IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 22 April 2024, accepted 19 May 2024, date of publication 22 May 2024, date of current version 31 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3404268

RESEARCH ARTICLE

Wind Speed Multi-Step Forecasting Using High-Resolution and Multi-Objective Ensemble Model With Error Correction

SHENGCAI ZHANG[1,2], CHANGSHENG ZHU[1], AND XIUTING GUO[1], (Member, IEEE)
[1]School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China
[2]School of Cyber Security, Gansu University of Political Science and Law, Lanzhou 730070, China

Corresponding author: Shengcai Zhang (zsc6731@gsupl.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51661020, in part by the Foundation Project of Gansu Provincial Department of Education under Grant 2020C-29 and Grant 2022CYZC-57, and in part by the University-Level Innovative Research Team of Gansu University of Political Science and Law.

ABSTRACT Accurate wind forecasting is essential for enhancing the stability and efficiency of wind power systems. However, the nonlinear and unstable characteristics of wind pose significant challenges to achieving high-performance predictions in this domain. To address the issues of insufficient wind speed data decomposition and the limited forecasting time range for high-resolution data in hybrid prediction models, this study proposes a short-term, multistep wind speed prediction model based on high-resolution data and multi-objective integration with error correction. Initially, the White Shark Optimizer (WSO) algorithm is employed to determine the optimal decomposition parameters for Variational Mode Decomposition (VMD), decomposing the original high-resolution wind speed data. Secondly, the decomposed data are processed using Convolutional Neural Networks (CNN) with three different sizes of convolutional kernels to capture multi-scale features, which are then fed into a Gated Recurrent Unit (GRU) model for three-step forecasting. Finally, three-step prediction results for features across different scales are input into an ensemble module, and the Non-Dominated Sorting Whale Optimization Algorithm (NSWOA) is utilized to tune the weight parameters of the ensemble and error correction (EC) module. Experimental results on different datasets indicate that the proposed approach not only leverages the detailed information from high-resolution data but also addresses the issue of low accuracy in multistep forecasting. Moreover, the hybrid model, which includes a multi-objective optimization-based integration module and error correction, not only provides highly accurate multistep forecasts but also ensures greater model stability.

INDEX TERMS Wind speed forecasting, high-resolution, multi-objective ensemble, error correction.

I. INTRODUCTION

A. BACKGROUND

The intensifying energy crisis and escalating environmental pollution have elevated the importance of wind power as a renewable and clean energy alternative, gaining increasing global attention. The Global Wind Energy Council (GWEC) reports a marked increase in global capacity, totaling 906 Gigawatts (GW), a year-on-year growth of 9%. The year 2023 is poised to be groundbreaking, as it is anticipated to be the first year where over 100 GW of new capacity will be added globally. GWEC Market Intelligence anticipates a remarkable 15% growth rate year-on-year. Their projections also suggest an astonishing addition of 1,221 GW of new capacity from 2023 to 2030 [1]. According to Zhang et al. [2], a mere 10% increase in prediction accuracy can lead to a substantial 30% increase in wind power generation. Nonetheless, achieving precise wind speed forecasts is a formidable task due to wind's inherent randomness, intermittence, and variability [3]. Thus, making accurate predictions is crucial for stable power system operation and competitiveness in the power market [4].

The associate editor coordinating the review of this manuscript and approving it for publication was Sajid Ali.

In recent decades, a variety of forecasting techniques for enhancing wind speed prediction have been introduced by researchers, broadly categorizing them into four groups: physical models, statistical models, artificial intelligence models, and combined models. Physical models, particularly relying on Numerical Weather Prediction (NWP), forecast wind speed by factoring in elements such as wind speed, temperature, humidity, terrain features, and barriers. Nonetheless, the application of physical models demands extensive resources and is more applicable for medium to long-term predictions [5]. On the other hand, statistical models, including autoregressive, moving average, and autoregressive moving average [6], utilize the timing and autocorrelation in wind speed data to develop their forecasts. These models, while simpler, are generally limited to handling wind speed data that exhibit linear characteristics.

Recently, the realm of wind speed prediction has witnessed a surge in the popularity of various machine learning models due to the rapid advancements in machine learning and artificial intelligence. Notably, support vector regression [7], long short-term memory network (LSTM) [8], [9], the bidirectional LSTM [10], the convolutional neural network (CNN) [11], and the gated recurrent unit (GRU) [12] have gained significant traction. These models offer the ability to capture non-linear features within time series data, leading to improved forecast accuracy. Through ongoing learning and training with the provided data, these models can adeptly replicate complex non-linear relationships between the input and output layers. This inherent self-organization and adaptive learning give them a competitive edge over traditional statistical methods. However, despite their advantages, these methods also suffer from drawbacks, such as susceptibility to local optima and the risk of overfitting [13]. To further enhance the predictive capabilities of these models, researchers have proposed the integration of intelligent algorithms into a hybrid prediction model [14]. Zhang et al. [15] introduced an advanced model that combines noise-processing techniques, optimization methods, statistical methods, and deep learning. These studies have shown that hybrid models, which utilize diverse algorithms, are markedly more accurate than individual models when it comes to prediction capabilities.

B. RELATED WORK

The wind farm has the capability to record wind speed in real-time with high-temporal resolutions, including intervals of 1-minute [16], 5-minute [17], 10-minute [18], 15-minute [19], among others. For forecasting over medium- to long-term periods, however, data of lower resolution become necessary. The sampling of such high-resolution data anew could substantially escalate the construction expenses, necessitating the investigation of resolution transformation methods [20]. A common approach in the realm of wind velocity forecasting is the use of the averaging technique, a conventional method for downscaling resolution [21]. Through consistent interval averaging, data with high resolution can be transformed into a format with lower resolution. This technique maintains the overall wind speed trends while simplifying the data resolution. To broaden the scope of forecasts, several researchers have directly downscaled high-resolution data to a lower resolution. For instance, Demolli et al. utilized a method to average data recorded every hour to produce daily wind speed figures [22], while Alonzo et al. developed daily wind speed data by averaging over every six hours [23]. Nevertheless, this method of averaging can result in the loss of detailed fluctuation data found in the high-resolution data, which can adversely affect the accuracy of forecasts. To enhance the prediction accuracy, it is crucial to make rational use of the high-resolution data. Given the rich array of information offered by high-resolution data, it is predicted that a high-resolution ensemble model could significantly boost forecasting capabilities. Consequently, it enables the exploitation of the extensive and critical wind speed data from the high-resolution datasets, potentially elevating forecast precision. Moreover, Liu et al. [24] proposed an ensemble model incorporating ten distinct moment vanishings. Peng et al. [25] introduced a regularized extreme learning machine ensemble model based on negative correlation learning. Wang et al. [26] developed an ensemble model employing LSTM, Random Forest, and Gaussian Process Regression to predict gust wind speeds. Memarzadeh et al. [27] proposed a hybrid model for short-term wind speed forecasting based on the crow search algorithm, wavelet transform (WT), feature selection, and LSTM. Therefore, delving into the high-resolution ensemble model could serve as a key step in addressing the existing research void, significantly improving the performance of wind speed prediction forecasts.

Additionally, optimization algorithms can be employed to integrate multiple modules. These algorithms are tailored to swiftly identify the best parameters for specific prediction models. Current research is increasingly focusing on the refinement of prediction models' accuracy through the adoption of optimization algorithms, which draw inspiration from diverse biological phenomena [28]. In their studies, Wu et al. [29] have demonstrated the application of multi-objective optimization algorithms in improving ELM parameter optimization. Similarly, Li et al. [30] advocated the utilization of synchronous optimization strategies to improve the performance of decomposition algorithms. Furthermore, Yang et al. [31] suggested utilizing the Water Cycle Algorithm (WCA) for calculating ensemble weights. Hence, several meta-heuristic optimization algorithms have been developed, including Particle Swarm Optimization [32], Grey Wolf Optimizer (GWO) [33], Multi-Objective Modified Seagull Optimization Algorithm [34], Multi-objective Multiverse Optimizer, (MOMVO) [35], Multi-Objective Cuckoo Search (MOCS) [36], and Multi-Objective Bat Algorithm (MOBA) [37]. Zhang et al. [38] successfully introduced a two-stage wind speed forecasting model, which synergizes

the strengths of VMD technology, improved multi-objective optimization algorithm, error correction, and nonlinear ensemble methods. The model demonstrates superior accuracy and stability in both one-step and multi-step predictions compared to other methods. Each of these algorithms has been effectively utilized in calculating the optimal ensemble weight coefficients for each base model within the realm of wind power forecasting. Currently, optimization algorithms are widely integrated into ensemble models. By assigning weights, multiple factors or methods can be incorporated into the model, thus achieving a more comprehensive model performance. Compared to ensemble methods like Adaboost, optimized ensembles provide greater flexibility. However, the improper setting of the objective function may hinder the formation of the Pareto front in multi-objective optimization algorithms, leading to failure. Furthermore, the ensemble's performance can be significantly influenced by the internal parameter settings of the optimization algorithms.

Despite the progress by these models, the challenge of achieving reliable wind power predictions with deep learning models is still formidable, primarily due to the non-stationary nature of meteorological factors. Various decomposition models are commonly used in research, including WT, empirical mode decomposition, VMD and similar models [39], [40], [41]. However, the performance of WT heavily relies on the specific function and threshold selections. Both empirical mode decomposition and its variant, ensemble empirical mode decomposition, grapple with issues pertaining to endpoint effects and are influenced by the absence of a robust mathematical foundation. Conversely, VMD effectively reduces data fluctuations and suppresses noise, making it the most widely recognized data decomposition technique [42]. The versatility of the VMD algorithm allows it to be applied to wind sequences in any geographical region. Instances of successful application include Rayi et al.'s deployment of VMD for wind power data analysis in regions like Sotavento, Spain, and in Wyoming and California, USA [43]. A combined model integrating VMD, Convolutional LSTM, and error analysis is developed by Sun et al. [44]. This model decomposes wind power signals, utilizes ConvLSTM for initial forecasting, and refines predictions by analyzing and integrating error trends, resulting in a comprehensive and accurate final forecast. Additionally, Duan et al. applied VMD for local feature extraction in wind power data from a wind farm in Shanxi Dingbian, China [45]. The effectiveness of VMD decomposition relies on the choice of decomposition parameters, which are typically determined through empirical knowledge. This often leads to suboptimal decomposition outcomes, thereby failing to meet expectations.

The significant application value is rooted in the efficient secondary development of the predictable elements in preliminary forecasts. The residual error correction technique enhances the model's forecasting ability by a significant margin [46]. In Ding et al.'s experiments, the Bidirectional GRU model was developed to address the residual errors in NWP forecasts [47]. Similarly, Liu et al. introduced the EWT-ORELM model, focusing on reducing residual errors [48]. The decomposition and forecasting framework enable the achievement of satisfactory correction results. However, the effectiveness of one-time error correction falls short of actual application requirements. The erratic and sporadic nature of wind speed time series data complicates the task of residual correction. A rudimentary error correction strategy might introduce unforeseen elements into the model, thus diminishing its predictive accuracy. Hence, addressing this issue necessitates the proposal of an adaptive multi-error correction technique [49]. Balancing error correction without overcompensating or underperforming is crucial, underlining the significance of ongoing research in this area.

Furthermore, the current error correction models predominantly utilize machine learning methods for predicting residual errors. Nonetheless, the inherent randomness and unique characteristics of error data present challenges in training models with strong generalization abilities. Advanced models like the decomposition-based error correction method [50] and the multi-loop error correction method [51] face difficulties related to ease of use, operational effectiveness, and the risk of excessive correction. Hence, it is imperative to undertake research on advanced error correction models that exhibit exceptional performance to improve the precision of wind speed forecasts.

C. MOTIVATION AND CONTRIBUTION

The above literature review reveals several existing gaps in wind power forecasting. Firstly, current artificial intelligence methods and decomposition algorithms for wind power forecasting struggle with finding appropriate parameters. This often relies on a time-consuming trial-and-error approach based on experience, which adversely affects the accuracy of the forecasts. The potential of an ensemble model that integrates high-resolution wind speed has not been totally explored. Currently, high-resolution data is simply averaged into low-resolution formats, causing a loss of crucial fluctuation details. Lastly, there are relatively few methods of using error correction to further improve performance. To address these gaps, a novel high-resolution and multi-objective ensemble model is suggested, consisting of three stages. In the first stage, the signal pre-processing module eliminates interference information from the original wind speed through decomposition techniques. The VMD technique is utilizing the WSO algorithm to optimize VMD parameters. The second phase, the foundational multi-step forecasts are constructed using high-resolution datasets, integrating models through the utilization of multi-objective optimization algorithms. The third phase employs ARIMA for adaptive multiple error corrections.

In this research, we introduce a new high-resolution, multi-objective ensemble model that incorporates error correction. This model leverages the fluctuation details present in high-resolution big data, enhancing the precision and stability

of the forecasts. The key contributions of this model are outlined as follows:

(1) Addressing the issues of pattern mixing and incomplete decomposition, the model employs the WSO technique to autonomously determine the most effective decomposition parameters for VMD. This approach effectively diminishes signal loss during the decomposition phase and enhances the overall decomposition performance of VMD.

