Power System Middle-term Load Forecasting Based on CEEMD with Fuzzy Entropy and Elman ANN

XIN-YOU WANG [1, a], CHEN-HUA WANG[2], LE-BIN BAI[1] , and ZHAN QU[1]
[1]Institute of Technology, Gansu Radio & TV University, Lanzhou 730030, P.R.CHINA
[2]GanSu Datang Renewable Power Co., Limited, Lanzhou 730030, P.R.CHINA
[a]wangxiny@gsrtvu.cn

Abstract: In this paper, a new combination prediction approach was proposed and applied to improve the middle-term electric load forecasting precision. Firstly, the load sequences were decomposed into a limited number of load sub-temporal sequences with different characteristics, which avoids the large computing scale problems of local analysis of load series. Then, the Elman prediction models were constructed by the feature analysis for each sub-temporal sequence respectively. And the final prediction valves were given by the superposition of each sub sequences prediction. The approach is applied to EUNITE's middle term load forecasting. A comparison of our approach to existing prediction methods is also given. Simulation results show that the new model proposed in this paper can significantly improve the load prediction accuracy.

Key-Words: middle-term load forecasting; Complementary Ensemble Empirical Mode Decomposition (CEEMD); Fuzzy Entropy; Elman Neural Network

0 Introduction

In fact, the essence of the load forecasting is modeling the history load data to express the load rules, in other words that is a prediction to power requirements. As the development of power system, more and more researchers realized that promoting the accuracy of the middle-term forecasting is an important guarantee to improve the economical efficiency and enhance the reliability of power system operation [1-2].

Until now, there are many traditional power load forecasting methods such as regression method, time series modeling method, trend analysis[3-4]. Those methods needs mathematics statistics and cannot satisfy the need of time-varying requirements of the power load sequences. Modern prediction methods including grey forecasting method, fuzzy prediction, and support vector machine (SVM) [5-8].Among this, back propagation(BP) neural network is the most popular algorithm[9]. However, due to the structure flaws of BP neural network, its slow convergence and local minimum limit their more widespread application in this area. Meanwhile, it's hard for the BP ANN to get better identification result for the power load forecasting, because the sample data of the prediction is a random variation.

In Elman network, context units save previous outputs values of hidden layer neurons, and all signals from neuron outputs are connected as network inputs through unit time delay, and thus achieve a low resolution memory. Therefore, both the stability and the ability to deal with dynamics information have been improved[10].

Moreover, in order to improve the load forecasting accuracy, the variation of the load forecasting should be found upon the analysis of the main factors that influence power prediction. Load sequences are kinds of nonstationary time series characterized by periodic components and random variation. In recent years, a new signal processing method, empirical mode decomposition (EMD), is suitable to analyze instability and non-linearity time series. But, mode mixing occurred in EMD sometimes, and that makes the physical meaning of individual IMF unclear[11-13].

The Ensemble Empirical Mode Decomposition (EEMD) is a substantial improvement over the original EMD. It utilizes the full advantage of the statistical characteristics of white noise to perturb the signal in its true solution neighborhood, and effectively overcome the scale separation problem without introducing a subjective intermittence test in EMD [14-17]. However, With the help of added noises, EEMD essentially resolved the mode mixing problem associated with EMD, but the resulting IMFs derived from EEMD would inevitably be contaminated by the added noise especially when the number of ensemble was relatively low. This is especially true in the reconstruction of the signal from the IMF components, though this affects can be reduced by large scale means, it's much time-consuming. To improve the efficiency of the original noise assisted algorithm of EEMD, the CEEMD approach was proposed [NORDEN E.HUANG *et al*. (2010)]. Unlike the EEMD, CEEMMD uses each noise in pairs with plus and minus signs [18]. The advantage for this new approach, however, is to have an exact cancellation of the residual noise in the reconstruction of the signal with low ensemble numbers.

