

# A Short-term Wind Power Prediction Framework based on Two-layer Decomposition and the Combination of Ensemble Model and Deep Network

Feichao Di <sup>1, 2, 3</sup>

<sup>1</sup> CCCC Wuhan Harbor Engineering Design and Research Institute Co., Ltd, Hubei Wuhan, 430040, China

<sup>2</sup> Hubei Provincial Key Laboratory of New Materials and Maintenance and Reinforcement Technology for Offshore Structures, China

<sup>3</sup> Transportation Industry Transportation Infrastructure Intelligent Manufacturing Technology R&D Center, China

**Abstract.** The time series of wind power is influenced by many external factors, showing strong volatility and randomness. Aiming at the problem of low prediction accuracy of wind power time series, this paper proposes a short-term wind power prediction framework based on two-layer decomposition and the combination of ensemble model and deep network, which is composed of complete ensemble empirical mode decomposition (CEEMD), sample entropy (SE), stacking ensemble, linear regression (LR), variational mode decomposition (VMD), long short term memory (LSTM) and multi-layer perceptron (MLP). Firstly, CEEMD is used to decompose the time series of wind power into different modes and then SE is used for reconstruction. Secondly, different models are applied to predict the different reconstruction components and select the optimal model. Subsequently, VMD is used to decompose the partially decomposed reconstruction components and a combined prediction model of stacking ensemble and LSTM is established. Finally, in order to further improve the prediction accuracy, MLP is applied to correct the error and the corrected error is superimposed with the prediction results and other reconstruction components to obtain the final predicted value. The simulation results show that the accuracy and effectiveness of this model is superior to the traditional model and the prediction accuracy of short-term wind power time series is improved effectively.

**Key words:** Wind power prediction; complete ensemble empirical mode decomposition; variational mode decomposition; stacking ensemble; long short term memory; multi-layer perceptron.

## 1. Introduction

With the development and continuous progress of society, people's demand and requirements for energy continue to increase. Traditional fossil fuels are far from meeting people's needs. The emergence of new energy can not only solve the problem of energy shortage, but also effectively alleviate the environmental issues such as the greenhouse effect and air pollution. Wind energy is the most widely used new energy with the largest application prospects. It is of great significance to use wind energy reasonably and realize wind power generation. However, due to the uncertainty and volatility of wind, when wind power generation is carried out, it will bring a severe test to the safe and stable operation of the power grid and the safety and reliability of the grid connection of the power system. In order to reduce the impact of wind power generation on the power system, it is of great significance to achieve accurate prediction of wind power in wind farms [1-4].

At present, domestic and foreign researchers have mainly divided the methods of wind power prediction into

physical method [5-8], statistical method [9-13] and combined prediction method [14-16]. The physical method is based on the wind speed, wind direction, longitude, temperature, air pressure and other impact factors to realize wind power prediction. Due to the large number of impact factors, it is usually necessary to perform feature selection; the statistical method is based on the time series of historical wind power, using methods such as data mining and deep feature analysis to realize wind power forecast; the combined forecasting method is currently the most popular method, which often manifested as the ensemble of multiple models such as the combination of shallow model and deep model, to improve the prediction accuracy to a certain extent. Currently, support vector machine [17-18], neural networks [19-21] and deep networks [22-26] are often used in wind power prediction. It proposes an ultra-short-term wind power forecasting method combining multi-cluster algorithm and hierarchical clustering algorithm, which clustering historical power sequences and historical meteorological sequences and establishing particle swarm optimization-back propagation (PSO-BP) neural network

prediction model separately to obtain the final prediction result in reference [27]. It reports a LSTM prediction model based on multiple meteorological features and error correction of the final prediction result to obtain the final prediction value in reference [28]. In reference [29], Ying-Yi Hong proposes a deep feature extraction of wind power time series by convolutional network (CNN) and then establishes an LSTM prediction model to realize the prediction of wind power. Reference [30] proposes a new short-term wind power prediction method consisting of partial direct prediction and iterative prediction through the combination of chaotic time series analysis and singular spectrum analysis (SSA).

## 2. Experimental Principle and Methods

### 2.1 COMPLETE ENSEMBLE EMPIRICAL MODE DECOMPOSITION(CEEMD)

Complete ensemble empirical mode decomposition (CEEMD) is an improvement of empirical mode decomposition (EMD) and integrated empirical mode decomposition (EEMD). EMD can decompose the fluctuation and change trend of the original signal, making the signal tends to be stable. CEEMD adds a set of positive and negative white noise to the original signal, which can not only suppress the EMD mode aliasing problem, but also make the residual noise maintain on a very small scale. The main steps are as follows:

Step 1: Add  $n$  sets of positive and negative white noises  $m^+, m^-$  to the original signal  $x(t)$  to form the composite signals  $P(t)$  and  $N(t)$ .

$$\begin{aligned} P(t) &= x(t) + m^+ \\ N(t) &= x(t) + m^- \end{aligned}$$

Step 2: Make empirical mode decomposition on the composite signals, where the  $j$ th IFM component of the  $i$ th signal is expressed as  $c_{ij}$ .

