Applied Energy 190 (2017) 291–305

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods

Zhang Yang [a,b,*], Li Ce [a], Li Lian [a]

[a] School of Information Science and Engineering, Lanzhou University, Lanzhou, China
[b] Network and communication Center, Lanzhou University, Lanzhou, China

HIGHLIGHTS

- Propose a new hybrid model for day-ahead electricity price forecasting.
- Analyze complex features of prices series by wavelet transform and stationarity test.
- The proposed model has linear and nonlinear prediction abilities.
- Three real electricity markets data are used to assess the forecasting performance.
- Improve the prediction accuracy of electricity price forecasting.

ARTICLE INFO

Article history:
Received 14 February 2016
Received in revised form 19 November 2016
Accepted 27 December 2016

Keywords:
Electricity price forecasting
ARMA
KELM
Wavelet transform
SAPSO

ABSTRACT

Electricity prices have rather complex features such as high volatility, high frequency, nonlinearity, mean reversion and non-stationarity that make forecasting very difficult. However, accurate electricity price forecasting is essential to market traders, retailers, and generation companies. To improve prediction accuracy using each model's unique features, this paper proposes a hybrid approach that combines the wavelet transform, the kernel extreme learning machine (KELM) based on self-adapting particle swarm optimization and an auto regressive moving average (ARMA). Self-adaptive particle swarm optimization (SAPSO) is adopted to search for the optimal kernel parameters of the KELM. After testing the wavelet decomposition components, stationary series as new input sets are predicted by the ARMA model and non-stationary series are predicted by the SAPSO-KELM model. The performance of the proposed method is evaluated by using electricity price data from the Pennsylvania-New Jersey-Maryland (PJM), Australian and Spanish markets. The experimental results show that the developed method has more accurate prediction, better generality and practicability than individual methods and other hybrid methods.

1. Introduction

In deregulated and competitive electricity markets, electricity price forecasting has become a very valuable tool for all market participants. Producers and consumers can use prediction information to adjust their production schedule and select the best bidding strategy to maximize their respective benefits. For managers of markets, the forecasting ensures healthy, stable and orderly operation of the power market. High-quality electricity price forecasting also has a very important role in power investment decisions and transmission expansion. However, accurate electricity price forecasting is a rather complex task because electricity prices have many exclusive features, such as high frequency, high volatility, nonlinearity, multiple seasonality, mean reversion, the calendar effect and price spikes [1,2].

In the recent years, many researchers have proposed different approaches to predict electricity prices. Some recently published papers and literature reviews can be found in [3–5]. According to [4], electricity price forecasting methods can be broadly summarized in five categories: multi-agent models, fundamental methods, reduced-form models, statistical approaches and computational intelligence techniques. Among the different methods, two widely used approaches are statistical approaches and computational intelligence. Statistical approaches such as exponential smoothing, auto regression (AR), moving average (MA), autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA), ARMA with exogenous variables (ARMAX) and generalized

* Corresponding author at: School of Information Science and Engineering, Lanzhou University, Lanzhou, China.

E-mail addresses: zhyang@lzu.edu.cn (Z. Yang), lil@lzu.edu.cn (L. Ce), ll@lzu.edu.cn (L. Lian).

http://dx.doi.org/10.1016/j.apenergy.2016.12.130
0306-2619/

autoregressive conditional heteroskedastic (GARCH) methods use a mathematical combination of the previous electricity prices to predict the current electricity prices. However, statistical approaches are often criticized for their limited ability to capture the non-linear behavior of electricity prices and rapid changes in the price signal [2]. Recent computational intelligence methods have attracted considerable attention from many researchers. These methods can approximate any multivariate function to a desired degree of accuracy by adjusting weightings during online updates and can capture the complex, dynamic and non-linear features of electricity prices. Thus, the computational intelligence method has higher prediction accuracy than statistical approaches. Many authors have reported their work in electricity price forecasting using various computational intelligence methods such as artificial neural networks (ANNs), feed-forward neural networks, recurrent neural networks, fuzzy neural networks and support vector machines [6–13].

Although computational intelligence methods can process multivariable and non-linear problems, the selection of network structure and parameters is mainly dependent on experience. Furthermore, for the ANNs method, because the functional relationships among price series vary with time, some features captured by the ANNs method may lose their value as time changes [3]. Some other drawbacks associated with the ANNs method are fast convergence speed, excessive tunable parameters, slow learning rate, high possibility of entrapment in local minima, long computational time, and over-tuning [14]. These limitations often misrepresent parts of the non-linear input–output relationship and lead to unsatisfactory prediction accuracy. A new learning algorithm called the extreme learning machine (ELM) based on a single-hidden-layer feed-forward neural network (SLFN) was proposed [15]. This algorithm randomly generates the connection weight between the input and hidden layers and the threshold of neurons in the hidden layer, and parameters do not need to be adjusted during the training process. Unlike other neural network algorithms, the output matrix is obtained by minimizing the squared loss function of the least squares solution, a process that requires no iterations and greatly reduces the network parameter settling time. The ELM has been successfully applied to various prediction applications such as load, price, wind power, computer sales and bankrupt [16–27]. In kernel-based ELM (KELM), the kernel matrix replaces the randomness matrix of ELM. This method uses the kernel function to map the training samples into a high-dimensional space for training. In KELM, if the penalty factor and the kernel parameter are determinant, the output of the KELM is also fixed. Therefore, the KELM avoids the random fluctuations of the ELM. In [28], optimal kernel parameters can improve the stability and generalization of the KELM.

In newly deregulated electricity markets, uncontrollable and unpredictable contingencies increase the complexity of accurate price prediction. A single prediction model does not meet the needs of all market participants in terms of error and accuracy. The hybrid method is a better option in that it effectively combines linear and nonlinear modeling capabilities and is thus able to capture different patterns in the electricity price series data and improve the forecasting accuracy. The wavelet transform (WT) is a data preprocessing method that provides useful decomposition information in terms of time and frequency, making this method suitable for the analysis of non-stationary signals such as price series. The WT method in combination with other forecasting techniques, such as statistical models and neural network models, has been proposed in the literature for forecasting electricity prices. Tan et al. [29] proposed a novel price forecasting method based on the WT combined with the ARIMA and GARCH models. Mandal et al. [30] presented a novel hybrid technique based on the WT, the firefly (FF) algorithm, and a fuzzy ARTMAP (FA) network to forecast day-ahead electricity prices in the Ontario market. In [31], Zhang et al. presented a new method that included the WT, CLSSVM and EGARCH, and their model was applied to the marginal price (LMP) of the PJM market and market clearing price (MCP) of the Spanish market. Shayeghi and Ghasemi [32] suggested a hybrid methodology that combined the WT, the Gravitational Search Algorithm and LSSVM to predict Iran's, Ontario's and Spain's price markets. The author of [33] developed a hybrid method based on wavelet and ELM to improve the forecasting accuracy. Shafie-Khah et al. [34] presented a hybrid method based on WT, ARIMA and RBFNN; the proposed method was examined with respect to the electricity market of mainland Spain. Nguyen et al. [35] presented a forecasting model combining the WT with fixed and adaptive machine learning/time series models. Voronin and Partanen [36] used ARIMA, WT and neural networks techniques to forecast demand and price forecasting in the Nordic Power Pool. In two recent papers, Abedinia and Amjady [37] adopted a hybrid approach based on the WT, ARIMA and radial basis function neural networks to predict electricity prices in mainland Spain. Zhang et al. [38] proposed a novel hybrid method using a generalized regression neural network (GRNN) combined with the WT and a generalized GARCH model to forecast Spanish electricity prices.

Real electricity price data are nonlinear and non-stationary. Therefore, in an effective electricity price forecasting model, nonlinear and linear predictions must both be considered. After using wavelet transform, the decomposed series have more stable variances than the original series and can thus be more accurately forecasted. The ARMA is frequently applied in linear and stationary time series because of its high performance and robustness [39]. The KELM has many advantages, such as time efficiency, nonlinear capacity, and good generalization performance. Because the kernel parameters affect the generalization ability of the KELM, the KELM based on self-adapting particle swarm optimization (SAPSO-KELM) method is developed in this paper. The SAPSO-KELM is more stable and more effective than the KELM due to the optimization of its parameters by SAPSO. The integration of the WT with the ARMA and the SAPSO-KELM should be investigated to form a hybrid approach for better price forecasting. To the best of the author's knowledge, the integration of these three technologies has not been tried or applied to electricity price forecasting before. Therefore, this paper proposes a new hybrid method based on a combination of the WT, the ARMA, and the SAPSO-KELM models. The WT is used to decompose the electricity price series into stationary and volatility series. The stationary series are predicted by ARMA, whereas the non-stationary series are forecasted by the SAPSO-KELM model. The forecasted values generated separately by ARMA and SAPSO-KELM are combined so that the final predicted electricity price is generated. To illustrate the performance of the proposed model, price forecasts in the Pennsylvania-New Jersey-Maryland (PJM), the Australian and the Spanish electricity markets are calculated and compared with those obtained using the ARMA model, the KELM model, and other hybrid models, including combined models reported in the literature.

The remainder of this paper is organized as follows. Section 2 introduces the fundamental theories of the proposed method. Section 3 describes the proposed hybrid approach in detail. In Section 4, three case studies and prediction results are given. The conclusions are presented in Section 5.

2. Methodologies

In this section, we present the methodologies used for electricity price forecasting.

2.1. Wavelet transform

The WT method is similar to the Fourier transform method [40]. The original signal is decomposed by a group of functions. Those functions are called wavelet functions and are obtained by stretching and translating the mother wavelet. The WT provides good localization features in both the time and frequency domains [41]. Decomposition information is appropriate for analyzing short-time series that have no stationary features, such as electricity price data. The WT can be divided into two types: continuous wavelet transformation (CWT) and discrete wavelet transformation (DWT).

For signal y(t), CWT is defined as

$$CWT_y(\alpha,\tau) = \frac{1}{\sqrt{|\alpha|}}\int_{-\infty}^{+\infty} y(t)\psi^*\left(\frac{t-\tau}{\alpha}\right)dt \tag{1}$$

where α is the scale parameter, τ is the translation parameter, $\psi^*(x)$ is the complex conjugate function, and $\psi(x)$ is the mother wavelet.

DWT is defined as

$$DWT_y(m,n) = \alpha_0^{-\frac{m}{2}}\int_{-\infty}^{+\infty} y(t)\psi^*(\alpha_0^{-m}t - n\tau_0)dt \tag{2}$$

where m is the scaling constant (decomposition level) and n is the translating constant which is an integer.

