Research on Short-term Wind Speed Prediction Based on Adaptive Hybrid Neural Network with Error Correction

Hongyu Long, Yunlong He, Wei Xiang*, Zhenqi Guan, Hao Tan, Jianbo Yu

Abstract—As a clean and efficient renewable energy, wind power is increasingly valued by countries all over the world. However, due to the inherent volatility, intermittency, and uncontrollability of wind power, the secure, stable, and economical operation of the power grid will face unprecedented pressure with wind power integration on a vast scale. Wind power predictions and management are significantly impacted by wind speed. Therefore, it is crucial to propose an accurate, stable, and efficient wind speed prediction method. In this paper, the original wind speed data is denoised using singular spectrum analysis to solve the problem of strong randomness and volatility of wind speed data. Then, the denoised data is processed using wavelet analysis, and the high-frequency and low-frequency sequences are respectively predicted using CNN-BiLSTM and CNN-BP, while the weights are optimized using the PSO algorithm. Finally, the prediction accuracy is further improved by incorporating the error correction module. In the two datasets used for testing, the outcomes of the trial indicate that the wind speed prediction model proposed in this paper performs extremely well in short-term wind speed prediction, superior to other models. The five evaluation indices (MAE, MSE, MAPE, RMSE, and R^2) for the first dataset are respectively 0.166, 0.055, 0.05, 0.234, and 0.953. Similar results were obtained for the second dataset. Furthermore, The ablation tests also show that each module suggested in this research is successful in enhancing prediction accuracy.

Index Terms—wind speed prediction, artificial neural network, hybrid model, convolutional neural network

Hongyu Long is a professor level senior engineer of Chongqing Key Laboratory of Complex Systems and Bionic Control, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: longhongyu20@163.com).

Yunlong He is a postgraduate student of Chongqing Key Laboratory of Complex Systems and Bionic Control, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: heyunlong_cqupt@163.com).

Wei Xiang* is a senior engineer of Economic and Technology Research Institute, State Grid Chongqing Electric Power Company, Chongqing 401120, China (corresponding author to provide phone: +8613983887778; e-mail: weixiang_cqupt@163.com).

Zhenqi Guan is a senior engineer of Economic and Technology Research Institute, State Grid Chongqing Electric Power Company, Chongqing 401120, China (e-mail: zhengqiguan_cqupt@163.com).

Hao Tan is a senior engineer of Economic and Technology Research Institute, State Grid Chongqing Electric Power Company, Chongqing 401120, China (e-mail: 6112219@qq.com).

Jianbo Yu is a senior engineer of Economic and Technology Research Institute, State Grid Chongqing Electric Power Company, Chongqing 401120, China (e-mail: 11941954@qq.com).

I. INTRODUCTION

ENERGY is a crucial factor in the continued growth of the global financial system[1]. The misuse and growing scarcity of non-renewable energy sources like coal and oil have become more critical problems as the degree of global economic growth continues to rise[2]. Therefore, the development and use of clean energy have become major research topics for countries around the world[3]. As environmental pollution and greenhouse gas emissions continue to worsen, wind power has become a rapidly developing and widely used clean and efficient power generation method in various countries around the world[4]. Its proportion in the power supply system is increasing, and its widespread and efficient usage is critical to global growth and environmental conservation. By 2050, the International Energy Agency wants to encourage the growth of renewable energy by increasing the supply of renewable energy to two-thirds of the world's total energy supply[5]. Environmentally friendly and sustainable energy has received increasing attention and has become the most promising power generation method currently[6].

However, there are certain issues with the reliable operation of the nationwide grid as a consequence of the widespread use of wind power. [7]. The effects of wind power on electricity security, stability, and economic operation have become increasingly clear as a result of the grid's continued integration of wind power and the ongoing rise in the amount of wind output[8]. Wind speed is the main factor in wind power prediction, and accurate wind speed signals will make a significant contribution to wind power prediction, which promotes ensuring the electrical grid's secure and trustworthy operation[9]. The inherent features of unpredictability and volatility present in wind speed signals, however, present a significant obstacle to the precise forecasting of short-term wind speed. Increasing the dependability and security of grid operation requires ensuring the safe and reliable integration of wind power, and a precise and trustworthy wind speed forecast is one of the most effective approaches to do this[10].

In an effort to improve the preciseness of wind speed forecasts, scholars have developed a number of effective wind speed prediction systems in recent decades. These prediction strategies may be categorized into physical, linear, nonlinear, and hybrid prediction methods. Physical prediction methods do not require too much historical data and fully consider the geological and meteorological factors of the prediction area, which is one of its major advantages.

However, because it is difficult to collect and analyze comprehensive meteorological factors in a brief time frame, physical prediction methods are more effective for medium-to long-term wind speed prediction than they are for short-term wind speed prediction. Currently, numerous scholars are engaged in research in this field. For instance, in [11], to assess the forecast mistakes, or residuals, of forecasted wind speed and direction in numerical weather prediction models, they use a thorough statistical methodology. Using data from Arctic station sites, they successfully combine statistical inference, probabilistic modeling, and hypothesis testing to provide outstanding findings. In [12], They used a point prediction method and incorporated data from three mathematical models for weather forecasting to produce results that were satisfactory. In [13], they put out a novel approach based on rank integration and probabilistic fluctuation perception for forecasting wind energy in order for the accuracy of numerical weather prediction. The findings demonstrate that for a four-hour lead time, the recommended approach can minimize the root mean square error (RMSE) by 2.16%–4.36%.

Moving average (MA), autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA), and autoregressive (AR) are the four primary linear prediction models used in linear prediction methods. However, these linear prediction algorithms frequently struggle to achieve the necessary prediction accuracy due to the randomness and uncertainty of wind speed. Current research in this area is also very advanced. For example, in [14], it is suggested to use a Window-Sliding ARIMA (WS-ARIMA) method, and its efficacy and accuracy in predicting daily and weekly wind speeds are confirmed. When compared to the ARIMA approach, this method lowers the overall RMSE of daily wind speed data by up to 75% and for weekly data by 50%. In [15], It is suggested to use an integrated framework based on EEMD-ARIMA. After decomposing the series using the ensemble empirical mode decomposition (EEMD) approach, the components are predicted using the Autoregressive Integrated Moving Average method (ARIMA). Through example analysis, it can be established that the suggested prediction algorithm is better at forecasting short-term wind speeds than the conventional ARIMA model. Additionally, in [16], It is advised to calculate short-term wind speeds by utilizing an annual autoregressive integrated moving average model. This model also achieves excellent accuracy.

With the popularization of artificial intelligence, methods such as artificial neural networks (ANN)[17–19] and support vector machines (SVM)[20–22] have been proposed and widely used. Compared with linear prediction methods, nonlinear prediction methods generally have better results in wind speed prediction and have obvious advantages in dealing with nonlinear problems.

The randomness component also makes it more complex and challenging to anticipate the wind speed accurately as it relates to wind speed variations. Hybrid prediction is a common practice because previous prediction algorithms had difficulty meeting the accuracy standards of wind speed prediction. Wind speed prediction is an extremely complex process, and hybrid prediction, which integrates data preprocessing and multiple prediction models, is currently widely used by scholars. Reference [23–34] are all recent examples of wind speed prediction using hybrid prediction models. Among them, [23-24], [30], [32] and [34] all innovatively incorporate spatial information extraction into traditional one-dimensional time series processing for wind speed prediction. [26] and [28] primarily concentrate on utilizing optimization algorithms to improve the model's hyperparameters. [35] is a combination of the previous two methods. [27], [29], [31] and [36], on the other hand, emphasize the integration of multiple models for ensemble prediction. Finally, [25] and [33] are more interested in how data pre-processing will affect the outcome of the prediction.

A hybrid predictive model based on Singular Spectrum Analysis (SSA), Wavelet Analysis, Convolutional Neural Networks (CNN), Back-Propagation Neural Network (BP), Bidirectional Long Short-Term Memory (BiLSTM), and other techniques is proposed in this paper, which will provide accurate short-term wind speed predictions based on historical wind speed data. The primary contributions of this study are as follows:

(1) With the initial wind speed information subjected to noise reduction, wind speed information with more pronounced features are obtained, creating the groundwork for precise prediction.

(2) The paper proposes an adaptive hybrid predictive model based on CNN-BP and CNN-BiLSTM. The two sub-models are utilized to forecast the high-frequency detail sequence and the low-frequency major component sequence, respectively, and optimization algorithms are employed to determine the best overlay weights.

(3) The hybrid model's training and testing mistakes are used as data sources for error correction, and an LSTM-based error correction model is employed to correct the prediction outputs in order to increase prediction accuracy.

(4) In order to validate the prediction performance of the suggested method from several angles, an assessment system based on MAE, MSE, RMSE, and R^2 is constructed.

The essay's remaining sections are organized as follows: Section II provides a detailed explanation of the SSA, CNN, BP, and BiLSTM core theories and mathematical models. The Particle Swarm Optimization (PSO) technique, the primary structure, and the modules of the proposed model are mostly discussed in Section III. In Section IV, to independently confirm the successful application of each module in the model, an assessment system for performance has been set up and ablation experiments are conducted. The major thesis of this study is finally summarized and discussed in Section V.

II. METHOD

In this part, we are going to present an in-depth description of the implementation principles of the algorithms and approaches employed in this study. Specifically, we will focus on SSA, CNN, BP, and BiLSTM. The methods discussed here are an essential component of the algorithm for the prediction that we suggest and provide assurance for the accuracy and continuity of the final prediction results.