Step 3: Calculate the average value of all inherent modal function (IMF) components to obtain the decomposition result of CEEMD, as shown in Equation 1.

$$C_j = \frac{1}{2n} \sum_{i=1}^{2n} c_{ij} \quad (1)$$

### 2.2 SAMPLE ENTROPY(SE)

Sample entropy (SE) is a method of measuring the complexity of time series. The larger the entropy value, the higher the complexity of the time series. The SE does not depend on the data length and has better consistency. The calculation of the sample entropy value is shown in Equation 2.

$$\text{SampEn}(k, r, L) = -\ln \frac{B^{k+1}(r)}{B^k(r)} \quad (2)$$

Where  $k$  is the dimension,  $r$  is the threshold,  $L$  is the length, and  $B^k(r)$  is the probability of matching  $k$  points when the threshold is  $r$ .

### 2.3 VARIATIONAL MODE DECOMPOSITION(VMD)

Variational mode decomposition (VMD) is a new signal decomposition method. Its purpose is to decompose a signal into multiple discrete modes  $u_k$  ( $k = 1, 2, 3, \dots, K$ ).

For each mode, the state  $u_k$  surrounds the center frequency  $\omega_k$  and the VMD method mainly includes the solution and construction of the variational problem. To decompose the original time series into a series of modal functions with limited bandwidth, the main process is as follows:

Step 1: Hilbert transform is performed for each mode  $u_k$  to obtain a single-sided spectrum.

Step 2: For each mode  $u_k$ , according to the estimated center frequency, adjust the spectrum to the corresponding base band.

Step 3: Using the Gaussian smoothness of the demodulated signal to estimate the bandwidth of each mode, the constraint variation problem is expressed as Equation 3.

$$\begin{aligned} \min \{ & \sum_{k=1}^K \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \} \quad (3) \\ \text{s. t. } & \sum_{k=1}^K u_k = f(t) \end{aligned}$$

Step 4: Introduce the second penalty factor  $C$  and Lagrange multiplier  $\theta$  to transform the constrained optimization problem into an unconstrained optimization problem, as shown in Equation 4.

$$\begin{aligned} L(u_k, w_k, \theta) = & C \left\{ \sum_{k=1}^K \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} + \|f(t) - \\ & \sum_{k=1}^K u_k\|_2^2 + \langle \theta(t), f(t) - \sum_{k=1}^K u_k \rangle \quad (4) \end{aligned}$$

Step 5: Solve the problem by using the alternating direction multiplier algorithm (ADMM), and get the updated formulas of  $u_k$  and  $w_k$ . The specific expressions are shown in Equation 5 and Equation 6.

$$\hat{u}_k^{n+1} = \frac{\hat{f}(\omega) - \sum_{k=1}^K \hat{u}_k(\omega) + \frac{\hat{\theta}(\omega)}{2}}{1 + 2C(\omega - \omega_k)} \quad (5)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega} \quad (6)$$

### 2.4 STACKING ENSEMBLE

The Stacking ensemble model is currently the most widely used model in data mining competitions. It is a method of classification and regression using the fusion of multiple models. The key lies in the selection of the base model and the number of cross-validation. The selection requires that the effects of the base models are close and the correlation is low. The main flows of the Stacking model are as follows:

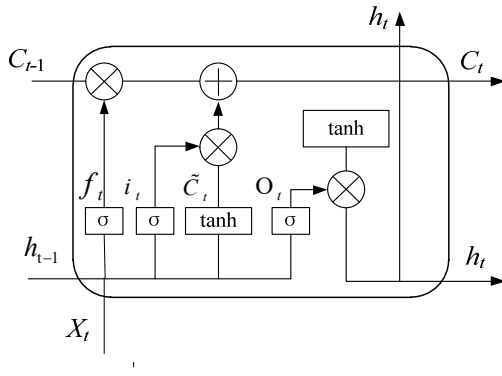
Step 1: Divide the dataset into training set and test set.

Step 2: Cross-validate the training set to obtain the prediction result of the validation set and the test set.

Step 3: Calculate the weighted average of the test set prediction results, the same operation is performed according to the number of base models, until a new training set and test set are constructed.

### 2.5 LONG SHORT TERM MEMORY(LSTM)

Long short term memory (LSTM) is an improvement of the recurrent neural network. It is an improved network proposed to solve the gradient disappearance or gradient explosion of the recurrent neural network (RNN). The LSTM consists of an input gate, a forget gate, and an output gate, as shown in Figure 1.



