

Article

A Short-Term Wind Speed Forecasting Framework Coupling a Maximum Information Coefficient, Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, Shared Weight Gated Memory Network with Improved Northern Goshawk Optimization for Numerical Weather Prediction Correction

Yanghe Liu ¹, Hairong Zhang ¹, Chuanfeng Wu ^{2,*}, Mengxin Shao ², Liting Zhou ¹ and Wenlong Fu ^{2,3,*} 

¹ Hubei Key Laboratory of Intelligent Yangtze and Hydroelectric Science, China Yangtze Power Co., Ltd., Yichang 443000, China; liu_yanghe@ctg.com.cn (Y.L.); zhang_hairong@ctg.com.cn (H.Z.); zhou_liting@ctg.com.cn (L.Z.)

² College of Electrical Engineering and New Energy, China Three Gorges University, Yichang 443002, China; 202308580121187@ctgu.edu.cn

³ Hubei Provincial Key Laboratory for Operation and Control of Cascaded Hydropower Station, China Three Gorges University, Yichang 443002, China

* Correspondence: 202208580121244@ctgu.edu.cn (C.W.); fuwenlong@ctgu.edu.cn (W.F.)



Citation: Liu, Y.; Zhang, H.; Wu, C.; Shao, M.; Zhou, L.; Fu, W. A Short-Term Wind Speed Forecasting Framework Coupling a Maximum Information Coefficient, Complete Ensemble Empirical Mode Decomposition with Adaptive Noise, Shared Weight Gated Memory Network with Improved Northern Goshawk Optimization for Numerical Weather Prediction Correction. *Sustainability* **2024**, *16*, 6782. <https://doi.org/10.3390/su16166782>

Academic Editors: Mohamed A. Mohamed and Bowen Zhou

Received: 20 June 2024

Revised: 1 August 2024

Accepted: 6 August 2024

Published: 7 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: In line with global carbon-neutral policies, wind power generation has received widespread public attention, which can enhance the security of supply and social sustainability. Since wind with non-stationarity and randomness makes power systems unstable, precise wind speed forecasting is an integral part of wind farm scheduling and management. Therefore, a compound short-term wind speed forecasting framework based on numerical weather prediction (NWP) is proposed coupling a maximum information coefficient (MIC), complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), shared weight gated memory network (SWGMN) with improved northern goshawk optimization (INGO). Firstly, numerical weather prediction is adopted to acquire the predicted variables with different domains, including the predicted wind speed and other predicted meteorological variables, after which the error is calculated using the predicted and actual wind speeds. Then, the correlation between the predicted variables and the error is obtained using the MIC to select the correlation factors. Subsequently, CEEMDAN is employed to decompose the correlation factors, corresponding the actual factors and the error into a series of subsequences, which are regarded as the input series. After that, the input series is fed into the proposed SWGMN to forecast each subsequent error, respectively, in which the shared gate is proposed to replace the input gate, the forgetting gate and the output gate. Meanwhile, the proposed INGO based on northern goshawk optimization (NGO), the levy flight disturbance strategy and the nonlinear contraction strategy is applied to calibrate the parameters of the SWGMN. Finally, the forecasting values are acquired by summing the forecasted error and the predicted wind speed from the NWP. The experimental results depict that the errors are small among all the models. Compared with the traditional method, the proposed framework achieves higher prediction accuracy and efficiency. The application of this framework not only assists in optimizing the operation and management of wind farms, but also reduces the dependence on fossil fuels, thereby promoting environmental protection and the sustainable use of resources.

Keywords: short-term wind speed forecasting; numerical weather prediction; maximum information coefficient; complete ensemble empirical mode decomposition with adaptive noise; shared weight gated memory network; improved northern goshawk optimization

1. Introduction

With ongoing global warming and rising fossil fuel prices, governments have vigorously implemented global carbon neutrality policies to reduce the reliance on fossil fuels and encourage the deployment of renewable energy [1]. As a mainstream energy source, wind energy is the key to energy transformation because it is recyclable, economical and abundant [2]. Despite some challenges in 2023, such as global supply chain disruptions and intensifying global energy crises, wind energy maintained high-quality and fast-paced development. For the first time, 1 TW of energy was generated, thereby providing a solid foundation for sustainable social development. However, the randomness of wind speed would result in fluctuations in the voltage and output power of the wind power generation system, resulting in negative effects on the stability of the power grid [3,4]. Therefore, it is crucial to research high-precision wind speed prediction models, which can improve power system stability, optimize the wind power generation plan, and reduce the dispatching cost of the power system [5].

Generally, wind speed forecasting models can be divided into three categories: statistical models, physical models and hybrid models [6]. Among them, the statistical models are dedicated to calculating the relationship between the input and output variables by mathematical statistical methods, such as the autoregressive moving average (ARMA) [7], the autoregressive integrated moving average (ARIMA) [8], and the fractional autoregressive integrated moving average (f-ARIMA) [9]. Although the statistical models have a high prediction performance for linear data, they are limited in processing nonlinear series [10]. Contrastively, numerical weather prediction (NWP), as a representative of physical models, is adopted to simulate wind speed trends using meteorological data and geographical parameters, especially for 48–72 h. With the development of NWP, there are now many models for forecasting wind speed, which mainly include high-resolution limited area models (HIRLAMs) [11], fifth-generation mesoscale models (MM5s) [12], and weather research and forecasting (WRF) models [13,14]. Nevertheless, the model structure, the inputs and the physical scheme of NWP have uncertainty, resulting in a certain error between the forecasting value and the actual data [15].

To correct the errors of NWP, hybrid models coupling data preprocessing methods, artificial intelligence (AI) models, and parameter optimization have been widely researched in recent years. Among them, data preprocessing consists of correlation analysis and a decomposition technique. Considering that the excessive input of meteorological variables collected by NWP will lead to information redundancy, correlation analysis is adopted to select the correlation factors. For instance, Chen et al. [16] adopted the Pearson coefficient to evaluate the correlation between meteorological factors and wind speed for NWP correction, thereby selecting the appropriate variables as inputs. Wu et al. [17] applied PCA to capture input data characteristics from NWP, where the experimental results demonstrate that PCA can reduce the computation complexity and improve the prediction accuracy. Moreover, due to the nonlinear nature of wind speed, decomposition methods are adopted to transform wind speed series into a set of subsequences. For example, Wang et al. [18] applied several subseries obtained by CEEMDAN and the predicted data from NWP as inputs for a prediction model. Among them, CEEMDAN introduces adaptive white noise to achieve a satisfactory decomposition performance compared with those of EMD, EEMD and CEEMD. However, few studies have applied decomposition techniques to correct NWP, especially in the field of multivariate wind speed forecasting. In our study, CEEMDAN is implemented to transform multivariate series into multiple subsequences.

Furthermore, AI models have been widely adopted to correct NWP due to its strong nonlinear adaptability and learning ability [19], such as artificial neural networks (ANNs) [20], support vector regression (SVR) [21], and long short-term memory (LSTM) [22]. Among them, ANN achieves nonlinear adaptability and has a short running time. Moosavi et al. [23] applied an ANN and random forest (RF) to study uncertainty quantification in NWP, which proved that ANNs outperform RF, and the running time of ANNs is shorter than that of RF. Nevertheless, the ANN prediction performance is unstable since the internal structure

randomly generates inherent parameters. Contrastively, SVR has a good generalization ability in dealing with small samples. Cai et al. [24] employed SVR to fuse the forecasting results obtained by NWP, where the experimental results affirm that SVR can effectively correct the error of NWP. However, the computational complexity of SVR surges with an increase in sample size [25], which is unsuitable for large sample prediction. In contrast, LSTM utilizes memory modules to effectively capture the important parts of time series information in a large sample, thereby overcoming the limited short-term memory ability aroused by recurrent neural networks [26]. For instance, Xu et al. [27] employed LSTM for NWP error correction, in which the experimental results depict that LSTM can reduce the wind speed prediction error of NWP significantly. Han et al. [15] applied bidirectional LSTM to extract the temporal correlation features from NWP, thereby improving accurate results and a better prediction effect. Although LSTM achieve high prediction accuracy, the prediction training time is longer than that of other AI models due to its complex internal structure and many weight parameters [28]. As an enhanced version of LSTM, a shared weight gated memory network (SWGMMN) is proposed for NWP correction in the field of wind speed forecasting, in which the proposed shared gate replaces the traditional forgetting gate, the input gate, and the output gate and shares the weights with different values, but of the same type, in the LSTM as a uniform weight, which leads to a simpler structure in the network and greatly reduces the forecasting time.

