

Elsevier required licence: © <2019>. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

The definitive publisher version is available online at [<http://doi.org/10.1016/j.jclepro.2019.03.036>]

Hybrid wind energy forecasting and analysis system based on divide and conquer scheme: a case study in China

Wendong Yang^{a,b}, Jianzhou Wang^{a,b*}, Haiyan Lu^b, Tong Niu^{a,b}, Pei Du^a

^a School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China

^b School of Software, Faculty of Engineering and Information Technology, University of Technology, Sydney, Australia

** Corresponding author. Address: School of Statistics, Dongbei University of Finance and Economics, Dalian, Liaoning 116025, China*

Tel.: +86 18340828108

E-mail address: wangjzh62@yeah.net

Abstract

Wind energy, acknowledged as a promising form of renewable energy and the fastest-growing clean method for electricity generation, has attracted considerable attention from many scientists and researchers in recent decades. However, wind energy forecasting is still a challenging task owing to its inherent features of non-linearity and randomness. Therefore, this study develops a hybrid wind energy forecasting and analysis system including a deterministic forecasting module and an uncertainty analysis module to mitigate the challenges in existing studies. In particular, these challenges are as follows: (1) It is difficult to guarantee that the data characteristics underlying the time series are effectively extracted; (2) in the modeling of each subseries, i.e., when the original data is decomposed into some time series, forecasting accuracy and stability are not simultaneously considered, and thus, they are not properly modeled; and (3) the best function to perform a deterministic forecasting and uncertainty analysis based on the forecaster of each subseries is unknown. The developed hybrid system consists of three steps: First, data preprocessing is conducted to capture and mine the main feature of the wind energy time series and weaken the noises' negative effects; second, multi-objective optimization is proposed to achieve the forecasting of each subseries with improvements in accuracy and stability; finally, search for the best function, which obtains the deterministic forecasting and uncertainty analysis results using an optimized extreme learning machine based on different modeling objectives, is conducted. Experimental simulations are performed using data from three sites in a real wind farm, which indicate that the developed system has a better performance in engineering applications than that of other methods. Furthermore, this system could not only be used as an effective tool for wind energy deterministic forecasting and uncertainty analysis, but also for other engineering application areas in the future.

Keyword: *hybrid system; deterministic forecasting; uncertainty analysis; optimization; wind energy*

1. Introduction

In recent years, China has faced enormous environmental and energy challenges, which have a significant influence on economic growth, public health, and welfare (Li et al., 2018a). With the decline of conventional energy sources and increasing environmental pollution, the energy structure based on fossil fuels will be replaced by a cleaner energy structure (Zendehboudi et al., 2018). Wind energy has become increasingly popular worldwide, and it can make a great contribution to reduce CO₂ emissions (Wang et al., 2008; Wang and Li, 2019). To the best of our knowledge, wind energy, as the renewable source with greatest potential around the world, has the highest growth rate in the electrical power system (EPS) owing to its environmental and ecological friendliness and because it is easily extractable with the rapid development of wind power technology (Wang et al., 2018a; Liang et al., 2016; Wang et al., 2018b). According to the Global Wind Report 2017 (Global Wind Energy Council, 2018), the global wind power market remained above 50 GW in 2017, and the total installations in 2017 were of 52,492 MW, bringing the global total to 539,123 MW. The annual market was down 3.8% compared with 2016, and the cumulative total increased 11% over 487,279 MW at the end of 2016. With a new added capacity to the power grid of 19,660 MW, China again led the global markets in 2017. It was down 15.9% compared with 2016, but China still represents 37% of the global installations. Moreover, it represents an increase of 11.7% over the previous year and cumulative installations of 188,392 MW. Therefore, to accelerate the wind power development, wind energy data is acknowledged as a real-world complex time series of great importance for the EPS, and its forecasting is presently one of the most important issues in clean production.

However, the large-scale integration of wind power could threaten the operation and planning of conventional power systems owing to the features of non-linearity and randomness inherent to wind energy, which severely hinder further development of wind power (Zhao et al., 2016a). One of the solutions that may help solve the above-mentioned problems is developing an effective forecasting technique, which could play a significant role in power plants to control the balancing, operation, and safety of the grid (Li et al., 2018b). A general, accurate wind energy forecasting can reduce the financial and technical risks of the uncertainty of wind power for all electricity market participants (Foley et al., 2012). However, because of the nature of wind energy, it is

not possible to ensure its availability when needed, and hence, wind energy forecasting is a challenging issue for wind power development (Zhao et al., 2016b). For this purpose, proposing an effective forecasting technique for wind energy that considers the relevant issues associated with the large-scale integration of wind power into the electricity grid has become particularly desirable and should not be delayed (Ma et al., 2009).

In recent decades, many forecasting methods for wind energy have been proposed, which mainly include three families, namely, the physical, conventional statistical, and intelligent forecasting approaches (Zhang et al., 2017; Jiang et al., 2019; Hao and Tian, 2019a). The physical approaches are developed according to some meteorological factors, such as temperature, humidity, surface roughness, and pressure, and they perform better for long-term wind energy forecasting (Su et al., 2014). Moreover, they are not suitable for conducting small-area and short-term wind energy forecasting because they need extensive simulation time and resources (Sun and Liu, 2016). In contrast, the conventional statistical approaches, such as the autoregressive moving average (ARMA) (Torres et al., 2005) and the autoregressive integrated moving average (ARIMA) model (Sfetsos, 2002), aim to forecast the future changes by historical data, and although they are more suitable for wind energy short-term forecasting than the physical models, they fail to deal with nonlinear conditions (Xiao et al., 2017). With the rapid development of artificial intelligence, intelligent forecasting models have been successfully developed and employed in wind energy forecasting (Jiang et al., 2016) and, among them, the artificial neural networks (ANNs) (Wang et al., 2018c; Tian et al., 2018), support vector machines (SVMs) (Hu et al., 2015), and fuzzy logic methods (Hong et al., 2010) are discussed mostly. Based on the various forecasting methods, many forecasting systems have been developed around the world. For example, the e WindTM system developed by AWS TruePower, USA, whose geographical locations of applications include the USA; the WPPT system developed by Eltra/Elsam in collaboration with Informatics and Mathematical Modeling at the Danmarks Tekniske Universitet (DTU), Denmark, with geographical locations of applications that include Denmark, Australia, Canada, Republic of Ireland, Holland, Sweden, Greece, and Northern Ireland; the Sipleóico system developed by the Universidad Carlos III, Madrid, Spain and Red Eléctrica de Espana, whose corresponding geographical locations of applications include Spain (Foley et al., 2012).

Moreover, Ma et al. (2019) point out that only the nonlinear models can obtain the desirable forecasting performance because the wind energy data usually present strong nonlinear phenomena. However, owing to the disadvantages that are inherent to the single methods, they cannot always obtain the desired forecasting results (Du et al., 2018, Ma et al., 2017a). In addition, owing to the increasing penetration of wind power in the power grids, it is necessary to have a better understanding of the forecasting error and to try to reduce it as much as possible (Bludszuweit et al., 2008).

Therefore, to remedy these weaknesses that are inherent to the single methods, many hybrid systems have been proposed for wind energy forecasting, which achieve better forecasting performance (Tian et al., 2018). In general, according to the data preprocessing strategy, the hybrid model could be developed by the following two strategies (Yang et al., 2017). (1) The decomposition and reconstruction strategy (Yang et al., 2017), which is a simple data preprocessing scheme that decomposes the original data into some subseries with different frequencies and removes the high frequency signals to obtain the reconstructed series. Then, the reconstructed data are used to develop the forecasting model. Ma et al. (2017b) developed a wind energy forecasting model based on singular spectrum analysis (SSA) to obtain a smoother sequence based on the original series. Similarly, Niu et al. (2018) also proposed a hybrid wind energy forecasting model based on SSA. Their case study proved that the developed model, based on the decomposition and reconstruction strategy, not only outperforms other methods, but also could be employed as an effective technique in engineering applications. (2) The divide and conquer strategy (Xu et al., 2017), another data preprocessing strategy, is the most widely used scheme in hybrid forecasting modeling. It employs certain data preprocessing technique to decompose the original data into some subseries and then constructs a predictor for each subseries. The final forecasting results can be obtained by aggregating all the predicted subseries. For instance, Wang et al. (2017) developed a forecasting system based on data preprocessing and optimization for wind energy. Similarly, Meng et al. (2016) developed a hybrid model for short term wind energy forecasting that combines data preprocessing, ANNs, and the optimization approach. According to the evaluation of the performance of the developed model, this method significantly enhances the forecasting accuracy compared with that of all the benchmark models. The results of the abovementioned

hybrid models based on these two schemes verified the forecasting ability of the hybrid models.

Based on the analysis and review of the abovementioned literature, it could be found that, in most cases, the divide and conquer strategy-based models are superior to the models using the decomposition and reconstruction strategy. The former methods improve the final forecasting performance to a certain extent and are now employed in mainstream applications. However, most of the previous studies emphasizing the contribution of the divide phase only improve the forecasting effectiveness to a certain degree. Most importantly, they ignore the significance of the conquer phase for the final forecasting performance, which includes two aspects. On the one hand, most of them adopted a traditional optimization algorithm to optimize the forecasting model for forecasting each subseries, which usually results in poor forecasting stability. On the other hand, and equally important, most of them employed a widely used integration way, named the direct integration method, to obtain the final forecasting results, whereas few of them considered other integration ways in the conquer phase, which leaves an area of potential improvement for future studies. Therefore, it is worth to conduct further research and studies on novel methods for the conquer phase.

Another issue of wind energy forecasting is the limited span of research directions. Specifically, most prior analyses have concentrated only on the deterministic wind energy prediction, which is insufficient for engineering applications and cannot ensure the reliability and controllability of the EPS. Fortunately, the probabilistic interval forecasting can provide more information, and its results will facilitate decision makers to conduct risk analysis and assessment. However, the research and applications of interval forecasting have not obtained enough attention from the relevant researchers, who are mainly centered on performing analyses based on the statistical model, such as the bootstrap methods (Errouissi et al., 2015), quantile regression method (Nielsen et al., 2016; Wang et al., 2016), and kernel density estimation method (Juban et al. 2007). In addition, forecasting models based on ANNs, the lower upper bound estimation (LUBE), and the nonparametric theory were developed for interval forecasting (Khosravi et al., 2011). Based on the review and analysis of the abovementioned interval forecasting methods, it can be found that: (1) the bootstrap method is adapted to a small sample, which can avoid possible discards in a quantile regression but with a heavy computational burden; (2) except for the heavy computational burden, the

quantile regression method needs a specific training sample for interval forecasting, with the probability of results being discarded in the process of resampling; (3) although the kernel density estimation method is easily performed for interval forecasting, it needs strict assumptions on distribution; (4) the LUBE method not only avoids the assumptions about distribution but also have high computational efficiency, although its complex objective function cannot be solved using the traditional mathematical method. In summary, there is not a single uniform method for interval forecasting, and further studies and investigations need to be conducted based on the currently available knowledge for obtaining more effective results (Moghram and Rahman, 1989). More specifically, only few of the past studies have performed in-depth analyses of the wind energy uncertainty and they have merely emphasized the modeling of wind energy centered on deterministic forecasting. More importantly, they have ignored the potential improvements and significance of obtaining high quality interval forecasting. Therefore, for the purpose of providing more information for decision makers in EPS, further studies of deterministic and probabilistic interval forecasting in the wind energy forecasting field are especially needed because of its significance.

The main contribution of this study is that we developed a hybrid forecasting and analysis system to perform wind energy deterministic and probabilistic interval forecasting, which compensates the insufficiency of existing studies. This system consists of a deterministic forecasting module and an uncertainty analysis module, which can provide more information for decision makers in EPS and improve the management and scheduling of EPS. Experimental simulations are performed using data from three sites in a real wind farm, which indicate that the developed system performs better in engineering applications than the other compared models.

In particular, the main differences between the present research work and previous studies involve four points, which can be considered as a strong contribution to the wind energy forecasting domain. The details can be summarized as follows: (1) Different forecast target: This research focuses on improving the model's accuracy and stability rather than only the accuracy target. Therefore, an optimized extreme learning machine (ELM) model is developed, which combines a multi-objective optimization algorithm with two targets (accuracy and stability) to provide a guarantee for the final forecasting performance. (2) Different integration method: By focusing on the conquer phase rather than on the divide phase when using the superior divide and conquer

strategy. In this study, another optimized ELM is developed as a new integration method to replace the simple direct integration method, which can obtain the final forecasting results by the forecasted future changes of each subseries. Owing to its outstanding forecasting ability, the effectiveness of the conquer stage and final forecasting performance can be further enhanced. (3) Different research directions: The focus is on providing both deterministic forecasting information and uncertainty analysis. Besides a deterministic forecasting module with two aims, a probabilistic interval forecasting model with three aims is developed to quantify the potential risks for the decision makers in EPS. The three aims of the fitness function are defined as follows: one is to maximize the forecast interval coverage probability; second is to minimize the forecast interval normalized average width; and third is to minimize the accumulated width deviation. (4) Apart from the abovementioned three differences, this research also provides a more comprehensive comparative study. The empirical mode decomposition (EMD) family is used to develop similar interval forecasting methods for comparative studies providing suggestions for future potential research directions of the EMD-based methods. Moreover, some insightful discussions are provided, such as the superiority of the developed system and a sensitivity analysis.

The remainder of the paper is organized as follows: Section 2 provides the required methods for the devised system. Section 3 constructs the proposed forecasting and analysis system. Section 4 presents the data and analyzes the results. Further discussions are conducted in Section 5. Finally, the conclusion is summarized in Section 6.

2. Methodology

This section provides the introduction of the required methods for the developed system, which include the variational mode decomposition (VMD), extreme learning machine (ELM), and multi-objective salp swarm algorithm (MSSA).

2.1 Variational Mode Decomposition (VMD)

VMD, developed by Dragomiretskiy and Zosso (2014), is a novel decomposition method that has been frequently used in many fields, such as biomedical image denoising (Lahmiri and Boukadoum, 2014), mechanical fault diagnosis (Huang et al., 2016), and seismic time–frequency analysis (Xue et al., 2016). Many previous studies have proven that the VMD method is superior to other decomposition approaches such

as wavelet transform and EMD for signal denoising (Lin et al., 2017). The details of VMD are presented as **Algorithm 1**.

