

RESEARCH

Open Access



A hybrid model for short term real-time electricity price forecasting in smart grid

Xing Luo^{1,2}, Xu Zhu^{1,3} and Eng Gee Lim^{2*}

*Correspondence:

enggee.lim@xjtlu.edu.cn

An earlier version of this paper was presented in part at IEEE International Conference on Ubiquitous Computing and Communications, Liverpool, UK, October 2015

²Department of Electrical and Electronic Engineering, Xi'an Jiaotong-Liverpool University, 215123 Suzhou, People's Republic of China

Full list of author information is available at the end of the article

Abstract

Background: With the prominent growth of power market, real-time electricity price has become a trend in smart grid as it enables moderation of power consumption of customers. Accurate forecast of real-time price (RTP) has much influence on customers' behaviors, such as better scheduling operating time of domestic appliances in order to maximize benefit. In this paper, an innovative hybrid RTP forecasting model considering linear and non-linear behaviors within input data, is proposed to forecast the short-term electricity prices in smart grid.

Results: The effectiveness of the proposed hybrid forecasting model is verified by numerical results in terms of forecasting performance evaluations. The results clearly demonstrate that our approach is effective in RTP forecasting with a high accuracy. The mean absolute percentage error (MAPE) is approximate to 3.5% and it also significantly outperforms the existing models.

Conclusion: Based on the achieved results, we can conclude that the proposed hybrid model is an accurate and efficient tool in short-term RTP forecasting and it is potentially effective to a variety of forecasting tasks.

Keywords: Power market, RTP forecasting, Hybrid model, Short-term electricity prices, Smart grid

Background

Real-time price (RTP), also referred to as dynamic tariff or spot price which was first introduced in the 1980s [1], nowadays is tentatively applied to the power system in many countries including the US, Australia, etc. The real-time price tariff is an inexorable trend in next generation of power system reforming [2, 3]. Unlike regulated markets which the companies determine prices independently, electricity prices are significantly dependent on a supply–demand relationship in a deregulated market. Generally, RTP offers higher prices during peak load demand periods and provides lower prices during off-peak load demand periods [4, 5]. In consideration of the manufacturing cost in different load levels, the dynamic tariff is a potential load management method for properly allocating incremental prices of electricity consumption to the time delivery, thus ensuring the overall economic rationality [6].

In addition, RTP tariff is broadly utilized as a basic control signal to support the demand response management (DRM) which is an excellent long-term solution to improving



energy efficiency and reducing wastage [7, 8]. On the one hand, RTP tariff is benefit to power grid as it offers specific price instructions for participants to average the power usage at different time so that alleviates the load burden of power grid especially in peak demand time. On the other hand, such an electricity tariff encourages consumption by price reduction during periods of abundance and allows customers to have multiple choices to determine the time of electricity consumption. The participants in electricity market can regulate the operating time of electrical devices automatically or manually during high-price periods and gain the benefits from low-price periods via DRM, thus achieving the aims of reducing energy usage and saving electric bills for themselves [5, 9–11]. Therefore, the research on RTP tariff is of interest to researchers, production companies, investors, independent market operators and large industrial consumers in recent years [12, 13].

Moreover, the real-time price is normally provided with the instantaneous property. Thus, it is a necessity to forecast RTP in advance in this competitive electricity market for electricity consumers and power suppliers in scheduling their operations and controlling the price risks. Over last two decades, much research has been conducted on RTP forecasting. In summary, the existing methods can be classified into two main categories: machine learning based methods like SVM (Support Vector Machine) and ANN (artificial neural network) [14–16], and statistical time series based methods like ARIMA (auto regressive integrated moving average) model and GARCH (generalized auto regressive conditional heteroskedasticity) model [17, 18].

Specifically, in [15], the authors proposed methods including hybrid networks of self-organized map (SOM) and support-vector machine (SVM) to predict short-term electricity price. With the trained network, one can predict the future hourly electricity prices in one day ahead. To confirm its feasibility, the proposed model had been trained and tested on the data of historical energy prices from the New England electricity market. In addition, in [16], a sensitivity analysis of similar days (SD) parameters to rise the accuracy of ANN model and SD-based short-term price forecasting model were presented. In order to train the network, a large sum of data were used. The model had been tested in Pennsylvania-New Jersey-Maryland (PJM) electricity market. The results showed that the mean absolute percentage error (MAPE) was around 11%. Furthermore, in [17], the authors introduced a method to predict next-day electricity prices based on the ARIMA methodology which was used to analyze the time series problem. The ARIMA model was tested in California electricity market. More than 30-day historical data samples were required to train the model.

However, the shared limitation of the mentioned studies above is that a large number of historical RTP data is required for training the model. The insufficient training data causes considerable estimation errors. Hence, our research in the paper mainly concentrates on building an effective estimation model for electricity price forecasting in smart grid with high accuracy by using limited sets of historical data. In order to evaluate the performance of methods, numerical error measures such as mean absolute error (MAE), means square error (MSE), root-mean square error (RMSE) and mean absolute percentage error (MAPE) are also used in this work.

The main contributions of this work can be summarized as follows.

- (1) A hybrid RTP forecasting model which is a consolidation of least-square (LS) fitting model, grey prediction (GP) model and artificial neural network (ANN), is proposed. The LS fitting model considers the linear behavior of the time series data and the GP model considers the non-linear behavior. However, the ANN model is an optional forecasting procedure and used in the error optimization.
- (2) Less historical RTP data is required, thereby improving the practicability. Since both LS and GP models can be established on the basis of a small number of data sets, the proposed hybrid forecasting model is easy to install and more practical compared with the previous methods.
- (3) The accuracy of time series RTP forecasting increases by using the hybrid model. The effectiveness of the hybrid forecasting model is verified by numerical results in terms of MAE, MSE, RMSE and MAPE evaluations. The results indicate that our method is an accurate and efficient tool to forecast the day-ahead RTP and it also significantly outperforms the previous methods.

To the best of our knowledge, this is the first work of combining above pieces together in RTP forecasting.

Method

This section introduces the methodology which includes the architecture of the proposed forecast strategy and the specific description of the proposed hybrid forecasting model in this work.

