

Integrated Forecasting Method for Wind Energy Management: A Case Study in China

Authors:

Yao Dong, Lifang Zhang, Zhenkun Liu, Jianzhou Wang

Date Submitted: 2020-02-03

Keywords: forecasting accuracy, wind energy forecasting, multi-objective optimization algorithm, data preprocessing technology, combined model

Abstract:

Wind speed forecasting helps to increase the efficacy of wind farms and prompts the comparative superiority of wind energy in the global electricity system. Many wind speed forecasting theories have been widely applied to forecast wind speed, which is nonlinear, and unstable. Current forecasting strategies can be applied to various wind speed time series. However, some models neglect the prerequisite of data preprocessing and the objective of simultaneously optimizing accuracy and stability, which results in poor forecast. In this research, we developed a combined wind speed forecasting strategy that includes several components: data pretreatment, optimization, forecasting, and assessment. The developed system remedies some deficiencies in traditional single models and markedly enhances wind speed forecasting performance. To evaluate the performance of this combined strategy, 10-min wind speed sequences gathered from large wind farms in Shandong province in China were adopted as a case study. The simulation results show that the forecasting ability of our proposed combined strategy surpasses the other selected comparable models to some extent. Thus, the model can provide reliable support for wind power generation scheduling.

Record Type: Published Article

Submitted To: LAPSE (Living Archive for Process Systems Engineering)

Citation (overall record, always the latest version):

LAPSE:2020.0160

Citation (this specific file, latest version):

LAPSE:2020.0160-1

Citation (this specific file, this version):

LAPSE:2020.0160-1v1

DOI of Published Version: <https://doi.org/10.3390/pr8010035>

License: Creative Commons Attribution 4.0 International (CC BY 4.0)

Article

Integrated Forecasting Method for Wind Energy Management: A Case Study in China

Yao Dong ^{1,2}, Lifang Zhang ^{3,*} , Zhenkun Liu ³ and Jianzhou Wang ³

¹ School of Statistics, Jiangxi University of Finance and Economics, Nanchang 330013, China; dongyao20051987@126.com

² Applied Statistics Research Center, Jiangxi University of Finance and Economics, Nanchang 330013, China

³ School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China; zhenkunliudufe@163.com (Z.L.); wangjz@dufe.edu.cn (J.W.)

* Correspondence: zhanglf_1106@163.com; Tel.: +86-1786-654-0740

Received: 23 October 2019; Accepted: 16 December 2019; Published: 30 December 2019



Abstract: Wind speed forecasting helps to increase the efficacy of wind farms and prompts the comparative superiority of wind energy in the global electricity system. Many wind speed forecasting theories have been widely applied to forecast wind speed, which is nonlinear, and unstable. Current forecasting strategies can be applied to various wind speed time series. However, some models neglect the prerequisite of data preprocessing and the objective of simultaneously optimizing accuracy and stability, which results in poor forecast. In this research, we developed a combined wind speed forecasting strategy that includes several components: data pretreatment, optimization, forecasting, and assessment. The developed system remedies some deficiencies in traditional single models and markedly enhances wind speed forecasting performance. To evaluate the performance of this combined strategy, 10-min wind speed sequences gathered from large wind farms in Shandong province in China were adopted as a case study. The simulation results show that the forecasting ability of our proposed combined strategy surpasses the other selected comparable models to some extent. Thus, the model can provide reliable support for wind power generation scheduling.

Keywords: combined model; data preprocessing technology; multi-objective optimization algorithm; forecasting accuracy; wind energy forecasting

1. Introduction

Wind energy, which is characterized by both government and researchers as having adequate supply, cleanliness, and a wide distribution, is an environmentally friendly and economical energy source that could help solve the energy shortage problem [1]. Enhance attention in all over the world has been paid recently to the utilization of wind energy, which occupies about 10% in Europe's energy consumption structure and over 15% in the United States and Spain's energy consumption [2]. Moreover, the global cumulatively installed wind capacity achieved nearly 591.55 GW at the end of 2018, among which China took up 209.53 GW with the proportion of 35.4% [3].

Wind power system operation is vulnerable to stochastic and unstable wind speeds [4], which may negatively affect energy transportation and power grid operation [5]. Thus, it is imperative to improve the accuracy and steadiness of wind speed forecasting. With refined forecasting results, the dispatching department could easily and effectively adjust the program, minimizing the negative impact of wind farms on the power grid and the use of wind power could be maximized in the global electricity market [6].

Researches have been devoted to developing and applying valid and precise wind speed forecasting technology to improve forecasting accuracy. These technologies could be categorized into four strategies

based on computational mechanisms: physical, statistical, intelligent, and combined strategies [7]. Physical strategies are always used to conduct large-scale wind speed forecasting that relies on numerical weather prediction (NWP) and atmosphere data. Some issues still occur when using these strategies, for example, with the forecasting accuracy, computing complexity, and simulation time [8]. Alessandrini et al. [9] used realistic data in Southern Italy to compare two ensemble methods for wind power forecasting. On the basis of frequently used assessment criteria, the higher horizontal resolution model produced a slightly better effect, especially for 27 to 48 h advanced forecasting. Sile et al. [10] supported the advantages of NWP strategies for meteorological forecasting, which were considered credible when we conducted our wind resource assessment. In this work, the forecasting results were verified based on data from May and November 2013. Statistical strategies usually adopt mathematical statistics to explore the relationship between every variable to identify the potential relationships between original data and forecasting data [11]. Classical statistical models, such as the autoregressive integrated moving average (ARIMA) model and autoregressive fractional integrated moving average (ARFIMA) model, are broadly used in wind speed forecasting. For example, Shukur and Lee [12] reported that ARIMA model cannot capture the nonlinearity of wind speed, thus Kalman filtering (KF) technology and an artificial neural network (ANN) were applied to improve forecasting capability. The simulation results revealed that the hybrid KF-ANN model can grasp the nonlinearity of wind speed. Yuan et al. [13] combined the ARFIMA method with the least square support vector machine method (LSSVM) to improve wind speed forecasting performance. By incorporating the two strategies, satisfactory forecasting results were obtained and measured using three performance indicators, which support the forecasting precision of the developed hybrid method. Understanding linear components is the primary objective of statistical strategies, which results in overlooking nonlinear components. Thus, artificial intelligence strategies, including ANNs, back propagation neural network (BPNN), and fuzzy logic (FL) methodologies, have been rapidly developed to decrease forecasting errors and improve forecasting performance [14–16].

Artificial intelligence strategies are different from physical and statistical strategies in terms of fault tolerance and robustness, that is, ANNs can accurately fit nonlinear sequences and provide adaptive control as well as solution forecasting with uncertainty and have been applied in various fields [17–19]. Hybrid models incorporating ANNs and intelligent algorithms are becoming increasingly popular because hybrid models more improve forecasting accuracy and are more reliable relative to single intelligence strategies [20]. For example, Pourmousavi et al. [21] combined ANN and Markov chain (MC) to create a new hybrid system for wind speed forecasting, in which ANN was used to grasp short-term patterns and MC was applied for the long-term patterns. Yang et al. [22] proposed a hybrid forecasting system integrating forecasting with both a certainty portion and an analysis with an uncertainty portion, which also contained data pretreatment technology and an optimization method to increase forecasting accuracy. The experimental results showed that the hybrid system was more accurate compared with other models and could be applied in other fields. Xiao et al. [23] synthesized the singular spectrum analysis (SSA), a novel modified cuckoo search (CS) algorithm, and the modified wavelet neural network (WNN) to a novel hybrid system. The developed system could conduct short-term forecasting for load, electricity price, and wind speed. The simulation results indicated that this system was superior to single forecasting models due to providing more accurate forecasting. Although intelligent strategies are powerful tools for depicting the nonlinearity of original wind speed data, some challenges remain when considering the volatility and instability of untreated time series [24]. Thus, data preprocessing strategies are broadly used to eliminate strong noise and extract the primary wind speed characteristics [25]. For instance, Wu et al. [26] combined the complete ensemble empirical mode decomposition (CEEMD), multi-objective grey wolf optimization (MOGWO) algorithm, and extreme learning machine (ELM) to establish a CEEMD-MOGWO-ANN system, which could grasp the fluctuation in original sequences with high levels of noise. Liu et al. [27] developed a new hybrid method called EMD-ANN relying on empirical model decomposition (EMD) to reduce prediction error and enhance forecasting performance. Based on multiple simulations,

the proposed system was found to be satisfactory and robust in managing jumping samples with erratic wind speed data.

We summarize the following several characteristics of previous wind speed forecasting strategies:

- (1) Forecasting short-term wind speed presents a challenge for physical models. Resources and time would be wasted to collect large amounts of physical information.
- (2) The statistical strategies assume that the wind speed series are linear; in reality, the wind speed time series are nonlinear along with certain trends, which may result in inferior forecasting compared with expectations.
- (3) The intelligent strategy does not assume linear raw wind speed sequences and can effectively capture nonlinear elements; however, some deficiencies remain in this strategy, like easily falling into the local optimum and over-fitting.
- (4) Various data pretreatment strategies have been used to improve forecasting ability by eliminating the noise in raw time series. However, previous de-noising methods have defects like mode mixing in empirical mode decomposition (EMD) and residue noise in ensemble empirical mode decomposition (EEMD).

Given previous research, combined wind speed forecasting strategies, driven by the combination of forecasting knowledge introduced by Bates and Granger in 1969 [28], have received attention and achieved satisfactory forecasting performance [29,30]. This strategy operates by determining the optimal weight in the situation where the minimum sum of squared errors of forecasting training sets are available [31,32]. In their review, Niu et al. [33] found that existing forecasting models do not provide sufficiently accurate forecasting; thus, they proposed a new method that combines data pretreatment technology and a multi-objective optimization algorithm with several well-performing ANNs and a linear model. The simulation results revealed that the developed model considerably improved forecasting capacity compared with previous forecasting models. Li et al. [34] introduced the idea of establishing a variable weighting combination model that includes three various hybrid models to improve wind speed forecasting ability. The experimental results showed that this innovatory strategy is better than benchmark strategies in wind speed forecasting.

In this paper, we propose a new combined strategy that incorporates three portions: data pretreatment, optimization, and forecasting portion. Singular spectrum analysis (SSA) was adopted to eliminate fluctuating noise from raw data sets. Then, a statistical model and three typical ANNs were applied to forecast wind speed, from which the combined strategy is structured with weight coefficients optimized by an efficient multi-objective dragonfly algorithm.

Our contributions with this research are as follows:

- (1) Data pretreatment technology is included in our method to reduce the volatility and randomness of historical wind speed sequences and improve forecasting accuracy. Original wind speed sequences are decomposed into some intrinsic mode functions (IMFs), from which the high-frequency IMFs are filtered, and the residuals are recombined to forecast wind speed. With this method, the characteristics of wind speed can be better extracted, so forecasting performance can be significantly improved.
- (2) Statistical models help grasp the linear characteristics of raw time series whereas ANNs can be used for nonlinear characteristics. To comprehensively control the linear and nonlinear features of original data, one widely used statistical model and three efficient ANNs were added to our combined model to improve forecasting accuracy.
- (3) The multi-objective dragonfly algorithm (MODA), as an effective weighting technology, was developed to determining the optimal weight coefficients of individual forecasting models. In the majority of situations, the multi-objective optimization algorithm, with an archive to reserve and search the optimal approximate value of the Pareto optimal solutions, can improve forecasting precision and forecasting steadiness; this algorithm can satisfactorily address complicated optimization problems.

- (4) A systematic assessment system was established to evaluate the forecasting ability of our developed combined model. Three experiments, four evaluation criteria, and six discussions are introduced in our study to compare and analyze the forecasting capacity of our developed combined strategy in every experiment.
- (5) The developed combined system provides a reference for the scheduling and management of smart grids. Referring to real wind speed data and comparative forecasting results, the combined forecasting model is confirmed to be effective and applicable to other forecasting fields.

The remainder of our study is organized as follows: The methodologies of our combined model are introduced in Section 2. Sections 3 and 4 describe the experiment and three experimental results, respectively. Further discussions, including the superiority and the stability of the forecasting method, are provided in Section 5. Section 6 concludes our study and provides an analysis of improvements to our research and recommendations for future study.

2. Methods

Our methods are introduced in this section, including the data pre-processing technology (singular spectrum analysis) and the multi-objective dragonfly algorithm; then, the workflow of our combined strategy is presented.