(2) To address the issue of limited multi-step prediction horizons for high-resolution data, this paper proposes an integrated predictive model for high-resolution datasets. Initially, decomposed high-resolution data are fed into a CNN model, which captures multi-scale features using three different sizes of convolutional kernels. Subsequently, the GRU model performs a three-step prediction on the extracted features. Thus, this model not only leverages the detailed information of high-resolution data effectively, but also resolves the challenge of its limited prediction horizon.

(3) To enhance the hybrid model's performance in optimizing weight parameters and error correction, this paper establishes a hybrid model that utilizes a multi-objective optimization algorithm to optimize parameters for both the ensemble module and error correction, significantly improving the accuracy of multi-step predictions.

(4) A detailed investigation is conducted to verify the effectiveness of the proposed hybrid model. Two separate wind speed time series are utilized for forecasting purposes, and the results of this model are benchmarked against other existing models. The empirical findings demonstrate that the application of high-resolution data significantly boosts the model's predictive capabilities. Particularly for multi-step prediction, the improvement rate is higher.

D. ORGANIZATION

The structure of the remainder of this paper is organized in the following manner: Section II offers a comprehensive introduction to the proposed model and associated methods. Section III presents an in-depth examination of the experimental outcomes from both the proposed model and those for comparison, complete with extensive analysis and validation. Finally, Section IV highlights the principal conclusions of this study and explores potential directions for future research in the domain of wind speed prediction.

II. METHODOLOGY

A. FLOWCHART OF THE PROPOSED MODEL

A novel composite forecasting model, as illustrated in Figure 1, has been proposed. This model amalgamates optimized modal decomposition, CNN-GRU forecasting, multi-objective optimization ensemble, and error correction (EC), succinctly termed as WSO-VMD-CNN-GRU-NSWOA-EC. The model is structured into three phases: In the first phase, the WSO algorithm is utilized to automatically determine the optimal parameters for VMD, leading to the division of the original wind speed into multiple Intrinsic Mode Functions (IMF). The second phase begins by inputting the decomposed data, now of higher resolution, into a CNN model. To capture features at diverse scales, three distinct convolutional kernels are applied. The extracted features are then processed through a GRU model for initial 3-step forecasting. Subsequently, the results from this tripartite multi-step forecasting are integrated into a multi-objective ensemble module, where the NSWOA algorithm is used to establish the most effective weights. The third phase involves the application of NSWOA-optimized ARIMA for error correction. The final forecast results are derived by an aggregate calculation that combines error correction with the integrated forecasting outcomes.

B. VMD

Variational Mode Decomposition (VMD), as a method in signal processing, estimates in adaptively determining the number of modal decompositions for a sequence, depending on current conditions [50]. VMD distinguishes itself by identifying the optimal central frequency and maintaining a limited bandwidth for each mode throughout the analysis and resolution process, thereby effectively isolating intrinsic mode elements and segmenting the frequency domain. When compared with empirical mode and wavelet decompositions, VMD showcases enhanced capabilities in reconstructing the original signal and exhibits stronger resistance to noise. This technique enables the decomposition of a signal into K distinct frequency bands and relatively stable sub-signals as required. At the core of the algorithm is the pursuit of an optimal solution to a variational problem, starting with a presumption that the original signal S is broken down into K components μ. This is predicated on ensuring that the decomposition comprises mode components with a finite bandwidth and central frequency. The goal is to balance the sum of all estimated bandwidths with that of the original signal while minimizing the bandwidths' sum for each mode. This is achieved by constructing a constrained variational formula, depicted in Equation (1) below:

$$\begin{cases} \min\limits_{\{u_k\},\{\omega_k\}} \left\{ \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \cdot u_k(t) \right] e^{-j\omega k^t} \right\|_2^2 \right\} \\ s.t. \sum_{k=1}^{K} u_k = S \end{cases} \tag{1}$$

where the kth mode component is denoted as u_k, while the central frequency of the kth ingredient is represented by $\{\omega_k\}$ r. Additionally, the Dirac distribution is symbolized as $\delta(t)$.

To address the constrained variational expression, a penalty parameter α and a Lagrange multiplier operator λ are incorporated. By introducing these elements, the initial constrained variational problem is transformed into an unconstrained one. Consequently, an augmented Lagrange expression (2) is obtained.

$$L\left[\{u_k\,(t)\}, \{\omega_k\}, \lambda\,(t)\right]$$

$$= \alpha \sum_{k=1} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \cdot u_t(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| S(t) - \sum_{k=1}^{K} u_k(t) \right\|_2^2 + \left\langle \lambda(t), S(t) - \sum_{k=1}^{K} u_k(t) \right\rangle \quad (2)$$

To achieve the optimal solution, the parameters u_1, ω_2, λ^1, and n are initialized. The starting value of n is established at 0. Subsequently, a looping procedure is initiated where n is incremented in each cycle. During each iteration, the values of u_1, ω_2, and λ^1 are updated accordingly.

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \hat{\lambda}(\omega)/2}{1 + 2\alpha(\omega - \omega_k)^2} \quad (3)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega \left| \hat{u}_k(\omega) \right|^2}{\int_0^\infty \left| \hat{u}_k(\omega) \right|^2} \quad (4)$$

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau \left(\hat{f}(\omega) - \sum_{k=1}^{K} \hat{u}_k^{n+1}(\omega) \right) \quad (5)$$

C. CNN

The convolutional neural network (CNN) model, according to reference [52], is a type of deep learning architecture that uses higher layers to extract local attributes from initial and lower layers, effectively synthesizing these into more intricate features.

Within the convolutional layers of a CNN, several convolutional units are tasked with isolating specific features from the input. This convolution operation is mathematically expressed in Equation (6).

$$y_{i,j}^{\text{conv}} = \sum_{i=1}^{k} \omega_{i,j}^{b} \times y_{l-1,i} + b_j^l \quad (6)$$

where $y_{l,j}^{\text{conv}}$ represents the output of channel j in convolutional layer l, $\omega_{i,j}^{b}$ denotes the weight parameter, and b_j^l signifies its bias value.

Following the convolutional layer, a feature vector of substantial dimensionality is typically produced. This vector undergoes segmentation into different sections, from which either the maximum or average values are computed, leading to the generation of new features with a lower dimensionality.

In CNN, the activation function commonly employed is Rectified Linear Units (ReLU) due to its rapid convergence and the simplicity of calculating its gradient. The initial formulation of ReLU can be described as:

$$y_{l,j}^{\text{Relu}} = f(y_{l,j}^{\text{conv}}) \quad (7)$$

where $y_{l,i}^{Relu}$ represents the output of channel j in the convolution layer l.

D. GRU

The GRU, as proposed by Cho et al. [53], is an advanced version of LSTM and a type of recurrent neural network (RNN). It is particularly adept at handling tasks involving time series data and natural language processing. When compared with traditional RNNs, GRUs show superior modeling prowess and improved efficiency. The GRU structure is simpler than LSTM, making it easier to implement. Figure 2 illustrates the structure of the GRU network.

Although GRU networks and standard RNNs exhibit certain similarities in their input-output structures, GRUs are distinct in that they employ only two gates: the update gate and the reset gate. These gates regulate the flow of information, enabling GRUs to address the vanishing gradient problem of standard RNNs more effectively. GRUs involve an input x_t at time t and a hidden state h_{t-1} from time t-1, which contains information from previous nodes. The output is the hidden state h_t at time t. GRUs utilize both the previous state h_{t-1} and the current input x_t to obtain two gated states, enabling efficient information processing.

In the GRU architecture, the reset gate is crucial for adjusting the mix of the current input with the past hidden state in the present time step. It employs a sigmoid function to merge the past hidden state and the present input, resulting in a value between 0 and 1. This value signifies the degree to which data from the preceding time step is maintained. When the reset gate's output nears 0, it neglects the previous hidden state, facilitating the autonomous processing of the current input. On the flip side, if the reset gate's output verges on 1, it fully incorporates the information from the last time step. The operational mechanism of the reset gate is outlined in Equation (8).

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \quad (8)$$

In the GRU framework, another key element is the update gate, designated as z_t. The update gate is used to balance the previous hidden state and the current input, as shown in Equation (9). An output near 0 from the update gate indicates less updating of the current timestep's hidden state, while an output close to 1 indicates more updating of the current timestep's hidden state.

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad (9)$$

The candidate hidden state of the GRU is computed from the current input combined with the previous hidden state and the reset gate, as outlined in Equation (10).

$$\tilde{h}_t = tanh(x_t W_{xh} + (r_t \odot h_{t-1}) W_{hh} + b_h) \quad (10)$$

where, W_r, W_z, W_{xh} and W_{hh} represent weight vectors, while b_r, b_z and b_h denote a bias vector. σ represents the sigmoid activation function, *tanh* represents the hyperbolic tangent activation function, $\odot$ represents element-wise multiplication.

The current hidden state of the GRU is determined by the previous hidden state h_{t-1}, the candidate hidden state $\tilde{h}_t$, and the output z_t of the update gate.

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (11)$$

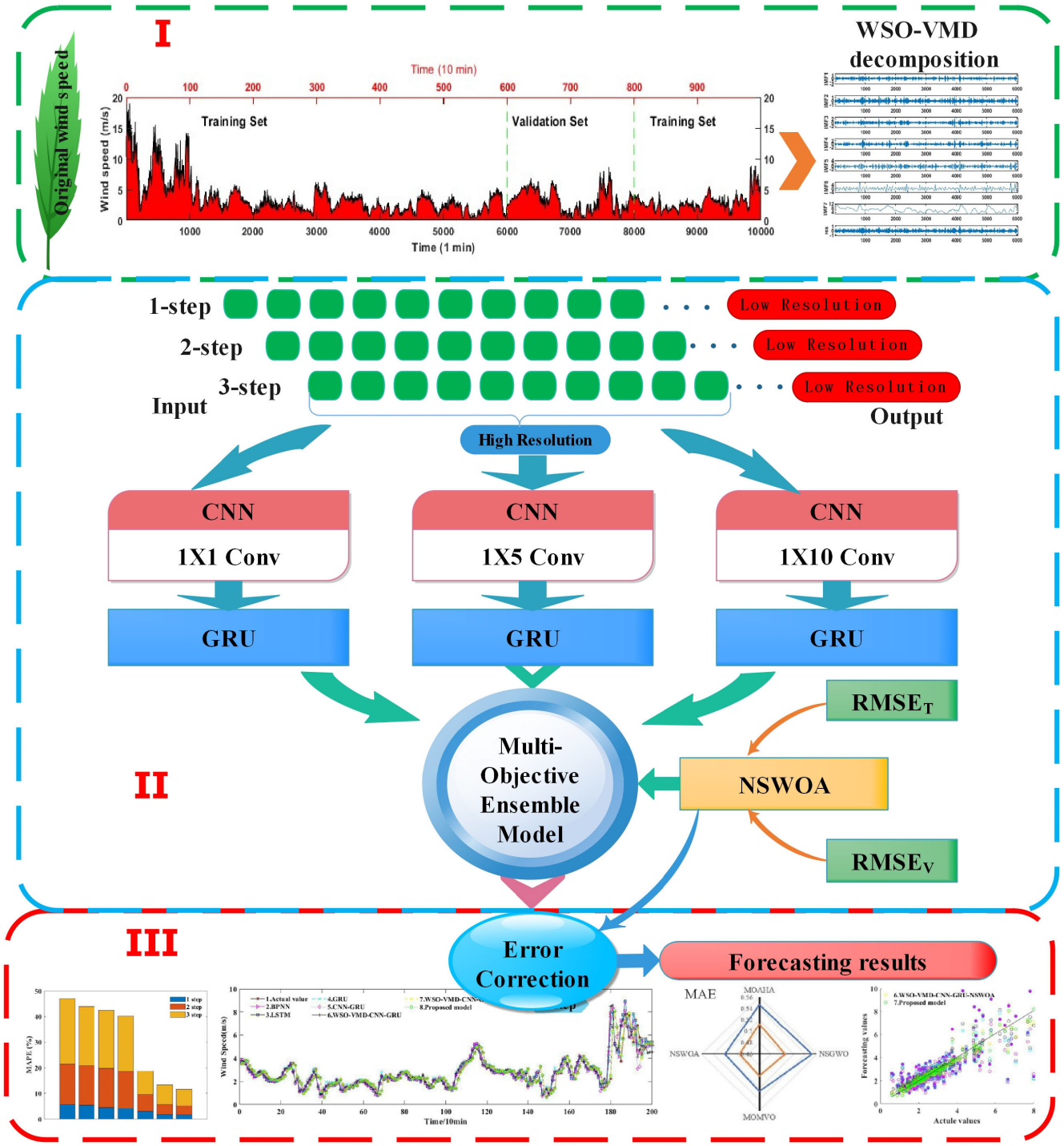

FIGURE 1. Flow chart of the proposed model.