A. Singular Spectrum Analysis

The SSA technique is effective for analyzing nonlinear

time series data. Colebrook first put forth the idea in his 1978 marine study. An observed time series is used to create a trajectory matrix, which is then broken down, determining the singular values associated with each component, and then reconstructing the matrix depending on the magnitude of the singular values. As a consequence, signals that characterize different components of the original time series, such as long-term trend signals, periodic signals, and noise signals, can be retrieved. Future trends can be predicted by delving into the time series' structure in this manner. The particular procedure is displayed below.

a. Embedding

The initial velocity of wind sequence is transformed into matrix Y, which is composed of an L-dimensional vector $Y_i = (x_1, \cdots, x_{i+L-1})^T$, with sample size N, as illustrated in the following:

$$Y = \begin{pmatrix} x_1 & x_2 & \cdots & x_K \\ x_2 & x_3 & \cdots & x_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & \cdots & x_N \end{pmatrix}, \qquad (1)$$

where, L represents the length of the embedding window, $K = N-L+1$.

b. Singular Value Decomposition

Calculate the matrix YY^T's eigenvalues, left singular matrix C_i, and right singular matrix D_i. The singular matrix as an entire thing and each of its vectors are now expressed as follows:

$$Y = Y_1 + Y_2 + \cdots + Y_d, \qquad (2)$$

$$Y_i = \sqrt{\lambda_i} C_i D_i, \qquad (3)$$

where, d represents the number of the rank of the matrix YY^T. λ_i represents the eigenvalues of the matrix YY^T, and C_i represents the orthonormal vectors corresponding to these eigenvalues.

c. Grouping

The decision-making process of the relevant and usable signal components is the central objective of the aforementioned method. The matrix Y is first divided into d submatrices. Any of the m submatrices with singular values greater than 0 is assumed to represent a component indicating long-term trends in the initial wind speed sequence. The long-term trend component matrix is defined as $Y_R = Y_{r1} + Y_{r2} + \ldots + Y_{rm}$, where Y_{r1}, Y_{r2}, ..., Y_{rm} indicate each sequence matching to each matrix. The component matrix consisting of these m submatrices is specified as $R = \{r_1, ..., r_m\}$. The noise component of the initial wind speed is therefore created through the remaining subsequences.

d. Reconstructing

The noise component will be turned into the noise time series X_{noise}, and the matrix Y_R will be transformed into the matching wind speed time series $X_R = \{X_{r1}, ..., X_{rm}\}$ via diagonal averaging. Here is the precise transformation procedure:

$$y_{ri} = \begin{cases} \dfrac{1}{i} \sum_{m=1}^{i} y^*_{m,k-m+1} & 1 \le i < l \\[2ex] \dfrac{1}{l} \sum_{m=1}^{i} y^*_{m,k-m+1} & l \le i \le k \\[2ex] \dfrac{1}{N-i+1} \sum_{m=i-k+1}^{N-k+1} y^*_{m,k-m+1} & k < i \le N \end{cases}, \qquad (4)$$

where, $l = min(L, K)$, $k = max(L, K)$.

Finally, the denoised results of the original wind speed time series by SSA are as follows:

$$X_N = X_{r1} + X_{r2} + \cdots + X_{rm} + X_{noise} \qquad (5)$$

B. Convolutional Neural Network (CNN)

CNN was initially developed for image classification and processing. It receives two-dimensional color-channeled input pictures to learn their attributes. Since the 1980s, this neural network model has been extensively employed by several researchers and has undergone ongoing strengthening and improvement. Its one-dimensional version has also been applied to time series processing and has achieved good results. The main components and functions of the CNN model are as follows:

a. Convolution Layer

The core of Convolutional Neural Networks (CNNs) lies in the convolutional layer, where convolution operations are applied to the input data. Convolution can be seen as a form of filtering operation in image processing, hence the convolutional kernel is also referred to as a filter. The construction of the convolutional layer is inspired by biological research, where humans have a receptive field to perceive information. By applying convolutional operations with appropriately sized kernels, information within the receptive field can be extracted from the raw data. Given an input data X, the feature map C of the convolutional layer can be represented as:

$$C = f(X \otimes W + b), \qquad (6)$$

where $\otimes$ is the convolution operation, W is the weight vector of the convolutional kernel, b represents bias, and f is the activation function. In this article, the ReLU function is used.

The convolution operation slides over the data with a certain interval, and performs element-wise multiplication with the kernel at each position, followed by a summation of the results. The detailed process of convolution is illustrated below.
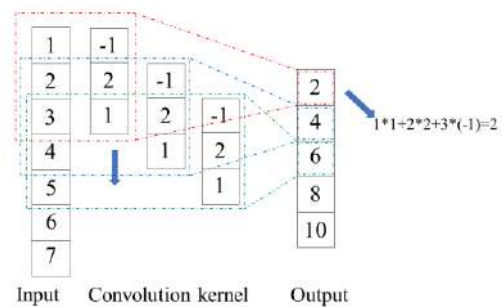

Fig. 1. Schematic diagram of one-dimensional convolution operation

b. Pooling Layer

Another crucial idea in convolutional neural networks is pooling. It is, in essence, a non-linear kind of downsampling that reduces the geographical size of the map of features. The difference between pooling and convolution layers is that feature maps in convolution layers require convolution with a kernel and the participation of parameters, which determine the error through forward propagation and are updated through backpropagation. Against this backdrop, pooling layers do not have any parameters to learn; they just extract the maximum or average value from the target region, depending on whether they use maximum or average pooling. Max pooling is more frequently used in practical applications, and this article also uses it. The calculation processes for max and average pooling are shown in the figure below.
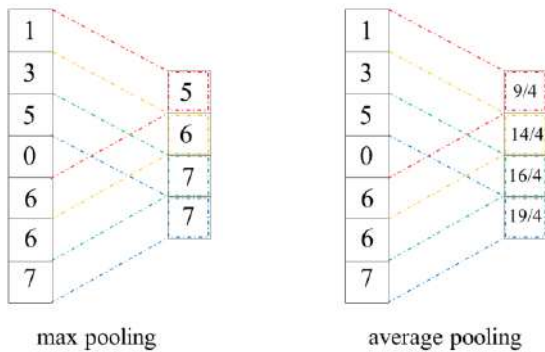
Fig. 2. Flowchart for calculating the maximum and average pooling operations

C. Backpropagation Neural Network (BP)

A typical kind of artificial neural network model is the BP neural network, sometimes referred to as the backpropagation neural network. It is a multi-layer feedforward neural network with an output layer, hidden layers, and input layer as a standard configuration. The BP neural network was initially proposed by American psychologists David Rumelhart and James McClelland in 1986 and is based on the backpropagation algorithm. The network receives input data at the input layer and transmits it via weighted connections to the output layer, where it can perform classifications or predictions on the input data. The weights of the connections are modified via the backpropagation method to reduce output error and training error. The error is transported backwards from the output layer in the backpropagation method, and each neuron's error gradient and weights are computed and modified using the gradient descent technique. The diagram that follows depicts the BP neural network's structure.

D. Bidirectional Long Short-Term Memory (BiLSTM)

BiLSTM, which is an improved version of LSTM neural network and can be used to process sequence data with long-term dependencies.

Schuster and Paliwal first presented the BiLSTM neural network in 1997. To improve the modeling capabilities of sequence data, two LSTM networks are combined—one in the forward direction and the other in the backward direction.

Information can only move forward in traditional LSTM networks. To better capture long-term dependencies and

contextual information in sequence data, BiLSTM networks, on the other hand, allow information to flow in both forward and backward directions at the same time and aggregate the outputs from both directions. Specifically, the BiLSTM network is made up of two LSTM networks, each of which processes the original sequence in a different way. The outputs from the two networks are then combined. The diagram that follows depicts the BiLSTM's construction.
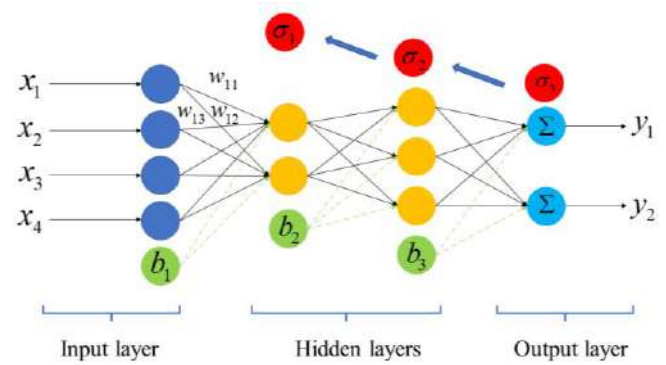

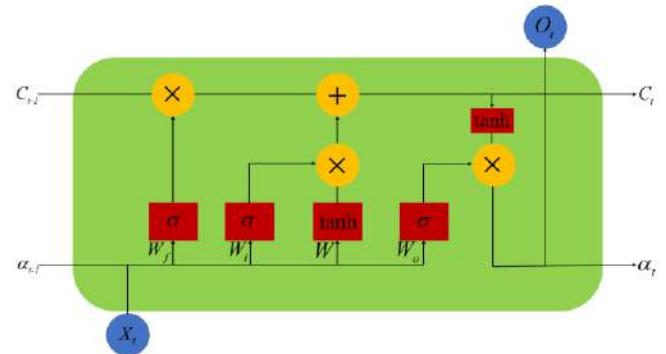Fig. 3. Schematic diagram of BP neural network structure

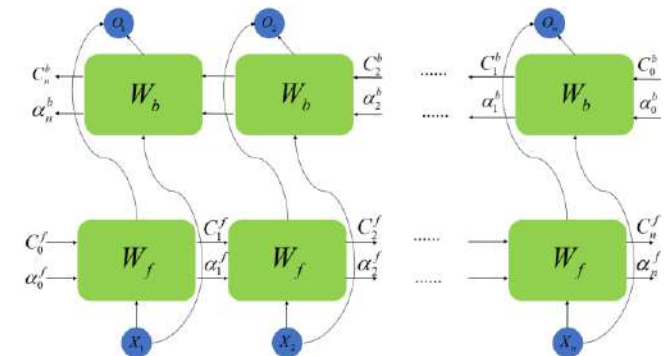Fig. 4. Schematic diagram of LSTM neural network structure

Fig. 5. Schematic diagram of BiLSTM neural network structure

III. Algorithm and Model Framework

The Particle Swarm Optimization (PSO) algorithm, which optimizes the overlay weights of the combination model in order to ultimately minimize the adaptive prediction error, is the algorithm for optimization that was used in this study and is the focus of this section. A thorough explanation of the wind speed prediction model, which is based on an adaptive hybrid neural network and correction of errors, is offered as well in this section.

A. Particle Swarm Optimization (PSO)

Kennedy and Eberhart first presented PSO, an optimization technique based on swarm intelligence, in 1995.

In order to find the best answer in a given search area, this algorithm continually modifies the position and velocity of each member of the swarm to mimic the behavior of a flock of birds or a school of fish.

The fundamental principle of PSO is to treat the optimization problem as a search problem in a multi-dimensional space, with each solution represented by a particle in the space that has a specific position and velocity. The method tries to discover the global optimum for the given issue by continually modifying the position and velocity of each particle in the swarm depending on its history and the successes of the entire swarm. Specifically, each particle uses its own history and the history of the entire swarm to update its velocity and position, gradually approaching the optimal solution over time. The general calculation process is shown below.