2.1. Singular Spectrum Analysis (SSA)

SSA, as a time series analysis tool used to decompose an original time sequence into interpretable components, has been applied in many fields, including biology, physics, climatology, and economics [35–38]. The decomposition process can be divided into four steps:

2.1.1. Embedding

Convert original time series $\mathbf{X} = (x_1, x_2, \dots, x_N)$ into a series $\mathbf{Z} = (z_1, z_2, \dots, z_K)$, which can be expressed as:

$$\mathbf{X} = (x_1, x_2, \dots, x_N) \rightarrow \mathbf{Z} = (z_1, z_2, \dots, z_K) \quad (1)$$

where $z_i = (x_i, x_{i+1}, \dots, x_{i+L-1})^T \in R^L$, $K = N - L + 1$, $L \in [2, N]$. The result of this mapping is provided as a trajectory matrix with the following mathematical expression:

$$\mathbf{Z} = [\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_K] = (\mathbf{z}_{ij})_{i,j=1}^{L,K} = \begin{pmatrix} x_1 & x_2 & \dots & x_K \\ x_2 & x_3 & \dots & x_{K+1} \\ \dots & \dots & \dots & \dots \\ x_L & x_{L+1} & \dots & x_N \end{pmatrix} \quad (2)$$

2.1.2. Singular Values Decomposition (SVD)

Given a covariance matrix $\mathbf{S} = \mathbf{X}\mathbf{X}^T$, singular values decomposition is used to generate L eigenvalues and eigenvectors, which are denoted by $\lambda_1, \lambda_2, \dots, \lambda_L$ and $\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_L$. Suppose $t = \max(i, \text{such that } \lambda_i > 0)$ and $\mathbf{V}_i = \mathbf{X}^T \mathbf{U}_i \sqrt{\lambda_i}$ ($i = 1, 2, \dots, t$), then the SVD of the trajectory matrix is:

$$\mathbf{Z} = \mathbf{E}_1 + \mathbf{E}_2 + \dots + \mathbf{E}_t \quad (3)$$

where $\mathbf{E}_i = \sqrt{\lambda_i} \mathbf{U}_i \mathbf{V}_i$, the rank of Z_i , is 1. Therefore, $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_t$ are the principle components and $(\sqrt{\lambda_i}, \mathbf{U}_i, \mathbf{V}_i)$ is the characteristic loop of the SVD of \mathbf{Z} .

2.1.3. Grouping

In this step, the interval ($i = 1, 2, \dots, t$) is decomposed into several subsets S_1, S_2, \dots, S_m without any connection between them. Suppose $\mathbf{S} = (s_1, s_2, \dots, s_p)$, then the consequence matrix \mathbf{Z}_S is defined as $\mathbf{Z}_S = \mathbf{Z}_{S_1} + \mathbf{Z}_{S_2} + \dots + \mathbf{Z}_{S_p}$ and the trajectory matrix can be decomposed as $\mathbf{Z} = \mathbf{Z}_{S_1} + \mathbf{Z}_{S_2} + \dots + \mathbf{Z}_{S_m}$.

2.1.4. Diagonal Averaging

In the diagonal averaging stage, the grouping results are switched into a time series of length N . Denote \mathbf{Z} as a $L \times K$ matrix, where $L^* = \min(L, K)$ and $K^* = \max(L, K)$. Once $L < K$, then $z_{ij}^* = z_{ij}$, or else, $z_{ij}^* = z_{ji}$. Subsequently, the \mathbf{Z} matrix can be transformed into a sequence $(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ with the following formula:

$$\mathbf{r}_k = \begin{cases} \frac{1}{k+1} \sum_{q=1}^{k+1} z_{q,k-q+1}^*, & 1 \leq k \leq L^* \\ \frac{1}{L^*} \sum_{q=1}^{L^*} z_{q,k-q+1}^* & L^* \leq k \leq K^* \\ \frac{1}{N-K+1} \sum_{q=1}^{N-K+1} z_{q,k-q+1}^* & K^* \leq k \leq N \end{cases} . \quad (4)$$

2.2. Multi-Objective Dragonfly Algorithm (MODA)

The MODA, proposed by Mirjalili in 2015, was developed by imitating the static and dynamic behavior of dragonflies when hunting and migrating [39]. During these two behaviors, five principles must be obeyed, including the separation of dragonflies S_i , the alignment A_i , the cohesion C_i , the behaviors of attraction to food source F_i , and the distraction from enemy source E_i . To update the dragonflies' position, the mathematical form is:

$$\Delta \mathbf{X}_{t+1} = (sS_i + aA_i + cC_i + fF_i + eE_i) + w\Delta \mathbf{X}_t \quad (5)$$

$$\mathbf{X}_{t+1} = \mathbf{X}_t + \Delta \mathbf{X}_{t+1} \quad (6)$$

where s, a, c, f , and e are weight coefficients that are obtained randomly; w is the inertia weight; \mathbf{X}_t and \mathbf{X}_{t+1} denote the current and the next population individual positions, respectively; and the $\Delta \mathbf{X}_{t+1}$ is the next population position update velocity.

To conduct MODA optimization based on DA, an archive was used to save the non-inferior solutions produced during optimization; the adaptive grid method and a roulette-wheel mechanism were used to choose food source and enemy locations from the archive set with the probability for each segment of $P_i = d/N_i$, where d denotes a constant within $(0, +\infty)$, and N_i is the quantity of Pareto optimal solutions in the i th segment. However, the storage space of the archive is restrictive; thus, the unsatisfactory results were eliminated from the archive with the probability of $P_i = N_i/d$.

Based on these mechanisms, the improved algorithm addresses multi-objective problems with and without constraints concurrently, and is adaptable, economical, and has good optimization ability [40]. For optimization objectives, accuracy $obj_1(\mathbf{y})$ and stability $obj_2(\mathbf{y})$ are considered simultaneously in our study, as indicated by:

$$\mathbf{min} \begin{cases} obj_1(\mathbf{y}) = MAPE = \frac{1}{N} \sum_{i=1}^N \left(\left| \frac{\hat{\mathbf{e}}_i - \mathbf{e}_i}{\mathbf{e}_i} \right| \right) \times 100\% \\ obj_2(\mathbf{y}) = \text{std}(\hat{\mathbf{e}}_i - \mathbf{e}_i), i = 1, 2, \dots, N \end{cases} , \quad (7)$$

where \mathbf{e}_i and $\hat{\mathbf{e}}_i$ denote the real value and the forecasting value, respectively. Algorithm 1 provides the MODA pseudocode.

Algorithm 1: Multi-Objective Dragonfly Algorithm (MODA)**Objective functions:**

$$\min \begin{cases} obf_1(\mathbf{y}) = MAPE = \frac{1}{N} \sum_{i=1}^N \left(\left| \frac{\hat{e}_i - e_i}{e_i} \right| \right) \times 100\% \\ obf_2(\mathbf{y}) = \text{std}(\hat{e}_i - e_i), i = 1, 2, \dots, N \end{cases}$$

Parameters:

$Iter_{Max}$, the maximum number of iterations
 \mathbf{x}_i , the position of the i th dragonfly
 $\Delta \mathbf{x}_i$, the step vectors of the i th dragonfly
 t , the current iteration number

```

1: /* Initialize the population of dragonflies  $\mathbf{x}_i(i = 1, 2, \dots, n)$ . */
2: /* Initialize step vectors  $\Delta \mathbf{x}_i(i = 1, 2, \dots, n)$ .*/
3: /* Confirm the maximum value of segments. */
4: /* Confirm the archive size. */
5: WHILE ( $t < Iter_{Max}$ ) DO
6: /* Compute the objective values of every dragonflies. */
7: /* Seek out solutions that are not in dominant position. */
8: /* Update the archive in accordance with the achieved non-dominant results. */
9: IF the archive reaches the maximum number
10: /* Run the archive maintenance system to remove one member from the existing archive. */
11: /* Save the novel solution to the archive. */
12: END IF
13: IF any of the novel augmented solutions is situated outside the segments
14: /* Update and relocation the segments to incorporate the novel solution(s) */
15: END IF
16: /* Discover the best solution as a food source. Discover the worst solution as an enemy. */
17: /* Update step vectors using Equation (5). */
18:  $\Delta \mathbf{x}_{t+1} = (sS_i + aA_i + cC_i + fF_i + eE_i) + w\Delta \mathbf{x}_t$ 
19: /* Update position vectors through Equation (6). */
20:  $\mathbf{x}_{t+1} = \mathbf{x}_t + \Delta \mathbf{x}_{t+1}$ 
21: /* Check and adjust the new positions according to the boundaries of the variables. */
22:  $t = t + 1$ 
23: END WHILE
24: RETURN archive
25: Obtain  $X^* = \text{Select Leader}(\text{archive})$ , and input  $X^*$ 

```

2.3. The Workflow of Developed Combined Model

According to the combination theories proposed by Bates and Granger in the late 1960s, two or more models exceed the forecasting capacity of one model [28]. Thus, in our study, we developed a combined strategy to offset the limitations of single forecasting models and incorporate both the linear and nonlinear features of wind speed. The workflow of our study is provided in Figure 1 and the further explanations are provided below.

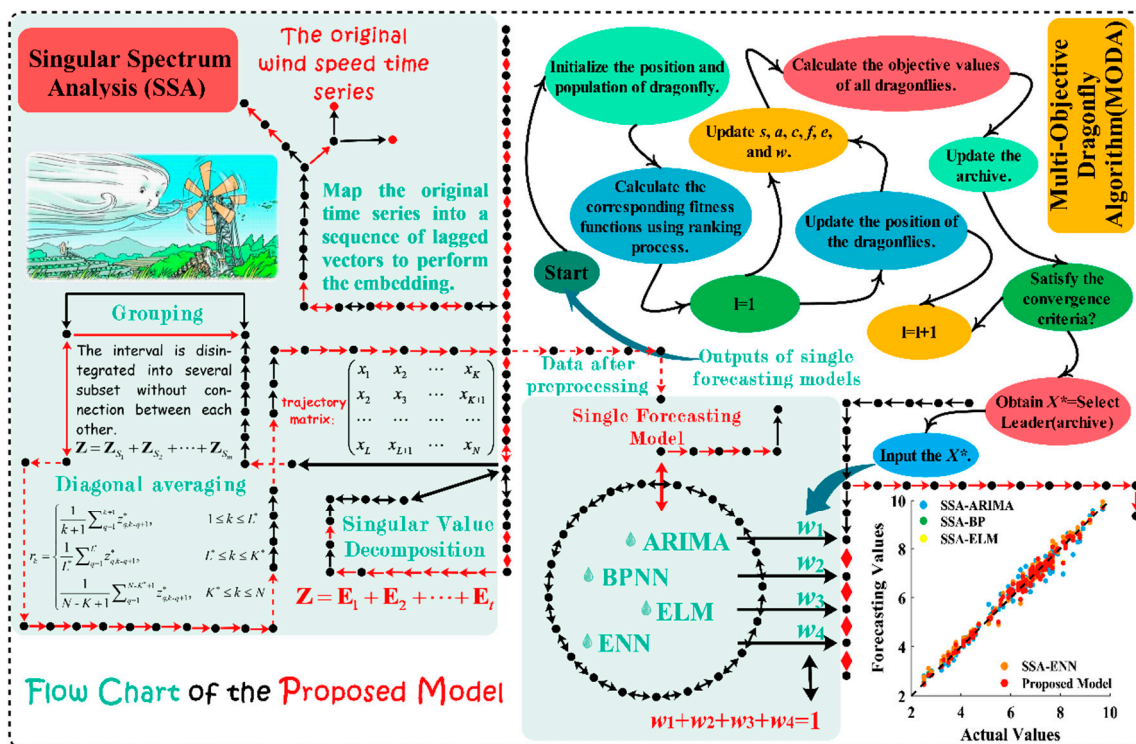


Figure 1. Flow chart of the developed model.

2.3.1. Data Preprocessing

As noisy data result in poor accuracy, the SSA technique was used to remove fluctuating components and retain the valid information in original wind speed sequences. The original data were decomposed into some IMFs, from which the volatile components were removed and the remainder were reconstructed for effective wind speed forecasting.

2.3.2. Hybrid Models Forecasting

To consider the linear and nonlinear components of short-term wind speed sequences, a classical statistical method (ARIMA) and three effective neural networks, BPNN, ELM, and Elman neural network (ENN), were used to form the basic forecasting model of our system. After integrating these models with the data preprocessing method, hybrid models SSA-ARIMA, SSA-BPNN, SSA-ELM, and SSA-ENN were established to forecast short-term wind speed.

2.3.3. Proposed Combined Forecasting Strategy

In this stage, a weight selection method relying on the MODA was used to find the optimal weight of hybrid strategies, which is crucial for improving forecasting precision and steadiness simultaneously. The data for the last three days in the training set were used to determine the weight coefficient to integrate the hybrid strategies. During the operation of the algorithm, the process is terminated when the iterations number reaches the maximum or the fitness function reaches the minimum. The forecasting results were obtained by combining these hybrid models using the optimized weight coefficients.

2.3.4. Rolling Forecasting

Rolling forecasting mechanism is adopted in our paper, that is, according to the results of trial and error, the input datasets are $\{x(t - 5), x(t - 4), x(t - 3), x(t - 2), x(t - 1)\}$ and the output datasets is $\{x(t)\}$, x represents wind speed time series, and the input and output sets change with the change of t .

