

# A hybrid wavelet-ELM based short term price forecasting for electricity markets



Nitin Anand Shrivastava\*, Bijaya Ketan Panigrahi

Department of Electrical Engineering, Indian Institute of Technology, New Delhi, India

## ARTICLE INFO

### Article history:

Received 12 March 2013

Received in revised form 11 July 2013

Accepted 23 August 2013

### Keywords:

Artificial Neural Network (ANN)

Deregulation

Ensemble

Extreme Learning Machine (ELM)

Price forecasting

Wavelet Transforms

## ABSTRACT

Accurate electricity price forecasting is a formidable challenge for market participants and managers owing to high volatility of the electricity prices. Price forecasting is also the most important management goal for market participants since it forms the basis of maximizing profits. This study investigates the performance of a novel neural network technique called Extreme Learning Machine (ELM) in the price forecasting problem. Keeping in view the risk associated with electricity markets with highly volatile prices, relying on a single technique is not so profitable. Therefore ELM has been coupled with the Wavelet technique to develop a hybrid model termed as WELM (wavelet based ELM) to improve the forecasting accuracy as well as reliability. In this way, the unique features of each tool are combined to capture different patterns in the data. The robustness of the model is further enhanced using the ensembling technique. Performances of the proposed models are evaluated by using data from Ontario, PJM, New York and Italian Electricity markets. The experimental results demonstrate that the proposed method is one of the most suitable price forecasting techniques.

© 2013 Elsevier Ltd. All rights reserved.

## 1. Introduction

Deregulation of electricity sector has led to the development of a competitive market structure where the participants compete for the market share through spot and bilateral markets. Electricity prices in such markets are directly or indirectly driven by a number of factors which are interlinked to each other in a complex fashion. Uncertainty in factors such as weather, equipment outages, fuel prices, and transmission bottlenecks result in extreme price volatility or even price spikes of electricity market. The complex, uncertain movement of electricity prices over different hours of the day is of great interest to the market participants. The market participants need reliable forecasted prices for either bidding or hedging against price volatility in the market. Driven by the importance of the future prices and the complexities involved in determining them, detailed modeling and forecasting of electricity prices has become a major research field in electrical engineering.

A significant number of research papers have addressed the problem of accurate price forecasting through different approaches. The most commonly observed one's are the time series based [1–4] and the artificial intelligence based approaches [5–8] which are basically regression models as they relate electricity prices variations to historical prices and other explanatory variables such as demand, fuel prices, temperature, time of day etc.

Some of the recent works which have addressed the problem of accurate electricity price forecasting are presented in [9–12]. In [9], the authors have based their research on the Grey system theory [GST] which considers uncertain systems which are partially known and have small samples. They proposed a new grey model in which the reference series is determined using the correlation method and the model parameters are identified using Particle Swarm Optimization (PSO) instead of Least Square method (LSM). A hybrid model comprising of Wavelet transforms (WT), Auto Regressive Integrated Moving Average (ARIMA) and Radial Basis Function Neural Network (RBFN) was proposed in [10]. Further, the network structure of RBFN used in this methodology is optimized using the PSO technique and the method works well even for lesser input data. Authors in [11] also combined well know techniques such as wavelet transforms, PSO and Adaptive-network-based fuzzy inference system (ANFIS) to develop a hybrid methodology. Wavelet transform is used initially for decomposing the non-linear price signal into independent, smooth signals which are later on recombined to form the individual signals. The prediction is done by the ANFIS module and the parameters of the network structure are optimized using PSO. In [12], a RBF-NN-GARCH model was proposed where the traditional RBF-NN model is extended by using GARCH specifications for modeling the variability of price signals. The model parameters are tuned using a maximum likelihood function which is further optimized using a derivative free genetic algorithm.

Many of the available papers including the ones discussed above have proposed hybrid methodologies and have some kind of network structure which needs to be tuned using optimization techniques

\* Corresponding author.

E-mail addresses: [anandnitin26@gmail.com](mailto:anandnitin26@gmail.com) (N.A. Shrivastava), [bijayaketan.panigrahi@gmail.com](mailto:bijayaketan.panigrahi@gmail.com) (B.K. Panigrahi).

### Nomenclature

|             |   |             |                                |
|-------------|---|-------------|--------------------------------|
| $A_m$       | Approximate series at resolution $m$      | $\tilde{N}$ | number of hidden nodes         |
| $D_m$       | detail series at resolution $m$           | $P_t$       | actual price value at hour $t$ |
| $E$         | number of ensembles                       | $\hat{P}_t$ | forecast price at hour $t$     |
| $g(x)$      | activation function                       | $T$         | target matrix                  |
| $H$         | hidden layer output matrix                | $W$         | input weight matrix            |
| $H^\dagger$ | Moore Pensrose generalized inverse of $H$ | $\beta$     | output weight matrix           |
| $L(\cdot)$  | wavelet function                          | $\varphi$   | father wavelet function        |
| $N$         | data samples                              | $\psi$      | mother wavelet function        |

requiring additional computational effort. Neural network based structures are most commonly employed and the efficiency of NN based methods is highly dependent on appropriate tuning of their adjustable parameters, e.g., the number of hidden layers, nodes, weights and transfer function etc. Most ANN based forecasting methods use gradient-based learning algorithms, such as the back propagation neural network (BPNN), and are plagued with problems such as over-tuning and long computation time. Recently, a novel learning algorithm for single-hidden-layer feedforward neural networks (SLFN) called extreme learning machine (ELM) has been proposed in [13,14]. In the proposed algorithm, the input weights and hidden biases are randomly chosen, and the output weights are determined analytically by using the Moore–Penrose (MP) generalized inverse. ELM surpasses the traditional gradient-based learning methods in terms of faster learning speed with a higher generalization and it also avoids many difficulties faced by gradient-based learning methods such as stopping criteria, learning rate, learning epochs, local minima, and the over-tuning problems [15–17].

Up to now, the ELM has been successfully applied in various areas such as classification [18], terrain reconstruction [19] and protein structure prediction [20] etc. The ELM technique has been applied for the case of electricity price forecasting in [21] where the advantages of ELM over traditional NN structures were highlighted and also the uncertainties related to prediction were quantified using the bootstrapping technique. In this paper the ELM, in conjunction with other techniques, is selected to forecast Day Ahead electricity prices for some of the existing markets. When making a decision, participants usually consider results from many types of techniques that help them to achieve their objective. Relying on a single technique can be very risky particularly in case of power markets where volatility is very high and millions of dollars are at stake. Combining several techniques together to form a hybrid tool has become a common practice to improve the forecasting accuracy where each models unique feature are combined to capture different patterns in the data. Theoretical as well as empirical findings suggest that hybrid methods can be effective and efficient in improving forecasts. Therefore ELM method has been coupled with the Wavelet techniques to develop a hybrid model termed as WELM (wavelet based ELM) and it is shown to give faster and more accurate results. The robustness of the model is further enhanced using the ensembling technique. The rest of the paper is organized as follows. The fundamental principles of the proposed method are introduced in Section 2. The various steps of model development are discussed in Section 2. Case studies and experimental results are presented in Sections 4 followed by the conclusion of the work in Section 5.

## 2. Methodology

In this section, we present the methodology employed for electricity price forecasting using a model consisting of ELM and data is preprocessed using a wavelet decomposition technique. Electric-

ity price signal is highly volatile, corrupted by occasional spikes and follows a weekly-daily cycle with each sample of one hour interval. Fig. 1 shows the price profile of the Ontario electricity market for the year 2004 which testifies the above statement. A price signal exhibits much richer structure than the load series and signal processing techniques like Fourier Transform (FT), Wavelet Transform (WT) are good tools to bring out the hidden patterns in the prices. WT is used for multi-scale analysis of the signal and decomposes the time series signal into the low-frequency sub-series (approximation part) and some high-frequency sub-series (detailed part) in the wavelet domain. These constitutive series have better statistical properties than original price series and hence better forecasting accuracy can be achieved by their appropriate utilization. In WT based models, first of all WT is applied to the price series, prediction is made in the wavelet domain using a ELM model and then inverse WT is applied to obtain the actual predicted value in time domain. A brief introduction to the ELM, wavelet technique, and other aspects relevant to the proposed methodology is presented in this section.