**Figure 1.** The block diagram of LSTM

Among the three gates, the forget gate determines which information will be discarded, as shown in Equation 7.

$$f_t = \sigma(w_f \cdot [h_{t-1}, X_t] + b_f) \quad (7)$$

The input gate determines which information is retained, and receives new input at the same time, and updates the control parameter  $C_t$ , as shown in Equations 8,9,10.

$$i_t = \sigma(w_i \cdot [h_{t-1}, X_t] + b_i) \quad (8)$$

$$\tilde{C}_t = \tanh(w_c \cdot [h_{t-1}, X_t] + b_c) \quad (9)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (10)$$

The output gate generates the output of the LSTM at the current time according to the control parameter  $C_t$ , as shown in Equations 11 and 12.

$$O_t = \sigma(w_o \cdot [h_{t-1}, X_t] + b_o) \quad (11)$$

$$h_t = O_t * \tanh(C_t) \quad (12)$$

Where  $X_t$  is the input,  $h_t$  is the output,  $h_{t-1}$  is the output at the last time,  $C_{t-1}$  is the control parameter at the last time,  $C_t$  is the latest control parameter,  $i_t$  is the information to be saved.

### 2.6 MULTILAYER PERCEPTRON(MLP)

Multi-layer perceptron (MLP) is a simple multi-layer neural network, the hierarchy of MLP is a directed graph. Each layer is fully connected to the next layer, and the output of the neuron in the previous layer is the input of the neuron in the next layer. The key to MLP lies in the number of hidden layers, the number of neurons in the hidden layer, and the connection weights and offsets, which are often adjusted by grid search and other methods.

## 3. Establishment Of Prediction Model

### 3.1 DATA PREPROCESSING

In order to eliminate the dimensional influence between the data, we can make the data be normalized. After the original data is normalized, the data will be in the same dimension. When the input of the model is the same dimension, the operating efficiency of the model will be greatly improved.

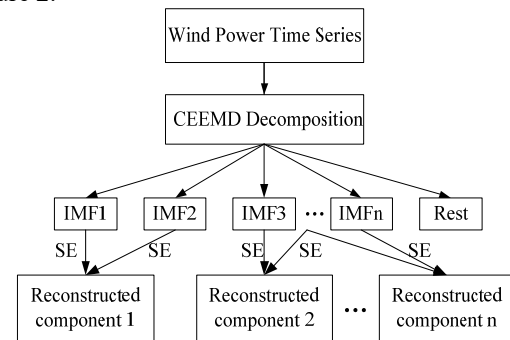
The equation for data normalization is shown in the Equation 13:

$$X^* = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (13)$$

Where  $X_{max}$  is the maximum value of the sample data,  $X_{min}$  is the minimum value of the sample data.

### 3.2 ONE-LAYER DECOMPOSITION

Due to the large volatility of wind power time series, in order to improve the prediction accuracy, many signal decomposition techniques are widely used in wind power prediction. At present, EMD, EEMD, WT, VMD and so on are widely used. In the one-layer decomposition, this paper uses complete ensemble empirical model decomposition (CEEMD), which can effectively suppress the modal aliasing problem of EMD and make the residual component always keep at a very small level. Also, the original wind power data can be completely decomposed without loss. Calculating the sample entropy value (SE) of each IMF and reconstructing the components which the SE are closed to form a new reconstructed component. The process of one-layer decomposition is shown in Figure 2.



**Figure 2.** The structure diagram of CEEMD + SE

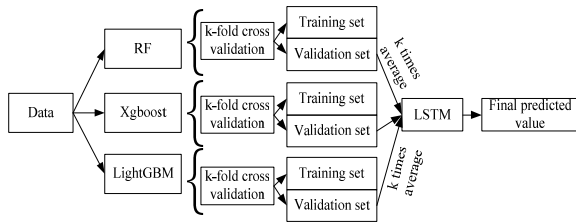
After obtaining different reconstructed components, Analyzing the reconstructed components and selecting different models for prediction according to the characteristics of the different components, Finally, we select the optimal model .

### 3.3 TWO-LAYER DECOMPOSITION

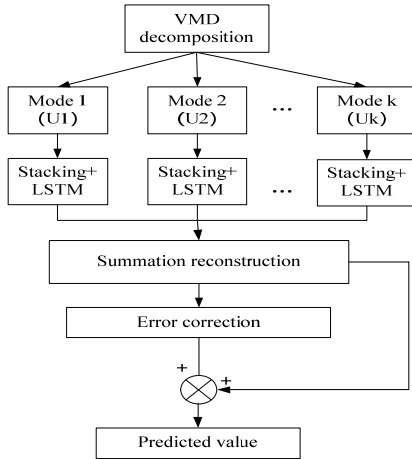
When the volatility of one of the reconstructed components is large, a two-layer decomposition is performed on the basis of the one-layer decomposition. The method of two-layer decomposition in this paper adopts the variational mode decomposition (VMD) technology. After the two-layer decomposition, the reconstructed component obtains k modes and the mode number k is determined by the center frequency  $\omega_k$  of each mode.

