison is also made with Phase C

where only ANN-PSO training is carried out with the random initialization of weights for ANN training using PSO.

7 Results and discussion

All the simulations are carried out using MATLAB programming using MATLAB R2011b version. The hardware details are as follows, Intel® Core i5-3210 M CPU @ 2.50 GHz, 4.00 GB RAM, 64-bit OS, x64-based processor using Windows 8.1 OS. Various samples are taken in order to validate the better performance of the proposed model. All trainings are carried out for 1000 epochs and 100 trial runs. The following are the five types of ANN-based forecast models used for the performance study.

- (a) ANN
- (b) ANN-ANN-PSO
- (c) Wavelet-based ANN
- (d) Wavelet-based ANN-ANN-PSO
- (e) Wavelet-based ANN-PSO (random initialization)

The same data used in finding the optimal architecture of ANN are again compared for all the five models (a–e). Here, the source and the target for the training data are taken from 27 September to 3 October 2014 and from 28 September to 4 October 2014, respectively. The source and the target for the validation data are the same as those of the training data where the validation error is evaluated based on MAPE. The verification is carried out by giving a test input from 4 to 10 October

Table 5 PSO parameter settings

Parameters	No. of population	Acceleration constant (local)	Acceleration constant (global)	Max. velocity	Inertial weights	
					Initial iteration	Final iteration (3500)
Value	25	0.0005	0.9999	100000	0.9	0.1

2014, and the forecasted resultant output is compared with the actual MCP from 05 to 11 October 2014 for a period of 1 week using day-ahead forecasting. The performance of the accuracy in forecasting is evaluated using the performance indices explained in Sect. 5. Table 6 summarizes the results of all the ANN-based models. Figures 7 and 8 show the convergence plot and validation plot of the best trial for the ANN-based forecast models, i.e. b, d and e, respectively. The wavelet-based ANN–ANN-PSO training is found to converge better when compared with the other models b and e. The convergence of ANN-PSO training with initial random initialization of weights is found to be non-satisfactory. The random initialization of weights of ANN-PSO will not give an immediate state of orientation of the weights or the particles towards the near-global optima. However, during Phase B operation, the recorded final weights of the batch mode ANN training for all trials give the initialization of particles for the PSO search to orient the weights towards the near-global optimum. Thereby, fine tuning of weight updation gives a better trained model without getting trapped in the local minima. Hence, the performance due to ANN–ANN-PSO model is better than the embedded ANN-PSO model. The same performance is also reflected while validation during training. The validation error is evaluated in terms of MAEP, and the WT-based ANN–ANN-PSO model is able to give a lower validation error of 9.25 % for the best trial.

Figures 9 and 10 show the forecasted MCP with respect to the actual MCP for the IEX from 5 to 11 October 2014 for the ANN models without wavelet pre-processing (i.e. a and b) and with wavelet pre-processing (i.e. c and d), respectively. However, the resultant forecasted ANN-PSO model with random initialization (see Fig. 11) is found to perform poorly with an average MAEP of around 70 %. The total time (Phase A + Phase B) taken for training of WT-based ANN (no. of trials = 100; no. of epochs = 1000) is 949 s. The average time (Phase C) taken for training of WT-based ANN-PSO (no. of trials = 1) is 650 s. Therefore, the total approximate execution time for the proposed model is 1599 s. In order to validate the proposed forecast model, some more samples are considered and the statistical results are tabulated in “Appendix” as Tables 7, 8 9. Also Figs. 12, 13, 14 give the forecasting market clearing price for all these samples. The resultant forecasted MCP validates the accuracy of the proposed WT-based ANN–ANN-PSO hybrid model.

