Research of a novel short-term wind forecasting system based on multi-objective Aquila optimizer for point and interval forecast

Qianyi Xing[a], Jianzhou Wang[a,*], Haiyan Lu[b], Shuai Wang[a]

[a] *School of Statistics, Dongbei University of Finance and Economics, Dalian, 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 116025, China*

Tel.: +86 13130476286

E-mail address: wjz@lzu.edu.cn

Abstract

Facing the increasing depletion of traditional energy resources and the worsening environmental issues, wind energy sources have been widely considered. As an essential renewable energy resource, wind energy features abundant deposits, extensive distribution, non-pollution, etc. In recent years, wind power generation occupies a non-negligible position in the electric power industry. Stable and reliable power system operation demands accurate wind speed prediction (**WSP**), but the inherent randomness of wind speed sequences complicates their fluctuations and causes them to be uncontrollable. In this paper, an innovative **WSP** system is proposed, which combines data pre-processing technique, benchmark model selection, an advanced optimizer for point forecast and interval forecast. Furthermore, this paper theoretically demonstrates that the weights allocated by this optimizer are Pareto optimal solutions. Six interval data from two sites in China are utilized to validate the forecasting performance of our developed model. The experimental results indicate that the developed model can achieve superior accuracy compared to the tested models in all cases for point forecast, and also obtains the forecasting interval with high coverage and low width error, which is an extremely crucial instruction to guarantee the security and stability of the power system.

Keywords: *wind speed forecasting; data pre-processing; optimal benchmark model selection strategy; multi-objective optimization algorithm*

1. Introduction

With environmental pollution and excessive consumption of resources, wind energy is increasingly in the spotlight. Compared to the traditional energy sources, wind power has no risk of fuel prices and is also without environmental costs. Moreover, the available wind energy is extensively distributed all over the world. These unique benefits show that wind power is gradually becoming an indispensable component of sustainable development strategies in many countries[1]. In the latest Wind Energy 2021 report, the Global Wind Energy Council (GWEC) declared that newly installed wind power facilities exceed 90GW, increased by 53% compared to

2019. Fig 1 illustrates the distribution of the top five wind power markets and regions in the world[2].

Wind turbine is an essential component of wind power system, which is applied to convert wind energy into mechanical energy. Therefore, the wind turbine not only determines the output power of the whole wind power system, and it directly affects the performance of the wind turbine in terms of operational safety, stability and reliability, etc. Therefore, the research of wind turbines is particularly important[3]. Duan et al. [4] used the Turbowinds T600-48 upwind turbine airfoil data and experimental data, based on the blade element momentum theory of wind turbine modeling and numerical simulation of the aerodynamic performance to obtain the output power, and compared with the experimental data to validate the correctness and practicality of the model. Zhao et al. [5] designed a novel wind turbine based on the blade element momentum theory and conducted wind tunnel tests, and the fabricated wind turbine has more ideal output torque and power characteristics. Since the 1960s, wind power generation has received more and more attention, and precise **WSP** has become one of the necessary conditions for the stable guarantee of the power supply. However, the intermittency and instability of wind speed can have adverse effects on the stability and effectiveness of the grid systems[6]. Therefore, it is highly advisable to find reliable forecasting techniques to predict wind speed.[7]. After decades of research on **WSP** methods, they can be generally categorized into four types: physical strategies, traditional statistical strategies, artificial intelligence strategies, and combined optimization strategies[4]. The physical approach employs intricate physical parameters (temperature, humidity, barometric pressure, topographic, etc.) to perform **WSP**[9]. Specifically, numerical weather prediction (NWP) is a well-known technique in physical strategy[10]. This technique has proven to produce outstanding performance in long-term **WSP**[11]. However, it demands historical data and geographic information and also requires a lot of calculations[12]. Wang Han [13]used the sequence transfer correction algorithm to optimize the NWP model, which consequently yields comparatively satisfactory results. Nevertheless, this model failed to result in satisfactory forecasting results in the short-term forecasting field. However, conventional statistical models can achieve better performance in this field. This type of statistical model mainly contains ARMA and ARIMA[14]. Liu et al. [15] employed ARIMA to forecast wind speed, which produced satisfactory forecasting outcomes. However, due to the linear assumption of statistical methods and volatility and intermittency of wind speed, the approaches mentioned above are inadequate[16]. Hence experts started working with artificial neural networks (ANNS) in **WSP**[17]. Younes Noorollahi et al.[18] used BPNN to perform **WSP** in Iran and obtained promising results, which provided a reliable and valuable prediction system to Iran. Erasmo Cadenas et al. [19] altered the structure of the ANNS, which revealed that the most effective forecasting was realized when the ANNS had only two input and one output neuron and provided a reliable guarantee for Mexico Oaxaca energy supply. Moreover, many classical ANNS, such as BPNN[20], ELM[21], GRNN[22], ENN[23] have been broadly utilized to solve economic forecasting and power load forecasting problems and obtained superior effectiveness. Nevertheless, AI models still possess their intrinsic deficiencies, such as easily falling into local optimal solutions and overfitting[24]. Despite some intelligent optimization algorithms are proposed to optimize the hyperparameters in neural networks, such as the employment of GA to optimize the thresholds and weights of ANNS[25], the deficiencies mentioned above cannot be solved adequately. In recent years, deep learning has been widely applied to time series forecasting due to its excellent

adaptability and portability[8]. Compared with conventional shallow ANNS, it achieves better approaches to complex functions and extracts more valuable data traits by constructing networks with more hidden layers[26]. Farah Shahid et al.[27] performed **WSP** using LSTM on seven sites data in Europe and achieved good prediction results. Wang et al.[28] performed **WSP** based on the GRU neural network, the findings indicated that this method had better performance and shorter training time than the traditional models. A new forecasting technique for **WSP** based on TCN was employed by Gan et al. [29], the results demonstrated this technique had superior predictive validity in **WSP**.

However, the characteristics of wind speed change with the sites, the single forecasting model may not achieve satisfactory results in all situations, hence scholars proposed combined optimization models[30]. The combined optimization model determines the weight coefficients for several single forecasting models to achieve higher stability and accuracy[31]. Lv et al. [32] designed a MOALO, Zhang et al.[33] developed a MODA to eliminate deficiencies in **WSP**, Liu et al. [34] devised a MOGOA to improve the precision of **WSP**. For the above-mentioned studies, the combined model can achieve satisfactory forecasting results. However, as the site and forecasting interval change, the optimal individual model consequently shifts. If the single forecasting models own inferior forecasting performance, the corresponding combined model will also fail to achieve satisfactory results[35].

Moreover, the variability and fluctuations existed in the wind speed tend to reduce the reliability and effectiveness of **WSP** [36]. Therefore, the original sequence needs to be processed to increase the forecasting accuracy. Fuzzy information granulation (FIG)[37] is a widely popular pre-processing approach to address the volatility and intermittency existing in the raw sequence[38]. Cheng et al.[39] used FIG technique to preprocess the raw wind speed sequence and used SVM to forecast the granularized data, and the experimental results demonstrated that the granularized data could effectively keep the features of original data and reduce the redundant information. Pan et al. [40] adopted the FIG technique to extract features from the original sequences and obtained the minimum, mean and maximum values that can represent the original interval, and employed a regularized ELM optimized by gravity search algorithm to perform gradient prediction, and the findings indicated that the model can effectively enhance the forecasting effectiveness.

Based on the research of the above literature, the strengths and weaknesses of these approaches for **WSP** are summarized in Table 1, the overview of the above methods are as follows:

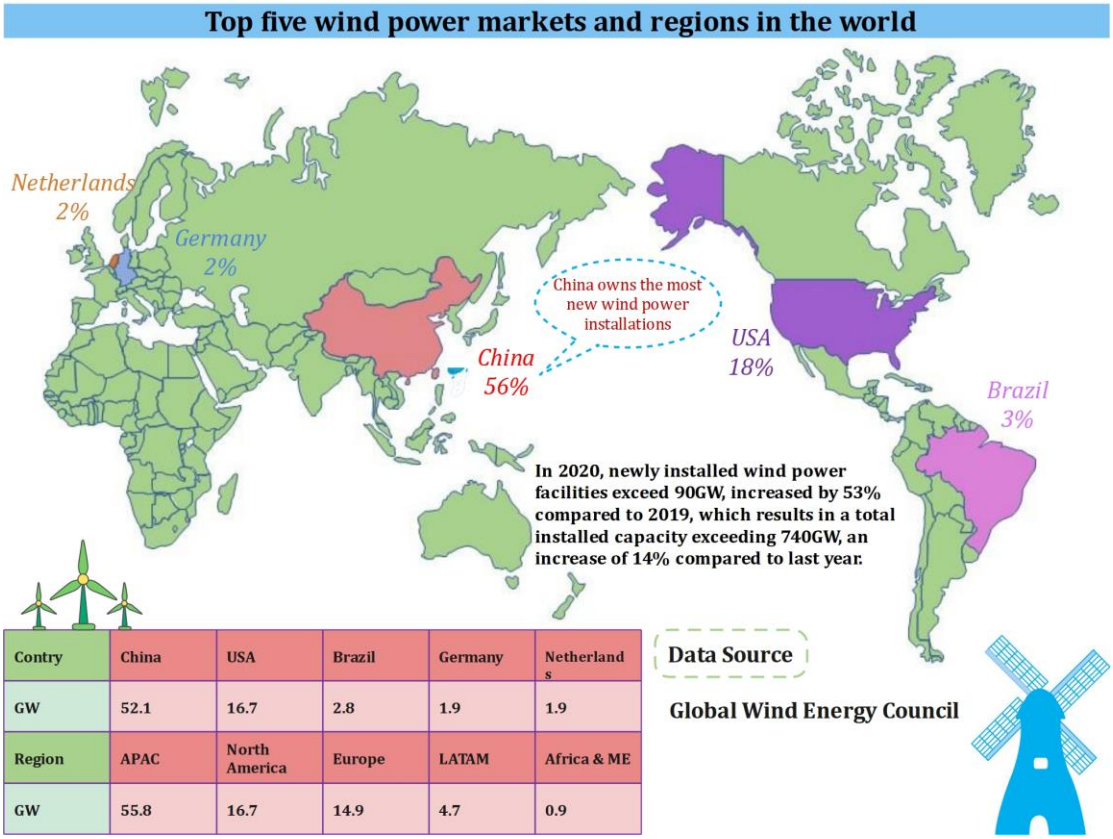
(1) The physical strategies are generally unsatisfactory in forecasting short-term wind speed and also need massive historical data and excessive calculations time.

(2) It is impractical for statistical methods to perform **WSP** with nonlinear features, since they require prior assumptions about the data distribution.

(3) Shallow neural networks can catch the nonlinear features more than other statistical methods, hence improving the forecasting accuracy. However, these methods still have shortcomings, such as overfitting, getting into local optimal.

(4) Deep learning model has multi-layer architecture, which results in a better approach to complex function and enhances forecasting efficiency.

(5) The forecasting performance of combined model is largely determined by individual forecasting models. Moreover, forecasting time and site variation can affect the selection of the benchmark models.



Contry	China	USA	Brazil	Germany	Netherlands
GW	52.1	16.7	2.8	1.9	1.9
Region	APAC	North America	Europe	LATAM	Africa & ME
GW	55.8	16.7	14.9	4.7	0.9

Fig. 1. Top five wind-generating countries and regions worldwide in 2020 form GWEC.

Table 1

Four types of models in summary.

Models	Reference	Strengths	Weaknesses
Physical strategies			
NWP system	[10,41]	It is well suited to solve global meteorological problems and is extensively applied in long-term **WSP**.	It requires significant financial expenditure and long-term calculations and is inefficient compared to other models.
Traditional statistical method			
ARIMA	[14,15]	In short-term **WSP**, these models can achieve relatively precise predictions and have shorter calculations.	Non-linear features of wind speed make these methods ineffective.
ARMA	[42,43]		
GM	[44,45]	Achieve better performance for some uncertain systems.	The gray correlation coefficient may change.
Artificial intelligence models			
Shallow neural networks			
BPNN	[20,46]	Unlike traditional statistical models, these models are effective in fitting the non-linear features of wind speed sequences, which results in more satisfactory forecasting performance.	Numerous parameters require adjustments. Despite the presence of optimization modules for these parameters in some hybrid models, they still suffer from overfitting and trapping in local optimal, etc.
ELM	[21,47]		
GRNN	[22,48]		
ENN	[34,49]		
Deep neural networks			
LSTM	[50,51]	Compared to shallow neural networks, these methods can increase the depth of networks, which allows them to possess superior learning and generalization capabilities.	Costly methods, tend to possess higher modeling complexity and require higher computing equipment capacity.
TCN	[29,52]		
GRU	[28,53]		
Combined optimization models			
MOALO	[32]	These models leverage the benefits of each forecasting model, which can significantly improve forecasting performance.	The forecasting effectiveness of combined models largely relies on single models, which implies these models are determined by the validity of single models.
MODA	[33]		
MOGOA	[34]		

Nowadays, most of **WSP** systems involve in point forecast (**PF**). However, **PF** provides no information to the distribution of real data, which makes it difficult for administrators to develop optimal strategies based on **PF**[54]. Therefore, to minimize the decision-making risk, it is imperative to apply interval forecast (**IF**) to quantify the uncertainty of **PF**[55]. Li et al. [56]proposed a novel upper and lower bound estimation (LUBE) method to extract the features of the original sequences, the findings suggested the proposed model was superior to the conventional LUBE model. Li et al. [57] proposed an improved LSSVM to quantify the uncertainty of wind speed, the results revealed the coverage width criterion of the improved LSSVM was superior to the tested models. While the above-mentioned scholars have quantified the uncertainty of wind speed sequences, it is still insufficient.

The above literature review indicates the deficiencies of the current forecasting methods and uncertainty in **PF**. Therefore, it is imperative to propose improvements and implement uncertainty measurements to the existing **WSP** system. In our paper, we propose a novel wind speed combined forecasting system (WSCFS), which involves data pre-processing, optimal benchmark model selection (OBMS) strategy, a novel multi-objective optimizer, **PF** and **IF**. **Our significant contributions are as follows:**

(1) ***The main characteristics of the original wind speed sequence are extracted and the redundant components can be effectively eliminated.*** The original sequences are decomposed into several components, and each component is fuzzified, which simplifies the sample size and retains primary information.

(2) ***The OBMS strategy is applied to select the suitable single forecasting models for the characteristics of different sites.*** Comprehensive evaluation metrics (CEM) objectively reflect the performance of different single forecasting models in different situations, which can significantly improve the forecasting performance of our proposed WSCFS.

(3) ***A multi-objective variant of the Aquila optimizer is proposed.*** MOAO can satisfy the accuracy and stability of forecasting demand simultaneously by the pareto optimality criterion. Moreover, by incorporating the roulette and archive mechanism, the global search ability and search accuracy are significantly improved, which enables our WSCFS to possess superior forecasting performance and generalization.

(4) ***We theoretically demonstrate that the weights assigned by this optimizer are pareto optimal solutions.*** This allows the combined model to have superior forecasting performance than the single model.

(5) ***We consider the fitting distribution of forecasting sequences and propose a valid probabilistic forecasting method to quantify and analyze the uncertainty in PF.*** The optimal distributions of forecasting sequences are applied to determine the forecasting intervals, which allows the managers to quantify the risk and instability of the power system operation.

In the upcoming studies. The methodology and framework of our proposed WSCFS are illustrated in Sections 2 and Sections 3. In Section 4, we describe the data sources, evaluation metrics, and simulated experimental analysis. Moreover, to further validate the stability and validity of the WSCFS, some specific discussions will be illustrated in Section 5. Finally, the experimental conclusions will be presented in Section 6.

Table 2
List of abbreviations

Abbreviations	Nomenclatures
AI	Artificial intelligence
ANNs	Artificial neural networks
AIS	Average interval score
ALO	Ant lion optimizer
AO	Aquila optimizer
ARMA	Autoregressive moving average
ARIMA	Auto-regressive integrated moving average
AWD	Accumulated width deviation
AWNN	Adaptive Wavelet Neural Network
BPNN	Back propagation network
CEM	Comprehensive evaluation metrics
CFM	Combined forecasting model
DA	Dragonfly algorithm
DF	Distribution function
ELM	Extreme Learning Machine
ENN	Elman neural network
FIG	Fuzzy information granulation
GA	Genetic algorithms
GM	Grey model
GOA	Grasshopper optimization algorithm
GRNN	General regression neural network
GRU	Gated recurrent unit
IF	Interval forecast
IR	Improvement ratio
LSSVM	Least square support vector machines
LSTM	Long short-term memory network
MAPE	Mean absolute percent error
MAE	Mean absolute error
MOAO	Multi-objective aquila optimizer
MODA	Multi-objective dragonfly algorithm
MOGOA	Multi-objective grasshopper optimization algorithm
MOALO	Multi-objective ant lion optimization algorithm
NWP	Numerical weather prediction
OBMS	Optimal benchmark models selection
PF	Point forecast
PI	Prediction interval
PICP	Prediction interval coverage probability
RMSE	Root mean square error
SDE	Standard deviation of the error
TCN	Temporal convolutional network
WSCFS	Wind speed combined forecasting system
WSP	Wind speed prediction

2. Methodology

In this section, we will illustrate the methodologies involved in our WSCFS.

The Fuzzy Information Granulation Technique

In our study, we adopt FIG to process the original wind speed sequences. The procedure is split into three steps，dividing the original sequences, determining the window length $\tilde{\tilde{\mathbf{w}}}'(l)$ and building fuzzy information particles. The original wind speed sequences $\tilde{\tilde{\mathbf{x}}}_i(t)$ $(i=1,2,...,6)$ are divided into subsequences with equal length as the operational windows, then the fuzzy particle $\tilde{\mathbf{P}}_s'$ is established through the affiliate function, which constructs a fuzzy set $\tilde{\tilde{\mathbf{G}}}_\kappa''$ to replace the original sequence.

$$\tilde{\mathbf{P}}_s' \triangleq \tilde{\tilde{\mathbf{x}}}_1(t) \rightarrow \tilde{\tilde{\mathbf{G}}}_\kappa'' \tag{1}$$

The data in the operational windows are fuzzified by the affiliate function. In this paper, the triangular fuzzy particle is adopted and its affiliate function is defined as:

$$\mathbf{A}[\tilde{\tilde{\mathbf{x}}}_1(t),\hat{\boldsymbol{\lambda}}_{min},\hat{\boldsymbol{\lambda}}_{mean},\hat{\boldsymbol{\lambda}}_{max}]=\begin{cases}0, & \tilde{\tilde{\mathbf{x}}}_1(t)<\hat{\boldsymbol{\lambda}}_{min}\\ \dfrac{\tilde{\tilde{\mathbf{x}}}_1(t)-\hat{\boldsymbol{\lambda}}_{min}}{\hat{\boldsymbol{\lambda}}_{mean}-\hat{\boldsymbol{\lambda}}_{min}}, & \hat{\boldsymbol{\lambda}}_{min}\le\tilde{\tilde{\mathbf{x}}}_1(t)\le\hat{\boldsymbol{\lambda}}_{mean}\\ \dfrac{\hat{\boldsymbol{\lambda}}_{max}-\tilde{\tilde{\mathbf{x}}}_1(t)}{\hat{\boldsymbol{\lambda}}_{max}-\hat{\boldsymbol{\lambda}}_{mean}}, & \hat{\boldsymbol{\lambda}}_{mean}<\tilde{\tilde{\mathbf{x}}}_1(t)\le\hat{\boldsymbol{\lambda}}_{max}\\ 0, & \tilde{\tilde{\mathbf{x}}}_1(t)>\hat{\boldsymbol{\lambda}}_{max}\end{cases} \tag{2}$$

Where $\tilde{\tilde{\mathbf{x}}}_1(t)$ refers to the original sequences, $\hat{\boldsymbol{\lambda}}_{min}$, $\hat{\boldsymbol{\lambda}}_{mean}$ and $\hat{\boldsymbol{\lambda}}_{max}$ are the minimum, average, and maximum values of each fuzzy particle, respectively.

Optimal Benchmark Model Selection (*OBMS*)

In our study, we choose seven single forecasting models as benchmark models to improve the forecasting performance of the WSCFS in different situations, which are BPNN, ELM, ENN, GRNN, LSTM, TCN and GRU, respectively.

In our proposed WSCFS, the OBMS strategy is applied to choose the optimal individual models, for enhancing the generalizability and robustness of our proposed WSCFS[58]. The process of constructing the CEM criterion and its formula is detailed as follows.

(1) Calculating the **SDE**, **RMSE**, **MAE**, and **MAPE** of the forecasting values based on each benchmark model respectively.

(2) The calculated error metrics are normalized by using Eq. (3).

$$\mathbf{Metrics}_i^* = \frac{\{\mathbf{Metrics}_i - \min\limits_{1\le i\le N_j}(\mathbf{Metrics})\}}{\{\max\limits_{1\le i\le N_j}(\mathbf{Metrics}) - \min\limits_{1\le i\le N_j}(\mathbf{Metrics})\}} \tag{3}$$

Where N_j represents the number of individual models, $\mathbf{Metrics}_i^*$ refers to the normalized values of **SDE**, **RMSE**, **MAE**, and **MAPE** based on individual models.

(3) Calculating **CEM** values for each model with Eq. (4).

$$\mathbf{CEM}_i^* = \left(0.25*\mathbf{SDE}_i^* + 0.25*\mathbf{RMSE}_i^* + 0.25*\mathbf{MAE}_i^* + 0.25*\mathbf{MAPE}_i^*\right) \tag{4}$$

(4) The five optimal benchmark models are chosen according to the smallest **CEM** values.

Meta-heuristic Optimization Algorithm

In recent years, meta-heuristic optimization algorithms have become an effective solution to complicated optimization projects for their non-dependence on gradient information, the avoidance of falling into local optimality significantly [59].

(a) ***Aquila Optimizer*** **(*AO*)**

Laith Abualigah et al.[60] proposed a novel intelligence optimizer, which was incited by the hunting behavior of Aquila. This algorithm searches for optimal solutions through the four methods of Aquila's hunting process. The principle of the algorithm is shown below.

- ***Expanded exploration: high soar with a vertical stoop*** ($\boldsymbol{\Psi}_1$)

In the first method ($\boldsymbol{\Psi}_1$), Aquila soars high with a vertical stoop to hunt airborne birds. This hunting behavior is presented through Eq. (5).

$$\boldsymbol{\Psi}_1\left(t'_{current}+1\right)=\boldsymbol{\Psi}_{best}\left(t'_{current}\right)\times\left(1-\frac{t'_{current}}{T''_{max}}\right)+\left(\boldsymbol{\Psi}_{mean}\left(t'_{current}\right)-\boldsymbol{\Psi}_{best}\left(t'_{current}\right)*\boldsymbol{\varepsilon}_0^{1i}\right) \tag{5}$$

Where T''_{max} and $t'_{current}$ represent the maximal and current iterations. $\boldsymbol{\Psi}_1\left(t'_{ccurrent}+1\right)$ represents the next iterative solution, $\boldsymbol{\Psi}_{best}\left(t'_{ccurrent}\right)$ refers to the optimal solution at $t'_{ccurrent}$. $\boldsymbol{\Psi}_{mean}\left(t'_{ccurrent}\right)$ refers to the mean of the solution position at $t'_{ccurrent}$, which is calculated by $\boldsymbol{\Psi}_{mean}\left(t'_{current}\right)=mean\left(\boldsymbol{\Psi}_i\left(t'_{current}\right)\right)$. $\boldsymbol{\varepsilon}_0^{1i}$ is a random number in$[0,1]$.