### 2.1. Extreme learning machines

The theoretical foundation behind the Extreme learning machine (ELM) architecture is explained in this section. ELM is an improved learning algorithm for the single feed-forward neural network architecture. ELM is different from the traditional neural network methodology in the sense that all the parameters of the feed-forward networks (input weights and hidden layer biases) are not required to be tuned in its case. The capability of SLFNs with randomly chosen input weights, hidden layer biases and a nonzero activation function to approximate any continuous functions on any input set has been demonstrated in [17]. The SLFN with randomly chosen input weights and the hidden layer biases can be considered as a linear system. For this linear system, the

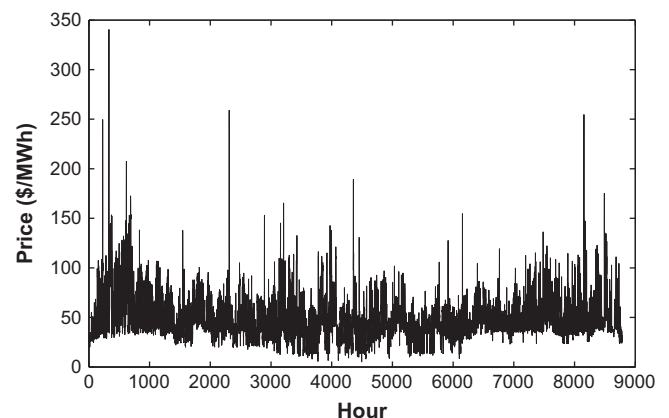


Fig. 1. Price profile of Ontario electricity market for the year 2004.

output weights which link the hidden layer to the output layer can be analytically determined through simple generalized inverse operation of the hidden layer output matrices. This simple approach makes ELM very efficient and many times faster than the traditional feedforward learning algorithms.

The structure of ELM consists of a single hidden-layer feedforward neural network (SLFN) in which the input weight matrix  $W$  is randomly chosen and the output weight matrix  $\beta$  is analytically determined. Suppose we are given a data set with  $N$  arbitrary distinct samples  $(x_i, t_i)$  where  $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in R^n$  and  $t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in R^m$ . The mathematical model of a standard SLFN with  $\tilde{N}$  hidden nodes and activation function  $g(x)$  for the given data can be formulated as [13]:

$$\sum_{i=1}^{\tilde{N}} \beta_i g_i(x_j) = \sum_{i=1}^{\tilde{N}} \beta_i g_i(w_i x_j + b_i) = y_j, \quad j = 1, \dots, N \quad (1)$$

where  $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$  denotes the weight vector which connects the input nodes to the  $i$ th hidden node and  $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$  is the weight vector which connects the output nodes with the  $i$ th hidden node. Also,  $b_i$  is the threshold of the  $i$ th hidden node. The inner product of  $w_i$  and  $x_j$  is denoted by the operation  $w_i \cdot x_j$  in (1). Let us consider that the standard SLFNs with  $\tilde{N}$  hidden nodes with activation function  $g(x)$  can approximate these  $N$  samples with zero error. In such a situation, we have

$$\sum_{j=1}^N \|y_j - t_j\| = 0 \quad (2)$$

where  $y$  denotes the actual output value of the SLFN. This indicates the existence of  $\beta_i$ ,  $w_i$ , and  $b_i$  such that

$$\sum_{i=1}^{\tilde{N}} \beta_i g_i(w_i x_j + b_i) = t_j, \quad j = 1, \dots, N \quad (3)$$

A succinct expression of the above  $N$  equations can be written as

$$H\beta = T \quad (4)$$

where  $H$  is the hidden layer output matrix.

$$H = \begin{bmatrix} h(x_1) \\ \dots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} h_1(x_1) & \dots & h_{\tilde{N}}(x_1) \\ \dots & \dots & \dots \\ h_1(x_N) & \dots & h_{\tilde{N}}(x_N) \end{bmatrix} \quad (5)$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \dots \\ \beta_{\tilde{N}}^T \end{bmatrix} \quad (6)$$

$$T = \begin{bmatrix} t_1^T \\ \dots \\ t_N^T \end{bmatrix} \quad (7)$$

As discussed earlier, the input weights and hidden biases are randomly generated and do not require any tuning as in the case of traditional SLFN methodology. The evaluation of the output weights linking the hidden layer to the output layer is equivalent to determining the least-square solution to the given linear system. The minimum norm least-square (LS) solution to the linear system (4) is

$$\hat{\beta} = H^\dagger T \quad (8)$$

The  $H^\dagger$  in the above equation is the Moore–Penrose (MP) generalized inverse of matrix  $H$  [22]. The minimum norm LS solution is unique and has the smallest norm among all the LS solutions. The MP inverse method based ELM is shown to obtain a good generalization performance with a radically increased learning speed. A general Algorithm for ELM can be stated as follows. For a given a training set, activation function  $g(x)$ , and hidden neuron number  $L$ :

Step 1: Assign random input weight  $w_i$  and bias  $b_i$ ,  $i = 1, \dots, L$ .

Step 2: Calculate the hidden layer output matrix  $H$ .

Step 3: Calculate the output weight  $\beta$ :  $\beta = H^\dagger T$

## 2.2. Wavelet transforms

Wavelet transform is an important tool used for analyzing the frequency components of signals and it has overcome the limitations of Fourier and Short-time Fourier transform. It has exceptional capacity to extract the relevant time–frequency information from non-periodic and transient signals. Wavelets functions disintegrate the data into different frequency components, and then study each component with a resolution matched to its scale [23]. The wavelet technique has made its way into a number of different fields such as image compression [24], fault classification [25], and hydrological prediction [26].

While implementing the wavelet transform technique, a limited number of positions and resolution levels (discrete wavelet transform) are considered. In this work, the wavelet transform is used to decompose the electricity price series into a set of better-behaved constitutive series. Predictions for the constitutive series are separately made and reverse wavelet transform is performed to generate actual predicted prices. The decomposition coefficients of the wavelet transform of the hourly price series determined by technique are given by [27]

$$\begin{aligned} p_{mn}^L &= 2^{-\left(\frac{n}{2}\right)} \sum_{t=0}^{T-1} p_t L\left(\frac{t-n \cdot 2^m}{2^m}\right) \\ &= 2^{-\left(\frac{n}{2}\right)} \sum_{t=0}^{T-1} p_t L_{mn}(t) \end{aligned} \quad (9)$$

where  $L(\cdot)$  is the selected wavelet function,  $p_t$  is the value of the price at hour  $t$ ,  $T$  is the length of the series,  $p_{mn}^L$  is the decomposition coefficient corresponding to resolution level  $m$  and position  $n$ . The number of coefficients at each resolution level is given by  $T/2^m$  provided the number of observations,  $T$ , is divisible by  $2^m$ . Faster calculations can be made by treating expression (9) as a convolution, and using the efficient Fast Fourier Transform [28].