8 Conclusion

The pre-processing of wavelet by smoothening of raw data by removing higher-frequency components helps the neural network to train better. The weights obtained from the various trials of ANN training in Phase B helps a better start for PSO search of weights for ANN training in Phase

Table 6 Statistical result

Model	Average of all trials						Best/min. of trials					
	MAEP	MSE	NMSE	EV	VE	Training error or avg MSE	Best trial no.	Min. MAEP	EV	VE	Training error or avg MSE	
No. of trials = 100, no. of epochs = 1000												
Training data												
Validation data												
Test data												
ANN	18.2229	9.7596E-03	9.9614E-07	3.3364E+02	2.2257E+01	6	10.1464	4.0563E-07	1.0091E+02	2.4122E+01	5.4926E-03	
ANN-ANN-PSO	11.3774	4.9872E-03	3.3543E-08	1.2779E+02	1.4699E+01	3	9.9469	2.8291E-08	9.6982E+01	1.4589E+01	4.9609E-03	
Wavelet-based ANN	18.2688	9.3942E-03	1.0004E-06	3.3844E+02	2.1441E+01	78	10.0097	4.1598E-07	9.8211E+01	2.0511E+01	5.8185E-03	
Wavelet-based ANN-ANN-PSO	11.0744	5.2619E-03	3.3566E-08	1.2100E+02	1.4981E+01	97	9.2523	2.5696E-08	8.3911E+01	1.4265E+01	4.7761E-03	
Wavelet-based ANN-PSO (random initialization)	70.0946	1.1219E-01	1.0550E-06	4.8520E+03	7.0092E+01	43	56.3747	7.3554E-07	3.1150E+03	5.5449E+01	7.5872E-02	

MAEP mean absolute error percentage, MSE mean square error, NMSE normalized mean square error, EV error variance, VE validation error in MAEP

