Electric Load Combination Forecast Method Based on EEMD

Yunfeng Shao¹, Yajing Wang², Yuanming Sun², Zhongjing Ma²⁺, Yang Zhao¹ and Yongqiang

Liu¹

¹ Lvliang Power Supply Company, State Grid Shanxi Electric Power Company, China ² Beijing Institute of Technology, China

Abstract. Load forecasting is of great significance to improve power system safety and reliability. Aiming at the problems of low electric load forecast accuracy and strong randomness, a combined load forecast method based on ensemble empirical mode decomposition is proposed. First, ensemble empirical mode decomposition is used to decompose the load data into intrinsic mode functions with different frequencies, and the sample matrix is formed according to decomposed components. Then, principal component analysis is used to construct a transformation matrix which is used to reduce the noise of the sample matrix, unit root test is used to judge the stability of each component of the sample matrix after noise reduction. If the component is judged to be stationary, multiple linear regression is used to forecast. If the component is greated to get the final load forecast result. Based on the proposed method, the load of a certain area in Shanxi is forecasted and compared with other methods. The results show that this method can forecast the load more effectively while reducing the noise of the load.

Keywords: load forecasting, EEMD, PCA, LSTM

1. Introduction

Load forecasting has an important position in the operation of modern power systems, and accurate forecasting is of great significance for ensuring the safe and stable operation of the power grid and improving economic benefits [1]. The electrical load has a certain periodicity, but the collected load data often contain noise. Noise is generated by factors such as weather, holidays, and equipment accuracy, which affects the trend of load to a certain extent, making load exhibits strong randomness and becomes no longer stable, which brings certain difficulties to short-term forecasting.

At present, there are a variety of methods used in load forecasting, which can be roughly divided into three categories: traditional methods, machine learning methods, and combined methods. Traditional forecasting methods based on statistical methods, such as time series [2], linear regression [3], etc., have simple models and fast prediction speeds but are susceptible to random noise interference [4], therefore it is particularly important to reduce the noise of the data before forecasting. Machine learning forecasting methods, such as neural networks [5-7], support vector machines [8], etc., have relatively complex models, slow running speeds, and high resource consumption, and the forecasting results largely depend on the quality of the data [9]. Because traditional forecasting methods and machine learning forecasting methods have certain limitations, combined forecasting methods have been produced, such as the Shapley value method [10], decision theory [11], etc., which use weight distribution to combine multiple forecasting methods to improve accuracy, but this method is only for a certain data, generalization ability is weak; another combined forecasting method is load decomposition, such as wavelet decomposition [7], empirical mode decomposition (EMD) [6,12], etc., the load data is decomposed into components of different time scales, and forecasted separately. This method can often improve the forecasting effect. EMD decomposition solves the problem of artificial selection of wavelet base and decomposition layer in wavelet decomposition, but the result is prone to modal aliasing. Literature [13] uses an ensemble empirical mode decomposition (EEMD) to eliminate modal aliasing by adding white noise. The decomposed high-frequency components

⁺ Corresponding author. Tel.: + 86 13146916588.

E-mail address: mazhongjing@bit.edu.cn.

usually contain a lot of noise. If they are directly discarded, the high-frequency detail information of the load may be lost, and the noise in other components cannot be eliminated.

Given the above situation, literature [14] proposed a method of empirical mode decomposition based on principal component analysis (PCA) to reduce noise, using PCA transform to reconstruct intrinsic mode function (IMF) to achieve the purpose of noise reduction. Literature [15] uses gated recurrent unit neural network (GRU) for high-frequency components containing a large amount of noise, and multiple linear regression (MLR) for low-frequency components containing a small amount of noise, making full use of the accuracy of the machine learning forecasting methods in a large amount of noise and the speed of the traditional forecasting methods in a small amount of noise.

In this regard, this article proposes a combined load forecasting method based on EEMD, which overcomes the shortcomings of traditional methods for data stability and the long training time of machine learning methods. Firstly, the load is decomposed into IMF components and residual function of different time scales through EEMD, and construct a sample matrix based on the decomposed components; then PCA is used to construct a transformation matrix according to the contribution rate to achieve noise reduction of the sample matrix; next the unit root test is used to judge the stationarity of each component, and different forecasting methods are selected according to the stationarity. When it is a stationary component, using traditional forecasting method MLR, when it is a non-stationary component, the machine learning forecasting method long short term memory (LSTM) is used; finally, the forecasting results of each component are superimposed to obtain the total forecasting result. The method in this article takes into account the influence of high-frequency component, the forecasting effect is better and accuracy is higher.

