Applied Energy

A Novel Ensemble Probabilistic Forecasting System for Uncertainty in Wind Speed

--Manuscript Draft--

Manuscript Number:	APEN-D-21-10889R1
Article Type:	Research Paper
Keywords:	Wind speed forecasts; Multi-objective optimization algorithm; Deep learning; ensemble probabilistic strategy; Forecast uncertainty
Corresponding Author:	Shuai Wang CHINA
First Author:	Jianzhou Wang
Order of Authors:	Jianzhou Wang
	Shuai Wang
	Bo Zeng
	Haiyan Lu
Abstract:	The quantification of wind speed uncertainty is of great significance for real-time control of wind turbines and power grid dispatching. However, the intermittence and fluctuation of wind energy present great challenges in modeling its uncertainty; research in this field is limited. A quantile regression bi-directional long short-term memory network (QrBiLStm) and a novel ensemble probabilistic forecasting strategy are proposed in this study to explore ensemble probabilistic forecasting. To verify the reliability of the proposed ensemble probabilistic forecasting system, the uncertainties of wind speed at wind farms in China were modeled as a case study. The results of comparative experiments including 15 other models demonstrate the superiority of this ensemble probabilistic forecasting system in terms of sharpness while maintaining high interval coverage. The forecasting interval coverage probability obtained by the proposed system is above 97%, and the sharpness is improved by at least 24.21% as compared with the commonly used single models. The proposed ensemble probabilistic forecasting system can accurately quantify the uncertainty of wind speed and reduce the operation cost of power systems by improving the efficiency of wind energy utilization.

Dear Editor,

We would like to submit the enclosed manuscript entitled '**A Novel Ensemble Probabilistic Forecasting System for Uncertainty in Wind Speed**', which we revised as the suggestions and we also wish to be considered for publication in **Applied Energy**. The original title was " Uncertainty Wind Speed Modeling Based on Quantile Regression Bi-directional Long Short Term Memory Network, Ensemble Probabilistic Forecasting Strategy, and Improved Swarm Intelligence Optimization " and has been revised to " A Novel Ensemble Probabilistic Forecasting System for Uncertainty in Wind Speed " based on the comments of the reviewers. Thank you very much for emailing us the comments raised by the respected reviewers. This manuscript is our own work and the content of this paper has not been copied from elsewhere. This manuscript has not been published before nor submitted to another journal for the consideration of publication and all data measurements are genuine results and have not been manipulated. In addition, none of the authors have any financial or scientific conflicts of interest with regard to the research described in this manuscript. And this manuscript was supported by the **National Natural Science Foundation of China** (No. 71671029).
All authors have seen the revised manuscript and approved to submit to your journal. Thank you very much for your attention and consideration.

Sincerely yours,
Shuai Wang

Dear editors and reviewers

Thank you very much for e-mailing us the comments raised by the respected editors/reviewers. The manuscript **No. APEN-D-21-10889 "A Novel Ensemble Probabilistic Forecasting System for Uncertainty in Wind Speed**" has been revised taking into account all of the helpful comments and suggestions. The details of the comments raised, the answers and the actions taken are presented here. All the changes made in the new revised manuscript have been marked in yellow. We appreciate for respected editors/reviewers' warm work earnestly, and hope that the correction will meet with approval. We look forward to hearing from you.

Best regards,
Shuai Wang

Comment raised by respected Reviewer #1:

Comment 1: The title is too long. Please shorten it.

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. We changed the title to "A Novel Ensemble Probabilistic Forecasting System for Uncertainty in Wind Speed"

****

Comment 2: Abstract needs to modify and to be revised to be more quantitative. You can absorb readers' consideration by having some numerical results in this section.

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. We revised the abstract and added the quantitative results of evaluation metric values. Please see the revised abstract (the modified places are marked in yellow).

The revised abstract:

The quantification of wind speed uncertainty is of great significance for real-time control of wind turbines and power grid dispatching. However, the intermittence and fluctuation of wind energy present great challenges in modeling its uncertainty; research in this field is limited. A quantile regression bi-directional long short-term memory network (QrBiLStm) and a novel ensemble probabilistic forecasting strategy are proposed in this study to explore ensemble probabilistic forecasting. To verify the reliability of the proposed ensemble probabilistic forecasting system, the uncertainties of wind speed at wind farms in China were modeled as a case study. The results of comparative experiments including 15 other models demonstrate the superiority of this ensemble probabilistic forecasting system in terms of sharpness while maintaining high interval coverage. The forecasting interval coverage probability obtained by the proposed system is above 97%, and the sharpness is improved by at least 24.21% as compared with the commonly used single models. The proposed ensemble probabilistic forecasting system can accurately quantify the uncertainty of wind speed and reduce the operation cost of power systems by improving the efficiency of wind energy utilization.

****

Comment 3: Novelty / Contribution of the paper should be more clear.

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. We modified the contribution part, combining the advantages of the proposed model with wind energy utilization and power system management. Please see the revised contributions (the modified places are marked in yellow).

The revised contributions:

The main innovations and contributions of this study are summarized as follows:

(1) **The deep QrBiLStm model for wind speed uncertainty modeling was successfully designed, implemented, and tested.** The proposed QrBiLStm model can obtain interval forecasting results with high interval coverage probability and narrower interval width, and provide more accurate information for wind energy utilization.

(2) **A pseudo interval was proposed, and pseudo-interval evaluation indicators were successfully designed** as a foundation for the ensemble probabilistic forecasting system (EPFS). The pseudo-interval training approach enables separate optimization of the upper and lower bounds of the interval. Optimized wind-speed interval forecasting results are more accurate which means less wind estimation fluctuation and less uncertainty. Thus, the proposed EPFS is of great significance to the safety dispatch and operation of wind power generation.

(3) **An ensemble probabilistic forecasting system was proposed, and optimization objective functions for ensemble forecasting were designed.** The experimental results show that the proposed EPFS based on QrBiLStm is a significant improvement over the single models. The EPFS overcomes the limitations of the single model forecast, making the wind speed forecasting results more stable and practical.

(4) **The tunicate swarm algorithm (TSa) was improved and used to perform interval ensemble optimization. The Tsa with the addition of archiving and a roulette wheel can output Pareto optimal solutions.** Comparative experiments show that the TSa with three improved strategies has better global optimization ability and more stable optimization. The improved Tsa can ensure more stable wind-speed interval forecasting results at a faster speed.

(5) Based on singular spectral analysis (SSa) and phase space reconstruction (PSr), the original wind speed sequence was decomposed and reconstructed, enabling the ensemble forecasting model to solve the chaos phenomenon and eliminate small fluctuations, with better forecasting results.

****

Comment 4: The authors have to add the state-of-the art references in the manuscripts.

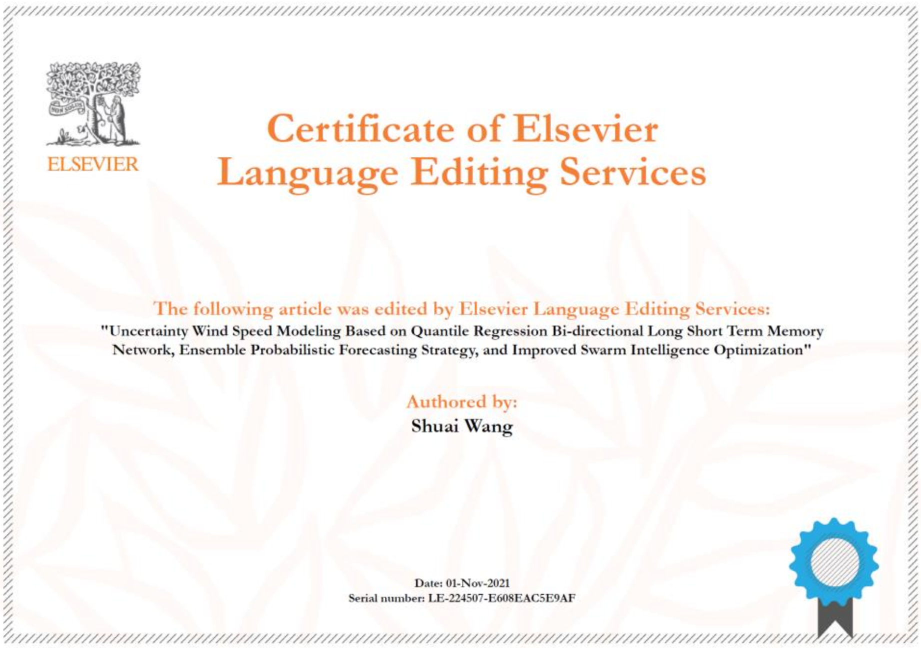
(1)A novel decomposition-ensemble learning framework for multi-step ahead wind energy forecasting
(2)Ultra-short-term wind speed forecasting based on EMD-VAR model and spatial correlation
(3)Hybrid multi-stage decomposition with parametric model applied to wind speed forecasting in Brazilian Northeast
(4)Data-augmented sequential deep learning for wind power forecasting
(5)Efficient bootstrap stacking ensemble learning model applied to wind power generation forecasting
(6)Wind turbines anomaly detection based on power curves and ensemble learning

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. We have carefully read all of the latest literature you have provided and considered its relevance to our article. We found all of the articles mentioned to be very informative and added these references where appropriate. The numbers of the references are: [57], [56], [51], [14], [32], and [12]. Please see the revised manuscript.

****

Comment 5: There are some occasional grammatical problems within the text. It may need the attention of someone fluent in English language to enhance the readability.

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. We have adopted three approaches to make our article more readable. (1) This article has been edited by Elsevier Language Editing Services.


ELSEVIER

Certificate of Elsevier Language Editing Services

The following article was edited by Elsevier Language Editing Services:

"Uncertainty Wind Speed Modeling Based on Quantile Regression Bi-directional Long Short Term Memory Network, Ensemble Probabilistic Forecasting Strategy, and Improved Swarm Intelligence Optimization"

Authored by:

Shuai Wang

Date: 01-Nov-2021

Serial number: LE-224507-E608EAC5E9AF

(2) We checked the language of the full article, and revised the incoherent places.
(3) We checked the grammar of the revised contents and additions.

****

Comment 6: The discussion section in the present form is relatively weak and should be strengthened with more details and justifications.

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. We mainly through the following three aspects to improve the discussion of this article.

(1) Our comparative experiment also includes the discussions and remarks of the model performance. In order to enhance readability, we combine the "**4. Experimental results and analysis**" and "**5. Discussion**" into "**4. Experimental results and discussion**".
(2) We added DM test in the discussion to verify whether the proposed model is significantly different from other models.
(3) We added the advantages and disadvantages of the proposed EPFS and future research directions in the discussion.
The current "Discussion" section includes: (1) 4.1~4.2 comparative discussion of different models; (2) 4.3. Statistical test (DM-test); (3) 4.4.Fist-order and second-order forecasting effectiveness evaluation； (4) 4.5. Improvement ratio； (5) 4.6. Stability analysis； (6) 4.7. Advantages and disadvantages compared to the existing studies.

Please see pages 14-28 of the revised manuscript (the modified places are marked in yellow).

****

Comment 7: In tables 2 and 3 the best results should be in bold font.

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. We have highlighted the best results in Table 2 and Table 3 in bold. Please see the revised manuscript (the modified places are marked in yellow).

****

Comment 8: Was cross validation performed on the data? How many folds were used?

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. We did not perform cross-validation on the data. Because conventional K-fold cross-validation is not suitable for time series data. For example, when we do the K-fold cross-validation, we may use the future data as the training set and the past data as the testing set, so that the observed error indicator values and model performance have no practical significance. Therefore, instead of doing the K-fold cross-validation, (1) we determined the hyperparameter values by the performance of the model on the validation set, and (2) we did a stability analysis in "Discussion" to illustrate the changes in the results of our model in multiple predictions.

****

Comment 9: What about the hyperparameters of optimization techniques? How were they used?

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. All the hyperparameters required by the model, the determined parameter values and the methods of determining the hyperparameter values are described in detail in Table A1 of Appendix A (page 30).

****

Comment 10: What about prediction data, how was it divided into training, validation, and testing? Presented in Section 3, item 1, explain better.

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. We added the "3.1 Data description" section. Here we describe the data collection information and the division of datasets in detail. In addition, we share the datasets behind the 'conclusion' for readers to use.
Please see page 12 of the revised manuscript (the modified places are marked in yellow).

****

Comment 11: What are the limitations behind this study? This topic should be highlighted somewhere in the text of manuscript.

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. In the "Conclusion" section, we added the limitations of this study and described the future work. Please see page 29 of the revised manuscript (the modified places are marked in yellow).

****

Comment 12: What are the advantages and disadvantages of this study compared to the existing studies in this area? This topic should be highlighted somewhere in the text of manuscript.

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. Section 4.7 "Advantages and disadvantages compared to the existing studies" (pages 24-25) is set up to highlight this aspect.
Please see the revised manuscript (the modified places are marked in yellow).

****

Comment 13: In conclusion section, limitations and recommendations of this research should be highlighted.

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. In the conclusion section, we highlight the limitations and the contribution of this paper. The revised part of conclusion reads as follows:

The main limitations of the proposed EPFS are as follows: (1) Because quantile loss is discontinuous and nondifferentiable around 0 point, this EPFS has not been applied to the field of deterministic forecasts; (2) It is not combined with other linear models or interval forecasting models based on distribution in ensemble forecasting. However, this EPFS can optimize the upper and lower bounds of the interval separately and does not need to assume the distribution in advance. This study provides a novel approach for wind speed ensemble probabilistic forecasting and can be used as a powerful decision tool in the power system scheduling process.

Comment raised by respected Reviewer #2:

Comment: The paper is well written with detailed technical content. I would recommend publication of this paper after some minor corrections are made and/or clarification questions raised above are properly addressed.

Response: Thank you for your comments and they are all valuable and very helpful for revising and improving our paper. We have studied the comments carefully and made corrections in the revised manuscript which we hope to meet with approval.

****

Comment 1: The acronym IMOTa, which was mentioned in the Introduction first, was later defined in Line 138 as Improved Multi-Objective Tunicate swarm algorithm (IMOTa). It would be good to define the acronym when it first appears in the manuscript. I also notice that all acronyms are defined in Table A2 (Line 526).

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. In the Introduction section (page 3), we redefined the Tunicate swarm algorithm as a TSa. In Section 2.3, Part B (page 9), we define the Improved Tunicate Swarm Optimization Algorithm as IMOTA.
Please see the revised manuscript (the modified places are marked in yellow).

****

Comment 2: Fig. 5: Although it is obvious that SPICP is the same as std(PICP), it should be defined. The same argument applies to SPINAW, SAIS.

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. We have changed S(PICP), S(PINAW), and S(AIS) in Fig.5 to std(PICP),std(PINAW) and std(AIS), respectively.
Please refer to Fig.5 in the revised manuscript.

****

Comment 3: Table 5, I think PIMWP should be PINAW. In this table, only PINAW and AIS are considered. Why not also considering PICP?

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. We changed the PIMWP in Table 6 to PINAW. For the problem that the PICP is not considered in the lifting rate: we show in the paper that the proposed EPFS can significantly reduce the interval width, but in some cases it will slightly reduce the interval coverage probability. Therefore, in order to comprehensively measure the final interval prediction effect, we observe the improvement rate of AIS, a comprehensive index, to reflect the total improvement effect of the proposed EPFS, and use the improvement rate of PINAW to reflect the effect of EPFS on shortening the interval width. In addition, in Section 4.5 (page 23, lines 484-487) of the revised manuscript, we explained why we only consider the promotion rate of PINAW and AIS. It reads as follows:

In this paper, the improvement rate of AIS is calculated to reflect the overall improvement ratio of the proposed EPFS compared with other models, and the improvement rate of PINAW reflects the contribution of the proposed EPFS to shortening the interval width.

****

Comment 4: Lines 261-262, it was mentioned that the parameter settings for all models are presented in Table A1 in Appendix A. Which machine-learning software was used in this study? Will the dataset be made available for interested readers to reproduce the results in this paper?

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. (1) We explained before section 4.1 (page 14, line 277) that the experiments were based on Matlab2020a. It reads as follows:

"The proposed EPFS and comparative models were implemented on Matlab2020a."

(2) We added a data availability section to the appendix to make our data publicly available for replication by other readers.

Data Availability

10-minute wind speed data of three Sites in Shandong Peninsula:

https://data.mendeley.com/datasets/sjyf2nhzdt/draft?a=af12330a-125b-499a-9473-6840ed7044f9

Please refer to the revised manuscript (the modified places are marked in yellow).

****

Comment 5: Lines 270-271, Eq. (25),

a. N is the length of the interval to be predicted. Can you elaborate N (an integer) here or how N is determined?

b. α is the confidence level. How many α were used in this study? I presume α=0.1 and 0.05 (see Lines 460-461).

c. Is PINC=1- α (in %)? For example, is α=0.1 the same as PINC=90% (or 1-0.1)?

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper.

a. N is the length of the divided testing set. In section 3, we added a more detailed description to the division of the data set.

b. We used $\alpha = 0.05$ and $\alpha = 0.1$ in this study.

c. $PINC = (1-\alpha)\times 100\%$.

In addition, we also explained the above problems in this paper, and the contents (page 14, lines 288-292) are as follows:

a. where N is the length of the testing set.

b. In this paper, interval prediction is implemented based on $\alpha = 0.05$ and $\alpha = 0.1$

c. confidence levels. And $PINC = (1-\alpha)\times 100\%$.

Please see page 14 of the revised manuscript (revised places are marked in yellow).

****

Comment 6: The symbol $\xi(i)$ was used twice in Eqs. (28) and (29), respectively, with different meanings. The authors could use, e.g., $\xi(i)$ and $\tilde{\xi}(i)$ to be different.

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. We have changed $\xi(i)$ to $\tilde{\xi}(i)$ in formula (29). Accordingly, we have also modified the text description of symbols.

Please refer to the revised manuscript (the modified places are marked in yellow).

****

Comment 7: Lines 280-281, Eq. (26), I believe that N here is the same N as in Eq. (25). R is the range of observation values. Is R an integer? How is R determined?

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. N in the Eq. (26) is the same as N in Eq. (25), we explained this point in section 4.1 (page 14, lines 298-299) of the revised manuscript. R is an integer, it is determined by the maximum of the observation values on the testing set minus its minimum.

Please refer to the revised manuscript (the modified places are marked in yellow).

****

Comment 8: Line 282, Sharpness, a reference about the definition of sharpness is needed.

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. We had added a reference about the definition of sharpness in section 4.1 (page 14, line 303). It reads as follows:

The sharpness combines the two metrics for assessment of prediction intervals [50];

[50] Cui M, Krishnan V, Hodge BM, Zhang J. A Copula-Based Conditional Probabilistic Forecast Model for Wind Power Ramps. IEEE Trans Smart Grid 2019. https://doi.org/10.1109/TSG.2018.2841932.

Please refer to the revised manuscript (the modified places are marked in yellow).

****

Comment 9: Lines 312-313, “A smaller PINAW indicates a narrower interval width with more uncertainty information.” Can you explain a narrower interval width indicates more uncertainty information?

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. First, let's explain it by an example. Case1: Tomorrow's temperature is between 10 and 20 degrees; Case2: Tomorrow's temperature will be between 12 and 18 degrees. Obviously, the uncertainty of the case 2 is smaller, and the more (uncertainty) information it contains.

In addition, we added " The PINAW metric indicates the interval width which determines the practicality and informative of interval [51]" and added references to it in the revised version.

[51] Quan H, Srinivasan D, Khosravi A. Uncertainty handling using neural network-based prediction intervals for electrical load forecasting. Energy 2014. https://doi.org/10.1016/j.energy.2014.06.104.

Please refer to page 15 (lines 332-333) of the revised manuscript.

****

Comment 10: Lines 328-329, “The AIS value is generally less than 0; a higher AIS value indicates a more effective forecasting interval.” Do you mean a higher AIS value (<0) is for a lower absolute value of AIS?

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. Yes, a higher AIS value (<0) is for a lower absolute value of AIS. In order to facilitate readers to check the optimal results, the optimal results in Table 2 and Table 3 are marked in bold. Please see the revised manuscript.

****

Comment 11: Fig. 3, AIS bar charts, colors for QrBiLStm and Proposed EPFS are the same (Blue), and they should be different to be distinguishable.

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. We've changed the color of the Proposed EPFS to a different color to make it easier to distinguish.

Please refer to Fig. 3 of the revised manuscript.

****

Comment 12: Lines 399-400,

a. “…the forecasting effectiveness (FE) approach [52] was modified for uncertainty forecasting.” How was FE in [52] modified?

b. “The required bias in FE is defined as AIS…”. I believe AIS here is the same as that defined in Eq. (27). Please explain.

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper.

a. Literature [52] is the original FE, which is mainly used to evaluate the point prediction results. In this paper, we modified it to measure the results of interval prediction. How to modified it is explained in b.

b. In the FE formula of point prediction, the bias (that is, the real value minus the predicted value) is needed. In the interval prediction, we also need a bias, so we employ “interval score” (part of the AIS), which can measure both the interval coverage and the interval width, as bias of interval forecast. We also revised the original description as “The required bias in FE is modified as interval score which is defined as Eq. (31)”.

Please see section 4.4 of the revised manuscript (page 22, lines 455-456).

****

Comment 13: Line 408, $Q_i = 1/n, i = 1, 2, \ldots, n$. I think $Q_i = i/n$. Is n here same as N in Eq. (27)? Again, are N in Eqs. (25)-(27) the same?

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. Firstly, $Q_i = 1/n, i = 1, 2, \ldots, n$ is correct; because $\sum_{i=1}^{n} \boldsymbol{Q}_i = 1, \boldsymbol{Q}_i > 0$, Because the prior information cannot be obtained, equal probability is taken for each observation point on the test set, so $Q_i = 1/n$. Secondly, n is same as N in Eq. (27). We also added "where n is the length of the testing set" (page 22, lines 459-460) to explain this point in section 4.4 of the revised manuscript.

****

Comment 14: Line 413, "…the second-order FE is defined as $FE\left(\boldsymbol{g}^1, \boldsymbol{g}^2\right) = \boldsymbol{g}^1\left(1 - \sqrt{\boldsymbol{g}^2 - \left(\boldsymbol{g}^1\right)^2}\right)$

a. Please explain this definition.

b. How can this definition be extended to $FE\left(\boldsymbol{g}^1, \boldsymbol{g}^2, \boldsymbol{g}^3\right)$

c. Is there any advantage to include the 3rd-order FE in the analysis to assess the forecasting performance?

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper.

a. This definition is given by reference [52], and this paper only applies this index to evaluate the prediction results from another angle. Specific information can refer to these two articles: "Research on model for combination forecasting based on second-order forecast effective measure" and "Research on superior combination forecasting model based on forecasting effective measure.".

b.c. At present, the research only defines the first-order and second-order forecasting effectiveness. Please refer to two articles in "a." for specific information.

****

Comment 15: Lines 439-441, "…increased by more than 30% compared with EPFMs based on other denoising methods. Compared with EPFMs based on other optimization algorithms, the improvement ratio is approximately 5–10%..."

a. How was 30% above related to the four denoising methods calculated? IR at Site 1 for PINC=90% and 95% is between 0.1894 and 0.4774.

b. How was 5-10% above related to the four optimization algorithms calculated? IR at Site 1 for PINC=90% and 95% is between 0.0208 (MOAa for PINC=90%) and 0.2360 (MODa for PINC=95%).

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper.

a. The 30% value is the average of the AIS values for the four denoising methods at the three Sites with PINC=90% and PINC=95%. We found the description a bit vague, so we revised the sentence. It reads as follows:

Taking into account $PINC = 90\%$ and $PINC = 95\%$ of the three Sites, the proposed EPFS represents a minimum of 12.96% improvement in AIS metrics compared to the EPFMs based on the other four denoising methods.

b. 5% to 10% is determined by the average of the AIS values for the four optimization algorithms at the three Sites (considering PINC=90% and PINC=95%). We found the description a bit vague, so we revised the sentence. It reads as follows:

Compared with EPFMs based on other optimization algorithms, the improvement ratio of the AIS metric is 1.57% to 23.60%.

Please refer to Section 4.5 (page 24, lines 496-500) of the revised manuscript (the modified places are marked in yellow).

****

Comment 16: Line 482, Note, what is P_{MAPE}?

Response: Thank you for the valuable advice and it is quite helpful for improving the quality of our paper. This is our writing mistake. We modified the Note in Table 6 (page 26), which reads as follows:

Note: The table above reports the IR of the proposed EPFS from other twelve models. The AIS and PINAW are used to measure the IR, and the corresponding indicator can be defined as $\overline{\boldsymbol{IR}}_{Metric} = \left[\left(\boldsymbol{Metric}^{com} - \boldsymbol{Metric}^{pro} \right) / \boldsymbol{Metric}^{com} \right]$, where $\boldsymbol{Metric}^{com}$ is the metric values of compared model, and the $\boldsymbol{Metric}^{pro}$ indicates the metric value of the proposed **EPFS**.

****

- Developed a novel deep QrBiLStm that can perform probabilistic forecasts.
- An ensemble probabilistic forecasting system was proposed.
- A pseudo-interval training method for ensemble probabilistic forecast is designed.
- Improved the optimizer with three strategies.

A Novel Ensemble Probabilistic Forecasting System for Uncertainty in Wind Speed

Jianzhou Wang[a], Shuai Wang[b,*], Bo Zeng[c], Haiyan Lu[d]
[a] *Institute of Systems Engineering, Macau University of Science and Technology, Macau 999078, China*
[b] *School of Statistics, Dongbei University of Finance and Economics, Dalian, China*
[c] *Collaborative Innovation Center for Chongqing's Modern Trade Logistics & Supply Chain, Chongqing Technology and Business University, Chongqing 400067, People's Republic of China*
[d] *School of Computer Science, 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 18742018422.
E-mail address: vvs09061513@163.com.

Abstract

The quantification of wind speed uncertainty is of great significance for real-time control of wind turbines and power grid dispatching. However, the intermittence and fluctuation of wind energy present great challenges in modeling its uncertainty; research in this field is limited. A quantile regression bi-directional long short-term memory network (QrBiLStm) and a novel ensemble probabilistic forecasting strategy are proposed in this study to explore ensemble probabilistic forecasting. To verify the reliability of the proposed ensemble probabilistic forecasting system, the uncertainties of wind speed at wind farms in China were modeled as a case study. The results of comparative experiments including 15 other models demonstrate the superiority of this ensemble probabilistic forecasting system in terms of sharpness while maintaining high interval coverage. The forecasting interval coverage probability obtained by the proposed system is above 97%, and the sharpness is improved by at least 24.21% as compared with the commonly used single models. The proposed ensemble probabilistic forecasting system can accurately quantify the uncertainty of wind speed and reduce the operation cost of power systems by improving the efficiency of wind energy utilization.

Keywords: *Wind speed forecasts; Multi-objective optimization algorithm; Deep learning; ensemble probabilistic strategy; Forecast uncertainty*

1. Introduction

Wind energy has attracted extensive attention as an inexhaustible, clean, and inexpensive form of renewable energy. According to the Global Wind Report released by GWEC in 2021, the 93GW of new installations brings global cumulative wind power capacity up to 743 GW [1]. However, volatility and randomness of wind energy pose great challenges to wind energy grid connection and grid scheduling [2]. Decision-makers must calculate and process the forecasted wind speed to obtain corresponding energy information [3]. Thus, wind-speed forecasting is critical for wind energy utilization.

Wind speed forecasting approximates or extracts the potential relationship behind the data; point-oriented forecasting is the most common form [4]. A data-driven model of point forecasting can use traditional statistical models and artificial intelligence models. Traditional statistical models include autoregressive moving average (ARMA) [5], autoregressive integrated moving average (ARIMA) [6], and Kalman filtering [7], etc. These models are based on a linear assumption, and produce forecasting results that are not accurate with nonlinear sequences [8]. With continuous development of artificial intelligence technology, researchers have begun to apply artificial neural networks to wind speed forecasts. Shallow models including the back propagation neural network (BPNN) [9], the extreme learning machine (ELM) [10], and support vector regression (SVR) [11] were first used. This type of supervised AI model can capture the nonlinear characteristics of wind speed series and manage long series [12]; the forecasting accuracy is higher than that of traditional statistical models [13]. However, there are some defects such as under-fitting, over-fitting, and long training time. With the development of deep learning technology, variants of recurrent neural networks (RNNs) [14] such as long short-term memory (LStm), gated recurrent units (GRu), and BiLStm networks have demonstrated excellent performance in time series forecasting [15]. This type of model can store historical information and facilitate capture of nonlinear features in wind speed series. These models often have many hyperparameters that must be set and weights to be updated; thus, they are subject to long training time and difficulty in parameter optimization [16].

As wind speed data usually fluctuate, point-oriented forecasting can be inaccurate for grid scheduling purposes; thus, interval forecasting has become popular [17]. Interval forecasting approaches include mean-variance estimates (MVE) [18], bootstrap [19], Bayesian [20], and the lower upper bound estimation (LUBE) [21]. These methods have advantages and disadvantages, summarized in **Table 1**. Quantile regression (Qr) [22] is usually used in uncertainty forecasting for its strong interpretability in estimating the conditional distribution of the dependent variable. With the limitations of Qr with nonlinear series, research has begun to focus on combining Qr and artificial intelligence models to expand uncertainty forecasting ability [23,24]. In 2000, Taylor [25] proposed a method for combining Qr with a neural network that could solve both linear and nonlinear problems. Based on QrNN, researchers began to combine Qr with other single models. Support Vector Quantile Regression (SVQr) [26] was developed to forecast the probability density of short-term wind power, and can effectively quantify the uncertainty of time series data. He et al [27] forecasted the probability density of electricity consumption based on QrLASso. This method can better learn high-dimensional data, with more accurate forecasting results. As RNNs have more advantages in time series forecasting, researchers have combined Qr with LStm and GRU, proposing QrLStm [28] and QrGRu [29], which further improve forecasting accuracy. Wang et al. [30] incorporated Qr into a convolution-simplified long-term and short-term memory network. This improved model shortened the training time without reducing the accuracy. Based on these studies, we incorporated Qr with BiLStm, proposing QrBiLStm to quantify the uncertainty of wind speed.

The shortcomings of a single model are obvious. In practical applications, the forecasting accuracy of a single model can be high or low in different regions. Thus, another focus of this study is the ensemble forecasting strategy [31]. The ensemble model weighs several well-performing models according to errors using an intelligent optimization algorithm [32]; forecasting is more stable and accurate than with single models [33]. Ensemble forecasting research focuses mostly on point-oriented forecasting. Liu et.al. [34] developed a multi-objective version of the mayfly optimization algorithm, combining several accurate single models to achieve more accurate forecasting. Wang et.al. [35] proposed the addition of two deep learning models to the ensemble forecasting framework, and used the improved dragonfly optimization algorithm to obtain more accurate point forecasting results. However, research on ensemble probabilistic forecasting has received little attention, limiting further development. Niu et al. [36] proposed the use of multiple single models for interval forecasting based on the distribution assumption, and used an optimization algorithm to integrate the results of the single models to obtain the final forecasting results. This approach provides ideas for ensemble probabilistic forecasting. However, with the need to fit the data distribution and estimate the parameters, its usability is limited in practice. The accuracy of the ensemble model depends on the forecasting performance of single models; thus, we propose two QrBiLStm models with excellent performance as benchmark models and use an improved optimization algorithm to realize ensemble probabilistic forecasting.

The main innovations and contributions of this study are summarized as follows:

(1) **The deep QrBiLStm model for wind speed uncertainty modeling was successfully designed, implemented, and tested.** The proposed QrBiLStm model can obtain interval forecasting results with high interval coverage probability and narrower interval width, and provide more accurate information for wind energy utilization.

(2) **A pseudo interval was proposed, and pseudo-interval evaluation indicators were successfully designed** as a foundation for the ensemble probabilistic forecasting system (EPFS). The pseudo-interval training approach enables separate optimization of the upper and lower bounds of the interval. Optimized wind-speed interval forecasting results are more accurate which means less wind estimation fluctuation and less uncertainty. Thus, the proposed EPFS is of great significance to the safety dispatch and operation of wind power generation.

(3) **An ensemble probabilistic forecasting system was proposed, and optimization objective functions for ensemble forecasting were designed.** The experimental results show that the proposed EPFS based on QrBiLStm is a significant improvement over the single models. The EPFS overcomes the limitations of the single model forecast, making the wind speed forecasting results more stable and practical.

(4) **The tunicate swarm algorithm (TSa) was improved and used to perform interval ensemble optimization. The Tsa with the addition of archiving and a roulette wheel can output Pareto optimal solutions.** Comparative experiments show that the TSa with three improved strategies has better global optimization ability and more stable optimization. The improved Tsa can ensure more stable wind-speed interval forecasting results at a faster speed.

(5) Based on singular spectral analysis (SSa) and phase space reconstruction (PSr), the original wind speed sequence was decomposed and reconstructed, enabling the ensemble forecasting model to solve the chaos phenomenon and eliminate small fluctuations, with better forecasting results.

Table 1

Advantages and disadvantages of wind speed forecasting models.

Models	References	Advantages	Disadvantages
ARMA, ARIMA, and Kalman filtering	[6, 7]	The model is simple and only needs endogenous variables; Accurately forecast the linear sequences.	Low forecasting accuracy in nonlinear data; The data is required to be stable or differentially stable.
AI Model (BPNN, ELM, SVR, LStm, and GRu)	[9–11], [37]	Strong robustness and fault tolerance to noise data; Have the ability of association, and can approximate any nonlinear relationship.	The calculation burden is high, and the interpretability is poor; It is difficult to determine the hyperparameter values.
Ensemble Model	[34], [38,39]	The forecasting accuracy on different data types can be ensured; Take advantages of each single model.	Need to train multiple models and choose efficient empowerment technique.
MVE	[18,40]	The computational burden is relatively small.	The accuracy is largely affected by the effect of numerical predictions associated with it; the underestimation of data variance will result in low coverage of real data by prediction intervals.
Bootstrap	[41,42]	High efficiency in small-scale data.	Is a resampling method that requires significant computational cost for large data sets.
Bayesian	[20]	Improve the generalization ability of model.	The calculation burden is large, which requires the calculation of the Hessian matrix. When the data size is not large enough, the accuracy largely depends on prior knowledge.
LUBE	[21,43]	It avoids the problem of numerical calculation of the Jacobian matrix and Hessian matrix.	Heavy computational burden. No suitable parameter initialization method.
Quantile Regression (Qr)	[25,44]	Ability to resolve heterogeneity issues; Tail features of the distribution can be captured.	Traditional Qr model can't solve nonlinear problems, so it is necessary to select a suitable neural network to combine with Qr.

2. Ensemble probabilistic forecasting system (EPFS)

In the EPFS, SSa [42] is used to decompose the reconstructed sequence, and PSr [43] is used to reconstruct an one-dimensional sequence into a dynamic chaotic space. The processed sequences are forecasted in two QrBiLStm units. The proposed IMOTa algorithm is used to aggregate the two QrBiLStm units to generate an effective wind speed forecasting interval. The details of the QrBiLStm, SSa, PSr, and IMOTa algorithms are described as follows.

2.1. Quantile Regression Bi-directional Long Short Term Memory Network

This section introduces the basic structure of BiLStm and the generation of QrBiLStm.

2.1.1. Bi-directional Long Short Term Memory Network

LStm proposed by S. Hochreiter [44], and is an RNN variant [45]. Owing to its cell structure, LStm can solve the problems of gradient disappearance and gradient explosion in long-sequence training. The cell structure consists of an input gate ($\boldsymbol{\Theta}_t$), a forgetting gate ($\boldsymbol{\Pi}_t$), and an output gate ($\boldsymbol{\Omega}_t$); the structure is shown in **Figure.1A**.

$$\begin{cases} \boldsymbol{\Lambda}_t = f\left(\boldsymbol{\Lambda}_{t-1}, \boldsymbol{\mathcal{O}}_t\right) \\ \boldsymbol{\Theta}_t = simoid\left(\boldsymbol{W}_{\boldsymbol{\Theta}} \times \left[\boldsymbol{\Lambda}_{t-1}, \boldsymbol{\mathcal{O}}_t\right] + \boldsymbol{B}ias_{\boldsymbol{\Theta}}\right) \\ \boldsymbol{\Pi}_t = sigmoid\left(\boldsymbol{W}_{\boldsymbol{\Pi}} \times \left[\boldsymbol{\Lambda}_{t-1}, \boldsymbol{\mathcal{O}}_t\right] + \boldsymbol{B}ias_{\boldsymbol{\Pi}}\right) \\ \boldsymbol{\Omega}_t = sigmoid\left(\boldsymbol{W}_{\boldsymbol{\Omega}} \times \left[\boldsymbol{\Lambda}_{t-1}, \boldsymbol{\mathcal{O}}_t\right] + \boldsymbol{B}ias_{\boldsymbol{\Omega}}\right) \\ \overline{\overline{\boldsymbol{\Phi}}} = \tanh\left(\boldsymbol{W}_{\boldsymbol{\Phi}} \times \left[\boldsymbol{\Lambda}_{t-1}, \boldsymbol{\mathcal{O}}_t\right] + \boldsymbol{B}ias_{\boldsymbol{\Phi}}\right) \\ \boldsymbol{\Phi}_t = \boldsymbol{\Pi}_t \times \boldsymbol{\Phi}_{t-1} + \boldsymbol{\Theta}_t \times \overline{\overline{\boldsymbol{\Phi}}} \\ \boldsymbol{\Lambda}_t = sigmoid\left(\boldsymbol{W}_{\boldsymbol{\Omega}} \times \left[\boldsymbol{\Lambda}_{t-1}, \boldsymbol{\mathcal{O}}_t\right] + \boldsymbol{B}ias_{\boldsymbol{\Omega}}\right) \times \tanh\left(\boldsymbol{\Phi}_t\right) \end{cases} \quad (1)$$

In Eq. (1), $\overline{\overline{W}}$ and $\overline{\overline{B}}ias$ represent the weight and bias of LStm cells, respectively; $\boldsymbol{\Phi}_t$ is the current cell state, $\overline{\overline{\boldsymbol{\Phi}}}$ is the candidate cell state, and $\tanh(\bullet)$ represents a hyperbolic tangent function.