Suppose there exists an *N*-dimensional target search space, where a population of *M* particles constitute a swarm. Each particle in the swarm is represented by an *N*-dimensional vector, with P_i denoting the position of the *i*th particle, as in (7).

$$P_i = (p_{i1}, p_{i2}, \ldots, p_{iN}), i = 1, 2, \ldots, M \qquad (7)$$

The velocity of each particle is represented by an *N*-dimensional vector, as shown in (8).

$$V_i = (v_{i1}, v_{i2}, \ldots, v_{iN}), i = 1, 2, \ldots, M \qquad (8)$$

The optimal position currently found by the *i*th particle is called the individual extreme by E_{best}, as in (9).

$$E_{best} = (p_{i1}, p_{i2}, \ldots, p_{iN}), i = 1, 2, \ldots, M \qquad (9)$$

While the global optimal position found by the entire swarm is denoted by G_{best}, its definition is as follows.

$$G_{best} = (p_{g1}, p_{g2}, \ldots, p_{gN}), g = 1, 2, \ldots, M \qquad (10)$$

Using individual extremes and the global optimum, the following formula is used to update particle velocities and positions:

$$V_i = WV_i + C_1 R_1 (E_{best} - P_i) + C_2 R_2 (G_{best} - P_i) , \qquad (11)$$

$$P_i = P_i + V_i , \qquad (12)$$

where C_1 and C_2 are acceleration constants or learning rates, W is the inertia constant, and R_1 and R_2 are random numbers in the range of [0,2]. The flow chart of the algorithm is shown in Fig 6.

B. Adaptive Hybrid Neural Network and Error Correction Forecasting Model

The proposed adaptive hybrid neural network and error correction wind speed prediction model in this article comprises of three subsections, namely the portion of data processing based on SSA and wavelet analysis, the adaptive hybrid neural network prediction model part, which is based on CNN-BP, CNN-BiLSTM and PSO algorithm, and the error correction part based on LSTM. Fig. 7 displays the prediction model's general structure., and the structure and implementation steps of the main parts or modules will be introduced separately below.
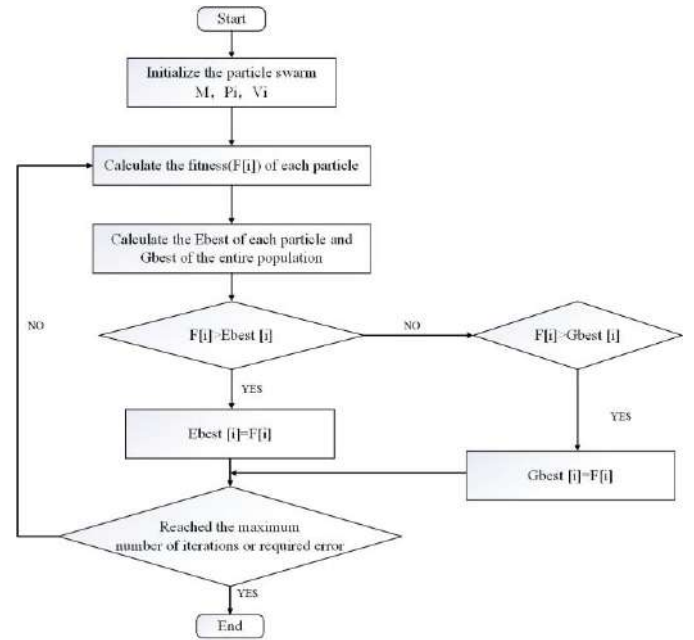

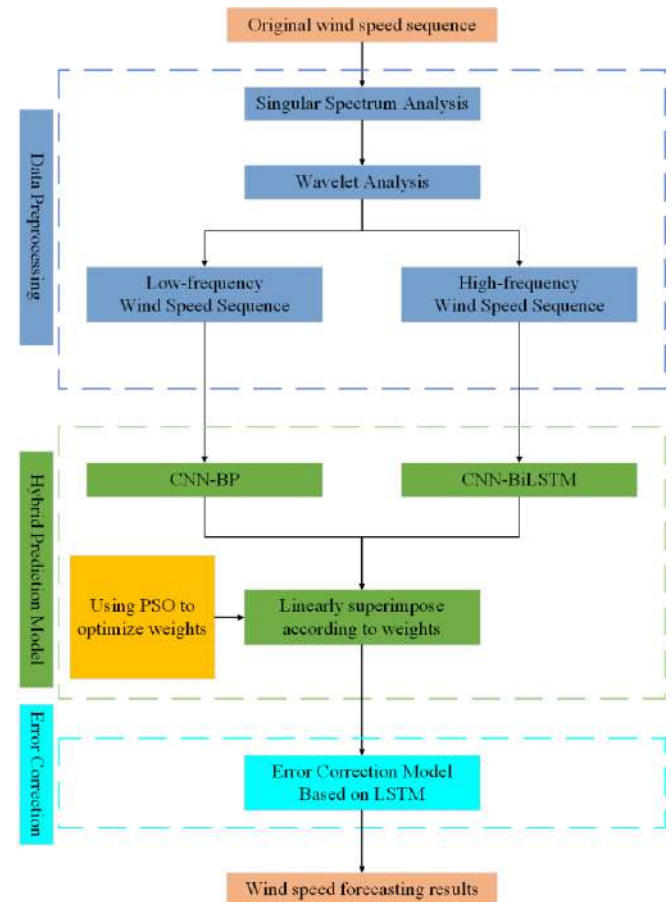Fig. 6. Calculation process of particle swarm optimization algorithm (PSO)

Fig. 7. The overall framework of the prediction model proposed in this article

a. Data Preprocessing

The initial wind speed information has strong uncertainty and volatility, it contains a lot of noise information, which is very unfavorable for subsequent prediction. Therefore, it is necessary to perform denoising processing on initial wind speed information. The data preparation module of the prediction model suggested in this article employs SSA to denoise the input raw data, reducing the noise and other information that will interfere with subsequent predictions in

initial wind speed information, and retaining the effective components as much as possible. The denoised data is then divided into high-frequency and low-frequency sequences by using wavelet analysis. Several prediction models are used to anticipate the data with different frequencies in order to make better use of the implicit features in the data.

b. Main Prediction Model of Adaptive Hybrid Neural Network

To predict the information at different frequencies, we use different models for prediction in this part. To verify the effectiveness of this approach, we selected a low-frequency sequence and a high-frequency sequence from the experimental data, and trained and predicted them respectively using CNN-BP and CNN-BiLSTM. For ease of presentation, the enlarged images of the test results are shown in Fig. 8 and Fig. 9. Apart from that, Table I represents the contrasting evaluation indicators of the two models for various frequency data, which further strengthens the conclusions. It can be seen that for low-frequency sequences, CNN-BP has a better advantage, while for high-frequency sequences, the more complex CNN-BiLSTM has higher accuracy. As a consequence, we employ several models in this article to forecast data at various frequencies before superimposing the predictions from the two models. To find the ideal weights and improve the accuracy of the prediction outcomes, the PSO algorithm optimizes the best weights of each model during the overlapping process. The benefit of this strategy is that it can more effectively utilize the implicit properties of various sequences, allowing for the adoption of more appropriate models for prediction and raising the model's overall prediction accuracy.
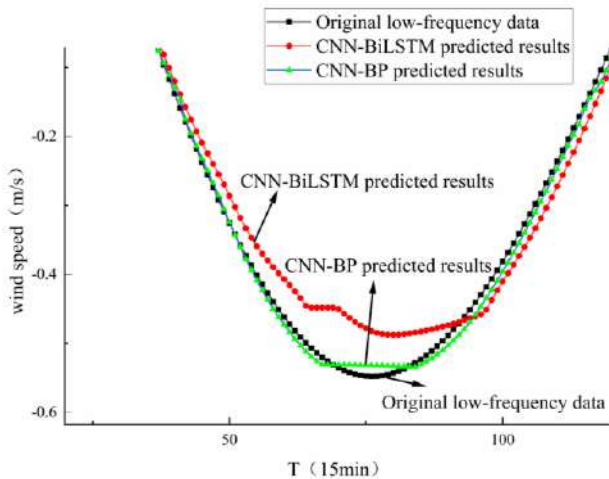

Fig. 8. Comparison between CNN-BiLSTM and CNN-BP in original low-frequency data

c. Error Correction

This article provides an error correction module to increase forecast accuracy further. The module is added after the model prediction is complete. Utilizing an LSTM network, the error value needed for rectification is obtained. The mistakes produced during the training and testing phases of the prediction model in the previous module provided the initial data for training and testing the LSTM network. In section II, the LSTM network's model structure was discussed.
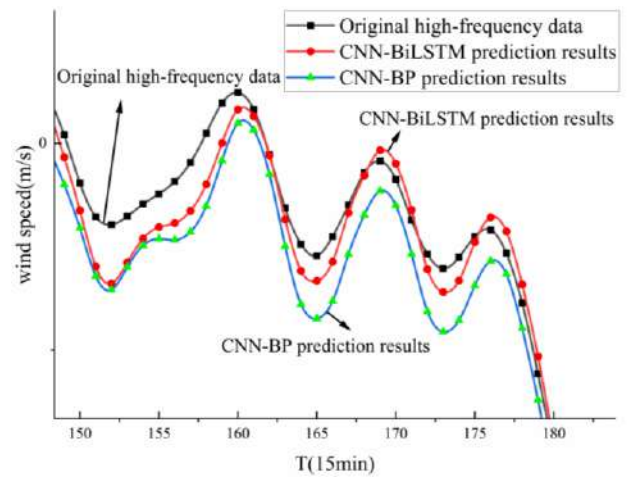

Fig. 9. Comparison between CNN-BiLSTM and CNN-BP in original high-frequency data

TABLE I
COMPARISON OF MODELS IN TEST DATA

	Model	RMSE	MAPE	MAE
Low-frequency data	CNN-BP	0.016	0.077	0.013
	CNN-BiLSTM	0.031	0.212	0.026
High-frequency data	CNN-BP	0.371	0.073	0.282
	CNN-BiLSTM	0.257	0.049	0.193

