frontiers | Research Topics

The impact of complex terrain on wind power output and the mechanism of improving prediction accuracy

Edited by
Linfei Yin, Lorenzo Ferrari, Hui Hwang Goh,
Unai Fernandez-Gamiz, Xun Shen, Lefeng Cheng
and David Carvalho

Published in
Frontiers in Energy Research

FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-5381-7
DOI 10.3389/978-2-8325-5381-7

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

The impact of complex terrain on wind power output and the mechanism of improving prediction accuracy

Topic editors
Linfei Yin — Guangxi University, China
Lorenzo Ferrari — University of Pisa, Italy
Hui Hwang Goh — Guangxi University, China
Unai Fernandez-Gamiz — University of the Basque Country, Spain
Xun Shen — Osaka University, Japan
Lefeng Cheng — Guangzhou University, China
David Carvalho — University of Aveiro, Portugal

Citation
Yin, L., Ferrari, L., Goh, H. H., Fernandez-Gamiz, U., Shen, X., Cheng, L., Carvalho, D., eds. (2025). *The impact of complex terrain on wind power output and the mechanism of improving prediction accuracy*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-5381-7

Table of contents

frontiers | Frontiers in Energy Research

TYPE Editorial
PUBLISHED 20 August 2024
DOI 10.3389/fenrg.2024.1468799

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
David Howe Wood,
University of Calgary, Canada

*CORRESPONDENCE
Lefeng Cheng,
✉ chenglefeng@gzhu.edu.cn

RECEIVED 22 July 2024
ACCEPTED 05 August 2024
PUBLISHED 20 August 2024

CITATION
Cheng L and Yin L (2024) Editorial: Impacts of complex terrain on wind power output and mechanisms to improve prediction accuracy.
Front. Energy Res. 12:1468799.
doi: 10.3389/fenrg.2024.1468799

COPYRIGHT
© 2024 Cheng and Yin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Impacts of complex terrain on wind power output and mechanisms to improve prediction accuracy

Lefeng Cheng[1]* and Linfei Yin[2]

[1]School of Mechanical and Electrical Engineering, Guangzhou University, Guangzhou, Guangdong, China, [2]College of Electrical Engineering, Guangxi University, Nanning, China

KEYWORDS

accurate wind power forecasting, renewable-energy grid connection and consumption, wind turbine parameter optimization, data-driven approach, economic scheduling considering wind power fluctuations

Editorial on the Research Topic
Impacts of complex terrain on wind power output and mechanisms to improve prediction accuracy

Random and intermittent shocks generated by large-scale wind farms continue to affect the safe and stable operations of power systems significantly (Hong et al., 2024; Li et al., 2024); hence, it is necessary to investigate the impacts of complex terrains on wind power outputs and mechanisms to improve their prediction accuracies. To this end, we host this Research Topic that contains nine final articles. Among these works, Kristianti et al. investigated the influences of air flow features on alpine wind energy potential. Man et al. proposed a multidevice wind turbine power generation forecasting model aimed at wind farms. Zhang et al. present a meta reservoir computing method while Konstantinou and Hatziargyriou establish a model combining convolutional neural networks and DeepSHAP to enhance the accuracy of wind power forecasting. Wang et al. present an incremental feedforward collective pitch control method for the wind turbine. Zhou et al. propose an interval model for the wind turbine power curve. Xu et al. survey some energy management strategies for a loop microgrid with wind energy prediction and energy storage systems. Wang and Liao propose a short-term hybrid prediction model for wind speed prediction. Finally, Gao et al. present a detailed review of the interval reservoir computing approach and examine some case studies. Overall, these articles cover a wide range of research topics and provide highly valuable research methods and models that are expected to serve as excellent references for researchers working on related research topics, particularly those related to the impacts of complex terrains on wind power outputs and mechanisms to improve their prediction accuracies.

Author contributions

LC: formal analysis, funding acquisition, project administration, supervision, writing–original draft, and writing–review and editing. LY: formal analysis, validation, writing–original draft, and writing–review and editing.