BiLStm [46] is composed of a forward LStm layer and a backward LStm layer. In the forward layer, the sequence $\mathcal{O}_t$ is input into the LStm model to calculate the output state $\overrightarrow{\overline{\overline{\boldsymbol{\Lambda}}}}_{t,i}$. In the backward layer, the inverse form of the input sequence is input into the LStm model to calculate the reverse layer output state $_{t,i}\overleftarrow{\overline{\overline{\boldsymbol{\Lambda}}}}$. This structure can extract the forward and backward relations of the wind speed series and connect them to the same output. The network structure is illustrated in **Figure 1B**.

The output of the BiLStm layer at time *t* is $\boldsymbol{\Lambda}_t = \left[\boldsymbol{\Lambda}_{t,1}, \boldsymbol{\Lambda}_{t,2}, \cdots, \boldsymbol{\Lambda}_{t,i}, \cdots, \boldsymbol{\Lambda}_{t,T}\right]^{\mathbf{T}}$, where $\boldsymbol{\Lambda}_{t,i}$ contains $\overrightarrow{\overline{\overline{\boldsymbol{\Lambda}}}}_{t,i}$ and $_{t,i}\overleftarrow{\overline{\overline{\boldsymbol{\Lambda}}}}$ which can be expressed as Eq. (2).

$$\begin{cases} \overrightarrow{\overline{\overline{\boldsymbol{\Lambda}}}}_{t,i} = \overrightarrow{\overline{\overline{F}}}_{LStm}\left(\overrightarrow{\overline{\overline{\boldsymbol{\Lambda}}}}_{t,i-1}, \mathcal{O}_t, \overrightarrow{\overline{\overline{\boldsymbol{\Phi}}}}_{t,i-1}\right); \quad i \in [1,T] \\ {}_{t,i}\overleftarrow{\overline{\overline{\boldsymbol{\Lambda}}}} = {}_{LStm}\overleftarrow{\overline{\overline{B}}}\left({}_{t,i+1}\overleftarrow{\overline{\overline{\boldsymbol{\Lambda}}}}, \mathcal{O}_t, {}_{t,i+1}\overleftarrow{\overline{\overline{\boldsymbol{\Phi}}}}\right); \quad i \in [T,1] \\ \boldsymbol{\Lambda}_{t,i} = \left[\overrightarrow{\overline{\overline{\boldsymbol{\Lambda}}}}_{t,i} \oplus {}_{t,i}\overleftarrow{\overline{\overline{\boldsymbol{\Lambda}}}}\right] \end{cases} \tag{2}$$

where $\overrightarrow{\overline{\overline{\boldsymbol{\Phi}}}}_{t,i-1}$ indicates the cell state of the *(i-1)*th input time step in the forward LStm layer at time *t*; ${}_{t,i+1}\overleftarrow{\overline{\overline{\boldsymbol{\Phi}}}}$ is the cell state of the *(i+1)*th input time step in the backward LStm layer at time *t*.

2.1.2. Quantile Regression

Quantile regression (Qr) can explore the relationship between the conditional quantiles of the independent and dependent variables. The linear Qr can be expressed as Eq. (3).

$$\boldsymbol{Q}_{Y_t}^{\boldsymbol{linear}}\left(\boldsymbol{\tau} \mid \boldsymbol{X}_t\right) \triangleq \boldsymbol{F}\left(X_t, \overline{\overline{\boldsymbol{\varepsilon}}}\left(\boldsymbol{\tau}\right)\right) = X_t \overline{\overline{\boldsymbol{\varepsilon}}}\left(\boldsymbol{\tau}\right), \ t = 1, 2, \cdots, n \tag{3}$$

where $\boldsymbol{Q}_{Y_t}^{\boldsymbol{linear}}\left(\tau \mid \boldsymbol{X}_t\right)$ is the τ th condition quantile of the dependent variable $\boldsymbol{Y}_t$ and $\boldsymbol{\tau} \in (0,1)$. Regression coefficients $\overline{\overline{\boldsymbol{\varepsilon}}}\left(\boldsymbol{\tau}\right) = \left\langle \boldsymbol{\varepsilon_0}\left(\boldsymbol{\tau}\right), \left(\boldsymbol{\varepsilon_1}\left(\boldsymbol{\tau}\right), \cdots, \boldsymbol{\varepsilon_m}\left(\boldsymbol{\tau}\right)\right)\right\rangle$.
The estimated value $\overline{\hat{\boldsymbol{\varepsilon}}}\left(\boldsymbol{\tau}\right)$ of $\overline{\overline{\boldsymbol{\varepsilon}}}\left(\boldsymbol{\tau}\right)$ can be obtained by minimizing Eq. (4).

$$\overline{\hat{\boldsymbol{\varepsilon}}}\left(\boldsymbol{\tau}\right) = \boldsymbol{argmin}\left(\sum_{t=1}^{n} \Phi_{\boldsymbol{\tau}}\left(\boldsymbol{Y}_t - \boldsymbol{X}_t \overline{\overline{\boldsymbol{\varepsilon}}}\left(\boldsymbol{\tau}\right)\right)\right) \tag{4}$$

where $\Phi_{\boldsymbol{\tau}}(\bullet)$ indicates an asymmetric function that can be written as

$$\Phi_{\boldsymbol{\tau}}\left(\boldsymbol{Y}_t - \boldsymbol{X}_t \overline{\overline{\boldsymbol{\varepsilon}}}\left(\boldsymbol{\tau}\right)\right) = \begin{cases} \boldsymbol{\tau}\left(\boldsymbol{Y}_t - \boldsymbol{X}_t \overline{\overline{\boldsymbol{\varepsilon}}}\left(\boldsymbol{\tau}\right)\right), & \boldsymbol{Y}_t - \boldsymbol{X}_t \overline{\overline{\boldsymbol{\varepsilon}}}\left(\boldsymbol{\tau}\right) \geq 0 \\ \left(1 - \boldsymbol{\tau}\right)\left(\boldsymbol{Y}_t - \boldsymbol{X}_t \overline{\overline{\boldsymbol{\varepsilon}}}\left(\boldsymbol{\tau}\right)\right), & \boldsymbol{Y}_t - \boldsymbol{X}_t \overline{\overline{\boldsymbol{\varepsilon}}}\left(\boldsymbol{\tau}\right) < 0 \end{cases} \tag{5}$$

From these equations, the τ th condition quantile of $\boldsymbol{Y}_t$ can be estimated as

$$\boldsymbol{Q}_{Y_t}^{linear}\left(\boldsymbol{\tau} \mid \boldsymbol{X}_t\right) \sim \boldsymbol{X}_t \overline{\hat{\boldsymbol{\varepsilon}}}\left(\boldsymbol{\tau}\right) \tag{6}$$

2.1.3. Quantile Regression BiLStm (QrBiLStm)

Based on the BiLStm and Qr, QrBiLStm was used for uncertainty modeling by modifying the cell structure and loss function of BiLStm. The loss function can be modified as $L_{\boldsymbol{QrBiLStm}}^{Pinball-loss} = \sum_{t=1}^{n} \Phi_{\boldsymbol{\tau}}\left(\boldsymbol{Y}_t - \boldsymbol{X}_t \overline{\overline{\boldsymbol{\varepsilon}}}\left(\boldsymbol{\tau}\right)\right)$. The condition quantile of $\boldsymbol{Y}_t$ obtained by QrBiLStm can be formulated as

$$\boldsymbol{Q}_{\boldsymbol{Y}_t}^{BiLStm}\left(\tau \mid \boldsymbol{X}_t\right) \triangleq f\left(\boldsymbol{X}_t, \varepsilon(\tau)\right) = \sigma\left(\boldsymbol{W}_{\boldsymbol{\Omega}}(\tau) \times \boldsymbol{\Lambda}_t(\tau)\right) \tag{7}$$

where $\boldsymbol{W}_{\boldsymbol{\Omega}}(\tau)$ indicates the weight matrix of τ, and $\boldsymbol{\Lambda}_t(\boldsymbol{\tau}) = \left[\overrightarrow{\overline{\overline{\boldsymbol{\Lambda}}}}_{t,i}(\boldsymbol{\tau}) \oplus {}_{t,i}\overleftarrow{\overline{\overline{\boldsymbol{\Lambda}}}}(\boldsymbol{\tau})\right]$.

The novel QrBiLStm network combines quantile regression with bi-directional data processing, and can effectively learn the hidden correlation between the pre- and post-time-step data in a time series, with better uncertainty modeling.

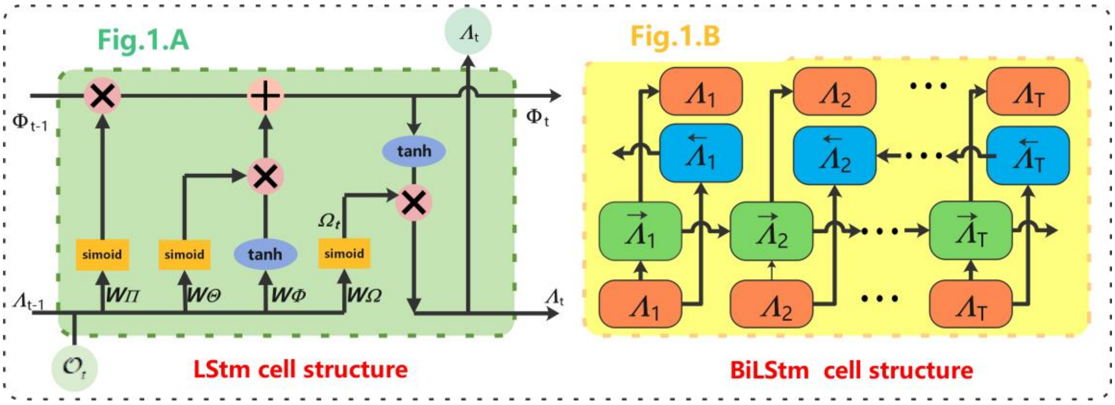


Fig.1. Cell structures of LStm and BiLStm

2.2. Original signal preprocessing

In this section, the principles of SSa and PSr are introduced in decomposing and reconstructing sequences.

2.2.1. Singular Spectral analysis (SSa)

The principal objective of SSa is to decompose the original series into a sum of series, identified as either a trend, a periodic or quasi-periodic component, or noise [47]. The flow of SSa can be summarized as follows.

(A). Embedding procedure

Based on the original sequence $\overline{\overline{\mho}}_{ini}^{T}$ and Karhunene–Loeve decomposition of the covariance matrix, the sequence $\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}^{i}=\left[\mho_{i-1},\cdots,\mho_{i+Y-2}\right]^{\boldsymbol{T}}$ of the L-dimensional vector is constructed.

$$\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}=\begin{bmatrix}\mho_0 & \cdots & \mho_M\\ \vdots & \ddots & \vdots\\ \mho_{Y-1} & \cdots & \mho_{T-1}\end{bmatrix} \tag{8}$$

where, $M=T-Y+1$ and $\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}$ is a Hankel matrix with equal elements on the diagonals.

(B) Singular value disintegration

The matrix $\overline{\overline{\boldsymbol{SS}}}^{T}$ is calculated to determine its eigenvalues using triples $\left(x_i,\boldsymbol{E}_i,\boldsymbol{F}_i\right)$ by SVD [51]**.** The eigenvalues of $\overline{\overline{\boldsymbol{SS}}}^{T}$ are defined as ζ_i, $i=1,2,\ldots,Y$ in descending order. $\boldsymbol{E}_i$ and $\boldsymbol{F}_i$ are the i^th^ left and right eigenvectors, respectively, of $\overline{\overline{\boldsymbol{SS}}}^{T}$. Assuming $r=rank\left(\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}^{i}\right)$, the trajectory matrix $\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}$ can be expressed as

$$\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}=\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}^{1}+\cdots+\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}^{r},\quad \overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}^{i}=x_i\boldsymbol{E}_i\boldsymbol{F}_i^{T} \tag{9}$$

where x_i is the singular value of $\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}$, and $\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}^{i}\left(i=1,2,\ldots,r\right)$ are matrices of $rank=1$.

(C). Reconstruction

Step 3.1 (Grouping): The indices $K = 1, 2, \ldots, r$ are grouped into V disjoint subsets $\{G_1^k, G_2^k, \cdots, G_V^k\}$ corresponding to splitting the elementary matrices $\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}^{i}\ (i = 1, 2, \ldots, r)$ into V groups. Each group contains a set of indices as $\overline{\overline{\boldsymbol{G}}}^{k} = \{d_1, \cdots, d_{\mathrm{p}}\}$. The resultant matrix is defined as $\boldsymbol{S}_{\overline{\overline{G}}^{k}} = \boldsymbol{S}_{d,1} + \boldsymbol{S}_{d,2} + \cdots + \boldsymbol{S}_{d,p}$. Thus, $\overline{\overline{\boldsymbol{S}_{\boldsymbol{H}}}} = \boldsymbol{S}_{\overline{\overline{G}}^{k},1} + \boldsymbol{S}_{\overline{\overline{G}}^{k},2} + \cdots \boldsymbol{S}_{\overline{\overline{G}}^{k},V}$, where $\overline{\overline{\boldsymbol{S}_{\boldsymbol{H}}}}$ is the sum of $\overline{\overline{\boldsymbol{G}}}^{k}$ resultant matrices.

Step 3.2 (Diagonal averaging): Each matrix $\boldsymbol{S}_{\overline{\overline{G}}^{k},j}$ $j = 1, 2, \ldots, V$ is transferred into a time series. Let $\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}$ be a $(Y \times M)$ matrix with elements $s_{i,j}$, with $Y^* = \boldsymbol{min}(Y, M)$ and $M^* = \boldsymbol{max}(Y, M)$. Define $\boldsymbol{S}_{i,j}{}^{*} = \begin{cases} \boldsymbol{S}_{i,j}, & Y < M \\ \boldsymbol{S}_{j,i}, & otherwise \end{cases}$.

Matrix $\overline{\overline{\boldsymbol{S}_{\boldsymbol{H}}}}$ is transformed into sequence $\Lambda_0, \Lambda_1, \cdots, \Lambda_{T-1}$ using Eq. (10).

$$\Lambda_i = \begin{cases} (b+1)^{-1} \cdot \sum_{c=1}^{b+1} \boldsymbol{S}^{*}{}_{c,b-c+2}, & \boldsymbol{b} \in [0, Y^* - 1) \\ Y^{*-1} \cdot \sum_{c=1}^{L^*} \boldsymbol{S}^{*}{}_{c,b-c+2}, & \boldsymbol{b} \in [Y^* - 1, M^*) \\ (T-b)^{-1} \cdot \sum_{c=b-b^*+2}^{N-b+1} \boldsymbol{S}^{*}{}_{c,b-c+2}, & \boldsymbol{b} \in [M^* - 1, T) \end{cases} \tag{10}$$

The averaging of the elements along the diagonal $i + j = b + 2$, applied to a resultant matrix $\boldsymbol{S}_{\overline{\overline{G}}^{k},j}$, produces a time series $\overline{\overline{\mho^{T}}}$ of length *T*. Thus, the original series $\overline{\overline{\mho_{\boldsymbol{ini}}^{T}}}$ is decomposed into the sum of *V* sequences. Defining the decomposed series as $\overline{\overline{\mho_{\boldsymbol{de}}^{T}}}$, it can be expressed as $\overline{\overline{\mho_{\boldsymbol{de}}^{T}}} = \overline{\overline{\mho_{1}^{T}}} + \cdots + \overline{\overline{\mho_{V}^{T}}}$

2.2.2. Phase Space reconstruction (PSr)

In the prediction of chaotic time series, the phase space reconstruction (PSr) method can be used to reconstruct a one-dimensional series into a dynamic chaotic space to obtain better forecasting results [48]. In this study, the C–C method was used to determine two important parameters of the PSr algorithm: delay time $\boldsymbol{\omega}$ and embedding dimension $\boldsymbol{\theta}$. The PSr process is expressed as

$$\overline{\overline{\boldsymbol{\Phi}}}_{P}^{ini} = \left[\overline{\overline{\boldsymbol{\Phi}}}_1, \overline{\overline{\boldsymbol{\Phi}}}_2, \cdots, \overline{\overline{\boldsymbol{\Phi}}}_P\right]^T = \begin{bmatrix} \xi_1 & \xi_{1+\tau} & \cdots & \xi_{1+(\theta-1)\omega} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_i & \xi_{i+\tau} & \cdots & \xi_{i+(\theta-1)\omega} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_P & \xi_{P+\tau} & \cdots & \xi_{P+(\theta-1)\omega} \end{bmatrix} \tag{11}$$

where $\{\xi_i \mid i = 1, 2, \cdots, Z\}$ signifies the samples of the sequence; Z indicates the length of the initial sequences, and $P = Z - (\theta - 1) \cdot \omega$. Accordingly, the target matrix $\overline{\overline{\boldsymbol{R}}}_P^{tar}$ corresponding to $\overline{\overline{\boldsymbol{\Phi}}}_P^{ini}$ can be expressed as Eq. (12).

$$\overline{\overline{\boldsymbol{R}}}_P^{tar} = \left[\overline{\overline{\boldsymbol{R}}}_1, \overline{\overline{\boldsymbol{R}}}_1, \cdots, \overline{\overline{\boldsymbol{R}}}_P \right]^{\boldsymbol{T}} = \left[\xi_{1+(\theta-1)\lambda}, \xi_{2+(\theta-1)\lambda}, \cdots, \xi_Z \right]^{\boldsymbol{T}} \tag{12}$$

2.3. Improved Tunicate Swarm Optimization Algorithm (IMOTa)

This section illustrates the mechanism of the original optimizer, the multi-objective optimization, and three improved optimization strategies.

A. Tunicate swarm algorithm (TSa)

The TSa was proposed by Kaur et al. [49], who regarded the optimal solution as the food source in the ocean, and the process of finding the optimal solution as the movement behavior combination of the capsule animals looking for food. The comprehensive mathematical principle of the TSa is presented as following.

Behavior 1. (avoidance) This behavior of tunicates aims to avoid collisions between individuals, and is defined by $\overline{\overline{\boldsymbol{A}}}vo = \overline{\overline{\boldsymbol{G}}}raf / \overline{\overline{\boldsymbol{S}}}ocf$. $\overline{\overline{\boldsymbol{A}}}vo$ is driven mainly by gravity and social forces. The gravity force is counteracted by the water flow $\overline{\overline{\boldsymbol{W}}}atf = 2 \cdot \vec{\mathrm{R}}(1)$, and is defined as $\overline{\overline{\boldsymbol{G}}}raf = \vec{\mathrm{R}}(1) + \vec{\mathrm{R}}(2) - \overline{\overline{\boldsymbol{W}}}atf$. The social force is driven mainly by initial speed $IniS$ and subordinate speed $SubS$. The social force is defined as $\overline{\overline{\boldsymbol{S}}}ocf = \boldsymbol{Inis} + \vec{\mathrm{R}}(1) \cdot (\boldsymbol{Subs} - \boldsymbol{Inis})$. In the definition, $IniS$ is preset as 1, $SubS$ is preset as 4, and $\vec{\boldsymbol{R}}$ is a matrix with elements that are all random values ranging from $[0,1]$.

Behavior 2. (movement) Tunicates move in the direction of their best neighbors. This behavior can be mathematically defined as $\overline{\overline{\boldsymbol{D}}}is = \left| \overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_{\boldsymbol{f}} - rand \cdot \overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_t(k) \right|$. $\overline{\overline{\boldsymbol{Dis}}}$ measures the absolute distance between the optimal solution and the agent. In this behavior, $\overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_{\boldsymbol{f}}$ indicates the position of the optimal solution, and $\overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_t(k)$ refers to the position of the k^{th} individual.

Behavior 3. (convergence) Tunicates begin to advance toward food sources by means of $\overline{\overline{\boldsymbol{A}}}\boldsymbol{vo}$ and $\overline{\overline{\boldsymbol{D}}}\boldsymbol{is}$. The tunicates update their positions according to Eq. (13).

$$\overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_u(k) = \begin{cases} \overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_{\boldsymbol{f}} + \overline{\overline{\boldsymbol{A}}}\boldsymbol{vo} \cdot \overline{\overline{\boldsymbol{D}}}\boldsymbol{is}, & rand \geq 0.5 \\ \overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_{\boldsymbol{f}} - \overline{\overline{\boldsymbol{A}}}\boldsymbol{vo} \cdot \overline{\overline{\boldsymbol{D}}}\boldsymbol{is}, & rand < 0.5 \end{cases} \tag{13}$$

Behavior 4. (swarm behavior) The best two solutions are retained, and the positions of other individuals relative to the food source are updated. The swarm behavior can be mathematically expressed as Eq. (14).

$$\overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_t(k+1) = \left[\overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_t(k) + \overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_u(k) \right] \Big/ \left(2 + \vec{\boldsymbol{R}}(1) \right) \tag{14}$$

B. Improved multi-objective tunicate swarm algorithm (IMOTa)

This study developed three improvement strategies: the multi-objective approach

(MOJ), the elite opposition learning approach (EOLA), and the exponential function step approach (EFSA). The MOJ produces multiple objective functions in the optimization algorithm to achieve better optimization results. The EOLA can improve the convergence speed of the algorithm. The EFSA can improve the global optimization and robustness.

a. Multi-objective tunicate swarm algorithm (MOTa)

To achieve multi-objective optimization, this section introduces the dominant strategy, the Pareto optimal solution and archiving with a roulette wheel. The ability of the MOTa system to find the Pareto optimal solution is demonstrated using the definitions.

Definition 1. Let $\bar{\bar{\boldsymbol{J}}} = \left(J_1, J_2, \cdots, J_i\right)$ and $\bar{\bar{\boldsymbol{K}}} = \left(K_1, K_2, \cdots, K_i\right)$ be two vectors; $\bar{\bar{\boldsymbol{J}}}$ strictly dominates $\bar{\bar{\boldsymbol{K}}}$, if $\forall \boldsymbol{n} \in \{1,2,...,N\}$, $\boldsymbol{f}_n\left(\bar{\bar{\boldsymbol{J}}}\right) \geq \boldsymbol{f}_n\left(\bar{\bar{\boldsymbol{K}}}\right)$; $\bar{\bar{\boldsymbol{J}}}$ partially dominates $\bar{\bar{\boldsymbol{K}}}$, if $\exists \boldsymbol{n} \in \{1,2,...,N\}$, $\boldsymbol{f}_n\left(\bar{\bar{\boldsymbol{J}}}\right) > \boldsymbol{f}_n\left(\bar{\bar{\boldsymbol{K}}}\right)$. $\bar{\bar{\boldsymbol{J}}}$ dominates $\bar{\bar{\boldsymbol{K}}}$, if

$$\left[\forall \boldsymbol{n} \in \{1,2,...,N\}, \boldsymbol{f}_n\left(\bar{\bar{\boldsymbol{J}}}\right) \geq \boldsymbol{f}_n\left(\bar{\bar{\boldsymbol{K}}}\right)\right] \wedge \left[\exists \boldsymbol{n} \in \{1,2,...,N\}, \boldsymbol{f}_n\left(\bar{\bar{\boldsymbol{J}}}\right) > \boldsymbol{f}_n\left(\bar{\bar{\boldsymbol{K}}}\right)\right] \quad (15)$$

where $f_{\boldsymbol{n}}(\bullet)$ indicates the *n*-th objective function and *N* is the number of functions.

Definition 2. If $\forall n \in \{1,2,...,N\} : \nexists K \in \bar{\bar{\boldsymbol{K}}} \mid \boldsymbol{f}_n(\boldsymbol{K}) \succsim \boldsymbol{f}_n(\boldsymbol{J})$, that is, none of the obtained solutions dominates $\bar{\bar{\boldsymbol{J}}}$, then $\bar{\bar{\boldsymbol{J}}}$ is the Pareto optimal solution.

Definition 3. Archiving with a roulette wheel is a matrix used to store the optimal solutions. When the archive is full, the individuals with the most adjacent solutions are eliminated by the roulette wheel. The probability that an individual is eliminated is $Pe_i = Ns_i / cq, cq > 1$, where Ns_i indicates the number of adjacent solutions, and cq is a constant.

Suppose that the fitness function corresponding to the objective function is $\boldsymbol{fit}(\cdot)$, and the optimal position $\boldsymbol{P}^*$ of the individual in MOTA is the weight of two QrbiLStm units, $\boldsymbol{We}\left(\boldsymbol{P}^*\right)$. It is proved that $\boldsymbol{P}^*$ is the optimal weight of two QrbiLStm units through reduction to absurdity.

Proof

If there exists at least one adjacent position $\boldsymbol{Q}^* = \boldsymbol{P}^* + \theta$, the weights satisfy $\left[\forall n \in \{1,2,...,N\},\ fit_{\boldsymbol{n}}\left(\boldsymbol{We}\left(\boldsymbol{Q}^*\right)\right) \geq fit_{\boldsymbol{n}}\left(\boldsymbol{We}\left(\boldsymbol{P}^*\right)\right)\right] \wedge \left[\exists n \in \{1,2,...,N\},\ fit_{\boldsymbol{n}}\left(\boldsymbol{We}\left(\boldsymbol{Q}^*\right)\right) > fit_{\boldsymbol{n}}\left(\boldsymbol{We}\left(\boldsymbol{P}^*\right)\right)\right]$, and $\boldsymbol{Q}^*$ is stored in the Archive. As $\boldsymbol{We}\left(\boldsymbol{Q}^*\right)$ dominates $\boldsymbol{We}\left(\boldsymbol{P}^*\right)$, and the capacity of Archive with Roulette-Wheel is limited, $\boldsymbol{P}^*$ is deleted from Archive with the $prob = Ns_i / cq$ or ranked behind $\boldsymbol{Q}^*$. The position with the highest fitness value in Archive is selected as the optimal position. Corresponding, the optimal weights of QrBiLStm is $\boldsymbol{We}\left(\boldsymbol{Q}^*\right)$ instead of $\boldsymbol{We}\left(\boldsymbol{P}^*\right)$. ■

b. Elite opposition learning (EOLA)-MOTa

The IMOTa based on EOLA was proposed to improve the convergence performance of the optimizer. The principle of EOLA is to calculate and evaluate the opposition solution of a feasible solution, and select the better solution as the next generation. In this study, the elite tunicate is defined as the individual that obtains the highest fitness value.

Definition 4. (opposition point) Let $\vec{\bar{X}_j} = \left(x_{j,1}, x_{j,2}, \cdots, x_{j,d}\right)$ be a point in *d*-dimensional space (regarded as a feasible solution), $x_i \in \left[lb_i, ub_i\right]$, and its corresponding opposition point $\vec{X_j} = \left(\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_d\right)$ are defined in Eq. (16)

$$\tilde{x}_i = lb_i + ub_i - x_i \tag{16}$$

Definition 5. (elite opposition solution) Suppose that $\overline{\overline{X_j}} = \left(x_{j,1}, x_{j,2}, \cdots, x_{j,d}\right)$ is a common tunicate, and the corresponding extreme value of itself is the elite tunicate $\overline{\overline{X_j^{elite}}} = \left(x_{j,1}{}^e, x_{j,2}{}^e, \cdots, x_{j,d}{}^e\right)$. The elite opposition solution $X_j^{elite} = \left(\tilde{x}_{j,1}{}^e, \tilde{x}_{j,2}{}^e, \cdots, \tilde{x}_{j,d}{}^e\right)$ can be defined as formula (17).

$$\tilde{x}_{j,i}{}^e = \varpi \cdot \left(dlb_j + dub_j\right) - x_{j,i}{}^e \tag{17}$$

where $\tilde{x}_{j,i}{}^e \in \left[dlb_j, dub_j\right]$; $c \in U(0,1)$; $\left[dlb_j, dub_j\right]$ is the dynamic boundary of the ith dimension search space, which can be calculated according to Eq. (18).

$$dlb_j = min\left(x_{j,i}\right), dub_j = max\left(x_{j,i}\right) \tag{18}$$

Replacing the fixed boundary with the dynamic boundary of the search space is conducive to preserving the search experience, such that the generated opposition solution can be located in the gradually reduced search space. However, it has the possibility of causing $x_{j,i}{}^e$ to exit $\left[dlb_j, dub_j\right]$. If $x_{j,i}{}^e < dlb_j$ or $x_{j,i}{}^e > dub_j$, then $\tilde{x}_{j,i}{}^e = \varpi \cdot \left(dub_j - dlb_j\right) + dub_j$, where ϖ is a random value between 0 and 1.

c. Exponential function steps (EFSA)-MOTa

In **Behaviors** 3 and 4 of the original TSA, the approach to promote the location update is random linear. This updating approach cannot guarantee individuals to find the optimal solution, which ultimately leads to poor optimization and robustness. Thus, improving the piecewise linear random step using EFSA is proposed. The new location update strategy can be mathematically expressed as Eq. (19).

$$\overline{\overline{Pos_u}}\left(k\right) = \overline{\overline{Pos_f}} + \left(rand - 0.5\right) \cdot 2^{rand} \cdot \overline{\overline{Avo}} \cdot \overline{\overline{Dis}} \tag{19}$$

C. Design of the multi-objective optimization function of EPFS

To simultaneously optimize the reliability and interval width of the forecasting system, two pseudo-interval indicators were designed. The purpose of constructing pseudo-intervals is to optimize the upper and lower bounds of the intervals, respectively, to achieve better optimization results. Based on the two pseudo-interval indicators that measure reliability and resolution, the objective functions for multi-objective optimization are developed.

a. Pseudo-interval indicators

The pseudo-interval is a half-interval composed of the observed values and the upper or lower bound of the interval. Thus, the indicators for evaluating the reliability and resolution of the pseudo-interval can be designed as $\boldsymbol{PICP}^{half}(\alpha)$, and $\boldsymbol{PINAW}^{half}(\alpha)$.

Two indicators for evaluating the upper pseudo-interval can be defined as

$$\boldsymbol{PICP}_{upper}^{half}(\alpha)=\frac{1}{M}\sum_{i=1}^{M}\boldsymbol{\Gamma}_i^{half},\quad \boldsymbol{\Gamma}_i^{half}=\begin{cases}1, & \boldsymbol{UB}_i^{half}(\alpha)\geq Obse\boldsymbol{V}_i\\ 0, & \boldsymbol{UB}_i^{half}(\alpha)< Obse\boldsymbol{V}_i\end{cases} \tag{20}$$

$$\boldsymbol{PINAW}_{upper}^{half}(\alpha)=\frac{1}{MR}\sum_{i=1}^{M}\left[\boldsymbol{UB}_i^{half}(\alpha)-Obse\boldsymbol{V}_i\right] \tag{21}$$

where $\boldsymbol{UB}_i^{half}(\alpha)$ is the upper bound of the forecasting interval corresponding to α. $ObseV_i$ is the observation value, M is the number of observation values, and R is the range of observation values.

Accordingly, the two indicators for evaluating the lower pseudo interval can be defined as:

$$\underrightarrow{\boldsymbol{PICP}_{lower}^{half}(\alpha)}=\frac{1}{M}\sum_{i=1}^{M}\boldsymbol{\Gamma}_i'^{half},\quad \boldsymbol{\Gamma}_i'^{half}=\begin{cases}1, & \boldsymbol{LB}_i^{half}(\alpha)\leq Obse\boldsymbol{V}_i\\ 0, & \boldsymbol{LB}_i^{half}(\alpha)> Obse\boldsymbol{V}_i\end{cases} \tag{22}$$

$$\underrightarrow{\boldsymbol{PINAW}_{lower}^{half}(\alpha)}=\frac{1}{MR}\sum_{i=1}^{M}\left[Obse\boldsymbol{V}_i-\boldsymbol{LB}_i^{half}(\alpha)\right] \tag{23}$$

where $\boldsymbol{LB}_i^{half}(\alpha)$ is the lower bound of the forecasting interval corresponding to α.

b. Multi-objective optimization function

The objective functions for multi-objective optimization can be determined as:

$$\boldsymbol{min}\begin{cases}\boldsymbol{Of}_1=(1-\alpha/2)-\boldsymbol{PICP}^{half}(\alpha)\\ \boldsymbol{Of}_2=\boldsymbol{PINAW}^{half}(\alpha)\cdot\left[1+\exp\left(-\Phi\cdot\left(\boldsymbol{PICP}^{half}(\alpha)-1+\alpha\right)\right)\right]\end{cases} \tag{24}$$

where $\Phi>0$ is the penalty coefficient. A larger Φ indicates a greater degree of punishment.

3. Framework of the proposed Ensemble probabilistic forecasting system

This section presents the description of the material analyzed (section 3.1) and the ensemble probabilistic forecasting system applied in this study (section 3.2).

3.2. Dataset description

The three experimental datasets were collected from Shandong Peninsula with an interval of 10 minutes. In each dataset, extract 2880 points as the experimental sequence and select 75% of the total length as training set 1, with a length of 2160. The remaining 720 points are divided into training set 2 and testing set. Training set 2 accounts for 75% of the remaining length, with a length of 540 and a testing set of 180.

3.3. Flow of the proposed Ensemble probabilistic forecasting system

In accordance with the aforementioned data processing approaches and forecasting models, the proposed ensemble probabilistic forecasting system includes SSa decomposition and reconstruction, phase space reconstruction of the C–C method, principal forecasts based on two QrBiLStm units, and construction of pseudo-intervals to optimize the upper and lower bounds. The process is presented; the complete system structure and procedure are shown in **Fig. 2**.

Step 1: Segment the original wind speed sequence into two training sets and a testing set. A total of 2,880 data were collected. Training set 1 included 2,160 points, Training set 2 included 540 points, and Test set included 180 points.

Step 2: Use SSa to decompose and reconstruct the wind speed sequence $\left\{ \boldsymbol{X}'_j \mid j=1,2,\cdots n \right\}$ to obtain the denoised wind speed sequence $\left\{ \boldsymbol{X}''_j \mid j=1,2,\cdots n \right\}$.

Step 3: Use the C–C method to find the optimal parameter values of PSr and reconstruct the sequence to adapt to the chaotic system.

Step 4: Implement uncertainty forecasting quantization on Training Set 1, based on two QrBiLStm network units.

Step5: Construct the pseudo-intervals based on the forecasting values and observation values of Training Set 2 obtained in Step 4.

Step 6: Use the IMOTa and the designed interval optimization objective function; the pseudo-interval is input for optimization, and the final probabilistic forecasting results on the testing set are obtained.

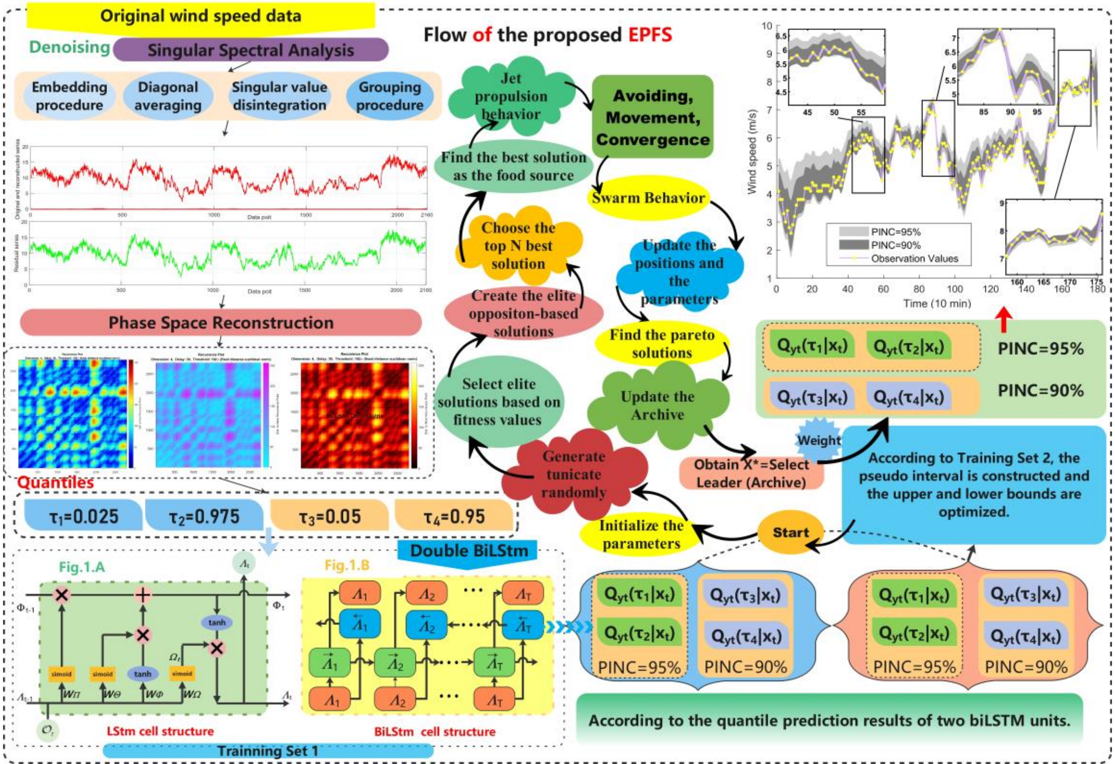


Fig.2. Flow of the proposed EPFS.

4. Experimental results and discussion

This section describes the evaluation metrics of uncertainty modeling. Three comparative experiments and their corresponding analyses are described to verify the forecasting effectiveness of the proposed EPFS. The whole experiment is implemented on the personal computer with AMD Ryzen 5 5600H six-core processor with Radeon Graphics 3.30 GHz, 16 GB of RAM and a single NVIDIA GeForce GTX 1650 of GPU. The proposed EPFS and comparative models were implemented on Matlab2020a.

The parameter settings for all models designed for the experiments are presented in Table A1 in Appendix A.

4.1. Evaluation metrics

The wind speed prediction interval was evaluated based on reliability, resolution, and sharpness.

Reliability

Reliability is based on the significance level α to evaluate the coverage probability of the forecasting interval. In this study, the reliability was characterized by the PICP metric which is expressed as Eq. (25).

$$\boldsymbol{PICP}(\alpha)=\frac{1}{N}\sum_{i=1}^{N}\Phi_i\ ,\ \ \Phi_i=\begin{cases}1, & ObseV_i\in\left[\boldsymbol{LB}_i(\alpha),\boldsymbol{UB}_i(\alpha)\right]\\ 0, & ObseV_i\notin\left[\boldsymbol{LB}_i(\alpha),\boldsymbol{UB}_i(\alpha)\right]\end{cases}\tag{25}$$

where N is the length of the testing set, α is the confidence level, and $Obse\boldsymbol{V}_i$ indicates the observation value. $\boldsymbol{LB}_i(\alpha)$ and $\boldsymbol{UB}_i(\alpha)$ are the lower and upper bounds of the prediction interval, respectively, corresponding to α. In this paper, interval prediction is implemented based on $\alpha=0.05$ and $\alpha=0.1$ confidence levels. And $PINC=(1-\alpha)\times 100\%$.

Resolution

Measurement of interval resolution (interval width) [50] is important for effective interval prediction. A prediction interval that is too broad contains a small amount of valuable information, which is less practical [51]. Thus, PINAW largely reflects the information contained in the forecast interval and can be represented by Eq. (26). R is the range of observation values, it is determined by the maximum of the observation values on the testing set minus its minimum. And N is the length of the testing set.

$$\boldsymbol{PINAW}(\alpha)=\frac{1}{NR}\sum_{i=1}^{N}\left[\boldsymbol{UB}_i(\alpha)-\boldsymbol{LB}_i(\alpha)\right]\tag{26}$$

Sharpness

Both PINAW and PICP are one-sided in evaluating the quality of forecast intervals. The sharpness combines the two metrics for assessment of prediction intervals [54]; the

AIS metric [19] meets the requirements, and is expressed as Eq. (27):

$$\boldsymbol{AIS}(\alpha) = \sum_{i=1}^{N} \boldsymbol{S}_{\rightarrow}^{(\alpha)} \tag{27}$$

$$\boldsymbol{S}_{\rightarrow}^{(\alpha)} = \begin{cases} -2\alpha\xi_i(\alpha) - 4(\boldsymbol{LB}_i(\alpha) - Obse\boldsymbol{V}), & if\ Obse\boldsymbol{V}_i < \boldsymbol{LB}_i(\alpha) \\ -2\alpha\xi_i(\alpha), & if\ Obse\boldsymbol{V}_i \in [\boldsymbol{LB}_i(\alpha), \boldsymbol{UB}_i(\alpha)] \\ -2\alpha\xi_i(\alpha) - 4(Obse\boldsymbol{V}_i - \boldsymbol{UB}_i(\alpha)), & if\ Obse\boldsymbol{V}_i > \boldsymbol{UB}_i(\alpha) \end{cases} \tag{28}$$

where $\xi_i(\alpha) = \boldsymbol{UB}_i(\alpha) - \boldsymbol{LB}_i(\alpha)$.

4.2. Comparative experiments

This section presents a comparison of the proposed EPFS with commonly used single and ensemble probabilistic forecasting models (EPFMs). Probabilistic forecasting single models include the QrLASso, QrLStm, Qr convolution neural network (QrCNn), QrGRu, Gaussian process regression (GPr), the Bayesian regression model (BLgm), and the proposed QrBiLStm. The EPFMs include models with different denoising methods and optimizers. The interval forecasting chart of the proposed EPFS and the metric values of the other single models is shown in **Fig. 3**.

The compared denoising methods include empirical mode decomposition (Emd) [56], ensemble Emd (Eemd), complete Eemd with adaptive noise (CeemdAN) [57], and wavelet transform (Wt). The compared optimizers include MOTa, MO dragonfly algorithm (MODa), MO grasshopper algorithm (MOGa), and MO antlion algorithm (MOAa).

4.2.1. Comparative Experimental Analysis with Single Models

Reliability

PICP shows the reliability of intervals; when the PICP is higher than the prediction interval nominal confidence (PINC), the forecasting interval is considered to be reliable. As shown in **Table 2**, all single models and EPFS except QrCNN are valid for all Sites.

The proposed QrBiLStm obtains $PICP = 1$ for both $PINC = 90\%$ and $PINC = 95\%$ for all three sites, indicating that it has better reliability for interval prediction. The EPFS optimized based on the two QrBiLStm benchmark models also has high PICP values, and is also reliable. For Site 1, when $PINC = 90\%$, EPFS obtains $PICP = 0.9833$; when $PINC = 95\%$, EPFS obtains $PICP = 0.9944$.

Resolution

The PINAW metric indicates the interval width which determines the practicality and informative of interval [55]. A smaller PINAW indicates a narrower interval width with more uncertainty information. Not considering the invalid model ($\boldsymbol{PICP}(\alpha) < \boldsymbol{PINC}(\alpha)$), the interval width of QrBiLStm is the narrowest of all single models. This is reflected in the PINAW values. For $PINC = 95\%$, QrBiLStm is obtained at three sites: $PINAW_{\boldsymbol{PINC=95\%}}^{Site1} = 0.2674$, $PINAW_{\boldsymbol{PINC=95\%}}^{Site2} = 0.2471$, and $PINAW_{\boldsymbol{PINC=95\%}}^{Site3} = 0.3306$. Thus, compared with other single models, the proposed QrBiLStm has the narrowest interval width and is more effective in interval forecasting than the model based on distribution hypothesis and other single models not based on distribution.

Compared with the single models, EPFS is greatly optimized in terms of the interval width. The experimental results show that QrBiLStm has the narrowest interval width. The interval width forecasted by EPFS was $27.1387\% \sim 56.6055\%$ smaller than that obtained by QrBiLStm, indicating that the forecasting of EPFS is greatly improved compared with that of a single model.

Sharpness

The AIS metric simultaneously considers coverage and interval width, punishing an interval that does not contain observation values and interval width. The AIS value is generally less than 0; a higher AIS value indicates a more effective forecasting interval. The AIS value of the proposed EPFS is the lowest at all sites and all PINCs, indicating that the proposed EPFS can provide more uncertainty information.

Compared with single models, EPFS is greatly optimized in interval width; its coverage rate is also high, resulting in the best interval prediction. For $PINC = 95\%$, the AIS values obtained by EPFS at the three sites are $AIS_{PINC=95\%}^{Site1} = -0.1273$, $AIS_{PINC=95\%}^{Site2} = -0.1216$, and $AIS_{PINC=95\%}^{Site3} = -0.1324$.

Remark: Compared with other single models, the proposed QrBiLStm can obtain the highest interval coverage and the narrowest interval width. Thus, the proposed QrBiLStm is more reliable and effective than the model based on the distribution hypothesis and other Qr–deep learning models not based on distribution. The proposed EPFS optimized using two QrBiLStm units can further reduce the interval width while ensuring high interval coverage. The proposed EPFS can provide more uncertainty information.

Table 2

Interval prediction metric values of single models with PINC=90% and PINC=95%.

Dataset	Models	PICP		PINAW		AIS	
		PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%
Site 1	**SSa-PSr-QrLASso**	0.9556	0.9778	0.4554	0.5341	-0.5439	-0.3055
	SSa-PSr-QrLStm	**1.0000**	**1.0000**	0.5160	0.5835	-0.5152	-0.2928
	SSa-PSr-QrGRu	**1.0000**	**1.0000**	0.5337	0.5986	-0.5334	-0.2964
	SSa-PSr-QrCNN	0.9611	0.8944	0.3147	0.7572	-0.3726	-0.4743
	SSa-PSr-QrBiLStm	1.0000	**1.0000**	0.3489	0.4249	-0.3500	-0.2086
	SSa-PSr-GPr	0.9944	**1.0000**	0.4457	0.5311	-0.4557	-0.2712
	SSa-PSr-BLgm	**1.0000**	**1.0000**	0.4460	0.5317	-0.4556	-0.2716
	Proposed EPFS	0.9833	0.9944	**0.1893**	**0.2674**	**-0.1805**	**-0.1273**
Site 2	**SSa-PSr-QrLASso**	0.8611	0.9111	0.3028	0.4918	-0.4409	-0.3389
	SSa-PSr-QrLStm	0.9778	0.9833	0.4770	0.5417	-0.5063	-0.2884
	SSa-PSr-QrGRu	0.9778	0.9889	0.4710	0.5729	-0.4905	-0.2930
	SSa-PSr-QrCNN	0.9833	**1.0000**	0.3734	1.1707	-0.3954	-0.5868
	SSa-PSr-QrBiLStm	**1.0000**	**1.0000**	0.3734	0.3494	-0.3802	-0.1748
	SSa-PSr-GPr	**1.0000**	**1.0000**	0.4547	0.5419	-0.4685	-0.2791
	SSa-PSr-BLgm	**1.0000**	**1.0000**	0.4555	0.5428	-0.4695	-0.2797
	Proposed EPFS	0.9778	0.9889	**0.1620**	**0.2471**	**-0.1590**	**-0.1216**
Site 3	**SSa-PSr-QrLASso**	0.8944	0.9889	0.4122	0.6835	-0.4157	-0.3049
	SSa-PSr-QrLStm	0.9889	**1.0000**	0.5568	0.7498	-0.4633	-0.3110
	SSa-PSr-QrGRu	0.9833	**1.0000**	0.5502	0.7138	-0.4539	-0.2933
	SSa-PSr-QrCNN	0.9111	0.9222	0.4586	0.4362	-0.4508	-0.2351
	SSa-PSr-QrBiLStm	**1.0000**	**1.0000**	0.3384	0.4537	-0.2771	-0.1835
	SSa-PSr-GPr	0.9889	**1.0000**	0.5029	0.5992	-0.4248	-0.2519
	SSa-PSr-BLgm	0.9944	**1.0000**	0.5080	0.6047	-0.4286	-0.2544
	Proposed EPFS	0.9833	0.9889	**0.2180**	**0.3306**	**-0.1784**	**-0.1324**

Note: The table above presents the reliability, resolution and comprehensive information of intervals obtained by different models. When PICP>PINC, the forecasting intervals are reliable. When the intervals are reliable, the narrower the interval width, the better the interval forecasting effect, which can be measured by $PINAW(\alpha)=\frac{1}{NR}\sum_{i=1}^{N}\left[UB_i(\alpha)-LB_i(\alpha)\right]$. Metric $AIS(\alpha)=\sum_{i=1}^{N}S^{(\alpha)}$ can comprehensively evaluate the interval forecasting effect.

4.2.2. Comparative Experimental Analysis with Ensemble Models

Reliability

The EPFM based on other denoising methods has high interval coverage, whether at $PINC = 90\%$ or at $PINC = 95\%$. However, the reliability of the prediction interval is not always guaranteed when other algorithms are used to perform the optimization. Of the four other optimization algorithms, only MOTa–EPFM and MODa–EPFM can obtain reliable interval prediction results in all cases (three sites and two confidence levels), similar to the proposed EPFS. For Site 1, the interval coverage rates of MOTa and MODa are $PICP^{MOTA}_{PINC=90\%} = 0.9833$, $PICP^{MOTa}_{\mathbf{PINC=95\%}} = 1$, $PICP^{MODa}_{\mathbf{PINC=90\%}} = 0.9778$, and $PICP^{MODa}_{\mathbf{PINC=95\%}} = 1$, respectively. The Emd, Eemd, CeemdAN, Wt-based PSr–EPFM, MOTa–EPFM, MODa–EPFM, and the proposed EPFS are better in terms of reliability.

Resolution

The interval width obtained by EPFMs using other denoising methods is significantly greater than that of the proposed EPFS, which is evident from the PINAW values. For Site 2, the PINAW values of Emd–EPFM, Eemd–EPFM, CeemdAN–EPFM, Wt–EPFM, and the proposed EPFS are $\overline{\mathbf{PINAW}}^{Site1}_{PINC=90\%} = \left[0.2921, 0.3027, 0.3577, 0.2517, 0.1620\right]$ and $\overline{\mathbf{PINAW}}^{Site1}_{PINC=95\%} = \left[0.3371, 0.3335, 0.3980, 0.2844, 0.2471\right]$ when $PINC = 90\%$ and $PINC = 95\%$, respectively.

The predicted interval widths of EPFMs using other optimization algorithms are similar to that of the proposed EPFS. However, EPFMs based on other algorithms have less reliability in their prediction results (manifested as low PICP values), with narrow interval widths. For Site 1, the PINAW values for MOGa–EPFM are 0.1755 at $PINC = 90\%$ and 0.2414 at $PINC = 95\%$; the corresponding PICP values are 0.9167 and 0.9611, respectively, and lower than those of all other models.

Sharpness

The interval prediction results after optimization are better than those of the single models, and other ensemble models cannot obtain interval reliability and interval width simultaneously. To better evaluate the effect of interval prediction, the AIS value was used to evaluate the uncertainty information contained in the prediction interval. A larger AIS value (AIS is generally a negative number), produces a better interval prediction. Considering all cases (three sites and two confidence levels), the proposed EPFS obtained the lowest AIS values, indicating that the proposed EPFS has better global optimization ability and better ability to optimize multiple objectives. The metric values of all three sites are presented in **Table 3** and **Fig. 4**.

Remark: Although EPFMs denoised by other methods have strong reliability, their resolution is not high; thus, the interval forecasting effect is not satisfying. EPFMs using other optimization algorithms cannot simultaneously optimize the interval reliability and interval resolution. The proposed EPFS has the best interval prediction results.

Table 3
Interval prediction metric values of ensemble models with PINC=90% and PINC=95%.

Dataset	Models	PICP		PINAW		AIS	
		PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%
Site 1	SSa-PSr-MOTa-EPFM	0.9833	1.0000	0.2048	0.2828	-0.1997	-0.1353
	SSa-PSr-MODa-EPFM	0.9778	1.0000	0.1979	0.3413	-0.1899	-0.1666
	SSa-PSr-MOGa-EPFM	0.9167	0.9611	**0.1755**	0.2414	-0.1887	-0.1223
	SSa-PSr-MOAa-EPFM	0.9333	1.0000	0.1769	0.2805	-0.1843	-0.1336
	Emd-PSr-EPFM	0.9778	1.0000	0.2538	0.3903	-0.2527	-0.1892
	Eemd-PSr-EPFM	0.9889	1.0000	0.2276	0.3824	-0.2226	-0.1845
	Wt-PSr-EPFM	0.9944	0.9944	0.3513	0.4031	-0.3453	-0.1941
	CeemdAN-PSr-EPFM	**1.0000**	**1.0000**	0.2868	0.3848	-0.2795	-0.1832
	Proposed EPFS	0.9833	0.9944	0.1893	**0.2674**	**-0.1805**	**-0.1273**
Site 2	SSa-PSr-MOTa-EPFM	0.9778	0.9556	0.1692	0.2476	-0.1627	-0.1291
	SSa-PSr-MODa-EPFM	0.8833	0.9500	**0.1454**	0.2512	-0.1569	-0.1364
	SSa-PSr-MOGa-EPFM	0.8833	0.9500	0.1570	**0.2342**	-0.1707	-0.1248
	SSa-PSr-MOAa-EPFM	0.8500	0.8778	0.1475	0.2385	-0.1737	-0.1518
	Emd-PSr-EPFM	**1.0000**	**1.0000**	0.2921	0.3371	-0.2887	-0.1633
	Eemd-PSr-EPFM	0.9944	0.9944	0.3027	0.3335	-0.2991	-0.1612
	Wt-PSr-EPFM	**1.0000**	**1.0000**	0.3577	0.3980	-0.3523	-0.1945
	CeemdAN-PSr-EPFM	0.9944	0.9778	0.2517	0.2844	-0.2505	-0.1445
	Proposed EPFS	0.9778	0.9889	0.1620	0.2471	**-0.1590**	**-0.1216**
Site 3	SSa-PSr-MOTa-EPFM	0.9667	0.9833	0.2489	0.3605	-0.2035	-0.1465
	SSa-PSr-MODa-EPFM	0.8833	0.9722	**0.2170**	0.3337	-0.1810	-0.1378
	SSa-PSr-MOGa-EPFM	0.9500	0.9611	0.2003	0.3431	-0.1846	-0.1448
	SSa-PSr-MOAa-EPFM	0.9833	0.9889	0.2172	**0.3179**	-0.1876	-0.1345
	Emd-PSr-EPFM	0.9611	0.9778	0.3240	0.3798	-0.2707	-0.1572
	Eemd-PSr-EPFM	0.9778	0.9833	0.3300	0.3843	-0.2692	-0.1581
	Wt-PSr-EPFM	0.9667	0.9833	0.3726	0.4324	-0.3200	-0.1873
	CeemdAN-PSr-EPFM	**0.9889**	**0.9944**	0.3253	0.3935	-0.2650	-0.1583
	Proposed EPFS	0.9833	0.9889	0.2180	0.3306	**-0.1784**	**-0.1324**

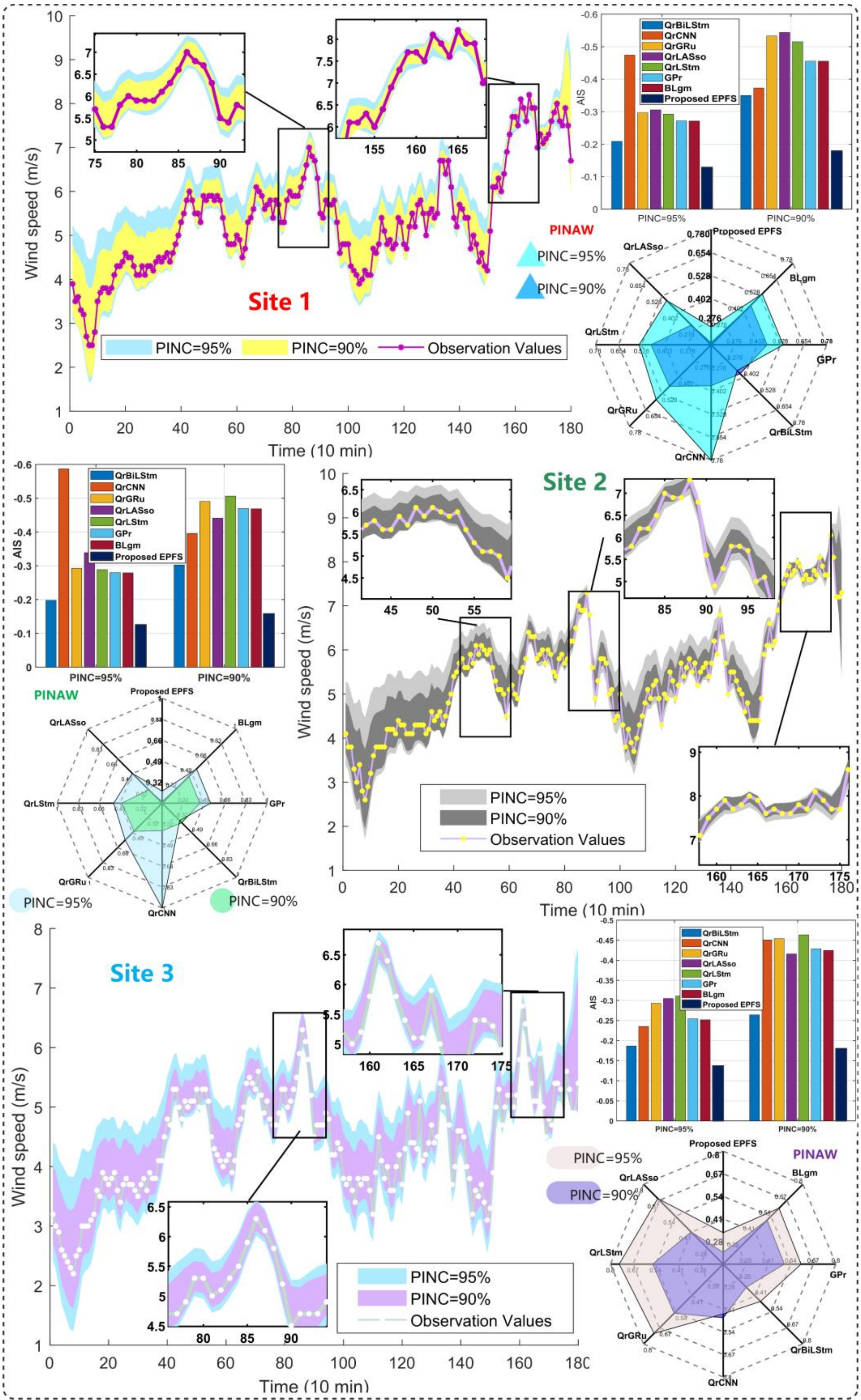


Fig.3. Forecasting results of proposed EPFS and other single models.

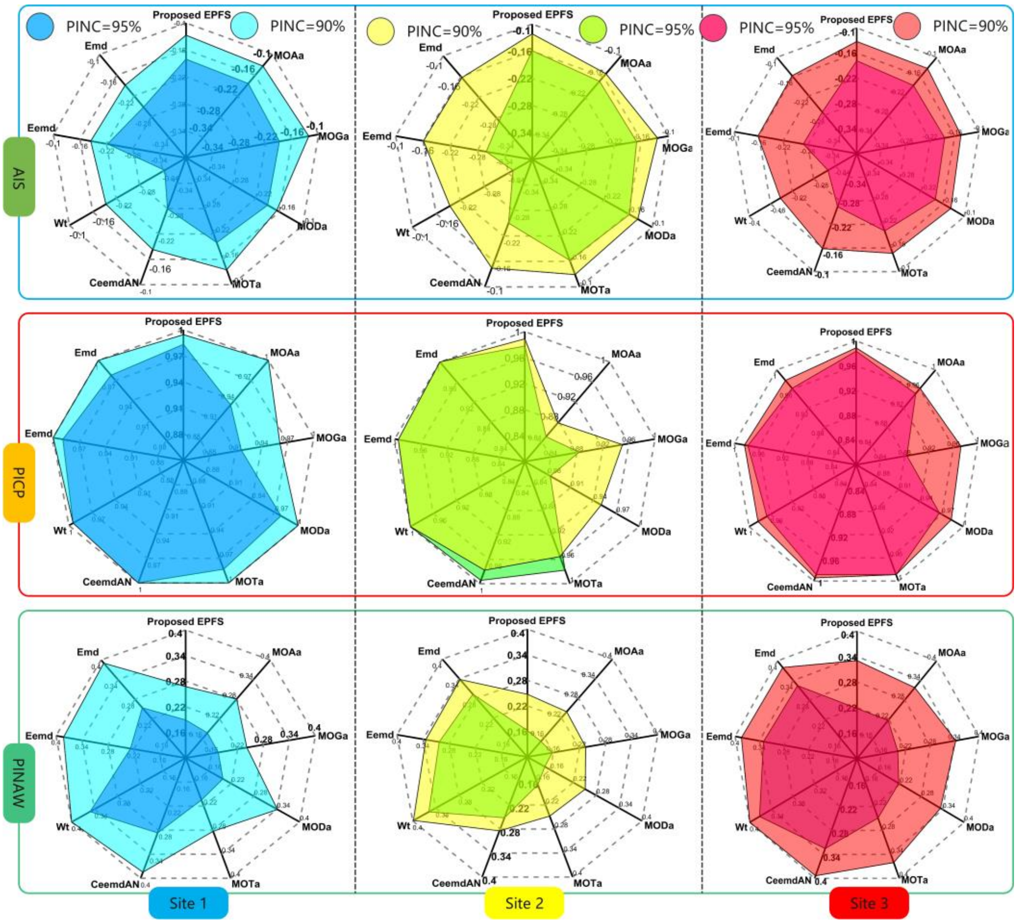


Fig.4. Forecasting results of proposed EPFS and other ensemble models.

4.3. Statistical tests

A Diebold-Mariano (DM) test [58] was implemented to validate whether the time-series forecasting results of two different models exist for the same or significantly different accuracy. A hypothesis test is conducted, where the null hypothesis says that there is no difference between the forecasting accuracy of the models compared, and the alternative hypothesis says that the forecasting error of the model proposed is different with the compared one. In this study, the hypothetical form can be defined as Eq. (29):

$$\begin{aligned} H_0 &: E\left[L\left(\tilde{\xi}_i^{(a)}\right)\right] = E\left[L\left(\tilde{\xi}_i^{(b)}\right)\right] \\ H_1 &: E\left[L\left(\tilde{\xi}_i^{(a)}\right)\right] \neq E\left[L\left(\tilde{\xi}_i^{(b)}\right)\right] \end{aligned} \tag{29}$$

in which $L(\cdot)$ is the loss function of forecasting bias. In point-oriented forecasting, $\tilde{\xi}_i$ is defined as the ith forecasting value minus the corresponding actual value. In interval forecast, $\tilde{\xi}_i$ represented by Eq. (30). is defined as the interval score of the ith point.

$$\xi_i(\alpha)=\begin{cases} -2\alpha\Phi_i(\alpha)-4\left(\boldsymbol{LB}_i(\alpha)-\boldsymbol{Obse}V_i\right), & \boldsymbol{Obse}V_i<\boldsymbol{LB}_i(\alpha) \\ -2\alpha\Phi_i(\alpha), & \boldsymbol{Obse}V_i\in\left[\boldsymbol{LB}_i(\alpha),\boldsymbol{UB}_i(\alpha)\right] \\ -2\alpha\Phi_i(\alpha)-4\left(\boldsymbol{Obse}V_i-\boldsymbol{UB}_i(\alpha)\right), & \boldsymbol{Obse}V_i>\boldsymbol{UB}_i(\alpha) \end{cases} \quad (30)$$

If the null hypothesis is rejected, it is possible to say that there is statistical evidence that there exists a significant difference between the proposed model regarding the compared model at the α level of significance. The DM test results are shown in Table 4.

Table 4
Statistics of DM test for statistical comparison of proposed approach versus other models

Model	**Site 1**		**Site 2**		**Site 3**	
	PINC=95%	**PINC=90%**	**PINC=95%**	**PINC=90%**	**PINC=95%**	**PINC=90%**
SSa-PSr-QrLASso	-10.2893***	-6.7006***	-2.9187***	-3.8964***	-19.8818***	-4.2333***
SSa-PSr-QrLStm	-68.6604***	-66.8613***	-4.5813***	-6.6625***	-42.4881***	-23.2736***
SSa-PSr-QrGRu	-45.0164***	-65.5619***	-11.1758***	-7.9645***	-35.1313***	-20.3249***
SSa-PSr-QrCNN	-4.0386***	-4.5426***	-41.1140***	-9.2481***	-2.9096***	-7.3515***
SSa-PSr-QrBiLStm	-49.3264***	-48.2301***	-2.4558**	-33.6391***	-14.4810***	-9.2998***
SSa-PSr-GPr	-49.1107***	-59.7660***	-11.7846***	-44.0720***	-30.8004***	-21.7912***
SSa-PSr-BLgm	-48.4703***	-60.4071***	-11.8552***	-44.0617***	-31.3836***	-23.5181***
Emd-PSr-EPFM	-30.5495***	-5.8206***	-1.7031*	-17.1494***	-3.0018***	-8.0663***
Eemd-PSr-EPFM	-28.5660***	-13.4367***	-1.6236	-21.1340***	-3.1871***	-6.6468***
Wt-PSr-EPFM	-23.4755***	-28.0409***	-4.0643***	-23.3569***	-2.2902**	-3.8406***
CeemdAN-PSr-EPFM	-18.1712***	-32.4127***	-1.6707*	-9.6597***	-5.0214***	-6.5898***
SSa-PSr-MOTa-EPFM	-7.2410***	-5.0958***	-1.3035	0.3255	-6.5390***	-3.5052***
SSa-PSr-MODa-EPFM	-39.8391***	-5.0069***	-1.2672	-0.6341	-1.7147*	-1.0823
SSa-PSr-MOGa-EPFM	-0.6505	-1.6847*	-1.4715	-2.6176***	-1.9806**	-1.9018*
SSa-PSr-MOAa-EPFM	-2.0946**	-1.5301	-1.7845*	-2.7302***	-1.3332	-1.5073

Note: ***, **, and * indicate that the results are significant at the 1%, 5%, and 10% confidence levels, respectively.

The proposed EPFS is significantly different from the commonly used single models in that the absolute values of the DM test results are all greater than the threshold $Z_{0.01/2}=2.58$. Compared with the EPFM based on different denoising methods, only Site 2 when PINC=95%, the EPFM based on Eemd is not significantly different from the proposed EPFS; In other forecasting scenarios, the proposed EPFS is significantly different from EPFM based on other denoising methods. Compared with EPFM based on different optimization algorithms, the accuracy was not significantly different in a few forecasting situations. Take Site 2 as an example, when PINC=95%, the DM value of SSa-PSr-MOGa-EPFM is -0.6505. Table 4 shows that the EPFM based on different optimization algorithms is significantly different in most forecasting situations, so it is important to use the improved optimization algorithm, i.e., IMOTA to perform the optimization task.

4.4. Fist-order and second-order forecasting effectiveness evaluation

In this study, the forecasting effectiveness (FE) approach [52] was modified for uncertainty forecasting; its first-order and second-order values were used to measure the availability of models. The required bias in FE is modified as interval score which is defined as Eq. (31), and the first-order FE and second-order FE are obtained. The details of this indicator are provided as follows.

The element of the k^{th} order FE can be calculated as $\boldsymbol{g}^k = \sum_{\boldsymbol{i}=1}^{\boldsymbol{n}} \boldsymbol{Q}_i \boldsymbol{A}_i^k$, where $\boldsymbol{A}_i$ refers to the forecasting accuracy that can be measured by $\boldsymbol{A}_i = 1 - |\xi_i|$, where n is the length of the testing set, and ξ_i can be mathematically expressed by Eq. (29). $\boldsymbol{Q}_i$ indicates the discrete probability distribution, and $\sum_{\boldsymbol{i}=1}^{\boldsymbol{n}} \boldsymbol{Q}_i = 1,\ \boldsymbol{Q}_i > 0$. As prior information for $\boldsymbol{Q}_i$ cannot be obtained, it is commonly determined as $\boldsymbol{Q}_i = 1/\boldsymbol{n},\ \boldsymbol{i} = 1, 2, \ldots, \boldsymbol{n}$.

$$\xi_i(\alpha) = \begin{cases} -2\alpha\Phi_i(\alpha) - 4\left(\boldsymbol{LB}_i(\alpha) - \boldsymbol{Obse}V_i\right), & \boldsymbol{Obse}V_i < \boldsymbol{LB}_i(\alpha) \\ -2\alpha\Phi_i(\alpha), & \boldsymbol{Obse}V_i \in \left[\boldsymbol{LB}_i(\alpha), \boldsymbol{UB}_i(\alpha)\right] \\ -2\alpha\Phi_i(\alpha) - 4\left(\boldsymbol{Obse}V_i - \boldsymbol{UB}_i(\alpha)\right), & \boldsymbol{Obse}V_i > \boldsymbol{UB}_i(\alpha) \end{cases} \quad (31)$$

where $\Phi_i(\alpha) = \boldsymbol{UB}_i(\alpha) - \boldsymbol{LB}_i(\alpha)$.

Thereafter, a function $FE\left(\boldsymbol{g}^1, \boldsymbol{g}^2, \cdots, \boldsymbol{g}^k\right)$ that contains k elements is designed to assess the k^{th} order ***FE***. Recall that $\boldsymbol{g}^k = \sum_{\boldsymbol{i}=1}^{\boldsymbol{n}} \boldsymbol{Q}_i \boldsymbol{A}_i^k$, the first-order FE, can be defined as $FE\left(\boldsymbol{g}^1\right) = \boldsymbol{g}^1$, and the second-order ***FE*** is designed as $FE\left(\boldsymbol{g}^1, \boldsymbol{g}^2\right) = \boldsymbol{g}^1\left(1 - \sqrt{\boldsymbol{g}^2 - \left(\boldsymbol{g}^1\right)^2}\right)$. The ***FE*** values are presented in **Table 5**.

According to the design mechanism of ***FE***, a higher ***FE*** value indicates greater availability of models. By comparing the first- and second-order ***FE*** values with reference models, we can determine that, apart from some results for Site 2, the proposed EPFS has the strongest availability. For Site 1, considering both *Order* = 1 and *Order* = 2, the ***FE*** values of the proposed EPFS are higher than those of other forecasting models. The ***FE*** values obtained using the proposed EPFS were $FE_{\boldsymbol{1st\text{-}order}}^{Site1} = 0.8195$ and $FE_{\boldsymbol{2\text{-}order}}^{Site1} = 0.7548$, indicating that the proposed EPFS can fully assimilate the merits of benchmark units and achieve the most satisfactory results at all three sites.

4.5. Improvement ratio

The indicator $\overline{\boldsymbol{IR}}_{\boldsymbol{Metric}}$ was used to assess the improvement in forecasting accuracy of the EPFS. $\overline{\boldsymbol{IR}}_{\boldsymbol{Metric}}$ can be defined as

$$\overline{\boldsymbol{IR}}_{\boldsymbol{Metric}} = \left[\left(\boldsymbol{Metric}^{com} - \boldsymbol{Metric}^{pro}\right) / \boldsymbol{Metric}^{com}\right] \quad (32)$$

where $\boldsymbol{Metric}^{com}$ is the metric value of the compared model, and $\boldsymbol{Metric}^{pro}$ indicates the metric value of the proposed EPFS. In this paper, the improvement rate of AIS is calculated to reflect the overall improvement ratio of the proposed EPFS compared with other models, and the improvement rate of PINAW reflects the contribution of the proposed EPFS to shortening the interval width. **Table 6** presents the improvement ratio results. According to the $\overline{\boldsymbol{IR}}_{\boldsymbol{Metric}}$, the following conclusions can be drawn.

Compared with all other single models, the forecasting performance of the proposed QrBiLStm is improved, indicating that the proposed QrBiLStm is effective.

The EPFS obtained by optimizing two efficient QrBiLStm units can further improve forecasting, manifested as a lower interval width and higher resolution. At Site 1, when $PINC = 90\%$ and $PINC = 95\%$, the improvement ratios of the proposed EPFS to the AIS value of a single model are $\overline{\overline{IR}}_{AIS}^{PINC=90\%}$ = $[66.8235\%, 64.9752\%, 66.1685\%, 51.5732\%, 48.4414\%]$ and $\overline{\overline{IR}}_{PINC=95\%}^{AIS}$ = $[58.3401\%, 56.5286\%, 57.0647\%, 73.1684\%, 39.0008\%]$. Taking into account $PINC = 90\%$ and $PINC = 95\%$ of the three Sites, the proposed EPFS represents a minimum of 12.96% improvement in AIS metrics compared to the EPFMs based on the other four denoising methods. Compared with EPFMs based on other optimization algorithms, the improvement ratio of the AIS metric is 1.57% to 23.60% and the forecasting results are more stable. This indicating that the proposed IMOTa has better global optimization ability and better ability to optimize multiple objectives.

4.6. Stability analysis

As most swarm intelligence optimization methods incorporate randomness or probabilistic mechanisms into their operation, the forecasting results for each trial are generally different, even with the same parameters and conditions. Thus, the stability of swarm intelligence optimization is one of the most important factors affecting prediction performance.

The stability is measured by the standard deviations of three evaluation metrics; the equation is expressed as $\overline{\overline{Std}}(Metric) = \sum_{t=1}^{N} \left(Metric_t^k - Metric_t \right)^2 / N$, where N is the number of testing trials; $Metric_k$ indicates the k^{th} error metric (PINAW and AIS) values in the testing trial, and $Metric_t$ refers to the average metric values of all testing trials. A lower $Std(Metric)$ value indicates a greater degree of stability.

Fig. 5a, b, and d show the distribution of different results from ten trials using IMOTA for the three indicators at $PINC = 90\%$; **Fig. 5** e, f, and h show the distribution of different results at $PINC = 95\%$; **Fig. 5** c and g show the $Std(Metric)$ results for ten trials at $PINC = 90\%$ and $PINC = 95\%$. The stability analysis is conducted using Site 1 as an example. Generally, the $Std(Metric)$ values are small for both $\alpha = 0.1$ and $\alpha = 0.05$, manifested as $\overline{\overline{Std}}(PICP)_{\alpha=0.1,\alpha=0.05}^{\text{Site1}} = [0.0240, 0.0228]$, $\overline{\overline{std}}(PINAW)_{\alpha=0.1,\alpha=0.05}^{\text{Site1}} = [0.0117, 0.0104]$, and $\overline{\overline{std}}(AIS)_{\alpha=0.1,\alpha=0.05}^{\text{Site1}} = [0.0082, 0.0048]$. In ten trials, the values of the three indicators fluctuated little. The values of the three indicators were analyzed. For $PINC = 90\%$, the minimum value of PICP in the ten trials was 0.9389, and the maximum value was 0.9889. PINAW had a maximum of 0.2086, and a minimum of 0.1752. The minimum AIS value was -0.2021, and the maximum value was -0.1783. For PINC = 95%, the minimum value of PICP in the ten trials was 0.9222, and the maximum value was 0.9944. The PINAW maximum was 0.2657, and the minimum was 0.2303. The minimum AIS value was -0.1311, and the maximum value was -0.1163.

In practical applications, future values are not available to calculate the metrics for comparison. However, through the stability analysis, we found that the proposed EPFS can achieve accurate forecasting results in all trials, which shows that the proposed EPFS is highly available.

4.7. Advantages and disadvantages compared to the existing studies

The advantages and disadvantages of this study compared with the existing models in this field and the future work are analyzed in this section.

Advantages

(1) Firstly, compared with the existing QR- machine learning models such as QrGRu [29], QrLStm [30], , QrLASso [27], QrCNN, etc., this study adopts the cell structure of BiLStm, which can train the network by inputting historical information forward and backward. Based on this network structure, more accurate interval forecasting results can be obtained. This conclusion can be drawn from the results of comparative experiment 1 (section 4.2.1). Secondly, this study proposes a probabilistic ensemble forecasting system, which can combine the forecasting results of two well-behaved single models to get more accurate forecasting results, specifically by reducing the interval width under the condition of ensuring high interval coverage.