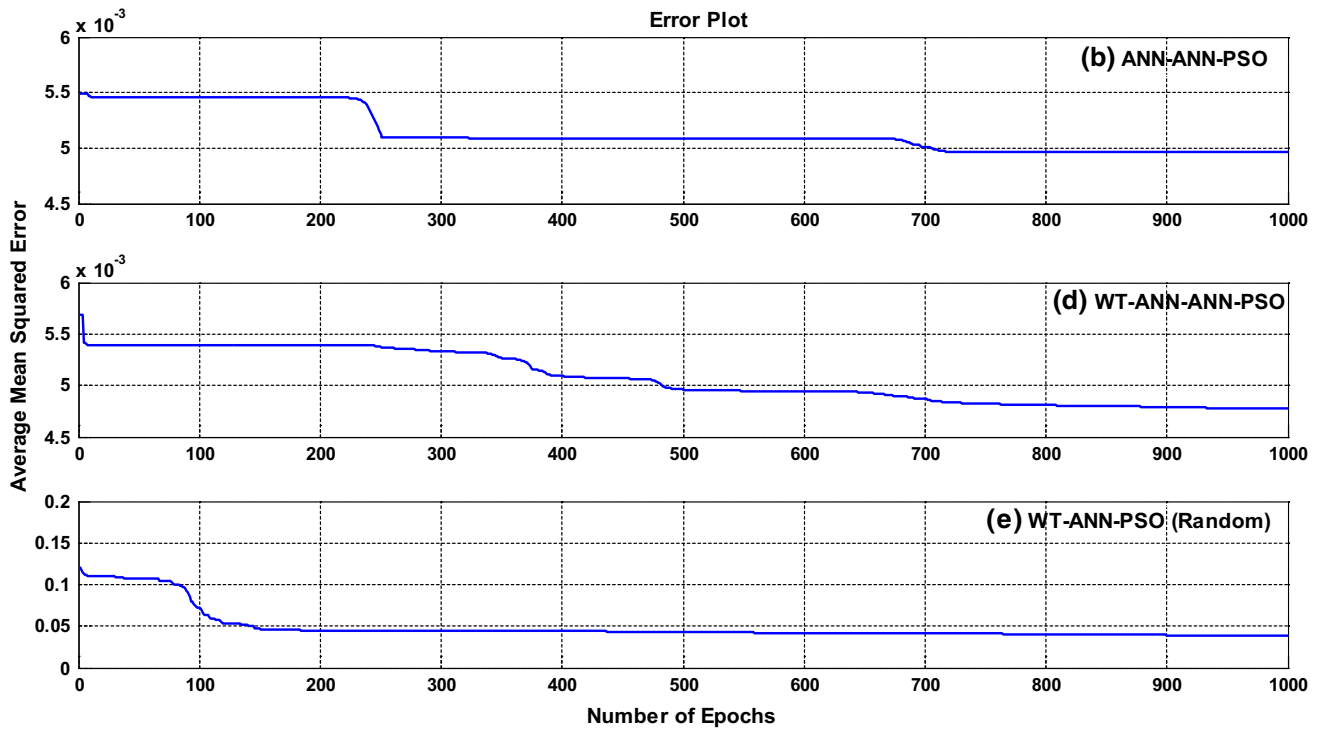


Fig. 7 Error convergence plot for the ANN-ANN-PSO-based models (b, d and e)

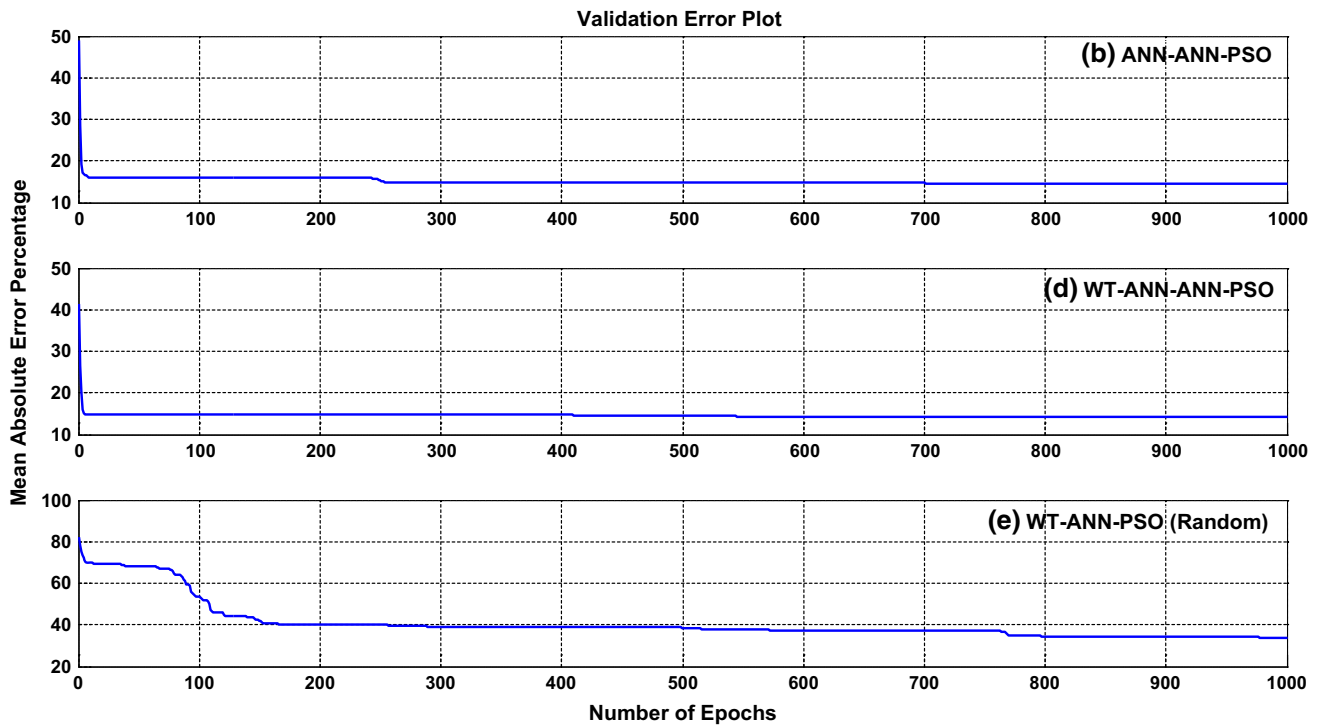


Fig. 8 Validation plot for the ANN-ANN-PSO-based models (b, d and e)