E. WSO

The White Shark Optimizer (WSO) algorithm, introduced by Braik et al. [54] in 2022, draws its inspiration from the predatory tactics of great white sharks in the ocean depths. This metaheuristic algorithm is based on a population model and simulates the swimming prowess, as well as the advanced smell and hearing capabilities, of these formidable sea creatures. The WSO algorithm mathematically models the entire hunting strategy of the white shark. It evaluates the shark's prey acquisition through three distinct behavioral patterns. The first involves detecting prey via the waves and water pressure changes it causes, with the shark leveraging its auditory and olfactory senses to move in an undulating manner towards the prey. In the second pattern, the shark

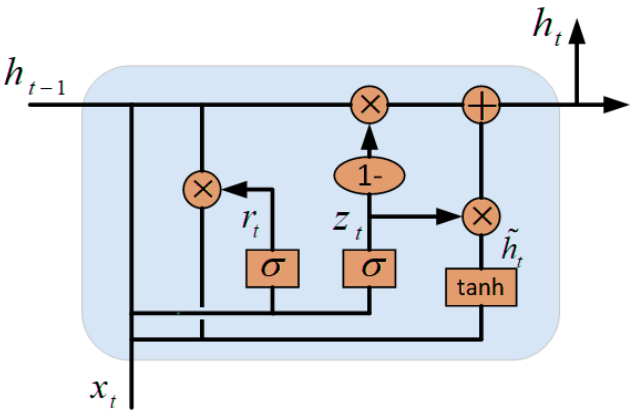

FIGURE 2. GRU structure.

conducts a random search within its marine hunting grounds to locate the most suitable prey. The third pattern features schooling behavior, where the shark positions itself close to another shark that is nearest to the ideal prey, continuing the hunt in the vicinity. If no prey is detected using these methods, the shark alters its position to discover an optimal solution. The WSO optimization process encompasses five phases: initialization, determining movement speed, progressing towards the best prey, identifying the optimal shark, and exhibiting schooling behavior. Auditory cues represent exploration, while olfactory cues denote exploitation in the WSO algorithm, with six control parameters managing the shift between these two states.

F. NSWOA

In 2016, Mirjalili Seyedali and his team introduced the Whale Optimization Algorithm, a nature-inspired stochastic population-based algorithm with a sole objective [55]. The concept of this algorithm is based on the unique hunting strategy of humpback whales, which includes the formation of spiral bubbles to trap prey, following these bubbles as they move upwards in the water [56]. The evolved version, known as the Non-dominated Sorting Whale Optimization Algorithm (NSWOA), incorporates a leader selection process. This process involves selecting solutions from an archive and forming links among them. The process begins by calculating the crowding distance between each archived solution, then assessing the number of solutions in close proximity. The best solution is then identified from the Pareto-optimal set using a crowding distance approach. This approach focuses on solution convergence and utilizes a bubble-net hunting technique, directing the humpback whale model towards the dominated region in the multi-objective search space [57].

Figure 3 presents a detailed flowchart depicting the NSWOA process. In the flow chart depicting NSWOA, the system executes two if-else statements to achieve the desired goal. When a randomly generated number, $p \geq 0.5$, the search agent's position is modified according to Equation (12). Conversely, if $p < 0.5$ and $|\vec{A}| < 1$, then the update is carried out using Equations (13) and (14). However, in situations where $p < 0.5$ and $|\vec{A}| \geq 1$, a random search agent from the current population is chosen, represented by the position vector $\vec{X}_r$. For this selected search agent, Equation (15) is used to calculate the distance from the prey. Subsequently, Equation (16) is employed to update the position of the search agent.

$$\vec{X}(t+1) = \vec{D} \cdot e^{\mathrm{hl}} \cdot \cos(2\pi l) + \vec{X}^*(t) \tag{12}$$

where, b represents the constant that shapes the logarithmic spiral path of the whales., l is a variable randomly assigned a value from the range of $[-1,1]$, and $\vec{D} = \left|\vec{X}^*(t) - \vec{X}(t)\right|$ signifies the distance between the prey's position $(\vec{x^*})$ and the search agent's position $(\vec{X})$.

$$\vec{D} = \left|\vec{C} \cdot \vec{X}^*(t) - \vec{X}(t)\right| \tag{13}$$

$$\vec{X}(t+1) = \vec{X}^*(t) - \vec{A} \cdot \vec{D} \tag{14}$$

Here $\vec{X}$ denotes the vector indicating position. t stands for the iteration count, while $\vec{A}$ and $\vec{C}$ are coefficient vectors that are continually updated to reflect the most recent position of the prey relative to the search agent.

$$\vec{D} = \left|\vec{C} \cdot \vec{X}_r - \vec{X}\right| \tag{15}$$

$$\vec{X}(t+1) = \vec{X}_r - \vec{A} \cdot \vec{D} \tag{16}$$

III. EXPERIMENT AND ANALYSIS

A. DATASET DESCRIPTION

For evaluating the efficacy of the suggested model, two distinct wind speed datasets were chosen, each from different locations and featuring varying resolutions. The initial dataset originates from the National Renewable Energy Laboratory Wind Technology Center and is available to the public. The tower is located at 39° 54' 38.34'' N and 105° 14' 5.28'' W, with its base at an elevation of 1855 meters above mean sea level. The data measurement height is 80 meters, with value resolution of 1 minute. In this paper, 10,000 records from December 2020 are utilized, denoted as dataset A. The second dataset originates from a wind farm in Jiuquan City, Gansu Province, featuring a data measurement height of 70 meters and a resolution of 5 minutes. This study utilizes 6,000 records from April to September 2019, referred to as dataset B. The data from dataset A were mapped to 10-minute intervals, and the data from dataset B were mapped to 30-minute intervals, with each used for one-, two-, and three-step short-term wind speed forecasting using five sets of data. Each dataset was segmented into three parts: 60% for training, 20% for validation, and 20% for testing. Figure 4 illustrates the variation curves of both high-resolution and low-resolution data. Moreover, Table 1 details the statistical attributes of these datasets, highlighting the distinct statistical properties between the two, with the high-resolution dataset showing greater standard deviation and maximum values in comparison to the low-resolution dataset.

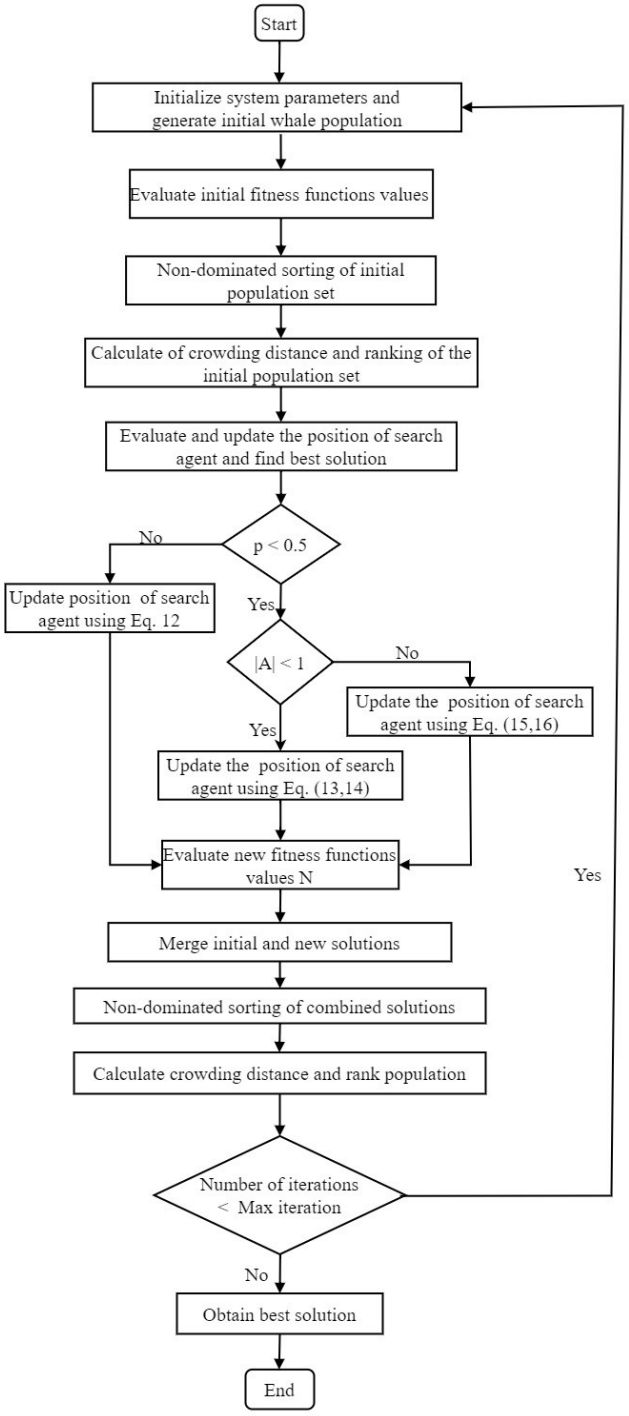

FIGURE 3. Flow chart of NSWOA.

B. EVALUATION METRICS

To comprehensively demonstrate the performance of the proposed model, four widely used metrics were utilized in the case studies. These metrics include the mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and the R-squared (R^2) score. The improvement rate of the evaluation metrics is calculated using Equation (21), which represents the improvement rate of the proposed model relative to the comparative models. For the MAE, RMSE, and MAPE metrics, a greater reduction indicates a higher improvement rate, whereas for the R^2 metric, a greater increase signifies a higher improvement rate. The tests were carried out on a system equipped with an Intel Core i7-10700 CPU and 32 GB of RAM. In order to reduce the impact of random variables, each experimental model was run 20 times, and their average values were calculated. The parameter settings for the proposed model and the comparison models are shown in Appendix Table 13, while other parameters are set to their default values.

TABLE 1. The statistical information.

Dataset		Min	Max	Mean	Std
Dataset A	Original Series	0.35	19	3.21	2.36
	Averaged Series	0.35	14.9	3.21	2.3
Dataset B	Original Series	0.01	17.1	6.09	2.99
	Averaged Series	0.83	14.7	6.09	2.91

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2 \tag{17}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(\hat{y}_i - y_i\right)^2}{N}} \tag{18}$$

$$MAPE = \frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{19}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2} \tag{20}$$

$$P_{INDEX} = \left|\frac{INDEX_{compared} - INDEX_{proposed}}{INDEX_{compared}}\right| \tag{21}$$

where y_i is the actual data, $\hat{y}_i$ is the prediction results, and N is the number of the actual data.

C. COMPARISON OF THE PROPOSED MODEL WITH OTHER COMPARABLE MODELS

To evaluate the developed model's accuracy and stability, six models were selected for comparison with WSO-VMD-CNN-GRU-NSWOA-EC, including Back Propagation Neural Network (BPNN), LSTM, GRU, CNN-GRU, WSO-VMD-CNN-GRU, WSO-VMD-CNN-GRU-NSWOA. Additionally, the forecast results and error metrics of each model were determined. To confirm the precision of the multi-step predictions made by these seven models, Tables 2 to 3 detail the prediction outcomes for the two distinct datasets. Additionally, Figures 5 to 6 illustrate the fit plots and scatter plots, showcasing the predictive performance of these seven models across the two datasets. The optimal outcomes are shown in green. The comparative analysis will be conducted from the following aspects:

(1) Single models versus hybrid models: As indicated in Tables 2 and 3, the BPNN model performs the worst in

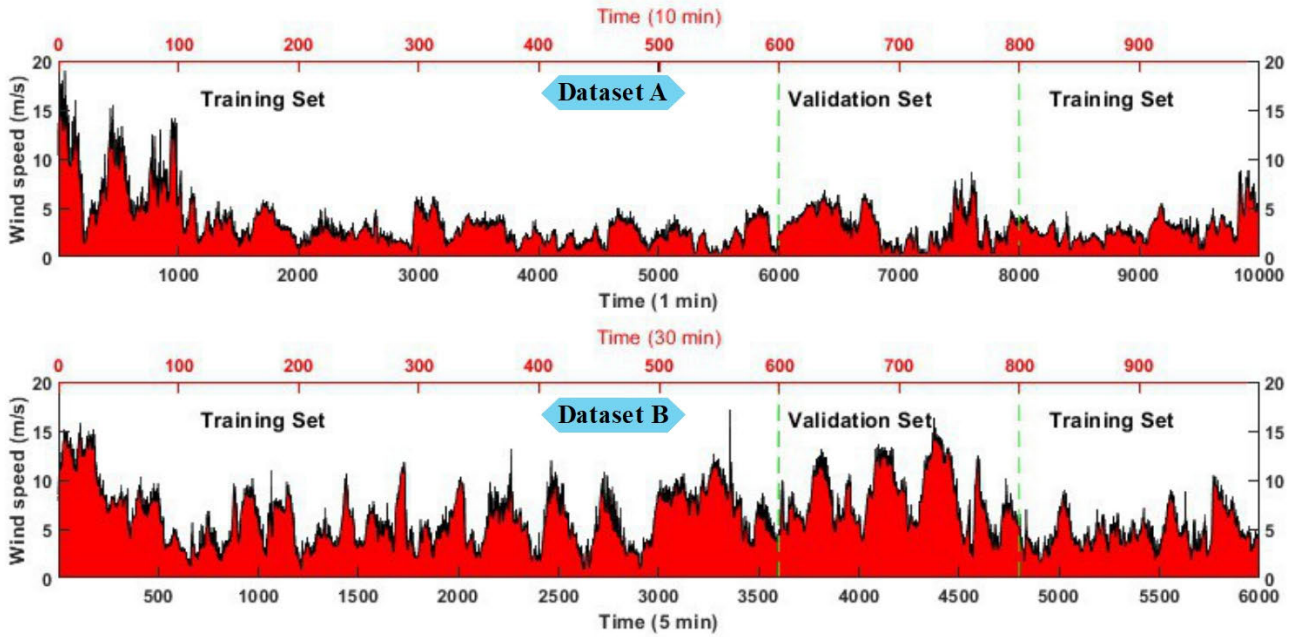

FIGURE 4. The original and averaged wind speed.

datasets A and B, with MAPEs as high as 25.5208% and 26.3331% for three-step forecasts. Additionally, the average MAPEs for 1-2-3 step forecasts for the LSTM and GRU models are 6.156%, 17.968%, 24.524% for the LSTM and 5.557%, 17.851%, 24.222% for the GRU, respectively. Thus, single models still exhibit significant differences when compared with other hybrid models. In the hybrid model, the WSO-VMD-CNN-GRU has the highest average MAPE, with values for the 1-2-3 steps being 3.830%, 6.838%, and 12.485% respectively. Therefore, it is evident that the hybrid forecasting model significantly outperforms the single prediction models in two-step and three-step forecasts.

