A Sample Entropy Parsimonious Model Using Decomposition-ensemble with SSA and CEEMDAN for Short-term Wind Speed Prediction

Jinxing Che, Yu Ye*, Heping Wang, Wenwei Huang

Abstract—**Data processing and integrated forecasting strategy has always been a major obstacle to the development of wind power forecasting system. In view of this, a novel decomposition method with SSA and CEEMDAN is constructed to decompose the original data, and also a sample entropy parsimonious integration model is applied to achieve ultra-short term wind speed prediction. Considering the respective data characteristics of each subsequence, we divide the decomposed multiple subsequences into three parts: high complexity group, low complexity group and residual group. PSO-ELM, IHOA-LSSVR, and IHOA-LSTM are applied to predict them respectively. Compared with other models with high accuracy in this paper, our model has higher prediction accuracy.**

Index Terms—**data processing, a sample entropy parsimonious integration model, ultra-short-term wind speed prediction, decomposition-ensemble, uncertainty analysis**

I. INTRODUCTION

WITH the sustained growth of power demand, and the shortage of fossil energy, the single use of fossil fuel for power generation fails to comply with the development. As one of the renewable energy sources, wind energy has broad application prospects [1]. However, wind power generation technology is still facing the problem of grid connection and "Abandoned wind " at the moment. Even if some areas are rich in wind resources, it is also a serious waste of wind energy [2]. Wind power prediction technology plays a considerable role in solving the problem of wind power grid connection and "Abandoned wind ".

There are four common methods in wind power prediction: physical model, statistical model, artificial intelligence model and probability model [3]. The physical prediction method based the wind power prediction curve constructed by numerical weather prediction [4]. Statistical prediction method analyzes the relatively complete historical data with mathematical methods to predict the future trends. Artificial intelligence prediction method processes nonlinear data by constructing neural network model [5]. Statistical model and artificial intelligence model are used widely. Generally, they can be divided into three categories: individual model, hybrid model and integrated model. Individual model uses a single independent model to analyze and model. The hybrid model combines multiple single models to avoid the limitations of a single model. It is able to learn the characteristics of time series completely, and reduce training error. There are two ways to integrate: Combine with decomposition technology or Integrate prediction results by optimization algorithms [6].

Generally, the subsequence is acquired by decomposition model, and then the segmented prediction is carried out by a single prediction model or a mixed prediction model [7]-[8]. Finally, the final prediction value is obtained by adding all predictions directly or calculating weight coefficient of all predictions by optimization algorithm [9]-[13].

Although a single decomposition model can denoise or decompose data to a certain extent, at present, few models can perfectly combine the two operations [14]. In order to fully learn the characteristics of time series, some scholars use decomposition models with different properties to perform combinatorial decomposition [15]-[17]. Considering the respective data characteristics of the individual subsequence, we propose an SSA-CEEMDAN mixed decomposition model. Then, we classify the complexity of each decomposition subsequence according to the principle of sample entropy calculation, use different mixed models for predictions, and finally merge an integrated prediction model. Experiments show that our proposed model has better adaptability and higher accuracy.

II. FUNDAMENTALS OF THE MODEL

A. Apply SSA to Denoise Wind Speed Sequence

The reconstructed sequence is selected by the contribution

Manuscript received August 30, 2022; revised December 28, 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).

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

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

Wenwei Huang is an undergraduate student in the School of Science, Nanchang Institute of Technology, Nanchang 330099, Jiangxi, China (e-mail: 3051508374@qq.com).

rate of the subsequence, the lowest contribution rate is discarded as the noise sequence. It mainly includes two parts: decomposition and reconstruction [18]. The calculation process of SSA is as follows:

Step 1. Suppose a time series has M samples, and the Embedding dimension is L, $L \le \frac{1}{3}M$. The sequence and the trajectory matrix can be expressed as (1-2), and then solve the SVD of the matrix G.

$$S=\left[s_1,s_2,...,s_M\right] \tag{1}$$

$$G=\begin{bmatrix} s_1 & s_2 & \cdots & s_{M-L+1} \\ s_2 & s_3 & \cdots & s_{M-L+2} \\ \vdots & \ddots & \ddots & \vdots \\ s_L & s_{L+1} & \cdots & s_M \end{bmatrix} \tag{2}$$

Step 2. Reconstruct time series. We rank all subsequences from large to small by the contribution rate, and select the first $M-1$ subsequences as refactor by diagonal average. Then we can get a new sequence $S^*=\left[s_1^*,s_2^*,..,s_M^*\right]$. s_t^* is expressed as (3-6). The detailed derivation process can be seen in [19]-[20].

$$s_t^*=\begin{cases} \sum_{j=0}^{p1-2}\sum_{i=1}^{j+1} s\left(j+1\right)+\frac{1}{j+1}\times T(i,j\text{-}i+2) \\ \sum_{j=p1-1}^{p2-1}\sum_{i=1}^{p1} s\left(j+1\right)+\frac{1}{p1}\times T(i,j\text{-}i+2) \\ \sum_{j=p2}^{M}\sum_{i=j-p2+2}^{M-p2+1} s\left(j+1\right)+\frac{1}{M-j}\times T(i,j\text{-}i+2) \end{cases} \tag{3}$$

$$K=M-L+1 \tag{4}$$

$$p1=\min\left(L,K\right) \tag{5}$$

$$p2=\max\left(L,K\right) \tag{6}$$

B. Decompose of New Data by CEEMDAN

Get a new time series through part A, and decompose the new sequence by CEEMDAN. IMF_1-IMF_n are obtained. Then, through processing, we can get the residual RES. CEEMDAN is grateful for its sustainable in nonlinear time series. It adaptively adds Gaussian white noise in the EMD decomposition process for many times, which can decrease the effects of modal aliasing and noise. It improves the operation efficiency of the model and becomes one of the more popular decomposition models at present [21]. The wind speed decomposition process as below:

