Optimal Component IGSCV-SVR Ensemble Model Improved by VMD for Ultra-short-term Wind Speed Forecasting

Yu Ye, Jinxing Che, Heping Wang

Abstract—**The chaotic nature of wind speed will damage power system seriously, and cause economic losses. Therefore, timely wind prediction is crucial for the safety of power system. However, the traditional prediction method is hard to fully learn the characteristic of wind speed. This paper proposes an optimal component IGSCV-SVR ensemble model to predict ultra-short-term wind speed. It changes the traditional single parameter optimization method of time series prediction. Firstly, the VMD based component correlation is applied to decomposing the original wind speed dataset to obtain multiple subsequences. Our model can find the dissimilarity of each subsequence, and then the model fully learns the feature of each subsequence. It can help improve the overall efficiency of ultra-short-term wind speed prediction accuracy. Finally, estimates are obtained by summing the prediction of all components. The case study proves the feasibility of our method through the comparative experiments with some previous prediction models in MSE, MAE, MAPE and running time in the experimental part of this paper.**

Index Terms—**ultra-short-term wind speed prediction, Decomposition, Improved grid search cross-validation, Support vector regression, component correlation**

I. Introduction

TODAY, the shortage of fossil fuels and environmental problems are on a mounting crisis. To reduce the consumption of fossil fuels, there is an upsurge of advocating clean energy around the world. Such as wind energy. Wind power generation is developing rapidly in the world [1-2]. According to the website of Forbes magazine in 2021, China's offshore wind power installed capacity exceeds that of any other country in five years. The data of China's national energy administration shows, China's offshore wind power installed capacity was nearly 17,000,000 kW in 2021, which expresses that China has nearly half of the world's offshore wind power installed capacity currently. We can see that wind energy will be considerable in the future.

Manuscript received December 14, 2021; revised August 2, 2022. The research is supported by the Jiangxi Provincial Education Department (Program No. GJJ190961), the National Natural Science Foundation of China (Grant No. 71971105 and 12161058), the National Statistical Science Research Project of China (Grant No. 2020LZ03), and the Jiangxi Provincial Natural Science Foundation (Program No. 20212BAB201020).

Yu Ye is a postgraduate student in School of Information Engineering of Nanchang Institute of Technology, Nanchang 330099, Jiangxi, China. E-mail address: (e-mail: 2980618232@qq.com).

Jinxing Che is an associate professor in the School of Science, Nanchang Institute of Technology, Nanchang 330099, Jiangxi, China (corresponding author, e-mail: jinxingche1@163.com).

Heping Wang is a lecturer in the School of Science, Nanchang Institute of Technology, Nanchang 330099, Jiangxi, China (e-mail: hpwang@nit.edu.cn).

In pace with the expansion of wind power grid connection scale, the defects of wind power generation are sticking out.

A. Brief Introduction of Prediction Model of Wind Speed

The nonlinear characteristic of ultra-short-term wind speed has a great destructive effect on the security of the power system [3-5]. The quality of wind energy directly determines the power quality of wind power generation, which leads to the research on wind power has been carried out all over the world. Some scholars have implemented the machine learning algorithm to predict wind speed [6-7]. Generally, wind speed prediction can be divided into three categories: physical model, statistical model [8] and intelligent model [9]. The physical method uses the meteorological elements such as wind speed and wind direction to predict [10], yet the statistical law realizes it through the statistical analysis of the historical data. The statistical method regards the wind speed to be predicted as a time series and predicts it by discovering the potential law of the time series. Owing to the low accuracy of the statistical prediction model, we adopt the intelligent model normally. SVR, ANN, ELM are the common intelligent models.

Although the above model is easy to implement, the single model shows low accuracy in prediction universally [11]. Some models depend on the selection of parameters, such as SVR. For SVR, different parameter selection may lead to great differences and affect the training results of the model. In view of this, researchers begin to study the mixed model [12-13]. In a nutshell, there are two methods in raising a single traditional prediction accuracy: processing the original data or applying algorithm to optimizing the parameters of prediction model. The first method is used for signal denoising through the decomposition and reconstruction of the signal. EMD and wavelet decomposition methods are commonly used in signal processing [14-18].

The decomposition method mentioned above can be able to improve the prediction accuracy to some extent. However, EMD is easy to produce mode aliasing, while WPD is greatly affected by the threshold and has weak applicability. In order to solve the above aporias, VMD is proposed, which is a non-recursive decomposition method in raw signal decomposition [19-21].

The second method consists of a variety of techniques, which can be used to optimize the parameters of a single prediction model, and the weight parameters of a combination forecasting model. Due to its strong adaptability, this kind of method has been widely concerned by scholars at home and abroad. Common optimization algorithms include PSO algorithm, GA, CS, ABC algorithm and so on [22-27].

Most of the above literatures only use one of the two methods, which improves the prediction results of the model to a certain extent. On this basis, some scholars combine two methods to improve the fitting effect of the model [28-29]. Above all, we propose the optimal on the hybrid optimization algorithm IGSCV and the component correlation based VMD. The optimal component ensemble model aims at balancing the information and variance of the multiple subsequences. What is more, a novel hybrid optimization algorithm is proposed for fully learn the characteristic of wind speed.

B. The Main Arrangements for This Article

On the limitations of some current models, we propose an optimal component IGSCV-SVR ensemble model based VMD to forecast ultra-short-term wind speed. To verify the reliability of the model, three groups of comparative experiments are carried out in our experiment. There are comparisons among undecomposed models, comparisons among decomposed models, and comparisons between decomposed models and undecomposed models. In general, the main content of this paper is summarized as six parts:
The first part introduces the related research on wind speed prediction. The second part introduces the basic principles of some models and algorithms involved in this paper. VMD and IGSCV are introduced in detail. The third part is our experiment, which briefly introduces the content of our experiment. The fourth part summarizes the experimental results. The part of reference describes literature reference, lists all the literature cited in this paper, and we introduce all authors' information at the end of the paper.

II. THE THEORETICAL BASIS OF THE MODEL

A. Variational Modal Decomposition (VMD)

VMD differs from recursive decomposition mode, pertains to a kind of variational mode decomposition problems. The key to realize VMD is to determine the modal decomposition times *K* [20], which is introduced in four parts below:

a. Model bandwidth evaluation

Firstly, the basic data can be decomposed into *k* parts. we call each component $s_k(t)$. We can obtain the unilateral frequency spectrum through convoluted $s_k(t)$ with the Hilbert transform.

$$\left[\delta(t)+\frac{j}{\pi t}\right]*s_k(t) \tag{1}$$

Then, we can transfer the spectrum to its baseband, in this process, we require estimating the center frequency $e^{-jw_k t}$ of each mode firstly. The operation formula is as in (2).

$$\left\{\left[\delta(t)+\frac{j}{\pi t}\right]*s_k(t)\right\}\cdot e^{-jw_k t} \tag{2}$$

Finally, (3) is obtained by Gaussian smoothing from (2). We can also get the model bandwidth from (3). Under the constraints of $\sum_k s_k = S(t)$, $S(t)$ is the original signal, the model bandwidth takes the minimum value of the sum of the estimated bandwidth of each mode.

$$\begin{cases} \min\left\{\sum_{k=1}^{k}\left\|\ell_t\left\{\left[\delta(t)+\frac{j}{\pi t}\right]*s_k(t)\right\}\cdot e^{-jw_k t}\right\|_2^2\right\} \\ s.t.\sum_k s_k = S(t) \end{cases} \tag{3}$$

In the equation, $\delta(t)$ represents Dirac distribution. $S(t)$ represents the raw signal. $\left[\delta(t)+\frac{j}{\pi t}\right]$ represents Hilbert transform function. s_k, w_k represent the assemble of modes and the center frequency.

b. Solutions for Variational Problem

Generally, compared with solving constrained variational problems directly, we are more used to transforming it into unconstrained variational problems by introducing Lagrange penalty factor ρ and multiplication factor η, as shown in (4).

$$L(\{s_k\}\{w_k\},\eta)=\rho\sum_{k=1}^{k}\left\|\ell_t\left\{\left[\delta(t)+\frac{j}{\pi t}\right]*s_k(t)\right\}\cdot e^{-jw_k t}\right\|_2^2 + \left\|f(t)-\sum_k s_k(t)\right\|_2^2+\left\langle \eta(t) f(t)-\sum_k s_k(t)\right\rangle \tag{4}$$

In which, ρ is used to reduce Gaussian signal interference. The above problem can be solved by Alternate Direction Method of Multipliers. We can minimize the extended Lagrangian expression by iteratively updating the s_k, w_k, λ. The iterative update formulas of s_k, w_k, λ are as in (5-7):

$$\hat{s}_k^{n+1}(w)=\frac{\hat{f}(w)-\sum_{i<k}\hat{s}_k^n(w)-\sum_{i>k}\hat{s}_i^n(w)+\frac{\hat{\lambda}^n(w)}{2}}{1+2\alpha\left(w-w_k^n\right)^2} \tag{5}$$

$$w_k^{n+1}=\frac{\int_o^\infty w\left|s_k^{n+1}(w)\right|^2 d_w}{\int_o^\infty \left|s_k^{n+1}(w)\right|^2 d_w} \tag{6}$$

$$\hat{\lambda}^{n+1}(w)=\hat{\lambda}^n(w)+\varsigma\left(\hat{s}(w)-\sum_{k=1}^{k}\hat{s}_k^{n+1}(w)\right) \tag{7}$$

In the above formula, ς is noise tolerance, meeting the signal fidelity requirements. $\hat{s}_k^{n+1}(w)$, $\hat{f}(w)$, $\hat{s}_i^n(w)$ and $\hat{\lambda}^n(w)$ represent Fourier transform of $s_k^{n+1}(w)$, $f(w)$, $s_i^n(w)$, $\lambda^n(w)$ respectively.

c. The process of VMD

The result is output at the end of the whole iterative period, and *K* narrowband IMFs components can be obtained.

d. Determination of modal decomposition times K

The number of modal components in this paper is acquired based component correlation, which is also known as Pearson correlation [30], the judgment principle as follows:

There are two modal components X and Y, and the component correlation of X and Y can be defined in (8).

$$r_{X_iY_i}=\frac{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)\left(Y_i-\bar{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}\sqrt{\sum_{i=1}^{n}\left(Y_i-\bar{Y}\right)^2}} \tag{8}$$

In the above formula, X_i and Y_i represent two modal components. $\overline{X}_i$ and $\overline{Y}_i$ represent the mean value of two modal components respectively, $r_{X_iY_i}$ is component correlation. The calculation results are in Table I. It shows the selection of modal decomposition number K by component correlation. It can be seen from Table I that there is a very weak correlation between any two modal components when $K = 1, 2, 3, 4, 5$. However, $K = 6$ shows weak correlation. Thus, we come to the conclusion that $K = 5$ is the best modal decomposition number of this experiment [31].

TABLE I
NUMBER K BASED COMPONENT CORRELATION

K	C_{12}	C_{23}	C_{34}	C_{45}	C_{56}	C_{67}
2	0.0872					
3	0.1274	0.0950				
4	0.1466	0.0920	0.1123			
5	0.1681	0.1073	0.1117	0.0979		
6	0.2170	0.1217	0.1137	0.1127	0.0983	
7	0.2359	0.1205	0.1148	0.1159	0.1186	0.1112

B. Support Vector Regression (SVR)

Previously, we perform VMD on the original data to obtain five modal components. For each component, we use SVR to predict separately. SVR is a derivative of SVM, using the decision boundary of the optimal hyperplane in support vector classification to solve the regression problems [32]. It converts the raw space into a higher dimensional feature space to solve complex nonlinear problems [33]. Take a single input-output system as an example.

Given a set of data with $\{(x_1,y_1),(x_2,y_2),...,(x_n,y_n)\}$, x_i is the sample input, y_i is the sample output, n is the number of data points, and $\varphi(x_i)$ denotes high dimensional eigenvector corresponding to x_i. Its optimal hyperplane formula as in (9).

$$\gamma(x_i) = \alpha^T \varphi(x_i) + \beta \tag{9}$$

where α is a weight vector and β is the bias. The essence of the training process of the SVR model is to find the optimal α and β, and make the $\gamma(x_i)$ approach to y_i steadily. Differs from the general regression problem, SVR allows a tolerance deviation between the model output and the real value. We introduce relaxation variable as given in (10).

$$\min_{\alpha,\beta} \frac{1}{2}\|\alpha\|^2 + C\sum_{i=1}^{m}(\xi_i + \xi_i^*) \tag{10}$$

$C > 0$ is the penalty or regularization factor, used to balance the interrelation between $\frac{1}{2}\|\alpha\|^2$ and $\sum_{i=1}^{m}(\xi_i + \xi_i^*)$. ξ_i *and* ξ_i^* represent the relaxation variables. Furthermore, the constraint conditions are defined as in (11).

$$s.t\begin{cases} \gamma(x_i) - y_i \le \varepsilon + \xi_i \\ y_i - \gamma(x_i) \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0, i = 1,..,m \\ C > 0 \end{cases} \tag{11}$$

Where ε is insensitive risk function. Then, Lagrange multiplier a_i, a_i^* are introduced to handle the convex optimization function with constraints. Meanwhile, α and regression function $\gamma(x)$ can be expressed in (12). Kernel function $\kappa(x, x_i)$ can better solve the nonlinear problem.

$$\alpha = \sum_{i=1}^{m}(a_i^* - a_i)\varphi(x_i)$$

$$\begin{cases} \gamma(x) = \sum_{i=1}^{m}(a_i^* - a_i)\kappa(x, x_i) + \beta \\ \kappa(x, x_i) = exp\{-g \,||\, x - x_i \,||^2\} \end{cases} \tag{12}$$

“Radial Basis Function (RBF)” is generally used as kernel function in SVR model. This paper also uses “RBF”.

When modeling, we ought to determine the penalty parameter C and kernel parameter g. They can balance the model complexity. ε is tolerance error, indicating the approximation degree of training data points. g is the width parameter of kernel function, controlling the local information of data.

C. Improved Particle Swarm Optimization (IPSO)

PSO is based search optimization technique [34]. It updates each candidate solution by velocity and position of particle. It both has demonstrated good convergence and better performance [35]. The update formula is in (13):

$$\begin{cases} V_{i(j+1)} = w \cdot v_{ij} + c_1 r_1 (p_{ij} - x_{ij}) + c_2 r_2 (p_{gj} - x_{ij}) \\ X_{i(j+1)} = x_{ij} + v_{ij} \end{cases} \tag{13}$$

v_{ij} represents the velocity of the particles, $1 \le i \le n$, $1 \le j, (j+1) \le m$. w is the inertia weight, used to maintain the inertia of particle motion, dynamically search for local and global optimal solutions. c_1, c_2 is the acceleration constant. $r_1, r_2 \in (0,1)$. The limit values of X_i and velocity V_i are as in (14).

$$\begin{aligned} -X_{max} \le X_i \le X_{max} \\ -V_{max} \le V_i \le V_{max} \end{aligned} \tag{14}$$

$-X_{max}, X_{max}, -V_{max}, V_{max}$ in (14) are the minimum and maximum values of position and velocity respectively.

For improving the convergence speed and global search ability of the algorithm, this section improves w, $c1$ and $c2$ of the traditional PSO algorithm. The updated weights and learning factors are as in (15-16).

$$w = w_1 - (w_1 - w_2) * (t / \max\_itera) \tag{15}$$

$$\begin{aligned} c_1 = c_1\_in + (c_1\_f - c_1\_in) * t / \max\_itera \\ c_2 = c_2\_in + (c_2\_f - c_2\_in) * t / \max\_itera \end{aligned} \tag{16}$$

where $w_1 = 0.9$, $w_2 = 0.3$, $c_1\_in = c_2\_in = 0.1$, $c_1\_f = c_2\_f = 2$. t represents current iterations, *max_itera* represents total iterations.

D. Improved Whale Optimization (IWOA)

WOA is a new heuristic algorithm proposed in 2016. It has good global search ability and local search ability. See the reference for the detailed process [33]. To improve the optimization effect of the algorithm, we adopt the opposite number inverse approximation method for initialization in the initialization stage, and the solutions can be evenly distributed in the range [37]. The mathematical expression is in (17).

$$s_i^o = rand\left(l_i, s_i\right) \tag{17}$$

where $s_i \in \left[a_i, b_i\right]$, $l_i = \left(a_i, b_i\right)/2$, $rand(\bullet)$ is uniformly distributed random number.

E. K-Fold Cross Validation (CV)

CV is one of the common optimization algorithms, which can reduce the phenomenon of under learning and over learning. Generally speaking, the amount of training data is inversely proportional to K. The general process of K-fold cross validation is as follows:

Firstly, divide the training data into K equal parts of the same size, taking one of them as validation data, and the remaining as training data. K is the number of iterations. The performance index of regression is the mean value of K operation results. when K= 3, the scenario looks like this:

K=3 is a typical choice in CV, which is also called as leave-one-out CV. Let we divide dataset into K copies with similar size. Let $\chi:\{1,2,3,....,N\} \rightarrow \{1,2,3,..,K\}$ be an indexing function, and $\hat{f}^{-\chi(i)}\left(x_i, \alpha\right)$ denote αth model fit function with the i-th part of the dataset removed. We can define the CV estimate of prediction error from (18):

$$CV\left(\hat{f}, \alpha\right) = \frac{1}{N}\sum_{1}^{N} L\left(y_i, \hat{f}^{-\chi(i)}\left(x_i, \alpha\right)\right) \tag{18}$$

The key for us is finding the tuning parameter $\hat{\alpha}$ to minimize (18).

F. Grid Search (GS)

GS is a method of specifying parameter values, combined with CV generally. It divides the grid into a particular range and traverses all points in the network with the values of parameters. The values of each parameter are accepted, and each of possible combined values is called as a "grid" in grid. Then the funded optimal value is employed for SVR training.

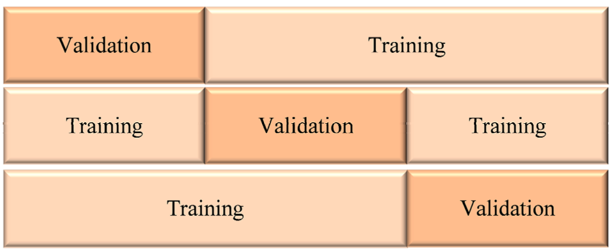

Fig. 1. three-fold cross validation

G. Improved Grid Search Cross Validation (IGSCV)

GS essentially divides the parameters to be searched into grid by fixed intervals in a specific space, and each grid point represents a set of parameter solutions. When the step size of GS is small, the search accuracy is high, but the process is quite long. Therefore, GS algorithm occupies a disadvantage in large-scale optimization. For this, we propose IGS algorithm embedded by IWOA.

Firstly, we use GS find the optimal solution with large step size. Then, carry out GS with IWOA when it is closed to the optimal solution, we name it as IGSCV. As shown in Figure 2, the red dashed box represents the better interval searched by GSCV in part (a), yellow dots indicate evenly distributed points in IWOA. The blue triangle indicates the local process point of the search, and the red pentagram indicates the best point of the search in part (b).

For the convenience of observation, we only draw the values of partial uniform distribution.

The specific flow chart of this paper is shown in Figure 3. Decompose the original wind speed into five sub parts by VMD. Raw data are shown in Figure 4 and the decomposed subsequences are shown in Figure 5, Figure 6, Figure 7, and Figure 8.

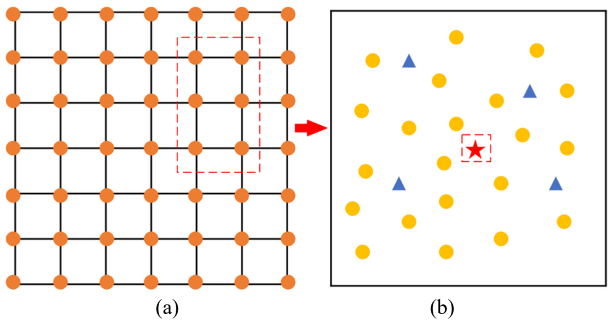

Fig. 2. GSCV-IWOA

Then feature transformation is carried out through the time series feature transformation matrix. We can get a simplified characteristic input matrix by Max-Relevance and Min-Redundancy (mRMR). A variety of models designed in this paper are used for modeling and prediction. There are some prediction models, such as SVR, LSTM, ELM, LSSVR, IPSO-SVR, GSCV-SVR, IGSCV-SVR etc.

III. Experiment

A. Ultra-Short-Term Wind Speed Prediction Based on VMD

This paper proposes a novel prediction model combined with VMD. Raw data show the large volatility and non-determinacy, which is shown in Figure 4.

We disintegrate the original dataset by VMD firstly, and get five components in each season, which is shown in Figure 5, Figure 6, Figure 7, and Figure 8. Then we use prediction models to predict each component partly.

In Figure 4, the green line represents spring data, the red line represents summer data, the yellow line represents autumn and the blue line represents winter.

B. Experimental Design

This section introduces the experimental part of this article. The data come from a wind farm in Penglai, Shandong

Province. Programming language is the python3.6, and the simulation software is Anaconda3.

Considering the dataset of the experiment is large, we divide the data into four seasons, and select one month of each season for experiment respectively. There are about 1000 data per season, and the proportion of training set and testing set is 3:1. Firstly, the missing values of the dataset are processed. The missing values were filled with the median of two adjacent points.

We created four time series characteristic matrices for model training. Each matrix dimension is $N \times M$. We use mRMR for feature selection, and select the top 20 of most relevant features as input features. A variety of optimization algorithms are used to optimize the parameters of SVR. Then mean square error (MSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) of the four months are calculated respectively.

To highlight the effectiveness of the proposed method, the parameters are set as follows: in GS search process, the step of C is 5 and the step of gamma is 1, and $C \in [1,505]$, $g \in [0,10]$. The step of GS in IGSCV and GSCV-IPSO are both 50. The feasibility of our proposed algorithm is further verified by controlling the external conditions.

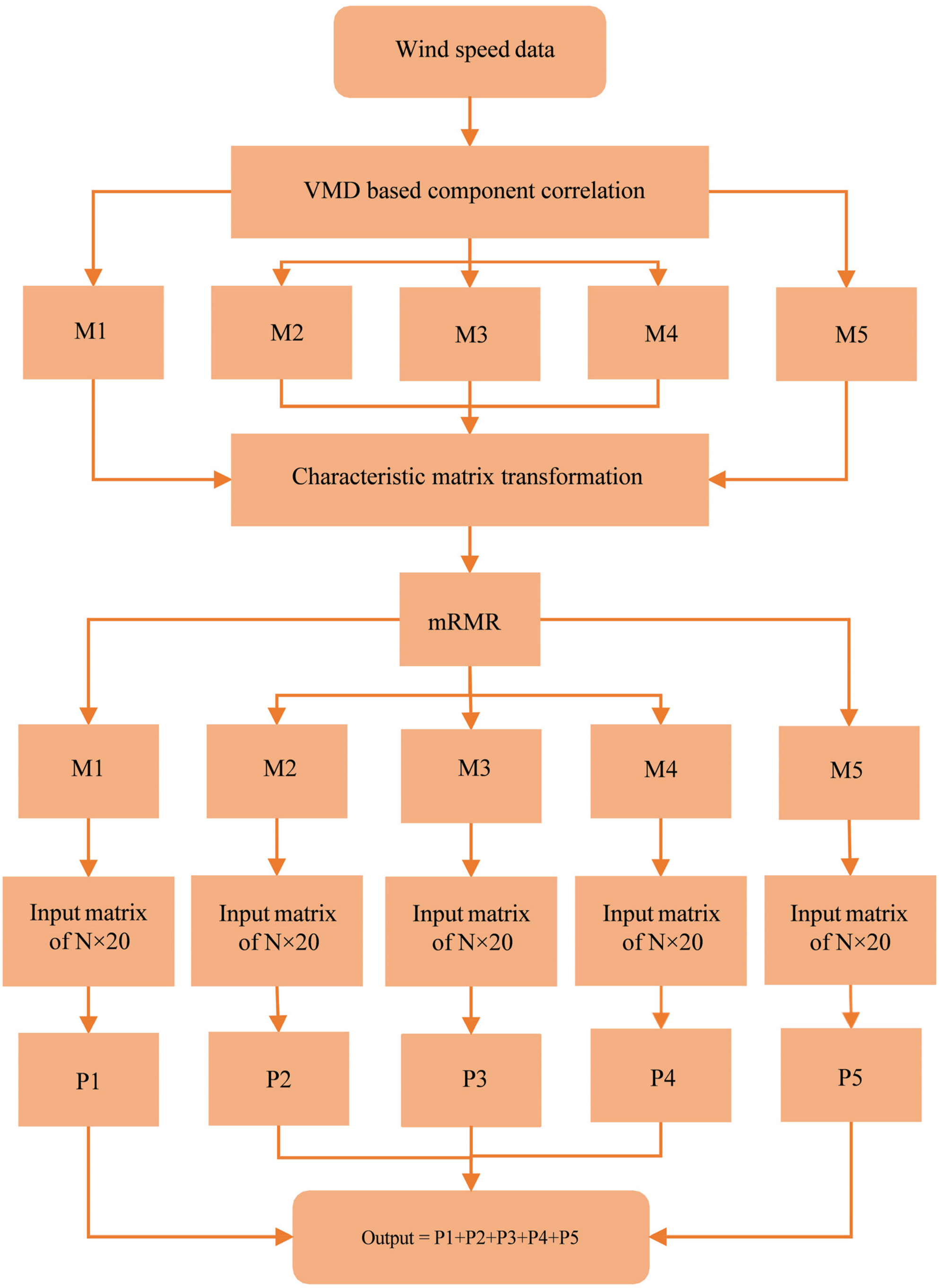

Fig. 3. Forecast flow chart

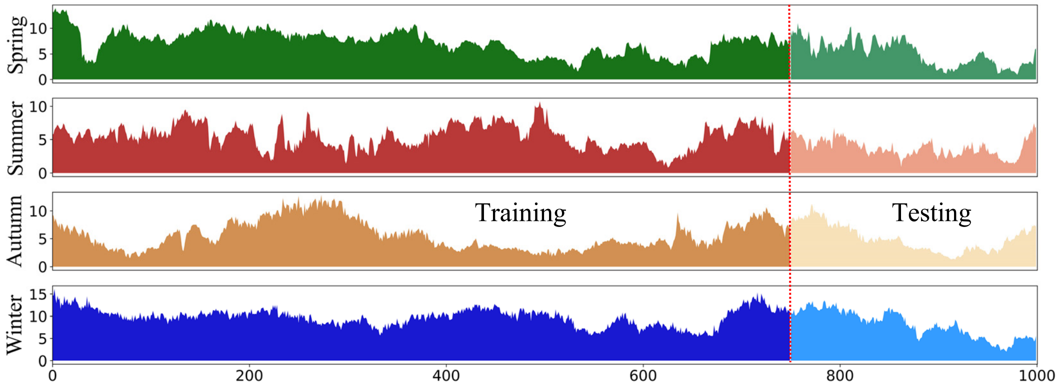

Fig. 4. Map of original wind speed data

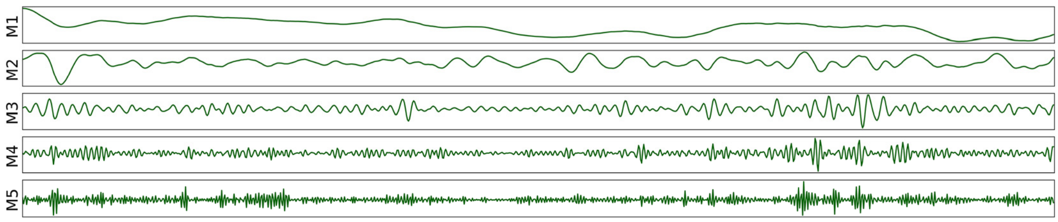

Fig. 5. VMD decomposition sequence diagram of original wind speed in Spring

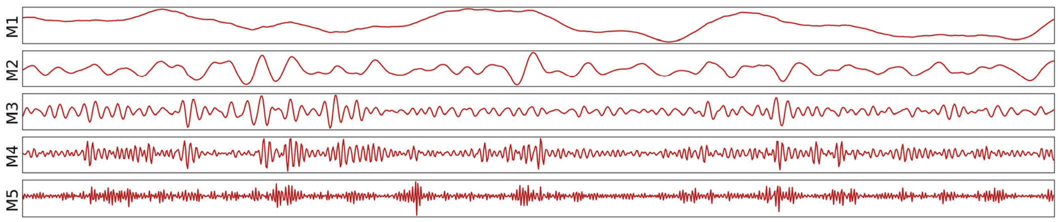

Fig. 6. VMD decomposition sequence diagram of original wind speed in Summer

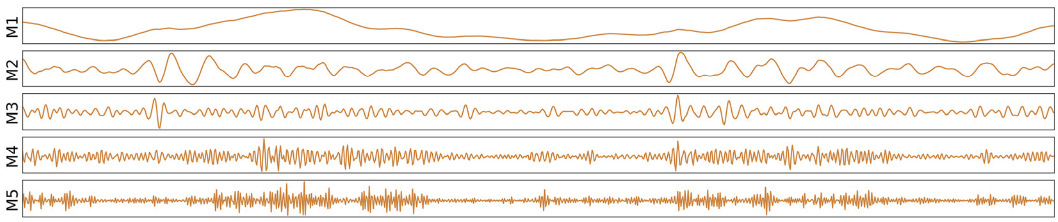

Fig. 7. VMD decomposition sequence diagram of original wind speed in Autumn

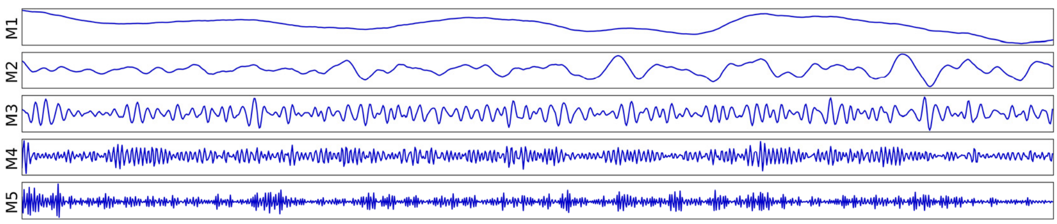

Fig. 8. VMD decomposition sequence diagram of original wind speed in Winter

IV. Analysis of Experimental Results

A. Description of Experimental Results

This article uses the common error measurement method like MSE, MAE, MAPE and calculate the running time. The experimental results will be given in Table II. We rank the errors of each quarter from small to large by MSE for facilitate observation.

It can be seen from Table II, VMD-IGSCV-SVR has the highest prediction accuracy in four seasons. The second is VMD-GSCV-IPSO-SVR.

The prediction accuracy of GSCV-SVR is higher than that of GSCV-IPSO-SVR. However, the prediction accuracy of VMD-GSCV-IPSO-SVR is higher than VMD-GSCV-SVR. The model proposed in this paper performs well both in decomposition and non-decomposition. At the same time, our model can achieve the effect of parameter optimization and the dynamic balance of step size by adjusting the step size in the early stage to improve the optimization efficiency.

Next, we make a detailed comparison of the prediction errors in four seasons. The prediction accuracy of our first ranked model (our model) is compared with that of the second ranked model in detail, MSE is reduced by about 76%, MAE is reduced by about 51%, for summer, MSE is reduced by about 77%, MAE is reduced by about 47%, for autumn,

TABLE II
COMPARISON TABLE OF MODEL TRAINING ERROR

	Test error	*MSE*	*MAE*	*MAPE*	*Runtime(/s)*	*Parameter C*	*Parameter g*
Spring	VMD-IGSCV-SVR	0.015361	0.088338	0.016680	77	\	\
	VMD-GSCV-IPSO-SVR	0.063590	0.181543	0.034161	20	\	\
	VMD-GSCV-SVR	0.126107	0.240937	0.046554	170	\	\
	IGSCV -SVR	0.225964	0.343136	0.062325	35	100.64023	0.000158
	VMD-ELM	0.330780	0.526029	0.319686	\	\	\
	GSCV-SVR	0.348863	0.429759	0.081857	31	13	0.1
	VMD-LSTM	0.427874	0.467325	0.087073	\	\	\
	ELM	0.465942	0.500802	0.313111	8	\	\
	VMD-IPSO-SVR	0.528769	0.494225	0.105795	2	\	\
	SVR	0.642475	0.585320	0.106551	0	\	\
	VMD-LSSVR	0.687422	0.669996	0.382713	\	\	\
	LSTM	0.761093	0.638784	0.116798	3	\	\
	LSSVR	0.954161	0.773398	0.406592	\	\	\
	GSCV-IPSO-SVR	2.507992	1.185972	0.268218	8	51.128782	0.512017
	IPSO-SVR	5.562264	1.983545	0.442213	1	251.088208	5.035342
Summer	VMD-IGSCV-SVR	0.022137	0.117363	0.027807	77	\	\
	VMD- GSCV-IPSO-SVR	0.098989	0.221926	0.052678	20	\	\
	VMD-GSCV-SVR	0.220001	0.315339	0.074821	2	\	\
	VMD-LSTM	0.306177	0.426575	0.096299	32	\	\
	IGSCV-SVR	0.378352	0.446412	0.103168	34	251.856389	0.001073
	VMD-LSSVR	0.467334	0.544415	0.364410	\	\	\
	VMD-ELM	0.532074	0.695463	0.400520	\	\	\
	GSCV-SVR	0.602732	0.588050	0.136953	8	5	0.1
	SVR	0.605088	0.595809	0.138516	792	\	\
	LSTM	0.668749	0.626650	0.151189	3s	\	\
	ELM	0.713201	0.647450	0.380716	\	\	\
	VMD-IPSO-SVR	0.735591	0.617533	0.150856	2	\	\
	LSSVR	0.755397	0.699235	0.409503	\	\	\
	GSCV-IPSO-SVR	2.210991	1.186094	0.285273	8	51.128782	0.512016
	IPSO-SVR	3.947300	1.684393	0.430644	1	286.080023	4.917442
Autumn	VMD-IGSCV-SVR	0.018084	0.099414	0.018581	72	\	\
	VMD-GSCV-IPSO-SVR	0.050020	0.162117	0.030042	19	\	\
	IGSCV-SVR	0.070011	0.383653	0.502677	164	264.260843	0.000633
	VMD-GSCV-SVR	0.074608	0.190437	0.035211	1	\	\
	VMD-LSTM	0.219033	0.361725	0.072838	31	\	\
	VMD-IPSO-SVR	0.316530	0.394118	0.073146	2	\	\
	VMD-ELM	0.364887	0.537344	0.354806	\	\	\
	SVR	0.381215	0.462334	0.094436	0	\	\
	GSCV-SVR	0.464245	0.487045	0.085156	8	5	0.1
	ELM	0.465683	0.516569	0.325013	\	\	\
	LSTM	0.472137	0.516734	0.107872	3	\	\
	VMD-LSSVR	0.585295	0.629840	0.380843	\	\	\
	LSSVR	0.747631	0.706959	0.400140	\	\	\
	GSCV-IPSO-SVR	3.118461	1.257256	0.222091	8	51.128782	0.512017
	IPSO-SVR	6.420130	2.128546	0.447715	1	246.004159	5.592757
Winter	VMD-IGSCV-SVR	0.017357	0.103964	0.012500	78	\	\
	VMD- GSCV-IPSO-SVR	0.050842	0.176892	0.021804	21	\	\
	VMD-GSCV-SVR	0.099323	0.237621	0.030412	1	\	\
	VMD-ELM	0.222765	0.430863	0.230331	\	\	\
	VMD-LSTM	0.336692	0.462563	0.053878	31	\	\
	IGSCV-SVR	0.414995	0.503628	0.058594	33	82.998774	0.0001
	ELM	0.562748	0.584176	0.262379	\	\	\
	SVR	0.589371	0.592745	0.070665	0	\	\
	VMD-LSSVR	0.610243	0.611581	0.297353	\	\	\
	LSTM	0.657169	0.646589	0.076337	8	\	\
	GSCV-SVR	0.797666	0.682068	0.080921	3	5	0.1
	VMD-IPSO-SVR	0.838433	0.594826	0.083120	2	\	\
	LSSVR	0.844353	0.709130	0.318508	\	\	\
	GSCV-IPSO-SVR	4.430394	1.721483	0.239507	8	51.128782	0.512016
	IPSO-SVR	5.680668	1.939931	0.281138	1	219.745971	5.567753

MSE is reduced by about 64%, MAE is reduced by about 39%, and for winter, MSE is reduced by about 66% and MAE is reduced by about 41%.

From the above, we can see that our model has been greatly improved both in MSE and MAE, which is very helpful for accurate prediction of wind speed. From the running time of the three models with high accuracy, we can find that the running time of VMD-GSCV-SVR is the longest, and VMD-IPSO-SVR is the shortest. The time of our model is medium among them, which is about three times more than that of VMD-IPSO-SVR, but this time is also acceptable.

To be more intuitive, we draw the histogram of MSE and MAE. When drawing MSE histogram, we omit the two models with the worst errors, as shown in Figure 10. From the perspective of MSE, our proposed model performs best in all models in this article. The same is true in Figure 11.

In a word, the prediction effect of IGSCV-SVR is better than GSCV-SVR, IPSO-SVR, and GSCV-IPSO-SVR. The fitting effect of our model on this dataset is the best, which can be seen in the Figure 10 and Figure 11. This also proves the accuracy and reliability of the model proposed in this paper.

B. Conclusions

The random fluctuation of ultra-short-term wind speed has inherent uncertainty, and the traditional SVR prediction model has the disadvantage of large prediction error. Experiments show that use intelligent optimization algorithm to optimize the parameters of SVR can reduce the prediction error to a certain extent. However, the existing WOA algorithms are not so perfect, and there are always some problems in the process of parameter optimization, such as WOA algorithm falls into local extremum and slow convergence easily, and the time of running GSCV is too long. In this paper, IWOA algorithm and GS algorithm are combined to optimize the super parameters of SVR, enhancing their strengths and avoiding their weaknesses. Thus, an improved optimization algorithm is proposed, the effectiveness of the algorithm is also proved by experiments. The experimental results show the prediction effect of the IWOA algorithm applied to parameter optimization is better in this experiment.

Taking summer as an example, when the search range is large, IPSO is easy to fall into local optimization, resulting in poor optimization results. The combination effect of GSCV and IPSO is affected by IPSO, the combined prediction accuracy is not high. The result of neural network model is better in the initial stage. However, if the optimization algorithm is used to optimize the neural layer, it will run for a very long time, exceeding our set time range. Therefore, we do not consider optimizing the neural network algorithm.

Secondly, by comparing the prediction accuracy between the optimization algorithm and the hybrid model of SVR, the advantages of IGSCV-SVR are highlighted whether in the decomposition process or not. Which can be prove from Figure 9. We find that VMD-IGSCV-SVR is closest to black true curve in four seasons, which is indicated by red dotted line. It shows that the actual fitting effect of our model is better. Above all, our model is feasible.

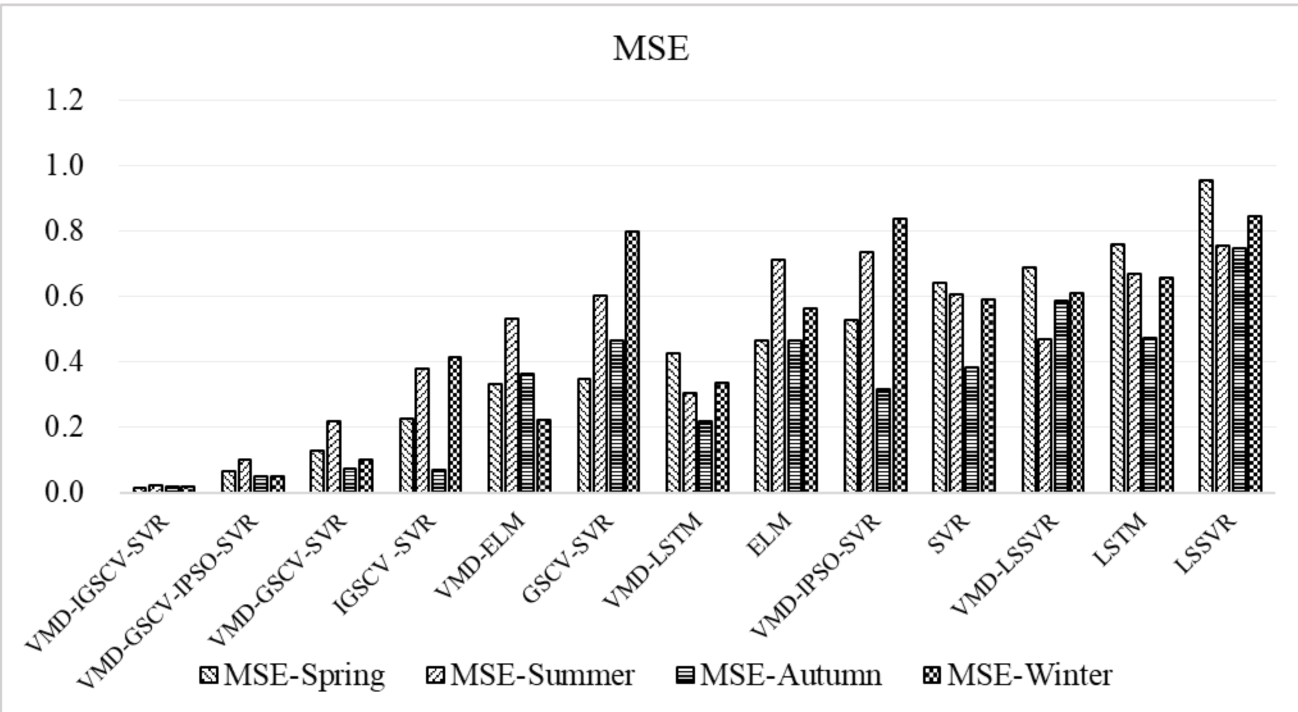

Fig. 10. MSE of all models in four seasons

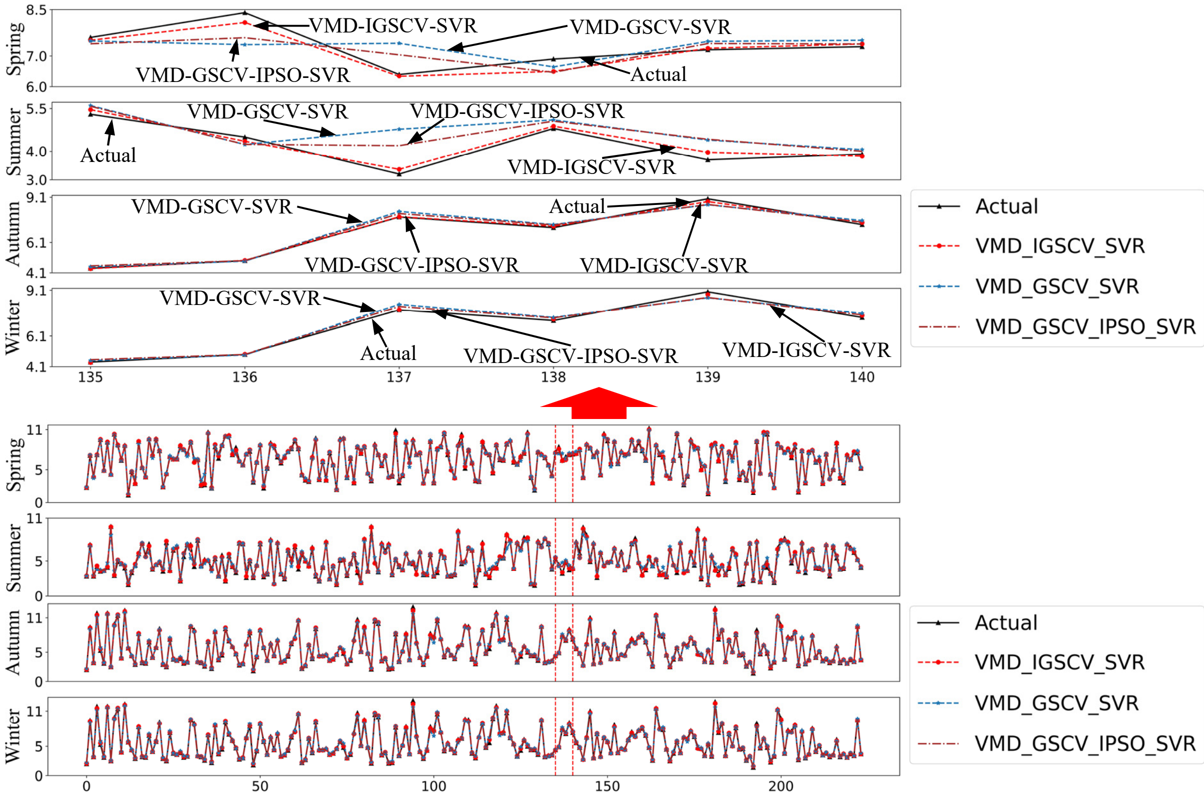

Fig. 9. Prediction curve of three top models in this paper

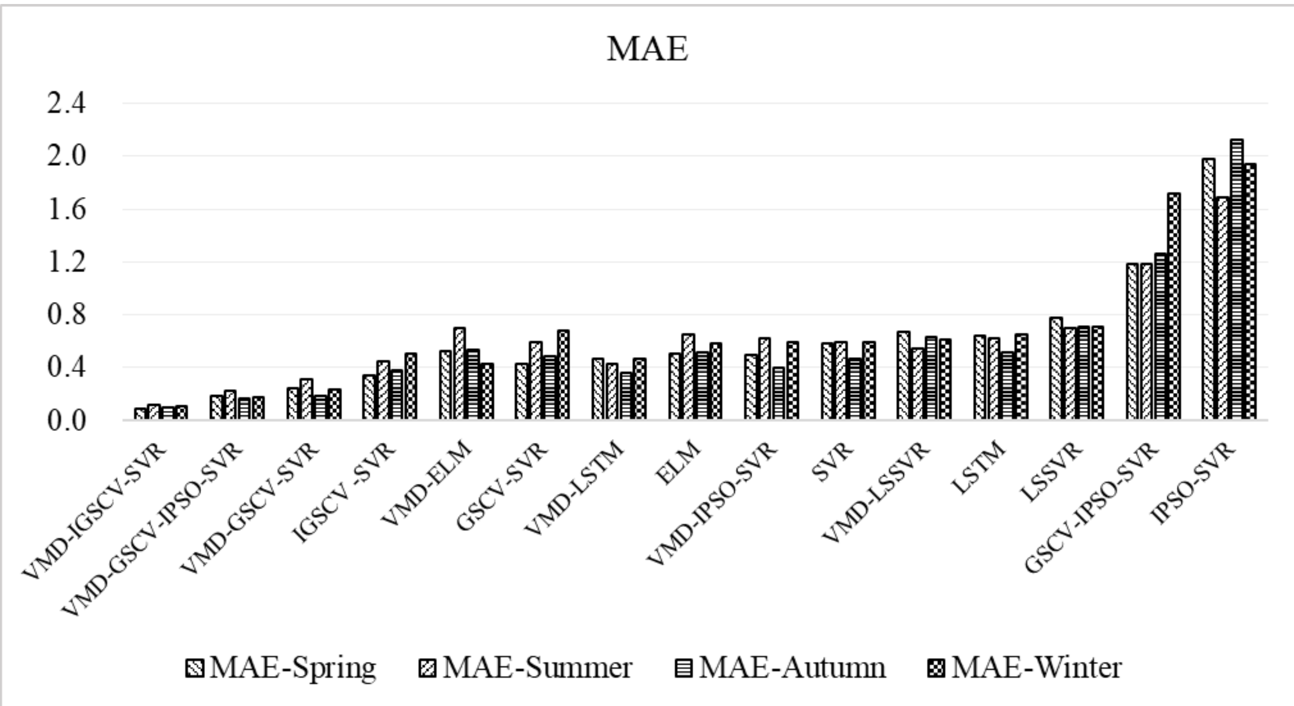

Fig. 11. MAE of all models in four seasons

REFERENCES

[1] D. Pei, W. Jianzhou, Y. Wendong, and N. Tong, "A novel hybrid model for short-term wind power forecasting," Applied Soft Computing, vol. 80, pp. 93-106, 2019.

[2] N. a. M. Safari, S. M. and Chung, C. Y., "Very Short-Term Wind Power Prediction Interval Framework via Bi-Level Optimization and Novel Convex Cost Function," IEEE Transactions on Power Systems, vol. 34, pp. 1289-1300, 2019, Art. no. 8477144.

[3] H. Haize, L. Yunyi, Z. Xiangping, and F. Mengge, "A novel hybrid model for short-term prediction of wind speed," Pattern Recognition, vol. 127, p. 108623, 2022.

[4] L. Hui, D. Zhu, and C. Chao, "Wind speed big data forecasting using time-variant multi-resolution ensemble model with clustering auto-encoder," Applied Energy, vol. 280, p. 115975, 2020.

[5] D. a. B. Lee, Ross, "Short-Term Wind Power Ensemble Prediction Based on Gaussian Processes and Neural Networks," IEEE Transactions on Smart Grid, vol. 5, pp. 501-510, 2014, Art. no. 6606922.
[6] F. Ümmühan Başaran and F. Tansu, "Wind Speed Prediction Using Artificial Neural Networks Based on Multiple Local Measurements in Eskisehir," Energy Procedia, vol. 107, pp. 264-269, 2017.
[7] J. Yan and H. Guoqing, "Short-term wind speed prediction: Hybrid of ensemble empirical mode decomposition, feature selection and error correction," Energy Conversion and Management, vol. 144, pp. 340-350, 2017.
[8] J. Tianyao, J. Yuzi, L. Mengshi, and W. Qinghua, "Ultra-short-term wind speed and wind power forecast via selective Hankelization and low-rank tensor learning-based predictor," International Journal of Electrical Power & Energy Systems, vol. 140, p. 107994, 2022.
[9] P. Mahum, K. Tariq, and M. F.-R. Luis, "A novel switched model predictive control of wind turbines using artificial neural network-Markov chains prediction with load mitigation," Ain Shams Engineering Journal, vol. 13, no. 2, p. 101577, 2022.
[10] W. W. Shuanglei Feng, Chun Liu, Huizhu Dai, "Study on physical method of wind farm power prediction [J]. Chinese Proceedings of Electrical Engineering," Chinese Proceedings of Electrical Engineering, vol. 30 (02), pp. 1-6., 2010.
[11] L. Zheng, L. Xiaorui, L. Mengjie, C. Xin, D. Shenhui, and S. Hexu, "Wind power prediction based on EEMD-Tent-SSA-LS-SVM," Energy Reports, vol. 8, pp. 3234-3243, 2022.
[12] Y. Chuanjin, L. Yongle, X. Huoyue, and Z. Mingjin, "Data mining-assisted short-term wind speed forecasting by wavelet packet decomposition and Elman neural network," Journal of Wind Engineering and Industrial Aerodynamics, vol. 175, pp. 136-143, 2018.
[13] L. Ming-De, D. Lin, and B. Yu-Long, "Application of hybrid model based on empirical mode decomposition, novel recurrent neural networks and the ARIMA to wind speed prediction," Energy Conversion and Management, vol. 233, p. 113917, 2021.
[14] W. Jian and Y. Zhongshan, "Ultra-short-term wind speed forecasting using an optimized artificial intelligence algorithm," Renewable Energy, vol. 171, pp. 1418-1435, 2021.
[15] M. Xi-wei, L. Hui, and L. Yan-fei, "Wind speed forecasting method using wavelet, extreme learning machine and outlier correction algorithm," Energy Conversion and Management, vol. 151, pp. 709-722, 2017.
[16] T. Akin, M. S. Borhan, P. Kameshwar, and V. Pravin, "Exploiting sparsity of interconnections in spatio-temporal wind speed forecasting using Wavelet Transform," Applied Energy, vol. 165, pp. 735-747, 2016.
[17] L. Sheng-Xiang and W. Lin, "Deep learning combined wind speed forecasting with hybrid time series decomposition and multi-objective parameter optimization," Applied Energy, vol. 311, p. 118674, 2022.
[18] L. Hui, T. Hong-qi, and L. Yan-fei, "An EMD-recursive ARIMA method to predict wind speed for railway strong wind warning system," Journal of Wind Engineering and Industrial Aerodynamics, vol. 141, pp. 27-38, 2015.
[19] L. Chaoshun, X. Zhengguang, X. Xin, Z. Wen, and Z. Chu, "A hybrid model based on synchronous optimisation for multi-step short-term wind speed forecasting," Applied Energy, vol. 215, pp. 131-144, 2018.
[20] Z. Yagang, P. Guifang, C. Bing, H. Jingyi, Z. Yuan, and Z. Chenhong, "Short-term wind speed prediction model based on GA-ANN improved by VMD," Renewable Energy, vol. 156, pp. 1373-1388, 2020.
[21] A. Ali Akbar, "A new intelligent method based on combination of VMD and ELM for short term wind power forecasting," Neurocomputing, vol. 203, pp. 111-120, 2016.
[22] X. Liye, Q. Feng, and S. Wei, "Multi-step wind speed forecasting based on a hybrid forecasting architecture and an improved bat algorithm," Energy Conversion and Management, vol. 143, pp. 410-430, 2017.
[23] W. Yun, W. Jianzhou, and W. Xiang, "A hybrid wind speed forecasting model based on phase space reconstruction theory and Markov model: A case study of wind farms in northwest China," Energy, vol. 91, pp. 556-572, 2015.
[24] L. W. DONG Lijiang, WANG Yu, QIAN Baiyun, SHEN Zhongxin, MA Qinyong, KONG Xiaoye, "Multi-parameter synchronous optimization of genetic algorithm for RVM short-term wind speed prediction model," CHINA MEASUREMENT & TEST, vol. (07), pp. 13-18, 2018.
[25] L. D.-s. WANG Yu- xin, GAO Yang, "Parameter s Optimization of SVM Based on Modified Flower Pollinate Algorithm," Fire Control & Command Control, vol. (10), pp. 8-13, 2018.
[26] W. L. Jing Wang, "Ultra-short term wind speed prediction based on CEEMD and GWO," Power system protection and control, vol. (09), pp. 69-74, 2018.
[27] D. Jiandong, W. Peng, M. Wentao, F. Shuai, and H. Zequan, "A novel hybrid model based on nonlinear weighted combination for short-term wind power forecasting," International Journal of Electrical Power & Energy Systems, vol. 134, p. 107452, 2022.
[28] L. Hui, W. Haiping, and L. Yanfei, "Multi-step wind speed forecasting model based on wavelet matching analysis and hybrid optimization framework," Sustainable Energy Technologies and Assessments, vol. 40, p. 100745, 2020.
[29] Z. Linyue, W. Jianzhou, and N. Xinsong, "Wind speed prediction system based on data pre-processing strategy and multi-objective dragonfly optimization algorithm," Sustainable Energy Technologies and Assessments, vol. 47, p. 101346, 2021.
[30] Y. L. Zhang Bowen, Lu Linghui, Huang Hongjun, Feng Jianhua, Zhang Enjie, "Single-pole-to-ground fault identification of photovoltaic power station collection system based on Pearson correlation coefficient," Proceedings of the CSU-EPSA, vol. 34, pp. 116-121, 2022.
[31] W. Patrik, "On the use of the Pearson correlation coefficient for model evaluation in genome-wide prediction," Frontiers in Genetic, pp. 10-899, 2019.
[32] J. Z. Jianxiu Hu, "Second order particle swarm optimization," Computer research and development, vol. (11), pp. 1825-1831, 2007.
[33] W. F. e. al, "A hybrid approach for measuring the vibrational trend of hydroelectric unit with enhanced multi-scale chaotic series analysis and optimized least squares support vector machine," Transactions of the Institute of Measurement and Control, vol. 41(15), pp. 4436-4449, 2019.
[34] Z. Wei, Hima, and H. Mahdi, "A chaos recurrent ANFIS optimized by PSO to predict ground vibration generated in rock blasting," Applied Soft Computing, vol. 108, p. 107434, 2021.
[35] J. Salman, M. Mojtaba, A. Rouzbeh, G. Vikram, G. Mohammadmahdi, and S. Fatemeh, "A hybrid SVR-PSO model to predict a CFD-based optimised bubbling fluidised bed pyrolysis reactor," Energy, vol. 191, p. 116414, 2020.
[36] D. Yunfei, C. Zijun, Z. Hongwei, W. Xin, and G. Ying, "A short-term wind power prediction model based on CEEMD and WOA-KELM," Renewable Energy, vol. 189, pp. 188-198, 2022.
[37] P. X.-j. LI Ling, MAI Xiong-fa, "A Research into PSO Algorithm Based on Stochastic Opposition Learning," Journal of Guangxi Teachers Education University：Natural Science Edition, vol. (29), pp. 1002-8743, 2012.

Yu Ye is studying for a master's degree in Nanchang Institute of Technology.

Jinxing Che received his B. S. degree from Jiujiang University in 2007 and his M. S. degree in applied mathematics from Lanzhou University in 2010, as well as his Ph. D. degree in mathematical statistics from Xidian University, China in 2019. He is currently an associate professor and Master's Supervisor in School of Science, Nanchang Institute of Technology, China. His main research interest is data analysis theory and application, hydrological information processing as well as prediction theory and method.

Heping Wang received his B. S. degree from Anhui Normal University in 2004 and his M. S. degree in applied mathematics from Nanning Normal University in 2010, as well as his Ph. D. degree in computational physics from Northwestern Polytechnical University, China in 2018. He is currently an lecturer in School of Science, Nanchang Institute of Technology, China. His main research interest is numerical calculation and simulation, hydrological information processing as well as prediction theory and method.