This paper proposes a model of middle-term load forecasting based on CEEED,Fuzzy Entropy and Elman neural network. Firstly, we applied the method of CEEED to decompose the power load sequences into a number of stationary series and a certain trend in weight based on its characteristics. Then, the complexity of each IMF was measured by Fuzzy Entropy, and new sub-sequences were combined by the Fuzzy Entropy value. Considering the weather, day types and other factors, different Elman neural network model will be constructed and applied to the analysis of those sub-sequences. Finally, we will get the forecasting results by synthesizing the prediction of each component. The simulation results show that the method is of high precision and strong adaptability.

1 EEMD and CEEMD

1.1 EEMD Theory

The Ensemble Empirical Mode Decomposition (EEMD) [ZHAOHUA WU *et al*. (2009)] is a noise assisted method designed for eliminating the mode-mixing problem caused by intermittence signals in Empirical Mode Decomposition (EMD) method.To decompose the true signal of any complex data, this approach effectively resolved this problem by utilizing the full advantage of the statistical characteristics of white noise. That makes the signal continuous in different scales, and different scales of the signal region will be automatically corresponded to the suitable scale, so the effect of the data analysis is enhanced by EEMD.Undoubtedly, EEMD is a breakthrough in the development of EMD algorithm in recent years, its process is as follow:

(1) Assign the EMD algorithm running circles N, assign the amplitude coefficient of white noise k,n=1. And execute m times EMD.

(2) Add a white noise $l_n(t)$ to the targeted data $x(t)$,gets the noise added signal $x_n(t)$;Decompose the data $x_n(t)$ with added white noise into I $(\mathrm{IMF})_{j,n}$, where j,n is the j-th IMF in n-th decomposition; repeat(2)if $n<N$, let $n=n+1$.

(3) Obtain the means of corresponding IMFs of the decompositions as the final result.$C_j(t)$ is the j-th IMF decomposed from the original signal by EEMD, and N is the number of white noise added.

1.2 CEEMD Theory

In contract with EEMD, the decomposition process is similarly, the difference are using each noise in pairs with plus and minus signs， the final IMF is the average of the IMFs with positive and negative noises.

Practically, the two complementary signals can be derived by

$$\begin{bmatrix} \boldsymbol{M}_+ \\ \boldsymbol{M}_- \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \boldsymbol{S} \\ \boldsymbol{v} \end{bmatrix} \tag{1}$$

where S is the original data; v is the added white noise; $\boldsymbol{M}_+$ is the sum of the original data with positive noise, and $\boldsymbol{M}_-$ is the sum of the original data with the negative noise; $\boldsymbol{M}_+$ and $\boldsymbol{M}_-$ have the same amplitude, with a 180° phase angle difference.

2 Fuzzy entropy

Fuzzy entropy (FuzzyEn) [Chen et al. (2007)] is a measure of time series regularity,which import the concept of fuzzy sets[19-20].In FuzzyEn,the removing means method algorithm are used in vector reconstruction formula, and the family of exponential functions $\exp(-d_{ij}^m)/r$ are employed as similarity degree measurement formulas. In this way, entropy values of FuzzyEn will change continuously and gracefully with parameters, and it makes the similarity definition fuzzier. The vectors' similarity in FuzzyEn is determined by the distance d_{ij}^m between the vectors rather than their absolute coordinates in ApEn and SampEn, thus the poor analytic accuracy in case signals with mild fluctuations or baseline deviation can be avoided. The definition of FuzzyEn is as follows:

For an sample time series $\{u(i):1\le i\le N\}$ with N samples, given m, form vector sequences in phase space as follows:

$$\mathrm{X}_i^m = \{u(i), u(i+1), u(i+m-1)\} - u_0(i) \tag{2}$$
$$i = 1,2,...n+m-1$$

where $\boldsymbol{X}_i^m$ represents m consecutive u values, commencing with the i-th point and generalized by removing a baseline

$$u_0(i)=\frac{1}{m}\sum_{j=0}^{m-1}u(i+j) \tag{3}$$

For fixed vector X_i^m , define the distance d_{ij}^m between X_i^m and X_j^m as the maximum absolute difference of the corresponding scalar components:

$$d_{ij}^m=d\left[X_i^m,X_j^m\right]= \max_{k\in(0,m-1)}\left(\left|u(i+k)-u_0(i)-u(j+k)-u_0(j)\right|\right) \tag{4}$$
$$i,j=1,2,\ldots,N-m; i\neq j$$