(2) Comparison with and without multi-objective optimization integration: As depicted in Tables 2 to 3, in three-step forecasts, the WSO-VMD-CNN-GRU and WSO-VMD-CNN-GRU-NSWOA models show MAPEs of 9.2329% and 7.6174% respectively in dataset A, reflecting a reduction of 21.2%. Similarly, in dataset B, the MAPEs are 15.7368% and 13.723%, indicating a decrease of 12.8%. Furthermore, in the two-step forecasts for dataset A, the MAE values for the WSO-VMD-CNN-GRU and WSO-VMD-CNN-GRU-NSWOA models are 0.178 and 0.1104, respectively. In the one-step forecasts for dataset B, the MAE values are 0.2301 and 0.1795 for the WSO-VMD-CNN-GRU and WSO-VMD-CNN-GRU-NSWOA models, respectively. Therefore, across one-, two-, and three-step forecasts, the integration of multi-objective optimization methods effectively enhances prediction performance.

(3) Comparison of the proposed model with other models: The performance indicators for each model on datasets A and B are comprehensively delineated in Tables 2 and 3, illustrating that the proposed model surpasses its counterparts. In datasets A and B, the CNN-GRU model consistently demonstrates lower error values compared to the GRU, LSTM, and BPNN models. For instance, in datasets A and B the one-step forecast MAPE values for the CNN-GRU model are 4.2375% and 5.8748%, respectively. Furthermore, in the three-step forecasts for dataset B, the MAPE for the WSO-VMD-CNN-GRU-NSWOA-EC model is 12.722%. In contrast, the MAPEs for the BPNN, LSTM, GRU, CNN-GRU, and WSO-VMD-CNN-GRU models are 26.3331%, 25.981%, 25.8441%, 23.7415%, and 15.7368% respectively. Additionally, from the forecast of Dataset B in Figure 6, it is evident that the fitting curve of the presented model (marked with green circles) closely aligns with the actual values. In the scatter plot, the BPNN is more dispersed, whereas the model proposed in this article is more concentrated. This data clearly demonstrates that the proposed model significantly enhances the accuracy of multi-step forecasts.

Tables 4 and 5 display the improvement rates in forecasting performance of the proposed prediction model compared to six other models. These tables show that the WSO-VMD-CNN-GRU-NSWOA-EC model exhibits varying degrees of improvement across four metrics relative to the other models. Therefore, the hybrid prediction model proposed in this study demonstrates substantial effectiveness in enhancing short-term wind speed forecasting performance at wind farms.

As shown in Figure 7, the stacked bar chart presents the three-step forecasting evaluation metrics for datasets A and B. It is evident that the hybrid forecasting model significantly outperforms the single prediction models in terms of overall

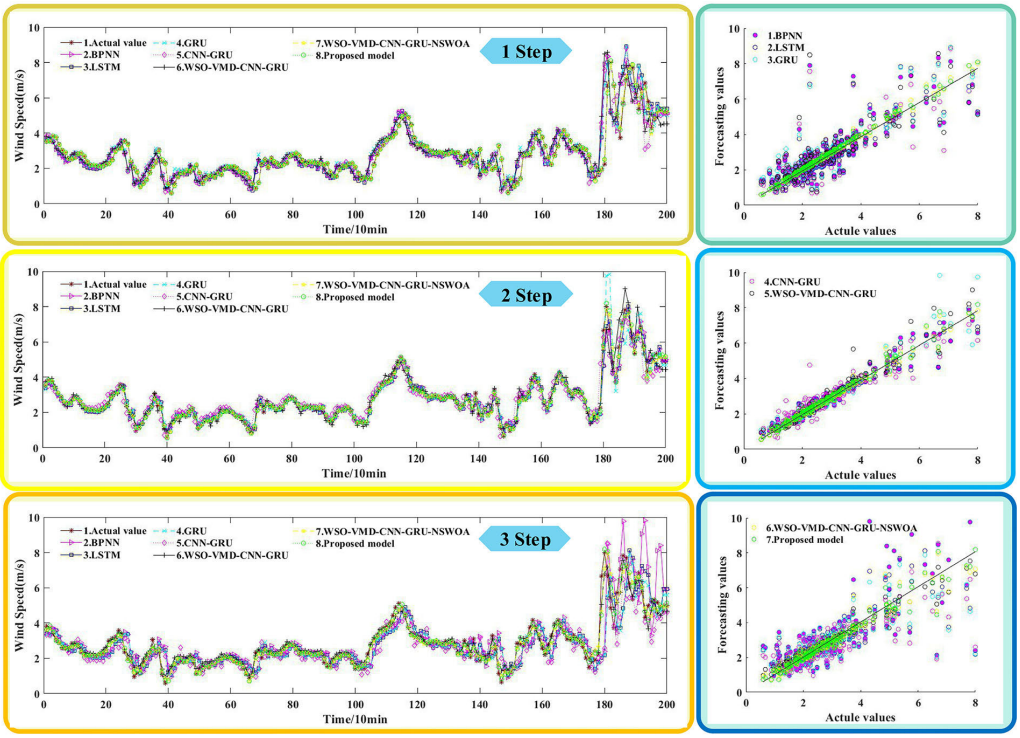

FIGURE 5. The forecasting results of dataset A.

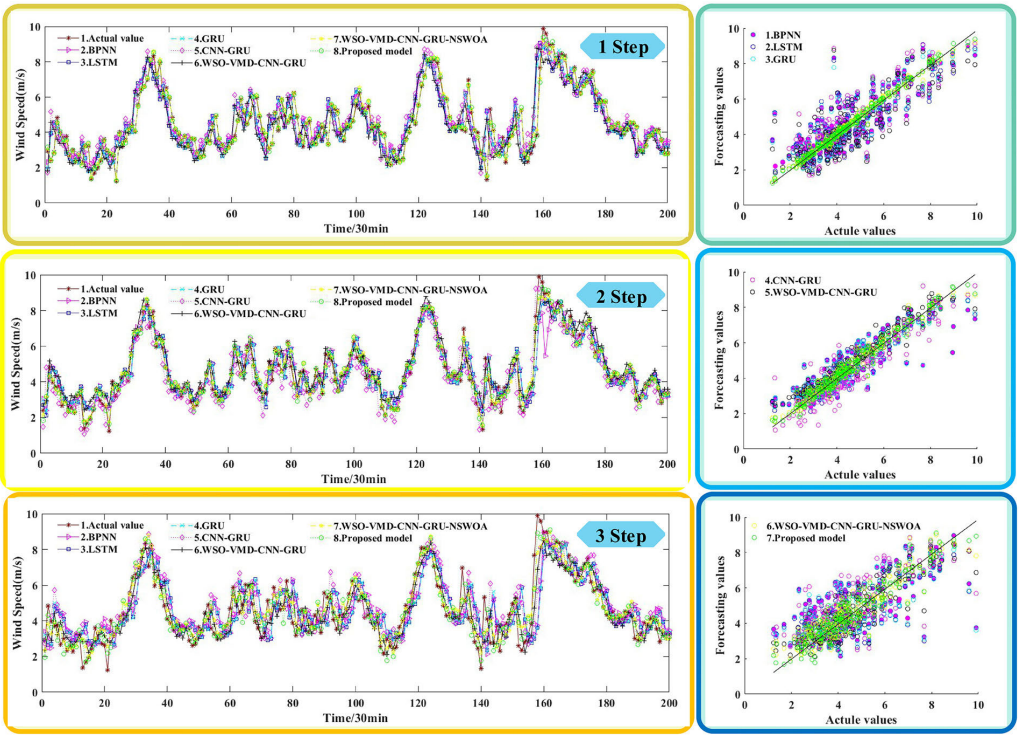

FIGURE 6. The forecasting results of dataset B.

performance. Considering the MAE, RMSE, MAPE, and 1-R2 metrics, the indicators for the seven models depicted in Figure 7 demonstrate a progressive downward trend from left to right. For example, in the three-step forecast for dataset A,

TABLE 2. The performance for dataset A.

Step	Models	MAE	RMSE	MAPE(%)	R^2
1 step	BPNN	0.1184	0.1543	5.6364	0.9874
	LSTM	0.1495	0.2287	5.4383	0.9722
	GRU	0.1112	0.1935	4.53	0.9801
	CNN-GRU	0.1331	0.2856	4.2375	0.9567
	WSO-VMD-CNN-GRU	0.0863	0.1463	3.0584	0.9886
	WSO-VMD-CNN-GRU-NSWOA	0.0553	0.0978	1.8773	0.9949
	WSO-VMD-CNN-GRU-NSWOA-EC	**0.0486**	**0.0873**	**1.7227**	**0.9959**
2 step	BPNN	0.4031	0.7001	15.8176	0.7405
	LSTM	0.3802	0.6202	15.334	0.7964
	GRU	0.3833	0.6243	15.2069	0.7937
	CNN-GRU	0.3409	0.5148	14.3854	0.8641
	WSO-VMD-CNN-GRU	0.178	0.284	6.5475	0.9573
	WSO-VMD-CNN-GRU-NSWOA	0.1104	0.1794	3.9101	0.983
	WSO-VMD-CNN-GRU-NSWOA-EC	**0.0956**	**0.1678**	**3.532**	**0.9852**
3 step	BPNN	0.6499	1.0616	25.5208	0.4085
	LSTM	0.5884	0.9383	23.0675	0.5379
	GRU	0.5766	0.9183	22.6007	0.5574
	CNN-GRU	0.5537	0.8771	21.4939	0.5962
	WSO-VMD-CNN-GRU	0.2153	0.3129	9.2329	0.9486
	WSO-VMD-CNN-GRU-NSWOA	0.2064	0.3488	7.6174	0.9366
	WSO-VMD-CNN-GRU-NSWOA-EC	**0.1747**	**0.3093**	**6.4325**	**0.9502**

TABLE 3. The performance for dataset B.

Step	Models	MAE	RMSE	MAPE(%)	R^2
1 step	BPNN	0.2778	0.3735	7.5129	0.9536
	LSTM	0.2759	0.3839	6.8737	0.9512
	GRU	0.263	0.3594	6.5842	0.9572
	CNN-GRU	0.2486	0.3171	5.8748	0.9666
	WSO-VMD-CNN-GRU	0.1772	0.2301	4.6019	0.9824
	WSO-VMD-CNN-GRU-NSWOA	0.1247	0.1795	3.011	0.9893
	WSO-VMD-CNN-GRU-NSWOA-EC	**0.1098**	**0.1597**	**2.6988**	**0.9915**
2 step	BPNN	0.8221	1.0898	21.3588	0.6063
	LSTM	0.777	1.0356	20.6026	0.6455
	GRU	0.7783	1.0409	20.4954	0.6419
	CNN-GRU	0.6265	0.8341	17.2648	0.7694
	WSO-VMD-CNN-GRU	0.2904	0.3345	7.1283	0.9629
	WSO-VMD-CNN-GRU-NSWOA	0.2414	0.3276	6.0507	0.9632
	WSO-VMD-CNN-GRU-NSWOA-EC	**0.2402**	**0.3238**	**6.1136**	**0.965**
3 step	BPNN	0.9926	1.2851	26.3331	0.4478
	LSTM	0.998	1.3012	25.981	0.4358
	GRU	0.9987	1.3007	25.8441	0.4363
	CNN-GRU	0.8965	1.1771	23.7415	0.5367
	WSO-VMD-CNN-GRU	0.6174	0.796	15.7368	0.7882
	WSO-VMD-CNN-GRU-NSWOA	0.5199	0.6636	13.723	0.8525
	WSO-VMD-CNN-GRU-NSWOA-EC	**0.4946**	**0.6362**	**12.722**	**0.8644**

the BPNN model recorded MAE, RMSE, and MAPE values of 0.6499, 1.0616, and 25.5208%, respectively. In comparison, the proposed model significantly outperforms the former in these metrics, with values of 0.1747, 0.3093, and 6.4325%, respectively. The corresponding improvements are reductions of 73.1%, 70.9%, and 74.8%. For the three-step forecast of dataset B, the LSTM model registered MAE, RMSE, and MAPE values of 0.9926, 1.2851, and 26.3331%, respectively. In contrast, the proposed model demonstrates superior performance, with corresponding values of 0.4946, 0.6362, and 12.722%, achieving reductions of 50.4%, 51.1%, and 51.0%, respectively. Thus, the hybrid model introduced in this study effectively extracts the hidden details within high-resolution data, reduces model error, and significantly enhances the multi-step forecasting accuracy of the hybrid model.