Since the model prediction accuracy is greatly affected by its own hyperparameters, it is essential for researchers to select appropriate hyperparameters for models using theoretical methods, which mainly include manual methods and optimization algorithms. Although the operating principle of manual methods are simple, they have strong subjectivity and limited experience, which easily lead to one calibrating a sub-optimal solution; thus, it should not be used in the field of actual wind speed prediction. Conversely, optimization algorithms employ the principle of gradient descent to calibrate the optimal solution route of the inherent parameters and the approaches to the optimal solution iteratively, thus avoiding the subjective deviation of human judgment. In recent years, a large number of optimization algorithms inspired by various biological behaviors have been proposed as research hotspots to improve prediction model accuracy, such as particle swarm optimization (PSO), beetle antennae search (BAS), and northern goshawk optimization (NGO). Among them, NGO has the best performance among all the classical optimization algorithms on 68 different objective functions, proving that NGO is highly capable of solving real-world problems [29]. Although optimization algorithms have a strong search ability and a fast convergence rate, they cannot always determine the local optimal solution. Therefore, many researchers have studied some improved optimization algorithms to find a global optimal solution. Fu et al. [30] proposed a combined optimization algorithm based on DE and a slime mold algorithm, which demonstrate that the proposed algorithm can enhance the global optimization ability. Referring to the previous studies, we propose improved northern goshawk optimization (INGO) based on a levy flight disturbance strategy and a nonlinear contraction strategy, in which the nonlinear contraction strategy is employed to speed up the convergence of this algorithm, and the levy flight disturbance strategy is used to enhance the ability of the algorithm to determine the local optimal solution.

In conclusion, a novel hybrid short-term wind speed forecasting framework is proposed based on the MIC, CEEMDAN, and the SWGMMN with INGO for NWP correction. Firstly, the MIC is employed to acquire the correlation between the predicted variables and the error to select the correlation factors, in which the predicted variables with different domains, including the predicted wind speed and other meteorological variables, are obtained by NWP, and the error is calculated using the predicted and actual wind speeds. Then, the selected correlation factors and the error are decomposed into multiple subsequences by CEEMDAN. Subsequently, the multiple subsequences are input into the proposed SWGMMN to forecast each subsequent error, in which the shared gate is proposed to replace the input gate, the forgetting gate and the output gate in the SWGMMN. Furthermore, the proposed INGO coupling NGO, the levy flight disturbance strategy and the nonlinear contraction

strategy is employed to optimize the parameters of the SWGMN. Ultimately, the wind speed forecasting values are obtained by accumulating the forecasted error of all the subsequences and the predicted wind speed from NWP. The framework optimizes the utilization of wind energy resources by improving the accuracy of wind speed prediction, thereby contributing to the sustainable development of society. The principal contributions are described as follows:

- (1) The MIC is deployed to select the meteorological factors with different time domains. By eliminating the irrelevant variables and retaining the main components, the influence of the irrelevant factors on the SWGMN can be avoided to improve the prediction accuracy.
- (2) The meteorological factors, the historical data and the error are decomposed into multiple subsequences by CEEMDAN to reduce data non-stationarity and boost the prediction performance.
- (3) An improved network, the SWGMN, as a variant of LSTM, is proposed by replacing the forgetting gate, the input gate and the output gate with the shared gate, which achieves a good prediction accuracy and can avoid the long training process caused by LSTM; thus, it is more suitable for NWP correction in the field of short-term wind speed forecasting.
- (4) The proposed INGO is developed to optimize the parameters of the SWGMN by combining the levy flight disturbance strategy and the nonlinear contraction strategy, which can determine the local optimal solution and accelerate the convergence speed, contributing to improving the generalization, prediction performance and stability of the SWGMN.

The rest of the paper is ordered as follows: Data preprocessing, the shared weight gated memory network, and improved northern goshawk optimization are described in Section 2. Section 3 shows the architecture of the proposed framework. The experimental results and analysis are presented in Section 4. Section 5 summarizes the conclusions.

2. Methodology

2.1. Data Preprocessing

2.1.1. Maximum Information Coefficient (MIC)

The MIC measures the correlation between variables as proposed by Reshef et al. [31], which can identify the complex functional relationship among large nonlinear samples. Compared with the other variable selection methods, the MIC achieves strong robustness and low computational complexity, which is widely adopted in different fields [32–34]. The essential thought behind the MIC is to divide the scatter plots of the selected variables and the target variables into grids, and then normalize the maximum mutual information obtained from all the different partition schemes.

For a given set C , there are two variables, including $A = \{a_1, a_2, a_3, \dots, a_n\}$ and $B = \{b_1, b_2, b_3, \dots, b_n\}$. The mutual information calculation between variables A and B is expressed as follows:

$$MI(C, A, B) = \sum_{a \in A} \sum_{b \in B} p(a, b) \lg \frac{p(a, b)}{p(a)p(b)} \quad (1)$$

where n is the number of variable samples; $p(a, b)$ represents the joint probability density function of A and B ; and $p(a)$ and $p(b)$ denote the edge probability density function of A and B , respectively.

Since scatter plots can be divided in many different ways, there are many mutual information values between the variables. Then, the maximum value in mutual information is selected and normalized to $[0, 1]$, which is shown as follows:

$$MIC(A, B) = \max_{a*b < G} \frac{I(A, B)}{\lg \min(A, B)} \quad (2)$$

where G denotes maximum grid size. It can be found from a large number of comparative experiments that the MIC achieves the highest operational efficiency and the most reliable results when $G = n^{0.6}$ [31].

2.1.2. Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN)

CEEMDAN is an adaptively decomposition technique [35], which is adopted to convert complex nonlinear wind speed time series into subsequences composed of different frequency domains. As a variant of EMD, CEEMDAN has the following advantages: (1) the subsequent noise levels are controlled by a noise coefficient vector; (2) the reconstructed subsequences are complete and have no noise; (3) compared with EMD, EEMD, and CEEMD, the number of iterations is lower. The detailed procedure for CEEMDAN is as follows:

Step 1: The original signal $x(t)$ with adaptive white noise is tested N times, and then the first mode component IMF_1 is calculated using the following formula:

$$IMF_1(t) = \frac{1}{N} \sum_{i=1}^N IMF_1^i(t) = \overline{IMF_1}(t) \quad (3)$$

Step 2: The first residual subsequence r_1 is calculated:

$$r_1(t) = x(t) - IMF_1(t) \quad (4)$$

Step 3: Adaptive white noise is added to r_1 , which is adopted to calculate the second modal component IMF_2 :

$$IMF_2(t) = \frac{1}{N} \sum_{i=1}^N E_1 \left(r_1(t) + \varepsilon_1 E_1(w^i(t)) \right) \quad (5)$$

where E_j is j -th mode generated by EMD, w^i is white noise with a normal distribution, and ε is the allowable noise deviation. Similarly, IMF_{k+1} is calculated as follows:

$$r_k(t) = r_{k-1}(t) - IMF_k(t) \quad (6)$$

$$IMF_{k+1}(t) = \frac{1}{N} \sum_{i=1}^N E_1 \left(r_k(t) + \varepsilon_k E_k(w^i(t)) \right) \quad (7)$$

Step 4: Step 3 is repeated until the residual sequence cannot be decomposed. The results of the final residual sequence and the decomposition results are successively expressed as follows:

$$R(t) = x(t) - \sum_{k=1}^M IMF_k(t) \quad (8)$$

$$x(t) = \sum_{k=1}^M IMF_k(t) + R(t) \quad (9)$$

where M is the number of subsequences.

2.2. The Proposed SWGMN for NWP Correction

Shared Weight Gated Memory Network (SWGMN)

As an improved type of RNN, LSTM introduces a core module based on the memory cell, which consists of three gate structures with different functions, including the forgetting gate, the input gate and the output gate [36,37]. The unit structure of LSTM is depicted in Figure 1.

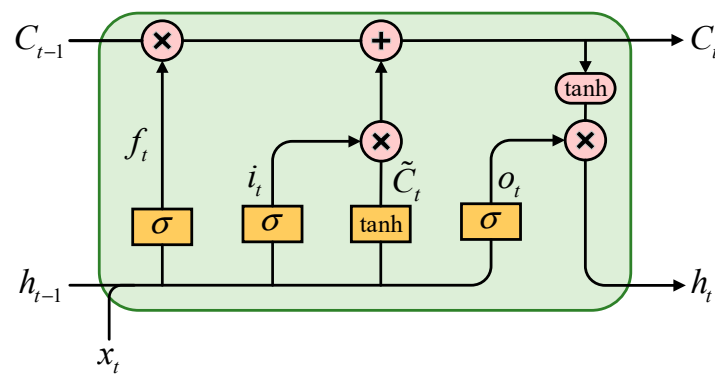


Figure 1. Structure of LSTM cell.