Given n and r, calculate the similarity degree D_{ij}^m of X_j^m to X_i^m by a fuzzy function $\mu(d_{ij}^m,n,r)$

$$D_{ij}^m(n,r)=\mu(d_{ij}^m,n,r)=\exp(-(d_{ij}^m)^n/r) \tag{5}$$

where the fuzzy function $\mu(d_{ij}^m,n,r)$ is the exponential function; n and r are the gradient and width of the boundary of the exponential function respectively.

Define the function ϕ^m as

$$\phi^m(n,r)=\frac{1}{N-m}\sum_{i=1}^{N-m}\left(\frac{1}{n-m-1}\sum_{j=1,j\neq i}^{N-m}D_{ij}^m\right) \tag{6}$$

Similarly, form $\left\{\mathbf{X}_i^{m+1}\right\}$ and get the function ϕ^{m+1}

$$\phi^{m+1}(n,r)=\frac{1}{N-m}\sum_{i=1}^{N-m}\left(\frac{1}{n-m-1}\sum_{j=1,j\neq i}^{N-m}D_{ij}^{m+1}\right) \tag{7}$$

Finally, we can define the parameter *FuzzyEn* (m,n,r) of the sequence as the negative natural logarithm of the deviation of ϕ^m from ϕ^{m+1}

$$FuzzyEn(m,n,r)=\lim_{N\to\infty}\left(\ln\varphi^m(n,r)-\ln\varphi^{m+1}(n,r)\right) \tag{8}$$

which, for finite datasets, can be estimated by the statistic

$$FuzzyEn(m,n,r,N)=\ln\phi^m(n,r)-\ln\phi^{m+1}(n,r) \tag{9}$$

Except N,the length of dataset ,there are three parameters m,r,n that are relevant to each calculation of FuzzyEn.Typically, a larger m allows a longer dataset, that is a v ery large N($N=10^m\sim 30^m$) is needed.

Experimentally, r is set as $0.1\sim 0.5SD$ standard deviation (SD), n is choose as 2or 3.

3 The Elman Network Theory

As a typically globally feed forward locally recurrent network ,the Elman network [Jeffrey L. Elman *et al.* (1990)] contains inputs layer, the hidden layer, the network output, and the particular context layer, its expression in non-linear state space is as follows:

$$y(k)=g(w^3x(k)) \tag{10}$$

$$x(k)=f(w^1x_c(k)+w^2(u(k-1))) \tag{11}$$

$$x_c(k)=x(k-1) \tag{12}$$

Where k stands for time, y stands for m-dimensional vectors of output node, x stands for n-dimensional element vector of interlayer, u stands for r-dimensional input vectors,x_c stands for n-dimensional feedback state vectors.ω^3 stands for the connection weight between the interlayer and the output layer,ω^2 stands for the connection weight between the input layer and the interlayer,ω^1 stands for the connection weight between the context units layer and the interlayer. $g(\ldots)$ stands for the transfer function of output neurons, $u(\ldots)$ stands respects for the linear combination of the output of the interlayer. $f(\ldots)$ stands for the transfer function of interneuron, generally, s-function are employed.

By the delay and store of the context layer, every hidden layer output is connected to every hidden unit input in this neural networks model. Therefore, this model is more activate to history data and dynamics modeling achieved, and its ability to deal with dynamics information has enhanced for interact feedback network exclusively engaged. The architecture of the Elman network is shown in figure 1.

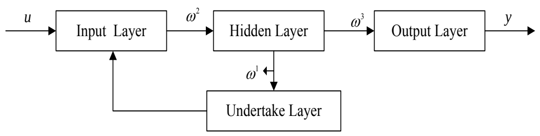


Fig.1 The Structure of Elman ANN

The weights in Elman networks can be revised(updated) by back- propagation (B P)BP algorithm.Error sum squares functions was taken as learning indicator function

$$E(\omega)=\sum_{k}^{n}\left[y_k(\omega)-y_k^*(\omega)\right]^2 \tag{13}$$

Where $y^*\,k(\omega)$ is target output vector.