- ***Narrowed exploration: contour flight with short glide attack*** ($\boldsymbol{\Psi}_2$)

Aquila hovers low around the hunting area and then glides to hit the target. This hunting behavior is presented through Eq. (6).

$$\boldsymbol{\Psi}_2\left(t'_{current}+1\right)=\boldsymbol{\Psi}_{best}\left(t'_{current}\right)\times\mathbf{Levy}_f^{\varsigma}(D)+\boldsymbol{\Psi}_r\left(t'_{current}\right)+(\overleftrightarrow{\mathbf{y}}-\overleftrightarrow{\mathbf{x}})*\boldsymbol{\varepsilon}_0^{2i} \tag{6}$$

Where $\boldsymbol{\Psi}_2\left(t'_{current}+1\right)$ represents the next iterative solution, D is the dimension of the proposed problem, $\mathbf{Levy}_f^{\varsigma}(D)$ is the *Lévy* flight function, the detailed calculation is given in [60], and $\boldsymbol{\Psi}_r\left(t'_{current}\right)$ is the random solution generated in the range $\left[1,...,\boldsymbol{N}\right]$ at $t'_{current}$. $\overleftrightarrow{\boldsymbol{x}}$ and $\overleftrightarrow{\boldsymbol{y}}$ represent the spiral search space, which can be calculated in $\overleftrightarrow{\mathbf{x}}=\tilde{\mathbf{r}}_j^{\zeta}\times\mathbf{sin}(\theta)$ and $\overleftrightarrow{\mathbf{y}}=\tilde{\mathbf{r}}_j^{\zeta}\times\mathbf{cos}(\theta)$, where $\tilde{\mathbf{r}}_j^{\zeta}=\mathbf{r}_k^{\gamma}+\boldsymbol{\rho}_c^{*}\times\mathbf{D}_d^{\delta}$ and $\boldsymbol{\theta}=-\boldsymbol{\omega}_c^{*}\times\mathbf{D}_d^{\delta}+3\boldsymbol{\pi}/2$. $\mathbf{r}_k^{\gamma}$ is a value from 1 to 20 for specific search periods, $\mathbf{D}_d^{\delta}$ is an integer, $\boldsymbol{\rho}_c^{*}$ and $\boldsymbol{\omega}_c^{*}$ are both constants.

- ***Expanded exploration: Slow-flying attack with vertically descent*** ($\boldsymbol{\Psi}_3$)

In the third method ($\boldsymbol{\Psi}_3$), the Aquila detects the prey and then swoops down and flies low to attack the target. This hunting behavior is presented through Eq. (7).

$$\boldsymbol{\Psi}_3\left(t'_{current}+1\right)=\left(\boldsymbol{\Psi}_{best}\left(t'_{current}\right)-\boldsymbol{\Psi}_{mean}\left(t'_{current}\right)\right)\times\boldsymbol{\alpha}_{\tau}^{\vartheta}-\boldsymbol{\varepsilon}_0^{3i}+\left(\left(\overline{\overline{\mathbf{UB}}}^{\bullet}-\overline{\overline{\mathbf{LB}}}^{\bullet}\right)*\boldsymbol{\varepsilon}_0^{3i}+\overline{\overline{\mathbf{LB}}}^{\bullet}\right)\times\boldsymbol{\delta}_{\tau}^{\vartheta} \tag{7}$$

Where $\boldsymbol{\Psi}_3\left(t'_{current}+1\right)$ represents the next iterative solution, $\boldsymbol{\varepsilon}_0^{3i}$ is a random number, $\boldsymbol{\alpha}_{\tau}^{\vartheta}$ and $\boldsymbol{\delta}_{\tau}^{\vartheta}$ are fixed at 0.1. $\overline{\overline{\mathbf{UB}}}^{\bullet}$ and $\overline{\overline{\mathbf{LB}}}^{\bullet}$ refer to the upper and lower bound for

the proposed problem, respectively.

- ***Narrowed exploration: Walk and grab prey*** ($\boldsymbol{\Psi}_4$)

The fourth method is walking and seizing prey. This hunting behavior is presented through Eq. (8).

$$\boldsymbol{\Psi}_4\left(t'_{current}+1\right)=\mathbf{QF}_1\times\boldsymbol{\Psi}_{best}\left(t'_{current}\right)-\left(\mathbf{G}_1\times\boldsymbol{\Psi}_4\left(t'_{current}\right)*\boldsymbol{\varepsilon}_0^{4i}\right)-\mathbf{G}_2\times\mathbf{Levy}_f^{\varsigma}(D)+\boldsymbol{\varepsilon}_0^{4i}*\mathbf{G}_1 \tag{8}$$

Where $\boldsymbol{\Psi}_4\left(t'_{current}+1\right)$ represents the next iterative solution, $\mathbf{QF}_1$ denotes quality function, which is calculated by $\mathbf{QF}_1=t'^{\,2\times\varepsilon_0^{4i}-1/(T''_{\max}-1)^2}_{current}$. $\mathbf{G}_1$ is a random value in $[-1,1]$. $\mathbf{G}_2$ means the flight gradient, which descends from 2 to 0 and its calculation equation is $\mathbf{G}_2=2\times\left(1-t'_{current}/T''_{\max}\right)$.

(b) ***Multi-Objective Aquila Optimizer*** (***MOAO***)

In this section, the proof of pareto optimality, objective function setting, and archive with roulette wheel will be presented, respectively. The flow chart and pseudo-code of **MOAO** will be presented in Fig 2 and Table 3.

- ***Pareto Optimality***

Multi-objective optimized problems require optimizing one or more two objective functions, and in most scenarios, the objective functions conflict with each other. Hence, scholars introduced the pareto optimality to address this problem[61]. Considering two vectors $\overrightarrow{\overrightarrow{\mathbf{M}}}=\left(\overrightarrow{\overrightarrow{\mathrm{M}_1}},\overrightarrow{\overrightarrow{\mathrm{M}_2}},\cdots,\overrightarrow{\overrightarrow{\mathrm{M}_k}}\right)$ and $\overrightarrow{\overrightarrow{\mathbf{N}}}=\left(\overrightarrow{\overrightarrow{\mathrm{N}_1}},\overrightarrow{\overrightarrow{\mathrm{N}_2}},\cdots,\overrightarrow{\overrightarrow{\mathrm{N}_k}}\right)$, if $[\forall s\in\{1,2,...,S\}:\mathbf{H}_s(\overrightarrow{\overrightarrow{M}})\geq\mathbf{H}_s(\overrightarrow{\overrightarrow{N}})]\cap[\exists s\in\{1,2,...,S\}:\mathbf{H}_s(\overrightarrow{\overrightarrow{M}})>\mathbf{H}_s(\overrightarrow{\overrightarrow{N}})]$,where $\mathbf{H}_s(\overrightarrow{\overrightarrow{M}})$ is the s_{th} objective function, we refer to $\overrightarrow{\overrightarrow{\mathbf{M}}}$ precedes $\overrightarrow{\overrightarrow{\mathbf{N}}}$ ($\overrightarrow{\overrightarrow{\mathbf{M}}}\succsim\overrightarrow{\overrightarrow{\mathbf{N}}}$).

If none gained solutions precedes $\overrightarrow{\overrightarrow{\mathbf{M}}}$, we refer to $\overrightarrow{\overrightarrow{\mathbf{M}}}\in\mathbf{M}$ is the optimal solution, i.e., $\forall s=1,2,...,S:/\exists\overrightarrow{\overrightarrow{\mathbf{N}}}\in\mathbf{M}\,|\,\mathbf{H}_s(\overrightarrow{\overrightarrow{N}})\succsim\mathbf{H}_s(\overrightarrow{\overrightarrow{M}})$.

- ***Archive with Roulette Wheel***

To improve the local and global search capability of **MOAO**, the pareto optimal results are archived. Once the archive is full, the solution with the most adjacent individuals will be replaced with a new one based on the roulette wheel mechanism, the probability of being replaced is defined by $\mathbf{P_i}=n_i/q$, $q>1$. n_i means the attachment individuals' number, and q is obtained through the pareto mechanism.

- ***Objective Functions***

Meanwhile, the **MOAO** objective function is defined as follows:

$$\boldsymbol{Minimize}\begin{cases}\mathbf{Obj}_1(P)=\dfrac{1}{K}\sum_{i=1}^{K}|\,(P_i^{act}-P_i^{pre})/P_i^{act}\,|\times100\%\\ \mathbf{Obj}_2(P)=std(P_i^{act}-P_i^{pre}),\ i=1,2,...,K\end{cases} \tag{9}$$

Where P_i^{act} and P_i^{pre} denote to the i_{th} actual value and predictive value.

- ***Proof***

If the optimal solution $\boldsymbol{Wt}(\boldsymbol{Q}^{\delta})$ obtained by **MOAO** doesn't have the minimum fitness value, there should be at least one neighboring solution $\boldsymbol{P}^{\delta}=\boldsymbol{Q}^{\delta}+\chi^{*}$,($\chi^{*}>0$). Its weight satisfies $[\forall\boldsymbol{s}\in\{\mathbf{1},\mathbf{2},...,\boldsymbol{S}\}:\boldsymbol{fit}_s(\boldsymbol{P}^{\delta})\geq\boldsymbol{fit}_s(\boldsymbol{Q}^{\delta})]\cap[\exists\boldsymbol{s}\in\{\mathbf{1},\mathbf{2},...,\boldsymbol{S}\}:\boldsymbol{fit}_s(\boldsymbol{P}^{\delta})>\boldsymbol{fit}_s(\boldsymbol{Q}^{\delta})]$, so $\boldsymbol{P}^{\delta}$ is saved in the archive, in which $\boldsymbol{Wt}(\boldsymbol{P}^{\delta})$ precedes $\boldsymbol{Wt}(\boldsymbol{Q}^{\delta})$. Due to the limited storage capacity of the archive, $\boldsymbol{Q}^{\delta}$ is removed from the archive with probability n_i/q or ranked behind $\boldsymbol{P}^{\delta}$. The position with the highest fitness value is

considered to be the optimal position, and the optimal weight matrix is $Wt(P^{\delta})$ instead of $Wt(Q^{\delta})$ .The Pareto optimal solutions are proved by the proof of contradiction.

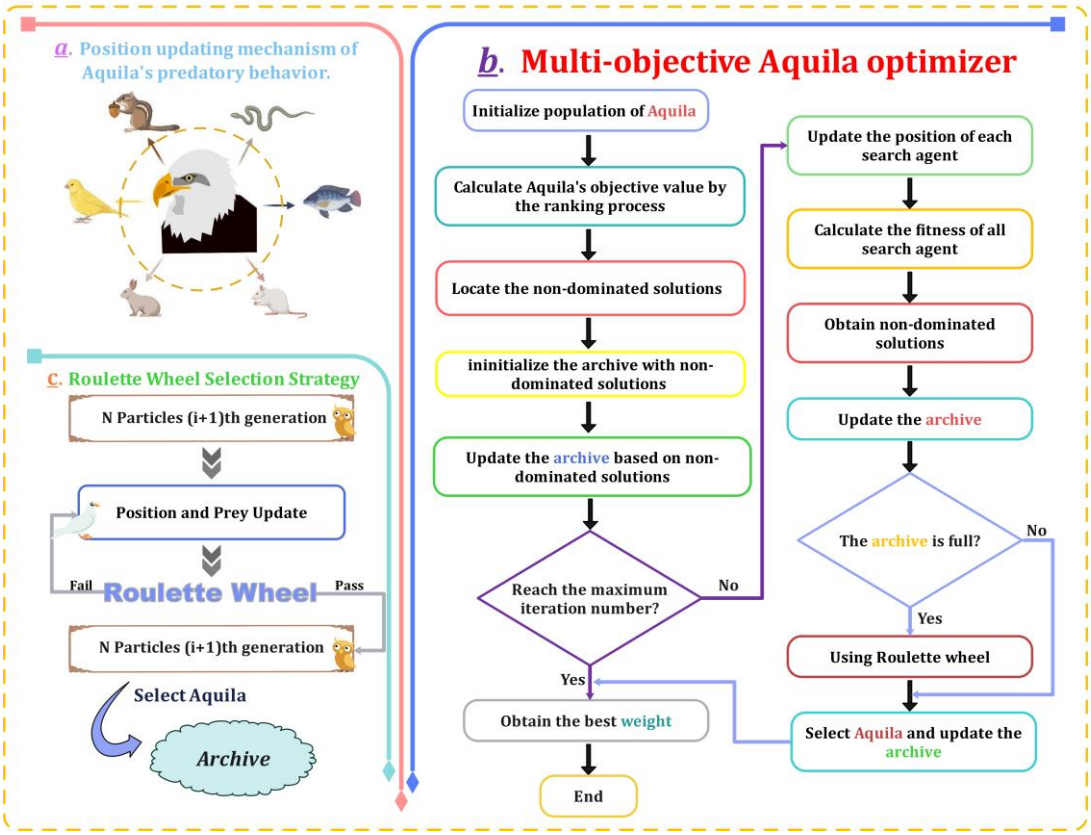


Fig. 2. The flow chart of MOAO.

Table 3
Pseudo-code for MOAO

***Algorithm 1*: MOAO**

Objective functions:

$$\textbf{\textit{Minimize}}\begin{cases}\textbf{Obj}_1(P)=\dfrac{1}{K}\sum_{i=1}^{K}|\,(P_i^{act}-P_i^{pre})/P_i^{act}\,|\times 100\% \\ \textbf{Obj}_2(P)=std(P_i^{act}-P_i^{pre})\ ,\ i=1,2,...,K\end{cases}$$

Input: $\boldsymbol{Y}_{train}^{(0)}=(\tilde{\boldsymbol{x}}^{(0)}(1),\tilde{\boldsymbol{x}}^{(0)}(2),\cdots,\tilde{\boldsymbol{x}}^{(0)}(n))$ ——Training Data

$\boldsymbol{Y}_{test}^{(0)}=(\tilde{\boldsymbol{x}}^{(0)}(n),\tilde{\boldsymbol{x}}^{(0)}(n+1),\cdots,\tilde{\boldsymbol{x}}^{(0)}(m))$ ——Testing Data

Output: $\boldsymbol{W}^*$——Optimal weights via optimization

Parameters:

$Inter_{max}$ - the maximum iteration number
n - the Aquila's number
$\boldsymbol{F}_i$ - the i_{th} Aquila's fitness function
$\boldsymbol{x}_i$- the i_{th} Aquila's position
t – the number of current iterations
d - the dimensions' number

/* Initialize the parameters of **MOAO** (i.e., $\boldsymbol{\alpha}_\tau^\vartheta$, $\boldsymbol{\delta}_\tau^\vartheta$,etc.). ***/**
/* Initialize candidate solution of **MOAO** randomly. ***/**
For each Aquila
Calculate objective value by the ranking process.
End for
/* Identify the best search agent X^* ***/**
While the criteria are not fulfilled
  For each Aquila
    Update the $\ddot{\mathbf{x}}$, $\ddot{\mathbf{y}}$, $\mathbf{G}_1$, $\boldsymbol{G}_2$, $\mathbf{Levy}_f^{\zeta}(D)$, etc.
    If $t'_{current}\le(2/3)*T''_{max}$ **then**
      If $rand\le 0.5$ **then**
        ▷ **step 1: Expanded exploration ($\boldsymbol{\Psi}_1$)**
        /* Update the Aquila's position by Eq.(5).*/
        If $rand>0.5$ **then**
          ▷ **step 2: Narrowed exploration ($\boldsymbol{\Psi}_2$)**
          /* Update the Aquila's position by Eq.(6).*/
        End if
      End if
    Else $t'_{current}>(2/3)*T''_{max}$ **then**
      If $rand\le 0.5$ **then**
        ▷ **step 3: Expanded exploration ($\boldsymbol{\Psi}_3$)**
        /* Update the Aquila's position by Eq.(7).*/
        If $rand>0.5$ **then**
          ▷ **step 4: Narrowed exploration ($\boldsymbol{\Psi}_4$)**
          /* Update the Aquila's position by Eq.(8).*/
        End if
      End if
    End else
  End for
  /* Check whether Aquila's position exceeds the search boundary, if so, correct it.*/
  /* Calculate all Aquila's objective values. */
  /* Locate the non-dominated solutions. */
  /* Update the archive based on non-dominated solutions. */
  If the archive is filled
    /* Using Roulette wheel with $\mathbf{P}_i=n_i/q$ to remove poorer solutions and add better solutions */
  End if
End while
Return archive
Acquire the best $\boldsymbol{W}^*$

3. Framework of Our Proposed WSCFS

In this part, we will introduce our proposed WSCFS, which includes data fuzzy granulation, OBMS, **PF** and **IF**. The detailed processes of WSCFS are shown below, and the framework is depicted in Fig 3.

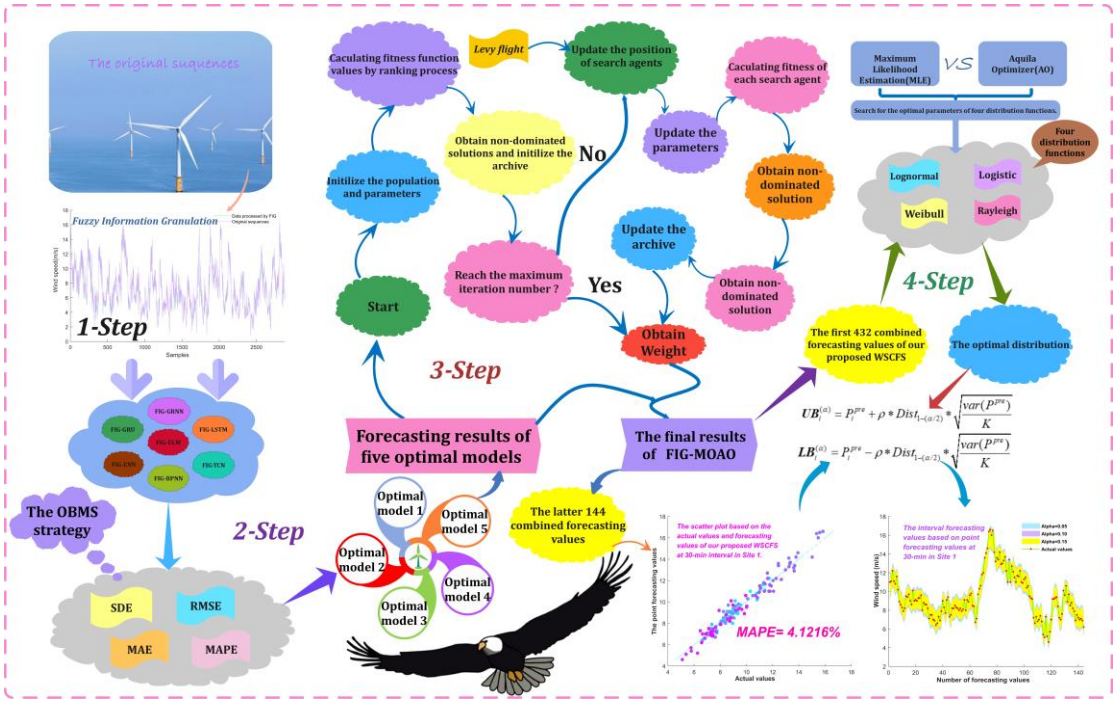


Fig. 3. The flow of our WSCFS.

3.1 Data Fuzzy Granulation

The FIG technique reconstructs and granulates the chaotic original sequences. After applying FIG to the original sequence $\tilde{\tilde{\mathbf{x}}}_i(t)$, we can extract three main parameters $\hat{\mathbf{\Phi}}'_{low}(t)$, $\hat{\mathbf{\Phi}}'_{upper}(t)$ and $\hat{\mathbf{\Phi}}'_{trend}(t)$, which respectively refer to the lower, upper bound, and trend. In our study, we employ $\hat{\mathbf{\Phi}}'_{trend}(t)$ for the following prediction in our proposed WSCFS.

3.2 Optimal Forecasting Model Selection

Considering none of any individual forecasting models can achieve satisfactory results in all cases, we formulate the OBMS strategy. Based on the rolling prediction mechanism, which will be utilized in one-step and multi-step forecasts. Six granulated sequences $\hat{\mathbf{\Phi}}'_{trend}(1\ th-2880\ th)$ are applied to the benchmark models for **WSP**. The $\hat{\mathbf{\Phi}}'_{trend}(1\ th-2160\ th)$ serve as the training set for the benchmark models, the $\hat{\mathbf{\Phi}}'_{trend}(2160\ th-2304\ th)$ serve as the validation set to determine the hyperparameters in each benchmark model and validate the training effectiveness of the model, and the $\hat{\mathbf{\Phi}}'_{trend}(2305\ th-2880\ th)$ serve as the testing set to verify the performance of candidate models. The optimal sequences $\hat{\tilde{\mathbf{y}}}_{ij}(2305\ th-2880\ th)\ (i=1,2,...,6\ ;\ j=1,2,...,5)$ will be integrated by **MOAO** to obtain the combined forecasting sequences $\hat{\tilde{\mathbf{Y}}}_i^{combined}(2737\ th\text{-}2880\ th)\ (i=1,2,...,6)$. Therefore, the **CEM** values are calculated based on the $\hat{\tilde{\mathbf{y}}}_{ij}(2305\ th-2880\ th)$ and the candidate models with the smallest **CEM** are

considered as the benchmark models. Specially, the optimal benchmark models in different situations are illustrated in Table 4.

Table 4

Results of optimal benchmark model in different situations.

Interval	Site 1			Site 2		
	1-Step	2-Step	3-Step	1-Step	2-Step	3-Step
30 min	FIG-ELM	FIG-LSTM	FIG-LSTM	FIG-LSTM	FIG-ELM	FIG-ELM
	FIG-ENN	FIG-ELM	FIG-ELM	FIG-ELM	FIG-BP	FIG-GRU
	FIG-BP	FIG-ENN	FIG-BP	FIG-BP	FIG-ENN	FIG-ENN
	FIG-LSTM	FIG-BP	FIG-ENN	FIG-ENN	FIG-LSTM	FIG-BP
	FIG-GRU	FIG-GRU	FIG-GRU	FIG-GRU	FIG-GRU	FIG-LSTM
60 min	FIG-ENN	FIG-BP	FIG-ENN	FIG-ENN	FIG-ENN	FIG-ENN
	FIG-ELM	FIG-ELM	FIG-LSTM	FIG-ELM	FIG-ELM	FIG-ELM
	FIG-BP	FIG-ENN	FIG-ELM	FIG-BP	FIG-BP	FIG-GRU
	FIG-LSTM	FIG-LSTM	FIG-BP	FIG-LSTM	FIG-LSTM	FIG-BP
	FIG-GRU	FIG-GRU	FIG-GRU	FIG-GRU	FIG-GRU	FIG-LSTM
120 min	FIG-ELM	FIG-ENN	FIG-ENN	FIG-LSTM	FIG-ENN	FIG-ENN
	FIG-ENN	FIG-ELM	FIG-ELM	FIG-ELM	FIG-ELM	FIG-GRU
	FIG-BP	FIG-BP	FIG-BP	FIG-ENN	FIG-LSTM	FIG-ELM
	FIG-LSTM	FIG-LSTM	FIG-LSTM	FIG-BP	FIG-BP	FIG-BP
	FIG-GRNN	FIG-GRNN	FIG-GRNN	FIG-GRU	FIG-GRU	FIG-LSTM

Note: In this table, the five optimal benchmark models are selected based on the OBMS strategy, and the higher-ranked ones imply that they have better forecasting performance.