To assess the forecasting ability of the developed strategy, both one-step and multi-step forecasting are used in our paper. Similarly, three-step ahead forecasting can be determined as follows: the input datasets are $\{x(t - 5), x(t - 4), x(t - 3), x(t - 2), x(t - 1)\}$ and the output datasets are $\{x(t + 2)\}$ [41]. A more detailed introduction of rolling forecasting mechanism is also presented in Figure 2.

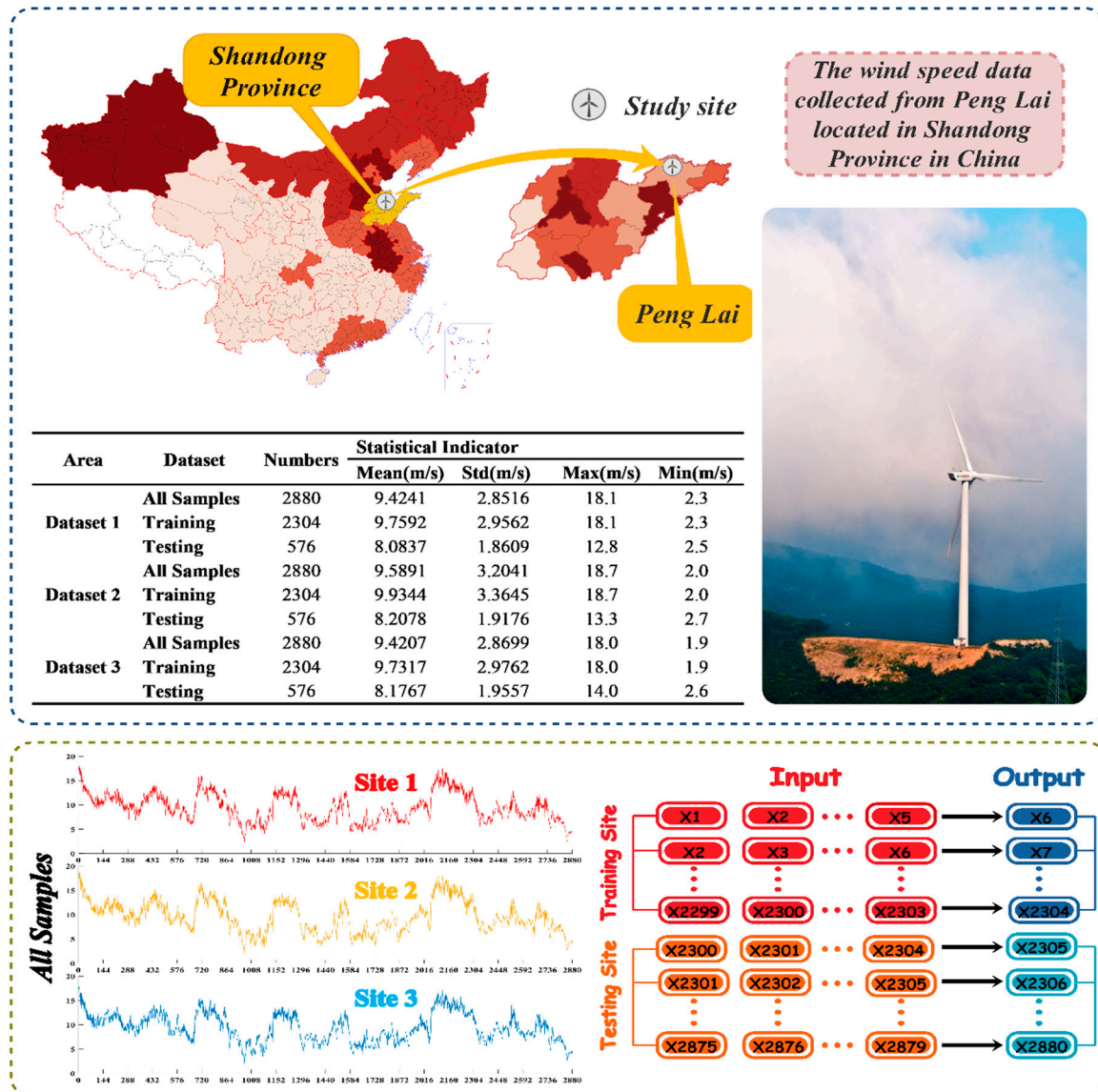


Figure 2. Observation sites of wind farms and data information.

3. Experimentation Setup

In this section, the setup of our three experiments are introduced, including data selection, performance evaluation criterion, and operating environment.

3.1. Datasets Selection

Raw wind speed data were chosen from four observation points on the Shandong Peninsula in China. Due to its long coastline, Shandong has adequate wind energy resources. All observation sites were located in mountainous and hilly regions with the altitude ranging from 100 m to 240 m above sea level. Generally, the rated power of the wind power generators is 1500 KW, and the height of measurement is 70 m. The time interval in our datasets is 10 min, for a total of 144 records per day.

We chose three datasets as the study cases, with each dataset containing 2880 records with a time span of 1–20 January 2011. The ratio of training set to testing set was 4:1, which included 2304 training samples and 576 testing samples in each group. A rolling forecasting mechanism was applied to one-step and multi-step forecasting to ultimately obtain the forecast of the sixth period based on the first five real time series. The data structure of our proposed model, the research sites, and several statistical indicators of raw wind speed series are provided in Figure 2.

3.2. Performance Evaluation Criteria

Many error indicators have been used in past research; however, no specific standard exists for model evaluation [42]. Thus, multiple evaluation criteria are usually used to compare the developed model and other models with regard to forecasting capacity [43]. In our study, four widely-used error criteria—mean absolute error (MAE), root mean square error (RMSE), mean absolute percent error (MAPE), and sum of squared errors (SSE)—were used as assessment indicators of forecasting performance, whose definitions and equations are listed in Table 1. It must be noted that the measure unit of MAPE in our paper is %.

Table 1. Performance of assessment criteria.

Metrics	Definition	Equation
MAE	The average absolute error obtained from n times forecasting consequences	$\text{MAE} = \frac{1}{n} \sum_{i=1}^n \hat{e}_i - e_i $
MAPE	The mean absolute percentage error obtained from n times forecasting consequences	$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left \frac{\hat{e}_i - e_i}{e_i} \right \times 100\%$
RMSE	The root mean square error obtained from n times forecasting consequences	$\text{RMSE} = \sqrt{\frac{1}{n} \times \sum_{i=1}^n (\hat{e}_i - e_i)^2}$
SSE	The sum square error obtained from n times forecasting consequences	$\text{SSE} = \sum_{i=1}^n (e_i - \hat{e}_i)^2$

Note: e_i and \hat{e}_i denote the real value and the forecasting value, respectively. The measure unit of MAPE in our paper is %.

3.3. Operating Environment

Our experiments were conducted using the Windows 7 professional operating system (Microsoft, Redmond, WA, USA). We used Matlab2016a (MathWorks, Natick, MA, USA) to operate the developed model. The detailed hardware information is: Intel (R) Core i5-4590 3.30 GHz CPU, and 8 GB RAM (Intel Corporation, Santa Clara, CA, USA).

4. Three Experiments and Relative Analysis

In this section, numerical experiments and the corresponding forecasting results are compared and analyzed in detail to provide evidence for the superior forecasting capacity of our developed combined strategy. The experiment setup and results are presented below.

4.1. Experimental Setup

Using actual wind speed time series, we performed three experiments to compare the forecasting ability of the proposed model and other comparable models. Experiment 1 compared the accuracy of our combined model with that of four hybrid models. Experiment 2 compared the differences in the forecasting accuracy of our proposed combined model with other combined models using various data preprocessing technologies. Experiment 3 compared our combined model with four benchmark models to investigate the differences in forecasting capacities. The one-step, two-step, and three-step ahead forecasting capability of the different models were analyzed using the four calculated error criteria. The smaller the value of the error criteria, the better the forecasting performance.

Appropriate parameters are crucial for enhancing wind speed forecasting accuracy and may affect the quality of forecasting results to a certain extent. In our study, the parameters contained in the different models and the different experiments were selected by referring the literature and the actual status of this article to provide useful reference for practical research in the future.

To compare the forecasts of the developed strategy with those of four hybrid strategies without MODA optimization, we conducted Experiment 1. The SSA parameter setting in hybrid strategies is identical to that of the proposed combined strategy, with the specific parameters listed in Table 2. In the rolling forecasting mechanism, the rolling number was 5 and the training-to-testing ratio in all forecasting strategies was set to 3:1.

Table 2. Parameters in experiment 1.

Algorithm	Parameters	Value
Multi-Objective Dragonfly Algorithm (MODA)	Maximum Iterations	200
	Maximum Archive	500
	Solution Numbers nearby the Solution	40
Singular Spectrum Analysis (SSA)	Window Length	48
	Principal Component Decomposition Number	20

Experiment 2 was performed to confirm that the forecasting performance of the combined strategy with SSA is superior to those of the other methods using other data preprocessing tools, such as EMD, CEEMD, and wavelet domain de-noising (WDD). In CEEMD, the ratio of the standard deviation of the added noise to the sequences was 0.05 and the realization values and maximum sifting iterations were set to 50 and 500, respectively. The decomposition layer number in WDD was set to 9. For the SSA used in our combined model, the window length and the principal component decomposition number were set to 48 and 20, respectively.

Experiment 3 compared the developed model and four widely-used benchmark models—ARIMA, BPNN, ELM, and ENN—in terms of forecasting performance. Without loss of generality, statistical models and artificial intelligence were considered to show that the forecasting capability far surpasses that of all individual models.

4.2. Experiment 1: Comparison with SSA-Based Hybrid Strategies

The parameters involved in our combined model are provided in Table 2. Based on four error indicators from one- to multi-step forecasting, the proposed combined strategy was confirmed to be the most accurate among the examined models. The comparison results are shown in Table 3, where bold elements indicate the best forecasting performance.

For Site 1, the developed strategy produced the best forecasting results in one- and multi-step forecasting with MAPE values of 2.209%, 2.935%, and 3.221% from one-step to three-step, respectively. The other three evaluation criteria obtained for the developed model are also the lowest among all compared models. The differences in the models at Site 1 are provided in Figure 3.

For Site 2, we forecasted the trend in wind speed using the developed combined model as the MAE, MAPE, RMSE, and SSE results for our combined strategy are superior to those of the other strategies. For instance, the MAPE value of our proposed model is 2.352% and the corresponding MAE, RMSE, and SSE values are 0.145, 0.193, and 5.371, respectively, for one-step forecasting.

For Site 3, the SSA-MODA-based combined model performed the best with a 2.804% MAPE in one-step, 3.193% MAPE in two-step, and 3.588% MAPE in three-step forecasting. Compared with the other hybrid models that improve prediction accuracy to a certain extent, the SSA-MODA-based model is superior.

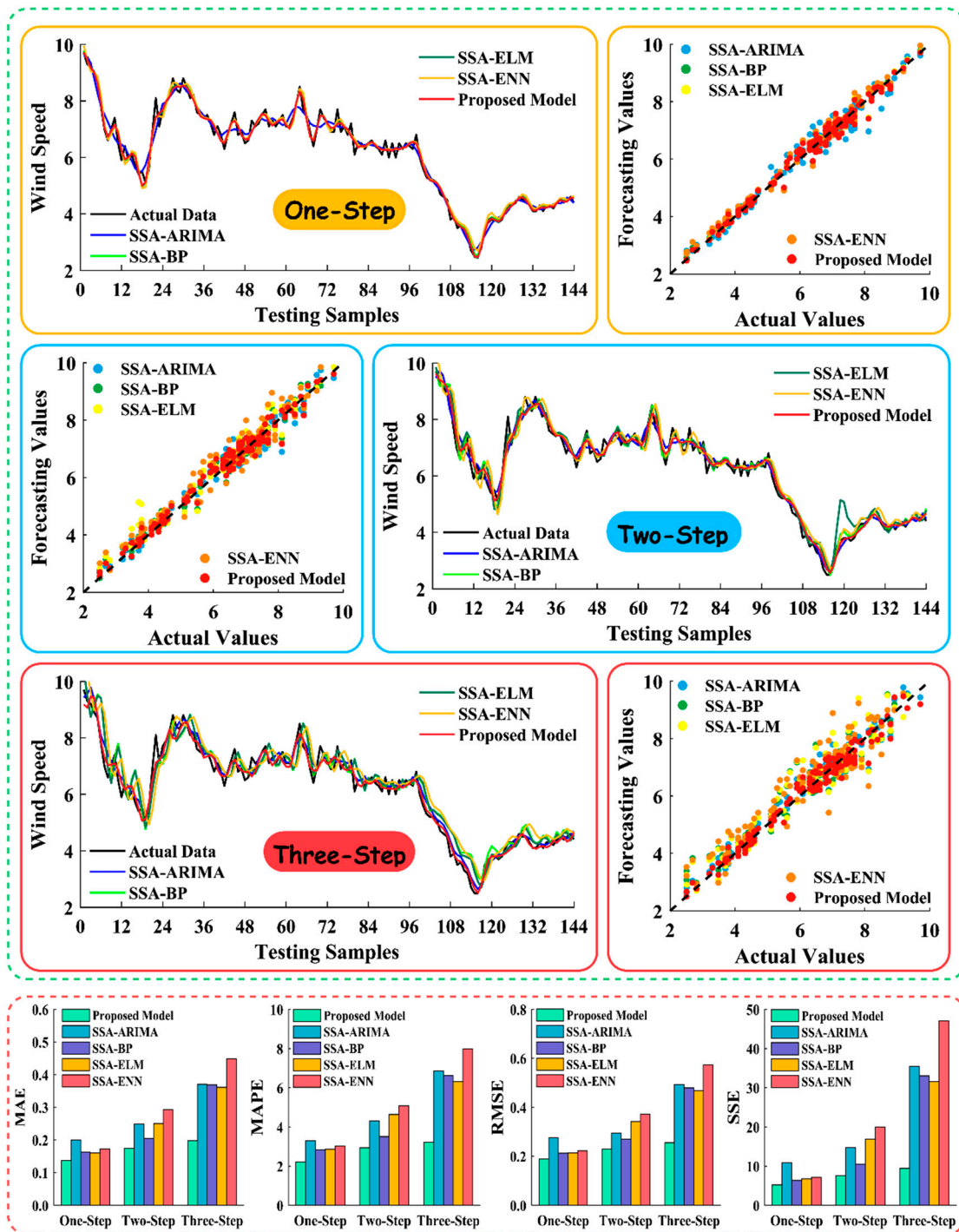


Figure 3. Comparison of experiment 1 results in site 1.