IV. EXPERIMENTS AND RESULTS ANALYSIS

A. Original Data Set and Data processing.

Two wind speed datasets obtained from a wind farm in Jiangsu Province, China, are used in this work, each with 2000 data sets and a time interval of 15 minutes. We ensured that there was no missing data or any other abnormal situations when conducting our experiments, and used the first 1,968 sets of data for our experiments, with the training set being the first 1,392, the validation set being sets 1,393 to 1,680, and the testing set being sets 1,681 to 1,968. In the experiments, the prediction interval was set to 48 sets, meaning that we used the previous 48 sets to predict the data for the next time point. Fig. 10 and Fig. 11 display the two datasets, respectively.
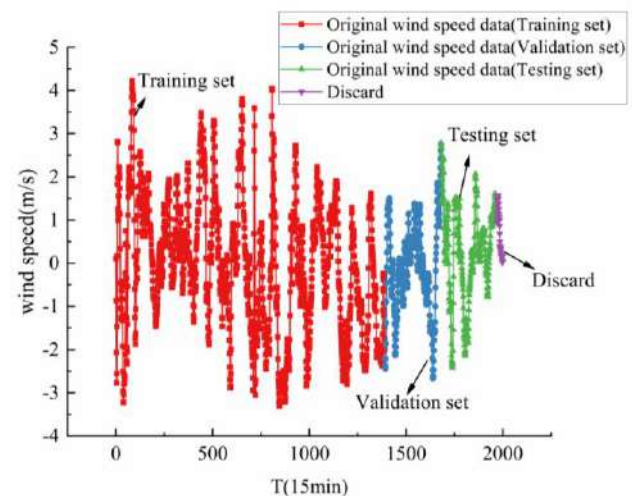

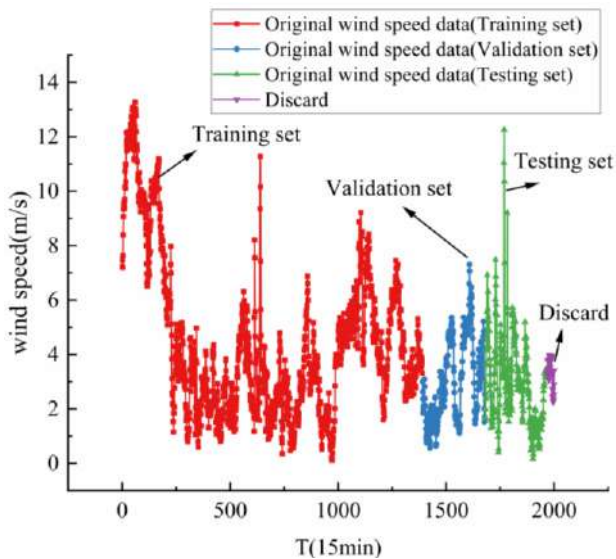Fig. 10. Original wind speed dataset 1 and grouping information

Fig. 11. Original wind speed dataset 2 and grouping information

As shown in Section II, in order to reduce the influence of noisy data in wind speed information on subsequent prediction and use mixed neural network for prediction, we used SSA and wavelet analysis to denoise and group the original data by frequency. In the original two wind speed datasets, we used SSA to divide the original data into ten groups and arrange them according to singular values. The processing results of the two datasets are shown in Fig. 12 and Fig. 13. From these two figures, we can conclude that the lower the wind speed signal, the higher the complexity and the stronger the volatility. many. Therefore, in the subsequent processing, we need to further analyze these data to remove as many noise signals as possible, so as to lay the foundation for the final accurate prediction.

In both datasets we took the first of the 10 sets of subsequences arranged according to the singular values, which is the primary component of wind speed data, as the reference and calculated Pearson correlation coefficients between each of the remaining groups and the first group. Based on these coefficients, we cleaned the data for denoising, using the magnitude difference of Pearson coefficients as the selection criterion for the data. Data with large discrepancies will be discarded as noise. The reason for doing so is to retain the main components contained in the original wind speed data as much as possible and to avoid excessive prediction errors caused by inappropriate denoising. Heatmaps of Pearson correlation coefficients between the decomposed groups of the two datasets are depicted in Fig. 14 and Fig. 15, respectively. Observing these two pictures separately, we can see that among the remaining nine groups, the differences in the Pearson correlation coefficients between the second, third and fourth groups and the first group are smaller than those of the other groups, and both datasets roughly show this phenomenon. Therefore, we respectively retain the first four groups in the two datasets as the wind speed data after denoising. Fig. 16 and Fig. 17 depict the comparisons of datasets 1 and 2 before and after denoising, respectively.
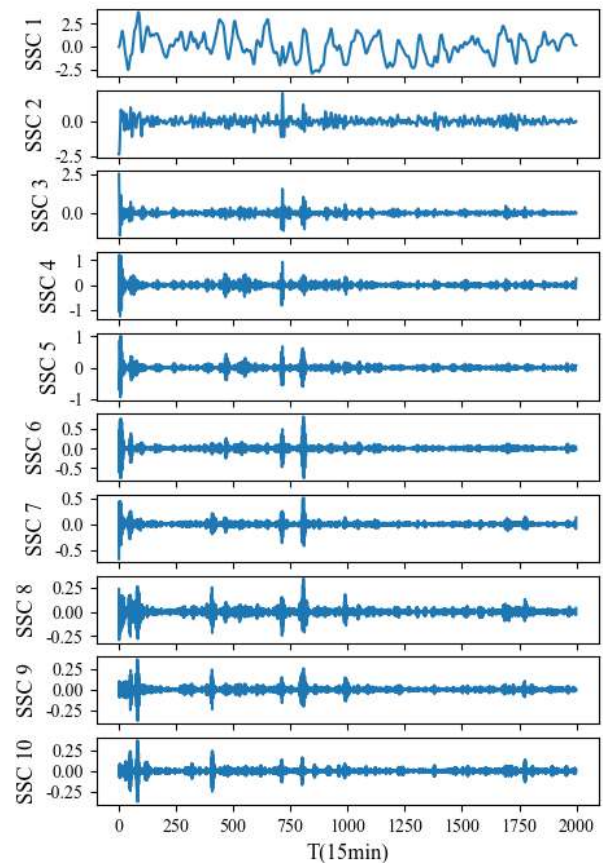

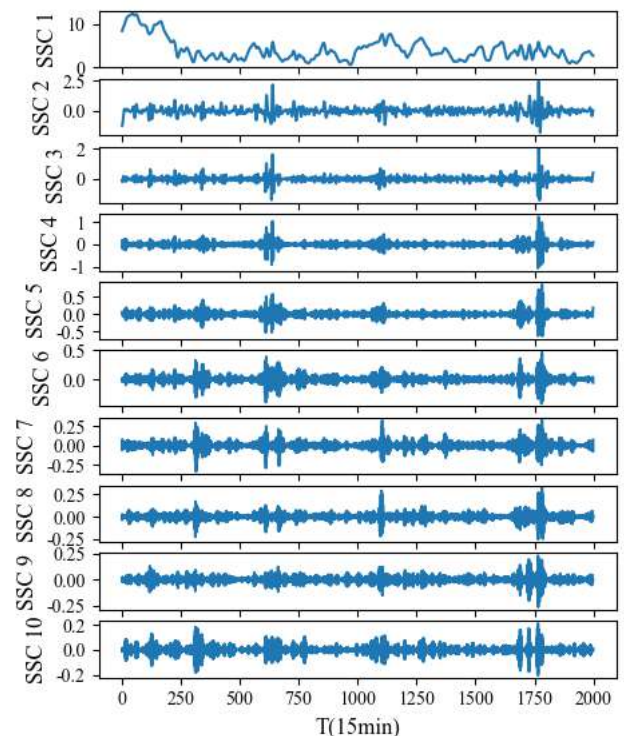Fig. 12. Sequence of ten components for dataset 1 sorted by singular value

Fig. 13. Sequence of ten components for dataset 2 sorted by singular value

Fig. 14. The Pearson correlation coefficients between each group in dataset 1 after sorting by singular values. From the figure, we can see that the Pearson correlation coefficients of groups 2, 3, and 4 with the first group have smaller order of magnitude differences compared with the remaining nine groups, which are 0.285, 0.03, and 0.017, respectively. Therefore, these three groups together with the first group as the benchmark are retained as the main components in the original wind speed data, and the remaining part is discarded as noise.

Fig. 15. The Pearson correlation coefficients between each group in dataset 2 after sorting by singular values. From the figure, we can see that the Pearson correlation coefficients of groups 2, 3, and 4 with the first group have smaller order of magnitude differences compared with the remaining nine groups, which are 0.139, 0.024, and 0.011, respectively. Therefore, these three groups together with the first group as the benchmark are retained as the main components in the original wind speed data, and the remaining part is discarded as noise. This is approximately the same as the processing result of dataset 1.

As mentioned in the description of the hybrid prediction network in the previous section, we will employ two different models to predict the denoised data and using wavelet analysis to achieve improved accuracy. The denoised wind speed data will be split into five sub-sequences in this step based on frequency. The results of the two datasets after processing with wavelet analysis are shown in Fig. 18 and Fig. 19, from which we can see that the frequencies increase from component 1 to component 5, indicating that the low-frequency sequence constitutes the approximate components of the original data, while the high-frequency sequence constitutes the detail components of the original data. Next, we manually integrated the first three subsequences of the two datasets into low-frequency sequences and the last two subsequences into high-frequency sequences according to the above principles, respectively, to obtain the final prediction sequences for subsequent prediction using different prediction models. In Fig. 20, the processing results for low and high frequencies of the two datasets are displayed.
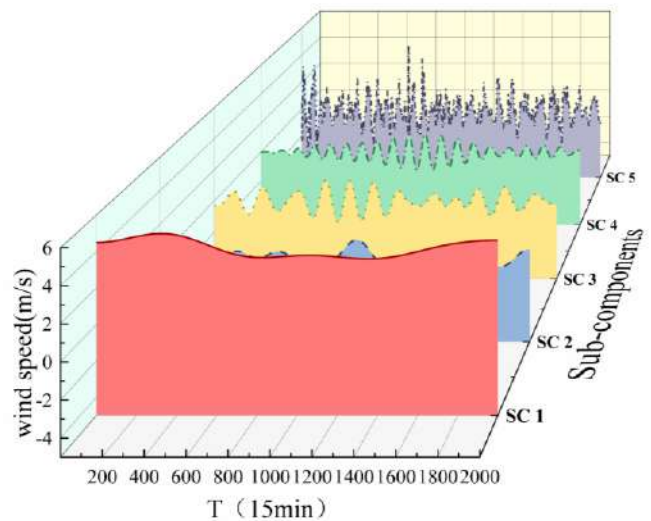

