Short-term wind speed prediction based on EEMD-LSTM

Yuansheng Huang¹, Xiangrong Zhang² and Lei Yang³

¹Department of Economic Management, North China Electric Power University, Baoding 071003, China
²North China Electric Power University, Baoding 071003, China
³Corresponding author:1346565944@qq.com

Abstract. Accurate wind speed prediction is of great significance for the smooth output of wind farms. To this end, this paper proposes a short-term wind speed prediction model based on the combination of ensemble empirical mode decomposition (EEMD) and long-term and short-term memory model (LSTM). Firstly, in order to reduce the nonlinearity and volatility of wind speed, the ensemble empirical mode decomposition technique is used to decompose the original wind speed time series into a plurality of different sub-sequences; then LSTM is used to predict each sub-sequence to obtain multiple prediction results. Finally, the prediction results of each LSTM model are superimposed to obtain the final wind speed prediction result of the combined model. The prediction model is verified by historical wind speed data of a wind farm in Hebei Province, and compared with ARIMA, GRNN and LSTM models. The simulation results show that the combined wind speed prediction model based on ensemble empirical mode decomposition and long-term and short-term memory model proposed in this paper has a higher prediction accuracy.

1. Introduction

Wind energy is an environmentally friendly and widely distributed renewable energy source. The development and utilization of wind energy has been highly valued by governments. Large-scale wind power can reduce carbon dioxide emissions and slow down the greenhouse effect. Due to the volatility and randomness of wind speed, the uncertainty of wind power generation increases. Ensuring the stability of the power system during its normal operation is a key factor in ensuring the overall operation level of the system. When controlling its stability, it is necessary to strengthen the internal anti-jamming system and the overall strength of the system to ensure the overall strength of the system. The abundance of the power system ensures that the power system can guarantee the power transmission to the power users during operation. However, due to the uncertainty of wind power, it will have a certain impact on its abundance and defense system. The most direct performance is the instability of the frequency of electricity[1].

Domestic wind speed prediction research started late, but in recent years, under the correct guidance of the national power sector, more and more wind farms began to use China's self-developed short-term wind speed forecasting system, and its development situation is more optimistic. However, due to the rapid development, the wind power industry has produced weightless results. Although the cumulative installed capacity of wind power is still developing continuously, its growth rate has slowed down remarkably. In the future power generation strategy, China's development from speed-scale-oriented development to quality-efficiency-oriented development has become the trend of the
times. Domestic scholars have also made remarkable achievements in the research work on short-term wind speed prediction, and a large number of research methods have emerged.

2. Summary of short-term wind speed prediction model

In recent years, domestic and foreign scholars have done a lot of research work on wind speed prediction. According to different processing methods of sample data, the prediction methods are mainly divided into time series method, signal decomposition method and multivariate method.

The time series method is to analogize the trend and law reflected by the historical wind speed data to predict the wind speed of the next period. For example, Chang C et al.\[^2\] first reconstruct the original information based on the phase space reconstruction of the chaotic wind speed time series. In the above, the wind speed is reconstructed into a multi-dimensional sequence, and then the Gaussian process model is used for prediction. The signal decomposition method introduces a signal decomposition algorithm, which decomposes the wind speed into multiple components with different frequencies, and then predicts each component separately, and combines the prediction results to obtain the final predicted value. For example, Wang S et al.\[^3\] proposed a composite prediction model based on empirical mode decomposition (EMD) and BP neural network, and the prediction effect is better. The multivariate method predicts the correlation between various meteorological information and wind speed, and emphasizes the physical meaning of predictive behavior. For example, Zhang et al.\[^4\] proposed the use of temperature, humidity, wind speed, wind direction, air pressure, rainfall, wind speed standard deviation and wind speed extreme value to predict the future wind speed; Yan Hongwen et al.\[^5\] proposed using six attributes such as temperature, local pressure, sea level pressure, wind direction and wind shear to predict wind speed.