An effective way to apply the wavelet functions is the multi-resolution technique based on using a father wavelet function and its complementary, a mother wavelet function. The father function provides for extracting the low frequency components, while the mother function allows extracting the high frequency components of the series. Orthogonal wavelet functions are preferably chosen because of their appropriate mathematical properties. Hence, the “approximation series”,  $A_m(m = 1, \dots, M)$ , and the “detail series”,  $D_m(m = 1, \dots, M)$ , are defined as

$$A_m = \sum_n \phi_{mn}^\phi \phi_{mn}(t); \quad m = 1, \dots, M \quad (10)$$

and

$$D_m = \sum_n \psi_{mn}^\psi \psi_{mn}(t); \quad m = 1, \dots, M \quad (11)$$

where  $\phi_{mn}(t)$  and  $\psi_{mn}(t)$  are the father and mother wavelet functions, and  $\phi_{mn}^\phi$  and  $\psi_{mn}^\psi$  are the coefficients obtained through (9). The father wavelet function  $\phi(t)$  is one solution of a functional equation

$$\varphi(t) = \sum_{k=-\infty}^{\infty} \alpha_k \sqrt{2} \varphi(2t - k) \quad (12)$$

and the mother wavelet function  $\psi(t)$  is given by

$$\psi(t) = \sum_{k=-\infty}^{\infty} (-1)^k \alpha_{-k+1} \sqrt{2} \varphi(2t - k) \quad (13)$$

Further details of the above functions can be obtained from [29]. The expression of the original price series  $p_t(t = 1, \dots, T)$  can now be reconstructed by

$$p_t = D_1 + \dots + D_M + A_M \quad (14)$$

which is the denominated multi-resolution decomposition of the price series.

Daubechies wavelets are most appropriate for treating a non-stationary series and have been considered in this work also.

### 2.3. Ensemble based decision making

Ensemble methods are sometimes employed in statistics and machine learning where multiple models are developed to obtain better predictive performance than could be obtained from individual models. They have been shown to produce better results compared to those of single-expert systems for a large number of applications [30]. Ensemble methods focus on modifying the training process with the intention that the resulting model will give different predictions. Methods based on neural network techniques bring about such modifications in the training process by means of different topologies, parameters, initial weights or by training only a portion of the training set [31]. Bagging and boosting are some of the popular ensembling methods. Opitz and Maclin [31] investigated the creation of a simple neural network ensemble where each network used the full training set and differed only in its random initial weight setting. The methodology was found to be very effective and at par with the Bagging method. Similar findings were presented by Ali and Pazzani [32] using randomized decision tree algorithms. An important aspect of the ensemble methods is the selection of appropriate combinational strategy for the integration of individual models. The commonly used strategies are arithmetic averaging, weighted averaging, voting, etc. The advantage of arithmetic averaging technique is its simplicity and has also been shown to give better results [33]. Therefore we have adopted this strategy in our work also. Extreme Learning Machines have been shown to have many advantages compared to other learning techniques but its application in important decision making problems can be limited due to its fluctuating output caused by random initializations of weights and biases. Every time the model is run, its output will be slightly different from the previous one. However, it should be mentioned that the fluctuations observed in ELM output is quite lower compared to that of ANN because of higher generalization capacity. In order to make the predictions of the proposed model more robust and consistent, the simple network ensemble technique is applied here. An ensemble of different WELM models based on different weight initializations is created and the arithmetic average of the ensemble member's prediction is used as finally predicted output.

$$\bar{y}_i = \frac{1}{E} \sum_{j=1}^E y_i^j \quad (15)$$

where  $E$  refers to the number of Ensembles taken into consideration,  $y_i^j$  is the output of the  $i$ th hour from the  $j$ th ensemble and  $\bar{y}_i$  is the final aggregate output.

### 2.4. Forecast error measures

To assess the prediction performance of the models, different statistical measures can be utilized [34]. The most widely used measures are those based on absolute errors, i.e. absolute values of differences between the actual price,  $P$ , and predicted price,  $\hat{P}$  for a given hour  $t$ . The Mean Absolute Error (MAE) is a typical example. For hourly prices it is given by:

$$MAE = \frac{1}{N} \sum_{t=1}^N |P_t - \hat{P}_t| \quad (16)$$

where  $N$  is the number of hour for forecasting. MAE is quite a sensible index when evaluating model performances for a single time series. However its usage for comparing across different datasets for different forecasting horizons could be quite misleading as the index is scale dependent [35,36]. The relative or percentage difference, which is a scale independent, is sometimes more informative than the absolute errors particularly when comparing results for two distinct data sets. In such cases the Mean Absolute Percentage Error (MAPE) is a better choice. For hourly prices the daily MAPE takes the form:

$$MAPE = \frac{1}{N} \sum_{t=1}^N \frac{|P_t - \hat{P}_t|}{P_t} \quad (17)$$

MAPE is known to have the disadvantage that it may be infinite or undefined for  $P_t = 0$ . Excessively large values can also pose a problem by reducing the index to a very small value [37]. They also tend to put a heavier penalty on positive errors than on negative errors. Some alternative performance measures have also been suggested in the literature for price forecasting due to the unique characteristics of electricity prices. For instance, the absolute error  $|P_t - \hat{P}_t|$  can be normalized by the average price attained during the day [38]. The resulting measure, commonly known as the Mean Daily Error, is given by:

$$MDE = \frac{1}{24} \sum_{t=1}^{24} \frac{|P_t - \hat{P}_t|}{\bar{P}_t} \quad (18)$$

where

$$\bar{P}_t = \frac{1}{24} \sum_{t=1}^{24} P_t \quad (19)$$

In this case, MDE compared to MAPE gives more emphasis to errors in the high-price range. Analogous to MDE, the Mean Weekly Error can be computed for  $N = 168$ .

The median price can also be used for normalization. Since median is more robust to outliers (or spikes), the resulting measures – Median Daily Error (MeDE) and Median Weekly Error (MeWE) – in some cases exhibit better performance. Apart from absolute value-type norms, square-type norms like Daily Root Mean Square Error (DRMSE) and the Weekly Root Mean Square Error (WRMSE) are also very popular.

$$WRMSE = \sqrt{\frac{1}{168} \sum_{t=1}^{168} |P_t - \hat{P}_t|^2} \quad (20)$$

RMSE has been a popular measure historically because it is on the same scale as the data and is quite relevant in statistical modeling. However, researchers have warned against its use as it is more sensitive to outliers than MAE and others [39]. In this work, the above mentioned five type of error indices have been taken into consideration.



### 3. Model development

The proposed methodology is tested for the short term price forecast of the Ontario, PJM, New York and Italian Electricity markets and compared with the recently published works. Data of Ontario electricity market has been obtained from their website [40]. To illustrate the behavior of the proposed technique and for the sake of fair comparison with other works, results comprising six weeks corresponding to the three prominent seasons of year 2004 are presented. The first test period is from April 26 to May 9, 2004 which represents spring's low demand period. The second test period from July 26 to August 8, 2004 corresponds to summer peak-demand period. Winter high-demand period, from December 13–26, 2004, has been selected as third test period. The same model has been tested for the PJM Market data to verify its robustness and efficacy. Four test weeks corresponding to different seasons of the year 2004 have been taken into consideration similar to other works for a fair comparison. The test dates are February 23–February 29, May 17–May 23, August 23–August 29 and November 22–November 28 corresponding to Winter, Spring, Summer and Autumn seasons. Data of PJM market has been obtained from website [41]. Similarly, the model has been tested for a selected period of the Italian and New York Electricity Market data series as considered in other works and the data has been obtained from the market's official website [42,43].