Step 1. For simplicity, we write the new sequence S^* as $s\left(t\right)$, $t=1,...,M$, and set adaptive Gaussian noise with different signal-to-noise ratios as $\tau g_i\left(t\right)$. The original data is transformed into (7).

$$x_i\left(t\right)=s\left(t\right)+\tau g_i\left(t\right), i=1,2,3,..,n \tag{7}$$

Step 2. The subsequence is obtained by the decomposition of $x_i\left(t\right)$. Calculate the mean of n first components as the first component. Get the first residual component in (9).

$$IMF_1\left(t\right)=\frac{1}{n}\sum_{i=1}^{n} IMF_{1,i}\left(t\right) \tag{8}$$

$$r_1\left(t\right)=s\left(t\right)-IMF_1\left(t\right) \tag{9}$$

Step 3. Set $D_l\left(\bullet\right)$ as the EMD decomposition process in j phase. As in step 2, we can get the component of $j+1$ phase in (10).

$$IMF_{j+1}\left(t\right)=\frac{1}{n}\sum_{i=1}^{n} D_1\left\{r_j\left(t\right)+\tau\times D_j\left\{g_i\left(t\right)\right\}\right\} \tag{10}$$

Step 4. Repeat step 3 until the residual component cannot be decomposed, so we can get the residual in (11).

$$RES=s\left(t\right)-\sum_{i=1}^{j} IMF_j\left(t\right) \tag{11}$$

C. Introduction to Sample Entropy Parsimonious Model

Sample entropy is one of the measurement forms of entropy, which both has good robustness and strong anti-interference ability. It is applied to calculating the complexity of time series and measuring the probability of the system generating new patterns. The calculation processes of sample entropy are as follows.

Step 1. Suppose a time series as $s\left(t\right)$, and transform it into a matrix of $N\times M$ dimension.

$$S_{N\times M}=\begin{bmatrix} s(1) & s(2) & \cdots & s(M) \\ s(M+1) & s(M+2) & \cdots & s(2M) \\ \vdots & \ddots & \ddots & \vdots \\ s(N-M) & s(N-M+1) & \cdots & s(N) \end{bmatrix}_{N\times M} \tag{12}$$

Step 2. Calculate the distance between two vectors. Take $S\left(a\right)$ and $S\left(b\right)$ as examples, and the distance formula is expressed as follows.

$$\begin{gathered} D\left(a,b\right)=\max\left\{\left|S\left(a+M\right)-S\left(b+M\right)\right|\right\} \\ \left(1\le a,b\le N-M, a\ne b\right) \end{gathered} \tag{13}$$

Step 3. Set a threshold r_0, and record $D\left(a,b\right)\le r_0$ as Q_i. Get the mean of the ratio of Q_i to the total number of vectors, recording as $Q^M\left(r_0\right)$.

$$Q^M\left(r_0\right)=\frac{1}{N-M}\sum_{a=1}^{N-M} Q_a^M\left(r_0\right) \tag{14}$$

Step 4. Suppose the reconstruction dimension is $M+1$, and repeat step 1-3 to get the result $Q^{M+1}\left(r_0\right)$.

Step 5. The sample entropy of the finite time series in M-dimension is calculated as follows.

$$SE\left(M,r_0\right)=-\ln\frac{Q^{M+1}\left(r_0\right)}{Q^M\left(r_0\right)} \tag{15}$$

We propose the parsimonious criteria: suppose the sample entropy of k-th subsequence is SE_k, and the sample entropy of raw data is SE_{raw}. We get a percentage value by the following formula.

$$se_k=\frac{\left(SE_k-SE_{raw}\right)}{SE_{raw}}\times 100\% \tag{16}$$

It is expressed as high complexity sequence when se_k is positive. It is expressed as low complexity sequence when se_k is negative. Because the nature of the residual is unstable, we cognizance it as a special sequence for special treatment. As shown in Table I and Table II.

According to the fitness principle of models for the nonlinear complexity of subsequence, the prediction models are matched for different types of complexity.

PSO-ELM is suitable for high complexity sequences, IHOA-LSTM is suitable for residual sequence, and IHOA-LSSVR is suitable for low complexity sequences. Next, we describe the general calculation process of models.

D. Introduction to PSO-ELM

PSO algorithm is a population intelligent evolutionary algorithm proposed by Kennedy and Eberhart [22]. It is used to fall into local optimization. Therefore, many scholars are committed to exploring the methods to improve PSO algorithm and make PSO algorithm have better global convergence. However, Suganthan means a better solution can be obtained when c1 and c2 are constants [23]. The velocity position update formula of particle i is expressed in (17-19).

Where weight is a variable weight coefficient, $w_{\max}=0.9$, $w_{\min}=0.3$, $iter_{\max}=100$, and $c1=c2=1.8$.

$$\begin{aligned} V_i\left(t+1\right) &= weight*V_i\left(t\right) \\ &+c1*rand1*\left(X_{pbest}\left(t\right)-X_i\left(t\right)\right) \\ &+c2*rand2*\left(X_{gbest}\left(t\right)-X_i\left(t\right)\right) \end{aligned} \tag{17}$$

$$X_i\left(t+1\right)=X_i\left(t\right)+V_i\left(t+1\right) \tag{18}$$

$$weight = w_{\max} - ((w_{\max} - w_{\min})/iter_{\max})*iter \tag{19}$$

ELM is an improved version of feedforward perceptron neural network. It is composed of input layer, hidden layer and output layer. ELM has higher convergence and better generalization of gradient than BPNN. It can effectively prevent over fitting and local optimization. Weight and threshold are two key factors affecting network structure [24]-[25]. The objective function of ELM can be expressed as (20-24).