(2) Compared with the existing probabilistic ensemble forecasting models. Firstly, the proposed EPFS is based on Qr theory, which can optimize the upper and lower bounds of the interval respectively. This optimization strategy is more flexible, ensuring that both the upper bound and the lower bound are optimal results. For example, Niu's model [36] is based on data distribution, and the upper and lower bound forecasting results are optimized simultaneously. Secondly, the objective functions of probabilistic ensemble optimization are designed to find the solution with high coverage and narrow interval width as the optimal solution.

Disadvantages

(1) The calculation burden is increased while the upper and lower bounds of the forecasting interval are optimized respectively. Every interval forecasting result obtained by EPFS needs to be optimized twice, which increases the operation time.

(2) The distribution information of data is not used to construct the interval. In this paper, loss function of the neural network is designed as pinball loss to obtain different quantile forecasting results, which is a supervised machine learning method without data distribution information. Although the EPFS based on Qr can get accurate interval forecasting results after optimization, it is possible to get better forecasting results if the distribution information of historical data can be fully utilized. This is the future work.

(3) The EPFS is based on historical wind-speed data without considering other influence factors, such as pressure and temperature. Probabilistic ensemble forecast of multivariate time series is also the future research direction.

Table 5

The forecasting effectiveness of the proposed EPFS and other existing models.

Model	Site 1				Site 2				Site 3			
	PINC=90%		PINC=95%		PINC=90%		PINC=95%		PINC=90%		PINC=95%	
	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
SSa-PSr-QrLASso	0.4561	0.3296	0.6945	0.6149	0.5591	0.3384	0.6611	0.4336	0.5843	0.4315	0.6951	0.6476
SSa-PSr-QrLStm	0.4848	0.4591	0.7072	0.6889	0.4937	0.3964	0.7116	0.6228	0.5367	0.4978	0.6890	0.6643
SSa-PSr-QrGRu	0.4666	0.4407	0.7036	0.6714	0.5095	0.4227	0.7070	0.6611	0.5461	0.4979	0.7067	0.6758
SSa-PSr-QrCNN	0.6274	0.4947	0.5257	0.3193	0.6046	0.5458	0.4132	0.3823	0.5492	0.4249	0.7649	0.6015
SSa-PSr-GPr	0.5443	0.5407	0.7288	0.7285	0.5315	0.5311	0.7209	0.7206	0.5707	0.5574	0.7481	0.7472
SSa-PSr-BLgm	0.5444	0.5426	0.7284	0.7271	0.5305	0.5287	0.7203	0.7188	0.5734	0.5532	0.7456	0.7443
SSa-PSr-QrBiLStm	0.6500	0.6131	0.7914	0.7577	0.6198	0.5771	0.8252	0.7941	0.7229	0.6775	0.8165	0.7815
Emd-PSr-EPFM	0.7473	0.6834	0.8108	0.7721	0.7113	0.6645	0.8367	0.7951	0.7293	0.6630	0.8428	0.7864
Eemd-PSr-EPFM	0.7774	0.7330	0.8155	0.7761	0.7009	0.6523	0.8388	0.7960	0.7308	0.6770	0.8419	0.7839
Wt-PSr-EPFM	0.6547	0.6034	0.8059	0.7627	0.6477	0.5934	0.8055	0.7660	0.6800	0.5679	0.8127	0.7041
CeemdAN-PSr-EPFM	0.7205	0.6742	0.8168	0.7691	0.7495	0.7019	0.8555	0.7854	0.7351	0.6874	0.8417	0.8008
SSa-PSr-MOTa-EPFM	0.8003	0.7309	0.8647	0.8266	0.8373	**0.7728**	0.8710	0.7661	0.7965	0.7357	0.8535	0.8073
SSa-PSr-MODa-EPFM	0.8101	0.7372	0.8334	0.8017	**0.8431**	0.7561	0.8636	0.7278	0.8190	0.7211	0.8622	0.7936
SSa-PSr-MOGa-EPFM	0.8113	0.7006	**0.8777**	0.8111	0.8293	0.7378	0.8752	0.7773	0.8154	0.6985	0.8552	0.7766
SSa-PSr-MOAa-EPFM	0.8157	0.7135	0.8578	0.7401	0.8264	0.7228	0.8482	0.6991	0.8124	0.7082	0.8655	0.7893
Proposed EPFS	**0.8195**	**0.7548**	0.8727	**0.8302**	0.8410	0.7660	**0.8784**	**0.8072**	**0.8216**	**0.7477**	**0.8676**	**0.8232**

Note: This table shows the circumstantial values of the *first*-order and *second*-order ***FE*** of sixteen models. The *first*-order ***FE*** is defined as $FE\left(g^1\right)=g^1$. When this continues function contains two variables, the *2nd*-order ***FE*** can be denoted by $FE\left(g^1,g^2\right)=g^1\left(1-\sqrt{g^2-\left(g^1\right)^2}\right)$.

Table 6

The improvement ratio of the proposed EPFS.

Models	Site 1				Site 2				Site 3			
	PINAW		AIS		PINAW		AIS		PINAW		AIS	
	PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%
SSa-PSr-QrLASso	0.5844	0.4994	0.6682	0.5834	0.4650	0.4976	0.6394	0.6412	0.4712	0.5164	0.5708	0.5657
SSa-PSr-QrLStm	0.6332	0.5417	0.6498	0.5653	0.6604	0.5439	0.6860	0.5784	0.6085	0.5591	0.6149	0.5743
SSa-PSr-QrGRu	0.6454	0.5533	0.6617	0.5706	0.6560	0.5687	0.6759	0.5850	0.6038	0.5369	0.6069	0.5486
SSa-PSr-QrCNN	0.3984	0.6469	0.5157	0.7317	0.5662	0.7890	0.5980	0.7928	0.5247	0.2421	0.6042	0.4370
SSa-PSr-GPr	0.5753	0.4966	0.6040	0.5307	0.6437	0.5440	0.6607	0.5643	0.5665	0.4483	0.5800	0.4745
SSa-PSr-BLgm	0.5756	0.4971	0.6039	0.5314	0.6443	0.5448	0.6614	0.5653	0.5709	0.4533	0.5837	0.4797
SSa-PSr-QrBiLStm	0.4574	0.3708	0.4844	0.3900	0.5661	0.2928	0.5819	0.3042	0.3558	0.2714	0.3562	0.2787
SSa-PSr-MOTa-EPFM	0.0758	0.0545	0.0965	0.0596	0.0423	0.0020	0.0229	0.0578	0.1243	0.0829	0.1231	0.0961
SSa-PSr-MODa-EPFM	0.0436	0.2166	0.0500	0.2360	-0.1143	0.0164	0.0132	0.1083	-0.0047	0.0093	0.0141	0.0395
SSa-PSr-MOGa-EPFM	-0.0787	-0.1075	0.0438	0.0407	-0.0317	-0.0550	0.0690	0.0256	-0.0883	0.0364	0.0334	0.0860
SSa-PSr-MOAa-EPFM	-0.0702	0.0466	0.0208	0.0471	-0.0983	-0.0360	0.0846	0.1988	-0.0038	-0.0399	0.0487	0.0157
Emd-PSr-EPFM	0.2542	0.3150	0.2859	0.3274	0.4453	0.2670	0.4494	0.2553	0.3273	0.1296	0.3408	0.1580
Eemd-PSr-EPFM	0.1685	0.3008	0.1894	0.3104	0.4647	0.2592	0.4686	0.2457	0.3394	0.1397	0.3372	0.1625
Wt-PSr-EPFM	0.4612	0.3368	0.4774	0.3442	0.5470	0.3792	0.5488	0.3747	0.4150	0.2354	0.4424	0.2930
CeemdAN-PSr-EPFM	0.3399	0.3051	0.3544	0.3052	0.3563	0.1311	0.3654	0.1584	0.3299	0.1599	0.3264	0.1639

Note: The table above reports the IR of the proposed EPFS from other twelve models. The AIS and PINAW are used to measure the IR, and the corresponding indicator can be defined as $\overline{IR}_{Metric} = \left[\left(Metric^{com} - Metric^{pro}\right) / Metric^{com}\right]$, where $Metric^{com}$ is the metric values of compared model, and the $Metric^{pro}$ indicates the metric value of the proposed **EPFS**.

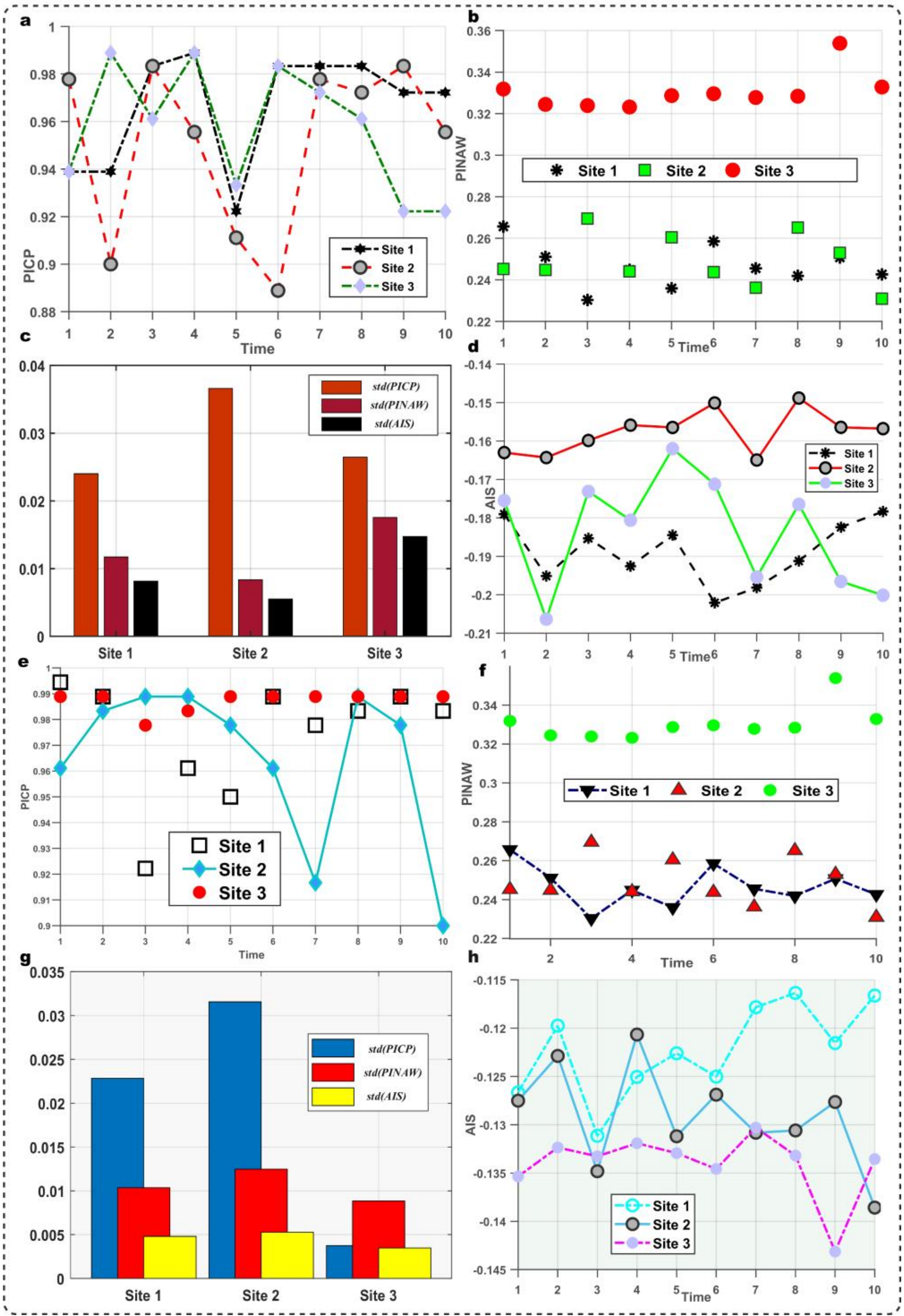


Fig.5. Stability analysis results of the proposed EPFS.

5. Conclusion

To generate high-quality wind speed prediction intervals and obtain comprehensive potential uncertainty information, an EPFS that combines SSa, PSr, QrBiSLStm, pseudo-interval construction, and multi-objective optimization was proposed. SSa and PSr were implemented successively, and time-frequency decomposition and reconstruction of the sequence were performed. The conditional quantiles of the sequences were obtained using the proposed QrBiLStm, and prediction intervals with different confidence levels were constructed. A pseudo-interval was constructed as the training set based on the forecasting results of two QrBiLStm units, and the proposed IMOTa was used for combinatorial optimization to obtain the final interval forecasting results. Comparison experiments were performed on three datasets, and the forecasting results were comprehensively evaluated in terms of reliability, resolution, and sharpness. Based on the analysis, we can draw the following conclusions: (1) a decomposition and reconstruction mechanism based on SSa and PSr can significantly improve forecasting performance, which can greatly improve uncertainty forecasting performance of the EPFS; (2) reliable uncertainty forecasts can be obtained from newly constructed QrBiLStm units; the forecasting results of this model far exceed those of other single models based on distribution hypothesis, and those of other Qr–deep learning models; (3) an ensemble probabilistic forecasting strategy based on pseudo-interval construction can effectively optimize the upper and lower bounds of the interval and further improve the forecasting performance of the main forecasting model; (4) the improved MOTa has better global optimization ability and stability, and produces more effective and stable prediction results. The main limitations of the proposed EPFS are as follows: (1) Because quantile loss is discontinuous and nondifferentiable around 0 point, this EPFS has not been applied to the field of deterministic forecasts; (2) It is not combined with other linear models or interval forecasting models based on distribution in ensemble forecasting. However, this EPFS can optimize the upper and lower bounds of the interval separately and does not need to assume the distribution in advance. This study provides a novel approach for wind speed ensemble probabilistic forecasting and can be used as a powerful decision tool in the power system scheduling process.

Acknowledgements

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

Appendix A

Table A1

Parameters setting.

Models	Symbol	Meaning	Determination method	determined value
Eemd	Ns	STD of added noise	Preset	0.05
	NR	Realization Number	Preset	50
	Ms	Maximum Sifting Iteration	Preset	500
CeemdAN	Ns	STD of added noise	Preset	0.05
	NR	Realization Number	Preset	50
	Ms	Maximum Sifting Iteration	Preset	500
Wt	D1	Decomposition Layer Number	Trial and error approach	5
SSa	W	Window Length	Trial and error approach	50
	D2	Primary Ingredient Disintegration Number	Karhunene Loeve decomposition	20
QrLStm	Nl	Number of hidden layers	Trial and error approach	2
	Nn	Number of hidden nodes	Trial and error approach	50
QrBILStm	Nl	Number of hidden layers	Trial and error approach	1
	Nn	Number of hidden nodes	Trial and error approach	50
QrGRu	Nl	Number of hidden layers	Trial and error approach	2
	Nn	Number of hidden nodes	Trial and error approach	50
QrCNN	Ck	Convolution kernel size	Trial and error approach	2*2
	Nn	Number of hidden layer nodes	Trial and error approach	35
MOGa, MOAa, MODa	As	Archive Size	Preset	10
	In	Iteration Number	Preset	200
	Ni	Individual Number	Trial and error approach	40
MOTa, IMOTa	Si	Initial Speed	Preset	1
	Nt	Number of tunicate	Trial and error approach	40
	Ss	Subordinate Speed	Preset	4
	As	Archive Size	Trial and error approach	10
	In	Iteration Number	Trial and error approach	200
PSr	τ	Embedded dimension	C-C method	Site 1~3: 4, 5,5
	M	Delay time	C-C method	Site 1~3: 31, 28, 28

Table A2
List of abbreviations.

ARIMA	Auto Regressive Integrated Moving Average	**LStm**	Long short-term memory neural network
ARMA	Autoregressive moving average	**LUBE**	lower upper bound estimation
BiLStm	Bi-directional Long Short Term Memory Network	**MO**	Multi-objective
BLgm	Bayesian regression model	**MOAa**	MO Antlion Algorithm
BPNN	Back propagation neural network	**MODa**	MO Dragonfly Algorithm
CeemdAN	Complete ensemble Empirical Mode Decomposition with Adaptive Noise	**MOGa**	MO Grasshopper Algorithm
Eemd	Ensemble Empirical Mode Decomposition	**MVE**	mean-variance estimates
EFS	Exponential function steps	**PINC**	prediction interval nominal confidence
ELM	Extreme learning machine	**PSr**	Phase space reconstruction
Emd	Empirical Mode Decomposition	**Qr**	quantile regression
EOL	Elite opposition learning	**QrCNN**	Quantile regression convolution neural network
EPFM	Ensemble probabilistic forecasting model	**QrGRu**	Quantile regression gated recurrent unit
EPFS	Ensemble probabilistic forecasting system	**QrLStm**	Quantile regression Long Short-Term Memory Network
GPr	Gaussian Process Regression	**RNN**	Recurrent neural network
GRu	Gated recurrent unit	**SSa**	Singular Spectral Analysis
GW	Gigawatt	**SVQr**	Support Vector Quantile Regression
GWEC	Global wind energy councile	**SVR**	Support Vector Regression
IMOTa	Improved multi-objective tunicate swarm algorithm	**Wt**	Wavelet Transform

Data Availability

10-minute wind speed data of three Sites in Shandong Peninsula: https://data.mendeley.com/datasets/sjyf2nhzdt/draft?a=af12330a-125b-499a-9473-6840ed7044f9

References

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

[2] Jiang P, Liu Z, Niu X, Zhang L. A combined forecasting system based on statistical method, artificial neural networks, and deep learning methods for short-term wind speed forecasting. Energy 2020:119361. https://doi.org/10.1016/j.energy.2020.119361.

[3] Jung J, Broadwater RP. Current status and future advances for wind speed and power forecasting. Renew Sustain Energy Rev 2014. https://doi.org/10.1016/j.rser.2013.12.054.

[4] Sheng C, Zhao J, Wang W, Leung H. Prediction intervals for a noisy nonlinear time series based on a bootstrapping reservoir computing network ensemble. IEEE Trans Neural Networks Learn Syst 2013;24:1036–48. https://doi.org/10.1109/TNNLS.2013.2250299.

[5] Niu X, Wang J, Zhang L. Carbon price forecasting system based on error correction and divide-conquer strategies. Appl Soft Comput 2021:107935. https://doi.org/10.1016/j.asoc.2021.107935.

[6] Aasim, Singh SN, Mohapatra A. Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting. Renew Energy 2019. https://doi.org/10.1016/j.renene.2019.01.031.

[7] Shukur OB, Lee MH. Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA. Renew Energy 2015. https://doi.org/10.1016/j.renene.2014.11.084.

[8] Manero J, Béjar J, Cortés U. Wind energy forecasting with neural networks: A literature review. Comput y Sist 2018. https://doi.org/10.13053/CyS-22-4-3081.

[9] Du P, Wang J, Yang W, Niu T. Multi-step ahead forecasting in electrical power system using a hybrid forecasting system. Renew Energy 2018. https://doi.org/10.1016/j.renene.2018.01.113.

[10] 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. https://doi.org/10.1016/J.RENENE.2021.07.113.

[11] 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. https://doi.org/10.1016/j.apenergy.2020.115561.

[12] Moreno SR, Coelho LDS, Ayala HVH, Mariani VC. Wind turbines anomaly detection based on power curves and ensemble learning. IET Renew Power Gener 2020. https://doi.org/10.1049/iet-rpg.2020.0224.

[13] 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. https://doi.org/10.1016/j.apenergy.2019.03.097.

[14] Chen H, Birkelund Y, Zhang Q. Data-augmented sequential deep learning for wind power forecasting. Energy Convers Manag 2021;248:114790.

https://doi.org/10.1016/J.ENCONMAN.2021.114790.

[15] Khodayar M, Wang J, Manthouri M. Interval Deep Generative Neural Network for Wind Speed Forecasting. IEEE Trans Smart Grid 2019. https://doi.org/10.1109/TSG.2018.2847223.

[16] Du P, Wang J, Yang W, Niu T. A novel hybrid model for short-term wind power forecasting. Appl Soft Comput J 2019. https://doi.org/10.1016/j.asoc.2019.03.035.

[17] Salkind N. American Statistical Association. Encycl Res Des 2012;92:748–57. https://doi.org/10.4135/9781412961288.n9.

[18] Khosravi A, Nahavandi S. An optimized mean variance estimation method for uncertainty quantification of wind power forecasts. Int J Electr Power Energy Syst 2014. https://doi.org/10.1016/j.ijepes.2014.03.060.

[19] Lian C, Zhu L, Zeng Z, Su Y, Yao W, Tang H. Constructing prediction intervals for landslide displacement using bootstrapping random vector functional link networks selective ensemble with neural networks switched. Neurocomputing 2018. https://doi.org/10.1016/j.neucom.2018.02.046.

[20] Khosravi A, Mazloumi E, Nahavandi S, Creighton D, Van Lint JWC. Prediction intervals to account for uncertainties in travel time prediction. IEEE Trans Intell Transp Syst 2011;12:537–47. https://doi.org/10.1109/TITS.2011.2106209.

[21] Quan H, Srinivasan D, Khosravi A. Short-term load and wind power forecasting using neural network-based prediction intervals. IEEE Trans Neural Networks Learn Syst 2014;25:303–15. https://doi.org/10.1109/TNNLS.2013.2276053.

[22] He Y, Li H. Probability density forecasting of wind power using quantile regression neural network and kernel density estimation. Energy Convers Manag 2018. https://doi.org/10.1016/j.enconman.2018.03.010.

[23] Wang HZ, Wang GB, Li GQ, Peng JC, Liu YT. Deep belief network based deterministic and probabilistic wind speed forecasting approach. Appl Energy 2016. https://doi.org/10.1016/j.apenergy.2016.08.108.

[24] Das K, Krzywinski M, Altman N. Quantile regression. Nat Methods 2019. https://doi.org/10.1038/s41592-019-0406-y.

[25] Taylor JW. A quantile regression neural network approach to estimating the conditional density of multiperiod returns. J Forecast 2000. https://doi.org/10.1002/1099-131x(200007)19:4<299::aid-for775>3.3.co;2-m.

[26] He Y, Liu R, Li H, Wang S, Lu X. Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory. Appl Energy 2017. https://doi.org/10.1016/j.apenergy.2016.10.079.

[27] He Y, Qin Y, Lei X, Feng N. A study on short-term power load probability density forecasting considering wind power effects. Int J Electr Power Energy Syst 2019. https://doi.org/10.1016/j.ijepes.2019.05.063.

[28] Chen T, Qian Z, Jing B, Wan J, Zhang F. Probabilistic Wind Speed Forecasting based on Minimal Gated Unit and Quantile Regression. J. Phys. Conf. Ser., 2020. https://doi.org/10.1088/1742-6596/1659/1/012039.

[29] Zhang Z, Qin H, Liu Y, Yao L, Yu X, Lu J, et al. Wind speed forecasting based on Quantile Regression Minimal Gated Memory Network and Kernel Density Estimation. Energy Convers Manag 2019.

https://doi.org/10.1016/j.enconman.2019.06.024.
[30] Wang K, Fu W, Chen T, Zhang B, Xiong D, Fang P. A compound framework for wind speed forecasting based on comprehensive feature selection, quantile regression incorporated into convolutional simplified long short-term memory network and residual error correction. Energy Convers Manag 2020;222:113234. https://doi.org/10.1016/j.enconman.2020.113234.
[31] Jiang P, Liu Z, Wang J, Zhang L. Decomposition-selection-ensemble forecasting system for energy futures price forecasting based on multi-objective version of chaos game optimization algorithm. Resour Policy 2021;73:102234. https://doi.org/10.1016/j.resourpol.2021.102234.
[32] Ribeiro MHDM, da Silva RG, Moreno SR, Mariani VC, Coelho L dos S. Efficient bootstrap stacking ensemble learning model applied to wind power generation forecasting. Int J Electr Power Energy Syst 2022;136:107712. https://doi.org/10.1016/J.IJEPES.2021.107712.
[33] Wang J, Du P, Lu H, Yang W, Niu T. An improved grey model optimized by multi-objective ant lion optimization algorithm for annual electricity consumption forecasting. Appl Soft Comput J 2018. https://doi.org/10.1016/j.asoc.2018.07.022.
[34] Liu Z, Jiang P, Wang J, Zhang L. Ensemble forecasting system for short-term wind speed forecasting based on optimal sub-model selection and multi-objective version of mayfly optimization algorithm. Expert Syst Appl 2021;177:114974. https://doi.org/10.1016/j.eswa.2021.114974.
[35] 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 2021:121275. https://doi.org/10.1016/j.energy.2021.121275.
[36] Wang J, Niu T, Du P, Yang W. Ensemble probabilistic prediction approach for modeling uncertainty in crude oil price. Appl Soft Comput J 2020;95:106509. https://doi.org/10.1016/j.asoc.2020.106509.
[37] Miao K chao, Han T ting, Yao Y qing, Lu H, Chen P, Wang B, et al. Application of LSTM for short term fog forecasting based on meteorological elements. Neurocomputing 2020. https://doi.org/10.1016/j.neucom.2019.12.129.
[38] Wang Y, Wang J, Li Z, Yang H, Li H. Design of a combined system based on two-stage data preprocessing and multi-objective optimization for wind speed prediction. Energy 2021;231:121125. https://doi.org/10.1016/j.energy.2021.121125.
[39] Wang J, Li Q, Zeng B. Multi-layer cooperative combined forecasting system for short-term wind speed forecasting. Sustain Energy Technol Assessments 2021. https://doi.org/10.1016/j.seta.2020.100946.
[40] Khosravi A, Nahavandi S, Creighton D, Jaafar J. Wind farm power uncertainty quantification using a mean-variance estimation method. 2012 IEEE Int. Conf. Power Syst. Technol. POWERCON 2012, 2012. https://doi.org/10.1109/PowerCon.2012.6401280.
[41] Efron B, Tibshirani R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Stat Sci 1986. https://doi.org/10.1214/ss/1177013815.
[42] DiCiccio TJ, Efron B. Bootstrap confidence intervals. Stat Sci 1996.

https://doi.org/10.1214/ss/1032280214.

[43] Lian C, Zeng Z, Wang X, Yao W, Su Y, Tang H. Landslide displacement interval prediction using lower upper bound estimation method with pre-trained random vector functional link network initialization. Neural Networks 2020;130:286–96. https://doi.org/10.1016/j.neunet.2020.07.020.

[44] Waldmann E. Quantile regression: A short story on how and why. Stat Modelling 2018. https://doi.org/10.1177/1471082X18759142.

[45] Afshar K, Bigdeli N. Data analysis and short term load forecasting in Iran electricity market using singular spectral analysis (SSA). Energy 2011;36:2620–7. https://doi.org/10.1016/j.energy.2011.02.003.

[46] Xu Z, Zhong L, Zhang A. Phase Space Reconstruction-Based Conceptor Network for Time Series Prediction. IEEE Access 2019. https://doi.org/10.1109/ACCESS.2019.2952365.

[47] Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Comput 1997. https://doi.org/10.1162/neco.1997.9.8.1735.

[48] Caterini AL, Chang DE. Recurrent neural networks. SpringerBriefs Comput. Sci., 2018. https://doi.org/10.1007/978-3-319-75304-1_5.

[49] Schuster M, Paliwal KK. Bidirectional recurrent neural networks. IEEE Trans Signal Process 1997. https://doi.org/10.1109/78.650093.

[50] Tyapkin VN, Ishchuk IN, Kabulova EG, Semenov ME, Meleshenko PA. Singular spectral analysis in filtration of noisecontaminated signals of pseudolite navigation. Indian J Sci Technol 2016. https://doi.org/10.17485/ijst/2016/v9i46/107567.

[51] Moreno SR, Mariani VC, Coelho L dos S. Hybrid multi-stage decomposition with parametric model applied to wind speed forecasting in Brazilian Northeast. Renew Energy 2021. https://doi.org/10.1016/j.renene.2020.10.126.

[52] wang shan, shuai hui, Liu Q. Phase Space Reconstruction Driven Spatio-Temporal Feature Learning for Dynamic Facial Expression Recognition. IEEE Trans Affect Comput 2020. https://doi.org/10.1109/TAFFC.2020.3007531.

[53] Kaur S, Awasthi LK, Sangal AL, Dhiman G. Tunicate Swarm Algorithm: A new bio-inspired based metaheuristic paradigm for global optimization. Eng Appl Artif Intell 2020. https://doi.org/10.1016/j.engappai.2020.103541.

[54] Cui M, Krishnan V, Hodge BM, Zhang J. A Copula-Based Conditional Probabilistic Forecast Model for Wind Power Ramps. IEEE Trans Smart Grid 2019. https://doi.org/10.1109/TSG.2018.2841932.

[55] Quan H, Srinivasan D, Khosravi A. Uncertainty handling using neural network-based prediction intervals for electrical load forecasting. Energy 2014. https://doi.org/10.1016/j.energy.2014.06.104.

[56] Jiang Z, Che J, Wang L. Ultra-short-term wind speed forecasting based on EMD-VAR model and spatial correlation. Energy Convers Manag 2021;250:114919. https://doi.org/10.1016/J.ENCONMAN.2021.114919.

[57] da Silva RG, Ribeiro MHDM, Moreno SR, Mariani VC, Coelho L dos S. A novel decomposition-ensemble learning framework for multi-step ahead wind energy forecasting. Energy 2021. https://doi.org/10.1016/j.energy.2020.119174.

[58] Diebold FX, Mariano RS. Comparing predictive accuracy. J Bus Econ Stat 1995. https://doi.org/10.1080/07350015.1995.10524599.

[59] Song J, Wang J, Lu H. A novel combined model based on advanced

optimization algorithm for short-term wind speed forecasting. Appl Energy 2018;215:643–58. https://doi.org/10.1016/j.apenergy.2018.02.070.

A Novel Ensemble Probabilistic Forecasting System for Uncertainty in Wind Speed

Jianzhou Wang[a], Shuai Wang[b,*], Bo Zeng[c], Haiyan Lu[d]
[a] *Institute of Systems Engineering, Macau University of Science and Technology, Macau 999078, China*
[b] *School of Statistics, Dongbei University of Finance and Economics, Dalian, China*
[c] *Collaborative Innovation Center for Chongqing's Modern Trade Logistics & Supply Chain, Chongqing Technology and Business University, Chongqing 400067, People's Republic of China*
[d] *School of Computer Science, 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 18742018422.
E-mail address: vvs09061513@163.com.

Abstract

The quantification of wind speed uncertainty is of great significance for real-time control of wind turbines and power grid dispatching. However, the intermittence and fluctuation of wind energy present great challenges in modeling its uncertainty; research in this field is limited. A quantile regression bi-directional long short-term memory network (QrBiLStm) and a novel ensemble probabilistic forecasting strategy are proposed in this study to explore ensemble probabilistic forecasting. To verify the reliability of the proposed ensemble probabilistic forecasting system, the uncertainties of wind speed at wind farms in China were modeled as a case study. The results of comparative experiments including 15 other models demonstrate the superiority of this ensemble probabilistic forecasting system in terms of sharpness while maintaining high interval coverage. The forecasting interval coverage probability obtained by the proposed system is above 97%, and the sharpness is improved by at least 24.21% as compared with the commonly used single models. The proposed ensemble probabilistic forecasting system can accurately quantify the uncertainty of wind speed and reduce the operation cost of power systems by improving the efficiency of wind energy utilization.

Keywords: *Wind speed forecasts; Multi-objective optimization algorithm; Deep learning; ensemble probabilistic strategy; Forecast uncertainty*

1. Introduction

Wind energy has attracted extensive attention as an inexhaustible, clean, and inexpensive form of renewable energy. According to the Global Wind Report released by GWEC in 2021, the 93GW of new installations brings global cumulative wind power capacity up to 743 GW [1]. However, volatility and randomness of wind energy pose great challenges to wind energy grid connection and grid scheduling [2]. Decision-makers must calculate and process the forecasted wind speed to obtain corresponding energy information [3]. Thus, wind-speed forecasting is critical for wind energy utilization.

Wind speed forecasting approximates or extracts the potential relationship behind the data; point-oriented forecasting is the most common form [4]. A data-driven model of point forecasting can use traditional statistical models and artificial intelligence models. Traditional statistical models include autoregressive moving average (ARMA) [5], autoregressive integrated moving average (ARIMA) [6], and Kalman filtering [7], etc. These models are based on a linear assumption, and produce forecasting results that are not accurate with nonlinear sequences [8]. With continuous development of artificial intelligence technology, researchers have begun to apply artificial neural networks to wind speed forecasts. Shallow models including the back propagation neural network (BPNN) [9], the extreme learning machine (ELM) [10], and support vector regression (SVR) [11] were first used. This type of supervised AI model can capture the nonlinear characteristics of wind speed series and manage long series [12]; the forecasting accuracy is higher than that of traditional statistical models [13]. However, there are some defects such as under-fitting, over-fitting, and long training time. With the development of deep learning technology, variants of recurrent neural networks (RNNs) [14] such as long short-term memory (LStm), gated recurrent units (GRu), and BiLStm networks have demonstrated excellent performance in time series forecasting [15]. This type of model can store historical information and facilitate capture of nonlinear features in wind speed series. These models often have many hyperparameters that must be set and weights to be updated; thus, they are subject to long training time and difficulty in parameter optimization [16].

As wind speed data usually fluctuate, point-oriented forecasting can be inaccurate for grid scheduling purposes; thus, interval forecasting has become popular [17]. Interval forecasting approaches include mean-variance estimates (MVE) [18], bootstrap [19], Bayesian [20], and the lower upper bound estimation (LUBE) [21]. These methods have advantages and disadvantages, summarized in **Table 1**. Quantile regression (Qr) [22] is usually used in uncertainty forecasting for its strong interpretability in estimating the conditional distribution of the dependent variable. With the limitations of Qr with nonlinear series, research has begun to focus on combining Qr and artificial intelligence models to expand uncertainty forecasting ability [23,24]. In 2000, Taylor [25] proposed a method for combining Qr with a neural network that could solve both linear and nonlinear problems. Based on QrNN, researchers began to combine Qr with other single models. Support Vector Quantile Regression (SVQr) [26] was developed to forecast the probability density of short-term wind power, and can effectively quantify the uncertainty of time series data. He et al [27] forecasted the probability density of electricity consumption based on QrLASso. This method can better learn high-dimensional data, with more accurate forecasting results. As RNNs have more advantages in time series forecasting, researchers have combined Qr with LStm and GRU, proposing QrLStm [28] and QrGRu [29], which further improve forecasting accuracy. Wang et al. [30] incorporated Qr into a convolution-simplified long-term and short-term memory network. This improved model shortened the training time without reducing the accuracy. Based on these studies, we incorporated Qr with BiLStm, proposing QrBiLStm to quantify the uncertainty of wind speed.

The shortcomings of a single model are obvious. In practical applications, the forecasting accuracy of a single model can be high or low in different regions. Thus, another focus of this study is the ensemble forecasting strategy [31]. The ensemble model weighs several well-performing models according to errors using an intelligent optimization algorithm [32]; forecasting is more stable and accurate than with single models [33]. Ensemble forecasting research focuses mostly on point-oriented forecasting. Liu et.al. [34] developed a multi-objective version of the mayfly optimization algorithm, combining several accurate single models to achieve more accurate forecasting. Wang et.al. [35] proposed the addition of two deep learning models to the ensemble forecasting framework, and used the improved dragonfly optimization algorithm to obtain more accurate point forecasting results. However, research on ensemble probabilistic forecasting has received little attention, limiting further development. Niu et al. [36] proposed the use of multiple single models for interval forecasting based on the distribution assumption, and used an optimization algorithm to integrate the results of the single models to obtain the final forecasting results. This approach provides ideas for ensemble probabilistic forecasting. However, with the need to fit the data distribution and estimate the parameters, its usability is limited in practice. The accuracy of the ensemble model depends on the forecasting performance of single models; thus, we propose two QrBiLStm models with excellent performance as benchmark models and use an improved optimization algorithm to realize ensemble probabilistic forecasting.

The main innovations and contributions of this study are summarized as follows:

(1) **The deep QrBiLStm model for wind speed uncertainty modeling was successfully designed, implemented, and tested.** The proposed QrBiLStm model can obtain interval forecasting results with high interval coverage probability and narrower interval width, and provide more accurate information for wind energy utilization.

(2) **A pseudo interval was proposed, and pseudo-interval evaluation indicators were successfully designed** as a foundation for the ensemble probabilistic forecasting system (EPFS). The pseudo-interval training approach enables separate optimization of the upper and lower bounds of the interval. Optimized wind-speed interval forecasting results are more accurate which means less wind estimation fluctuation and less uncertainty. Thus, the proposed EPFS is of great significance to the safety dispatch and operation of wind power generation.

(3) **An ensemble probabilistic forecasting system was proposed, and optimization objective functions for ensemble forecasting were designed.** The experimental results show that the proposed EPFS based on QrBiLStm is a significant improvement over the single models. The EPFS overcomes the limitations of the single model forecast, making the wind speed forecasting results more stable and practical.

(4) **The tunicate swarm algorithm (TSa) was improved and used to perform interval ensemble optimization. The Tsa with the addition of archiving and a roulette wheel can output Pareto optimal solutions.** Comparative experiments show that the TSa with three improved strategies has better global optimization ability and more stable optimization. The improved Tsa can ensure more stable wind-speed interval forecasting results at a faster speed.

(5) Based on singular spectral analysis (SSa) and phase space reconstruction (PSr), the original wind speed sequence was decomposed and reconstructed, enabling the ensemble forecasting model to solve the chaos phenomenon and eliminate small fluctuations, with better forecasting results.

Table 1
Advantages and disadvantages of wind speed forecasting models.