The various components of the model are depicted graphically in the flowchart presented in Fig. 2 for better understanding and explained further in this section. After collecting the relevant data, the first step is to select a suitable training duration for the model. The training data enables the model to develop a generalized network structure using which it can estimate the unknown data accurately. In order to create the training and testing data sets, it is essential to decide which input variables should be taken into consideration. Electricity prices are influenced by a number of factors such as historical load, system load rate, imports/exports, weather, fuel prices, generation outages, bidding strategies, demand elasticity, and holidays. However, inclusion of all these factors in a prediction model can complicate the process as a separate module of relevant feature selection is necessary to determine and include the most relevant input features. Also the redundancy in the input features can degrade the model performance rather than improving it [44]. For example, in [45] it was observed that inclusion of load demand slightly degrades the performance of the model for Spain market. In order to keep the methodology simple and for comparison purposes, we have considered historical prices as the only input variable as considered in most of the referred publications. A detailed feature selection technique considering other factors would be considered in our future work. In this work a dynamic feature selection strategy has been employed to capture the movement of electricity prices in different scenarios. Selection of input features is based on their high correlation with the price at the current hour. The dynamics of electricity prices is different for all the markets thereby leading to different kinds of trends and seasonality. These differences can be captured by the autocorrelation factors. Keeping in view this fact, a fixed set of input features will not be applicable for all the markets and corresponding seasons. After performing the correlation analysis for different lag periods, those periods which have high correlation values are separated out and a fraction of them are finally selected as the input features. Next step is to decompose the price times series into different component series. In this work, Daubechies wavelets (dB4) have been considered and different case studies are performed. The Wavelet transform decomposes the original prices series into a approximation series and a detailed series. Approximation series correspond to low frequency bands and rep-

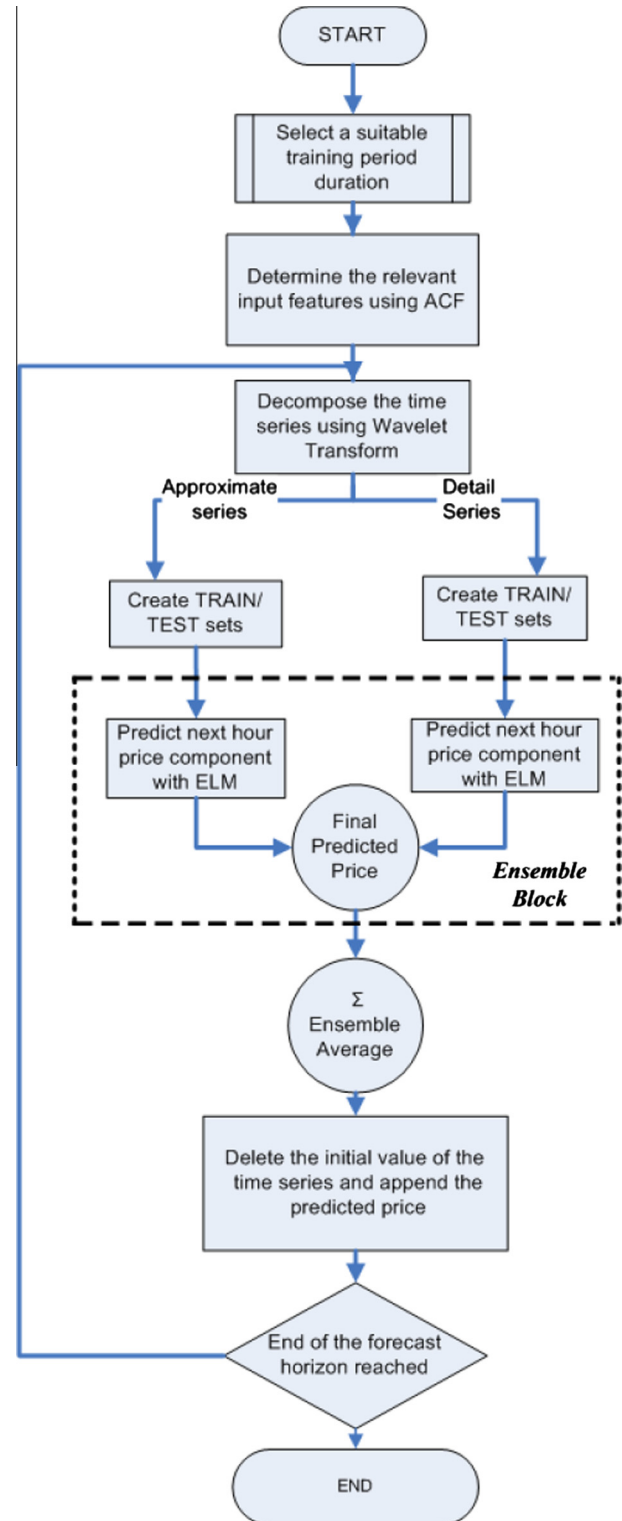


Fig. 2. Structure of the proposed WELM model.

resent the trend of the price signal whereas the detailed series correspond to high frequency bands and contain the local short-period discrepancies in the price signal due to bidding strategies adopted by the participants as well as other known and unknown factors.

The approximate and the detailed series is now processed to form training and testing data sets. Using the selected input features, training and testing data sets are created and they are

further processed to get normalized data sets using the normalizing technique. The normalized data sets corresponding to both the approximate and the detailed series are separately given to ELM model for prediction. The resultant predicted values corresponding to the approximate and detailed sets are first unnormalized and then added to give the final predicted value of the future prices. The steps confined in the box labeled as *Ensemble block* is repeated for different ensembles and the aggregate of all the model outcomes is deemed to be the final predicted output for practical considerations. Ensemble based decision making requires additional computational time owing to the result of more than one model being taken into consideration. The number of models that should be taken into consideration depends on the sensitivity of the problem and time considerations. Based on their research, Hanson and Salamon [46] recommended that the number of ensemble members required to reduce the test set error adequately could be as few as ten. Alternate suggestions have also been made by researchers, however we have considered the ensemble member number to be ten for computational efficiency and time considerations. The forecasting horizon of the proposed methodology is 24 h therefore we have employed the recursive forecasting methodology in this work. Once the price of an hour is forecasted, it is given as an input to the model to predict for the next consecutive hour while the initial value of the historical price series is eliminated. The model acts like a moving window and the process is repeated till the prices for all the 24 h of the day are predicted.

#### 4. Case studies and results

With the above developed model in hand, a number of case studies are performed and performance of the model under different conditions is investigated. The major advantage of ELM over traditional neural network models is that the parameters do not require any tuning in the former case thus eliminating the need for repetitive tuning and cross validation to come up with the best model. The learning capacity of ELM is very fast and it gives excellent performance when the training data set is shown even once to the network. For each forecast day, we considered a historical data of seven weeks as taken in [47], where 49 days were considered for training, one day for validation and one testing day. However, in

this work, the requirement of a separate validation set is eliminated because the ELM model does not require any iterative tuning based on validation set. The trained model was directly validated using the testing data. The input features for the forecasting model are determined using the autocorrelation analysis for each forecasting period as explained in Section 3. The autocorrelation function curves corresponding to the Ontario and PJM Day-ahead market electricity prices for summer and winter seasons are depicted in Fig. 3. From the figure, we can observe that different relation between the past and present prices exist in both the markets. The daily and weekly seasonality is more prominent in the PJM Day-Ahead market and comparatively less significant seasonality is observed for the case of the Ontario market. Such variations are not only seen across different markets, but they can also be observed across different seasons of the year corresponding to the same market. The selected features using this method are depicted in Table 1. The number of input features required for appropriate generalization from the training data is also case dependent and it is shown in the Table.

The price series was decomposed till one level into approximate and detailed series and corresponding training and testing data sets were created. Corresponding data sets were fed to the ELM model and prices were predicted for a duration of week. In the first case study, a basic ELM network was tuned with a simple data without any decomposition and with different activation function to determine the most suitable activation function for the given data. The performance with sigmoid, sinusoidal and hardlim function was evaluated for the first test week of the spring season. Fig. 4 shows the performance comparison and it was observed that hardlim function gives a poor performance in comparison to sigmoid and sin. The performance of sigmoid function was still on the better side and is used as a default activation function in all the further analysis. Now a detailed experiment is performed with decomposed price series for all the test weeks.