Architecture of the proposed forecasting strategy

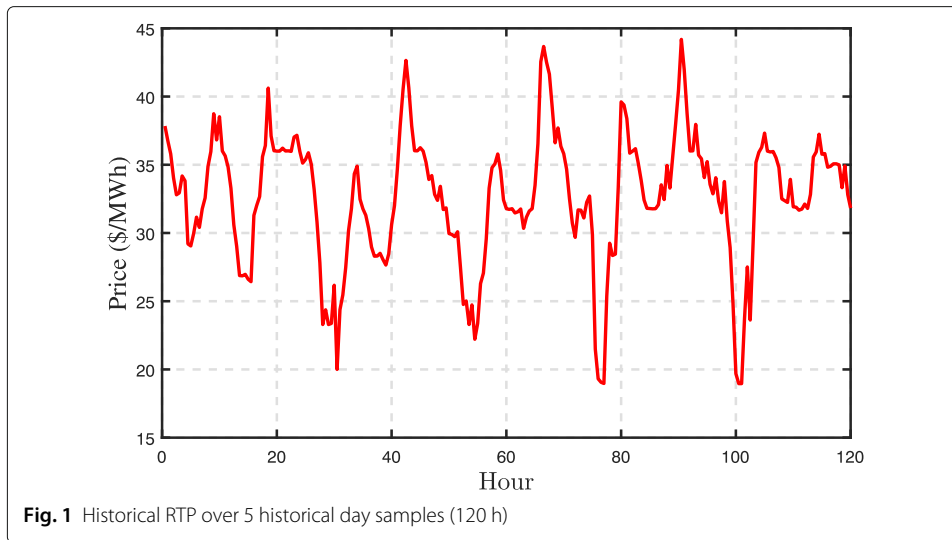
Considering time scales, the RTP forecasting is classified into ultra-short term, short term, medium term and long term [19]. The Ultra-short term is from several minutes to 1 h ahead forecasting. The short term means the forecasting values from 1 h to several hours. From a few hours to 1 week ahead forecasting is defined as the medium term and beyond that it is the long term forecasting. However, we focus on the day (24 h) ahead RTP forecasting with a resolution of 0.5 h in this work, which belongs to the short term forecasting.

Figure 1 shows the historical RTP data over 5 historical day samples which is provided by Australia Energy Market Operator (AEMO) [20]. The time series dynamic electricity prices vary dependent on load demand at different time periods. Based on the variations of historical RTP samples in Fig. 1, electricity prices exhibit a prominent regularity apparently and it consists of linear and non-linear information along with the prices varying. According to these, the characteristics of the linear and non-linear properties of time series data have to be incorporated into the forecasting model. Therefore, the proposed forecasting model can be formulated as:

$$P_t = L_t + N_t + E_t^* \quad (1)$$

where P_t is the forecasting RTP at time t . L_t and N_t represent the estimations of linear behavior and non-linear behavior, respectively, of the input data. Additionally, E_t^* which is an optional forecasting component, denotes the error optimization procedure.

In order to present the architecture of the proposed hybrid forecasting model, Fig. 2 illustrates the flow chart of forecasting day-ahead real-time electricity prices based on

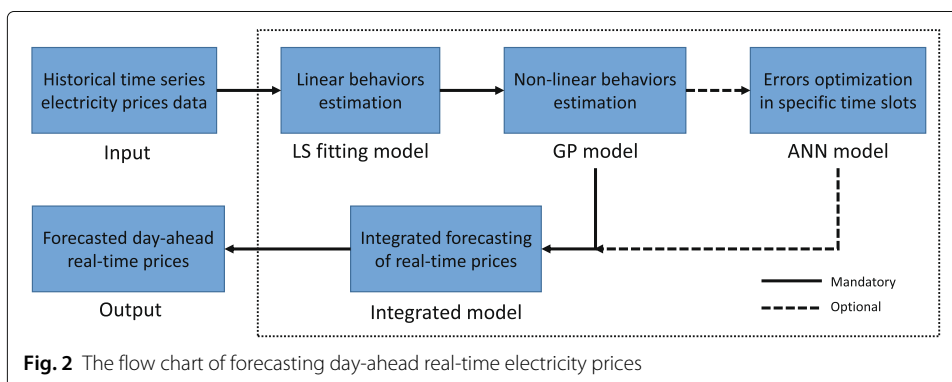


several days' historical RTP data. Specifically, the historical data is input as the basis to establish the model. Then, the linear behavior of the data are estimated by using the LS fitting model. Afterwards, the GP model is applied in the estimation of the non-linear behavior within the data. After that, the ANN model based error optimization procedure will be determined if it is necessary to be executed on this stage in accordance with the spot error rate (ER) of the initial forecasting result. The ANN model will be executed to improve the specific forecasting accuracy if the spot ER exceeds the maximum tolerable ER. Finally, the forecasted day-ahead RTP coming from the integrated model is the desired output in this study.

In next subsections, the specific descriptions of the relevant forecasting components in the hybrid model are introduced in details.

Least squares fitting model for linear behavior forecasting

Based on Eq. (1), the LS fitting model is employed to obtain the linear behavior L_t within the input data. The least square fitting for data is a standard approach in regression analysis to the approximate solution of over determined systems. It is one of the fitting algorithms [21–24]. On the stage of linear behavior forecasting, the LS fitting model can



be used to build a fitting function to express the main stream variation among the historical data. Assume the input data set H consisting of $n = N$ days' historical RTP data, H can be formulated as:

$$H = \{D_1, D_2, \dots, D_n\} \tag{2}$$

However, the historical RTP of a day can be treated as a number of discrete values with an interval. In this study, a time interval of 0.5 h is adopted, which means $t = 48$ fixed values are included in an individual day sample. Hence, D_n is represented as:

$$D_n = [y_{n,1}, y_{n,2}, \dots, y_{n,t}] \tag{3}$$

In addition, the fitting function $L(t)$ is taken to model the main stream variation in the linear behavior estimation. However, the general formats of the fitting function include Fourier, Gaussian, polynomial, sum of sine, etc. and they can be formulated in Eqs. (4) - (7), respectively, as follows.

$$\text{Gaussian format: } f(x) = \sum_{i=1}^d a_i \cdot e^{-\left(\frac{x-b_i}{c_i}\right)^2} \tag{4}$$

$$\text{Fourier format: } f(x) = \sum_{i=0}^d a_i \cdot \cos(i \cdot \omega \cdot x) + b_i \cdot \sin(i \cdot \omega \cdot x) \tag{5}$$

$$\text{Polynomial format: } f(x) = \sum_{i=0}^d p_i \cdot x^i \tag{6}$$

$$\text{Sum of sine format: } f(x) = \sum_{i=1}^d a_i \cdot \sin(b_i \cdot x + c_i) \tag{7}$$

where $d \in \mathbb{N}^+$ is the degree of the adopted function. Additionally, a_i, b_i, c_i, ω_i and p_i are undetermined constant parameters in model. Although all the proposed fitting function formats are effective in modeling the linear behavior within the data, the Fourier format is adopted in this study due to its better fitting performance. Therefore, the objective function on this stage can be formulated as determining a group of appropriate parameters (a_i, b_i and ω_i) to minimize the total square errors J . The objective function is presented in Eq. (8).

$$\arg \min_{a,b,\omega} J = \sum_{n=1}^N \sum_{t=1}^{48} (L_t - y_{n,t})^2 \tag{8}$$

On the one side, a higher value of the fitting degree d leads to a better performance of the estimation when J is in a reasonable range. On the other hand, it results more complexity of the calculation and more CPU wastage. Therefore, selecting an appropriate fitting degree in the fitting model is significant and may lead to a better linear behavior estimation performance.

Table 1 presents the total square errors J with different values of fitting degree which range from 1 to 7. Apparently, when $d \in [1, 3], J$ decreases quickly with the fitting degree increasing. However, J becomes stable when $d > 3$. For example, there is only $|J_{d=5} - J_{d=6}| = |2.995 - 2.981| = 0.015$ differences between the cases $d = 5$ and $d = 6$ based on the obtained results in Table 1.

Table 1 Total square errors with different values of fitting degree; $d = 1$ to 7 are evaluated

d	1	2	3	4	5	6	7
$J \times 10^2$	5.601	4.152	3.223	2.988	2.995	2.981	2.950

Although different fitting degree values can work in the linear behavior estimation on this stage, considering the estimation accuracy and efficiency, and avoiding the over fitting, $d = 4$ is taken as a proper fitting degree in this study. Afterwards, the relevant parameters can be determined as in Table 2.

Grey prediction model for non-linear behavior forecasting

The second stage in the proposed hybrid RTP forecasting model is to estimate N_t which denotes the non-linear behavior within the input data. Obviously, the non-linear information within the data is included in the forecasting errors after using the LS fitting model. Thus, the non-linear behavior within the historical data at time t can be expressed in Eq. (9) and the initial records are used to estimate the next record ($D_{n+1,t} - L_t$) by using the GP model.

$$M_t = \{D_{1,t} - L_t, D_{2,t} - L_t, \dots, D_{n,t} - L_t\} \tag{9}$$

The GP model or GM(1,1) was first proposed to deal with the data in grey system. It is able to analyze system that includes insufficient information and unapparent relationship [25–27]. Hence, the GP model is often used in predicting data in non-linear system based on limited information. It transforms the forms of the irregular discrete sequences and displays the potential regularities within the sequences. Transforming the forms of the sequences can make the properties of stochastic and randomness get weaker thereby turning irregular sequences to regular ones [28–30]. Since only a few non-linear data proceeded from LS fitting model are used, it is quite appropriate to employ the GP model to estimate the non-linear behavior within the input data on this stage.

The GP model is established by using generalized series data. The primitive sequence data is defined as $X^{(0)}$ and it can be presented as:

$$X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\} \tag{10}$$

where $x^{(0)}(n) = D_{n,t} - L_t$ and $x^{(0)}(n) \geq 0$. However, if any $x^{(0)}(n) < 0$ in the primitive sequence, all the candidates in the sequence have to be improved until $\forall x^{(0)}(n) \geq 0$. Afterwards, the first accumulated generating data $X^{(1)}$ can be obtained in Eq. (11).

$$X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\} \tag{11}$$

where

$$x^{(1)}(n) = \sum_{i=1}^n x^{(0)}(i) \tag{12}$$

Table 2 Values of the parameters in L_t . Fourier format is selected and $d = 4$

Parameters	a_0	a_1	a_2	a_3	a_4	b_1	b_2	b_3	b_4	ω
Values	32.640	0.570	2.153	0.022	-0.831	-2.950	-4.158	1.738	1.185	0.289

In addition, $Z^{(1)}$ which is determined by $X^{(1)}$, is defined as the background factors and $Z^{(1)}$ can be presented as:

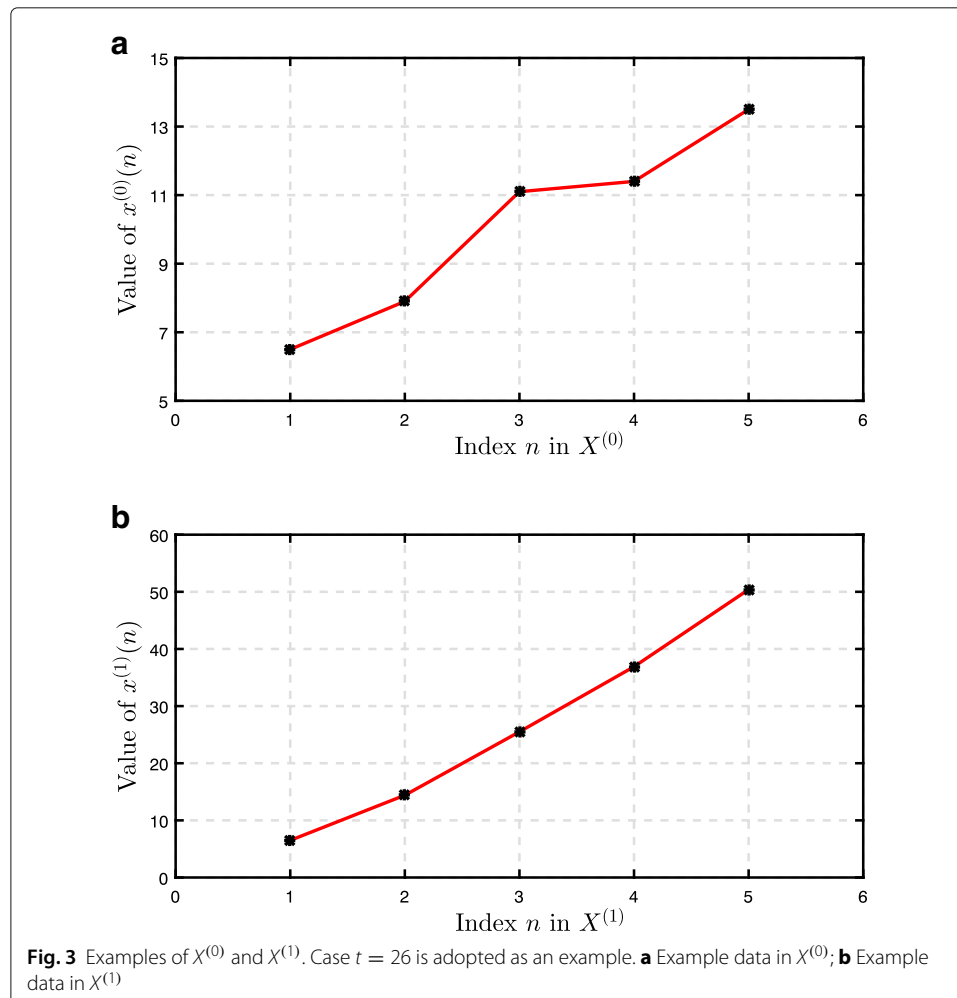
$$Z^{(1)} = \{z^{(1)}(1), z^{(1)}(2), \dots, z^{(1)}(n)\} \tag{13}$$

where

$$z^{(1)}(n) = \frac{1}{2} [x^{(1)}(n) + x^{(1)}(n - 1)] \tag{14}$$

For example, when $t = 26$ in our case, the initial sequence (after data preprocessing) is $X^{(0)} = \{6.5, 7.9, 11.1, 11.4, 13.5\}$ as shown in Fig. 3a. Then, the first accumulated generating data $X^{(1)}$ can be calculated as $X^{(1)} = \{6.5, 14.4, 25.5, 36.9, 50.4\}$ as shown in Fig. 3b. Obviously, there are not any prominent regularities between the numbers in $X^{(0)}$. However, after first accumulated generating operation (AGO) towards to $X^{(0)}$, the new sequence $X^{(1)}$ is provided with the quasi exponential property (otherwise, the second AGO will be executed towards to $X^{(1)}$). Therefore, $X^{(1)}$ can be regarded as being satisfied with the first order ordinary differential equation which is shown in (15).

$$\frac{dX^{(1)}}{dt} + aX^{(1)} = u \tag{15}$$



where a is treated as a development coefficient that describes the increasing speed of numbers in $X^{(0)}$ and u is an endogenous control coefficient in system. The parameters $U = [a, u]^T$ can be determined in Eq. (16).

$$U = (B^T B)^{-1} B^T Y \tag{16}$$

where Y is a $(n - 1) \times 1$ matrix and B is a $(n - 1) \times 2$ matrix. Y and B can be presented as:

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} \text{ and } B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}.$$

According to these, Eq. (15) can be resolved by using the obtained parameters a and u , so that the forecasting formula of $\hat{X}^{(1)}$ can be denoted as shown in Eq. (17).

$$\hat{x}^{(1)}(n + 1) = \left(x^{(1)}(1) - \frac{u}{a}\right) e^{-ak} + \frac{u}{a} \tag{17}$$

Based on Eq. (17), when $n = 1, 2, \dots, N - 1$, $\hat{x}^{(1)}(n + 1)$ is a fixed value. When $n \geq N$, $\hat{x}^{(1)}(n + 1)$ is a predicted value of $X^{(1)}$. Afterwards, the predicted formula of $\hat{X}^{(0)}$ can be achieved through the inverse accumulated generating operation (IAGO) as shown in Eq. (18).

$$\begin{aligned} \hat{x}^{(0)}(n + 1) &= \hat{x}^{(1)}(n + 1) - \hat{x}^{(1)}(n) \\ &= (1 - e^a) \left[x^{(0)}(1) - \frac{u}{a}\right] e^{-ak} \end{aligned} \tag{18}$$

According to Eq. (18), when $n = N$, $\hat{x}^{(0)}(n + 1)$ is the objective forecasting value. After that, the GP model can be utilized periodically to achieve forecasting values on all time spot t . Furthermore, the forecasting result is the combination of the linear behavior and the non-linear behavior within the input data, which can be described in Fig. 4. In general, the variation of the obtained RTP by using LS model+GP model is in line with actual RTP in an overview. The result indicates that the error rates are lower than 10% at most of

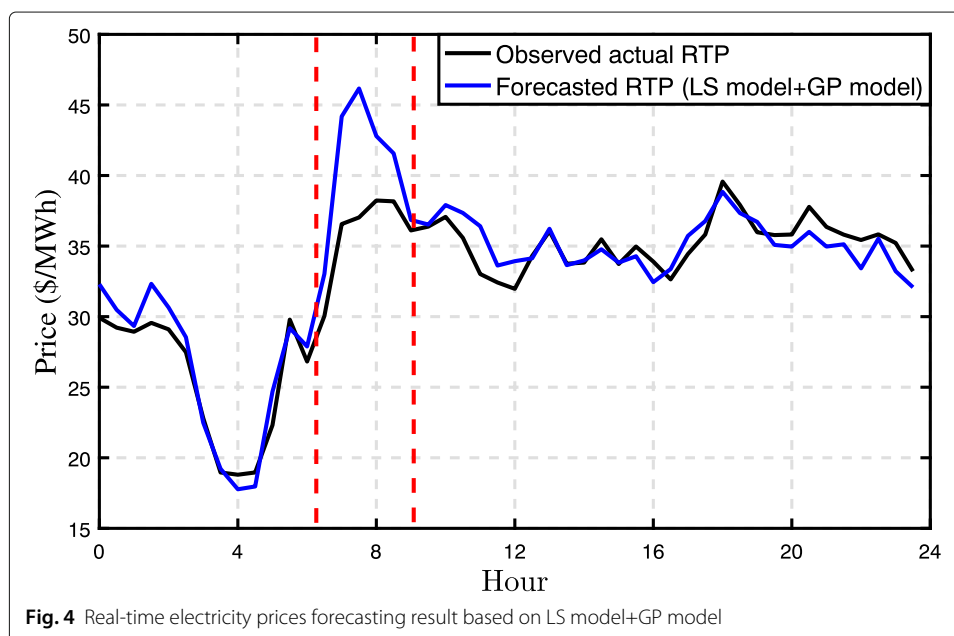


Fig. 4 Real-time electricity prices forecasting result based on LS model+GP model

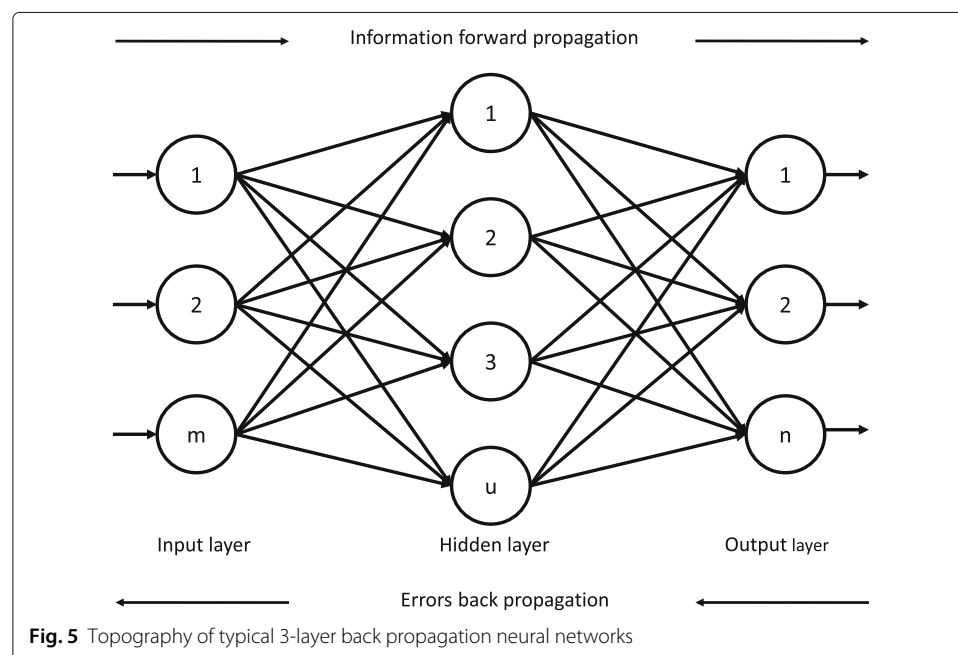
the time, particularly between 12:30 - 22:00 (lower than 5%). Nonetheless, the error rates are higher than 10% during time period 6:30 - 8:30, as unexpected. These unexpected errors may be caused by the defects of the current forecasting models. Since limited data sets were used for improving the practicality in the current models, the random error increases when there are great differences among the input data sets. Therefore, in order to improve the forecasting accuracy at specific time slots based on the initial forecasting result, the ANN model based error optimization procedure is required.

Artificial neural network for error optimization

The artificial neural network (ANN) is also a non-linear modeling where any prior knowledge of relationship between input and output is needed [31]. It gives great results for forecasting problems [16]. To establish the model, only sufficient data is required to assimilate the connection between inputs and outputs. The main parameters of ANN model are the number of the input vectors, the number of layers and the number of neurons in each layer [32–34]. However, the large and sudden spikes in the input data will lead to less accuracy in the output using ANN. In this study, the back propagation (BP) algorithm is utilized to train the ANN model.

Figure 5 shows the topography of a typical 3-layer back propagation neural network [35, 36]. A 3-layer back propagation neural network is a typical multiple-layer network and it includes input layer (LA), hidden layer (LB), and output layer (LC). There are no connections between nodes that belong to the same layer. LA has m nodes that correspond to the m inputs of the network. LC consists of n nodes that correspond to the n output of the network. The node number of LB can be varied to fit the task.

Define W_{ir} as the connection weight between node a_i of the LA layer and node b_r of the LB layer. Similarly, let V_{rj} be the connection weight between node b_r of the LB layer and node c_j of the LC layer. Set T_r and θ_j as the bias of node b_r of the LB layer and the bias of node c_j of the LC layer, respectively. Then the output function of the LB layer node



should be:

$$b_r = f \left(\sum_{i=1}^m W_{ir} \cdot a_i - T_r \right) \quad (r = 1, 2, \dots, u) \quad (19)$$

The output function of the LC layer node should be:

$$c_j = f \left(\sum_{r=1}^u V_{rj} \cdot b_r - \theta_j \right) \quad (j = 1, 2, \dots, n) \quad (20)$$

where $f(x)$ is a sigmoid function and it can be expressed as:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (21)$$

In addition, the BP learning algorithm which is a typical error-revised learning algorithm is used to learn and store knowledge in the 3-layer back propagation neural networks. The learning procedures can be illustrated as follows.

- (1) Initialize the variables W_{ir} , T_r , V_{rj} and θ_j with small random values.
- (2) For each model pair $(A^{(k)}, C^{(k)})$ ($k = 1, 2, \dots, p$), take the following steps.
 - Input the values of $A^{(k)}$ at layer LA, then calculate b_r and c_j by Eqs. (19) and (20).
 - Calculate the bias d_j of the desired value and calculate the value c_j of the layer LC nodes and let

$$d_j = c_j \cdot (1 - c_j) \cdot (c_j^{(k)} - c_j) \quad (22)$$

- Back propagate the errors to the layer LB nodes and let

$$e_r = b_r \cdot (1 - b_r) \cdot \left(\sum_{j=1}^n V_{rj} \cdot d_j \right) \quad (23)$$

- Adjust the connection weights V_{rj} and the bias of the layer LC nodes θ_j :

$$V_{rj} = V_{rj} + \alpha \cdot d_j + \beta \cdot \Delta V'_{rj} \quad (24)$$

$$\theta_j = \theta_j + \alpha \cdot d_j + \beta \cdot \Delta \theta'_j \quad (25)$$

where $\Delta V'_{rj}$ and $\Delta \theta'_j$ are the adjusting values of the previous learning loop. α is the learning ratio and $0 < \alpha < 1$. β is the momentum factor.

- Adjust the connection weights W_{ir} and the bias of the layer LB nodes T_r :

$$W_{ir} = W_{ir} + \alpha \cdot e_r + \beta \cdot \Delta W'_{ir} \quad (26)$$

$$T_r = T_r + \alpha \cdot e_r + \beta \cdot \Delta T'_r \quad (27)$$

where $\Delta W'_{ir}$ and $\Delta T'_r$ are the adjusting values of the previous learning loop.

- (3) Repeat step (2), until d_j becomes adequately small.

In accordance with the analysis in previous sections, the ANN model is used on this stage to improve the accuracy of the RTP forecasting further in particular time slots, such as between 6:30 - 8:30 as shown in Fig. 4. In this case, 2 hidden layers with 20 and 40 neurons are designed and 10-day historical data is adopted. In next section, a number of simulations are carried out to prove the effectiveness of the proposed hybrid model and the forecasting quality is also evaluated in terms of some evaluation criteria.

Results

This section demonstrates the real-time electricity prices forecasting results by using the proposed hybrid forecasting model. Limited data sets (5 days) of the historical RTP with a time interval of 0.5 h in Australia is adopted. The achieved results are also compared with the previous methods (e.g., ARIMA model, independent BP-ANN model, etc.) in this work.

In addition, a number of evaluation criteria including MAE, MSE, RMSE and MAPE [37–39] are proposed to evaluate the forecasting performances. To begin with, we define x as the forecasting values of a model, ref as the observed true values and $t = 48$ which means 48 points are required to predict within a day. Hence, the mathematical formulations of MAE, MSE, RMSE and MAPE can be expressed in Eqs. (28) - (31), respectively.

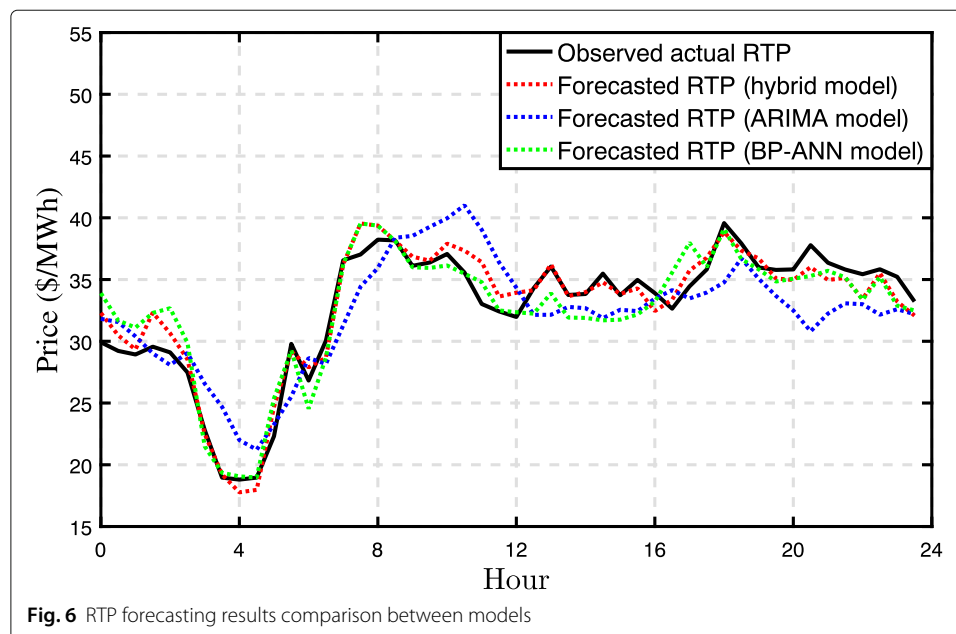
$$MAE = \frac{1}{t} \cdot \sum_{i=1}^t |ref_i - x_i| \quad (28)$$

$$MSE = \frac{1}{t} \cdot \sum_{i=1}^t (ref_i - x_i)^2 \quad (29)$$

$$RMSE = \sqrt{\frac{1}{t} \cdot \sum_{i=1}^t (ref_i - x_i)^2} \quad (30)$$

$$MAPE = \frac{1}{t} \cdot \sum_{i=1}^t \left| \frac{ref_i - x_i}{ref_i} \right| \quad (31)$$

Figure 6 shows the RTP forecasting results comparison between the proposed hybrid model and some typical models. According to the obtained results, it is obvious that all three compared models, i.e., the hybrid model, the ARIMA model and the BP-ANN model are able to accomplish the task of forecasting RTP in advance. The forecasting price variations are in line with the observed actual RTP in general. Additionally, the forecasting RTP by the hybrid model is slightly better than the other two models on the basis of the results.



However, comparing the RTP variations in Figs. 4 and 6, it is apparent that the forecasting errors are significantly reduced at the time period 6:30 - 8:30, due to the contribution of the BP-ANN model in the error optimization procedure.

In addition to these, the RTP forecasting quality evaluation comparison between the hybrid model and different models in state-of-the-arts are also presented in Table 3. Based on the achieved results, the proposed hybrid model performs best in forecasting quality evaluation in an overview, which confirms the advantages of our approach. Specifically, the MAE, MSE, RMSE and MAPE of the hybrid model are 1.06, 1.72, 1.31 and 3.38%, respectively, which are the lowest among all models. However, the ARIMA model performs the worst in the MAE evaluation and the independent LS model performs worst in MSE, RMSE and MAPE evaluations.

Discussion

The hybrid model analyzes the input data in views of linear behavior, non-linear behavior and errors optimization within the data. The advantage of our approach is that the hybrid model is more robust in dealing with forecasting tasks based on insufficient data compared with the traditional models such as ARIMA which needs a large number of historical data for training. The RTP forecasting quality evaluation results in Table 3 also indicate that the ARIMA model is not effective in the case with limited input data. In addition, the individual LS model did not perform well in this case, as the LS model only extracts the main stream within the input data. Therefore, using a LS model independently to forecast the RTP will lead to considerable errors as expected. It is much more interesting to see that the LS model cooperated with the GP model performs a bit worse than the independent GP model in overall evaluation. This is because the combined model (LS model + GP model) perform not well in a specific time period, *i.e.*, 6:30 to 8:30 in this case, so that the errors improved significantly in overall, although it has higher forecasting accuracies in other time periods compared with the GP model.