Fig. 9 Forecasted MCP versus actual MCP for the IEX—without wavelet transform

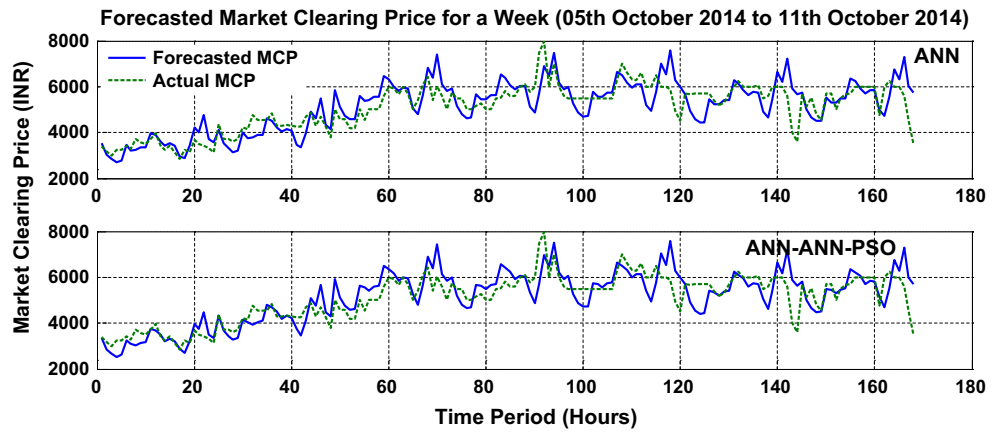


Fig. 10 Forecasted MCP versus actual MCP for the IEX—with wavelet transform

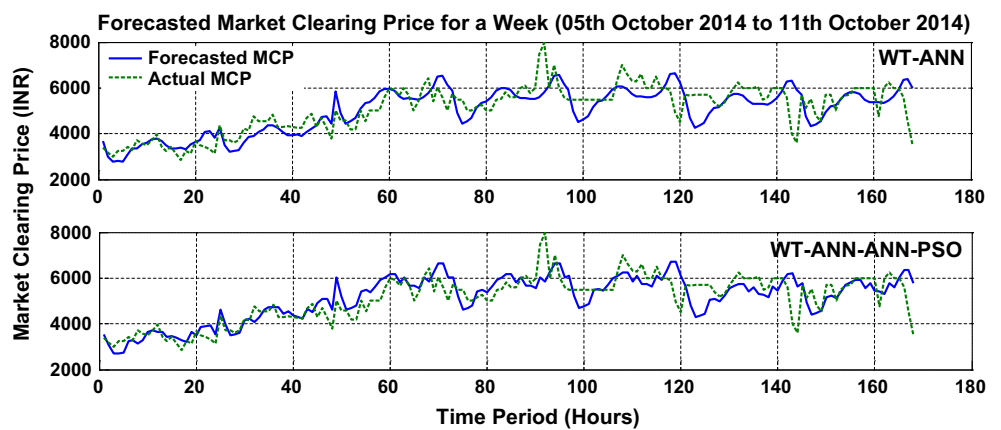
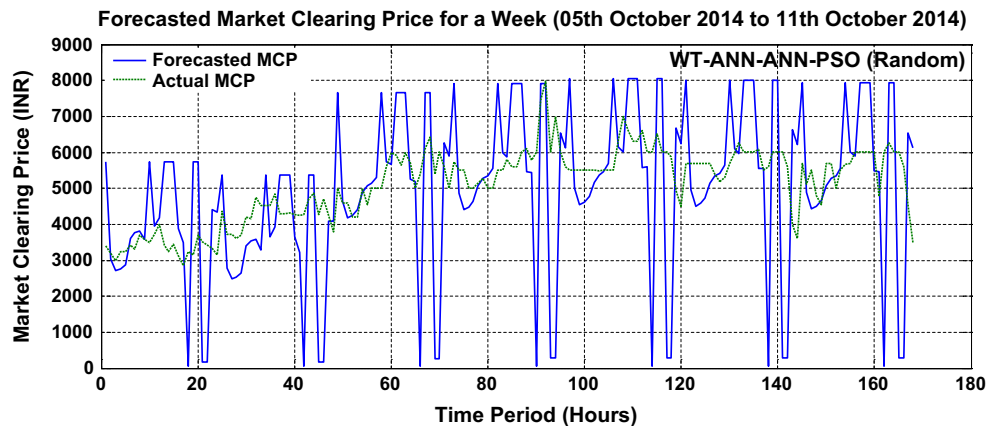