This paper proposes a short-term wind speed prediction model based on the combination of ensemble empirical mode decomposition (EEMD) and long-term and short-term memory model (LSTM), which provides a new solution for short-term wind speed prediction. Taking a wind farm in Hebei Province as an example, the model is trained with historical wind speed data of wind farms, and short-term wind speed prediction is carried out. Through experiments and comparison with other prediction methods, the simulation results show that the combined wind speed prediction model has higher prediction accuracy.

3. Establishment of combined prediction model based on EEMD-LSTM

3.1. EEMD

The ensemble empirical mode decomposition (EEMD) is an improvement based on the empirical mode decomposition (EMD) method. By adding white noise to the initial sequence and extracting the original signal using EMD, the numerator calculation is finally performed on the sub-sequences after several extractions. EEMD handles the shortcomings of frequency mixing caused by EMD in the noise elimination process\[^6\]. Therefore, this paper uses EEMD to decompose the wind speed time series to achieve the purpose of denoising. The specific steps are as follows:

(1) Add white noise to the wind speed time series to obtain a new sequence;
(2) Decomposing the wind speed time series with white noise by EMD to obtain each IMF component;
(3) Repeat steps (1) and (2), each time adding a new white noise sequence of the same magnitude, taking the average of the IMF obtained each time as the final result.

3.2. LSTM

Long short-term memory (LSTM) is an improved recurrent neural network (RNN) model\[^7\]. RNN consists of three parts: Input Units, Output Units and Hidden Units. The input unit inputs the time series, and the RNN calculates the state sequence of the hidden unit by calculation, and then repeat the iteration to obtain the output sequence. RNN uses the back propagation through time (BPTT) algorithm to perform gradient calculation on the expansion of the RNN model. However, with the
continuous iteration of time steps, the BPTT training method will bring about the problem of gradient disappearance.

In order to improve the defects of RNN, LSTM designs the hidden layer into a more complex structure, mainly through the forgetting gate, input gate and output gate to add or delete information of the memory unit.

3.3. Modeling process
First, the original wind speed time series is decomposed into a number of different subsequences by EEMD technology; then each subsequence is predicted by LSTM prediction method. Finally, the predicted values of all subsequences are superimposed to obtain the actual prediction results, and the error analysis is performed using the evaluation indicators. The modeling flow chart is shown in Figure 1. The specific modeling steps are as follows:
(1) Using EEMD technology to reduce the nonlinearity of the original wind speed time series, and obtain a plurality of different sub-sequences;
(2) Predicting all subsequences separately by LSTM to obtain multiple predicted values;
(3) Superimposing the predicted values of all subsequences to obtain actual prediction results;
(4) Perform error analysis on the prediction results.

![Modeling flow chart](image)

4. Case study
4.1. Raw data
The verification case in this paper is the actual wind speed data collected by a wind farm in Hebei Province in May 2018. The sample set has a total of 900 wind speed data, and the interval between each two sample points is 5 min. Among them, the first to the 800th wind speed data in the sample set are selected to form a training sample set of the prediction model, and the 801th to the 900th data are used as the verification set of the prediction model. The original wind speed data is shown in Figure 2.
4.2. Data preprocessing

In this paper, the input dimension of the prediction model is 10 dimensions, that is, the 10 historical data before the prediction point is used as the input of the prediction model to predict the future wind speed data. For example, the 791th to the 800th data constitute a prediction model input, thereby predicting the predicted value of the 801th data point. By adding new data and erasing the old data, the prediction model input is formed by rolling forward every 10 data groups, and finally the prediction of all the data in the verification set is completed.

In this paper, we use EEMD to decompose the original wind speed data and extract a series of sub-sequences with different bandwidths. The original wind speed data decomposition results are shown in Figure 3 below.

Figure 2 Raw wind speed data

Figure 3 Decomposition wind speed data
4.3. Analysis of prediction results
Using the comparison model mentioned above, all subsequences extracted from the original sequence by EEMD are predicted separately. In this verification, the number of EEMD sets \( N \) is 100, the standard deviation of adding white noise is set to 0.01; the input layer node in LSTM is 10, the number of hidden layer nodes is 200, and the output layer node is 1, the initial learning rate is 0.05, and other parameters are based on Matlab system defaults. The prediction results of each model are shown in Figure 4.