D. MULTI-OBJECTIVE OPTIMIZATION ENSEMBLE AND ERROR CORRECTION

This evaluation particularly focuses on the NSWOA's proficiency in generating superior Pareto front solutions, and its effectiveness in improving prediction accuracy. The NSWOA algorithm's optimization efficiency is crucial for determining the predictive accuracy of the ensemble model. The Pareto front of NSWOA and the selected solutions are depicted in Figure 8. It is observable that the Pareto front is well-structured, resembling a continuous hyperbolic curve. This curve's hyperbolic nature signifies that the solutions on the Pareto front are non-dominated, suggesting a balanced compromise between multi-objective. The results indicate that the predictions closely follow the actual trends and fluctuations of wind speed. For a comprehensive and fair assessment, three renowned multi-objective optimization algorithms were selected: Multi-Objective Grey Wolf Optimizer (MOGWO), Multi-objective Multi-verse Optimizer (MOMVO), and Multi-objective Artificial Hummingbird Algorithm (MOAHA). To ensure a basis for this comparative study, the maximum number of iterations was set to 500, and the population size was established at 100 for each multi-objective algorithm. By adhering to these consistent parameters, the study endeavors to provide a transparent and impartial assessment of the NSWOA's capabilities relative to its counterparts.

As presented in Tables 6 and 7, the optimal outcomes are highlighted in bold, and the 'Model' column omits the shared prefix WSO-VMD-CNN-GRU-, listing only the respective

TABLE 4. The improvement rates of the evaluation metrics for the proposed model in dataset A compared to other models.

Step	WSO-VMD-CNN-GRU-NSWOA-EC VS. The following models	P_{MAE}	P_{RMSE}	P_{MAPE}	P_{R2}
1 step	BPNN	59.0%	43.4%	69.4%	0.9%
	LSTM	67.5%	61.8%	68.3%	2.4%
	GRU	56.3%	54.9%	62.0%	1.6%
	CNN-GRU	63.5%	69.4%	59.3%	4.1%
	WSO-VMD-CNN-GRU	43.7%	40.3%	43.7%	0.7%
	WSO-VMD-CNN-GRU-NSWOA	12.1%	10.7%	8.2%	0.1%
2 step	BPNN	76.3%	76.0%	77.7%	33.0%
	LSTM	74.9%	72.9%	77.0%	23.7%
	GRU	75.1%	73.1%	76.8%	24.1%
	CNN-GRU	72.0%	67.4%	75.4%	14.0%
	WSO-VMD-CNN-GRU	46.3%	40.9%	46.1%	2.9%
	WSO-VMD-CNN-GRU-NSWOA	13.4%	6.5%	9.7%	0.2%
3 step	BPNN	73.1%	70.9%	74.8%	132.6%
	LSTM	70.3%	67.0%	72.1%	76.6%
	GRU	69.7%	66.3%	71.5%	70.5%
	CNN-GRU	68.4%	64.7%	70.1%	59.4%
	WSO-VMD-CNN-GRU	18.9%	1.2%	30.3%	1.5%
	WSO-VMD-CNN-GRU-NSWOA	15.4%	11.3%	15.6%	0.2%

TABLE 5. The improvement rates of the evaluation metrics for the proposed model in dataset B compared to other models.

Step	WSO-VMD-CNN-GRU-NSWOA-EC VS. The following models	P_{MAE}	P_{RMSE}	P_{MAPE}	P_{R2}
1 step	BPNN	60.5%	57.2%	64.1%	4.0%
	LSTM	60.2%	58.4%	60.7%	4.2%
	GRU	58.3%	55.6%	59.0%	3.6%
	CNN-GRU	55.8%	49.6%	54.1%	2.6%
	WSO-VMD-CNN-GRU	38.0%	30.6%	41.4%	0.9%
	WSO-VMD-CNN-GRU-NSWOA	11.9%	11.0%	10.4%	0.2%
2 step	BPNN	70.8%	70.3%	71.8%	59.2%
	LSTM	69.1%	68.7%	70.8%	49.5%
	GRU	69.1%	68.9%	70.7%	50.3%
	CNN-GRU	61.7%	61.2%	65.2%	25.4%
	WSO-VMD-CNN-GRU	17.3%	3.2%	15.6%	0.2%
	WSO-VMD-CNN-GRU-NSWOA	0.5%	1.2%	0.6%	0.2%
3 step	BPNN	50.2%	50.5%	51.7%	93.0%
	LSTM	50.4%	51.1%	51.0%	98.3%
	GRU	50.5%	51.1%	50.8%	98.1%
	CNN-GRU	44.8%	46.0%	46.4%	61.1%
	WSO-VMD-CNN-GRU	19.9%	20.1%	19.2%	9.7%
	WSO-VMD-CNN-GRU-NSWOA	4.9%	4.1%	7.3%	1.4%

multi-objective optimization algorithms. The tables reveal that the WSO-VMD-CNN-GRU-NSWOA model exhibits superior performance in multi-objective optimization ensemble forecasting across two datasets for 1-2-3 step predictions, outshining the other three ensemble models dedicated to multi-objective optimization. Specifically, in the 1-step forecast for Dataset A, the NSWOA model registers an average enhancement of 16.67% across three metrics when compared to MOAHA. Furthermore, the R^2 value of NSWOA averages a 0.15% improvement relative to the trio of alternative multi-objective optimization methods. In the 3-step forecast for Dataset B, NSWOA demonstrates an average increment of 12.24% over MOAHA across the same three metrics, coupled with an average R^2 advancement of 1.51% when juxtaposed with the other algorithms. It is noteworthy that the minor rate of enhancement observed in the 1-step forecasts, where R^2 values exceed 0.99, can be attributed to the already high baseline accuracy. However, in the 3-step forecasts, where R^2 values fluctuate between 0.92 and 0.94, the fitting results are comparatively less precise, thereby presenting substantial scope for further refinement.

In Figure 9, the radar charts depict the 3-step forecast results for two distinct datasets, both utilizing the base module WSO-VMD-CNN-GRU, with variations only in the multi-objective optimization algorithms used. The blue line represents the corresponding indicator results without error correction, and the orange line represents the corresponding indicator results with error correction. For the R2 metric,

TABLE 6. The performance of different multi-objective models for dataset A.

step	Models	MSE	RMSE	MAPE(%)	R^2
1	MOAHA	0.0634	0.1203	2.1126	0.9923
1	MOGWO	0.0587	0.1121	1.9715	0.9933
1	MOMVO	0.0562	0.1013	1.8941	0.9946
1	NSWOA	**0.0553**	**0.0978**	**1.8773**	**0.9949**
2	MOGWO	0.1064	0.1803	3.9714	0.9829
2	MOAHA	0.1095	0.1806	4.0771	0.9828
2	MOMVO	0.1105	0.1831	3.9253	0.9823
2	NSWOA	**0.1104**	**0.1794**	**3.9101**	**0.983**
3	MOAHA	0.23	0.3959	8.5198	0.9183
3	MOGWO	0.2264	0.3933	8.2118	0.9194
3	MOMVO	0.2141	0.3672	7.9822	0.9297
3	NSWOA	**0.2064**	**0.3488**	**7.6174**	**0.9366**

Note: The 'Models' column omits the shared prefix WSO-VMD-CNN-GRU-.

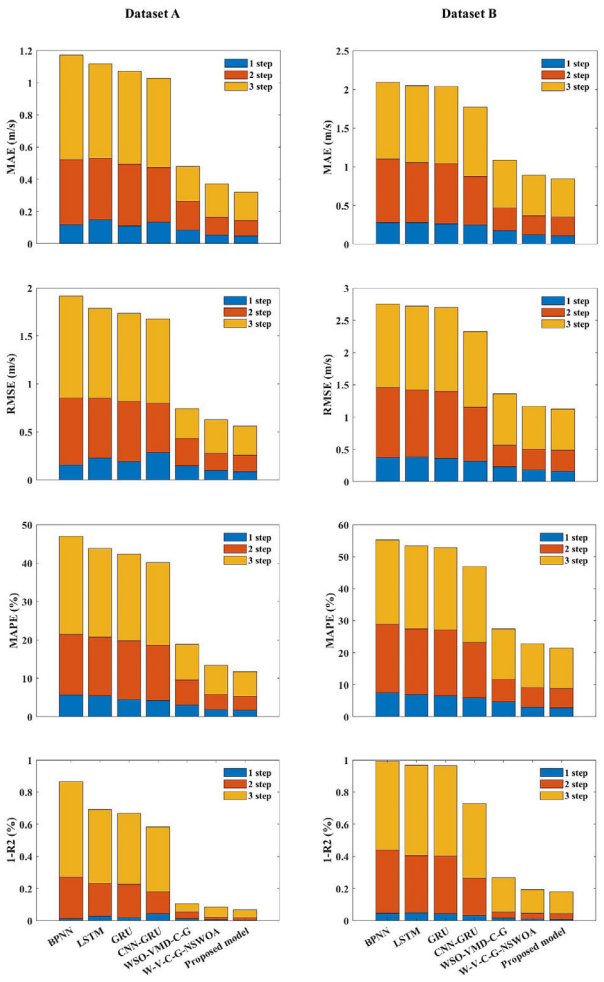

FIGURE 7. Columnar stacked chart of MAE, RMSE, MAPE, and 1-R^2 by various models.

where values closer to 1 are indicative of higher accuracy, the orange line, denoting the error-corrected results, is positioned on the outer circle, signifying an enhancement compared to the blue line. Regarding the MAE, RMSE, and MAPE metrics, where lower values are preferable, a noticeable reduction in these metrics is observed for Dataset B in comparison to Dataset A. This observation is supported by the 3-step prediction results: Dataset A shows 0.2064, 0.3488, and 7.6174, while Dataset B exhibits 0.5199, 0.6636, and 13.723. The results demonstrate a clearly superior performance in Dataset A and a relatively poorer outcome for Dataset B. Hence, Dataset B exhibits more significant potential for improvement via error correction. To summarize, the incorporation of multi-objective optimization for error correction unequivocally enhances the results, surpassing those achieved without such correction. This not only improves the predictive precision of the composite model but also ensures closer alignment with actual values. Moreover, in scenarios involving datasets with inherently poorer multi-step forecast results, the application of error correction through multi-objective optimization is particularly advantageous, offering substantial improvements.

TABLE 7. The performance of different multi-objective models for dataset B.

step	Models	MSE	RMSE	MAPE(%)	R^2
1	MOAHA	0.1283	0.1798	3.1306	0.9893
1	MOGWO	0.123	0.1755	2.9839	0.9898
1	MOMVO	0.1269	0.1817	3.0336	0.989
1	NSWOA	**0.1247**	**0.1795**	**3.011**	**0.9893**
2	MOAHA	0.2534	0.3391	6.3493	0.9617
2	MOGWO	0.2459	0.3318	6.1313	0.9633
2	MOMVO	0.2419	0.3289	6.084	0.9639
2	NSWOA	**0.2414**	**0.3276**	**6.0507**	**0.9632**
3	MOAHA	0.5463	0.7021	14.2826	0.8349
3	NSGWO	0.5509	0.7139	14.4169	0.8292
3	MOMVO	0.5251	0.6704	13.8564	0.8494
3	NSWOA	**0.5199**	**0.6636**	**13.723**	**0.8525**

Note: The 'Models' column omits the shared prefix WSO-VMD-CNN-GRU-.

TABLE 8. The performance of different multi-objective with error correction models for dataset A.

step	Models	MSE	RMSE	MAPE(%)	R^2
1	MOAHA	0.0614	0.1114	2.1601	0.9934
1	MOGWO	0.0591	0.1073	2.0318	0.9939
1	MOMVO	0.0501	0.0899	1.7769	0.9957
1	NSWOA	**0.0486**	**0.0873**	**1.7227**	**0.9959**
2	MOGWO	0.1059	0.1841	3.9863	0.9821
2	MOAHA	0.1061	0.1799	3.9722	0.9829
2	MOMVO	0.1004	0.1736	3.6809	0.9841
2	NSWOA	**0.0956**	**0.1678**	**3.532**	**0.9852**
3	MOAHA	0.1893	0.3251	7.3235	0.9449
3	MOGWO	0.1901	0.3375	6.9939	0.9407
3	MOMVO	0.1761	0.3108	6.5827	0.9497
3	NSWOA	**0.1747**	**0.3093**	**6.4325**	**0.9502**

Note: The 'Models' column omits the common portion WSO-VMD-CNN-GRU-MO-EC, where 'MO' denotes the respective multi-objective optimization algorithm.

