Article

Wind Speed Prediction Based VMD-BLS and Error Compensation

Xuguo Jiao , Daoyuan Zhang [*] , Dongran Song , Dongdong Mu , Yanbing Tian , Haotian Wu

Article

Wind Speed Prediction Based VMD-BLS and Error Compensation

Xuguo Jiao [1,2], **Daoyuan Zhang** [1,*], **Dongran Song** [3], **Dongdong Mu** [4], **Yanbing Tian** [1] and **Haotian Wu** [1]

1. School of Information and Control Engineering, Qingdao University of Technology , Qingdao 266520, China
2. State Key Laboratory of Industrial Control Technology , College of Control Science and Engineering, Zhejiang University , Hangzhou 310027, China
3. School of Automation, Central South University, Changsha, China
4. College of Marine Electrical Engineering, Dalian Maritime University, Dalian, China
* Correspondence: zdy431877849@gmail.com

Abstract: As one of the fastest-growing new energy sources, wind power technology has attracted widespread attention from all over the world. In order to improve the quality of wind power generation, wind speed prediction is an indispensable task. In this paper, an error correction based Variational Mode Decomposition and Broad Learning System (VMD-BLS) hybrid model is proposed for wind speed prediction. First, the wind speed is decomposed into multiple components by the VMD algorithm, and then an ARMA model is established for each component to find the optimal number of sequence divisions. Second, the BLS model is used to predict each component, and the prediction results are summed to obtain the wind speed forecast value. However, in some traditional methods, there is always time lag, which will reduce the forecast accuracy. To deal with this, a novel error correction technique is developed by utilizing BLS. Through verification experiment with actual data, it proves that the proposed method can reduce the phenomenon of prediction lag, and can achieve higher prediction accuracy than traditional approaches, which shows our method's effectiveness in practice.

Keywords: wind speed prediction; Variational Mode Decomposition (VMD); Autoregressive Moving Average (ARMA); Broad Learning System (BLS); error compensation

1. Introduction

With the development of the times, traditional energy is no longer able to meet the rapidly growing energy consumption[1], and the fossil fuels are non renewable resources, and their combustion can also bring serious environmental pollution. Therefore, developing renewable and pollution-free energy is a hot issue in today's society. In recent years, various new energy sources have been constantly emerging, among which wind power is one of the fastest growing renewable energy sources due to its advantages such as friendly to environment, zero carbon dioxide emissions, abundant resources, and cheap prices[2,3]. In 2021, renewable energy accounts for 38.3% of global electricity generation with wind power accounting for 6.7%[4]. In 2022, 78GW wind capacity is installed globally, and this index can be expected to reach 115GW in 2023[5].

However, although wind power has many advantages, it is a natural resource with intermittent, volatility, and unstable characteristics[6]. Therefore, it cannot generate sustained and stable electricity like traditional energy sources such as thermal power. If the wind power is connected to the power grid directly without any optimization behaviors, it will introduce huge interference to the power system and leads to instable issues[7]. In particular, it may cause equipment damage or even large-scale power outages[8].

Wind speed prediction is one of the most effective methods to address the instability issues that arise with wind power generation[9]. By accurately forecasting wind speeds, wind farms can lower their operating costs, increase the competitiveness of wind power[10], and generate positive impacts. Specifically, wind speed prediction can allow for reasonable adjustments to the power generation

plan[11], optimize torque control to maximum utilization of wind power[12], and optimize layout of wind turbines, which will improve economic benefits for wind farms[13].

Early wind speed prediction methods generally use linear models, such as Autoregressive Moving Average (ARMA)[14], which assumes that future values are linearly correlated with historical data. However, linear models may not be sufficient to characterize the real world, because the real world is usually composed of linear and nonlinear components[15]. After the 1990s, Artificial Intelligence (AI) gradually emerges, and various machine learning models are also emerging, such as Artificial Neural Networks (ANN)[16], Support Vector Machine (SVM)[17], Extreme Learning Machine (ELM)[18], etc. Many researchers have integrated these methods and proposed some hybrid wind prediction models[19,20]. However, wind speed is affected by many factors, such as temperature, air pressure, humidity, terrain, etc. If these data can not be collected beacuse of some reasons such as costs, we need to use historical data as the basis to predict wind speed. But how much historical data to choose as input is also worth considering. Too much data selection may cause waste of machine performance and overfitting problem, while too little selection may lead to poor fitting results due to insufficient utilization of data information. This will lead to poor wind speed prediction results.

To improve the wind speed prediction accuracy, deep learning algorithms are introduced by researchers. In fact, it is widely used to deal with prediction problems. Convolutional Neural Network (CNN) is employed to forecast ships' electric consumption[21]. Long Short-term Memory (LSTM) model is used to predict short-term wind power[22]. Recurrent Neural Network (RNN)-LSTM model is used to enhance the intra-day stock market prediction[23], and so on[24,25]. However, Deep Learning requires setting many parameters and long training time, which narrows its applicaitons in practice. Moreover, most of the existing methods ignore the pediction lad which is brought by sharp changes in natural wind speed[24]. Obviously, this will increase the error between the actual wind speed and predicted values[2].

In order to select the appropriate data as input, reduce the occurrence of forecast lag, and improve the forecast accuracy of the univariate wind speed time series, this paper proposes a hybrid forecasting system of ARMA-VMD-BLS. First, the Variational Mode Decomposition (VMD) algorithm is used to decompose the original wind speed sequence into multiple components. Second, the ARMA model is used to select the appropriate amount of input data, and then the predictiong results of each componentusing can be obtained by using the Broad Learning System (BLS) model. After getting the the predicted results of each component, they are added up to obtain the wind speed prediction result. Moreover, in order to further improve the prediction accuracy, an error compensation method is proposed in this paper, which uses the BLS model to give the predcition error, and the final wind speed prediction result can be obtained by adding the error prediction result to the wind speed prediction result. Through the verification of real wind speed data, the proposed method can effectively alleviate the prediction lag phenomenon and improve the prediction accuracy. Compared with various machine learning models, the superiority of the BLS is proved.

The major contributions of this paper can be summarized as follows:

- The VMD method is used to decompose data into multiple components and the ARMA model is used for each components to choose the optimal number of inputs. By this way, hidden features in the wind speed series can be extracted, and the forecast lags can be reduced.
- The BLS model is utilized to predict each component, and the wind speed prediction result can be obtained by adding them. The BLS is a fast and efficient method, which is very suitable for wind speed forecasting.
- An error compensation method is proposed to further improve the prediction accuracy. The BLS model is used to predict error, and the final wind speed prediction result is obtained by adding the error prediction result to the wind speed prediction result.

The rest of this article is organized as follows: Section 2 introduces the proposed method and mathematical models used; The Section 3 is the experimental, which uses the real data validation of a

wind farm in China to verify the feasibility of proposed method. The Section 4 summarizes and draws conclusions.

2. The Proposed Method and Mathematical Models

2.1. The Framework of the Proposed Method

The flowchart is shown in Figure 1. In order to predict wind speed more accurately, this paper proposes a hybrid prediction model based on ARMA-VMD-BLS. As shown in Figure 1, the framework of proposed method includes the following content.
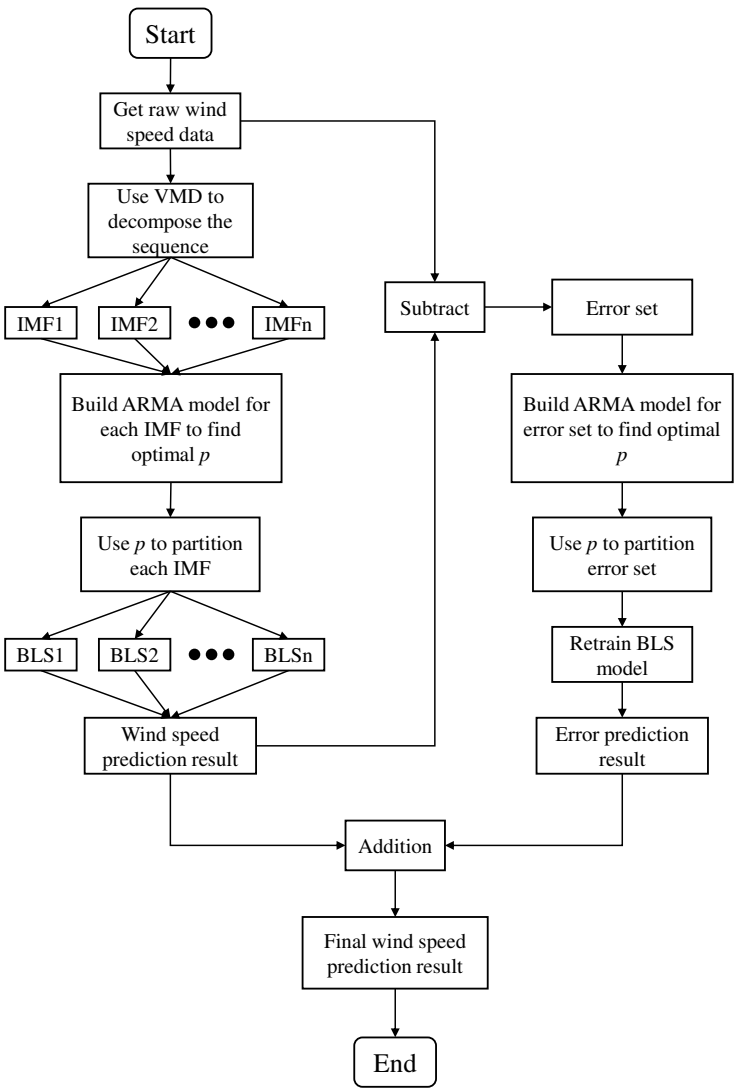

Figure 1. The flowchart of proposed method.