Funding

The authors declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Guangdong Basic and Applied Basic Research Foundation (no. 2022A1515010699).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations or those of the publisher, editors, and reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.

References

Hong, S., McMorland, J., Zhang, H., Collu, M., and Halse, K. H. (2024). Floating offshore wind farm installation, challenges and opportunities: a comprehensive survey. *Ocean. Eng.*, 304, 117793. doi:10.1016/j.oceaneng.2024.117793

Li, B., Dong, Y., Jiao, X., Chen, X., Li, B., and Ji, L. (2024). Study on the calculation method of electrical quantity for connection line open-phase operation of wind farm connected to MMC-HVDC Considering negative sequence current suppression. *Int. J. Electr. Power and Energy Syst.*, 159, 110056.doi:10.1016/j.ijepes.2024.110056

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 21 November 2023
DOI 10.3389/fenrg.2023.1298088

A short-term hybrid wind speed prediction model based on decomposition and improved optimization algorithm

Lu Wang[1,2] and Yilan Liao[1]*

[1]State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China, [2]School of Sciences, Guangxi University of Science and Technology, Liuzhou, China

OPEN ACCESS

EDITED BY
Linfei Yin,
Guangxi University, China

REVIEWED BY
Ning Li,
Xi'an University of Technology, China
Mehdi Neshat,
University of South Australia, Australia

*CORRESPONDENCE
Yilan Liao,
✉ liaoyl@lreis.ac.cn

RECEIVED 21 September 2023
ACCEPTED 30 October 2023
PUBLISHED 21 November 2023

CITATION
Wang L and Liao Y (2023), A short-term hybrid wind speed prediction model based on decomposition and improved optimization algorithm.
Front. Energy Res. 11:1298088.
doi: 10.3389/fenrg.2023.1298088

COPYRIGHT
© 2023 Wang and Liao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Introduction: In the field of wind power generation, short-term wind speed prediction plays an increasingly important role as the foundation for effective utilization of wind energy. However, accurately predicting wind speed is highly challenging due to its complexity and randomness in practical applications. Currently, single algorithms exhibit poor accuracy in short-term wind speed prediction, leading to the widespread adoption of hybrid wind speed prediction models based on deep learning techniques. To comprehensively enhance the predictive performance of short-term wind speed models, this study proposes a hybrid model, VMDAttention LSTM-ASSA, which consists of three stages: decomposition of the original wind speed sequence, prediction of each mode component, and weight optimization.

Methods: To comprehensively enhance the predictive performance of short-term wind speed models, this study proposes a hybrid model, VMDAttention LSTM-ASSA, which consists of three stages: decomposition of the original wind speed sequence, prediction of each mode component, and weight optimization. Firstly, the model incorporates an attention mechanism into the LSTM model to extract important temporal slices from each mode component, effectively improving the slice prediction accuracy. Secondly, two different search operators are introduced to enhance the original Salp Swarm Algorithm, addressing the issue of getting trapped in local optima and achieving globally optimal short-term wind speed predictions.

Result: Through comparative experiments using multiple-site short-term wind speed datasets, this study demonstrates that the proposed VMD-AtLSTM-ASSA model outperforms other hybrid prediction models (VMD-RNN, VMD-BPNN, VMD-GRU, VMD-LSTM) with a maximum reduction of 80.33% in MAPE values. The experimental results validate the high accuracy and stability of the VMD-AtLSTM-ASSA model.

Discussion: Short-term wind speed prediction is of paramount importance for the effective utilization of wind power generation, and our research provides strong support for enhancing the efficiency and reliability of wind power generation systems. Future research directions may include further improvements in model performance and extension into other meteorological and environmental application domains.