3.3 Point Forecast

After using OBMS strategy, the five optimal sequences $\hat{\tilde{\mathbf{y}}}_{ij}(2305\ th-2880\ th)$ are integrated into the combined model. In **MOAO**, the $\hat{\tilde{\mathbf{y}}}_{ij}(2305\ th\text{-}2736\ th)$ are served as training sets to obtain $\left\{\boldsymbol{w}_j^i\right\}$ $(i=1,2,...,6;\ j=1,2,...5)$, the detailed procedure for calculating $\left\{\boldsymbol{w}_j^i\right\}$ is as follows.

$$\left\{\boldsymbol{w}_j^i\right\} = arg\ min_{\left\{\boldsymbol{w}_j^i\right\}}\left\{Obj_1, Obj_2\right\}$$

$$= arg\ min_{\left\{\boldsymbol{w}_j^i\right\}}\left\{\begin{array}{c}\text{mean}\left(\left|\dfrac{\tilde{\tilde{\boldsymbol{x}}}_i(2305th-2736th)-\hat{\tilde{\mathbf{Y}}}_i^{train}(2305\ th\text{-}2736\ th)}{\tilde{\tilde{\boldsymbol{x}}}_i(2305th-2736th)}\right|\right)\times 100\% \\ std\left[\tilde{\tilde{\boldsymbol{x}}}_i(2305th-2736th)-\hat{\tilde{\mathbf{Y}}}_i^{train}(2305\ th\text{-}2736\ th)\right]\end{array}\right\} \tag{10}$$

$$s.t.\ \sum_{j=1}^{J}\boldsymbol{w}_j^i = 1,\ -2 \le \boldsymbol{w}_j^i \le 2\ \forall \boldsymbol{J}$$

$$\hat{\tilde{\mathbf{Y}}}_i^{train}(2305\ th\text{-}2736\ th) = \sum_{j=1}^{J}\boldsymbol{w}_j^i * \left[\hat{\tilde{\mathbf{y}}}_{ij}(2305\ th\text{-}2736\ th)\right] \tag{11}$$

Furthermore, $\hat{\tilde{\mathbf{Y}}}_i^{train}(2305\ th\text{-}2736\ th)$ will also be utilized in the interval forecast to measure the uncertainty features of $\hat{\tilde{\mathbf{Y}}}_i^{combined}(2737\ th\text{-}2880\ th)$. Then $\hat{\tilde{\mathbf{y}}}_{ij}(2305\ th\text{-}2736\ th)$ are served as testing sets to be multiplied with $\left\{\boldsymbol{w}_j^i\right\}$ to get $\hat{\tilde{\mathbf{Y}}}_i^{combined}(2737\ th\text{-}2880\ th)$. The detailed equation for this process is as follows.

$$\hat{\tilde{\mathbf{Y}}}_i^{combined}(2737\ th\text{-}2880\ th) = \sum_{j=1}^{J}\boldsymbol{w}_j^i * \left[\hat{\tilde{\mathbf{y}}}_{ij}(2737\ th\text{-}2880\ th)\right] \tag{12}$$

3.4 Interval Forecasting Based on Point Forecasting Results

In our study, four distribution functions (DF), including Lognormal, Weibull, Logistic, and Rayleigh distributions, are selected to determine the optimal distribution

of $\hat{\tilde{\mathbf{Y}}}_i^{train}(2305\ th\text{-}2736\ th)$. Moreover, in practical applications, since the distribution of $\hat{\tilde{\mathbf{Y}}}_i^{combined}(2737\ th\text{-}2880\ th)$ is unknown, these DFs are supposed to be fitted to the real-time $\hat{\tilde{\mathbf{Y}}}_i^{train}(2305\ th\text{-}2736\ th)$. The fitting performance largely relies on the distribution parameters. Therefore, we adopt two methods to search the optimal parameters of four distributions, which are the MLE and **AO**. The effectiveness of the fitting is also evaluated by goodness-of-fit ($\mathbf{R^2}$). Furthermore, the upper and lower bounds of the forecasting interval are calculated by the following equations.

$$Upperbound = \hat{\tilde{\mathbf{Y}}}_i^{combined}(2737\ th\text{-}2880\ th) + \rho * Dist_{1-(\alpha/2)} * \sqrt{\frac{var(\hat{\tilde{\mathbf{Y}}}_i^{train}(2305\ th\text{-}2736\ th))}{432}} \tag{13}$$

$$Lowerbound = \hat{\tilde{\mathbf{Y}}}_i^{combined}(2737\ th\text{-}2880\ th) - \rho * Dist_{1-(\alpha/2)} * \sqrt{\frac{var(\hat{\tilde{\mathbf{Y}}}_i^{train}(2305\ th\text{-}2736\ th))}{432}} \tag{14}$$

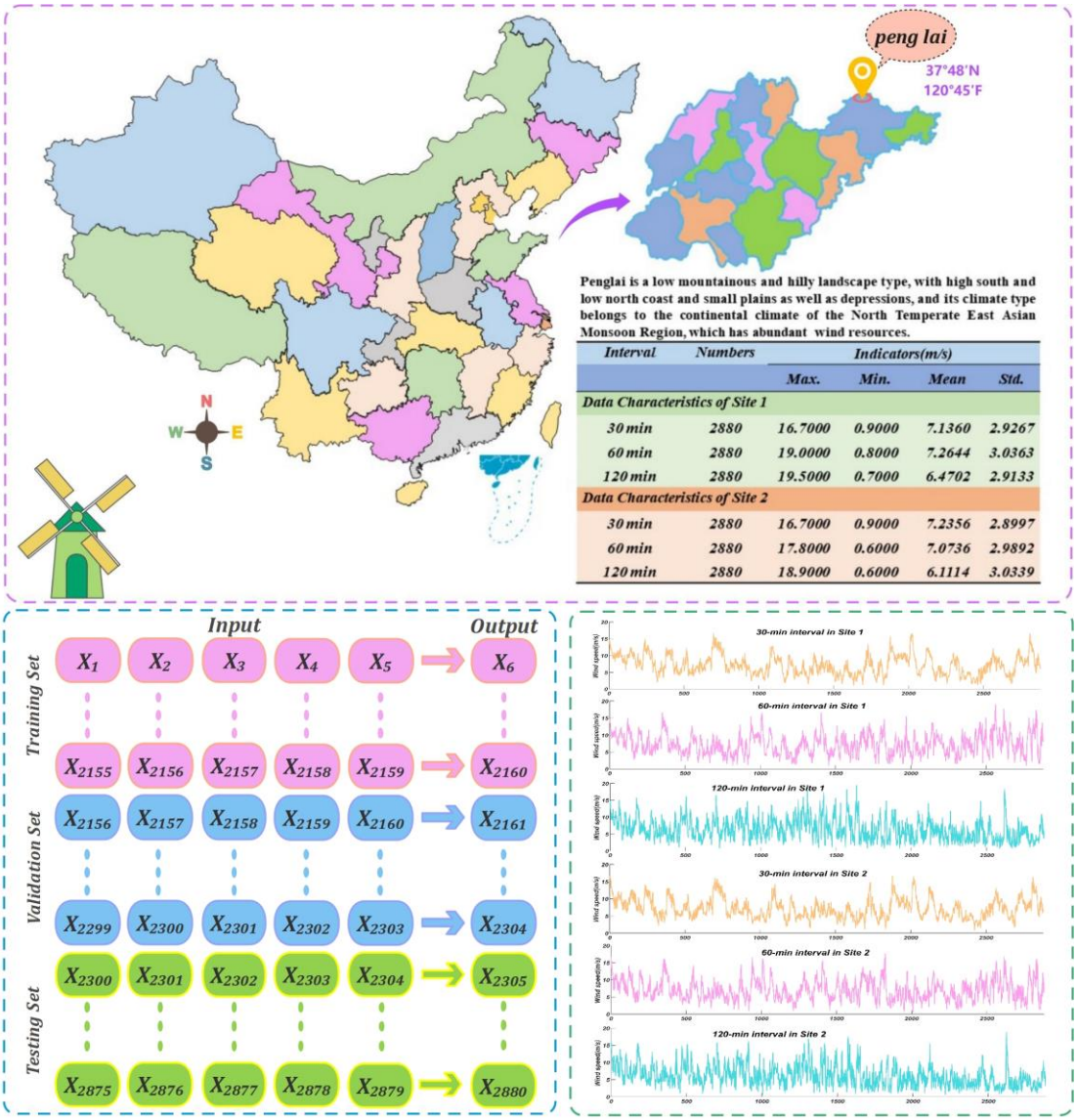
Where ρ represents the width adjusted coefficient for the **IF**, which is fixed to 0.6 in our study. $Dist_{1-(\alpha/2)}$ refers to the quantile of DFs.

4. Experiment and Analysis

We performed five experiments to validate the effectiveness of our proposed WSCFS in this section. Moreover, the dataset source, some statistical metrics of the samples, and the indicators employed to evaluate the **PF** and **IF** performance are also shown in this section.

4.1 Data source presentation

In our study, we adopt two datasets from Penglai, Shandong, China. Each site contains three groups with different time intervals of wind speed sequences. The specific data characteristics and the regions of research are shown in Fig 4.



Interval	Numbers	Indicators(m/s)			
		Max.	Min.	Mean	Std.
Data Characteristics of Site 1					
30 min	2880	16.7000	0.9000	7.1360	2.9267
60 min	2880	19.0000	0.8000	7.2644	3.0363
120 min	2880	19.5000	0.7000	6.4702	2.9133
Data Characteristics of Site 2					
30 min	2880	16.7000	0.9000	7.2356	2.8997
60 min	2880	17.8000	0.6000	7.0736	2.9892
120 min	2880	18.9000	0.6000	6.1114	3.0339

Fig. 4. The specific data information and the location of the wind farm.

4.2 Evaluation metrics

The establishment of evaluation metrics is an essential component of our research, hence we applied four indicators to evaluate the performance of **PF**, and also employed three indicators for **IF**. The **SDE**, **RMSE**, **MAE**, **MAPE** are adopted as the criteria to assess the performance of **PF**, the **PICP**, **AWD**, and **AIS** are applied to verify the performance of **IF**. Specifically, it is emphasized that **PICP** and prediction interval (PI) are intrinsically conflicting, the PICP increases, the width of PI also increases, whereas a wide PI is typically pointless. Therefore, the AIS is the comprehensive metric for evaluating the **IF** capability, by rewarding valid PI and penalizing invalid PI. The definitions of these metrics and the formulas are detailed in Table 5.

Table 5
evaluation indicators applied in our study.

Metric	Definition	Expression
SDE	The standard deviation of *N* forecasting error	$\mathbf{SDE}=\sqrt{\frac{1}{K}\times\sum_{i=1}^{K}(P_i^{act}-P_i^{pre})^2}$
RMSE	The square root of the average error squares	$\mathbf{RMSE}=\sqrt{\frac{1}{K}\times\sum_{i=1}^{K}\left(P_i^{act}-P_i^{pre}\right)^2}$
MAE	The mean absolute error of *N* forecasting results	$\mathbf{MAE}=\frac{1}{K}\sum_{i=1}^{K}\left\|P_i^{act}-P_i^{pre}\right\|$
MAPE	The average absolute percentage error	$\mathbf{MAPE}=\frac{1}{K}\sum_{i=1}^{K}\left\|\frac{P_i^{act}-P_i^{pre}}{P_i^{act}}\right\|\times 100\%$
$\mathbf{UB}_i^{(\alpha)}$	The upper bound of the prediction interval	$\mathbf{UB}_i^{(\alpha)}=P_i^{pre}+\rho*Dist_{1-(\alpha/2)}*\sqrt{\frac{var(P^{pre})}{K}}$
$\mathbf{LB}_i^{(\alpha)}$	The lower bound of the prediction interval	$\mathbf{LB}_i^{(\alpha)}=P_i^{pre}-\rho*Dist_{1-(\alpha/2)}*\sqrt{\frac{var(P^{pre})}{K}}$
$\mathbf{PICP}^{(\alpha)}$	Prediction interval coverage probability	$\mathbf{PICP}^{(\alpha)}=\sum_{i=1}^{K}\mathrm{C}_i^{(\alpha)}\Big/K,\mathrm{C}_i^{(\alpha)}=\begin{cases}1, & P_i^{act}\in[LB_i^{(\alpha)},UB_i^{(\alpha)}]\\0, & P_i^{act}\notin[LB_i^{(\alpha)},UB_i^{(\alpha)}]\end{cases}$
$\mathbf{AWD}^{(\alpha)}$	Accumulated width deviation	$\mathbf{AWD}^{(\alpha)}=\sum_{i=1}^{K}\mathrm{AWD}_i^{(\alpha)}\Big/\mathrm{AR}$ $\mathrm{AWD}_i^{(\alpha)}=\begin{cases}(LB_i^{(\alpha)}-P_i^{act})/(UB_i^{(\alpha)}-LB_i^{(\alpha)}), & P_i^{act}<LB_i^{(\alpha)}\\0, & P_i^{act}\in[LB_i^{(\alpha)},UB_i^{(\alpha)}]\\(P_i^{act}-UB_i^{(\alpha)})/(UB_i^{(\alpha)}-LB_i^{(\alpha)}), & P_i^{act}>UB_i^{(\alpha)}\end{cases}$
$\mathbf{AIS}^{(\alpha)}$	The average interval score of the model	$\mathbf{AIS}^{(\alpha)}=\sum_{i=1}^{K}\mathrm{S}^{(\alpha)}\Big/K$ $\mathrm{S}^{(\alpha)}=\begin{cases}-2\alpha\gamma_i^{\alpha}-4\left(LB_i^{(\alpha)}-P_i^{act}\right), & P_i^{act}<LB_i^{(\alpha)}\\-2\alpha\gamma_i^{\alpha}, & P_i^{act}\in[LB_i^{(\alpha)},UB_i^{(\alpha)}]\\-2\alpha\gamma_i^{\alpha}-4\left(P_i^{act}-UB_i^{(\alpha)}\right), & P_i^{act}>UB_i^{(\alpha)}\end{cases}$

Note: p_i^{act}, p_i^{pre} and K refer to the actual value, predicted value, and sample size, respectively, α denotes the significance level, ρ represents the width adjusted coefficient for the PI, which is fixed to 0.6 in our study. $Dist_{1-(\alpha/2)}$ refers to the quantile of DFs, AR means the range of actual values and γ_i^{α} can be defined as $\gamma_i^{\alpha}=UB_i^{(\alpha)}-LB_i^{(\alpha)}$.

4.3 Experimental settings and presentation

Five experiments are implemented to validate the superiority of our proposed WSCFS in this section. The detailed parameter settings of tested models involved in the experiments will be presented in Table 6.

Table 6
Parameter details of key models.

Model	Parameter	Implication	Value or Function
BPNN	n_h	Number of nodes in the hidden layer	10
	n_i	Number of nodes in the input layer	5
	l_r	Learning rate	0.0001
	A_f	Activation function	Sigmoid
ELM	n_h	Number of nodes in the hidden layer	16
	n_i	Number of nodes in the input layer	5
	A_f	Activation function	Sin
ENN	n_h	Number of nodes in the hidden layer	15
	n_i	Number of nodes in the input layer	5
	A_f	Activation function	Tansig

GRNN	$\boldsymbol{n}_i$	Number of nodes in the input layer	5
	$\boldsymbol{s}_p$	Spread	1
LSTM	$\boldsymbol{n}_h$	Number of nodes in the hidden layer	15
	$\boldsymbol{n}_i$	Number of nodes in the input layer	5
	$\boldsymbol{e}_h$	Training number	100
	$\boldsymbol{o}_p$	Optimization algorithm	Adam
GRU	$\boldsymbol{n}_h$	Number of nodes in the hidden layer	64
	$\boldsymbol{n}_i$	Number of nodes in the input layer	5
	$\boldsymbol{e}_h$	Training number	200
	$\boldsymbol{o}_p$	Optimization algorithm	Adam
TCN	$\boldsymbol{n}_h$	Number of nodes in the hidden layer	4
	$\boldsymbol{n}_i$	Number of nodes in the input layer	5
	$\boldsymbol{e}_h$	Training number	500
	$\boldsymbol{o}_p$	Optimization algorithm	Adam
MOAO	$\boldsymbol{i}_{er}$	Number of iterations	200
	$\boldsymbol{a}_{ch}$	Archive size	500
	$\boldsymbol{n}_{er}$	Number of individuals	40
MOALO、MODA、MOGOA	$\boldsymbol{i}_e$	Number of iterations	200
	$\boldsymbol{a}_c$	Archive size	500
	$\boldsymbol{n}_r$	Number of individuals	40

Experiment I: Model selection based on OBMS strategy

We conducted this experiment to select five optimal models from seven benchmark models to improve forecasting effectiveness in our proposed WSCFS, which are FIG-BP ($\boldsymbol{n}_h$=10, $\boldsymbol{n}_i$=5, $\boldsymbol{l}_r$=0.0001), FIG-ELM ($\boldsymbol{n}_h$=16, $\boldsymbol{n}_i$=5), FIG-ENN ($\boldsymbol{n}_h$=15, $\boldsymbol{n}_i$=5), FIG-GRNN ($\boldsymbol{n}_i$=5, $\boldsymbol{s}_p$=1), FIG-LSTM (n_h=15, $\boldsymbol{n}_i$=5, e_h=100), FIG-TCN ($\boldsymbol{n}_h$=4, $\boldsymbol{n}_i$=5, $\boldsymbol{e}_h$=1000) and FIG-GRU ($\boldsymbol{n}_h$=64, $\boldsymbol{n}_i$=5, $\boldsymbol{e}_h$=200) respectively. Detailed results of the OBMS strategy are shown in Table 7. More specifics regarding this experiment are as follows:

(a) For 30-min intervals, in the one-step forecasting, we can notice that FIG-ELM has the best forecasting performance at Site 1 with the **CEM** value of 0.0000, and FIG-LSTM achieves the most satisfactory results at Site 2 with the **CEM** value of 0.0000. As the forecasting steps increase, we can find that FIG-BP, FIG-ELM, FIG-ENN, FIG-LSTM and FIG-GRU always yield desirable forecasting results.
(b) For 60-min intervals, the optimal models are similar to the 30-min intervals, which are FIG-BP, FIG-ELM, FIG-ENN, FIG-LSTM and FIG-GRU. At Site 2, we can find that FIG-ENN achieves the optimal forecasting results in one-step and multi-step forecasting with the $\mathbf{CEM}_{site\,2}^{1-step} = 0.0023$, $\mathbf{CEM}_{site\,2}^{2-step} = 0.0087$ and $\mathbf{CEM}_{site\,2}^{3-step} = 0.0003$ respectively.
(c) For 120-min intervals, we observe that the optimal models vary for different sites. At Site 2, the optimal models remain the same as the previous 30-minute intervals and 60-minute intervals, but at Site 1, the optimal benchmark models change to FIG-BP, FIG-ELM, FIG-ENN, FIG-GRNN, FIG-LSTM. This indicates that as the site and time intervals change, the benchmark models will also change, which implies the OBMS strategy can be applied to different forecasting step.

Remark. According to the OBMS strategy, we select the optimal models based on **CEM**, which can significantly enhance the forecasting performance of the WSCFS.

Table 7
The CEM results of optimal benchmark model selection.

Datasets	Models	30 min interval			60 min interval			120 min interval		
		1-Step	2-Step	3-Step	1-Step	2-Step	3-Step	1-Step	2-Step	3-Step
Site 1	FIG-BP	**0.1002**	**0.0465**	**0.0456**	**0.1011**	**0.0070**	**0.0366**	**0.0943**	**0.2113**	**0.3353**
	FIG-ELM	**0.0000**	**0.0300**	**0.0429**	**0.0368**	**0.0174**	**0.0253**	**0.0029**	**0.1660**	**0.2486**
	FIG-ENN	**0.0444**	**0.0369**	**0.0491**	**0.0000**	**0.0195**	**0.0000**	**0.0611**	**0.0000**	**0.0613**
	FIG-GRNN	0.6949	0.1595	0.2125	0.5893	0.1269	0.2272	**0.4752**	**0.3644**	**0.4260**
	FIG-LSTM	**0.1018**	**0.0111**	**0.0000**	**0.1249**	**0.0282**	**0.0094**	**0.1410**	**0.2741**	**0.3374**
	FIG-TCN	0.9992	1.0000	1.0000	1.0000	0.8016	1.0000	1.0000	1.0000	1.0000
	FIG-GRU	**0.1099**	**0.0903**	**0.0606**	**0.2799**	**0.0474**	**0.1545**	0.4775	0.5779	0.5320
Site 2	FIG-BP	**0.0187**	**0.0422**	**0.0865**	**0.0495**	**0.0356**	**0.0291**	**0.1617**	**0.1224**	**0.1831**
	FIG-ELM	**0.0086**	**0.0152**	**0.0226**	**0.0144**	**0.0097**	**0.0167**	**0.1154**	**0.0738**	**0.1638**
	FIG-ENN	**0.0214**	**0.0222**	**0.0594**	**0.0023**	**0.0087**	**0.0003**	**0.0571**	**0.0000**	**0.0000**
	FIG-GRNN	0.5285	0.2965	0.1737	0.6149	0.2582	0.1525	0.4534	0.3342	0.4825
	FIG-LSTM	**0.0000**	**0.0454**	**0.1543**	**0.1330**	**0.0696**	**0.0612**	**0.0108**	**0.1066**	**0.2166**
	FIG-TCN	1.0000	1.0000	1.0000	1.0000	1.0000	0.7917	1.0000	1.0000	0.7672
	FIG-GRU	**0.0937**	**0.1084**	**0.0309**	**0.3348**	**0.2184**	**0.0261**	**0.2937**	**0.1686**	**0.1466**

Note: The above table gives the optimal model options under different situations based on $\boldsymbol{CEM}_i^* = \left(0.25 * \boldsymbol{SDE}_i^* + 0.25 * \boldsymbol{RMSE}_i^* + 0.25 * \boldsymbol{MAE}_i^* + 0.25 * \boldsymbol{MAPE}_i^*\right)$.

The **CEM** values for the optimal benchmark models for different situations are bolded.

Experiment II: Comparison to other benchmark models based on FIG technique

The purpose of this experiment aims to verify the WSCFS owns superior forecasting performance and effectiveness compared to seven benchmark models. The detailed results of this experiment are presented in Table 8, 9. More specifics regarding this experiment are as follows:

(a) For 30-min intervals, the WSCFS ($\boldsymbol{i}_{er}$=200, $\boldsymbol{a}_{ch}$=*500*, $\boldsymbol{n}_{er}$=*40*) shows the best performance in two sites. The values of $\mathbf{SDE}_{site\,1}^{1-step}$, $\mathbf{RMSE}_{site\,1}^{1-step}$, $\mathbf{MAE}_{site\,1}^{1-step}$, $\mathbf{MAPE}_{site\,1}^{1-step}$ are 0.4912, 0.4897, 0.3821 and 4.1216% respectively. Meanwhile, the WSCFS still achieves the optimal forecasting performance in Site 2 in all cases. For instance, the $\mathbf{MAPE}_{site\,2}^{1-step}$ is 5.0643%, which is decreased by $\bar{\bar{\mathbf{D}}}_{FIG\text{-}BP}^{1-step} = 2.4505\%$, $\bar{\bar{\mathbf{D}}}_{FIG\text{-}ELM}^{1-step} = 2.4204\%$, $\bar{\bar{\mathbf{D}}}_{FIG\text{-}ENN}^{1-step} = 2.4737\%$, $\bar{\bar{\mathbf{D}}}_{FIG\text{-}GRNN}^{1-step} = 3.2359\%$, $\bar{\bar{\mathbf{D}}}_{FIG\text{-}LSTM}^{1-step} = 2.4422\%$, $\bar{\bar{\mathbf{D}}}_{FIG\text{-}TCN}^{1-step} = 4.8343\%$ and $\bar{\bar{\mathbf{D}}}_{FIG\text{-}GRU}^{1-step} = 2.5424\%$ respectively. Take one-step at 30-min intervals for instance, the comparative results are depicted in Fig. 5.

(b) For 60-min intervals, our proposed WSCFS demonstrates the best-evaluated metric, which are $\mathbf{SDE}_{site\,1}^{1-step} = 0.9801$, $\mathbf{RMSE}_{site1}^{1-step} = 0.9050$, $\mathbf{MAE}_{site\,1}^{1-step} = 0.7219$ and $\mathbf{MAPE}_{site\,1}^{1-step} = 13.1232\%$ respectively. At Site 2, the WSCFS as before performs better than the single models in all situations.

(c) The situation for the 120-minute intervals is analogous to the 30-minute and 60-minute intervals. The WSCFS still achieves the best forecasting effectiveness and stability in any situation.

Remark. According to the analysis of experiment two, the performance of our proposed WSCFS outshines the benchmark models.

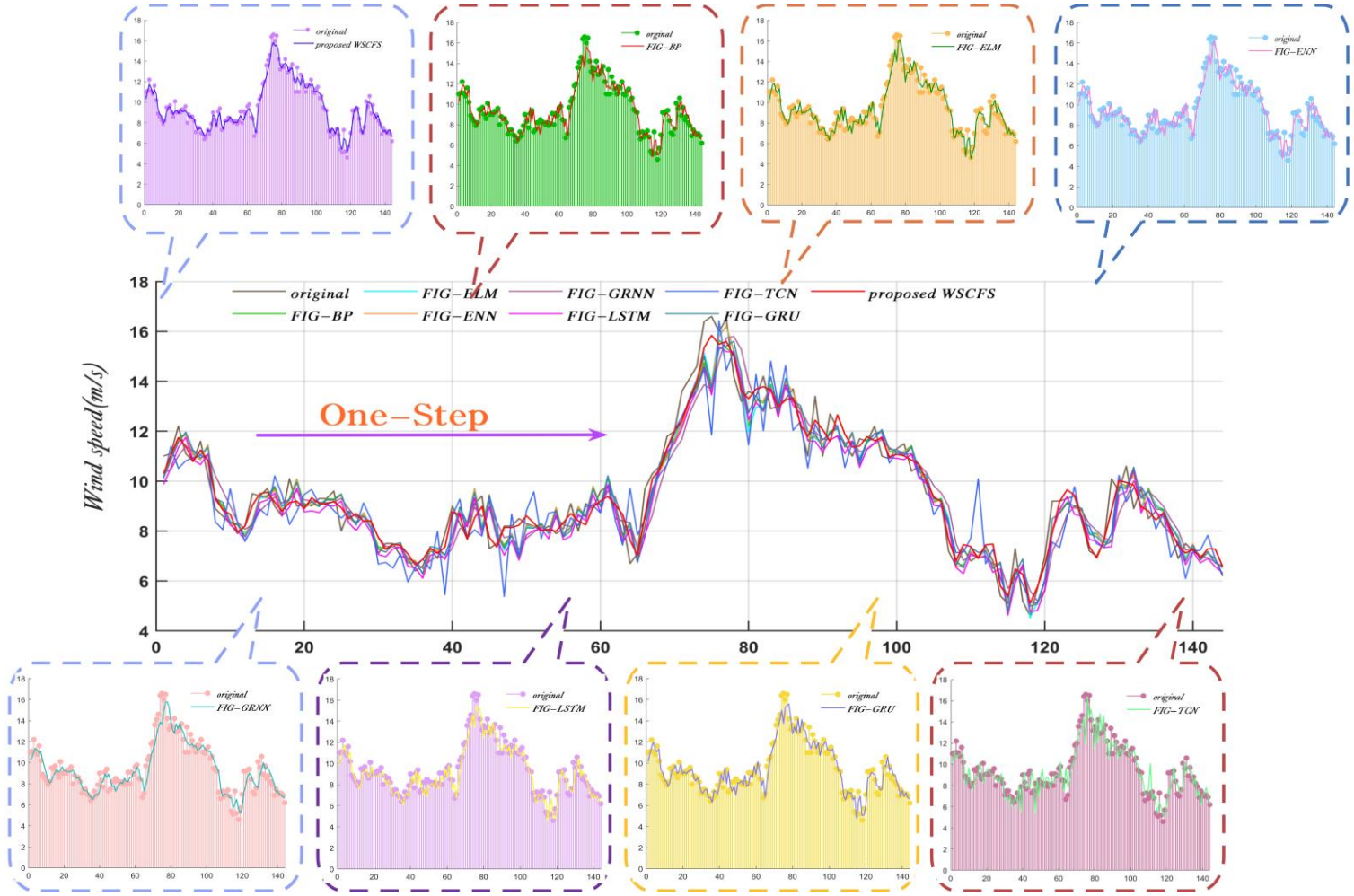


Fig. 5. A comparison of our proposed WSCFS and benchmark models in Site1 at 30-min intervals.