2. Load Data Processing

2.1. Ensemble Empirical Mode Decomposition

The EMD [16] decomposes the nonlinear and non-stationary time series data according to the time scale of the series itself and obtains a limited number of IMF components with different time scales and a residual function. Each IMF component and the residual component must meet:

- The difference between the number of extreme points and the number of zero-crossing points is less than or equal to 1
- The mean value of the sum of the upper and lower envelopes is 0

Due to equipment failures, noise interference and other factors, using the EMD to decompose the load data usually produce modal aliasing, resulting in a decrease in forecasting accuracy. The EEMD can reduce the occurrence of this phenomenon due to the addition of white noise [17].

By adding white noise s_j that follows the Gaussian distribution into load x whose length is m, we can obtain a new time series x_j (line 2 in Algorithm 1). Then perform EEMD on x_j to obtain some IMF components I_{ji} and a residual component r_j (line 3 in Algorithm 1). The final EEMD decomposition result is shown in line 5 in Algorithm 1. According to the literature [18], the standard deviation of the white noise amplitude is selected as 0.2 times the standard deviation of the load data; the number of repetitions k is 100.

Algorithm 1 Ensemble Empirical Mode Decomposition (EEMD)				
Input: x, k				
1: for $j = 1, 2,, k$ do				
2: $x_j = x + s_j$				
3: $\boldsymbol{x}_j = \sum_{i=1}^n \boldsymbol{I}_{ji} + \boldsymbol{r}_j$				
4: end for				
5: $x = \sum_{i=1}^{n} \bar{I}_i + \bar{r}$, where $\bar{I}_i = \frac{1}{k} \sum_{j=1}^{k} I_{ji}$, $\bar{r} = \frac{1}{k} \sum_{j=1}^{k} r_j$				
Output: $\bar{I}_{i=1,\dots,n}$, \bar{r}				

2.2. Principal Component Analysis

PCA [19] is commonly used for data dimensionality reduction and noise reduction, which converts multiple correlated components into unrelated principal components through orthogonal transformation, and calculates the contribution rate of each principal component. The smaller the contribution rate, the less useful information it contains. If the noise does not have a fixed direction, the noise can be reduced by discarding

the components with a relatively small contribution rate. The specific steps of PCA noise reduction are as follows:

Algorithm 2 Principal Component Analysis (PCA) Input: $\overline{I}_{i=1,...,n}$, \overline{r} , ε_0 1: $\varepsilon = 0, j = 0$ 2: $I'_{i} = \overline{I}_{i} - \mathbb{E}(\overline{I}_{i}), \, \overline{r}' = \overline{r} - \mathbb{E}(\overline{r})$ 3: $C = \frac{X'^T X'}{m-1}$, where $X = [\bar{I}_1, \bar{I}_2, \dots, \bar{I}_n, \bar{r}], X' = [\bar{I}_1', \bar{I}_2', \dots, \bar{I}_n', \bar{r}']$ 4: $C = U\Lambda U^{-1}$, where $\Lambda = diag(\lambda_1, \ldots, \lambda_m)$. 5: rearrange eigenvectors U according to eigenvalues Λ 6: while $\varepsilon > \varepsilon_0$ do i = i + 17: $p_i = \frac{1}{\sum_{i=1}^{m}}$ 8: 9٠ $\varepsilon = \varepsilon +$ 10: end while 11: select eigenvectors corresponding to the first j eigenvalues to form transformation matrix P12: $Y = XPP^{T} + \mathbb{E}(X)$, where $Y = \left| \tilde{I}_{1}, \tilde{I}_{2}, \dots, \tilde{I}_{n}, \tilde{r} \right|, \mathbb{E}(X) = \left[\mathbb{E}(\bar{I}_{1}), \mathbb{E}(\bar{I}_{2}), \dots, \mathbb{E}(\bar{I}_{n}), \mathbb{E}(\bar{r}) \right]$ **Output:** $\tilde{I}_{i=1,\ldots,n}$, \tilde{r}

Usually, when cumulative contribution rate ε_0 is equal to 0.85, the corresponding eigenvectors can express almost all original information.