The fast DWT algorithm, which was developed by Mallat [42] uses low-pass and high-pass filters instead of father and mother wavelets. This algorithm allows irregular information to be extracted from the original signal. The low-pass filter is called the scaling function and uses the analysis of low-frequency components, whereas the high-pass filter is called the wavelet function and uses the analysis of high-frequency components. From a given signal F(t), one approximation series d and detail series (c_1, c_2, c_3, c_4) are obtained by Mallat's algorithm (Fig. 1). The approximation series represents the low-frequency components, which contain the trend information of the original signal. In contrast, the detail series represents the high-frequency components, which contain the characteristics of influential factors associated with the original signal. The mother wavelet has an important effect on the results obtained. In this paper, Daubechies 5 is used as the mother wavelet, which offers an appropriate tradeoff between wavelength and smoothness. The decomposition results exhibit appropriate behavior for the electricity price. Similar wavelets have been considered by previous researchers [36,43]. Five levels of decomposition are considered because this wavelet describes the electricity price series in a more thorough and meaningful manner.

2.2. ARMA

The ARMA model is a random time series model and is one of the most common models in time series forecasting analysis [44]. Time series forecasting analysis consists of three basic types: AR, MA and ARMA. An ARMA (p, q) model of order p and q is defined by

$$y_t = \varnothing_1 y_{t-1} + \varnothing_1 y_{t-2} + \cdots + \theta_p y_{t-p} + v_t - \theta_1 v_{t-1} - \theta_2 v_{t-2} - \cdots - \theta_q v_{t-q} \tag{3}$$

where Y_t is the ARMA series, the stochastic term v_t is the independent white noise series, Φ_p is the autocorrelation coefficient, and θ_q is the moving average coefficient. If p = 0, the equation becomes an MA model of order q; if q = 0, it becomes an AR model of order p.

The ARMA model is constructed as follows:

Step 1: Data stationary analysis: via an autocorrelation graph and unit root test [45].
Step 2: Identification of the ARMA (p, q) model structure: through the autocorrelation function (ACF) and partial autocorrelation function (PACF).
Step 3: Estimation of the model parameters and construction of the model.
Step 4: Model testing, in which the estimated residual is white noise.
Step 5: Forecast of the future values.

2.3. Kernel-based extreme learning machine

The ELM is a type of SLFN. This model, which has been widely used in many fields, has fast learning speed and generalization ability. The advantage of the ELM is that the hidden layer parameters do not require tuning and the input weights and biases are randomly produced. Thus, the computation time is low. For N arbitrary samples (x_i, t_i), $x_i \in R^n$, $t_i \in R^n$, i = 1,2 ... n. If the activation function of the hidden layer is $g(x)$, the output matrix T of the SLFN with L hidden nodes is

$$T = \begin{bmatrix} \sum_{i=1}^{l}\beta_{i1}g(w_i x_j + b_i) \\ \sum_{i=1}^{l}\beta_{i2}g(w_i x_j + b_i) \\ \vdots \\ \sum_{i=1}^{l}\beta_{im}g(w_i x_j + b_i) \end{bmatrix},\ j = (1,2,\ldots n) \tag{4}$$

This formula can also be written as

$$H\beta = T' \tag{5}$$

H in (5) can be represented as follows:

$$H(w_1, w_2, \ldots, w_l, b_1, b_2, \ldots, b_l, x_1, x_2, \ldots, x_n) = \begin{bmatrix} g(w_1\cdot x_1 + b_1) & g(w_2\cdot x_1 + b_2) & g(w_l\cdot x_1 + b_l) \\ g(w_1\cdot x_2 + b_1) & g(w_2\cdot x_2 + b_2) & g(w_l\cdot x_2 + b_l) \\ & \vdots & \\ g(w_1\cdot x_n + b_1) & g(w_2\cdot x_n + b_2) & g(w_l\cdot x_n + b_l) \end{bmatrix} \tag{6}$$

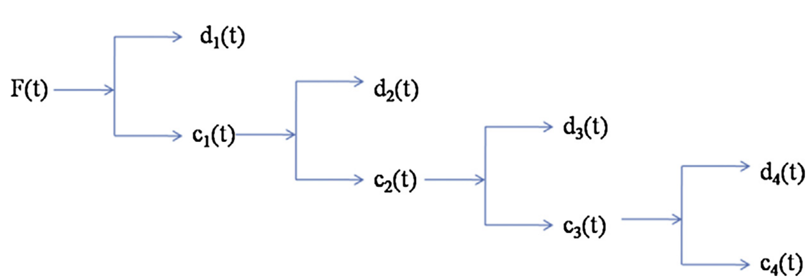

Fig. 1. Mallat algorithm.

where H is the output matrix of the hidden layer, the hidden layer threshold is b and g is the activation function. Then, the typical SLFN model is expressed as

$$f_L(x) = \sum_{i=1}^{L} \beta_i G(a_i, b_i, x) \tag{7}$$

where β is the output weight between the hidden and output layers, G is the output function of a hidden node.

If the training data set is $(x_i, t_i), i = 1, 2, \ldots, N$, where $x \in R^d$ and $t_i \in \{-1, 1\}$, and t_j is the same vector as O_j. The SLFN with L hidden nodes can approximate these N samples with zero error, as follows:

$$\sum_{i=1}^{n} \|O_i - t_i\| = 0 \tag{8}$$

where o is the actual output value of the SLFN. If g, β, and b exist, then

$$\sum_{i=1}^{N} g(w_i + x_i + b_i)\beta_i = t_j, \ (j = 1, 2, \ldots, N) \tag{9}$$

The matrix can also be expressed as

$$H\beta = T,\ T \in R^{N\times m},\ \ \beta \in R^{N\times m},\ \ H = H(w, b) = g(w_i \cdot x_i + b) \tag{10}$$

From Huang's two theorems [46,47], the input weights and hidden biases are randomly generated instead of tuned. The only unknown parameter is the output weight β, which can be solved by the ordinary least-squares solution method.

The solution of the equation is given by

$$\beta = H^{\dagger} T \tag{11}$$

where $H^{\dagger}$ is the Moore-Penrose generalized inverse of matrix H [48]. Using this method, the ELM tends to obtain a good generalization and a fast learning speed. From the orthogonal projection method and ridge regression theory [49,50], the output weight β can be calculated by adding a positive value $1/\lambda$ as follows:

$$\beta = H^T \left(\frac{1}{\lambda} + HH^T\right)^{-1} T \tag{12}$$

The output function of ELM is:

$$f(x) = h(x)\beta = h(x)H^T \left(\frac{1}{\lambda} + HH^T\right)^{-1} T \tag{13}$$

Huang et al. proposed the kernel ELM [51]. If the user does not known the feature mapping h(x), a kernel matrix for ELM can be defined. The formula (14) can be written as follows:

$$f(x) = h(x)H^T \left(\frac{1}{\lambda} + HH^T\right)^{-1} T = \begin{bmatrix} k(x,x_1) \\ k(x,x_2) \\ \vdots \\ k(x,x_n) \end{bmatrix}^T \left(\frac{1}{\lambda} + HH^T\right)^{-1} T \tag{14}$$

In Eq. (15), users do not need to know the hidden layer feature mapping h(x), and the function does not specify the number of hidden nodes L. The kernel k(u, v) that replaces h(x) and L is given by the user. λ is the penalty factor. The stable kernel function replaces the random mapping of the ELM, and the output weight is more stable. Therefore, the KELM has better generalization ability than the ELM.

2.4. KELM based on self-adapting particle swarm optimization

The kernel method was first used in support vector machines [52], and it can be widely used with problems of classification and regression. Classification of nonlinear data in the original space is mapped into a high-dimensional space. This operation is often less computationally burdensome than the explicit computation of the coordinates. All functions that can satisfy Mercer's theorem can be taken as kernel functions. The different kernel functions have various learning abilities and generalization performances. Therefore, the three types of kernel functions, the polynomial kernel, the multiquadric kernel and the Gaussian kernel are used.

Two parameters, the penalty factor and the kernel parameter affect the generalization of the KELM. Thus, in this paper, the KELM based on self-adapting particle swarm optimization (SAPSO) is developed. In the process of training the neural network, the SAPSO algorithm is adopted to adjust the penalty factor and kernel parameter of the KELM to achieve stability and more effective regression performance. PSO, which was first presented by Kennedy and Eberhart in 1995 [53], is the most widely used optimization approach. PSO can efficiently find optimal or near optimal solutions to optimization problems. The PSO algorithm is based on the biological and sociological behavior of birds. In PSO, every potential solution of the optimization problem is a bird called a "particle". Every particle has a fitness value determined by its objective function, and is flown through the multi-dimensional search space with random and adaptable velocity to find the lower function values. The swarm X consists of N particles in an m-dimensional search space. Each particle has two vectors, position $x_i = (x_{i1}, x_{i2}, \ldots, x_{im})$ and velocity $v_i = (v_{i1}, v_{i2}, \ldots, v_{im})$, where i = 1, 2, 3, ... m. When x_i is substituted for the objective function, it is calculated as a corresponding fitness value, which is the measurement criterion for determining the quality of the particle. Each particle is accelerated while searching for the particles in the locally best and globally best positions. P_i is the locally best position, whereas P_i is the globally best position. During the iteration process, each particle updates its velocity and location through P_i and P_j. According to the fitness, if the current position is better than the best position P_i, P_i takes the place of the current position. Otherwise, P_i remains unchanged. P_j is changed similarly. This process continues with further iterations until the maximum number of iterations is reached or the objective function's minimum error is achieved. The velocity and position can be updated using the following formula:

$$V_{id}(k+1) = w \times V_{id}(k) + c1 \times r1 \times (p_{id\_}x_{id}) + c2 \times r2 \times (p_{id\_}x_{id}), \quad i = 1, 2, \ldots, N \tag{15}$$

$$x_{id}(k+1) = x_{id}(k) + V_{id}(k+1) \tag{16}$$

where w is the inertia weight, which provides a balance between local and global exploration (a larger w will facilitate the global search, whereas a smaller w will facilitate the local search), k is the number of iterations, d = 1, 2, 3 ..., m, c1 and c2 are acceleration coefficients, which are positive numbers, r1 and r2 are uniform distribution numbers in the range [0, 1], and N is the number of particles. The position of particle i in dimension d and iteration k is denoted by $x_{id}(k)$. The self-adapting inertia weight update strategy [54] was adopted to balance the search ability of the algorithm and is set according to the following equation:

$$w = w1 \times (w1 - w2) \times (T - K) \div T \tag{17}$$

where $w1$ and $w2$ are the maximum and final values of the inertia weight, respectively. K is the present iteration number and T is the maximum number of allowable iterations. This strategy can strike a good balance between global and local searching in the entire search process [55].

The main steps of the SAPSO-KELM can be described as:

(1) Initialization of the size, velocity and position of the particles.
(2) Calculation of the fitness value by the objective function.
(3) Updating of the local best position P_i, the global best position P_j, the velocity and the position of each particle according to the fitness value.
(4) Repetition of the steps 3 until the termination condition is satisfied.
(5) Output of the optimal penalty factor and kernel parameter to KELM.