Table 8
Predictive performance comparison of our proposed WSCFS and benchmark models in Site1.

Interval	Models	SDE			RMSE			MAE			MAPE (%)		
		Step-1	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**
30 min	FIG-BP	0.8062	1.1035	1.3198	0.8047	1.1082	1.3285	0.6174	0.8694	1.0126	6.7328	9.5449	11.0923
	FIG-ELM	0.7848	1.1009	1.3500	0.7842	1.0995	1.3493	0.6039	0.8646	1.0408	6.5955	9.5710	11.3828
	FIG-ENN	0.7851	1.1299	1.3515	0.7843	1.1302	1.3514	0.6056	0.8802	1.0376	6.6511	9.6962	11.3408
	FIG-GRNN	0.9075	1.1991	1.4242	0.9078	1.2004	1.4281	0.7089	0.9241	1.0842	7.6451	10.0346	11.8657
	FIG-LSTM	0.8041	1.1013	1.3618	0.8453	1.1310	1.3606	0.6619	0.8872	1.0373	7.0978	9.7167	11.4604
	FIG-TCN	1.0638	1.6504	1.5563	1.0736	1.6544	1.5514	0.8077	1.3115	1.2024	8.7643	14.6475	13.7019
	FIG-GRU	0.8070	1.1315	1.3591	0.8050	1.1294	1.3618	0.6282	0.8957	1.0406	6.8732	9.9472	11.3816
	Proposed WSCFS	**0.4912**	**0.6276**	**0.8776**	**0.4897**	**0.6261**	**0.8757**	**0.3821**	**0.4980**	**0.6847**	**4.1216**	**5.5656**	**7.7479**
60 min	FIG-BP	1.2699	2.0497	2.4985	1.2661	2.0481	2.4903	0.9670	1.5385	1.8760	18.9698	30.6460	36.8597
	FIG-ELM	1.2720	2.0214	2.5468	1.2708	2.0188	2.5422	0.9501	1.5279	1.9547	17.6122	29.5009	36.9795
	FIG-ENN	1.2333	2.0477	2.5080	1.2360	2.0499	2.5047	0.9114	1.5482	1.9247	16.5335	29.7458	36.3279
	FIG-GRNN	1.4661	2.1386	2.7406	1.4649	2.1345	2.7479	1.1193	1.6859	2.1801	21.9121	32.8771	45.4528
	FIG-LSTM	1.2748	2.0887	2.5819	1.2977	2.0819	2.5797	0.9981	1.5850	1.9667	19.9990	32.2531	37.5999
	FIG-TCN	1.7106	2.6897	3.1287	1.7056	2.6824	:3.1179	1.3266	2.0074	2.4072	24.6015	41.4330	45.8778
	FIG-GRU	1.4490	2.0823	2.6452	1.4566	2.0774	2.6477	1.0698	1.5899	2.0609	22.4586	32.6330	43.4168
	Proposed WSCFS	**0.9081**	**1.6160**	**2.1641**	**0.9050**	**1.6120**	**2.1616**	**0.7219**	**1.2466**	**1.6603**	**13.1232**	**23.8782**	**30.7028**
120 min	FIG-BP	0.8977	1.3494	1.5166	0.8982	1.4574	1.7104	0.7192	1.2002	1.4509	19.6284	36.1213	44.7397
	FIG-ELM	0.8444	1.3315	1.5290	0.8809	1.4500	1.7119	0.7111	1.2277	1.4513	19.4964	36.6110	44.5304
	FIG-ENN	0.8512	1.2607	1.4510	0.9028	1.35991	1.6212	0.7431	1.1437	1.3739	20.9691	33.3771	41.3651
	FIG-GRNN	0.9256	1.3289	1.5406	0.9814	1.4568	1.7479	0.8143	1.2545	1.4801	23.7753	37.6344	45.4528
	FIG-LSTM	0.8698	1.3671	1.5468	0.9184	1.4873	1.8697	0.7382	1.2440	1.5497	21.1383	37.7367	50.2111
	FIG-TCN	1.2860	2.0066	2.2928	1.3231	1.6767	2.4660	1.0214	1.6767	1.9543	27.6763	47.1806	56.9219
	FIG-GRU	1.0123	1.5891	1.7543	1.0580	1.7658	1.8734	0.8326	1.4213	1.5416	23.3034	43.8049	47.5398
	Proposed WSCFS	**0.6217**	**1.0634**	**1.2514**	**0.6252**	**1.0942**	**1.3649**	**0.4803**	**0.8890**	**1.1459**	**12.7251**	**24.3349**	**31.9833**

Note: The above table specifically shows the **PF** evaluation index values of our proposed WSCFS and seven benchmark models in Site 1. The lower the values of the evaluation index, the better the performance of the model. The optimal index values are bolded.

Table 9
Predictive performance comparison of our proposed WSCFS and benchmark models in Site 2.

Interval	Models	SDE			RMSE			MAE			MAPE (%)		
		Step-1	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**
30 min	FIG-BP	0.9115	1.2259	1.5005	0.9131	1.2294	1.5042	0.6875	0.9528	1.1747	7.5184	10.7117	13.1650
	FIG-ELM	0.9068	1.2277	1.4933	0.9072	1.2313	1.4955	0.6808	0.9496	1.1602	7.4847	10.6360	13.0312
	FIG-ENN	0.9176	1.2253	1.5009	0.9178	1.2283	1.5018	0.6887	0.9530	1.1717	7.5380	10.6905	13.1745
	FIG-GRNN	0.9983	1.2852	1.4866	1.0039	1.2923	1.4945	0.7653	1.0048	1.1518	8.3002	11.0896	12.7831
	FIG-LSTM	0.9135	1.2127	1.5181	0.9138	1.2625	1.5514	0.6891	0.9728	1.2084	7.5065	10.6369	13.3217
	FIG-TCN	1.1497	1.4709	1.7814	1.1458	1.4717	1.7752	0.8763	1.1567	1.4046	9.8986	13.1590	16.0614
	FIG-GRU	0.9195	1.2403	1.4751	0.9164	1.2362	1.5046	0.6914	0.9721	1.1699	7.6067	10.9455	12.9143
	Proposed WSCFS	**0.5635**	**0.8499**	**1.0545**	**0.5619**	**0.8472**	**1.0538**	**0.4539**	**0.6318**	**0.8210**	**5.0643**	**7.1645**	**9.4055**
60 min	FIG-BP	1.2130	1.9611	2.4642	1.2123	1.9575	2.4559	0.9009	1.4731	1.8103	17.9064	30.6906	37.6699
	FIG-ELM	1.2211	1.9770	2.4517	1.2209	1.9775	2.4456	0.9103	1.4916	1.8072	18.2885	30.6571	37.5283
	FIG-ENN	1.2032	1.9752	2.4572	1.2012	1.9705	2.4500	0.8842	1.4753	1.8351	17.2521	29.9598	36.8603
	FIG-GRNN	1.3968	2.0634	2.5146	1.3965	2.0597	2.5095	1.0636	1.5234	1.8992	22.9530	32.9678	40.8140
	FIG-LSTM	1.2474	1.9646	2.4683	1.2590	1.9688	2.4622	0.9359	1.4811	1.8381	19.2904	32.0748	39.4295
	FIG-TCN	1.3965	2.3897	2.7959	1.3956	2.4035	2.8663	1.1189	1.8141	2.2569	22.3284	37.5376	48.9463
	FIG-GRU	1.3552	2.0663	2.4612	1.3684	2.0869	2.4542	1.0148	1.5375	1.8223	22.7513	34.3327	37.7232
	Proposed WSCFS	**1.0360**	**1.6704**	**2.4261**	**1.0333**	**1.6647**	**2.4219**	**0.7805**	**1.2486**	**1.7404**	**15.2888**	**24.0599**	**33.7250**
120 min	FIG-BP	0.8027	1.1595	1.3449	0.8437	1.2877	1.5640	0.6882	1.0730	1.3112	19.9194	34.5306	43.7706
	FIG-ELM	0.7652	1.1565	1.3432	0.8148	1.2833	1.5749	0.6661	1.0726	1.3132	19.6929	34.6125	44.2253
	FIG-ENN	0.7622	1.1402	1.3342	0.7987	1.2722	1.5503	0.6546	1.0515	1.2843	18.9548	33.2935	42.6051
	FIG-GRNN	0.8557	1.2360	1.3993	0.9308	1.4049	1.6573	0.7947	1.1741	1.3901	25.2919	39.4474	47.5824
	FIG-LSTM	0.7657	1.2126	1.3418	0.7736	1.2797	1.5319	0.6296	1.0724	1.2893	17.5104	34.3864	43.5440
	FIG-TCN	1.2323	1.6919	2.2928	1.3133	1.8141	2.4660	1.0483	1.4589	1.9543	29.2509	42.9915	56.9219
	FIG-GRU	0.7845	1.1671	1.3289	0.8201	1.3231	1.5788	0.6809	1.1131	1.3190	19.3648	36.2661	44.1340
	Proposed WSCFS	**0.5757**	**0.9825**	**1.2432**	**0.5812**	**0.9935**	**1.3735**	**0.4537**	**0.7256**	**1.1275**	**13.1076**	**24.7976**	**36.6291**

Note: The above table presents the **PF** evaluation index values of our proposed WSCFS and seven benchmark models in Site 2. The smaller the values of the evaluation index, the better the performance of the model. The optimal index values are bolded.

Experiment III: Compared to the individual models without pre-treatment techniques

The purpose of this experiment is to verify the effectiveness of our WSCFS compared with other classical individual models. The evaluation indicators for these models are presented in Table 10, 11. The details of the experimental finds are illustrated below:

(a) For 30-min intervals, the WSCFS always outperformed the conventional neural networks. the values of the evaluation indicators are $\mathbf{SDE}_{site\,2}^{1-step} = 0.5635$, $\mathbf{RMSE}_{site\,2}^{1-step} = 0.5619$, $\mathbf{MAE}_{site\,2}^{1-step} = 0.4539$ and $\mathbf{MAPE}_{site\,2}^{1-step} = 5.0630\%$ respectively. In multi-step forecasting, the WSCFS with the $\mathbf{MAPE}_{site\,2}^{2-step} = 7.1645\%$ and $\mathbf{MAPE}_{site\,2}^{3-step} = 9.4055\%$. By contrast, the ranks of the BP, ELM, ENN, GRNN, LSTM, TCN and GRU are $\mathbf{R}_{site\,2}^{2-step} = [2, 6, 4, 3, 5, 7, 1]$. Take Site 2 at 30-min intervals for example, the comparative results are depicted in Fig. 6.

(b) For 60-min intervals, our proposed WSCFS still possesses satisfactory effectiveness concerning the individual models in any situation, the error metrics are $\mathbf{SDE}_{site\,1}^{2-step} = 1.6160$, $\mathbf{RMSE}_{site\,1}^{2-step} = 1.6120$, $\mathbf{MAE}_{site\,1}^{2-step} = 1.2466$ and $\mathbf{MAPE}_{site\,1}^{2-step} = 23.8782\%$ respectively. In contrast, the remaining seven single models show a significantly inferior forecasting performance compared to our proposed WSCFS in all cases.

(c) For 120-min intervals, the most satisfactory model continues to be the WSCFS in all cases, the **MAPE** metrics are $\mathbf{MAPE}_{site\,2}^{1-step} = 13.1076\%$, $\mathbf{MAPE}_{site\,2}^{2-step} = 24.7976\%$ and $\mathbf{MAPE}_{site\,2}^{3-step} = 36.6291\%$ respectively, the indicators are declined by $\bar{\bar{\mathbf{D}}}_{BP}^{1-step} = 16.5796\%$, $\bar{\bar{\mathbf{D}}}_{ELM}^{1-step} = 18.7527\%$, $\bar{\bar{\mathbf{D}}}_{ENN}^{1-step} = 15.0606\%$, $\bar{\bar{\mathbf{D}}}_{GRNN}^{1-step} = 16.3941\%$, $\bar{\bar{\mathbf{D}}}_{LSTM}^{1-step} = 15.6194\%$, $\bar{\bar{\mathbf{D}}}_{TCN}^{1-step} = 21.2055\%$ and $\bar{\bar{\mathbf{D}}}_{GRU}^{1-step} = 19.3815\%$ respectively.

Remark. As shown in this experiment, our proposed WSCFS still achieves the best forecasting effectiveness in all situations.

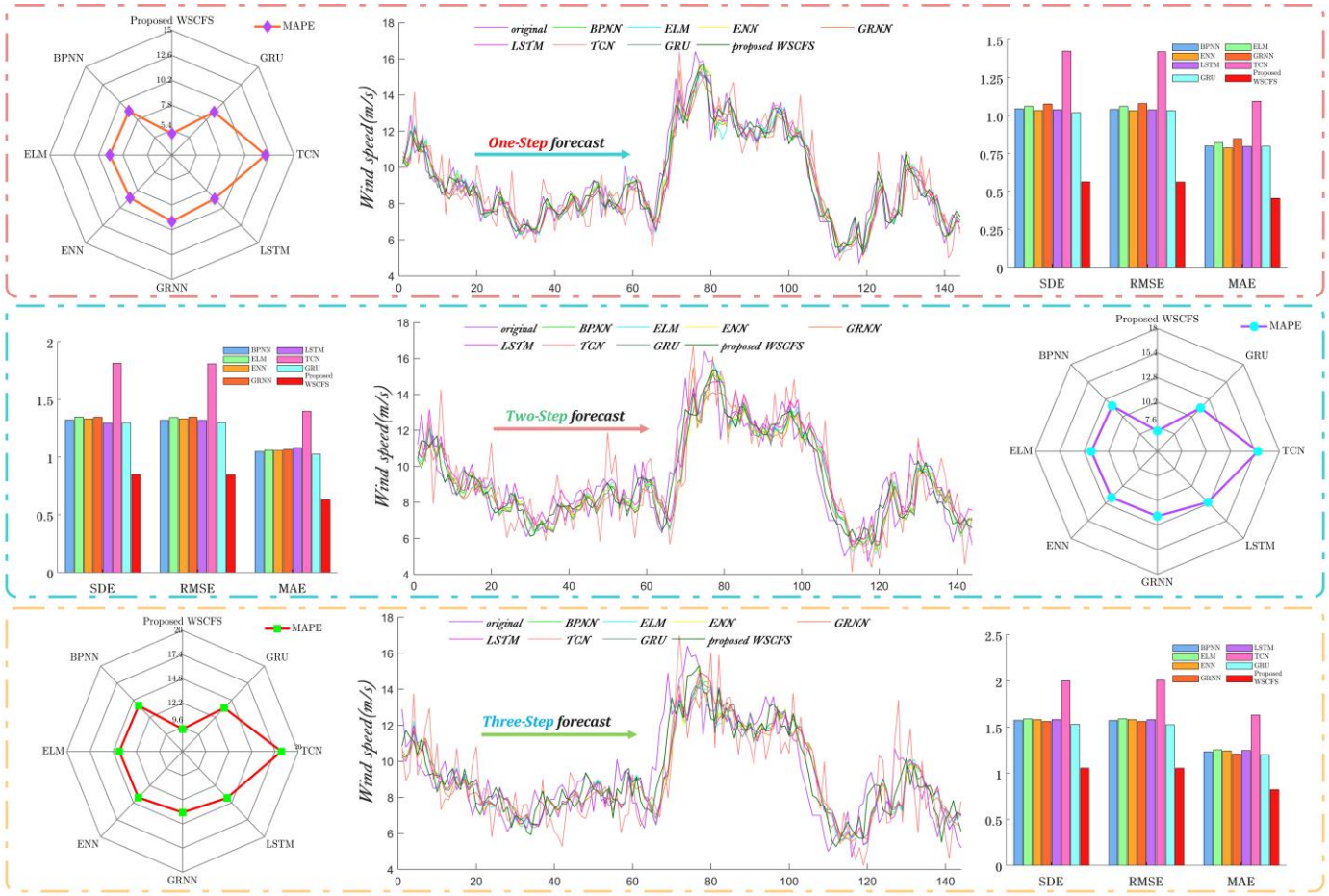


Fig. 6. The comparative results of our proposed WSCFS and classical models in Site 2 at 30-min intervals.

Table 10
Comparative results of our proposed WSCFS with the classical models in Site1.

Interval	Models	SDE			RMSE			MAE			MAPE (%)		
		Step-1	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**
30 min	BP	0.9672	1.2362	1.4676	0.9678	1.2366	1.4648	0.7492	0.9614	1.1360	8.1852	10.6598	12.5024
	ELM	1.0214	1.2452	1.4684	1.0183	1.2436	1.4679	0.7816	0.9634	1.1191	8.6749	10.6791	12.2839
	ENN	0.9302	1.2154	1.4690	0.9288	1.2144	1.4678	0.7271	0.9451	1.1168	8.0287	10.4680	12.3047
	GRNN	1.0401	1.3136	1.5488	1.0437	1.3180	1.5583	0.8090	1.0065	1.1768	8.8197	11.0245	12.8097
	LSTM	0.9422	1.2134	1.4382	0.9731	1.2344	1.5240	0.7618	0.9551	1.1358	8.2555	10.4712	12.0484
	TCN	1.2575	1.7362	1.9675	1.2611	1.7304	1.9608	0.9594	1.3362	1.4752	10.6185	14.9521	15.9010
	GRU	0.9573	1.2356	1.4445	0.9539	1.2318	1.4409	0.7321	0.9514	1.1256	8.1722	10.6849	12.5684
	Proposed WSCFS	**0.4912**	**0.6276**	**0.8776**	**0.4897**	**0.6261**	**0.8757**	**0.3821**	**0.4980**	**0.6847**	**4.1216**	**5.5656**	**7.7479**
60 min	BP	2.0145	2.3730	2.8610	2.0154	2.3681	2.8595	1.4811	1.8783	2.2707	27.4307	37.1303	42.0051
	ELM	1.8467	2.8284	2.8695	1.8468	2.8186	2.8755	1.4145	2.2282	2.3549	26.2315	40.4399	43.0242
	ENN	1.8972	2.3510	2.7871	1.8940	2.3512	2.7917	1.4237	1.8457	2.2269	26.2529	34.2205	40.0391
	GRNN	1.9849	2.4703	2.6366	1.9851	2.4773	2.6333	1.5126	1.9774	2.1000	27.9322	36.7102	40.3824
	LSTM	1.9523	2.3975	2.7864	1.9474	2.3909	2.7813	1.4854	1.8763	2.2329	27.9202	36.8345	44.0615
	TCN	2.2119	2.6852	3.3151	2.2046	2.6773	3.3053	1.7040	2.1654	2.6705	31.7809	40.3754	50.2771
	GRU	1.9335	2.4467	2.8681	1.9288	2.4389	2.8896	1.4685	1.8985	2.2867	28.1648	37.0898	46.9865
	Proposed WSCFS	**0.9081**	**1.6160**	**2.1641**	**0.9050**	**1.6120**	**2.1616**	**0.7219**	**1.2466**	**1.6603**	**13.1232**	**23.8782**	**30.7028**
120 min	BP	1.2116	1.4326	1.5833	1.2757	1.5886	1.8142	1.0182	1.3481	1.5526	29.1523	41.0618	47.8477
	ELM	1.2639	1.5949	1.6962	1.3627	1.7981	1.9564	1.0970	1.5360	1.6681	31.9733	46.3124	52.1950
	ENN	1.2090	1.4577	1.5915	1.2596	1.5833	1.7934	1.0024	1.3634	1.5468	28.0191	40.0604	47.1335
	GRNN	1.2238	1.4446	1.5524	1.2816	1.5552	1.8098	1.0278	1.3224	1.5608	29.4321	39.3955	48.1273
	LSTM	1.1836	1.4724	1.5962	1.2548	1.5632	1.8822	1.0102	1.3086	1.5945	28.8072	38.5880	49.8817
	TCN	1.5563	1.9943	2.2933	1.6070	2.1757	2.4708	1.2991	1.7482	2.0002	35.6633	49.8283	56.3275
	GRU	1.2357	1.6747	1.8743	1.2881	1.8350	2.0485	1.0183	1.4828	1.6369	28.1520	44.7245	50.2641
	Proposed WSCFS	**0.6217**	**1.0634**	**1.2514**	**0.6252**	**1.0942**	**1.3649**	**0.4803**	**0.8890**	**1.1459**	**12.7251**	**24.3349**	**31.9833**

Note: The above table gives the detailed **PF** results of the evaluation metrics for our proposed WSCFS and seven classical models in Site 1. The smaller the values of the evaluation metrics, the better the performance of the model. The best evaluation metrics are bolded.

Table 11
Comparative results of our proposed WSCFS with the classical models in Site 2.