However, given a group of historical data, there are several forecasting models can be used and each model may be able to complete the task of forecasting with different accuracies. After a great number of tests, we realize that the forecasting performance is crucially dependent on both selecting an appropriate model and the data correlations. A forecasting model works well to one group of data, but can be not effective for another group of data. Therefore, a hybrid forecasting model is generally more efficient than an independent model.

Table 3 RTP Forecasting quality evaluation comparison between models

Models	Evaluation criteria			
	MAE (\$/MWh)	MSE (\$/MWh)	RMSE (\$/MWh) ^{1/2}	MAPE (%)
ARIMA model	2.61	9.25	3.04	8.29
LS model	2.52	9.41	3.07	8.51
GP model	1.35	3.01	1.74	4.29
LS+GP models	1.53	5.33	2.31	4.65
BP-ANN model	1.49	3.50	1.87	4.69
Hybrid model	1.06	1.72	1.31	3.38

The best performance metrics are marked in bold

Conclusion

In this paper, a hybrid model consisting of LS model, GP model and BP-ANN model, is proposed to forecast the day-ahead real-time prices based on limited historical data. The achieved forecasting performance evaluation results clearly demonstrate that our approach is an accurate and efficient tool in RTP forecasting and it also significantly outperforms the previous forecasting models. As RTP tariff is a trend for smart grid in next decade, the theories in this paper have bright prospects not only in RTP forecasting, but also in applications in other industrial fields, such load forecasting, wind forecasting, GDP forecasting, etc.

Acknowledgement

This work is partially supported by the XJTLU Key Programme Special Fund (KSF-P-02).

Availability of data and materials

All data generated or analyzed during this study are included in this paper and reference list. The data is available from the corresponding author on reasonable request.

Authors' contributions

XL proposed this topic, carried out numerical experiments, and drafted the manuscript. XZ and EGL checked and clarified the manuscript carefully. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Electrical Engineering and Electronics, University of Liverpool, L69 3GJ Liverpool, UK. ²Department of Electrical and Electronic Engineering, Xi'an Jiaotong-Liverpool University, 215123 Suzhou, People's Republic of China.

³School of Electronic and Information Engineering, Harbin Institute of Technology (Shenzhen), 518055 Shenzhen, People's Republic of China.

Received: 20 June 2018 Accepted: 26 September 2018

Published online: 17 October 2018

References

1. Garcia EV, Runnels JE. The utility perspective of spot pricing. *IEEE Power Eng Rev.* 1985;PER-5(6):45. <https://doi.org/10.1109/MPER.1985.5526640>.
2. Saele H, Grande OS. Demand response from household customers: Experiences from a pilot study in norway. *IEEE Trans Smart Grid.* 2011;2(1):102–9. <https://doi.org/10.1109/TSG.2010.2104165>.
3. Vlachos AG, Biskas PN. Simultaneous clearing of energy and reserves in multi-area markets under mixed pricing rules. *IEEE Trans Power Syst.* 2011;26(4):2460–71. <https://doi.org/10.1109/TPWRS.2011.2126025>.
4. Yu R, Yang W, Rahardja S. A statistical demand-price model with its application in optimal real-time price. *IEEE Trans Smart Grid.* 2012;3(4):1734–42. <https://doi.org/10.1109/TSG.2012.2217400>.
5. Luo X, Zhu X, Lim EG. Load scheduling based on an advanced real-time price forecasting model. In: *Proc. 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing (CIT/IUCC/DASC/PICOM)*. 2015. p. 1252–7. <https://doi.org/10.1109/CIT/IUCC/DASC/PICOM.2015.186>.
6. Chang CS, Yi M. Real-time pricing related short-term load forecasting. In: *Proc. 1998 International Conference on Energy Management and Power Delivery*. 1998. p. 411–4162. <https://doi.org/10.1109/EMPD.1998.702588>.
7. Luo X, Zhu X, Lim EG. Dynamic pricing based and electric vehicle assisted demand response strategy. In: *2017 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. 2017. p. 357–62. <https://doi.org/10.1109/SmartGridComm.2017.8340691>.
8. Qian K, Zhou C, Allan M, Yuan Y. Modeling of load demand due to ev battery charging in distribution systems. *IEEE Trans Power Syst.* 2011;26(2):802–10. <https://doi.org/10.1109/TPWRS.2010.2057456>.
9. O'Dwyer C, Duignan R, O'Malley M. Modeling demand response in the residential sector for the provision of reserves. In: *2012 IEEE Power and Energy Society General Meeting*. 2012. p. 1–8. <https://doi.org/10.1109/PESGM.2012.6344757>.