After using the optimal kernel parameters, SAPSO-KELM has more stable and excellent generalization performance.

2.5. Forecast error measures

To assess the forecasting performance of the proposed model, different error measures are employed as the criteria. The most widely used measures are the mean absolute error (MAE), the mean absolute percent error (MAPE) and the root mean square error (RMSE). A low error measure indicates a better prediction performance. Many spikes or outliers in the electricity price series greatly affect the distribution of the price. The most common approach is to normalize the absolute error by the average price obtained in the evaluation interval (a day). Thus, the daily weighted mean absolute errors (DMAE) and the weekly weighted mean absolute errors (WMAE) are used. In additional, Theil U statistic 1 and Theil U statistic 2 performance metrics are also adopted [8]. The Theil U statistic 1 value is between 0 (indicating good accuracy) and 1 (indicating poor accuracy). If the value of Theil U statistic 2 is greater than 1, it indicates poor forecasting accuracy. The seven error measures are defined as follows:

$$\mathrm{DMAE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i - y_i|}{Y} \tag{18}$$

$$\mathrm{WMAE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i - y_i|}{Y} \tag{19}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|Y_i(t) - y_i(t)| \tag{20}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_i(t) - y_i(t)}{Y_i(t)}\right| \tag{21}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n}(Y_i(t) - y_i(t))^2} \tag{22}$$

$$\mathrm{U1} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i(t) - y_i(t))^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}Y_i(t)^2} + \sqrt{\frac{1}{n}\sum_{i=1}^{n}y_i(t)^2}} \tag{23}$$

$$\mathrm{U2} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{Y_{i+1}(t) - y_{i+1}(t)}{Y_i(t)}\right)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{Y_{i+1}(t) - y_i(t)}{Y_i(t)}\right)^2}} \tag{24}$$

where Y_i and y_i denote the actual and forecast values at time i, respectively, Y is the average actual value for the day, and n is the number of data samples. For the PJM and Spanish electricity market cases, n is 24 for the DMAE and 168 for the WMAE; for the Australian electricity market case, n is 48 for the DMAE and 336 for the WMAE.

3. Proposed hybrid method

Electricity price forecasting is difficult due to the influence of many factors, including load, weather, bidding strategies, and historical prices. In addition, because electricity is non-storable and requires maintaining a balance between demand and supply and due to the inelastic nature of the electrical short time period and its oligopolistic generation side features [3], electricity prices show complex behavior. To achieve the accurate price predictions, the forecasting model must be able to capture different patterns in the price series. Normally, the price series present not only stationary linearity but also high fluctuation and irregularity. The WT can convert abnormal price series into a set of normal constitutive series, and the decomposed series can be forecast more accurately than can be achieved through direct forecasting by a model. However, information on the wavelet decomposition series cannot be completely reflected by any single model; therefore, if the wavelet decomposition series as the input sets are predicted directly by a single model, the prediction performance might not be satisfied. A better option is to use a hybrid method that can utilize each model's unique characteristics to extract linear and nonlinear information in the decomposed price series. This method generates good forecasting results. In the hybrid method, the selection of a single model becomes particularly important. The ARMA model is a quantitative model that is based on the data sequence of the autocorrelation function and the partial correlation function, and it can reflect the internal connection of the data between past and current activities. ARMA is commonly used to resolve the linear time series problem. In particular, it shows good prediction performance for stationary series. The advantages of KELM are its fast learning speed and good generalization capabilities, both of which make it suitable for non-linear prediction. Because the penalty factor and the kernel parameters of KELM affect its prediction performance, the SAPSO algorithm is adopted as a training process to adjust these parameters of KELM. The use of SAPSO-KELM method makes it easier to obtain accurate predictions. Hence, a new hybrid method based on WT, ARMA and SAPSO-KELM is proposed to predict electricity prices. Through this hybrid structure, the different patterns hidden in prices can be adequately extracted.

The operation of the hybrid model of electricity price forecasting can be described as follows:

- Data preparation. Each price dataset is divided into training data and testing data. The training data are selected for model development, and the test data are selected to assess the established model.
- Wavelet transform. Two factors significantly affect the prediction results when the WT is involved in the prediction application: the mother wavelet and the number of decomposition levels. In this paper, Daubechies 5, which offers an appropriate tradeoff between wavelength and smoothness, is used as the mother wavelet. It can reflect the uncertain factors associated with price features. Five levels of decomposition are considered because this wavelet describes the electricity prices series in a more thorough and meaningful manner. Using the WT, the original electricity price data are decomposed into one approximation series A and detail series d1, d2, d3 ..., d_n.
- Checking of the stationarity assumption for each series. Stationarity is used to assess the subseries as new input sets that are predicted by either ARMA or SAPSO-KELM. Two methods, the autocorrelation map and the unit root test, are used together to test the stationarity of the time series. If the autocorrelation

coefficient decays exponentially and approaches zero, this is an indication that the time series meets the stationary assumption. The unit root test includes the ADF test, the DF test, and the Said-Dickey test. In this paper, the ADF test is used to test the time series' stationarity. The results of ADF are the ADF statistic and critical values below the 1%, 5% and 10% levels, which are the three significant levels of the corresponding critical values. If the ADF statistic is lower than the critical values, the unit root hypothesis is rejected, and this time series is stationary.

- Construction of the ARMA model and prediction of the stationary series using the ARMA model,
- Processing of the non-stationary series as an input to SAPSO-KELM, optimization of the penalty factor and kernel parameter by SAPSO, and obtaining of the predicted values. The SAPSO parameters are set as follows: the particle size is 30, the acceleration coefficients c1 and c2 are both set at 2, the initial inertia weight w is 0.9, and the maximum number of iterations is 500. The mean square root (MSE) is used as the fitness function. The MSE is defined as follows:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - y_i)^2 \tag{25}$$

where Y_i represents the actual values and y_i represents the forecasted values.

The kernel functions are the polynomial kernel, the multi-quadric kernel and the Gaussian kernel.

- Combine the prediction results of the ARMA and SAPSO-KELM models.

The components of the proposed method are depicted graphically in the flowchart presented in Fig. 2.

4. Case studies and results

In this paper, three cases are used to illustrate the prediction performance of the proposed method.

Case (1): the proposed method is applied to predict electricity prices for the PJM market in 2006 and 2004. First, five single days from January to May 2006 and two weeks between February 1–7 2006 and February 22–28, 2006 are selected to forecast and compared with those in [56]. Second, four weeks of February, April, October and November are selected for the winter, spring, summer and autumn seasons, respectively, in the year 2004, which is the test period considered by Pindoriya et al. [57] and Shrivastava et al. [33].

Case (2): the proposed method is tested for price forecasting of the Australian electricity market. First, the daily forecasted prices for the twelve months of the year 2014 in New South Wales (NSW) and Victoria (VIC) are carried out. The last day of the month is used to validate the performance of the proposed method, whereas the training set comprised 30 or 29 days of data prior to the beginning of the test day. Second, four one-week periods corresponding to the four seasons of 2013 (January 28 to February 3, April 29 to May 5, July 29 to August 4 and October 28 to November 3) are selected for VIC. The price data from the four weeks prior to the first day of each test week are used as the training samples. The data on NSW and VIC used in this case study are obtained from website [58].

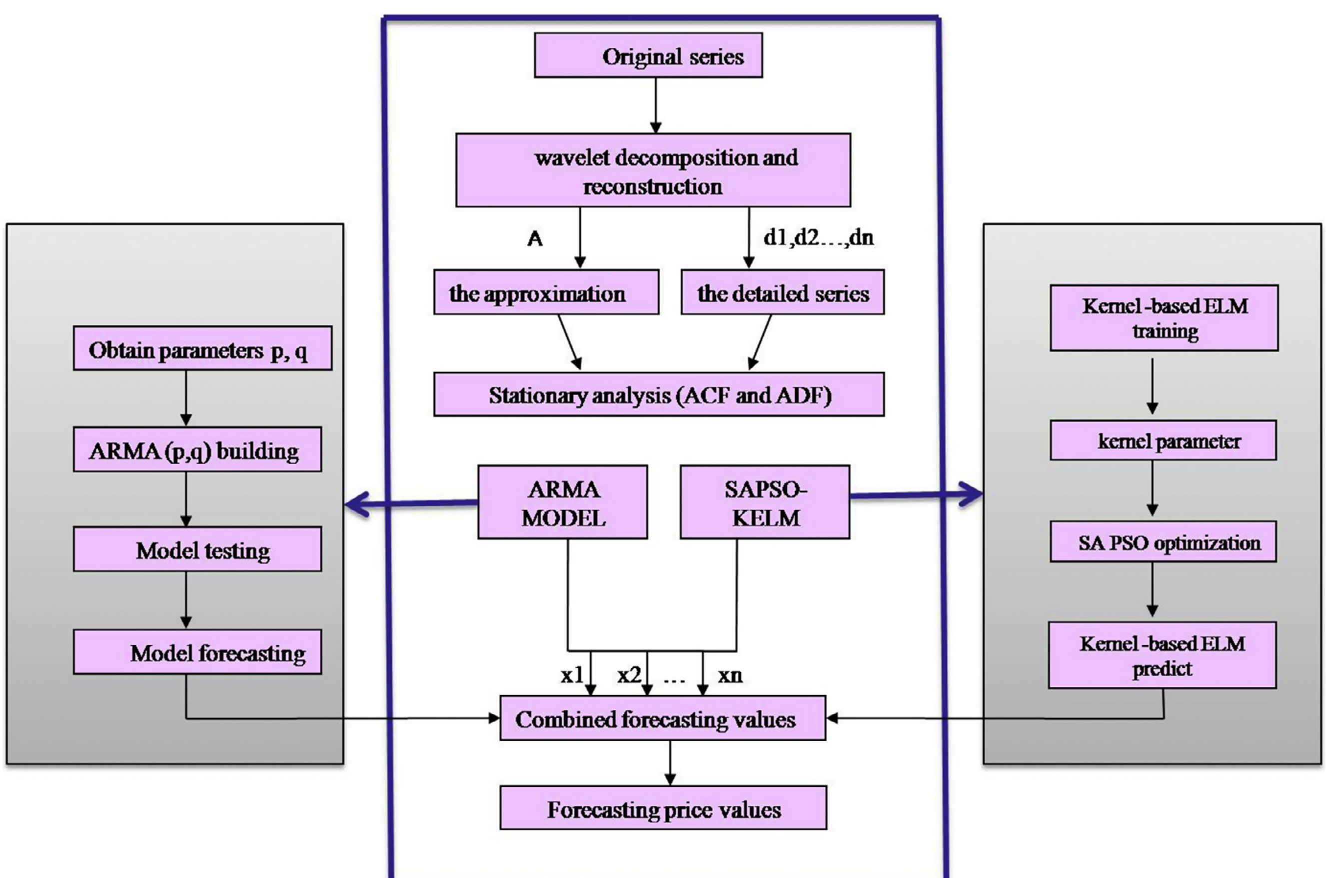

Fig. 2. The overall framework of the proposed model.

