Ultra-Short-Term Wind Power Prediction Based on Double Decomposition and LSSVM

BIN QIN [a], XUN HUANG [b], XIN WANG [a*], LINGZHONG GUO [c]

[a]*School of Electrical & Information Engineering, Hunan University of Technology, Zhuzhou, Hunan, 412007, China*

[b]*College of Railway Transportation, Hunan University of Technology, Zhuzhou, Hunan, 412007, China*

[c]*Department of Automatic and Systems Engineering, The University of Sheffield, S1 3JD, UK*

Abstract

In order to reduce the influence of the random fluctuation on wind power prediction, a new ultra-short term wind power prediction model, based on wavelet decomposition (WD)-variational mode decomposition (VMD)-least square support vector machine (LSSVM), is proposed in this paper. The method is based on the double decomposition and LSSVM, where the wind power sequence is decomposed by WD into low and high frequency components which are further decomposed by VMD to obtain many modal components with tendency and periodicity. Multiple LSSVM prediction models are then established with historical wind power data and weather data as the inputs to obtain the predicted values of the multiple modal components. The final predicted values of wind power are achieved by data fusion of outputs of these LSSVM models. The experimental results show that the MAPE (Mean Absolute Percentage Error) of the combined prediction model is 4.66%, which is the best compared with eight benchmark models. This demonstrates the high performance of the proposed WD-VMD-LSSVM model for short-term prediction of wind power.

Keywords: Wind power prediction; Wavelet decomposition; Variational modal decomposition; Data fusion; Least squares support vector machine

1. Introduction

With the development of new energy technologies, more and more attention has been paid to wind power generation due to its being cost-effective and environmentally friendly. Wind power prediction plays a crucial role during wind power generation (Wang et al., 2019). However, wind is considered highly intermittent, non-stationary, highly random, and difficult to predict (Sun et al., 2020), which brings challenges to the accurate prediction of wind power output. In general, wind power prediction is to process and analyze historical wind power data and external influencing factors, and establish a specific mathematical model to predict future wind power output, which can be long-term (one year to ten years), medium-term (one month to one year), short-term (one day to one week), or ultra-short-term (seconds to hours) (Li et al., 2020). Among them, the ultra-short-term wind power prediction is of great significance to the safety and stability of the power systems as well as economical energy saving, and helps to improve their operating performance (Li and Li, 2019). Therefore, in this paper, we will focus on the ultra-short-term prediction of wind power output.

In the past few decades, many mathematical models and methods have been used to predict wind power. These methods can be generally divided into three categories. The first category is to construct mechanism models of wind power through physical information, such as the use of spatial correlation (Alexiadis et al., 1999) and numerical weather forecast (Zhao et al., 2016), and other information. The second category is to use statistical methods or artificial intelligence methods to find the correlation between historical data and then establish data-driven models of wind power, such as autoregressive integrated moving average model (AIMAM) (Korprasertsak and Leephakpreeda, 2016), neural networks (NN) (Mason et al., 2018), and support vector machine (SVM) (Wu and Gao, 2018). The third category is the combined prediction methods, such as the probabilistic WPF ensemble of ensemble empirical mode decomposition (EEMD)-sample entropy (SE)-extreme learning machine(ELM) proposed by (Zhang, et al., 2014), which combines the signal decomposition technology with the prediction model to obtain better prediction performance. Wang et al. (Wang et al., 2020) proposed a complete EEMD with adaptive noise (CEEMDAN) in combination with neural network-induced ordered weighted averaging (IOWA) algorithm. Compared with EEMD and EMD, the decomposition process of EEMDAN is more complete and outperforms the combined prediction model of the single prediction model ELM and EEMD. To enhance the predictive performance of single mechanism models or single data-driven models, a fusion of multiple predictive models is desirable, which will be the focus of the study in this paper. In view of the random fluctuation of wind power series, it is difficult to fully extract data feature information (Abedinia et al., 2020). Therefore, it is necessary to stabilize the series first, so that some statistics in the stabilized series do not change with time, enhance the regularity of the series, and then improve the prediction accuracy (Sun and Zhao, 2020). The sequence stabilization can be implemented by using signal decomposition techniques, such as EMD, wavelet decomposition (WD) and variational mode decomposition (VMD) techniques. Among them, WD can be used to obtain unstable and nonlinear features from the signals and analyze the signals with intermittent and instability (Liu et al., 2021). The problem of modal mixing is appeared in the application of EMD (Lu et al., 2021), and the EEMD obtained by improving EMD has been used to solve this problem (Liu et al., 2016), while VMD not only has no problem of modal mixing, but also its number of components is far smaller than EEMD (Zhang et al., 2019). Given that the single signal decomposition methods do not fully consider the mutual influence of wind power sequence across different frequencies (Wu et al., 2020), a combined WD-VMD decomposition of the wind power sequence is adopted in this paper.

Regarding the nonlinear characteristics of wind power prediction, NN and SVM have been shown to be powerful in dealing with them (Xu et al., 2021, Khosravi, et al., 2018). The results of wind speed prediction by Mohandes et al. (Mohandes et al., 2014) using SVM and multi-layer perceptron (MLP) have shown that SVM has higher prediction accuracy than MLP. However, SVM is often replaced by least squares support vector machines (LSSVM) due to its complex computation (Zhongda, 2020). De Giorgi et al., 2014 used LSSVM and ANN to compare and predict wind power, and the results showed that the LSSVM model has higher prediction accuracy. Considering that there will be some information loss after the decomposition of the wind power sequence and error superposition caused by multiple predictors, only summing all modal components to obtain the final predicted values could have a large error. Therefore, the final predicted values are obtained by data fusion of all modal components (Xu et al., 2019).