Fig. 18. Result of comparative analysis of sub-components of dataset 1 sorted by frequency after wavelet analysis
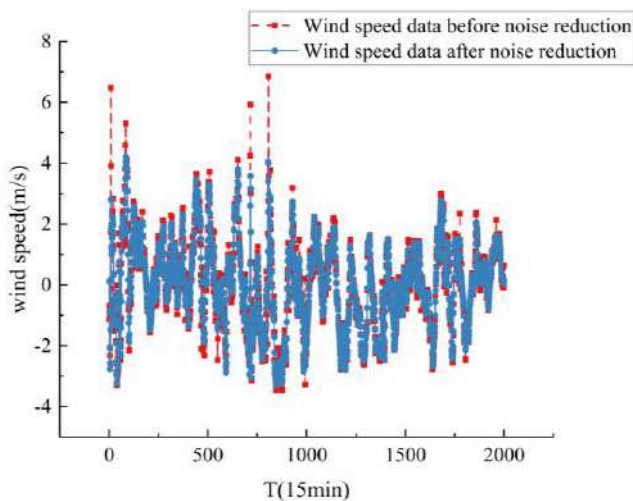

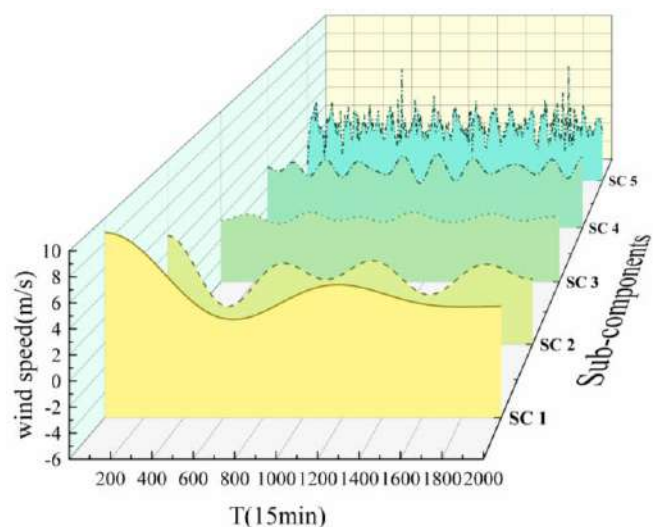Fig. 16. Comparison of the original wind speed series in dataset 1 before and after noise reduction

Fig. 19. Result of comparative analysis of sub-components of dataset 2 sorted by frequency after wavelet analysis
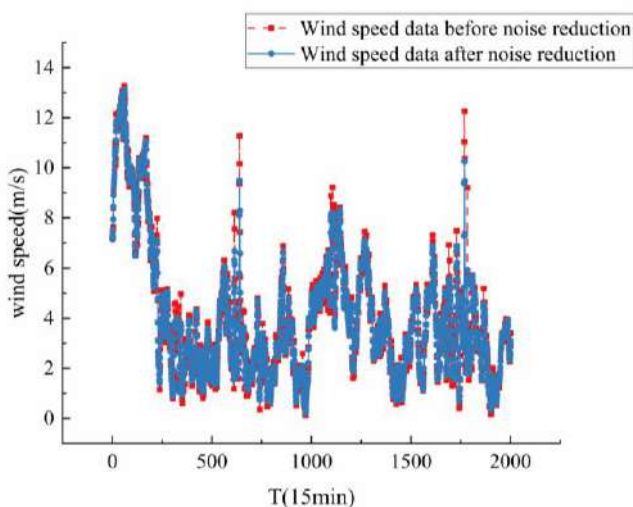

Fig. 17. Comparison of the original wind speed series in dataset 2 before and after noise reduction
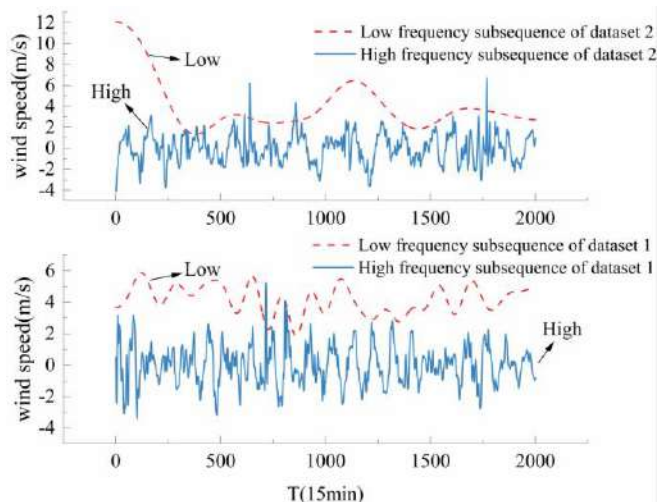

Fig. 20. Low-frequency and high-frequency reconstruction results for two datasets

B. Performance evaluation index

For confirmation that the prediction model that was presented in this study is superior, we apply five evaluation indicators, namely MSE, MAE, RMSE, MAPE, and R^2, to

confirm and assess the effectiveness of our suggested prediction model. These assessment metrics are used often in pertinent research to assess how well the model predicts outcomes. We may more thoroughly assess the performance of our suggested model by comparing these parameters. Below are the expressions in mathematics for each evaluation metric:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}[y(i) - \hat{y}(i)]^2 \ , \qquad (13)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y(i) - \hat{y}(i)\right| , \qquad (14)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[y(i) - \hat{y}(i)]} \ , \qquad (15)$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y(i) - \hat{y}(i)}{y(i)}\right| , \qquad (16)$$

$$R^2 = 1 - \frac{\sum_i (y(i) - \hat{y}(i))^2}{\sum_i (\bar{y} - \hat{y}(i))^2} \ , \qquad (17)$$

where $y(i)$ represents the wind speed measurement data at time i, $\hat{y}(i)$ is the predicted value of wind speed, N is the number of data points, and the values for MSE, MAE, RMSE, and MAPE are all better when smaller, while the value of R^2 is better when closer to 1.

These metrics may effectively and completely illustrate how well a model predicts outcomes. Based on these assessment indices, we will assess the model's performance in the next section.

C. Prediction results and comparative analysis

For the purpose of to confirm the accurateness and stability of the suggested model, a comparative analysis will be conducted between it and six other prediction models, namely SSA-CNN-BiLSTM, SSA-CNN-BP, SSA-CNN-LSTM, SSA-BiLSTM, SSA-BP and SSA-LSTM. Each model will be separately run five times, with the average value being used as the final forecast result, to reduce the influence of randomness on the comparison findings. Fig. 21 and Fig.22 compare predictions made by several models and show how they compare. We validate the efficacy of the model described in this work by comparing and evaluating the projected outcomes of various models.

Moreover, for the proposed of rigorously evaluate the accuracy of our proposed model in wind speed prediction, we have compared the predictive performance of seven models using the aforementioned five evaluation indexes. Table II and Table III present the comparative results from two datasets. Among them, the data representing the best performance are also bolded for easy observation. And for easy visualization, we also compare the evaluation metrics of the different models in the two datasets by using the bar charts in Fig. 23 and Fig. 24, respectively.

Based on Table II and Fig. 23, it is evident that the proposed predictive model exhibits high accuracy in wind speed prediction. Among all the contenders that were compared, the proposed model had the best performance, with MAE, MSE, MAPE, RMSE, and R^2 values of 0.166, 0.055, 0.05, 0.234, and 0.953, respectively, for dataset 1 used in our experiments.

Apart from that, the evaluation indices of the various models in the dataset 2 we employed show findings that are quite consistent. With Table III and Fig. 24, we can easily conclude that the values of MAE, MSE, MAPE, RMSE and R^2 of the proposed method are 0.325, 0.276, 0.126, 0.525 and 0.910, respectively, which are the best results among all the models used in the experiment. This further proves the good performance of the proposed model in this paper for wind speed prediction. It is worth noting that there is a certain gap between the evaluation indexes of dataset 1 and dataset 2, which is caused by the volatility and uncertainty of wind speed data.

Furthermore, we found that incorporating 1dCNN is very successful in predicting short-term wind speeds by contrasting the predictive outcomes of the fundamental models without 1dCNN integration with the same models with 1dCNN integration. For instance, the MAE, MSE, MAPE, RMSE, and R^2 values of SSA-BP for dataset 1 are, respectively, 0.348, 0.228, 0.089, 0.477, and 0.806. However, for SSA-CNN-BP, the addition of 1dCNN significantly increased the predictive performance, resulting in optimized performance indexes of 0.211, 0.077, 0.054, 0.277, and 0.934, respectively. This trend was similarly observed in the other two models, LSTM and BiLSTM. The same phenomenon is also reflected in dataset 2. Thus, we can strongly conclude that 1dCNN integration is highly effective in enhancing predictive accuracy.

Additionally, by comparing the prediction results and evaluation indexes of the models based on BP and BiLSTM, we observed that BP is better suited for low-frequency signal prediction, while BiLSTM is better suited for high-frequency signal prediction. which was previously mentioned in the text. This largely illustrates the effectiveness and reliability of our method of using two models to separately train and test the low-frequency subsequences and high-frequency subsequences obtained after processing. This also lays the foundation for the accuracy and stability of our subsequent forecasts.

Furthermore, the approach we propose surpasses the comparison models in terms of prediction results and evaluation indicators, and it performs better than other different prediction models in terms of prediction precision and consistency, further demonstrating the value of using hybrid models for wind speed prediction.