Case (3): the proposed method is tested for price forecasting of the Spanish electricity market. First, four one-week periods representing different seasons in 2009 (January 26 to February 1, April 27 to May 3, July 27 to August 2 and October 26 to November 1) are selected as the test period. The price data from the four weeks prior to the first day of each test week are used as the training samples. Second, the proposed approach is applied to the twelve months of 2015. Similarly, the last day of the month is used to validate the performance of the proposed method, whereas the training set comprised 30 or 29 days of data prior to the beginning of the test day. The data are obtained from the official Spanish electricity website [59].

4.1. Analysis of the case study results

After wavelet decomposition and reconstruction, the approximation and detail series are produced. These series represent the general trend component and high-frequency component, respectively. Taking the data of VIC as an example, stationarity analysis is first used for the wavelet decomposition series. As shown in Fig. 3, the ACF coefficient of the D1 experiences exponential decay and becomes close to zero. The ACF coefficients of the other series do not fall into a random area. The ADF results of D1 show that the t-statistics are smaller than the 1% significance levels; thus, the unit root hypothesis is rejected. However, the ADF results of other decomposition series are larger than the 10% significance levels; thus, the unit root hypothesis cannot be rejected. Both methods display slightly more confidence to indicate that the D1 is a stationary series and the other series are non-stationary. Thus, forecasting of the D1 series is appropriate for the ARMA model.

Second, the ARMA model structure and parameters are identified. Through correlograms (ACF and PACF), the ACF coefficient shows an exponential decaying pattern after 1 lag, whereas the PACF shows a cutoff value after 11 or 12 lag. For brevity, the PACF graph is not provided here. When estimating the ARMA model and the two parameters p and q, two general information criteria are available for justification, i.e., the Akaike information criterion (AIC) and the Schwarz criterion (SC). The best model is characterized by the smallest AIC and SC values. ARMA (11, 1) is established by testing the values of the AIC and the SC.

Further analyses are conducted to determine whether the model residual is white noise. The ACF and PACF are also used for identification. Finally, future values are forecasted based on the final estimated model. The D1 series forecasted values are obtained.

The same method is applied to the other wavelet decomposition series. The non-stationary series, such as approximation series A and the remainder of the detailed series, are predicted by the SAPSO-KELM. The stationary series are predicted by ARMA.

4.2. Comparison and discussion

4.2.1. Case 1: Prediction of electricity prices in the PJM market

To enable a fair comparison, the five test days include January 20, February 10, March 5, April 7 and May 13 in 2006, and two weeks, February 1–7 and February 22–28 in 2006 are selected, which are exactly the test periods considered by Mandal et al. [56], whose error measure is also adopted to assess the prediction capacity of the proposed method. The obtained results are shown in Table 1, indicating that the prediction accuracies of the proposed method for five test days and two test weeks are notably higher than the results of [56]. As an additional explanation, Mandal et al. [56] state that his prediction for April 7 is particularly inaccurate due to the morning and evening peaks. However, as Table 1 shows, the proposed methods improve the prediction accuracy for the morning and evening peaks, because the DMAE generated by the proposed methods is reduced from 9.02% to 5.32%, with 63.96% improvement. In the weeklong test periods, the average MAPE is obtained by the proposed method for February 1 to 7 decreases from 7.66 to 6.35, representing a 17.10% improvement. Similarly, compared with the results of [56], the average MAPE is obtained by the proposed method for February 22 to 28 decrease from 8.88 to 4.10, representing a 53.82% improvement. All prediction results reveal that the proposed approach has not only process ARMA's linear prediction ability but also displays SAPSO-KELM's good generalization performances, thus it can better capture the complex behavior of electricity prices compared to comparison methods.

Table 2 shows the prediction results generated by the methods used by Pindoriya et al. [57], Shrivastava et al. [33] and the proposed method for the four test weeks. It is observed that the

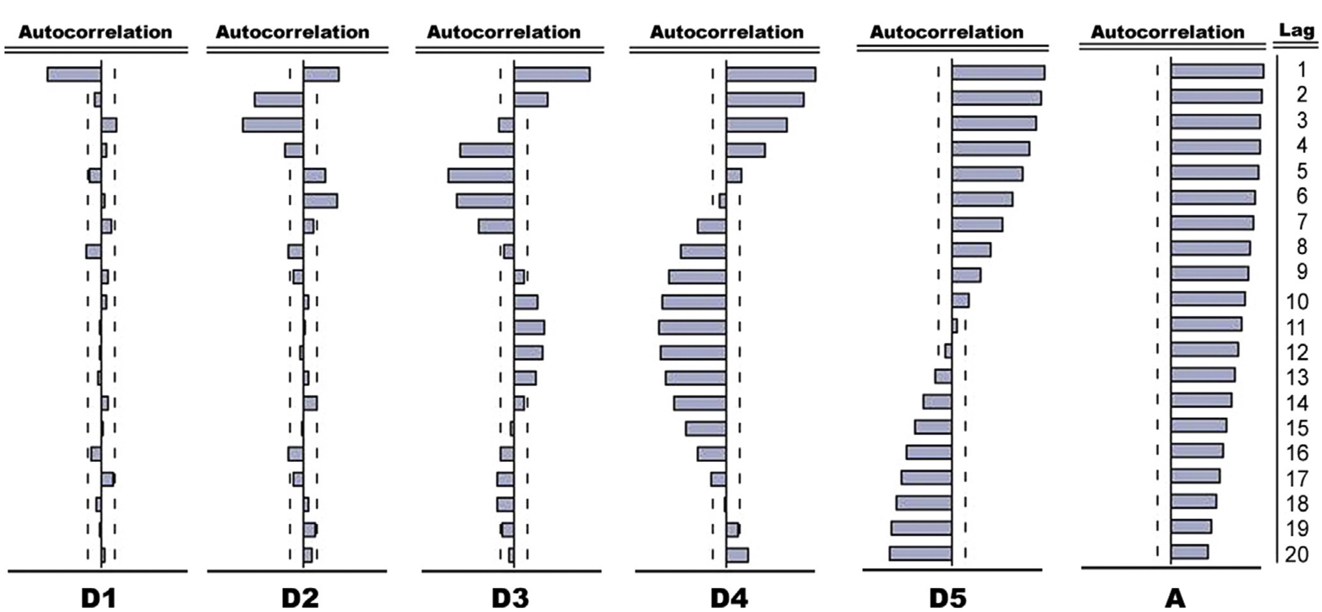

Fig. 3. Autocorrelation function graph of the decomposition series of Victoria in March.

Table 1
Comparison of the DMAE (%) and WMAE (%) for PJM of 2006.

Year 2006	ANN method [56] DMAE			Proposed method DMAE		Improvement DMAE
January 20	6.93			**5.10**		**16.88%**
February 10	7.96			**4.09**		**48.62%**
March 05	7.88			**2.84**		**63.96%**
April 07	9.02			**5.32**		**41.02%**
May 13	6.91			**3.94**		**42.98%**
	Max WMAE	Min WMAE	Ave. WMAE	Max WMAE	Min WMAE	Ave.WMAE
February 01–07	11.32	5.94	7.66	**8.28**	**4.79**	**6.35**
February 22–28	12.37	5.66	8.88	**5.28**	**2.98**	**4.10**

Table 2
Comparison of the WMAE (%) for PJM in 2004.

Test week	Results in [57]	Results in [33]	Proposed method
Feb. 23–Feb. 29	6.362	6.010	**5.278**
May. 17–May. 23	5.976	4.937	**4.805**
Aug. 23–Aug. 29	5.954	5.843	**5.332**
Nov. 22–Nov. 28	6.648	6.056	**5.492**
Average	6.235	5.712	**5.227**

WMAE values generated by the proposed method for all test weeks are lower than generated by the methods used by Pindoriya et al. [57] and Shrivastava et al. [33], with the average WMAE being respectively reduced by 16.16% and 8.49%. Although the differences in the WMAE values predicted by the three methods are small, the generalization ability of the proposed method is better than that of the methods used by Pindoriya et al. [57] and Shrivastava et al. [33]. This comparison also indicates that the accuracy of the proposed method is due to its linear and non-linear prediction ability, which can capture the complex characteristics of electricity prices.

4.2.2. Case 2: Prediction of electricity prices in Australian market

To illustrate the efficiency of the proposed method, data for twelve months of 2014 from the NSW and VIC are tested, and the six types of error indices—MAPE, WMAE, RMSE, MAE, U1 and U2—are used. Firstly, the performance of the proposed method is compared with that of the individual ARMA and individual KELM methods both of which lack a hybrid structure; secondly, the proposed method is compared with other hybrid methods, such as the W-ARMA-BP, which is the wavelet transform in combination with ARMA and the back propagation neural network (BPNN), and the W-ARMA-LSSVM, which is the wavelet transform in combination with ARMA and a least squares support vector machine. The reason for selecting BPNN and LSSVM is that they are both not only often applied to electricity price forecasting but also often selected as benchmarks and compared with ELM [60–62].

The results in Table 3 are based on the data obtained from NSW. It can be seen that when the ARMA and KELM models are not combined n a hybrid method, the accuracy of the prediction is less than that obtained with the proposed method; the average MAPE obtained with the two individual methods increases from 3.74% to7.09% and from 3.74% to 15.61%, respectively. This result shows that single ARMA and KELM models cannot effectively capture complex electricity price features. Because the ARMA cannot effectively solve nonlinear problems, its predictions are less accurate than the KELM. Compared to the other hybrid method, the prediction performance of the proposed method is still superior, with an average MAPE value of 3.74%, whereas the average MAPE values of the W-ARMA-BP and W-ARMA-LSSVM methods are 7.9% and 6.58%, respectively. Thus, the prediction errors can be significantly reduced when the proposed method is used. The main reason for this is that use of the WT technique causes the decomposed price series to have a more stable variance and no outliers; the ARMA and the SAPSO-KELM models can then better capture the patterns of the electricity price, and the prediction performance is improved. Moreover, because both BPNN and LSSVM have problems in the choice of parameters, the KELM model optimized by SAPSO can achieve better rate of accuracy. The results obtained from W-ARMA-BP and W-ARMA-LSSVM are better compared to the ARMA and the KELM in most cases, which shows that hybrid structure can be improve prediction accuracy. On the two test days August 30 and December 31, the MAPE values are comparatively higher due to the morning and evening peaks and to significant changes in the power market. The prediction behavior of the proposed method is the best in May. In September and November, some special events take place in the power market and lower prices are conducted; therefore, the MAPE values obtained from this two month are unsatisfied. However, relative to the four comparison methods, the prediction accuracy of the proposed method is still the best.

Table 3 also shows the U1 and U2 statistical measures calculated from the proposed method and from the comparison methods. The U1 values of the proposed method are the closest to zero, denoting good accuracy of the proposed method. For the U2 statistic in August, September and December, the U2 values of KELM and ARMA are greater than or close to 1, and these two methods lead to poor performance. In September and December in particular, the U2 values of W-ARMA-BP and W-ARMA-LSSVM are also greater than 1. Again, these results demonstrate that the proposed method has the best prediction performance.

To verify that the proposed method had good adaptability in different environments, the hybrid model is used to predict the electricity price in VIC for the year 2014. Based on Table 4 and without using wavelet transform and SAPSO, the MAPE values of the KELM and ARMA models for a day in March are 5.2% and 18.92%, respectively. However, the MAPE value of the proposed method reaches 2.19% and decreases by 58% and 88%, respectively, for the two models. The reason for the lower prediction accuracy of December 31 is the presence of morning and evening peaks on that day. For the test days in June and September, the proposed method performed better than the single models and the other hybrid models, which also confirms that the proposed method overcomes the limitations of the single models and exhibits good prediction ability. Similarly, the U1 and U2 statistical values for the different methods show different forecasting accuracies. The performance of ARMA is the worst; the U2 statistical values for eight test days are greater than 1. In August, due to the frequent abnormal price spikes, the U2 statistical values for the KELM, ARMA, W-ARMA-BP and W-ARMA-LSSVM methods are all greater than 1; thus, these methods lead to poor forecasting performance, whereas the U2 value of the proposed method is 0.64, and good accuracy is still obtained from the proposed method. The average MAPEs of the proposed method, KELM, ARMA, W-ARMA-BP and W-ARMA-LSSVM are 4.65%, 10.3%, 23.15%, 13.09 and 10.47%, respectively. These results also show that the proposed method has better

Table 3
Methods comparison of MAPE (%), U1 and U2 in New South Wales, 2014.

Test day	Error	KELM	ARMA	W-ARMA-BP	W-ARMA-LSSVM	Proposed method
January	MAPE	4.50	7.31	4.63	4.50	**3.21**
	U1	0.03	0.06	0.03	0.03	**0.02**
	U2	0.36	0.67	0.38	0.35	**0.27**
February	MAPE	2.68	3.05	2.61	2.47	**1.53**
	U1	0.02	0.02	0.02	0.02	**0.01**
	U2	0.22	0.26	0.21	0.20	**0.13**
March	MAPE	7.40	8.04	5.26	4.73	**3.82**
	U1	0.05	0.05	0.03	0.03	**0.02**
	U2	0.50	0.67	0.42	0.41	**0.31**
April	MAPE	4.75	4.88	3.89	3.46	**2.08**
	U1	0.03	0.03	0.02	0.02	**0.01**
	U2	0.37	0.42	0.33	0.30	**0.18**
May	MAPE	2.86	7.85	2.96	2.30	**1.51**
	U1	0.02	0.04	0.02	0.01	**0.01**
	U2	0.26	0.61	0.25	0.19	**0.13**
June	MAPE	9.52	10.81	7.80	4.62	**2.43**
	U1	0.07	0.08	0.06	0.04	**0.02**
	U2	0.81	0.80	0.72	0.45	**0.22**
July	MAPE	6.05	11.31	5.33	4.80	**3.84**
	U1	0.07	0.17	0.04	0.02	**0.02**
	U2	0.88	1.12	0.46	0.31	**0.30**
August	MAPE	15.06	43.08	14.93	13.31	**6.15**
	U1	0.15	0.17	0.08	0.11	**0.04**
	U2	1.79	2.60	0.97	0.99	**0.42**
September	MAPE	18.81	27.34	12.15	11.75	**5.33**
	U1	0.11	0.12	0.07	0.07	**0.03**
	U2	1.18	1.77	1.06	0.96	**0.40**
October	MAPE	5.41	29.47	4.87	4.68	**1.95**
	U1	0.03	0.19	0.03	0.03	**0.01**
	U2	0.44	2.09	0.44	0.45	**0.17**
November	MAPE	12.44	16.88	11.64	11.06	**5.67**
	U1	0.07	0.15	0.07	0.07	**0.03**
	U2	0.89	1.26	0.84	0.73	**0.42**
December	MAPE	12.21	17.28	18.76	11.22	**7.35**
	U1	0.11	0.29	0.31	0.10	**0.05**
	U2	1.05	1.14	1.33	0.98	**0.54**

prediction accuracy due to its hybrid structure. The linear and nonlinear prediction ability of the proposed method is superior to that of the comparison methods. From the results shown in Tables 3 and 4, the prediction performance of the proposed method is not as good as for the fall and spring test days in 2014. The reason for this is that the high demand for power in summer and winter causes the NSW and VIC electricity markets to show complex price variation.

Because the accuracy of electricity price forecasting varies across different periods, the proposed method is also applied to the four weeks of 2013 in VIC. From Table 5, the MAPE, WMAE, MAE and RMSE values obtained using the proposed method in all of the test weeks are significantly lower than those obtained using any of the other methods. For the summer week, the prediction performance of the proposed method is the most accurate; the MAPE value is 2.34%. The prediction accuracy is improved by 68.16%, 64.81%, 58.87% and 47.29% compared with the KELM, ARIMA, W-ARMA-BP and W-ARMA-LSSVM methods, respectively. The performance of the proposed approach is unsatisfactory in the autumn week. The main reason for this is that some outliers appear in this month. The MAPE for the proposed method reaches 5.88%. Many instances of abnormal power market behavior occurred during the winter test week, and the training and testing data include some unexpected electricity prices. The MAPE for the proposed method is 4.59%. However, the MAPEs for the KELM, ARIMA, W-ARMA-BP and W-ARMA-LSSVM during this period are 9.78%, 10.52%, 10.32% and 8.51%, respectively; thus, the proposed method still outperforms the other methods. For the spring week, the prediction performance of the proposed method is accurate. The MAPE value for the proposed method is 3.61%. Compared with the KELM, ARIMA, W-ARMA-BP and W-ARMA-LSSVM methods, the proposed model leads to reductions in the total MAPE of 52.37%, 71.77%, 46.60% and 43.77%, respectively. According to the U1 and U2 indicators, the comparison of the performance of the proposed model with that of the ARMA, KELM, W-ARMA-BP and W-ARMA-LSSVM models is clearer. The best prediction accuracy is still obtained from the proposed method. The above results prove that our idea of combining the WT, the ARMA and the KELM-SAPSO methods is rational and effective. The proposed method is able to capture the features of linearity and nonlinearity in electricity prices. Therefore, compared with other methods, the proposed method yields better prediction accuracy.

4.2.3. Case 3: Prediction of electricity prices in the Spanish market

To better verify the prediction efficiency of the proposed method, the developed model is applied to electricity price forecasting for the Spanish market. First, four weeks in January, April, July and October, representing four seasons in the year 2009, are studied. The MAPE, WMAE, and RMSE error measures and the U1 and U2 statistics indicators are used to evaluate the prediction performance of the proposed method. Second, the daily forecast prices for 2015 are also again measured to illustrate the forecasting efficiency of the proposed method. Similarly, in case 2, different price forecast techniques such as the individual ARMA and KELM

Table 4
Methods comparison of MAPE (%), U_1 and U_2 in Victoria, 2014.

Test day	Error	KELM	ARMA	W-ARMA-BP	W-ARMA-LSSVM	Proposed method
January	MAPE	12.10	12.91	11.92	8.70	**6.40**
	U_1	0.11	0.11	0.09	0.05	**0.04**
	U_2	0.85	0.98	0.70	0.43	**0.42**
February	MAPE	5.0	5.06	4.60	4.39	**2.86**
	U_1	0.03	0.03	0.03	0.03	**0.02**
	U_2	0.44	0.43	0.40	0.36	**0.26**
March	MAPE	15.20	18.92	13.28	12.58	**2.19**
	U_1	0.10	0.12	0.09	0.09	**0.01**
	U_2	1.11	1.41	1.05	1.01	**0.19**
April	MAPE	4.78	7.37	3.64	3.46	**2.58**
	U_1	0.03	0.04	0.02	0.02	**0.02**
	U_2	0.39	0.57	0.30	0.03	**0.21**
May	MAPE	5.63	9.17	4.82	4.57	**1.75**
	U_1	0.03	0.05	0.03	0.03	**0.01**
	U_2	0.45	0.70	0.45	0.43	**0.18**
June	MAPE	10.54	15.05	7.86	6.67	**3.02**
	U_1	0.07	0.11	0.05	0.48	**0.02**
	U_2	0.86	1.15	0.63	0.53	**0.24**
July	MAPE	14.80	33.00	12.94	11.77	**7.92**
	U_1	0.11	0.15	0.11	0.11	**0.06**
	U_2	0.73	1.51	0.70	0.70	**0.50**
August	MAPE	28.47	74.37	34.30	22.19	**9.32**
	U_1	1.99	0.24	0.16	0.11	**0.05**
	U_2	13.09	4.58	2.44	1.50	**0.64**
September	MAPE	22.63	36.2	15.19	10.45	**5.39**
	U_1	0.15	0.16	0.09	0.06	**0.03**
	U_2	1.79	2.32	1.00	0.74	**0.40**
October	MAPE	5.40	29.61	4.87	4.68	**1.95**
	U_1	0.03	0.18	0.03	0.03	**0.01**
	U_2	0.45	2.09	0.44	0.42	**0.17**
November	MAPE	12.44	16.84	11.63	11.05	**5.67**
	U_1	0.11	0.14	0.06	0.07	**0.03**
	U_2	0.97	1.26	0.89	0.84	**0.42**
December	MAPE	20.11	19.34	18.77	14.94	**6.76**
	U_1	0.15	0.15	0.18	0.13	**0.04**
	U_2	1.45	1.42	1.25	1.15	**0.53**

Table 5
Comparison of the MAPE (%), WMAE (%), MAE, RMSE, U1 and U2 in Victoria, 2013.

Test week	Error	KELM	ARMA	W-ARMA-BP	W-ARMA-LSSVM	Proposed method
Summer 28/01/-3/02/	MAPE	7.35	6.65	5.69	4.44	**2.34**
	WMAE	7.21	6.60	5.66	4.40	**2.33**
	MAE	3.27	2.94	2.54	1.96	**1.04**
	RMSE	3.81	3.52	2.95	2.26	**1.40**
	U1	0.05	0.04	0.04	0.03	**0.02**
	U2	0.15	0.12	0.18	0.16	**0.11**
Autumn 28/04/-4/05/	MAPE	9.78	10.52	10.32	8.51	**5.88**
	WMAE	10.46	11.34	11.28	9.23	**6.66**
	MAE	5.58	6.09	6.00	4.93	**3.55**
	RMSE	9.01	9.23	8.94	7.54	**6.26**
	U1	0.08	0.09	0.09	0.07	**0.06**
	U2	0.35	0.33	0.40	0.29	**0.25**
Winter 28/07/-3/08/	MAPE	10.03	14.93	7.54	6.57	**4.59**
	WMAE	10.28	14.74	7.72	6.73	**4.84**
	MAE	5.70	8.03	4.26	3.72	**2.68**
	RMSE	8.09	9.40	5.59	4.91	**3.89**
	U1	0.08	0.09	0.05	0.05	**0.04**
	U2	0.19	0.23	0.14	0.17	**0.12**
Spring 27/10/-3/11/	MAPE	7.58	12.79	6.76	6.42	**3.61**
	WMAE	7.16	12.07	6.36	5.96	**3.34**
	MAE	3.26	5.52	2.91	2.74	**1.52**
	RMSE	4.28	6.36	3.66	3.47	**2.18**
	U1	0.05	0.08	0.04	0.04	**0.03**
	U2	0.24	0.43	0.36	0.23	**0.16**

Table 6
Methods comparison of MAPE (%), WMAE (%), RMSE, U1 and U2 the Spanish electricity market, 2009.

Test week	Error	KELM	ARMA	W-ARMA-BP	W-ARMA-LSSVM	Proposed method
Winter 26/01/-1/02	MAPE	10.24	21.11	15.33	12.59	**5.87**
	WMAE	10.11	19.41	15.27	12.82	**6.12**
	RMSE	6.42	12.07	9.44	7.85	**4.76**
	U1	0.07	0.13	0.10	0.08	**0.05**
	U2	0.76	1.72	1.34	1.25	**0.51**
Spring 27/04/-3/05	MAPE	7.62	12.81	9.86	8.30	**4.01**
	WMAE	7.50	11.83	9.54	7.82	**4.07**
	RMSE	3.60	5.55	4.29	3.58	**2.55**
	U1	0.05	0.11	0.05	0.05	**0.03**
	U2	0.81	1.4	0.71	0.68	**0.48**
Summer 27/07-2/08/	MAPE	6.62	11.78	8.91	7.42	**3.91**
	WMAE	6.53	10.31	8.50	6.57	**3.86**
	RMSE	2.83	4.52	3.69	3.02	**1.75**
	U1	0.05	0.11	0.05	0.04	**0.03**
	U2	0.72	1.35	0.68	0.61	**0.44**
Autumn 26/10/-1/11	MAPE	7.94	14.51	8.72	7.48	**3.60**
	WMAE	7.90	13.98	8.76	7.08	**3.52**
	RMSE	3.71	6.13	4.32	3.33	**1.70**
	U1	0.05	0.06	0.05	0.04	**0.02**
	U2	0.99	1.00	0.67	0.60	**0.49**

Table 7
Methods comparison of MAPE (%), U1 andU2 the Spanish electricity market, 2015.

Test day	Error	KELM	ARMA	W-ARMA-BP	W-ARMA-LSSVM	Proposed method
January	MAPE	60.84	68.29	41.42	36.93	**12.86**
	U_1	0.42	0.46	0.28	0.17	**0.06**
	U_2	2.02	1.29	1.23	1.05	**0.52**
February	MAPE	11.93	11.71	9.20	7.45	**5.87**
	U_1	0.08	0.08	0.06	0.04	**0.03**
	U_2	0.65	0.63	0.47	0.40	**0.30**
March	MAPE	8.03	8.51	9.54	8.03	**4.75**
	U_1	0.04	0.05	0.06	0.04	**0.03**
	U_2	0.43	0.51	0.53	0.43	**0.31**
April	MAPE	5.60	7.03	4.66	4.62	**2.46**
	U_1	0.04	0.05	0.03	0.03	**0.02**
	U_2	0.33	0.42	0.26	0.27	**0.14**
May	MAPE	5.82	4.63	4.49	4.18	**3.47**
	U_1	0.05	0.04	0.03	0.03	**0.02**
	U_2	0.40	0.35	0.26	0.31	**0.25**
June	MAPE	5.83	6.77	5.64	5.25	**2.28**
	U_1	0.03	0.04	0.03	0.03	**0.01**
	U_2	0.32	0.36	0.34	0.34	**0.13**
July	MAPE	3.16	3.62	3.02	2.84	**1.86**
	U_1	0.02	0.02	0.02	0.02	**0.01**
	U_2	0.22	0.22	0.17	0.18	**0.13**
August	MAPE	4.66	7.47	5.91	5.62	**3.67**
	U_1	0.04	0.08	0.03	0.03	**0.03**
	U_2	0.28	0.51	0.33	0.32	**0.23**
September	MAPE	6.97	8.82	5.25	4.90	**2.36**
	U_1	0.04	0.05	0.03	0.03	**0.01**
	U_2	0.39	0.50	0.30	0.28	**0.14**
October	MAPE	14.21	14.92	12.83	10.07	**4.36**
	U_1	0.08	0.10	0.06	0.07	**0.03**
	U_2	0.80	0.82	0.76	0.58	**0.29**
November	MAPE	5.49	5.76	5.27	4.86	**3.07**
	U_1	0.03	0.03	0.03	0.03	**0.02**
	U_2	0.30	0.30	0.32	0.28	**0.19**
December	MAPE	6.13	8.36	5.72	5.13	**3.64**
	U_1	0.03	0.06	0.03	0.03	**0.02**
	U_2	0.37	0.47	0.32	0.27	**0.25**

methods and the hybrid W-ARMA-BP and W-ARMA-LSSVM methods are compared with the proposed method. The obtained results in this case are presented in Tables 6 and 7.

From Table 6, we can observe that the proposed method outperforms all other comparative methods. The smallest WMAE, MAPE and RMSE values occur in the week of October; the largest occurs

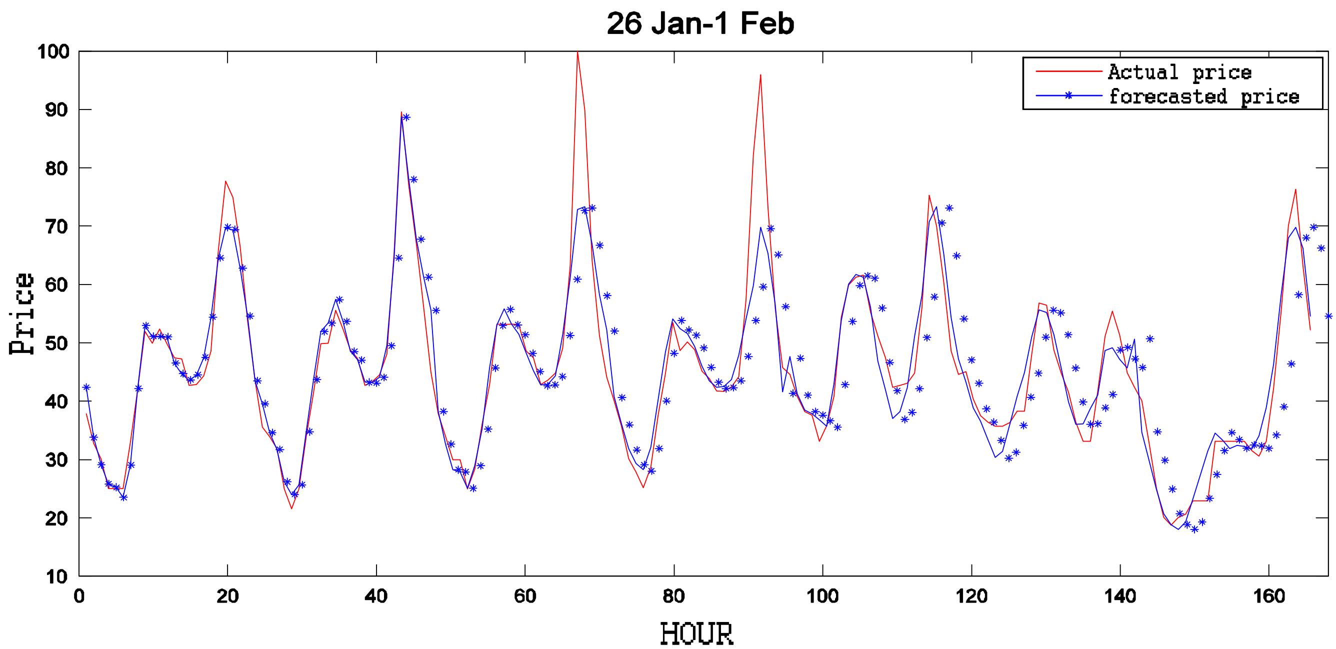

Fig. 4. Actual electricity price and forecasting values for 26 Jan-2 Feb for the Spanish electricity market.

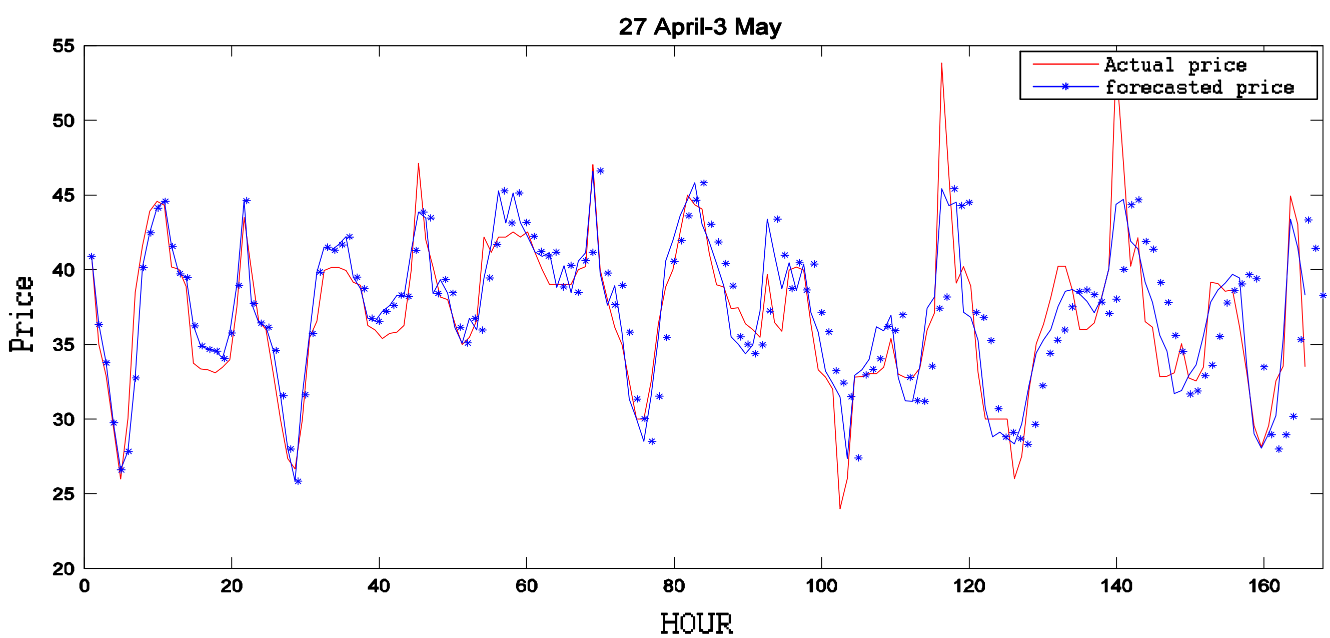

Fig. 5. Actual electricity price and forecasting values for 27 April-3 May for the Spanish electricity market.

in the week of January, mainly due to the occurrence of many spikes that week. Another reason for this trend is that the winter high load demand causes unstable price behavior. The average WMAE of the proposed method in the four weeks is 4.39, which is appropriate for electricity price forecasting for the Spanish market. From the MAPE list in Table 6, we can also observe the electricity price variation pattern for the spring, summer, fall and winter seasons of 2009 in the Spanish market. The highest MAPE appears during the winter peak demand, and the lowest MAPE appears in the autumn low peak period. The main reason for this is that some uncertainties and market strategies occurred during the summer power market. However, when the proposed method is used, the prediction results are more accurate and reliable than the results obtained using the single ARMA or KELM method or either of the hybrid methods W-ARMA-BP and W-ARMA-LSSVM. The scatter plots of the actual and predicted prices of 2009 for the Spanish market are presented in Figs. 4–7. It can be seen that the results obtained from the proposed method agree with the actual electricity price except for a few spikes.

It can also be seen from Table 7 that the prediction accuracy of the proposed method is the best of all the compared methods. On the January test day, some other energy sources such as wind and gas are adopted due to the market strategy, and a few lower prices appear in the morning and at night. Thus, the prediction

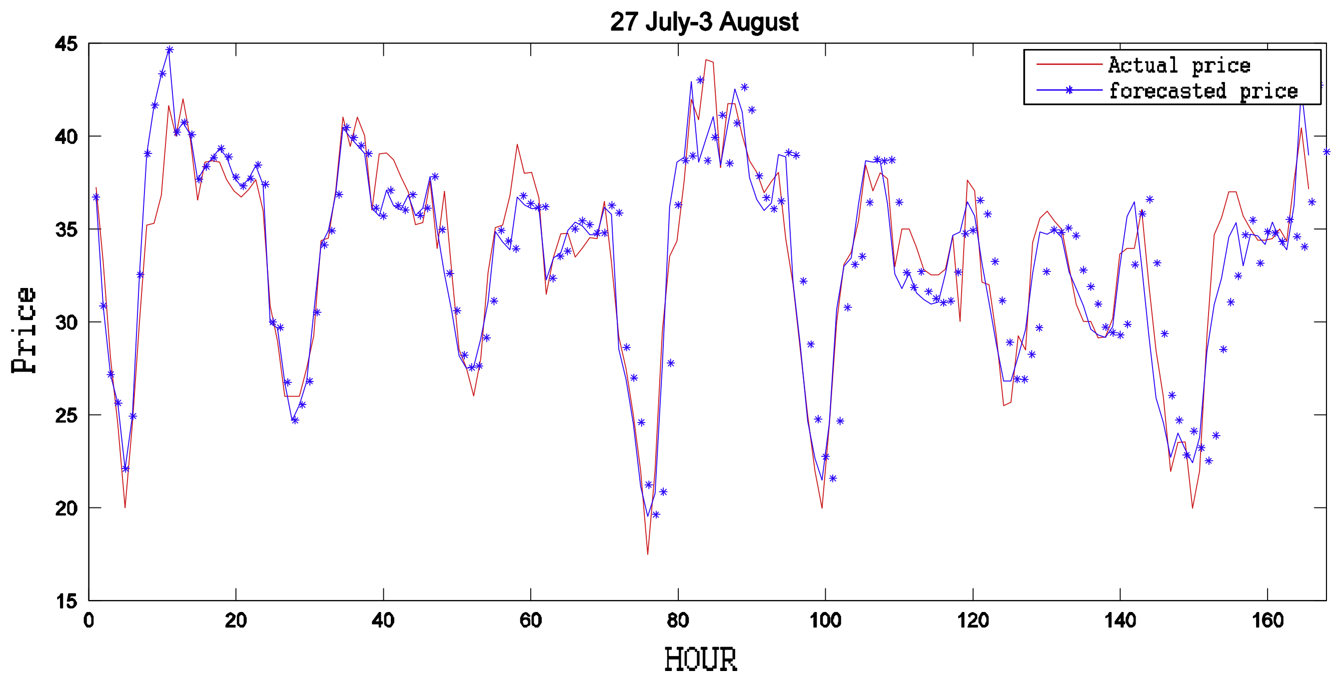

Fig. 6. Actual electricity price and forecasting values for 27 July-3 August for the Spanish electricity market.

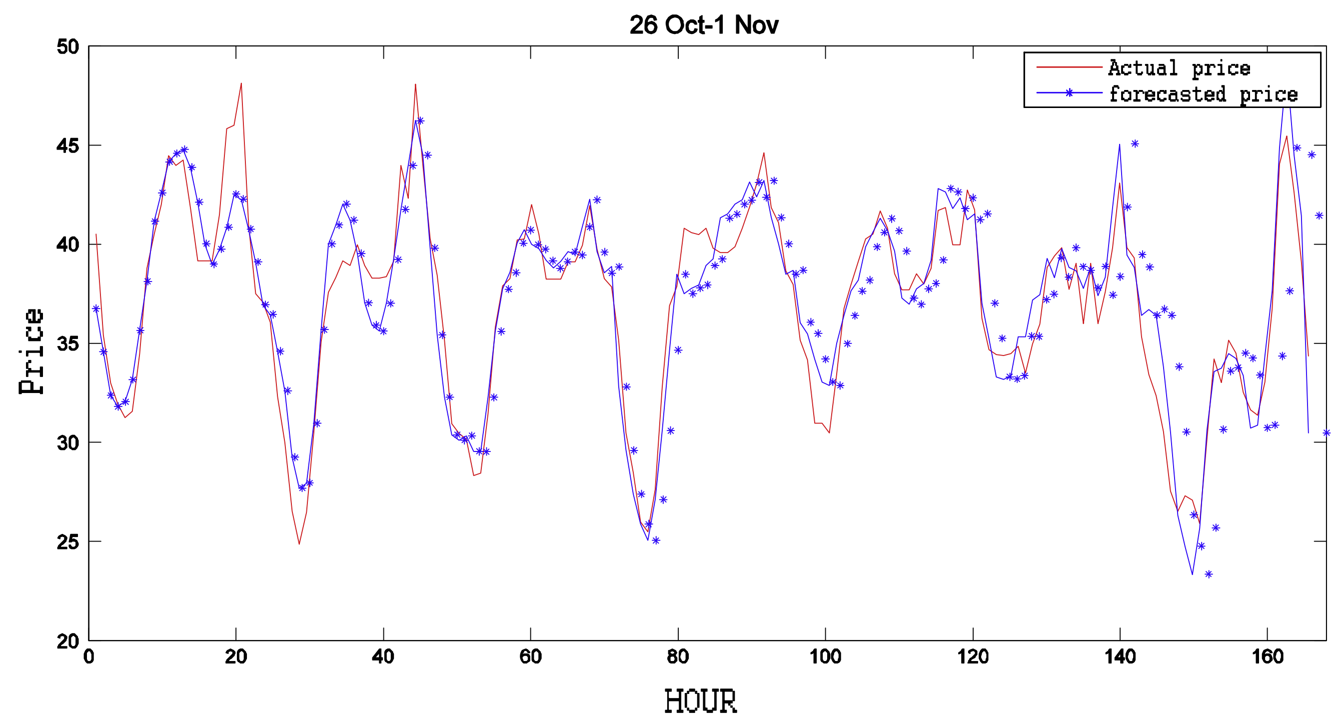

Fig. 7. Actual electricity price and forecasting values for 26 Oct-1Nov for the Spanish electricity market.

results for January 31 are not accurate. The MAPE values for the KELM model range from 60.84% to 3.16%; the ARMA values range from 68.29% to 3.62%, those for the W-ARMA-BP method range from 41.42% to 3.02% and those for the W-ARMA-LSSVM method range from 36.93% to 2.84%, whereas the proposed method yields values ranging from 1.86% to 12.86%. The results show that by using a hybrid of the WT, ARMA and SAPSO-KELM methods, the prediction errors can be significantly reduced. The developed methods can be used to obtain accurate predictions for the Spanish electricity market.

5. Conclusions

Electricity price forecasting has become indispensable to power consumers and producers in the deregulated electricity market. Developing an effective and accurate forecasting model has thus become very important. In this paper, a new hybrid electricity price prediction method combining WT with ARMA, KELM and SAPSO is proposed. After the original price series is decomposed by WT, the ARMA model predicts the linear part of the electricity price, while the KELM model together with SAPSO forecasts the nonlinear part. The SAPSO algorithm is adopted as a training process of the KELM to adjust the penalty factor and kernel parameters of the KELM. To verify the predictive ability of the proposed method, this hybrid model is applied to price forecasting of the PJM, Australian and Spanish power markets. The three cases discussed in this paper show that the hybrid model outperforms the individual ARMA and KELM models. Moreover, due to the fast learning speed and stable generalization capabilities possessed by SAPSO-KELM, the proposed model also has better prediction accuracy than other hybrid methods, such as those that involve the integration of WT with ARMA and BPNN or the integration of WT with ARMA and LSSVM. All the empirical results suggest that the proposed method provides an effective and feasible approach to electricity price forecasting.

The prediction performance of the proposed hybrid method can be attributed to three factors. First, wavelet transform produces a good local representation of the original electricity price in the time and frequency domains. Second, the ARMA model can predict the linear component of electricity prices from the stationary series, and KELM can more easily capture the nonlinear component of electricity prices. Third, the SAPSO algorithm can help identify suitable kernel parameters of KELM that can be used to obtain more accurate predictions. This paper focused solely on improving the prediction accuracy by a hybrid approach. Although some prices are not matched well, the proposed method do improved the prediction accuracy for electricity price. In the future, several influential factors, such as load and weather, should be taken into consideration when the hybrid method is used in the electricity price forecasting.

References

[1] Mandal P, Haque AU, Meng J, Martinez R, Srivastava AK. A hybrid intelligent algorithm for short-term energy price forecasting in the Ontario market. In: 2012 IEEE power and energy society general meeting. IEEE; 2012. p. 1–7.

[2] Chan S-C, Tsui KM, Wu H, Hou Y, Wu Y-C, Wu FF. Load/price forecasting and managing demand response for smart grids: methodologies and challenges. Signal Process Mag, IEEE 2012;29:68–85.

[3] Cerjan M, Krželj I, Vidak M, Delimar M. A literature review with statistical analysis of electricity price forecasting methods. In: EUROCON, 2013 IEEE. IEEE; 2013. p. 756–63.

[4] Weron R. Electricity price forecasting: a review of the state-of-the-art with a look into the future. Int J Forecast 2014;30:1030–81.

[5] Aggarwal SK, Saini LM, Kumar A. Electricity price forecasting in deregulated markets: a review and evaluation. Int J Electr Power Energy Syst 2009;31:13–22.

[6] Lin W-M, Gow H-J, Tsai M-T. An enhanced radial basis function network for short-term electricity price forecasting. Appl Energy 2010;87:3226–34.

[7] Anbazhagan S, Kumarappan N. Day-ahead deregulated electricity market price forecasting using recurrent neural network. Syst J, IEEE 2013;7:866–72.

[8] Panapakidis IP, Dagoumas AS. Day-ahead electricity price forecasting via the application of artificial neural network based models. Appl Energy 2016;172:132–51.

[9] Yan X, Chowdhury NA. Mid-term electricity market clearing price forecasting: a multiple SVM approach. Int J Electr Power Energy Syst 2014;58:206–14.

[10] Hung Y-C, Lin F-J, Hwang J-C, Chang J-K, Ruan K-C. Wavelet fuzzy neural network with asymmetric membership function controller for electric power steering system via improved differential evolution. IEEE Trans Power Electron 2015;30:2350–62.

[11] Anbazhagan S, Kumarappan N. Day-ahead deregulated electricity market price forecasting using neural network input featured by DCT. Energy Convers Manage 2014;78:711–9.

[12] Khosravi A, Nahavandi S, Creighton D. Quantifying uncertainties of neural network-based electricity price forecasts. Appl Energy 2013;112:120–9.

[13] Keles D, Scelle J, Paraschiv F, Fichtner W. Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks. Appl Energy 2016;162:218–30.

[14] Taylor JA, Mathieu JL, Callaway DS, Poolla K. Price and capacity competition in balancing markets with energy storage. arXiv preprint arXiv: 150306851; 2015.

[15] Huang G-B, Zhu Q-Y, Siew C-K. Extreme learning machine: a new learning scheme of feedforward neural networks. In: Neural networks, 2004 Proceedings 2004 IEEE international joint conference on. IEEE; 2004. p. 985–90.

[16] Chen X, Dong ZY, Meng K, Xu Y, Wong KP, Ngan H. Electricity price forecasting with extreme learning machine and bootstrapping. IEEE Trans Power Syst 2012;27:2055–62.

[17] Wan C, Xu Z, Pinson P, Dong ZY, Wong KP. Probabilistic forecasting of wind power generation using extreme learning machine. IEEE Trans Power Syst 2014;29:1033–44.

[18] Lu C-J, Shao YE. Forecasting computer products sales by integrating ensemble empirical mode decomposition and extreme learning machine. Mathe Probl Eng 2012;2012.

[19] Wan C, Xu Z, Wang Y, Dong ZY, Wong KP. A hybrid approach for probabilistic forecasting of electricity price. IEEE Trans Smart Grid 2014;5:463–70.

[20] Li S, Wang P, Goel L. Short-term load forecasting by wavelet transform and evolutionary extreme learning machine. Electric Power Syst Res 2015;122:96–103.

[21] Teo T, Logenthiran T, Woo W. Forecasting of photovoltaic power using extreme learning machine. In: Smart grid technologies-Asia (ISGT ASIA), 2015 IEEE innovative. IEEE; 2015. p. 1–6.

[22] Wang X, Han M. Online sequential extreme learning machine with kernels for nonstationary time series prediction. Neurocomputing 2014;145:90–7.

[23] Yu Q, Miche Y, Séverin E, Lendasse A. Bankruptcy prediction using extreme learning machine and financial expertise. Neurocomputing 2014;128:296–302.

[24] Deo RC, Şahin M. Application of the extreme learning machine algorithm for the prediction of monthly Effective Drought Index in eastern Australia. Atmos Res 2015;153:512–25.

[25] Yin J-C, Zou Z-J, Xu F, Wang N-N. Online ship roll motion prediction based on grey sequential extreme learning machine. Neurocomputing 2014;129:168–74.

[26] Yadav B, Ch S, Mathur S, Adamowski J. Discharge forecasting using an Online Sequential Extreme Learning Machine (OS-ELM) model: a case study in Neckar River, Germany. Measurement 2016;92:433–45.

[27] Xia M, Zhang Y, Weng L, Ye X. Fashion retailing forecasting based on extreme learning machine with adaptive metrics of inputs. Knowl-Based Syst 2012;36:253–9.

[28] Li B, Rong X, Li Y. An improved kernel based extreme learning machine for robot execution failures. Sci World J 2014;2014.

[29] Tan Z, Zhang J, Wang J, Xu J. Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models. Appl Energy 2010;87:3606–10.

[30] Mandal P, Haque AU, Meng J, Srivastava AK, Martinez R. A novel hybrid approach using wavelet, firefly algorithm, and fuzzy ARTMAP for day-ahead electricity price forecasting. IEEE Trans Power Syst 2013;28:1041–51.

[31] Zhang J, Tan Z. Day-ahead electricity price forecasting using WT, CLSSVM and EGARCH model. Int J Electr Power Energy Syst 2013;45:362–8.

[32] Shayeghi H, Ghasemi A. Day-ahead electricity prices forecasting by a modified CGSA technique and hybrid WT in LSSVM based scheme. Energy Convers Manage 2013;74:482–91.

[33] Shrivastava NA, Panigrahi BK. A hybrid wavelet-ELM based short term price forecasting for electricity markets. Int J Electr Power Energy Syst 2014;55:41–50.

[34] Shafie-Khah M, Moghaddam MP, Sheikh-El-Eslami M. Price forecasting of day-ahead electricity markets using a hybrid forecast method. Energy Convers Manage 2011;52:2165–9.

[35] Nguyen HT, Nabney IT. Short-term electricity demand and gas price forecasts using wavelet transforms and adaptive models. Energy 2010;35:3674–85.

[36] Voronin S, Partanen J. Forecasting electricity price and demand using a hybrid approach based on wavelet transform, ARIMA and neural networks. Int J Energy Res 2014;38:626–37.

[37] Abedinia O, Amjady N. Day-ahead price forecasting of electricity markets by a new hybrid forecast method. Model Simulat Electr Electron Eng 2015;1:1–8.

[38] Zhang J, Tan Z, Li C. A novel hybrid forecasting method using GRNN combined with wavelet transform and a GARCH model. Energy Sources Part B: Econ, Plann, Policy 2015;10:418–26.

[39] Rojas I, Valenzuela O, Rojas F, Guillén A, Herrera LJ, Pomares H, et al. Soft-computing techniques and ARMA model for time series prediction. Neurocomputing 2008;71:519–37.

[40] Bracewell R. The fourier transform and its applications, vol. 5. New York; 1965.

[41] Morlet J, Arens G, Fourgeau E, Glard D. Wave propagation and sampling theory-Part I: complex signal and scattering in multilayered media. Geophysics 1982;47:203–21.

[42] Mallat SG. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Mach Intell 1989;11:674–93.

[43] Conejo AJ, Plazas M, Espinola R, Molina AB. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. IEEE Trans Power Syst 2005;20:1035–42.

[44] Box GE, Jenkins GM, Reinsel GC. Time series analysis: forecasting and control. John Wiley & Sons; 2011.

[45] Dickey DA, Fuller WA. Distribution of the estimators for autoregressive time series with a unit root. J Am Stat Assoc 1979;74:427–31.

[46] Huang G-B, Chen L. Convex incremental extreme learning machine. Neurocomputing 2007;70:3056–62.

[47] Huang G-B, Zhu Q-Y, Siew C-K. Extreme learning machine: theory and applications. Neurocomputing 2006;70:489–501.

[48] Huang G-B, Wang DH, Lan Y. Extreme learning machines: a survey. Int J Mach Learn Cybern 2011;2:107–22.
[49] Hoerl AE, Kennard RW. Ridge regression: biased estimation for nonorthogonal problems. Technometrics 1970;12:55–67.
[50] Penrose R. A generalized inverse for matrices. Mathematical proceedings of the Cambridge philosophical society. Cambridge Univ Press; 1955. p. 406–13.
[51] Huang G-B, Chen L. Enhanced random search based incremental extreme learning machine. Neurocomputing 2008;71:3460–8.
[52] Burges CJ. A tutorial on support vector machines for pattern recognition. Data Min Knowl Disc 1998;2:121–67.
[53] Eberhart RC, Kennedy J. A new optimizer using particle swarm theory. In: Proceedings of the sixth international symposium on micro machine and human science: New York, NY; 1995. p. 39–43.
[54] Shi Y, Eberhart R. A modified particle swarm optimizer. In: Evolutionary computation proceedings, 1998 IEEE world congress on computational intelligence, the 1998 IEEE international conference on. IEEE; 1998. p. 69–73.
[55] Han F, Ling Q-H. A new approach for function approximation incorporating adaptive particle swarm optimization and a priori information. Appl Math Comput 2008;205:792–8.
[56] Mandal P, Senjyu T, Urasaki N, Funabashi T, Srivastava AK. A novel approach to forecast electricity price for PJM using neural network and similar days method. IEEE Trans Power Syst 2007;22:2058–65.
[57] Pindoriya N, Singh S, Singh S. An adaptive wavelet neural network-based energy price forecasting in electricity markets. IEEE Trans Power Syst 2008;23:1423–32.
[58] AEMO. <http://www.aemo.com.au/Electricity/Data/Price-and-Demand/Aggregated-Price-and-Demand-Data-Files>.
[59] Spanish OMEL. <http://www.omelholding.es/omel-holding/>.
[60] Zhu W, Miao J, Hu J, Qing L. Vehicle detection in driving simulation using extreme learning machine. Neurocomputing 2014;128:160–5.
[61] Zhong H, Miao C, Shen Z, Feng Y. Comparing the learning effectiveness of BP, ELM, I-ELM, and SVM for corporate credit ratings. Neurocomputing 2014;128:285–95.
[62] Kaytez F, Taplamacioglu MC, Cam E, Hardalac F. Forecasting electricity consumption: a comparison of regression analysis, neural networks and least squares support vector machines. Int J Electr Power Energy Syst 2015;67:431–8.