Taking all the discussions above, an ultra-short-term wind power prediction model, based on WD-VMD and LSSVM, is proposed in this paper. WD is used to decompose the wind power sequence into low and high frequency components, and VMD is then used to further decompose the low frequency components to obtain multiple sub-components. An LSSVM model for each sub-component is constructed with historical wind power data and weather data and the predicted values of each sub-component are further fused to obtain the final predicted values by an LSSVM model. The main contributions of this work are as follows:

1) The double decomposition technology of WD and VMD is used to decompose the wind power sequence to stabilize it and enhance its regularity, based on which a new combined prediction model WD-VMD-LSSVM is proposed for ultra-short-term prediction of wind power.

2) The proposed model takes into consideration both historical wind power data and historical weather data and a data fusion method is proposed to obtain the final wind

power prediction values, reducing the influence of the information loss and error superposition from multiple LSSVM predictors.

The paper is organized as follows: Section 2 presents the double decomposition method and the combined prediction model WD-VMD-LSSVM; Section 3 compares and analyzes the experimental results; conclusions are drawn in Section 4.

2. Ultra-short-term wind power prediction modeling based on WD-VMD-LSSVM (data fusion)

2.1. WD-VMD-LSSVM (data fusion) combined prediction model

The structure of the proposed WD-VMD-LSSVM (data fusion) combined prediction model is shown in Figure 1.

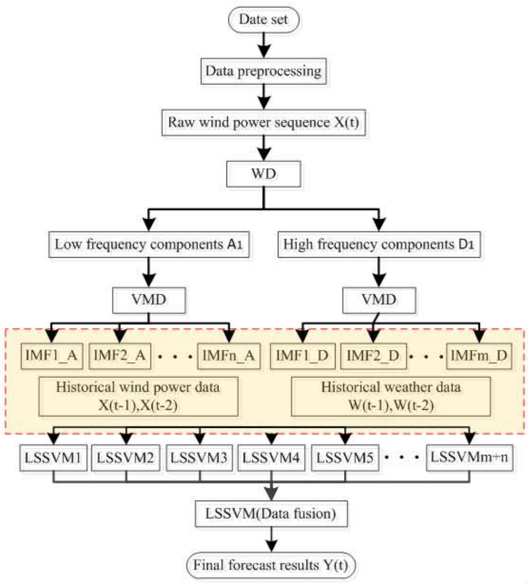

Fig. 1. Structure of WD-VMD-LSSVM prediction model.

2.2. WD

The wavelet transform is a time-frequency transform, which is ideal for analyzing nonstationary signals such as wind power output. Discrete wavelet transform separates the original signal into sub-sets at a smaller size, capturing signals with more information (Gürsoy and Engin, 2019). The Daubechies wavelet functions (Azimi et al., 2016), being effective for dealing with irregular signals such as wind power sequences, are adopted to implement the wind power signal decomposition.

In particular, the db10 wavelet (Adly et al, 2016) is used as the basic wavelet to perform discrete wavelet transform (DWT) on the wind power data, as shown in Eq. (1) (Azimi et al., 2016):

$$\psi_{p,q}(t) = \frac{1}{\sqrt{s_0^p}} \psi\left(\frac{t - q\tau_0 s_0^p}{s_0^p}\right) \tag{1}$$

where s_0^p is a fixed dilation step and τ_0 is the translation factor which depends on the dilation step; both p and q are integers.

The scaling and wavelet functions are used by DWT, as shown in Eqs. (2-3):

$$\varphi(2^p t) = \sum_{i=0}^{q} h_{p+1}(q)\varphi(2^{p+1}t - q) \tag{2}$$

$$\psi(2^p t) = \sum_{i=0}^{q} g_{p+1}(q)\varphi(2^{p+1}t - q) \tag{3}$$

where g and h are the wavelet and scaling coefficients, respectively.

A wind power signal $W(t)$ can be therefore written as:

$$W(t)=\sum_{i=0}^{q}h_{p+1}(q)\varphi(2^{p+1}t-q)+\sum_{i=0}^{q}g_{p+1}(q)\varphi(2^{p+1}t-q) \tag{4}$$

The cascade decomposition of a wind power time series signal $W(t)=[w_1,w_2,w_3,...,w_n]$ can be defined as follows:

$$w(t)=f_A(t)+f_D(t)$$

(5)

$$f_A(t)=f_{A_n}(t)+\sum\nolimits_{m=1}^{n}f_{D_m}(t) \tag{6}$$

where $f_{A_n}(t)$ and $f_{D_n}(t)$ are approximation and detail coefficients, respectively. $f_A(t)$ and $f_D(t)$ are the low and high frequency components, respectively.

2.3. VMD

VMD decomposition is an adaptive and non-recursive signal decomposition proposed by (Dragomiretskiy and Zosso, 2013). Compared with WD and EMD, the signal can be better restored by VMD (Zhou et al., 2022). Using this method, the frequency components are decomposed to obtain a series of natural mode components at low and high frequencies:

Low frequency modal components: [IMF1_A, IMF2_A, …, IMFn_A];

High frequency modal components: [IMF1_D, IMF2_D, …, IMFm_D];