$$O=H\left(Input*w+b\right)*\rho \tag{20}$$

where ρ is the weight factor between hidden layer and output layer, $H(\bullet)$ is the activation function, and $*$ is the convolution. Such as "sigmoid" and "Relu" function, etc. "Sigmoid" is selected in our article. The expression is shown in (25).

$$X_{l,p}=\begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & \cdots & x_{p,1} \\ x_{2,1} & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & x_{l-1,p} \\ x_{l,1} & \cdots & \cdots & x_{l,p-1} & x_{l,p} \end{bmatrix} \tag{21}$$

$$Input=\begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p-1} \\ x_{2,1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ x_{l,1} & \cdots & \cdots & x_{l,p-1} \end{bmatrix} \tag{22}$$

$$w=\begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,\mathrm{k}} \\ w_{2,1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ w_{p,1} & \cdots & \cdots & w_{l,\mathrm{k}} \end{bmatrix} \tag{23}$$

$$\rho=\left[\rho_{1,1},....,\rho_{k,1}\right]^T \tag{24}$$

$$H\left(t\right)=\frac{1}{1+e^{-t}} \tag{25}$$

The procedures of ELM calculation are listed [26]. Randomly assign weights and thresholds to fixed shapes, and select the appropriate activation function firstly, and then determine the weight matrix by Moore Penrose inverse function. We can see ELM has good prediction performance. However, the weight and threshold of the model will be great limitations under the condition of manual selection. Thus, we apply PSO to adaptively selecting the parameters of ELM.

E. Introduction to IHOA-LSTM

IHOA is an improved hybrid optimization algorithm, which greatly makes up for the shortcomings of a single algorithm. It draws on the excellent characteristics of PSO and GA algorithms. Therefore, IHOA has obvious advantages in both local search and global search. It can better conduct global search in the initial period, and will not fall into local optimization too early. In the later period, it can converge better even the scope of the search is very small. The overall process is shown in Figure 1.

TABLE I
SEQUENCE COMPLEXITY CALCULATION BY CEEMDAN

Subsequence	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10	RES
CEEMDAN	1.35	1.18	1.36	0.83	0.63	0.56	0.24	0.24	0.05	0.01	\
Raw	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	\
se_k (%)	41.60	24.04	42.10	-12.81	-33.77	-41.06	-75.14	-75.01	-94.71	-98.84	\

TABLE II
SEQUENCE COMPLEXITY CALCULATION BY SSA-CEEMDAN

Subsequence	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10	IMF11	RES
SSA-CEEMDAN	1.14	1.32	1.35	0.90	0.63	0.54	0.21	0.29	0.05	0.03	0.00	\
Raw	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	\
se_k (%)	19.62	38.43	41.17	-6.04	-34.01	-43.87	-78.38	-69.52	-94.85	-96.41	-1	\

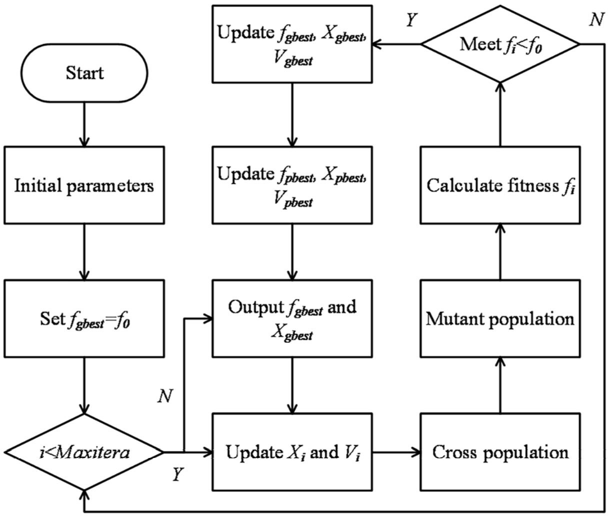

Fig.1. Flow chart of IHOA

The key optimization steps of IHOA are as follows.

Step 1. Initialize the parameters randomly.

Step 2. Update the speed according to (17), and update the location, which is shown in (26).

$$X_i(t+1)=X_i(t)+p*V_i(t+1) \tag{26}$$

The inheritance probability p is added here, and it can be set as needed. Keep formula (19) inconvenient for weight coefficient.

Step 3. Calculate the fitness of the population as f_i, and perform crossover and mutation.

Step 4. Evaluate the optimal fitness of the population, and update the individual extreme value X_{pbest} and population extreme value X_{gbest} of the population.

Step 5. Judge whether the iteration termination conditions are attained. When the conditions are met, it is output as the best fitness value. Otherwise, return to step 2.

LSTM is an improved model of RNN, which can effectively solve the problems of gradient disappearance and gradient explosion, and the network structure diagram is shown in Figure 2.

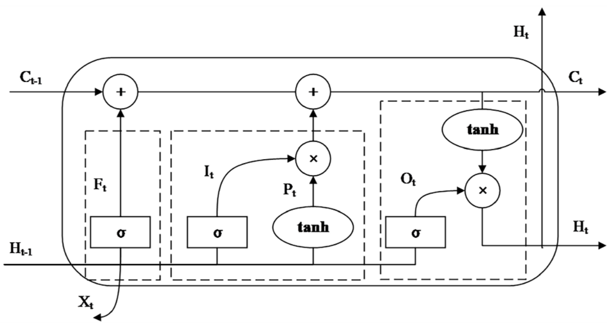

Fig.2. Neuron structure diagram of LSTM

LSTM can learn the long-term dependent information in the sequence. That is why LSTM has been widely used in various fields since its emergence [27].