2.3. Unit Root Test

The unit root test also called the extended Dick Fuller test, is used to test the stationarity of the series. If the series is stationary, there is no unit root; otherwise, there will contain a unit root.

3. Load Forecasting Methods

3.1. Multiple Linear Regression

MLR is a traditional statistical forecasting method [3]. For periodic data, compared with other forecasting methods, MLR does not require parameter adjustment and iterative training, and the obtained forecasting values are more accurate. The matrix expression is as follows:

$$Z = X_{\mu} \times \theta + \mu \tag{1}$$

where *Z* is the forecasting value of stationary IMF component; X_b is the influencing factor of the forecasting value, here is the previous stationary IMF component; θ is the regression coefficient that is estimated by the least square method $\theta = (X_b^T X_b)^{-1} X_b^T Z$; μ is the random disturbance.

3.2. Long Short Term Memory

LSTM [20] is a recurrent neural network (RNN) with a special structure. The gate and cell state can be used to avoid the disappearance of the RNN gradient to solve the problem of long-term dependence. LSTM is suitable for timing series problems and its neuron structure is shown in Figure 1 below.



where h_{t-1} is the output of the LSTM neuron at time *t*-1, h_t is the output of the LSTM neuron at time t; C_{t-1} is the cell state at time *t*-1, C_t is the cell state at time t; x_t is the input of the LSTM neuron at time t; σ is the sigmoid function.

LSTM uses three gates to control the cell state:

• Forget gate: determine the information that should be discarded or retained in the cell state

$$f_{t} = \sigma \left(W_{f} \cdot \left[h_{t-1}, x_{t} \right] + b_{f} \right)$$
⁽²⁾

• Input gate: update the cell state

$$i_{t} = \sigma \left(W_{i} \bullet \left[h_{t-1}, x_{t} \right] + b_{i} \right)$$
(3)

$$C'_{t} = \tanh\left(W_{C} \cdot \left[h_{t-1}, x_{t}\right] + b_{C}\right) \tag{4}$$

$$C_{t} = f_{t} * C_{t-1} + i_{t} * C_{t}'$$
(5)

• Output gate: determine the final output result of the LSTM neuron

$$o_t = \sigma \left(W_o \bullet \left[h_{t-1}, x_t \right] + b_o \right) \tag{6}$$

$$h_t = o_t * \tanh\left(C_t\right) \tag{7}$$

where W, b are the weight and bias of each gate.

3.3. Load Combination Forecast Method Based on EEMD

The combined load forecasting method based on EEMD reduces the noise of each IMF component while retaining the load information in the high-frequency IMF component. According to the stationarity of each component, different forecasting methods are adopted to makes the forecasting result more accurate.

Algorithm 3 EEMD-based Power Load Combination Forecasting Method

```
Input: x
 1: apply EEMD to obtain \overline{I}_{i=1,...,n}, \overline{r}
 2: apply PCA to obtain I_{i=1,...,n}, \tilde{r}
 3: for i = 1, 2, ..., n do
          if \tilde{I}_i is stationary then
 4 \cdot
 5:
               apply MLR to forecast to obtian \hat{I}_i
          else
 6:
               apply LSTM to forecast to obtian \hat{I}_i
 7:
          end if
 8:
 9: end for
10: apply MLR to forecast to obtian \hat{r}
11: \hat{x} = \sum_{i=1}^{n} I_i + \hat{r}
Output: \hat{x}
```

4. Case Study

4.1. Data Display and Processing

The load data in this article comes from a certain area in Shanxi from June 1 to June 30, 2018. The sampling interval is 1h and each collection point collects 24 data per day. Since the actual sampling process is affected by other factors such as equipment failure, noise interference, etc., there will be missing and abnormal data, so noise reduction processing is required. The load data is shown in Figure 2. It can be seen that the periodicity of the load is poor, and there are fluctuations and noises. It is difficult to obtain accurate results if the forecasting is made directly.

The load data is decomposed by EEMD, and 9 IMF components and 1 residual component with frequency from high to low are obtained, as shown in Figure 3. It can be seen that through EEMD, there is no modal aliasing phenomenon in each component, and the frequency is relatively stable.