For each different mode, the combined model of stacking and LSTM is used to make predictions and the final prediction results are accumulated to obtain the wind power prediction results. Among them, the selection of the basic model of the stacking model are random forest (Random Forest), Xgboost and LightGBM three ensemble models, making full use of the characteristics of the three basic models and constructing new features as the input of the LSTM model to predict wind power, which is shown in Figure 3. All prediction results are superimposed to obtain the final prediction value. Since certain prediction errors will occur, in order to further improve the prediction accuracy, the prediction errors are corrected by the MLP model. The specific process of the

two-layer decomposition of the reconstructed components is shown in Figure 4.

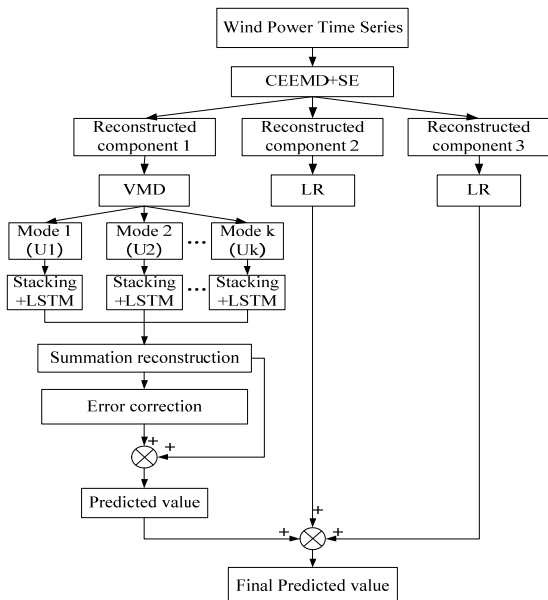


**Figure 3.** The flowchart of Stacking ensemble model



**Figure 4.** The flowchart of two-layer decomposition

A short-term wind power prediction framework based on two-layer decomposition and the combination of ensemble model and deep network is shown in Figure 5.



**Figure 5.** short-term wind power prediction model based on two-layer decomposition and the combination of ensemble model and deep network

The main steps of the short-term wind power prediction model based on two-layer decomposition and the combination of ensemble model and deep network are:

Step 1: Perform CEEMD decomposition and SE reconstruction on the original wind power time series to form reconstructed component 1, reconstructed component 2, reconstructed component 3, and the residual component can be directly ignored because of its small magnitude.

Step 2: For reconstructed component 2 and reconstructed component 3, analyzing their data characteristics and selecting different models for comparison and prediction, we decide to use linear regression (LR) model on the reconstructed component 2 and reconstructed component 3.

Step 3: Because the volatility of reconstructed component 1 is relatively large, VMD is used for two-layer decomposition and the combined prediction models of stacking and LSTM are used for each mode, and error correction is performed using MLP to obtain the final predicted value of reconstructed component 1.

Step 4: Accumulate the predicted results of the reconstructed components 1, 2 and 3 to obtain the final predicted value of wind power.

### 3.4 D. EVALUATION FUNCTION

The evaluations function of the prediction model selected in this paper are:

a. Root mean square error:

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2}$$

b. Mean absolute error:

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i|$$

c. R-Square:

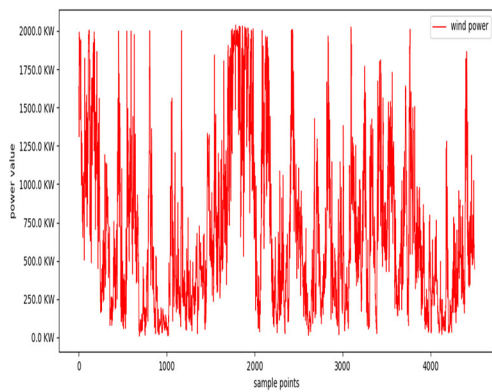
$$R^2 = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y})^2}$$

Where  $y_i$  is the actual value of the wind power at the  $i$ -th moment,  $\hat{y}_i$  is the predicted value of the wind power at the  $i$ -th moment, and  $\bar{y}$  is the average value of the  $m$  wind power values.

## 4. Case studies

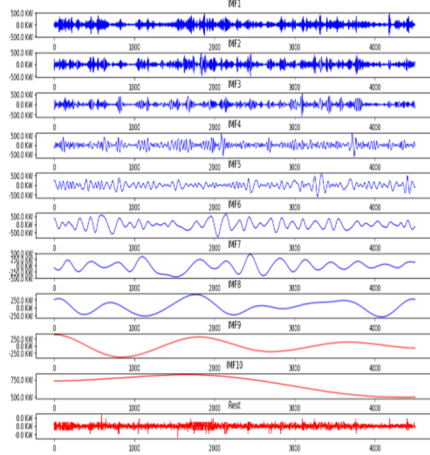
In this paper, we use a wind farm data in the UK for short-term wind power prediction. In order to prove the validity of the model, this paper selects a total of 4,490 wind power data from the wind farm at 1:00 on March 1, 2018 to 8:00 on April 14, 2018.

In the experiment, the sampling interval between the experimental data points is 10 minutes and the time series of the wind power obtained is shown in Figure 6. Among them, the first 3995 data points are used as the training set and the last 495 data points are used as the test set. The previous 5 wind power data predict the wind power at the next moment to achieve rolling prediction.