The derivation formula is in (27-32).

$$F_t=\sigma\left(W_F\bullet[H_{t-1},X_t]+B_F\right) \tag{27}$$

$$I_t=\sigma\left(W_I\bullet[H_{t-1},X_t]+B_I\right) \tag{28}$$

$$P_t=\tanh\left(W_P\bullet[H_{t-1},X_t]+B_P\right) \tag{29}$$

$$C_t=F_t*C_{t-1}+I_t*P_t \tag{30}$$

$$O_t=\sigma\left(W_O\bullet[H_{t-1},X_t]\right)+B_O \tag{31}$$

$$H_t=O_t*\tanh\left(C_t\right) \tag{32}$$

F_t, I_t, C_t, O_t represent forgetting gate, input gate, cell state and output gate respectively. W_F, W_I, W_P, W_O represent the corresponding weights, and B_F, B_I, B_P, B_O represent the bias term. X_t is the input variables, and H_t is the output of the hidden layer at time t. Learning rate is an important factor, affecting the prediction results of LSTM. In this paper, IHOA is used to optimize the learning rate and the drop factor of learning rate in LSTM network. The prediction results of LSTM after parameter optimization are compared with those of LSTM. The experimental results show the prediction accuracy of IHOA-LSTM is better and more suitable for wind speed prediction [28].

F. Introduction to IHOA-LSSVR

SVR is widely used in classification and regression prediction tasks. That is attributed to its better generalization ability. SVR converts the complex calculation process in low-dimensional space to high-dimensional space, so we only need to calculate the linear model [29]. LSSVR is a variant model of the combination of SVR and PLS. It can effectively transform the quadratic programming problem into a linear equation to solve, which reduces the computational complexity of the traditional SVR [30].

The constrained optimization problem of LSSVR can be expressed as (33), and the regression formula established is shown in (34).

$$\begin{cases}\min\ J\langle w,\varepsilon\rangle=\frac{1}{2}\|w\|^2+\frac{1}{2}\eta\sum_{k=1}^{K}\varepsilon_k^2\\ s.t.\qquad O=w^T\phi(x_k)+b+\varepsilon_k\end{cases} \tag{33}$$

$$O(x)=\sum_{k=1}^{K}l_k\kappa\langle x,x_k\rangle+b \tag{34}$$

Where $J\langle\bullet\rangle$ is loss function, w is weight vector, ε is error, and η is regularization. $O(x)$ is output function of regression equation, $\kappa(\bullet)$ is kernel function, and b is bias. The RBF kernel function is selected as the kernel function, and the expression is shown in (35).

$$\kappa(x_a,x_b)=\exp\left(-\frac{\|x_a-x_b\|^2}{2\sigma^2}\right) \tag{35}$$

For more detailed mathematical derivation processes of LSSVR is shown in [31]-[32]. Regularization parameter η and kernel function C are two hyper parameters of LSSVR [33]. They are also the target parameters of our optimization. Experiments show that the prediction result of IHOA-LSSVR is better than LSSVR. Through the above steps, we get a sample entropy parsimonious model. There are two differences: the number of subsequences decomposed before and after denoising is distinct, and the Figure 3 shows the

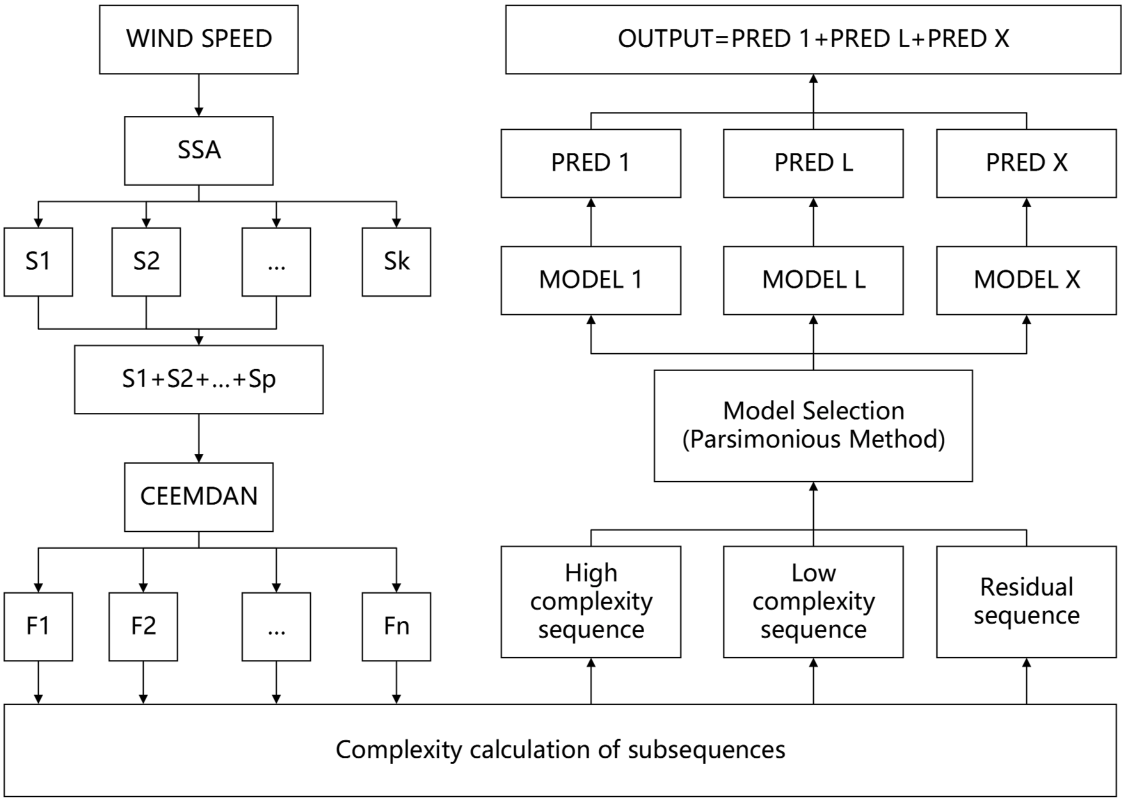

Fig.3. Flow chart

overall planning. Where the MODEL 1, ..., MODEL X represent three of the prediction models in this experiment. See Table III for details.