![Figure 4 Individual model prediction results](image)

4.4. Error Analysis
In order to ensure an effective and comprehensive evaluation of the prediction accuracy of each model, three different evaluation indicators are used in this paper: mean absolute error (MAE), mean absolute percentage error (MAPE) and the root of the mean squared error (RMSE) to evaluate the prediction results related to the model. The prediction error analysis of each model is shown in Table 1.

\[
E_{MAE} = \frac{1}{N} \sum_{i=1}^{N} | y_i - \hat{y}_i | \quad (1)
\]

\[
E_{MAPE} = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (2)
\]

\[
E_{RMSE} = \left( \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \right)^{\frac{1}{2}} \quad (3)
\]

Where \( N \) is the number of test samples; \( y_i \) is the actual wind speed value at time \( t \); \( \hat{y}_i \) is the predicted wind speed value at time \( t \).

| Model        | \( E_{MAE} \) | \( E_{MAPE} \) | \( E_{RMSE} \) |
|--------------|---------------|----------------|----------------|
| ARIMA        | 1.0157        | 0.0940         | 1.5088         |
| BPNN         | 0.7633        | 0.0825         | 1.0506         |
| LSTM         | 1.0315        | 0.0937         | 1.3678         |
| EEMD-LSTM    | 0.6668        | 0.0684         | 0.9289         |
It can be seen from Table 1 that the EEMD-LSTM prediction model proposed in this paper has the highest prediction accuracy in all prediction models. Compared with ARIMA, EEMD-LSTM increased $E_{MAE}$, $E_{MAPE}$ and $E_{RMSE}$ by 34.89%, 2.56% and 57.99% respectively. It proves that the proposed method is better than single statistical analysis model. Compared with BPNN, EEMD-LSTM increased $E_{MAE}$, $E_{MAPE}$ and $E_{RMSE}$ by 9.65%, 1.41% and 12.17% respectively, indicating that the prediction accuracy of the proposed method is higher than that of a single intelligent algorithm. Compared with LSTM, EEMD-LSTM increased $E_{MAE}$, $E_{MAPE}$ and $E_{RMSE}$ by 36.47%, 2.53% and 43.89% respectively. It can be seen that the data preprocessing method EEMD has a positive effect in promoting prediction accuracy.

5. Conclusion
Wind speed prediction can not only optimize, secure, and control power grid scheduling, but also has important significance for reducing the operating cost of the power grid and increasing the value of power in the power market. This paper proposes a short-term wind speed prediction model based on EEMD and LSTM. The case study shows that the EEMD is used to decompose the original wind speed data to provide more input information for the prediction model. The preprocessing method based on EEMD can improve the prediction effect of combined forecasting model. The EEMD-LSTM model can sharply identify the trend of wind speed and provide more accurate wind speed prediction data for decision makers. It is a predictive model with better predictive performance.

Acknowledgments
This work was supported by the Foundation of Hebei Province (grant number 2014BD0065).

References
[1] Kong X M. Exploring the influence of wind power uncertainty on power system. 2018 Technology Information. 16 53-55
[2] Chang C and Li D S. Short-term wind speed prediction based on phase space reconstruction and evolutionary Gaussian process. 2016 Computer and Modernization. 43 33-36
[3] Wang S, Zhang N and Wu L. Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method. 2016 Renewable Energy. 94 629-36
[4] Zhang Y and Han P. Short-term wind speed prediction of wind farm based on variational mode decomposition and LSSVM. 2018 Journal of Solar Energy. 39 194-202
[5] Yan H W and Lu G Y. Application of CEEMD-WT and CNN in short-term wind speed prediction. 2018 Computer Engineering and Applications. 54 224-30
[6] Fu T L. Short-term wind speed prediction model based on set empirical mode decomposition and support vector machine. 2017 Journal of Yanbian University (Natural Science Edition). 43 205-09
[7] Li B, Zhang Y and Liu S. Short-term wind speed prediction based on LSTM. 2018 Computer Simulation. 35 456-61