KEYWORDS

variational modal decomposition, attention, long short-term memory, salp swarm algorithm, short-term wind speed prediction

1 Introduction

Wind energy plays an important role in many new energy sources. According to the latest report released by the Global Wind Energy Council (GWEC) (Guliyev, 2020), the global installed capacity of wind power will reach 743 GW in 2020, with a 53% year-on-year growth in new installations. However, the stochastic, fluctuating and intermittent nature of wind farms poses significant challenges to the operation and control of the entire power system including wind farms (Lacal-Arantegui, 2019). Among them, short-term wind speed prediction is an indispensable factor for the development of daily scheduling plans. Therefore, proposing a method to accurately predict the short-term wind speed has an important impact on the economic and reliable operation of the power system (Rizwan-ul-Hassan et al., 2021).

Currently, short-term wind speed prediction methods are divided into two main categories: physical process-driven models (Higashiyama et al., 2018) and data-driven models (Yuan et al., 2017). Data-driven models are divided into statistical models (Liu et al., 2010) and artificial intelligence models (Khodayar et al., 2017). Physical process-driven models are mostly numerical weather prediction (NWP) models (Lowery and O'Malley, 2012), which make predictions based on local environmental information, such as, temperature, humidity, and geography. These methods are usually time-consuming and unsuitable for short-term and ultrashort-term wind speed forecasting due to excessive model considerations and model over-complexity (Wang and Li, 2018). In contrast, statistical models are more suitable for short-term wind speed forecasting. Statistical models learn the patterns of historical wind speed data and establish non-linear mapping relationships between the data, thus realizing time series forecasting (Rodrigues Moreno et al., 2020). Commonly used statistical methods are time series modeling (Liu et al., 2020b), Kalman filtering (Paliwal and Basu, 1987), Markov chain (Sahin and Sen, 2001), Bayesian method (Liu et al., 2020a) and so on. Statistical models have high prediction accuracy for static time series, but when facing highly nonlinear and complex wind speed data, these methods are less scalable and less effective in fitting.

In recent years, artificial intelligence models, including machine learning and deep learning models, have become increasingly popular in the field of short-term wind speed prediction (Scutaru et al., 2020). Compared with physical and statistical models, artificial intelligence models have greatly improved the accuracy of predicting wind speed. Among these, artificial neural network-based models seem to be the best choice because they can learn directly from historical data of wind speed without any a priori concepts and are more adaptable to practical applications (Tascikaraoglu and Uzunoglu, 2014). The most basic artificial neural network model is the back propagation neural network (BPNN) (Wang et al., 2015). Theoretically, as long as the number of neural units in the hidden layer of a BPNN reaches a certain number, then any nonlinear function can be fitted. However, BPNNs also have obvious shortcomings, firstly, it is easy to fall into the local optimum rather than obtaining the global optimal solution, and secondly, the learning efficiency caused by the need for too many trainings is low, and the convergence speed is not ideal. The other artificial neural network model, recurrent neural network (RNN) (Zaremba et al., 2015), is better at finding local correlations compared to BPNN. It can pass previous state information to neurons at the current time step. This mechanism allows the RNN to deal with dependencies in long sequences and also allows the same parameters to be shared between each time step, which gives it a smaller number of parameters and faster training speed, which fits well with the temporal continuum of wind speed prediction (Tanaka et al., 2015; Yu et al., 2018; Duan et al., 2021). However, the problem of backpropagation in the network architecture of RNNs leads to the problem of gradient vanishing and gradient explosion. This means that there are difficulties with very long sequences and the gradient decreases to near zero in hard-to-handle iterations. In order to overcome this problem, "gate control" techniques are used in RNN models, such as the long short-term memory (LSTM) (Hochreiter and Schmidhuber, 1997) and gate recurrent unit (GRU) (Niu et al., 2020), The GRU model adopts a simplified gating mechanism to prevent overfitting, but its prediction results are more logically correlated with recent time steps, which may lead to the loss of useful information from distant time steps. On the other hand, the LSTM model can effectively handle long-term dependencies, avoiding the issues of gradient vanishing or exploding. Experimental validation using multiple wind speed datasets has demonstrated the superior predictive performance of LSTM (Altan et al., 2021; Jaseena and Kovoor, 2021; Shahid et al., 2021). However, the computational structure of LSTM is relatively complex and it has a larger number of parameters, which could potentially lead to overfitting. Therefore, the key focus of research lies in effectively capturing important information based on the data conditions within the LSTM network, aiming to improve the prediction accuracy and robustness of wind speed forecasts.

