



# Combined Short-Term Load Forecasting Method Based on HHT

Yuan Zhang<sup>1</sup>, Shu Xia<sup>1</sup>, Chihao Chen<sup>2</sup>, Fan Yang<sup>1</sup>, and Xing He<sup>2</sup>(✉)

<sup>1</sup> State Grid Shanghai Municipal Electric Power Company, Shanghai 200122, China

<sup>2</sup> Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China

xhe@sjtu.edu.cn

**Abstract.** Short-term load forecasting of the power grid can realize the optimal configuration of power generation and dispatch of the power grid which saves energy to the greatest extent and ensures the stable operation of the power system. The power load data is affected by many factors and presents complex volatility. It is difficult for a single prediction method to obtain accurate prediction results. In this paper, a combined optimization prediction method based on Hilbert-Huang transform (HHT) is proposed. By acquiring more regular component sequences of load data, its essential characteristics are explored and then combined with different neural network models for prediction to improve the accuracy and stability of short-term load forecasting. Simulation experiment results verify the prediction accuracy of the combined prediction method.

**Keywords:** Short-term Load Forecasting · Hilbert-Huang Transform · Neural Network

## 1 Introduction

With the continuous popularization and development of smart grid, the accuracy of load forecasting is vital for improving the scientificity of power generation and distribution and dispatching in the power system [1]. The short-term load forecasting is used to forecast the load data in the next day or next few days, which is affected by many factors. The forecasting method is continuously improved, and the forecasting accuracy also remains to be further improved. The method for short-term load forecasting is mainly classified as classical forecasting method and modern forecasting method. The classical forecasting method includes empirical forecasting method and traditional forecasting method, of which the machine learning algorithm is used in the modern forecasting method for analysis modeling through historical load data and relevant factors of impact load, such as air temperature and special event. The research methods include support vector machine, random forest and neural network at present [2].

In recent years, the dynamic equilibrium of power supply and demand is provided with higher requirement with the development and reform of power market, so that the requirement for accuracy of short-term load forecasting becomes higher and higher;

at the same time, with the continuous renewal of research method in short-term load forecasting field, many experts put forward the method for combination forecasting. The combination forecasting model not only overcomes the limitation of single model algorithm, but also realizes the complementary advantages between different neural networks by virtue of multi-feature characteristics of load data. It is widely applied and deeply researched in the short-term load forecasting field at present. The common thoughts in combination forecasting model include: firstly, combine with different neural network models to realize combination forecasting for load data, such as CNN-LSTM and GRU-NN combination forecasting model [3, 4]; secondly, preprocess the original load sequence, take different methods to extract features, then establish the forecasting model respectively, such as variational mode decomposition (VMD) and local mean decomposition (LMD) selected [5, 6], and execute load forecasting in combination with different forecasting models.

On the basis of existing scientific research achievements of combination forecasting method and characteristics of periodicity, non-linearity, non-stationarity and strong randomness of short-term load forecasting sequence, a short-term load combination forecasting method for power grid based on Hilbert-Huang Transform is proposed in this paper. HHT is a time-frequency analysis method characterized by self-adaptability and decomposing signal locally. Firstly, the load data is decomposed by empirical mode decomposition (EMD) algorithm to get IMF component and get different components of instantaneous frequency by Hilbert conversion; secondly, the different load forecasting models are selected to forecast the high-frequency, medium-frequency and low-frequency components respectively. Such a combination model can not only make full use of the most of character of different components, but also realize the complementary advantage between different forecasting models. The rationality and effectiveness of the combination model are validated by test in this paper.

## 2 Theoretical Basis of Combination Model

### 2.1 Empirical Mode Decomposition (EMD)

HHT include EMD and Hilbert conversion. EMD decomposition is the core part of HHT conversion, as well as the truly innovative point. EMD process is an iterative process of data processing by envelope fit by extreme value. After the data column iterated meets the certain conditions, it will become the intrinsic mode function (IMF), calling as screening. IMF is characterized by two points: 1) In sequence, the number of extreme point is equal to that of zero crossing point or the difference is not greater than 1; 2) At any time, the local value of upper envelope and lower envelope defined by extreme point is zero [7]. The general steps of EMD decomposition include:

- 1) The maximum value and minimum value of input signal  $x(t)$  are evaluated to fit the envelope. The curve fitting mentioned here is an important issue in *EMD*. The effect of the interpolation method selected on envelope fitting will have a direct influence on *EMD* decomposition result, and the cubic spline interpolation method is used in this paper.