Models	References	Advantages	Disadvantages
ARMA, ARIMA, and Kalman filtering	[6, 7]	The model is simple and only needs endogenous variables; Accurately forecast the linear sequences.	Low forecasting accuracy in nonlinear data; The data is required to be stable or differentially stable.
AI Model (BPNN, ELM, SVR, LStm, and GRu)	[9–11], [37]	Strong robustness and fault tolerance to noise data; Have the ability of association, and can approximate any nonlinear relationship.	The calculation burden is high, and the interpretability is poor; It is difficult to determine the hyperparameter values.
Ensemble Model	[34], [38,39]	The forecasting accuracy on different data types can be ensured; Take advantages of each single model.	Need to train multiple models and choose efficient empowerment technique.
MVE	[18,40]	The computational burden is relatively small.	The accuracy is largely affected by the effect of numerical predictions associated with it; the underestimation of data variance will result in low coverage of real data by prediction intervals.
Bootstrap	[41,42]	High efficiency in small-scale data.	Is a resampling method that requires significant computational cost for large data sets.
Bayesian	[20]	Improve the generalization ability of model.	The calculation burden is large, which requires the calculation of the Hessian matrix. When the data size is not large enough, the accuracy largely depends on prior knowledge.
LUBE	[21,43]	It avoids the problem of numerical calculation of the Jacobian matrix and Hessian matrix.	Heavy computational burden. No suitable parameter initialization method.
Quantile Regression (Qr)	[25,44]	Ability to resolve heterogeneity issues; Tail features of the distribution can be captured.	Traditional Qr model can't solve nonlinear problems, so it is necessary to select a suitable neural network to combine with Qr.

2. Ensemble probabilistic forecasting system (EPFS)

In the EPFS, SSa [42] is used to decompose the reconstructed sequence, and PSr [43] is used to reconstruct an one-dimensional sequence into a dynamic chaotic space. The processed sequences are forecasted in two QrBiLStm units. The proposed IMOTa algorithm is used to aggregate the two QrBiLStm units to generate an effective wind speed forecasting interval. The details of the QrBiLStm, SSa, PSr, and IMOTa algorithms are described as follows.

2.1. Quantile Regression Bi-directional Long Short Term Memory Network

This section introduces the basic structure of BiLStm and the generation of QrBiLStm.

2.1.1. Bi-directional Long Short Term Memory Network

LStm proposed by S. Hochreiter [44], and is an RNN variant [45]. Owing to its cell structure, LStm can solve the problems of gradient disappearance and gradient explosion in long-sequence training. The cell structure consists of an input gate ($\boldsymbol{\Theta}_t$), a forgetting gate ($\boldsymbol{\Pi}_t$), and an output gate ($\boldsymbol{\Omega}_t$); the structure is shown in **Figure.1A**.

$$\begin{cases} \boldsymbol{\Lambda}_t = f\left(\boldsymbol{\Lambda}_{t-1}, \mathcal{O}_t\right) \\ \boldsymbol{\Theta}_t = simoid\left(\boldsymbol{W}_{\Theta} \times \left[\boldsymbol{\Lambda}_{t-1}, \mathcal{O}_t\right] + \boldsymbol{B}ias_{\Theta}\right) \\ \boldsymbol{\Pi}_t = sigmoid\left(\boldsymbol{W}_{\Pi} \times \left[\boldsymbol{\Lambda}_{t-1}, \mathcal{O}_t\right] + \boldsymbol{B}ias_{\Pi}\right) \\ \boldsymbol{\Omega}_t = sigmoid\left(\boldsymbol{W}_{\Omega} \times \left[\boldsymbol{\Lambda}_{t-1}, \mathcal{O}_t\right] + \boldsymbol{B}ias_{\Omega}\right) \\ \overline{\overline{\boldsymbol{\Phi}}} = \tanh\left(\boldsymbol{W}_{\Phi} \times \left[\boldsymbol{\Lambda}_{t-1}, \mathcal{O}_t\right] + \boldsymbol{B}ias_{\Phi}\right) \\ \boldsymbol{\Phi}_t = \boldsymbol{\Pi}_t \times \boldsymbol{\Phi}_{t-1} + \boldsymbol{\Theta}_t \times \overline{\overline{\boldsymbol{\Phi}}} \\ \boldsymbol{\Lambda}_t = sigmoid\left(\boldsymbol{W}_{\Omega} \times \left[\boldsymbol{\Lambda}_{t-1}, \mathcal{O}_t\right] + \boldsymbol{B}ias_{\Omega}\right) \times \tanh\left(\boldsymbol{\Phi}_t\right) \end{cases} \quad (1)$$

In Eq. (1), $\overline{\overline{W}}$ and $\overline{\overline{B}}ias$ represent the weight and bias of LStm cells, respectively; $\boldsymbol{\Phi}_t$ is the current cell state, $\overline{\overline{\boldsymbol{\Phi}}}$ is the candidate cell state, and $\tanh(\bullet)$ represents a hyperbolic tangent function.

BiLStm [46] is composed of a forward LStm layer and a backward LStm layer. In the forward layer, the sequence $\mathcal{O}_t$ is input into the LStm model to calculate the output state $\overrightarrow{\overline{\overline{\boldsymbol{\Lambda}}}}_{t,i}$. In the backward layer, the inverse form of the input sequence is input into the LStm model to calculate the reverse layer output state ${}_{t,i}\overleftarrow{\overline{\overline{\boldsymbol{\Lambda}}}}$. This structure can extract the forward and backward relations of the wind speed series and connect them to the same output. The network structure is illustrated in **Figure 1B**.

The output of the BiLStm layer at time *t* is $\boldsymbol{\Lambda}_t = \left[\boldsymbol{\Lambda}_{t,1}, \boldsymbol{\Lambda}_{t,2}, \cdots, \boldsymbol{\Lambda}_{t,i}, \cdots, \boldsymbol{\Lambda}_{t,T}\right]^{\mathbf{T}}$, where $\boldsymbol{\Lambda}_{t,i}$ contains $\overrightarrow{\overline{\overline{\boldsymbol{\Lambda}}}}_{t,i}$ and ${}_{t,i}\overleftarrow{\overline{\overline{\boldsymbol{\Lambda}}}}$ which can be expressed as Eq. (2).

$$\begin{cases} \overrightarrow{\overline{\overline{\boldsymbol{\Lambda}}}}_{t,i} = \overrightarrow{\overline{\overline{F}}}_{LStm}\left(\overrightarrow{\overline{\overline{\boldsymbol{\Lambda}}}}_{t,i-1}, \mathcal{O}_t, \overrightarrow{\overline{\overline{\boldsymbol{\Phi}}}}_{t,i-1}\right); & i \in [1, T] \\ {}_{t,i}\overleftarrow{\overline{\overline{\boldsymbol{\Lambda}}}} = {}_{LStm}\overleftarrow{\overline{\overline{B}}}\left({}_{t,i+1}\overleftarrow{\overline{\overline{\boldsymbol{\Lambda}}}}, \mathcal{O}_t, {}_{t,i+1}\overleftarrow{\overline{\overline{\boldsymbol{\Phi}}}}\right); & i \in [T, 1] \\ \boldsymbol{\Lambda}_{t,i} = \left[\overrightarrow{\overline{\overline{\boldsymbol{\Lambda}}}}_{t,i} \oplus {}_{t,i}\overleftarrow{\overline{\overline{\boldsymbol{\Lambda}}}}\right] \end{cases} \tag{2}$$

where $\overrightarrow{\overline{\overline{\boldsymbol{\Phi}}}}_{t,i-1}$ indicates the cell state of the *(i-1)*$^{\text{th}}$ input time step in the forward LStm layer at time *t*; ${}_{t,i+1}\overleftarrow{\overline{\overline{\boldsymbol{\Phi}}}}$ is the cell state of the *(i+1)*$^{\text{th}}$ input time step in the backward LStm layer at time *t*.

2.1.2. Quantile Regression

Quantile regression (Qr) can explore the relationship between the conditional quantiles of the independent and dependent variables. The linear Qr can be expressed as Eq. (3).

$$\boldsymbol{Q}_{Y_t}^{\boldsymbol{linear}}\left(\boldsymbol{\tau} \mid \boldsymbol{X}_t\right) \triangleq \boldsymbol{F}\left(X_t, \overline{\overline{\boldsymbol{\varepsilon}}}(\boldsymbol{\tau})\right) = X_t \overline{\overline{\boldsymbol{\varepsilon}}}(\boldsymbol{\tau}),\ t = 1, 2, \cdots, n \tag{3}$$

where $\boldsymbol{Q}_{Y_t}^{\boldsymbol{linear}}\left(\tau \mid \boldsymbol{X}_t\right)$ is the τ^{th} condition quantile of the dependent variable $\boldsymbol{Y}_t$ and $\boldsymbol{\tau} \in (0,1)$. Regression coefficients $\overline{\overline{\boldsymbol{\varepsilon}}}(\boldsymbol{\tau}) = \left\langle \boldsymbol{\varepsilon_0}(\boldsymbol{\tau}), \left(\boldsymbol{\varepsilon_1}(\boldsymbol{\tau}), \cdots, \boldsymbol{\varepsilon_m}(\boldsymbol{\tau})\right)\right\rangle$.

The estimated value $\overline{\hat{\boldsymbol{\varepsilon}}}(\boldsymbol{\tau})$ of $\overline{\overline{\boldsymbol{\varepsilon}}}(\boldsymbol{\tau})$ can be obtained by minimizing Eq. (4).

$$\overline{\hat{\boldsymbol{\varepsilon}}}(\boldsymbol{\tau}) = \boldsymbol{argmin}\left(\sum_{t=1}^{n} \Phi_{\boldsymbol{\tau}}\left(\boldsymbol{Y}_t - \boldsymbol{X}_t \overline{\overline{\boldsymbol{\varepsilon}}}(\boldsymbol{\tau})\right)\right) \tag{4}$$

where $\Phi_{\tau}(\bullet)$ indicates an asymmetric function that can be written as

$$\Phi_{\boldsymbol{\tau}}\left(\boldsymbol{Y}_t - \boldsymbol{X}_t \overline{\overline{\boldsymbol{\varepsilon}}}(\boldsymbol{\tau})\right) = \begin{cases} \boldsymbol{\tau}\left(\boldsymbol{Y}_t - \boldsymbol{X}_t \overline{\overline{\boldsymbol{\varepsilon}}}(\boldsymbol{\tau})\right), & \boldsymbol{Y}_t - \boldsymbol{X}_t \overline{\overline{\boldsymbol{\varepsilon}}}(\boldsymbol{\tau}) \geq 0 \\ (1 - \boldsymbol{\tau})\left(\boldsymbol{Y}_t - \boldsymbol{X}_t \overline{\overline{\boldsymbol{\varepsilon}}}(\boldsymbol{\tau})\right), & \boldsymbol{Y}_t - \boldsymbol{X}_t \overline{\overline{\boldsymbol{\varepsilon}}}(\boldsymbol{\tau}) < 0 \end{cases} \tag{5}$$

From these equations, the τ^{th} condition quantile of $\boldsymbol{Y}_t$ can be estimated as

$$\boldsymbol{Q}_{Y_t}^{linear}\left(\boldsymbol{\tau} \mid \boldsymbol{X}_t\right) \sim \boldsymbol{X}_t \overline{\hat{\boldsymbol{\varepsilon}}}(\boldsymbol{\tau}) \tag{6}$$

2.1.3. Quantile Regression BiLStm (QrBiLStm)

Based on the BiLStm and Qr, QrBiLStm was used for uncertainty modeling by modifying the cell structure and loss function of BiLStm. The loss function can be modified as $L_{\boldsymbol{QrBiLStm}}^{Pinball-loss} = \sum_{t=1}^{n} \Phi_{\boldsymbol{\tau}}\left(\boldsymbol{Y}_t - \boldsymbol{X}_t \overline{\overline{\boldsymbol{\varepsilon}}}(\boldsymbol{\tau})\right)$. The condition quantile of $\boldsymbol{Y}_t$ obtained by QrBiLStm can be formulated as

$$\boldsymbol{Q}_{\boldsymbol{Y}_t}^{BiLStm}\left(\tau \mid \boldsymbol{X}_t\right) \triangleq f\left(\boldsymbol{X}_t, \varepsilon(\tau)\right) = \sigma\left(\boldsymbol{W}_{\boldsymbol{\Omega}}(\tau) \times \boldsymbol{\Lambda}_t(\tau)\right) \tag{7}$$

where $\boldsymbol{W}_{\boldsymbol{\Omega}}(\tau)$ indicates the weight matrix of τ, and $\boldsymbol{\Lambda}_t(\boldsymbol{\tau}) = \left[\overrightarrow{\overline{\overline{\boldsymbol{\Lambda}}}}_{t,i}(\boldsymbol{\tau}) \oplus {}_{t,i}\overleftarrow{\overline{\overline{\boldsymbol{\Lambda}}}}(\boldsymbol{\tau})\right]$.

The novel QrBiLStm network combines quantile regression with bi-directional data processing, and can effectively learn the hidden correlation between the pre- and post-time-step data in a time series, with better uncertainty modeling.

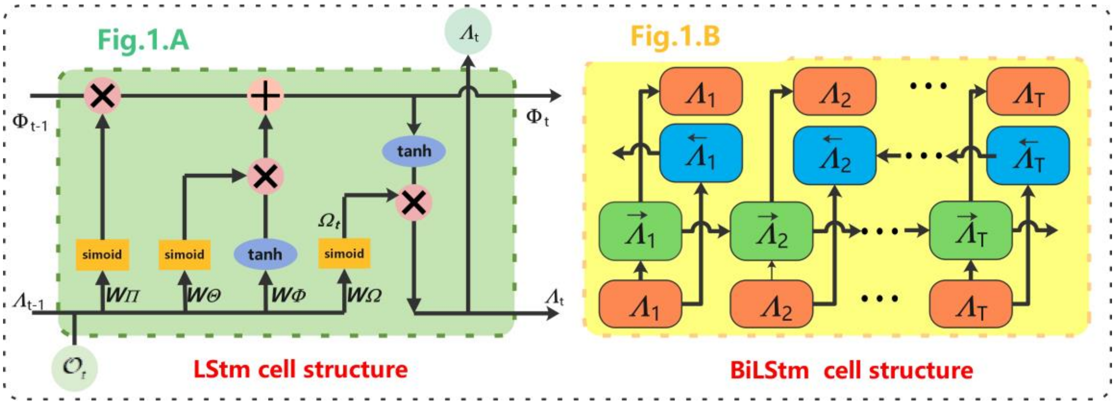


Fig.1. Cell structures of LStm and BiLStm

2.2. Original signal preprocessing

In this section, the principles of SSa and PSr are introduced in decomposing and reconstructing sequences.

2.2.1. Singular Spectral analysis (SSa)

The principal objective of SSa is to decompose the original series into a sum of series, identified as either a trend, a periodic or quasi-periodic component, or noise [47]. The flow of SSa can be summarized as follows.

(A). Embedding procedure

Based on the original sequence $\overline{\overline{\mho}}_{ini}^{T}$ and Karhunene–Loeve decomposition of the covariance matrix, the sequence $\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}^{i} = \left[\mho_{i-1}, \cdots, \mho_{i+Y-2}\right]^{\boldsymbol{T}}$ of the L-dimensional vector is constructed.

$$\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}} = \begin{bmatrix} \mho_0 & \cdots & \mho_M \\ \vdots & \ddots & \vdots \\ \mho_{Y-1} & \cdots & \mho_{T-1} \end{bmatrix} \tag{8}$$

where, $M = T - Y + 1$ and $\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}$ is a Hankel matrix with equal elements on the diagonals.

(B) Singular value disintegration

The matrix $\overline{\overline{\boldsymbol{SS}}}^{T}$ is calculated to determine its eigenvalues using triples $\left(x_i, \boldsymbol{E}_i, \boldsymbol{F}_i\right)$ by SVD [51]**.** The eigenvalues of $\overline{\overline{\boldsymbol{SS}}}^{T}$ are defined as ζ_i, $i = 1, 2, \ldots, Y$ in descending order. $\boldsymbol{E}_i$ and $\boldsymbol{F}_i$ are the i[th] left and right eigenvectors, respectively, of $\overline{\overline{\boldsymbol{SS}}}^{T}$. Assuming $r = rank\left(\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}^{i}\right)$, the trajectory matrix $\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}$ can be expressed as

$$\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}} = \overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}^{1} + \cdots + \overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}^{r}, \quad \overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}^{i} = x_i \boldsymbol{E}_i \boldsymbol{F}_i^{T} \tag{9}$$

where x_i is the singular value of $\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}$, and $\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}^{i}\left(i = 1, 2, \ldots, r\right)$ are matrices of $rank = 1$.

(C). Reconstruction

Step 3.1 (Grouping): The indices $K = 1, 2, \ldots, r$ are grouped into V disjoint subsets $\{G_1^k, G_2^k, \cdots, G_V^k\}$ corresponding to splitting the elementary matrices $\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}^{i}$ $(i = 1, 2, \ldots, r)$ into V groups. Each group contains a set of indices as $\overline{\overline{\boldsymbol{G}^k}} = \{d_1, \cdots, d_p\}$. The resultant matrix is defined as $\boldsymbol{S}_{\overline{\overline{\boldsymbol{G}}}^k} = \boldsymbol{S}_{d,1} + \boldsymbol{S}_{d,2} + \cdots + \boldsymbol{S}_{d,p}$. Thus, $\overline{\overline{\boldsymbol{S}_{\boldsymbol{H}}}} = \boldsymbol{S}_{\overline{\overline{\boldsymbol{G}}}^k,1} + \boldsymbol{S}_{\overline{\overline{\boldsymbol{G}}}^k,2} + \cdots \boldsymbol{S}_{\overline{\overline{\boldsymbol{G}}}^k,V}$, where $\overline{\overline{\boldsymbol{S}_{\boldsymbol{H}}}}$ is the sum of $\overline{\overline{\boldsymbol{G}}}^k$ resultant matrices.

Step 3.2 (Diagonal averaging): Each matrix $\boldsymbol{S}_{\overline{\overline{\boldsymbol{G}}}^k,j}$ $j = 1, 2, \ldots, V$ is transferred into a time series. Let $\overline{\overline{\boldsymbol{S}}}_{\boldsymbol{H}}$ be a $(Y \times M)$ matrix with elements $s_{\boldsymbol{i,j}}$, with $Y^* = \boldsymbol{min}(Y, M)$ and $M^* = \boldsymbol{max}(Y, M)$. Define $\boldsymbol{S}_{\boldsymbol{i,j}}\boldsymbol{*} = \begin{cases} \boldsymbol{S}_{\boldsymbol{i,j}}, & Y < M \\ \boldsymbol{S}_{\boldsymbol{j,i}}, & otherwise \end{cases}$.

Matrix $\overline{\overline{\boldsymbol{S}_{\boldsymbol{H}}}}$ is transformed into sequence $\Lambda_0, \Lambda_1, \cdots, \Lambda_{T-1}$ using Eq. (10).

$$\Lambda_i = \begin{cases} (b+1)^{-1} \cdot \sum_{c=1}^{b+1} \boldsymbol{S}^*_{\boldsymbol{c,b-c+2}}, & \boldsymbol{b} \in [0, Y^* - 1) \\ Y^{*-1} \cdot \sum_{c=1}^{L^*} \boldsymbol{S}^*_{\boldsymbol{c,b-c+2}}, & \boldsymbol{b} \in [Y^* - 1, M^*) \\ (T-b)^{-1} \cdot \sum_{c=b-b^*+2}^{N-b+1} \boldsymbol{S}^*_{\boldsymbol{c,b-c+2}}, & \boldsymbol{b} \in [M^* - 1, T) \end{cases} \quad (10)$$

The averaging of the elements along the diagonal $i + j = b + 2$, applied to a resultant matrix $\boldsymbol{S}_{\overline{\overline{\boldsymbol{G}}}^k,j}$, produces a time series $\overline{\overline{\mho^T}}$ of length *T*. Thus, the original series $\overline{\overline{\mho^T_{\boldsymbol{ini}}}}$ is decomposed into the sum of *V* sequences. Defining the decomposed series as $\overline{\overline{\mho^T_{\boldsymbol{de}}}}$, it can be expressed as $\overline{\overline{\mho^T_{\boldsymbol{de}}}} = \overline{\overline{\mho^T_1}} + \cdots + \overline{\overline{\mho^T_V}}$

2.2.2. Phase Space reconstruction (PSr)

In the prediction of chaotic time series, the phase space reconstruction (PSr) method can be used to reconstruct a one-dimensional series into a dynamic chaotic space to obtain better forecasting results [48]. In this study, the C–C method was used to determine two important parameters of the PSr algorithm: delay time $\boldsymbol{\omega}$ and embedding dimension $\boldsymbol{\theta}$. The PSr process is expressed as

$$\overline{\overline{\boldsymbol{\Phi}}}_P^{ini} = \left[\overline{\overline{\boldsymbol{\Phi}_1}}, \overline{\overline{\boldsymbol{\Phi}_2}}, \cdots, \overline{\overline{\boldsymbol{\Phi}_P}}\right]^T = \begin{bmatrix} \xi_1 & \xi_{1+\tau} & \cdots & \xi_{1+(\theta-1)\omega} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_i & \xi_{i+\tau} & \cdots & \xi_{i+(\theta-1)\omega} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_P & \xi_{P+\tau} & \cdots & \xi_{P+(\theta-1)\omega} \end{bmatrix} \quad (11)$$

where $\{\xi_i \mid i = 1, 2, \cdots, Z\}$ signifies the samples of the sequence; Z indicates the length of the initial sequences, and $P = Z - (\theta - 1) \cdot \omega$. Accordingly, the target matrix $\overline{\overline{\boldsymbol{R}}}_P^{tar}$ corresponding to $\overline{\overline{\boldsymbol{\Phi}}}_P^{ini}$ can be expressed as Eq. (12).

$$\overline{\overline{\boldsymbol{R}}}_P^{tar} = \left[\overline{\overline{\boldsymbol{R}}}_1, \overline{\overline{\boldsymbol{R}}}_1, \cdots, \overline{\overline{\boldsymbol{R}}}_P \right]^{\boldsymbol{T}} = \left[\xi_{1+(\theta-1)\lambda}, \xi_{2+(\theta-1)\lambda}, \cdots, \xi_Z \right]^{\boldsymbol{T}} \tag{12}$$

2.3. Improved Tunicate Swarm Optimization Algorithm (IMOTa)

This section illustrates the mechanism of the original optimizer, the multi-objective optimization, and three improved optimization strategies.

A. Tunicate swarm algorithm (TSa)

The TSa was proposed by Kaur et al. [49], who regarded the optimal solution as the food source in the ocean, and the process of finding the optimal solution as the movement behavior combination of the capsule animals looking for food. The comprehensive mathematical principle of the TSa is presented as following.

Behavior 1. (avoidance) This behavior of tunicates aims to avoid collisions between individuals, and is defined by $\overline{\overline{\boldsymbol{A}}}vo = \overline{\overline{\boldsymbol{G}}}raf \big/ \overline{\overline{\boldsymbol{S}}}ocf$. $\overline{\overline{\boldsymbol{A}}}vo$ is driven mainly by gravity and social forces. The gravity force is counteracted by the water flow $\overline{\overline{\boldsymbol{W}}}atf = 2 \cdot \vec{\mathrm{R}}(1)$, and is defined as $\overline{\overline{\boldsymbol{G}}}raf = \vec{\mathrm{R}}(1) + \vec{\mathrm{R}}(2) - \overline{\overline{\boldsymbol{W}}}atf$. The social force is driven mainly by initial speed $IniS$ and subordinate speed $SubS$. The social force is defined as $\overline{\overline{\boldsymbol{S}}}ocf = \boldsymbol{Inis} + \vec{\mathrm{R}}(1) \cdot (\boldsymbol{Subs} - \boldsymbol{Inis})$. In the definition, $IniS$ is preset as 1, $SubS$ is preset as 4, and $\vec{\boldsymbol{R}}$ is a matrix with elements that are all random values ranging from $[0,1]$.

Behavior 2. (movement) Tunicates move in the direction of their best neighbors. This behavior can be mathematically defined as $\overline{\overline{\boldsymbol{D}}}is = \left| \overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_{\boldsymbol{f}} - rand \cdot \overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_t(k) \right|$. $\overline{\overline{\boldsymbol{Dis}}}$ measures the absolute distance between the optimal solution and the agent. In this behavior, $\overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_{\boldsymbol{f}}$ indicates the position of the optimal solution, and $\overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_t(k)$ refers to the position of the k^{th} individual.

Behavior 3. (convergence) Tunicates begin to advance toward food sources by means of $\overline{\overline{\boldsymbol{A}}}\boldsymbol{vo}$ and $\overline{\overline{\boldsymbol{D}}}\boldsymbol{is}$. The tunicates update their positions according to Eq. (13).

$$\overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_u(k) = \begin{cases} \overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_{\boldsymbol{f}} + \overline{\overline{\boldsymbol{A}}}\boldsymbol{vo} \cdot \overline{\overline{\boldsymbol{D}}}\boldsymbol{is}, & rand \geq 0.5 \\ \overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_{\boldsymbol{f}} - \overline{\overline{\boldsymbol{A}}}\boldsymbol{vo} \cdot \overline{\overline{\boldsymbol{D}}}\boldsymbol{is}, & rand < 0.5 \end{cases} \tag{13}$$

Behavior 4. (swarm behavior) The best two solutions are retained, and the positions of other individuals relative to the food source are updated. The swarm behavior can be mathematically expressed as Eq. (14).

$$\overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_t(k+1) = \left[\overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_t(k) + \overline{\overline{\boldsymbol{P}}}\boldsymbol{os}_u(k) \right] \Big/ \left(2 + \vec{\boldsymbol{R}}(1) \right) \tag{14}$$

B. Improved multi-objective tunicate swarm algorithm (IMOTa)

This study developed three improvement strategies: the multi-objective approach

(MOJ), the elite opposition learning approach (EOLA), and the exponential function step approach (EFSA). The MOJ produces multiple objective functions in the optimization algorithm to achieve better optimization results. The EOLA can improve the convergence speed of the algorithm. The EFSA can improve the global optimization and robustness.

a. Multi-objective tunicate swarm algorithm (MOTa)

To achieve multi-objective optimization, this section introduces the dominant strategy, the Pareto optimal solution and archiving with a roulette wheel. The ability of the MOTa system to find the Pareto optimal solution is demonstrated using the definitions.

Definition 1. Let $\bar{\bar{\boldsymbol{J}}}=\left(J_1, J_2, \cdots, J_i\right)$ and $\bar{\bar{\boldsymbol{K}}}=\left(K_1, K_2, \cdots, K_i\right)$ be two vectors; $\bar{\bar{\boldsymbol{J}}}$ strictly dominates $\bar{\bar{\boldsymbol{K}}}$, if $\forall \boldsymbol{n} \in\{1,2,\ldots,N\}$, $\boldsymbol{f}_n\left(\bar{\bar{\boldsymbol{J}}}\right) \geq \boldsymbol{f}_n\left(\bar{\bar{\boldsymbol{K}}}\right)$; $\bar{\bar{\boldsymbol{J}}}$ partially dominates $\bar{\bar{\boldsymbol{K}}}$, if $\exists \boldsymbol{n} \in\{1,2,\ldots,N\}$, $\boldsymbol{f}_n\left(\bar{\bar{\boldsymbol{J}}}\right)>\boldsymbol{f}_n\left(\bar{\bar{\boldsymbol{K}}}\right)$. $\bar{\bar{\boldsymbol{J}}}$ dominates $\bar{\bar{\boldsymbol{K}}}$, if

$$\left[\forall \boldsymbol{n} \in\{1,2,\ldots,N\}, \boldsymbol{f}_n\left(\bar{\bar{\boldsymbol{J}}}\right) \geq \boldsymbol{f}_n\left(\bar{\bar{\boldsymbol{K}}}\right)\right] \wedge\left[\exists \boldsymbol{n} \in\{1,2,\ldots,N\}, \boldsymbol{f}_n\left(\bar{\bar{\boldsymbol{J}}}\right)>\boldsymbol{f}_n\left(\bar{\bar{\boldsymbol{K}}}\right)\right] \quad (15)$$

where $f_{\boldsymbol{n}}(\bullet)$ indicates the n-th objective function and N is the number of functions.

Definition 2. If $\forall n \in\{1,2,\ldots,N\}: \not\exists K \in \bar{\bar{\boldsymbol{K}}} \mid \boldsymbol{f}_n(\boldsymbol{K}) \succsim \boldsymbol{f}_n(\boldsymbol{J})$, that is, none of the obtained solutions dominates $\bar{\bar{\boldsymbol{J}}}$, then $\bar{\bar{\boldsymbol{J}}}$ is the Pareto optimal solution.

Definition 3. Archiving with a roulette wheel is a matrix used to store the optimal solutions. When the archive is full, the individuals with the most adjacent solutions are eliminated by the roulette wheel. The probability that an individual is eliminated is $Pe_i = Ns_i / cq, cq>1$, where Ns_i indicates the number of adjacent solutions, and cq is a constant.

Suppose that the fitness function corresponding to the objective function is $\boldsymbol{fit}(\cdot)$, and the optimal position $\boldsymbol{P}^*$ of the individual in MOTA is the weight of two QrbiLStm units, $\boldsymbol{We}\left(\boldsymbol{P}^*\right)$. It is proved that $\boldsymbol{P}^*$ is the optimal weight of two QrbiLStm units through reduction to absurdity.

Proof

If there exists at least one adjacent position $\boldsymbol{Q}^*=\boldsymbol{P}^*+\theta$, the weights satisfy $\left[\forall n \in\{1,2,\ldots,N\}, fit_{\boldsymbol{n}}\left(\boldsymbol{We}\left(\boldsymbol{Q}^*\right)\right) \geq fit_{\boldsymbol{n}}\left(\boldsymbol{We}\left(\boldsymbol{P}^*\right)\right)\right] \wedge\left[\exists n \in\{1,2,\ldots,N\}, fit_{\boldsymbol{n}}\left(\boldsymbol{We}\left(\boldsymbol{Q}^*\right)\right)>fit_{\boldsymbol{n}}\left(\boldsymbol{We}\left(\boldsymbol{P}^*\right)\right)\right]$, and $\boldsymbol{Q}^*$ is stored in the Archive. As $\boldsymbol{We}\left(\boldsymbol{Q}^*\right)$ dominates $\boldsymbol{We}\left(\boldsymbol{P}^*\right)$, and the capacity of Archive with Roulette-Wheel is limited, $\boldsymbol{P}^*$ is deleted from Archive with the $prob = Ns_i / cq$ or ranked behind $\boldsymbol{Q}^*$. The position with the highest fitness value in Archive is selected as the optimal position. Corresponding, the optimal weights of QrBiLStm is $\boldsymbol{We}\left(\boldsymbol{Q}^*\right)$ instead of $\boldsymbol{We}\left(\boldsymbol{P}^*\right)$. ■

b. Elite opposition learning (EOLA)-MOTa

The IMOTa based on EOLA was proposed to improve the convergence performance of the optimizer. The principle of EOLA is to calculate and evaluate the opposition solution of a feasible solution, and select the better solution as the next generation. In this study, the elite tunicate is defined as the individual that obtains the highest fitness value.

Definition 4. (opposition point) Let $\vec{X_j} = (x_{j,1}, x_{j,2}, \cdots, x_{j,d})$ be a point in *d*-dimensional space (regarded as a feasible solution), $x_i \in [lb_i, ub_i]$, and its corresponding opposition point $\vec{X_j} = (\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_d)$ are defined in Eq. (16)

$$\tilde{x}_i = lb_i + ub_i - x_i \tag{16}$$

Definition 5. (elite opposition solution) Suppose that $\overline{\overline{X_j}} = (x_{j,1}, x_{j,2}, \cdots, x_{j,d})$ is a common tunicate, and the corresponding extreme value of itself is the elite tunicate $\overline{\overline{X_j^{elite}}} = (x_{j,1}{}^e, x_{j,2}{}^e, \cdots, x_{j,d}{}^e)$. The elite opposition solution $X_j^{elite} = (\tilde{x}_{j,1}{}^e, \tilde{x}_{j,2}{}^e, \cdots, \tilde{x}_{j,d}{}^e)$ can be defined as formula (17).

$$\tilde{x}_{j,i}{}^e = \varpi \cdot (dlb_j + dub_j) - x_{j,i}{}^e \tag{17}$$

where $\tilde{x}_{j,i}{}^e \in [dlb_j, dub_j]$; $c \in U(0,1)$; $[dlb_j, dub_j]$ is the dynamic boundary of the i^{th} dimension search space, which can be calculated according to Eq. (18).

$$dlb_j = min(x_{j,i}), \ dub_j = max(x_{j,i}) \tag{18}$$

Replacing the fixed boundary with the dynamic boundary of the search space is conducive to preserving the search experience, such that the generated opposition solution can be located in the gradually reduced search space. However, it has the possibility of causing $x_{j,i}{}^e$ to exit $[dlb_j, dub_j]$. If $x_{j,i}{}^e < dlb_j$ or $x_{j,i}{}^e > dub_j$, then $\tilde{x}_{j,i}{}^e = \varpi \cdot (dub_j - dlb_j) + dub_j$, where ϖ is a random value between 0 and 1.

c. Exponential function steps (EFSA)-MOTa

In **Behaviors** 3 and 4 of the original TSA, the approach to promote the location update is random linear. This updating approach cannot guarantee individuals to find the optimal solution, which ultimately leads to poor optimization and robustness. Thus, improving the piecewise linear random step using EFSA is proposed. The new location update strategy can be mathematically expressed as Eq. (19).

$$\overline{\overline{Pos_u}}(k) = \overline{\overline{Pos_f}} + (rand - 0.5) \cdot 2^{rand} \cdot \overline{\overline{Avo}} \cdot \overline{\overline{Dis}} \tag{19}$$

C. Design of the multi-objective optimization function of EPFS

To simultaneously optimize the reliability and interval width of the forecasting system, two pseudo-interval indicators were designed. The purpose of constructing pseudo-intervals is to optimize the upper and lower bounds of the intervals, respectively, to achieve better optimization results. Based on the two pseudo-interval indicators that measure reliability and resolution, the objective functions for multi-objective optimization are developed.

a. Pseudo-interval indicators

The pseudo-interval is a half-interval composed of the observed values and the upper or lower bound of the interval. Thus, the indicators for evaluating the reliability and resolution of the pseudo-interval can be designed as $\boldsymbol{PICP}^{half}(\alpha)$, and $\boldsymbol{PINAW}^{half}(\alpha)$.

Two indicators for evaluating the upper pseudo-interval can be defined as

$$\boldsymbol{PICP}_{upper}^{half}(\alpha)=\frac{1}{M}\sum_{i=1}^{M}\boldsymbol{\Gamma}_{i}^{half}, \quad \boldsymbol{\Gamma}_{i}^{half}=\begin{cases}1, & \boldsymbol{UB}_{i}^{half}(\alpha)\geq Obse\boldsymbol{V}_{i}\\ 0, & \boldsymbol{UB}_{i}^{half}(\alpha)< Obse\boldsymbol{V}_{i}\end{cases} \tag{20}$$

$$\boldsymbol{PINAW}_{upper}^{half}(\alpha)=\frac{1}{MR}\sum_{i=1}^{M}\left[\boldsymbol{UB}_{i}^{half}(\alpha)-Obse\boldsymbol{V}_{i}\right] \tag{21}$$

where $\boldsymbol{UB}_{i}^{half}(\alpha)$ is the upper bound of the forecasting interval corresponding to α. $ObseV_i$ is the observation value, M is the number of observation values, and R is the range of observation values.

Accordingly, the two indicators for evaluating the lower pseudo interval can be defined as:

$$\boldsymbol{PICP}_{lower}^{half}(\alpha)=\frac{1}{M}\sum_{i=1}^{M}\boldsymbol{\Gamma'}_{i}^{half}, \quad \boldsymbol{\Gamma'}_{i}^{half}=\begin{cases}1, & \boldsymbol{LB}_{i}^{half}(\alpha)\leq Obse\boldsymbol{V}_{i}\\ 0, & \boldsymbol{LB}_{i}^{half}(\alpha)> Obse\boldsymbol{V}_{i}\end{cases} \tag{22}$$

$$\boldsymbol{PINAW}_{lower}^{half}(\alpha)=\frac{1}{MR}\sum_{i=1}^{M}\left[Obse\boldsymbol{V}_{i}-\boldsymbol{LB}_{i}^{half}(\alpha)\right] \tag{23}$$

where $\boldsymbol{LB}_{i}^{half}(\alpha)$ is the lower bound of the forecasting interval corresponding to α.

b. Multi-objective optimization function

The objective functions for multi-objective optimization can be determined as:

$$\boldsymbol{min}\begin{cases}\boldsymbol{Of}_{1}=(1-\alpha/2)-\boldsymbol{PICP}^{half}(\alpha)\\ \boldsymbol{Of}_{2}=\boldsymbol{PINAW}^{half}(\alpha)\cdot\left[1+\exp\left(-\Phi\cdot\left(\boldsymbol{PICP}^{half}(\alpha)-1+\alpha\right)\right)\right]\end{cases} \tag{24}$$

where $\Phi>0$ is the penalty coefficient. A larger Φ indicates a greater degree of punishment.

3. Framework of the proposed Ensemble probabilistic forecasting system

This section presents the description of the material analyzed (section 3.1) and the ensemble probabilistic forecasting system applied in this study (section 3.2).

3.2. Dataset description

The three experimental datasets were collected from Shandong Peninsula with an interval of 10 minutes. In each dataset, extract 2880 points as the experimental sequence and select 75% of the total length as training set 1, with a length of 2160. The remaining 720 points are divided into training set 2 and testing set. Training set 2 accounts for 75% of the remaining length, with a length of 540 and a testing set of 180.

3.3. Flow of the proposed Ensemble probabilistic forecasting system

In accordance with the aforementioned data processing approaches and forecasting models, the proposed ensemble probabilistic forecasting system includes SSa decomposition and reconstruction, phase space reconstruction of the C–C method, principal forecasts based on two QrBiLStm units, and construction of pseudo-intervals to optimize the upper and lower bounds. The process is presented; the complete system structure and procedure are shown in **Fig. 2**.