Algorithm 1: VMD

```

1  /*Set the parameters of VMD. */
2  /*Initialize the  $\{\hat{u}_k^1\}, \{w_k^1\}, \hat{\lambda}^1$  and  $n$ . */
3  REPEAT THE ITERATION
4       $n = n + 1$ 
5      FOR  $k = 1 : k$  DO
6          /* Update  $\hat{u}_k$  for all  $w \geq 0$  */
7           $\hat{u}_k^{n+1}(w) = \frac{f(w) - \sum_{i < k} \hat{u}_i^{n+1}(w) - \sum_{i > k} \hat{u}_i^n(w) + \hat{\lambda}^n / 2}{1 + 2\alpha(w - w_k^n)^2}$ 
8          /* Update  $w_k$  */
9           $w_k^{n+1} = \int_0^\infty w |\hat{u}_k^{n+1}(w)|^2 dw / \int_0^\infty |\hat{u}_k^{n+1}(w)|^2 dw$ 
10         END FOR
11         /*Dual ascent for all  $w \geq 0$ . */
12          $\hat{\lambda}^{n+1}(w) = \hat{\lambda}^n(w) + \tau(\hat{f}(w) - \sum_k \hat{u}_k^{n+1}(w))$ 
13     END UNTIL MEET TERMINATION CONDITION
14     RETURN a series of band-limited modes

```

2.2 Extreme learning machine (ELM)

ELM, developed by Huang et al. (2004), is a type of single hidden layer feedforward neural network (Jiang et al., 2017), which has a simple structure, high forecasting accuracy, fast calculation speeds, and fewer training sample requirements. Compared with some well-known neural networks, it has a good prediction ability and real-time learning capability (Huang et al., 2012; Huang et al., 2015; Zhang et al., 2012).

The ELM predictor between x_i and y_i can be presented as follows:

$$f_L = \sum_{i=1}^L \beta_i G(a_i, b_i, x_j) = y_j \quad j = 1, 2, \dots, N \quad (1)$$

where a_i is the weight vector of hidden node i , b_i is polarization of node i , G is the excitation function, β denotes the weight vector of the input node and output node, L is the hidden layer number, and N is the number of training samples.

For simplicity, **Eq. (1)** is transformed as follows:

$$H\beta = Y \quad (2)$$

$$H(a_1, \dots, a_L, b_1, \dots, b_L, x_1, \dots, x_N) = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_L, b_L, x_1) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_N) & \cdots & G(a_L, b_L, x_N) \end{bmatrix}_{N \cdot L} \quad (3)$$

$$\begin{cases} \beta = [\beta_1^Y, \dots, \beta_L^Y]_{L \cdot m} \\ Y = [y_1^Y, \dots, y_N^Y]_{N \cdot m} \end{cases} \quad (4)$$

where H is the output matrix of the hidden layer, and the output weight can be obtained as follows:

$$\|H\beta - Y\| = \|HH^+Y - Y\| = \min_{\beta} \|H\beta - Y\| \quad (5)$$

$$\beta = H^+Y \quad (6)$$

where H^+ is the Moore-Penros generalized inverse matrix of the hidden layer.

2.3 Multi-objective salp swarm algorithm (MSSA)

The SSA, developed by Mirjalili et al. (2017), is a new bio-inspired optimization algorithm that is inspired by the swarming behavior of salps.

Similar to other swarm intelligence optimization algorithms, the salps' position is defined as a two-dimensional matrix x . Furthermore, the food source named F is defined as the target in the search space.

The leader's position is updated as follows:

$$x_j^1 = \begin{cases} F_j + c_1((ub_j - lb_j)c_2 + lb_j) & c_3 \geq 0 \\ F_j - c_1((ub_j - lb_j)c_2 + lb_j) & c_3 < 0 \end{cases} \quad (7)$$

where x_j^1 denotes the leader's position in the j th dimension, F_j is the food source's position in the j th dimension, ub_j and lb_j are the upper and lower bounds, respectively, c_1 , c_2 , and c_3 are random numbers. c_1 is used to balance exploration and exploitation and is calculated as follows:

$$c_1 = 2e^{-(4l/L)^2} \quad (8)$$

where L is the maximum number of iterations and l is the current iteration number.

The followers's position is updated according to the following equation:

$$x_j^i = \frac{1}{2}(x_j^i + x_j^{i-1}) \quad (9)$$

where $i \geq 2$ and x_j^i shows the i th follower salp's position in j th dimension.

Different from solving the single objective problem, one cannot compare different solutions by the arithmetic relational operators for the multi-objective optimization problem. For this aim, the definitions include the Pareto dominance, Pareto optimality, Pareto optimal set, and Pareto optimal front (Coello Coello, 2009).

Definition 1 Pareto dominance

Vector $\vec{x} = (x_1, x_2, \dots, x_k)$ dominates $\vec{y} = (y_1, y_2, \dots, y_k)$, denoted by $x \succ y$, if

$$\forall i \{1, 2, \dots, k\}, [f(x_i) \geq f(y_i)] \wedge [i \in 1, 2, \dots, k : f(x_i)] \quad (10)$$

Definition 2 Pareto optimality

$\vec{x} \in X$ is the Pareto optimal if

$$\nexists \vec{y} \in X | F(\vec{y}) \succ F(\vec{x}) \quad (11)$$

Definition 3 Pareto optimal set

The Pareto optimal set consists of all the non-dominated solutions:

$$P_s := \{x, y \in X | \exists F(y) \succ F(x)\} \quad (12)$$

Definition 4 Pareto optimal front

The Pareto optimal front consists of the Pareto optimal solutions:

$$P_f := \{F(x) | x \in P_s\} \quad (13)$$

The multi-objective version of the SSA is developed to solve multi-objective issues, and it is named multi-objective SSA (MSSA). Firstly, in a similar way with the archives in multi-objective particle swarm optimization (MOPSO), the repository of food sources is equipped for the SSA algorithm. Secondly, the food source is chosen from a set of non-dominated solutions with the least crowded neighborhood. Finally, the pseudo-code of the MSSA algorithm is presented in **Algorithm 2**.

Algorithm 2: MSSA

Parameters:

- L —maximum iterations
- n —number of salps
- F_i —fitness of the i -th salp
- $[lb_j, ub_j]$ —boundaries of the j -th dimension
- x_i —position of the i -th salp
- l —current iterations

```

1  /*Set the basic parameters of MSSA. */
2  /*Initialize the salp population  $x_i$  ( $i = 1, 2 \dots n$ ) randomly. */
3  WHILE ( $l < L$ ) DO
4      /*Calculate the objective values of each salp. */
5      /*Find the non-dominated salps. */
6      /*Update the repository in regard to the obtained non-dominated salps. */
7      IF the repository is full DO
8          /* Remove one repository resident using repository maintenance. */
9          /* Add the non-dominated salp to the repository. */
10     END IF
11     /* Choose a source of food from repository:  $F = \text{SelectFood}(\text{repository})$ . */
12     /*Update  $c_1$ .*/
13      $c_1 = 2e^{-(4l/L)^2}$ 
14     FOR EACH salp ( $X_i$ )
15         IF  $i == 1$  DO
16             /*Update the position of the leading salp. */
17              $x_j^1 = \begin{cases} F_j + c_1((ub_j - lb_j)c_2 + lb_j) & c_3 \geq 0 \\ F_j - c_1((ub_j - lb_j)c_2 + lb_j) & c_3 < 0 \end{cases}$ 
18         ELSE IF
19             /*Update the position of the follower salp. */
20              $x_j^i = \frac{1}{2}(x_j^i + x_j^{i-1})$ 
21         END IF
22     END FOR
23     /* Amend the salps based on the boundaries  $[lb_j, ub_j]$ . */
24      $l = l + 1$ 
25 END WHILE
26 RETURN respository
27 /* Obtain  $X^*$  based on the returned respository */

```

Algorithm 2 shows that the main procedure is as follows: the MSSA algorithm first initializes the salp population randomly, then computes the objective values of each salps, and then the non-dominated salps are obtained. If the repository is not full, the non-dominated solutions are added to it; on the contrary, if it is full, the repository

maintenance is performed to remove the solutions with crowded neighborhood. A food source is chosen from the non-dominated solutions in the repository with the least crowded neighborhood after updating the repository. Next, c_1 is updated by *Eq. (8)*. Then, the positions of the leading and follower salps are updated by *Eq. (7)* and *(9)*, respectively. It should be noted that the algorithm will amend the salps based on the boundaries to avoid that a salp goes outside of the boundaries. Finally, the MSSA is terminated when it satisfactorily meets the end condition.

3. Construction of the wind energy forecasting and analysis system

This section describes the development of the forecasting and analysis system, which includes system design and system evaluation.

3.1 System design

In this section, the overall structure of the system devised in this study is presented in **Fig. 1**, which consists of two modules: deterministic forecasting module and uncertainty analysis module.

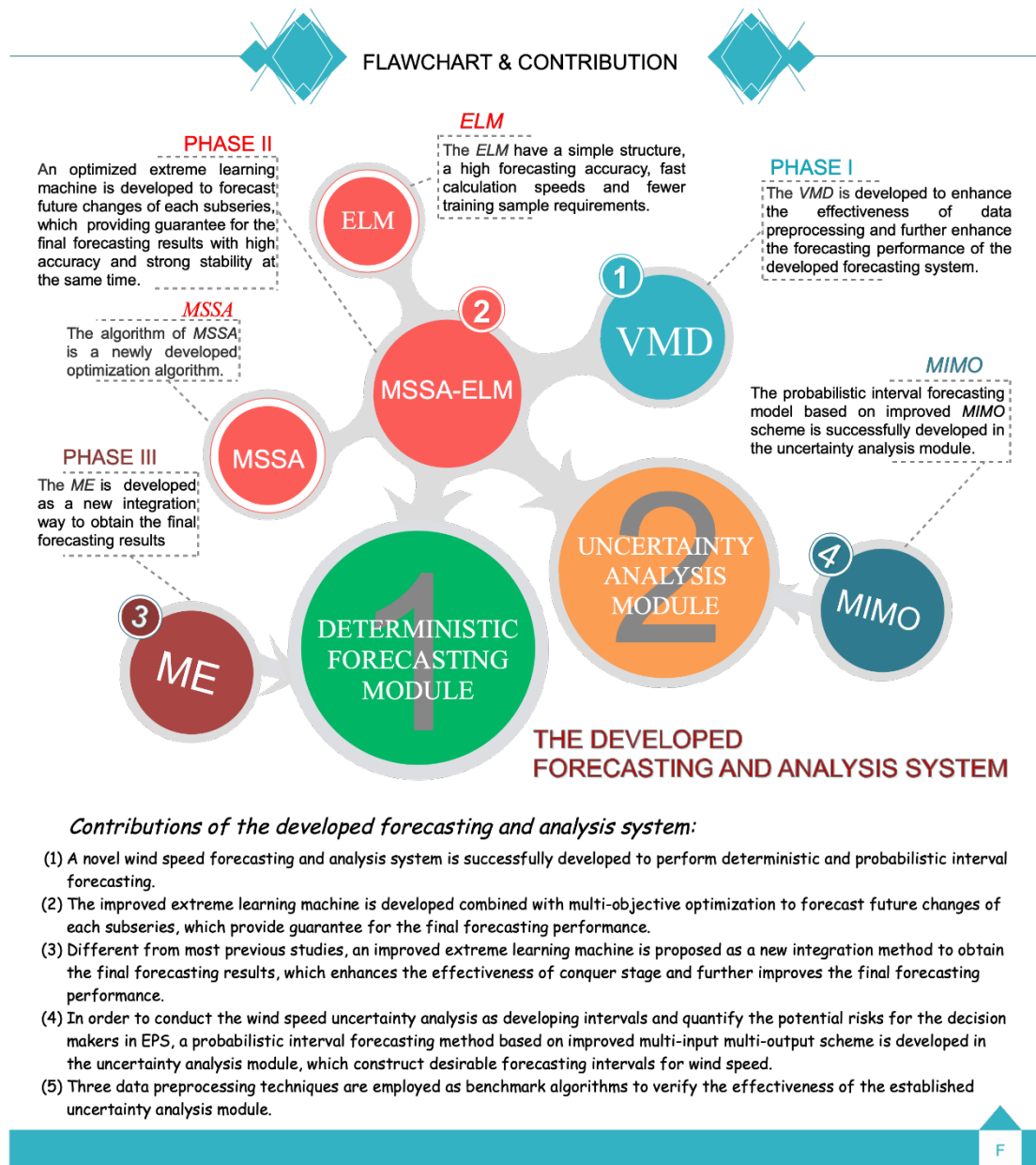


Fig. 1. Flowchart and contribution of the devised system

3.1.1 Deterministic forecasting module

The flowchart of the deterministic forecasting module of the devised system is shown in **Fig. 1**, which is composed of three phases, and the detailed information of each phase is as follows:

- ◆ **Phase I:** Owing to the wind energy data's inherent features of non-linearity and randomness, the utilization of wind energy is restricted by the wind energy forecasting performance. To weaken the noise's negative effects in the original wind energy data and further improve the final forecasting performance, data preprocessing is conducted to decompose the original data into a number of

subseries for capturing and mining the main feature of the wind energy time series, which has made important contributions to the system's forecasting performance.

- ◆ **Phase II:** As discussed above, the forecasting accuracy and stability are indispensable for an effective and robust forecasting model in engineering applications. For this aim, two objectives (accuracy and stability) are defined for deterministic forecasting. More specifically, the forecasting accuracy is represented by the mean squared error of the forecasting results, which is one of the most widely used fitness functions to improve the model's accuracy. Moreover, it is well known that the forecasting errors are one of the most important indicators for evaluating the performance of a model. Therefore, in this study, the stability of the forecasting errors is represented by the standard deviation of the forecasting errors, which is defined as another fitness function for improving the deterministic forecasting stability. Based on this, an optimized ELM is developed to forecast future changes of each subseries, which involves multi-objective optimization and ensures accuracy and stability of the final forecasting results. In particular, the forecasting is made using the previous data for one step ahead.
- ◆ **Phase III:** Most of the previous studies based on the divide and conquer strategy only have employed the widely used direct integration method to obtain the final forecasting results at the third phase, while few of them have considered other integration ways in the conquer phase. This offers potential for improvements in the final forecasting performance. Therefore, in this study, different from most of the previous studies, another optimized ELM is developed as a new integration method, instead of a direct integration method. In particular, the one step ahead forecasting results of each subseries are inputted to the new developed integration method to obtain the final one step ahead wind energy forecasting results, which enhances the effectiveness of the conquer stage and further improves the final forecasting performance.