Fig. 11 Forecasted MCP VS actual MCP for the IEX—with wavelet transform (random)



C. The method is simple, and the performance of forecasting improves accuracy. Therefore, the proposed novel sequential wavelet-ANN with embedded ANN-PSO hybrid model can be used in the Indian energy exchange for a better estimate of market clearing price.

9 Appendix

See Tables 7, 8 9 and Figs. 12, 13, 14.

Table 7 Statistical result

No. of trials = 100, no. of epochs = 1000

Model	Average of all trials						Best/min. of trials			Training error or Avg MSE	
	MAEP	MSE	NMSE	EV	VE	Best trial no.	Min. MAEP	EV	VE		
ANN	13.5464	7.5794E-04	1.2472E-06	1.7995E+02	7.9616E+00	80	13.0963	1.1638E-06	1.6811E+02	9.3348E+00	7.7090E-04
ANN-ANN-PSO	13.1390	6.9308E-04	3.5291E-07	1.6921E+02	7.3978E+00	51	12.9346	3.4333E-07	1.6398E+02	7.3854E+00	6.9407E-04
Wavelet-based ANN	13.3383	6.1047E-04	1.2678E-06	1.7454E+02	7.5503E+00	36	13.0329	1.2132E-06	1.6648E+02	7.4727E+00	5.6076E-04
Wavelet-based ANN-ANN-PSO	12.9990	5.3454E-04	3.5992E-07	1.6562E+02	6.6630E+00	3	12.8764	3.5436E-07	1.6251E+02	6.8253E+00	5.4678E-04

Source range
Target range
Input range
Output range
Input range
Output range

13–19 September 2014
14–20 September 2014
13–19 September 2014
14–20 September 2014
20–26 September 2014
21–27 September 2014

MAEP mean absolute error percentage, MSE mean square error, NMSE normalized mean square error, EV error variance, VE validation error in MAEP

Table 8 Statistical result

No. of trials = 100, no. of epochs = 1000

Model	Average of all trials						Best/min. of trials			Training error or Avg MSE	
	MAEP	MSE	NMSE	EV	VE	Best trial no.	Min. MAEP	EV	VE		
ANN	15.0974	1.3000E-03	5.9059E-06	2.2345E+02	1.0868E+01	11	14.7888	5.9369E-06	2.1437E+02	1.0902E+01	1.3000E-03
ANN-ANN-PSO	15.0035	1.2822E-03	2.0526E-07	2.2085E+02	1.0866E+01	43	14.7534	2.0740E-07	2.1335E+02	1.0885E+01	1.2847E-03
Wavelet-based ANN	14.4359	9.7042E-04	4.2314E-06	2.0429E+02	1.0118E+01	53	14.1191	4.1001E-06	1.9539E+02	1.0191E+01	9.7856E-04
Wavelet-based ANN-ANN-PSO	14.3077	9.7048E-04	1.4538E-07	2.0075E+02	1.0103E+01	33	14.0773	1.4129E-07	1.9424E+02	1.0098E+01	9.7057E-04

Source range
Target range
Input range
Output range
Input range
Output range

18–24 October 2014
19–25 October 2014
18–24 October 2014
19–25 October 2014
25–31 October 2014
26 October 2014–01 November 2014

MAEP mean absolute error percentage, MSE mean square error, NMSE normalized mean square error, EV error variance, VE validation error in MAEP