**Figure 6.** Time series diagram of wind power

It can be seen from Figure 6 that the time series of wind power has obvious fluctuation, so CEEMD decomposition is performed on it, and the specific decomposition results are shown in Figure 7.



**Figure 7.** The graph of CEEMD decomposition results

Because that the residual component Rest is always on a very small scale, only the sample entropy values of the decomposed components (IMF1-IMF10) need to be calculated, the sample entropy threshold  $r$  is set to 0.2, the length  $L$  is set to 2, and each component The sample entropy values are shown in Table 1.

**Table 1.** Sample entropy of each component

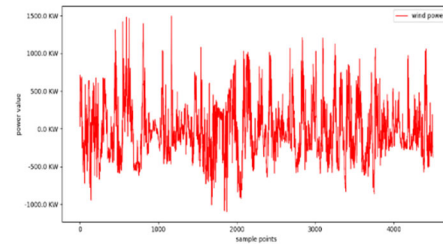
	IM F1	IM F2	IM F3	IM F4	IM F5	IM F6	IM F7	IM F8	IM F9	IM F10
S	5.7	5.9	5.7	5.4	4.7	3.9	2.5	1.3	1.0	0.1
E	57	43	00	02	97	77	90	00	53	00

It can be drawn from Table 1 that the sample entropy values of IMF1-6 are close, so the components are superimposed to obtain the reconstructed component 1; IMF6-9 is used as the reconstructed component 2 and IMF10 is used as the reconstructed component 3.

For the reconstructed component 1, the graph is shown in Figure 8.

It can be seen from Figure 8 that the time series of wind power still exhibits a certain degree of non-stationarity, so

VMD decomposition is performed on the reconstructed component 1 and the center frequency method is used to determine the number of modes  $k$ . The general value range of  $k$  is 3~ 8. The center frequency when  $k$  takes different values is shown in Table 2.



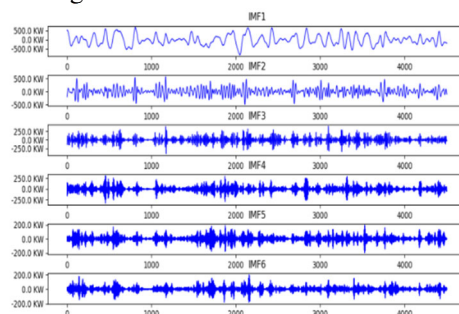
**Figure 8.** The graph of reconstruction component 1

**Table 2.** Center frequencies for different values of  $k$

	u1	u2	u3	u4	u5	u6	u7
K=3	9.147	48.973	139.975	-	-	-	-
K=4	7.787	31.679	78.593	156.950	-	-	-
K=5	7.720	30.964	76.114	150.410	263.141	-	-
K=6	7.673	30.466	74.291	145.673	235.520	354.132	-
K=7	7.202	25.636	54.285	93.459	156.300	251.171	359.550

It can be seen from the center frequency in Table 2 that when  $k=7$ , the center frequency appears 25.636 and 54.285, the two center frequencies are relatively close and the number of modes of VMD decomposition can be determined to be 6. As is shown

in Figure 7, Comparing with the six modes before decomposition, the wind power is decomposed two times, which reduces the fluctuation of wind power and has better stability. The results of VMD decomposition are shown in Figure 9.



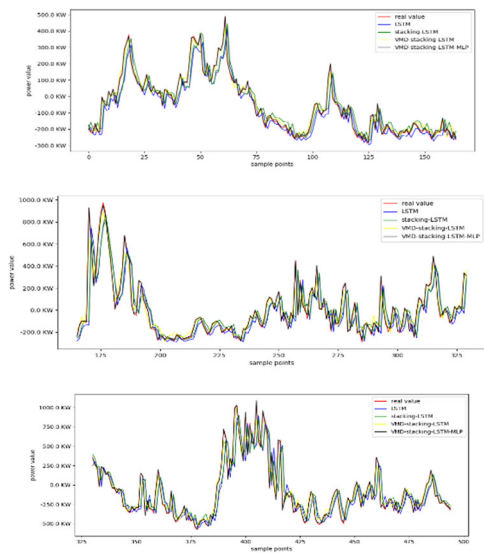
**Figure 9.** The graph of VMD decomposition results

When training and predicting each mode, we build a combined prediction model of stacking and LSTM. The three base models are random forest (RF), xgboost and LightGBM. Among them, the parameters of each model are determined by grid search, where the max depth of three base models is 4, the learning rate is 0.1, LSTM is a 4-layer network structure, the number of neurons in each layer is 50, dropout is 0.1, the optimizer is adam and The number of iterations is 50. when MLP is used for error correction, the number of MLP network layers is 4 and the number of neurons in each layer is 100, 100, 100, and 50 respectively.

Table 3 lists the predicted evaluation indexes of four different models and Figure 10 shows the predicted results of the four models.