As indicated in Tables 8 and 9, the best results are displayed in bold, and the 'Model' column omits the common portion WSO-VMD-CNN-GRU-MO-EC, where 'MO' denotes the respective multi-objective optimization algorithm. Tables 8-9 clearly show that error correction contributes less to 1-step predictions and more significantly to 2-step and 3-step predictions, with the latter seeing an improvement of up to 46.9%. This is attributed to the inherently higher accuracy of 1-step predictions, where errors are typically smaller, leaving limited scope for error correction. In contrast, for 2-step and 3-step forecasting, the longer forecast horizon correlates with

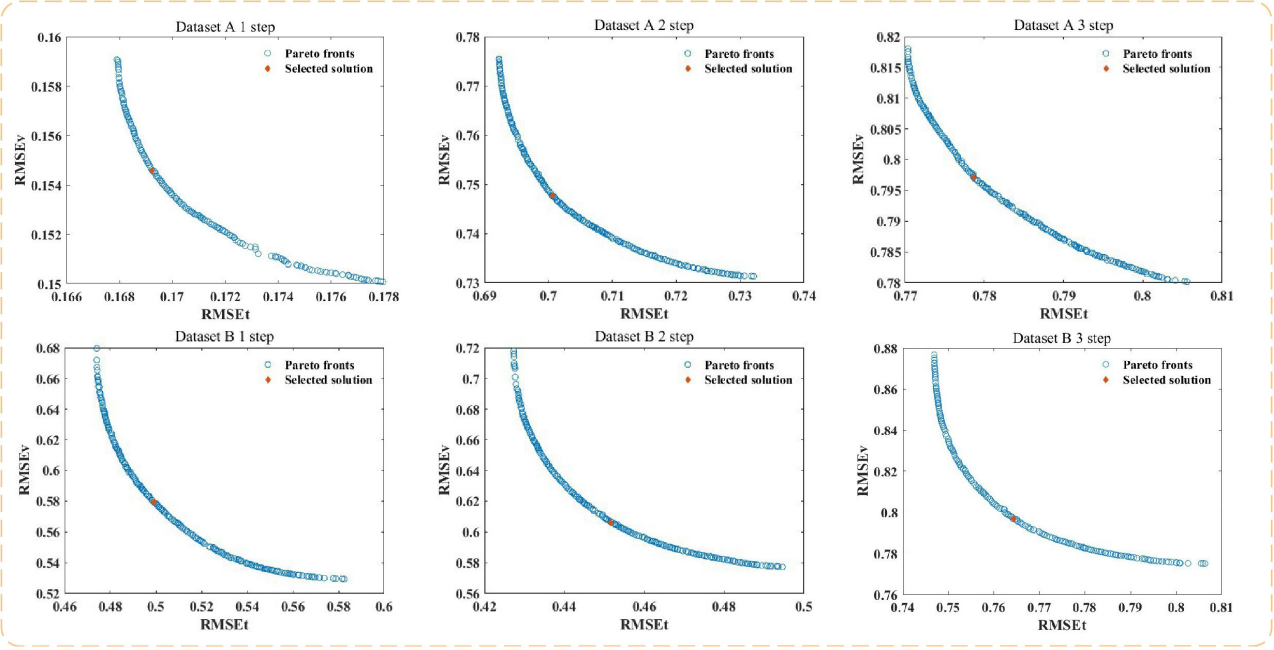

FIGURE 8. The Pareto front of NSWOA.

TABLE 9. The performance of different multi-objective with error correction models for dataset B.

step	Models	MSE	RMSE	MAPE(%)	R^2
1	MOAHA	0.1128	0.157	2.8113	0.9918
1	MOGWO	0.1106	0.1545	2.7838	0.9921
1	MOMVO	0.1116	0.16	2.7698	0.9915
1	NSWOA	**0.1098**	**0.1597**	**2.6988**	**0.9915**
2	MOAHA	0.2615	0.3446	6.5626	0.9604
2	MOGWO	0.2598	0.3417	6.5399	0.9611
2	MOMVO	0.257	0.3433	6.4424	0.9607
2	NSWOA	**0.2402**	**0.3238**	**6.1136**	**0.965**
3	MOAHA	0.5125	0.6619	13.1603	0.8532
3	NSGWO	0.5045	0.6616	12.9203	0.8533
3	MOMVO	0.4978	0.6409	12.8529	0.8624
3	NSWOA	**0.4946**	**0.6362**	**12.722**	**0.8644**

Note: The 'Models' column omits the common portion WSO-VMD-CNN-GRU-MO-EC, where 'MO' denotes the respective multi-objective optimization algorithm.

larger errors, thereby offering greater potential for improvement through error correction. Consequently, it is advisable to incorporate an error correction module in multi-step predictions to further enhance their accuracy.

E. DECOMPOSITION METHODS AND DIFFERENT OPTIMIZATION FOR VMD

To further enhance data quality and mitigate the impact of noise, the Wind Energy Sequence was decomposed using the WSO-VMD method. Initially, the number of variables was fixed at 2, the number of whales was set to 10, and with a maximum of 20 iterations. The penalty factor was chosen within the range of 500 to 3000, and the value of K ranged from 3 to 10, including only integers. Subsequently, WSO was employed to optimize the VMD parameters. As observed from Figure 10, the optimal number of IMF was identified as 9. The time domain of the modal components obtained through WSO-VMD decomposition is shown in Figure 10. Figure 11 exhibits the spectral distribution of the modal components. It is evident from Figure 11 that each mode is independent, effectively avoiding the issue of modal mixing.

The performance of the WSO-VMD method is assessed in comparison with GWO-VMD, VMD, and Improved Complementary Ensemble Empirical Mode Decomposition with Adaptive Noise (ICEEMDAN). It is important to note that the primary distinction among these four models lies in their respective optimization and decomposition methods. Table 10 illustrates the forecast results for two datasets using hybrid methods based on the two decomposition techniques: VMD and ICEEMDAN. A comprehensive review of Table 10 demonstrates that the VMD method achieves lower values in RMSE, MAE, and MAPE. For instance, in the case of dataset A, the MAE, RMSE, MAPE, and R2 values for the VMD method are 0.1109, 0.2112, 3.9029%, and 0.9763, respectively. These results imply that VMD is a relatively more effective model for data decomposition.

Subsequently, to assess the influence of distinct optimization algorithms on the predictive efficacy following VMD decomposition, two comparative models were employed: GWO-VMD-CNN-GRU and WSO-VMD-CNN-GRU. The

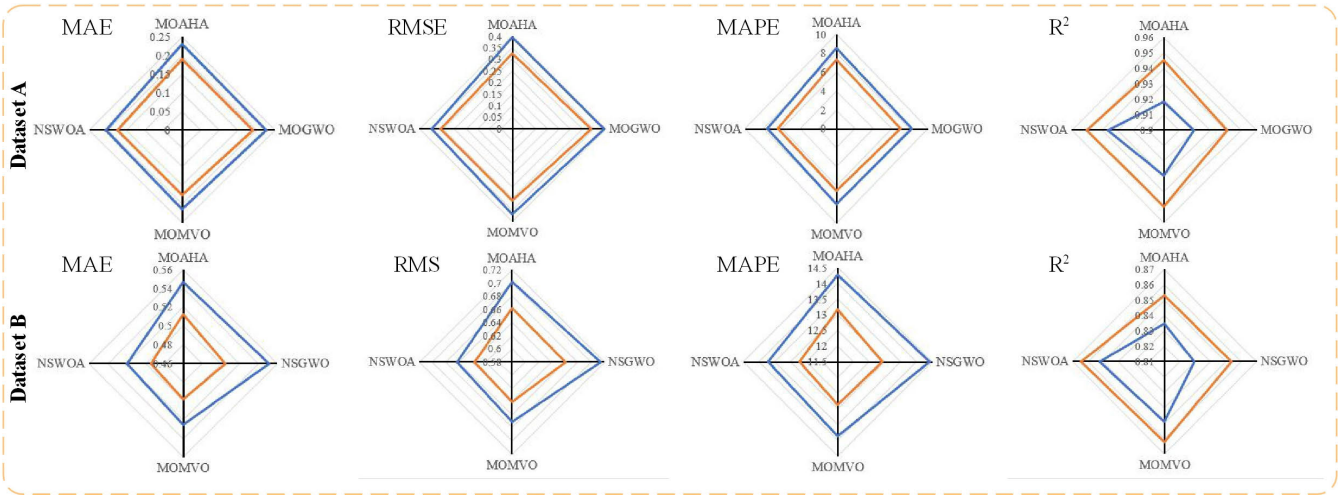

FIGURE 9. The radar chart of 3-step forecast results.

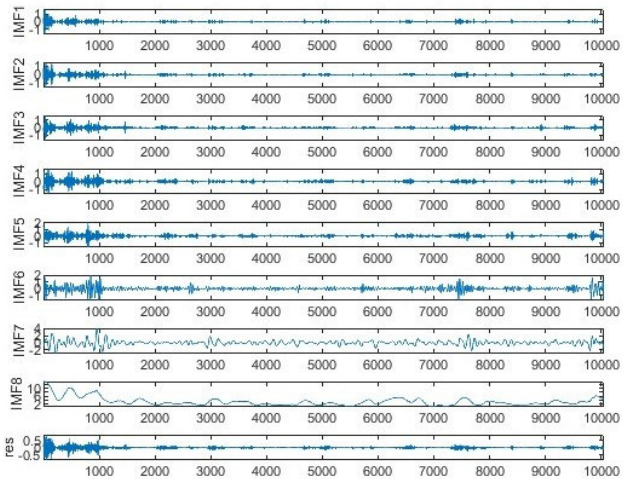

FIGURE 10. The IMF of WSO-VMD decomposition.

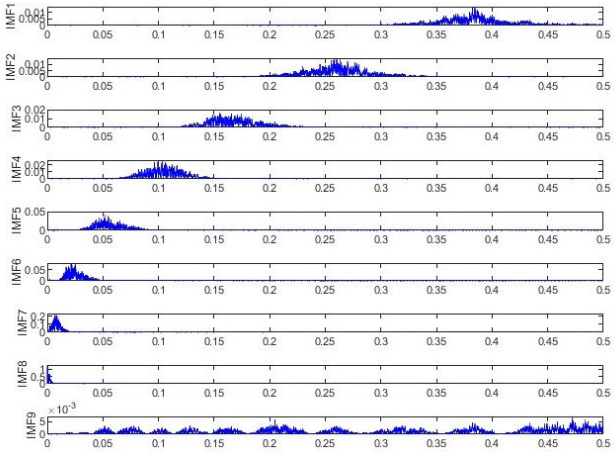

FIGURE 11. The spectrum of WSO-VMD decomposition.

forecasting effectiveness of these models across various datasets is quantitatively presented in Table 11. A comprehensive examination of the table reveals a clear edge of the WSO-VMD-CNN-GRU model in achieving superior evaluation metrics, such as MAPE, RMSE, MAE, across almost all datasets. Delving into a comparative assessment, the forecast results of the GWO-VMD and WSO-VMD-based hybrid models indicate that the minimum reduction in MAPE was 47.7%. This significant metric highlights the enhanced performance of the hybrid model that employs WSO-optimized VMD parameters over the GWO-optimized VMD model.

In Figure 12, a radar chart illustrates the performance metrics for various decomposition methods and optimization algorithms in Dataset A. The metrics include MAE, RMSE, and MAPE, where lower values signify best performance. The chart clearly shows that WSO-VMD-CNN-GRU achieves the lowest values for these indicators. For the R^2 metric, indicative of higher accuracy as it approaches 1,

TABLE 10. The performance of different decomposition methods.

Dataset	Models	MSE	RMSE	MAPE(%)	R^2
Dataset A	ICEEMDAN-CNN-GRU	0.1299	0.2477	4.0791	0.9674
	VMD-CNN-GRU	**0.1109**	**0.2112**	**3.9029**	**0.9763**
Dataset B	ICEEMDAN-CNN-GRU	0.2178	0.2759	5.3844	0.9747
	VMD-CNN-GRU	**0.2145**	**0.2718**	**5.3232**	**0.9754**

WSO-VMD-CNN-GRU is notably the closest to this ideal value. Figure 13 presents two scatter plots comparing real and forecasted values for the hybrid models that utilize various decomposition and optimization techniques across the two datasets. A detailed analysis of Figure 13 indicates that the WSO-VMD based hybrid model tends to exhibit least scattered points and best fitting accuracy. This observation underscores the effectiveness of the WSO-VMD approach in

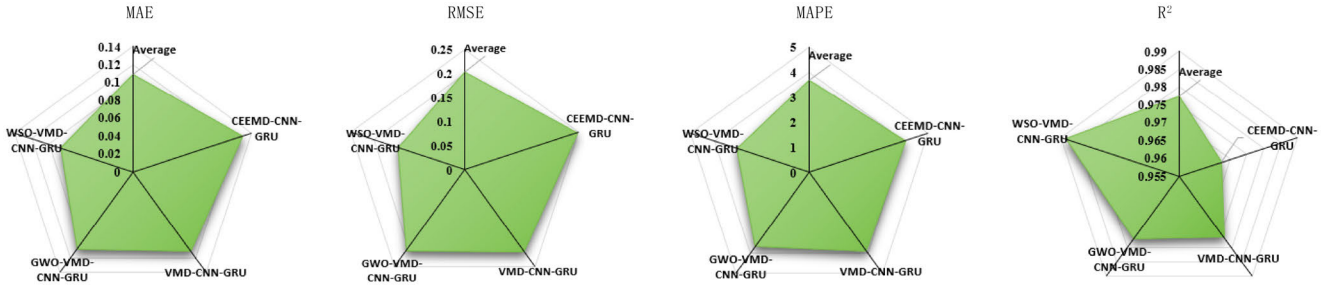

FIGURE 12. The radar chart of dataset A.