The parsimonious method is mainly to calculate the sample entropy of each series, divide all the series into three categories according to the threshold of sample entropy. Then, we randomly select a series from each category as the representative of this category, and select an appropriate prediction model for it. Finally, we can use this prediction model as the prediction model of all the sub series of this category.

TABLE III
NAME OF MAIN FORECAST MODEL

MODEL 1	PSO-ELM
MODEL L	IHOA-LSSVR
MODEL X	IHOA-LSTM

III. EXPERIMENT

In part II, we introduce the main idea of this paper. This part describes experiments to verify the reliability of our model. The experimental dataset is from Penglai wind farm, Shandong Province. It is the wind speed data at ten-minute intervals. In order to verify the model effectively, we select a small amount of data for our experimental. This experiment is simulated in MATLAB 2021a.

A. Experimental Design

In this paper, four groups of contrast experiments are designed: the first group of experiments is to predict directly without using decomposition model; The second group of experiments is to predict directly with a single decomposition model; The third group of experiments is to use SSA-CEEMDAN model to predict after decomposition; The fourth group of experiments is to select the models that perform well in the first three groups for comparison. Here we focus on the fourth group of experiments and present the models with high prediction accuracy in the form of graphs.

B. Error Calculation Method

It is also used to evaluate the prediction performance of the integrated model. We assume that the predicted value is y and the real value is x in i-th, $i \in N$, MSE expression is as (36).

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - y_i\right)^2 \tag{36}$$

Of course, it is far from judging the prediction effect of our model only according to MSE, on this basis, this paper adopts two other measures: mean absolute error (MAE) and mean average percentage error (MAPE).

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|\left(y_i - \hat{y}_i\right)\right| \tag{37}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\left(\frac{y_i - \hat{y}_i}{y_i}\right)\right| \tag{38}$$

where N is number of the sequence. We can use these evaluation indicators to measure the wind speed prediction effect of the model, which are also commonly used measurement indicators. The higher the value of these three indicators, the higher the prediction accuracy of the model.

C. Analysis of Experimental Results

To prove the rationality of the re-decomposition model proposed in this paper. The same prediction model

IHOA-LSSVR is selected, and the following four operations are performed on the data: no decomposition, with SSA decomposition, with CEEMDAN decomposition, and with SSA-CEEMDAN decomposition. Based on the statistical table of prediction errors of TABLE IV to TABLE VII, the four prediction models were analyzed and comprehensively sorted according to MSE, MAE and MAPE. The errors are sorted from large to small as follows: IHOA-LSSVR, SSA-IHOA-LSSVR, CEEMDAN-IHOA-LSSVR, SSA -CEEMDAN-IHOA-LSSVR. Looking at other models, we can also see that the decomposition effect of a single decomposition model is not good sometimes, but the prediction accuracy of the model after SSA-CEEMDAN re-decomposition can steadily decrease. Therefore, the re-decomposition model proposed in this paper plays a very important role in reducing the prediction model error.

TABLE IV
DENOISED EXPERIMENTAL RESULTS OF RAW DATA

Model	SSA-Raw data		
	MSE	MAE	MAPE
IHOA-LSSVR	0.378191	0.480192	0.233888
IHOA-LSTM	0.674732	0.639986	0.282527
LSTM	0.808759	0.709656	0.291590
ELM	2.404935	1.178880	0.378644
PSO-ELM	0.794808	0.716776	0.298603
LSSVR	5.415059	1.887088	0.511766

TABLE V
NONDENOISED EXPERIMENTAL RESULTS OF RAW DATA

Model	Raw data		
	MSE	MAE	MAPE
IHOA-LSSVR	0.233888	0.384914	0.483938
IHOA-LSTM	0.282527	0.640103	0.649365
LSTM	0.291590	0.745826	0.679293
ELM	0.378644	3.097865	1.335095
PSO-ELM	0.298603	0.907867	0.770949
LSSVR	0.511766	5.417281	1.887729

TABLE VI
FORECAST RESULTS AFTER SSA-CEEMDAN DECOMPOSITION

Model	SSA-CEEMDAN		
	MSE	MAE	MAPE
SampSE	0.189589	0.351794	0.204401
IHOA-LSSVR	0.221827	0.374722	0.211115
IHOA-LSTM	0.482889	0.556050	0.261559
LSTM	0.578778	0.620219	0.271190
ELM	2.197941	1.150787	0.386133
PSO-ELM	0.306853	0.444027	0.231996
LSSVR	2.592575	1.243551	0.413146

TABLE VII
FORECAST RESULTS AFTER CEEMDAN DECOMPOSITION

Model	CEEMDAN		
	MSE	MAE	MAPE
SampSE	0.206116	0.361672	0.205433
IHOA-LSSVR	0.327163	0.454635	0.231166
IHOA-LSTM	0.853059	0.757642	0.300487
LSTM	0.778552	0.705615	0.29373
ELM	3.439253	1.400244	0.453066
PSO-ELM	0.402416	0.502124	0.248672
LSSVR	2.331407	1.157242	0.397885

From TABLE VI and TABLE VII, we can find that the reduced sample entropy integration model is also critical to improve the accuracy of wind speed prediction. Since SSA-CEEMDAN decomposition is more effective than CEEMDAN decomposition, we will only discuss the SSA-CEEMDAN decomposition corresponding to the left half of TABLE V. From it, we can find that the prediction error of SampSE and IHOA-LSSVR models is very small, the error of SampSE is the smallest. Specifically, compared with IHOA-LSSVR, the MSE of SampSE decreases by about 14%, the MAE decreases by about 6%, and the MAPE decreases by about 3%. In conclusion, the reduced sample entropy integration model proposed in this paper can effectively reduce the prediction error. It is of great practical significance for ultra-short term wind speed prediction.