Fig. 2: Load data



Fig. 3: EEMD decomposition results

4.2. Noise Reduction Effect Comparison

The IMF components and the residual component are formed into the original sample matrix, and PCA is performed to reduce noise. When the cumulative contribution rate is more than 85%, it can be considered that these components can express almost all the original information. The contribution rate of each component is shown in Table 1. The first three principal components are selected to form the transformation matrix to reduce the noise of the original sample matrix. The result of noise reduction is shown in Figure 4. The solid blue line is the original sample, and the red dashed line is the sample after noise reduction. It can be seen that by performing PCA noise reduction, the volatility of each component is significantly reduced and the trend is more obvious while retaining the detailed information related to the load in the high-frequency components.

Table 1: Contribution rate of Each Component

Component	Contribution rate / %	Cumulative contribution rate / %
First	52.20	52.20
Second	30.57	82.77
Third	9.97	92.74
Fourth	4.75	97.49
Remaining	2.51	100



Fig. 4: Noise reduction effect

Another method of noise reduction is used to directly remove the high-frequency components of IMF1 after EEMD decomposition. The comparison of the results of the two noise reduction methods is shown in Figure 5. In order to show the noise reduction effect more clearly, the data from June 1st to June 5th is selected to show. It can be seen that directly removing the IMF1 component will lose the load information related to the high-frequency components in the original data, making the subsequent forecasting results inaccurate.



Fig. 5: Comparison of effects of two methods

In order to further measure the noise reduction effect, the signal-to-noise ratio (SNR) and the correlation coefficient R are used to compare the effects of the two noise reduction methods. The larger the SNR and R, the better the noise reduction effect. The metrics of the two noise reduction methods are shown in Table 2. It can be seen that the effect of PCA noise reduction is better than removing the IMF1.

Method	SNR	R
PCA	21.3209	0.9887
Remove IMF1	-22.0787	0.8865

Table 2: Comparison of metrics of the two noise reduction methods

4.3. Forecast Result Analysis

Select the data from June 1st to June 29th, 2018 as the training set, and select the data from June 30th, 2018 as the test set. In order to verify the effectiveness and superiority of the combined forecasting method, the MLR, LSTM and BP are used for comparative analysis, and the control variable method is used to set the same parameters to forecast the load on June 30, 2018. The result is shown in Figure 6. It can be seen that the forecasting results of the method in the article basically restore the trend of the real load, especially between 9:00 and 17:00. The MLR basically forecasts the trend of the real load, but the load forecast value around 5:00 to 7:00 has a large deviation. The deviation of the LSTM load forecast value at around 6:00 is relatively large. The BP is overall worse than other methods, which is caused by the BP neural network easily falling into the local optimum.

In order to measure the forecasting accuracy of the method more accurately, mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) are used for further comparative analysis. The four indicators all reflect the difference between the forecasting data and the real data, the smaller the better.

The forecasting accuracy indicators of each method are shown in Table 3. It can be seen that the method in the article has reduced various indicators compared with other methods, indicating that the forecasting accuracy of the method in the article has improved compared with other methods, and the BP's indicators are obviously larger than other methods, the image performance is that the curve fit is poor.



Fig. 6: Comparison of forecasting results



Fig. 7: The relative error of the method in this article

Method	MSE	RMSE	MAE	MAPE
MLR	0.0243	0.1559	0.1250	0.1604
LSTM	0.0310	0.1761	0.1330	0.1647
BP	0.0682	0.2612	0.2029	0.2699
Method in this article	0.0062	0.0785	0.0636	0.0839

Table 3: Comparison of Different Forecasting Method

Figure 7 shows the relative error value of the method in the article. It can be seen that the maximum relative error value is around 21:00, and most of the forecasting values at other times are less than 10%, indicating that the EEMD decomposition and PCA noise reduction help reflect the true load trend. The MLR forecasting method is used for the stationary component and the LSTM forecasting method is used for the non-stationary component, which can improve the forecasting accuracy and obtain more accurate forecasting results.

5. Conclusion

Aiming at the non-stationary characteristics of load affected by noise, a combined load forecasting method based on EEMD is proposed. Due to the influence of factors such as equipment failure and electromagnetic interference in the actual sampling process of load, the data obtained often contain noise, which has an adverse effect on subsequent research. Therefore, the sampled load data should first be denoised. The comparison with other forecasting methods shows that the combined forecasting method of

load based on EEMD can improve the forecasting accuracy while reducing the noise of load. Accurate load forecasting can provide guidance for grid planning and dispatch, rationally allocate resources, improve resource utilization, and reduce unnecessary equipment investment.