Table 9 Statistical result

Model	Average of all trials						Best/min. of trials					
	MAEP	MSE	NMSE	EV	VE	Best trial no.	Min. MAEP	NMSE	EV	VE	Training error or avg MSE	
No. of trials = 100, no. of epochs = 1000												
Training data												
					Source range						14–20 June 2014	
					Target range						15–21 June 2014	
					Input range						14–20 June 2014	
					Output range						15–21 June 2014	
Validation data												
					Input range						21–27 June 2014	
					Output range						22–28 June 2014	
Test data												
ANN	24.4069	6.5482E-04	7.1342E-07	5.8387E-07	6.5433E+00	94	24.3832	7.1227E-07	5.8273E+02	6.5437E+00	6.5487E-04	
ANN-ANN-PSO	24.3927	6.5480E-04	6.5920E-07	5.8319E+02	6.5304E+00	64	24.3855	6.5888E-07	5.8284E+02	6.5377E+00	6.5481E-04	
Wavelet-based ANN	24.4082	6.5482E-04	7.1347E-07	5.8393E+02	6.5433E+00	77	24.3809	7.1219E-07	5.8262E+02	6.5437E+00	6.5487E-04	
Wavelet-based ANN-ANN-PSO	24.3842	6.5480E-04	6.5877E-07	5.8278E+02	6.5315E+00	36	24.3802	6.5859E-07	5.8259E+02	6.5408E+00	6.5481E-04	

MAEP mean absolute error percentage, MSE mean square error, NMSE normalized mean square error, EV error variance, VE validation error in MAEP