A unique feature of ELM is that their generalization performance is independent of the number of hidden nodes if the number of hidden nodes is considerably large [13]. This fact is also verified by testing the model against different number of hidden nodes varying from 5 to 100 and also different number of training periods varying from 7 to 70. The testing period is one week i.e., 168 h of a week. The performance variation is depicted in three

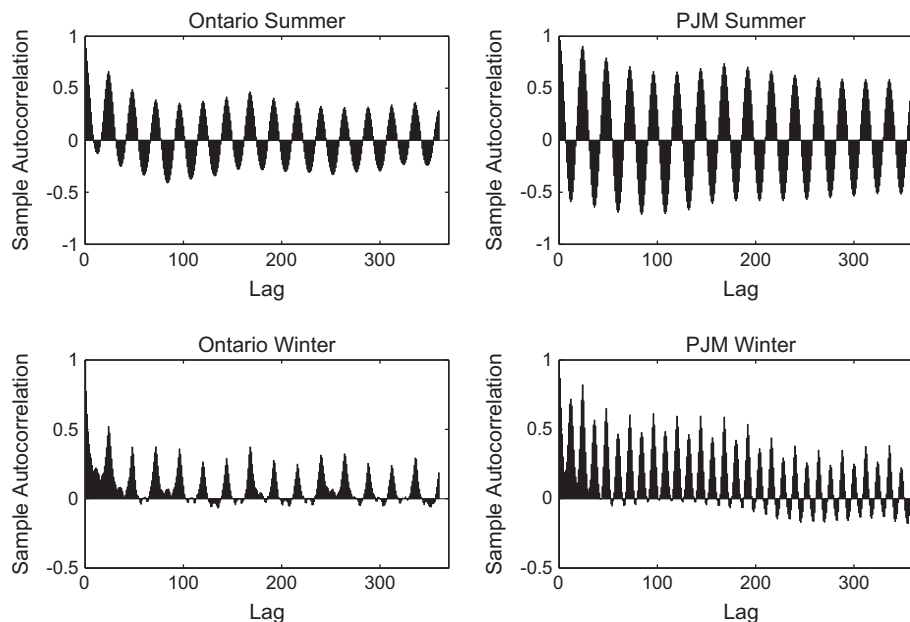
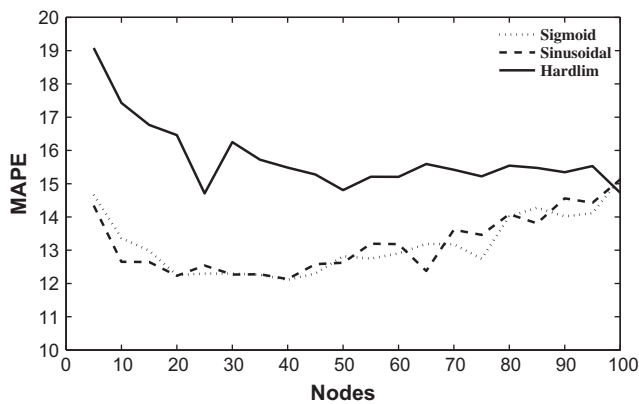
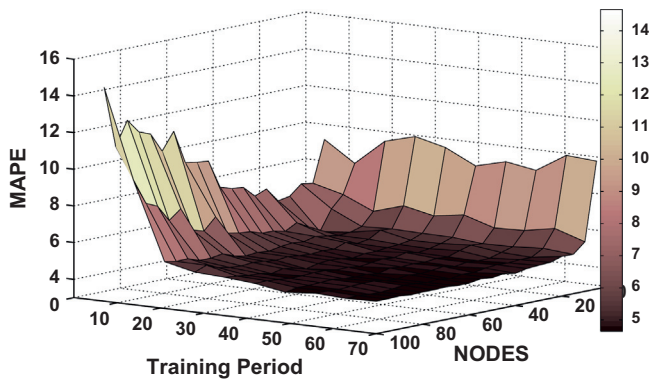


Fig. 3. ACF plots for Ontario and PJM day-ahead market.

**Table 1**  
Selected features using ACF.

| Test week            | Input features, $P_{t-x}$   |
|----------------------|---|
| <i>ONTARIO</i>       |   |
| 26 April–2 May       | $x = 1, 2, 3, 22, 23, 24, 25, 26, 47, 48, 49, 72, 96, 143, 144, 145, 167, 168, 169, 192$  |
| 3–9 May              | $x = 1, 2, 3, 22, 23, 24, 25, 26, 47, 48, 49, 72, 143, 144, 145, 167, 168, 169, 192, 336$ |
| 26 July–1 August     | $x = 1, 2, 3, 4, 21, 22, 23, 24, 25, 26, 27, 47, 48, 49, 144, 167, 168, 169, 192, 193$    |
| 2–8 August           | $x = 1, 2, 3, 4, 21, 22, 23, 24, 25, 26, 27, 47, 48, 84, 85, 144, 167, 168, 169, 192$     |
| 13–19 December       | $x = 1, 2, 23, 24, 25, 47, 48, 49, 71, 72, 73, 95, 96, 97, 144, 168, 169, 192, 216, 336$  |
| 20–26 December       | $x = 1, 2, 3, 23, 24, 25, 48, 71, 72, 73, 95, 96, 97, 120, 144, 145, 168, 169, 336, 337$  |
| <i>PJM day-ahead</i> |   |
| Winter               | $x = 1, 2, 11, 12, 12, 23, 24, 25, 48, 72$  |
| Spring               | $x = 1, 2, 23, 24, 25, 48, 49, 72, 96$  |
| Summer               | $x = 1, 2, 22, 23, 24, 25, 26, 47, 48, 168$   |
| Autumn               | $x = 1, 23, 24, 25, 168$  |

**Fig. 4.** Performance comparison with different activation functions.**Fig. 5.** Performance comparison with different nodes and training data.

dimensional Fig. 5. The MAPE values are comparatively higher for very low training periods and less number of hidden nodes which

can be seen in the form of peaks at the ends. The major portion of the Fig. 5 is flat which shows that the model performs equally well for different combinations of training periods (data) and hidden nodes if they have high values.

Once the activation function, the number of hidden nodes and the training period duration is fixed for the ELM model as explained above, training and testing data sets corresponding to different markets are given to the model and next hour prices are predicted. Once the entire forecasted price profile is obtained from the model, different forecast error measures given in Section 2.4 are evaluated. The MAPE values for Ontario Market data are compared with other techniques such as ARIMA, transfer function (TF) and Dynamic Regression (DR) [48] and modified relief and hybrid neural network technique (MR + HNN) [47] and the numerical results are enlisted in Table 2. The forecasting accuracy of the WELM model is seen to be better than the results of previous works and it is further improved with the ensembling technique at a slightly higher computational cost. A performance improvement of 20.1%, 21.1%, 59.9%, 41.7%, 26.8% respectively is obtained for the first five test weeks in comparison to the best results quoted from previous works. The performance was lower only for the last week by a margin of 2.5%. Ensembling technique is shown to further enhance the prediction performance over a single model run and its results can be considered to be more reliable. From the MAPE values enlisted in the table, we can observe a cyclical pattern in the MAPE index obtained for different seasons. The lowest MAPE index is obtained for the summer peak demand period and highest for the winter peak demand period and intermediate MAPE indices are obtained for the spring low demand period for the proposed method. Higher error levels can be generally attributed to higher uncertainty in the data particularly in the winter season which is more severe compared to other seasons in this region. However, the ensembling technique used in this work, where models with different network weights are used, ensures that the results are more reliable compared to single models. The general range of MAPE's obtained in this region is around 3–9% for the considered periods of study. The forecasted price profile of the proposed