6. References

- [1] Jiang Jiandong, Zhang Haojie, Wang Jing. Research and application of HHT-based power load combination forecasting[J]. Journal of Zhengzhou University (Engineering Science), 2015,36(04):1-5.
- [2] Wan Kun, Liu Ruiyu. Application of interval time-series vector autoregressive model in short-term load forecasting[J]. Power System Technology, 2012,36(11):77-81.
- [3] Amral N, Ozveren C S, King D. Short term load forecasting using multiple linear regression[C]//: International Universities Power Engineering Conference, 2007.
- [4] Wang Yanling, Wu Mengkai. Medium and long term load forecasting model based on partial least-square regression under new normal economy[J]. Electric Power Automation Equipment, 2018,38(03):133-139.
- [5] Quan H, Srinivasan D, Khosravi A. Short-term load and wind power forecasting using neural network-based prediction intervals[J]. IEEE Transactions on Neural Networks and Learning Systems, 2014,25(2):303-315.
- [6] Li Xianglong, Zhang Baoqun, Zhang Yu, et al. Short-term load forecasting of distribution transformer with electric heating based on EEMD-BP neutral network[J]. Electrical Measurement & Instrumenta- tion, 2018,55(10):101-107.
- [7] Chang Yufang, Zhang Li, Xie Hao, et al. Short-term wind speed combined forecasting using wavelet analysis and neural network[J]. Journal of Huaqiao University (Natural Science), 2019,40(04):556-560.
- [8] Fan G, Peng L, Hong W, et al. Electric load forecasting by the SVR model with differential empirical mode decomposition and auto regression[J]. Neurocomputing, 2016,173:958-970.
- [9] Li Long, Wei Jing, Li Canbing, et al. Prediction of load model based on artificial neural network[J]. Transactions of China Electrotechnical Society, 2015,30(08):225-230.
- [10] Chen Qiming, Chen Huayou. Shapley value approach to determining the weights of a kind of combined forecasting model[J]. Journal of Anhui University (Natural Science Edition), 2012,36(02):29-34.
- [11] Yan Feng, Miao Qian, Liu Yanan. A combination short-term load forecasting based on decision-making[J]. Electrical Measurement & Instrumentation, 2013,50(02):64-67+111.
- [12] Qiu X, Ren Y, Suganthan P N, et al. Empirical mode decomposition based ensemble deep learning for load demand time series forecasting[J]. Applied Soft Computing, 2017,54:246-255.
- [13] Guo Jianpeng, She Yingling, Wen Buying. Regional monthly load forecast Based on EEMD-ARIMA Model[J]. Electric Power Engineering Technology, 2018,37(06):28-32.
- [14] Wang Wenbo, Zhang Xiaodong, Wang Xiangli. Empirical Mode Decomposition De-noising Method Based on Principal Component Analysis[J]. Acta Electronica Sinica, 2013,41(07):1425-1430.
- [15] Deng Daiyu, Li Jian, Zhang Zhenyuan, et al. Short-term Electric Load Forecasting Based on EEMD-GRU-MLR[J]. Power System Technology, 2020,44(02):593-602.
- [16] Tang Qingfeng, Liu Nian, Zhang Jianhua, et al. A short-term load forecasting method for micro-grid based on EMD-KELM-EKF and parameter optimization[J]. Power System Technology, 2014,38(10):2691-2699.
- [17] Wu Z, Huang N E. Ensemble empirical mode decomposition: a noise-assisted data analysis method[J]. Advances in Adaptive Data Analysis, 2009,01(01):1-41.
- [18] 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[J]. Proceedings of the Royal Society A-Mathematical Physical and Engineering Sciences, 1998,454(1971):903-995.
- [19] Zhou Xianying, Gao Chengwen, Cao Jianhua. Principal component analysis method and its application in data noise reducing[J]. Ordnance Industry Automation, 2014,33(09):55-58.
- [20] Chen Liang, Wang Zhen, Wang Gang. Application of LSTM networks in short-term power load forecasting under the deep learning framework[J]. Electric Power Information and Communication Technology, 2017,15 (05):8-11.