4 Simulation of Middle-term Load Forecasting

The power system load data, which is used to forecast the middle-term, comes from World-wide competition within the EUNITE network http://neuron.tuke.sk/competition/index.php. That's contains the every 30 minutes electricity load data, daily average temperature, and the holidays information in East-Slovakia Power Distribution

Company in years 1997 and 1998.The 31 max. Load values for January 1999 are expected to be predicted.

In this section, calendars, holidays etc. as t he characteristic information were used as prediction input, forecasting model was made by time series modeling method. To avoid the difficulty that the temperature prediction bring to the whole forecasting, the training data set were chosen by data segmentation method to make up the influence of weather to the prediction. Thereby, the data in January to March, October to December in years 1997 and 1998 were employed to prediction.

Simulations were done with ELMAN network model, which takes the form as

$$y(t) = f_{i(t)}(x_t), \forall t = \Delta \ldots l \tag{14}$$

Where *i(t)* represents different time period after data decomposition, Δ represents embedding dimension, x_t contains historical load data, calendar and holidays in formation. For given y_t , where t=1…l, the task is modeling with different f_i algorithm in each time period, and segment the data in accordance with the prediction day membership of climatic season.

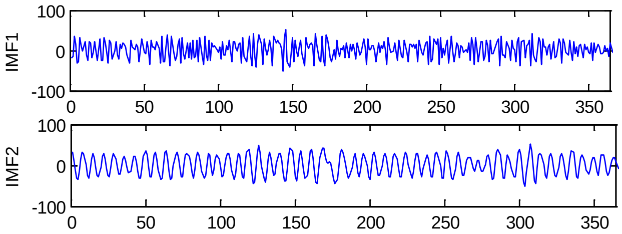
Maximal error (ME), mean absolute percentage error (MAPE) and mean square error (MSE) were used as performance measures.

$$\mathrm{MAPE} = 100 \frac{\sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right|}{n}, \qquad \mathrm{ME} = \max \left| y_i - \hat{y}_i \right|$$

Where y_i represents the actual load value, $\hat{y}_i$ represents the prediction value, n represents the prediction days.

Multi-step iterative prediction was employed in the experiments. Therefore, the one-step predictive model, model (14) which is constructed by ELMAN network, was used for training. In addition, the output is used as the input for next prediction recursively until the well test finished in testing process.

Consider the instability of the load time series, the original load time series were decomposed by CEEMD, and 200 pairs white noise sequences which are set at a standard deviation of 0.2 were added. The IMFs decomposed are shown in Figure.2



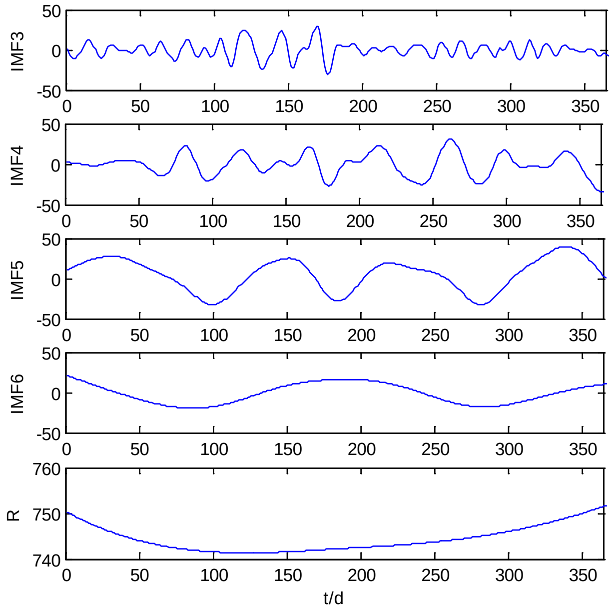


Fig.2. The decomposed IMFs by CEEMD

Figure 3 shows the reconstruction error of EEMD and CEEMD algorithm respectively. Through Fig3, the reconstruction errors of CEEMD are close to 0(of the order of 10^{-7}, determined by the calculation accuracy of computers) with 200 pairs added white noise. Whereas, due to the affects of the ensemble times with small *M*, the reconstruction errors of EEMD is much larger than those in CEEMD algorithm, which will partly affect the integrity of the decomposition.