Then a variational problem is solved to minimize the sum of the bandwidths of the low/high frequency components with the constraint that it is equal to the original low/high frequency components, as shown in Eq. (7) (Dragomiretskiy and Zosso, 2013):

$$\begin{cases} \min\limits_{\{u_k\}\{\omega_k\}} \left\{ \sum_{k=1}^{k} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\tau t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ s.t. \sum_{k=1}^{k} u_k = f_{A/D} \end{cases} \tag{7}$$

where k is the modal number, $\delta(t)$ is the Dirac function, $\{u_k\}$ is the k modal component, $\{\omega_k\}$ is the center frequency corresponding to the k modal component, $f_{A/D}$ is the original low/high frequency component from Eqs.(5-6), t is the time, and j is the imaginary unit. Introducing the quadratic penalty factor α and Lagrange multiplier λ, Eq. (7) can be re-formulated as Eq. (8):

$$\begin{aligned} L(\{u_k\},\{\omega_k\},\lambda) = \alpha \sum_{k=1}^{k} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\tau t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f_{A/D}(t) - \sum_{k=1}^{k} u_k(t) \right\|_2^2 \\ + \left\langle \lambda(t), f_{A/D}(t) - \sum_{k=1}^{k} u_k(t) \right\rangle \end{aligned} \tag{8}$$

The final update process of VMD is shown in the Eqs. (9-11):

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}_{A/D}(\omega) - \sum_{i \neq k}^{k} \hat{u}_i^n(\omega) + \frac{\hat{\lambda}^n(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \tag{9}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega \left| \hat{u}_k^{n+1}(\omega) \right|^2 d\omega}{\int_0^\infty \left| \hat{u}_k^{n+1}(\omega) \right|^2 d\omega} \tag{10}$$

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau(\hat{f}_{A/D}(\omega) \sum_{k=1}^{k} \hat{u}_k^{n+1}(\omega)) \tag{11}$$

where $\hat{f}_{A/D}(\omega)$, $\hat{u}_i^n(\omega)$, $\hat{u}_k^n(\omega)$, $\hat{\lambda}^n(\omega)$ are the Fourier transform of $f_{A/D}(t)$, $u_i(t)$, $u_k(t)$, $\lambda^n(t)$ respectively, τ is the update parameter, and n is the number of iterations.

In general, a sequence is said to be convergent if it approaches some limit and furthermore every convergent sequence is a Cauchy sequence. Therefore, let $\{\hat{u}_k^n\}$ be the sequence, if there exists $\{\hat{u}_k^m\}$, for any given number $q > 0$, there is always $\left|\hat{u}_k^n - \hat{u}_k^M\right| < q$ when $n > M$ and M is a positive integer, that is, $\{\hat{u}_k^n\}$ converges to $\{\hat{u}_k^m\}$.

Thus, we will use the following criterion for convergence: for a given precision convergence criterion $\varepsilon > 0$, if Eq. (12) is satisfied, thenVMD converges, that is, Eqs (9-11) stops updating.

$$\sum_k \frac{\left\|\hat{u}_k^{n+1} - \hat{u}_k^n\right\|_2^2}{\left\|\hat{u}_k^n\right\|_2^2} < \varepsilon \tag{12}$$

2.4. LSSVM

Compared with SVM, the LSSVM facilitates the solution of Lagrange multiplier α and improves the convergence speed (Xie et al, 2021). Different kernel functions have different effects on the regression performance of LSSVM, and radial basis functions (RBF) (Zhang et al, 2018) are used here as shown in Eq. (13) (Zhang et al, 2018):

$$K(x_i, x_j) = \exp(-\frac{\left\|x_i - x_j\right\|}{2\sigma^2}) \tag{13}$$

where σ is the kernel function parameter.

Set $\{(x_1, u_{k_1}),...,(x_n, u_{k_n})\}$ as the input training sample set and use the following high-dimensional linear mapping to fit the training sample set, as shown in Eq. (14) (Zhang et al, 2018):

$$f_k(x) = w^T \varphi(x) + b \tag{14}$$

where w is the weight, b is the bias, and $\varphi(x)$ is the nonlinear mapping function. According to the structured risk minimization principle, the LSSVM regression problem can be transformed into the following constrained optimization problem:

$$\begin{cases} \min \frac{1}{2} w^T w + \gamma \sum_{i=1}^{n} e_i^2 \\ s.t. u_{k_i} = w^T \varphi(x_i) + b + e_i, i = 1,2,...,n \end{cases} \tag{15}$$

where e_i is the training error; γ is the regularization factor, u_{k_i} is the outputs of the *kth* LSSVM model from Eq.(9); x_i includes historical wind power data and weather data is the inputs of LSSVM model.

Eq. (15) is transformed into the dual problem of Eq. (16) by a Lagrange multiplier method:

$$L = \frac{1}{2} w^T w + \gamma \sum_{i=1}^{n} e_i^2 + \sum_{i=1}^{n} \left[\lambda_i (w^T \varphi(x_i) + b + e_i - u_{k_i}) \right] \tag{16}$$

where λ_i is the Lagrange multiplier.

The partial derivative of Eq. (16) is obtained by KKT conditions:

$$\begin{cases} \partial L / \partial w = 0 \\ \partial L / \partial b = 0 \\ \partial L / \partial e_i = 0 \\ \partial L / \partial \lambda_i = 0 \end{cases} \Rightarrow \begin{cases} w = \sum_{i=1}^{n} \lambda_i \varphi(x_i) \\ \sum_{i=1}^{n} \lambda_i = 0 \\ \lambda = \gamma e_i \\ u_{k_i} = w^T \varphi(x_i) + b + e_i \end{cases} \tag{17}$$

Eq. (17) is arranged into Eq. (18), and the linear Eq. (18) is solved to obtain λ and b of the LSSVM model

$$\begin{cases} \begin{bmatrix} 0 & 1^{\mathrm{T}} \\ 1 & K + I / \gamma \end{bmatrix} \begin{bmatrix} b \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ u_k \end{bmatrix} \\ \begin{cases} 1 = (1,1,...1)^{\mathrm{T}} \\ \lambda = (\lambda_1, \lambda_2, ...\lambda_n)^{\mathrm{T}} \\ u_k = (u_{k_1}, u_{k_2}, ...u_{k_n})^{\mathrm{T}} \\ K(x_i, x_j) = \varphi(x_i)^{\mathrm{T}} \varphi(x_j) \end{cases} \end{cases} \tag{18}$$

where K is the kernel function matrix and I is the nth-order unit matrix.