Interval	Models	SDE			RMSE			MAE			MAPE (%)		
		Step-1	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**
30 min	BP	1.0445	1.3211	1.5741	1.0415	1.3189	1.5724	0.7989	1.0471	1.2329	8.9736	11.8106	13.9606
	ELM	1.0609	1.3470	1.5872	1.0606	1.3429	1.5896	0.8210	1.0596	1.2532	9.1109	12.0340	14.1043
	ENN	1.0341	1.3312	1.5818	1.0329	1.3314	1.5820	0.7868	1.0559	1.2414	8.8258	11.9053	14.0107
	GRNN	1.0767	1.3451	1.5617	1.0796	1.3458	1.5616	0.8468	1.0652	1.2070	9.4004	11.8507	13.5725
	LSTM	1.0396	1.2947	1.5817	1.0381	1.3189	1.5808	0.7969	1.0801	1.2460	8.9436	12.6059	14.0687
	TCN	1.4254	1.8133	2.0000	1.4208	1.8081	2.0079	1.0943	1.3971	1.6294	12.1940	15.7037	18.1002
	GRU	1.0185	1.2958	1.5292	1.0327	1.2993	1.5244	0.7978	1.0246	1.2016	8.8702	11.5022	13.6272
	Proposed WSCFS	**0.5635**	**0.8499**	**1.0545**	**0.5619**	**0.8472**	**1.0538**	**0.4539**	**0.6318**	**0.8210**	**5.0643**	**7.1645**	**9.4055**
60 min	BP	1.8547	2.3375	2.5822	1.8488	2.3298	2.5752	1.3545	1.7910	2.0041	26.2618	36.5613	41.4711
	ELM	1.7854	2.3756	2.5476	1.7468	2.3673	2.5395	1.4371	1.8069	1.9801	28.0806	36.0594	39.1714
	ENN	1.8348	2.3134	2.5246	1.8286	2.3054	2.5158	1.3291	1.7661	1.9826	25.8142	34.9473	39.3005
	GRNN	1.9498	2.3046	2.5525	1.9438	2.2973	2.5440	1.4742	1.7888	1.9960	29.3929	36.5520	41.2269
	LSTM	1.8677	2.3157	2.5422	1.8613	2.3080	2.5369	1.3804	1.7619	1.9575	26.6025	36.6265	40.1665
	TCN	2.2575	2.7497	3.1538	2.2808	2.7402	3.1634	1.6933	2.1339	2.4260	35.2821	40.4581	50.0731
	GRU	1.9219	2.4768	2.5080	1.9284	2.4933	2.4994	1.4120	1.9158	1.9464	28.2885	41.6243	38.7396
	Proposed WSCFS	**1.0360**	**1.6704**	**2.4261**	**1.0333**	**1.6647**	**2.4219**	**0.7805**	**1.2486**	**1.7404**	**15.2888**	**24.0599**	**33.7250**
120 min	BP	1.1327	1.3534	1.4886	1.1930	1.5002	1.7743	0.9592	1.2809	1.4979	28.8660	41.3772	51.2367
	ELM	1.0854	1.3846	1.4506	1.1468	1.5533	1.7126	0.9371	1.3199	1.4255	28.0806	43.5503	48.0741
	ENN	1.0723	1.3364	1.4384	1.1259	1.4761	1.6769	0.9094	1.2417	1.3965	27.2525	39.8582	46.8440
	GRNN	1.1079	1.3787	1.4919	1.1680	1.5154	1.6879	0.9617	1.2678	1.3781	29.6098	41.1917	47.0450
	LSTM	1.1047	1.3474	1.5280	1.1340	1.4957	1.8071	0.9114	1.2589	1.5114	26.8936	40.4170	51.6688
	TCN	1.5084	1.8416	2.2213	1.5409	1.8978	2.3389	1.2225	1.5650	1.8422	34.6531	46.0031	56.0811
	GRU	1.1345	1.5410	1.8743	1.1891	1.7214	2.0485	0.9829	1.3970	1.6369	29.8600	44.1791	50.2641
	Proposed WSCFS	**0.5757**	**0.9825**	**1.2432**	**0.5812**	**0.9935**	**1.3735**	**0.4537**	**0.7256**	**1.1275**	**13.1076**	**24.7976**	**36.6291**

Note: The above table gives the detailed **PF** results of the evaluation metrics for our proposed WSCFS and seven classical models in Site 2. The smaller the values of the evaluation metrics, the better the performance of the model. The best evaluation metrics are bolded.

Experiment Ⅳ: Comparison of our WSCFS and the combined models employing other optimization algorithms.

This trial is aim to validate the performance of the WSCFS and combined forecasting models (CFM) utilizing other optimization algorithms, which are ALO (i_{er}=200, a_{ch}=*500*, n_{er}=*40*), DA (i_{er}=200, a_{ch}=*500*, n_{er}=*40*) and GOA (i_{er}=200, a_{ch}=*500*, n_{er}=*40*). The calculated evaluation metrics are shown in Table 12 and 13, and more details will be discussed in the following:

(a) For 30-min intervals, there is no wide discrepancy between the CFM employing different optimization algorithms with the average value of $\mathbf{MAPE}_{site1}^{1-step} = 5.4782\%$. The **MAPE** value of our proposed WSCFS is 4.1216% in the same situation, which is the minimum value among all the CFMs. The error metrics are declined by $\bar{\bar{\mathbf{D}}}_{ALO}^{1-step} = 1.1865\%$, $\bar{\bar{\mathbf{D}}}_{DA}^{1-step} = 1.3133\%$ and $\bar{\bar{\mathbf{D}}}_{GOA}^{1-step} = 1.5701\%$ respectively. As for the multi-step prediction, the WSCFS is always the dominant version compared to CFMs employing various optimization algorithms. Take Site 1 at 30-min intervals for instance, the comparative results are depicted in Fig. 7.

(b) For 60-min intervals, the discrepancy between the WSCFS and the CFMs using different optimization algorithms is not significant in one-step forecasting. However, as the forecasting steps increase, this gap is increasing. Taking three-step forecasting in Site1 for instance, our proposed WSCFS owns the smallest $\mathbf{MAPE}_{site1}^{3-step} = 30.7028\%$, simultaneously, the relatively inferior forecasting performance is ALO-CFM with the $\mathbf{MAPE}_{site1}^{3-step} = 31.7808\%$.

(c) As similar to the 60-min intervals, in the one-step forecasting, there is no noticeable performance gap between the WSCFS and the CFMs, the WSCFS still achieves satisfactory forecasting accuracy and stability among all CFMs.

Remark. Compared to CFMs employing different optimization algorithms, the WSCFS still owns the highest effectiveness in all cases. Moreover, the differences in forecasting performance between ALO-CFM, DA-CFM, and GOA-CFM are not significant.

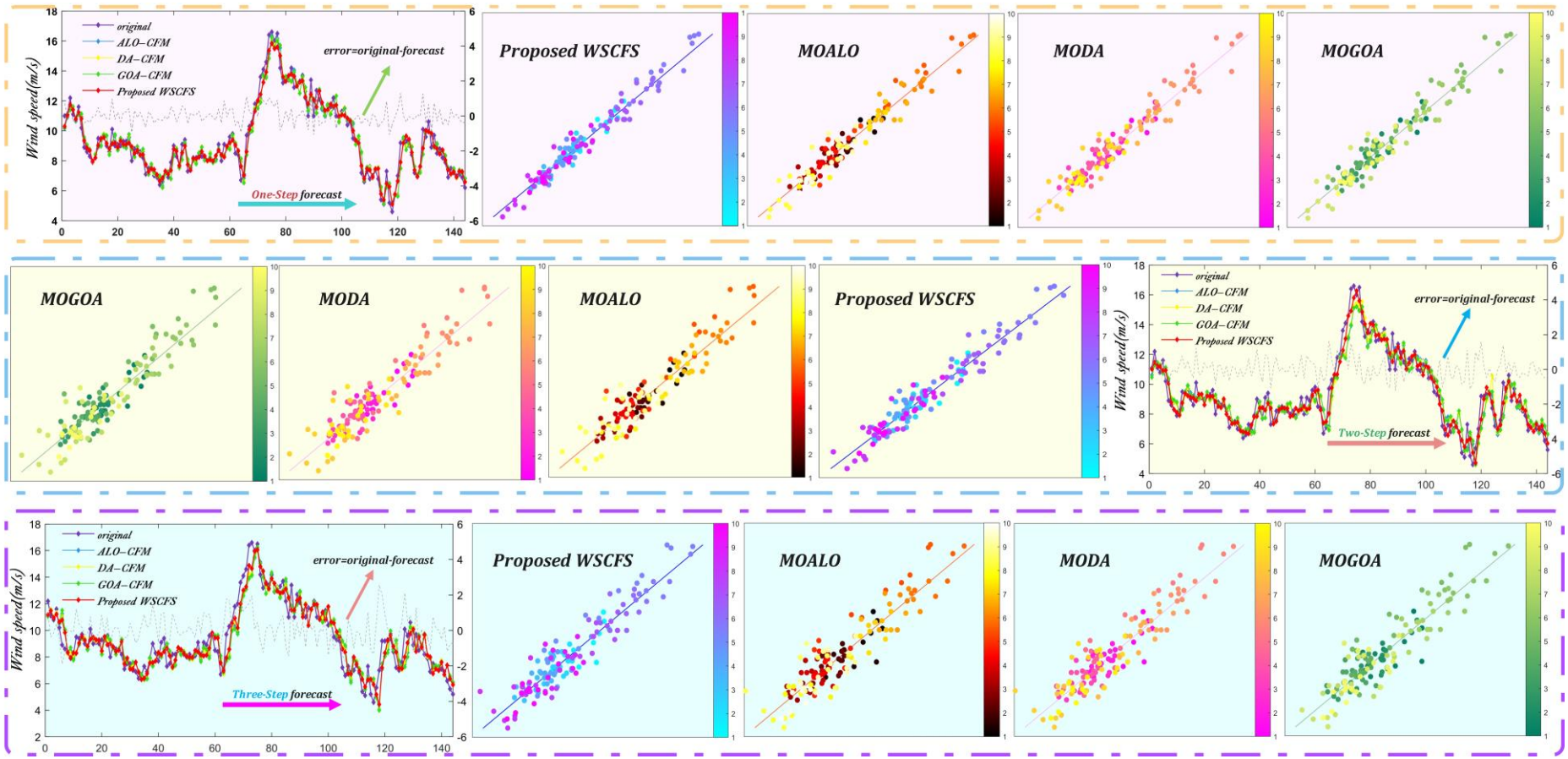


Fig. 7 A comparison of our proposed WSCFS and combined models in Site1 at 30-min interval.

Table 12
Comparative results of our proposed WSCFS and CFMs using different optimization algorithms in Site 1.

Interval	Models	SDE			RMSE			MAE			MAPE (%)		
		Step-1	Step-2	Step-3	Step-1	Step-2	Step-3	Step-1	Step-2	Step-3	Step-1	Step-2	Step-3
30 min	ALO-CFM	0.6177	0.8589	1.0038	0.6162	0.8565	1.0020	0.4802	0.6711	0.7767	5.3081	7.4983	8.7722
	DA-CFM	0.6356	0.9041	0.9740	0.6335	0.9015	0.9721	0.4941	0.7030	0.7628	5.4349	7.8804	8.6308
	GOA-CFM	0.6605	0.9123	1.0577	0.6584	0.9095	1.0553	0.5173	0.7162	0.8105	5.6917	8.0190	9.2020
	Proposed WSCFS	**0.4912**	**0.6276**	**0.8776**	**0.4897**	**0.6261**	**0.8757**	**0.3821**	**0.4980**	**0.6847**	**4.1216**	**5.5656**	**7.7479**
60 min	ALO-CFM	1.1490	1.7509	2.1901	1.1466	1.7474	2.1857	0.8990	1.3279	1.6978	15.7373	25.2472	31.7807
	DA-CFM	1.1492	1.8537	2.2262	1.1457	1.8498	2.2295	0.9084	1.4364	1.8047	15.8979	27.5108	34.7349
	GOA-CFM	1.1325	1.8105	2.4260	1.1313	1.8048	2.4247	0.8911	1.3707	1.8740	15.5690	26.7575	35.1170
	Proposed WSCFS	**0.9081**	**1.6160**	**2.1641**	**0.9050**	**1.6120**	**2.1616**	**0.7219**	**1.2466**	**1.6603**	**13.1232**	**23.8782**	**30.7028**
120 min	ALO-CFM	0.6486	1.1329	1.2788	0.6641	1.1777	1.3753	0.5193	0.9665	1.1628	13.8439	27.3083	33.5137
	DA-CFM	0.6530	1.1062	1.4294	0.6729	1.1362	1.5060	0.5333	0.9374	1.2703	14.5150	25.8559	36.8066
	GOA-CFM	0.6617	1.1218	1.3729	0.6700	1.1682	1.4484	0.5323	0.9568	1.2329	13.9879	26.7779	34.5747
	Proposed WSCFS	**0.6217**	**1.0634**	**1.2514**	**0.6252**	**1.0942**	**1.3649**	**0.4803**	**0.8890**	**1.1459**	**12.7251**	**24.3349**	**31.9833**

Note: The above table shows the detailed contents of our proposed WSCFS and the CFMs using different optimization algorithms in Site 1. The most satisfactory index values are bolded.

Table 13
Comparative results of our proposed WSCFS and CFMs using different optimization algorithms in Site 2.

Interval	Models	SDE			RMSE			MAE			MAPE (%)		
		Step-1	Step-2	Step-3	Step-1	Step-2	Step-3	Step-1	Step-2	Step-3	Step-1	Step-2	Step-3
30 min	ALO-CFM	0.7023	1.0827	1.0945	0.7007	1.0798	1.0927	0.5700	0.8430	0.8485	6.3507	9.6921	9.7308
	DA-CFM	0.6844	1.0952	1.1923	0.6841	1.0945	1.1921	0.5553	0.8601	0.8656	6.1846	9.8395	9.9290
	GOA-CFM	0.6720	0.9199	1.0943	0.6703	0.9172	1.0907	0.5385	0.7132	0.8514	6.0273	8.0603	9.8075
	Proposed WSCFS	**0.5635**	**0.8499**	**1.0545**	**0.5619**	**0.8472**	**1.0538**	**0.4539**	**0.6318**	**0.8210**	**5.0643**	**7.1645**	**9.4055**
60 min	ALO-CFM	1.0549	1.6929	2.4519	1.0514	1.6898	2.4377	0.8183	1.2625	1.8377	16.1666	26.4274	35.0431
	DA-CFM	1.0688	1.7044	2.4542	1.0657	1.7017	2.4484	0.8216	1.2728	1.8898	15.6689	25.0570	33.9303
	GOA-CFM	1.0506	1.6805	2.4496	1.0473	1.6860	2.4498	0.8031	1.2542	1.8456	16.1927	24.9170	35.3620
	Proposed WSCFS	**1.0360**	**1.6704**	**2.4261**	**1.0333**	**1.6647**	**2.4219**	**0.7805**	**1.2486**	**1.7404**	**15.2888**	**24.0599**	**33.7250**
120 min	ALO-CFM	0.7895	1.1319	1.2572	0.7871	1.1943	1.4023	0.6260	0.9403	1.1586	16.6680	27.7421	37.9108
	DA-CFM	0.6796	1.0787	1.2517	0.6896	1.0776	1.4226	0.5603	0.9125	1.1810	15.7409	25.9288	39.1962
	GOA-CFM	0.6754	1.2701	1.2753	0.6862	1.1834	1.4057	0.5589	0.9787	1.1616	16.0133	29.4898	38.2319
	Proposed WSCFS	**0.5757**	**0.9825**	**1.2432**	**0.5812**	**0.9935**	**1.3735**	**0.4537**	**0.7256**	**1.1275**	**13.1076**	**24.7976**	**36.6291**

Note: The above table shows the detailed contents of our proposed WSCFS and the CFMs using different optimization algorithms in Site 2. The most satisfactory index values are bolded.

Experiment V: Comparative results of interval forecast according to point forecast and fitting distribution.

In this experiment, we employ AO and MLE to search the optimal parameters of four DFs. Moreover, based on $\mathbf{R}^2$, the DFs are fitted to the forecasting sequence to determine the optimal distribution in different situations. To validate the supremacy of the WSCFS in terms of **IF**, three CFMs are employed for comparison. The $\mathbf{R}^2$ of the WSCFS are shown in Table 14, 15. Furthermore, it can be noticed that the AO strategy possesses better fitting performance than the MLE. Take one-step for instance, the results of distribution fitting are depicted in Fig. 8. The significance level $\boldsymbol{\alpha}$ is set at 0.05, 0.1 and 0.15, and the detailed results are detailed in Table 16, 17, the specific experimental conclusions are as follows:

(a) For 30-min intervals, the **PICP** of the WSCFS outperforms all CFMs in all situations. Taking Site 1 for example, the **PICP** of WSCFS are $\mathbf{PICP}_{\boldsymbol{1\text{-}step}}^{\alpha=0.05}=100.000$, $\mathbf{PICP}_{\boldsymbol{1\text{-}step}}^{\alpha=0.1}=100.000$ and $\mathbf{PICP}_{\boldsymbol{1\text{-}step}}^{\alpha=0.15}=98.6111$,which are increased by $\tilde{\tilde{\boldsymbol{I}}}_{\boldsymbol{ALO}}^{\alpha=0.05}=4.8611$, $\tilde{\tilde{\boldsymbol{I}}}_{\boldsymbol{ALO}}^{\alpha=0.1}=7.6389$ and $\tilde{\tilde{\boldsymbol{I}}}_{\boldsymbol{ALO}}^{\alpha=0.15}=8.3333$ respectively. Simultaneously, the **AWD** values of the WSCFS are also the lowest in all cases. Take Site 2 for instance, the **AWD** values of the WSCFS are $\mathbf{AWD}_{\boldsymbol{1\text{-}step}}^{\alpha=0.05}=0.0000$, $\mathbf{AWD}_{\boldsymbol{2\text{-}step}}^{\alpha=0.05}=0.0041$ and $\mathbf{AWD}_{\boldsymbol{3\text{-}step}}^{\alpha=0.05}=0.0094$ respectively, which are declined by $\bar{\bar{\boldsymbol{D}}}_{\boldsymbol{DA}}^{1-step}=0.0027$, $\bar{\bar{\boldsymbol{D}}}_{\boldsymbol{DA}}^{2-step}=0.0078$ and $\bar{\bar{\boldsymbol{D}}}_{\boldsymbol{DA}}^{3-step}=0.0035$ respectively. As for **AIS**, it can provide a comprehensive evaluation on the performance of the PI, and the lower the absolute value, the better the effectiveness of the PI. The **AIS** acquired from the WSCFS is lower than the three CFMs.

(b) As for 60-min intervals, despite the IP performance of the WSCFS is not as better as 30-minute intervals, which is attributed to the increasing uncertainty embedded in the wind speed sequence as the forecasting time intervals increase. However, compared to the other three CFMs, the WSCFS is still the optimal model in terms of **IF**.

(c) For 120-min intervals, the results are compatible with the above two scenarios, the WSCFS is still the optimal model compared to the remaining three CFMs. Take Site 1 for instance, the **IF** results of our proposed WSCFS are depicted in Fig. 9.

Remark. Through fitting the optimal DFs to our proposed WSCFS, significantly improves the effectiveness of the WSCFS concerning IF.

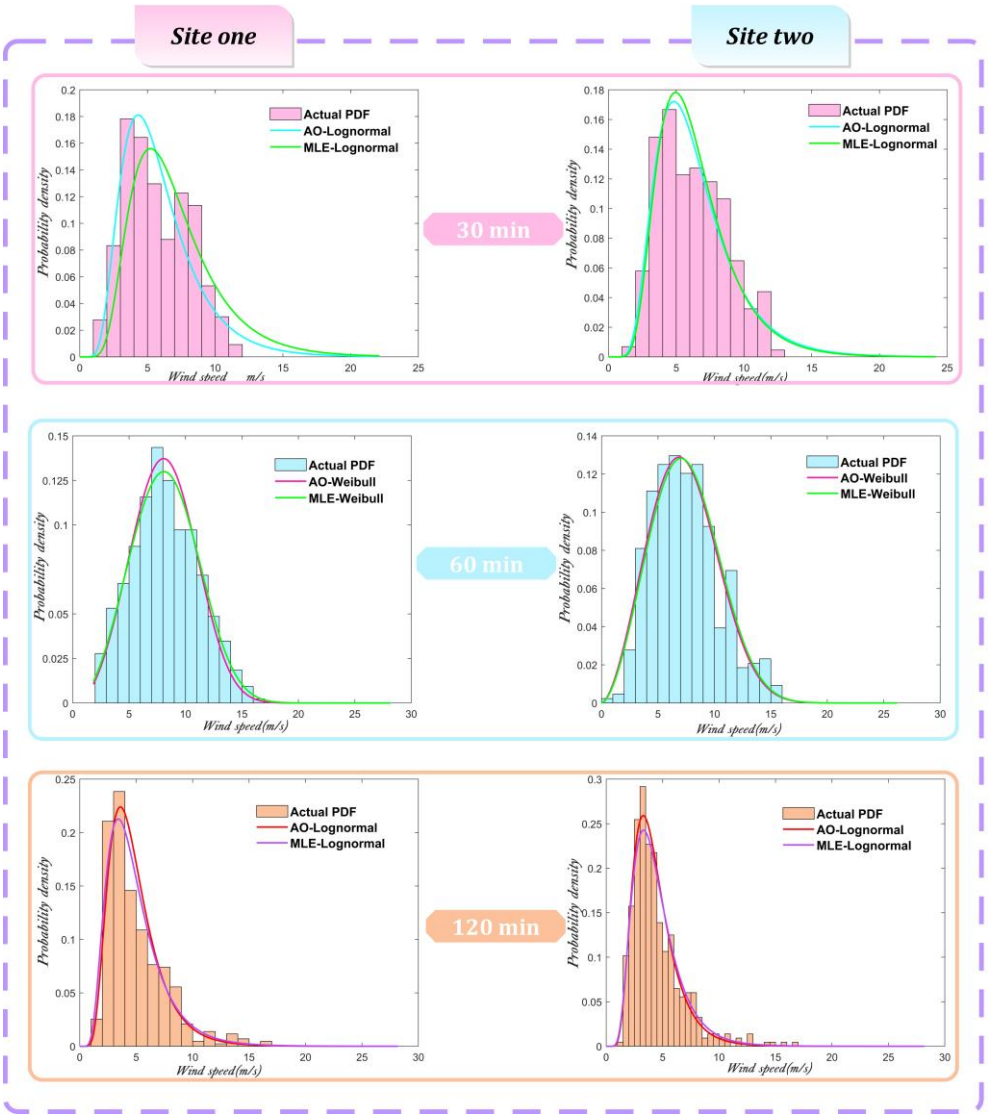


Fig. 8 The comparative results of distribution fitting for two methods.

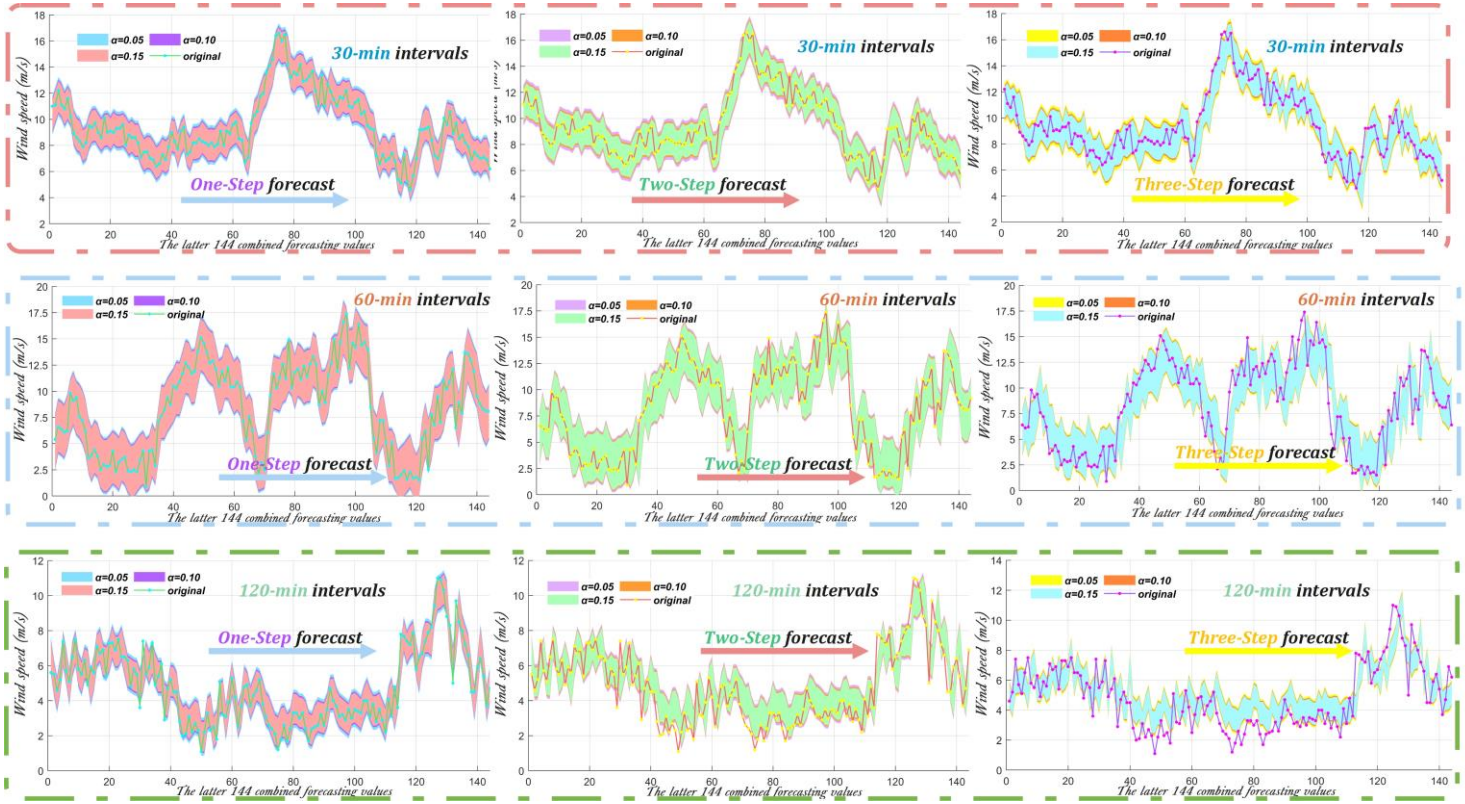


Fig. 9. The **IF** results of our proposed WSCFS in Site 1.