**Table 3.** Evaluation indicators of the four models

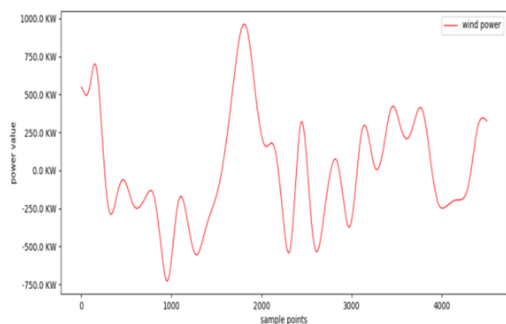
Model	RMSE(KW)	MAE(KW)	R-Square
LSTM	155.71	97.57	0.687
stacking-LSTM	153.42	96.62	0.801
VMD-stacking-LSTM	62.35	43.21	0.949
VMD-stacking-LSTM-MLP	<b>12.21</b>	<b>8.87</b>	<b>0.998</b>



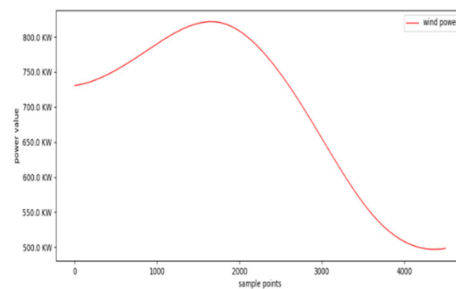
**Figure 10.** The graph of four model prediction curves

According to the predicted results in Figure 10 and Table 3, we can conclude that for the reconstructed component 1, the four models can both predict the wind power at the next moment. Compared with the one-layer decomposition and the two-layer decomposition, the MSE is reduced by 91.07 KW and MAE are reduced by 53.41KW. In terms of predicted accuracy, VMD-stacking-LSTM-MLP has the highest predicted accuracy and R-Square achieves 0.998, indicating that this model has better fitting ability and better predicted effect than other models.

For the reconstructed component 2 and the reconstructed component 3, the graphs of the two reconstructed components are shown in Figure 11 and Figure 12.

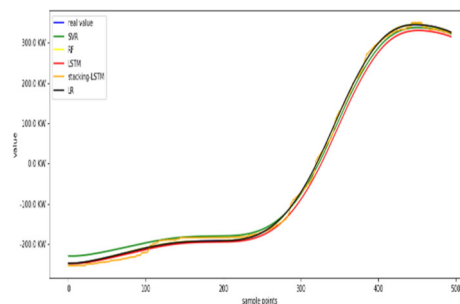


**Figure 11.** The graph of reconstruction component 2 curve

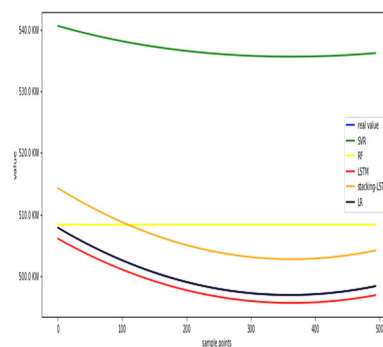


**Figure 12.** The graph of reconstruction component 3 curve

As can be seen from Figure 11 and Figure 12, the curve relationship between the reconstructed component 2 and the reconstructed component 3 has a certain linear relationship, because of its potential data characteristics, a linear regression (LR) model is selected for prediction and the predicted effect is the best. As shown in Figure 13, Figure 14.



**Figure 13.** The graph of reconstruction component 2 prediction curve



**Figure 14.** The graph of reconstruction component 3 prediction curve

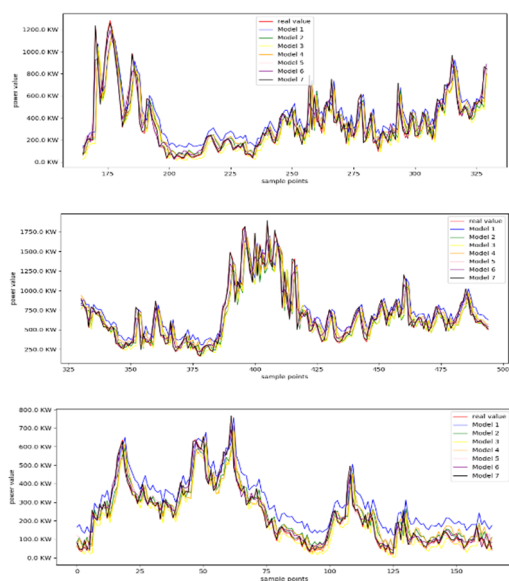
It can be drawn from Figure 13 and Figure 14, compared with other prediction models, the predicted effect of the linear regression (LR) model on the reconstructed component 2 and the reconstructed component 3 is almost completely fitted with the true values, which shows the predicted effect of the model is the best.