The LSSVM prediction model is given by:

$$f_k(x) = \sum_{i=1}^{n} (\lambda_i K(x, x_i)) + b \tag{19}$$

Using Eq. (19), the predicted values of k modal components are obtained. In the same way, A LSSVM data fusion model is constructed according to Eqs. (13-19) with the predicted values of k modal components as inputs and the actual value of wind power as the output, thus the final wind power predicted value $Y(t)$ is given by:

$$Y(t) = \sum_{i=1}^{n} (\lambda_{1i} K(f_k(x), f_{ki}(x))) + b_1 \tag{20}$$

2.5. Experimental data and its preprocessing

The dataset used in the paper contains onshore wind power data and weather data (wind direction and speed) in Spain from January 1, 2015 to 2016, with a sampling time of 1 hour, and a total of 8782 pieces of data. In this paper, the data mining method was used to find out that the data missing rate of the wind power data set is 0.0018%, and the missing values were filled by means of mean interpolation. The data on the 1st January 2016 was used as the test set (24h).

The predicted horizon was set to be two hours, which means the current prediction is made based on the wind power and weather data two hours ago. The input data were normalized for consistency. The dataset was taken from the Kaggle website (Nicholas Jhana, 2019).

2.6. Evaluation indicators

Three evaluation indicators are used to evaluate the performance of the proposed model: Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and fitting coefficient R-square as shown in Eqs. (21-23):

$$MAPE = \frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{21}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{22}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (\hat{y}_i - \overline{y})^2}{\sum_{i=1}^{N} (\hat{y}_i - y_i)^2} \tag{23}$$

where N is the total number of wind power prediction samples, $y_i(t)$ is the original wind power data at time, $\hat{y}_i(t)$ is the predicted value of wind power at time, and $\hat{y}$ represents the mean value of the original wind power data.

2.7. Ultra-short-term wind power prediction based on WD-VMD-LSSVM (data fusion)

3.3.1 WD

Because the length of the wind power sequence is 8758, there are six layers from the wavelet decomposition. DWT was performed on the wind power sequence by db10 wavelet, and then the A_1 and D_1 were obtained as Eq. (1). The signals before and after the WD decomposition of the wind power sequence are shown in Figure 2.

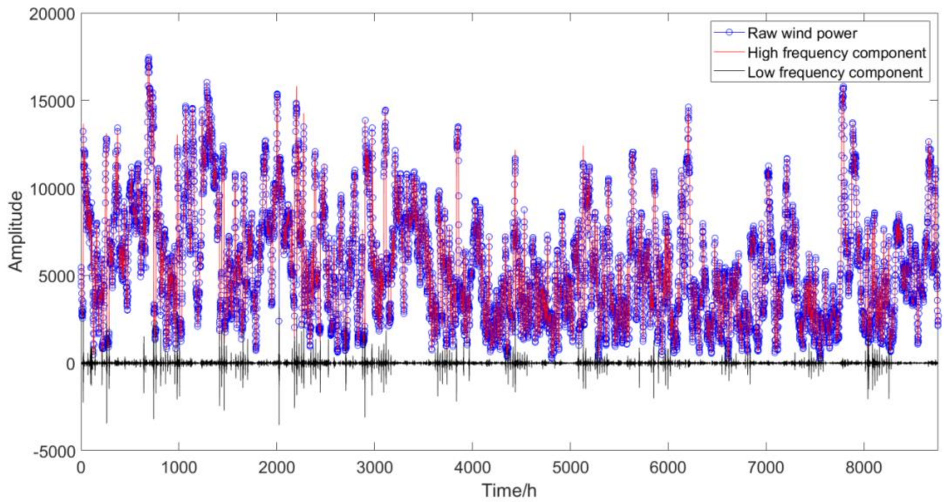

Fig. 2. Signal comparison

3.3.2 VMD

The VMD algorithm needs to set the number of modal decompositions K first. The selection of the K value is very important. Different K values will produce different decomposition effects. The selection principle is to avoid modal aliasing and preserve the main feature information in the signal as much as possible (Xinghua et al., 2019). Here the method of observing the centre frequency of each mode after decomposition was used to determine the K value (Fu et al., 2021). The centre frequencies of A_1 and D_1 after VMD with different K values are shown in Table 1.

It can be observed from Table 1 that when $K = 9$, the two centre frequencies in the low frequency component 2233.14×10^{-4} and the high frequency component 2172.45×10^{-4} are too close, and the phenomenon of modal aliasing may occur. When K=7, the difference between the two centre frequencies of the low frequency component 1867.83×10^{-4} and the high frequency component 2305.28×10^{-4} is too large, and the phenomenon of under-decomposition may occur, so it is more appropriate to choose K=8. Therefore, the time-domain waveforms and spectrograms of the A_1 and D_1 were obtained by using the VMD of K=8, as shown in Figure 3. It can be seen from Figure 3 that the signal decomposition results are satisfactory.