Table 14
The results of fitting distributions for our proposed WSCFS in Site 1.

Forecasting Interval	Step	Method	Lognormal	Weibull	Rayleigh	Logistic	Optimal Distribution
30min	1-step	AO	**0.9927**	0.9880	0.9816	0.9861	**Lognormal**
		MLE	**0.9814**	0.9876	0.9698	0.9804	**Lognormal**
	2-step	AO	**0.9916**	0.9911	0.9782	0.9866	**Lognormal**
		MLE	**0.9890**	0.9874	0.9649	0.9807	**Lognormal**
	3-step	AO	**0.9902**	0.9798	0.9692	0.9798	**Lognormal**
		MLE	**0.9897**	0.9832	0.9545	0.9771	**Lognormal**
60min	1-step	AO	0.9859	**0.9993**	0.9435	0.9985	**Weibull**
		MLE	0.9820	**0.9990**	0.9113	0.9970	**Weibull**
	2-step	AO	0.9899	**0.9988**	0.9382	0.9967	**Weibull**
		MLE	0.9856	**0.9978**	0.9032	0.9945	**Weibull**
	3-step	AO	0.9937	**0.9972**	0.9226	0.9958	**Weibull**
		MLE	0.9897	**0.9956**	0.8815	0.9933	**Weibull**
120min	1-step	AO	**0.9911**	0.9731	0.9682	0.9617	**Lognormal**
		MLE	**0.9850**	0.9575	0.9597	0.9599	**Lognormal**
	2-step	AO	**0.9882**	0.9684	0.9307	0.9643	**Lognormal**
		MLE	**0.9813**	0.9419	0.9306	0.9606	**Lognormal**
	3-step	AO	**0.9811**	0.9632	0.8873	0.9668	**Lognormal**
		MLE	**0.9738**	0.9248	0.8824	0.9598	**Lognormal**

Note: The above table presents the fitting distributions of the forecasting sequences for **PF** at different time intervals in Site 1. The fitting effectiveness is calculated by R^2, the detailed calculation equation of $\mathbf{R^2}$ is defined as: $\mathbf{R^2} = \sum_{i=1}^{K}\left(P_i^{pre} - \bar{P}_i^{act}\right)^2 \Big/ \sum_{i=1}^{K}\left(P_i^{act} - \bar{P}_i^{act}\right)^2$, where P_i^{pre}, P_i^{act} and $\bar{P}_i^{act}$ refer to the predicted values, actual values and the mean of actual values respectively. The best-fitting $\mathbf{R^2}$ is bolded in the table.

Table 15
The results of fitting distributions for our proposed WSCFS in Site 2.

Forecasting Interval	Step	Method	Lognormal	Weibull	Rayleigh	Logistic	Optimal Distribution
30min	**1-step**	**AO**	**0.9951**	0.9945	0.9677	0.9906	**Lognormal**
		MLE	**0.9913**	0.9916	0.9512	0.9866	**Lognormal**
	2-step	**AO**	**0.9945**	0.9934	0.9640	0.9885	**Lognormal**
		MLE	**0.9928**	0.9892	0.9470	0.9847	**Lognormal**
	3-step	**AO**	**0.9944**	0.9912	0.9512	0.9889	**Lognormal**
		MLE	**0.9942**	0.9881	0.9314	0.9858	**Lognormal**
60min	**1-step**	**AO**	0.9932	**0.9986**	0.9683	0.9962	**Weibull**
		MLE	0.9916	**0.9969**	0.9540	0.9949	**Weibull**
	2-step	**AO**	0.9859	**0.9977**	0.9600	0.9958	**Weibull**
		MLE	0.9822	**0.9970**	0.9368	0.9931	**Weibull**
	3-step	**AO**	0.9555	**0.9951**	0.9119	0.9943	**Weibull**
		MLE	0.9550	**0.9911**	0.8752	0.9927	**Weibull**
120min	**1-step**	**AO**	**0.9972**	0.9826	0.9630	0.9773	**Lognormal**
		MLE	**0.9940**	0.9563	0.9570	0.9721	**Lognormal**
	2-step	**AO**	**0.9780**	0.9606	0.8301	0.9635	**Lognormal**
		MLE	**0.9673**	0.9018	0.8147	0.9589	**Lognormal**
	3-step	**AO**	**0.9723**	0.9480	0.8055	0.9533	**Lognormal**
		MLE	**0.9550**	0.8973	0.7795	0.9500	**Lognormal**

Note: The above table presents the fitting distributions of the forecasting sequences for **PF** at different time intervals in Site 2. The fitting effectiveness is calculated by $\boldsymbol{R^2}$, the detailed calculation equation of $\mathbf{R^2}$ is defined as: $\mathbf{R^2} = \sum_{i=1}^{K}\left(P_i^{pre} - \bar{P}_i^{act}\right)^2 \Big/ \sum_{i=1}^{K}\left(P_i^{act} - \bar{P}_i^{act}\right)^2$, where P_i^{pre}, P_i^{act} and $\bar{P}_i^{act}$ refer to the predicted values, actual values and the mean of actual values respectively. The best-fitting $\mathbf{R^2}$ is bolded in the table.

Table 16
The comparative results of interval forecasting for our proposed WSCFS and other tested models in Site 1.

Forecasting Interval	**Forecasting Step**	**Alpha**	**Proposed WSCFS**			FIG-MOALO			FIG-MODA			FIG-MOGOA		
			AIS	PICP	AWD	AIS	PICP	AWD	AIS	PICP	AWD	AIS	PICP	AWD
30min	one	0.05	-0.3007	**100.000**	**0.0000**	**-0.2946**	95.1389	0.0026	-0.3062	93.7500	0.0032	-0.3330	91.6667	0.0048
		0.1	**-0.5204**	**100.000**	**0.0000**	-0.5387	92.3611	0.0047	-0.5555	91.6667	0.0048	-0.5690	90.9722	0.0053
		0.15	**-0.7156**	**98.6111**	**0.0012**	-0.7534	90.2778	0.0056	-0.7713	90.9722	0.0057	-0.7758	86.1111	0.0086
	two	0.05	**-0.3255**	**98.6111**	**0.0012**	-0.4965	84.7222	0.0092	-0.5476	79.1667	0.0119	-0.5829	81.9444	0.0118
		0.1	**-0.5758**	**96.5278**	**0.0031**	-0.7639	79.8611	0.0119	-0.8400	76.3889	0.0148	-0.8638	77.0833	0.0153
		0.15	**-0.7940**	**95.8333**	**0.0039**	-0.9996	73.6111	0.0162	-1.0851	74.3056	0.0169	-1.1068	72.9167	0.0171
	three	0.05	**-0.4245**	**91.6667**	**0.0066**	-0.6726	79.1667	0.0133	-0.6351	75.4129	0.0187	-0.8117	76.3889	0.0154
		0.1	**-0.7311**	**85.4167**	**0.0091**	-0.9866	73.6111	0.0158	-0.9427	72.9167	0.0156	-1.1291	69.4444	0.0200
		0.15	**-0.9973**	**82.6389**	**0.0110**	-1.2534	70.1389	0.0187	-1.2041	70.1389	0.0179	-1.3945	67.3611	0.0220
60min	one	0.05	**-0.4620**	**97.9167**	**0.0004**	-0.5431	95.1389	0.0019	-0.5311	97.2222	0.0005	-0.5366	97.2222	0.0014
		0.1	**-0.8731**	**96.5278**	**0.0013**	-0.9942	92.3611	0.0033	-0.9634	95.1389	0.0015	-0.9796	94.4444	0.0024
		0.15	**-1.2544**	**96.5278**	**0.0014**	-1.3970	92.3611	0.0035	-1.3548	93.7500	0.0025	-1.3932	91.6667	0.0038
	two	0.05	**-1.0546**	**83.9444**	**0.0073**	-1.1707	83.3333	0.0075	-1.3722	77.7778	0.0108	-1.2256	79.3792	0.0078
		0.1	**-1.5279**	**81.8611**	**0.0091**	-1.7068	81.2500	0.0089	-1.9127	76.3889	0.0122	-1.7796	77.7317	0.0099
		0.15	**-1.9479**	**80.7778**	**0.0108**	-2.1546	79.5556	0.0115	-2.3591	73.6111	0.0142	-2.2473	75.3691	0.0114
	three	0.05	-2.2622	**67.8889**	**0.0184**	**-2.1408**	66.6667	0.0186	-2.5978	57.6389	0.0237	-2.6976	63.1944	0.0207
		0.1	-2.7669	**62.5000**	**0.0200**	**-2.7168**	61.8056	0.0204	-3.1445	54.8611	0.0267	-3.2754	61.1111	0.0230
		0.15	-3.1969	**61.1111**	**0.0218**	**-3.1891**	58.3333	0.0225	-3.5788	50.6944	0.0304	-3.7328	59.7222	0.0253
120min	one	0.05	**-0.3115**	**92.3611**	**0.0046**	-0.3547	87.5000	0.0060	-0.3606	87.5000	0.0062	-0.3971	84.7222	0.0099
		0.1	**-0.5291**	**88.1944**	**0.0070**	-0.5799	83.3333	0.0084	-0.5768	85.4167	0.0085	-0.6060	82.6389	0.0112
		0.15	**-0.7137**	**83.3333**	**0.0095**	-0.7727	81.9444	0.0099	-0.7666	79.8611	0.0101	-0.7804	78.4722	0.0130
	two	0.05	**-0.9998**	**65.9722**	**0.0196**	-1.4911	50.0000	0.0339	-1.3959	49.3056	0.0350	-1.7410	40.9722	0.0486
		0.1	**-1.3497**	**56.9444**	**0.0272**	-1.7739	45.8333	0.0395	-1.6817	43.7500	0.0419	-1.9962	38.8889	0.0547
		0.15	**-1.6163**	**52.0833**	**0.0318**	-2.0004	43.0556	0.0434	-1.9078	40.9722	0.0472	-2.1882	38.8889	0.0585
	three	0.05	**-1.8561**	**45.8333**	**0.0387**	-2.2803	36.8056	0.0529	-2.7351	29.8611	0.0655	-2.7674	25.0000	0.0756
		0.1	**-2.2383**	**40.9722**	**0.0487**	-2.5559	34.7222	0.0584	-3.0117	27.7778	0.0725	-3.0468	23.6111	0.0838
		0.15	**-2.5173**	**36.8056**	**0.0542**	-2.7726	33.3333	0.0625	-3.2297	25.6944	0.0786	-3.2507	22.9167	0.0895

Note: The above table shows the specific **IF** results of our WSCFS and the tested models in Site 1. The significance level α is set as $\alpha = 0.05$, $\alpha = 0.1$ and $\alpha = 0.15$ respectively. The optimal results are bolded in the table.

Table 17
The comparative results of interval forecasting for our proposed WSCFS and other tested models in Site 2.

Forecasting Interval	Forecasting Step	Alpha	Proposed WSCFS			FIG-MOALO			FIG-MODA			FIG-MOGOA		
			AIS	PICP	AWD	AIS	PICP	AWD	AIS	PICP	AWD	AIS	PICP	AWD
30min	one	0.05	-0.3308	**100.000**	**0.0000**	-0.3279	93.0556	0.0041	**-0.3146**	95.1389	0.0027	-0.3174	95.8333	0.0020
		0.1	-0.5816	**97.2222**	**0.0024**	-0.5871	89.5833	0.0063	**-0.5736**	92.3611	0.0041	-0.5780	92.3611	0.0041
		0.15	-0.8095	**97.2222**	**0.0025**	-0.8160	89.5833	0.0066	**-0.8058**	90.9722	0.0044	-0.8112	90.9722	0.0050
	two	0.05	**-0.4344**	**93.7500**	**0.0041**	-0.6909	84.7222	0.0107	-0.7401	82.6389	0.0119	-0.5287	85.4167	0.0092
		0.1	**-0.7505**	**91.6667**	**0.0046**	-1.0279	79.1667	0.0118	-1.0740	75.0000	0.0156	-0.8243	82.6389	0.0113
		0.15	**-1.0200**	**89.5833**	**0.0069**	-1.3149	76.3889	0.0140	-1.3582	71.5278	0.0191	-1.0774	79.1667	0.0126
	three	0.05	**-0.6265**	**86.1111**	**0.0094**	-0.6614	85.4167	0.0096	-0.7231	81.2500	0.0129	-0.6700	83.3333	0.0104
		0.1	**-0.9742**	**81.8611**	**0.0124**	-1.0084	80.9444	0.0129	-1.0654	75.6944	0.0166	-1.0206	78.4722	0.0127
		0.15	**-1.2662**	**77.7778**	**0.0141**	-1.3046	76.3889	0.0140	-1.3560	72.9167	0.0177	-1.3212	74.3056	0.0143
60min	one	0.05	**-0.5144**	**97.2222**	**0.0010**	-0.5928	96.6111	0.0011	-0.5594	96.6111	0.0023	-0.4690	95.1389	0.0030
		0.1	-0.9554	**95.8333**	**0.0011**	-1.0362	95.6111	0.0012	-0.9562	91.6667	0.0044	**-0.8661**	92.3611	0.0041
		0.15	-1.3574	**95.1389**	**0.0017**	-1.4400	94.8333	0.0027	-1.3620	90.9722	0.0047	**-1.2132**	92.3611	0.0043
	two	0.05	-1.0870	**84.0278**	**0.0075**	**-0.9595**	82.6389	0.0077	-1.2270	81.2500	0.0093	-1.0734	79.8611	0.0096
		0.1	-1.5772	**82.6389**	**0.0085**	**-1.4526**	78.4722	0.0094	-1.7261	77.0833	0.0114	-1.5233	75.6944	0.0124
		0.15	-2.0067	**81.2500**	**0.0099**	**-1.8728**	77.7778	0.0104	-2.1447	75.0000	0.0127	-1.9090	72.2222	0.0146
	three	0.05	**-2.9611**	**61.1111**	**0.0304**	-3.1254	56.2500	0.0311	-3.6316	45.8333	0.0438	-3.5751	44.4444	0.0431
		0.1	**-3.4424**	**56.9444**	**0.0340**	-3.6408	52.0833	0.0354	-4.1231	41.6667	0.0494	-4.0228	39.5833	0.0474
		0.15	**-3.8600**	**52.0833**	**0.0367**	-4.0448	45.8333	0.0405	-4.4963	40.9722	0.0521	-4.3768	36.8056	0.0508
120min	one	0.05	**-0.3155**	**89.5833**	**0.0062**	-0.6080	74.3056	0.0233	-0.5459	75.6944	0.0166	-0.7211	70.8333	0.0332
		0.1	**-0.5119**	**84.7222**	**0.0082**	-0.8306	64.5833	0.0277	-0.7738	69.4444	0.0193	-0.9331	63.8889	0.0365
		0.15	**-0.6818**	**78.4722**	**0.0124**	-1.0223	58.3333	0.0319	-0.9679	63.8889	0.0231	-1.1111	61.8056	0.0393
	two	0.05	**-0.9821**	**60.5295**	**0.0136**	-1.3068	48.6111	0.0300	-2.0088	40.2778	0.0754	-1.9941	34.7222	0.0656
		0.1	**-1.2390**	**56.8917**	**0.0196**	-1.5919	44.4444	0.0361	-2.2129	33.3333	0.0853	-2.2210	30.5556	0.0731
		0.15	**-1.5672**	**50.3296**	**0.0239**	-1.8229	41.6667	0.0392	-2.3760	31.2500	0.0908	-2.4001	29.1667	0.0783
	three	0.05	**-2.2313**	**38.1944**	**0.0501**	-2.5245	36.1111	0.0573	-2.7202	32.6389	0.0661	-2.6272	35.4167	0.0643
		0.1	**-2.5040**	**36.8056**	**0.0557**	-2.7690	32.6389	0.0655	-2.9494	29.8611	0.0732	-2.8578	31.2500	0.0727
		0.15	**-2.7132**	**34.7222**	**0.0612**	-2.9640	31.9444	0.0700	-3.1322	29.1667	0.0780	-3.0423	29.8611	0.0781

Note: The above table shows the specific **IF** results of our WSCFS and the tested models in Site 2. The significance level α is set as $\alpha = 0.05$, $\alpha = 0.1$ and $\alpha = 0.15$ respectively. The optimal results are bolded in the table.

5. Discussion

In this section, we further discuss the **PF** and **IF** results to validate the superior performance of the WSCFS, which involves the Diebold-Mariano (**DM**) test, metrics improvement ratio (**IR**), and forecasting sensitivity analysis. Moreover, the practical applications of our proposed WSCFS will be further illustrated.

5.1 Diebold-Mariano (DM)-test

It is insufficient to use the error evaluation indicators to illustrate the WSCFS has superior performance than the tested models. Therefore, the **DM** test is applied to verify the significant difference between the WSCFS and tested models. Moreover, we also tested the discrepancy between the forecasting sequences obtained by the proposed WSCFS and the original sequences.

We define $\hat{\tilde{\boldsymbol{\Omega}}}''^{i}_{var}(\hat{\tilde{\sigma}}^{a}_{ij},\hat{\tilde{\sigma}}^{b}_{ij})$ as a variance function. The $\hat{\tilde{\sigma}}^{a}_{ij}$ and $\hat{\tilde{\sigma}}^{b}_{ij}$ represent the **PF** error or **IF** score of the WSCFS and other tested models. As for the test of forecasting and original sequences, $\hat{\tilde{\tilde{\sigma}}}^{a}_{ij}$ and $\hat{\tilde{\tilde{\sigma}}}^{b}_{ij}$ refer to the forecasting sequences and original sequences.

The DM statistic is calculated by:

$$DM=\frac{\sum_{t=1}^{T}\left[\hat{\tilde{\boldsymbol{\Omega}}}''^{i}_{var}(\hat{\tilde{\sigma}}^{a}_{ij},\hat{\tilde{\sigma}}^{b}_{ij})\right]/T}{\sqrt{\chi^{2}/T}} \tag{15}$$

Where T represents the length of the sequence, χ^{2} represents the variance of $\hat{\tilde{\boldsymbol{\Omega}}}''^{i}_{var}(\hat{\tilde{\sigma}}^{a}_{ij},\hat{\tilde{\sigma}}^{b}_{ij})$.

For the test of **PF**, the formulation of the null hypothesis $\boldsymbol{H_0}$ and alternative hypothesis $\boldsymbol{H}_1$ are as follows.

$$\boldsymbol{H_0}:\hat{\tilde{\tilde{\boldsymbol{E}}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{error}^{a}_{ij})\right]=\hat{\tilde{\tilde{\boldsymbol{E}}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{error}^{b}_{ij})\right],\forall n \tag{16}$$

$$\boldsymbol{H}_1:\hat{\tilde{\tilde{\boldsymbol{E}}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{error}^{a}_{ij})\right]\neq\hat{\tilde{\tilde{\boldsymbol{E}}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{error}^{b}_{ij})\right],\exists n \tag{17}$$

While for **IF**, the null hypotheses $\boldsymbol{H_0}$ and alternative hypotheses $\boldsymbol{H}_1$ are defined as follows.

$$\boldsymbol{H_0}:\hat{\tilde{\tilde{\boldsymbol{E}}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{score}^{a}_{ij})\right]=\hat{\tilde{\tilde{\boldsymbol{E}}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{score}^{b}_{ij})\right],\forall n \tag{18}$$

$$\boldsymbol{H}_1:\hat{\tilde{\tilde{\boldsymbol{E}}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{score}^{a}_{ij})\right]\neq\hat{\tilde{\tilde{\boldsymbol{E}}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{score}^{b}_{ij})\right],\exists n \tag{19}$$

For the test of forecasting sequences and original sequences, the null hypotheses $\boldsymbol{H_0}$ and alternative hypotheses $\boldsymbol{H}_1$ are determined by the following formulas.

$$\boldsymbol{H_0}:\hat{\tilde{\tilde{\boldsymbol{E}}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{forecast}^{a}_{ij})\right]=\hat{\tilde{\tilde{\boldsymbol{E}}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{original}^{b}_{ij})\right],\forall n \tag{20}$$

$$\boldsymbol{H}_1:\hat{\tilde{\tilde{\boldsymbol{E}}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{forecast}^{a}_{ij})\right]\neq\hat{\tilde{\tilde{\boldsymbol{E}}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{original}^{b}_{ij})\right],\exists n \tag{21}$$

The specific values of **DM** statistics are presented in Table 18-21 respectively. The detailed analysis is illustrated below.

(a) In **PF**, comparing with the benchmark models, it is apparent that almost all the models pass the significance test at the confidence level $\boldsymbol{\alpha}=0.1$ in all cases. As for CFMs, it can be noticed that all the tests are significant, excluding one model in Site 1. Moreover, by comparing with some classical neural network models, almost all **DM**

statistics values are higher than $Z_{0.01/2} = 2.58$. The **DM** test further validates the performance of our **WSCFS-PF** compared to other models for **PF**.

(b) As for **IF**, we adopt the PI score to gauge the discrepancy between the WSCFS and three CFMs. It can be observed that most of the **DM** statistics values are higher than $Z_{0.1/2} = 1.645$. The largest **DM** statistic is 6.7688, which is higher than $Z_{0.01/2} = 2.58$. Therefore, the **DM** test applied to the **IF** performance suggests that there are significant differences between the **WSCFS-IF** and CFMs-IF.

(c) As can be seen in Table 21, the DM statistic values all fall into the acceptance field in any case, which indicates that we cannot reject the null hypothesis and believe that there is no discrepancy between the forecasting sequences of the proposed WSCFS and the original sequences, further demonstrating that our proposed WSCFS possesses superior forecasting performance.

Table 18
The **PF** results of the **DM** test in Site 1.

Models	30 min			60 min			120 min		
	Step-1	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**
FIG-BP	4.8069*	6.6153*	4.9453*	4.1605*	4.0514*	2.5018**	5.6094*	4.8962*	5.2656*
FIG-ELM	5.0913*	6.4587*	5.0701*	3.9130*	3.6994*	3.0884*	5.8253*	5.8001*	5.2337*
FIG-ENN	4.9675*	6.5512*	5.0094*	3.5176*	4.4731*	2.9866*	6.4704*	4.7871*	4.1269*
FIG-GRNN	6.2571*	6.6309*	5.3096*	5.2212*	4.4654*	3.8063*	6.8819*	6.0859*	5.5002*
FIG-LSTM	5.2985*	6.5344*	4.8920*	4.1322*	4.1009*	3.2287*	6.2717*	5.6762*	6.3872*
FIG-TCN	4.6131*	5.5526*	5.4445*	5.1759*	4.9250*	4.4938*	6.3758*	6.7711*	5.6427*
FIG-GRU	5.2170*	6.8575*	5.1144*	4.2008*	3.9963*	3.5885*	5.0377*	4.8411*	5.2878*
FIG-MOALO	4.7894*	5.2557*	5.5525*	3.9745*	3.0325*	1.5246	2.7709*	5.3328*	1.6533***
FIG-MODA	5.1672*	5.6233*	5.8747*	4.2097*	4.4635*	1.7628***	2.9573*	4.1293*	4.8743*
FIG-MOGOA	5.6550*	5.8277*	6.2171*	3.8810*	4.1342*	3.0048*	2.1405**	5.3409*	3.4527*
BP	7.0053*	6.4438*	5.5229*	5.5431*	4.7418*	5.0753*	7.0114*	6.0385*	5.8358*
ELM	6.3900*	6.6786*	5.2952*	7.2432*	5.0789*	5.4092*	8.1483*	7.1326*	6.4676*
ENN	7.1011*	6.5845*	5.1564*	6.9830*	4.7390*	4.8114*	6.9825*	6.0441*	5.7519*
GRNN	7.5154*	6.7468*	5.0146*	6.7302*	5.7557*	5.7209*	7.2967*	5.8791*	5.9056*
LSTM	7.0818*	6.4827*	4.9988*	7.0594*	5.0796*	4.4145*	7.5658*	5.3803*	5.7864*
TCN	6.7735*	7.5707*	5.3454*	7.2342*	6.0160*	6.5255*	8.2909*	6.7912*	6.7813*
GRU	6.9132*	6.5086*	5.6774*	6.5178*	5.3960*	4.3287*	6.9837*	5.5074*	4.6251*

Note: The **PF** results of **DM** test values for our proposed WSCFS and tested models in Site 1 are detailed in the above table, and the $\boldsymbol{H_0}$ and $\boldsymbol{H_1}$ are $\boldsymbol{H_0}: \tilde{\tilde{\boldsymbol{E}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{error}_{ij}^{a})\right] = \tilde{\tilde{\boldsymbol{E}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{error}_{ij}^{b})\right], \forall n$ and $\boldsymbol{H_1}: \tilde{\tilde{\boldsymbol{E}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{error}_{ij}^{a})\right] \neq \tilde{\tilde{\boldsymbol{E}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{error}_{ij}^{b})\right], \exists n$, where $\boldsymbol{error}_{ij}^{a}$ and $\boldsymbol{error}_{ij}^{b}$ refer to the forecasting error of our proposed WSCFS and tested models respectively. The **DM** statistic is calculated as $DM = \left(\sum_{t=1}^{T}\left[\hat{\tilde{\boldsymbol{\Omega}}}'^{i}_{var}(error_{ij}^{a}, error_{ij}^{b})\right]/T\right) \Big/ \sqrt{\chi^2/T}$. Moreover, the asterisks *, **, and *** denote a significance level of 1%, 5%, and 10% respectively.