First, the VMD algorithm is used to decompose the original wind speed sequence into multiple Intrinsic Mode Functions (IMF). These IMF are extracted from the original data and have a fixed central frequency, which make them easy to be predicted. Second, the ARMA (p, q) model is built for each IMF to find the optimal number of inputs, and the p-values are used to partition each IMF. Third, the BLS is used to predict each IMF. Through the above steps, the wind speed prediction result can be obtained by adding up the prediction results of each IMF. Moreover, to deal with prediction lag phenomenon and further improve the forecast accuracy, a novel error compensation method is

proposed in the paper. The error set is obtained by subtracting the predicted value from the real value. Then, the ARMA (p, q) model is built for the error set and the p-value is used to partition the error set. Finally, the BLS is used to predict the error, and the final wind speed prediction result can be obtained by adding the wind speed prediction result to the error prediction result.

2.2. The Detail Steps of the Proposed Method

Step 1: Use VMD to decompose the original wind speed sequence.

In this step, the VMD algorithm is used to decompose the original wind speed time series, which can extract potential patterns and remove noise from the original data, and obtain multiple components with fixed center frequencies. In this way, the more wind speed features will be obtained, which is beneficial for improving prediction accuracy[6].

The VMD algorithm was proposed by Konstantin Dragomiretskiy and Dominique Zosso in 2014[26]. Because it can effectively overcome problems such as the dependence of Empirical Mode Decomposition (EMD) on extreme point selection and frequency overlap, the VMD has attracted increasing attention from researchers[27].

The VMD can decompose the wind speed time series into several IMF. Its expression is:

$$u_k(t) = A_k(t) \cos\left(\Phi_k(t)\right) \tag{1}$$

where $A_k(t)$ is the non-negative envelope, $\Phi_k(t)$ is the phase and $\Phi_k'(t) \geq 0$, $u_k(t)$ is the raw wind speed data decomposed to the k-th IMF.

The corresponding constrained variational model is as:

$$\begin{cases} \min_{\{u_k\}\{\omega_k\}} \left\{ \sum_K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ \text{s.t. } \sum_k u_k(t) = f \end{cases} \tag{2}$$

where $\{u_k\} = \{u_1, \ldots, u_k\}$ is the set of all modes, $\{\omega_k\} = \{\omega_1, \ldots, \omega_k\}$ is the set of center frequencies of all modes and f is the raw wind speed data.

To solve the constrained optimization problem (2), the augmented lagrangian function is used to convert the above equation constraint problem to an unconstrained problem:

$$L\left(\{u_k\}, \{\omega_k\}, \lambda\right) = \alpha \sum_K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \|f(t) - \sum_k u_k(t)\|_2^2 + [\lambda(t), f(t) - \sum_k u_k(t)] \tag{3}$$

where α is the penalty parameter, $*$ represents convolution operation, and λ is the Lagrangian multiplier.

Then, the Alternate Direction Method of Multipliers (ADMM) is used to update u_k, ω_k and λ:

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \tag{4}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega} \tag{5}$$

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau \left[\hat{f}(\omega) - \sum_{k=1}^K \hat{u}_k^{n+1}(\omega) \right] \tag{6}$$

Step 2: Use the ARMA model to select the optimal the optimal number of inputs.

In the Step 2, we use training sets to establish ARMA model to find the optimal number of inputs. The training sets are the multiple IMF decomposed through VMD in Step 1, and then an ARMA(p,

q) model is established for each IMF. Since the fact that the decomposed data belongs to a relatively stable sequence, there is no need for stationarity test and an ARMA model can be directly established. For each p and q in ARMA (p, q), we set a loop and use the Akaike Information Criterion (AIC) to find the optimal combination of (p, q). The resulting p-value is the optimal number of inputs we are looking for. It is worth noting that the optimal p value corresponding to each IMF may be different, so in the final summation, the same length needs to be selected for addition, which will be explained in detail later.

The ARMA(p, q) model is one of the earliest models used for time series prediction[28].It consists of two models: Autoregressive(AR) model and Moving Average(MA) model. The mathematical model is:

$$X_t = \sum_{i=1}^{p} \alpha_i X_{t-i} + \sum_{j=1}^{q} \beta_j \epsilon_{t-j} + \epsilon_t \tag{7}$$

where $X = \{X_1, X_2, \ldots, X_t\}$ is wind speed time series, X_t is the current wind speed, α_i and β_j are the autoregressive and moving average parameter, p and q are important parameter called Partial Autocorrelation Coefficient (PAC) and Autocorrelation Coefficient(AC), ϵ_t is white nose of time t.

The ARMA model describes the relationship between current value, historical error and historical value. From (7), one can see that the current value X_t is composed of a linear combination of p historical values and q historical errors. It also means that if every p data is taken as a group, the internal correlation of this group is very strong. Thus, it is very reasonable to select p as the basis for dividing the univariate series.

Step 3: Use the p-value to partition data sequence.

In Step 2, the optimal p-values of each IMF was obtained, and in this step, we will use them to partition data for each IMF. The specific method is that, for each IMF, starting from the first data, every p data is a group, the first p-1 data is used as input, and the pth data is output. That is, the first p-1 data is used to predict the pth value. Then, starting from the second data, the previous steps are repeated until all the data is divided, which is shown as (8).

$$
\begin{aligned}
x_p &= f\left(x_1, x_2, \ldots, x_{p-1}\right) \\
x_{p+1} &= f\left(x_2, x_3, \ldots, x_p\right) \\
&\vdots \\
x_{\text{end}} &= f\left(x_{\text{end}-p+1}, x_{\text{end}-p}, \ldots, x_{\text{end}-1}\right)
\end{aligned}
\tag{8}
$$

where x_t represent the wind speed value at time t and $f(\cdot)$ represent the relationship between historical and current value.

Step 4: Use the BLS model to predict each IMF.

In this step, the divided data in Step 3 is used to train the BLS model. For each IMF, triple nested loops are used to find the optimal number of mapping nodes, windows in mapping feature layers, and enhancement nodes. The first 75% of data set is used as the training set, and the rest is used as the testing set. The result will be shown in the following section.

Broad Learning System (BLS) was proposed by C.L. Philip Chen, Dean of the School of Science and Technology at the University of Macau, and his student Zhulin Liu in January 2018. Compared with the deep learning system, BLS has a faster training process, and can increase the number of feature nodes and enhancement nodes through incremental learning, so that the system can continuously update parameters without retraining, which is an alternative to deep learning[29].

The basic structure of BLS model is shown in Figure 2. It is based on the Random Vector Functional Link Neural Network (RVFLNN), which maps the input to the mapping layer through multiple sets of mappings. The mapping is $\phi\left(XW_{e_i} + \beta_{e_i}\right)$ in Figure 2, where X is the wind speed IMF decomposed by the VMD algorithm, W_{e_i} and β_{e_i} is respectively the random weights and bias with the proper

6 of 21

dimensions. Each mapping corresponds to a mapping feature: $Z_1, Z_2, \ldots, Z_n$. The nodes in the mapping feature are called mapping nodes. Each mapping feature is passed to the output layer and the enhancement layer respectively. In the process of passing from the mapping layer to the enhancement layer, there will also be a layer of mapping ,which is $\xi\left(XW_{e_j} + \beta_{e_j}\right)$ in Figure 2, where W_{e_j} and β_{e_j} is the random weights and bias with the proper dimensions respectively. The nodes in the enhancement layer are called enhancement nodes: $H_1, H_2, \ldots, H_n$, and each enhancement node is finally passed to the output layer
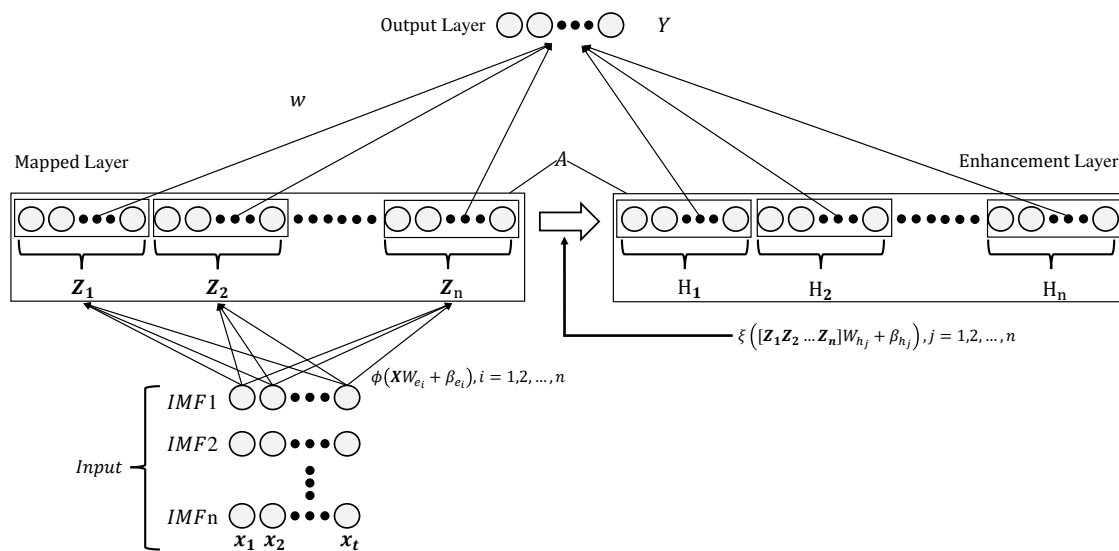

Figure 2. The BLS model.