**Table 2**  
Weekly MAPE for HOEP forecast in the Ontario electricity market.

| Test week        | Arima | TF   | DR   | MR + HNN | WELM | WELM with ensembles |
|------------------|-------|------|------|----------|------|---------------------|
| 26 April–2 May   | 15.9  | 15.6 | 15.9 | 9.46     | 9.22 | 7.56                |
| 3–9 May          | 18.6  | 18   | 18.1 | 9.72     | 8.88 | 7.67                |
| 26 July–1 August | 13.6  | 12.3 | 13   | 9        | 4.68 | 3.61                |
| 2–8 August       | 21.5  | 18.3 | 19   | 9.04     | 6.56 | 5.27                |
| 13–19 December   | 15.4  | 14.8 | 14.7 | 9.36     | 7.63 | 6.85                |
| 20–26 December   | 20.8  | 17.5 | 18.5 | 8.8      | 10.3 | 9.02                |
| Average          | 17.6  | 16.1 | 16.5 | 9.23     | 7.88 | 6.66                |

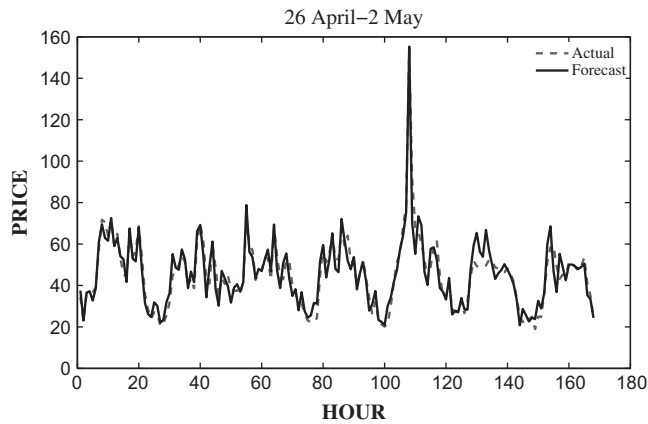


Fig. 6. Actual and forecast prices for 26 April–2 May: Ontario.

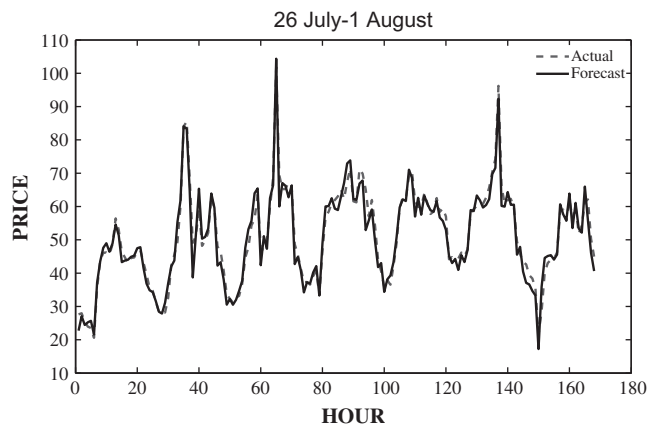


Fig. 7. Actual and forecast prices for 26 July–1 August: Ontario.

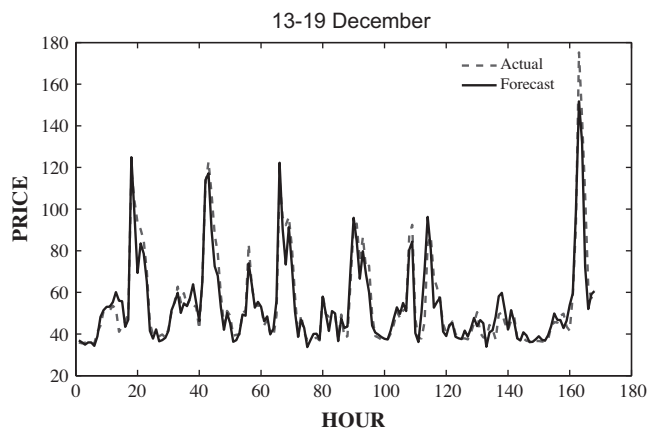


Fig. 8. Actual and forecast prices for 13–19 December: Ontario.

Table 3

Weekly MAPE for PJM market.

| Test week | AWNN  | WELM  | WELM + ensembles |
|-----------|-------|-------|------------------|
| Winter    | 6.362 | 6.108 | 6.010            |
| Spring    | 5.976 | 5.128 | 4.937            |
| Summer    | 5.954 | 5.872 | 5.843            |
| Autumn    | 6.648 | 6.428 | 6.056            |
| Average   | 6.235 | 5.884 | 5.712            |

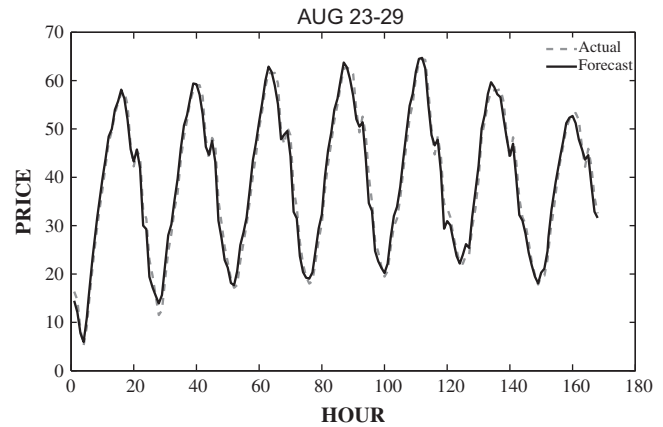


Fig. 9. Actual and forecast prices for 23–29 August: PJM.

market for one test weeks is graphically shown in Fig. 9. The range of errors for PJM market for the four seasons is around 4–6% which is comparatively more compact compared to that of Ontario Market. This discrepancy is quite reasonable due to high volatility of the Ontario Market compared to PJM-Day Ahead Market. The model was also tested for the price data corresponding to months May-2004 to March-2005 of the Italian Electricity Market and the results are presented in Table 4. Since the Italian Market data has certain instances of prices being zero, therefore an alternate MAPE definition is used in previous works to assess the forecasting accuracy. In the alternate MAPE definition, the mean price of the considered period is used in the denominator in place of the actual price and this alternate definition is same as the MDE index which we have considered in our work. The performance of the model is compared with a dynamic price forecast method [49] and modified relief and hybrid neural network technique (MR + HNN) [47] and it is seen to be quite comparable with that of the previous works. In six out of the eleven test months, the performance of WELM model with ensembling techniques is found to be better than the other works and the overall average MDE index for all the months is lower (8.57) compared to 13.57 and 9.37 obtained through other models. The range of error lies between 6% and 13% which is quite high compared to the error ranges for other markets. The performance of the model for data corresponding to months February-2004 to December-2004 of the New York Market are also presented in Table 5 and compared with the dynamic price forecast methodology [49]. Significantly lower error indices (2.76–3.98%) are obtained for all the months compared to the dynamic forecast technique. Amongst all the markets, the Italian market is observed to have higher ranges of error. This observation can be attributed to the fact that Italian market is the youngest amongst all as it started its operation on 1st April 2004. Since the market was not fully consolidated during the studied period, therefore its predictability is quite low. On the other hand, PJM, New York and Ontario Markets started operating in 1997, 1999 and 2002 respectively. Since these markets are mature during the studied period, therefore their

method for the Ontario Market for all the test weeks is graphically shown in Figs. 6–8.