To further illustrate how each module of our suggested approach works to increase the precision of wind speed predictions, we conducted corresponding ablation experiments. We compared different models by gradually adding sub-modules of our proposed method. We started from a single model with BiLSTM as the main prediction network and sequentially added the hybrid prediction model, the adaptive weight optimization algorithm, and the error correction module to compare the final predictive performance. Fig. 25 and Fig. 26 depict the predictions made by the four models using the two datasets, while Table IV and Table V compare the performance indices.
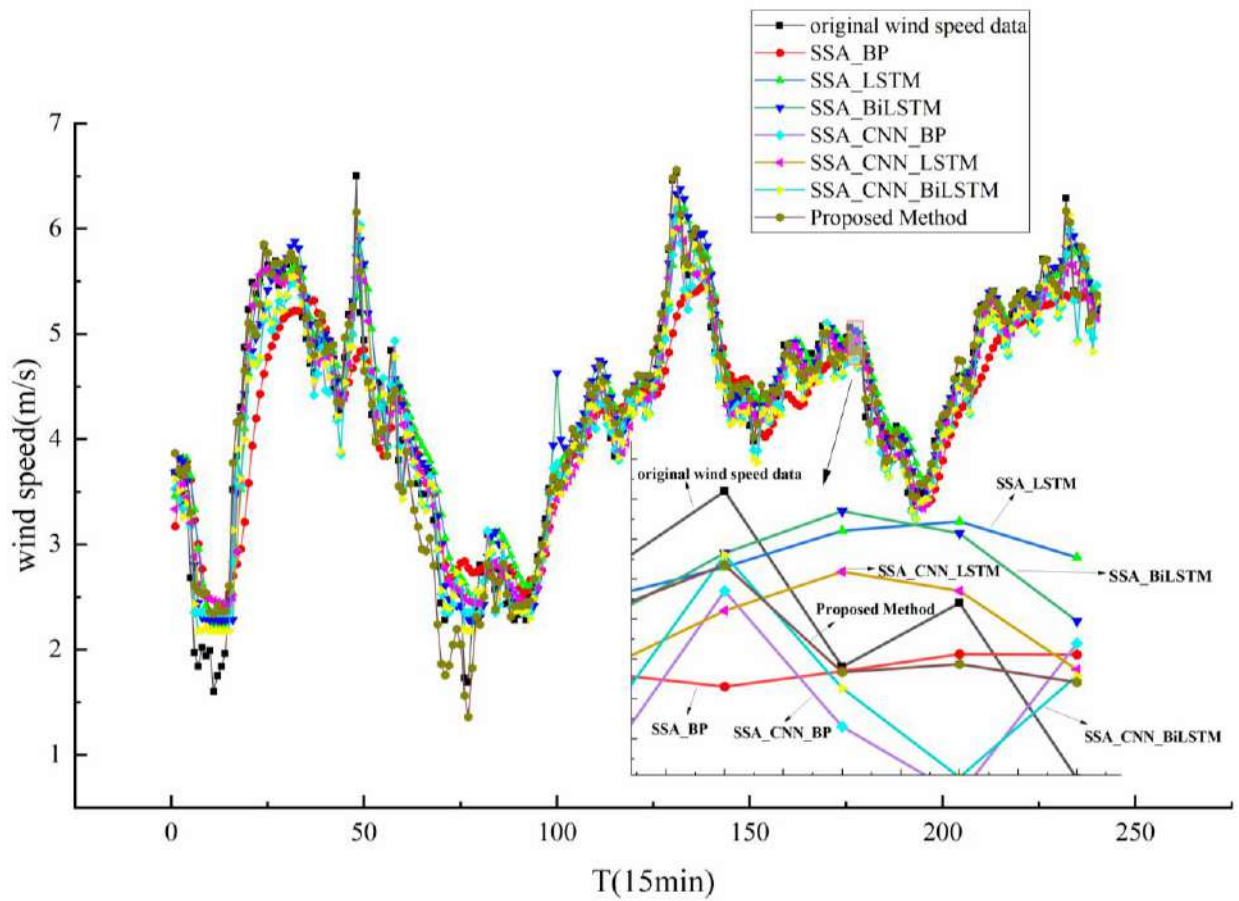
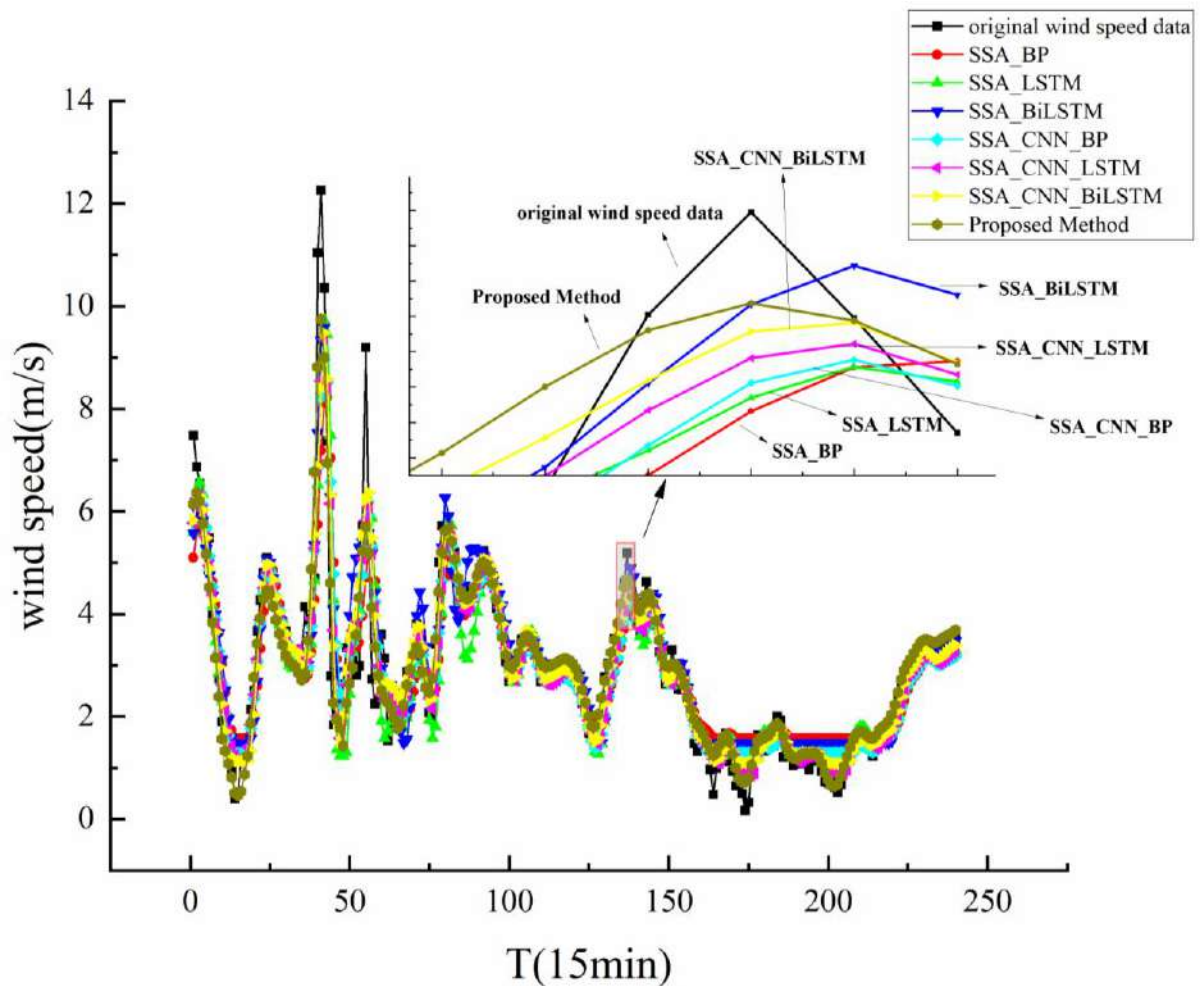Fig. 21. Comparison of predictive results of different models for dataset 1

Fig. 22. Comparison of predictive results of different models for dataset 2

TABLE II
COMPARISON OF FIVE EVALUATION INDEX IN DATASET 1

Model	MAE	MSE	MAPE	RMSE	R^2
SSA-BP	0.348	0.228	0.089	0.477	0.806
SSA-LSTM	0.269	0.133	0.068	0.364	0.887
SSA-BiLSTM	0.244	0.091	0.057	0.301	0.923
SSA-CNN-BP	0.211	0.077	0.054	0.277	0.934
SSA-CNN-LSTM	0.208	0.081	0.056	0.285	0.931
SSA-CNN-BiLSTM	0.193	0.066	0.05	0.257	0.944
Proposed Method	**0.166**	**0.055**	**0.05**	**0.234**	**0.953**

TABLE III
COMPARISON OF FIVE EVALUATION INDEX IN DATASET 2

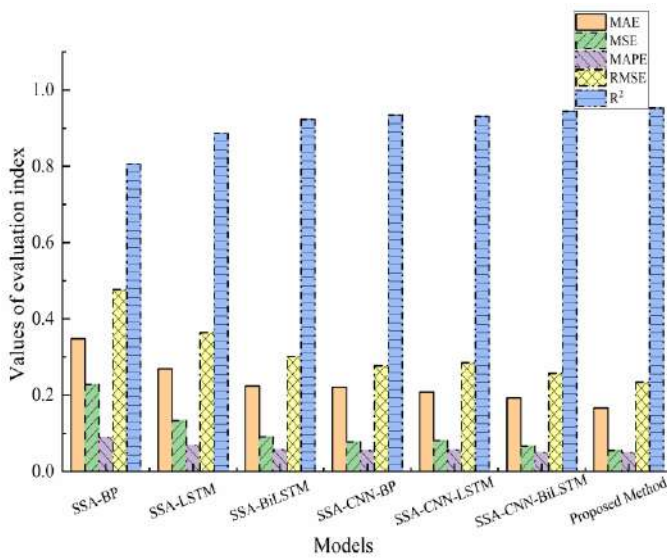| Model	MAE	MSE	MAPE	RMSE	R^2
SSA-BP	0.533	0.813	0.192	0.901	0.734
SSA-LSTM	0.490	0.670	0.182	0.819	0.781
SSA-BiLSTM	0.464	0.572	0.176	0.756	0.813
SSA-CNN-BP	0.407	0.523	0.152	0.722	0.829
SSA-CNN-LSTM	0.376	0.443	0.143	0.666	0.855
SSA-CNN-BiLSTM	0.355	0.428	0.133	0.654	0.861
Proposed Method	**0.313**	**0.276**	**0.126**	**0.525**	**0.910**

Fig. 23. Comparison of evaluation index results of different models in dataset 1
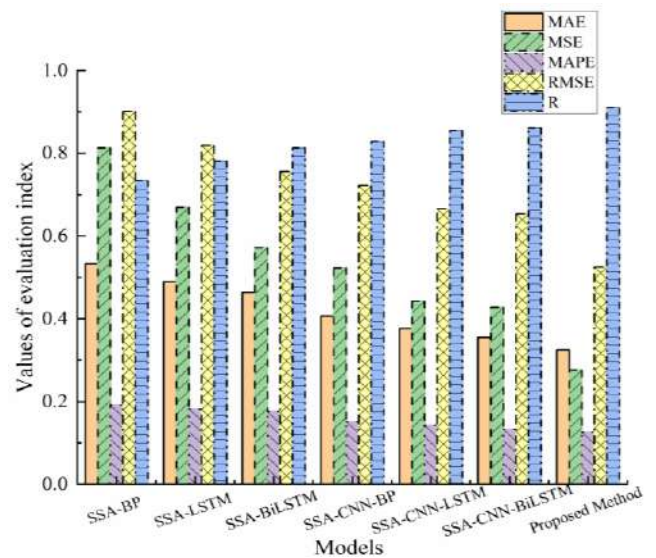