TABLE 11. The performance of different optimization algorithms.

Dataset	Models	MSE	RMSE	MAPE(%)	R^2
Dataset A	GWO-VMD-CNN-GRU	0.108	0.209	3.6647	0.9768
	WSO-VMD-CNN-GRU	**0.0863**	**0.1463**	**3.0584**	**0.9886**
Dataset B	GWO-VMD-CNN-GRU	0.2122	0.2684	5.0613	0.976
	WSO-VMD-CNN-GRU	**0.1772**	**0.2301**	**4.6019**	**0.9824**

TABLE 12. The DM test results.

Models	Dataset A			Dataset B		
	1 step	2 step	3 step	1 step	2 step	3 step
BPNN	3.5^a	2.68^a	4.12^a	6.13^a	5.46^a	5.33^a
LSTM	4.28^a	2.53^a	3.84^a	6.93^a	6.01^a	5.14^a
GRU	4.63^a	2.37^b	3.69^a	6.9^a	5.99^a	5.17^a
CNN-GRU	3.82^a	3.94^a	3.72^a	6.9^a	6.27	6.52^a
WSO-VMD-CNN-GRU	3.3^a	3.13^a	2.38^b	6.99^a	5.71^a	3.56^a
WSO-VMD-CNN-GRU-NSWOA	1.15	1.05	1.92^c	2.01^c	1.36	1.08

[a] indicates the 1% significance level (DM> 2.58), [b] indicates the 5% significance level (DM> 1.96), and [c] indicates the 10% significance level (DM> 1.65).

wind speed prediction, especially in terms of data fitting and prediction accuracy.

F. DM TEST

The Diebold Mariano (DM) test operates similarly to a t-test, where it examines whether the means of loss series generated by two different forecasting models are equal. In scenarios where autocorrelation is present, the DM test effectively estimates the standard deviation of the time series of loss differences in a way that accounts for autocorrelation. This feature makes the DM test particularly suitable for assessing models that forecast time series data. The basic assumption, or the null hypothesis (H0), of the test is that there is no discernible difference in the forecasting accuracy between the two compared models. However, as indicated in Table 12 of the study, the forecasting accuracies of the models exhibit

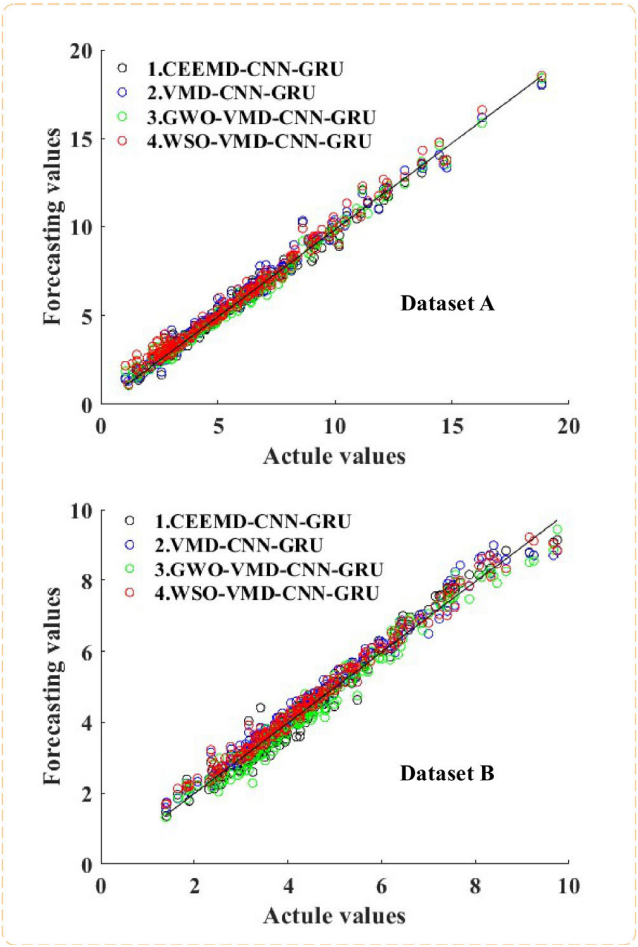

FIGURE 13. The scatter plots of different decomposition and optimization methods.

notable differences at significance levels of 1%, 5%, and 10%. Consequently, the hybrid forecasting model proposed in this research shows significantly superior forecasting performance compared to the reference models.

IV. CONCLUSION

The utilization and development of wind energy can mitigate the current energy crisis and reduce environmental pollution. Accurate forecasting of wind speed is crucial

TABLE 13. Parameter settings of the models.

Model	Model parameters	Reference values
CNN	Convolutional Kernel	1×1,1×5,1×10
	Activation Function	ReLU
	Max Pooling	2×1
	Optimizer	Adam
GRU	Input Layer Neurons	5
	Hidden Layer Neurons	128
	Output Layer Neurons	1/2/3
	Training Epochs	200
	Optimizer	Adam
ICEEMDAN	Noise Amplitude	0.2
	Noise Addition Quantity	200
VMD	Decomposition Quantity K	3-10
	Penalty Factor α	500-3000
WSO	Population Size	10
	Maximum Iterations	30
NSWOA/ MOGWO/ MOMVO/ MOAHA	Population Size	100
	Maximum Iterations	500
	Archive Library	100
	Objective Function Number	2
BPNN	Size of Input Units	5
	Size of Hidden Units	1
	Maximum Training Iterations	1000
	Learning Rate	0.01
	Training Accuracy	0.00004
LSTM	Size of input units	5
	Size of hidden units	30
	Size of output units	1
	Training Epochs	1000
	Learning rate	0.01
	Optimizer	Adam

for enhancing the efficiency of wind energy utilization. However, the inherent instability and intermittency of wind speed present significant challenges to its precise prediction. Therefore, this paper proposes a novel hybrid model termed WSO-VMD-CNN-GRU-NSWOA-EC. The model's performance has been validated and analyzed through comparisons with six other models and using four evaluation metrics across two distinct datasets. Consequently, this study arrives at the following conclusions: (1) Compared with six benchmark models, the proposed model demonstrates superior performance, outperforming both single and other hybrid models. After optimization with WSO, extraction of high-resolution information, integration of multi-objective optimization, and the inclusion of an error-correction module, the model increased prediction accuracy by at least 30% in steps 1-2-3. (2) A comparative analysis between VMD and ICEEMDAN, as well as different optimization algorithms, reveals that the predictive effectiveness following VMD decomposition surpasses that of ICEEMDAN. Utilizing WSO for optimal parameter selection in VMD mitigates the effects of volatility and non-stationarity in wind speed forecasting. (3) The multi-objective optimization integrated model significantly enhances the multi-step prediction results of CNN-GRU, and efficiency comparisons with other popular multi-objective algorithms show that NSWOA improves three key indicators by more than 10%. (4) Incorporating an error-correction module into multi-step forecasting significantly enhances its accuracy. Experimental comparisons indicate that the accuracy of error correction results for two-step and three-step predictions improved by 46.9%. (4) The final DM test confirms that, compared to other benchmark models, the developed model statistically exhibits significant differences and outperforms these benchmarks.

This study focuses on the issue of short-term wind speed forecasting at wind farms, with potential future research extending to a broader range of wind speed forecasting scenarios and types. Specifically, future studies could consider the spatio-temporal dependencies of multiple wind farms for joint uncertainty forecasting; real-time dynamic rolling decomposition and noise reduction based on online data collection for wind speed forecasting; and expanding the research scope to include medium- and long-term wind speed predictions.

APPENDIX

See Table 13.

REFERENCES

[1] Global Wind Energy Council. *Global Wind Report 2023-Global Wind Energy Council*. Accessed: Nov. 1, 2023. [Online]. Available: https://gwec.net/globalwindreport2023/

[2] W. Zhang, Z. Lin, and X. Liu, "Short-term offshore wind power forecasting—A hybrid model based on discrete wavelet transform (DWT), seasonal autoregressive integrated moving average (SARIMA), and deep-learning-based long short-term memory (LSTM)," *Renew. Energy*, vol. 185, pp. 611–628, Feb. 2022, doi: 10.1016/j.renene.2021.12.100.

[3] P. Lu, L. Ye, M. Pei, Y. Zhao, B. Dai, and Z. Li, "Short-term wind power forecasting based on meteorological feature extraction and optimization strategy," *Renew. Energy*, vol. 184, pp. 642–661, Jan. 2022, doi: 10.1016/j.renene.2021.11.072.

[4] S. M. H. D. Perera, G. Putrus, M. Conlon, M. Narayana, and K. Sunderland, "Wind energy harvesting and conversion systems: A technical review," *Energies*, vol. 15, no. 24, p. 24, Jan. 2022, Art. no. 24, doi: 10.3390/en15249299.

[5] E. P. James, S. G. Benjamin, and M. Marquis, "Offshore wind speed estimates from a high-resolution rapidly updating numerical weather prediction model forecast dataset," *Wind Energy*, vol. 21, no. 4, pp. 264–284, Apr. 2018, doi: 10.1002/we.2161.

[6] A. Jahanshahi, D. Jahanianfard, A. Mostafaie, and M. Kamali, "An auto regressive integrated moving average (ARIMA) model for prediction of energy consumption by household sector in euro area," *AIMS Energy*, vol. 7, no. 2, pp. 151–164, 2019, doi: 10.3934/energy.2019.2.151.

[7] Y. Ye, J. Che, and H. Wang, "Optimal component IGSCV-SVR ensemble model improved by VMD for ultra-short-term wind speed forecasting," *Eng. Lett.*, vol. 30, no. 3, pp. 1166–1175, 2022.

[8] Y.-L. Hu and L. Chen, "A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and differential evolution algorithm," *Energy Convers. Manage.*, vol. 173, pp. 123–142, Oct. 2018, doi: 10.1016/j.enconman.2018.07.070.