The performance of the proposed model for PJM market is also evaluated and the results are depicted in Table 3. The results obtained with the proposed methodology is compared with a adaptive wavelet neural network (AWNN) based technique presented in [45] and better results are obtained for all the test weeks with WELM as well as WELM with ensembling technique. The forecasted price profile of the proposed method for PJM Day-Ahead



**Table 4**  
Monthly MDE for Italian electricity market.

| Test month | Dynamic price forecast | MR + HNN | WELM + ensembles |
|------------|------------------------|----------|------------------|
| 1          | 36.12                  | 12.12    | 6.54             |
| 2          | 18.46                  | 11.23    | 9.31             |
| 3          | 12.76                  | 10.14    | 10.28            |
| 4          | 12.42                  | 10.64    | 6.97             |
| 5          | 13.27                  | 9.75     | 10.25            |
| 6          | 9.36                   | 10.56    | 6.11             |
| 7          | 11.63                  | 9.54     | 8.8              |
| 8          | 11.96                  | 9.73     | 12.92            |
| 9          | 12.81                  | 9.25     | 8.36             |
| 10         | 4.97                   | 4.81     | 8.16             |
| 11         | 5.49                   | 5.35     | 6.55             |
| Average    | 13.57                  | 9.37     | 8.57             |

**Table 5**  
Monthly MDE for New York electricity market.

| Test month | Dynamic Price forecast | WELM | WELM + ensembles |
|------------|------------------------|------|------------------|
| 1          | 11.54                  | 3.57 | 3.49             |
| 2          | 6.48                   | 3.69 | 3.61             |
| 3          | 7.32                   | 3.28 | 3.20             |
| 4          | 8.91                   | 3.56 | 3.47             |
| 5          | 10.59                  | 3.57 | 3.45             |
| 6          | 10.7                   | 2.75 | 2.76             |
| 7          | 7.8                    | 2.94 | 2.87             |
| 8          | 6.05                   | 2.63 | 2.58             |
| 9          | 9.62                   | 3.70 | 3.42             |
| 10         | 8.8                    | 4.00 | 3.98             |
| 11         | 9.03                   | 3.90 | 3.79             |

**Table 6**  
Performance indices of WELM for Ontario market.

| Test week        | MAE  | MAPE | MDE   | MeDE  | DRMSE |
|------------------|------|------|-------|-------|-------|
| 26 April–2 May   | 3.95 | 9.22 | 8.7   | 8.75  | 5.32  |
| 3–9 May          | 3.79 | 8.88 | 8.16  | 8.21  | 5.03  |
| 26 July–1 August | 2.29 | 4.68 | 4.47  | 4.63  | 3.04  |
| 2–8 August       | 2.35 | 6.56 | 5.04  | 5.29  | 3.1   |
| 13–19 December   | 4.49 | 7.63 | 8.18  | 9.42  | 6.86  |
| 20–26 December   | 5.26 | 10.3 | 10.19 | 12.78 | 7.18  |

predictability is higher and consequent error index ranges are comparatively lower.

All the performance indices discussed in Section 2.4 are listed in Table 6 corresponding to Ontario Market with WELM model and without using the ensemble technique. Most of the research works conducted so far in the area of forecasting have considered MAPE as the standard index for evaluating as well as comparing the performance of different models. It has been specifically used to evaluate the performance of load forecasting. But, it has been observed that, MAPE may not be a reasonable criterion for the problem of price forecasting [38,50]. The justification given here is that the actual value of price may be very small or even zero sometimes and this may shoot up the MAPE index to a very high value. This problem does not occur in load forecasting as the loads generally have a high value. In the alternative definition of MAPE proposed in [38], the error is normalized by the Mean of the actual price profile instead of the actual price in that hour. Although this alternative definition prevents the error index from shooting up to very high values, but still there are issues related to its application for the price forecasting problem. The price series is not just composed of very low and sometimes zero values but it is also notable for spikes and outliers. Both, the Mean and the Median are measures of the central tendency of the sample distribution. The Mean gives equal weightage to all the sample values in determining the center

and therefore it is prone to get misled by extreme sample values (outliers). On the other hand, Median discounts the effect of extreme sample values and therefore is considered to be a robust statistic for skewed or outlier prone distributions like Electricity price series. If the errors are normalized by individual sample values, the final index will be biased by very large or very small sample values. On the other hand median would remain constant for a given sample and would not be biased by extreme values. Therefore it would be an appropriate, stable normalizing basis for the errors and can be safely used to compare performances of different data sets with different forecasting models. We also observed some peculiarities related to error indices in our experiments which further strengthen the need for alternate error indices and suitability of MeDE over MAPE and other indices used contemporarily. For the same magnitudes of actual and forecasted points, there are significant and sometimes peculiar differences between the performance indices. When comparing the forecasting results across different data sets corresponding to different markets, normalized error indices such as MAPE, MDE and MeDE are generally preferred. There is a positive correlation between these indices in four out of six test week i.e., if the value of MAPE decreases from one test week to another test week, then the values of MDE and MeDE also decrease in certain proportion. However a peculiar behavior is seen for the first test week (26 April–2 May) and the fifth test week (13–19 December). Here it is seen that the MAPE index for fifth week is lower (7.63) in comparison to the first week (9.22). Contrary to this, the MeDE index for the fifth week is higher (9.42) in comparison to the first week (8.75). This happens because the MAPE index is normalized by the actual price in that hour and it can be observed from the Figs. (6 and 8) that the individual hour prices for the fifth week are higher when compared to that of the first week; thereby leading to the lower MAPE. The peaks in the fifth week are more prominent which further bring down the MAPE index. The MeDE index, on the other hand, is normalized by the median which remains constant for the entire price profile. Therefore the MeDE index is less biased by the individual hour prices and unexpected peaks. The MDE index also behaves similar to the MeDE index although its normalization factor is the Mean price which also remains constant for a given price profile. The difference occurs when the Mean and the Median values are quite disparate depending on the distribution of the prices leading to different index magnitudes. Median is considered to be a better measure of central tendency than Mean in case of outlier prone distributions like electricity prices and therefore MeDE index can be considered to be a more reliable representation of the forecasting accuracy.

The average computational time of the proposed approach for a 24 h forecasting horizon with one level decomposition is around 27.73 s using MATLAB on a PC with 16 GB of RAM and a 2.93-GHz based processor. When the ensembling technique is applied, the computational time for 10 ensemble members turns out to be 59.55 seconds on average. Hence the remarkable features of the proposed approach are lower modeling complexity, better forecasting accuracy and low computational time which makes it very suitable for short term price forecasting requirements.

## 5. Conclusions

In the present scenario where competition and globalization is overtaking all fields of life, electricity price forecasting plays a more and more prominent role in a decision support system of power market participants as well as consumers. Developing more accurate and timely price forecasting methods has become an important research topic. In this paper, we apply a relatively novel neural network technique, extreme learning machine (ELM) in collaboration with a powerful multi-scale resolution technique known as Wavelets to electricity price forecasting. ELM is claimed

to have a higher generalization performance than the traditional gradient-based learning algorithms and it also avoids many difficulties faced by gradient-based learning methods such as stopping criteria, learning rate, learning epochs, local minima, and the over-tuning problem. This fact is verified in this work by using ELM for price forecasting of Ontario, PJM, Italy and New York Electricity market. The effectiveness of ELM is further enhanced by coupling it with wavelet and ensembling techniques and different case studies are performed. Extensive experiments are performed to determine appropriate number of hidden nodes and the training data requirement. Different performance indices are studied and it is inferred that instead of contemporary measurement indices like MAPE and MAE, MeDE is more suitable for the outlier prone electricity price series. Our experiments have also successfully demonstrated that the hybrid wavelet-ELM model can produce smaller predicting errors than the existing techniques.