- 2) The mean  $m(t)$  of upper envelope and lower envelope is evaluated, and  $h(t)$  is evaluated by  $h(t) = x(t) - m(t)$ ;
- 3) *IMF* end condition of  $h(t)$  is judged. If two features of *IMF* mentioned above are met,  $h(t)$  is the first *IMF* evaluated by decomposition, recorded as  $h_1(t)$ ; if not, Step 1 and Step 2 are cycled and repeated through assuming  $h(t) = x(t)$  as the new sequence;
- 4) The new sequence  $r(t)$  is gotten by  $r(t) = x(t) - h_n(t)$ , and the decomposition end condition is judged. If it is met, the decomposition is ended;  $r(t)$  is the residual component, and  $n$  *IMF* components are gotten by decomposition; if not, the above steps are cycled and repeated through assuming  $r(t) = x(t)$ .

A certain number of *IMF* and residual component  $r$  are gotten through decomposition of original sequence signal  $x(t)$ , of which  $x(t)$  is expressed as:

$$x(t) = \sum_{i=1}^n h_i(t) + r(t) \quad (1)$$

wherein,  $h_i(t)$  is the  $i$ *IMF* component, and  $r(t)$  is the residual component. Each *IMF* component gotten here is the independent data sequence of a characteristic scale [8].

## 2.2 Hilbert Spectrum Analysis

After *EMD* decomposition screening, *Hilbert* conversion is applied in each independent *IMF* component, so as to get the instantaneous frequency and instantaneous amplitude to analyze the components. *Hilbert* conversion can realize the  $90^\circ$  phase shift of base frequency and harmonic wave accurately, and remain constant amplitude [9]. With regard to the given signal  $x(t)$ , *Hilbert* conversion may be defined as:

$$H[x(t)] = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(\tau)}{t - \tau} d\tau = x(t) * \frac{1}{\pi t} \quad (2)$$

The analytic signal is gotten through *Hilbert* conversion for signal  $x(t)$ :

$$z(t) = x(t) + jH[x(t)] = a(t)e^{j\theta(t)} \quad (3)$$

So the instantaneous frequency  $\omega$ :

$$f_i = \frac{1}{2\pi} \omega_i(t) = \frac{1}{2\pi} \frac{d\theta_i(t)}{dt} \quad (4)$$

Through *Hilbert* conversion, the analytic function of each *IMF* component  $h_i(t)$ :

$$z_i(t) = h_i(t) + j\tilde{h}_i(t) = a_i(t)e^{j\theta_i(t)} \quad (5)$$

The instantaneous frequency and instantaneous amplitude of each *IMF* component are gotten from Eq. (4, 5). In essence, *Hilbert* conversion shows the optimal approximation degree of local signal and sine function, and its localization feature is further reinforced in the differential operation of solving instantaneous frequency [10].

### 2.3 RBF Neural Network

*RBF* neural network can approximate arbitrary nonlinear function with good function approximation function, and is widely applied in the load forecasting aspect by virtue of characteristics of simple structure and rapid learning convergence rate. *RBF* neural network based on Gaussian kernel is used [11]. It is assumed that the input vector is  $n$  dimension, and is denoted as  $x = (x_1, x_2, \dots, x_n)^T$ . In addition, there are  $k$  hidden nodes and  $m$  outputs in the model.  $h_i(x)$  represents the  $i$  hidden layer node. The Gaussian function is used for conversion of space mapping of input information as the kernel function of hidden layer neuron:

$$h_i(x) = \Phi_i(x) = e^{-\frac{x^2}{\delta_i^2}} \quad (6)$$

$\delta$  is the extended constant. When the vector is input to neural network through *Gaussian* radial basis function, the output of the  $j$  node of hidden layer:

$$\Phi_i(x) = \exp\left(\frac{-\|x_i - c_j\|}{2\sigma_j^2}\right) \quad (7)$$

wherein,  $c_j$  is the center of Gaussian function of the  $j$  hidden layer;  $\|\cdot\|$  is Euclidean norm, and  $\sigma_j$  is the width of Gaussian function of the  $j$  hidden layer. The output of *RBF* neural network:

$$Y = (y_1, y_2, \dots, y_m)^T = \sum_{j=1}^m w_j \Phi_j(x) \quad (8)$$

wherein,  $w_j$  is the network connection weight between the  $j$  hidden layer node and output layer [18]. *RBF* neural network forecasting model can turn the nonlinear mapping from input layer to hidden layer into the linear mapping on the other space, and forecast the signal with high frequency, large volatility and strong randomness better [19].