If one wants to train a model and the input X and labels Y are known, one can figure out w in Figure 2 by equation:

$$\text{Y} = Aw \tag{9}$$

where A is the augmented matrix of the mapping layer matrix and enhancement layer matrix, that is, $A = [Z|H]$, Z and H can be figured by X. Then the w can be calculated by:

$$w = A^+\text{Y} \tag{10}$$

where A^+ is the pseudo-inverse matrix of A.

From equation (9) and (10) we know that, the BLS inputs all the data into the system for training. There is no weight update process, and the weights are directly calculated, while deep learning needs to pass the data into the neural network one by one to update the weights continuously. Therefore, the training speed of BLS is significantly faster than that of deep learning.

In order to solve the problem that the model training cannot achieve the required accuracy, the author also proposed the BLS with incremental learning, including increasing the number of mapping nodes and enhancing nodes, and increasing the input data. In this paper, the incremental learning is used to find the best number of mapping nodes and enhancing nodes, which can improve the prediction accuracy.

Step 5: Add up prediction results of each IMF.

The predictions for each IMF using BLS model in Step 4 are not the true wind speed, we need to add all the predictions to get the wind speed prediction.

$$R_{ws} = R_{IMF1} + R_{IMF2} + \ldots + R_{IMFn} \tag{11}$$

where R_{ws} is the wind speed prediction result, and R_{IMF} is the prediction result of each IMF.

Step 6: Get the prediction error of wind speed.

The error time series can also be predicted due to its non-linear characteristics[30]. If the predicted error can be added to the predicted wind speed, a more accurate prediction result can be obtained. In Step 4, the training model is gotten, in Step 5, we get the prediction result of the testing set. And in this step, we subtract the predicted value from the real value of the testing set to get the error testing set. For the error training set, the training set is used as the input of the BLS model trained in Step 4 to get the prediction result of the training set, and then the error training set can be gotten by subtracting the predicted value from the real value of the training set.

$$\begin{aligned} E_{tr} &= R_{trrv} - R_{trpv} \\ E_{te} &= R_{terv} - R_{tepv} \end{aligned} \tag{12}$$

where E_{tr} and E_{te} represent error training set and error testing set, and R_{trrv} and R_{terv} represent the real value of training and testing set, R_{trpv} and R_{tepv} are the prediction value of training and testing set.

Step 7: Build the ARMA model for error set and partition error set.

This step is very similar to Step 2 and Step 3. The error set obtained in Step 6 is also a univariate time series. Since its fluctuation is very small, there is no need to perform VMD like the original data. The ARMA model is used to select the optimal number of inputs p for the error set, and its p-value is used to divide the error training set and testing set. The way of division is the same as in Step 3, and only the value of p may be different.

$$\begin{aligned} e_p &= g\left(e_1, e_2, \ldots, e_{p-1}\right) \\ e_{p+1} &= g\left(e_2, e_3, \ldots, e_p\right) \\ &\vdots \\ e_{\text{end}} &= g\left(e_{\text{end}-p+1}, e_{\text{end}-p}, \ldots, e_{\text{end}-1}\right) \end{aligned} \tag{13}$$

where x_t represent the error value at time t and $g(\cdot)$ represent the relationship between historical and current error.

Step 8: Use the BLS model to predict error.

In this step, divided data in Step 7 are used to retrain the BLS model. For error training set, triple nested loops are used to find the optimal number of mapping nodes, windows in mapping feature layers and enhancement nodes. After that, the error testing set is used to check the prediction effect of the model. The result will be shown in the following section.

Step 9: Add the wind speed prediction result to the error prediction result.

This is the final step. The final wind speed prediction result will be obtained by adding the wind speed prediction result to the error prediction result.

$$R_{fws} = R_{ws} + R_{err} \tag{14}$$

where R_{dws} is the final wind speed prediction result, and R_{err} is the prediction result of error.

3. Experiment

This section mainly introduces the process and results of the experiment. Subsection 3.1 shows the wind speed prediction, subsection 3.2 shows the error prediction. And in order to prove the superiority of this method, subsection 3.3 uses direct prediction and selects three models to replace the BLS model for comparison.

The raw wind speed data from a wind farm in China is shown in the Figure 3. Among them, the data is recorded every 15 minutes, a total of 671 data are selected, the first 500 data are used as the training set, and the last 171 data are used as the testing set. The experiment is tested on a PC with AMD Ryzen 5800H, NVIDIA RTX 3070 Laptop and 16GB memory. The following is the experiment result.

3.1. Wind Speed Prediction

3.1.1. Use VMD to Decompose the Raw Data

It can be seen from the Figure 3 that the fluctuation of this sequence is very large. If direct prediction is carried out, the problem of prediction lag may occur, which seriously affects the prediction accuracy. Therefore, 5 IMF are obtained by using the VMD algorithm to decompose the original data, and the results are as Figure 4.
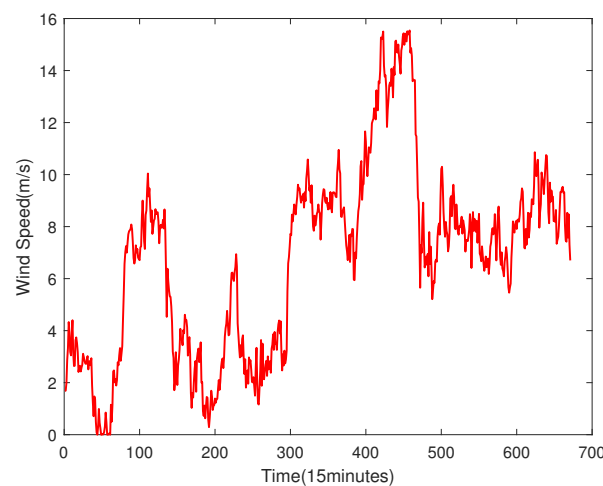

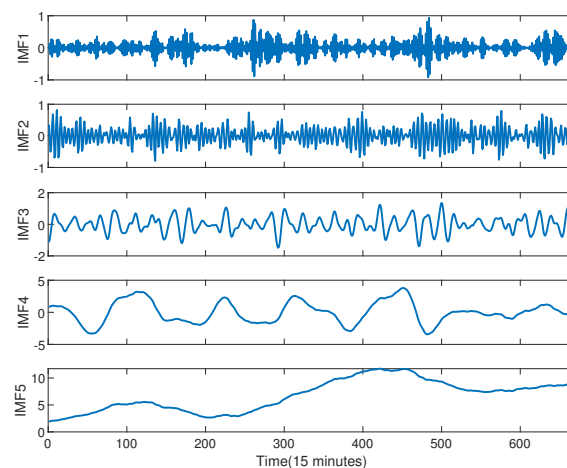**Figure 3.** The raw wind speed data.

Figure 4. The results of VMD algorithm.

Here, the obtained IMF can be added and compared with the original data, as shown in the Figure 5. It can be seen that the result of the decomposition fits well with the original data, which proves that the decomposition effect is good.
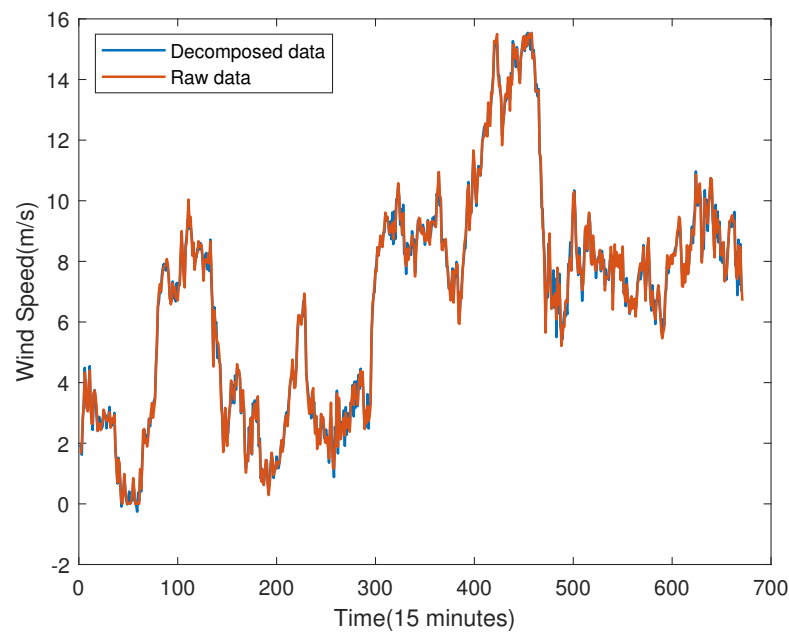

Figure 5. Comparison of decomposed data and raw data.