Although the prediction accuracy of the LSTM network is higher, its structure is more complex, which leads to more time taken in the training process [38,39]. Therefore, an improved network named SWGMN is proposed in this paper, which achieves a good prediction accuracy and can greatly reduce running time; thus, it is more suitable for NWP correction in the field of short-term wind speed forecasting. As a variant of LSTM, the SWGMN changes the gate structure by reorganizing the forgotten, input and output gates of LSTM into a shared gate, thereby reducing the running time significantly. The cell structure of SWGMN is shown in Figure 2, and the mathematical expressions are depicted as follows:

- (1) Calculate the shared gate and information status:

$$\tilde{x}_t = \tanh(W \cdot x_t + b) \quad (10)$$

$$r_t = \sigma(W \cdot x_t + b) \quad (11)$$

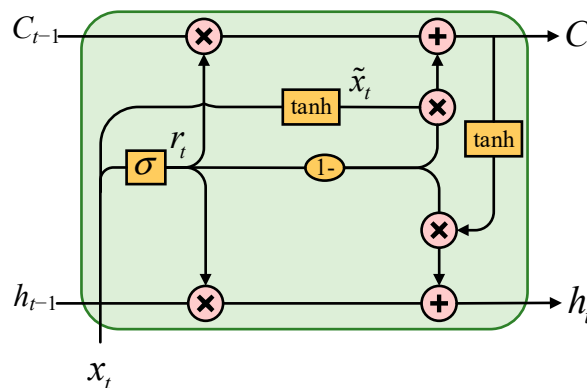


Figure 2. SWGMN structure.

- (2) Update the current module status:

$$C_t = r_t \otimes C_{t-1} + (1 - r_t) \otimes \tilde{x}_t \quad (12)$$

- (3) Calculate the output of the current module:

$$h_t = r_t \otimes h_{t-1} + (1 - r_t) \otimes \tanh(C_t) \quad (13)$$

where $\tilde{x}_t$ denotes the current input memory cell, r_t represents the shared gate output, W is the shared weight, and b is the shared bias.

As can be seen from the above formula and Figure 2, the SWGMN changes the gate structure of LSTM to make the network structure simpler. Specifically, shared gates in the

SWGMN are adopted to replace the input gate, the forget gate and the output gate in LSTM. Meanwhile, the SWGMN replaces the four different value weights W_f , W_i , W_C and W_o and the biases b_f , b_i , b_C and b_o in LSTM with one shared weight W and one shared bias b in turn. Therefore, the proposed SWGMN with a shared gate can control the discarding of useless historical information and retain current useful information, which achieves a satisfactory forecasting accuracy and short model training time, thus providing strong support for the correction NWP in the field of wind speed forecasting.

2.3. The Proposed INGO

2.3.1. The Proposed Improved Northern Goshawk Optimization (INGO)

Although NGO has a strong global search ability and a fast convergence rate [29], it may only determine the local optimal solution in the middle and late iterations. Moreover, the hunting radius of the northern goshawk does not decrease linearly with the increase in the number of them in nature. Therefore, improved northern goshawk optimization is proposed based on the levy flight disturbance strategy and the nonlinear contraction strategy.

For the development phase, INGO employs the proposed nonlinear contraction strategy to capture prey instead of the traditional linear contraction strategy, which is similar to the predatory behavior of the northern goshawk in nature, thereby accelerating the convergence rate. The proposed nonlinear contraction strategy is depicted as follows:

$$R = 0.01 \times \left[1 + \cos\left(\frac{\pi t}{T}\right) \right]^2 \quad (14)$$

where t represents current iteration, and T denotes maximum number of iterations.

Furthermore, the third stage, named the perturbation stage, is proposed in INGO to avoid finding local optimal solutions. In the perturbation stage, the levy flight disturbance strategy is introduced into the position update process of the northern goshawk, which can improve the local search ability and the convergence speed. Furthermore, perturbation factor r and judgment factor p are introduced in the levy flight disturbance strategy, where r is a random number in the range (0, 1), and p decreases as the number of iterations increases as follows:

$$p = 1 - \sqrt{t/T} \quad (15)$$

As the operating condition, the levy flight disturbance strategy is implemented in the current iteration when r is greater than p .

In the early stages of iteration, the levy flight disturbance strategy is not performed in early iteration since p is close to 1. Moreover, p decreases rapidly in the middle and late iterations, which employ the levy flight disturbance strategy to prevent INGO from finding local optimal solutions. The disturbed update of the northern goshawk position is depicted as follows:

$$X_i^{\text{new},S3} = X_i + (X_i - X_{\text{best}}) \otimes \text{Levy}(d) \quad (16)$$

$$X_i = \begin{cases} X_i^{\text{new},S3}, & F_i^{\text{new},S3} < F_i \\ X_i, & F_i^{\text{new},S3} \geq F_i \end{cases} \quad (17)$$

where $X_i^{\text{new},S3}$ is the i -th individual new position, $\text{Levy}(d)$ is levy flight disturbance strategy, d is dimensionality, $\otimes$ is multiplication, and $F_i^{\text{new},S3}$ is the i -th individual position objective function.

2.3.2. INGO Evaluation

In this section, some benchmark functions are selected to assess the performance of INGO, which include single-peak benchmark functions (F_1 , F_2 , F_3 , F_4 and F_7) and multi-peak benchmark functions (F_{10} , F_{11} , F_{15} and F_{23}). Among these, single-peak benchmark functions have a unique optimal solution, which are adopted to evaluate convergence,

while each multi-peak benchmark function has multiple local optimal solutions, which can verify the global searching ability. The selected benchmark functions are depicted in Table 1. To enhance this experiment's persuasiveness, the proposed INGO is compared with PSO, DE, GWO, BAS and NGO, whose parameters, including the running time, the population size and the maximum iterations, are set 20, 30 and 200 in turn. The convergence curves and evaluation indicators are depicted in Figure 3 and Table 2, respectively, where the evaluation indicators include the average value (Ave.) and the standard deviation (Std.) obtained from the results of 20 runs.

Table 1. Benchmark functions to verify the proposed INGO.

Function	Dim	Range	F _{min}
$F_1 = \sum_{i=1}^n x_i^2$	30	$[-100, 100]$	0
$F_2 = \sum_{i=1}^n x_i + \prod_{i=1}^n x_i $	30	$[-10, 10]$	0
$F_3 = \sum_{i=1}^n \left(\sum_{j=1}^i x_j \right)^2$	30	$[-100, 100]$	0
$F_4 = \max_i \{ x_i , 1 \leq i \leq n\}$	30	$[-100, 100]$	0
$F_7 = \sum_{i=1}^n ix_i^4 + \text{rand}[0, 1]$	30	$[-1.28, 1.28]$	0
$F_{10} = -20 \sum_{i=1}^n \exp(-0.2\sqrt{1/n} \sum_{i=1}^n x_i^2) - \exp(\sum_{i=1}^n \cos(2\pi x_i)/n) + 20 + e$	30	$[-32, 32]$	0
$F_{11} = \sum_{i=1}^n x_i^2/4000 - \prod_{i=1}^n \cos(x_i/\sqrt{i}) + 1$	30	$[-600, 600]$	0
$F_{15} = \sum_{i=1}^{11} [a_i - x_1(b_i^2 + b_i x_2)/(b_i^2 + b_i x_3 + x_4)]^2$	4	$[-5, 5]$	0.0003
$F_{23} = -\sum_{i=1}^{10} [(X - a_i)(X - a_i)^T + c_i]^{-1}$	4	$[0, 10]$	−10.536