Step 1: Segment the original wind speed sequence into two training sets and a testing set. A total of 2,880 data were collected. Training set 1 included 2,160 points, Training set 2 included 540 points, and Test set included 180 points.

Step 2: Use SSa to decompose and reconstruct the wind speed sequence $\left\{ X'_j \mid j = 1, 2, \cdots n \right\}$ to obtain the denoised wind speed sequence $\left\{ X''_j \mid j = 1, 2, \cdots n \right\}$.

Step 3: Use the C–C method to find the optimal parameter values of PSr and reconstruct the sequence to adapt to the chaotic system.

Step 4: Implement uncertainty forecasting quantization on Training Set 1, based on two QrBiLStm network units.

Step5: Construct the pseudo-intervals based on the forecasting values and observation values of Training Set 2 obtained in Step 4.

Step 6: Use the IMOTa and the designed interval optimization objective function; the pseudo-interval is input for optimization, and the final probabilistic forecasting results on the testing set are obtained.

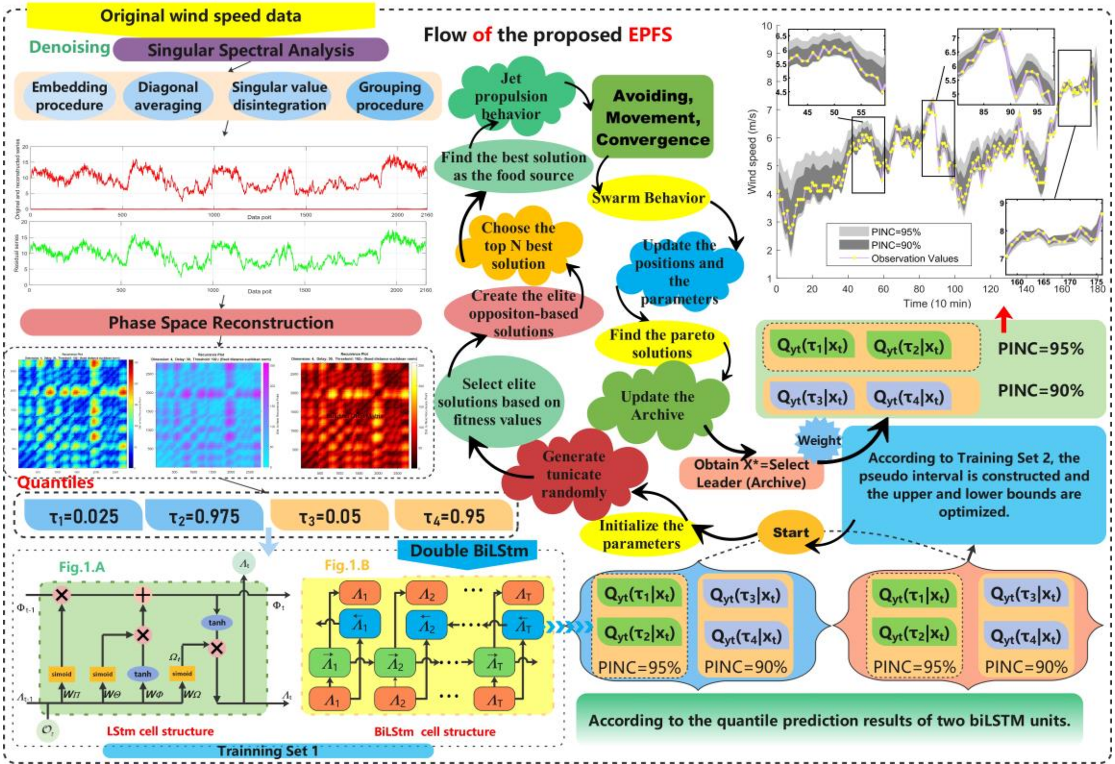


Fig.2. Flow of the proposed EPFS.

4. Experimental results and discussion

This section describes the evaluation metrics of uncertainty modeling. Three comparative experiments and their corresponding analyses are described to verify the forecasting effectiveness of the proposed EPFS. The whole experiment is implemented on the personal computer with AMD Ryzen 5 5600H six-core processor with Radeon Graphics 3.30 GHz, 16 GB of RAM and a single NVIDIA GeForce GTX 1650 of GPU. The proposed EPFS and comparative models were implemented on Matlab2020a.

The parameter settings for all models designed for the experiments are presented in Table A1 in Appendix A.

4.1. Evaluation metrics

The wind speed prediction interval was evaluated based on reliability, resolution, and sharpness.

Reliability

Reliability is based on the significance level α to evaluate the coverage probability of the forecasting interval. In this study, the reliability was characterized by the PICP metric which is expressed as Eq. (25).

$$\boldsymbol{PICP}(\alpha)=\frac{1}{N}\sum_{i=1}^{N}\Phi_i\ ,\ \ \Phi_i=\begin{cases}1, & Obse\boldsymbol{V}_i\in\left[\boldsymbol{LB}_i(\alpha),\boldsymbol{UB}_i(\alpha)\right]\\ 0, & Obse\boldsymbol{V}_i\notin\left[\boldsymbol{LB}_i(\alpha),\boldsymbol{UB}_i(\alpha)\right]\end{cases} \tag{25}$$

where N is the length of the testing set, α is the confidence level, and $Obse\boldsymbol{V}_i$ indicates the observation value. $\boldsymbol{LB}_i(\alpha)$ and $\boldsymbol{UB}_i(\alpha)$ are the lower and upper bounds of the prediction interval, respectively, corresponding to α. In this paper, interval prediction is implemented based on $\alpha=0.05$ and $\alpha=0.1$ confidence levels. And $PINC=(1-\alpha)\times 100\%$.

Resolution

Measurement of interval resolution (interval width) [50] is important for effective interval prediction. A prediction interval that is too broad contains a small amount of valuable information, which is less practical [51]. Thus, PINAW largely reflects the information contained in the forecast interval and can be represented by Eq. (26). R is the range of observation values, it is determined by the maximum of the observation values on the testing set minus its minimum. And N is the length of the testing set.

$$\boldsymbol{PINAW}(\alpha)=\frac{1}{NR}\sum_{i=1}^{N}\left[\boldsymbol{UB}_i(\alpha)-\boldsymbol{LB}_i(\alpha)\right] \tag{26}$$

Sharpness

Both PINAW and PICP are one-sided in evaluating the quality of forecast intervals. The sharpness combines the two metrics for assessment of prediction intervals [54]; the

AIS metric [19] meets the requirements, and is expressed as Eq. (27):

$$\boldsymbol{AIS}(\alpha) = \sum_{i=1}^{N} \boldsymbol{S}_{\rightarrow}^{(\alpha)} \quad (27)$$

$$\boldsymbol{S}_{\rightarrow}^{(\alpha)} = \begin{cases} -2\alpha\boldsymbol{\xi}_i(\alpha) - 4\left(\boldsymbol{LB}_i(\alpha) - Obse\boldsymbol{V}\right), & if\ Obse\boldsymbol{V}_i < \boldsymbol{LB}_i(\alpha) \\ -2\alpha\boldsymbol{\xi}_i(\alpha), & if\ Obse\boldsymbol{V}_i \in \left[\boldsymbol{LB}_i(\alpha), \boldsymbol{UB}_i(\alpha)\right] \\ -2\alpha\boldsymbol{\xi}_i(\alpha) - 4\left(Obse\boldsymbol{V}_i - \boldsymbol{UB}_i(\alpha)\right), & if\ Obse\boldsymbol{V}_i > \boldsymbol{UB}_i(\alpha) \end{cases} \quad (28)$$

where $\xi_i(\alpha) = \boldsymbol{UB}_i(\alpha) - \boldsymbol{LB}_i(\alpha)$.

4.2. Comparative experiments

This section presents a comparison of the proposed EPFS with commonly used single and ensemble probabilistic forecasting models (EPFMs). Probabilistic forecasting single models include the QrLASso, QrLStm, Qr convolution neural network (QrCNn), QrGRu, Gaussian process regression (GPr), the Bayesian regression model (BLgm), and the proposed QrBiLStm. The EPFMs include models with different denoising methods and optimizers. The interval forecasting chart of the proposed EPFS and the metric values of the other single models is shown in **Fig. 3**.

The compared denoising methods include empirical mode decomposition (Emd) [56], ensemble Emd (Eemd), complete Eemd with adaptive noise (CeemdAN) [57], and wavelet transform (Wt). The compared optimizers include MOTa, MO dragonfly algorithm (MODa), MO grasshopper algorithm (MOGa), and MO antlion algorithm (MOAa).

4.2.1. Comparative Experimental Analysis with Single Models

Reliability

PICP shows the reliability of intervals; when the PICP is higher than the prediction interval nominal confidence (PINC), the forecasting interval is considered to be reliable. As shown in **Table 2**, all single models and EPFS except QrCNN are valid for all Sites.

The proposed QrBiLStm obtains $PICP = 1$ for both $PINC = 90\%$ and $PINC = 95\%$ for all three sites, indicating that it has better reliability for interval prediction. The EPFS optimized based on the two QrBiLStm benchmark models also has high PICP values, and is also reliable. For Site 1, when $PINC = 90\%$, EPFS obtains $PICP = 0.9833$; when $PINC = 95\%$, EPFS obtains $PICP = 0.9944$.

Resolution

The PINAW metric indicates the interval width which determines the practicality and informative of interval [55]. A smaller PINAW indicates a narrower interval width with more uncertainty information. Not considering the invalid model ($\boldsymbol{PICP}(\alpha) < \boldsymbol{PINC}(\alpha)$), the interval width of QrBiLStm is the narrowest of all single models. This is reflected in the PINAW values. For $PINC = 95\%$, QrBiLStm is obtained at three sites: $PINAW_{\boldsymbol{PINC=95\%}}^{Site1} = 0.2674$, $PINAW_{\boldsymbol{PINC=95\%}}^{Site2} = 0.2471$, and $PINAW_{\boldsymbol{PINC=95\%}}^{Site3} = 0.3306$. Thus, compared with other single models, the proposed QrBiLStm has the narrowest interval width and is more effective in interval forecasting than the model based on distribution hypothesis and other single models not based on distribution.

Compared with the single models, EPFS is greatly optimized in terms of the interval width. The experimental results show that QrBiLStm has the narrowest interval width. The interval width forecasted by EPFS was $27.1387\% \sim 56.6055\%$ smaller than that obtained by QrBiLStm, indicating that the forecasting of EPFS is greatly improved compared with that of a single model.

Sharpness

The AIS metric simultaneously considers coverage and interval width, punishing an interval that does not contain observation values and interval width. The AIS value is generally less than 0; a higher AIS value indicates a more effective forecasting interval. The AIS value of the proposed EPFS is the lowest at all sites and all PINCs, indicating that the proposed EPFS can provide more uncertainty information.

Compared with single models, EPFS is greatly optimized in interval width; its coverage rate is also high, resulting in the best interval prediction. For $PINC = 95\%$, the AIS values obtained by EPFS at the three sites are $AIS_{PINC=95\%}^{Site1} = -0.1273$, $AIS_{PINC=95\%}^{Site2} = -0.1216$, and $AIS_{PINC=95\%}^{Site3} = -0.1324$.

Remark: Compared with other single models, the proposed QrBiLStm can obtain the highest interval coverage and the narrowest interval width. Thus, the proposed QrBiLStm is more reliable and effective than the model based on the distribution hypothesis and other Qr–deep learning models not based on distribution. The proposed EPFS optimized using two QrBiLStm units can further reduce the interval width while ensuring high interval coverage. The proposed EPFS can provide more uncertainty information.

Table 2

Interval prediction metric values of single models with PINC=90% and PINC=95%.

Dataset	**Models**	**PICP**		**PINAW**		**AIS**	
		PINC=90%	**PINC=95%**	**PINC=90%**	**PINC=95%**	**PINC=90%**	**PINC=95%**
Site 1	**SSa-PSr-QrLASso**	0.9556	0.9778	0.4554	0.5341	-0.5439	-0.3055
	SSa-PSr-QrLStm	**1.0000**	**1.0000**	0.5160	0.5835	-0.5152	-0.2928
	SSa-PSr-QrGRu	**1.0000**	**1.0000**	0.5337	0.5986	-0.5334	-0.2964
	SSa-PSr-QrCNN	0.9611	0.8944	0.3147	0.7572	-0.3726	-0.4743
	SSa-PSr-QrBiLStm	1.0000	**1.0000**	0.3489	0.4249	-0.3500	-0.2086
	SSa-PSr-GPr	0.9944	**1.0000**	0.4457	0.5311	-0.4557	-0.2712
	SSa-PSr-BLgm	**1.0000**	**1.0000**	0.4460	0.5317	-0.4556	-0.2716
	Proposed EPFS	0.9833	0.9944	**0.1893**	**0.2674**	**-0.1805**	**-0.1273**
Site 2	**SSa-PSr-QrLASso**	0.8611	0.9111	0.3028	0.4918	-0.4409	-0.3389
	SSa-PSr-QrLStm	0.9778	0.9833	0.4770	0.5417	-0.5063	-0.2884
	SSa-PSr-QrGRu	0.9778	0.9889	0.4710	0.5729	-0.4905	-0.2930
	SSa-PSr-QrCNN	0.9833	**1.0000**	0.3734	1.1707	-0.3954	-0.5868
	SSa-PSr-QrBiLStm	**1.0000**	**1.0000**	0.3734	0.3494	-0.3802	-0.1748
	SSa-PSr-GPr	**1.0000**	**1.0000**	0.4547	0.5419	-0.4685	-0.2791
	SSa-PSr-BLgm	**1.0000**	**1.0000**	0.4555	0.5428	-0.4695	-0.2797
	Proposed EPFS	0.9778	0.9889	**0.1620**	**0.2471**	**-0.1590**	**-0.1216**
Site 3	**SSa-PSr-QrLASso**	0.8944	0.9889	0.4122	0.6835	-0.4157	-0.3049
	SSa-PSr-QrLStm	0.9889	**1.0000**	0.5568	0.7498	-0.4633	-0.3110
	SSa-PSr-QrGRu	0.9833	**1.0000**	0.5502	0.7138	-0.4539	-0.2933
	SSa-PSr-QrCNN	0.9111	0.9222	0.4586	0.4362	-0.4508	-0.2351
	SSa-PSr-QrBiLStm	**1.0000**	**1.0000**	0.3384	0.4537	-0.2771	-0.1835
	SSa-PSr-GPr	0.9889	**1.0000**	0.5029	0.5992	-0.4248	-0.2519
	SSa-PSr-BLgm	0.9944	**1.0000**	0.5080	0.6047	-0.4286	-0.2544
	Proposed EPFS	0.9833	0.9889	**0.2180**	**0.3306**	**-0.1784**	**-0.1324**

Note: The table above presents the reliability, resolution and comprehensive information of intervals obtained by different models. When PICP>PINC, the forecasting intervals are reliable. When the intervals are reliable, the narrower the interval width, the better the interval forecasting effect, which can be measured by $PINAW(\alpha)=\frac{1}{NR}\sum_{i=1}^{N}\left[UB_i(\alpha)-LB_i(\alpha)\right]$. Metric $AIS(\alpha)=\sum_{i=1}^{N}S^{(\alpha)}$ can comprehensively evaluate the interval forecasting effect.

4.2.2. Comparative Experimental Analysis with Ensemble Models

Reliability

The EPFM based on other denoising methods has high interval coverage, whether at $PINC = 90\%$ or at $PINC = 95\%$. However, the reliability of the prediction interval is not always guaranteed when other algorithms are used to perform the optimization. Of the four other optimization algorithms, only MOTa–EPFM and MODa–EPFM can obtain reliable interval prediction results in all cases (three sites and two confidence levels), similar to the proposed EPFS. For Site 1, the interval coverage rates of MOTa and MODa are $PICP_{PINC=90\%}^{MOTA} = 0.9833$, $PICP_{\boldsymbol{PINC=95\%}}^{MOTa} = 1$, $PICP_{\boldsymbol{PINC=90\%}}^{MODa} = 0.9778$, and $PICP_{\boldsymbol{PINC=95\%}}^{MODa} = 1$, respectively. The Emd, Eemd, CeemdAN, Wt-based PSr–EPFM, MOTa–EPFM, MODa–EPFM, and the proposed EPFS are better in terms of reliability.

Resolution

The interval width obtained by EPFMs using other denoising methods is significantly greater than that of the proposed EPFS, which is evident from the PINAW values. For Site 2, the PINAW values of Emd–EPFM, Eemd–EPFM, CeemdAN–EPFM, Wt–EPFM, and the proposed EPFS are $\overline{\mathbf{PINAW}}_{PINC=90\%}^{Site1} = \left[0.2921, 0.3027, 0.3577, 0.2517, 0.1620\right]$ and $\overline{\mathbf{PINAW}}_{PINC=95\%}^{Site1} = \left[0.3371, 0.3335, 0.3980, 0.2844, 0.2471\right]$ when $PINC = 90\%$ and $PINC = 95\%$, respectively.

The predicted interval widths of EPFMs using other optimization algorithms are similar to that of the proposed EPFS. However, EPFMs based on other algorithms have less reliability in their prediction results (manifested as low PICP values), with narrow interval widths. For Site 1, the PINAW values for MOGa–EPFM are 0.1755 at $PINC = 90\%$ and 0.2414 at $PINC = 95\%$; the corresponding PICP values are 0.9167 and 0.9611, respectively, and lower than those of all other models.

Sharpness

The interval prediction results after optimization are better than those of the single models, and other ensemble models cannot obtain interval reliability and interval width simultaneously. To better evaluate the effect of interval prediction, the AIS value was used to evaluate the uncertainty information contained in the prediction interval. A larger AIS value (AIS is generally a negative number), produces a better interval prediction. Considering all cases (three sites and two confidence levels), the proposed EPFS obtained the lowest AIS values, indicating that the proposed EPFS has better global optimization ability and better ability to optimize multiple objectives. The metric values of all three sites are presented in **Table 3** and **Fig. 4**.

Remark: Although EPFMs denoised by other methods have strong reliability, their resolution is not high; thus, the interval forecasting effect is not satisfying. EPFMs using other optimization algorithms cannot simultaneously optimize the interval reliability and interval resolution. The proposed EPFS has the best interval prediction results.

Table 3
Interval prediction metric values of ensemble models with PINC=90% and PINC=95%.

Dataset	Models	PICP		PINAW		AIS	
		PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%
Site 1	**SSa-PSr-MOTa-EPFM**	0.9833	1.0000	0.2048	0.2828	-0.1997	-0.1353
	SSa-PSr-MODa-EPFM	0.9778	1.0000	0.1979	0.3413	-0.1899	-0.1666
	SSa-PSr-MOGa-EPFM	0.9167	0.9611	**0.1755**	0.2414	-0.1887	-0.1223
	SSa-PSr-MOAa-EPFM	0.9333	1.0000	0.1769	0.2805	-0.1843	-0.1336
	Emd-PSr-EPFM	0.9778	1.0000	0.2538	0.3903	-0.2527	-0.1892
	Eemd-PSr-EPFM	0.9889	1.0000	0.2276	0.3824	-0.2226	-0.1845
	Wt-PSr-EPFM	0.9944	0.9944	0.3513	0.4031	-0.3453	-0.1941
	CeemdAN-PSr-EPFM	**1.0000**	**1.0000**	0.2868	0.3848	-0.2795	-0.1832
	Proposed EPFS	0.9833	0.9944	0.1893	**0.2674**	**-0.1805**	**-0.1273**
Site 2	**SSa-PSr-MOTa-EPFM**	0.9778	0.9556	0.1692	0.2476	-0.1627	-0.1291
	SSa-PSr-MODa-EPFM	0.8833	0.9500	**0.1454**	0.2512	-0.1569	-0.1364
	SSa-PSr-MOGa-EPFM	0.8833	0.9500	0.1570	**0.2342**	-0.1707	-0.1248
	SSa-PSr-MOAa-EPFM	0.8500	0.8778	0.1475	0.2385	-0.1737	-0.1518
	Emd-PSr-EPFM	**1.0000**	**1.0000**	0.2921	0.3371	-0.2887	-0.1633
	Eemd-PSr-EPFM	0.9944	0.9944	0.3027	0.3335	-0.2991	-0.1612
	Wt-PSr-EPFM	**1.0000**	**1.0000**	0.3577	0.3980	-0.3523	-0.1945
	CeemdAN-PSr-EPFM	0.9944	0.9778	0.2517	0.2844	-0.2505	-0.1445
	Proposed EPFS	0.9778	0.9889	0.1620	0.2471	**-0.1590**	**-0.1216**
Site 3	**SSa-PSr-MOTa-EPFM**	0.9667	0.9833	0.2489	0.3605	-0.2035	-0.1465
	SSa-PSr-MODa-EPFM	0.8833	0.9722	**0.2170**	0.3337	-0.1810	-0.1378
	SSa-PSr-MOGa-EPFM	0.9500	0.9611	0.2003	0.3431	-0.1846	-0.1448
	SSa-PSr-MOAa-EPFM	0.9833	0.9889	0.2172	**0.3179**	-0.1876	-0.1345
	Emd-PSr-EPFM	0.9611	0.9778	0.3240	0.3798	-0.2707	-0.1572
	Eemd-PSr-EPFM	0.9778	0.9833	0.3300	0.3843	-0.2692	-0.1581
	Wt-PSr-EPFM	0.9667	0.9833	0.3726	0.4324	-0.3200	-0.1873
	CeemdAN-PSr-EPFM	**0.9889**	**0.9944**	0.3253	0.3935	-0.2650	-0.1583
	Proposed EPFS	0.9833	0.9889	0.2180	0.3306	**-0.1784**	**-0.1324**

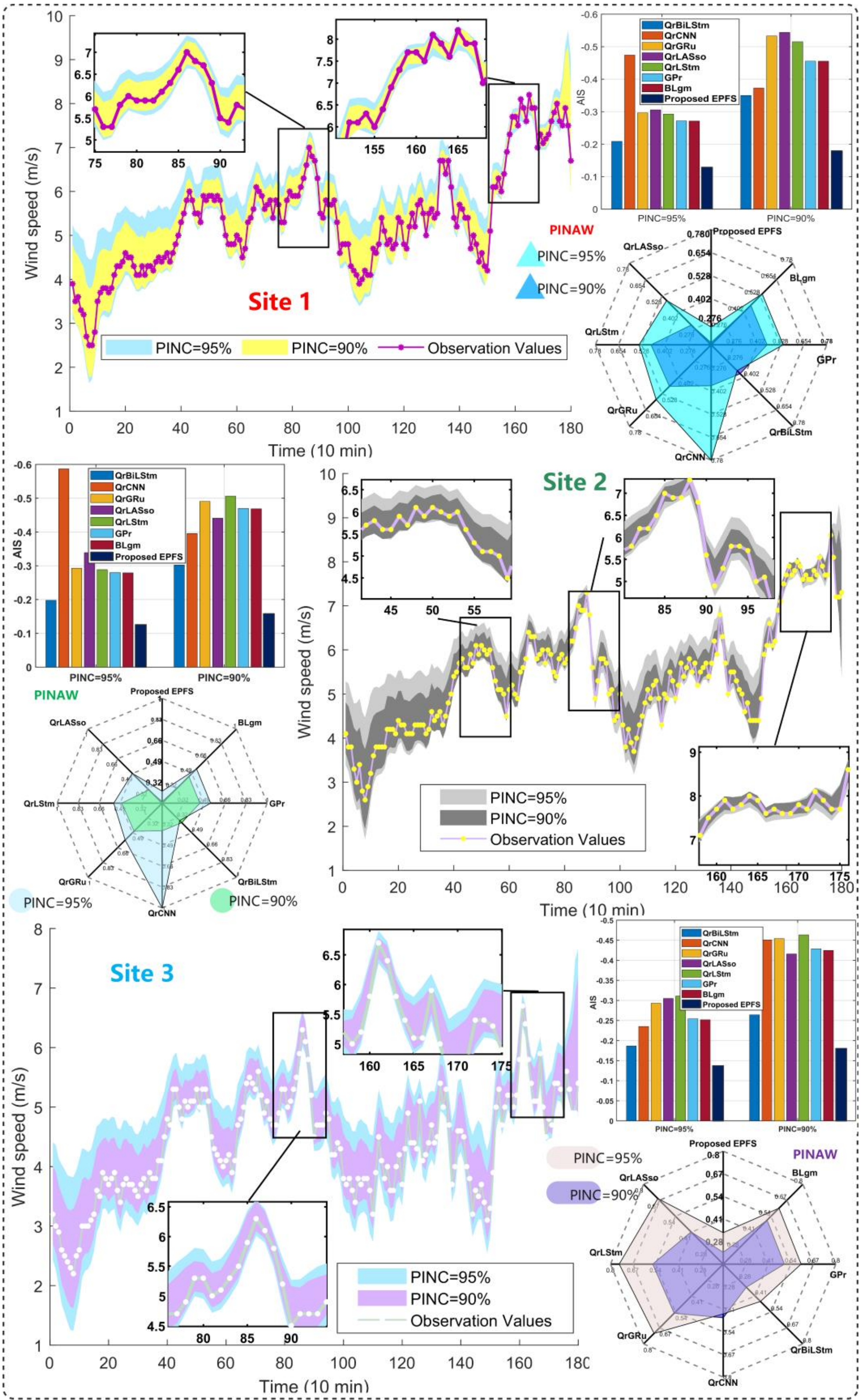


Fig.3. Forecasting results of proposed EPFS and other single models.

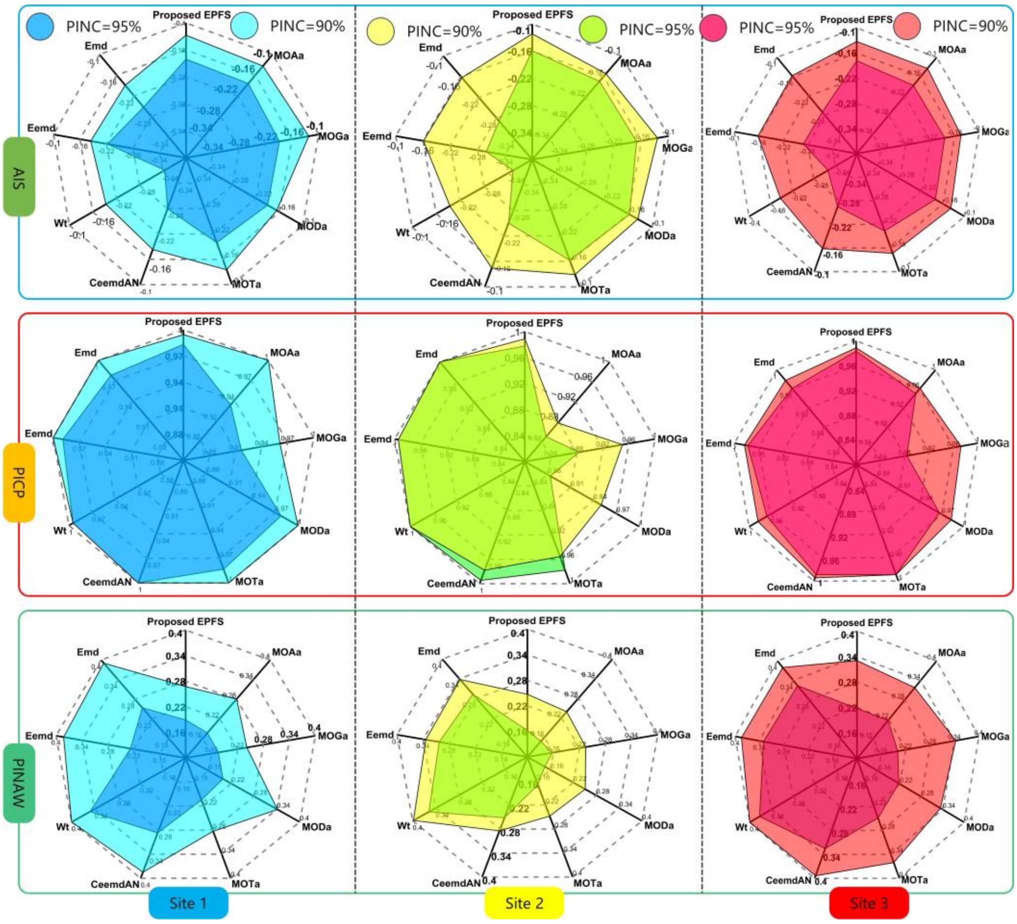


Fig.4. Forecasting results of proposed EPFS and other ensemble models.

4.3. Statistical tests

A Diebold-Mariano (DM) test [58] was implemented to validate whether the time-series forecasting results of two different models exist for the same or significantly different accuracy. A hypothesis test is conducted, where the null hypothesis says that there is no difference between the forecasting accuracy of the models compared, and the alternative hypothesis says that the forecasting error of the model proposed is different with the compared one. In this study, the hypothetical form can be defined as Eq. (29):

$$\begin{aligned} \boldsymbol{H}_0 &: \boldsymbol{E}\left[\boldsymbol{L}\left(\tilde{\xi}_{\boldsymbol{i}}^{(\boldsymbol{a})}\right)\right]=\boldsymbol{E}\left[\boldsymbol{L}\left(\tilde{\xi}_{\boldsymbol{i}}^{(\boldsymbol{b})}\right)\right] \\ \boldsymbol{H}_1 &: \boldsymbol{E}\left[\boldsymbol{L}\left(\tilde{\xi}_{\boldsymbol{i}}^{(\boldsymbol{a})}\right)\right]\neq\boldsymbol{E}\left[\boldsymbol{L}\left(\tilde{\xi}_{\boldsymbol{i}}^{(\boldsymbol{b})}\right)\right] \end{aligned} \tag{29}$$

in which $\boldsymbol{L}(\cdot)$ is the loss function of forecasting bias. In point-oriented forecasting, $\tilde{\xi}_{\boldsymbol{i}}$ is defined as the ith forecasting value minus the corresponding actual value. In interval forecast, $\tilde{\xi}_{\boldsymbol{i}}$ represented by Eq. (30). is defined as the interval score of the ith point.

$$\xi_i(\alpha)=\begin{cases}-2\alpha\Phi_i(\alpha)-4\left(\boldsymbol{LB}_i(\alpha)-\boldsymbol{Obse}V_i\right), & \boldsymbol{Obse}V_i<\boldsymbol{LB}_i(\alpha)\\ -2\alpha\Phi_i(\alpha), & \boldsymbol{Obse}V_i\in\left[\boldsymbol{LB}_i(\alpha),\boldsymbol{UB}_i(\alpha)\right]\\ -2\alpha\Phi_i(\alpha)-4\left(\boldsymbol{Obse}V_i-\boldsymbol{UB}_i(\alpha)\right), & \boldsymbol{Obse}V_i>\boldsymbol{UB}_i(\alpha)\end{cases} \quad (30)$$

If the null hypothesis is rejected, it is possible to say that there is statistical evidence that there exists a significant difference between the proposed model regarding the compared model at the α level of significance. The DM test results are shown in Table 4.

Table 4
Statistics of DM test for statistical comparison of proposed approach versus other models

Model	Site 1		Site 2		Site 3	
	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%
SSa-PSr-QrLASso	-10.2893^{***}	-6.7006^{***}	-2.9187^{***}	-3.8964^{***}	-19.8818^{***}	-4.2333^{***}
SSa-PSr-QrLStm	-68.6604^{***}	-66.8613^{***}	-4.5813^{***}	-6.6625^{***}	-42.4881^{***}	-23.2736^{***}
SSa-PSr-QrGRu	-45.0164^{***}	-65.5619^{***}	-11.1758^{***}	-7.9645^{***}	-35.1313^{***}	-20.3249^{***}
SSa-PSr-QrCNN	-4.0386^{***}	-4.5426^{***}	-41.1140^{***}	-9.2481^{***}	-2.9096^{***}	-7.3515^{***}
SSa-PSr-QrBiLStm	-49.3264^{***}	-48.2301^{***}	-2.4558^{**}	-33.6391^{***}	-14.4810^{***}	-9.2998^{***}
SSa-PSr-GPr	-49.1107^{***}	-59.7660^{***}	-11.7846^{***}	-44.0720^{***}	-30.8004^{***}	-21.7912^{***}
SSa-PSr-BLgm	-48.4703^{***}	-60.4071^{***}	-11.8552^{***}	-44.0617^{***}	-31.3836^{***}	-23.5181^{***}
Emd-PSr-EPFM	-30.5495^{***}	-5.8206^{***}	-1.7031^{*}	-17.1494^{***}	-3.0018^{***}	-8.0663^{***}
Eemd-PSr-EPFM	-28.5660^{***}	-13.4367^{***}	-1.6236	-21.1340^{***}	-3.1871^{***}	-6.6468^{***}
Wt-PSr-EPFM	-23.4755^{***}	-28.0409^{***}	-4.0643^{***}	-23.3569^{***}	-2.2902^{**}	-3.8406^{***}
CeemdAN-PSr-EPFM	-18.1712^{***}	-32.4127^{***}	-1.6707^{*}	-9.6597^{***}	-5.0214^{***}	-6.5898^{***}
SSa-PSr-MOTa-EPFM	-7.2410^{***}	-5.0958^{***}	-1.3035	0.3255	-6.5390^{***}	-3.5052^{***}
SSa-PSr-MODa-EPFM	-39.8391^{***}	-5.0069^{***}	-1.2672	-0.6341	-1.7147^{*}	-1.0823
SSa-PSr-MOGa-EPFM	-0.6505	-1.6847^{*}	-1.4715	-2.6176^{***}	-1.9806^{**}	-1.9018^{*}
SSa-PSr-MOAa-EPFM	-2.0946^{**}	-1.5301	-1.7845^{*}	-2.7302^{***}	-1.3332	-1.5073

Note: ***, **, and * indicate that the results are significant at the 1%, 5%, and 10% confidence levels, respectively.

The proposed EPFS is significantly different from the commonly used single models in that the absolute values of the DM test results are all greater than the threshold $Z_{0.01/2}=2.58$. Compared with the EPFM based on different denoising methods, only Site 2 when PINC=95%, the EPFM based on Eemd is not significantly different from the proposed EPFS; In other forecasting scenarios, the proposed EPFS is significantly different from EPFM based on other denoising methods. Compared with EPFM based on different optimization algorithms, the accuracy was not significantly different in a few forecasting situations. Take Site 2 as an example, when PINC=95%, the DM value of SSa-PSr-MOGa-EPFM is -0.6505. Table 4 shows that the EPFM based on different optimization algorithms is significantly different in most forecasting situations, so it is important to use the improved optimization algorithm, i.e., IMOTA to perform the optimization task.

4.4. Fist-order and second-order forecasting effectiveness evaluation

In this study, the forecasting effectiveness (FE) approach [52] was modified for uncertainty forecasting; its first-order and second-order values were used to measure the availability of models. The required bias in FE is modified as interval score which is defined as Eq. (31), and the first-order FE and second-order FE are obtained. The details of this indicator are provided as follows.

The element of the k^{th} order FE can be calculated as $\boldsymbol{g}^k = \sum_{i=1}^{n} \boldsymbol{Q}_i \boldsymbol{A}_i^k$, where $\boldsymbol{A}_i$ refers to the forecasting accuracy that can be measured by $\boldsymbol{A}_i = 1 - \left|\xi_i\right|$, where n is the length of the testing set, and $\boldsymbol{\xi}_i$ can be mathematically expressed by Eq. (29). $\boldsymbol{Q}_i$ indicates the discrete probability distribution, and $\sum_{i=1}^{n} \boldsymbol{Q}_i = 1, \boldsymbol{Q}_i > 0$. As prior information for $\boldsymbol{Q}_i$ cannot be obtained, it is commonly determined as $\boldsymbol{Q}_i = 1/\boldsymbol{n}, \boldsymbol{i} = 1, 2, \ldots, \boldsymbol{n}$.

$$\xi_i(\alpha) = \begin{cases} -2\alpha\Phi_i(\alpha) - 4\left(\boldsymbol{LB}_i(\alpha) - \boldsymbol{Obse}V_i\right), & \boldsymbol{Obse}V_i < \boldsymbol{LB}_i(\alpha) \\ -2\alpha\Phi_i(\alpha), & \boldsymbol{Obse}V_i \in \left[\boldsymbol{LB}_i(\alpha), \boldsymbol{UB}_i(\alpha)\right] \\ -2\alpha\Phi_i(\alpha) - 4\left(\boldsymbol{Obse}V_i - \boldsymbol{UB}_i(\alpha)\right), & \boldsymbol{Obse}V_i > \boldsymbol{UB}_i(\alpha) \end{cases} \quad (31)$$

where $\Phi_i(\alpha) = \boldsymbol{UB}_i(\alpha) - \boldsymbol{LB}_i(\alpha)$.

Thereafter, a function $FE\left(\boldsymbol{g}^1, \boldsymbol{g}^2, \cdots, \boldsymbol{g}^k\right)$ that contains k elements is designed to assess the k^{th} order ***FE***. Recall that $\boldsymbol{g}^k = \sum_{i=1}^{n} \boldsymbol{Q}_i \boldsymbol{A}_i^k$, the first-order FE, can be defined as $FE\left(\boldsymbol{g}^1\right) = \boldsymbol{g}^1$, and the second-order ***FE*** is designed as $FE\left(\boldsymbol{g}^1, \boldsymbol{g}^2\right) = \boldsymbol{g}^1\left(1 - \sqrt{\boldsymbol{g}^2 - \left(\boldsymbol{g}^1\right)^2}\right)$. The ***FE*** values are presented in **Table 5**.

According to the design mechanism of ***FE***, a higher ***FE*** value indicates greater availability of models. By comparing the first- and second-order ***FE*** values with reference models, we can determine that, apart from some results for Site 2, the proposed EPFS has the strongest availability. For Site 1, considering both *Order* = 1 and *Order* = 2, the ***FE*** values of the proposed EPFS are higher than those of other forecasting models. The ***FE*** values obtained using the proposed EPFS were $FE_{\boldsymbol{1st\text{-}order}}^{Site1} = 0.8195$ and $FE_{\boldsymbol{2\text{-}order}}^{Site1} = 0.7548$, indicating that the proposed EPFS can fully assimilate the merits of benchmark units and achieve the most satisfactory results at all three sites.