3.1.2. ARMA Modeling and Partitioning Data

For each IMF, an ARMA (p, q) model needs to be established separately to find its corresponding optimal p value. Among them, p sets the cycle from 5 to 10, and q sets the cycle from 1 to 15. The p-value determines the number of historical data inputs. Using p-value to select the number of wind speed data inputs can solve some problems result from wrong input numbers, such as underfitting caused by too little data selection and overfitting problem caused by too much selection. The q-value is also an important parameter of the ARMA model, and therefore it also needs to be selected carefully based on AIC. It is worth noting that finding the optimal p and q requires using the training set, not the entire dataset, since the testing set is considered unknown.

The values of p and q are related to the Autocorrelation Functions (ACF) and Partial Autocorrelation Functions (PACF) of each IMF, which are shown in Figure 6. And the AIC criterion is used to find the optimal value.[31].

The optimal (p, q) of each IMF obtained by using the AIC criterion is shown in the Table 1.

Table 1. The optimal p and q.

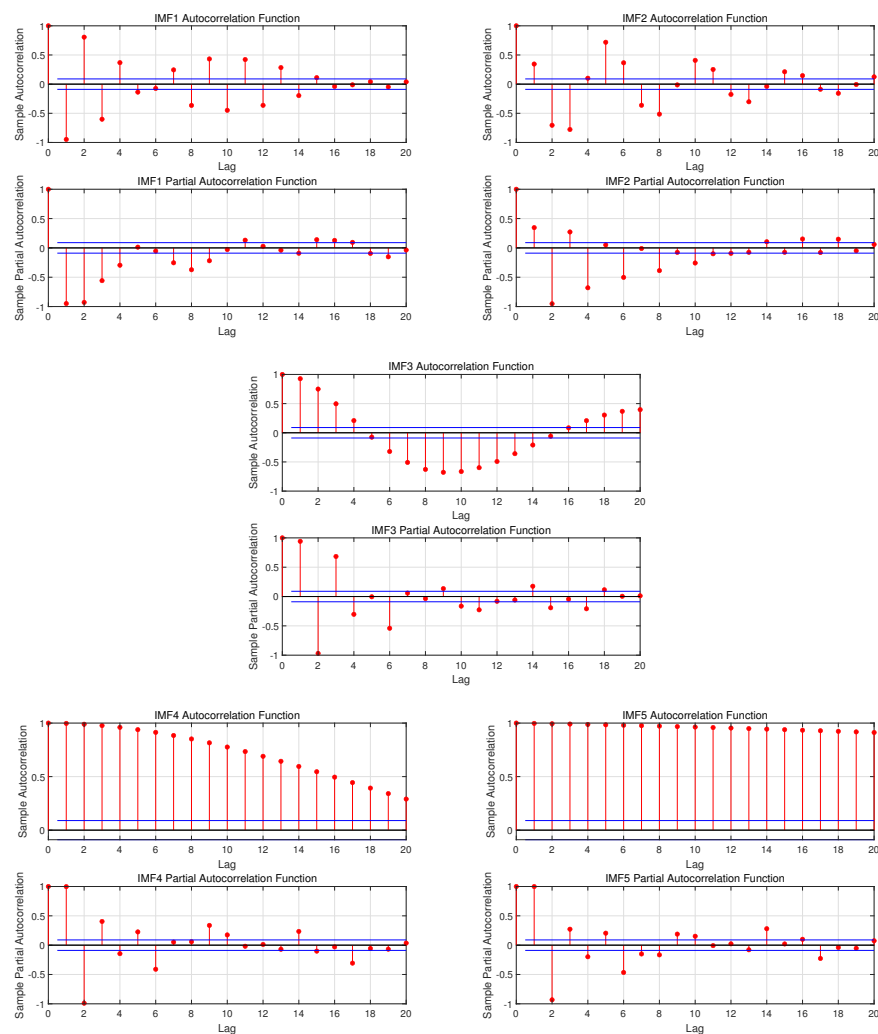| IMF	p	q	AIC
IMF1	6	14	-7.959024352430473
IMF2	8	15	-6.530289111317059
IMF3	10	11	-7.877437434400358
IMF4	9	9	-8.787895064787708
IMF5	9	15	-9.172089217393752

Figure 6. The ACF and PACF of each IMF.

From Table 1, the optimal p of each IMF are obtained according to its minimum AIC value, and they are used to partition data training and testing set like Step3. By this way, the inputs for the BLS model can be found.

3.1.3. Train the BLS Model to Predict Wind Speed

For each IMF, it is necessary to establish a BLS model for prediction. In this step, a triple nested loop is set up to find the optimal number of mapping nodes, windows in mapping feature layers and enhancement nodes for each model. The results are shown in the Table 2 below.

Table 2. The number of mapping nodes, windows in mapping feature layers, and enhancement nodes.

IMF	Set	Mapping nodes	Windows in feature layer	Enhancement nodes
IMF1	Training	4	9	50
	Testing	7	10	14
IMF2	Training	6	9	50
	Testing	6	7	6
IMF3	Training	8	5	50
	Testing	8	6	3
IMF4	Training	8	8	50
	Testing	8	10	16
IMF5	Training	4	9	50
	Testing	4	10	8

The trained optimal BLS model is used to predict each IMF, and the prediction results are shown in the following figures. From Figure 7 and Figure 8, it can be seen that for each IMF, the BLS model can predict well. Next, each prediction result is added and the wind speed prediction result can be obtained, which as shown in the Figure 9.
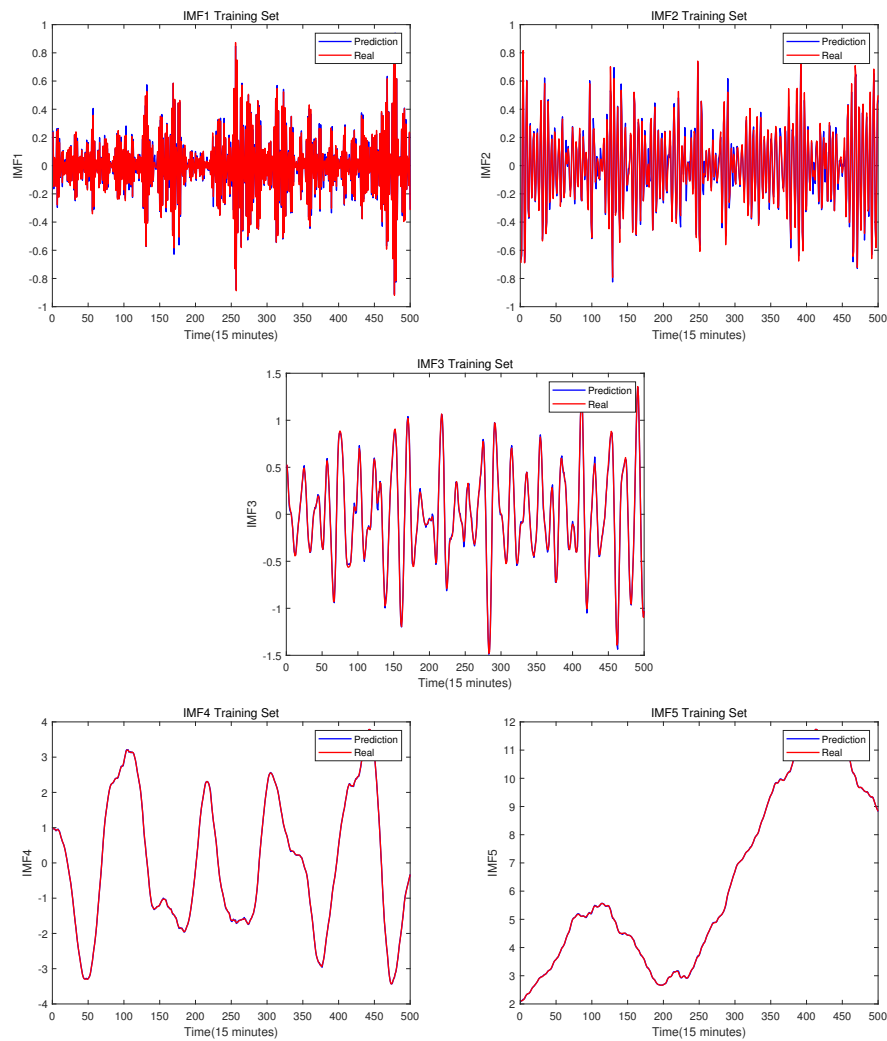