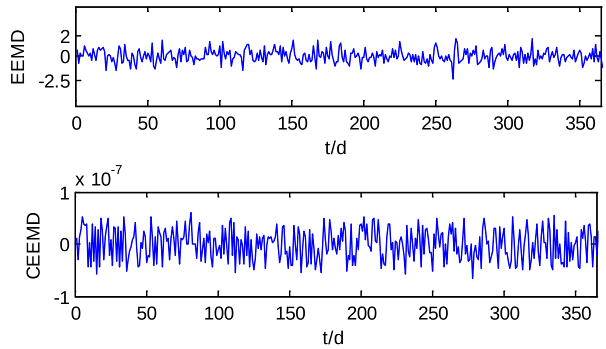


Fig.3. The reconstruction errors of EEMD and CEEMD (a)EEMD ; (b)CEEMD

Through Figure.3, each decomposed IMFs shows strong regularity and non-stationary. However, much IMF component generated, the computation scale increased badly if the ELMAN network was used to forecast each IMF component. In order to effectively predict the electrical load, FuzzyEn theory were employed to analyze the complexity of each IMF component in this research.FuzzyEn entropy values of each IMF are shown in Figure.4.

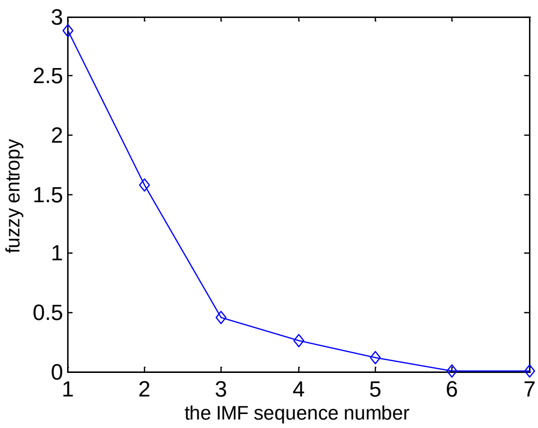


Fig.4. FuzzyEn entropy values of each IMF

As shown in Figure.4, the FuzzyEn entropy values of each IMF decrease with the reducing of frequency of IMF component, illustrate the complexity of IMF component goes down from high frequency to low frequency, thus the effectiveness of FuzzyEn theory get verified. Moreover, the differences of partly neighboring FuzzyEn entropy values were little. Therefore, those IMF components can be combined to reduce the computation effort. The combination results are listed in Table 1. The reconstruction sequences are shown in Figure 5.

Tab.1. The new IMF component after combination with FuzzyEn theory

SN of the New IMF components	1	2	3	4
SN of the original IMF components	1	2	3,4	5,6,7

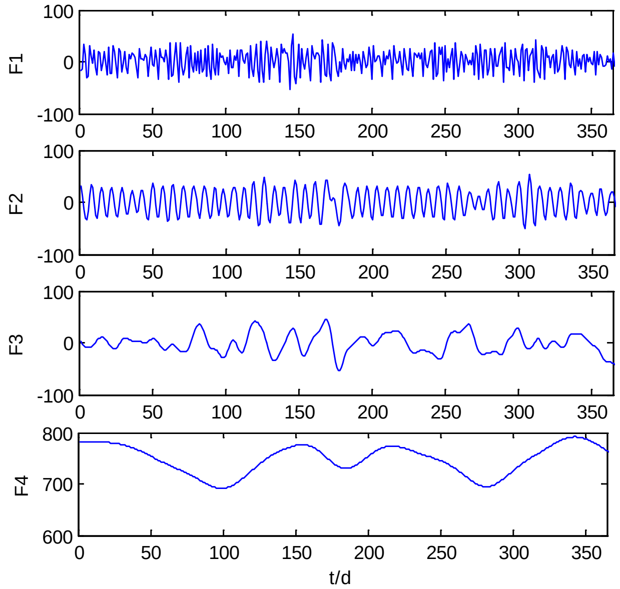


Fig.5. The reconstruction component after CEEMD-FuzzyEn treating

In view of the four new sequences after combination, four ELMAN forecasting model were built for prediction. In the sub sequence model, calendar and holidays information were considered. Calendar information was represented by 7 binary code, and the holidays information were represented by 0 or 1. The steps forecasting model proposed in the paper are shown in Fig.6.