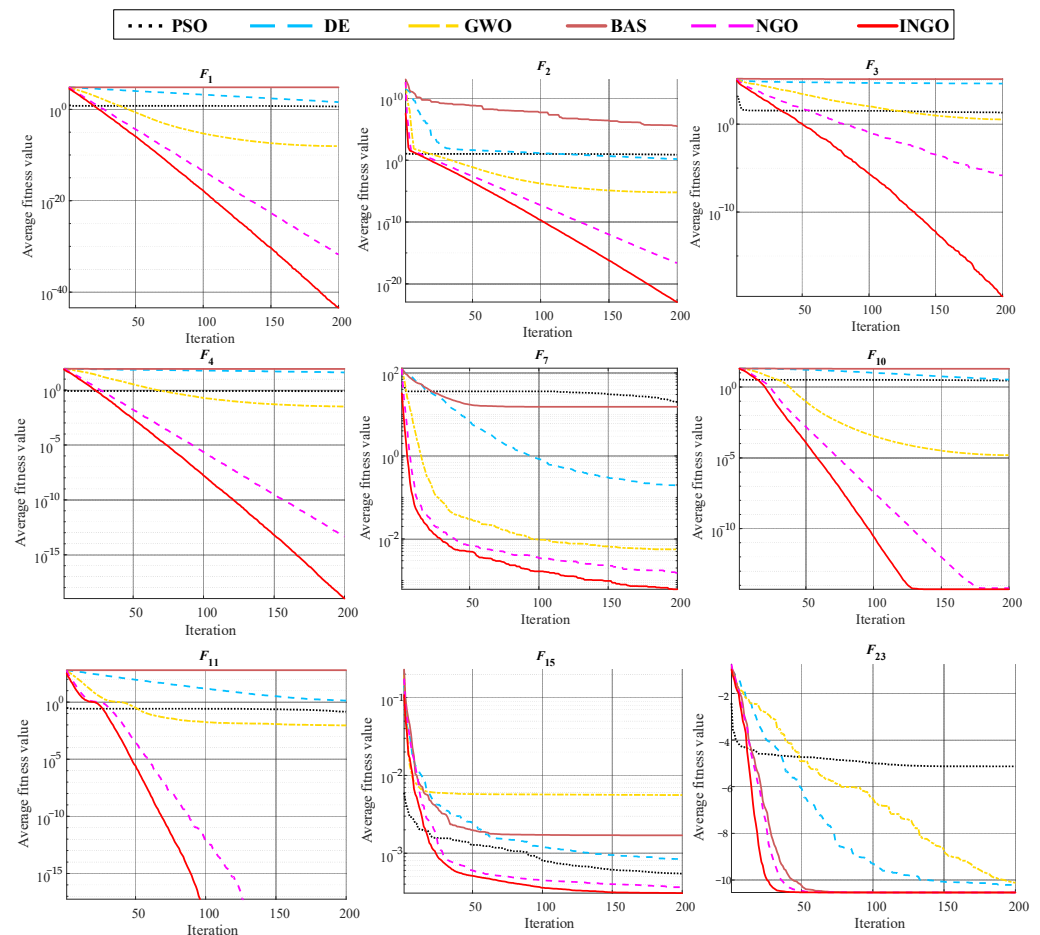


Figure 3. Convergence curves of INGO and comparative ones.

Table 2. Evaluation indicators of INGO and comparative ones.

Function	Metric	INGO	NGO	PSO	DE	GWO	BAS
F1	Ave.	7.78×10^{-45}	2.20×10^{-32}	3.50	38.40	6.44×10^{-9}	6.24×10^4
	Std.	1.58×10^{-44}	3.84×10^{-32}	0.87	12.00	8.13×10^{-9}	7.91×10^3
F2	Ave.	6.85×10^{-24}	2.10×10^{-17}	7.6	1.65	7.26×10^{-6}	3.02×10^5
	Std.	4.75×10^{-24}	1.66×10^{-17}	1.37	0.23	3.19×10^{-6}	8.18×10^5
F3	Ave.	3.68×10^{-19}	1.95×10^{-7}	17.90	4.12×10^4	3.9	1.38×10^5
	Std.	1.07×10^{-18}	2.27×10^{-7}	8.71	5.54×10^3	4.94	5.42×10^4
F4	Ave.	7.70×10^{-20}	3.75×10^{-14}	0.82	40.80	0.03	85.90
	Std.	4.87×10^{-20}	2.53×10^{-14}	0.04	5.54	0.02	2.10
F7	Ave.	9.07×10^{-4}	0.14×10^{-2}	18.40	0.19	0.66×10^{-2}	15.20
	Std.	3.61×10^{-4}	5.86×10^{-4}	6.83	0.04	0.37×10^{-2}	5.03
F10	Ave.	4.61×10^{-15}	5.51×10^{-15}	2.9	3.2	2.08×10^{-5}	19.60
	Std.	7.94×10^{-16}	1.67×10^{-15}	0.2	0.2	1.29×10^{-5}	0.21
F11	Ave.	0	0	0.14	1.3	0.01	6.19×10^2
	Std.	00	0	0.04	0.10	1.49×10^{-2}	57.70
F15	Ave.	3.07×10^{-4}	3.51×10^{-44}	5.93×10^{-4}	9.39×10^{-4}	0.36×10^{-2}	0.17×10^{-2}
	Std.	8.82×10^{-10}	4.27×10^{-5}	1.58×10^{-4}	3.44×10^{-4}	0.72×10^{-2}	6.17×10^{-4}
F23	Ave.	−10.5364	−10.2587	−5.1285	−10.2142	−9.7121	−10.5364
	Std.	2.47×10^{-15}	1.24	1.27×10^{-5}	0.59	2.49	2.60×10^{-8}

It can be seen from Figure 3 that for the single-peak benchmark functions (F_1 , F_2 , F_3 , F_4 and F_7), the proposed INGO optimization algorithm has a faster convergence speed. Although each algorithm can eventually approach 0, the final result of the INGO algorithm is significantly smaller than those of the other algorithms and closer to 0. This shows that the INGO optimization algorithm has a faster and stronger optimization ability. For the multi-peak benchmark functions (F_{10} , F_{11} , F_{15} and F_{23}), as they are more complex than the single-peak benchmark functions, each of which has multiple different local optimal solutions. From Figure 3, it can be seen that some optimization algorithms will find local optimal solutions and search slowly, while the proposed INGO optimization algorithm can find the global optimal solutions of each multi-peak benchmark function with the fastest speed, proving the proposed INGO achieves a stronger global optimization ability and better performance. As can be seen from Table 2, the Ave. and Std. of the proposed INGO for all the benchmark functions are the smallest among all the comparison algorithms. Specifically, the average value of the optimal solution obtained by INGO is closest to the optimal solution of the benchmark function in 20 runs, which demonstrates that the deviation of all the running results is within a satisfactory range, proving that INGO achieves a stronger optimization ability than those of the other algorithms. On the other hand, the standard deviation of INGO is significantly lower than those of the other optimization algorithms, which indicates that the dispersion degree of the optimal solution is minimal, proving that the proposed INGO achieves strong robustness and satisfactory stability.

3. Architecture of the Proposed Framework Coupling MIC, CEEMDAN, Shared Weight Gated Memory Network with Improved Northern Goshawk Optimization for NWP Correction

In this section, a compound short-term wind speed forecasting framework for NWP correction coupling the MIC, CEEMDAN and the SWGMN with INGO is proposed for NWP correction, as illustrated in Figure 4. The implementation steps are depicted as follows:

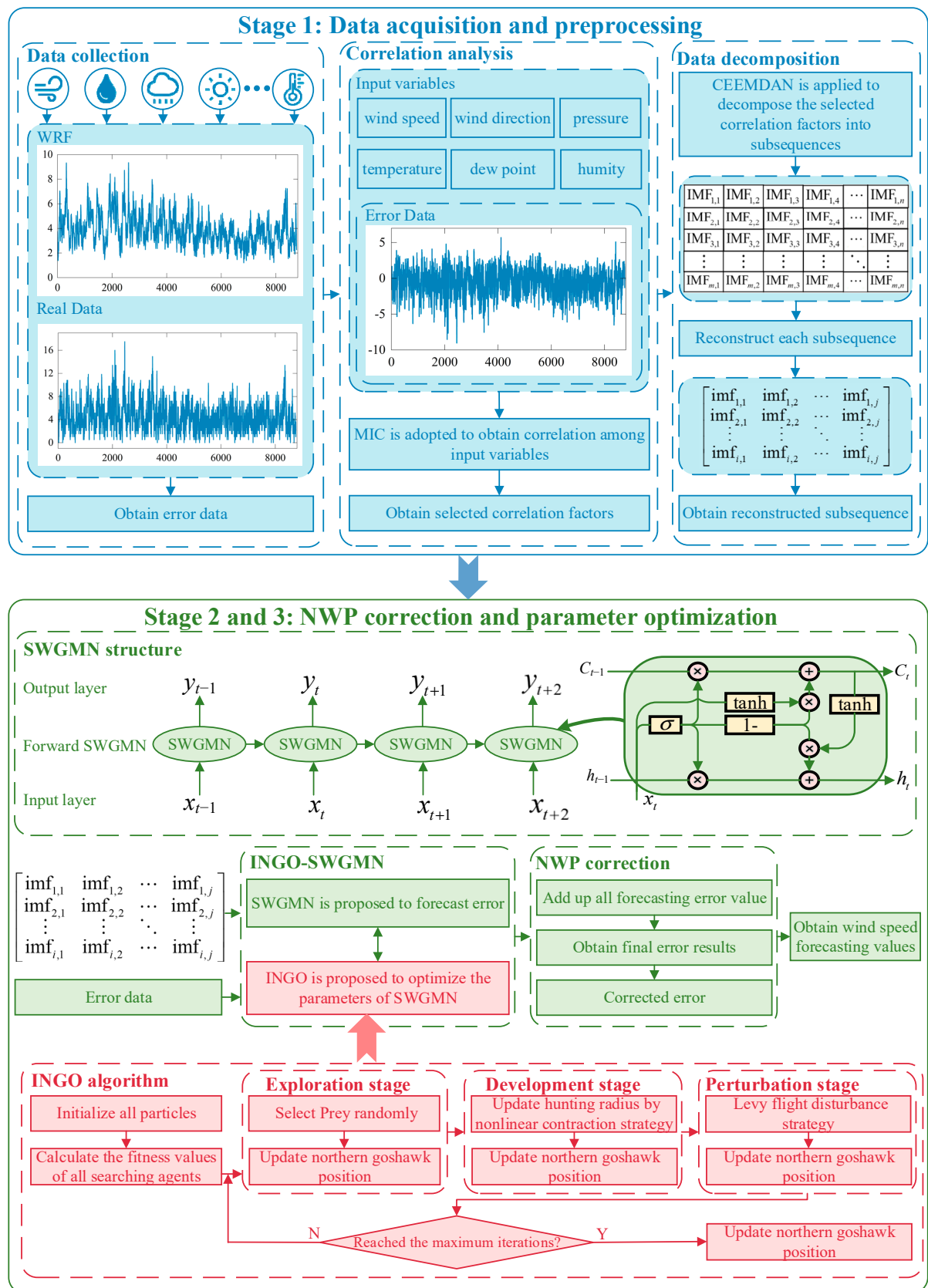


Figure 4. The implementation steps of the proposed compound framework.