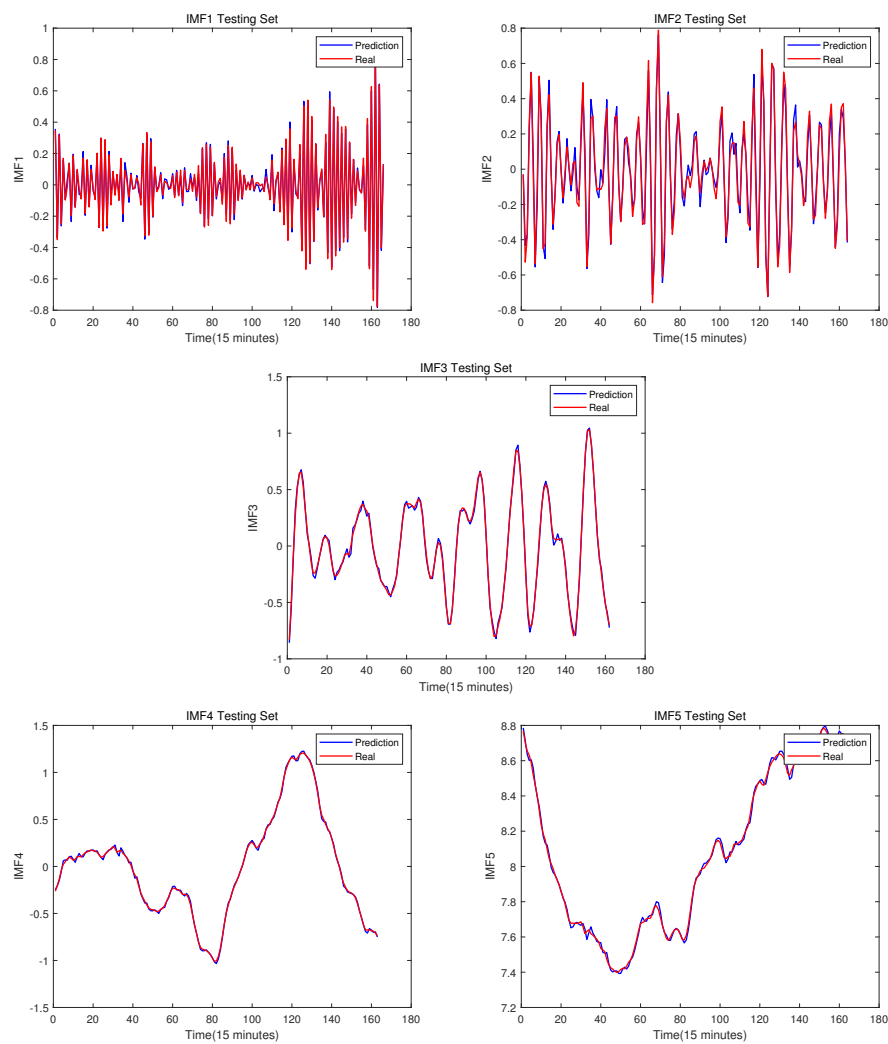**Figure 7.** The training set prediction results of each IMF.

Figure 8. The testing set prediction results of each IMF.

From the Figure 9, we can draw a conclusion that after decomposing the original sequence by using the VMD method, the BLS model performs well in predicting wind speed.
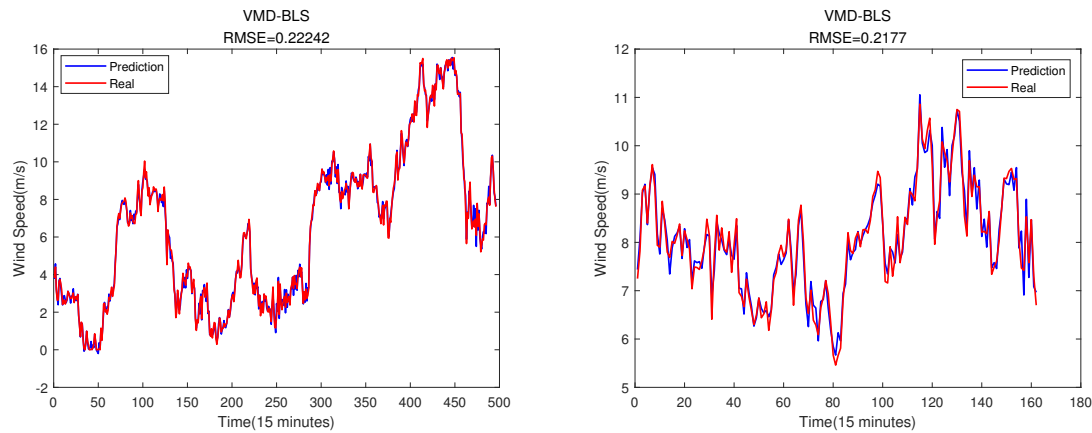

Figure 9. The wind speed prediction result. The training set is on the left, and the testing set is on the right.

3.2. Error Compensation

In order to achieve a better prediction effect, this paper uses error compensation method. The BLS model is trained by the historical error to predict future errors. Then, a more accurate wind speed prediction result can be obtained by adding the error prediction result to the wind speed prediction result.

3.2.1. Obtain Error Set

Error is the difference between the predicted value and the true value. According to Figure 7 and Figure 8, we can obtain the error training set and error testing set, which is shown in Figure 10.
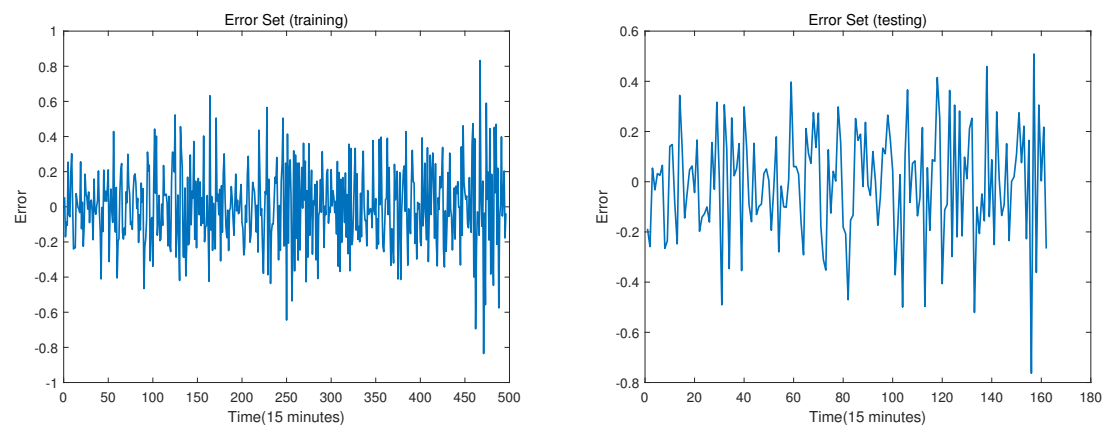

Figure 10. The error set. The training set is on the left, and the testing set is on the right.

3.2.2. ARMA Modeling and Partitioning Error Data

The fluctuation of the error set is small, so there is no need to use VMD method, and the prediction can be made directly. However, the error set is still a univariate sequence. Therefore, it is also necessary to select the optimal p and q values through the ARMA model to divide the data to train model.

The ARMA model is built by using the error training set, since the testing set is considered unknown. The optimal p and q results are shown below.
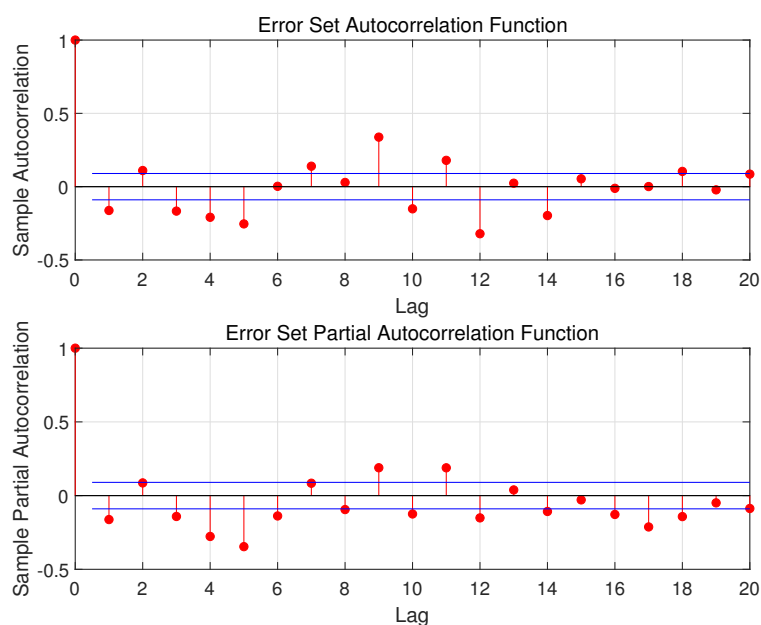

Figure 11. The ACF and PACF of error set.