Fig. 24. Comparison of evaluation index results of different models in dataset 2
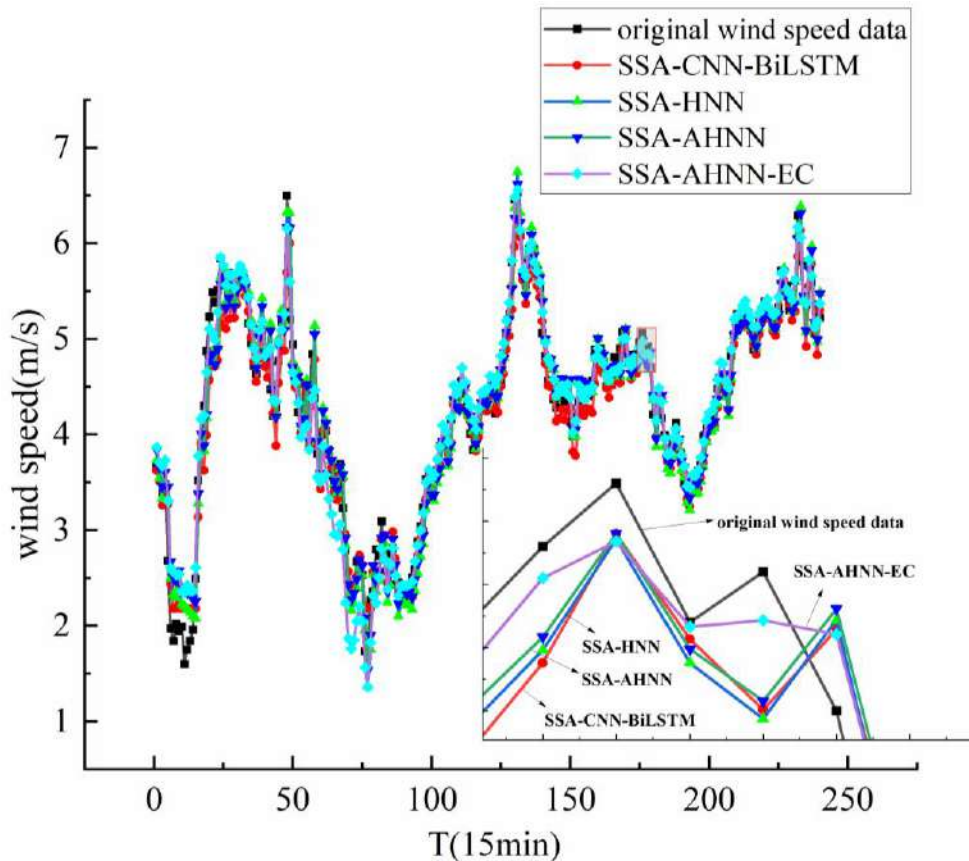

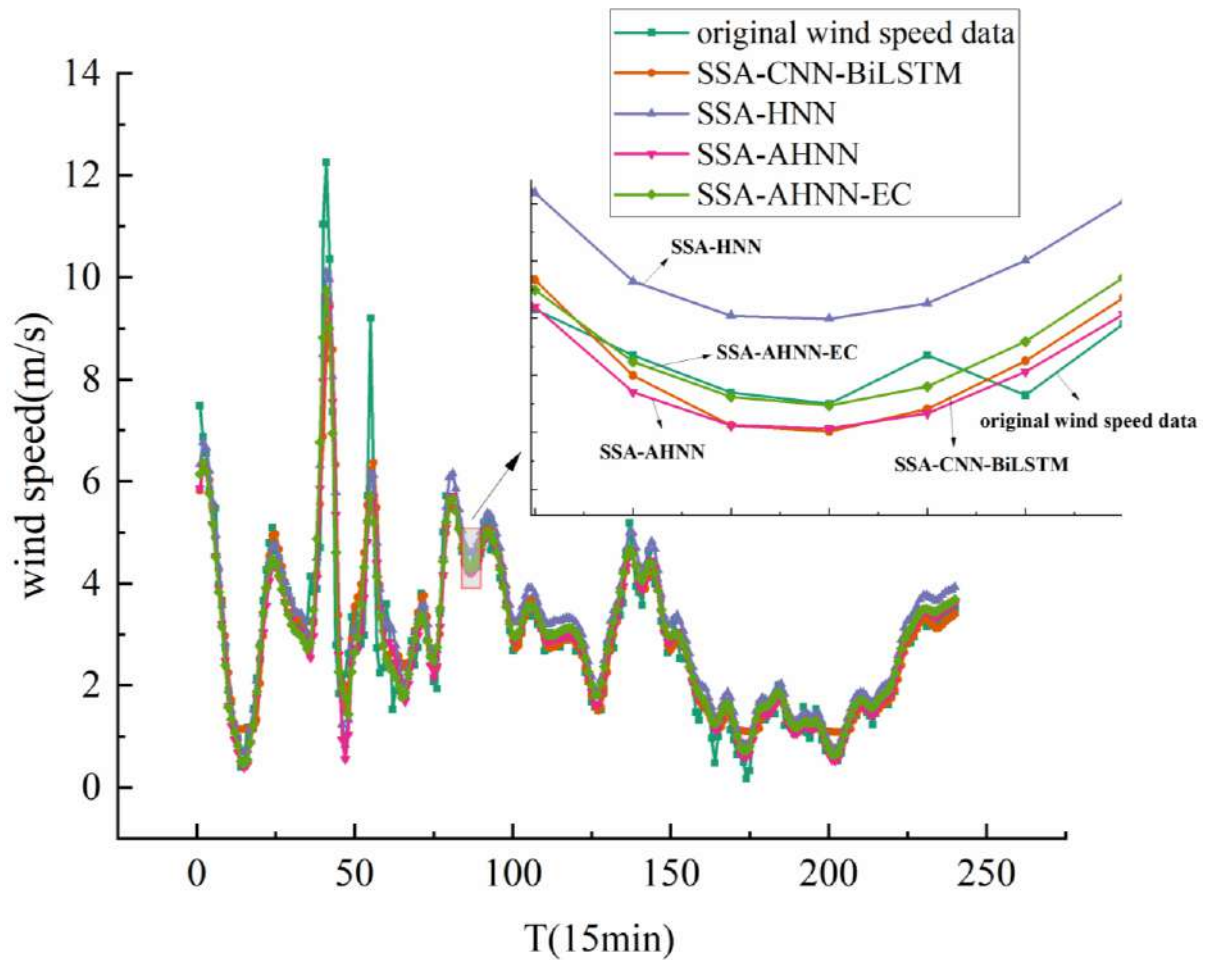Fig. 25. Comparison of ablation experiment results for dataset 1

Fig. 26. Comparison of ablation experiment results for dataset 2

TABLE IV
COMPARISON OF EVALUATION INDEX RESULTS OF ABLATION EXPERIMENTS FOR DATASET 1

Hybrid predictive models	ADAPTIVE WEIGHTING OPTIMIZATION	ERROR CORRECTION	MAE	MSE	MAPE	RMSE	R²
			0.193	0.066	0.05	0.257	0.944
√			0.184	0.061	0.05	0.247	0.948
√	√		0.176	0.059	**0.046**	0.243	0.950
√	√	√	**0.166**	**0.055**	0.05	**0.234**	**0.953**

TABLE V
COMPARISON OF EVALUATION INDEX RESULTS OF ABLATION EXPERIMENTS FOR DATASET 2

Hybrid predictive models	ADAPTIVE WEIGHTING OPTIMIZATION	ERROR CORRECTION	MAE	MSE	MAPE	RMSE	R²
			0.355	0.428	0.133	0.654	0.861
√			0.333	0.362	0.162	0.601	0.882
√	√		0.325	0.316	0.139	0.562	0.897
√	√	√	**0.313**	**0.276**	**0.126**	**0.525**	**0.910**

From Fig. 25 and Table IV, it is evident that the various sub-modules employed in this study have effectively improved the accuracy of wind speed prediction. Taking the single prediction model based on SSA-CNN-BiLSTM in dataset 1 as an example, its MAE, MSE, MAPE, RMSE, and R² are 0.193, 0.066, 0.05, 0.257, and 0.944, respectively. As we continuously added mixed prediction models, adaptive weight optimization, and error correction modules, all the values showed an optimizing trend except for MAPE, which remained stable. The final values were 0.166, 0.055, 0.05, 0.234, and 0.953, respectively.

However, in dataset 2, the single prediction model based on SSA-CNN-BiLSTM is also used as the benchmark. From Figure 26 and Table V, we can conclude that the values of MAE, MSE, MAPE, RMSE and R² are 0.355, 0.428, 0.133, 0.654 and 0.861, also as we continue to increase the hybrid prediction model, adaptive weight optimization and error correction modules, all values show an obvious optimization trend. The final evaluation index values are 0.313, 0.276, 0.126, 0.525 and 0.91 respectively.

Additionally, in order to further demonstrate the model's applicability in predicting short-term wind speeds, we take the MSE results of ten independent runs with the seven models mentioned above as an example to compare the

upward and downward fluctuations of the results of each model, and the fluctuations of the minimum and maximum values of each model run independently for ten times are shown in Table VI.

TABLE VI
COMPARISON OF MSE VALUE FLUCTUATION OF DIFFERENT MODELS

Model	Fluctuations in MSE values
SSA-BP	55.8%
SSA-LSTM	43.5%
SSA-BiLSTM	38.2%
SSA-CNN-BP	42.4%
SSA-CNN-LSTM	35.7%
SSA-CNN-BiLSTM	27.4%
Proposed Method	**12.3%**

From the data comparisons in Table VI, we can draw the conclusion that, when compared to the other six comparison models, the suggested approach is the most stable for predicting short-term wind speeds. The maximum and minimum MSE values of the suggested model have the least difference after ten independent repeats of the tests and the removal of one maximum and one minimum value, which has a very evident benefit over the other models.

Therefore, we may infer that the several sub-modules used in this study were successful in increasing prediction accuracy. It is important to note that in dataset 1, the value of MAPE stayed largely constant, with no significant upward or downward trend. Considering the significant decrease or increase in other evaluation indexes, we believe that this cannot be used to indicate the ineffectiveness of the optimization measures employed. Meanwhile, in dataset 2, with the continuous increase of optimization modules, all evaluation indicators showed an obvious optimization trend, which further illustrates how well the additional approach improved prediction accuracy. Additionally, the suggested approach provides blatant advantages in terms of stability.