Table 1. Corresponding centre frequency to different K values in the low and high frequency components

K	Centre frequency($\times10^{-4}$ Hz)								
5(low)	1.53	144.85	400.99	738.19	1281.96	-	-	-	-
6(low)	1.25	129.92	308.02	566.94	965.69	1583.58	-	-	-
7(low)	1.05	113.92	259.21	489.69	802.70	1225.83	1867.83	-	-

8(low)	1.00	110.02	248.13	466.94	734.59	1051.87	1496.93	2061.20	-
9(low)	0.97	107.53	241.32	453.02	694.58	963.55	1305.73	1728.55	2233.14
5(high)	2534.48	2977.58	3432.99	3986.31	4671.78	-	-	-	-
6(high)	2339.96	2723.00	3108.89	3545.31	4085.54	4721.23			
7(high)	2305.28	2635.78	2937.41	3286.29	3688.88	4194.96	4744.32		
8(high)	2291.74	2614.09	2865.27	3135.04	3509.78	3874.27	4325.68	4784.49	
9(high)	2172.45	2429.41	2659.92	2893.22	3141.12	3480.44	3892.02	4330.02	4789.93

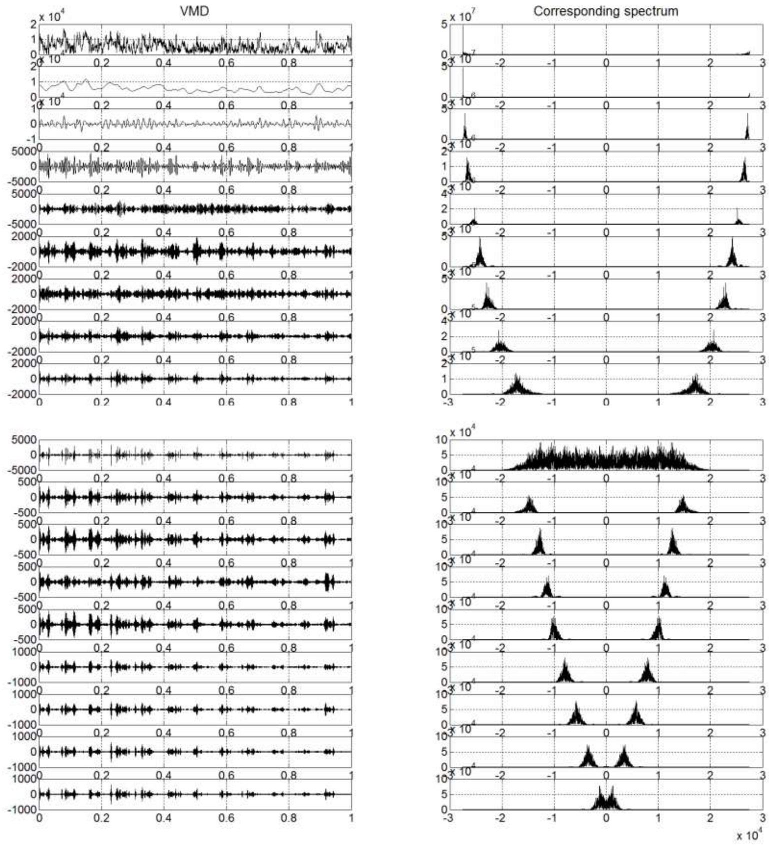

Fig. 3. VMD decomposition results and spectrum of low and high frequency components

3.3.3 WD-VMD-LSSVM (data fusion) combined prediction model

Taking the 16 modal components in Figure 3 as outputs, and the wind power and weather data (wind direction and speed) in the past two hours as inputs, 16 LSSVM prediction models were established by Eqs. (8-14). The data of 2015 (8760h) was used for training, and the test data of January 1, 2016 (24h) was then input into the trained 16 prediction models to obtain the predicted values of 16 modal components. Finally, an LSSVM model for data fusion of 16 modal components was performed to obtain the predicted values of wind power after signal reconstruction.

In order to verify the feasibility of the WD-VMD-LSSVM (data fusion) combined prediction model, a comparison study was conducted with 9 benchmark models, which are shown in Table 2. The description of the parameters of the models and their values are shown in Table 3.

Table 2. Comparing experimental model with benchmark models

Serial number	Predictive model	Describe
1	ELM	Extreme Learning Machine
2	LSSVM	Least Squares Support Vector Machine
3	WD-VMD-ELM	Small wave decomposition + Variational Mode Decomposition +ELM
4	WD-VMD-LSSVM	WD+VMD+LSSVM(component prediction)
5	VMD-LSSVM-ARMA-BPNN	VMD+LSSVM+Auto Regression Moving Average+BP Nerual Network
6	VMD-LSSVM(Data Fusion)	VMD+ LSSVM(Component prediction and data fusion)
7	EMD-LSSVM (Data Fusion)	Empirical Mode Decomposition +LSSVM (Component prediction and data fusion)
8	EEMD-LSSVM (Data Fusion)	Ensemble empirical mode decomposition + LSSVM (Component prediction and data fusion)
9	CEEMDAN-LSSVM(Data Fusion)	Complete EEMD with adaptive noise +LSSVM (Component prediction and data fusion)
10	WD-VMD-LSSVM (Data Fusion)	WD+VMD+LSSVM(Component prediction and data fusion)

Table 3. Main model parameters

Main model	Main parameters	Parameter description
ARMA	$p = 5$	Autoregressive
	$q = 4$	Moving average
BPNN	$n_i = 22$	Input layer nodes
	$n_{hbp} = 10$	Hidden layer nodes
	$n_o = 1$	Output layer nodes
ELM	$n_i = 22$	Input layer nodes

	$n_h = 100$	Hidden layer nodes
	$n_o = 1$	Output layer nodes
LSSVM	$\gamma = 100$	Regularization parameter
	$\sigma = 20$	Kernel function parameters
WD	$n = 6$	Decomposition layers
VMD	$K = 16$	Decomposition number
	$a = 1000$	Penalty factor
EEMD CEEMDAN	$Ntsd = 0.2$	Signal to noise ratio
	$NR = 500$	Number of noise additions
	$MaxIter = 5000$	Maximum number of envelopes