10. Yao L, Lim WH, Tsai TS. A real-time charging scheme for demand response in electric vehicle parking station. *IEEE Trans Smart Grid*. 2017;8(1):52–62. <https://doi.org/10.1109/TSG.2016.2582749>.
11. Tsui KM, Chan SC. Demand response optimization for smart home scheduling under real-time pricing. *IEEE Trans Smart Grid*. 2012;3(4):1812–21. <https://doi.org/10.1109/TSG.2012.2218835>.
12. Mazengia DH, Tuan LA. Forecasting spot electricity market prices using time series models. In: 2008 IEEE International Conference on Sustainable Energy Technologies. 2008. p. 1256–61. <https://doi.org/10.1109/ICSET.2008.4747199>.
13. Tahmasebifar R, Sheikh-El-Eslami MK, Kheirollahi R. Point and interval forecasting of real-time and day-ahead electricity prices by a novel hybrid approach. *IET Gener Transm Distrib*. 2017;11(9):2173–83. <https://doi.org/10.1049/iet-gtd.2016.1396>.
14. Amjady N. Day-ahead price forecasting of electricity markets by a new fuzzy neural network. *IEEE Trans Power Syst*. 2006;21(2):887–96. <https://doi.org/10.1109/TPWRS.2006.873409>.
15. Fan S, Mao C, Chen L. Next-day electricity-price forecasting using a hybrid network. *IET Gener Transm Distrib*. 2007;1(1):176–82. <https://doi.org/10.1049/iet-gtd:20060006>.
16. Mandal P, Srivastava AK, Park JW. An effort to optimize similar days parameters for ann-based electricity price forecasting. *IEEE Trans Ind Appl*. 2009;45(5):1888–96. <https://doi.org/10.1109/TIA.2009.2027542>.
17. Contreras J, Espinola R, Nogales FJ, Conejo AJ. Arima models to predict next-day electricity prices. *IEEE Trans Power Syst*. 2003;18(3):1014–20. <https://doi.org/10.1109/TPWRS.2002.804943>.
18. Garcia RC, Contreras J, van Akkeren M, Garcia JBC. A garch forecasting model to predict day-ahead electricity prices. *IEEE Trans Power Syst*. 2005;20(2):867–74. <https://doi.org/10.1109/TPWRS.2005.846044>.
19. Nair KR, Vanitha V, Jisma M. Forecasting of wind speed using ann, arima and hybrid models. In: 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT). 2017. p. 170–5. <https://doi.org/10.1109/ICICT1.2017.8342555>.
20. In: AEMO-Australian Energy Market Operator. Available: <http://www.aemo.org.au/>. Accessed 25 Dec 2017.
21. Miao Z, Fan L, Aghamolki HG, Zeng B. Least squares estimation based sdp cuts for socp relaxation of ac opf. *IEEE Trans Autom Control*. 2018;63(1):241–8. <https://doi.org/10.1109/TAC.2017.2719607>.
22. Kesaniemi M, Virtanen K. Direct least square fitting of hyperellipsoids. *IEEE Trans Pattern Anal Mach Intell*. 2018;40(1):63–76. <https://doi.org/10.1109/TPAMI.2017.2658574>.
23. Renczes B, Kollar I, Daboczi T. Efficient implementation of least squares sine fitting algorithms. *IEEE Trans Instrum Meas*. 2016;65(12):2717–24. <https://doi.org/10.1109/TIM.2016.2600998>.
24. Guilin T, Yunming Q. Improved least square method apply in ship performance analysis. In: Proc. 2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE). 2010. p. 5–5945596. <https://doi.org/10.1109/ICACTE.2010.5579420>.
25. Haiping H, Shichao H, Jiutian C, Ruchuan W. Image hiding algorithm in discrete cosine transform domain based on grey prediction and grey relational analysis. *China Commun*. 2013;10(7):57–70. <https://doi.org/10.1109/CC.2013.6571289>.
26. Yang SH, Huang WJ, Tsai JF, Chen YP. Symbiotic structure learning algorithm for feedforward neural-network-aided grey model and prediction applications. *IEEE Access*. 2017;5:9378–88. <https://doi.org/10.1109/ACCESS.2017.2702340>.
27. Wang Y, Liu Q, Tang J, Cao W, Li X. Optimization approach of background value and initial item for improving prediction precision of gm(1,1) model. *J Syst Eng Electron*. 2014;25(1):77–82. <https://doi.org/10.1109/JSEE.2014.00009>.
28. Lee JS, Lee YC. An application of grey prediction to transmission power control in mobile sensor networks. *IEEE Internet Things J*. 2018;5(3):2154–62. <https://doi.org/10.1109/JIOT.2018.2826008>.
29. Hui P, Qu W, Tang J, Chen J. Traffic indexes prediction based on grey prediction model. In: Proc. 2013 Sixth International Symposium on Computational Intelligence and Design (ISCID). 2013. p. 244–7. <https://doi.org/10.1109/ISCID.2013.68>.
30. Xu N, Zhang X-R. Traffic volume prediction based on improved grey self-adaptable prediction formula. In: Proc. 2010 International Conference on Machine Learning and Cybernetics (ICMLC). 2010. p. 1027–30. <https://doi.org/10.1109/ICMLC.2010.5580624>.
31. Khashhei M, Bijari M. An artificial neural network (p,d,q) model for timeseries forecasting. *Expert Syst Appl*. 2010;37(1):479–89. <https://doi.org/10.1016/j.eswa.2009.05.044>.
32. Zhang LD, Jia L, Zhu WX. Overview of traffic flow hybrid ann forecasting algorithm study. In: Proc. 2010 International Conference on Computer Application and System Modeling (ICCASM 2010). 2010. p. 1–6151619. <https://doi.org/10.1109/ICCASM.2010.5620414>.
33. Zhao H, Peng L, Takahashi T, Hayashi T, Shimizu K, Yamamoto T. Ann based data integration for multi-path ultrasonic flowmeter. *IEEE Sensors J*. 2014;14(2):362–70. <https://doi.org/10.1109/JSEN.2013.2282466>.
34. Lian T, Xie M, Xu J, Chen L, Gao H. Modified bp neural network model is used for oddeven discrimination of integer number. In: Optoelectronics and Microelectronics (ICOM), 2013 International Conference On. 2013. p. 67–70. <https://doi.org/10.1109/ICoOM.2013.6626492>.
35. Zhu D, Zhang T, Mao G. Back-propagation artificial neural networks for water supply pipeline model. *Tsinghua Sci Technol*. 2002;7(5):527–31.
36. Methaprayoon K, Yingvivanapong C, Lee WJ, Liao JR. An integration of ann wind power estimation into unit commitment considering the forecasting uncertainty. *IEEE Trans Ind Appl*. 2007;43(6):1441–8. <https://doi.org/10.1109/TIA.2007.908203>.
37. Zhao X, Wang S, Li T. Review of evaluation criteria and main methods of wind power forecasting. *Energy Procedia*. 2011;12:761–9. <https://doi.org/10.1016/j.egypro.2011.10.102>. The Proceedings of International Conference on Smart Grid and Clean Energy Technologies.
38. Mehdiyev N, Enke D, Fettek P, Loos P. Evaluating forecasting methods by considering different accuracy measures. *Procedia Comput Sci*. 2016;95:264–71. <https://doi.org/10.1016/j.procs.2016.09.332>. Complex Adaptive Systems Los Angeles, CA November 2–4, 2016.
39. Vihinen M. How to evaluate performance of prediction methods? measures and their interpretation in variation effect analysis. *BMC Genomics*. 2012;13(4):2. <https://doi.org/10.1186/1471-2164-13-54-52>.