4.5. Improvement ratio

The indicator $\overline{\boldsymbol{IR}_{\boldsymbol{Metric}}}$ was used to assess the improvement in forecasting accuracy of the EPFS. $\overline{\boldsymbol{IR}_{\boldsymbol{Metric}}}$ can be defined as

$$\overline{\boldsymbol{IR}_{\boldsymbol{Metric}}} = \left[\left(\boldsymbol{Metric}^{com} - \boldsymbol{Metric}^{pro}\right) \middle/ \boldsymbol{Metric}^{com}\right] \quad (32)$$

where $\boldsymbol{Metric}^{com}$ is the metric value of the compared model, and $\boldsymbol{Metric}^{pro}$ indicates the metric value of the proposed EPFS. In this paper, the improvement rate of AIS is calculated to reflect the overall improvement ratio of the proposed EPFS compared with other models, and the improvement rate of PINAW reflects the contribution of the proposed EPFS to shortening the interval width. **Table 6** presents the improvement ratio results. According to the $\overline{\boldsymbol{IR}_{\boldsymbol{Metric}}}$, the following conclusions can be drawn.

Compared with all other single models, the forecasting performance of the proposed QrBiLStm is improved, indicating that the proposed QrBiLStm is effective.

The EPFS obtained by optimizing two efficient QrBiLStm units can further improve forecasting, manifested as a lower interval width and higher resolution. At Site 1, when $PINC = 90\%$ and $PINC = 95\%$, the improvement ratios of the proposed EPFS to the AIS value of a single model are $\overline{\overline{IR}}_{AIS}^{PINC=90\%}$ = $[66.8235\%, 64.9752\%, 66.1685\%, 51.5732\%, 48.4414\%]$ and $\overline{\overline{IR}}_{PINC=95\%}^{AIS}$ = $[58.3401\%, 56.5286\%, 57.0647\%, 73.1684\%, 39.0008\%]$. Taking into account $PINC = 90\%$ and $PINC = 95\%$ of the three Sites, the proposed EPFS represents a minimum of 12.96% improvement in AIS metrics compared to the EPFMs based on the other four denoising methods. Compared with EPFMs based on other optimization algorithms, the improvement ratio of the AIS metric is 1.57% to 23.60% and the forecasting results are more stable. This indicating that the proposed IMOTa has better global optimization ability and better ability to optimize multiple objectives.

4.6. Stability analysis

As most swarm intelligence optimization methods incorporate randomness or probabilistic mechanisms into their operation, the forecasting results for each trial are generally different, even with the same parameters and conditions. Thus, the stability of swarm intelligence optimization is one of the most important factors affecting prediction performance.

The stability is measured by the standard deviations of three evaluation metrics; the equation is expressed as $\overline{\overline{Std}}(Metric) = \sum_{t=1}^{N}\left(Metric_t^k - Metric_t\right)^2 / N$, where N is the number of testing trials; $Metric_k$ indicates the k^{th} error metric (PINAW and AIS) values in the testing trial, and $Metric_t$ refers to the average metric values of all testing trials. A lower $Std(Metric)$ value indicates a greater degree of stability.

Fig. 5a, b, and d show the distribution of different results from ten trials using IMOTA for the three indicators at $PINC = 90\%$; **Fig. 5** e, f, and h show the distribution of different results at $PINC = 95\%$; **Fig. 5** c and g show the $Std(Metric)$ results for ten trials at $PINC = 90\%$ and $PINC = 95\%$. The stability analysis is conducted using Site 1 as an example. Generally, the $Std(Metric)$ values are small for both $\alpha = 0.1$ and $\alpha = 0.05$, manifested as $\overline{\overline{Std}}(PICP)_{\alpha=0.1,\alpha=0.05}^{\text{Site1}} = [0.0240, 0.0228]$, $\overline{\overline{std}}(PINAW)_{\alpha=0.1,\alpha=0.05}^{\text{Site1}} = [0.0117, 0.0104]$, and $\overline{\overline{std}}(AIS)_{\alpha=0.1,\alpha=0.05}^{\text{Site1}}$ $= [0.0082, 0.0048]$. In ten trials, the values of the three indicators fluctuated little. The values of the three indicators were analyzed. For $PINC = 90\%$, the minimum value of PICP in the ten trials was 0.9389, and the maximum value was 0.9889. PINAW had a maximum of 0.2086, and a minimum of 0.1752. The minimum AIS value was -0.2021, and the maximum value was -0.1783. For PINC = 95%, the minimum value of PICP in the ten trials was 0.9222, and the maximum value was 0.9944. The PINAW maximum was 0.2657, and the minimum was 0.2303. The minimum AIS value was -0.1311, and the maximum value was -0.1163.

In practical applications, future values are not available to calculate the metrics for comparison. However, through the stability analysis, we found that the proposed EPFS can achieve accurate forecasting results in all trials, which shows that the proposed EPFS is highly available.

4.7. Advantages and disadvantages compared to the existing studies

The advantages and disadvantages of this study compared with the existing models in this field and the future work are analyzed in this section.

Advantages

(1) Firstly, compared with the existing QR- machine learning models such as QrGRu [29], QrLStm [30], , QrLASso [27], QrCNN, etc., this study adopts the cell structure of BiLStm, which can train the network by inputting historical information forward and backward. Based on this network structure, more accurate interval forecasting results can be obtained. This conclusion can be drawn from the results of comparative experiment 1 (section 4.2.1). Secondly, this study proposes a probabilistic ensemble forecasting system, which can combine the forecasting results of two well-behaved single models to get more accurate forecasting results, specifically by reducing the interval width under the condition of ensuring high interval coverage.

(2) Compared with the existing probabilistic ensemble forecasting models. Firstly, the proposed EPFS is based on Qr theory, which can optimize the upper and lower bounds of the interval respectively. This optimization strategy is more flexible, ensuring that both the upper bound and the lower bound are optimal results. For example, Niu's model [36] is based on data distribution, and the upper and lower bound forecasting results are optimized simultaneously. Secondly, the objective functions of probabilistic ensemble optimization are designed to find the solution with high coverage and narrow interval width as the optimal solution.

Disadvantages

(1) The calculation burden is increased while the upper and lower bounds of the forecasting interval are optimized respectively. Every interval forecasting result obtained by EPFS needs to be optimized twice, which increases the operation time.
(2) The distribution information of data is not used to construct the interval. In this paper, loss function of the neural network is designed as pinball loss to obtain different quantile forecasting results, which is a supervised machine learning method without data distribution information. Although the EPFS based on Qr can get accurate interval forecasting results after optimization, it is possible to get better forecasting results if the distribution information of historical data can be fully utilized. This is the future work.
(3) The EPFS is based on historical wind-speed data without considering other influence factors, such as pressure and temperature. Probabilistic ensemble forecast of multivariate time series is also the future research direction.

Table 5
The forecasting effectiveness of the proposed EPFS and other existing models.

Model	Site 1				Site 2				Site 3			
	PINC=90%		PINC=95%		PINC=90%		PINC=95%		PINC=90%		PINC=95%	
	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
SSa-PSr-QrLASso	0.4561	0.3296	0.6945	0.6149	0.5591	0.3384	0.6611	0.4336	0.5843	0.4315	0.6951	0.6476
SSa-PSr-QrLStm	0.4848	0.4591	0.7072	0.6889	0.4937	0.3964	0.7116	0.6228	0.5367	0.4978	0.6890	0.6643
SSa-PSr-QrGRu	0.4666	0.4407	0.7036	0.6714	0.5095	0.4227	0.7070	0.6611	0.5461	0.4979	0.7067	0.6758
SSa-PSr-QrCNN	0.6274	0.4947	0.5257	0.3193	0.6046	0.5458	0.4132	0.3823	0.5492	0.4249	0.7649	0.6015
SSa-PSr-GPr	0.5443	0.5407	0.7288	0.7285	0.5315	0.5311	0.7209	0.7206	0.5707	0.5574	0.7481	0.7472
SSa-PSr-BLgm	0.5444	0.5426	0.7284	0.7271	0.5305	0.5287	0.7203	0.7188	0.5734	0.5532	0.7456	0.7443
SSa-PSr-QrBiLStm	0.6500	0.6131	0.7914	0.7577	0.6198	0.5771	0.8252	0.7941	0.7229	0.6775	0.8165	0.7815
Emd-PSr-EPFM	0.7473	0.6834	0.8108	0.7721	0.7113	0.6645	0.8367	0.7951	0.7293	0.6630	0.8428	0.7864
Eemd-PSr-EPFM	0.7774	0.7330	0.8155	0.7761	0.7009	0.6523	0.8388	0.7960	0.7308	0.6770	0.8419	0.7839
Wt-PSr-EPFM	0.6547	0.6034	0.8059	0.7627	0.6477	0.5934	0.8055	0.7660	0.6800	0.5679	0.8127	0.7041
CeemdAN-PSr-EPFM	0.7205	0.6742	0.8168	0.7691	0.7495	0.7019	0.8555	0.7854	0.7351	0.6874	0.8417	0.8008
SSa-PSr-MOTa-EPFM	0.8003	0.7309	0.8647	0.8266	0.8373	**0.7728**	0.8710	0.7661	0.7965	0.7357	0.8535	0.8073
SSa-PSr-MODa-EPFM	0.8101	0.7372	0.8334	0.8017	**0.8431**	0.7561	0.8636	0.7278	0.8190	0.7211	0.8622	0.7936
SSa-PSr-MOGa-EPFM	0.8113	0.7006	**0.8777**	0.8111	0.8293	0.7378	0.8752	0.7773	0.8154	0.6985	0.8552	0.7766
SSa-PSr-MOAa-EPFM	0.8157	0.7135	0.8578	0.7401	0.8264	0.7228	0.8482	0.6991	0.8124	0.7082	0.8655	0.7893
Proposed EPFS	**0.8195**	**0.7548**	0.8727	**0.8302**	0.8410	0.7660	**0.8784**	**0.8072**	**0.8216**	**0.7477**	**0.8676**	**0.8232**

Note: This table shows the circumstantial values of the *first*-order and *second*-order ***FE*** of sixteen models. The *first*-order ***FE*** is defined as $\boldsymbol{FE}\left(\boldsymbol{g}^1\right)=\boldsymbol{g}^1$. When this continues function contains two variables, the *2nd*-order ***FE*** can be denoted by $\boldsymbol{FE}\left(\boldsymbol{g}^1,\boldsymbol{g}^2\right)=\boldsymbol{g}^1\left(1-\sqrt{\boldsymbol{g}^2-\left(\boldsymbol{g}^1\right)^2}\right)$.

Table 6
The improvement ratio of the proposed EPFS.

Models	Site 1				Site 2				Site 3			
	PINAW		AIS		PINAW		AIS		PINAW		AIS	
	PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%
SSa-PSr-QrLASso	0.5844	0.4994	0.6682	0.5834	0.4650	0.4976	0.6394	0.6412	0.4712	0.5164	0.5708	0.5657
SSa-PSr-QrLStm	0.6332	0.5417	0.6498	0.5653	0.6604	0.5439	0.6860	0.5784	0.6085	0.5591	0.6149	0.5743
SSa-PSr-QrGRu	0.6454	0.5533	0.6617	0.5706	0.6560	0.5687	0.6759	0.5850	0.6038	0.5369	0.6069	0.5486
SSa-PSr-QrCNN	0.3984	0.6469	0.5157	0.7317	0.5662	0.7890	0.5980	0.7928	0.5247	0.2421	0.6042	0.4370
SSa-PSr-GPr	0.5753	0.4966	0.6040	0.5307	0.6437	0.5440	0.6607	0.5643	0.5665	0.4483	0.5800	0.4745
SSa-PSr-BLgm	0.5756	0.4971	0.6039	0.5314	0.6443	0.5448	0.6614	0.5653	0.5709	0.4533	0.5837	0.4797
SSa-PSr-QrBiLStm	0.4574	0.3708	0.4844	0.3900	0.5661	0.2928	0.5819	0.3042	0.3558	0.2714	0.3562	0.2787
SSa-PSr-MOTa-EPFM	0.0758	0.0545	0.0965	0.0596	0.0423	0.0020	0.0229	0.0578	0.1243	0.0829	0.1231	0.0961
SSa-PSr-MODa-EPFM	0.0436	0.2166	0.0500	0.2360	-0.1143	0.0164	0.0132	0.1083	-0.0047	0.0093	0.0141	0.0395
SSa-PSr-MOGa-EPFM	-0.0787	-0.1075	0.0438	0.0407	-0.0317	-0.0550	0.0690	0.0256	-0.0883	0.0364	0.0334	0.0860
SSa-PSr-MOAa-EPFM	-0.0702	0.0466	0.0208	0.0471	-0.0983	-0.0360	0.0846	0.1988	-0.0038	-0.0399	0.0487	0.0157
Emd-PSr-EPFM	0.2542	0.3150	0.2859	0.3274	0.4453	0.2670	0.4494	0.2553	0.3273	0.1296	0.3408	0.1580
Eemd-PSr-EPFM	0.1685	0.3008	0.1894	0.3104	0.4647	0.2592	0.4686	0.2457	0.3394	0.1397	0.3372	0.1625
Wt-PSr-EPFM	0.4612	0.3368	0.4774	0.3442	0.5470	0.3792	0.5488	0.3747	0.4150	0.2354	0.4424	0.2930
CeemdAN-PSr-EPFM	0.3399	0.3051	0.3544	0.3052	0.3563	0.1311	0.3654	0.1584	0.3299	0.1599	0.3264	0.1639

Note: The table above reports the IR of the proposed EPFS from other twelve models. The AIS and PINAW are used to measure the IR, and the corresponding indicator can be defined as $\overline{IR}_{Metric} = \left[\left(Metric^{com} - Metric^{pro}\right)/Metric^{com}\right]$, where $Metric^{com}$ is the metric values of compared model, and the $Metric^{pro}$ indicates the metric value of the proposed **EPFS**.

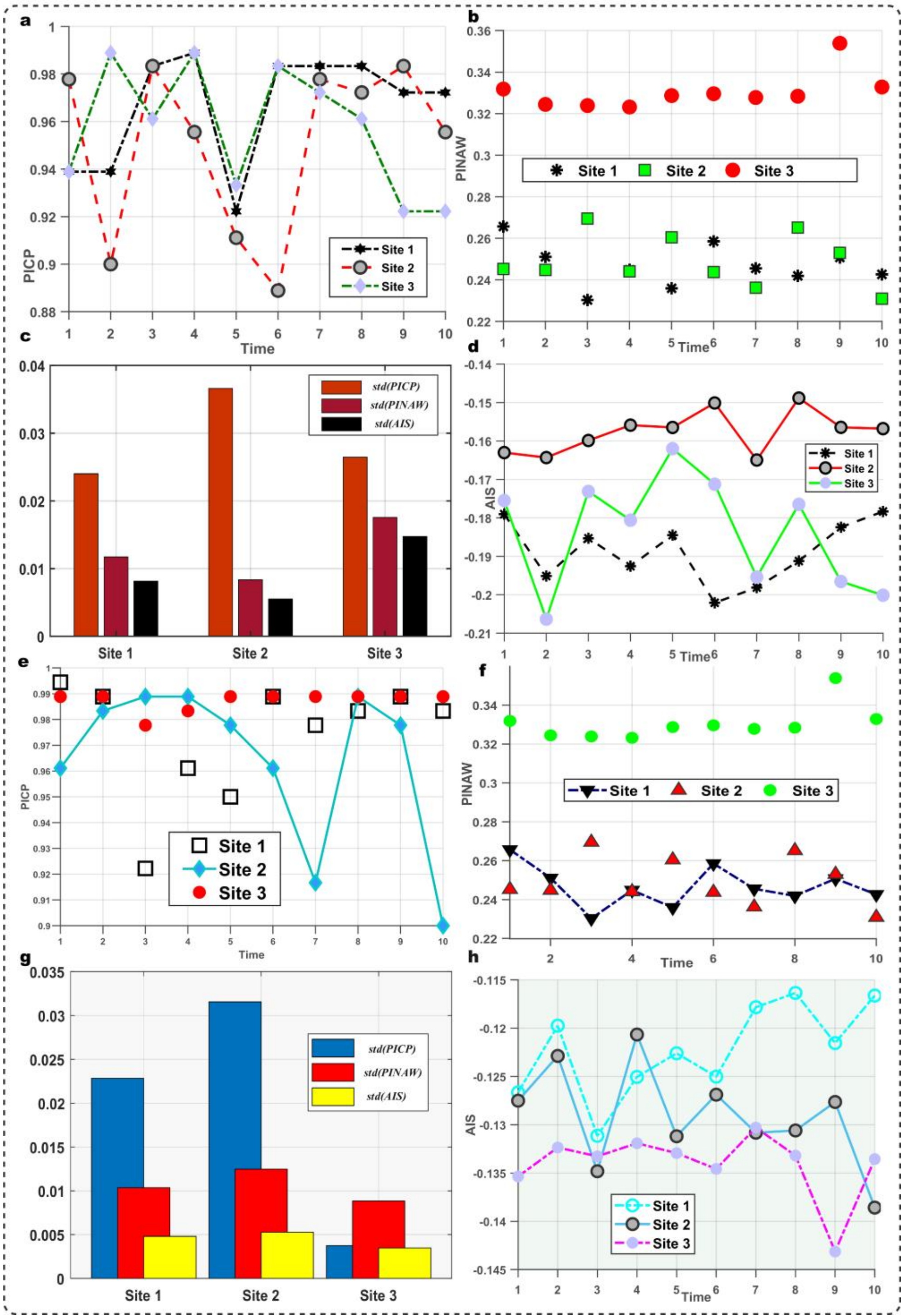


Fig.5. Stability analysis results of the proposed EPFS.

5. Conclusion

To generate high-quality wind speed prediction intervals and obtain comprehensive potential uncertainty information, an EPFS that combines SSa, PSr, QrBiSLStm, pseudo-interval construction, and multi-objective optimization was proposed. SSa and PSr were implemented successively, and time-frequency decomposition and reconstruction of the sequence were performed. The conditional quantiles of the sequences were obtained using the proposed QrBiLStm, and prediction intervals with different confidence levels were constructed. A pseudo-interval was constructed as the training set based on the forecasting results of two QrBiLStm units, and the proposed IMOTa was used for combinatorial optimization to obtain the final interval forecasting results. Comparison experiments were performed on three datasets, and the forecasting results were comprehensively evaluated in terms of reliability, resolution, and sharpness. Based on the analysis, we can draw the following conclusions: (1) a decomposition and reconstruction mechanism based on SSa and PSr can significantly improve forecasting performance, which can greatly improve uncertainty forecasting performance of the EPFS; (2) reliable uncertainty forecasts can be obtained from newly constructed QrBiLStm units; the forecasting results of this model far exceed those of other single models based on distribution hypothesis, and those of other Qr–deep learning models; (3) an ensemble probabilistic forecasting strategy based on pseudo-interval construction can effectively optimize the upper and lower bounds of the interval and further improve the forecasting performance of the main forecasting model; (4) the improved MOTa has better global optimization ability and stability, and produces more effective and stable prediction results. The main limitations of the proposed EPFS are as follows: (1) Because quantile loss is discontinuous and nondifferentiable around 0 point, this EPFS has not been applied to the field of deterministic forecasts; (2) It is not combined with other linear models or interval forecasting models based on distribution in ensemble forecasting. However, this EPFS can optimize the upper and lower bounds of the interval separately and does not need to assume the distribution in advance. This study provides a novel approach for wind speed ensemble probabilistic forecasting and can be used as a powerful decision tool in the power system scheduling process.

Acknowledgements

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

Appendix A

Table A1

Parameters setting.

Models	Symbol	Meaning	Determination method	determined value
Eemd	Ns	STD of added noise	Preset	0.05
	NR	Realization Number	Preset	50
	Ms	Maximum Sifting Iteration	Preset	500
CeemdAN	Ns	STD of added noise	Preset	0.05
	NR	Realization Number	Preset	50
	Ms	Maximum Sifting Iteration	Preset	500
Wt	D1	Decomposition Layer Number	Trial and error approach	5
SSa	W	Window Length	Trial and error approach	50
	D2	Primary Ingredient Disintegration Number	Karhunene Loeve decomposition	20
QrLStm	Nl	Number of hidden layers	Trial and error approach	2
	Nn	Number of hidden nodes	Trial and error approach	50
QrBILStm	Nl	Number of hidden layers	Trial and error approach	1
	Nn	Number of hidden nodes	Trial and error approach	50
QrGRu	Nl	Number of hidden layers	Trial and error approach	2
	Nn	Number of hidden nodes	Trial and error approach	50
QrCNN	Ck	Convolution kernel size	Trial and error approach	2*2
	Nn	Number of hidden layer nodes	Trial and error approach	35
MOGa, MOAa, MODa	As	Archive Size	Preset	10
	In	Iteration Number	Preset	200
	Ni	Individual Number	Trial and error approach	40
MOTa, IMOTa	Si	Initial Speed	Preset	1
	Nt	Number of tunicate	Trial and error approach	40
	Ss	Subordinate Speed	Preset	4
	As	Archive Size	Trial and error approach	10
	In	Iteration Number	Trial and error approach	200
PSr	τ	Embedded dimension	C-C method	Site 1~3: 4, 5,5
	M	Delay time	C-C method	Site 1~3: 31, 28, 28

Table A2
List of abbreviations.

ARIMA	Auto Regressive Integrated Moving Average	**LStm**	Long short-term memory neural network
ARMA	Autoregressive moving average	**LUBE**	lower upper bound estimation
BiLStm	Bi-directional Long Short Term Memory Network	**MO**	Multi-objective
BLgm	Bayesian regression model	**MOAa**	MO Antlion Algorithm
BPNN	Back propagation neural network	**MODa**	MO Dragonfly Algorithm
CeemdAN	Complete ensemble Empirical Mode Decomposition with Adaptive Noise	**MOGa**	MO Grasshopper Algorithm
Eemd	Ensemble Empirical Mode Decomposition	**MVE**	mean-variance estimates
EFS	Exponential function steps	**PINC**	prediction interval nominal confidence
ELM	Extreme learning machine	**PSr**	Phase space reconstruction
Emd	Empirical Mode Decomposition	**Qr**	quantile regression
EOL	Elite opposition learning	**QrCNN**	Quantile regression convolution neural network
EPFM	Ensemble probabilistic forecasting model	**QrGRu**	Quantile regression gated recurrent unit
EPFS	Ensemble probabilistic forecasting system	**QrLStm**	Quantile regression Long Short-Term Memory Network
GPr	Gaussian Process Regression	**RNN**	Recurrent neural network
GRu	Gated recurrent unit	**SSa**	Singular Spectral Analysis
GW	Gigawatt	**SVQr**	Support Vector Quantile Regression
GWEC	Global wind energy councile	**SVR**	Support Vector Regression
IMOTa	Improved multi-objective tunicate swarm algorithm	**Wt**	Wavelet Transform

Data Availability

10-minute wind speed data of three Sites in Shandong Peninsula: https://data.mendeley.com/datasets/sjyf2nhzdt/draft?a=af12330a-125b-499a-9473-6840ed7044f9

References

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

[2] Jiang P, Liu Z, Niu X, Zhang L. A combined forecasting system based on statistical method, artificial neural networks, and deep learning methods for short-term wind speed forecasting. Energy 2020:119361. https://doi.org/10.1016/j.energy.2020.119361.

[3] Jung J, Broadwater RP. Current status and future advances for wind speed and power forecasting. Renew Sustain Energy Rev 2014. https://doi.org/10.1016/j.rser.2013.12.054.

[4] Sheng C, Zhao J, Wang W, Leung H. Prediction intervals for a noisy nonlinear time series based on a bootstrapping reservoir computing network ensemble. IEEE Trans Neural Networks Learn Syst 2013;24:1036–48. https://doi.org/10.1109/TNNLS.2013.2250299.

[5] Niu X, Wang J, Zhang L. Carbon price forecasting system based on error correction and divide-conquer strategies. Appl Soft Comput 2021:107935. https://doi.org/10.1016/j.asoc.2021.107935.

[6] Aasim, Singh SN, Mohapatra A. Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting. Renew Energy 2019. https://doi.org/10.1016/j.renene.2019.01.031.

[7] Shukur OB, Lee MH. Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA. Renew Energy 2015. https://doi.org/10.1016/j.renene.2014.11.084.

[8] Manero J, Béjar J, Cortés U. Wind energy forecasting with neural networks: A literature review. Comput y Sist 2018. https://doi.org/10.13053/CyS-22-4-3081.

[9] Du P, Wang J, Yang W, Niu T. Multi-step ahead forecasting in electrical power system using a hybrid forecasting system. Renew Energy 2018. https://doi.org/10.1016/j.renene.2018.01.113.

[10] 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. https://doi.org/10.1016/J.RENENE.2021.07.113.

[11] 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. https://doi.org/10.1016/j.apenergy.2020.115561.

[12] Moreno SR, Coelho LDS, Ayala HVH, Mariani VC. Wind turbines anomaly detection based on power curves and ensemble learning. IET Renew Power Gener 2020. https://doi.org/10.1049/iet-rpg.2020.0224.

[13] 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. https://doi.org/10.1016/j.apenergy.2019.03.097.

[14] Chen H, Birkelund Y, Zhang Q. Data-augmented sequential deep learning for wind power forecasting. Energy Convers Manag 2021;248:114790.

https://doi.org/10.1016/J.ENCONMAN.2021.114790.
[15] Khodayar M, Wang J, Manthouri M. Interval Deep Generative Neural Network for Wind Speed Forecasting. IEEE Trans Smart Grid 2019. https://doi.org/10.1109/TSG.2018.2847223.
[16] Du P, Wang J, Yang W, Niu T. A novel hybrid model for short-term wind power forecasting. Appl Soft Comput J 2019. https://doi.org/10.1016/j.asoc.2019.03.035.
[17] Salkind N. American Statistical Association. Encycl Res Des 2012;92:748–57. https://doi.org/10.4135/9781412961288.n9.
[18] Khosravi A, Nahavandi S. An optimized mean variance estimation method for uncertainty quantification of wind power forecasts. Int J Electr Power Energy Syst 2014. https://doi.org/10.1016/j.ijepes.2014.03.060.
[19] Lian C, Zhu L, Zeng Z, Su Y, Yao W, Tang H. Constructing prediction intervals for landslide displacement using bootstrapping random vector functional link networks selective ensemble with neural networks switched. Neurocomputing 2018. https://doi.org/10.1016/j.neucom.2018.02.046.
[20] Khosravi A, Mazloumi E, Nahavandi S, Creighton D, Van Lint JWC. Prediction intervals to account for uncertainties in travel time prediction. IEEE Trans Intell Transp Syst 2011;12:537–47. https://doi.org/10.1109/TITS.2011.2106209.
[21] Quan H, Srinivasan D, Khosravi A. Short-term load and wind power forecasting using neural network-based prediction intervals. IEEE Trans Neural Networks Learn Syst 2014;25:303–15. https://doi.org/10.1109/TNNLS.2013.2276053.
[22] He Y, Li H. Probability density forecasting of wind power using quantile regression neural network and kernel density estimation. Energy Convers Manag 2018. https://doi.org/10.1016/j.enconman.2018.03.010.
[23] Wang HZ, Wang GB, Li GQ, Peng JC, Liu YT. Deep belief network based deterministic and probabilistic wind speed forecasting approach. Appl Energy 2016. https://doi.org/10.1016/j.apenergy.2016.08.108.
[24] Das K, Krzywinski M, Altman N. Quantile regression. Nat Methods 2019. https://doi.org/10.1038/s41592-019-0406-y.
[25] Taylor JW. A quantile regression neural network approach to estimating the conditional density of multiperiod returns. J Forecast 2000. https://doi.org/10.1002/1099-131x(200007)19:4<299::aid-for775>3.3.co;2-m.
[26] He Y, Liu R, Li H, Wang S, Lu X. Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory. Appl Energy 2017. https://doi.org/10.1016/j.apenergy.2016.10.079.
[27] He Y, Qin Y, Lei X, Feng N. A study on short-term power load probability density forecasting considering wind power effects. Int J Electr Power Energy Syst 2019. https://doi.org/10.1016/j.ijepes.2019.05.063.
[28] Chen T, Qian Z, Jing B, Wan J, Zhang F. Probabilistic Wind Speed Forecasting based on Minimal Gated Unit and Quantile Regression. J. Phys. Conf. Ser., 2020. https://doi.org/10.1088/1742-6596/1659/1/012039.
[29] Zhang Z, Qin H, Liu Y, Yao L, Yu X, Lu J, et al. Wind speed forecasting based on Quantile Regression Minimal Gated Memory Network and Kernel Density Estimation. Energy Convers Manag 2019.

https://doi.org/10.1016/j.enconman.2019.06.024.
[30] Wang K, Fu W, Chen T, Zhang B, Xiong D, Fang P. A compound framework for wind speed forecasting based on comprehensive feature selection, quantile regression incorporated into convolutional simplified long short-term memory network and residual error correction. Energy Convers Manag 2020;222:113234. https://doi.org/10.1016/j.enconman.2020.113234.
[31] Jiang P, Liu Z, Wang J, Zhang L. Decomposition-selection-ensemble forecasting system for energy futures price forecasting based on multi-objective version of chaos game optimization algorithm. Resour Policy 2021;73:102234. https://doi.org/10.1016/j.resourpol.2021.102234.
[32] Ribeiro MHDM, da Silva RG, Moreno SR, Mariani VC, Coelho L dos S. Efficient bootstrap stacking ensemble learning model applied to wind power generation forecasting. Int J Electr Power Energy Syst 2022;136:107712. https://doi.org/10.1016/J.IJEPES.2021.107712.
[33] Wang J, Du P, Lu H, Yang W, Niu T. An improved grey model optimized by multi-objective ant lion optimization algorithm for annual electricity consumption forecasting. Appl Soft Comput J 2018. https://doi.org/10.1016/j.asoc.2018.07.022.
[34] Liu Z, Jiang P, Wang J, Zhang L. Ensemble forecasting system for short-term wind speed forecasting based on optimal sub-model selection and multi-objective version of mayfly optimization algorithm. Expert Syst Appl 2021;177:114974. https://doi.org/10.1016/j.eswa.2021.114974.
[35] 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 2021:121275. https://doi.org/10.1016/j.energy.2021.121275.
[36] Wang J, Niu T, Du P, Yang W. Ensemble probabilistic prediction approach for modeling uncertainty in crude oil price. Appl Soft Comput J 2020;95:106509. https://doi.org/10.1016/j.asoc.2020.106509.
[37] Miao K chao, Han T ting, Yao Y qing, Lu H, Chen P, Wang B, et al. Application of LSTM for short term fog forecasting based on meteorological elements. Neurocomputing 2020. https://doi.org/10.1016/j.neucom.2019.12.129.
[38] Wang Y, Wang J, Li Z, Yang H, Li H. Design of a combined system based on two-stage data preprocessing and multi-objective optimization for wind speed prediction. Energy 2021;231:121125. https://doi.org/10.1016/j.energy.2021.121125.
[39] Wang J, Li Q, Zeng B. Multi-layer cooperative combined forecasting system for short-term wind speed forecasting. Sustain Energy Technol Assessments 2021. https://doi.org/10.1016/j.seta.2020.100946.
[40] Khosravi A, Nahavandi S, Creighton D, Jaafar J. Wind farm power uncertainty quantification using a mean-variance estimation method. 2012 IEEE Int. Conf. Power Syst. Technol. POWERCON 2012, 2012. https://doi.org/10.1109/PowerCon.2012.6401280.
[41] Efron B, Tibshirani R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Stat Sci 1986. https://doi.org/10.1214/ss/1177013815.
[42] DiCiccio TJ, Efron B. Bootstrap confidence intervals. Stat Sci 1996.

https://doi.org/10.1214/ss/1032280214.
[43] Lian C, Zeng Z, Wang X, Yao W, Su Y, Tang H. Landslide displacement interval prediction using lower upper bound estimation method with pre-trained random vector functional link network initialization. Neural Networks 2020;130:286–96. https://doi.org/10.1016/j.neunet.2020.07.020.
[44] Waldmann E. Quantile regression: A short story on how and why. Stat Modelling 2018. https://doi.org/10.1177/1471082X18759142.
[45] Afshar K, Bigdeli N. Data analysis and short term load forecasting in Iran electricity market using singular spectral analysis (SSA). Energy 2011;36:2620–7. https://doi.org/10.1016/j.energy.2011.02.003.
[46] Xu Z, Zhong L, Zhang A. Phase Space Reconstruction-Based Conceptor Network for Time Series Prediction. IEEE Access 2019. https://doi.org/10.1109/ACCESS.2019.2952365.
[47] Hochreiter S, Schmidhuber J. Long Short-Term Memory. Neural Comput 1997. https://doi.org/10.1162/neco.1997.9.8.1735.
[48] Caterini AL, Chang DE. Recurrent neural networks. SpringerBriefs Comput. Sci., 2018. https://doi.org/10.1007/978-3-319-75304-1_5.
[49] Schuster M, Paliwal KK. Bidirectional recurrent neural networks. IEEE Trans Signal Process 1997. https://doi.org/10.1109/78.650093.
[50] Tyapkin VN, Ishchuk IN, Kabulova EG, Semenov ME, Meleshenko PA. Singular spectral analysis in filtration of noisecontaminated signals of pseudolite navigation. Indian J Sci Technol 2016. https://doi.org/10.17485/ijst/2016/v9i46/107567.
[51] Moreno SR, Mariani VC, Coelho L dos S. Hybrid multi-stage decomposition with parametric model applied to wind speed forecasting in Brazilian Northeast. Renew Energy 2021. https://doi.org/10.1016/j.renene.2020.10.126.
[52] wang shan, shuai hui, Liu Q. Phase Space Reconstruction Driven Spatio-Temporal Feature Learning for Dynamic Facial Expression Recognition. IEEE Trans Affect Comput 2020. https://doi.org/10.1109/TAFFC.2020.3007531.
[53] Kaur S, Awasthi LK, Sangal AL, Dhiman G. Tunicate Swarm Algorithm: A new bio-inspired based metaheuristic paradigm for global optimization. Eng Appl Artif Intell 2020. https://doi.org/10.1016/j.engappai.2020.103541.
[54] Cui M, Krishnan V, Hodge BM, Zhang J. A Copula-Based Conditional Probabilistic Forecast Model for Wind Power Ramps. IEEE Trans Smart Grid 2019. https://doi.org/10.1109/TSG.2018.2841932.
[55] Quan H, Srinivasan D, Khosravi A. Uncertainty handling using neural network-based prediction intervals for electrical load forecasting. Energy 2014. https://doi.org/10.1016/j.energy.2014.06.104.
[56] Jiang Z, Che J, Wang L. Ultra-short-term wind speed forecasting based on EMD-VAR model and spatial correlation. Energy Convers Manag 2021;250:114919. https://doi.org/10.1016/J.ENCONMAN.2021.114919.
[57] da Silva RG, Ribeiro MHDM, Moreno SR, Mariani VC, Coelho L dos S. A novel decomposition-ensemble learning framework for multi-step ahead wind energy forecasting. Energy 2021. https://doi.org/10.1016/j.energy.2020.119174.
[58] Diebold FX, Mariano RS. Comparing predictive accuracy. J Bus Econ Stat 1995. https://doi.org/10.1080/07350015.1995.10524599.
[59] Song J, Wang J, Lu H. A novel combined model based on advanced

optimization algorithm for short-term wind speed forecasting. Appl Energy 2018;215:643–58. https://doi.org/10.1016/j.apenergy.2018.02.070.

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

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:

Table 1

Advantages and disadvantages of wind speed forecasting models.

Models	References	Advantages	Disadvantages
ARMA, ARIMA, and Kalman filtering	[6, 7]	The model is simple and only needs endogenous variables; Accurately forecast the linear sequences.	Low forecasting accuracy in nonlinear data; The data is required to be stable or differentially stable.
AI Model (BPNN, ELM, SVR, LStm, and GRu)	[9–11], [37]	Strong robustness and fault tolerance to noise data; Have the ability of association, and can approximate any nonlinear relationship.	The calculation burden is high, and the interpretability is poor; It is difficult to determine the hyperparameter values.
Ensemble Model	[34], [38,39]	The forecasting accuracy on different data types can be ensured; Take advantages of each single model.	Need to train multiple models and choose efficient empowerment technique.
MVE	[18,40]	The computational burden is relatively small.	The accuracy is largely affected by the effect of numerical predictions associated with it; the underestimation of data variance will result in low coverage of real data by prediction intervals.
Bootstrap	[41,42]	High efficiency in small-scale data.	Is a resampling method that requires significant computational cost for large data sets.
Bayesian	[20]	Improve the generalization ability of model.	The calculation burden is large, which requires the calculation of the Hessian matrix. When the data size is not large enough, the accuracy largely depends on prior knowledge.
LUBE	[21,43]	It avoids the problem of numerical calculation of the Jacobian matrix and Hessian matrix.	Heavy computational burden. No suitable parameter initialization method.
Quantile Regression (Qr)	[25,44]	Ability to resolve heterogeneity issues; Tail features of the distribution can be captured.	Traditional Qr model can't solve nonlinear problems, so it is necessary to select a suitable neural network to combine with Qr.

Table 2

Interval prediction metric values of single models with PINC=90% and PINC=95%.