The combined prediction model proposed in this paper uses WD once for wind power series, and its computational complexity is $O(N\log(N))$, and uses VMD respectively for low and high frequency components, and its computational complexity is $O(2*N\lg(N))$. Finally, sixteen LSSVM predictors and one LSSVM fusion are constructed, and their computational complexity is $O(17*N^3)$, so the computational complexity of the proposed model is $O(N\log(N)+2*N\lg(N)+17*N^3)$. Here N is the number of training samples, namely 8760. However, the higher the computational complexity of the algorithm, the more time will be spent on model training and prediction, which can not quickly verify the idea and improve the model, and can not achieve rapid prediction. Therefore, this paper makes a comparative analysis of the time spent on experimental model training and prediction. In addition, Eq. (21) and Eq. (22) were used to evaluate the prediction results of the models, as shown in Table 4.

Table 4. Comparison of prediction results

Predictive model	MAPE（%）	RMSE（KWh）	Time/s
ELM	8.07	623.77	1.62
LSSVM	7.32	612.63	19.52
WD-VMD-ELM	9.76	689.74	138.95

WD-VMD-LSSVM	7.60	626.48	491.28
VMD-LSSVM-ARMA-BPNN	9.51	850.69	195.07
VMD-LSSVM(Data Fusion)	5.96	729.72	483.21
EMD-LSSVM(Data Fusion)	7.19	693.82	264.16
EEMD-LSSVM (Data Fusion)	5.72	531.80	656.43
CEEMDAN-LSSVM(Data Fusion)	5.44	550.66	934.63
WD-VMD-LSSVM (Data Fusion)	4.66	489.63	501.96

In order to analyze and compare the prediction performance of the 10 models, the results are visualized as shown in Figures 4 and 5 where the blue solid line is the regression line of the predicted result, the red dotted line is the ±5% error bar, and the black solid line is the “predicted value=actual value” baseline.

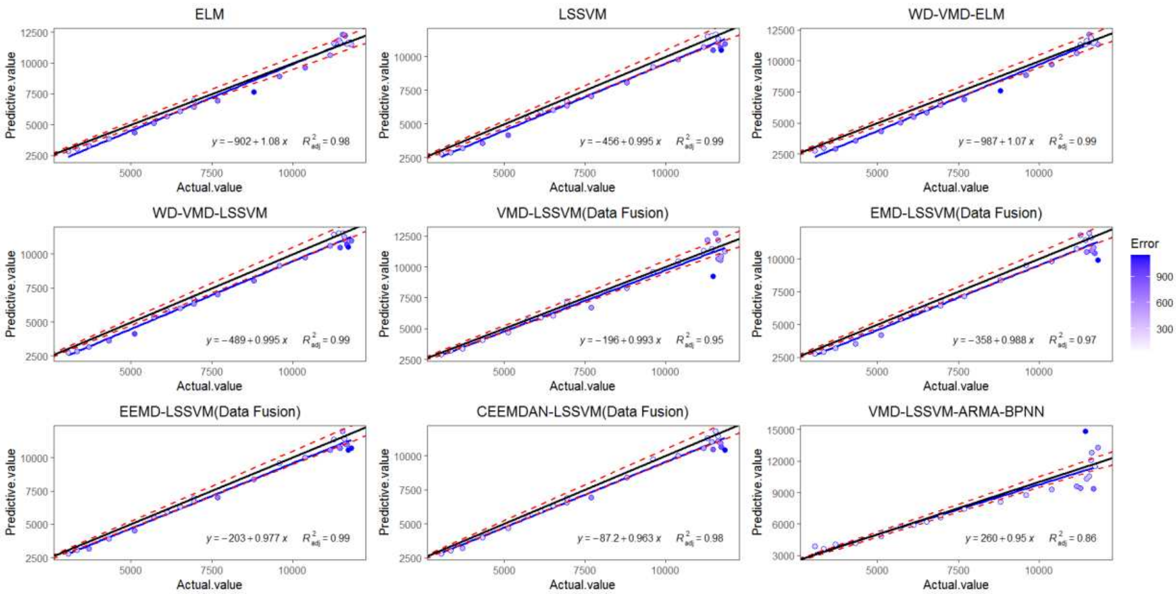

Fig. 4. Wind power prediction performance of benchmark models

It can be observed from Figure 4 and the left panels in Figure5 that the points predicted by the WD-VMD-LSSVM (data fusion) model are closer to the black "predicted value = actual value" baseline than the points predicted by the other 9 benchmark models. At the same time, its blue regression line is also nearly coincident with the black reference line, and the blue regression line is basically within the range of the ±5% error red line. It indicates that the predicted value is close to the actual value.