3.1.2 Uncertainty analysis module

The framework of the uncertainty analysis module of the devised system is shown in **Fig. 1**, which is established based on the deterministic forecasting module to conduct the wind energy uncertainty analysis as developing intervals and quantify the potential risks for the decision makers in EPS. Moreover, to obtain the high-quality interval forecasting results, the multi-objective optimization algorithm is also incorporated in

the uncertainty analysis module. Therefore, the optimization objective, named the fitness function, is defined in this study.

The forecast interval coverage probability (FICP) and forecast interval normalized average width (FINAW) are two widely employed metrics for evaluating the interval forecasting quality. However, only considering the FICP and FINAW is not enough for some special conditions. In detail, if the actual value is outside the forecasted interval, and when the values of FICP and FINAW are approximated to the benchmark model, the effect of two forecasted intervals with different levels of upper and lower bounds of the actual target value deviating from the forecast interval cannot be reasonably evaluated. To solve this issue, a new metric, named the accumulated width deviation (AWD), is also employed to validate the performance of the uncertainty analysis module. Therefore, these three objectives should be considered simultaneously and applied in the development of the uncertainty analysis module. More specifically, different from the two objectives (accuracy and stability) of the deterministic forecasting process, another fitness function with three objectives is designed in the uncertainty analysis module. The first objective is to maximize the FICP (i.e., $\min 1 - \alpha - \text{FICP}$), the second is to minimize the FINAW (i.e., $\min \text{FINAW}$), and the third is to minimize the accumulated width deviation (i.e., $\min \text{AWD}$).

In this section, based on the multi-objective optimization algorithm with the abovementioned three aims, a probabilistic interval forecasting model based on the optimized multi-input multi-output ELM is successfully developed to forecast the future wind energy interval based on the forecasted future changes of each subseries. In particular, the one step ahead forecasting results of each subseries are inputted to the optimized multi-input multi-output ELM to obtain the final one step ahead wind energy interval forecasting results, which provide desirable forecasting intervals for wind energy.

3.2 System evaluation

This section introduces a comprehensive evaluation of the devised system, which comprises eight metrics for the deterministic forecasting and three metrics for the uncertainty analysis.

3.2.1 Deterministic forecasting evaluation

To test the performance of the deterministic forecasting module, the multiple error criteria (MEC), grey relational analysis (GRA), and Pearson's test are used from

different perspectives, and the details of each perspective are as follows.

◆ **Multiple error criteria**

Many error metrics have been widely employed in recent studies, but there is no universal standard for model evaluation. Therefore, researchers usually adopt multiple error criteria to compare the performance of the developed methods and of other compared models (Xu et al., 2017). By referring to the relevant literature in forecasting fields (Xiao et al., 2016a; Wang et al., 2018d; Xiao et al., 2016b), five common error criteria, including the mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), median absolute percentage error (MdAPE), Theil U statistic 1 (U1), and Theil U statistic 2 (U2), are adopted as multiple error criteria in this study and presented in **Table 1**. Smaller values for these metrics reveal a better forecasting performance (Jiang, 2018).

Table 1

Performance metric rules for the deterministic forecasting module

Metric	Definition	Equation
MAE	Mean absolute error of N forecasting results	$\mathbf{MAE} = \frac{1}{N} \sum_{i=1}^N F_i - A_i $
RMSE	Square root of average of the error squares	$\mathbf{RMSE} = \sqrt{\frac{1}{N} \times \sum_{i=1}^N (F_i - A_i)^2}$
MAPE	Average of N absolute percentage error	$\mathbf{MAPE} = \frac{1}{N} \sum_{i=1}^N \left \frac{A_i - F_i}{A_i} \right \times 100\%$
MdAPE	Median of N absolute percentage error	$\mathbf{MdAPE} = \text{median} \left(\left \frac{A_i - F_i}{A_i} \right \times 100\% \right)$
U1	Theil U statistic 1 of forecasting results	$\mathbf{U1} = \sqrt{\frac{1}{N} \times \sum_{i=1}^N (F_i - A_i)^2} / \left(\sqrt{\frac{1}{N} \times \sum_{i=1}^N A_i^2} + \sqrt{\frac{1}{N} \times \sum_{i=1}^N F_i^2} \right)$
U2	Theil U statistic 2 of forecasting results	$\mathbf{U2} = \sqrt{\frac{1}{N} \times \sum_{i=1}^N ((A_{i+1} - F_{i+1}) / A_i)^2} / \sqrt{\frac{1}{N} \times \sum_{i=1}^N ((A_{i+1} - F_i) / A_i)^2}$

◆ **Grey relational analysis (GRA)**

The GRA is considered to test whether the forecasted values' curve has a higher similarity to the observed values' curve. The main procedure is defined as follows

(Yang et al., 2019):

The reference series:

$$X_0 = (X_0(1), X_0(2), \dots, X_0(n)) \quad (14)$$

The comparison series:

$$X_i = (X_i(1), X_i(2), \dots, X_i(n)) \quad (15)$$

The series are standardized as follows:

$$x_i(t) = \frac{X_i(t) - \frac{1}{n} \sum_{t=1}^n X_i(t)}{\sqrt{\frac{1}{n-1} \sum_{t=1}^n (X_i(t) - \frac{1}{n} \sum_{t=1}^n X_i(t))^2}} \quad (16)$$

The correlation coefficient is calculated as follows:

$$\xi_i(k) = \frac{\min_i \min_k |x_0(k) - x_i(k)| + \rho \max_i \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \rho \max_i \max_k |x_0(k) - x_i(k)|}, \rho \in (0, \infty) \quad (17)$$

The grey relation degree (GRD) between x_0 and x_i is given as follows:

$$r_i = \frac{1}{n} \sum_{k=1}^n \xi_i(k) \quad (18)$$

If $r_a < r_b$, the series b has a higher similarity to the reference curve than series a .

◆ Pearson's test

Pearson's test, developed by Karl Pearson, can be used to present the association strength of the observed and forecasted values (Wang et al., 2018b). In this study, the evaluation of association strength is performed based on Pearson's test for further elaborating the superiority of the proposed deterministic forecasting module. Pearson's test result can be presented by the metric named Pearson's correlation coefficient (PCC). If $PCC = 1$, it means that there is a linear relationship between the observed and forecasted values and, on the contrary, if $PCC = 0$, there is no relationship between these two time series.

3.2.2 Uncertainty analysis evaluation

As discussed in Section 3.1.2, three metrics, including the forecast interval coverage probability (FICP), forecast interval normalized average width (FINAW), and accumulated width deviation (AWD), are studied simultaneously and implemented in the development of the uncertainty analysis module. Owing to the significance of these

three indexes for the quality of interval forecasting, these three metrics are also adopted for uncertainty analysis evaluation, and they are defined as follows:

Table 2

Performance metric rules for the uncertainty analysis module

Metric	Definition	Equation
FICP	Forecast interval coverage probability of testing dataset	$\text{FICP} = \frac{1}{N} \sum_{i=1}^N c_i \times 100\%$
FINAW	Forecast interval normalized average width of testing dataset	$\text{FINAW} = \frac{1}{NR} \sum_{i=1}^N (U_i - L_i) \times 100\%$
AWD_i	Accumulated width deviation of testing sample <i>i</i>	$\text{AWD}_i = \begin{cases} (L_i - A_i)/((U_i - L_i)), & A_i < L_i \\ 0, & A_i \in [L_i, U_i] \\ (L_i - A_i)/((U_i - L_i)), & A_i > U_i \end{cases}$
AWD	Accumulated width deviation of testing dataset	$\text{AWD} = \frac{1}{NR} \sum_{i=1}^N \text{AWD}_i$

Here, U_i and L_i , respectively, denote the upper and lower bounds of the forecasted interval; if the actual value $A_i \in [L_i, U_i]$, $c_i = 1$, otherwise $c_i = 0$; and N is the length of the testing dataset.

4. Data and results

To verify the performance of the devised forecasting and analysis system, wind energy datasets collected from a certain wind farm situated in Chengde, Hebei province of China, are used in this study.

4.1 Data description

The wind measuring data were collected by the meteorological observation tower. The current wind measuring devices generally record data at sampling intervals of 2 s to 3 s. Some parameters were considered during the collection process, such as time date, wind speed, wind direction, temperature, and pressure. The observation tower can automatically calculate and record the average wind speed every 10 min. In this study, wind energy data from three sites in the Chengde wind farm in Hebei province, China, with a 10 min period are randomly assembled to propose a wind energy forecasting and analysis system for wind farms. In total, 5040 data points covering 35 days are selected for the comparison study from each dataset, which are denoted as dataset A, dataset B, and dataset C in this study, and the first 4032 points covering 28 days are adopted as

the training set, and the other 1008 data covering 7 days are used as the testing set. The proportion of the training and testing datasets is 4:1 in this study. The details of the experimental data used in the case study are listed in **Table 3**. The standard deviations are all above 3 m/s, which indicate significant fluctuations of the wind energy data when taken the maximum/minimum of datasets A to C, which are 22.40/0.10 m/s, 22.50/0.10 m/s, and 23.40/0.10 m/s, respectively. It shows an evident fluctuation in frequency and amplitude between the series, indicating that these three series have different characteristics and the experiments based on them are appropriate for validation of the devised system.

Table 3

Statistical values of each data dataset used in this study

Data Set	Number	Statistic values				
		Maximum	Minimum	Median	Mean	Std.
Dataset A						
All Samples	5040	22.4000	0.1000	6.6000	7.4676	3.9102
Training set	4032	22.4000	0.1000	6.7000	7.6284	3.9150
Testing set	1008	19.3000	0.2000	5.9000	6.8241	3.8256
Dataset B						
All Samples	5040	22.5000	0.1000	6.5000	7.3006	3.8739
Training set	4032	22.5000	0.2000	6.7000	7.3818	3.7512
Testing set	1008	22.0000	0.1000	5.7000	6.9759	4.3168
Dataset C						
All Samples	5040	23.4000	0.1000	6.6000	7.2519	3.8179
Training set	4032	20.6000	0.2000	6.9000	7.3353	3.6348
Testing set	1008	23.4000	0.1000	5.8000	6.9182	4.4626

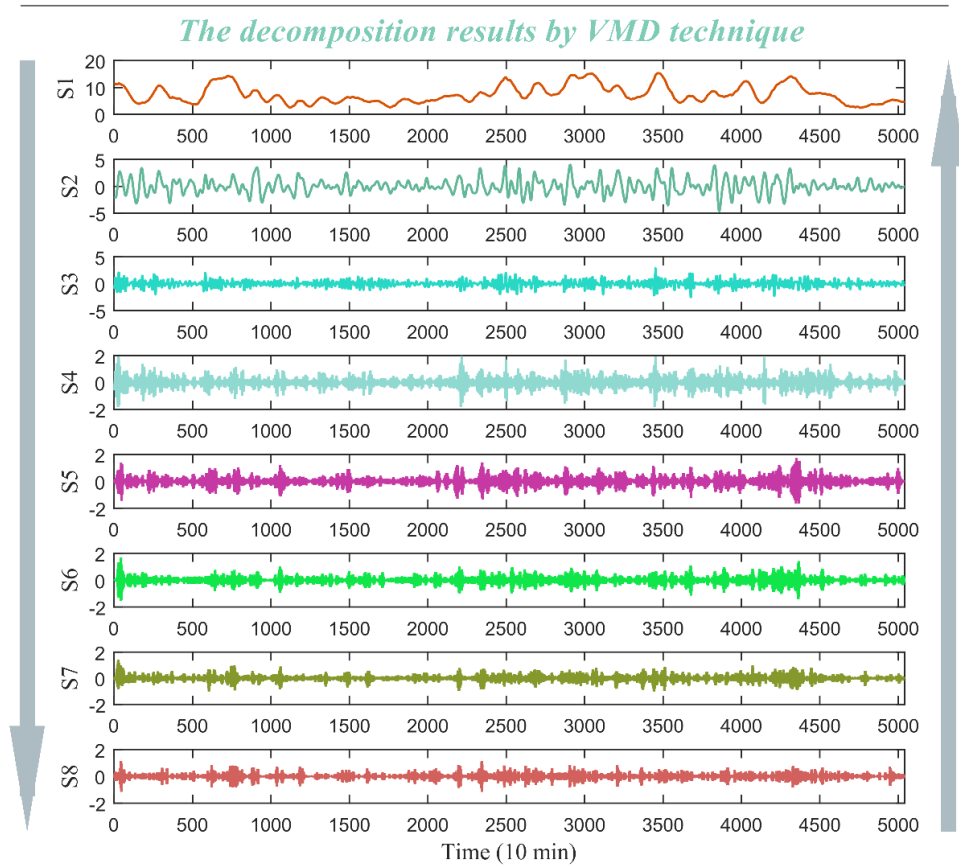


Fig. 2. Decomposition results by VMD technique for dataset A

4.2 Analysis of decomposition results by VMD technique

In phase I, according to the construction scheme of the developed wind energy forecasting and analysis system, to lower the negative influence of noise on wind energy forecasting and improve the final forecasting performance, the VMD technique is employed to decompose the original data into various subseries. Taking dataset A as an example, the decomposition results are presented in **Fig. 2**. As shown in **Fig. 2**, the original data, including the training and testing sets, are respectively decomposed into 8 subseries, which are named as S1, S2, ..., S7, S8. The VMD technique ensures that the main data characteristics underlying the time series are effectively mined and extracted. As mentioned above, after decomposition by the VMD technique, the two objectives (accuracy and stability) are defined for deterministic forecasting, which is combined with the multi-objective optimization algorithm MSSA to achieve these two specific aims. Based on this, an optimized ELM is developed based on multi-objective optimization with two aims (accuracy and stability) as a predictor, which ensures the final forecasting results' accuracy and stability. For the previous example shown in **Fig.**

2, each subseries can be properly modeled based on the developed MSSA–ELM model, and the forecasting results of each subseries are shown in **Fig. 3**. Besides the forecasting results of each subseries shown in **Fig. 3**, the fitted results of the corresponding training set can also be obtained by the proposed MSSA-ELM model. Based on the fitted and forecasting results of each subseries, the final deterministic forecasting can be obtained by phase III of the deterministic forecasting module. Meanwhile, the probabilistic interval forecasting results can also be obtained using the developed uncertainty analysis module. To verify the effectiveness of the developed system in both deterministic forecasting and interval forecasting, the two following validation studies are conducted in more detail.