[9] F. Shahid, A. Zameer, and M. Muneeb, "A novel genetic LSTM model for wind power forecast," *Energy*, vol. 223, May 2021, Art. no. 120069, doi: 10.1016/j.energy.2021.120069.
[10] L. P. Joseph, R. C. Deo, R. Prasad, S. Salcedo-Sanz, N. Raj, and J. Soar, "Near real-time wind speed forecast model with bidirectional LSTM networks," *Renew. Energy*, vol. 204, pp. 39–58, Mar. 2023, doi: 10.1016/j.renene.2022.12.123.
[11] M. Nazemi, S. Chowdhury, and X. Liang, "A novel two-dimensional convolutional neural network-based an hour-ahead wind speed prediction method," *IEEE Access*, vol. 11, pp. 118878–118889, 2023, doi: 10.1109/ACCESS.2023.3327135.
[12] C. Cai, Y. Li, Z. Su, T. Zhu, and Y. He, "Short-term electrical load forecasting based on VMD and GRU-TCN hybrid network," *Appl. Sci.*, vol. 12, no. 13, p. 6647, Jun. 2022, doi: 10.3390/app12136647.
[13] H. Liu, R. Yang, and Z. Duan, "Wind speed forecasting using a new multi-factor fusion and multi-resolution ensemble model with real-time decomposition and adaptive error correction," *Energy Convers. Manage.*, vol. 217, Aug. 2020, Art. no. 112995, doi: 10.1016/j.enconman.2020.112995.
[14] W. Yang, Z. Tian, and Y. Hao, "A novel ensemble model based on artificial intelligence and mixed-frequency techniques for wind speed forecasting," *Energy Convers. Manage.*, vol. 252, Jan. 2022, Art. no. 115086, doi: 10.1016/j.enconman.2021.115086.
[15] Z. Zhang, J. Wang, D. Wei, T. Luo, and Y. Xia, "A novel ensemble system for short-term wind speed forecasting based on two-stage attention-based recurrent neural network," *Renew. Energy*, vol. 204, pp. 11–23, Mar. 2023, doi: 10.1016/j.renene.2022.12.120.
[16] Z. Kong, B. Tang, L. Deng, W. Liu, and Y. Han, "Condition monitoring of wind turbines based on spatio-temporal fusion of SCADA data by convolutional neural networks and gated recurrent units," *Renew. Energy*, vol. 146, pp. 760–768, Feb. 2020, doi: 10.1016/j.renene.2019.07.033.
[17] X. Guo, C. Zhu, J. Hao, S. Zhang, and L. Zhu, "A hybrid method for short-term wind speed forecasting based on Bayesian optimization and error correction," *J. Renew. Sustain. Energy*, vol. 13, no. 3, Jun. 2021, Art. no. 036101, doi: 10.1063/5.0048686.
[18] J. Dai, Y. Tan, and X. Shen, "Investigation of energy output in mountain wind farm using multiple-units SCADA data," *Appl. Energy*, vol. 239, pp. 225–238, Apr. 2019, doi: 10.1016/j.apenergy.2019.01.207.
[19] S. Zhang, C. Zhu, and X. Guo, "A novel combined model based on hybrid data decomposition, MSWOA and ENN for short-term wind speed forecasting," *IAENG Int. J. Comput. Sci.*, vol. 50, no. 3, pp. 1000–1008, 2023.
[20] F. Bilendo, A. Meyer, H. Badihi, N. Lu, P. Cambron, and B. Jiang, "Applications and modeling techniques of wind turbine power curve for wind farms—A review," *Energies*, vol. 16, no. 1, p. 180, Dec. 2022, doi: 10.3390/en16010180.
[21] M. Lopez-Martin, A. Sanchez-Esguevillas, L. Hernandez-Callejo, J. I. Arribas, and B. Carro, "Novel data-driven models applied to short-term electric load forecasting," *Appl. Sci.*, vol. 11, no. 12, p. 5708, Jun. 2021, doi: 10.3390/app11125708.
[22] H. Demolli, A. S. Dokuz, A. Ecemis, and M. Gokcek, "Wind power forecasting based on daily wind speed data using machine learning algorithms," *Energy Convers. Manage.*, vol. 198, Oct. 2019, Art. no. 111823, doi: 10.1016/j.enconman.2019.111823.
[23] B. Alonzo, P. Tankov, P. Drobinski, and R. Plougonven, "Probabilistic wind forecasting up to three months ahead using ensemble predictions for geopotential height," *Int. J. Forecasting*, vol. 36, no. 2, pp. 515–530, Apr. 2020, doi: 10.1016/j.ijforecast.2019.07.005.
[24] H. Liu and Z. Duan, "A vanishing moment ensemble model for wind speed multi-step prediction with multi-objective base model selection," *Appl. Energy*, vol. 261, Mar. 2020, Art. no. 114367, doi: 10.1016/j.apenergy.2019.114367.
[25] T. Peng, C. Zhang, J. Zhou, and M. S. Nazir, "Negative correlation learning-based RELM ensemble model integrated with OVMD for multi-step ahead wind speed forecasting," *Renew. Energy*, vol. 156, pp. 804–819, Aug. 2020, doi: 10.1016/j.renene.2020.03.168.
[26] H. Wang, Y.-M. Zhang, J.-X. Mao, and H.-P. Wan, "A probabilistic approach for short-term prediction of wind gust speed using ensemble learning," *J. Wind Eng. Ind. Aerodynamics*, vol. 202, Jul. 2020, Art. no. 104198, doi: 10.1016/j.jweia.2020.104198.
[27] G. Memarzadeh and F. Keynia, "A new short-term wind speed forecasting method based on fine-tuned LSTM neural network and optimal input sets," *Energy Convers. Manage.*, vol. 213, Jun. 2020, Art. no. 112824, doi: 10.1016/j.enconman.2020.112824.
[28] H. Liu, Y. Li, Z. Duan, and C. Chen, "A review on multi-objective optimization framework in wind energy forecasting techniques and applications," *Energy Convers. Manage.*, vol. 224, Nov. 2020, Art. no. 113324, doi: 10.1016/j.enconman.2020.113324.
[29] C. Wu, J. Wang, X. Chen, P. Du, and W. Yang, "A novel hybrid system based on multi-objective optimization for wind speed forecasting," *Renew. Energy*, vol. 146, pp. 149–165, Feb. 2020, doi: 10.1016/j.renene.2019.04.157.
[30] C. Li, Z. Xiao, X. Xia, W. Zou, and C. Zhang, "A hybrid model based on synchronous optimisation for multi-step short-term wind speed forecasting," *Appl. Energy*, vol. 215, pp. 131–144, Apr. 2018, doi: 10.1016/j.apenergy.2018.01.094.
[31] Z. Yang and J. Wang, "A combination forecasting approach applied in multistep wind speed forecasting based on a data processing strategy and an optimized artificial intelligence algorithm," *Appl. Energy*, vol. 230, pp. 1108–1125, Nov. 2018, doi: 10.1016/j.apenergy.2018.09.037.
[32] 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, Jul. 2022, Art. no. 123848, doi: 10.1016/j.energy.2022.123848.
[33] W. Fu, K. Wang, C. Li, and J. Tan, "Multi-step short-term wind speed forecasting approach based on multi-scale dominant ingredient chaotic analysis, improved hybrid GWO-SCA optimization and ELM," *Energy Convers. Manage.*, vol. 187, pp. 356–377, May 2019, doi: 10.1016/j.enconman.2019.02.086.
[34] X. Guo, C. Zhu, J. Hao, and S. Zhang, "Multi-step wind speed prediction based on an improved multi-objective seagull optimization algorithm and a multi-kernel extreme learning machine," *Appl. Intell.*, vol. 53, no. 13, pp. 16445–16472, Jul. 2023, doi: 10.1007/s10489-022-04312-7.
[35] Y. Nie, N. Liang, and J. Wang, "Ultra-short-term wind-speed bi-forecasting system via artificial intelligence and a double-forecasting scheme," *Appl. Energy*, vol. 301, Nov. 2021, Art. no. 117452, doi: 10.1016/j.apenergy.2021.117452.
[36] L. Luo, H. Li, J. Wang, and J. Hu, "Design of a combined wind speed forecasting system based on decomposition-ensemble and multi-objective optimization approach," *Appl. Math. Model.*, vol. 89, pp. 49–72, Jan. 2021, doi: 10.1016/j.apm.2020.07.019.
[37] H. Liu, Z. Duan, Y. Li, and H. Lu, "A novel ensemble model of different mother wavelets for wind speed multi-step forecasting," *Appl. Energy*, vol. 228, pp. 1783–1800, Oct. 2018, doi: 10.1016/j.apenergy.2018.07.050.
[38] L. Zhang, Y. Dong, and J. Wang, "Wind speed forecasting using a two-stage forecasting system with an error correcting and nonlinear ensemble strategy," *IEEE Access*, vol. 7, pp. 176000–176023, 2019, doi: 10.1109/ACCESS.2019.2957174.
[39] K. U. Jaseena and B. C. Kovoor, "Decomposition-based hybrid wind speed forecasting model using deep bidirectional LSTM networks," *Energy Convers. Manage.*, vol. 234, Apr. 2021, Art. no. 113944, doi: 10.1016/j.enconman.2021.113944.
[40] Z. Jiang, J. Che, and L. Wang, "Ultra-short-term wind speed forecasting based on EMD-VAR model and spatial correlation," *Energy Convers. Manage.*, vol. 250, Dec. 2021, Art. no. 114919, doi: 10.1016/j.enconman.2021.114919.
[41] A. Meng, S. Chen, Z. Ou, W. Ding, H. Zhou, J. Fan, and H. Yin, "A hybrid deep learning architecture for wind power prediction based on bi-attention mechanism and crisscross optimization," *Energy*, vol. 238, Jan. 2022, Art. no. 121795, doi: 10.1016/j.energy.2021.121795.
[42] L. Han, H. Jing, R. Zhang, and Z. Gao, "Wind power forecast based on improved long short term memory network," *Energy*, vol. 189, Dec. 2019, Art. no. 116300, doi: 10.1016/j.energy.2019.116300.
[43] V. Krishna Rayi, S. P. Mishra, J. Naik, and P. K. Dash, "Adaptive VMD based optimized deep learning mixed kernel ELM autoencoder for single and multistep wind power forecasting," *Energy*, vol. 244, Apr. 2022, Art. no. 122585, doi: 10.1016/j.energy.2021.122585.
[44] Z. Sun and M. Zhao, "Short-term wind power forecasting based on VMD decomposition, ConvLSTM networks and error analysis," *IEEE Access*, vol. 8, pp. 134422–134434, 2020, doi: 10.1109/ACCESS.2020.3011060.
[45] J. Duan, P. Wang, W. Ma, S. Fang, and Z. Hou, "A novel hybrid model based on nonlinear weighted combination for short-term wind power forecasting," *Int. J. Electr. Power Energy Syst.*, vol. 134, Jan. 2022, Art. no. 107452, doi: 10.1016/j.ijepes.2021.107452.

[46] H. Liu, Z. Duan, C. Chen, and H. Wu, "A novel two-stage deep learning wind speed forecasting method with adaptive multiple error corrections and bivariate Dirichlet process mixture model," *Energy Convers. Manage.*, vol. 199, Nov. 2019, Art. no. 111975, doi: 10.1016/j.enconman.2019.111975.

[47] M. Ding, H. Zhou, H. Xie, M. Wu, Y. Nakanishi, and R. Yokoyama, "A gated recurrent unit neural networks based wind speed error correction model for short-term wind power forecasting," *Neurocomputing*, vol. 365, pp. 54–61, Nov. 2019, doi: 10.1016/j.neucom.2019.07.058.

[48] H. Liu and C. Chen, "Multi-objective data-ensemble wind speed forecasting model with stacked sparse autoencoder and adaptive decomposition-based error correction," *Appl. Energy*, vol. 254, Nov. 2019, Art. no. 113686, doi: 10.1016/j.apenergy.2019.113686.

[49] H. Liu and Z. Duan, "Corrected multi-resolution ensemble model for wind power forecasting with real-time decomposition and bivariate kernel density estimation," *Energy Convers. Manage.*, vol. 203, Jan. 2020, Art. no. 112265, doi: 10.1016/j.enconman.2019.112265.

[50] Y. Deng, B. Wang, and Z. Lu, "A hybrid model based on data preprocessing strategy and error correction system for wind speed forecasting," *Energy Convers. Manage.*, vol. 212, May 2020, Art. no. 112779, doi: 10.1016/j.enconman.2020.112779.

[51] H. Liu, R. Yang, T. Wang, and L. Zhang, "A hybrid neural network model for short-term wind speed forecasting based on decomposition, multi-learner ensemble, and adaptive multiple error corrections," *Renew. Energy*, vol. 165, pp. 573–594, Mar. 2021, doi: 10.1016/j.renene.2020.11.002.

[52] M. Chen, L. Yu, C. Zhi, R. Sun, S. Zhu, Z. Gao, Z. Ke, M. Zhu, and Y. Zhang, "Improved faster R-CNN for fabric defect detection based on Gabor filter with genetic algorithm optimization," *Comput. Ind.*, vol. 134, Jan. 2022, Art. no. 103551, doi: 10.1016/j.compind.2021.103551.

[53] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder–decoder for statistical machine translation," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, 2014, pp. 1724–1734, doi: 10.3115/v1/d14-1179.

[54] M. Braik, A. Hammouri, J. Atwan, M. A. Al-Betar, and M. A. Awadallah, "White shark optimizer: A novel bio-inspired meta-heuristic algorithm for global optimization problems," *Knowl.-Based Syst.*, vol. 243, May 2022, Art. no. 108457, doi: 10.1016/j.knosys.2022.108457.

[55] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Adv. Eng. Softw.*, vol. 95, pp. 51–67, May 2016, doi: 10.1016/j.advengsoft.2016.01.008.

[56] A. M. Abdelaziz, T. H. A. Soliman, K. K. A. Ghany, and A. A. E.-M. Sewisy, "A Pareto-based hybrid whale optimization algorithm with Tabu search for multi-objective optimization," *Algorithms*, vol. 12, no. 12, p. 261, Dec. 2019, doi: 10.3390/a12120261.

[57] Q. N. U. Islam, A. Ahmed, and S. M. Abdullah, "Optimized controller design for islanded microgrid using non-dominated sorting whale optimization algorithm (NSWOA)," *Ain Shams Eng. J.*, vol. 12, no. 4, pp. 3677–3689, Dec. 2021, doi: 10.1016/j.asej.2021.01.035.

SHENGCAI ZHANG was born in Lanzhou, Gansu, China, in 1982. He received the master's degree in signal and information processing from Lanzhou University of Technology, Lanzhou, in 2009, where he is currently pursuing the Ph.D. degree with the School of Computer and Communication. He is also an Associate Professor with the School of Cyber Security, Gansu University of Political Science and Law. His research interests include time series forecasting, artificial intelligence, and intelligent optimization.

CHANGSHENG ZHU is currently a Professor and a Ph.D. Supervisor with Lanzhou University of Technology. He is also the Director of the Traditional Chinese Medicine Big Data Engineering Research Center, Gansu. His research interests include high-performance computing, big data mining, and engineering manufacturing information systems. In 2010, he was selected as a Leading Talent in Gansu Province.

XIUTING GUO (Member, IEEE) received the master's degree in engineering. She is currently pursuing the Ph.D. degree with the School of Computer and Communication, Lanzhou University of Technology. Since December 2012, she has been teaching with the School of Science, Lanzhou University of Technology.