Table 19
The **PF** results of the **DM** test in Site 2.

Models	30 min			60 min			120 min		
	Step-1	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**
FIG-BP	4.4254*	5.9913*	5.3977*	2.0193**	2.9467*	3.0177*	4.9828*	4.4076*	3.3685*
FIG-ELM	4.3787*	5.9385*	5.3934*	2.0516**	3.1663*	3.4606*	4.8124*	4.4582*	3.7454*
FIG-ENN	4.2748*	5.8190*	5.6073*	1.9648**	3.0454*	3.1298*	4.3682*	4.5722*	3.2463*
FIG-GRNN	5.4917*	5.8003*	4.6431*	3.2547*	3.0237*	4.1300*	7.4829*	5.8788*	4.6510*
FIG-LSTM	4.4016*	5.5346*	4.7714*	2.3006**	3.0653*	2.7409*	4.0176*	3.5597*	5.5538*
FIG-TCN	4.8030	5.7069*	5.7786*	3.5926*	4.2360*	2.3196**	6.0826*	3.8629*	4.7172*
FIG-GRU	4.5204	6.0728*	5.2130*	3.0865*	3.4643*	1.9665**	5.0266*	3.0049*	3.6456*
FIG-MOALO	5.3844*	5.7663*	4.8182*	2.1437**	2.8885*	2.0204**	3.4931*	2.5555*	1.9550***
FIG-MODA	4.6077*	6.0283*	2.9489*	2.4395**	2.9693*	1.9743**	3.1049*	3.6878*	2.5814*
FIG-MOGOA	4.1890*	2.9185*	4.7814*	2.6859*	2.3177**	2.0956**	3.5637*	2.2473**	2.1666**
BP	6.1153*	6.3205*	5.3314*	4.7930*	4.2349*	3.4390*	7.4523*	6.8977*	5.8056*
ELM	6.5326*	6.7664*	5.5344*	4.6713*	4.1021*	4.1484*	7.9101*	7.2992*	4.8007*
ENN	6.0151*	6.3557*	5.2853*	4.6820*	3.9365*	3.0526*	7.6333*	6.6323*	4.5205*
GRNN	7.0896*	6.4258*	4.8786*	5.2427*	3.4779*	4.1679*	8.2145*	7.0945*	4.3096*
LSTM	6.2868*	7.4169*	5.4344*	4.9840*	3.9012*	3.1185*	7.3409*	5.7886*	5.7934*
TCN	7.5001*	6.1665*	6.6630*	5.5713*	5.4327*	4.6232*	6.9545*	6.5549*	5.7707*
GRU	6.0167*	5.9516*	5.3212*	5.4045*	4.6553*	2.1986**	7.9693*	5.1260*	5.7519*

Note: The **PF** results of **DM** test values for our proposed WSCFS and tested models in Site 2 are presented in the above table, and the $\boldsymbol{H_0}$ and $\boldsymbol{H_1}$ are $\boldsymbol{H_0}: \tilde{\tilde{\boldsymbol{E}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{error}^{a}_{ij})\right] = \tilde{\tilde{\boldsymbol{E}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{error}^{b}_{ij})\right], \forall n$ and $\boldsymbol{H}_1: \tilde{\tilde{\boldsymbol{E}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{error}^{a}_{ij})\right] \neq \tilde{\tilde{\boldsymbol{E}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{error}^{b}_{ij})\right], \exists n$, where $\boldsymbol{error}^{a}_{ij}$ and $\boldsymbol{error}^{b}_{ij}$ are refer to the forecasting errors of our proposed WSCFS and tested models respectively. The **DM** statistic is calculated as $DM = \left(\sum_{t=1}^{T}\left[\hat{\tilde{\boldsymbol{\Omega}}}'^{i}_{var}(error^{a}_{ij}, error^{b}_{ij})\right]/T\right) \Big/ \sqrt{\chi^2/T}$. Moreover, the asterisks *, **, and *** denote a significance level of 1%, 5%, and 10% respectively.

Table 20
The **IF** results of the **DM** test for our proposed WSCFS and tested models.

Forecasting interval	Model	1-step Forecasting			2-step Forecasting			3-step Forecasting		
		α = 0.05	α = 0.10	α = 0.15	α = 0.05	α = 0.10	α = 0.15	α = 0.05	α = 0.10	α = 0.15
The IF values of the DM test for our proposed WSCFS and tested models in Site1										
30 min	FIG-MOALO	1.4410	1.6424	2.0308**	2.4626**	2.6550*	2.8104*	3.4959*	3.8200*	4.0013*
	FIG-MODA	1.5864	2.0979**	2.4103**	2.7037*	2.9761*	3.1834*	3.3168*	3.6443*	3.8072*
	FIG-MOGOA	1.7169***	2.6073*	3.2926*	2.2582**	2.5947*	2.8922*	3.2695*	3.5781*	3.7743*
60 min	FIG-MOALO	2.4531**	2.1181**	2.1627**	1.4465	1.9033***	2.2132*	1.5618	1.6239	1.5181
	FIG-MODA	2.3063**	1.9466***	1.9111***	2.1756**	2.6565*	2.9277*	2.5207**	3.0282*	3.1291*
	FIG-MOGOA	3.0810*	2.5891*	2.6158*	1.3817	1.8457***	2.3604**	2.4365**	2.6769*	2.7950*
120 min	FIG-MOALO	2.3475**	1.5970	5.0016*	4.5607*	4.2857*	2.1730**	2.3678**	2.4294**	2.0074**
	FIG-MODA	1.8519***	1.9649**	4.6309*	3.6163*	2.9924*	3.4291*	5.4404*	4.9910*	4.7553*
	FIG-MOGOA	2.1713**	2.2157**	6.7370*	6.7688*	6.6303*	4.5902*	5.6047*	5.1642*	4.8428*
The IF values of the DM test for our proposed WSCFS and tested models in Site 2										
30 min	FIG-MOALO	1.6251	1.9602**	2.1427**	3.3494*	3.5785*	3.6946*	2.7961*	2.5963*	2.2700**
	FIG-MODA	1.6729***	1.5380	1.5464	3.6965*	3.8844*	3.9914*	3.0229*	2.9529*	2.7913*
	FIG-MOGOA	1.6611***	1.5912	1.7673***	2.3147**	2.0107**	1.7763***	3.1166*	2.7104*	2.2612**
60 min	FIG-MOALO	1.9821**	1.9384***	1.9065***	1.8767***	1.7140***	1.9966**	4.4019*	4.2506*	4.1110*
	FIG-MODA	1.6341	1.6896***	1.6935***	1.6551***	2.0076**	1.8001***	3.7186*	3.3922*	3.1782*
	FIG-MOGOA	1.7008***	1.8791***	1.5034	2.0956**	1.6697***	1.6662***	2.4514**	2.1743**	2.0797**
120 min	FIG-MOALO	3.2300*	3.9330*	4.0648*	4.1952*	4.0800*	3.9676*	1.9980**	1.9429***	1.8466***
	FIG-MODA	3.2319*	3.9888*	4.1195*	4.5244*	4.5170*	4.4786*	2.3016**	2.1489**	2.0811**
	FIG-MOGOA	3.4754*	3.8648*	3.9984*	3.3572*	3.2995*	3.2302*	1.9088***	1.7502***	1.6823***

Note: The **IF** results of **DM** test values for our proposed WSCFS and tested models are presented in the above table, and the $\boldsymbol{H_0}$ and $\boldsymbol{H_1}$ are $\boldsymbol{H_0}: \tilde{\tilde{\boldsymbol{E}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{score}_{ij}^{a})\right] = \tilde{\tilde{\boldsymbol{E}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{score}_{ij}^{b})\right], \forall n$ and $\boldsymbol{H_1}: \tilde{\tilde{\boldsymbol{E}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{score}_{ij}^{a})\right] \neq \tilde{\tilde{\boldsymbol{E}}}''\left[\hat{\tilde{\boldsymbol{\Omega}}}'_{var}(\boldsymbol{score}_{ij}^{b})\right], \exists n$, where $\boldsymbol{score}_{ij}^{a}$ and $\boldsymbol{score}_{ij}^{b}$ are refer to the interval score of our proposed WSCFS and tested models respectively. The DM statistic is calculated as $DM = \left(\sum_{t=1}^{T}\left[\hat{\tilde{\boldsymbol{\Omega}}}''^{i}_{var}(score_{ij}^{a}, score_{ij}^{b})\right]/T\right) \Big/ \sqrt{\chi^2/T}$. Moreover, the asterisks *, **, and *** denote a significance level of 1%, 5%, and 10% respectively.

Table 21
The **DM** results for the forecasting sequence of our proposed WSCFS and original sequence.

Interval	Site 1			Site 2		
	Step-1	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**
30 min	0.2611	0.5503	0.6152	0.4411	0.3115	0.8970
60 min	0.3436	0.5309	0.8228	1.4302	1.4208	1.5275
120 min	1.6206	1.0389	1.3830	1.5538	1.4795	1.5935

Note: The **DM** test values for the sequences of WSCFS and original sequences are presented in the above table, and the $\boldsymbol{H_0}$ and $\boldsymbol{H_1}$ are $\boldsymbol{H_0}: \tilde{\tilde{\boldsymbol{E}}}''\left[\hat{\boldsymbol{\Omega}}'_{var}(\boldsymbol{forecast}_{ij}^{a})\right] = \tilde{\tilde{\boldsymbol{E}}}''\left[\hat{\boldsymbol{\Omega}}'_{var}(\boldsymbol{original}_{ij}^{b})\right], \forall n$ and $\boldsymbol{H_1}: \tilde{\tilde{\boldsymbol{E}}}''\left[\hat{\boldsymbol{\Omega}}'_{var}(\boldsymbol{forecast}_{ij}^{a})\right] \neq \tilde{\tilde{\boldsymbol{E}}}''\left[\hat{\boldsymbol{\Omega}}'_{var}(\boldsymbol{original}_{ij}^{b})\right], \exists n$, where $\boldsymbol{forecast}_{ij}^{a}$ and $\boldsymbol{original}_{ij}^{b}$ are refer to the forecasting sequences of our proposed WSCFS original sequences. The DM statistic is calculated as $DM = \left(\sum_{t=1}^{T}\left[\hat{\boldsymbol{\Omega}}''^{i}_{var}(forecast_{ij}^{a}, original_{ij}^{b})\right]/T\right) \Big/ \sqrt{\chi^2/T}$.

5.2 Improvement ratio (IR) *analysis*

In this section, we adopt two vital metrics including **MAPE** and **PICP** to determine the improvement ratio (**IR**) of the WSCFS and tested models. The calculation formula for the **IR** is defined as follows:

$$\mathbf{P}_{\boldsymbol{metrics}} = \left|\frac{metrics_{\boldsymbol{compared}} - metrics_{\boldsymbol{proposed}}}{metrics_{\boldsymbol{compared}}}\right| * 100\% \tag{22}$$

Where $metrics_{\boldsymbol{compared}}$ and $metrics_{\boldsymbol{proposed}}$ refer to the evaluation metrics of the tested models and the WSCFS.

The contents of **IR** are shown in Table 22 and 23 and the analysis of the **IR** is shown below.

(a) As can be seen in Table 22, the **PF** performance of the WSCFS is superior to the tested models. We adopted a simple average strategy for two sites to gauge the average **IR**. For the benchmark models，the maximum **IR** is $\mathbf{P}_{\boldsymbol{MAPE}}^{1-step} = 54.6054\%$,the minimum **IR** is acquired by FIG-MOALO in 60-min intervals with the value of $\mathbf{P}_{\boldsymbol{MAPE}}^{3-step} = 3.5765\%$ .For classical neural networks, the **IR** values for BPNN are $\mathbf{P}_{\boldsymbol{MAPE}}^{1-step} = 46.6051\%$, $\mathbf{P}_{\boldsymbol{MAPE}}^{2-step} = 43.5636\%$ and $\mathbf{P}_{\boldsymbol{MAPE}}^{1-step} = 35.3285\%$ in 30-min intervals. The above **IR** analysis of WSCFS-PF once more illustrates that our proposed WSCFS achieves more satisfactory forecasting results than the tested models.

(b) Simultaneously, the performance of WSCFS-IF is also dominated over three CFMs-IF. The maximum **IR** is $\mathbf{P}_{\boldsymbol{PICP}}^{2-step} = 67.6710\%$, the minimum **IR** is obtained by FIG-MODA in 60-min intervals with the value of $\mathbf{P}_{\boldsymbol{PICP}}^{1-step} = 1.0818\%$ at $\boldsymbol{\alpha} = 0.05$. Furthermore, As the time interval and forecasting step increase, the values of **IR** increase in parallel. For the above **IR** analysis, it can be concluded that our proposed WSCFS owns more excellent forecasting effectiveness than the tested models regarding both **PF** and **IF**.

Table 22
The **PF** average values of **IR** for our adopted models.

Models	30 min			60 min			120 min		
	Step-1	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**	**Step-1**	**Step-2**	**Step-3**
FIG-BP	35.7123	37.4028	29.3537	22.7194	21.8444	13.5879	34.6834	30.4083	22.4141
FIG-ELM	34.9234	37.2442	29.8782	20.9451	20.2894	13.5540	34.0855	30.9438	22.6763
FIG-ENN	35.4239	37.7914	30.1447	16.0033	19.7093	11.9951	35.0816	26.3046	18.3535
FIG-GRNN	42.5371	39.9652	30.5629	36.7503	27.1957	24.9101	47.3262	36.2381	26.3268
FIG-LSTM	37.2329	37.6831	30.8957	27.5624	25.4772	16.4055	32.4723	31.6997	26.0913
FIG-TCN	50.9055	53.7787	42.4471	39.0922	39.1368	32.0875	54.6054	45.3708	39.7311
FIG-GRU	36.7285	39.2962	29.5480	37.1837	28.3747	19.9412	38.8530	38.0351	24.8640
FIG-MOALO	21.3043	25.9271	7.5098	11.0203	7.1904	3.5765	14.7211	10.7510	3.9737
FIG-MODA	21.1393	28.2802	7.7510	9.9395	8.5918	6.1066	14.5302	5.1227	9.8269
FIG-MOGOA	21.7815	20.8543	9.9505	10.6458	7.1003	8.5996	13.5867	12.5172	5.8437
BP	46.6051	43.5636	35.3285	46.9709	34.9419	22.7926	55.4706	40.4027	30.8330
ELM	48.4516	44.1739	35.1205	47.7627	37.1154	21.2712	56.7612	45.2574	31.2652
ENN	45.6418	43.3266	34.9511	45.3930	30.6881	18.7524	53.2437	38.5200	26.9747
GRNN	49.6975	44.5299	35.1086	50.5012	34.5656	21.0832	56.2484	39.0144	27.8423
LSTM	46.7248	45.0070	34.4197	47.7631	34.7422	23.1777	53.5440	37.7911	32.4948
TCN	59.8268	58.5771	49.6553	57.6871	40.6954	35.7907	63.2468	48.6291	38.9523
GRU	46.2361	42.8117	34.6670	49.6798	38.9090	23.8003	55.4509	44.7298	31.7481

Note: The above table presents the WSCFS-PF compared to the tested models. The $\mathbf{P}_{metrics}$ is aimed to measure the improvements between our proposed WSCFS and tested models, and the calculation of $\mathbf{P}_{metrics}$ is defined as $\mathbf{P}_{metrics} = \left| MAPE_{compared} - MAPE_{proposed} \right/ MAPE_{compared} \left| * 100\% \right.$ where $MAPE_{compared}$ and $MAPE_{proposed}$ represent the **MAPE** values of the tested models and our proposed WSCFS. ,

Table 23
The **IF** average values of **IR** for our adopted models.

Forecasting interval	Model	1-step Forecasting			2-step Forecasting			3-step Forecasting		
		α = 0.05	α = 0.10	α = 0.15	α = 0.05	α = 0.10	α = 0.15	α = 0.05	α = 0.10	α = 0.15
30 min	**FIG-MOALO**	6.2861	8.3989	8.8789	13.5246	18.3295	23.7307	8.3012	12.1365	9.8200
	FIG-MODA	5.8881	6.7669	7.6336	19.0033	24.2930	27.1072	13.7680	17.2552	12.2442
	FIG-MOGOA	6.7194	7.5934	10.6932	15.0475	18.0749	22.2931	11.6667	14.7720	13.6767
60 min	**FIG-MOALO**	1.7761	2.3719	2.4168	1.2070	3.0310	3.0003	5.2376	5.2284	9.1992
	FIG-MODA	1.0818	3.0026	3.7716	5.6737	7.1854	9.0346	25.5583	25.2953	23.8333
	FIG-MOGOA	1.4520	2.9827	4.1553	5.4843	7.2434	9.8382	22.4644	23.0662	21.9174
120 min	**FIG-MOALO**	13.0581	18.5081	18.1094	28.2311	26.1244	20.8792	15.1486	15.3830	9.5563
	FIG-MODA	11.9521	12.6260	13.5869	42.0414	50.4169	44.0867	35.2547	35.3779	31.1455
	FIG-MOGOA	17.7435	19.6656	16.5804	67.6710	66.3096	53.2434	45.5881	45.6537	38.4425

Note: The above table presents the WSCFS-IF compared to the tested models. The $\mathbf{P}_{metrics}$ is aimed to measure the improvements between our proposed WSCFS and tested models, and the calculation of $\mathbf{P}_{metrics}$ is defined as $\mathbf{P}_{metrics} = \left| PICP_{compared} - PICP_{proposed} \right/ PICP_{compared} \left| * 100\% \right.$,where $PICP_{compared}$ and $PICP_{proposed}$ represent the **PICP** values of the tested models and our proposed WSCFS.

5.3 Sensitivity analysis

To explore the sensitiveness of the WSCFS regarding the critical parameters varying, the sensitive mechanism is discussed in this section, we alter only one key parameter in the model while maintaining the other parameters invariant simultaneously. By observing the fluctuations of the evaluation metrics, and analyzing the sensitivity of the parameters. The sensitivity coefficients are defined by the standard deviation of the four evaluations metrics in **PF**.

$$\mathbf{S}(M)=\sum_{k=1}^{n}\frac{(M_k-\bar{M})^2}{n} \tag{23}$$

where n indicates the testing time, M_k represents the value of the error indicators which include **SDE**, **RMSE**, **MAE** and **MAPE** at the *k-th* time. $\bar{M}$ refers to the mean of all the testing times. The lower value of $\mathbf{S}$ implies that the model is more robust. In the WSCFS, there are three key parameters which are Aquila's number, iteration number, and archive number.

Regarding Aquila's number, we specify its variation range as 20-100, where the interval is 20, we define this pattern as $\bar{\mathbf{Q}}=\left[20,40^{\omega},60,80,100\right]$. Meanwhile, the iteration number and archive number are set to range from 50 to 250 and 100 to 500, with intervals of 50 and 100, respectively, we refer to these patterns as $\bar{\mathbf{T}}=\left[50,100,150,200^{\omega},250\right]$ and $\bar{\mathbf{A}}=\left[100,200,300,400,500^{\omega}\right]$. The symbol ω refers to the optimal parameter in the variation range, which is determined by a trial-and-error manner. We define $\bar{\mathbf{Q}}\Delta(\overline{\mathbf{TA}})$ as Aquila's number continuously increasing from 20 to 100 with the growth interval of 20, while other parameters are fixed. The fluctuations of these error indicators are shown in Table 23, and the comparative results of three key parameters for three-intervals forecast at Site 2 are depicted in Fig. 10.

For 30-min intervals, at Site 1, the archive number is the most influential factor for the forecasting performance among the three parameters. While for Site 2, the iteration number becomes the parameter that contributes to the greatest fluctuation to the WSCFS performance. Moreover, the evaluation indicators increase with the forecasting step, which is attributed to the existence of more uncertainty factors in multi-step forecasting. Nevertheless, almost all the metrics are not very high in all situations. Through the analysis of the parameter variation, we can conclude that the performance of the WSCFS is not significant for the changes of three critical parameters in many cases, which further validates the superior robustness of the WSCFS.

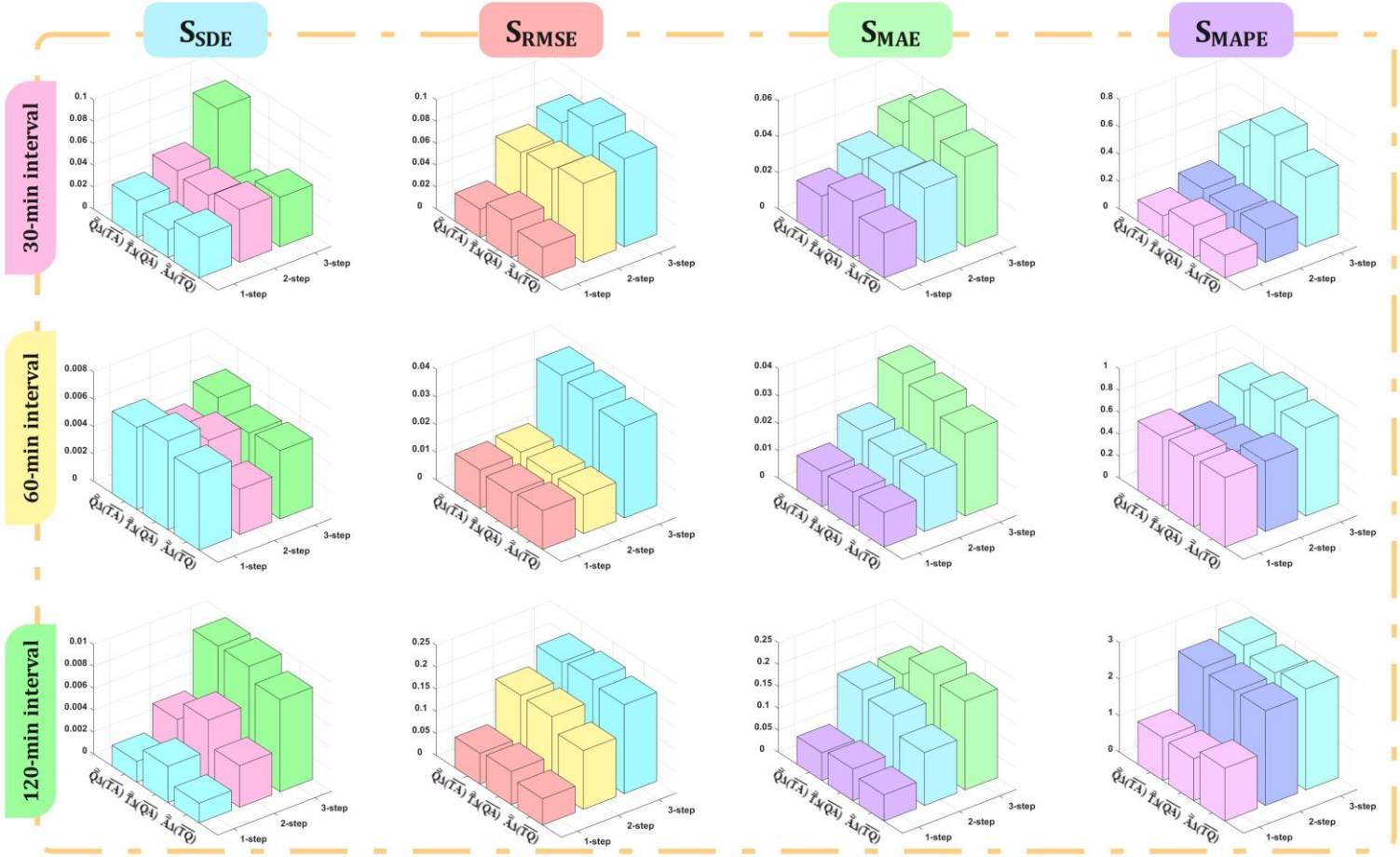


Fig. 10. Comparison of three key parameters for three-intervals forecast at site 2.

Table 24
The results of sensitivity analysis for our proposed WSCFS.