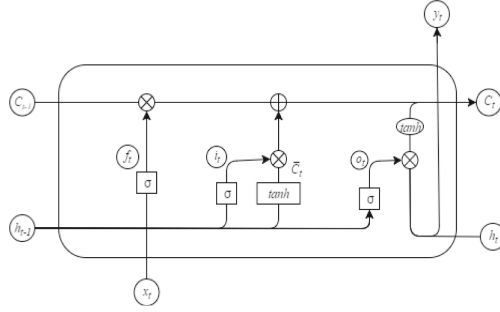
### 2.4 LSTM Recurrent Neural Network

LSTM is an improved structure proposed for easy gradient vanishing and gradient explosion of common *RNN* in practical training. It is a mechanism which leads in cell gate from neuron of standard *RNN* model, which consists of input gate, output gate and forget gate [12]. The forget gate is used to decide the forget and update of transitional information. LSTM cell structure is as shown in Fig. 1.

LSTM model can decide which information is forgotten and updated to constitute the long-term and short-term memory network through gate mechanism in cell structure. According to LSTM cell structure chart,  $C_t$  is the cell state at  $t$  time;  $x_t$  is the input at  $t$  time;  $h_t$  is the output at  $t$  time, and  $f_t$ ,  $i_t$  and  $o_t$  are output of forget gate, input gate and output gate respectively. The operation process of concrete cell structure:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (9)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (10)$$



**Fig. 1.** LSTM Unit Structure Diagram.

$$\bar{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (11)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (12)$$

$$C_t = f_t C_{t-1} + i_t \bar{C}_t \quad (13)$$

$$h_t = o_t \tanh(C_t) \quad (14)$$

wherein,  $W_f$ ,  $W_i$ ,  $W_c$  and  $W_o$  are weight matrixes, and  $b_f$ ,  $b_i$ ,  $b_c$  and  $b_o$  are bias vectors. *LSTM* model is optimized by “forget gate” additionally, so as to control the convergence of gradient during training data, and solve the gradient vanishing or gradient explosion better.

### 3 Short-Term Load Forecasting Combination Model Based on HHT

The short-term load data of power grid is affected by human production and life, change of meteorological condition, economic factor, political factor, etc. The system load data includes multiple characteristics for analysis and forecasting, and it is difficult to obtain the essential characteristics. In order to further explore the inherent law of load data, the short-term load forecasting combination model of power grid based on *HHT* is established to decompose the load data as a certain amount of *IMF* by *EMD* algorithm, and then convert and process each component alone by *Hilbert* conversion, so as to get different instantaneous frequencies and instantaneous amplitudes. According to different characteristics of *IMF*, the different neural network models are selected for forecasting, and the result is overlapped to get the forecasted value of load in the end. In the meantime, due to large influence of change of air temperature on fluctuation of load data, the accuracy of load forecasting is promoted in combination with the correlation between temperature data of the region and *IMF* component.

### 3.1 Hilbert-Huang Transform (HHT) of Load Data

The data sample in March 2021 of a region in East China is selected for test. The load data curve is shown in Fig. 2. Firstly, the load sequence is provided with *EMD* decomposition, then the envelope is fitted with cubic spline interpolation method, and a total of 7 *IMF* components and a residual component *r* are decomposed. The concrete result is as shown in Fig. 3.

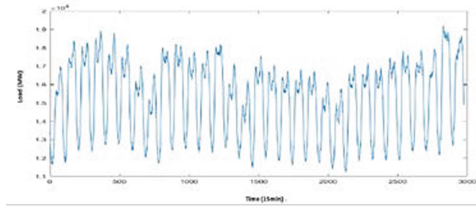


Fig. 2. Load data curve (March, 2021)