Fig. 12 Forecasted MCP versus actual MCP for the IEX

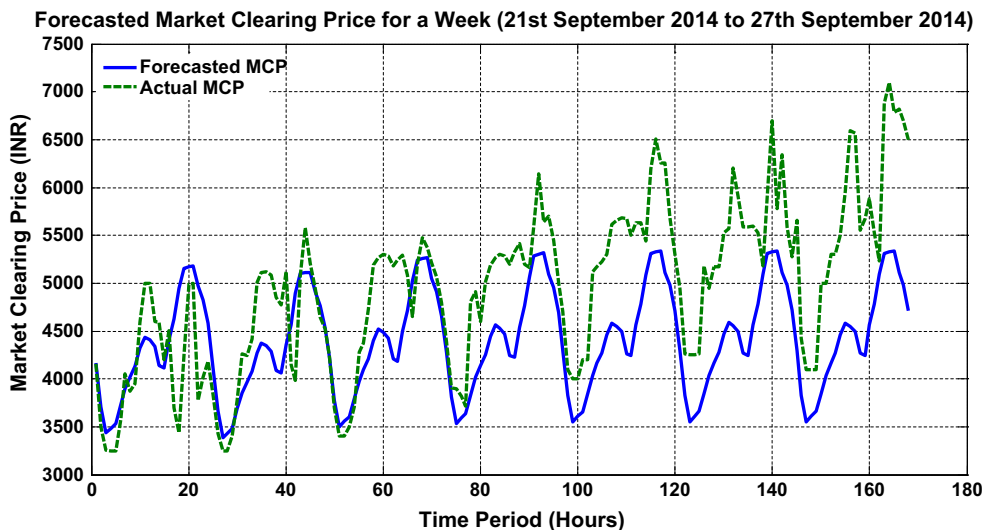


Fig. 13 Forecasted MCP versus actual MCP for the IEX

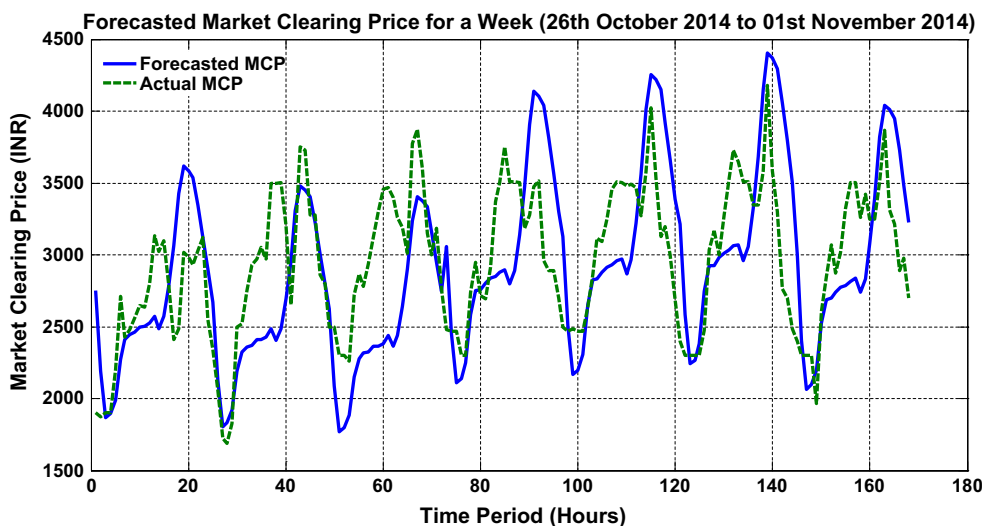
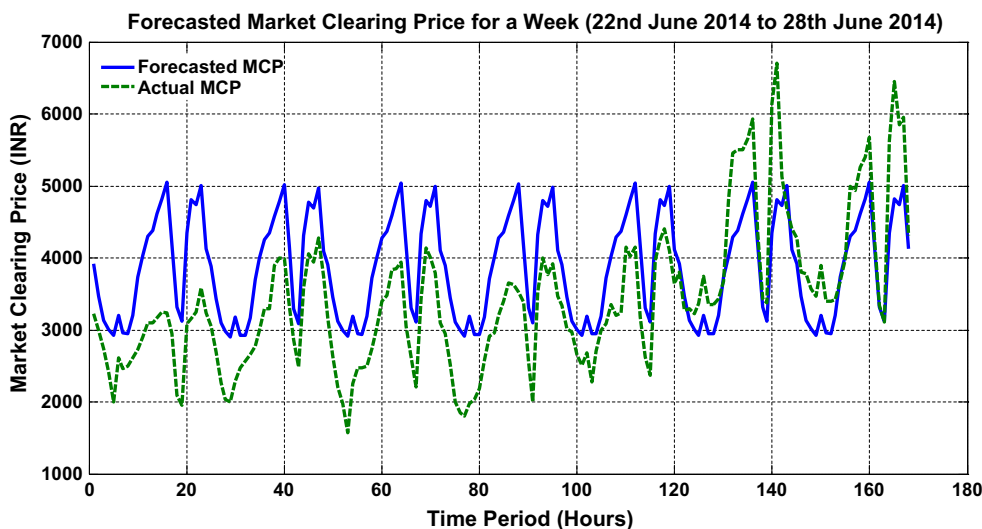