Table 3. The optimal p and q.

Data Set	p	q	AIC
Error	7	14	-3.527038720946614

After dividing the data with the optimal p value, the data set can be sent to the BLS model for training. The same training method as IMF is used for error data. One loop is set to find the optimal number of mapping nodes, windows in mapping feature layers and enhancement nodes. The result is as shown in the below.

Table 4. The number of mapping nodes, windows in mapping feature layers, and enhancement nodes.

Data Set	Mapping nodes	Windows in feature layer	Enhancement nodes
Error	8	10	44

The error prediction results are shown in the Figure 12 below. It can be seen from the figure that the prediction effect of BLS on the error is not as good as the prediction effect on IMF. However, the prediction result does not need to be very accurate. The final wind speed prediction accuracy will also improve as long as the prediction value can reflect the change trend of the error.

Next, we add the error prediction result to the error prediction result, and the final wind speed prediction result is shown in the Figure 13 below. It is worth noting that in the process of dividing the data, the number of data sets will inevitably decrease by p. This is because in the process of using the first p-1 data to predict the p-th data, the first p-1 data cannot be predicted. However, it has no effect on the prediction accuracy, since we always want to predict the later data, and do not need the missing data in the front.
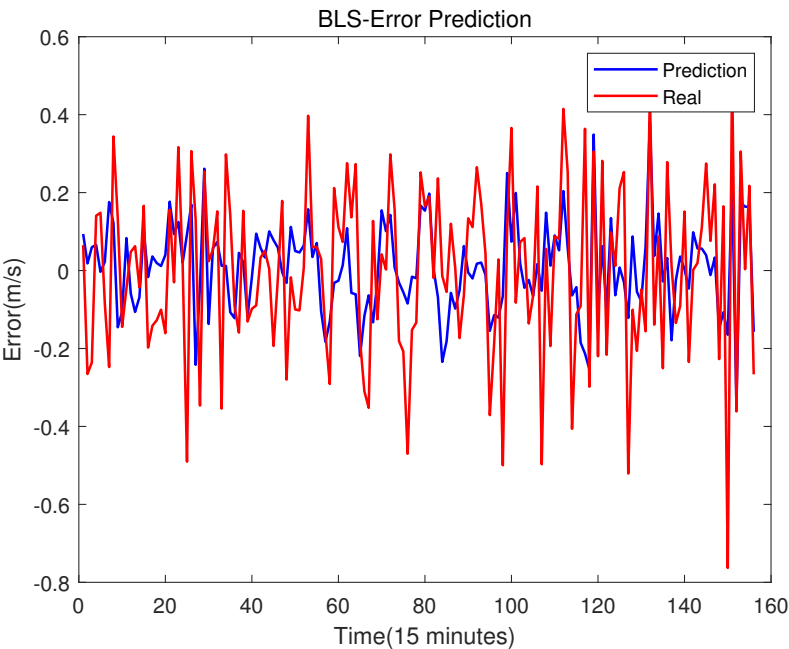

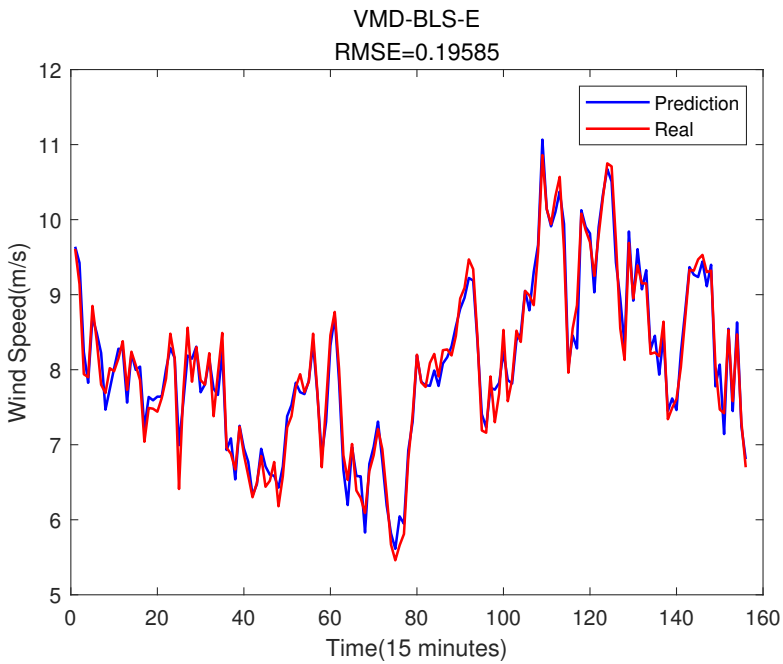**Figure 12.** The error prediction result.

Figure 13. The BLS with error compensation prediction result.

Here, the BLS prediction without error compensation result of the same window size are also shown as Figure 14 below.
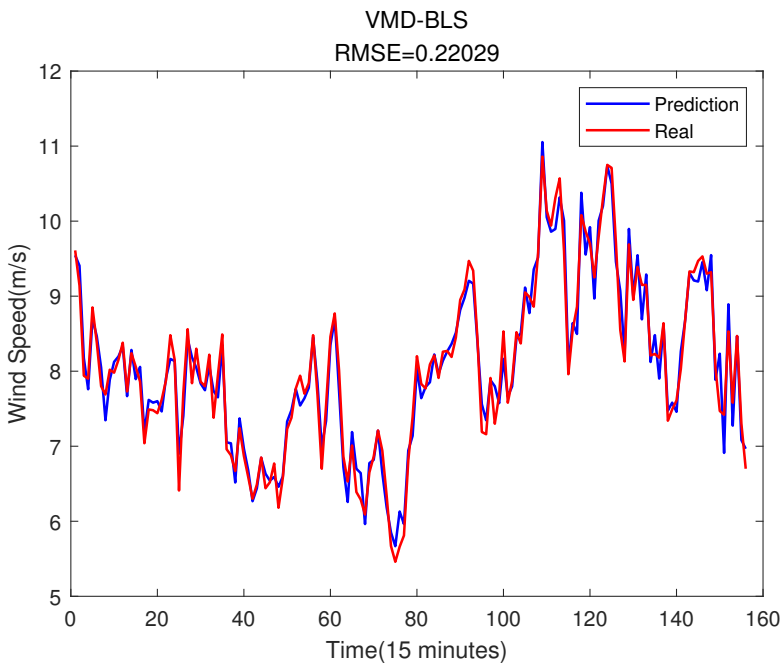

Figure 14. The same window size of BLS without error compensation prediction result.

From the above two figures, it can be seen that the prediction results after error compensation have been significantly improved, and the prediction accuracy is better in some periods of severe wind speed changes.

3.3. Comparison

In order to prove the superiority of the VMD-BLS hybrid wind speed forecasting system proposed in this paper, this subsection uses the BLS model to directly predict the wind speed without the VMD method to decompose. And three models of SVR, ELM, and BPNN are used to replace the BLS model for prediction.

3.3.1. BLS Direct Prediction

The direct prediction mentioned above means that the VMD method and error compensation are not used. The prediction results are shown in the Figure 15 below.
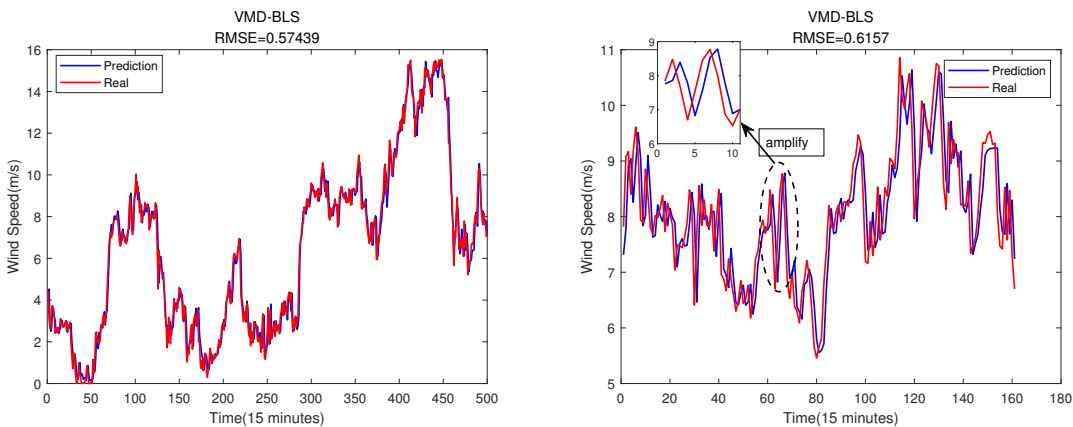

Figure 15. The BLS direct prediction result. The training set is on the left, and the testing set is on the right.

In the Figure 15, the prediction effect seems to be very good. However, if the figure is zoomed in, we will find that the predicted value always lags behind the real value, which will seriously affect the prediction accuracy. Therefore, it is necessary to use some methods such as VMD to process sequences.