Interval	Modes	1-step Forecasting				2-step Forecasting				3-step Forecasting			
		S_{SDE}	S_{RMSE}	S_{MAE}	S_{MAPE}	S_{SDE}	S_{RMSE}	S_{MAE}	S_{MAPE}	S_{SDE}	S_{RMSE}	S_{MAE}	S_{MAPE}
The results of sensitivity analysis in Site 1													
30 min	$\overline{\overline{Q}}_{A(TA)}$	0.0023	**0.0205**	**0.0198**	0.1469	**0.0033**	0.0580	0.0294	0.1981	0.0052	0.0971	0.0781	0.6088
	$\overline{\overline{T}}_{A(QA)}$	0.0021	0.0157	0.0171	0.1159	0.0024	**0.0684**	**0.0343**	**0.2288**	**0.0063**	0.0942	0.0706	0.5314
	$\overline{\overline{A}}_{A(QT)}$	**0.0026**	0.0114	0.0137	**0.3565**	0.0016	0.0602	0.0292	0.1655	0.0040	**0.1027**	**0.0827**	**0.6449**
60 min	$\overline{\overline{Q}}_{A(TA)}$	0.0122	**0.0161**	0.0127	0.3166	0.0130	0.0206	0.0234	0.6196	0.0121	0.0252	0.0237	0.5489
	$\overline{\overline{T}}_{A(QA)}$	**0.0125**	0.0141	**0.0161**	**0.6179**	0.0124	0.0163	0.0185	**0.7121**	**0.0221**	**0.0338**	**0.0309**	**0.5925**
	$\overline{\overline{A}}_{A(QT)}$	0.0085	0.0091	0.0114	0.4552	**0.0138**	**0.0261**	**0.0291**	0.4427	0.0109	0.0286	0.0288	0.5290
120 min	$\overline{\overline{Q}}_{A(TA)}$	0.0393	0.0797	0.0781	1.3988	**0.0646**	**0.1810**	**0.1758**	**2.1523**	0.0611	0.2132	0.1802	2.5533
	$\overline{\overline{T}}_{A(QA)}$	**0.0544**	0.1096	0.1139	1.5874	0.0529	0.1409	0.1396	2.0429	0.0637	0.2308	0.2363	**2.4634**
	$\overline{\overline{A}}_{A(QT)}$	0.0368	**0.1119**	**0.1343**	**1.9163**	0.0602	0.1563	0.1529	2.1913	**0.0643**	**0.2585**	**0.2440**	2.2614
The results of sensitivity analysis in Site 2													
30 min	$\overline{\overline{Q}}_{A(TA)}$	0.0019	0.0257	0.0228	0.1554	0.0043	0.0632	0.0344	0.2395	0.0096	0.0761	0.0463	0.4305
	$\overline{\overline{T}}_{A(QA)}$	**0.0033**	**0.0347**	**0.0311**	**0.2292**	**0.0061**	0.0664	0.0380	0.2208	**0.0096**	**0.0917**	**0.0608**	**0.6622**
	$\overline{\overline{A}}_{A(QT)}$	0.0018	0.0280	0.0248	0.1634	0.0038	**0.0721**	**0.0410**	**0.2416**	0.0084	0.0802	0.0498	0.5084
60 min	$\overline{\overline{Q}}_{A(TA)}$	0.0060	0.0139	**0.0128**	0.6336	0.0045	**0.0147**	**0.0219**	0.5607	**0.0059**	**0.0368**	**0.0371**	0.7701
	$\overline{\overline{T}}_{A(QA)}$	**0.0065**	0.0131	0.0127	**0.6453**	**0.0054**	0.0141	0.0202	0.5391	0.0048	0.0358	0.0346	**0.8719**
	$\overline{\overline{A}}_{A(QT)}$	0.0056	**0.0140**	0.0126	0.6341	0.0033	0.0140	0.0201	**0.6416**	0.0050	0.0332	0.0302	0.8053
120 min	$\overline{\overline{Q}}_{A(TA)}$	0.0334	0.0697	0.0633	1.1691	0.0465	**0.1651**	**0.1710**	**2.6684**	**0.0890**	0.2050	0.1706	**2.8510**
	$\overline{\overline{T}}_{A(QA)}$	0.0252	**0.0746**	**0.0725**	1.1523	0.0426	0.1635	0.1589	2.5798	0.0266	**0.2110**	**0.2170**	2.6483
	$\overline{\overline{A}}_{A(QT)}$	**0.0371**	0.0594	0.0598	**1.4464**	**0.0481**	0.1328	0.1208	2.5935	0.0452	0.2011	0.2025	2.7654

Note: The above table presents the detailed evaluation indicators for the sensitivity analysis of our proposed WSCFS. The critical parameters of the WSCFS are Aquila's number, iteration number, and archive number. $\bar{\mathbf{Q}}_{\mathbf{\Delta}(\overline{\mathbf{TA}})}$ represents the Aquila number increase from 20 to 100 with the interval of 20. $\tilde{\bar{\mathbf{T}}}_{\mathbf{\Delta}(\overline{\mathbf{QA}})}$ denotes the iteration number set to 50,100,150,200,250. $\bar{\mathbf{A}}_{\mathbf{\Delta}(\overline{\mathbf{QT}})}$ denotes the archive number set to 100, 200, 300, 400, and 500.

5.4 Practical applications in power systems

The above discussion indicates that our proposed WSCFS owns superior precision and robustness in different time intervals **WSP**. Accurate and reliable **WSP** is crucial to the distribution and dispatch of wind power facilities and the security of grid operation. The estimable benefits of our proposed WSCFS and practical application to power systems are shown below.

(1) Accurate and reliable forecasting results facilitate dispatchers to grasp the power variation of wind farms in advance, promptly develop dispatching operational schedules, enhance energy transformation efficiency, minimize risks, and boost power generation. If the **WSP** has an increase in the accuracy rate within 10%, it can enhance power generation by approximately 30%[62]. Therefore, accurate **WSP** contributes to the efficient deployment of wind resources and the operational efficiency of wind farms. Moreover, it also contributes to the wind power grid-connection and stable operation and provides timely alarms for risks that may affect the security and stability operation of the grid, which will prevent power loss or even grid collapse resulting from random fluctuations in wind power. The wind energy potential assessment is determined by the **WSP**, which is attributed to the fact that wind energy is proportional to the cube of the wind speed. Therefore, to ensure that we can exploit and utilize wind energy to the greatest extent feasible, it is crucial to assess wind energy potential for wind farms. Our proposed model has stable and reliable prediction capability, which can provide high accuracy results of wind speed and wind power, thereby contributing to the decision-making support for wind energy assessment and wind farm construction.

(2) Wind speed features intermittency and fluctuation, and as one of the most sensitive parameters of wind power systems, the value variation will generate a tremendous impact on wind farm power generation and grid operation. On the one hand, wind farms can reasonably operate, and overhaul wind turbines and other equipment based on **WSP**, which contributes to solving possible abnormal issues, further directing the installation capacity of wind farms, and consequently improving the effective utilization of wind farms equipment. Moreover, our proposed WSCFS assists the power system dispatching department to optimize the dispatching program timely and provides the power system operation and dispatching department with variations of wind power, which can be utilized to formulate or adapt the dispatching programs promptly, further reducing the spare capacity and operation cost of the power system, which is overwhelming guidance for the stable operation of wind power systems.

(3) In the power market, compared with other controllable power generation methods, such as hydroelectric power generation and nuclear power generation, the volatility and instability of wind speed and other characteristics lead to a considerable influence on the safety and stability of the power grid, which will inevitably make wind power less competitive in the power market and even be financially penalized for its unstable power supply. Moreover, the existence of uncertainties in the actual data, model structure, and model parameters frequently render **PF** unreliable and inaccurate. In our study the proposed WSCFS is capable of quantifying the uncertainty in wind speed, **IF** can significantly mitigate the adverse effects attributed to wind speed uncertainties and other features, which effectively enhances the competitiveness of wind energy in the market and promotes the development of wind power.

6. Conclusion

As a green and renewable energy resource, wind energy has been widely utilized

all around the world in recent years. Accurate **WSP** is indispensable for the high-efficiency utilization of wind energy and the stability as well as security of grid operation. However, in the previous research, the scholars merely achieved accurate **PF** for wind speed, while ignoring the importance of **IF** for **WSP**. Therefore, in our study, we propose a novel WSCFS based on fuzzy information granulation, optimal benchmark models selection, which also combines neural networks, deep learning, and an advanced multi-objective optimizer, to promote **PF** and **IF** capability in wind speed. We adopt six intervals of wind speed data from two sites to simulate the experiment with fourteen benchmark models and three CFMs. Furthermore, to verify the forecasting validity and generalization ability of our WSCFS, hypothesis testing, **IR** analysis, and sensitivity analysis are introduced in the discussion section to comprehensively evaluate the WSCFS. The summary of the simulated experiments is listed below.

(1) The FIG technique is adopted to establish fuzzy windows to sufficiently exploit the effective information in the wind speed data, hence it can enhance the forecasting effectiveness of the models.

(2) The OBMS strategy is applied to select the optimal five models with the best forecasting performance in different situations, which can greatly improve the performance of the combined model.

(3) A theoretical demonstration indicates that the optimal solutions acquired by **MOAO** are the optimal weights.

(4) Comparing AO with three optimizers (ALO, DA, GOA), AO can better improve the forecasting accuracy and stability of the combined model.

(5) Based on four distribution functions, the forecasting values are fitted to the distribution, which consequently quantifies the uncertainty and random fluctuations of wind speed.

Moreover, the DM test also affirms that the WSCFS is significantly different from the tested models for **PF** and **IF** as well as original sequences. The **IR** analysis demonstrates that the WSCFS exhibits considerable improvement compared to tested models. The sensitivity analysis verifies the WSCFS exhibits high stability towards the variation of key parameters. Overall, the experimental findings suggest that our proposed WSCFS can achieve superior forecasting effectiveness and excellent generalization as well as robustness in **PF** and **IF** compared to the tested models.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 71671029).

References

[1] Shao Y, Wang J, Zhang H, Zhao W. An advanced weighted system based on swarm intelligence optimization for wind speed prediction. Appl Math Model 2021;100:780–804. https://doi.org/10.1016/j.apm.2021.07.024.

[2] Lee J, Zhao F. Global Wind Report 2021. Glob Wind Energy Counc 2021:75.

[3] WANG L, WANG T. Wind turbine design and its aerodynamic problems. Sci Sin Phys Mech Astron 2013;43:1579–88. https://doi.org/10.1360/132013-167.

[4] Element B. Modeling of Wind Turbine Impeller Based on The Theory of Momentum Blade Element 2016:130–3.

[5] Duan W, Zhao F. Loading analysis and strength calculation of wind turbine blade based on blade element momentum theory and finite element method. Asia-Pacific Power Energy Eng Conf APPEEC 2010:42–4. https://doi.org/10.1109/APPEEC.2010.5448929.

[6] Jiang H, Tao C, Dong Y, Xiong R. Robust low-rank multiple kernel learning with compound regularization. Eur J Oper Res 2021. https://doi.org/10.1016/j.ejor.2020.12.024.

[7] Zhang W, Zhang L, Wang J, Niu X. Hybrid system based on a multi-objective optimization and kernel approximation for multi-scale wind speed forecasting. Appl Energy 2020;277:115561. https://doi.org/10.1016/j.apenergy.2020.115561.

[8] Wang S, Wang J, Lu H, Zhao W. A novel combined model for wind speed prediction – Combination of linear model, shallow neural networks, and deep learning approaches. Energy 2021;234:121275. https://doi.org/10.1016/j.energy.2021.121275.

[9] Niu X, Wang J. A combined model based on data preprocessing strategy and multi-objective optimization algorithm for short-term wind speed forecasting. Appl Energy 2019;241:519–39. https://doi.org/10.1016/j.apenergy.2019.03.097.

[10] Chen N, Qian Z, Nabney IT, Meng X. Wind power forecasts using gaussian processes and numerical weather prediction. IEEE Trans Power Syst 2014;29:656–65. https://doi.org/10.1109/TPWRS.2013.2282366.

[11] Sile T, Bekere L, Cepite-Frisfelde D, Sennikovs J, Bethers U. Verification of numerical weather prediction model results for energy applications in Latvia. Energy Procedia 2014;59:213–20. https://doi.org/10.1016/j.egypro.2014.10.369.

[12] Wang J, Wang S, Li Z. Wind speed deterministic forecasting and probabilistic interval forecasting approach based on deep learning, modified tunicate swarm algorithm, and quantile regression. Renew Energy 2021;179:1246–61. https://doi.org/10.1016/j.renene.2021.07.113.

[13] Wang H, Han S, Liu Y, Yan J, Li L. Sequence transfer correction algorithm for numerical weather prediction wind speed and its application in a wind power forecasting system. Appl Energy 2019;237:1–10. https://doi.org/10.1016/j.apenergy.2018.12.076.

[14] Wu C, Wang J, Chen X, Du P, Yang W. A novel hybrid system based on multi-objective optimization for wind speed forecasting. vol. 146. Elsevier B.V.; 2020. https://doi.org/10.1016/j.renene.2019.04.157.

[15] Liu M De, Ding L, Bai YL. Application of hybrid model based on empirical mode decomposition, novel recurrent neural networks and the ARIMA to wind speed prediction. Energy Convers Manag 2021;233:113917.

https://doi.org/10.1016/j.enconman.2021.113917.
[16] Du P, Wang J, Yang W, Niu T. A novel hybrid model for short-term wind power forecasting. Appl Soft Comput J 2019;80:93–106. https://doi.org/10.1016/j.asoc.2019.03.035.
[17] Wang J, Yang W, Du P, Niu T. A novel hybrid forecasting system of wind speed based on a newly developed multi-objective sine cosine algorithm. Energy Convers Manag 2018;163:134–50. https://doi.org/10.1016/j.enconman.2018.02.012.
[18] Noorollahi Y, Jokar MA, Kalhor A. Using artificial neural networks for temporal and spatial wind speed forecasting in Iran. Energy Convers Manag 2016;115:17–25. https://doi.org/10.1016/j.enconman.2016.02.041.
[19] Cadenas E, Rivera W. Short term wind speed forecasting in La Venta, Oaxaca, México, using artificial neural networks. Renew Energy 2009;34:274–8. https://doi.org/10.1016/j.renene.2008.03.014.
[20] Qu Z, Mao W, Zhang K, Zhang W, Li Z. Multi-step wind speed forecasting based on a hybrid decomposition technique and an improved back-propagation neural network. Renew Energy 2019;133:919–29. https://doi.org/10.1016/j.renene.2018.10.043.
[21] Zhang D, Peng X, Pan K, Liu Y. A novel wind speed forecasting based on hybrid decomposition and online sequential outlier robust extreme learning machine. Energy Convers Manag 2019;180:338–57. https://doi.org/10.1016/j.enconman.2018.10.089.
[22] Kumar G, Malik H. Generalized Regression Neural Network Based Wind Speed Prediction Model for Western Region of India. Procedia Comput Sci 2016;93:26–32. https://doi.org/10.1016/j.procs.2016.07.177.
[23] Liu H, Mi X wei, Li Y fei. Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network. Energy Convers Manag 2018;156:498–514. https://doi.org/10.1016/j.enconman.2017.11.053.
[24] Wang Y, Wang J, Li Z. A novel hybrid air quality early-warning system based on phase-space reconstruction and multi-objective optimization: A case study in China. J Clean Prod 2020;260:121027. https://doi.org/10.1016/j.jclepro.2020.121027.
[25] Wang S, Zhang N, Wu L, Wang Y. Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method. Renew Energy 2016;94:629–36. https://doi.org/10.1016/j.renene.2016.03.103.
[26] Aly HHH. A novel deep learning intelligent clustered hybrid models for wind speed and power forecasting. Energy 2020;213:118773. https://doi.org/10.1016/j.energy.2020.118773.
[27] Shahid F, Zameer A, Muneeb M. A novel genetic LSTM model for wind power forecast. Energy 2021;223:120069. https://doi.org/10.1016/j.energy.2021.120069.
[28] Wang R, Li C, Fu W, Tang G. Deep Learning Method Based on Gated Recurrent Unit and Variational Mode Decomposition for Short-Term Wind Power Interval Prediction. IEEE Trans Neural Networks Learn Syst 2020;31:3814–27. https://doi.org/10.1109/TNNLS.2019.2946414.
[29] Gan Z, Li C, Zhou J, Tang G. Temporal convolutional networks interval prediction model for wind speed forecasting. Electr Power Syst Res 2021;191:106865. https://doi.org/10.1016/j.epsr.2020.106865.
[30] Li H, Wang J, Lu H, Guo Z. Research and application of a combined model

based on variable weight for short term wind speed forecasting. Renew Energy 2018;116:669–84. https://doi.org/10.1016/j.renene.2017.09.089.
[31] Wang J, Du P, Niu T, Yang W. A novel hybrid system based on a new proposed algorithm—Multi-Objective Whale Optimization Algorithm for wind speed forecasting. Appl Energy 2017;208:344–60. https://doi.org/10.1016/j.apenergy.2017.10.031.
[32] Lv M, Wang J, Niu X, Lu H. A newly combination model based on data denoising strategy and advanced optimization algorithm for short-term wind speed prediction. J Ambient Intell Humaniz Comput 2022. https://doi.org/10.1007/s12652-021-03595-x.
[33] Zhang L, Wang J, Niu X. Wind speed prediction system based on data pre-processing strategy and multi-objective dragonfly optimization algorithm. Sustain Energy Technol Assessments 2021;47:101346. https://doi.org/10.1016/j.seta.2021.101346.
[34] Jiang P, Liu Z. Variable weights combined model based on multi-objective optimization for short-term wind speed forecasting. Appl Soft Comput J 2019;82:105587. https://doi.org/10.1016/j.asoc.2019.105587.
[35] Jiang H, Luo S, Dong Y. Simultaneous feature selection and clustering based on square root optimization. Eur J Oper Res 2021;289:214–31. https://doi.org/10.1016/j.ejor.2020.06.045.
[36] Yu Y, Wang J, Liu Z, Zhao W. A combined forecasting strategy for the improvement of operational efficiency in wind farm . J Renew Sustain Energy 2021;063304. https://doi.org/10.1063/5.0065937.
[37] Zadeh LA. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets Syst 1997;90:111–27. https://doi.org/10.1016/S0165-0114(97)00077-8.
[38] He Y, Yan Y, Xu Q. Wind and solar power probability density prediction via fuzzy information granulation and support vector quantile regression. Int J Electr Power Energy Syst 2019;113:515–27. https://doi.org/10.1016/j.ijepes.2019.05.075.
[39] Cheng X, Guo P. Short-term wind speed prediction based on support vector machine of fuzzy information granulation. 2013 25th Chinese Control Decis Conf CCDC 2013 2013:1918–23. https://doi.org/10.1109/CCDC.2013.6561247.
[40] Pan C, Wang D, Tan Q. Short-term wind speed interval prediction by improved regularized extreme learning machine based on attribute reduction. Wind Eng 2020;44:631–44. https://doi.org/10.1177/0309524X19862762.
[41] Johnson PL, Negnevitsky M, Muttaqi KM. Short term wind power forecasting using Adaptive Neuro-Fuzzy Inference Systems. 2007 Australas Univ Power Eng Conf AUPEC 2007:1–4. https://doi.org/10.1109/AUPEC.2007.4548099.
[42] Erdem E, Shi J. ARMA based approaches for forecasting the tuple of wind speed and direction. Appl Energy 2011;88:1405–14. https://doi.org/10.1016/j.apenergy.2010.10.031.
[43] Torres JL, García A, De Blas M, De Francisco A. Forecast of hourly average wind speed with ARMA models in Navarre (Spain). Sol Energy 2005;79:65–77. https://doi.org/10.1016/j.solener.2004.09.013.
[44] Xie NM, Yuan CQ, Yang YJ. Forecasting China's energy demand and self-sufficiency rate by grey forecasting model and Markov model. Int J Electr Power Energy Syst 2015;66:1–8. https://doi.org/10.1016/j.ijepes.2014.10.028.
[45] Zhang Y, Sun H, Guo Y. Wind power prediction based on pso-svr and grey

combination model. IEEE Access 2019;7:136254–67. https://doi.org/10.1109/ACCESS.2019.2942012.
[46] Guo ZH, Wu J, Lu HY, Wang JZ. A case study on a hybrid wind speed forecasting method using BP neural network. Knowledge-Based Syst 2011;24:1048–56. https://doi.org/10.1016/j.knosys.2011.04.019.
[47] Liu H, Mi X, Li Y. An experimental investigation of three new hybrid wind speed forecasting models using multi-decomposing strategy and ELM algorithm. Renew Energy 2018;123:694–705. https://doi.org/10.1016/j.renene.2018.02.092.
[48] Lee CY, He Y Lou. Wind prediction based on general regression neural network. Proc - 2012 Int Conf Intell Syst Des Eng Appl ISDEA 2012 2012:617–20. https://doi.org/10.1109/ISdea.2012.520.
[49] Wang J, Zhang W, Li Y, Wang J, Dang Z. Forecasting wind speed using empirical mode decomposition and Elman neural network. Appl Soft Comput J 2014;23:452–9. https://doi.org/10.1016/j.asoc.2014.06.027.
[50] Li Y, Wu H, Liu H. Multi-step wind speed forecasting using EWT decomposition, LSTM principal computing, RELM subordinate computing and IEWT reconstruction. Energy Convers Manag 2018;167:203–19. https://doi.org/10.1016/j.enconman.2018.04.082.
[51] Ibrahim M, Alsheikh A, Al-Hindawi Q, Al-Dahidi S, Elmoaqet H. Short-Time Wind Speed Forecast Using Artificial Learning-Based Algorithms. Comput Intell Neurosci 2020;2020. https://doi.org/10.1155/2020/8439719.
[52] Meka R, Alaeddini A, Bhaganagar K. A robust deep learning framework for short-term wind power forecast of a full-scale wind farm using atmospheric variables. Energy 2021;221:119759. https://doi.org/10.1016/j.energy.2021.119759.
[53] Li C, Tang G, Xue X, Saeed A, Hu X. Short-Term Wind Speed Interval Prediction Based on Ensemble GRU Model. IEEE Trans Sustain Energy 2020;11:1370–80. https://doi.org/10.1109/TSTE.2019.2926147.
[54] Qin S, Liu F, Wang J, Song Y. Interval forecasts of a novelty hybrid model for wind speeds. Energy Reports 2015;1:8–16. https://doi.org/10.1016/j.egyr.2014.11.003.
[55] Hu J, Wang J, Xiao L. A hybrid approach based on the Gaussian process with t-observation model for short-term wind speed forecasts. Renew Energy 2017;114:670–85. https://doi.org/10.1016/j.renene.2017.05.093.
[56] Li Q, Wang J, Zhang H. A wind speed interval forecasting system based on constrained lower upper bound estimation and parallel feature selection. Knowledge-Based Syst 2021;231:107435. https://doi.org/10.1016/j.knosys.2021.107435.
[57] Li R, Jin Y. A wind speed interval prediction system based on multi-objective optimization for machine learning method. Appl Energy 2018;228:2207–20. https://doi.org/10.1016/j.apenergy.2018.07.032.
[58] Jiang H, Zheng W, Dong Y. Sparse and robust estimation with ridge minimax concave penalty. Inf Sci (Ny) 2021;571:154–74. https://doi.org/10.1016/j.ins.2021.04.047.
[59] Liu Z, Jiang P, Wang J, Zhang L. Ensemble system for short term carbon dioxide emissions forecasting based on multi-objective tangent search algorithm. J Environ Manage 2022;302:113951. https://doi.org/10.1016/j.jenvman.2021.113951.
[60] Abualigah L, Yousri D, Abd Elaziz M, Ewees AA, Al-qaness MAA, Gandomi

AH. Aquila Optimizer: A novel meta-heuristic optimization algorithm. Comput Ind Eng 2021;157:107250. https://doi.org/10.1016/j.cie.2021.107250.
[61] Mirjalili S, Saremi S, Mirjalili SM, Coelho LDS. Multi-objective grey wolf optimizer: A novel algorithm for multi-criterion optimization. Expert Syst Appl 2016;47:106–19. https://doi.org/10.1016/j.eswa.2015.10.039.
[62] Ma X, Jin Y, Dong Q. 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 2017;54:296–312. https://doi.org/10.1016/j.asoc.2017.01.033.

Credit Author Statement

Qianyi Xing: Writing-review & editing, Visualization, Software, Methodology.
Jianzhou Wang: Supervision, Conceptualization, Funding acquisition.
Haiyan Lu: Validation, Formal analysis.
Shuai Wang: Validation, Visualization.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: