Applied Energy 87 (2010) 3606–3610

Contents lists available at ScienceDirect

Applied Energy

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

Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models

Zhongfu Tan [a], Jinliang Zhang [a,*], Jianhui Wang [b], Jun Xu [a]

[a] North China Electric Power University, Beijing 102206, China
[b] Argonne National Laboratory, Argonne, IL 60439, USA

ARTICLE INFO

Article history:
Received 12 November 2009
Received in revised form 7 April 2010
Accepted 10 May 2010
Available online 11 June 2010

Keywords:
Price forecasting
Wavelet transform
ARIMA
GARCH

ABSTRACT

This paper proposes a novel price forecasting method based on wavelet transform combined with ARIMA and GARCH models. By wavelet transform, the historical price series is decomposed and reconstructed into one approximation series and some detail series. Then each subseries can be separately predicted by a suitable time series model. The final forecast is obtained by composing the forecasted results of each subseries. This proposed method is examined on Spanish and PJM electricity markets and compared with some other forecasting methods.

1. Introduction

In a competitive electricity market, day ahead price forecasting is a key issue for all the market participants. An accurate day ahead price forecasting in the spot market helps the power suppliers to adjust their bidding strategies to achieve the maximum benefit and on the other hand, consumers can derive a plan to maximize their utilities using the electricity purchased from the pool, or use self production capability to protect themselves against high prices [1]. However, price series usually has complex features such as nonstationarity, nonlinearity and high volatility, which makes the price forecasting turns out to be very difficult.

Importance of price forecasting on one hand and its complexity on the other hand, researchers have done much work in this area and proposed many methods. Among these methods, two widely-used approaches for price forecasting are Artificial Neural Networks (ANN) techniques and time series methods. ANN methods can be found in Refs. [2–4]. Szkuta et al. [2] proposed a three-layered ANN with back-propagation using the Victorian electricity market data for training and testing. Wang et al. [3] proposed a neural-network-based approach to predict system marginal prices, taking into account weekend and public holidays. Zhang et al. [4] presented a cascaded neural network structure for market-clearing price (MCP) prediction in the New England market. Other techniques regarding model used, preprocessing employed have been also presented in Refs. [5–10].

Another popular technique of price forecasting is time series analysis. Autoregressive integrated moving average (ARIMA) models [11–13], multivariate transfer function (TF) models [11,14], dynamic models [11], and generalized autoregressive conditional heteroskedasticity (GARCH) models [15] have also been proposed. Besides, some more efficient time series models such as ARIMA–EGARCH and GIGARCH models are recently presented in Refs. [16,17]. In order to provide a robust price forecasting method, wavelet transform have been utilized because they can produce a good local representation of the signal in both time and frequency domains [1,13]. However, they suffer the following defects: in Ref. [1], the offline training process of the methods is complicated and time consuming. Besides, the methods are not applicable when the experts are not available, thus the time series of market-clearing prices can be break in some manner. In Ref. [13], not all the subseries are suited to be predicted by ARIMA model, which cannot capture the characteristics of high volatility. To overcome the above problems, a new hybrid method based on wavelet transform combined with ARIMA and GARCH models is proposed in this paper.

The main contribution of this paper is to use wavelet transform to decompose and reconstruct an ill-behaved price series into a set of better-behaved constitutive series. Then each constitutive series is separately predicted by a suitable time series model according to their features. This hybrid method can capture the complex characteristics of nonstationarity, nonlinearity and high volatility. To the best of our knowledge, this method has never been presented in the literature.

The remainder of the paper is organized as follows: Section 2 outlines the proposed method. Section 3 presents numerical

* Corresponding author. Tel.: +86 010 51963749; fax: +86 010 80796904.
E-mail address: zhangjinliang1213@163.com (J. Zhang).

0306-2619/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.apenergy.2010.05.012

results of the simulations and Section 4 provides some relevant conclusions.

2. The proposed method

Before proposing the hybrid prediction method, the wavelet transform and two time series models (ARIMA and GARCH) are introduced.

2.1. Wavelet transform

The basic concept in wavelet analysis begins with the selection of a proper wavelet (mother wavelet) and then performing an analysis on its translated and dilated versions. A wavelet can be defined as a function $\psi(t)$ with a zero mean [18].

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0 \tag{1}$$

A signal can be decomposed into many series of wavelets with different scales a and translation b:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \tag{2}$$

Thus, the wavelet transform of a signal $f(t)$ at translation b and scale a is defined by:

$$wf(b,a) = \frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} f(t)\psi\left(\frac{t-b}{a}\right)dt \tag{3}$$

The original signal $f(t)$ can be reconstructed by inverse wavelet transform:

$$f(t) = \int_0^{\infty}\int_{-\infty}^{+\infty} \frac{1}{a^2} wf(b,a)\psi_{b,a}(t)dbda \tag{4}$$

2.2. Arima

The ARIMA model is widely used in the areas of nonstationary time series forecasting, which can be written as:

$$\phi(B)(1-B)^d X_t = \theta(B)\varepsilon_t \tag{5}$$

where X_t represents a nonstationary time series at time t, ε_t is a white noise (zero mean and constant variance), d is the order of differencing, B is a backward shift operator defined by $BX_t = X_{t-1}$, $\phi(B)$ is the autoregressive operator defined as: $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$, and $\theta(B)$ is the moving average operator defined as: $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$.Generally, this method includes four phases. In the first step, a general ARIMA formulation is selected to model the electricity price data. If the time series contains multiple seasons, the form $(1 - B^s)$ should be included in the model, s is the order of seasonality. In the second step, a logarithmic transformation is applied to the time series to produce a more stable variance. After the underlying process is accepted as being stationary, the structure of $\phi(B)$ and $\theta(B)$ must be selected by autocorrelation and partial autocorrelation plots The third step is parameter estimation, which involves maximizing a likelihood function for the available data. The last step is diagnostic checking. If the residual term is a white noise process, then the model can be used for forecasting purposes. Otherwise, the process should be repeated until an adequate model is found. More details can be found in Ref. [12].

2.3. GARCH

As previously mentioned, electricity price series can be highly volatile. So a proper GARCH model can be used to predict electricity prices, since the model considers moments of a time series as variants. Generally, a GARCH(p,q) model is expressed as:

$$\varepsilon_t/\psi_{t-1} \sim N(0,\delta_t) \tag{6}$$

$$\varepsilon_t = \sqrt{\delta_t}u_t, \quad u_t \sim N(0,1) \tag{7}$$

$$\delta_t = \omega + \sum_{i=1}^{p}\alpha_i\varepsilon_{t-i}^2 + \sum_{j=1}^{q}\beta_j\delta_{t-j} \tag{8}$$

where $p > 0,\ q \geqslant 0,\ \omega > 0,\ \alpha_i \geqslant 0\ (i = 1,2,\ldots p),\ \beta_j \geqslant 0\ (j = 1, 2,\ldots q)$.

In general, GARCH(p,q) model is more suitable to capture the dynamics of a time series' conditional variance. The model also includes four phases: data preparation, model identification, parameter estimation and diagnostic checking. More details can be seen in Ref. [15].

Time series models have become popular in price forecasting. However, directly applying them to price prediction does not necessarily produce the best result. The combination of related models like the Wavelet–GARCH–ARIMA method described below has proven to have a higher accuracy.

2.4. Proposed method

Electricity price series usually has complex features such as nonstationarity, nonlinearity and high volatility. These features are often the most important parts of the original price series and must be taken into account. To track the complex behavior of electricity price series, wavelet transform has been utilized because it can decompose and reconstruct an ill-behaved price series into a set of better-behaved constitutive series. Thus, each constitutive series can be separately predicted by a suitable time series model according to their features.

In this paper, Daubechies wavelet of order 4 is used as the mother wavelet. This wavelet offers an appropriate trade off between wave-length and smoothness, resulting in an appropriate behavior for price forecast. Besides, we consider three decomposition levels, since it describes the price series in a more thorough and meaningful way than others [1]. Thus, the original price series is decomposed and reconstructed into one approximation series (general trend component) and three detail series (high frequency component). The approximation series is the low frequency component of the original price signal and follows trends of the signal [1]. In Ref. [13], the future values of approximation series are predicted by ARIMA model. When using this model, the error term includes zero mean and constant variance is assumed. However, this assumption has to be discarded most of the time. Our research has found that the heteroskedastic error specification is strongly supported by the data applied in this paper. So an ARIMA process with GARCH error components is applied to predict the future values of approximation series. It has been described that detail series is more related to the noisy part of the original signal. Thus, detail series is suited to be predicted by the GARCH model, which can capture the features of high volatility. In this way, the combinational forecast method becomes an effective technique to predict electricity prices. The procedure is outlined in Fig. 1, and a detailed explanation of above procedure is given below.

(1) By wavelet transform, an electricity price series P_t is decomposed into one approximation series defined as $a3_t$ and three detail series defined as $b1_t, b2_t, b3_t$, respectively.
(2) To duplicate the original series, it is important to reconstruct the approximate and detail series. By wavelet reconstruction, series of $a3_t, d1_t, d2_t, d3_t$ are denominated as

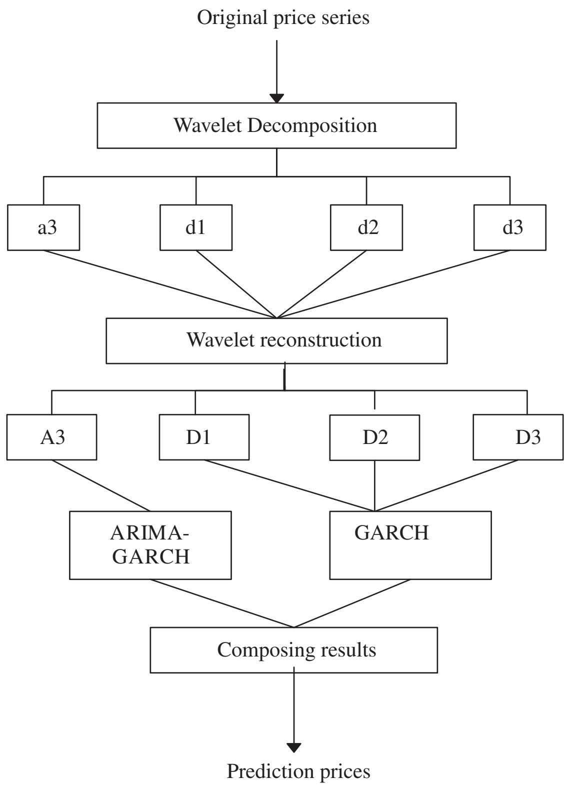

Fig. 1. Procedure of the proposed method.

$A3_t, D1_t, D2_t, D3_t$. So the original series can be defined as follows with less loss:

$$P_t = A3_t + D1_t + D2_t + D3_t \tag{9}$$

(3) As for price prediction, an ARIMA–GARCH model is built to forecast the future vaules of $A3_t$, while the GARCH model is used for detail series (including $D1_t$, $D2_t$ and $D3_t$).
(4) The original price prediction is obtained by composing the forecasted results of series $A3_t, D1_t, D2_t, D3_t$, which can be represented as:

$$\hat{P}_t = \hat{A3}_t + \hat{D1}_t + \hat{D2}_t + \hat{D3}_t \tag{10}$$

where $\hat{P}_t$ is the forecasted price, $\hat{A3}_t$, $\hat{D1}_t$, $\hat{D2}_t$, $\hat{D3}_t$ are the forecasted results of series $A3_t, D1_t, D2_t, D3_t$.

3. Numerical results

Generally, when there is no transmission congestion, market-clearing price (MCP) is the only price for the entire system. However, when there is congestion, the locational marginal price (LMP) could be employed [1]. Thus, the effectiveness of this proposed method is demonstrated on MCP prediction in Spanish electricity market for the year 2002 and for LMP forecasting in PJM electricity market for the year 2006.

3.1. Prediction of MCP in Spanish market

For the sake of a fair comparison, the fourth week of February, May, August, and November are selected for the winter, spring, summer, and fall seasons, respectively, which is the test period considered in Refs. [19–21]. To assess the prediction capacity of this proposed method, two types of error measures have been taken from Ref. [22] to avoid problems caused by spiking and near-zero prices.

The weekly absolute error is expressed as:

$$\mathrm{MAPE}_{week} = \frac{1}{168}\sum_{t=1}^{168}\frac{\left|P_t - \hat{P}_t\right|}{P_{week}-} \tag{11}$$

where P_t and $\hat{P}_t$ are actual and forecast price of hour t, respectively.

$$P_{week}- = \frac{1}{168}\sum_{t=1}^{168} P_t.$$

An index of the uncertainly of a model is the variability of what is still unexplained after fitting the model, which can be measured by the estimation of the variance of the error term. The smaller the variance, the more precise is the prediction of prices. The weekly error variance is expressed as:

$$EV_{week} = \frac{1}{168}\sum_{t=1}^{168}\left(\frac{\left|P_t - \hat{P}_t\right|}{P_{week}-} - \mathrm{MAPE}_{week}\right)^2 \tag{12}$$

Now the proposed method is compared with some of the most recent price forecast techniques [19–21]. Obtained results for 4 weeks are presented in Tables 1 and 2.

As can be seen from Table 1, the proposed method outperforms all other examined techniques. The average of the MAPE values of the proposed method is considerably less than all other techniques. However, the forecasting performance for test week of summer and fall seasons are poor than that of the winter and spring week for all represented methods. The reason is that Spanish market is more unstable in respect to price behavior in summer and fall seasons than winter and spring seasons [12]. Table 2 also shows that the variance value of the proposed method is less than all other methods. This finding demonstrates that the prediction accuracy can be significantly improved by the proposed method.

Since price prediction accuracy varies across different test periods, the proposed method is applied for the 12 months of 2002 in Spanish market and compared with some other methods, such as ARIMA, ARIMA–GARCH and WT–ARIMA. The last week of every month is studied to validate the performance of the proposed method, while the historical data is 50 days previous to the beginning of the weekly test period. Results of these comparisons are shown in Table 3.

From Table 3, it is evident that the accuracy of the proposed method is better than the ARIMA, ARIMA–GARCH and WT–ARIMA methods. It is clear that the ARIMA predictions are less accurate

Table 1
MAPE (%) for the four test week of Spanish electricity market in year 2002.

Test week	FNN [19]	CNN [20]	Mixed [21]	Proposed
Winter	4.62	4.21	4.22	0.63
Spring	5.30	4.76	4.39	0.65
Summer	9.84	6.01	5.55	1.19
Fall	10.32	5.88	5.66	2.18
Average	7.52	5.22	4.95	1.16

Table 2
Error variance for 4 weeks of the Spanish electricity market in 2002.

Test week	FNN [19]	CNN [20]	Mixed [21]	Proposed
Winter	0.0018	0.0014	0.0015	0.0002
Spring	0.0019	0.0033	0.0029	0.0002
Summer	0.0092	0.0045	0.0039	0.0009
Fall	0.0088	0.0048	0.0052	0.0008
Average	0.0054	0.0035	0.0034	0.0005

Table 3
MAPE (%) for 12 weeks of the Spanish market in 2002.

	ARIMA	ARIMA–GARCH	WT–ARIMA	Proposed
January	11.15	10.00	5.71	2.14
February	4.72	4.44	4.39	1.09
March	5.15	5.84	3.74	0.64
April	14.26	12.3	4.07	0.92
May	7.92	5.30	5.64	0.82
June	6.41	5.31	4.22	1.02
July	10.37	8.70	9.43	2.09
August	12.14	7.51	8.74	1.55
September	13.89	7.53	10.45	2.01
October	9.97	9.54	6.58	1.63
November	13.93	9.98	5.28	2.42
December	17.43	17.39	8.22	2.99
Average	10.61	8.65	6.37	1.61

than the proposed method with the average MAPE increasing from 1.61% to 10.61%. This result indicates that the wavelet transform produces constitutive series that can be predicted more accurately than the original price series, and the GARCH model can capture the hourly price volatility. Besides, the proposed method compared with ARIMA–GARCH methods shows better results, which also confirm the intuition that the wavelet transform can produce a good local representation of the price in both time and frequency domains. Although the performance of the WT–ARIMA technique is better than the performance of ARIMA and ARIMA–GARCH methods, however, the results obtained from it are worse as compared to the proposed method. The experimental results verify the assumption that there is heteroskedasticity in the approximation series. Additionally, detail series contains high frequency components of the original price, which is suitable to be predicted by GARCH model. All these comparisons reveal the forecast capability of the proposed technique. Thus, it is seen that the proposed method can give acceptable results in a long run for the whole year.

3.2. Prediction of LMP in PJM market

To illustrate the behavior of the proposed method, results comprising 4 weeks corresponding to the four seasons of year 2006 are also presented.

For the sake of a fair comparison, the four considered weeks are Feb. 15 to Feb. 21, May 15 to May 21, August 15 to August 21 and Nov. 15 to Nov. 21. We consider their 50 days ago as the training period. Obtained results of the proposed method and three other techniques for LMP prediction are shown in Table 3. Here, MLP with LM learning algorithm proposed in Ref. [23], PCA + CNN proposed in Ref. [20], and WT with NN proposed in Ref. [1] are selected as the compared techniques. The results of these techniques have been quoted from Refs. [1,20]. In this table, weekly mean error (WME) is error indicator:

$$WME = \frac{1}{168}\sum_{i=1}^{168}\frac{|P_i - P_i'|}{P_i} \tag{13}$$

where P_i and P_i' are actual and forecast LMP of hour i, respectively.

As seen from Table 4, the proposed method has considerably lower WME for LMP prediction in the PJM market. The average of WME values for MLP + LM, PCA + CNN, and WT + NN is 9.67%, 8.12%, and 4.57%, respectively, which are more than 0.74% obtained using the proposed method. The MLP + LM and PCA + CNN give maximum errors due to their limitations in dealing with the complex behavior of electricity price.

For a more detailed representation of the performance of the proposed method, its results for the period of the high volatility are presented in Table 5. High volatility in PJM market can be observed on August 4 with the prices varying from a minimum of 44.17$/MWh to a maximum of 157.43$/MWh. In this table, the error can be defined as follows:

$$Error = \frac{1}{t}\sum_{i=1}^{t}\frac{|P_i - P_i'|}{P_i} \tag{14}$$

As can be seen that the errors has an average value of 1.00% on August 4, where the minimum error is 0.60% occurs in hour 10, and the maximum error is 1.16% occurs in hour 1. Fig. 2 is quite illustrative in showing how the proposed method works for the day of August 4. It can be observed that the proposed technique is also able to follow the trend of the real price, even when the price

Table 4
WME (%) for the four test week of PJM electricity market in year 2006.

Test week	MLP + LM	PCA + CNN	WT + NN	Proposed
Winter	9.82	8.61	4.44	0.73
Spring	8.87	7.34	4.31	0.60
Summer	10.43	8.17	4.78	0.75
Fall	9.54	8.36	4.75	0.86
Average	9.67	8.12	4.57	0.74

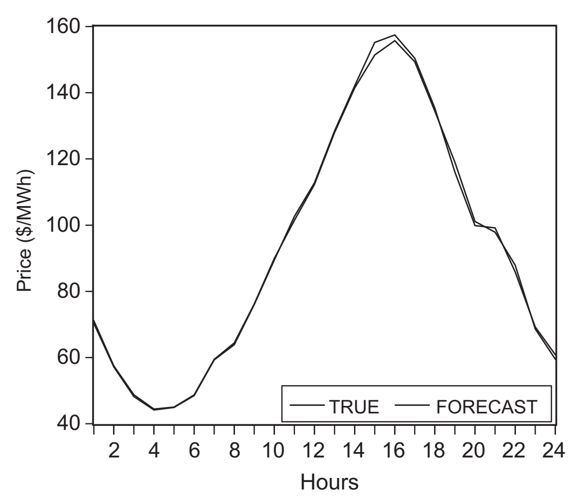

Fig. 2. Results of the price forecast for the day of August 4 in PJM market.

Table 5
Errors (%) during the day of August 4 in PJM market.

	1	2	3	4	5	6	7	8	9	10	11	12
True	70.30	57.22	48.25	44.17	44.93	48.52	59.30	63.89	76.01	89.47	102.6	112.8
Forecast	71.12	57.55	48.79	44.43	45.07	48.73	59.49	64.43	76.17	89.91	101.4	112.3
Error	1.16	0.86	0.95	0.86	0.74	0.69	0.64	0.67	0.62	0.60	0.66	0.64
	13	14	15	16	17	18	19	20	21	22	23	24
True	128.4	142.0	155.2	157.4	150.4	135.4	116.0	99.87	99.19	85.88	69.23	60.88
Forecast	127.9	141.4	151.4	155.7	149.3	134.4	118.9	101.08	97.84	87.95	68.64	59.54
Error	0.62	0.61	0.73	0.75	0.75	0.74	0.84	0.86	0.88	0.95	0.95	1.00

showed a high volatility for that period. These examinations indicate the accuracy of the proposed method due to its hybrid structure, which can capture the complex features of electricity price.

4. Conclusion

A new price forecasting method based on wavelet transform combined with ARIMA and GARCH models is proposed in this paper. The hybrid method is examined for MCP prediction in the Spanish market and LMP prediction in the PJM market and compared with some of the most recently published price forecast techniques. The results from the comparisons clearly show that the proposed method is far more accurate than the other forecast methods.

Acknowledgements

The authors want to thank China national science foundation for their financial support (No. 70971038).

References

[1] Amjady N, Keynia F. Day-ahead price forecasting of electricity markets by a mixed data model and hybrid forecast method. Electr Power Energy Syst 2008;30(9):533–46.
[2] Szkuta B, Sanavria L, Dillon T. Electricity price short term forecasting using artificial neural networks. IEEE Trans Power Syst 1999;14(3):851–7.
[3] Wang A, Ramsay B. A neural network based estimator for electricity spot-pricing with particular reference to weekend and public holidays. Neurocomputing 1998;23(1):47–57.
[4] Zhang L, Luh P, Kasiviswanathan K. Energy clearing price prediction and confidence interval estimation with cascaded neural networks. IEEE Trans Power Syst 2003;18(1):99–105.
[5] Hong YY, Hsiao CY. Locational marginal price forecasting in deregulated electric markets using a recurrent neural network. IEE Proc Gener Transm Distrib 2002;149:621–6.
[6] Yang H, Duan XZ. Chaotic characteristics of electricity price and its forecasting model. In: Proceedings of the IEEE Canadian conference on electrical and computer engineering, CCECE 2003. p. 659–62.
[7] Gonzalez AM, San Roque AM, Garcia-Gonzalez J. Modeling and forecasting electricity prices with input/output hidden Markov models. IEEE Trans Power Syst 2005;20(1):13–24.
[8] Vahidinasab V, Jadid S, Kazemi A. Day-ahead price forecasting in restructured power systems using artificial neural networks. Electr Power Syst Res 2008;78:1332–42.
[9] Amjady N, Keynia F. Day-ahead price forecasting of electricity markets by a new feature selection algorithm and cascaded neural network technique. Energy Conv Manage 2009;50(12):2976–82.
[10] Amjady N, Keynia F. Application of a new hybrid neuro-evolutionary system for day-ahead price forecasting of electricity markets. Appl Soft Comput 2010;10(3):784–92.
[11] Doulai P, Cahill W. Short-term price forecasting in electric energy market. In: Proceedings of international power engineering conference, Singapore, 2001. p. 749–54.
[12] Contreras J, Espinola R, Nogales F, Conejo A. ARIMA models to predict next-day electricity prices. IEEE Trans Power Syst 2003;18(3):1014–20.
[13] Conejo A, Plazas M, Espinola R, Molina A. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. IEEE Trans Power Syst 2005;20(2):1035–42.
[14] Nogales F, Conejo A. Electricity price forecasting through transfer function models. J Oper Res Soc 2006;57(4):1–7.
[15] Garcia R, Contreras J, Van Akkeren M, Garcia J. A GARCH forecasting model to predict day-ahead electricity prices. IEEE Trans Power Syst 2005;20(2):867–74.
[16] Nicholas B, James EP. Short term forecasting of electricity prices for MISO hubs: evidence from ARIMA–EGARCH models. Energy Econ 2008;30(6):3186–97.
[17] Abdou KD, Dominique G, Bertrand V. Forecasting electricity spot market prices with a *k*-factor GIGARCH process. Appl Energy 2009;86(4):505–10.
[18] Mallat S. A wavelet tour of signal processing. 2nd ed. California, USA: Academic Press; 1999.
[19] Amjady N. Day-ahead price forecasting of electricity markets by a new fuzzy neural network. IEEE Trans Power Syst 2006;21(2):887–96.
[20] Amjady N, Daraeepour A. Design of input vector for day-ahead price forecasting of electricity markets. Expert Syst Appl 2009;36:12281–94.
[21] Amjady N, Daraeepour A. Mixed price and load forecasting of electricity markets by a new iterative prediction method. Electr Power Syst Res 2009;79(9):1329–36.
[22] Conejo AJ, Contreras J, Plazas MA, Espinola R. Forecasting electricity prices for a day-ahead pool based energy market. Int J Forecast 2005;21:435–62.
[23] Catala JPS, Mariano SJPS, Mendes VMF, Ferreira LAFM. Short-term electricity prices forecasting in a competitive market: a neural network approach. Elect Power Syst Res 2007;77:1297–304.