Fig. 14 Forecasted MCP versus actual MCP for the IEX



References

1. Aggarwal SK, Saini LM, Kumar A (2008) Electricity price forecasting in Ontario electricity market using wavelet transform in artificial neural network based model. *Int J Control Autom Syst* 6(5):639–650
2. Aggarwal SK, Saini LM, Kumar A (2009) Electricity price forecasting in deregulated markets: a review and evaluation. *Int J Electr Power* 31:13–22
3. Amjady N (2006) Day-ahead price forecasting of electricity markets by a new fuzzy neural network. *IEEE Trans Power Syst* 21(2):887–896
4. Amjady N, Daraeepour A (2009) Design of input vector for day-ahead price forecasting of electricity markets. *Expert Syst Appl* 36(10):12281–12294
5. Amjady N, Keynia F, Zareipour H (2011) Short term wind forecasting using Ridgelet neural network. *Electr Power Syst Res* 81:2099–2107
6. Anbazhagan S, Kumarappan N (2012) Day-ahead deregulated electricity market price classification using neural network input featured by DCT. *Int J Electr Power Energy Syst* 37:103–109
7. Catalao JPS, Pousinho HMI, Mendes VMF (2011) Hybrid wavelet-PSO-ANFIS approach for short-term electricity prices forecasting. *IEEE Trans Power Syst* 26(1):137–144
8. Conejo AJ, Plazas MA, Espínola R, Molina AB (2005) Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. *IEEE Trans Power Syst* 20(2):1035–1042
9. Daubechies I (1988) Orthonormal bases of compactly supported wavelets. *Commun Pure Appl Math* 41:909–996
10. Dev P, Martin MA (2014) Using neural networks and extreme value distributions to model electricity pool prices: evidence from the Australian National Electricity Market 1998–2013. *Energ Convers Manag* 84:122–132
11. Diongue KA, Guegan D, Vignal B (2004), A k-factor GIGARCH process: estimation and application on electricity market spot prices. In: Proceedings of the 2004 international conference on probabilistic methods applied to power systems, pp 1–7
12. Gao F, Cao X, Papalexopoulos A (2000) Forecasting power market clearing price and quantity using a neural network method. *IEEE Power Eng Soc Summer Meet* 4:2183–2188
13. Kennedy J (1997) The particle swarm: social adaptation of knowledge. In: Proceedings of IEEE international conference on evolutionary computation, pp 303–308
14. Kennedy J, Eberhart R (1995) Particle swarm optimization. *Proc IEEE Int Conf Neural Netw* 4:1942–1948
15. Li G, Liu C-C, Mattson C, Lawarree J (2007) Day-ahead electricity price forecasting in a grid environment. *IEEE Trans Power Syst* 2(1):266–274
16. Liu H, Shi J (2013) Applying ARMA–GARCH approaches to forecasting short-term electricity prices. *Energy Econ* 37:152–166
17. Mandal P, Haque AU, Meng J, Srivastava AK, Martinez R (2012) A novel hybrid approach using wavelet, firefly algorithm and fuzzy ARTMAP for day-ahead electricity price forecasting. *IEEE Trans Power Syst* 28(2):1041–1051
18. Nogales FJ, Contreras J, Conejo AJ, Espinola R (2002) Forecasting next day electricity prices by time series models. *IEEE Trans Power Syst* 17(2):342–348
19. Pany PK, Ghoshal SP (2015) Dynamic electricity price forecasting using local linear wavelet neural network. *Neural Comput Appl*. doi:10.1007/s00521-015-1867-2 (Publisher Springer, London)
20. Peter SE, Raglend IJ, Simon SP (2014) An architectural framework of ANN based short term electricity price forecast engine for indian energy exchange using similar day approach. *IMPACT Int J Res Eng Technol* 2(4):111–122
21. Pindoriya NM, Singh SN, Singh SK (2008) An adaptive wavelet neural network-based energy price forecasting in electricity markets. *IEEE Trans Power Syst* 23(3):1423–1432
22. Pousinho HMI, Mendes VMF, Catalao JPS (2012) Short-term electricity prices forecasting in a competitive market by a hybrid PSO–ANFIS approach. *Int J Electr Power* 39:29–35
23. Sarkar PK (2009) Power market development—key issues and solutions. *Power Line*, pp 42–44
24. Shahidehpour M, Yamin H, Li Z (2002) Market operations in electric power systems: forecasting, scheduling, and risk management. Wiley, New York
25. Sharma V, Srinivasan D (2013) A hybrid intelligent model based on recurrent neural networks and excitable dynamics for price prediction in deregulated electricity market. *Eng Appl Artif Intell* 26:1562–1574
26. Shrivastava NA, Panigrahi BK (2014) A hybrid wavelet-ELM based short term price forecasting for electricity markets. *Int J Electr Power* 55:41–50
27. Zhang L, Luh PB (2005) Neural network-based market clearing price prediction and confidence interval estimation with an improved extended Kalman filter method. *IEEE Trans Power Syst* 20(1):59–66