Dataset	Models	PICP		PINAW		AIS	
		PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%
Site 1	**SSa-PSr-QrLASso**	0.9556	0.9778	0.4554	0.5341	-0.5439	-0.3055
	SSa-PSr-QrLStm	**1.0000**	**1.0000**	0.5160	0.5835	-0.5152	-0.2928
	SSa-PSr-QrGRu	**1.0000**	**1.0000**	0.5337	0.5986	-0.5334	-0.2964
	SSa-PSr-QrCNN	0.9611	0.8944	0.3147	0.7572	-0.3726	-0.4743
	SSa-PSr-QrBiLStm	1.0000	**1.0000**	0.3489	0.4249	-0.3500	-0.2086
	SSa-PSr-GPr	0.9944	**1.0000**	0.4457	0.5311	-0.4557	-0.2712
	SSa-PSr-BLgm	**1.0000**	**1.0000**	0.4460	0.5317	-0.4556	-0.2716
	Proposed EPFS	0.9833	0.9944	**0.1893**	**0.2674**	**-0.1805**	**-0.1273**
Site 2	**SSa-PSr-QrLASso**	0.8611	0.9111	0.3028	0.4918	-0.4409	-0.3389
	SSa-PSr-QrLStm	0.9778	0.9833	0.4770	0.5417	-0.5063	-0.2884
	SSa-PSr-QrGRu	0.9778	0.9889	0.4710	0.5729	-0.4905	-0.2930
	SSa-PSr-QrCNN	0.9833	**1.0000**	0.3734	1.1707	-0.3954	-0.5868
	SSa-PSr-QrBiLStm	**1.0000**	**1.0000**	0.3734	0.3494	-0.3802	-0.1748
	SSa-PSr-GPr	**1.0000**	**1.0000**	0.4547	0.5419	-0.4685	-0.2791
	SSa-PSr-BLgm	**1.0000**	**1.0000**	0.4555	0.5428	-0.4695	-0.2797
	Proposed EPFS	0.9778	0.9889	**0.1620**	**0.2471**	**-0.1590**	**-0.1216**
Site 3	**SSa-PSr-QrLASso**	0.8944	0.9889	0.4122	0.6835	-0.4157	-0.3049
	SSa-PSr-QrLStm	0.9889	**1.0000**	0.5568	0.7498	-0.4633	-0.3110
	SSa-PSr-QrGRu	0.9833	**1.0000**	0.5502	0.7138	-0.4539	-0.2933
	SSa-PSr-QrCNN	0.9111	0.9222	0.4586	0.4362	-0.4508	-0.2351
	SSa-PSr-QrBiLStm	**1.0000**	**1.0000**	0.3384	0.4537	-0.2771	-0.1835
	SSa-PSr-GPr	0.9889	**1.0000**	0.5029	0.5992	-0.4248	-0.2519
	SSa-PSr-BLgm	0.9944	**1.0000**	0.5080	0.6047	-0.4286	-0.2544
	Proposed EPFS	0.9833	0.9889	**0.2180**	**0.3306**	**-0.1784**	**-0.1324**

Note: The table above presents the reliability, resolution and comprehensive information of intervals obtained by different models. When PICP>PINC, the forecasting intervals are reliable. When the intervals are reliable, the narrower the interval width, the better the interval forecasting effect, which can be measured by $PINAW(\alpha)=\frac{1}{NR}\sum_{i=1}^{N}\left[UB_i(\alpha)-LB_i(\alpha)\right]$. Metric $AIS(\alpha)=\sum_{i=1}^{N}S^{(\alpha)}$ can comprehensively evaluate the interval forecasting effect.

Table 3

Interval prediction metric values of ensemble models with PINC=90% and PINC=95%.

Dataset	Models	PICP		PINAW		AIS	
		PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%
Site 1	**SSa-PSr-MOTa-EPFM**	0.9833	1.0000	0.2048	0.2828	-0.1997	-0.1353
	SSa-PSr-MODa-EPFM	0.9778	1.0000	0.1979	0.3413	-0.1899	-0.1666
	SSa-PSr-MOGa-EPFM	0.9167	0.9611	**0.1755**	0.2414	-0.1887	-0.1223
	SSa-PSr-MOAa-EPFM	0.9333	1.0000	0.1769	0.2805	-0.1843	-0.1336
	Emd-PSr-EPFM	0.9778	1.0000	0.2538	0.3903	-0.2527	-0.1892
	Eemd-PSr-EPFM	0.9889	1.0000	0.2276	0.3824	-0.2226	-0.1845
	Wt-PSr-EPFM	0.9944	0.9944	0.3513	0.4031	-0.3453	-0.1941
	CeemdAN-PSr-EPFM	**1.0000**	**1.0000**	0.2868	0.3848	-0.2795	-0.1832
	Proposed EPFS	0.9833	0.9944	0.1893	**0.2674**	**-0.1805**	**-0.1273**
Site 2	**SSa-PSr-MOTa-EPFM**	0.9778	0.9556	0.1692	0.2476	-0.1627	-0.1291
	SSa-PSr-MODa-EPFM	0.8833	0.9500	**0.1454**	0.2512	-0.1569	-0.1364
	SSa-PSr-MOGa-EPFM	0.8833	0.9500	0.1570	**0.2342**	-0.1707	-0.1248
	SSa-PSr-MOAa-EPFM	0.8500	0.8778	0.1475	0.2385	-0.1737	-0.1518
	Emd-PSr-EPFM	**1.0000**	**1.0000**	0.2921	0.3371	-0.2887	-0.1633
	Eemd-PSr-EPFM	0.9944	0.9944	0.3027	0.3335	-0.2991	-0.1612
	Wt-PSr-EPFM	**1.0000**	**1.0000**	0.3577	0.3980	-0.3523	-0.1945
	CeemdAN-PSr-EPFM	0.9944	0.9778	0.2517	0.2844	-0.2505	-0.1445
	Proposed EPFS	0.9778	0.9889	0.1620	0.2471	**-0.1590**	**-0.1216**
Site 3	**SSa-PSr-MOTa-EPFM**	0.9667	0.9833	0.2489	0.3605	-0.2035	-0.1465
	SSa-PSr-MODa-EPFM	0.8833	0.9722	**0.2170**	0.3337	-0.1810	-0.1378
	SSa-PSr-MOGa-EPFM	0.9500	0.9611	0.2003	0.3431	-0.1846	-0.1448
	SSa-PSr-MOAa-EPFM	0.9833	0.9889	0.2172	**0.3179**	-0.1876	-0.1345
	Emd-PSr-EPFM	0.9611	0.9778	0.3240	0.3798	-0.2707	-0.1572
	Eemd-PSr-EPFM	0.9778	0.9833	0.3300	0.3843	-0.2692	-0.1581
	Wt-PSr-EPFM	0.9667	0.9833	0.3726	0.4324	-0.3200	-0.1873
	CeemdAN-PSr-EPFM	**0.9889**	**0.9944**	0.3253	0.3935	-0.2650	-0.1583
	Proposed EPFS	0.9833	0.9889	0.2180	0.3306	**-0.1784**	**-0.1324**

Table 4

Statistics of DM test for statistical comparison of proposed approach versus other models

Model	Site 1		Site 2		Site 3	
	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%
SSa-PSr-QrLASso	-10.2893***	-6.7006***	-2.9187***	-3.8964***	-19.8818***	-4.2333***
SSa-PSr-QrLStm	-68.6604***	-66.8613***	-4.5813***	-6.6625***	-42.4881***	-23.2736***
SSa-PSr-QrGRu	-45.0164***	-65.5619***	-11.1758***	-7.9645***	-35.1313***	-20.3249***
SSa-PSr-QrCNN	-4.0386***	-4.5426***	-41.1140***	-9.2481***	-2.9096***	-7.3515***
SSa-PSr-QrBiLStm	-49.3264***	-48.2301***	-2.4558**	-33.6391***	-14.4810***	-9.2998***
SSa-PSr-GPr	-49.1107***	-59.7660***	-11.7846***	-44.0720***	-30.8004***	-21.7912***
SSa-PSr-BLgm	-48.4703***	-60.4071***	-11.8552***	-44.0617***	-31.3836***	-23.5181***
Emd-PSr-EPFM	-30.5495***	-5.8206***	-1.7031*	-17.1494***	-3.0018***	-8.0663***
Eemd-PSr-EPFM	-28.5660***	-13.4367***	-1.6236	-21.1340***	-3.1871***	-6.6468***
Wt-PSr-EPFM	-23.4755***	-28.0409***	-4.0643***	-23.3569***	-2.2902**	-3.8406***
CeemdAN-PSr-EPFM	-18.1712***	-32.4127***	-1.6707*	-9.6597***	-5.0214***	-6.5898***
SSa-PSr-MOTa-EPFM	-7.2410***	-5.0958***	-1.3035	0.3255	-6.5390***	-3.5052***
SSa-PSr-MODa-EPFM	-39.8391***	-5.0069***	-1.2672	-0.6341	-1.7147*	-1.0823
SSa-PSr-MOGa-EPFM	-0.6505	-1.6847*	-1.4715	-2.6176***	-1.9806**	-1.9018*
SSa-PSr-MOAa-EPFM	-2.0946**	-1.5301	-1.7845*	-2.7302***	-1.3332	-1.5073

Note: ***, **, and * indicate that the results are significant at the 1%, 5%, and 10% confidence levels, respectively.

Table 5
The forecasting effectiveness of the proposed EPFS and other existing models.

Model	**Site 1**				**Site 2**				**Site 3**			
	PINC=90%		**PINC=95%**		**PINC=90%**		**PINC=95%**		**PINC=90%**		**PINC=95%**	
	1st	**2nd**	**1st**	**2nd**	**1st**	**2nd**	**1st**	**2nd**	**1st**	**2nd**	**1st**	**2nd**
SSa-PSr-QrLASso	0.4561	0.3296	0.6945	0.6149	0.5591	0.3384	0.6611	0.4336	0.5843	0.4315	0.6951	0.6476
SSa-PSr-QrLStm	0.4848	0.4591	0.7072	0.6889	0.4937	0.3964	0.7116	0.6228	0.5367	0.4978	0.6890	0.6643
SSa-PSr-QrGRu	0.4666	0.4407	0.7036	0.6714	0.5095	0.4227	0.7070	0.6611	0.5461	0.4979	0.7067	0.6758
SSa-PSr-QrCNN	0.6274	0.4947	0.5257	0.3193	0.6046	0.5458	0.4132	0.3823	0.5492	0.4249	0.7649	0.6015
SSa-PSr-GPr	0.5443	0.5407	0.7288	0.7285	0.5315	0.5311	0.7209	0.7206	0.5707	0.5574	0.7481	0.7472
SSa-PSr-BLgm	0.5444	0.5426	0.7284	0.7271	0.5305	0.5287	0.7203	0.7188	0.5734	0.5532	0.7456	0.7443
SSa-PSr-QrBiLStm	0.6500	0.6131	0.7914	0.7577	0.6198	0.5771	0.8252	0.7941	0.7229	0.6775	0.8165	0.7815
Emd-PSr-EPFM	0.7473	0.6834	0.8108	0.7721	0.7113	0.6645	0.8367	0.7951	0.7293	0.6630	0.8428	0.7864
Eemd-PSr-EPFM	0.7774	0.7330	0.8155	0.7761	0.7009	0.6523	0.8388	0.7960	0.7308	0.6770	0.8419	0.7839
Wt-PSr-EPFM	0.6547	0.6034	0.8059	0.7627	0.6477	0.5934	0.8055	0.7660	0.6800	0.5679	0.8127	0.7041
CeemdAN-PSr-EPFM	0.7205	0.6742	0.8168	0.7691	0.7495	0.7019	0.8555	0.7854	0.7351	0.6874	0.8417	0.8008
SSa-PSr-MOTa-EPFM	0.8003	0.7309	0.8647	0.8266	0.8373	**0.7728**	0.8710	0.7661	0.7965	0.7357	0.8535	0.8073
SSa-PSr-MODa-EPFM	0.8101	0.7372	0.8334	0.8017	**0.8431**	0.7561	0.8636	0.7278	0.8190	0.7211	0.8622	0.7936
SSa-PSr-MOGa-EPFM	0.8113	0.7006	**0.8777**	0.8111	0.8293	0.7378	0.8752	0.7773	0.8154	0.6985	0.8552	0.7766
SSa-PSr-MOAa-EPFM	0.8157	0.7135	0.8578	0.7401	0.8264	0.7228	0.8482	0.6991	0.8124	0.7082	0.8655	0.7893
Proposed EPFS	**0.8195**	**0.7548**	0.8727	**0.8302**	0.8410	0.7660	**0.8784**	**0.8072**	**0.8216**	**0.7477**	**0.8676**	**0.8232**

Note: This table shows the circumstantial values of the *first*-order and *second*-order ***FE*** of sixteen models. The *first*-order ***FE*** is defined as $\boldsymbol{FE}\left(\boldsymbol{g}^1\right)=\boldsymbol{g}^1$. When this continues function contains two variables, the *2nd*-order ***FE*** can be denoted by $\boldsymbol{FE}\left(\boldsymbol{g}^1,\boldsymbol{g}^2\right)=\boldsymbol{g}^1\left(1-\sqrt{\boldsymbol{g}^2-\left(\boldsymbol{g}^1\right)^2}\right)$.

Table 6

The improvement ratio of the proposed EPFS.

Models	Site 1				Site 2				Site 3			
	PINAW		AIS		PINAW		AIS		PINAW		AIS	
	PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%	PINC=90%	PINC=95%
SSa-PSr-QrLASso	0.5844	0.4994	0.6682	0.5834	0.4650	0.4976	0.6394	0.6412	0.4712	0.5164	0.5708	0.5657
SSa-PSr-QrLStm	0.6332	0.5417	0.6498	0.5653	0.6604	0.5439	0.6860	0.5784	0.6085	0.5591	0.6149	0.5743
SSa-PSr-QrGRu	0.6454	0.5533	0.6617	0.5706	0.6560	0.5687	0.6759	0.5850	0.6038	0.5369	0.6069	0.5486
SSa-PSr-QrCNN	0.3984	0.6469	0.5157	0.7317	0.5662	0.7890	0.5980	0.7928	0.5247	0.2421	0.6042	0.4370
SSa-PSr-GPr	0.5753	0.4966	0.6040	0.5307	0.6437	0.5440	0.6607	0.5643	0.5665	0.4483	0.5800	0.4745
SSa-PSr-BLgm	0.5756	0.4971	0.6039	0.5314	0.6443	0.5448	0.6614	0.5653	0.5709	0.4533	0.5837	0.4797
SSa-PSr-QrBiLStm	0.4574	0.3708	0.4844	0.3900	0.5661	0.2928	0.5819	0.3042	0.3558	0.2714	0.3562	0.2787
SSa-PSr-MOTa-EPFM	0.0758	0.0545	0.0965	0.0596	0.0423	0.0020	0.0229	0.0578	0.1243	0.0829	0.1231	0.0961
SSa-PSr-MODa-EPFM	0.0436	0.2166	0.0500	0.2360	-0.1143	0.0164	0.0132	0.1083	-0.0047	0.0093	0.0141	0.0395
SSa-PSr-MOGa-EPFM	-0.0787	-0.1075	0.0438	0.0407	-0.0317	-0.0550	0.0690	0.0256	-0.0883	0.0364	0.0334	0.0860
SSa-PSr-MOAa-EPFM	-0.0702	0.0466	0.0208	0.0471	-0.0983	-0.0360	0.0846	0.1988	-0.0038	-0.0399	0.0487	0.0157
Emd-PSr-EPFM	0.2542	0.3150	0.2859	0.3274	0.4453	0.2670	0.4494	0.2553	0.3273	0.1296	0.3408	0.1580
Eemd-PSr-EPFM	0.1685	0.3008	0.1894	0.3104	0.4647	0.2592	0.4686	0.2457	0.3394	0.1397	0.3372	0.1625
Wt-PSr-EPFM	0.4612	0.3368	0.4774	0.3442	0.5470	0.3792	0.5488	0.3747	0.4150	0.2354	0.4424	0.2930
CeemdAN-PSr-EPFM	0.3399	0.3051	0.3544	0.3052	0.3563	0.1311	0.3654	0.1584	0.3299	0.1599	0.3264	0.1639

Note: The table above reports the IR of the proposed EPFS from other twelve models. The AIS and PINAW are used to measure the IR, and the corresponding indicator can be defined as $\overline{IR}_{Metric} = \left[\left(Metric^{com} - Metric^{pro} \right) / Metric^{com} \right]$, where $Metric^{com}$ is the metric values of compared model, and the $Metric^{pro}$ indicates the metric value of the proposed **EPFS**.

Table A1

Parameters setting.

Models	Symbol	Meaning	Determination method	determined value
Eemd	Ns	STD of added noise	Preset	0.05
	NR	Realization Number	Preset	50
	Ms	Maximum Sifting Iteration	Preset	500
CeemdAN	Ns	STD of added noise	Preset	0.05
	NR	Realization Number	Preset	50
	Ms	Maximum Sifting Iteration	Preset	500
Wt	D1	Decomposition Layer Number	Trial and error approach	5
SSa	W	Window Length	Trial and error approach	50
	D2	Primary Ingredient Disintegration Number	Karhunene Loeve decomposition	20
QrLStm	Nl	Number of hidden layers	Trial and error approach	2
	Nn	Number of hidden nodes	Trial and error approach	50
QrBILStm	Nl	Number of hidden layers	Trial and error approach	1
	Nn	Number of hidden nodes	Trial and error approach	50
QrGRu	Nl	Number of hidden layers	Trial and error approach	2
	Nn	Number of hidden nodes	Trial and error approach	50
QrCNN	Ck	Convolution kernel size	Trial and error approach	2*2
	Nn	Number of hidden layer nodes	Trial and error approach	35
MOGa, MOAa, MODa	As	Archive Size	Preset	10
	In	Iteration Number	Preset	200
	Ni	Individual Number	Trial and error approach	40
MOTa, IMOTa	Si	Initial Speed	Preset	1
	Nt	Number of tunicate	Trial and error approach	40
	Ss	Subordinate Speed	Preset	4
	As	Archive Size	Trial and error approach	10
	In	Iteration Number	Trial and error approach	200
PSr	τ	Embedded dimension	C-C method	Site 1~3: 4, 5,5
	M	Delay time	C-C method	Site 1~3: 31, 28, 28

Table A2
List of abbreviations.

ARIMA	Auto Regressive Integrated Moving Average	**LStm**	Long short-term memory neural network
ARMA	Autoregressive moving average	**LUBE**	lower upper bound estimation
BiLStm	Bi-directional Long Short Term Memory Network	**MO**	Multi-objective
BLgm	Bayesian regression model	**MOAa**	MO Antlion Algorithm
BPNN	Back propagation neural network	**MODa**	MO Dragonfly Algorithm
CeemdAN	Complete ensemble Empirical Mode Decomposition with Adaptive Noise	**MOGa**	MO Grasshopper Algorithm
Eemd	Ensemble Empirical Mode Decomposition	**MVE**	mean-variance estimates
EFS	Exponential function steps	**PINC**	prediction interval nominal confidence
ELM	Extreme learning machine	**PSr**	Phase space reconstruction
Emd	Empirical Mode Decomposition	**Qr**	quantile regression
EOL	Elite opposition learning	**QrCNN**	Quantile regression convolution neural network
EPFM	Ensemble probabilistic forecasting model	**QrGRu**	Quantile regression gated recurrent unit
EPFS	Ensemble probabilistic forecasting system	**QrLStm**	Quantile regression Long Short-Term Memory Network
GPr	Gaussian Process Regression	**RNN**	Recurrent neural network
GRu	Gated recurrent unit	**SSa**	Singular Spectral Analysis
GW	Gigawatt	**SVQr**	Support Vector Quantile Regression
GWEC	Global wind energy councile	**SVR**	Support Vector Regression
IMOTa	Improved multi-objective tunicate swarm algorithm	**Wt**	Wavelet Transform

Fig.1

Click here to access/download;Figure;Fig.1.tif

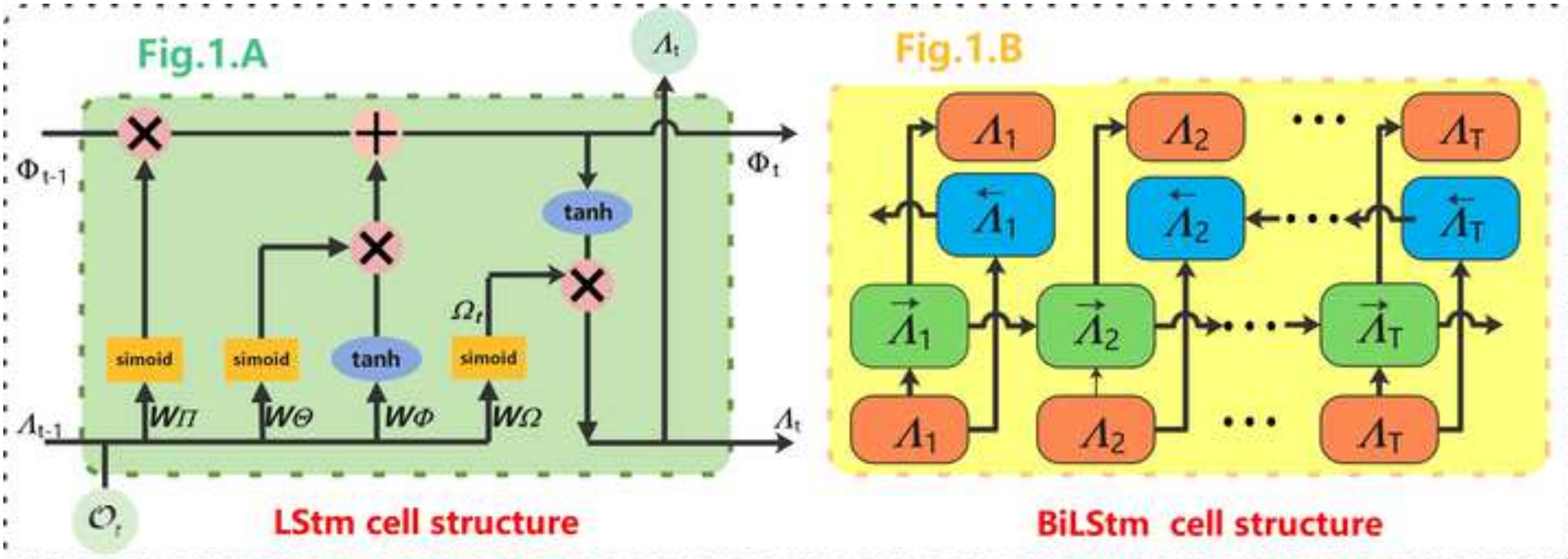


Fig.2

Click here to access/download;Figure;Fig.2.tif

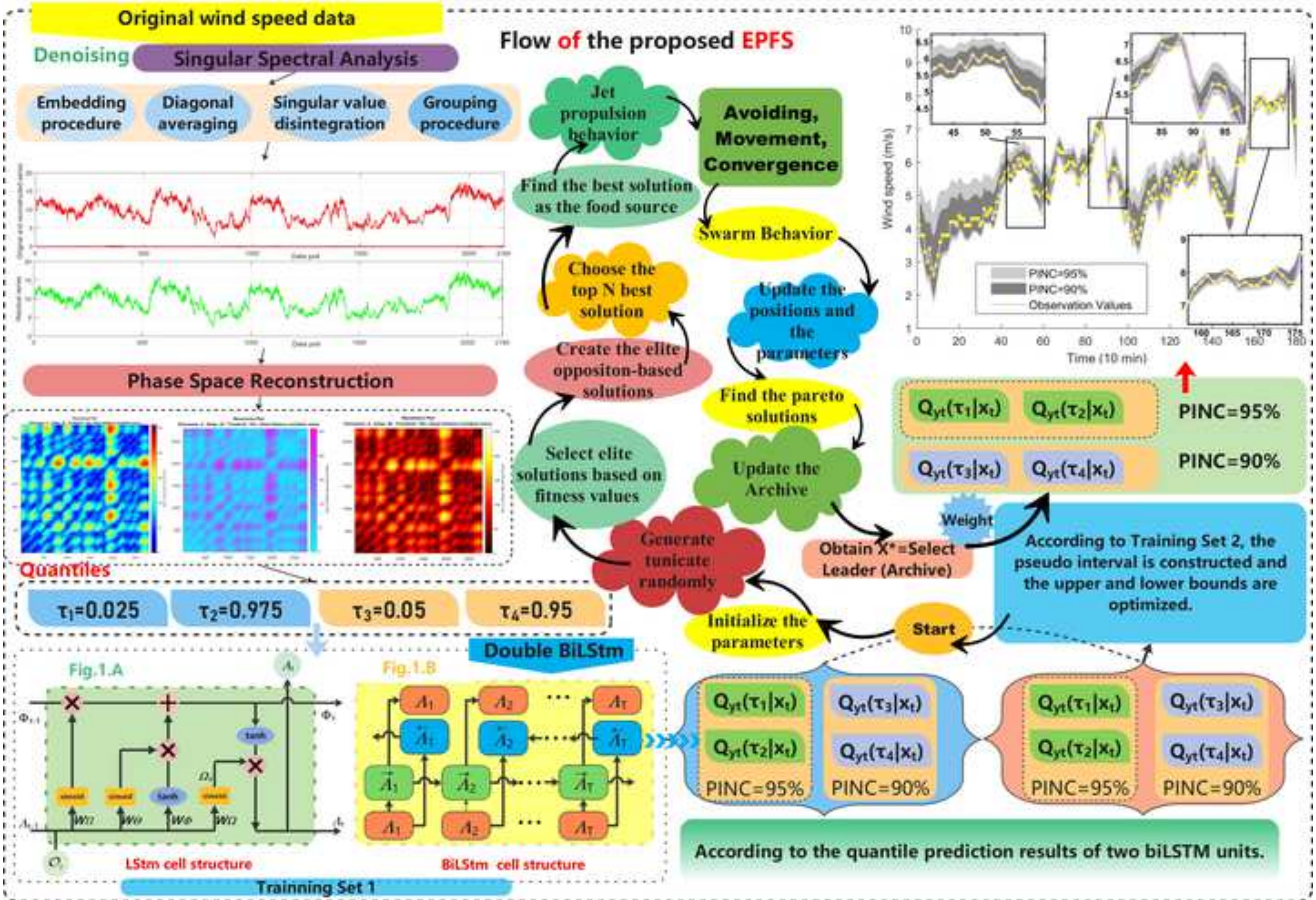


Fig.3 Click here to access/download;Figure;Fig.3.tif

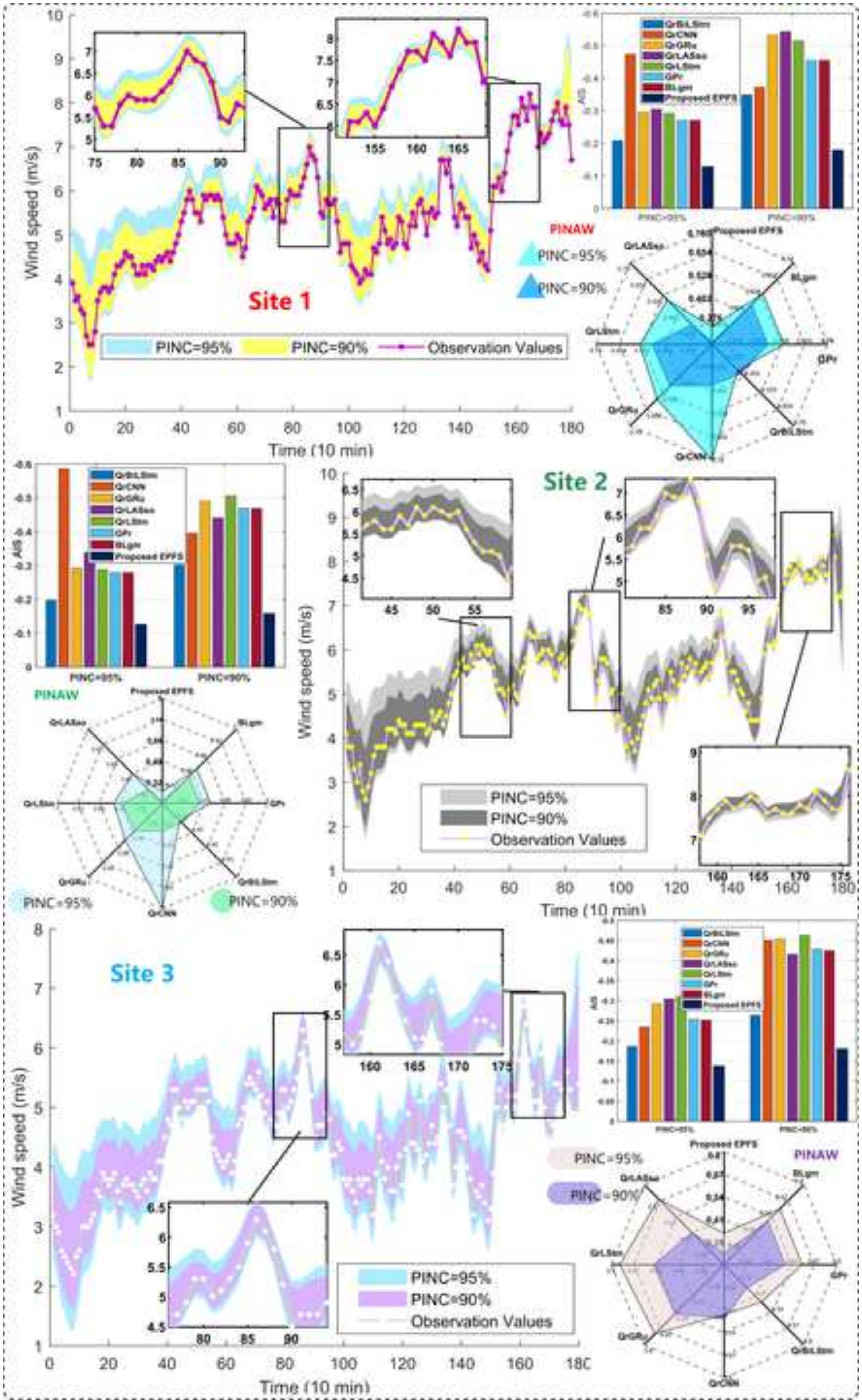


Fig.4

Click here to access/download;Figure;Fig.4.tif

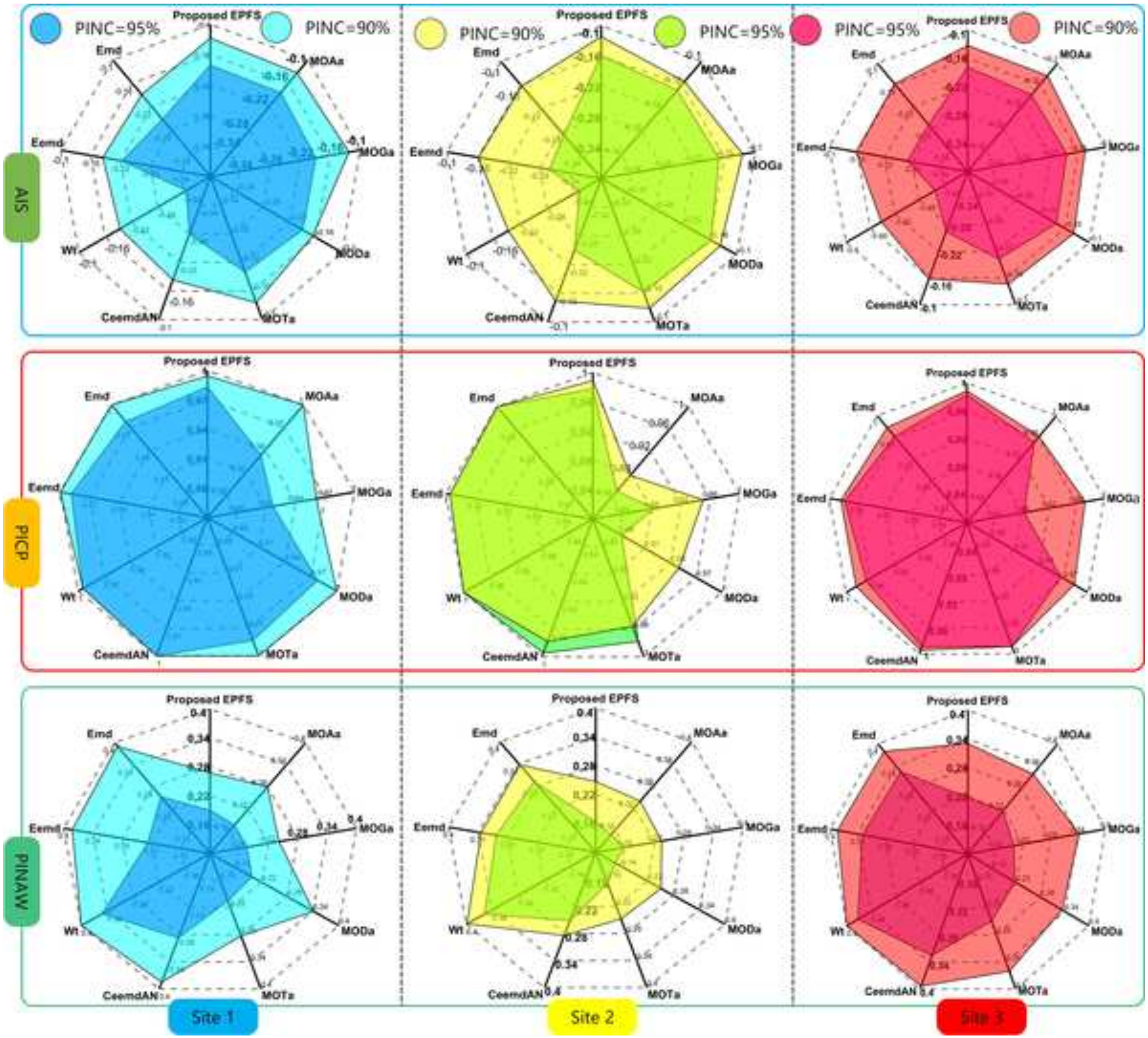
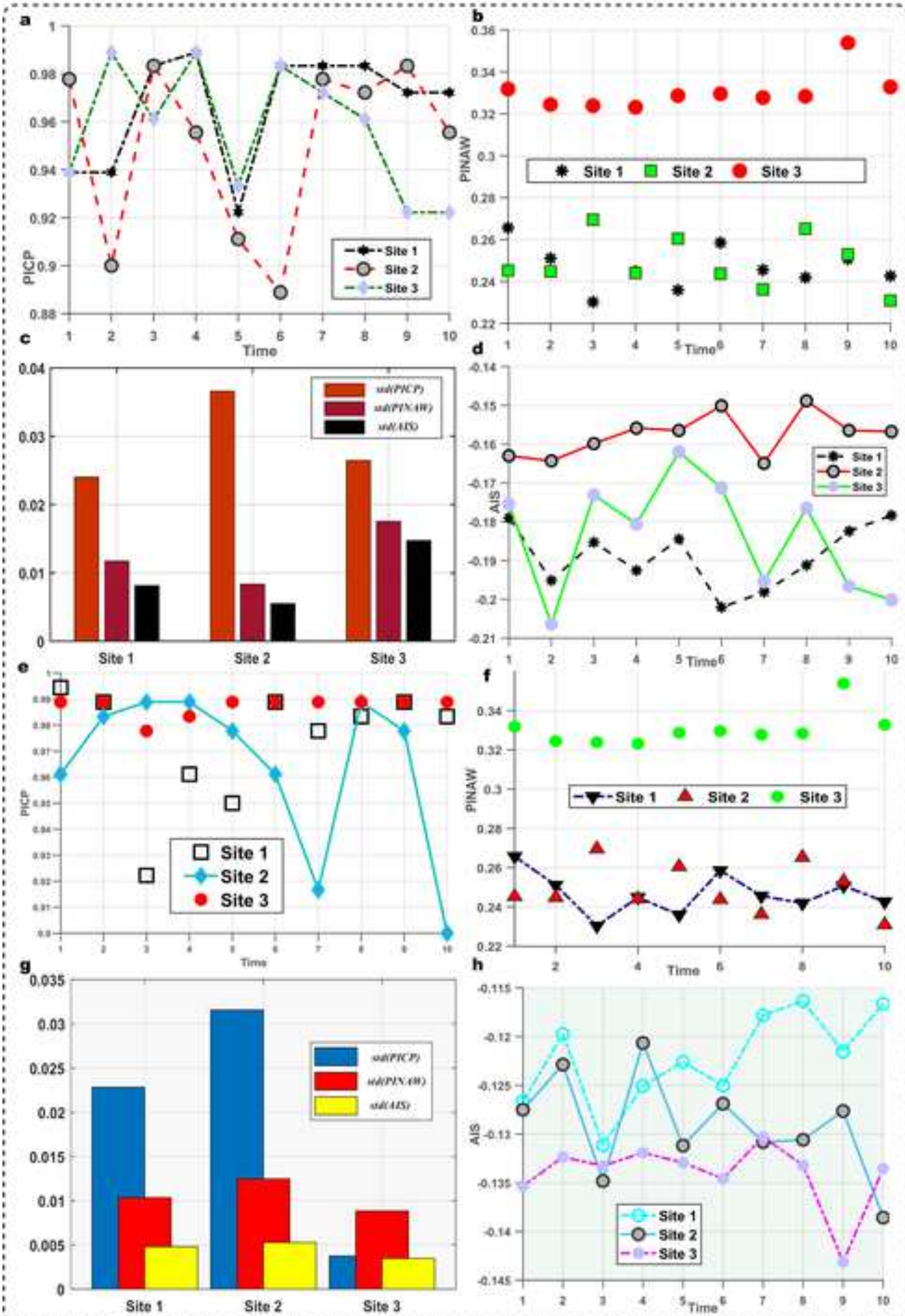


Fig.5



a
PICP
Time
Site 1
Site 2
Site 3
b
PINAW
Time
Site 1
Site 2
Site 3
c
std(PICP)
std(PINAW)
std(AIS)
Site 1
Site 2
Site 3
d
AIS
Time
Site 1
Site 2
Site 3
e
PICP
Time
Site 1
Site 2
Site 3
f
PINAW
Time
Site 1
Site 2
Site 3
g
std(PICP)
std(PINAW)
std(AIS)
Site 1
Site 2
Site 3
h
AIS
Time
Site 1
Site 2
Site 3