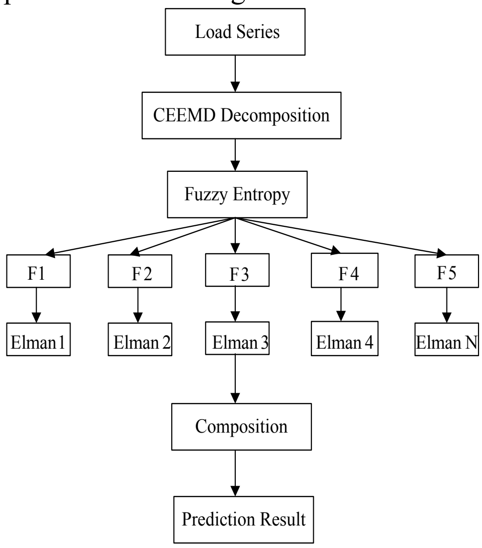


Fig.6. Hybrid forecasting model of CEEMD and ELMAN numerals network

Then, the treated data mentioned above are inputted in the trained ELMAN numerals network to get the relative predictive value. And, the ultimate forecasting value is given by the superposition of every forecasting result. Under the same situation, the ELMAN algorithm, EEMD-Elam algorithm were used for comparison, and their forecasting effects are shown in Fig.8.Through Fig.8, it is clear that the CEEMD-FuzzyEn combined with Elman achieves a better forecasting results.

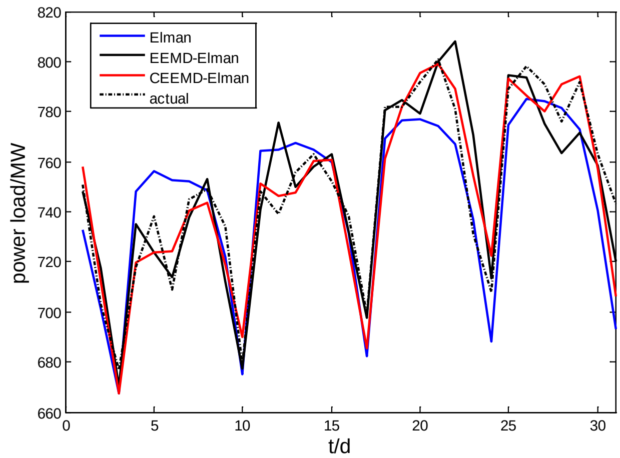


Fig.7 The forecasting result of different methods

In order to further evaluate the effectiveness of the method proposed in this paper, Table 2 demonstrates the evaluation index of MAPE and ME in the three forecasting approaches,respectively. It can be noticed that the CEEMD-Elman method has the minimum value of MAPE and ME, and the forecasting effect are better than the other two approaches.

Tab.2.Comparison of the forecasting error of the three models

Model	MAPE	ME
Elman	1.99	49.96
EEMD-Elman	1.51	39.53
CEEMD-Elman	1.34	36.80

5 Conclusions

A new power system load middle-term forecasting based on CEEMD-FuzzyEn and Elman numerals network is proposed in this paper. The load sequences are decomposed into a s eries of IMF components with different scales by CEEMD,aiming at its non-stationary property. Furthermore,the FuzzyEn entropy theory is employed in the reconstruction of IMF components to decrease the computation in the modeling of each sub-sequences. Then, the Elman forecasting model are built for each sub-sequences reconstructed,respectively. Finally,the predictive value of those components are summed to get the ultimate predictive result.The simulation results illustrate this hybrid forecasting algorithm has the best prediction precision performance,compared with Elamn algorithm and CEEMD-Elamn algorithm. Thus,this method proposed has good practical value and explore a n ew effective way for power system load middle-term forecasting.