Step 1: The predicted and actual variables are acquired by NWP and Open Weather, respectively, in which the predicted variables with different domains include the predicted wind speed and other meteorological variables.

Step 2: The MIC is employed to obtain the correlation between the predicted variables and the error, which is calculated using the actual and predicted wind speeds, after which the correlation factors strongly related to the error are selected from all the variables.

Step 3: The error and the correlation factors are decomposed into a series of subsequences by CEEMDAN.

Step 4: The proposed SWGMN is applied to forecast the error of each subsequence, respectively, in which the shared gate of the SWGMN is proposed to replace the input gate, the forgetting gate and the output gate of LSTM.

Step 5: The proposed INGO is employed to optimize the parameters of the SWGMN, which is composed of NGO, the levy flight disturbance strategy and the nonlinear contraction strategy.

Step 6: The wind speed forecasting results are attained by accumulating the forecasting error of all the subsequences and the predicted wind speed to achieve NWP correction.

4. Experimental Results and Analysis

4.1. Data Description and Processing

The experimental data are collected from Open Weather during the period from 8 October 2017 to 23 March 2018, including the actual and NWP values with a resolution of 1 h. Within this, the actual data contain the measured values of meteorological data, such as temperature, pressure and wind speed, and NWP includes the corresponding predicted values. Moreover, the first 80 percent of the data are applied as training set, and the last 20 percent of data are used as a test set. Furthermore, the time dimensions in NWP are 0:00 and 12:00. The predicted wind speed from NWP at 0:00, the actual wind speed, and the error calculated using the predicted wind speed and the actual wind speed are depicted in Figure 5. It can be seen that the predicted wind speed from NWP can roughly reflect the actual wind speed trend, but there is still a certain deviation, in which the maximum deviation is about 17 m/s. Therefore, it is necessary to correct the predicted wind speed. This correction task is completed using Python 3.7.

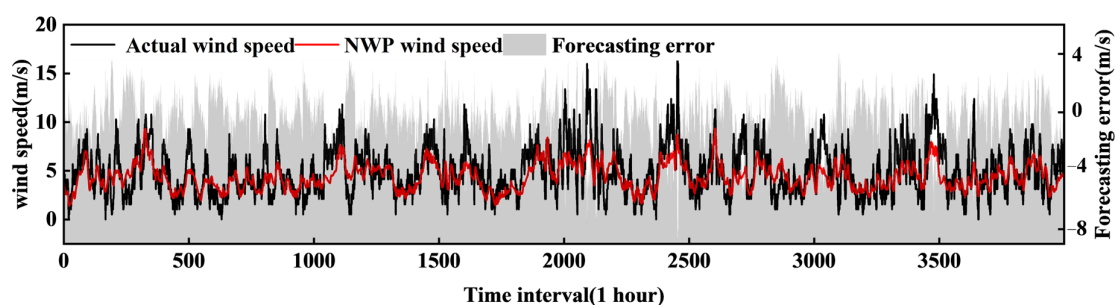


Figure 5. NWP predicting wind speed and error.

Moreover, the meteorological data contain 12 factors, among which there are unnecessary factors that increase the computational complexity and reduce the forecasting accuracy. To this end, the MIC is employed to select the input meteorological data. The correlation between the meteorological data and the error is shown as Figure 6. It can be found that the predicted wind speeds at 0:00 and 12:00 have the most significant correlation with the error. Therefore, the predicted wind speeds at 0:00 and 12:00, the prediction error, and the actual wind speed are considered as the input factors. Since the wind speed and the error have significant volatility, CEEMDAN is applied to decompose the input factors to improve the prediction accuracy. The decomposition results are shown in Figure 7.

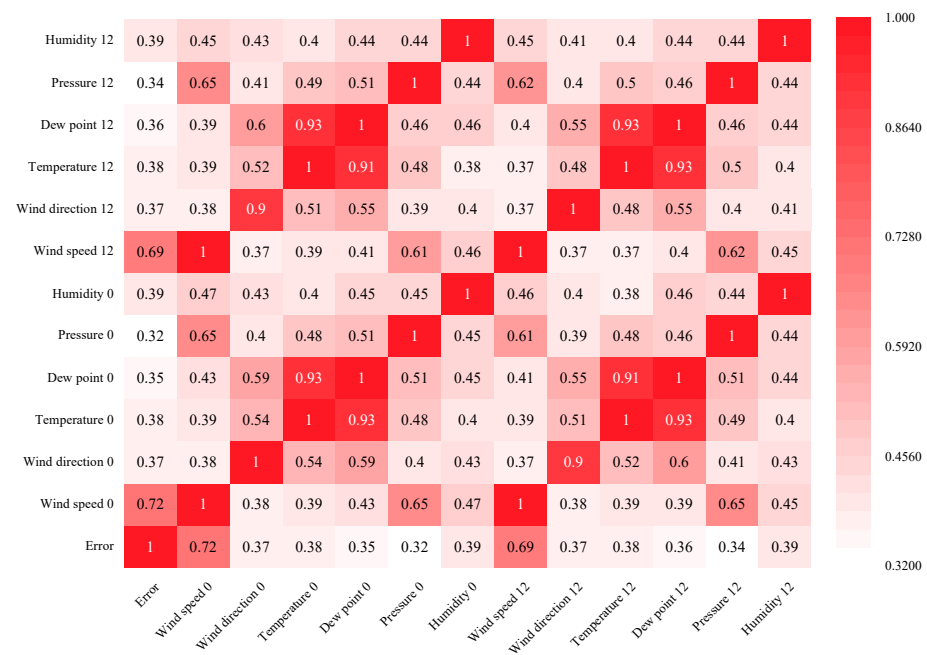


Figure 6. Correlation of various factors.

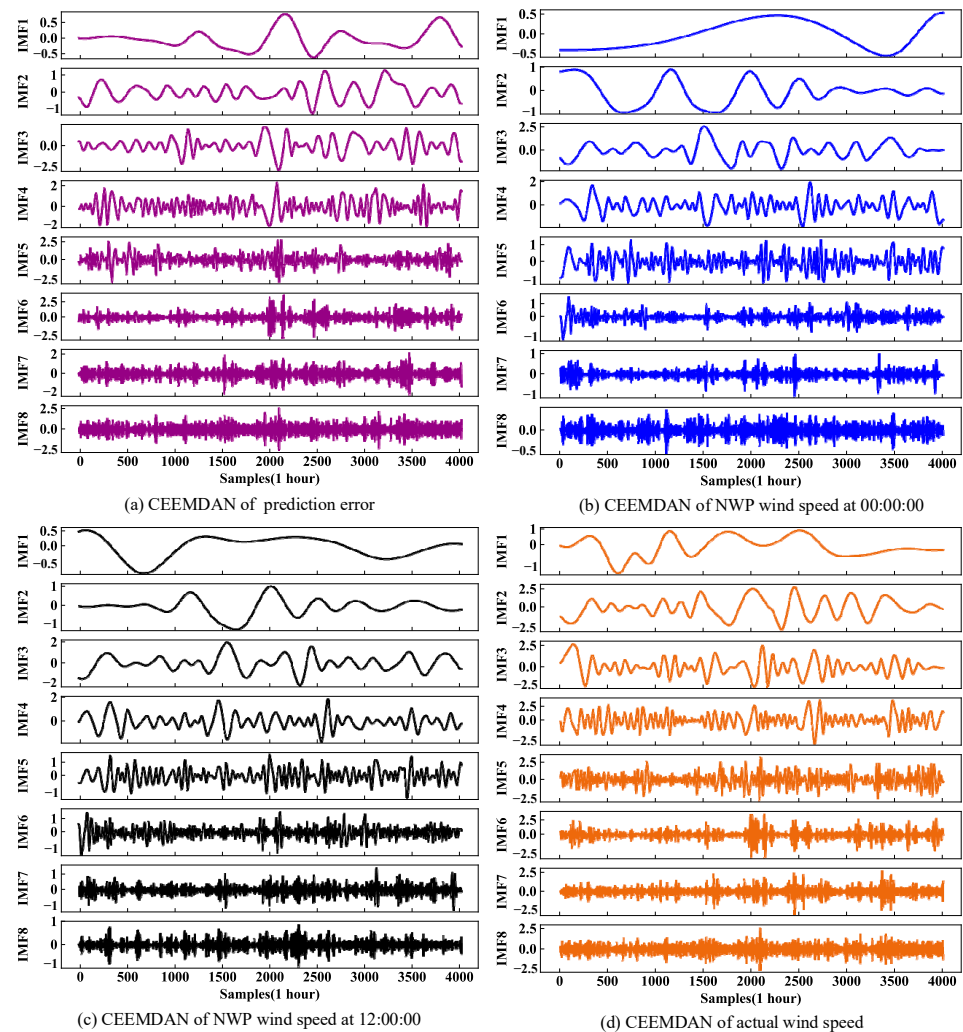


Figure 7. The results of the CEEMDAN of the input variables.