## References

- [1] Nogales FJ, Conejo AJ. Electricity price forecasting through transfer function models. *J Oper Res* 2006;57(4):350–6.
- [2] Contreras J, Espinola R, Nogales FJ, Conejo AJ. ARIMA models to predict nextday electricity prices. *IEEE Trans Power Syst* 2003;18(3):1014–20.
- [3] Garcia-Martos C, Rodriguez J, Sanchez MJ. Mixed models for short-run forecasting of electricity prices: application for the spanish market. *IEEE Trans Power Syst* 2007;22(2):544–52.
- [4] Garcia RC, Contreras J, van Akkeren M, Garcia JBC. A GARCH forecasting model to predict day-ahead electricity prices. *IEEE Trans Power Syst* 2005;20(2):867–74.
- [5] Rodriguez CP, Anders GJ. Energy price forecasting in the Ontario competitive power system market. *IEEE Trans Power Syst* 2004;19(1):366–74.
- [6] Anbazhagan S, Kumarappan N. Day-ahead deregulated electricity market price classification using neural network input featured by DCT. *Electr Power Energy Syst* 2012;37(1):103–9.
- [7] Zhao JH, Dong ZY, Xu Z, Wong KP. A statistical approach for interval forecasting of the electricity price. *IEEE Trans Power Syst* 2008;23(2):267–76.
- [8] Unsihuay-Vila C, Zambroni de Souza AC, Marangon-Lima JW, Balestrassi PP. Electricity demand and spot price forecasting using evolutionary computation combined with chaotic nonlinear dynamic model. *Electr Power Energy Syst* 2010;32(2):108–16.
- [9] Lei Mingli, Feng Zuren. A proposed grey model for short-term electricity price forecasting in competitive power markets. *Electr Power Energy Syst* 2012;43(2):531–8.
- [10] Shafie-khah M, Parsa Moghaddam M, Sheikh-El-Eslami MK. Price forecasting of day-ahead electricity markets using a hybrid forecast method. *Energy Convers Manag* 2011;52(5):2165–9.
- [11] Catalao JPS, Pousinho HM I, Mendes VMF. Hybrid wavelet-PSO-ANFIS approach for short-term electricity prices forecasting. *IEEE Trans Power Syst* 2011;26(1):137–44.
- [12] dos Santos Coelho Leandro, Santos André AP. A RBF neural network model with GARCH errors: application to electricity price forecasting. *Electr Power Syst Res* 2011;81(1):74–83.
- [13] Huang Guang-Bin, Zhu Qin-Yu, Siew Chee-Kheong. Extreme learning machine: theory and applications. *Neurocomputing* 2006;70(1–3):489–501.
- [14] Zhu Qin-Yu, Qin AK, Suganthan PN, Huang Guang-Bin. Evolutionary extreme learning machine. *Pattern Recogn* 2005;38(10):1759–63.
- [15] Huang Guang-Bin. Learning capability and storage capacity of two-hidden-layer feedforward networks. *IEEE Trans Neural Netw* 2003;14(2):274–81.
- [16] Huang Guang-Bin, Babri HA. Upper bounds on the number of hidden neurons in feedforward networks with arbitrary bounded nonlinear activation functions. *IEEE Trans Neural Netw* 1998;9(1):224–9.
- [17] Huang Guang-Bin, Chen Lei, Siew Chee-Kheong. Universal approximation using incremental constructive feedforward networks with random hidden nodes. *IEEE Trans Neural Netw* 2006;17(4):879–92.
- [18] Zhang Runxuan, Huang G-B, Sundararajan N, Saratchandran P. Multicategory classification using an extreme learning machine for microarray gene expression cancer diagnosis. *IEEE/ACM Trans Comput Biol Bioinform* 2007;4(3):485–95.
- [19] Yeu C-WT, Meng-Hiot Lim, Huang Guang-Bin, Agarwal A, Ong Yew-Soon. A new machine learning paradigm for terrain reconstruction. *IEEE Trans Geosci Remote Sens* 2006;3(3):382–6.
- [20] Wang Guoren, Zhao Yi, Wang Di. A protein secondary structure prediction framework based on the Extreme Learning Machine. *Neurocomputat* 2008;72(1–3):262–8.
- [21] Xia Chen, Zhao Yang Dong, Ke Meng, Yan Xu, Kit Po Wong, Ngan, HW. Electricity price forecasting with extreme learning machine and bootstrapping. *IEEE Trans Power Syst* 2012;27(4):2055–62.
- [22] Rao CR, Mitra SK. Generalized inverse of matrices and its applications. New York: Wiley; 1971.
- [23] Graps Amara. An introduction to wavelets. *IEEE Comput Sci Eng* 1995;2:50–61.
- [24] Boix Macarena, Canto Begona. Wavelet transform application to the compression of images. *Math Comput Modell* 2010;52(7–8):265–1270.
- [25] Jayabharata Reddy M, Mohanta DK. A wavelet-fuzzy combined approach for classification and location of transmission line faults. *Electr Power Energy Syst* 2007;29(9):669–78.
- [26] Maheswaran R, Khosa Rakesh. Comparative study of different wavelets for hydrologic forecasting. *Comput Geosci* 2012;46:284–95.
- [27] Conejo AJ, Plazas MA, Espinola R, Molina AB. Day-ahead electricity price forecasting using the wavelet transform and arima models. *IEEE Trans Power Syst* 2005;20(2):1035–42.
- [28] Nievergelt Y. Wavelets made easy. Cambridge, MA: Birkhäuser; 1999.
- [29] Lee Akio Yamamoto Daniel TL. Wavelet analysis: theory and applications. Hewlett-Packard Journal 1994.
- [30] Polikar R. Ensemble based systems in decision making. *IEEE Circuits Syst Magaz* 2006;6(3):21–45.
- [31] Opitz David, Maclin Richard. Popular ensemble methods: an empirical study. *J Artif Intell Res* 1999;11:169–98.
- [32] Ali Kamal M, Pazzani Michael J. Error reduction through learning multiple descriptions. *Mach Learn* 1996;24:173–202.
- [33] Luo Dingsheng, Chen Ke. A comparative study of statistical ensemble methods on mismatch conditions. *Proc Int Joint Conf Neural Netw* 2002;1:59–64.
- [34] Weron R. Modeling and forecasting electricity loads and prices: a statistical approach. New York: Wiley; 2006.
- [35] Chatfield Chris. Apples, oranges and mean square error. *Int J Forecast* 1988;4(4):515–8.
- [36] Scott Armstrong J, Collopy Fred. Error measures for generalizing about forecasting methods: empirical comparisons. *Int J Forecast* 1992;8(1):69–80.
- [37] Hyndman Rob J, Koehler Anne B. Another look at measures of forecast accuracy. *Int J Forecast* 2006;22(4):679–88.
- [38] Shahidepour M, Yamin H, Li Z. Market operations in electric power systems: forecasting, scheduling and risk management. New York: John Wiley; 2002.
- [39] Scott Armstrong J. Evaluating forecasting methods, Principles of forecasting: a handbook for researchers and practitioners. Massachusetts: Norwell publishers; 2001 [Chapter 14].
- [40] Website of Ontario electricity market <<http://www.ieso.ca>>.
- [41] PJM Website <<http://www.pjm.com>>.
- [42] Italy Electricity Market Website <<http://www.mercatoelettrico.org>>.
- [43] New York ISO Website <<http://www.nyiso.com>>.
- [44] Kohavi Ron, John George H. Wrappers for feature subset selection. *Artif Intell* 1997;97(1–2):273–324.
- [45] Pindoriya NM, Singh SN, Singh SK. An adaptive wavelet neural network-based energy price forecasting in electricity markets. *IEEE Trans Power Syst* 2008;23(3):1423–32.
- [46] Hansen L, Salamon P. Neural network ensembles. *IEEE Trans Pattern Anal Mach Intell* 1990;12:993–1001.
- [47] Amjadi N, Daraeepour A, Keynia F. Day-ahead electricity price forecasting by modified relief algorithm and hybrid neural network. *IET Gener Transm Distrib* 2010;4(3):432–44.
- [48] Zareipour H, Canizares CA, Bhattacharya K, Thomson J. Application of public-domain market information to forecast Ontario's wholesale electricity prices. *IEEE Trans Power Syst* 2006;21(4):1707–17.
- [49] Bompard E, Ciwei G, Napoli R, Torelli F. Dynamic price forecast in a competitive electricity market. *IET Gener Transm Distrib* 2007;1(5):776–83.
- [50] Zareipour H, Canizares C, Bhattacharya K. Economic impact of electricity market price forecasting errors: a demand-side analysis. *IEEE Trans Power Syst* 2010;25(1):254–62.