References:

[1] Nihuan L, Zhihong H, Yingying M, et al. Review of the short-term load forecasting methods of electric power system[J]. Power system protection and control, 2011, 39(1): 147-152.
[2] Hippert H S, Pedreira C E, Souza R C. Neural networks for short-term load forecasting: A review and evaluation[J]. Power Systems, IEEE Transactions on, 2001, 16(1): 44-55.
[3] Song K B, Baek Y S, Hong D H, et al. Short-term load forecasting for the holidays using fuzzy linear regression method[J]. Power Systems, IEEE Transactions on, 2005, 20 (1): 96-101.
[4] Amjady N. Short-term hourly load forecasting using time-series modeling with peak load estimation capability[J]. Power Systems, IEEE Transactions on, 2001, 16(3): 498-505.
[5] Yao A W L, Chi S C, Chen J H. An improved grey-based approach for electricity demand forecasting[J]. Electric Power Systems Research, 2003, 67(3): 217-224.
[6] Yu Haiyan，Zhang Fengling.Short-term load forecasting based on fuzzy neural network[J] . Power System Technology,2007,31(3)：68-72．
[7] Chen B -J.,Chang M –W.,Lin C -J..Load forecasting using support vector machines:A study on E UNITE competition 2001[J].IEEE Transactions on Power Systems,2004,19(4):1821-1830.
[8] Zhu Zhihui,Sun Yunlian,Ji Yu.Short-term load forecasting based on empirical mode decomposition and least square support vector machine.[J].Relay,2007,35(8):37-40．
[9] Xiao Z, Ye S J, Zhong B, et al. BP neural network with rough set for short term load forecasting[J]. Expert Systems with Applications, 2009, 36(1): 273-279.
[10] Wu J G, Lundstedt H. Prediction of geomagnetic storms from solar wind data using Elman recurrent neural networks[J]. Geophysical research letters, 1996, 23(4): 319-322.
[11] Huang N E，Shen Z，Long S R．The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis [C]．Proceedings A of the Royal Society，1998，454（1997）：903-995．
[12] Guo Z, Zhao W, Lu H, et al. Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model [J]. Renewable Energy, 2012, 37(1): 241-249.
[13] ZHU Zh-i hui, SUN Yun-lian, JI Yu. Short-term Load Forecasting Based on EMD and SVM[J]. High Voltage Engineering, 2007, 33(5): 118-122.
[14] Wu Z, Huang N E. Ensemble empirical mode decomposition: a noise-assisted data analysis method [J]. Advances in Adaptive Data Analysis, 2009, 1(01): 1-41.
[15] LIU Dai, PANG Song-Ling, LUO Wei. Power system short-term load forecasting based on EEMD and dynamic neural network[J].Journal of Northeast DianLi University, Natural Science Edition, 2009, 29(6): 20-26.
[16] WANG H, HU Z, CHEN Z, et al. A Hy brid Model for Wind Power Forecasting Based on

Ensemble Empirical Mode Decomposition and Wavelet Neural Networks [J]. Transactions of China electro technical Society, 2013, 2 8(9): 137-144.)

[17] Zhang X，Liang J. Chaotic time series prediction model of wind power based on ensemble empirical mode decomposition-approximate entropy and reservoir [J]．Acta Physica Sinica，2013，62（5）：1-10．

[18] Yeh J R, Shieh J S, Huang N E. Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method[J]. Advances in Adaptive Data Analysis, 2010, 2(02): 135-156.

[19] Chen W, Wang Z, Xie H, et al. Characterization of surface EMG signal based on fuzzy entropy [J]. Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 2007, 15(2): 266-272.

[20] Chen Xiao-jun, Li Zan. A New Complexity Metr ic of Chaotic Pseudorandom Sequences Based o n Fuzzy Entropy [J]. Journal of Electronics & In formation Technology,2011,33(5):1198-1203.