Table 3. Evaluation metrics comparison of proposed model with relative hybrid models.

Dataset	Model	One-Step				Two-Step				Three-Step			
		MAE	MAPE	RMSE	SSE	MAE	MAPE	RMSE	SSE	MAE	MAPE	RMSE	SSE
Site 1	SSA-ARIMA	0.200	3.296	0.276	10.911	0.249	4.313	0.295	14.746	0.371	6.848	0.493	35.385
	SSA-BP	0.163	2.836	0.213	6.343	0.204	3.506	0.270	10.530	0.369	6.615	0.479	33.014
	SSA-ELM	0.160	2.857	0.213	6.713	0.250	4.627	0.342	16.851	0.361	6.310	0.468	31.579
	SSA-ENN	0.173	3.018	0.222	7.119	0.293	5.083	0.372	19.916	0.448	7.972	0.572	47.124
	Proposed Model	0.138	2.209	0.189	5.165	0.174	2.935	0.229	7.563	0.199	3.221	0.256	9.403
Site 2	SSA-ARIMA	0.221	3.535	0.291	13.090	0.250	4.339	0.309	15.328	0.393	7.019	0.540	45.811
	SSA-BP	0.168	2.884	0.226	8.527	0.229	3.790	0.302	13.123	0.374	6.354	0.502	36.340
	SSA-ELM	0.178	2.987	0.230	8.543	0.235	3.925	0.308	13.688	0.378	6.455	0.504	36.595
	SSA-ENN	0.177	2.955	0.229	7.539	0.314	5.429	0.406	23.768	0.451	7.791	0.590	50.201
	Proposed Model	0.145	2.352	0.193	5.371	0.183	2.966	0.251	9.057	0.197	3.243	0.270	10.476
Site 3	SSA-ARIMA	0.223	3.612	0.319	16.439	0.249	4.464	0.336	14.977	0.447	8.127	0.598	45.523
	SSA-BP	0.220	3.491	0.289	14.597	0.238	3.969	0.309	13.717	0.383	6.531	0.497	35.503
	SSA-ELM	0.220	3.471	0.288	13.613	0.248	4.227	0.320	14.758	0.390	6.785	0.511	37.638
	SSA-ENN	0.224	3.648	0.321	21.389	0.309	5.357	0.387	21.509	0.474	8.455	0.609	53.330
	Proposed Model	0.175	2.804	0.228	7.500	0.197	3.193	0.261	9.830	0.216	3.588	0.285	11.674

Note: The best results are indicated in bold. The measure unit of MAPE in our paper is %.

Remark 1. We have two main findings here: the prediction of our proposed model is the most accurate, and as the number of forecasting steps increases, the variation in the evaluation criteria in the proposed model from one-step to three-step is minimal, which means the developed strategy provides the best forecasting precision and stability.

4.3. Experiment 2: Comparison of Different Combined Strategies Based on Four Data Preprocessing Technologies

The purpose of experiment 2 was to compare the developed combined model using SSA technology with other combined models employing different data preprocessing technologies. The parameters involved in every method are listed in Tables 4 and 5, and provide the comparative results for the three sites. The combined models based on EMD, CEEMD, and WDD technology are labeled as EMD-C-Model, CEEMD-C-Model, and WDD-C-Model, respectively.

Table 4. Parameters in experiment 2.

Algorithm	Parameters	Value
MODA	Maximum Iterations	200
	Maximum Archive	500
	Solution Numbers nearby the Solution	40
EMD	There are no parameters that need to be set subjectively	
CEEMD	Ratio of the standard deviation of the added noise and that of sequences	0.05
	Realization Numbers	50
	Maximum Sifting Iterations	500
WDD	Decomposition Layer Number	9
SSA	Window Length	48
	Principal Component Decomposition Number	20

Note: MODA is Multi-Objective Dragonfly Algorithm, EMD is empirical mode decomposition, CEEMD is combined the complete ensemble empirical mode decomposition, WDD is wavelet domain de-noising, and SSA is singular spectrum analysis.

For Site 1, the proposed combined strategy produced the best forecasting results for all assessment metrics, whereas the others also produced good forecasting results with MAPE values lower than 6%.

For Site 2, the proposed combined strategy also performed better than methods based on other data pretreatment techniques, indicated by MAPE values of 2.352%, 2.966%, and 3.243% for one- to three-step forecasting, respectively. Correspondingly, EMD-C-Model has the worst forecasting accuracy

which increases 1.562%, 2.223%, and 2.864%, respectively, relative to SSA based combined model. The comparison results of our SSA-based forecasting strategy and EMD-C-Model, CEEMD-C-Model, and WDD-C-Model are shown in Figure 4.

Table 5. Comparison of the evaluation metrics for the proposed strategy with models using different data pretreatment technologies.

Dataset	Model	One-Step				Two-Step				Three-Step			
		MAE	MAPE	RMSE	SSE	MAE	MAPE	RMSE	SSE	MAE	MAPE	RMSE	SSE
Site 1	EMD-C- Model	0.211	3.545	0.265	10.072	0.283	4.868	0.354	18.067	0.355	5.938	0.445	28.484
	CEEMD-C-Model	0.181	3.056	0.227	7.428	0.239	4.062	0.305	13.398	0.266	5.061	0.337	16.329
	WDD-C-Model	0.196	3.309	0.252	9.125	0.252	4.363	0.339	16.592	0.323	5.646	0.379	20.680
	Proposed Model	0.138	2.209	0.189	5.165	0.174	2.935	0.229	7.563	0.199	3.221	0.256	9.403
Site 2	EMD-C- Model	0.234	3.914	0.295	12.551	0.290	5.189	0.365	19.182	0.369	6.107	0.450	29.116
	CEEMD-C-Model	0.194	3.305	0.245	8.655	0.257	4.332	0.339	16.537	0.321	5.375	0.381	20.891
	WDD-C-Model	0.223	3.668	0.285	11.663	0.276	4.882	0.347	17.377	0.336	5.660	0.406	23.724
	Proposed Model	0.145	2.352	0.193	5.371	0.183	2.966	0.251	9.057	0.197	3.243	0.270	10.476
Site 3	EMD-C-Model	0.251	4.011	0.314	14.229	0.272	4.709	0.344	16.991	0.333	5.781	0.395	22.467
	CEEMD-C-Model	0.199	3.290	0.254	9.320	0.251	4.079	0.321	14.820	0.293	5.040	0.357	18.310
	WDD-C-Model	0.223	3.736	0.284	11.644	0.267	4.398	0.338	16.452	0.316	5.461	0.388	21.661
	Proposed Model	0.175	2.804	0.228	7.500	0.197	3.193	0.261	9.830	0.216	3.588	0.285	11.674

Note: The best results are indicated in bold. The measure unit of MAPE in our paper is %. The EMD-C-Model, CEEMD-C-Model, and WDD-C-Model represent combined models based on EMD, CEEMD, and WDD technology, respectively.

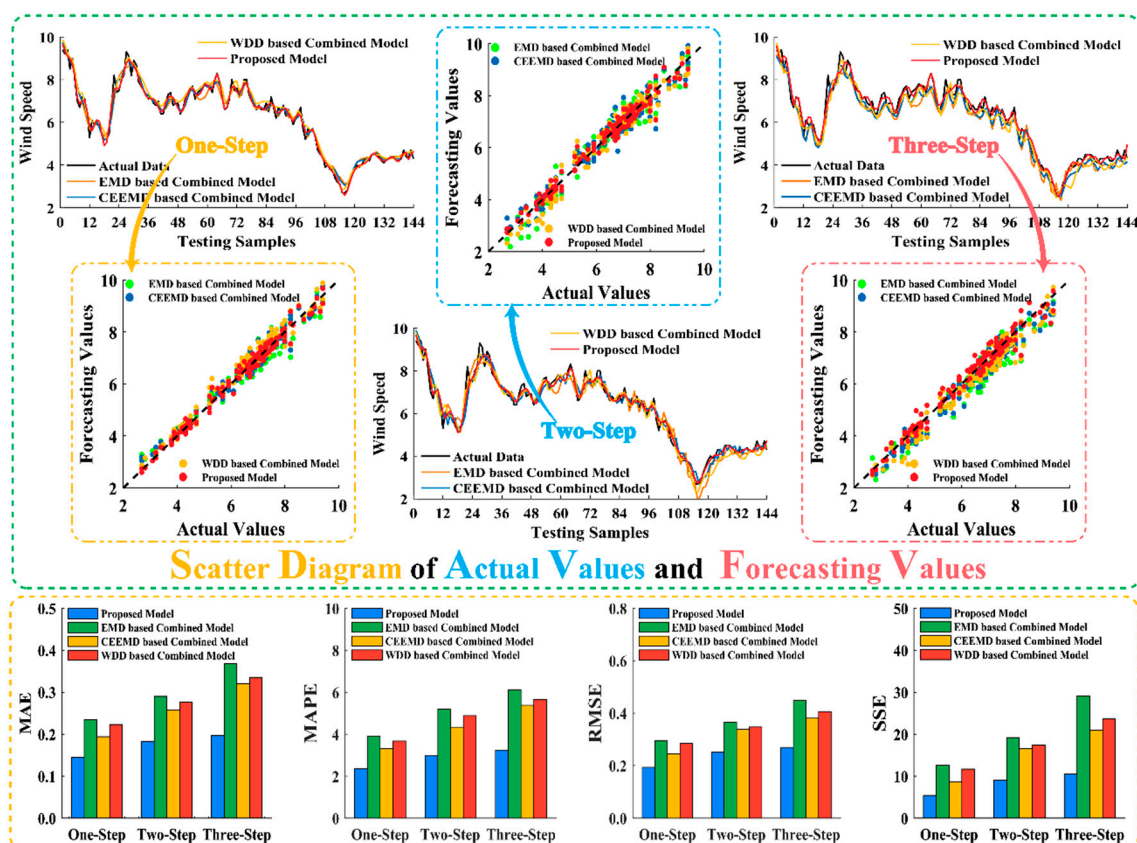


Figure 4. Comparison results of experiment 2 at site 2.

For Site 3, the SSA-based combined model produced more accurate and effective forecasting compared with the other combined strategies with MAPE value of 2.804% for one-step forecasting. In contrast, the CEEMD-C-Model, WDD-C-Model, and EMD-C-Model had MAPE values of 3.290%, 3.736%, and 4.011%, respectively, which are inferior to our developed combined model.

Remark 2. We concluded that using the SSA-based forecasting system would be preferable over forecasting systems using other data pretreatment techniques because the MAPE values of the developed strategy, with average MAPE values of 2.789%, 2.854%, and 3.195% for Sites 1 to 3, respectively, are the lowest compared to the other selected models.

4.4. Experiment 3: Comparison with Several Benchmark Strategies

Tables 6 and 7 present the parameters of the algorithms and compare the results of the developed combined model and several individual models, to demonstrate the forecasting improvement of the developed SSA-MODA combined model.

Table 6. Parameters in experiment 3.