To more intuitively see the prediction effect of each model after SSA-CEEMDAN decomposition, a multi curve fitting effect diagram is drawn as shown in Figure 4. From it, we can see that the curve called SampSE has better coincidence with the real curve, and the curve named IHOA-LSSVR has only second fitting effect to SampSE curve, while other curves have average fitting effect.

Then, select the models with the highest prediction accuracy in this experiment, and draw the curve as shown in Figure 5. It can be seen from the sub graph that the fitting effect of the curve with five pointed stars is still the best. Although the performance effect in the middle section is poor, the fitting effect of other models at this stage is not as good as that of the curve with five pointed stars.

Finally, we observe the MSE and MAE prediction error histograms of the seven selected models with high accuracy. We rank them from large to small, they are shown in Figure 6 and Figure 7. We can find that the MSE and MAE of M7, the model proposed in this paper is the smallest of the seven models. It has the highest accuracy. The model proposed in this paper has a good advantage in improving the accuracy of wind speed prediction.

This is roughly the same as we can see from the prediction curve. Through various comparisons, it is found that our model has high prediction accuracy in this training process, it is proved that our method is feasible.

IV. CONCLUSION

We compared the prediction effects of models in this part. Through the above experiments and the analysis of experimental results, we can summarize the following points:

The prediction effect of IHOA-LSSVR is the best when the experiment is conducted with undecomposed data.

The model has better prediction accuracy when the sequence decomposed by CEEMDAN. At this time, the prediction accuracy of IHOA-LSSVR is still high, but not as good as that of SampSE.

When SSA-CEEMDAN is used for re-decomposition, the prediction accuracy of SampSE is further improved and becomes the model with the highest accuracy in this experiment.

Through this experiment and the comparative analysis of the experimental results, the correctness and effectiveness of the proposed model are obtained.