Because of the CEEMD decomposition is a complete decomposition and almost no loss. After the predicted values of each reconstructed component are obtained, they are accumulated to obtain the final predicted value of

wind power. Table 4 shows the predicted indexes of the 7 models. The graphs of the final prediction results of the 7 models are shown in Figure 15. Among them, Model 1 uses the SVR model for prediction directly; Model 2 uses the LSTM model for prediction directly; Model 3 uses the LSTM model for prediction after one-layer decomposition; Model 4 uses a combination of stacking and LSTM for prediction after one-layer decomposition of CEEMD; Model 5 uses stacking and LSTM for combined prediction after one-layer decomposition of VMD; Model 6 uses stacking and LSTM combined prediction after two-layer decomposition; Model 7 uses LR, stacking, LSTM and MLP after two-layer decomposition combination forecast.

**Tab 4.** Evaluation indicators of the seven models

Model	RMSE(KW)	MAE(KW)	R-Square
Model 1	168.31	124.79	0.753
Model 2	165.75	115.64	0.768
Model 3	161.24	100.62	0.773
Model 4	153.42	98.68	0.795
Model 5	81.08	54.78	0.942
Model 6	62.34	43.22	0.966
Model 7	<b>13.15</b>	<b>9.44</b>	<b>0.998</b>



**Figure 15.** The graph of seven model prediction curves

Analyzing Table 4 and Figure 15, we can get:

- (1) Compared with the SVR model, The RMSE and MAE of the LSTM model is reduced by 2.56KW and 9.15KW respectively.
- (2) After one-layer decomposition of CEEMD, RMSE and MAE are reduced by 4.51KW and 15.02KW respectively and R-Square is increased to 0.773.
- (3) After one-layer decomposition of CEEMD, using stacking integration and LSTM combined prediction, both RMSE and MAE are reduced, which indicates that the combined prediction model has a strong fitting ability than the single model.

(4) The second decomposition of VMD based on the first decomposition of CEEMD greatly improves the prediction accuracy of the model. RMSE and MAE are reduced by 91.08KW and 55.46KW respectively and R-Square is also greatly improved to 0.966.

(5) After MLP correction of the error, RMSE and MAE were reduced to 13.15KW and 9.44KW respectively and R-Square is increased to 0.998, which indicates that the correction of the error can improve the prediction accuracy of the model.

Therefore, compared with other models, the prediction model proposed in this paper can better predict the short-term wind power changes and the predicted accuracy of the model has also been greatly improved, which verifies the feasibility and effectiveness of the proposed method.

## 5. Conclusion

This paper proposes a short-term wind power prediction based on the two-layer decomposition and the combination of ensemble model and deep network and draws the following conclusions:

- a. Signal decomposition techniques such as CEEMD, VMD can effectively improve the predicted accuracy of the model.
- b. The predicted results obtained by the combined prediction model are significantly better than the single prediction model.
- c. Perform feature analysis on the data to find the most suitable model for prediction, which can not only get better predicted results, but also speed up the operation efficiency of the model.
- d. When it is found that the prediction error shows a certain regularity, an appropriate model can be used to train and correct the error, which can further improve the predicted accuracy of the model.

At present, the extensive use of ensemble models and combined forecasting models, the combined use of shallow models and deep models and the correction of errors are the main development directions of forecasting models. In the future, we can do major research in these three areas.

## References

1. Liu Shuai. Research on Short-Term Power Forecasting Methods for Wind Farms [D]. Beijing: North China Electric Power University, 2019.
2. Liang Zhifeng, Wang Yan, Feng Shuanglei, et al. Ultra-short term forecasting method of wind power based on fluctuation law mining [J]. Power System Technology, 2020, 25 (10): 1-10.
3. Peng Xiaosheng, Xiong Lei, Wen Jinyu, et al. Summary of improved short-term and ultra-short-term power prediction accuracy methods for wind power clusters [J]. Chinese Journal of Electrical Engineering, 2016, 36 (23): 6315-6326 + 6596.
4. Lu Peng, Ye Lin, Tang Yong, et al. Multi-time scale active power optimization scheduling strategy for wind power cluster based on model predictive control