4.3 Validity of the deterministic forecasting module

To verify the validity of the deterministic forecasting module of the devised system, two cases, denoted as case 1 and case 2, are designed in this section. The main motivation of case 1 is to verify the superiority of the components of the developed deterministic forecasting module, such as the ELM model, multi-objective optimization algorithm MSSA, and newly proposed integration method, and then the effectiveness of the hybridization modeling of these methods can be proved reasonably. Another motivation of case 1 is to prove the effectiveness and superiority of the developed deterministic module compared with the traditional forecasting model. Moreover, the data preprocessing technique plays a vital and indispensable role in wind energy forecasting, as it makes a great contribution to improve the forecasting results. Therefore, choosing a reasonable data preprocessing technique is a crucial step in the process of designing the wind energy forecasting model. To verify the reasonability and superiority of the VMD technique used in this study, case 2 is designed by comparing the VMD technique with other commonly used data preprocessing techniques. The detailed analysis for each one of the cases is as follows:

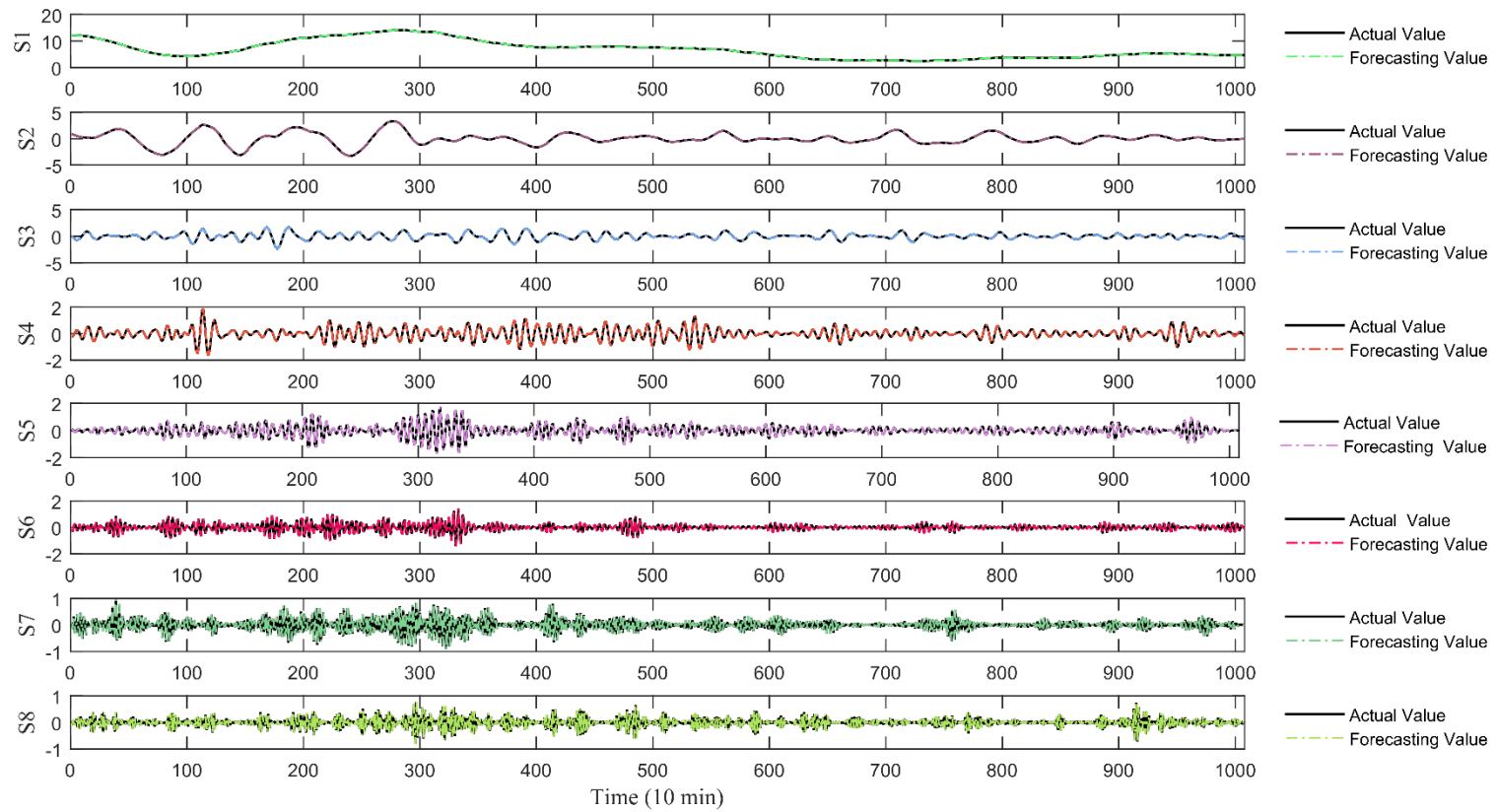


Fig. 3. Forecasting results of each subseries for dataset A

4.3.1 Case 1

In this case, all the experimental datasets, i.e., dataset A, dataset B, and dataset C, are used to evaluate the performance of the deterministic forecasting module. Moreover, the four following comparisons are established in this case, denoted as comparison I, comparison II, comparison III, and comparison IV. In comparison I, the GRNN, ENN, and ELM are used as the individual forecasting model's comparison for wind energy forecasting, which is aimed at obtaining the best individual forecasting model for developing an effective wind energy forecasting and analysis system. In comparison II, the comparison of the ELM-based model combined with different multi-objective optimization methods, i.e., MODA-ELM, MOGOA-ELM, and MSSA-ELM, is regarded as the study of the improved version of the model, which is used to emphasize the significance of the optimization for the improvement of the final forecasting accuracy and stability. In comparison III, the ELM and MSSA-ELM are employed as the compared methods to validate the performance of the deterministic forecasting module. Meanwhile, to further reveal the superiority of the deterministic forecasting module, two traditional models, i.e., the persistence model and ARIMA, are regarded as benchmark models in comparison IV.

Table 4

Results of the proposed deterministic forecasting module and other compared models in case 1

Data	Model	MAE	RMSE	MAPE	MdAPE	U1	U2	GRD	PCC
Site 1	ARIMA	0.8147	1.0891	16.4396	10.3502	0.0696	0.9427	0.7927	0.9590
Dataset A	Persistence model	0.7934	1.0921	15.1921	10.3195	0.0698	0.9781	0.7996	0.9593
	ENN	0.7766	1.0524	15.9522	9.9005	0.0674	0.9370	0.8011	0.9615
	GRNN	0.8403	1.1124	18.1713	10.6901	0.0713	0.8574	0.7895	0.9569
	ELM	0.7736	1.0505	15.6789	9.6814	0.0673	0.9573	0.8017	0.9616
	MODA-ELM	0.7708	1.0455	15.5108	9.5189	0.0670	0.9640	0.8012	0.9619
	MOGOA-ELM	0.7767	1.0524	15.5792	9.6934	0.0674	0.9460	0.7997	0.9614
	MSSA-ELM	0.7728	1.0486	15.4527	9.5657	0.0672	0.9698	0.8008	0.9617
	ME-VMD-MSSA-ELM	0.1000	0.1335	2.0195	1.2897	0.0085	0.0709	0.9647	0.9994
Site 2	ARIMA	0.7138	0.9409	14.0369	9.2273	0.0574	1.2640	0.8106	0.9762
Dataset B	Persistence model	0.7114	0.9483	14.6466	9.0854	0.0583	0.8560	0.8116	0.9757
	ENN	0.6931	0.9204	14.7438	8.9751	0.0564	0.7598	0.8156	0.9770
	GRNN	0.8018	1.0933	17.1262	9.9674	0.0678	0.7684	0.7986	0.9696
	ELM	0.6973	0.9229	14.5435	8.9272	0.0566	0.7895	0.8149	0.9771
	MODA-ELM	0.7013	0.9318	14.6705	8.9084	0.0572	0.7002	0.8137	0.9766
	MOGOA-ELM	0.7026	0.9364	14.4187	9.0608	0.0575	0.7405	0.8131	0.9765
	MSSA-ELM	0.7004	0.9287	14.2728	9.0610	0.0569	0.7436	0.8138	0.9767
	ME-VMD-MSSA-ELM	0.0895	0.1168	1.8574	1.1623	0.0071	0.0677	0.9683	0.9996
Site 3	ARIMA	0.7942	1.0697	15.9145	10.0000	0.0650	1.0961	0.8060	0.9712
Dataset C	Persistence model	0.7932	1.0524	16.8574	10.1129	0.0641	1.0956	0.8037	0.9718
	ENN	0.8091	1.0919	17.5684	10.5407	0.0674	1.0582	0.8053	0.9715
	GRNN	0.8822	1.2440	19.5654	11.0462	0.0771	1.0640	0.7931	0.9635
	ELM	0.7761	1.0410	17.2341	10.2149	0.0639	1.0448	0.8076	0.9730
	MODA-ELM	0.7757	1.0325	17.0915	10.1401	0.0632	1.0656	0.8096	0.9732
	MOGOA-ELM	0.7983	1.0814	17.0986	10.2821	0.0665	1.1058	0.8054	0.9713
	MSSA-ELM	0.7775	1.0453	16.9580	10.2305	0.0640	1.1538	0.8081	0.9726
	ME-VMD-MSSA-ELM	0.1026	0.1362	2.0656	1.3872	0.0083	0.0568	0.9657	0.9995

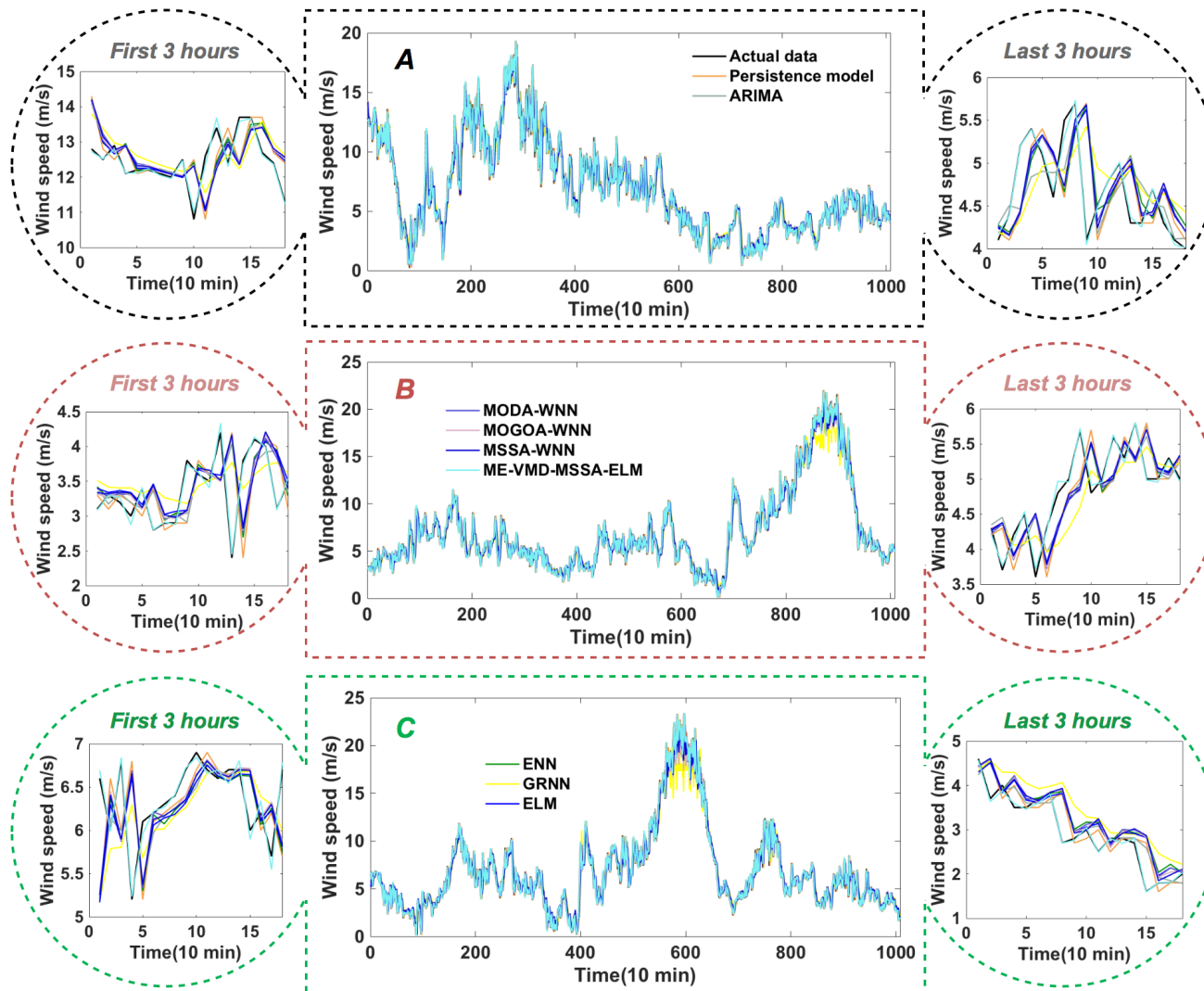


Fig. 4. Forecasting results of different models considered in case 1

In addition, the results of the developed deterministic forecasting module and other benchmark methods are presented in **Table 4**, where the bold values represent the best values of MAE, RMSE, MAPE, MdAPE, U1, U2, GRD, and PCC among all models. Furthermore, the forecasting results are also illustrated in **Fig. 4**. From **Table 4** and **Fig. 4**, it can be evidently observed that the proposed deterministic module is superior to other methods, which means that this module is more suitable for wind energy deterministic forecasting than others.

For wind energy deterministic forecasting in the Chengde wind farm, the results of case 1 are as follows:

- (a) Comparison I is designed to select the best individual forecasting model, i.e., ELM, among three commonly used models including the GRNN, ENN, and ELM. For example, the ELM model obtains the minimum value of MAPE for dataset A among all models, and the MAPE values of the GRNN, ENN, and ELM are 18.1713%, 15.9522%, and 15.6789%, respectively.
- (b) Comparison II is established to choose an effective and excellent multi-objective optimization algorithm for developing an optimized ELM model and further improving the forecasting performance. For dataset A, compared with the MODA-ELM and MOGOA-ELM, the MAPE values of the MSSA-ELM are decreased by 0.3746% and 0.8120%, respectively. The results show that the optimization of the ELM has a positive influence for wind energy forecasting, and the MSSA is superior to MODA and MOGOA in solving the optimization problem of the ELM or other complex non-linear issues.
- (c) In comparison III, the component models of the developed forecasting module, i.e., ELM and MSSA-ELM, are adopted as benchmark models for another comparison study. It can be found that the proposed forecasting module obtains better results than those obtained by the ELM and MSSA-ELM. By comparing the ELM and MSSA-ELM, we found that the MSSA is effective for improving the forecasting ability of ELM. Furthermore, the significance of optimization for wind energy forecasting has been proved. Moreover, by comparing the developed forecasting module with the MSSA-ELM, we can confirm the positive influence of the divide and conquer theory and the new integration method for improving the forecasting performance. Similarly, by comparing the proposed forecasting module with the ELM, it is possible to quantify clearly the contributions of the

VMD, MSSA, and integration way to the module's excellent performance. Thus, these techniques are able to effectively enhance the performance of the ELM, which can be widely employed in other areas.

- (d) In comparison IV, by comparing the devised forecasting module with two traditional models including the persistence model and ARIMA, it is obvious that the proposed model is superior to the typical traditional forecasting methods. For dataset A, the MAPE values of the developed module, persistence model, and ARIMA are 2.0195%, 15.1921%, and 16.4396%, respectively.

4.3.2 Case 2

This case aims to compare the performance of the VMD used in this study with other well-known decomposition techniques including EMD, EEMD, and CEEMD. For this aim, the deterministic forecasting modules based on different decomposition approaches, i.e., ME-EMD-MSSA-ELM, ME-EEMD-MSSA-ELM, and ME-CEEMD-MSSA-ELM, are considered as the compared models. Their only difference is the decomposition method used for data preprocessing in phase I. Therefore, through these comparisons we can not only test if the developed deterministic forecasting module is effective and superior, but also find the most effective and efficient way to decompose the original time series for enhancing the system's performance. Moreover, this case can be employed to illustrate the indispensability of the data preprocessing method for forecasting of complex time series.

Table 5

Results of the proposed deterministic forecasting module and other compared models in case 2

Data	Model	MAE	RMSE	MAPE	MdAPE	U1	U2	GRD	PCC
Site 1	ME-EMD-MSSA-ELM	0.4304	0.5590	9.1486	5.6283	0.0357	0.4388	0.8721	0.9893
Dataset A	ME-EEMD-MSSA-ELM	0.2237	0.3032	4.3507	2.8693	0.0195	0.2507	0.9301	0.9973
	ME-CEEMD-MSSA-ELM	0.1314	0.1812	2.5275	1.6722	0.0116	0.1545	0.9547	0.9989
	ME-VMD-MSSA-ELM	0.1000	0.1335	2.0195	1.2897	0.0085	0.0709	0.9647	0.9994
Site 2	ME-EMD-MSSA-ELM	0.4078	0.5193	8.2299	5.4140	0.0316	0.5517	0.8775	0.9930
Dataset B	ME-EEMD-MSSA-ELM	0.1779	0.2399	3.4370	2.2452	0.0146	0.2212	0.9401	0.9985
	ME-CEEMD-MSSA-ELM	0.1138	0.1513	2.1475	1.5181	0.0092	0.1566	0.9602	0.9994
	ME-VMD-MSSA-ELM	0.0895	0.1168	1.8574	1.1623	0.0071	0.0677	0.9683	0.9996
Site 3	ME-EMD-MSSA-ELM	0.4388	0.6069	9.9914	5.5011	0.0373	0.4627	0.8785	0.9917
Dataset C	ME-EEMD-MSSA-ELM	0.2079	0.2841	4.6130	2.5956	0.0173	0.1485	0.9361	0.9981
	ME-CEEMD-MSSA-ELM	0.1217	0.1697	2.7804	1.5071	0.0103	0.1127	0.9601	0.9993
	ME-VMD-MSSA-ELM	0.1026	0.1362	2.0656	1.3872	0.0083	0.0568	0.9657	0.9995

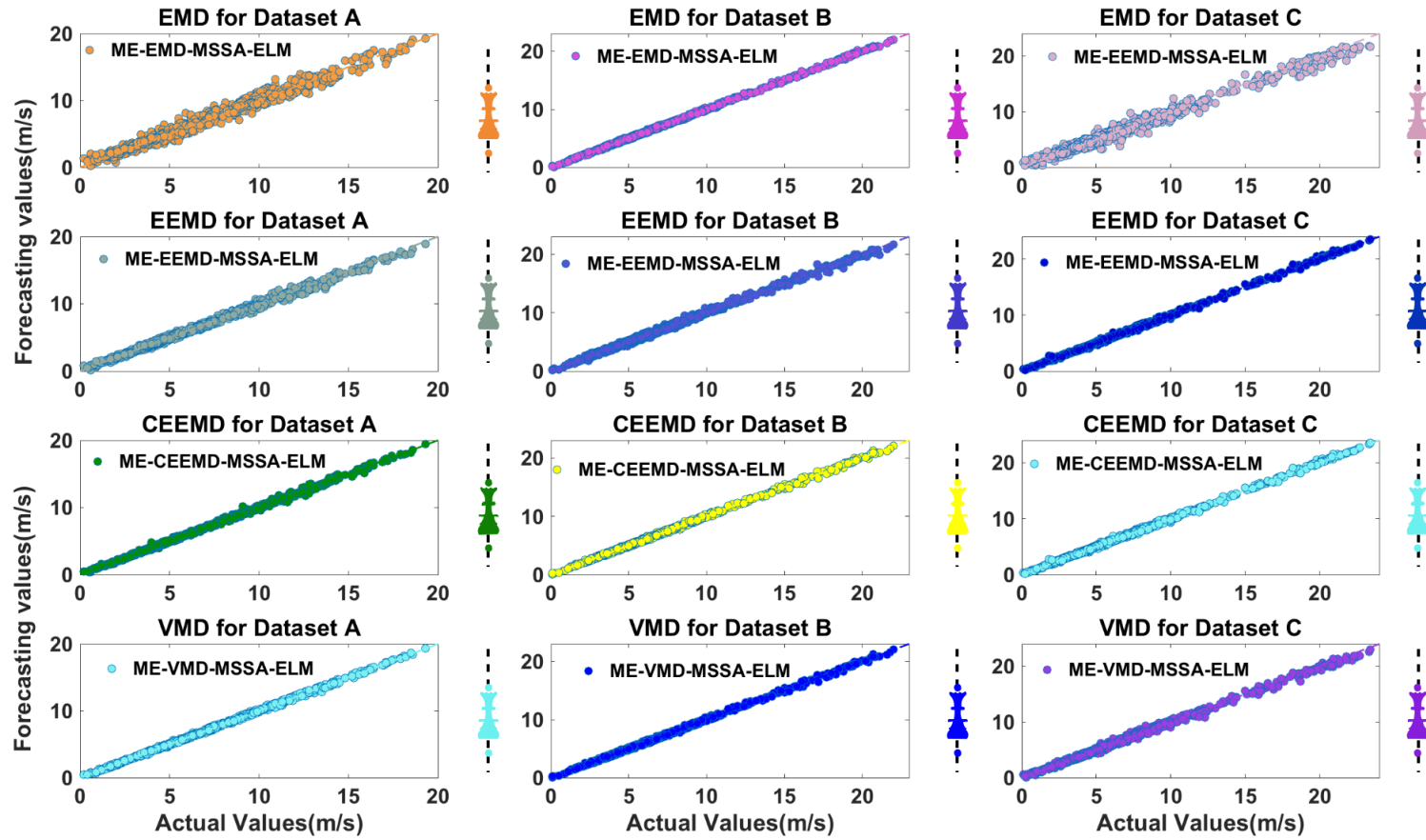


Fig. 5. Forecasting results of different models considered in Case 2

All the results obtained by the different data decomposition techniques considered in this study are presented in **Table 5**, where the bold values represent the best values of each metric. The results are also shown in **Fig. 5** for providing a clear comparison. As can be observed from **Table 5** and **Fig. 5**, the developed deterministic forecasting module, i.e., ME–VMD–MSSA–ELM, performs much better than the other three decomposition-based forecasting modules including ME–EMD–MSSA–ELM, ME–EEMD–MSSA–ELM, and ME–CEEMD–MSSA–ELM. For instance, the results show that the MAPE values of the ME–EMD–MSSA–ELM, ME–EEMD–MSSA–ELM, ME–CEEMD–MSSA–ELM, and the developed module for wind energy forecasting of dataset A are 9.1486%, 4.3507%, 2.5275%, and 2.0195%, respectively. And the ME–EMD–MSSA–ELM model presents the worst performance among all the benchmark models. The MAE errors of the developed module are decreased by 55.2973% compared with the ME–EMD–MSSA–ELM. Moreover, the forecasting accuracy and quality can be confirmed by U1 and U2, respectively. For example, the minimum values of U1 and U2 are obtained by the developed module for dataset A, which are equal to 0.0085 and 0.0709, respectively. This results also reveals that selecting a proper data preprocessing technique can considerably reduce the difficulty and uncertainty in wind energy forecasting and effectively improve the system’s forecasting performance. Moreover, the VMD technique is proposed to decompose the original wind energy data with features of randomness and non-linearity, as it can effectively address the forecasting difficulties through the proper decomposition approach.

Remark for the deterministic forecasting module

Case 1 and case 2 kept the focus on the validation of the devised deterministic forecasting module’s superiority, attempting to prove it from the perspectives of the basic individual forecasting model, optimization, integration method, traditional method, and decomposition. The results demonstrated that the proposed deterministic forecasting module outperforms all of the compared models considered in these two cases. The underlying reason is that the hybridization of data preprocessing, optimization, forecasting, and integration can effectively address the uncertainty, difficulty, and potential improvements in wind energy forecasting through the proper forecasting approach. Based on the excellent performance of the devised deterministic forecasting module, we believe that it is a promising alternative for wind energy forecasting.

4.4 Validity of the results of the uncertainty analysis module

Different from the deterministic wind energy forecasting, the probabilistic interval forecasting can provide more information and its results are conducive to risk analysis and assessment for the decision makers in the EPS, ensuring its reliability and controllability. In this section, the uncertainty analysis module was developed, which is named as MIMO–VMD–MSSA–ELM, to perform interval forecasting and was verified by three datasets with 10-min periods in different sites. To verify the superiority of the devised system for interval forecasting, three models based on different decompositions and the improved multi-input multi-output scheme, including the MIMO–EMD–MSSA–ELM, MIMO–EEMD–MSSA–ELM, and MIMO–CEEMD–MSSA–ELM, are developed as benchmark models to perform the same interval forecasting. We developed these three models mainly as comparison models because the deterministic forecasting modules based on these three different decomposition approaches perform better than the other compared models considered in the comparative case study of deterministic forecasting modules. Moreover, if the developed MIMO–VMD–MSSA–ELM module outperforms all these three competitive models, we can safely believe that the developed uncertainty analysis module is superior to other benchmark models, which fully proves the effectiveness and superiority of this module.

Table 6

Results of the proposed analysis module and other compared models

Site 1 Dataset A				Site 2 Dataset B			Site 3 Dataset C		
MIMO-EMD-MSSA-ELM									
Alpha	FICP	FINAW	AWD	PICP	PINAW	AWD	PICP	PINAW	AWD
0.1	58.4325	0.0733	0.0135	57.0437	0.0643	0.0081	51.7857	0.0554	0.0172
0.2	84.9206	0.1414	0.0034	85.9127	0.1249	0.0014	78.1746	0.1141	0.0030
MIMO-EEMD-MSSA-ELM									
Alpha	PICP	PINAW	AWD	PICP	PINAW	AWD	PICP	PINAW	AWD
0.1	62.2024	0.0735	0.0096	71.7262	0.0619	0.0053	58.3333	0.0563	0.0124
0.2	86.1111	0.1511	0.0017	92.2619	0.1286	0.0012	80.0595	0.1047	0.0102
MIMO-CEEMD-MSSA-ELM									
Alpha	PICP	PINAW	AWD	PICP	PINAW	AWD	PICP	PINAW	AWD
0.1	93.9484	0.0715	0.0020	94.8413	0.0638	0.0013	94.9405	0.0594	0.0085
0.2	98.1151	0.1430	0.0013	98.6111	0.1274	0.0011	97.6190	0.1185	0.0004
MIMO-VMD-MSSA-ELM									
Alpha	PICP	PINAW	AWD	PICP	PINAW	AWD	PICP	PINAW	AWD
0.1	95.4365	0.0712	0.0010	96.9246	0.0637	0.0007	95.8333	0.0589	0.0017
0.2	99.2063	0.1429	0.0001	99.3056	0.1264	0.0001	99.0079	0.1176	0.0001

Results for Dataset A

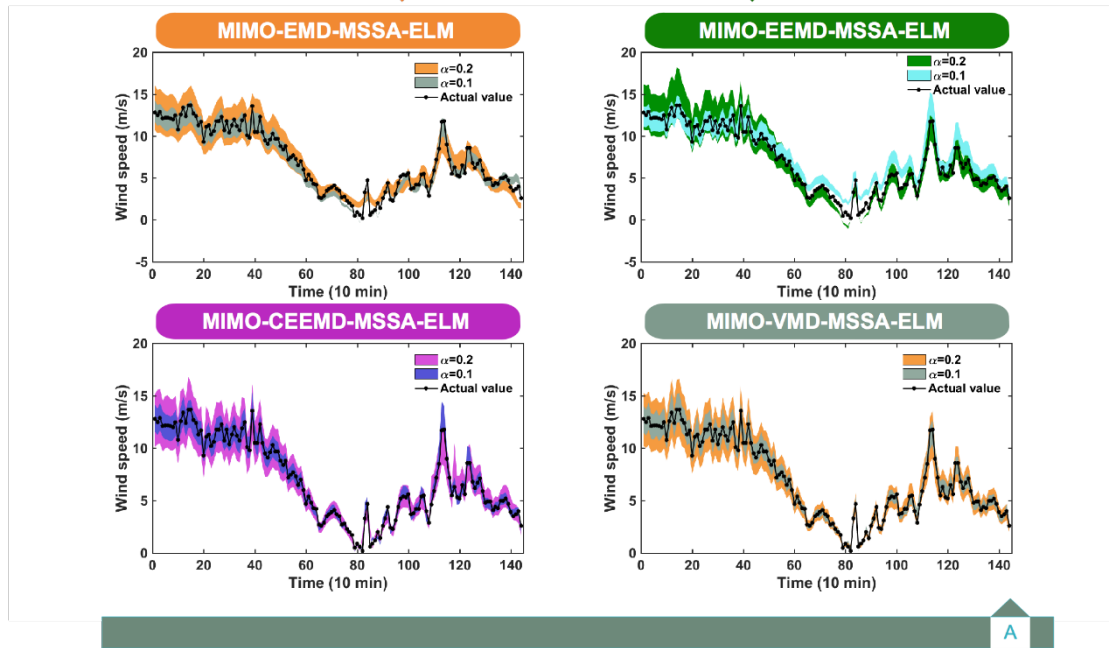


Fig. 6. Forecasting results of different models for dataset A

The results of the developed MIMO–VMD–MSSA–ELM module and the other three benchmark models are presented in **Table 6**, and they comprise the experiments based on three sites of the Chengde wind farm. Moreover, the expectation probability, i.e., $(1-a) \times 100\%$, is set to 90% and 80%, respectively, for evaluating the developed uncertainty analysis module. From **Table 6**, it can be observed that the proposed uncertainty analysis module obtains the best values of all metrics among all models, which demonstrates that the devised analysis module is superior to all benchmark methods. For example, as for 90% expectation probability, taking dataset A as an example, the FICP values offered by MIMO–EMD–MSSA–ELM, MIMO–EEMD–MSSA–ELM, MIMO–CEEMD–MSSA–ELM, and the developed uncertainty analysis module are 58.4325, 62.2024, 93.9484, and 95.4365, respectively, while the IFAW values are 0.0733, 0.0735, 0.0715, and 0.0712, respectively. In addition, the proposed analysis module leads to reductions of 0.0125, 0.0086, and 0.0010 in AWD when compared with the MIMO–EMD–MSSA–ELM, MIMO–EEMD–MSSA–ELM, and MIMO–CEEMD–MSSA–ELM models, respectively.

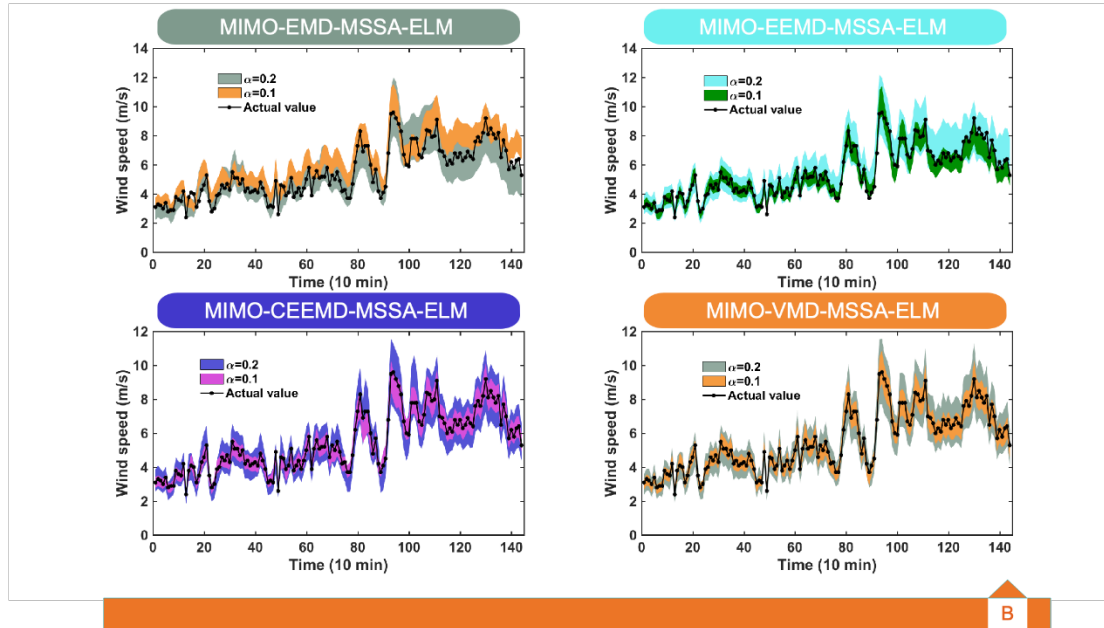


Fig. 7. Forecasting results of different models for dataset B

Meanwhile, for presenting the comparison clearly and intuitively, the results of the devised module and the other compared methods are shown in **Fig. 6–8**, which support the conclusions obtained from **Table 6** and provide intuitive evidence to verify the proposed system's superiority for wind energy uncertainty analysis. As shown in **Fig. 6–8**, the developed uncertainty analysis module obtains more effective interval forecasting results when compared with the other methods, and it is clear that its forecasted interval not only covers actual values of wind energy with a higher probability, but also becomes smoother than the other compared models, which means that the robustness of the developed uncertainty analysis module is more stable than that of the other models. Furthermore, the results reveal that the proposed uncertainty analysis module presents the same superiority for the three experiment datasets from different sites, which further demonstrates the effectiveness and robustness of the devised system for uncertainty analysis.

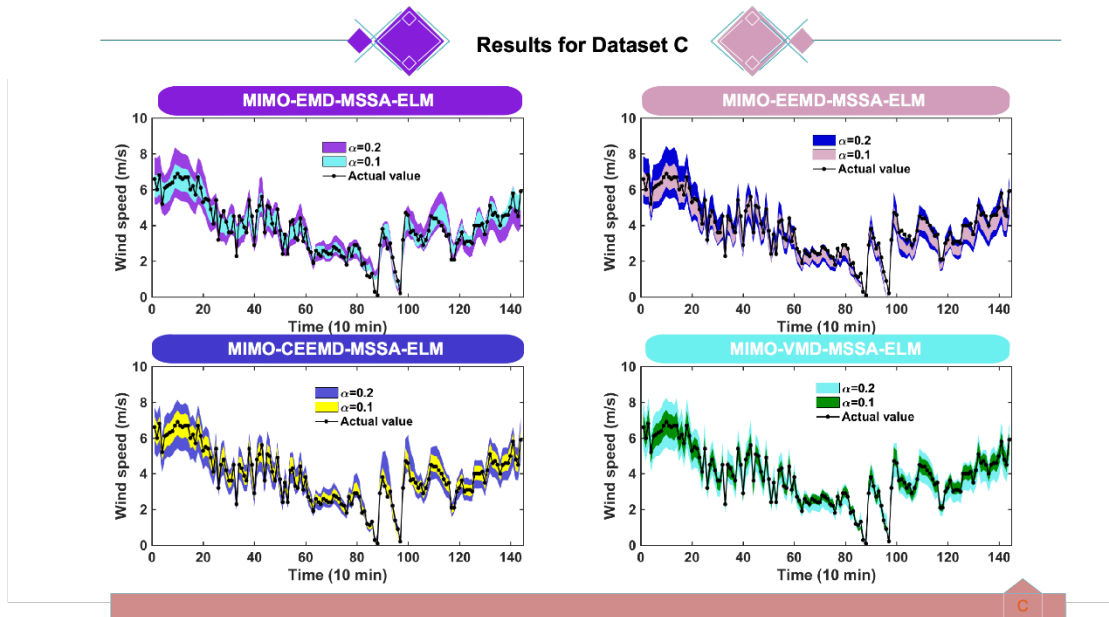


Fig. 8. Forecasting results of different models for dataset C

Remark for the uncertainty analysis module

The comparative studies are conducted based on data from different sites and three competitive models based on different decompositions and the improved multi-input multi-output scheme, and they show that the developed uncertainty analysis module outperforms all the compared models. Based on the excellent performance of the devised uncertainty analysis module, we believe that it is a promising alternative for wind energy probabilistic interval forecasting.

5. Discussion

To further verify the performance of the devised system and provide suggestions for future studies, some issues are discussed in this section.

5.1 Discussion of the proper integration way

Most of the prior studies only employed the simple direct integration (DI) method to obtain the final results based on the scheme of divide and conquer, and few of them considered other integration ways in the conquer phase, which offers the potential for improvements in future studies. Therefore, in this study, an optimized ELM is proposed as a new integration method for further improving the effectiveness of the conquer stage and final forecasting performance. For the aim of proving the superiority of the developed integration method and verifying the contribution of this study at the same time, the simple direct integration method is used as a benchmark model, denoted as DI-VMD-MSSA-ELM, to obtain the final forecasting results. The results, including eight metrics of the proposed deterministic forecasting module and of the other

compared models, are listed in **Table 7**. According to the comparative study presented in **Table 7**, we compare the MAE, RMSE, MAPE, MdAPE, U1, U2, GRD, and Pearson's test, and the developed module performs better in most tests. For example, compared with the DI-VMD-MSSA-ELM, the MAE errors of ME-VMD-MSSA-ELM are reduced by 10.9528%, 11.6549%, and 10.1576% for wind energy forecasting in datasets A, B, and C, respectively. The results show that the proper integration way plays an important role in wind energy forecasting based on the divide and conquer theory, and the newly proposed integration way has shown its significant advantage over the direct integration method in the conquer stage to forecast the final results based on the forecasted future changes of each subseries. More importantly, the developed deterministic forecasting module satisfactorily fits all data on the different sites, and thus, the module can be used as the forecasting part of the wind energy forecasting and analysis system.

Table 7

Results for the developed forecasting module and compared models

Model	MAE	RMSE	MAPE	MdAPE	U1	U2	GRD	PCC
Site 1 Dataset A								
DI-VMD-MSSA-ELM	0.1123	0.1489	2.2415	1.4267	0.0095	0.0884	0.9607	0.9992
ME-VMD-MSSA-ELM	0.1000	0.1335	2.0195	1.2897	0.0085	0.0709	0.9647	0.9994
Site 2 Dataset B								
DI-VMD-MSSA-ELM	0.1013	0.1312	2.1943	1.2967	0.0080	0.0882	0.9645	0.9995
ME-VMD-MSSA-ELM	0.0895	0.1168	1.8574	1.1623	0.0071	0.0677	0.9683	0.9996
Site 3 Dataset C								
DI-VMD-MSSA-ELM	0.1142	0.1517	2.3773	1.5061	0.0092	0.0637	0.9624	0.9994
ME-VMD-MSSA-ELM	0.1026	0.1362	2.0656	1.3872	0.0083	0.0568	0.9657	0.9995

5.2 Discussion of the system's statistical significance

Although the developed deterministic forecasting module has already been verified by eight evaluation metrics, the statistical significance needs to be tested from the statistical perspective using hypothesis testing. Therefore, an effective hypothesis testing method, i.e., the Diebold–Mariano (DM) test (Diebold and Mariano, 1995), is presented in this section based on the square error loss function. The DM test values are listed in **Table 8**. Furthermore, the smallest DM value is 4.2797, which is much bigger than $Z_{0.01/2} = 2.58$. Based on this, at the 1% significance level, we can reject the null hypothesis and consider that the devised deterministic forecasting module significantly performs better than the other benchmark models. Therefore, we believe that the proposed deterministic forecasting module not only shows excellent forecasting performance, but also presents a significant difference in forecasting performance level, which further verifies the devised system's superiority for wind energy forecasting.

Table 8

DM test results for the developed forecasting module and compared models

Model	Site 1 Dataset A	Site 2 Dataset B	Site 3 Dataset C
ARIMA	17.2217*	17.0170*	17.0194*
Persistence model	15.9187*	16.3703*	17.2768*
ENN	16.2131*	16.9979*	15.8565*
GRNN	17.2339*	14.9825*	12.9364*
ELM	16.2896*	17.1267*	16.6961*
MODA–ELM	16.0098*	16.8541*	16.9938*
MOGOA–ELM	16.1020*	16.5049*	15.9114*
MSSA–ELM	16.0519*	16.8759*	16.5526*
ME–EMD–MSSA–ELM	18.0716*	19.7921*	13.9840*
ME–EEMD–MSSA–ELM	13.2885*	11.4929*	11.2910*
ME–CEEMD–MSSA–ELM	5.9088*	6.4596*	4.5651*
DI–VMD–MSSA–ELM	5.8904*	7.0468*	4.2797*
ME–VMD–MSSA–ELM	-	-	-

* Indicates the 1% significance level.

5.3 Discussion of the system's forecasting stability

Mainly owing to the significance of achieving accuracy and stability at the same time

for engineering applications, the MSSA algorithm is employed to develop an effective system in this study and to predict future changes based on the achieved improvements in forecasting accuracy and stability. The abovementioned analysis is focused on the evaluation of the forecasting accuracy, which cannot prove the improvement of the proposed module in terms of stability. However, as discussed above, stability is one of the most important factors in the development of a model. Therefore, to perform a more comprehensive evaluation for the deterministic module, besides evaluating the performance from the forecasting accuracy perspective in Section 4, a discussion of the module's forecasting stability must be conducted in this study. It is well known that the forecasting performance can be measured by the forecasting error. Moreover, the variance can represent the stability of a time series (Tian and Hao, 2018). Therefore, in this study, the variance of the model's forecasting error is employed to measure its forecasting stability. Low variance values mean that the model performs better in terms of stability, while high variance values indicate that the model is more unstable than other models with low values (Hao and Tian, 2019b). The variances of the forecasting errors for different modes are listed in **Table 9**, which shows that the proposed deterministic forecasting module obtains the smallest variance values among all models in the different datasets. The results prove that the proposed deterministic forecasting module is more stable than the other compared models considered in this study and that it can perform wind energy forecasting with the desired stability in a real wind farm.

Table 9

Results of the variance of forecasting error for each model

Model	Site 1 Dataset A	Site 2 Dataset B	Site 3 Dataset C
ARIMA	1.1855	0.8862	1.1454
Persistence model	1.1937	0.8975	1.1087
ENN	1.1066	0.8479	1.1897
GRNN	1.2348	1.1878	1.5447
ELM	1.1032	0.8523	1.0838
MODA-ELM	1.0930	0.8688	1.0672
MOGOA-ELM	1.1070	0.8773	1.1692
MSSA-ELM	1.0995	0.8633	1.0937
ME-EMD-MSSA-ELM	0.3109	0.2629	0.3583

ME–EEMD–MSSA–ELM	0.0850	0.0576	0.0773
ME–CEEMD–MSSA–ELM	0.0328	0.0229	0.0288
DI–VMD–MSSA–ELM	0.0222	0.0172	0.0230
ME–VMD–MSSA–ELM	0.0178	0.0136	0.0185

5.4 Discussion of sensitivity analysis for the multi-objective optimization algorithm

Owing to the significance of multi-objective optimization in the developed system, more comments about the tuning of the control parameters must be presented. In this study, there is a key parameter, named search agents' number, which will affect the model's performance. Thus, a discussion of the sensitivity analysis of search agents' number in the multi-objective optimization algorithm is an important issue worth studying. In this section, the discussion on this important parameter is conducted based on dataset B, and the results measured by MAPE are shown in **Fig. 9**. It should be noted that only the search agents' number in phase III of the deterministic forecasting module is considered for the discussion of sensitivity analysis. The aim is to avoid the influence of other parameters, including other search agents' numbers, in phase I. Specifically, as shown in **Fig. 9**, the results prove that the search agent's number does have an impact on the model's forecasting. If the number is too large, it may result in poor forecasting performance and an increase in computation complexity, while if the number is too small, it might not achieve the best model's parameter and provide in poor forecasting results. Moreover, there will be one turning point, which can be set as the model's optimal parameter for wind energy forecasting. Thus, considering the model's performance and computation complexity, the search agents' number is defined as 40, which is the optimal parameter in this study.

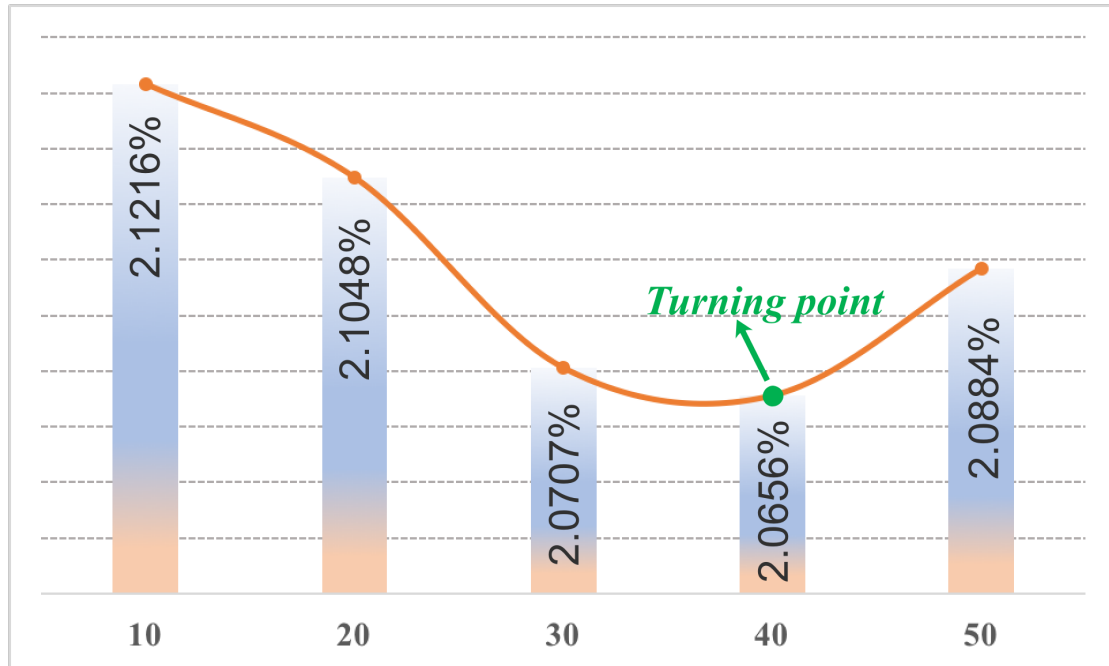


Fig. 9. Sensitivity analysis of different search agent numbers based on MSSA

5.5 Discussion of superiority of the multi-objective optimization algorithm

The comparison between different optimization-algorithm-based models has proven that the MSSA algorithm performs better than other two well-known algorithms. The theory named no free lunch (NFL) (Wolpert and Macready, 1997) **has** proven that no single algorithm can solve all the optimization problems. To achieve the better performance, the researchers should properly select the optimization algorithm for their studies. Moreover, the multi-objective optimization performance indices can serve as a reference standard. Therefore, in this study, multi-objective optimization performance indices, including the widely used metric named inverted generational distance (IGD) (Mirjalili et al., 2016), are presented and discussed, and they are employed to evaluate the optimization performance of the MSSA. In particular, IGD is the performance metric that shows the convergence of an algorithm and its statistical results can be used to measure the stability and robustness of the algorithm (Mirjalili et al., 2016; Mirjalili et al., 2017; Kusakci and Can, 2013). Four test functions including ZDT1, ZDT2, ZDT3, and ZDT1 with linear front are used. For each test function, thirty experiments are performed based on the following parameters: the iteration number equals 100, the search agents' number equals 40, and the archive size equals 100. The results are presented in **Table 10**. From **Table 10**, it can be observed that the MSSA obtain the best values of IGD for three test functions including ZDT1, ZDT2, and ZDT1 with linear front. A better performance is shown in both average and standard deviation, which

proves how good and robust the MSSA is when solving such problems. Therefore, the MSSA method is superior to MOGOA and MODA in terms of convergence, stability, and robustness, and it has the best potential ability for model's optimization. Moreover, that was the reason behind choosing a MSSA in the developed system.

Table 10

Statistic values of IGD for four test functions

ZDT1	Mean	Std.	Median	Best	Worse
MOGOA	0.015759	0.008718	0.013463	0.005627	0.044987
MODA	0.003673	0.001496	0.003468	0.002064	0.009450
MSSA	0.002284	0.000344	0.002243	0.001751	0.003129
ZDT2	Mean	Std.	Median	Best	Worse
MOGOA	0.015759	0.008718	0.013463	0.005627	0.044987
MODA	0.003673	0.001496	0.003468	0.002064	0.009450
MSSA	0.002284	0.000344	0.002243	0.001751	0.003129
ZDT3	Mean	Std.	Median	Best	Worse
MOGOA	0.023466	0.001199	0.023450	0.021290	0.026890
MODA	0.024921	0.000484	0.024821	0.024359	0.026596
MSSA	0.024799	0.000307	0.024793	0.024274	0.025535
ZDT1*	Mean	Std.	Median	Best	Worse
MOGOA	0.019879	0.012171	0.015506	0.007895	0.057791
MODA	0.003993	0.002180	0.003179	0.002013	0.011178
MSSA	0.002474	0.000355	0.002456	0.002001	0.003761

: ZDT1 is ZDT1 with linear front.

5.6 Discussion of the superiority and advancement of the developed system

To further prove the superiority of the developed system and emphasize the advancement of the paper compared with other models in the literature, a comparison between the devised system and other systems proposed in previous studies needs to be carried out. Therefore, in this section, three hybrid forecasting systems, named ICEEMDAN–MODA–ENN (Wang et al., 2018c), MCEEMD–MOSCA–WNN (Wang et al., 2018b), and CSV–MOSBO–ENN (Tian et al., 2018c), are adopted as benchmark models to further prove the superiority of the developed system for deterministic forecasting. The results of the proposed deterministic forecasting module and other benchmark models are presented in **Table 11**, where the bold values represent the smallest values of MAE, RMSE, MAPE, U1, and U2 and the highest values of GRD

and Pearson’s test among all models. From **Table 11**, it can be obviously observed that the proposed deterministic module performs better than the other compared models, which shows the advancement of this study with respect to the literature (Tian et al., 2018c; Wang et al., 2018b; Wang et al., 2018c) and further validates the superiority of the developed system for deterministic forecasting.

Table 11

Results of the developed deterministic forecasting module and other models

Data	Model	MAE	RMSE	MAPE	MdAPE	U1	U2	GRD	PCC
Site 1	ICEEMDAN–MODA–ENN	0.4132	0.5642	8.3086	5.2374	0.0361	0.3923	0.7857	0.9891
Dataset A	MCEEMD–MOSCA–WNN	0.1109	0.1591	2.4207	1.3765	0.0102	0.0817	0.9293	0.9992
	CSV–MOSBO–ENN	0.1850	0.2648	3.0776	2.2159	0.0170	0.0878	0.8858	0.9977
	ME–VMD–MSSA–ELM	0.1000	0.1335	2.0195	1.2897	0.0085	0.0709	0.9302	0.9994
Site 2	ICEEMDAN–MODA–ENN	0.3660	0.4988	7.5334	4.7055	0.0306	0.6786	0.8599	0.9937
Dataset B	MCEEMD–MOSCA–WNN	0.1030	0.1382	3.0295	1.2627	0.0084	0.4517	0.9533	0.9995
	CSV–MOSBO–ENN	0.1708	0.2422	3.9946	2.0818	0.0147	0.4702	0.9273	0.9986
	ME–VMD–MSSA–ELM	0.0895	0.1168	1.8574	1.1623	0.0071	0.0677	0.9589	0.9996
Site 3	ICEEMDAN–MODA–ENN	0.4039	0.6245	7.5831	4.6371	0.0377	0.2073	0.8529	0.9907
Dataset C	MCEEMD–MOSCA–WNN	0.1086	0.1507	3.1613	1.2783	0.0091	0.0711	0.9517	0.9994
	CSV–MOSBO–ENN	0.2132	0.3929	3.6605	2.1411	0.0241	0.1057	0.9029	0.9975
	ME–VMD–MSSA–ELM	0.1026	0.1362	2.0656	1.3872	0.0083	0.0568	0.9523	0.9995

6. Conclusion

An effective wind energy deterministic forecasting and uncertainty analysis provide more information for decision makers in the EPS, which can improve its management and scheduling. However, owing to the features of randomness and non-linearity associated with the wind energy, developing an effective wind energy forecasting and analysis system is not only a particularly challenging issue, but also a critical task for the whole society. The relevant research has deficiencies, despite its significance. Therefore, in this study, a novel wind energy forecasting and analysis system is successfully developed based on the idea of divide and conquer to compensate the insufficiency of existing studies and conduct an effective deterministic forecasting and uncertainty analysis. The experimental results illustrate that the devised system can produce more accurate and stable deterministic forecasting results and a higher quality

forecasted interval than other benchmark models. Thus, we can draw the conclusion that the devised system outperforms all the considered compared models in engineering applications.

The reasons why the devised system is superior to other benchmark models can be summarized as follows: (1) data preprocessing is conducted to decompose the original data into various subseries for capturing and mining the main features of wind energy data and eliminating the negative influence of noise, which has a positive effect on enhancing the system's performance; (2) an optimized ELM is developed based on multi-objective optimization to forecast future changes of each subseries, which ensures an effective forecasting performance with better accuracy and stability; (3) different from most of the previous studies, another optimized ELM is successfully proposed as a new integration method to obtain the final forecasting results, which further improves the effectiveness of the conquer stage and ultimately enhances the final forecasting performance; (4) the optimized multi-input multi-output ELM is developed with three aims, namely, to maximize the forecast interval coverage probability, to minimize the forecast interval normalized average width, and to minimize the accumulated width deviation, to forecast future wind energy intervals based on the future changes of each subseries, which construct high quality forecasting intervals for wind energy.

The results reveal that the devised system is an effective tool for wind energy deterministic forecasting and uncertainty analysis. More specifically, taking dataset A as an example, the comparative studies show that the developed system for deterministic forecasting has a lower MAPE value of 2.0195% compared with the MAPEs of 8.3086%, 2.4207%, and 3.0776% for the models in the previous literature, i.e., ICEEMDAN–MODA–ENN, MCEEMD–MOSCA–WNN, and CSV–MOSBO–ENN, respectively. This shows that the developed system leads to reductions of 6.2891%, 0.4012% and 1.0581% in MAPE when compared with these three models, respectively. Therefore, we can reasonably draw the conclusion that the devised system can also be widely employed in other areas of engineering applications in the future owing to its outstanding performance, which can reduce the financial and technical risk of uncertainty of wind power for all electricity market participants and finally bring enormous economic, social, and environmental benefits. Additionally, some implications based on the developed wind energy deterministic forecasting and

uncertainty analysis system from the perspective of the government and practitioners can be stated clearly as follows. (1) The government can establish and improve systematic and effective wind power policies and promote wind power to participate in market competition so as to overcome the economic obstacles to the acceptance of wind power by the power grid. Moreover, the governments can increase the support for relevant scientific research investment and promote the advancement of the wind energy deterministic forecasting and uncertainty analysis system. (2) The related practitioners can apply the effective wind energy forecasting to improve the performance of wind turbines, power grid scheduling ability, and other technologies, thus enabling the EPS to accept more wind power and breaking the technical constraints of the power grid to accept wind power capacity. The related practitioners can take practical measures to realize an effective allocation of power system resources and improve the economic benefit under the condition of ensuring the stability of the power market.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 71671029) and the China Scholarship Council (CSC) (File No. 201808210272).

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

- Bludszuweit, H., Domínguez-Navarro, J.A., Llombart, A., 2008. Statistical analysis of wind power forecast error. *IEEE Trans. Power Syst.* <https://doi.org/10.1109/TPWRS.2008.922526>
- Coello Coello, C.A., 2009. Evolutionary multi-objective optimization: Some current research trends and topics that remain to be explored. *Front. Comput. Sci. China.* <https://doi.org/10.1007/s11704-009-0005-7>
- Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. *J. Bus. Econ. Stat.* <https://doi.org/10.1080/07350015.1995.10524599>
- Dragomiretskiy, K., Zosso, D., 2014. Variational mode decomposition. *IEEE Trans. Signal Process.* <https://doi.org/10.1109/TSP.2013.2288675>
- Du, P., Wang, J., Yang, W., Niu, T., 2018. Multi-step ahead forecasting in electrical

power system using a hybrid forecasting system. *Renew. Energy*.
<https://doi.org/10.1016/j.renene.2018.01.113>

Errouissi, R., Cardenas-Barrera, J., Meng, J., Castillo-Guerra, E., Gong, X., Chang, L., 2015. Bootstrap prediction interval estimation for wind speed forecasting, in: 2015 IEEE Energy Conversion Congress and Exposition, ECCE 2015.
<https://doi.org/10.1109/ECCE.2015.7309931>

Foley, A.M., Leahy, P.G., Marvuglia, A., McKeogh, E.J., 2012. Current methods and advances in forecasting of wind power generation. *Renew. Energy*.
<https://doi.org/10.1016/j.neucom.2015.11.010>

Global Wind Energy Council, 2018. Global Wind Report 2017.
<http://files.gwec.net/register/success/?file=/files/GWR2017.pdf>

Hao, Y., Tian, C., 2019a. A novel two-stage forecasting model based on error factor and ensemble method for multi-step wind power forecasting. *Appl. Energy*.
<https://doi.org/10.1016/j.apenergy.2018.11.012>.

Hao, Y., Tian, C., 2019b. The study and application of a novel hybrid system for air quality early-warning. *Appl. Soft Comput. J.*
<https://doi.org/10.1016/j.asoc.2018.09.005>

Hong, Y.Y., Chang, H.L., Chiu, C.S., 2010. Hour-ahead wind power and speed forecasting using simultaneous perturbation stochastic approximation (SPSA) algorithm and neural network with fuzzy inputs. *Energy*.
<https://doi.org/10.1016/j.energy.2010.05.041>

Hu, J., Wang, J., Ma, K., 2015. A hybrid technique for short-term wind speed prediction. *Energy*. <https://doi.org/10.1016/j.energy.2014.12.074>

Huang, G. Bin, Zhu, Q.Y., Siew, C.K., 2004. Extreme learning machine: A new learning scheme of feedforward neural networks, in: IEEE International Conference on Neural Networks - Conference Proceedings. <https://doi.org/10.1109/IJCNN.2004.1380068>

Huang, G. Bin, Zhou, H., Ding, X., Zhang, R., 2012. Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. Man, Cybern. Part B Cybern.*
<https://doi.org/10.1109/TSMCB.2011.2168604>

Huang, G. Bin, Bai, Z., Kasun, L.L.C., Vong, C.M., 2015. Local receptive fields based extreme learning machine. *IEEE Comput. Intell. Mag.*
<https://doi.org/10.1109/MCI.2015.2405316>

Huang, N., Chen, H., Cai, G., Fang, L., Wang, Y., 2016. Mechanical fault diagnosis of

high voltage circuit breakers based on variational mode decomposition and multi-layer classifier. *Sensors (Switzerland)*. <https://doi.org/10.3390/s16111887>

Jiang, P., Ma, X., 2016. A hybrid forecasting approach applied in the electrical power system based on data preprocessing, optimization and artificial intelligence algorithms. *Appl. Math. Model.* <https://doi.org/10.1016/j.apm.2016.08.001>

Jiang, P., Liu, F., Song, Y., 2017. A hybrid forecasting model based on date-framework strategy and improved feature selection technology for short-term load forecasting. *Energy*. <https://doi.org/10.1016/j.energy.2016.11.034>

Jiang, H., 2018. Model forecasting based on two-stage feature selection procedure using orthogonal greedy algorithm. *Appl. Soft Comput. J.* <https://doi.org/10.1016/j.asoc.2017.11.047>

Jiang P, Yang H, Heng J., 2019. A hybrid forecasting system based on fuzzy time series and multi-objective optimization for wind speed forecasting. *Appl. Energy*. <https://doi.org/10.1016/j.apenergy.2018.11.012>.

Juban, J., Siebert, N., Kariniotakis, G.N., 2007. Probabilistic short-term wind power forecasting for the optimal management of wind generation, in: 2007 IEEE Lausanne POWERTECH, Proceedings. <https://doi.org/10.1109/PCT.2007.4538398>

Khosravi, A., Nahavandi, S., Creighton, D., Atiya, A.F., 2011. Lower upper bound estimation method for construction of neural network-based prediction intervals. *IEEE Trans. Neural Networks*. <https://doi.org/10.1109/TNN.2010.2096824>

Kusakci, A.O., Can, M., 2013. Performance Evaluation of Nature-Inspired Algorithms in constrained Optimization. *Southeast Eur. J. Soft Comput.* <https://doi.org/10.21533/scjournal.v2i1.50>

Lahmiri, S., Boukadoum, M., 2014. Biomedical image denoising using variational mode decomposition, in: IEEE 2014 Biomedical Circuits and Systems Conference, BioCAS 2014 - Proceedings. <https://doi.org/10.1109/BioCAS.2014.6981732>

Liang, Z., Liang, J., Wang, C., Dong, X., Miao, X., 2016. Short-term wind power combined forecasting based on error forecast correction. *Energy Convers. Manag.* <https://doi.org/10.1016/j.enconman.2016.04.036>

Li, C., Lin, S., Xu, F., Liu, D., Liu, J., 2018a. Short-term wind power prediction based on data mining technology and improved support vector machine method: A case study in Northwest China. *J. Clean. Prod.* <https://doi.org/10.1016/j.jclepro.2018.09.143>

Li, H., Wang, J., Lu, H., Guo, Z., 2018b. Research and application of a combined model

based on variable weight for short term wind speed forecasting. *Renew. Energy*.
<https://doi.org/10.1016/j.renene.2017.09.089>

Lin, Y., Luo, H., Wang, D., Guo, H., Zhu, K., 2017. An ensemble model based on machine learning methods and data preprocessing for short-term electric load forecasting. *Energies*. <https://doi.org/10.3390/en10081186>

Ma, L., Luan, S., Jiang, C., Liu, H., Zhang, Y., 2009. A review on the forecasting of wind speed and generated power. *Renew. Sustain. Energy Rev.*
<https://doi.org/10.1016/j.rser.2008.02.002>

Ma, X., Wang, Y., Wang, C., 2017a. Low-carbon development of China's thermal power industry based on an international comparison: Review, analysis and forecast. *Renew. Sustain. Energy Rev.* <https://doi.org/10.1016/j.rser.2017.05.102>

Ma, X., Jin, Y., Dong, Q., 2017b. A generalized dynamic fuzzy neural network based on singular spectrum analysis optimized by brain storm optimization for short-term wind speed forecasting. *Appl. Soft Comput. J.*
<https://doi.org/10.1016/j.asoc.2017.01.033>

Ma, X., Liu, Z., Wang, Y., 2019. Application of a novel nonlinear multivariate grey Bernoulli model to predict the tourist income of China. *J. Comput. Appl. Math.*
<https://doi.org/10.1016/j.cam.2018.07.044>

Meng, A., Ge, J., Yin, H., Chen, S., 2016. Wind speed forecasting based on wavelet packet decomposition and artificial neural networks trained by crisscross optimization algorithm. *Energy Convers. Manag.* <https://doi.org/10.1016/j.enconman.2016.02.013>

Mirjalili, S., Saremi, S., Mirjalili, S.M., Coelho, L.D.S., 2016. Multi-objective grey wolf optimizer: A novel algorithm for multi-criterion optimization. *Expert Syst. Appl.*
<https://doi.org/10.1016/j.eswa.2015.10.039>

Mirjalili, S., Gandomi, A.H., Mirjalili, S.Z., Saremi, S., Faris, H., Mirjalili, S.M., 2017. Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems. *Adv. Eng. Softw.* <https://doi.org/10.1016/j.advengsoft.2017.07.002>

Moghram, I., Rahman, S., 1989. Analysis and evaluation of five short-term load forecasting techniques. *IEEE Trans. Power Syst.* <https://doi.org/10.1109/59.41700>

Nielsen, H.A., Madsen, H., Nielsen, T.S., 2006. Using quantile regression to extend an existing wind power forecasting system With probabilistic forecasts, in: *Wind Energy*.
<https://doi.org/10.1002/we.180>

Niu, T., Wang, J., Zhang, K., Du, P., 2018. Multi-step-ahead wind speed forecasting

based on optimal feature selection and a modified bat algorithm with the cognition strategy. *Renew. Energy*. <https://doi.org/10.1016/j.renene.2017.10.075>

Sfetsos, A., 2002. A novel approach for the forecasting of mean hourly wind speed time series. *Renew. Energy*. [https://doi.org/10.1016/S0960-1481\(01\)00193-8](https://doi.org/10.1016/S0960-1481(01)00193-8)

Su, Z., Wang, J., Lu, H., Zhao, G., 2014. A new hybrid model optimized by an intelligent optimization algorithm for wind speed forecasting. *Energy Convers. Manag.* <https://doi.org/10.1016/j.enconman.2014.05.058>

Sun, W., Liu, M., 2016. Wind speed forecasting using FEEMD echo state networks with RELM in Hebei, China. *Energy Convers. Manag.* <https://doi.org/10.1016/j.enconman.2016.02.022>

Tian, C., Hao, Y., Hu, J., 2018. A novel wind speed forecasting system based on hybrid data preprocessing and multi-objective optimization. *Appl. Energy*. <https://doi.org/10.1016/j.apenergy.2018.09.012>

Tian, C., Hao, Y., 2018. A novel nonlinear combined forecasting system for short-term load forecasting. *Energies*. <https://doi.org/10.3390/en11040712>

Torres, J.L., García, A., De Blas, M., De Francisco, A., 2005. Forecast of hourly average wind speed with ARMA models in Navarre (Spain). *Sol. Energy*. <https://doi.org/10.1016/j.solener.2004.09.013>

Wang, J., Shahidehpour, M., Li, Z., 2008. Security-constrained unit commitment with volatile wind power generation. *IEEE Trans. Power Syst.* <https://doi.org/10.1109/TPWRS.2008.926719>

Wang, H.Z., Wang, G.B., Li, G.Q., Peng, J.C., Liu, Y.T., 2016. Deep belief network based deterministic and probabilistic wind speed forecasting approach. *Appl. Energy*. <https://doi.org/10.1016/j.apenergy.2016.08.108>

Wang, J., Du, P., Niu, T., Yang, W., 2017. A novel hybrid system based on a new proposed algorithm—Multi-Objective Whale Optimization Algorithm for wind speed forecasting. *Appl. Energy*. <https://doi.org/10.1016/j.apenergy.2017.10.031>

Wang, J., Niu, T., Lu, H., Guo, Z., Yang, W., Du, P., 2018a. An analysis-forecast system for uncertainty modeling of wind speed: A case study of large-scale wind farms. *Appl. Energy*. <https://doi.org/10.1016/j.apenergy.2017.11.071>

Wang, J., Yang, W., Du, P., Niu, T., 2018b. A novel hybrid forecasting system of wind speed based on a newly developed multi-objective sine cosine algorithm. *Energy Convers. Manag.* <https://doi.org/10.1016/j.enconman.2018.02.012>

- Wang, J., Yang, W., Du, P., Li, Y., 2018c. Research and application of a hybrid forecasting framework based on multi-objective optimization for electrical power system. *Energy*. <https://doi.org/10.1016/j.energy.2018.01.112>
- Wang, J., Du, P., Lu, H., Yang, W., Niu, T., 2018d. An improved grey model optimized by multi-objective ant lion optimization algorithm for annual electricity consumption forecasting. *Appl. Soft Comput. J.* <https://doi.org/10.1016/j.asoc.2018.07.022>
- Wang, Z.X., Li, Q., 2019. Modelling the nonlinear relationship between CO2 emissions and economic growth using a PSO algorithm-based grey Verhulst model. *J. Clean. Prod.* <https://doi.org/10.1016/j.jclepro.2018.10.010>
- Wolpert, D.H., Macready, W.G., 1997. No free lunch theorems for optimization. *IEEE Trans. Evol. Comput.* <https://doi.org/10.1109/4235.585893>
- Xiao, L., Shao, W., Wang, C., Zhang, K., Lu, H., 2016a. Research and application of a hybrid model based on multi-objective optimization for electrical load forecasting. *Appl. Energy*. <https://doi.org/10.1016/j.apenergy.2016.07.113>
- Xiao, L., Shao, W., Liang, T., Wang, C., 2016b. A combined model based on multiple seasonal patterns and modified firefly algorithm for electrical load forecasting. *Appl. Energy*. <https://doi.org/10.1016/j.apenergy.2016.01.050>
- Xiao, L., Shao, W., Yu, M., Ma, J., Jin, C., 2017. Research and application of a hybrid wavelet neural network model with the improved cuckoo search algorithm for electrical power system forecasting. *Appl. Energy*. <https://doi.org/10.1016/j.apenergy.2017.04.039>
- Xu, Y., Yang, W., Wang, J., 2017. Air quality early-warning system for cities in China. *Atmos. Environ.* <https://doi.org/10.1016/j.atmosenv.2016.10.046>
- Xue, Y.J., Cao, J.X., Wang, D.X., Du, H.K., Yao, Y., 2016. Application of the Variational-Mode Decomposition for Seismic Time-frequency Analysis. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* <https://doi.org/10.1109/JSTARS.2016.2529702>
- Yang, W., Wang, J., Wang, R., 2017. Research and Application of a Novel Hybrid Model Based on Data Selection and Artificial Intelligence Algorithm for Short Term Load Forecasting. *Entropy*. <https://doi.org/10.3390/e19020052>
- Yang, W., Wang, J., Niu, T., Du, P., 2019. A hybrid forecasting system based on a dual decomposition strategy and multi-objective optimization for electricity price forecasting. *Appl. Energy*. <https://doi.org/10.1016/j.apenergy.2018.11.034>
- Zendehboudi, A., Baseer, M.A., Saidur, R., 2018. Application of support vector

machine models for forecasting solar and wind energy resources: A review. *J. Clean. Prod.* <https://doi.org/10.1016/j.jclepro.2018.07.164>

Zhang, R., Lan, Y., Huang, G. Bin, Xu, Z. Ben, 2012. Universal approximation of extreme learning machine with adaptive growth of hidden nodes. *IEEE Trans. Neural Networks Learn. Syst.* <https://doi.org/10.1109/TNNLS.2011.2178124>

Zhang, W., Qu, Z., Zhang, K., Mao, W., Ma, Y., Fan, X., 2017. A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting. *Energy Convers. Manag.* <https://doi.org/10.1016/j.enconman.2017.01.022>

Zhao, Y., Ye, L., Li, Z., Song, X., Lang, Y., Su, J., 2016a. A novel bidirectional mechanism based on time series model for wind power forecasting. *Appl. Energy.* <https://doi.org/10.1016/j.apenergy.2016.03.096>

Zhao, J., Guo, Z.H., Su, Z.Y., Zhao, Z.Y., Xiao, X., Liu, F., 2016b. An improved multi-step forecasting model based on WRF ensembles and creative fuzzy systems for wind speed. *Appl. Energy.* <https://doi.org/10.1016/j.apenergy.2015.10.145>