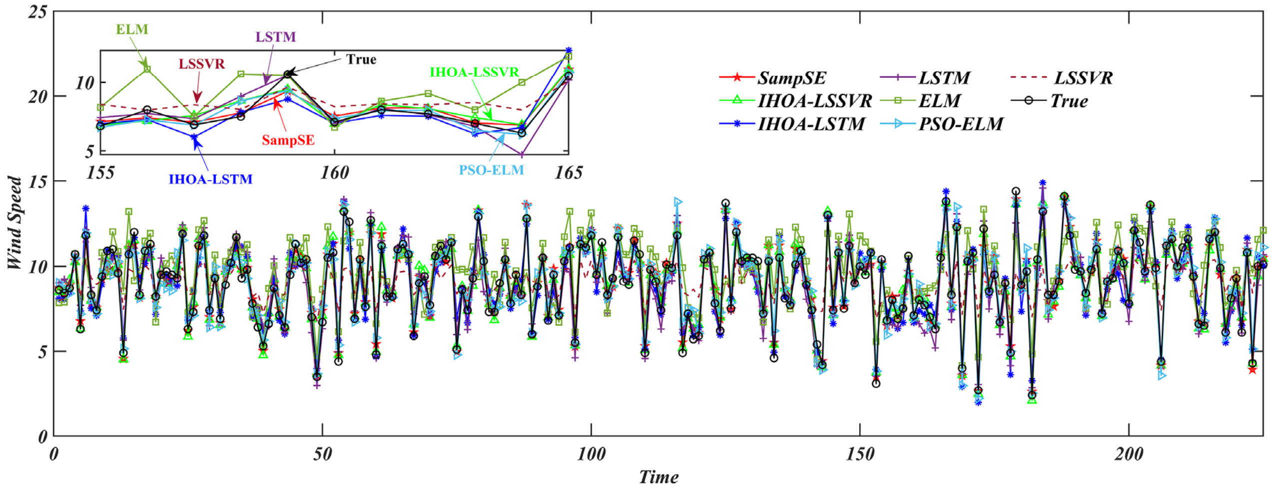

Fig.4. Comparison diagram of prediction model after SSA-CEEMDAN

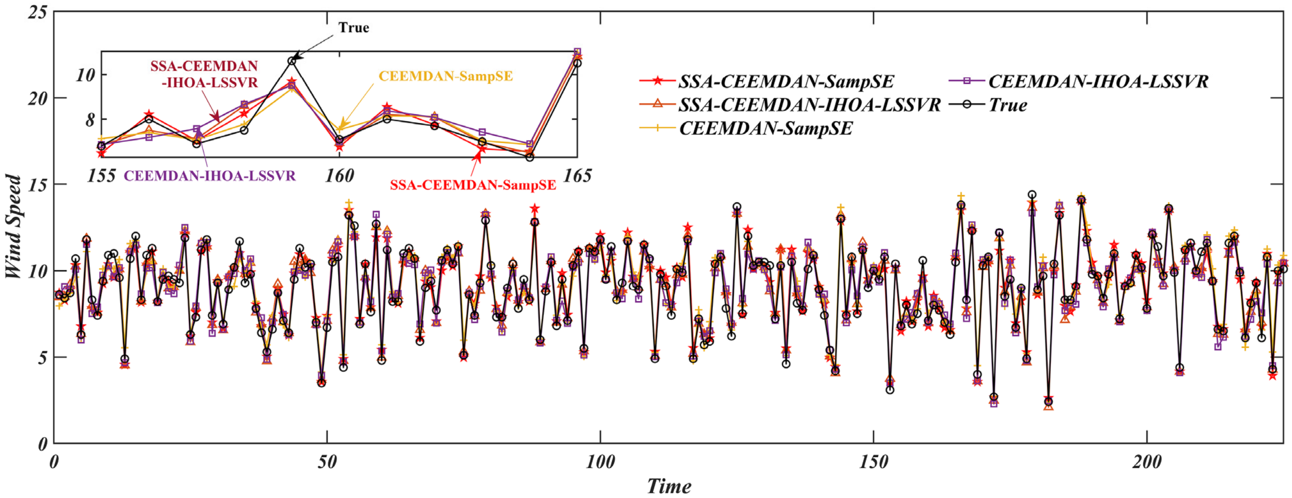

Fig.5. Comparison of prediction curves of four models with the highest accuracy

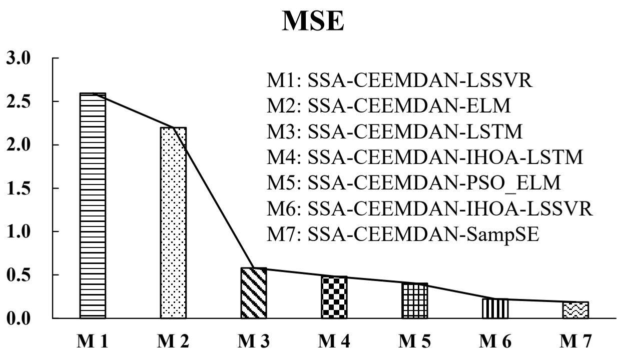

Fig.6. MSE of prediction model sorted in ascending order

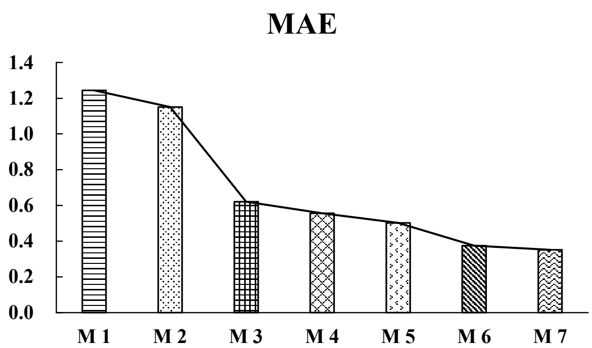

Fig.7. MAE of prediction model sorted in ascending order

REFERENCES

[1] W. Weigao, W. Yunbing, T. Xudong, and H. Yuan, "Short－term wind turbine generation power prediction based on sparrow search optimization support vector machine," Intelligent Computer and Applications, vol. 12, no. 01, pp. 119-123, 2022.

[2] G. Hao, L. Shan, L. Wenshuai, and e. al, "Research on wind speed prediction method based on EEMD-LSTSVR," Distribution & Utilization, vol. 39, no. 01, pp. 88-96, 2022.

[3] P. Chao, L. Runyu, C. Guowei, Y. Yuqing, and M. Tao, "Wind Speed Prediction with Self-tuning Convolutional Memory Based on Attribute Reduction Reconstruction," Proceedings of the CSEE, pp. 1-13.

[4] N. Dongxiao and J. Huizheng, "Quantitative Analysis Method for Errors Introduced by Physical Prediction Model of Wind Power,"