3.3.2. Different Models to Predict

The SVR, BPNN, and ELM are used to replace the BLS model in the VMD-BLS hybrid system to predict each IMF. The parameters of each model and prediction results are as follows.

Table 5. The set of SVR. * σ is the parameter of RBF, c is penalty factor, ϵ is the margin of SVR.

IMF	Kernel Function	σ	c	ϵ
IMF1	RBF	0.019	50.8524495	0.02204995
IMF2	RBF	0.1	40.8524495	0.02204995
IMF3	RBF	0.2	10.8524495	0.02204995
IMF4	RBF	0.019	30.8524495	0.02204995
IMF5	RBF	0.0019	50.8524495	0.02204995

Table 6. The set of BPNN.

IMF	Number of Layers	Number of Neurons	Activation Function
IMF1	3	10	Sigmoid
IMF2	3	10	Sigmoid
IMF3	3	10	Sigmoid
IMF4	3	10	Sigmoid
IMF5	3	10	Sigmoid

Table 7. The set of ELM.

IMF	Number of Hidden Neurons	Activation Function
IMF1	50	Sigmoid
IMF2	50	Sigmoid
IMF3	50	Sigmoid
IMF4	50	Sigmoid
IMF5	50	Sigmoid

Compared with Figure 9, the prediction effect of using SVR, BPNN, and ELM is slightly worse than that of BLS. The evaluation indicators of this three models is as shown in Table 8. It can be seen that BLS model performs well on most indicators, only MAPE is slightly lower than the ELM model.

Table 8. The evaluation indicators of this three models.

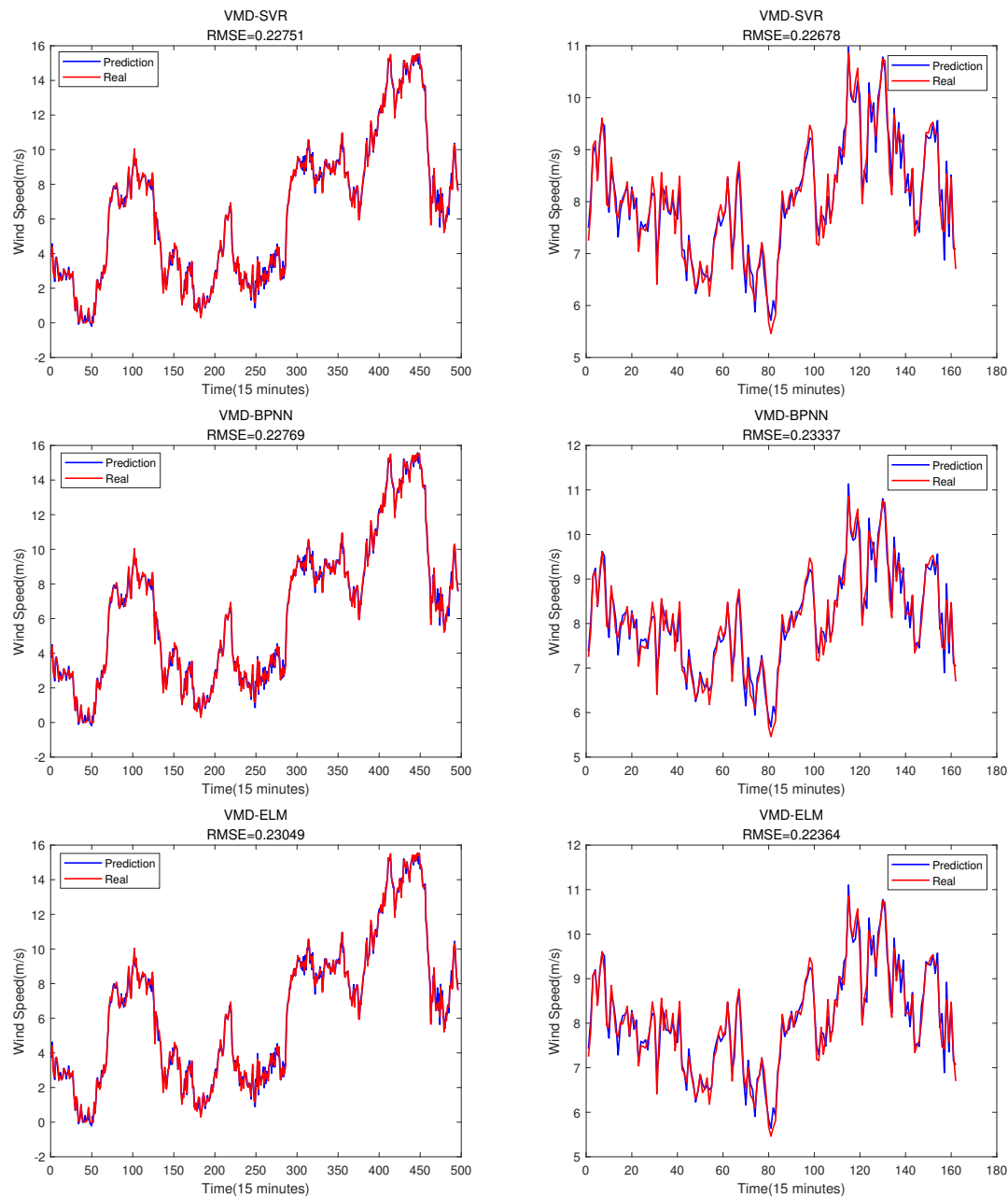| Model	RMSE	MAPE	R^2
BLS	0.22029	0.02195	0.96158
SVR	0.22678	0.02232	0.95863
BPNN	0.23337	0.02314	0.95619
ELM	0.22364	0.02154	0.95977

Figure 16. The SVR BPNN and ELM prediction results. The training set is on the left, and the testing set is on the right.

3.3.3. Different Models for Error Compensation

We have shown above that using the BLS model to predict the wind speed after VMD method is the best. Then, different models are used to predict the error, and the prediction result will be added to the wind speed prediction value. The wind speed prediction values after different model error compensations are as follows.

It can be seen from Figure 17 that after error compensation, the prediction accuracy has been improved, but the accuracy of these three models still has some gaps compared with BLS in Figure 13. The following are the performance indicators of each model after error compensation.
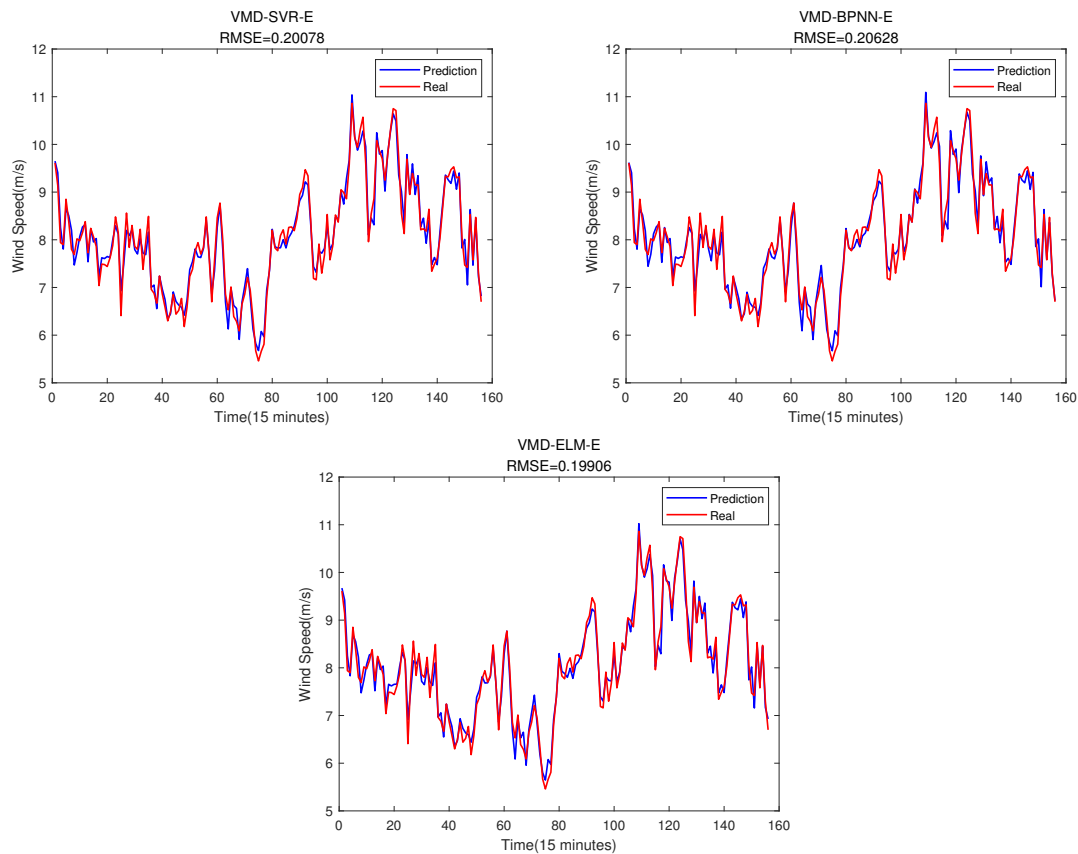
Figure 17. The SVR BPNN and ELM prediction results with error compensation.