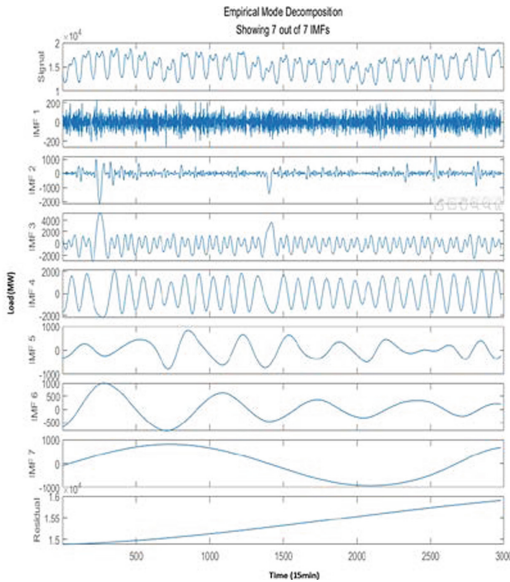
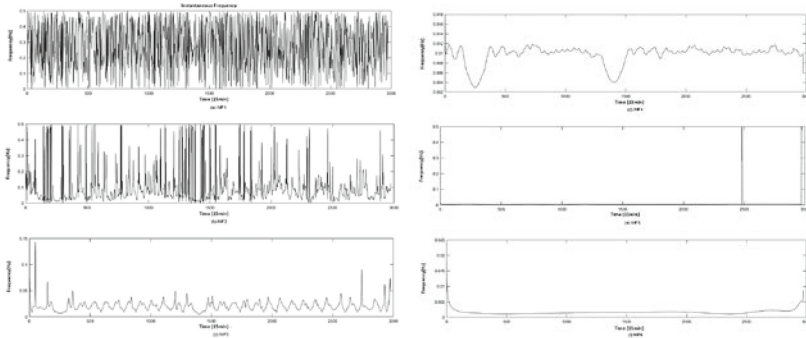


Fig. 3. EMD decomposition result

The frequency of *IMF*<sub>1</sub> component and *IMF*<sub>2</sub> component is high in Fig. 3, but that of *IMF*<sub>3</sub> to *IMF*<sub>7</sub> is decreased progressively in contrast. In order to further analyze each component, *Hilbert* conversion is also applied to obtain the concrete instantaneous frequency curve chart of each component, as shown in Fig. 4.

The mean frequency of each *IMF* component is further calculated in Table 1. It shows that each *IMF* component owns different frequency characteristics in Fig. 4 and



**Fig. 4.** Instantaneous frequency of *IMF* component

Table 1. It is decreased progressively, and the mean value calculated is also reduced in turn. *IMF*<sub>1</sub> to *IMF*<sub>3</sub> is characterized by large fluctuation, strong randomness and high frequency upon calculation as the random part of load; *IMF*<sub>4</sub> to *IMF*<sub>5</sub> is characterized by steady fluctuation trend and mean frequency decrease, which represents the periodicity of load; eventually, according to the calculation result of mean frequency in Table 1, *IMF*<sub>6</sub>, *IMF*<sub>7</sub> and residual term *r* approach to zero, which represents the trend component of load. The components are divided into random component, periodic component and trend component according to their characteristics. On one hand, they reduce the difficulty of building forecasting model; on the other hand, they emphasize on different characteristics of each component. The model is built by combination of *RBF* neural network model and recurrent neural network based on *LSTM*. It not only takes advantage of high learning rate of *RBF* neural network to process the data signal with large volatility and high frequency, but also processes the problems with strong periodicity and highly correlated with time sequence in combination with *LSTM* to effectively improve the forecasting accuracy.

**Table 1.** Average frequency of *IMF* component

Component	<i>IMF</i> <sub>1</sub>	<i>IMF</i> <sub>2</sub>	<i>IMF</i> <sub>3</sub>	<i>IMF</i> <sub>4</sub>	<i>IMF</i> <sub>5</sub>	<i>IMF</i> <sub>6</sub>	<i>IMF</i> <sub>7</sub>
Frequency (Hz)	0.2792	0.0992	0.0204	0.0097	0.005	0.0016	0.0006

### 3.2 Correlation Analysis of *IMF* Component Temperature

The power load data includes multiple properties of power utilizations, i.e., industrial load, appliance load and transportation load. Different *IMF* components represent different properties of power utilization data, and the meteorological influence is also different, so it should be analyzed in preliminary data processing in combination with meteorological factors. For example, the air temperature is selected as the representation of meteorological factors, the meteorological factor is integrated to adjust the input

data and model parameter through comparison with correlation between different *IMF* components and temperatures as well as neural network modeling for different *IMF* components.

The correlation coefficient of each *IMF* component and air temperature is defined as

$$r_i = \frac{\text{cov}(h_i(t), c_i(t))}{\sqrt{\text{cov}(h_i(t))}\sqrt{\text{cov}(c_i(t))}} \quad (15)$$

wherein,  $c_i(t)$  is air temperature of the corresponding point ( $i = 1, 2, \dots, n$ ;  $n$  is total number of *IMF* components). The difference in correlation of different seasons is obvious, of which the correlation coefficient of *IMF* component and temperature in summer and winter is high, and the load data in March is selected; the correlation of *IMF* component and air temperature data gotten by decomposition is small on the whole, and the curve chart of correlation coefficient in Fig. 5 is obtained. Specially, it is shown that the correlation between *IMF*<sub>1</sub> component and *IMF*<sub>2</sub> component and air temperature data scarcely exists. *IMF*<sub>3</sub> and *IMF*<sub>4</sub> are positively correlated with air temperature data, but *IMF*<sub>5</sub> and *IMF*<sub>7</sub> are negatively correlated with air temperature data, so the short-term load forecasting combination model is trained for different *IMF* components respectively. As an example of *IMF*<sub>4</sub> component, it shows that *IMF*<sub>4</sub> component is greatly affected by weather in contrast in Fig. 5, and the forecasting difficulty is high. In the short-term load forecasting combination modeling of power grid corresponding to *IMF*<sub>4</sub> component, the proportion of training data, validation data and test data is about 90%, 5% and 5% respectively.

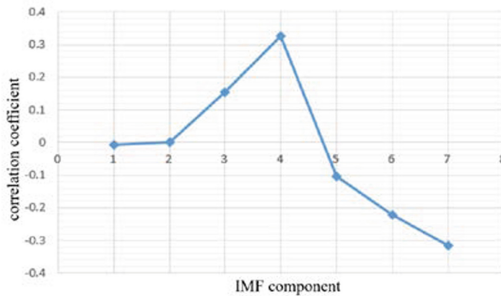


Fig. 5. Correlation between IMF components and temperature

### 3.3 Short-Term Load Forecasting Combination Model of Power Grid

The load sequence of power system is characterized by volatility and special periodicity, and it is greatly affected by actual scenes, for example, the difference of geographic position and living habit of the southern and northern China, economic and social difference of the first-tier and second-tier cities and the third-tier and fourth-tier cities will cause different periodicities and volatilities of load data due to climate, major events, electricity price fluctuation, etc. The short-term load combination forecasting model of power

grid based on *HHT* proposed herein is applied to study the essence of load data through decomposition of short-term load data of power grid, and then forecast in combination with the appropriate neural network forecasting model according to characteristics of different components, so as to improve the forecasting accuracy and stability.

The concrete steps of short-term load combination forecasting model of power grid based on *HHT* is:

- (1) Preprocess the historical load data and specify the evaluation index;
- (2) Decompose the load data by *EMD* algorithm, provide *Hilbert* conversion for *IMF* component, and get the instantaneous frequency;
- (3) Apply appropriate neural network model for forecasting respectively according to characteristics of different frequencies of each component;
- (4) Add the forecasted result of each component to get the final result;
- (5) Eventually, get the accuracy index through comparison with the non-compositional method.

### 4 Simulated Analysis

The short-term load forecasting model is applied, and the load data in March 2021 of one region in eastern China is selected to forecast the load value for 24 h on April 1, 2021 as the training sample, and analyze the accuracy of forecasted result. The curve chart of actual load and forecasted load is as shown in Fig. 6. The error of concrete forecasted result value and relative percentage of forecasting is as shown in Table 2, of which the forecasting percentage error is defined as:

$$APE = \left| \frac{A_t - P_t}{A_t} \right| \times 100\% \tag{16}$$

wherein,  $A_t$  is the real load value;  $P_t$  is the load value gotten by forecasting.

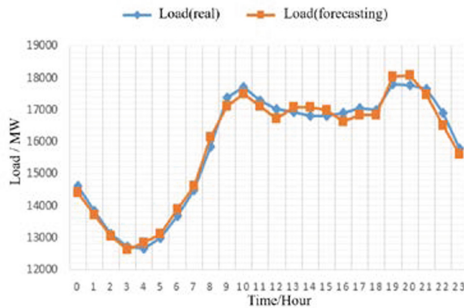


Fig. 6. Comparison between real load and forecasting load

**Table 2.** APE of forecasting load

Time/H	APE/%	Time/H	APE/%	Time/H	APE/%	Time/H	APE/%
0:00	1.52	1:00	0.78	2:00	0.64	3:00	0.67
4:00	1.34	5:00	0.96	6:00	1.48	7:00	0.82
8:00	1.89	9:00	1.45	10:00	1.17	11:00	0.97
12:00	1.78	13:00	0.91	14:00	1.74	15:00	1.05
16:00	1.62	17:00	1.20	18:00	0.95	19:00	1.34
20:00	1.74	21:00	1.06	22:00	2.27	23:00	1.13

In order to compare with other methods, the mean absolute percentage error (*MAPE*) is selected to measure the forecasted result as the evaluation index of short-term load forecasting of power grid.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - P_t}{A_t} \right| \times 100\% \quad (17)$$

The result which shows the higher forecasting accuracy and better effect of the method proposed herein is presented in Table 3 by comparison of the method proposed herein and forecasted result of *RNN* recurrent neural network and *LSTM* recurrent neural network forecasting model for mean load of one region.

**Table 3.** MAPE of short-term load forecasting

Date	Proposed method/%	LSTM/%	RNN/%
Apr. 1	1.27	2.26	2.53
Apr. 2	1.67	2.07	3.22
Apr. 3	1.36	2.72	3.05
Apr. 4	1.94	2.45	3.52
Apr. 5	1.44	2.31	2.47

In contrast with *MAPE* result based on the combination method and network model method based on *LSTM* and *RNN*, although the forecasting accuracy is fluctuated, *MAPE* index of combination forecasting method proposed in this paper is basically less than 2%, which is obviously superior to that of recurrent neural network forecasting model based on *LSTM* applied alone.

## 5 Conclusions

In this paper, the short-term load combination forecasting model of power grid based on *HHT* is studied. The original load sequence is decomposed by *EMD* decomposition

algorithm, and then each *IMF* component is changed along by *HT*. According to characteristics of different components and analysis on correlation of air temperature data, it is forecasted in combination with the forecasting model of neural network based on *RBF* and recurrent neural network based on *LSTM*. It not only takes advantage of *HHT* to process nonlinear and non-stable signal, but also realizes the complementary advantages between different neural networks. It is discovered that the accuracy of short-term load forecasting combination method of power grid based on *HHT* is higher through experimental result. Certainly, the characteristic analysis on *IMF* component should remain to be further studied and explored in contrast with correlation analysis of other factors, such as social experience factors and holiday activity factors. The combined neural network model also remains to be further improved and tried to promote the load forecasting accuracy to a greater extent.

## References

1. Yao, J.G., Yang, S.C., Gao, Z.H., et al.: Development trend prospects of power dispatching automation system. *Autom. Electr. Power Syst.* **31**(13), 7–11 (2007)
2. Osman, Z.H., Awad, M.L., Mahmoud, T.K.: Neural network based approach for short-term load forecasting. In: 2009 IEEE/PES Power Systems Conference and Exposition, pp. 1–8. Seattle, WA (2009)
3. Lu, J.X., Zhang, Q.P., Yang, Z.H., et al.: Short-term load forecasting method based on CNN-LSTM hybrid neural network model. *Autom. Electr. Power Syst.* **43**(8), 131–137 (2019)
4. Wu, L.Z., Kong, C., Hao, X.H., et al.: A short-term load forecasting method based on GRU-CNN hybrid neural network model. *Math. Probl. Eng.* **2020**, 1–10 (2020)
5. Liang, Z., Sun, G.Q., Li, H.C., et al.: Short-term load forecasting based on VMD and PSO optimized deep belief network. *Power Syst. Technol.* **42**(2), 598–606 (2018)
6. Kong, X.Y., Li, C., Zheng, F., et al.: Short-term load forecasting method based on empirical mode decomposition and feature correlation analysis. *Autom. Electr. Power Syst.* **43**(5), 46–51 (2019)
7. Huang, N.E., Shen, Z., Long, S.R., et al.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In: *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, pp. 903–995 (1998)
8. Zhu, Z.H., Sun, Y.L., Ji, Y.: Short-term load forecasting based on EMD and SVM. *High Volt. Eng.* **33**(5), 118–122 (2007)
9. Huang, C.G.: Hilbert transform and its applications. *J. Chengdu Inst. Meteorol.* **14**(3), 273–276 (1999)
10. Munir, B.S., Reza, M., Trisetayrso, A., et al.: Feature extraction using Hilbert-Huang transform for power system oscillation measurements. In: 2017 4th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), pp. 93–96. Semarang, Indonesia (2017)
11. Schilling, R.J., Carroll, J.J., Al-Ajlouni, A.F.: Approximation of nonlinear systems with radial basis function neural networks. *IEEE Trans. Neural Netw.* **12**(1), 1–15 (2001)
12. Gers, F.A., Schmidhuber, J., Cummins F.: Learning to forget: continual prediction with LSTM. In: *Ninth International Conference on Artificial Neural Networks (ICANN)*, pp. 850–855. Edinburgh, UK (1999)