Algorithm	Parameters	Value
Multi-Objective Dragonfly Algorithm (MODA)	Maximum Iterations	200
	Maximum Archive	500
	Solution Numbers nearby the Solution	40
Singular Spectrum Analysis (SSA)	Window Length	48
	Principal Component Decomposition Number	20

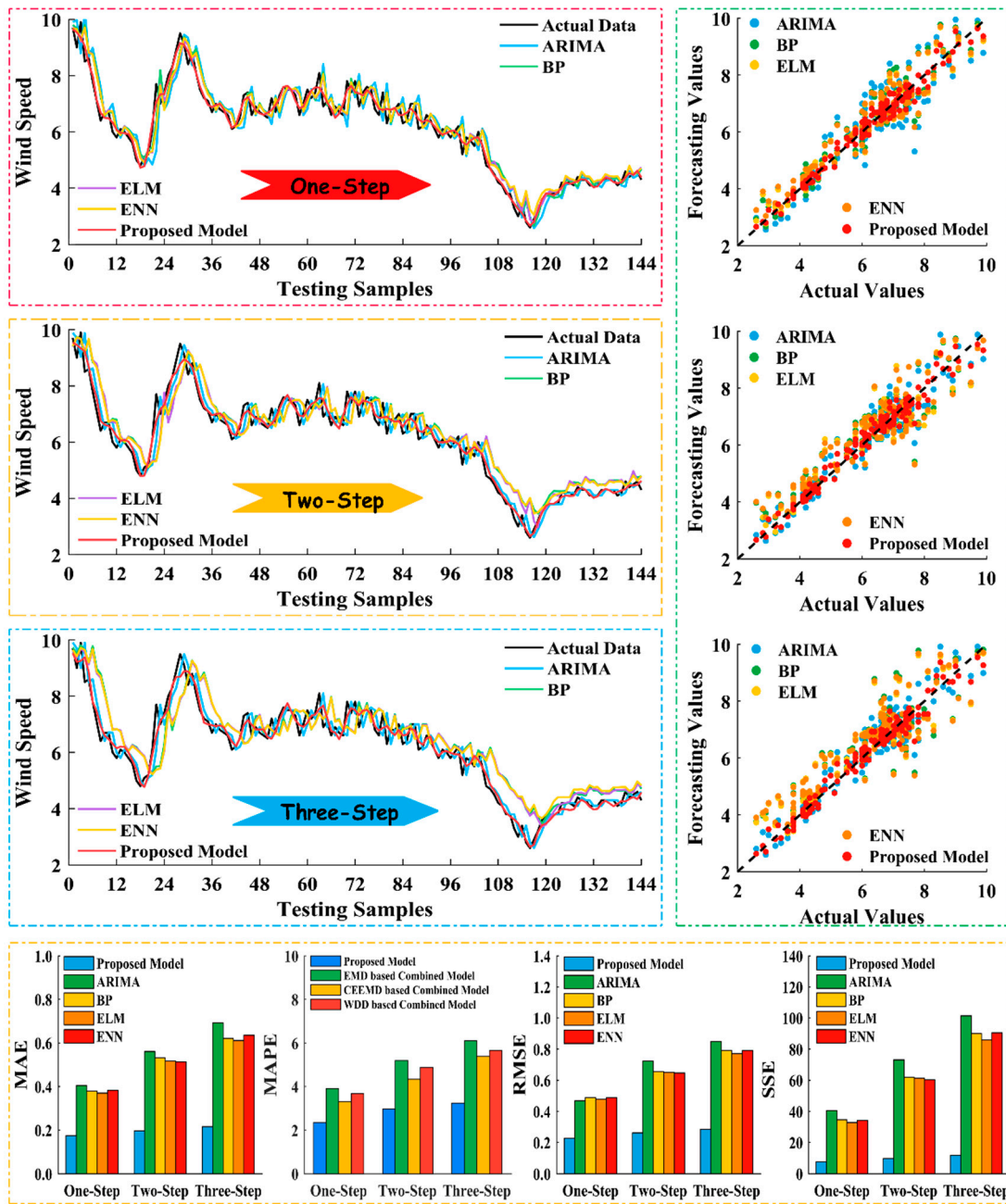
Table 7. Comparison of the evaluation metrics of the proposed model with benchmark models.

Dataset	Model	One-Step				Two-Step				Three-Step			
		MAE	MAPE	RMSE	SSE	MAE	MAPE	RMSE	SSE	MAE	MAPE	RMSE	SSE
Site 1	ARIMA	0.415	7.179	0.474	44.272	0.492	8.398	0.616	55.279	0.630	10.982	0.776	85.342
	BP	0.362	6.314	0.449	29.062	0.461	8.154	0.599	51.726	0.577	10.545	0.734	77.518
	ELM	0.340	5.687	0.431	26.704	0.449	7.846	0.591	50.212	0.575	10.492	0.731	76.942
	ENN	0.345	5.838	0.429	26.554	0.460	8.137	0.598	51.524	0.551	9.948	0.715	73.698
	Proposed Model	0.138	2.209	0.189	5.165	0.174	2.935	0.229	7.563	0.199	3.221	0.256	9.403
Site 2	ARIMA	0.373	6.246	0.471	41.912	0.504	8.558	0.655	64.033	0.669	11.548	0.816	91.082
	BP	0.349	5.821	0.454	29.619	0.485	8.265	0.639	58.774	0.585	10.219	0.753	81.687
	ELM	0.347	5.747	0.448	28.888	0.480	8.152	0.633	57.679	0.591	10.314	0.755	81.991
	ENN	0.347	5.741	0.449	29.079	0.490	8.446	0.646	60.161	0.593	10.319	0.754	81.923
	Proposed Model	0.145	2.352	0.193	5.371	0.183	2.966	0.251	9.057	0.197	3.243	0.270	10.476
Site 3	ARIMA	0.404	7.100	0.470	40.420	0.560	9.869	0.723	73.088	0.692	11.871	0.850	101.396
	BP	0.379	6.330	0.490	34.510	0.531	9.415	0.656	61.957	0.620	11.016	0.790	89.905
	ELM	0.369	6.176	0.478	32.838	0.518	9.031	0.652	61.193	0.611	10.861	0.773	85.930
	ENN	0.383	6.510	0.488	34.299	0.513	9.049	0.647	60.346	0.635	11.426	0.792	90.411
	Proposed Model	0.175	2.804	0.228	7.500	0.197	3.193	0.261	9.830	0.216	3.588	0.285	11.674

Note: The best results are indicated in bold. The measure unit of MAPE in our paper is %.

For Site 1, the proposed combined strategy exceeds the four benchmark strategies in one- and multi-step forecasting. With the increase in prediction steps, the MAPE values of individual models increase significantly, while the values in the proposed model remain stable and fluctuate around two and three. This was also observed for Sites 2 and 3, where the proposed model demonstrated an advantage over the other models.

Taking Site 3 as an example, the MAPE values of our proposed technique are 2.804%, 3.193%, and 3.588% for one- to three-step prediction, followed by the ELM model with MAPE values of 6.176%, 9.031%, and 10.861%, respectively. The ARIMA model produced the largest MAPE values of 7.100%, 9.869%, and 11.871%, respectively. Thus, our developed strategy enhances 3.372%, 5.838%, and 7.273% compared with ELM, and 4.296%, 6.676%, and 8.283% compared with ARIMA from one- to three-step forecasting, respectively. The results of the comparison between the proposed model and the four benchmark models at Site 3 are provided in Figure 5.



Experiment III

Figure 5. Comparison of results of experiment 3 at site 3.

Remark 3. The developed combined model significantly improves forecasting accuracy compared with the benchmark models, as proven by the lower-value evaluation criteria for our proposed model. The MAE, MAPE, RMSE, and SSE values of individual methods are approximate and increase markedly with increasing forecasting steps.

5. Discussion

In this section, seven aspects are discussed to support the forecasting capacity of our proposed combined strategy: forecasting significance, forecasting effectiveness, the degree of improvement,

sensitivity analysis, the performance for MODA optimization algorithm, the application of our model in day-ahead forecasts, and the practical application in power system.

5.1. Forecasting Significance: Diebold–Mariano Test

To investigate the differences in forecasting capability between the proposed strategy and the other methods, a typical statistical test, the Diebold–Mariano (DM) test [44] was conducted. The theory of this test is outlined below.

Firstly, formulate original hypothesis H_0 and alternative hypothesis H_1 . H_0 signifies that the forecasting capability of the developed strategy is analogous to that of the compared model, whereas H_1 is the opposite. The specific hypothetical formulas are written as follows:

$$\begin{aligned} H_0: E[L(e_i^1)] &= E[L(e_i^2)] \\ H_1: E[L(e_i^1)] &\neq E[L(e_i^2)] \end{aligned} \quad (8)$$

where L is the loss function of forecasting errors; e_i^j , $j = 1, 2$ represents the forecasting errors of the other methods.

Subsequently, set the DM statistic, which is expressed as

$$DM = \frac{\sum_{i=1}^n (L(e_i^1) - L(e_i^2)) / n}{\sqrt{S^2 / n}} \mathbf{s}^2 \quad (9)$$

where S^2 is the estimated variance of $\mathbf{d}_i = L(e_i^1) - L(e_i^2)$.

Given a certain level of significance α , the calculated statistics were compared with critical value $Z_{\alpha/2}$. If the DM statistic was not included in the interval $[-Z_{\alpha/2}, Z_{\alpha/2}]$, H_0 was rejected, which means that significant differences exist between our developed combined strategy and the compared model; otherwise, H_0 is accepted.

Table 8 lists the mean DM values from one- to multi-step forecasting. The proposed model is markedly different from individual models including ARIMA, BPNN, ELM, and ENN at the 1% significance level. Although the DM values obtained from the comparison between the developed strategy and SSA-based hybrid methods are not significant as that obtained from the comparison with individual models, the SSA-based combined model has a better forecasting ability compared with four hybrid models at the 10% significance level. Finally, when comparing the developed model with models applying different data pretreatment technologies, the degree of the difference was huge because the DM statistics from one- to three-step all surpass the critical values at the 1% significance level, which signifies a 99% possibility that H_1 will be accepted.

Table 8. Diebold–Mariano (DM) test results for the developed model and comparison models.

Model	One-Step	Two-Step	Three-Step
ARIMA	8.204 *	7.943 *	5.894 *
BPNN	9.448 *	6.820 *	5.130 *
ELM	9.462 *	6.417 *	6.245 *
ENN	8.780 *	7.294 *	6.285 *
SSA-ARIMA	1.971 **	1.916 ***	2.047 **
SSA-BPNN	1.759 ***	2.598 *	4.315 *
SSA-ELM	1.735 ***	2.507 **	4.276 *
SSA-ENN	1.728 ***	2.615 *	4.300 *
EMD-C-Model	8.993 *	6.877 *	4.421 *
CEEMD-C-Model	8.677 *	7.109 *	6.625 *
WDD-C-Model	7.856 *	7.135 *	4.918 *

Note: *, **, and *** denote significance at 1%, 5%, and 10% levels, respectively. The EMD-C-Model, CEEMD-C-Model, and WDD-C-Model represent combined models based on EMD, CEEMD, and WDD technology, respectively.

The DM statistics show that the developed combined strategy is distinguished from the benchmark models, SSA-based hybrid methods, as well as combined forecasting models using different data preprocessing methods. Thus, the developed combined model should be employed in practical wind speed forecasting.

5.2. Forecasting Effectiveness of the Developed Strategy

In this subsection, the forecasting effectiveness is examined to determine whether the developed strategy is superior to other compared models [45]. As for forecasting effectiveness, both first- and second-order are available, where first-order forecasting effectiveness is relative to the expected values of forecasting precision sequences and second-order is relative to the distinction between standard deviation and expected value of the forecasting precision sequences. The forecasting ability increased with increasing forecasting effectiveness. Table 9 lists the computed forecasting effectiveness values of all forecasting models examined in our study, which shows that our proposed model forecasts better than all the other models at the selected three sites. The forecasting effectiveness values obtained from the proposed model were the highest among all models, with first-order values of 0.974, 0.973, and 0.969 and second-order values of 0.952, 0.951, and 0.946 for Sites 1, 2, and 3, respectively. The forecasting effectiveness values obtained for SSA-based hybrid models were similar to those calculated for models based on different data preprocessing techniques. The forecasting effectiveness values for the benchmark models were the lowest.

Table 9. Forecasting effectiveness of our proposed combined strategy.

Model	Site 1		Site 2		Site 3	
	1-Order	2-Order	1-Order	2-Order	1-Order	2-Order
Proposed Model	0.974	0.952	0.973	0.950	0.969	0.946
ARIMA	0.910	0.853	0.919	0.858	0.914	0.845
BPNN	0.922	0.855	0.925	0.859	0.918	0.853
ELM	0.925	0.861	0.925	0.862	0.919	0.856
ENN	0.925	0.861	0.924	0.859	0.916	0.847
SSA-ARIMA	0.947	0.902	0.949	0.903	0.945	0.901
SSA-BPNN	0.962	0.927	0.962	0.929	0.959	0.927
SSA-ELM	0.962	0.927	0.961	0.927	0.957	0.921
SSA-ENN	0.950	0.908	0.951	0.909	0.947	0.903
EMD-C-Model	0.964	0.939	0.963	0.935	0.955	0.925
CEEMD-C-Model	0.968	0.941	0.967	0.942	0.965	0.941
WDD-C-Model	0.966	0.935	0.962	0.932	0.961	0.935

Note: The best results are indicated in bold. The EMD-C-Model, CEEMD-C-Model, and WDD-C-Model represent combined models based on EMD, CEEMD, and WDD technology, respectively.