From Table 9, the performance indicators after using the BLS model for error compensation are better than the other three models.

Subsection 3.3 proves two things: one is that VMD method can reduce the prediction lag phenomenon and the other thing is that using the BLS model to carry out error compensation can further improve the accuracy.

Table 9. The evaluation indicators of this three models after error compensation.

Model	RMSE	MAPE	R^2
BLS	0.19585	0.01963	0.96963
SVR	0.20078	0.02065	0.96808
BPNN	0.20628	0.02071	0.96631
ELM	0.19906	0.02019	0.96862

4. Conclusions

In this paper, a hybrid wind speed forecasting system based on VMD-BLS with error compensation is proposed, which is suitable for predicting univariate wind speed series. First, the wind speed sequence is decomposed into multiple IMF by the VMD algorithm. Second, an ARMA model is established for each IMF to find the optimal number of inputs. Then, the BLS model is used to predict the wind speed. In order to further improve the prediction accuracy, an error compensation method by using BLS model is proposed. The error is predicted by the BLS model, and the final wind speed prediction can be obtained by adding the error prediction result to the wind speed prediction result. After verification experiments with real data, it is proved that this method can reduce the phenomenon of prediction lag, and the prediction effect is better than most models, which can be applied in practice.

Although we have proposed some methods to improve forecast accuracy, there is still much room for improvement in this work. For example, a better prediction result can be obtained if a more accurate algorithm can be found to decompose the original data. And it is also worth considering to make this method an online prediction system, which can continuously update the data. The online system is more in line with practical applications, and we will study these workslater.

Abbreviations

The following abbreviations are used in this manuscript:

ARMA	Autoregressive Moving Average
BLS	Broad Learning System
BPNN	Backpropagation Neural Network
ELM	Extreme Learning Machine
IMF	Intrinsic Mode Functions
SVR	Support Vector Regression
VMD	Variational Mode Decomposition

References

1. Zhang, Y.; Zhao, Y.; Kong, C.; Chen, B. A new prediction method based on VMD-PRBF-ARMA-E model considering wind speed characteristic. *Energy Conversion and Management* **2020**, *203*, 112254.
2. Xu, X.; Hu, S.; Shi, P.; Shao, H.; Li, R.; Li, Z. Natural phase space reconstruction-based broad learning system for short-term wind speed prediction: Case studies of an offshore wind farm. *Energy* **2023**, *262*, 125342.
3. Sacie, M.; Santos, M.; López, R.; Pandit, R. Use of state-of-art machine learning technologies for forecasting offshore wind speed, wave and misalignment to improve wind turbine performance. *Journal of Marine Science and Engineering* **2022**, *10*, 938.
4. Fan, Q.; Wang, X.; Yuan, J.; Liu, X.; Hu, H.; Lin, P. A Review of the Development of Key Technologies for Offshore Wind Power in China. *Journal of Marine Science and Engineering* **2022**, *10*, 929.
5. GWEC GLOBAL WIND REPORT 2023. https://gwec.net/globalwindreport2023/.
6. Xu, W.; Liu, P.; Cheng, L.; Zhou, Y.; Xia, Q.; Gong, Y.; Liu, Y. Multi-step wind speed prediction by combining a WRF simulation and an error correction strategy. *Renewable Energy* **2021**, *163*, 772–782.
7. Sun, G.; Jiang, C.; Cheng, P.; Liu, Y.; Wang, X.; Fu, Y.; He, Y. Short-term wind power forecasts by a synthetical similar time series data mining method. *Renewable energy* **2018**, *115*, 575–584.
8. Du, Y.; Zhou, S.; Jing, X.; Peng, Y.; Wu, H.; Kwok, N. Damage detection techniques for wind turbine blades: A review. *Mechanical Systems and Signal Processing* **2020**, *141*, 106445.
9. Wang, X.; Guo, P.; Huang, X. A review of wind power forecasting models. *Energy procedia* **2011**, *12*, 770–778.
10. Jung, J.; Broadwater, R.P. Current status and future advances for wind speed and power forecasting. *Renewable and Sustainable Energy Reviews* **2014**, *31*, 762–777.
11. Zhang, Y.; Yang, J.; Wang, K.; Wang, Z. Wind power prediction considering nonlinear atmospheric disturbances. *Energies* **2015**, *8*, 475–489.
12. Jiao, X.; Zhou, X.; Yang, Q.; Zhang, Z.; Liu, W.; Zhao, J. An improved optimal torque control based on estimated wind speed for wind turbines. In Proceedings of the 2022 13th Asian Control Conference (ASCC). IEEE, 2022, pp. 2024–2029.
13. Tang, X.; Yang, Q.; Wang, K.; Stoevesandt, B.; Sun, Y. Optimisation of wind farm layout in complex terrain via mixed-installation of different types of turbines. *IET Renewable Power Generation* **2018**, *12*, 1065–1073.
14. Hannan, E.J. The estimation of the order of an ARMA process. *The Annals of Statistics* **1980**, pp. 1071–1081.
15. de Oliveira, J.F.; Silva, E.G.; de Mattos Neto, P.S. A hybrid system based on dynamic selection for time series forecasting. *IEEE Transactions on Neural Networks and Learning Systems* **2021**, *33*, 3251–3263.

16. Ordulj, M.; Šantić, D.; Matić, F.; Jozić, S.; Šestanović, S.; Šolić, M.; Veža, J.; Ninčević Gladan, Ž. Analysis of the Influence of Seasonal Water Column Dynamics on the Relationship between Marine Viruses and Microbial Food Web Components Using an Artificial Neural Network. *Journal of Marine Science and Engineering* **2023**, *11*, 639.

17. Alshahri, A.H.; Elbisy, M.S. Assessment of Using Artificial Neural Network and Support Vector Machine Techniques for Predicting Wave-Overtopping Discharges at Coastal Structures. *Journal of Marine Science and Engineering* **2023**, *11*, 539.

18. Zhang, C.; Ding, M.; Wang, W.; Bi, R.; Miao, L.; Yu, H.; Liu, L. An improved ELM model based on CEEMD-LZC and manifold learning for short-term wind power prediction. *IEEE access* **2019**, *7*, 121472–121481.

19. Liu, M.; Cao, Z.; Zhang, J.; Wang, L.; Huang, C.; Luo, X. Short-term wind speed forecasting based on the Jaya-SVM model. *International Journal of Electrical Power & Energy Systems* **2020**, *121*, 106056.

20. Jiao, X.; Yang, Q.; Xu, B. Hybrid intelligent feedforward-feedback pitch control for VSWT with predicted wind speed. *IEEE Transactions on Energy Conversion* **2021**, *36*, 2770–2781.

21. Kim, J.Y.; Oh, J.S. Electric Consumption Forecast for Ships Using Multivariate Bayesian Optimization-SE-CNN-LSTM. *Journal of Marine Science and Engineering* **2023**, *11*, 292.

22. Shahid, F.; Zameer, A.; Muneeb, M. A novel genetic LSTM model for wind power forecast. *Energy* **2021**, *223*, 120069.

23. Kumar, K.; Haider, M.T.U. Enhanced prediction of intra-day stock market using metaheuristic optimization on RNN–LSTM network. *New Generation Computing* **2021**, *39*, 231–272.

24. Zhang, Z.; Li, M.; Lin, X.; Wang, Y.; He, F. Multistep speed prediction on traffic networks: A deep learning approach considering spatio-temporal dependencies. *Transportation research part C: emerging technologies* **2019**, *105*, 297–322.

25. Chu, X.; Jin, H.; Li, Y.; Feng, J.; Mu, W. CDA-LSTM: an evolutionary convolution-based dual-attention LSTM for univariate time series prediction. *Neural Computing and Applications* **2021**, *33*, 16113–16137.

26. Dragomiretskiy, K.; Zosso, D. Variational mode decomposition. *IEEE transactions on signal processing* **2013**, *62*, 531–544.

27. Dibaj, A.; Ettefagh, M.M.; Hassannejad, R.; Ehghaghi, M.B. A hybrid fine-tuned VMD and CNN scheme for untrained compound fault diagnosis of rotating machinery with unequal-severity faults. *Expert Systems with Applications* **2021**, *167*, 114094.

28. McLeod, A.I.; Li, W.K. Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of time series analysis* **1983**, *4*, 269–273.

29. Chen, C.P.; Liu, Z. Broad learning system: An effective and efficient incremental learning system without the need for deep architecture. *IEEE transactions on neural networks and learning systems* **2017**, *29*, 10–24.

30. Fathian, F.; Mehdizadeh, S.; Sales, A.K.; Safari, M.J.S. Hybrid models to improve the monthly river flow prediction: Integrating artificial intelligence and non-linear time series models. *Journal of Hydrology* **2019**, *575*, 1200–1213.

31. Sakamoto, Y.; Ishiguro, M.; Kitagawa, G. Akaike information criterion statistics. *Dordrecht, The Netherlands: D. Reidel* **1986**, *81*, 26853.