4.2. Parameter Setting

To verify the prediction performance of the proposed model, called the MIC-CEEMDAN-INGO-SWGMN, seven models, including BP, the SWGMN, the MIC-SWGMN, MIC-CEEMDAN-LSTM, MIC-CEEMDAN-GRU, the MIC-CEEMDAN-SWGMN and the MIC-CEEMDAN-INGO-SWGMN, are established for comparison. The parameter settings of each model are shown in Table 3. For the proposed model, INGO is proposed to optimize the initial learning rate and the number of iterations of the SWGMN, in which the optimization range is between [0, 1] and [100, 300], in turn, where the maximum number of iterations and population size of INGO are set to 50 and 30, respectively. The maximum number of iterations and the total number of CEEMDAN are determined by the trial and error method.

Table 3. Parameter settings of all laboratorial models.

Models	Parameters	Determination Methods	Values/Research Range
BP	Number of hidden neurons	Trial and error method	128
	Batch size	Trial and error method	16
	Epochs of training	Trial and error method	100
LSTM	Epochs of training	Trial and error method	100
	Initial learning rate	Trial and error method	0.01
	Number of hidden neurons	Trial and error method	128
	Batch size	Trial and error method	16
GRU	Epochs of training	Trial and error method	100
	Initial learning rate	Trial and error method	0.01
	Number of hidden neurons	Trial and error method	128
	Batch size	Trial and error method	16
SWGMN	Epochs of training	INGO	[100, 300]
	Initial learning rate	INGO	[0, 1]
	Number of hidden layer nodes	Trial and error method	128
	Batch size	Trial and error method	16
INGO	Population number	Present	30
	Maximum iterations	Present	50
CEEMDAN	Total number of times	Trial and error method	50
	Maximum number of filtering iterations	Trial and error method	500

4.3. Evaluation Indicators

To quantitatively measure the performance of the prediction model, the mean absolute error (MAE), the root mean square error (RMSE) and the mean absolute percentage error (MAPE) are adopted to analyze the error [40]. The smaller the three indicators are, the smaller the prediction error of the model is. The expression is as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (18)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (19)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% \quad (20)$$

where n represents the number of prediction samples, y_i denotes the actual value of the i -th sample, and $\hat{y}_i$ denotes the predicted value for the i -th sample.

4.4. Comparative Analysis of Prediction Performance

Table 4 presents the evaluation indicator of the seven prediction models used to correct the errors. In addition, a correction curve fitting diagram is shown in Figure 8. Combining Table 4 and Figure 8, the following conclusions can be drawn: (1) It can be seen that the corrected wind speed has a remarkable fit with the actual wind speed. Table 4 shows that the wind speed forecast with correction is more accurate than that without correction, and its *MAE*, *RMSE* and *MAPE* values are reduced by 1.1367 m/s, 1.4264 m/s and 69.95%, respectively. (2) The proposed MIC-CEEMDAN-INGO-SWGMN achieves the best correction effect with the smallest *MAE*, *RMSE* and *MAPE* of 0.4371 m/s, 0.5743 m/s and 10.771%, respectively. (3) Compared with the prediction results of the SWGMN, the *MAE*, *RMSE* and *MAPE* of the MIC-SWGMN are better, which indicates the MIC can effectively select input variables to improve the prediction accuracy of the model. (4) By comparing the prediction results of the MIC-CEEMDAN-SWGMN and the MIC-SWGMN, it can be seen that data volatility can be significantly reduced by using CEEMDAN, thus improving the prediction accuracy. (5) Compared with the prediction results of MIC-CEEMDAN-LSTM and MIC-CEEMDAN-GRU, it is found that the *MAE*, *RMSE* and *MAPE* of the MIC-CEEMDAN-SWGMN are the smallest, indicating that the proposed SWGMN has the best performance. Moreover, Table 5 illustrates that the training and prediction times of the MIC-CEEMDAN-SWGMN are about 44.01% and 34.54% shorter compared with those of MIC-CEEMDAN-LSTM and MIC-CEEMDAN-GRU. (6) Comparing the MIC-CEEMDAN-INGO-SWGMN with the MIC-CEEMDAN-SWGMN, it can be seen that the model has a better prediction performance after the addition of INGO, indicating that the hyperparameters of the INGO model are effective.

Table 4. The evaluation results of the proposed model and the comparison ones.

Models	Evaluation Indicators		
	<i>MAE</i> (m/s)	<i>RMES</i> (m/s)	<i>MAPE</i> (%)
Without correction	1.5738	2.0007	35.8483
BP	1.5838	1.8675	28.0752
SWGMN	0.8933	1.1540	27.7089
MIC-SWGMN	0.8618	1.1135	19.8645
MIC-CEEMDAN-LSTM	0.4989	0.6435	12.2026
MIC-CEEMDAN-GRU	0.4733	0.6142	14.5071
MIC-CEEMDAN-SWGMN	0.4417	0.5839	11.5129
MIC-CEEMDAN-INGO-SWGMN	0.4371	0.5743	10.7717

Table 5. Training and prediction times for all subsequences.

Models	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	Total Time
MIC-CEEMDAN-LSTM	73.83	72.37	78.11	77.65	78.82	74.84	74.36	75.52	72.88	678.38
MIC-CEEMDAN-GRU	64.43	65.26	64.37	64.14	65.24	64.79	65.15	63.56	63.23	580.17
MIC-CEEMDAN-SWGMN	42.70	42.94	43.16	41.04	42.61	41.26	41.73	42.06	42.27	379.77

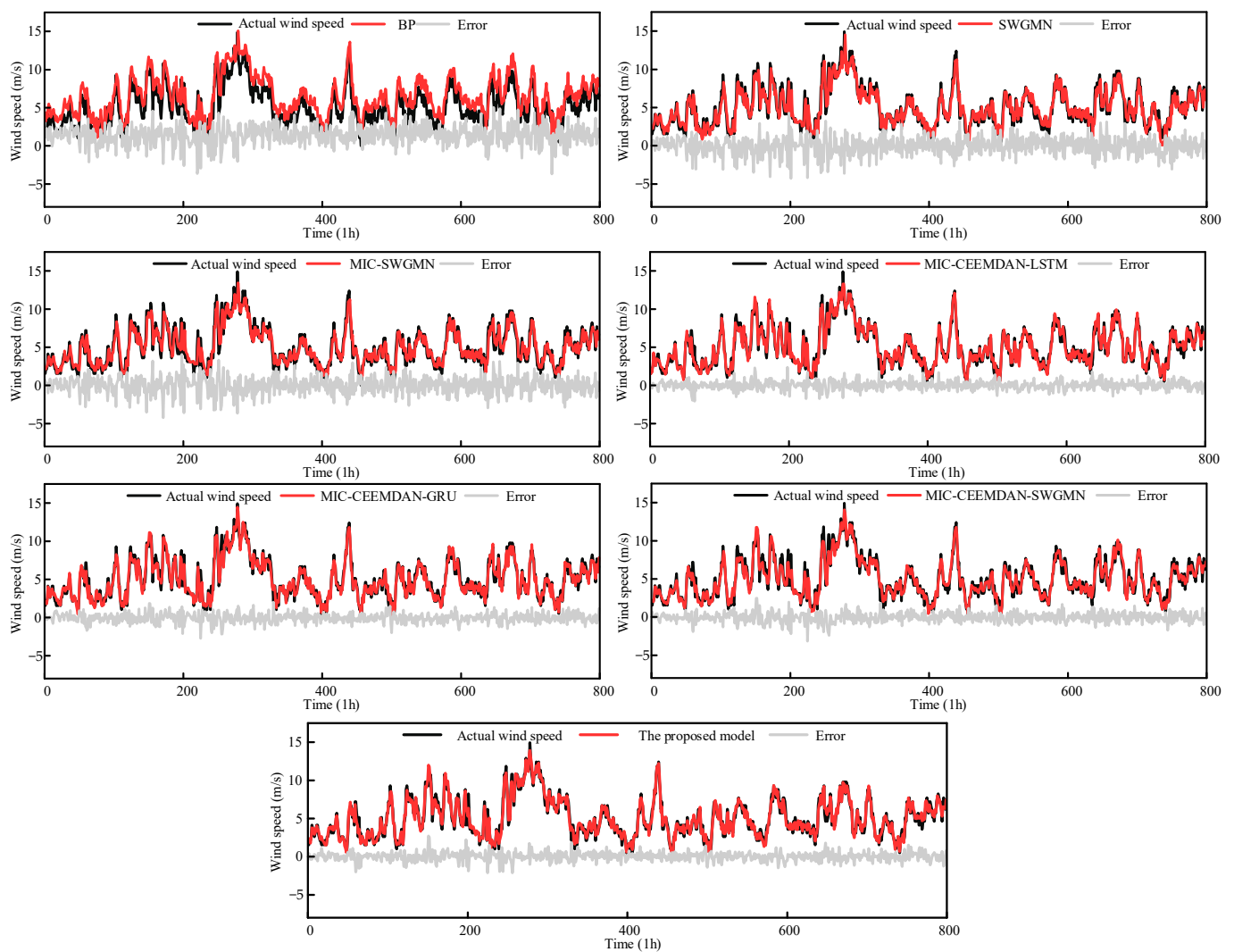


Figure 8. The corrected results of the proposed model and the other comparison models.