Thus, we conclude that the SSA-MODA combined model has potential to improve wind speed forecasting accuracy compared with the other selected models. Both the first- and second-order forecasting effectiveness values of the developed strategy are satisfactory, indicating that the proposed model can simultaneously improve forecasting accuracy and stability.

5.3. Improvements Percentage Relative to Other Involved Models

Based on the previous analysis, the proposed combined strategy exceeds the other strategies in terms of forecasting accuracy improvement. To further discuss and evaluate the degree improvement in forecasting when comparing a selected model with the proposed model, we examined four additional metrics: P_{MAE} , P_{MAPE} , P_{RMSE} , and P_{SSE} [46]. The four metrics are briefly described in Table 10; Table 11 provides the improvement percentages of the developed strategy over the selected methods. According to the definition the larger the improvement percentage values, the better the forecasting accuracy of our developed model relative to the selected models. From our analysis of the results in Table 11:

- (1) The developed combined model markedly increases wind speed forecasting accuracy relative to the selected models, with improvement percentages from more than 20% up to 80%.
- (2) When the proposed model was compared with four individual models, the P_{MAE} , P_{MAPE} , P_{RMSE} , and P_{SSE} values obtained for the ARIMA model are the highest. Similar to previous research, the improvement percentages for the majority of metrics for the SSA-based hybrid model and the combined models using other data preprocessing methods are similar, which vary from 30% to 50%.
- (3) These results reveal that the combination of a data pretreatment technique and an optimization algorithm significantly improves forecasting accuracy in our combined model; only adding a data pretreatment technique or only using hybrid model are insufficient to improving forecast precision and stability simultaneously for short-term wind speed forecasting.

Table 10. Brief definition of four metrics.

Metric	Definition	Equation
P_{MAE}	Improvement percentages of MAE	$P_{MAE} = \left \frac{MAE_1 - MAE_2}{MAE_1} \right \times 100\%$
P_{MAPE}	Improvement percentages of MAPE	$P_{MAPE} = \left \frac{MAPE_1 - MAPE_2}{MAPE_1} \right \times 100\%$
P_{RMSE}	Improvement percentages of RMSE	$P_{RMSE} = \left \frac{RMSE_1 - RMSE_2}{RMSE_1} \right \times 100\%$
P_{SSE}	Improvement percentages of SSE	$P_{SSE} = \left \frac{SSE_1 - SSE_2}{SSE_1} \right \times 100\%$

Table 11. Improvement percentages of the proposed model relative to selected models.

Model	Site 1				Site 2				Site 3			
	P_{MAE}	P_{MAPE}	P_{RMSE}	P_{SSE}	P_{MAE}	P_{MAPE}	P_{RMSE}	P_{SSE}	P_{MAE}	P_{MAPE}	P_{RMSE}	P_{SSE}
ARIMA	66.775	68.501	63.875	88.031	66.011	67.511	63.241	87.360	64.522	66.766	62.087	86.504
BPNN	63.493	66.556	62.178	86.020	62.947	64.773	61.332	85.358	61.616	64.183	60.002	84.438
ELM	62.545	65.179	61.527	85.616	62.930	64.641	61.117	85.225	60.792	63.230	59.293	83.883
ENN	62.339	65.030	61.326	85.419	63.238	65.065	61.425	85.450	61.623	64.481	59.839	84.327
SSA-ARIMA	37.691	42.135	36.607	63.746	39.184	42.515	37.365	66.450	36.066	40.843	38.234	62.303
SSA-BPNN	30.559	35.435	29.917	55.639	31.894	34.289	30.725	57.055	30.161	31.493	29.255	54.552
SSA-ELM	33.823	39.356	34.148	59.867	33.469	35.950	31.554	57.666	31.575	33.821	30.856	56.061
SSA-ENN	44.115	47.952	42.203	70.158	44.248	47.069	41.769	69.446	41.693	45.104	41.150	69.859
EMD-C-Model	39.871	41.706	36.611	60.916	41.140	43.711	35.701	59.073	31.346	33.903	26.464	45.976
CEEMD-C-Model	25.637	31.310	22.419	40.437	31.896	34.207	26.042	45.958	20.856	22.754	16.911	31.676
WDD-C-Model	33.641	37.183	30.516	52.302	37.035	39.749	31.240	52.801	27.064	29.497	23.361	41.709

Note: The units of all values revealed in the table are (%). The EMD-C-Model, CEEMD-C-Model, and WDD-C-Model represent combined models based on EMD, CEEMD, and WDD technology, respectively.

5.4. Sensitivity Analysis

To further study how varying the data pretreatment technique parameters and multi-objective optimization algorithm affects the performance of the proposed model, sensitivity analysis was conducted. In the calculation, we used the standard deviation of each error criterion as the novel assessment metric for sensitivity (S_{MAE} , S_{MAPE} , S_{RMSE} , and S_{SSE}) [47]. The robustness decreases with increasing novel assessment metrics. Table 12 provides the calculated results obtained by altering one parameter of SSA or MODA with the remaining parameters unchanged. The relevant parameters in SSA are window length and principal component decomposition number, whereas the parameters in MODA are dragonfly number, iteration number, and archive size.

Table 12. Results of sensitivity indicators obtained from SSA and MODA.

Step	Algorithm	Parameter	Site 1				Site 2				Site 3			
			S_{MAE}	S_{MAPE}	S_{RMSE}	S_{SSE}	S_{MAE}	$MAPE$	S_{RMSE}	S_{SSE}	S_{MAE}	$MAPE$	S_{RMSE}	S_{SSE}
One	SSA	Window Length	0.004	0.085	0.003	0.177	0.007	0.119	0.007	0.410	0.004	0.074	0.004	0.236
		PCDN	0.009	0.154	0.008	0.457	0.003	0.037	0.004	0.240	0.006	0.126	0.006	0.398
	MODA	Dragonfly Number	0.007	0.171	0.005	0.290	0.008	0.186	0.006	0.349	0.004	0.049	0.002	0.144
		Iteration Number	0.009	0.219	0.007	0.417	0.011	0.213	0.012	0.731	0.004	0.072	0.006	0.382
		Archive Size	0.008	0.200	0.006	0.329	0.012	0.210	0.013	0.761	0.005	0.094	0.004	0.279
			0.015	0.363	0.034	3.032	0.007	0.179	0.009	0.635	0.013	0.270	0.014	1.173
Two	SSA	Window Length	0.025	0.603	0.048	4.105	0.007	0.182	0.009	0.666	0.008	0.231	0.008	0.623
		PCDN	0.015	0.363	0.034	3.032	0.007	0.179	0.009	0.635	0.013	0.270	0.014	1.173
	MODA	Dragonfly Number	0.025	0.604	0.052	4.488	0.004	0.100	0.004	0.277	0.019	0.309	0.019	1.587
		Iteration Number	0.062	1.259	0.101	10.730	0.012	0.262	0.015	1.141	0.010	0.229	0.009	0.704
		Archive Size	0.044	1.055	0.096	9.660	0.007	0.166	0.007	0.527	0.013	0.299	0.012	0.975
			0.021	0.527	0.027	2.241	0.010	0.195	0.009	0.720	0.009	0.218	0.012	1.025
Three	SSA	Window Length	0.021	0.527	0.027	2.241	0.010	0.195	0.009	0.720	0.009	0.218	0.012	1.025
		PCDN	0.007	0.152	0.008	0.625	0.007	0.122	0.005	0.360	0.017	0.505	0.021	1.807
	MODA	Dragonfly Number	0.015	0.391	0.019	1.521	0.018	0.307	0.017	1.432	0.012	0.286	0.016	1.383
		Iteration Number	0.021	0.525	0.024	1.948	0.008	0.155	0.007	0.566	0.012	0.316	0.015	1.316
		Archive Size	0.019	0.441	0.022	1.824	0.026	0.475	0.025	2.195	0.014	0.384	0.018	1.532
			0.019	0.441	0.022	1.824	0.026	0.475	0.025	2.195	0.014	0.384	0.018	1.532

Note: The parameter principal component decomposition number of SSA is too long for display, thus, we use PCDN to replace this parameter in this table.

Without loss of generality, the analysis was divided into two parts: the parameters variation in SSA and in MODA. The window lengths considered in SSA were 32, 40, 48, 56, and 64, and the principal component decomposition numbers used were 10, 15, 20, 25, and 30. In the analysis of MODA parameter variation, the assigned dragonfly numbers were 20, 40, 60, 80, and 100; the iteration numbers were 50, 100, 150, 200, and 250; and the archive size was set to 200, 300, 400, 500, and 600. Our analysis of Table 12 provided the following:

- (1) When the parameters in SSA were changed, the S_{MAE} , S_{MAPE} , S_{RMSE} , and S_{SSE} values changed within acceptable limits. For instance, in one-step forecasting at Site 1, the S_{MAE} , S_{MAPE} , and S_{RMSE} , S_{SSE} values for window length were 0.004, 0.085, 0.003, and 0.177, respectively, and the principal component decomposition number values were 0.009, 0.154, 0.008, and 0.457, respectively, which are small and similar to the values obtained for Sites 2 and 3. In other words, the forecasting results of the proposed strategy are not considerably affected by parameter variation in SSA.
- (2) With altering the MODA parameter, the assessment criteria for sensitivity changes within acceptable limits. The results reveal that changing the optimization parameters has a minimal influence on the forecasting results of our developed model, which further verifies the stability of the proposed model.
- (3) The S_{MAE} , S_{MAPE} , S_{RMSE} , and S_{SSE} values in two- and three-step forecasting are larger than in one-step forecasting, revealing that the forecasting robustness declines with parameters variation when increasing the forecasting steps. We found that in two-step forecasting at Site 1, the S_{SSE} values of the five parameters were abnormal and far larger than the other assessment metrics, which can be attributed to the randomness and volatility of the selected dataset. Therefore, choosing suitable parameters in multi-step forecasting is crucial as parameters variation impacts the stability of wind speed forecasting to some extent.

5.5. Comparison in Terms of Optimization Algorithm Performance

To further verify the superiority of the MODA over other optimization algorithms in terms of short-term wind speed forecasting, three combined models based on three good performing optimization algorithms, including Elitist Nondominated Sorting Genetic Algorithm (NSGA II) [48], multi-objective particle swarm optimization (MOPSO) [49], and multi-objective ant lion optimization (MOALO) [50], were used to compare with the developed model with MODA algorithm. The comparative results are established in Table 13, based on which we can conduct the following discussions:

- (1) Regardless of the forecasting steps and observation sites, the employed MODA was always superior to NSGA II, MOPSO, and MOALO indicated by the lower MAE, MAPE, RMSE, and SSE values. For example, the mean MAPE values of the developed model are 2.789%, 2.854%, and 3.195% from Site 1 to Site 3, respectively, whereas the mean MAPE values of combined model using NSGA II algorithm (5.439%, 5.184%, 5.186%, respectively), combined model using MOPSO algorithm (4.735%, 4.605%, 4.574%, respectively), and combined model using the MOALO algorithm (4.151%, 4.373%, 4.313%, respectively) are obviously inferior to the proposed model based on MODA algorithm. Thus, the MODA algorithm can provide better forecasting results relative to NSGA II, MOPSO, and MOALO algorithm in combined short-term wind speed forecasting system.
- (2) From another perspective, with the increase of forecasting steps, the evaluation criteria values of the proposed model changes in a much smaller range compared with other comparative models, which reveals that the developed model with MODA algorithm can provide more stable and robust forecasting performance. In Site 1, the MAPE values of the proposed model from one-step to three-step forecasting are 2.209%, 2.935%, and 3.221%, and in the same condition, the corresponding values obtained from the MOALO based combined model whose forecasting capacity is second only to the MODA algorithm are 3.046%, 4.203%, and 5.203%, respectively. Thus, when the forecasting steps increase, the MAPE values of MOALO based combined model increase rapidly compared with the proposed MODA based combined model, which also applies to NAGA II and MOPSO algorithm. Ultimately, we can conclude that MODA is a powerful tool for solving multi-objective optimization issues in our work, significantly improving forecasting accuracy and stability.

Table 13. Comparison of the evaluation metrics for the proposed strategy with models using different optimization algorithms.

Dataset	Model	One-Step				Two-Step				Three-Step			
		MAE	MAPE	RMSE	SSE	MAE	MAPE	RMSE	SSE	MAE	MAPE	RMSE	SSE
Site 1	NSGA II-C-Model	0.249	4.047	0.322	14.961	0.335	5.541	0.469	31.630	0.362	6.728	0.474	32.408
	MOPSO-C-Model	0.213	3.487	0.277	11.047	0.251	4.220	0.312	14.021	0.398	6.497	0.530	40.378
	MOALO-C-Model	0.181	3.046	0.227	7.389	0.256	4.203	0.326	15.312	0.295	5.203	0.357	18.361
	Proposed model	0.138	2.209	0.189	5.165	0.174	2.935	0.229	7.563	0.199	3.221	0.256	9.403
Site 2	NSGA II-C-Model	0.284	4.575	0.373	20.051	0.304	5.145	0.397	22.665	0.355	5.833	0.475	32.528
	MOPSO-C-Model	0.228	3.757	0.288	11.903	0.260	4.533	0.333	16.007	0.321	5.526	0.427	26.293
	MOALO-C-Model	0.200	3.416	0.251	9.076	0.278	4.394	0.355	18.182	0.290	5.310	0.358	18.464
	Proposed model	0.145	2.352	0.193	5.371	0.183	2.966	0.251	9.057	0.197	3.243	0.270	10.476
Site 3	NSGA II-C-Model	0.292	4.715	0.383	21.076	0.303	5.127	0.395	22.466	0.348	5.716	0.454	29.613
	MOPSO-C-Model	0.216	3.527	0.261	9.817	0.271	4.788	0.359	18.603	0.311	5.407	0.406	23.782
	MOALO-C-Model	0.206	3.300	0.259	9.624	0.274	4.371	0.348	17.398	0.299	5.266	0.366	19.238
	Proposed model	0.175	2.804	0.228	7.500	0.197	3.193	0.261	9.830	0.216	3.588	0.285	11.674

Note: The best results are indicated in bold. NSGA II-C-Model, MOPSO-C-Model, MOALO-C-Model in the table represent the combined model based on Elitist Nondominated Sorting Genetic Algorithm (NSGA II), multi-objective particle swarm optimization (MOPSO)MOPSO, multi-objective ant lion optimization (MOALO) algorithm MOALO, respectively. The measure unit of MAPE in our paper is %.

5.6. The Forecasting Performance Verify of Our Model with Longer Testing Set

In order to further test the long-term prediction ability of our proposed model, we expanded the original data set to 5760 records with a time span of 1 January to 9 February 2011. The ratio of training set to testing set was still 4:1, indicating that 4608 training samples and 1152 testing samples in each dataset and the observation sites are the same as that in the previous experiments. Accordingly, the forecasting horizon of the combined model reaches 288 periods, that is, two days. If the prediction performance of our proposed model is good in this case, it will prove that after one training, our proposed model can maintain a high prediction accuracy and avoid training the proposed model again when rolling the prediction of future wind speed of 288 periods. SSA-BP, CEEMD-C-model, ELM,

and MOALO-C-model are adopted in this discussion as comparative models due to their relatively superior forecasting ability in previous experiments and discussion. The comparison results for the proposed model based on the testing set with two-day length and some good performing models based on the testing set with two-day length are presented in Table 14, where CEEMD-C-model and MOALO-C-model represent the CEEMD-MODA based combined model and SSA-MOALO based combined model. From the table, we can conduct that the developed model can always provide accurate and stable forecasting when the testing set length becomes two days. The detailed analysis in terms of the comparative results are presented in the following:

- (1) With longer duration of continuous prediction, the proposed combined model can achieve satisfactory forecasting performance in the selected observation sites whether in one-step forecasting or multi-step forecasting. For example, the mean MAPE of the proposed model in Site 1 is 3.133% and corresponding mean MAE, RMSE, and SSE values are 0.191, 0.253, and 8.294, which improve to a great extent compared with ELM forecasting model whose forecasting capacity is the worst among all selected models. Similar to the forecasting results based on the testing set with one-day length, the proposed model can always provide more stable forecasting results because the evaluation criteria values increase slightly with the increase of forecasting steps, whereas the values in other models increase to a great extent.
- (2) Compared with the forecasting results based on the testing set with one-day length, the simulation results based on the testing set with two-day length (intra-day and the next day's forecasting) display slightly higher errors, either one step or multi-step. This is understandable because more uncertain information is contained in forecasting process when the forecasting length increases. For example, the mean MAPE values of the proposed model in this section are 3.133%, 3.282%, and 3.604%, respectively and the corresponding values in the forecasting based on the testing set with one-day length are 2.789%, 2.854%, and 3.195% from Site 1 to Site 3, respectively. It does not mean that our developed model is invalid for wind speed forecasting based on the testing set with a longer length because the MAE, MAPE, RMSE, and SSE values are all better than the compared models with more accurate and stable forecasting performance. Thus, it is credible and reliable that the developed model can achieve great forecasting performance whether based on the testing set with one-day length or two-day length.

Table 14. Comparison of forecasting evaluation metrics based on the testing set with two days length for the proposed strategy with some good performing models.

Dataset	Model	One-Step				Two-Step				Three-Step			
		MAE	MAPE	RMSE	SSE	MAE	MAPE	RMSE	SSE	MAE	MAPE	RMSE	SSE
Site 1	SSA-BP	0.175	3.049	0.229	6.818	0.219	3.769	0.291	11.320	0.424	7.607	0.551	37.966
	CEEMD-C-model	0.195	3.285	0.244	7.985	0.275	4.671	0.351	15.407	0.286	5.440	0.362	17.554
	ELM	0.366	6.113	0.463	28.707	0.482	8.435	0.635	53.978	0.661	12.065	0.841	88.484
	MOALO-C-model	0.194	3.274	0.244	7.943	0.275	4.518	0.351	16.461	0.317	5.593	0.384	19.738
	Proposed model	0.158	2.541	0.218	5.940	0.187	3.155	0.246	8.130	0.229	3.704	0.294	10.813
Site 2	SSA-BP	0.181	3.101	0.243	9.166	0.263	4.359	0.347	15.092	0.403	6.831	0.540	39.066
	CEEMD-C-model	0.223	3.801	0.282	9.953	0.276	4.657	0.364	17.777	0.345	5.778	0.410	22.458
	ELM	0.373	6.178	0.482	31.055	0.552	9.375	0.728	66.331	0.679	11.861	0.868	94.289
	MOALO-C-model	0.230	3.929	0.289	10.438	0.277	4.669	0.398	39.645	0.333	6.106	0.412	21.234
	Proposed model	0.167	2.705	0.222	6.177	0.210	3.411	0.288	10.415	0.227	3.730	0.310	12.047
Site 3	SSA-BP	0.253	4.015	0.333	16.787	0.256	4.267	0.332	14.746	0.440	7.510	0.571	40.829
	CEEMD-C-model	0.214	3.536	0.274	10.019	0.270	4.385	0.345	15.932	0.336	5.795	0.410	21.057
	ELM	0.397	6.639	0.513	35.301	0.556	9.708	0.701	65.782	0.703	12.490	0.888	98.820
	MOALO-C-model	0.237	3.795	0.297	11.068	0.294	4.699	0.374	18.703	0.343	6.056	0.420	22.123
	Proposed model	0.188	3.014	0.245	8.063	0.226	3.672	0.301	11.304	0.248	4.126	0.327	13.425

Note: The best results are indicated in bold. The measure unit of MAPE in our paper is %. CEEMD-C-model represents combined models based on CEEMD technology and MOALO-C-Model represents the combined model based on multi-objective ant lion optimization (MOALO) algorithm.

5.7. Pragmatic Applications in Power Systems

Accurate wind speed forecasting is a prerequisite to improve the stability of wind power systems and the efficiency of wind power generation [51,52]. As an improvement to the existing wind power systems, the developed wind speed prediction strategy, with the target of improving prediction accuracy and stability, shows the ability to minimize the risks of wind power generation due to wind variation and achieve real-time forecasting in terms of wind power generation scheduling and safe operation [53]. The contributions of an effective wind speed forecasting system to power systems are listed below:

- Optimizing the potential wind energy in every wind farm is important for maintaining the upper limit of wind energy output. As wind energy is proportional to the cube of wind speed, precisely predicting wind speed is critical so that the wind energy capacity can be determined, and smart grids can be effectively planned by decision makers [54].
- With wind speed forecasting, decisions regarding the operation and administration of wind turbines can be credibly made. To ensure the largest output of wind energy, managers can alter the wind turbine potential without delay according to the predicted wind. Once the wind turbine capacity is under the forecasted wind speed, the output can be closed to prevent breakdown and minimize operation costs.
- The scheduling and management of power systems, to a large extent, rely on wind speed forecasting. We must consider the balance between power demand and supply to satisfy demand and ensure sustainable energy development. Excess power supply poses a problem in practical applications, which may result in supply quality degradation, power system insecurity, and operational cost increases [55]. Thus, it is imperative to formulate more exact and stable wind speed forecasting models to assist decision makers to make timely decisions so that these problems can be effectively addressed.

6. Conclusions

Short-term wind speed forecasting, an important tool in wind energy research and practical applications, has been growing in popularity, mostly due to its importance in the scheduling and operation of power grids [56]. However, wind speed time series are fluctuating and stochastic, which creates challenges for wind speed forecasting and generation. Anticipating this constraint is important because higher uncertainty requires more comprehensive forecasting models to meet specified accuracy and stability objectives [57]. In this study, we developed a combined strategy that effectively integrates data preprocessing technology and an optimization algorithm and significantly improves wind speed forecasting accuracy and stability. To overcome the passive impacts of high-frequency noise on forecasting performance, an effective de-noising method was firstly applied to decompose raw wind speed sequences and retain the fundamental characteristics of wind speed data. Then, a typical statistical model, ARIMA, and three classical neural networks, BPNN, ELM, and ENN, were used as wind speed forecasting system benchmarks. To combine these forecasting models, MODA was successfully used to confirm the optimal weight coefficient of our developed SSA-MODA system. Based on the above processes and actual wind speed sequences, three numerical experiments, four evaluation criteria, and seven discussions were conducted to verify the accuracy and stability of our developed model. In experiment 1, the MAPE values of one-step forecasting obtained at Sites 1 to 3 in our developed strategy were 2.209%, 2.935%, and 3.221%, respectively, whereas the lowest MAPE values in the compared models were 2.836%, 3.506%, and 6.310%, respectively, which are larger than in the developed model. In Experiments 2 and 3, regardless of the forecasting step or the observation site, the proposed combined strategy was always superior to all the selected methods because the performance metrics were the lowest and fluctuated over a small range, indicating that the proposed strategy accurately and stably predicts short-term wind speed. Moreover, in the discussions, we have verified the forecasting significance and forecasting effectiveness of our developed model and further

proved that our proposed model can provide accurate and stable wind speed forecasting performance. To verify the superiority of MODA, we conduct a comparison among MODA, NAGA II, MOPSO, and MOALO, and the results indicate that the MODA is a helpful choice for short-term wind speed forecasting in our work. Besides, the developed model is used to intra-day and next day forecasting and the simulation results have testified the effectiveness of our developed model in both intra-day and next day forecasting, which can help wind producers bid their offers in day-ahead schedule. Ultimately, the practical applications of the developed model in power system are illuminated in detail to verify the practical value of the developed model. Overall, the developed model significantly increases short-term wind speed forecasting accuracy and stability, providing a reference for smart grid planning.

The main limitation of this research is that only power systems including wind farms were considered, whereas other fields including stock price forecasting were not included. Recommendations for improvements in our work in the future include:

- (1) Applying more effective data pretreatment techniques to handle the stochastic and fluctuating wind speed data to increase the forecasting accuracy.
- (2) Improving existing optimization algorithms to enhance the global search ability and convergence speed and further improving optimization performance and forecasting accuracy of combined models.
- (3) Developing new neural network models and statistic methods to grasp the characteristic of wind speed series and constructing fundamental forecasting models.

Author Contributions: Software, Z.L.; Supervision, J.W.; Writing—original draft, Y.D.; Writing—review and editing, L.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the National Natural Science Foundation of China (Grant No. 71761016), Natural Science Foundation of Jiangxi (Grant No. 20171BAA218001), Scientific Research Fund of Jiangxi Provincial Education Department (Grant No. GJJ180287), Postdoctoral Foundation of Jiangxi Province (Grant No. 2018KY08) and Humanities and Social Sciences Foundation of Jiangxi University (Grant No. TJ19202).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

FL	fuzzy logic
KF	Kalman filtering
ENN	elman neural network
ANN	artificial neural network
RMSE	root mean square error
IMFs	intrinsic mode functions
NWP	numerical weather prediction
BP	back propagation neural network
EEMD	ensemble empirical mode decomposition
MOALO	multi-objective ant lion optimization
ARIMA	autoregressive integrated moving average model
ARFIMA	autoregressive fractional integrated moving average
CEEMD	complete ensemble empirical mode decomposition
MODA	multi-objective dragonfly optimization algorithm
MOGWO	multi-objective grey wolf optimization algorithm
MC	Markov chain
CS	cuckoo search
MAE	mean absolute error
SSE	error sum of square
WNN	wavelet neural network

DM Test	Diebold Mariano test
ELM	extreme learning machine
WDD	wavelet domain de-noising
SSA	singular spectrum analysis
EMD	empirical model decomposition
MAPE	mean absolute percentage error
DA	dragonfly optimization algorithm
MOPSO	multi-objective particle swarm optimization
NSGA II	Elitist Nondominated Sorting Genetic Algorithm

References

- Li, R.; Jin, Y. A wind speed interval prediction system based on multi-objective optimization for machine learning method. *Appl. Energy* **2018**, *228*, 2207–2220. [[CrossRef](#)]
- Salcedo-Sanz, S.; Pérez-Bellido, Á.M.; Ortiz-García, E.G.; Portilla-Figueras, A.; Prieto, L.; Correoso, F. Accurate short-term wind speed forecasting by exploiting diversity in input data using banks of artificial neural networks. *Neurocomputing* **2009**, *72*, 1336–1341. [[CrossRef](#)]
- Data on Global Onshore and Offshore Wind Power in 2018. Available online: <http://news.bjx.com.cn/html/20190516/980725.shtml> (accessed on 30 June 2019).
- Calif, R.; Schmitt, F. Multiscaling and joint multiscaling of the atmospheric wind speed and the aggregate power output from a wind farm. *Nonlinear Processes Geophys.* **2014**, *21*, 379–392. [[CrossRef](#)]
- Jiang, P.; Liu, Z. Variable weights combined model based on multi-objective optimization for short-term wind speed forecasting. *Appl. Soft Comput.* **2019**, *82*, 105587. [[CrossRef](#)]
- Zhou, Q.; Wang, C.; Zhang, G. Hybrid forecasting system based on an optimal model selection strategy for different wind speed forecasting problems. *Appl. Energy* **2019**, *250*, 1559–1580. [[CrossRef](#)]
- Bo, H.; Niu, X.; Wang, J. Wind Speed Forecasting System Based on the Variational Mode Decomposition Strategy and Immune Selection Multi-Objective Dragonfly Optimization Algorithm. *IEEE Access* **2019**, *7*, 178063–178081. [[CrossRef](#)]
- Zhao, X.; Liu, J.; Yu, D.; Chang, J. One-day-ahead probabilistic wind speed forecast based on optimized numerical weather prediction data. *Energy Convers. Manag.* **2018**, *164*, 560–569. [[CrossRef](#)]
- Alessandrini, S.; Sperati, S.; Pinson, P. A comparison between the ECMWF and COSMO Ensemble Prediction Systems applied to short-term wind power forecasting on real data. *Appl. Energy* **2013**, *107*, 271–275. [[CrossRef](#)]
- Sile, T.; Bekere, L.; Cepite-Frisfelde, D.; Sennikovs, J.; Bethers, U. Verification of numerical weather prediction model results for energy applications in Latvia. *Energy Procedia* **2014**, *59*, 213–220. [[CrossRef](#)]
- Liu, H.; Tian, H.-Q.; Chen, C.; Li, Y. A hybrid statistical method to predict wind speed and wind power. *Renew. Energy* **2010**, *35*, 1857–1861. [[CrossRef](#)]
- Shukur, O.B.; Lee, M.H. Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA. *Renew. Energy* **2015**, *76*, 637–647. [[CrossRef](#)]
- Yuan, X.; Tan, Q.; Lei, X.; Yuan, Y.; Wu, X. Wind power prediction using hybrid autoregressive fractionally integrated moving average and least square support vector machine. *Energy* **2017**, *129*, 122–125. [[CrossRef](#)]
- Jiang, H. Model forecasting based on two-stage feature selection procedure using orthogonal greedy algorithm. *Appl. Soft Comput.* **2018**, *63*, 10–15. [[CrossRef](#)]
- Du, P.; Wang, J.; Yang, W.; Niu, T. A novel hybrid model for short-term wind power forecasting. *Appl. Soft Comput. J.* **2019**, *80*, 93–106. [[CrossRef](#)]
- Jiang, H.; Dong, Y. Structural regularization in quadratic logistic regression model. *Knowl.-Based Syst.* **2019**, *163*, 842–857. [[CrossRef](#)]
- Wang, R.; Li, J.; Wang, J.; Gao, C. Research and Application of a Hybrid Wind Energy Forecasting System Based on Data Processing and an Optimized Extreme Learning Machine. *Energies* **2018**, *11*, 1712. [[CrossRef](#)]
- Hao, Y.; Tian, C.; Wu, C. Modelling of carbon price in two real carbon trading markets. *J. Clean. Prod.* **2020**, *244*, 118556. [[CrossRef](#)]
- Tian, C.; Hao, Y. Point and interval forecasting for carbon price based on an improved analysis-forecast system. *Appl. Math. Model.* **2019**. [[CrossRef](#)]

20. Wang, J.; Wu, C.; Niu, T. A Novel System for Wind Speed Forecasting Based on Multi-Objective Optimization and Echo State Network. *Sustainability* **2019**, *11*, 526. [[CrossRef](#)]
21. Pourmousavi Kani, S.A.; Ardehali, M.M. Very short-term wind speed prediction: A new artificial neural network–Markov chain model. *Energy Convers. Manag.* **2011**, *52*, 738–745. [[CrossRef](#)]
22. Yang, W.; Wang, J.; Lu, H.; Niu, T.; Du, P. Hybrid wind energy forecasting and analysis system based on divide and conquer scheme: A case study in China. *J. Clean. Prod.* **2019**, *222*, 942–959. [[CrossRef](#)]
23. Xiao, L.; Shao, W.; Yu, M.; Ma, J.; Jin, C. Research and application of a hybrid wavelet neural network model with the improved cuckoo search algorithm for electrical power system forecasting. *Appl. Energy* **2017**, *198*, 203–205. [[CrossRef](#)]
24. Zhao, X.; Wang, C.; Su, J.; Wang, J. Research and application based on the swarm intelligence algorithm and artificial intelligence for wind farm decision system. *Renew. Energy* **2019**, *134*, 681–697. [[CrossRef](#)]
25. Wang, J.; Niu, T.; Lu, H.; Yang, W.; Du, P. A Novel Framework of Reservoir Computing for Deterministic and Probabilistic Wind Power Forecasting. *IEEE Trans. Sustain. Energy* **2019**. [[CrossRef](#)]
26. Wu, C.; Wang, J.; Chen, X.; Du, P.; Yang, W. A Novel Hybrid System Based on Multi-objective Optimization for Wind Speed Forecasting. *Renew. Energy* **2019**. [[CrossRef](#)]
27. Liu, H.; Chen, C.; Tian, H.Q.; Li, Y.F. A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. *Renew. Energy* **2012**, *48*, 545–556. [[CrossRef](#)]
28. Bates, J.M.; Granger, C.W.J. The Combination of Forecasts. *J. Oper. Res. Soc.* **1969**, *20*, 451–468. [[CrossRef](#)]
29. Niu, T.; Wang, J.; Lu, H.; Du, P. Uncertainty modeling for chaotic time series based on optimal multi-input multi-output architecture: Application to offshore wind speed. *Energy Convers. Manag.* **2018**, *156*, 597–617. [[CrossRef](#)]
30. Zhang, S.; Wang, J.; Guo, Z. Research on combined model based on multi-objective optimization and application in time series forecast. *Soft Comput.* **2019**, *23*, 11493–11521. [[CrossRef](#)]
31. Hao, Y.; Tian, C. The study and application of a novel hybrid system for air quality early-warning. *Appl. Soft Comput.* **2019**, *74*, 729–746. [[CrossRef](#)]
32. Jiang, H. Sparse estimation based on square root nonconvex optimization in high-dimensional data. *Neurocomputing* **2018**, *282*, 122–125. [[CrossRef](#)]
33. Niu, X.; Wang, J. A combined model based on data preprocessing strategy and multi-objective optimization algorithm for short-term wind speed forecasting. *Appl. Energy* **2019**, *241*, 519–539. [[CrossRef](#)]
34. Li, H.; Wang, J.; Lu, H.; Guo, Z. Research and application of a combined model based on variable weight for short term wind speed forecasting. *Renew. Energy* **2018**, *116*, 669–684. [[CrossRef](#)]
35. Hassani, H.; Ghodsi, Z. A glance at the applications of Singular Spectrum Analysis in gene expression data. *Biomol. Detect. Quantif.* **2015**, *4*, 17–21. [[CrossRef](#)]
36. Krishnannair, S.; Aldrich, C.; Jemwa, G.T. Detecting faults in process systems with singular spectrum analysis. *Chem. Eng. Res. Des.* **2016**, *113*, 151–155. [[CrossRef](#)]
37. Unnikrishnan, P.; Jothiprakash, V. Daily rainfall forecasting for one year in a single run using Singular Spectrum Analysis. *J. Hydrol.* **2018**, *561*, 609–621. [[CrossRef](#)]
38. De Carvalho, M.; Rua, A. Real-time nowcasting the US output gap: Singular spectrum analysis at work. *Int. J. Forecast.* **2015**, *33*, 185–198. [[CrossRef](#)]
39. Mirjalili, S. Dragonfly algorithm: A new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. *Neural Comput. Appl.* **2015**, *27*, 1053–1073. [[CrossRef](#)]
40. Li, H.; Wang, J.; Li, R.; Lu, H. Novel analysis–forecast system based on multi-objective optimization for air quality index. *J. Clean. Prod.* **2019**, *208*, 1365–1383. [[CrossRef](#)]
41. Yang, W.; Wang, J.; Niu, T.; Du, P. A hybrid forecasting system based on a dual decomposition strategy and multi-objective optimization for electricity price forecasting. *Appl. Energy* **2019**. [[CrossRef](#)]
42. Liu, Z.; Jiang, P.; Zhang, L.; Niu, X. A combined forecasting model for time series: Application to short-term wind speed forecasting. *Appl. Energy* **2019**, 114137. [[CrossRef](#)]
43. Yang, H.; Zhu, Z.; Li, C.; Li, R. A novel combined forecasting system for air pollutants concentration based on fuzzy theory and optimization of aggregation weight. *Appl. Soft Comput.* **2019**, 105972. [[CrossRef](#)]
44. Diebold, F.X.; Mariano, R.S. Comparing predictive accuracy. *J. Bus. Econ. Stat.* **1995**, *13*, 253–265. [[CrossRef](#)]
45. Chen, H.; Hou, D. Research on superior combination forecasting model based on forecasting effective measure. *J. Univ. Sci. Technol. China* **2002**, *32*, 172–180.

46. Mi, X.; Liu, H.; Li, Y. Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine. *Energy Convers. Manag.* **2019**, *180*, 196–205. [[CrossRef](#)]
47. Wang, J.; Du, P.; Lu, H.; Yang, W.; Niu, T. An improved grey model optimized by multi-objective ant lion optimization algorithm for annual electricity consumption forecasting. *Appl. Soft Comput. J.* **2018**, *72*, 321–337. [[CrossRef](#)]
48. Sarshar, J.; Moosapour, S.S.; Joorabian, M. Multi-objective energy management of a micro-grid considering uncertainty in wind power forecasting. *Energy* **2017**. [[CrossRef](#)]
49. He, Z.; Chen, Y.; Shang, Z.; Li, C.; Li, L.; Xu, M. A novel wind speed forecasting model based on moving window and multi-objective particle swarm optimization algorithm. *Appl. Math. Model.* **2019**. [[CrossRef](#)]
50. Du, P.; Wang, J.; Guo, Z.; Yang, W. Research and application of a novel hybrid forecasting system based on multi-objective optimization for wind speed forecasting. *Energy Convers. Manag.* **2017**. [[CrossRef](#)]
51. Zhou, J.; Sun, N.; Jia, B.; Peng, T. A Novel Decomposition-Optimization Model for Short-Term Wind Speed Forecasting. *Energies* **2018**, *11*, 1752. [[CrossRef](#)]
52. Iversen, E.B.; Morales, J.M.; Møller, J.K.; Madsen, H. Short-term probabilistic forecasting of wind speed using stochastic differential equations. *Int. J. Forecast.* **2015**, *32*, 981–990. [[CrossRef](#)]
53. Sun, N.; Zhou, J.; Chen, L.; Jia, B.; Tayyab, M.; Peng, T. An adaptive dynamic short-term wind speed forecasting model using secondary decomposition and an improved regularized extreme learning machine. *Energy* **2018**. [[CrossRef](#)]
54. Zuluaga, C.D.; Álvarez, M.A.; Giraldo, E. Short-term wind speed prediction based on robust Kalman filtering: An experimental comparison. *Appl. Energy* **2015**. [[CrossRef](#)]
55. Zhang, X.; Wang, J.; Gao, Y. A hybrid short-term electricity price forecasting framework: Cuckoo search-based feature selection with singular spectrum analysis and SVM. *Energy Econ.* **2019**, *81*, 899–913. [[CrossRef](#)]
56. Wang, J.; Gao, Y.; Chen, X. A novel hybrid interval prediction approach based on modified lower upper bound estimation in combination with multi-objective salp swarm algorithm for short-term load forecasting. *Energies* **2018**, *11*, 1561. [[CrossRef](#)]
57. Zhang, L.; Dong, Y.; Wang, J. Wind Speed Forecasting Using a Two-Stage Forecasting System with an Error Correcting and Nonlinear Ensemble Strategy. *IEEE Access* **2019**, *7*, 176000–176023. [[CrossRef](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).