- [J]. Proceedings of the CSEE, 2019, 39 (22): 6572-6583.
5. Wang Yining, Xie Da, Wang Xitian, et al. Prediction of wind turbine network interaction based on PCA-LSTM model [J]. Chinese Journal of Electrical Engineering, 2019, 39(14): 4070-4081.
  6. Song Jiakang, Peng Yonggang, Cai Hongda, et al. Study on short-term wind power prediction considering multi-site NWP and atypical characteristics[J]. Power System Technology, 2018, 42 (10): 3234-3242.
  7. J. Yan, H. Zhang, Y. Liu, et al. Forecasting the High Penetration of Wind Power on Multiple Scales Using Multi-to-Multi Mapping [J]. *Transactions on Power Systems*, 2018, 3276-3284.
  8. Wang Youjia, Lu Zongxiang, Qiao Ying, et al. Statistical upscaling prediction of regional wind power based on feature clustering [J]. Power System Technology, 2017, 41 (05): 1383-1389.
  9. Wu Xiaomei, Lin Xiang, Xie Xuquan, et al. Short-term wind power prediction based on VMD-PE and optimized correlation vector machine [J]. Journal of Solar Energy, 2018, 39 (11): 3277-3285.
  10. Zhu Qiaomu, Li Hongyi, Wang Ziqi, et al. Ultra-short-term prediction of wind farm power generation based on long-short-term memory network [J]. Power System Technology, 2017, 41 (12): 3797-3802.
  11. Mu Gang, Yang Xiuyu, Yan Qian, et al. Prediction method of field group continuous power characteristics based on the evolution trend of wind farm group convergence [J]. Chinese Journal of Electrical Engineering, 2018, 38 (S1): 32-38.
  12. S.Tian, Y. Fu, P. Ling, et al. Wind Power Forecasting Based on ARIMA-LGARCH Model[J]. 2018 International Conference on Power System Technology, Guangzhou, 2018, pp. 1285-1289.
  13. E. Yatiyana, S. Rajakaruna, A. Ghosh. Wind speed and direction forecasting for wind power generation using ARIMA model[J]. 2017 Australasian Universities Power Engineering Conference, 2017, pp. 1-6.
  14. Yin Hao, Ou Zuhong, Chen De, et al. Ultra-short-term wind power prediction based on quadratic mode decomposition and cascading deep learning[J]. Power System Technology, 2020 (02): 445-453.
  15. Hao CHEN, Fangxing LI, Yurong WANG. Wind power forecasting based on outlier smooth transition autoregressive GARCH model[J]. Journal of Modern Power Systems and Clean Energy, 2018, 6(3).
  16. Yu Hua, Lu Jiping, Zeng Yanting, et al. Nonlinear combined prediction of wind power based on different optimization criteria and generalized regression neural network [J]. High Voltage Technology, 2019, 45 (03): 1002-1008.
  17. Y. Zhang, H. Sun, Y. Guo, Wind Power Prediction Based on PSO-SVR and Grey Combination Model[J], IEEE Access, 2019, pp. 136254-136267.
  18. Li Yanqing, Yuan Yanwu, Guo Tong. Ultra-short-term wind power combination prediction based on AMD-ICSA-SVM [J]. Power System Protection and Control, 2017, 45 (14): 113-120.
  19. H. J. Lu and G. W. Chang. Wind Power Forecast by Using Improved Radial Basis Function Neural Network, 2018 IEEE Power & Energy Society General Meeting, Portland, OR, 2018, pp. 1-5.
  20. Ding Ming, Zhang Chao, Wang Bo. Short-term prediction and error correction of wind power based on the power rate fluctuation process [J]. Automation of Electric Power Systems, 2019, 43 (03): 2-12.
  21. J. Varanasi, M. M. Tripathi. Artificial neural network based wind power forecasting in belgium, 2016 IEEE 7th Power India International Conference, 2016, pp. 1-6.
  22. Jinhua Zhang, Jie Yan, David Infield, et al. Short-term forecasting and uncertainty analysis of wind turbine power based on long short-term memory network and Gaussian mixture model[J]. Applied Energy, 2019, pp. 229-244.
  23. X. Peng, D. Deng, J. Wen, et al. A very short term wind power forecasting approach based on numerical weather prediction and error correction method[C]. 2016 China International Conference on Electricity Distribution , 2016, pp. 1-4.
  24. Xiong Yindi, Liu Kaipei, Qin Liang, et al. Short-term wind power prediction method based on dynamic weather classification of time series data [J]. Power System Technology, 2019, 43 (09): 3353-3359.
  25. B. Khorramdel, C. Y. Chung, N. Safari, G. C. D. Price, A Fuzzy Adaptive Probabilistic Wind Power Prediction Framework Using Diffusion Kernel Density Estimators, IEEE Transactions on Power Systems, 2018, pp. 7109-7121.
  26. T. Ouyang, H. Huang, Y. He, Ramp events forecasting based on long-term wind power prediction and correction, IET Renewable Power Generation, 2019, pp. 2793-2801.
  27. Peng Chenyu, Chen Ning, Gao. Ultra-short-term wind power prediction method combining multiple clustering and hierarchical clustering [J]. Automation of Electric Power Systems: 1-8 [2019-11-10].
  28. Shuang Han, Yan-hui Qiao, Jie Yan, et al. Mid-to-long term wind and photovoltaic power generation prediction based on copula function and long short-term memory network[J]. Applied Energy, 2019, pp. 181-191.
  29. Ying-Yi Hong, Christian Lian Paulo P. Rioflorido. A hybrid deep learning-based neural network for 24-h ahead wind power forecasting[J]. Applied Energy, 2019, pp. 530-539.
  30. N. Safari, C. Y. Chung, G.C.D. Price, Novel Multi-Step Short-Term Wind Power Prediction Framework Based on Chaotic Time Series Analysis and Singular Spectrum Analysis, IEEE Transactions on Power Systems, 2018, pp. 590-601.