4.5. Discussion

4.5.1. Discussion on the Effectiveness of MIC

According to Table 6, it can be seen that the forecasting performance achieves different degrees of improvement after using the MIC. Compared with the SWGMN, the *MAE*, *RMSE* and *MAPE* of the MIC-SWGMN are reduced by 3.52%, 3.5% and 28.31%, respectively, which indicates the MIC can effectively select the input variables.

Table 6. Percent performance improvement.

Improvement Percentages	<i>PI</i> _{index} (%)		
	<i>PI</i> _{MAE} (%)	<i>PI</i> _{RMSE} (%)	<i>PI</i> _{MAPE} (%)
SWGMN vs. MIC-SWGMN	3.52	3.50	28.31
MIC-SWGMN vs. MIC-CEEMDAN-SWGMN	48.74	47.56	42.04
MIC-CEEMDAN-LSTM vs. MIC-CEEMDAN-SWGMN	11.47	9.26	5.65
MIC-CEEMDAN-SWGMN vs. MIC-CEEMDAN-INGO-SWGMN	1.04	1.64	6.44

4.5.2. Discussion on the Effectiveness of CEEMDAN

Table 6 shows the percentage of performance improvement with the model after adding CEEMDAN. Compared with the MIC-SWGMN, the *MAE*, *RMSE* and *MAPE* of the MIC-CEEMDAN-SWGMN are increased by 48.74%, 47.56% and 42.04%, respectively, which indicates CEEMDAN can effectively decompose the original series into a series of subseries to reduce the volatility of the series.

4.5.3. Discussion on the Effectiveness of INGO

Figure 3 illustrates that the proposed INGO achieves superior optimization results within the shortest time compared with those of the other optimization algorithms. Table 3 demonstrates that the proposed INGO outperforms other optimization algorithms in terms of optimization effectiveness. From Table 6, the *MAPE* of the MIC-CEEMDAN-SWGMN is improved by 6.44% after the addition of INGO. The results show that INGO achieves a robust global search ability and can effectively calibrate the parameters of the SWGMN.

4.5.4. Discussion on the Effectiveness of SWGMN

From Table 6, it is apparent that the evaluation indicators of the MIC-CEEMDAN-SWGMN are better than those of MIC-CEEMDAN-LSTM, and the total training time of all the subsequences is reduced by 44.01%. It shows that the proposed SWGMN with a shared gate, a shared bias and a shared weight can effectively reduce the training time and improve the prediction accuracy.

5. Conclusions

To improve the NWP forecast accuracy, a compound framework is proposed by coupling the MIC, CEEMDAN and the SWGMN with INGO for NWP correction. Firstly, numerical weather prediction is employed to obtain the predicted variables, which include the predicted wind speed and other meteorological variables. Then, the MIC is applied to select the correlation factors based on the correlation between the predicted variables and the error, which is calculated using the predicted and actual wind speeds. Afterwards, the correlation factors are decomposed into a set of subsequences with CEEMDAN. Subsequently, the SWGMN, as a variant of LSTM, is proposed to predict the error of each subsequence, in which the shared gate of the SWGMN is proposed to replace the input gate, the forgetting gate and the output gate of LSTM. Meanwhile, the proposed INGO is adopted to calibrate the optimal parameters of the SWGMN. Lastly, the wind speed forecasting values are obtained using the predicted error of all the subsequences and the predicted wind speed from NWP. Through these experiments and comparative analysis, the following conclusions are drawn: (1) The MIC can effectively eliminate the irrelevant variables and retain the main factors of affect the error to reduce the training complexity of the model. (2) CEEMDAN can reduce data volatility, thereby improving the prediction accuracy of the model. (3) INGO is superior to PSO, DE, GWO, BAS and NGO and can provide sufficient support for the SWGMN. (4) The proposed model SWGMN simplifies the gate structure and shares the internal weights to achieve a higher prediction accuracy and a significantly reduced running time, which is more suitable for short-term wind speed prediction error correction.

Author Contributions: Conceptualization, Y.L. and H.Z.; methodology, C.W.; software, C.W.; validation, Y.L., H.Z. and C.W.; formal analysis, M.S.; investigation, L.Z.; resources, M.S.; data curation, C.W.; writing—original draft preparation, Y.L. and C.W.; writing—review and editing, W.F.; visualization, C.W.; supervision, W.F.; project administration, Y.L. and H.Z.; funding acquisition, W.F. and H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This paper was supported by the 2023 Open Research Fund of Hubei Key Laboratory of Intelligent Yangtze and Hydroelectric Science (No. 242202000901) and the Hubei Natural Science Foundation (No. 2022CFD170).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data of this paper are available at <https://openweathermap.org/api#current> (accessed on 13 April 2024).

Conflicts of Interest: Yanghe Liu, Hairong Zhang and Liting Zhou were employed by the company China Yangtze Power Co., Ltd. The remaining authors declare that this research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Abbreviations

AI	Artificial Intelligence
ANN	Artificial Neural Network
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
Ave.	Average Value
BAS	Beetle Antennae Search
CEEMD	Complementary Ensemble Empirical Mode Decomposition
CEEMDAN	Complete Ensemble Empirical Mode Decomposition
CEEMDAN	Complete Ensemble Empirical Mode Decomposition with Adaptive Noise
DE	Differential Evolution
EMD	Empirical Mode Decomposition
EEMD	Ensemble Empirical Mode Decomposition
MM5	Fifth-Generation Mesoscale Model
f-ARIMA	Fractional Auto Regressive Integrated Moving Average
GWO	Grey Wolf Optimizer
HIRLAM	High-Resolution Limited Area Model
INGO	Improved Northern Goshawk Optimization
LSTM	Long Short-Term Memory
MIC	Maximum Information Coefficient
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
NGO	Northern Goshawk Optimization
NWP	Numerical Weather Prediction
PSO	Particle Swarm Optimization
RF	Random Forest
RMSE	Root Mean Square Error
SWGMMN	Shared Weight Gated Memory Network
Std.	Standard Deviation
SVR	Support Vector Regression
WRF	Weather Research and Forecasting

References

1. Wang, J.; Wang, S.; Zeng, B.; Lu, H. A novel ensemble probabilistic forecasting system for uncertainty in wind speed. *Appl. Energy* **2022**, *313*, 118796. [CrossRef]
2. Agga, A.; Abbou, A.; Labbadi, M.; El Houm, Y. Short-term self consumption PV plant power production forecasts based on hybrid CNN-LSTM, ConvLSTM models. *Renew. Energy* **2021**, *177*, 101–112. [CrossRef]
3. Zhang, S.; Chen, Y.; Xiao, J.; Zhang, W.; Feng, R. Hybrid wind speed forecasting model based on multivariate data secondary decomposition approach and deep learning algorithm with attention mechanism. *Renew. Energy* **2021**, *174*, 688–704. [CrossRef]
4. Wang, C.; Wang, Z.; Chu, S.; Ma, H.; Yang, N.; Zhao, Z.; Lai, C.; Lai, L. A two-stage underfrequency load shedding strategy for microgrid groups considering risk avoidance. *Appl. Energy* **2024**, *367*, 123343. [CrossRef]
5. Wang, A.; Xu, L.; Li, Y.; Xing, J.; Chen, X.; Liu, K.; Liang, Y.; Zhou, Z. Random-forest based adjusting method for wind forecast of WRF model. *Comput. Geosci.* **2021**, *155*, 104842. [CrossRef]
6. Liu, H.; Chen, C. Data processing strategies in wind energy forecasting models and applications: A comprehensive review. *Appl. Energy* **2019**, *249*, 392–408. [CrossRef]
7. Huang, Z.; Gu, M. Characterizing nonstationary wind speed using the ARMA-GARCH model. *J. Struct. Eng.* **2019**, *145*, 04018226. [CrossRef]

8. Wang, L.; Li, X.; Bai, Y. Short-term wind speed prediction using an extreme learning machine model with error correction. *Energy. Convers. Manag.* **2018**, *162*, 239–250. [[CrossRef](#)]
9. Yatiyana, E.; Rajakaruna, S.; Ghosh, A. Wind speed and direction forecasting for wind power generation using ARIMA model. In Proceedings of the 2017 Australasian Universities Power Engineering Conference (AUPEC), Melbourne, VIC, Australia, 19–22 November 2017. [[CrossRef](#)]
10. Hua, L.; Zhang, C.; Peng, T.; Ji, C.; Nazir, M. Integrated framework of extreme learning machine (ELM) based on improved atom search optimization for short-term wind speed prediction. *Energy. Convers. Manag.* **2022**, *252*, 115102. [[CrossRef](#)]
11. Valkonen, T.; Schyberg, H.; Figa-Saldaña, J. Assimilating advanced scatterometer winds in a high-resolution limited area model over northern Europe. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2394–2405. [[CrossRef](#)]
12. Posada, R.; García-Ortega, E.; Sánchez, J.; López, L. Verification of the MM5 model using radiosonde data from Madrid-Barajas Airport. *Atmos. Res.* **2013**, *122*, 174–182. [[CrossRef](#)]
13. Qin, Y.; Liu, Y.; Jiang, X.; Yang, L.; Xu, H.; Shi, Y.; Huo, Z. Grid-to-point deep-learning error correction for the surface weather forecasts of a fine-scale numerical weather prediction system. *Atmosphere* **2023**, *14*, 145. [[CrossRef](#)]
14. Xiong, X.; Zou, R.; Sheng, T.; Zeng, W.; Ye, X. An ultra-short-term wind speed correction method based on the fluctuation characteristics of wind speed. *Energy* **2023**, *283*, 129012. [[CrossRef](#)]
15. Han, Y.; Mi, L.; Shen, L.; Cai, C.; Liu, Y.; Li, K.; Xu, G. A short-term wind speed prediction method utilizing novel hybrid deep learning algorithms to correct numerical weather forecasting. *Appl. Energy* **2022**, *312*, 118777. [[CrossRef](#)]
16. Chen, Y.; Bai, M.; Zhang, Y.; Liu, J.; Yu, D. Multivariable space-time correction for wind speed in numerical weather prediction (NWP) based on ConvLSTM and the prediction of probability interval. *Earth. Sci. Inform.* **2023**, *16*, 1953–1974. [[CrossRef](#)]
17. Wu, Y.-K.; Huang, C.-L.; Wu, S.-H.; Hong, J.-S.; Chang, H.-L. Deterministic and probabilistic wind power forecasts by considering various atmospheric models and feature engineering approaches. *IEEE Trans. Ind. Appl.* **2022**, *59*, 192–206. [[CrossRef](#)]
18. Wang, J.; Yang, Z. Ultrashort-term wind speed forecasting using an optimized artificial intelligence algorithm. *Renew. Energy* **2021**, *171*, 1418–1435. [[CrossRef](#)]
19. Huang, Y.; Wen, B.; Liao, W.; Shan, Y.; Fu, W.; Wang, R. Image enhancement based on dual-branch generative adversarial network combining spatial and frequency domain information for imbalanced fault diagnosis of rolling bearing. *Symmetry* **2024**, *16*, 512. [[CrossRef](#)]
20. Buhan, S.; Çadırcı, I. Multistage wind-electric power forecast by using a combination of advanced statistical methods. *IEEE Trans. Ind. Inform.* **2015**, *11*, 1231–1242. [[CrossRef](#)]
21. Cai, H.; Jia, X.; Feng, J.; Yang, Q.; Hsu, Y.; Chen, Y.; Lee, J. A combined filtering strategy for short term and long term wind speed prediction with improved accuracy. *Renew. Energy* **2019**, *136*, 1082–1090. [[CrossRef](#)]
22. Peng, X.; Wang, H.; Lang, J.; Li, W.; Xu, Q.; Zhang, Z.; Cai, T.; Duan, S.; Liu, F.; Li, C. EALSTM-QR: Interval wind-power prediction model based on numerical weather prediction and deep learning. *Energy* **2021**, *220*, 119692. [[CrossRef](#)]
23. Moosavi, A.; Rao, V.; Sandu, A. Machine learning based algorithms for uncertainty quantification in numerical weather prediction models. *J. Comput. Sci.* **2021**, *50*, 101295. [[CrossRef](#)]
24. Cai, H.; Jia, X.; Feng, J.; Li, W.; Hsu, Y.; Lee, J. Gaussian process regression for numerical wind speed prediction enhancement. *Renew. Energy* **2020**, *146*, 2112–2123. [[CrossRef](#)]
25. Liao, W.; Fu, W.; Yang, K.; Tan, C.; Huang, Y. Multi-scale residual neural network with enhanced gated recurrent unit for fault diagnosis of rolling bearing. *Meas. Sci. Technol.* **2024**, *35*, 056114. [[CrossRef](#)]
26. Fu, W.; Yang, K.; Wen, B.; Shan, Y.; Li, S.; Zheng, B. Rotating machinery fault diagnosis with limited multisensor fusion samples by fused attention-guided wasserstein GAN. *Symmetry* **2024**, *16*, 285. [[CrossRef](#)]
27. Xu, W.; Liu, P.; Cheng, L.; Zhou, Y.; Xia, Q.; Gong, Y.; Liu, Y. Multi-step wind speed prediction by combining a WRF simulation and an error correction strategy. *Renew. Energy* **2021**, *163*, 772–782. [[CrossRef](#)]
28. Zhang, Y.; Wang, H. Multi-head attention-based probabilistic CNN-BiLSTM for day-ahead wind speed forecasting. *Energy* **2023**, *278*, 127865. [[CrossRef](#)]
29. Dehghani, M.; Hubálovský, Š.; Trojovský, P. Northern goshawk optimization: A new swarm-based algorithm for solving optimization problems. *IEEE Access* **2021**, *9*, 162059–162080. [[CrossRef](#)]
30. Fu, W.; Fu, Y.; Li, B.; Zhang, H.; Zhang, X.; Liu, J. A compound framework incorporating improved outlier detection and correction, VMD, weight-based stacked generalization with enhanced DESMA for multi-step short-term wind speed forecasting. *Appl. Energy* **2023**, *348*, 121587. [[CrossRef](#)]
31. Reshef, D.N.; Reshef, Y.A.; Finucane, H.K.; Grossman, S.R.; McVean, G.; Turnbaugh, P.J.; Lander, E.S.; Mitzenmacher, M.; Sabeti, P.C. Detecting novel associations in large data sets. *Science* **2011**, *334*, 1518–1524. [[CrossRef](#)] [[PubMed](#)]
32. Zhang, C.; Hu, H.; Ji, J.; Liu, K.; Xi, X.; Nazir, M.; Peng, T. An evolutionary stacked generalization model based on deep learning and improved grasshopper optimization algorithm for predicting the remaining useful life of PEMFC. *Appl. Energy* **2023**, *330*, 120333. [[CrossRef](#)]
33. Tang, X.; Chen, H.; Xiang, W.; Wang, J.; Zou, M. Short-term load forecasting using channel and temporal attention based temporal convolutional network. *Electr. Power Syst. Res.* **2022**, *205*, 107761. [[CrossRef](#)]
34. Lin, G.; Lin, A.; Gu, D. Using support vector regression and K-nearest neighbors for short-term traffic flow prediction based on maximal information coefficient. *Inform. Sci.* **2022**, *608*, 517–531. [[CrossRef](#)]

35. Torres, M.; Colominas, M.; Schlotthauer, G.; Flandrin, P. A complete ensemble empirical mode decomposition with adaptive noise. In Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, 22–27 May 2011; pp. 4144–4147. [[CrossRef](#)]
36. Yu, Y.; Si, X.; Hu, C.; Zhang, J. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput.* **2019**, *31*, 1235–1270. [[CrossRef](#)] [[PubMed](#)]
37. Chen, T.; Wan, W.; Li, X.; Qin, H.; Yan, W. Flexible Load Multi-Step Forecasting Method Based on Non-Intrusive Load Decomposition. *Electronics* **2023**, *12*, 2842. [[CrossRef](#)]
38. Wang, Y.; Liu, J.; Jia, S. Research on wind turbine status monitoring methods based on improved PSO-LSTM algorithm. *Shandong Electr. Power* **2024**, *51*, 30–37. (In Chinese)
39. Greff, K.; Srivastava, R.; Koutnik, J.; Steunebrink, B.; Schmidhuber, J. Transactions on neural networks and learning systems 1 LSTM: A search space odyssey. *IEEE Trans. Neural. Netw. Learn. Syst.* **2017**, *28*, 2222–2232. [[CrossRef](#)] [[PubMed](#)]
40. Chen, X.; Wu, J.; Cai, J. Short-term load prediction based on BiLSTM optimized by hunter-prey optimization algorithm. *Shandong Electr. Power* **2024**, *51*, 64–71. (In Chinese)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.