Automation of Electric Power Systems, vol. 44, no. 08, pp. 57-65, 2020.
[5] P. Chao, L. Runyu, W. Dian, C. Guowei, and Z. Yonghui, "MULTI-STEP WIND SPEED PREDICTION METHOD BASED ON WIND SPEED SPATIAL-TIME CORRELATION," ACTA ENERGIAE SOLARIS SINICA, vol. 43, no. 02, pp. 458-464, 2022.
[6] Y. Wendong, T. Zhirui, and H. Yan, "A novel ensemble model based on artificial intelligence and mixed-frequency techniques for wind speed forecasting," Energy Conversion and Management, vol. 252, p. 115086, 2022.
[7] H. Yumeng, D. Xingyu, W. Qunwei, and Z. Dequn, "A hybrid model for carbon price forecastingusing GARCH and long short-term memory network," Applied Energy, vol. 285, p. 116485, 2021.
[8] W. Xiaowei, L. Wenjie, W. Yingnan, and Y. Guotian, "A hybrid NOx emission prediction model based on CEEMDAN and AM-LSTM," Fuel, vol. 310, p. 122486, 2022.
[9] Z. Haochen, P. Zhiyun, T. Junjie, D. Ming, W. Ke, and L. Wenyuan, "A multi-layer extreme learning machine refined by sparrow search algorithm and weighted mean filter for short-term multi-step wind speed forecasting," Sustainable Energy Technologies and Assessments, vol. 50, p. 101698, 2022.
[10] Z. Yagang and L. Ruixuan, "Short term wind energy prediction model based on data decomposition and optimized LSSVM," Sustainable Energy Technologies and Assessments, vol. 52, p. 102025, 2022.
[11] 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.
[12] F. Guo-Feng, Z. Liu-Zhen, Y. Meng, H. Wei-Chiang, and D. Song-Qiao, "Applications of random forest in multivariable response surface for short-term load forecasting," International Journal of Electrical Power & Energy Systems, vol. 139, p. 108073, 2022.
[13] S. Yuanyuan, W. Jianzhou, Z. Haipeng, and Z. Weigang, "An advanced weighted system based on swarm intelligence optimization for wind speed prediction," Applied Mathematical Modelling, vol. 100, pp. 780-804, 2021.
[14] Z. Suling, W. Xinlu, M. Dongshuai, W. Lin, and L. Mingming, "CEEMD-MR-hybrid model based on sample entropy and random forest for SO2 prediction," Atmospheric Pollution Research, vol. 13, no. 3, p. 101358, 2022.
[15] L. Yu, Z. Xi, J. Liwen, L. Xiaoqing, and F. Xianghua, "Estimation of the foetal heart rate baseline based on singular spectrum analysis and empirical mode decomposition," Future Generation Computer Systems, vol. 112, pp. 126-135, 2020.
[16] M. Xiwei, L. Hui, and L. Yanfei, "Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine," Energy Conversion and Management, vol. 180, pp. 196-205, 2019.
[17] Z. Shuai, C. Yong, X. Jiuhong, Z. Wenyu, and F. Ruijun, "Hybrid wind speed forecasting model based on multivariate data secondary decomposition approach and deep learning algorithm with attention mechanism," Renewable Energy, vol. 174, pp. 688-704, 2021.
[18] W. Xun, Z. Shuning, Z. Lingzhi, C. Si, and Z. Huichang, "Research on anti-Narrowband AM jamming of Ultra-wideband impulse radio detection radar based on improved singular spectrum analysis," Measurement, vol. 188, p. 110386, 2022.
[19] Y. Hao, Z. Yun, M. Anbo, and L. Zhe, "Short-term electricity price forecasting based on singular spectrum analysis," Power System Protection and Control, vol. 47, no. 01, pp. 115-122, 2019.
[20] W. Jian, X. Song, Y. Cheng, W. Xiaodan, M. Jitao, and L. Fusuo, "Multi-step prediction of super-short-term wind power based on singular spectrum analysis," Renewable Energy Resources, vol. 39, no. 11, pp. 1548-1555, 2021.
[21] Z. Feite, H. Zhehao, and Z. Changhong, "Carbon price forecasting based on CEEMDAN and LSTM," Applied Energy, vol. 311, p. 118601, 2022.
[22] L. Daoqing, Y. Xiaodong, L. Shulin, D. Xia, Z. Hongzhi, and X. Rui, "Wind power prediction based on PSO-Kalman," Energy Reports, vol. 8, pp. 958-968, 2022.
[23] W. Jun, C. Junxing, Y. Shan, and C. Ming, "Short-term forecasting of natural gas prices by using a novel hybrid method based on a combination of the CEEMDAN-SE-and the PSO-ALS-optimized GRU network," Energy, vol. 233, p. 121082, 2021.
[24] D. Sudeepa, S. Tirath Prasad, and J. Rekh Ram, "Stock market forecasting using intrinsic time-scale decomposition in fusion with cluster based modified CSA optimized ELM," Journal of King Saud University - Computer and Information Sciences, 2021.
[25] A. Rana Muhammad, Reham, K. Ozgur, Y. Zaher Mundher, S. Shamsuddin, and Z.-K. Mohammad, "Improving streamflow prediction using a new hybrid ELM model combined with hybrid particle swarm optimization and grey wolf optimization," Knowledge-Based Systems, vol. 230, p. 107379, 2021.
[26] B. Abidhan, G. Anasua, G. Shubham, P. Biswajeet, and G. Candan, "A novel integrated approach of ELM and modified equilibrium optimizer for predicting soil compression index of subgrade layer of Dedicated Freight Corridor," Transportation Geotechnics, vol. 32, p. 100678, 2022.
[27] L. Zhongshan and Y. Jianhua, "Ultra-short term power prediction of photovoltaic power generation system based on EEMD-LSTM method," China Measurement & Test, pp. 1-8.
[28] X. Yuanhao et al., "Research on particle swarm optimization in LSTM neural networks for rainfall-runoff simulation," Journal of Hydrology, vol. 608, p. 127553, 2022.
[29] C. Renyin, Y. Junqi, Z. Min, F. Chunyong, and Z. Wanhu, "Short-term hybrid forecasting model of ice storage air-conditioning based on improved SVR," Journal of Building Engineering, vol. 50, p. 104194, 2022.
[30] W. Mingxing, S. Hongwei, L. Ming, and W. Chaoquan, "Prediction of split-phase flow of low-velocity oil-water two-phase flow based on PLS-SVR algorithm," Journal of Petroleum Science and Engineering, vol. 212, p. 110257, 2022.
[31] M. Mojtaba, A. Pouria, and H.-M. Seyed-Mohammad, "Groundwater level simulation and forecasting using interior search algorithm-least square support vector regression (ISA-LSSVR)," Groundwater for Sustainable Development, vol. 11, p. 100447, 2020.
[32] Z. Ping, C. Weiqi, Y. Chengming, J. Zhaohui, Y. Tao, and C. Tianyou, "Fast just-in-time-learning recursive multi-output LSSVR for quality prediction and control of multivariable dynamic systems," Engineering Applications of Artificial Intelligence, vol. 100, p. 104168, 2021.
[33] Z. Jun, S. Xinyu, G. Liang, J. Ping, and Q. Haobo, "A prior-knowledge input LSSVR metamodeling method with tuning based on cellular particle swarm optimization for engineering design," Expert Systems with Applications, vol. 41, no. 5, pp. 2111-2125, 2014.

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.

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

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 a 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.

Wenwei Huang is studying for a bachelor degree in Nanchang Institute of Technology.