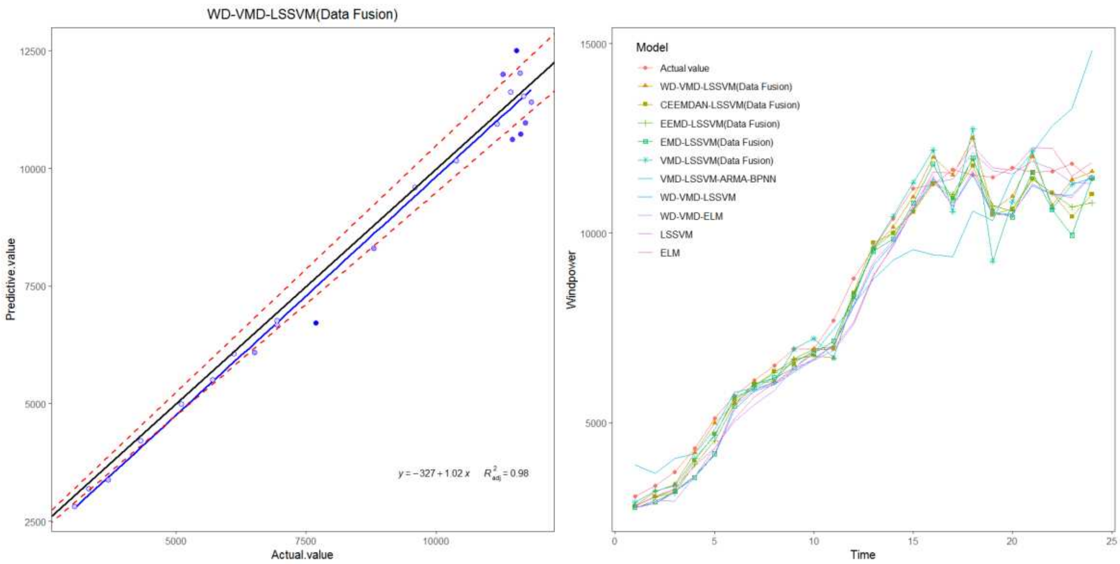

Fig. 5. Wind power prediction performance of different models

As can be seen from Table 4 and the right panels in Figure5, comparing Model 1 and Model 2, it can be found that LSSVM has better prediction effect than ELM, while ELM is faster than LSSVM. Compared with Models 1 and 2, Models 3 and 4 introduce a double decomposition method, but they cannot improve the prediction accuracy and improve the running time. The reason is that after the signal passes through WD and VMD, each modal component obtained has the superposition of two decomposition errors. Although the prediction sequence becomes stable, its information is lost, at the same time, errors caused by multiple LSSVM predictors are superimposed, which leads to the simple summation of sub-components and the failure to describe the wind power sequence well, thus eventually leading to the deterioration of prediction accuracy. Through the comparison of Model 2, Model 4, and Model 10, it can be observed that adding an LSSVM as a data fusion device can effectively eliminate the errors caused by the previous signal decomposition, thereby improving the accuracy of the prediction model. Comparing Models 6-10, it can be found that double decomposition can stabilize the sequence more thoroughly than other single decomposition techniques, and thus obtain better prediction accuracy. Meanwhile, compared with EEMD and CEEMDAN single decomposition technique, the double decomposition technique of WD-VMD has faster speed and better effect. The comparison between Model 5 and model 10 further shows that WD-VMD double decomposition has more advantages than single decomposition VMD and LSSVM has better effects than ARMA and BPNN as predictor and fusion. In summary, the wind power data step is 1h, and the slowest of the above models can be completed in 16min. The proposed model can be predicted in eight and a half minutes. For the processing power of current mainstream computers, the computational complexity of this algorithm is very low. Meanwhile the error indicators MAPE and RMSE of Model 10 are 4.66% and 489.63, respectively, which are the lowest among all models, demonstrating the effectiveness of the proposed method.

3. Conclusions

A new ultra-short-term wind power combined prediction model has been proposed by a combination of double decomposition and LSSVM. The experimental results show that:

1) The combination of WD and VMD decomposition techniques can well extract the effective information in the wind power time series. The two-level decompositions complement each other, resulting better stable modal components from the non-stationary wind power series, reducing the predictive difficulty.

2) By taking into consideration the historical wind power data and weather data, the proposed model improves the existing similar methods.

3) In order to eliminate the information loss caused by the signal decomposition and the errors caused by multiple LSSVM predictors, an LSSVM is introduced as a fusion device to reconstruct the modal components of the wind power sequence so as to further improve the prediction accuracy.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

Financial support was provided by the National Natural Science Foundation of China (61673166, 62033014), Natural Science Foundation of Hunan Province (2021JJ50006, 2022JJ50074).

DATA AVALIABILITY

The data that support the findings of this study are openly available in Kaggle at https://www.kaggle.com/datasets/nicholasjhana/energy-consumption-generation-prices-and-weather (Nicholas Jhana, 2019).

References

Abedinia O, Lotfi M, Bagheri M, et al. (2020) Improved EMD-based complex prediction model for wind power forecasting. *IEEE Transactions on Sustainable Energy*, 11(4): 2790-2802.

Adly AR, Aleem SHEA, Elsadd MA et al. (2020) Wavelet packet transform applied to a series-compensated line: a novel scheme for fault identification. *Measurement*, 151: 107156.

Alexiadis MC, Dokopoulos PS, Sahsamanoglou HS (1999) Wind speed and power forecasting based on spatial correlation models. *IEEE Transactions on Energy Conversion*, 14(3): 836-842.

Azimi R, Ghofrani M, Ghayekhloo M (2016) A hybrid wind power forecasting model based on data mining and wavelets analysis. *Energy Conversion and Managemen*t, 127: 208-225.

De Giorgi M G, Campilongo S, Ficarella A, et al. (2014) Comparison between wind power prediction models based on wavelet decomposition with least-squares support vector machine (LS-SVM) and artificial neural network (ANN). *Energies*, 7(8): 5251-5272.

Dragomiretskiy K, Zosso D (2013) Variational mode decomposition. *IEEE Transactions on Signal Processing*, 62(3): 531-544.

Fu W, Fang P, Wang K, et al. (2021) Multi-step ahead short-term wind speed forecasting approach coupling variational mode decomposition, improved beetle antennae search algorithm-based synchronous optimization and Volterra series model. *Renewable Energy*, 179: 1122-1139.

Gürsoy Ö, Engin SN (2019) A wavelet neural network approach to predict daily river discharge using meteorological data. *Measurement and Control*, 52(5-6): 599-607.

Khosravi A, Koury R N N, Machado L, et al. (2018) Prediction of wind speed and wind direction using artificial neural network, support vector regression and adaptive neuro-fuzzy inference system. *Sustainable Energy Technologies and Assessments*, 25: 146-160.

Korprasertsak N, Leephakpreeda T (2016) Robust short-term prediction of wind power generation under uncertainty via statistical interpretation of multiple forecasting models. *Energy*, 180: 387-397.

Li J, Li M (2019) Prediction of ultra-short-term wind power based on BBO-KELM method. *Journal of Renewable and Sustainable Energy*, 11(5): 056104.

Li LL, Zhao X, Tseng ML, et al. (2020) Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm. *Journal of Cleaner Production*, 242: 118447.

Liu K, Zhang Y, Qin L (2016) A novel combined forecasting model for short-term wind power based on ensemble empirical mode decomposition and optimal virtual prediction. *Journal of Renewable and Sustainable Energy*, 8(1): 013104.

Liu Z, Hara R, Kita H (2021) Hybrid forecasting system based on data area division and deep learning neural network for short-term wind speed forecasting. *Energy Conversion and Management*, 238: 114136.

Lu J, Qu X, Wang D, et al. (2021) Signal filtering method of variational mode decomposition and Euclidean distance based on optimizing parameters of classification particle swarm optimization algorithm. *Transactions of the Institute of Measurement and Control*, 43(9): 2018-2029.

Mason K, Duggan J, Howley E (2018) Forecasting energy demand, wind generation and carbon dioxide emissions in Ireland using evolutionary neural networks. *Energy*, 2155: 705-720.

Mohandes M A, Halawani T O, Rehman S, et al. (2004) Support vector machines for wind speed prediction. *Renewable Energy*, 29(6): 939-947.

Nicholas Jhana (2019). “Electrical demand, generation by type, prices and weather in Spain,” Kaggle, V.1.0, Hourly energy demand generation and weather, Dataset https://www.kaggle.com/datasets/nicholasjhana/energy-consumption-generation-prices-and-weather

Sun S, Fu J, Zhu F, et al. (2020) A hybrid structure of an extreme learning machine combined with feature selection, signal decomposition and parameter optimization for short-term wind speed forecasting. *Transactions of the Institute of Measurement and Control*, 42(1): 3-21.

Sun Z, Zhao M (2020) Short-term wind power forecasting based on VMD decomposition, ConvLSTM networks and error analysis. *IEEE Access*, 8: 134422-134434.

Wang H, Lei Z, Liu Y, et al. (2019) Echo state network based ensemble approach for wind power forecasting. *Energy Conversion and Management*, 201: 112188.

Wang J, Wang Y, Li Z, et al. (2020) A combined framework based on data preprocessing, neural networks and multi-tracker optimizer for wind speed prediction. *Sustainable Energy Technologies and Assessments*, 40: 100757.

Wu D, Gao C (2018) Short-term wind power generation forecasting based on the SVM-GM approach. *Electric Power Components and Systems*, 46(11-12): 1250-1264.

Wu Z, Xia X, Xiao L, et al. (2020) Combined model with secondary decomposition-model selection and sample selection for multi-step wind power forecasting. *Applied Energy*, 261: 114345.

Xie L, Wang B, Bao H, et al. Maimaitireyimu• Abulizi Ultra-short-term wind power prediction based on EEMD-WOA-LSSVM. Acta Solar Energy, 2021, 42(7): 290-296.

Xinghua X, Tinglong P, Dinghui W (2019) Wind Speed Noise Reduction in Wind Farm Based on Variational Mode Decomposition. *Journal of System Simulation*, 31(1): 87-93 (In Chinese)

Xu K, Yu J, Hu Y, et al. (2019) Security monitoring data fusion method based on ARIMA and LS-SVM. *IOP Conference Series: Earth and Environmental Science* (EES), Hangzhou, China, 2019.

Xu Y, Chen N, Shen X, et al. (2021) Proposal and experimental case study on building ventilating fan fault diagnosis based on cuckoo search algorithm optimized extreme learning machine. *Sustainable Energy Technologies and Assessments*, 45: 100975.

Zhang G, Liu H, Zhang J, et al. (2019) Wind power prediction based on variational mode decomposition multi-frequency combinations. *Journal of Modern Power Systems and Clean Energy*, 7(2): 281-288.

Zhang G, Wu Y, Wong KP, et al. (2014) An advanced approach for construction of optimal wind power prediction intervals. *IEEE Transactions on Power Systems*, 30(5): 2706-2715.

Zhang L, Wang B, Fang B, et al. (2018) Two-stage short-term wind speed prediction based on LSTM-LSSVM-CFA. *2nd IEEE conference on energy internet and energy system integration (EI2)*. IEEE, Beijing, China, 2018.

Zhao J, Guo ZH, Su ZY, et al. (2016) An improved multi-step forecasting model based on WRF ensembles and creative fuzzy systems for wind speed. *Applied Energy*, 162: 808-826.

Zhongda T (2020) Kernel principal component analysis-based least squares support vector machine optimized by improved grey wolf optimization algorithm and application in dynamic liquid level forecasting of beam pump. *Transactions of the Institute of Measurement and Control*, 42(6): 1135-1150..

Zhou Y, Wang J, Lu H, et al. (2022) Short-term wind power prediction optimized by multi-objective dragonfly algorithm based on variational mode decomposition. *Chaos, Solitons & Fractals*, 157: 111982.