V. CONCLUSION

In order to maximize the accuracy and stability of wind speed prediction as much as is practical, a novel prediction model based on an adaptive hybrid neural network model and error correction is proposed in this study. The key technique for preparing data in the model is SSA. As was previously said, using this approach to denoise the data preserves the primary elements of the initial wind speed information while excluding high-frequency noise components, making it simpler to extract characteristics from the data and accomplish accurate wind speed prediction. Experimental results showed that data denoising considerably improved wind speed forecast accuracy. We demonstrate the viability and efficacy of using different models to predict different frequency data after denoising the data, and then we decompose and reconstruct the denoised data using wavelet analysis to obtain low-frequency and high-frequency sub-sequences, which are subsequently used as input for training and prediction in the appropriate model. The effectiveness of this approach is also verified in subsequent comparative experiments. In addition, another improvement

of the method proposed in this paper is to use PSO algorithm to optimize the superposition weights of the prediction results of the two sub-models, instead of simple linear superposition. Finally, for the purpose of to further minimize the inaccuracy of wind speed prediction, this study additionally uses an error correction module that trains and predicts using LSTM and superimposes the final forecast result on the previous prediction results. The dataset for this module is made up of the model's training and testing mistakes.

According to experimental findings, our suggested model for predicting wind speed is more accurate than other single prediction models that were employed as a comparison. Taking SSA-CNN-BiLSTM, the best-performing model among the comparison models, as an example, in dataset 1, our method reduces the MAE, MSE, and RMSE by 16%, 17%, and 9%, respectively, while maintaining stability or slightly improving the MAPE and R^2. And in dataset 2, our method reduces the MAE, MSE, MAPE and RMSE by 12%, 36%, 5%, and 20%. There is also a 5% increase in the value of R^2. In addition, to evaluate the impact of each submodule of the model in improving accuracy, we also designed corresponding ablation experiments. From the comparison results, we can see that all evaluation indicators continue to improve or remain stable, indicating that the improvements we have made are effective and reliable.

As a consequence, the mixed prediction model proposed in this work greatly improves wind speed forecasting accuracy and serves as a helpful benchmark for managing and running wind farms in reality.

REFERENCES

[1] W. Dong, G. Zhao, S. Yüksel, H. Dinçer, and G. G. Ubay, "A novel hybrid decision making approach for the strategic selection of wind energy projects," *Renewable Energy*, vol. 185, pp. 321–337, 2022.

[2] G. Msigwa, J. O. Ighalo, and P.-S. Yap, "Considerations on environmental, economic, and energy impacts of wind energy generation: Projections towards sustainability initiatives," *Science of The Total Environment*, vol. 849, p. 157755, 2022.

[3] J. Han and H. Chang, "Development and Opportunities of Clean Energy in China," *Applied Sciences*, vol. 12, no. 9, Art. no. 9, 2022.

[4] C. Doblinger, K. Surana, D. Li, N. Hultman, and L. D. Anadón, "How do global manufacturing shifts affect long-term clean energy innovation? A study of wind energy suppliers," *Research Policy*, vol. 51, no. 7, p. 104558, 2022.

[5] L. Chang, H. B. Saydaliev, M. S. Meo, and M. Mohsin, "How renewable energy matter for environmental sustainability: Evidence from top-10 wind energy consumer countries of European Union," *Sustainable Energy, Grids and Networks*, vol. 31, p. 100716, 2022.

[6] J. Sun and F. Dong, "Decomposition of carbon emission reduction efficiency and potential for clean energy power: Evidence from 58 countries," *Journal of Cleaner Production*, vol. 363, p. 132312, 2022.

[7] M. Amir, A. K. Prajapati, and S. S. Refaat, "Dynamic Performance Evaluation of Grid-Connected Hybrid Renewable Energy-Based Power Generation for Stability and Power Quality Enhancement in Smart Grid," *Front. Energy Res.*, vol. 10, 2022.

[8] M. T. Rahman, K. N. Hasan, and P. Sokolowski, "Evaluation of wind farm aggregation using probabilistic clustering algorithms for power system stability assessment," *Sustainable Energy, Grids and Networks*, vol. 30, p. 100678, 2022.

[9] B. Xiong, X. Meng, G. Xiong, H. Ma, L. Lou, and Z. Wang, "Multi-branch wind power prediction based on optimized variational mode decomposition," *Energy Reports*, vol. 8, pp. 11181–11191, 2022.

[10] H. T. K. Abdelbadie, A. T. M. Taha, H. M. Hasanien, R. A. Turky, and S. M. Muyeen, "Stability Enhancement of Wind Energy Conversion Systems Based on Optimal Superconducting Magnetic Energy Storage Systems Using the Archimedes Optimization Algorithm," *Processes*, vol. 10, no. 2, Art. no. 2, 2022.

[11] H. Chen, "A comprehensive statistical analysis for residuals of wind speed and direction from numerical weather prediction for wind energy," *Energy Reports*, vol. 8, pp. 618–626, 2022.

[12] G. Yakoub, S. Mathew, and J. Leal, "Intelligent estimation of wind farm performance with direct and indirect 'point' forecasting approaches integrating several NWP models," *Energy*, vol. 263, p. 125893, 2023.

[13] C. Liu *et al.*, "Numerical weather prediction enhanced wind power forecasting: Rank ensemble and probabilistic fluctuation awareness," *Applied Energy*, vol. 313, p. 118769, 2022.

[14] S. Sheoran and S. Pasari, "Efficacy and application of the window-sliding ARIMA for daily and weekly wind speed forecasting," *Journal of Renewable and Sustainable Energy*, vol. 14, no. 5, p. 053305, Sep. 2022.

[15] J. Wang and J. Wang, "Short-term Wind Speed Forecast Using ARIMA Based on EEMD Decomposition," *J. Phys.: Conf. Ser.*, vol. 2450, no. 1, p. 012020, 2023.

[16] I. Tyass, A. Bellat, A. Raihani, K. Mansouri, and T. Khalili, "Wind Speed Prediction Based on Seasonal ARIMA model," *E3S Web Conf.*, vol. 336, p. 00034, 2022.

[17] M. Monfared, H. Rastegar, and H. M. Kojabadi, "A new strategy for wind speed forecasting using artificial intelligent methods," *Renewable Energy*, vol. 34, no. 3, pp. 845–848, 2009.

[18] T. Kaur, S. Kumar, and R. Segal, "Application of artificial neural network for short term wind speed forecasting," in *2016 Biennial International Conference on Power and Energy Systems: Towards Sustainable Energy (PESTSE)*, pp. 1–5, 2016.

[19] P. Ramasamy, S. S. Chandel, and A. K. Yadav, "Wind speed prediction in the mountainous region of India using an artificial neural network model," *Renewable Energy*, vol. 80, pp. 338–347, 2015.

[20] J. Zhou, J. Shi, and G. Li, "Fine tuning support vector machines for short-term wind speed forecasting," *Energy Conversion and Management*, vol. 52, no. 4, pp. 1990–1998, 2011.

[21] S. Salcedo-Sanz, E. G. Ortiz-Garcı́a, Á. M. Pérez-Bellido, A. Portilla-Figueras, and L. Prieto, "Short term wind speed prediction based on evolutionary support vector regression algorithms," *Expert Systems with Applications*, vol. 38, no. 4, pp. 4052–4057, 2011.

[22] M. A. Mohandes, T. O. Halawani, S. Rehman, and A. A. Hussain, "Support vector machines for wind speed prediction," *Renewable Energy*, vol. 29, no. 6, pp. 939–947, 2004.

[23] Y. Chen *et al.*, "2-D regional short-term wind speed forecast based on CNN-LSTM deep learning model," *Energy Conversion and Management*, vol. 244, p. 114451, 2021.

[24] J. Duan *et al.*, "A combined short-term wind speed forecasting model based on CNN–RNN and linear regression optimization considering error," *Renewable Energy*, vol. 200, pp. 788–808, 2022.

[25] Y. Zhang, Y. Zhao, X. Shen, and J. Zhang, "A comprehensive wind speed prediction system based on Monte Carlo and artificial intelligence algorithms," *Applied Energy*, vol. 305, p. 117815, 2022.

[26] M. Neshat *et al.*, "A deep learning-based evolutionary model for short-term wind speed forecasting: A case study of the Lillgrund offshore wind farm," *Energy Conversion and Management*, vol. 236, p. 114002, 2021.

[27] M. Lv, J. Wang, X. Niu, and H. Lu, "A newly combination model based on data denoising strategy and advanced optimization algorithm for short-term wind speed prediction," *J Ambient Intell Human Comput*, 2022.

[28] J. Li, Z. Song, X. Wang, Y. Wang, and Y. Jia, "A novel offshore wind farm typhoon wind speed prediction model based on PSO–Bi-LSTM improved by VMD," *Energy*, vol. 251, p. 123848, 2022.

[29] Z. Tian, "Modes decomposition forecasting approach for ultra-short-term wind speed," *Applied Soft Computing*, vol. 105, p. 107303, 2021.

[30] L. Ji, C. Fu, Z. Ju, Y. Shi, S. Wu, and L. Tao, "Short-Term Canyon Wind Speed Prediction Based on CNN—GRU Transfer Learning," *Atmosphere*, vol. 13, no. 5, Art. no. 5, 2022.

[31] G. Chen, B. Tang, X. Zeng, P. Zhou, P. Kang, and H. Long, "Short-term wind speed forecasting based on long short-term memory and improved BP neural network," *International Journal of Electrical Power & Energy Systems*, vol. 134, p. 107365, 2022.

[32] L. Ø. Bentsen, N. D. Warakagoda, R. Stenbro, and P. Engelstad, "Spatio-temporal wind speed forecasting using graph networks and novel Transformer architectures," *Applied Energy*, vol. 333, p. 120565, 2023.

[33] Q. Yang, C. Deng, and X. Chang, "Ultra-short-term / short-term wind speed prediction based on improved singular spectrum analysis," *Renewable Energy*, vol. 184, pp. 36–44, 2022.

[34] A. Lawal, S. Rehman, L. M. Alhems, and Md. M. Alam, "Wind Speed Prediction Using Hybrid 1D CNN and BLSTM Network," *IEEE Access*, vol. 9, pp. 156672–156679, 2021.

[35] G. Chen, M. Zhu, J. Huang, Y. Fu, X. Xie, and H. Long, "Short-term Wind Speed Prediction with Master-slave Performance Based on CNN- LSTM and Improved POABP," *Engineering Letters*, vol. 31, no. 2, pp. 848-861, 2023

[36] G. Chen, P. Qiu, X. Hu, F. Long, and H. Long, "Research of Short-Term Wind Speed Forecasting Based on the Hybrid Model of Optimized Quadratic Decomposition and Improved Monarch Butterfly," *Engineering Letters*, vol. 30, no. 1, pp. 73-90, 2022.