Due to the distinct characteristics exhibited by various single models, hybrid models can effectively leverage the advantages of different individual models to achieve enhanced wind speed prediction performance. Consequently, hybrid prediction models based on decomposition and optimization have emerged as a research hotspot in the field of wind speed prediction in recent years. In order to ensure the sufficiency and integrity of feature decomposition and reconstruction, some scholars have proposed a novel hybrid model based on singular spectrum analysis and temporal convolutional attention network with adaptive receptive field (ARFTCAN). The results demonstrate that the proposed model effectively supports the adaptability of short-term wind power forecasting (WPF) across all four seasons (Shao et al., 2022). Furthermore, another group of researchers have introduced a wind speed prediction method that combines quaternion convolutional neural network (QCNN), Bi-LSTM, and adaptive decomposition techniques. This approach offers highly accurate forecasting results for long-term wind speed prediction (Neshat et al., 2022). Short-term hybrid wind speed prediction models usually include three steps: decomposition, prediction, and optimization (Ma et al., 2009). In the signal decomposition step, the unstable original wind speed sequence is decomposed into multiple IMF components with significant frequency characteristics by data decomposition methods, which reduces the complexity of the original data and performs noise reduction, e.g., empirical mode decomposition (EMD) (Ren et al., 2016) performs adaptive decomposition of nonlinear and highly fluctuating data in the original wind speed sequence to improve the prediction performance of wind speed prediction models (Naik et al., 2018). However, the EMD has the problems of large reconstruction error, poor decomposition completeness, and large noise residuals. Therefore, Literature (Hu et al., 2021) proposed a method using variational mode decomposition to mine the features of the wind speed sequence and eliminate the noise to predict each intrinsic mode function (IMF), which has obvious accuracy advantages over other decomposition methods in wind speed prediction.

In constructing the wind speed hybrid model, usually after signal decomposition of the data, a parameter optimization algorithm is also used to optimize the weights of each IMF to improve the performance of the prediction algorithm. Among the parameter optimization algorithms, the swarm intelligence optimization algorithm is the most commonly used algorithm for wind speed prediction. The swarm intelligence optimization algorithm is a number of algorithms proposed for solving optimization problems through the simulation study of the behavior of animal groups, which overcomes the limitations of the traditional algorithms when dealing with some complex problems such as, nonlinear, multi-constraint, multi-variable, etc., and demonstrates a better optimization ability. Some common ones are grey wolf optimizer (GWO) (Fu et al., 2019), differential evolution (DE) (Storn and Price, 1997), particle swarm optimization (PSO) (Kennedy, 2011), covariance matrix adaptation evolution strategy (CMAES) (Hansen and Ostermeier, 2001), whale optimization algorithm (WOA) (Mirjalili and Lewis, 2016), salp swarm algorithm (SSA) (Mirjalili et al., 2017), etc. These algorithms have their respective advantages, but the salp swarm algorithm (SSA), as an algorithm that achieves parameter optimization by simulating the behavior of salp populations, exhibits significant advantages in terms of parameter configuration, robustness, and convergence speed. For example, SSA only requires adjusting the position and velocity of salp individuals to update the search space. It utilizes information transmission and competition mechanisms among salp individuals to promote diversity and convergence during the search process (Mirjalili et al., 2017), achieving a balance between global and local search. Furthermore, SSA demonstrates superior robustness and faster convergence speed in solving complex optimization problems. However, like other heuristic algorithms, the algorithm also suffers from problems such as, a high likelihood of falling into local optimum, low optimization accuracy, and unstable solution results (Faris et al., 2018; Kang et al., 2019). Therefore, many scholars have improved the deficiencies of the salp swarm algorithm accordingly. For example, Literature (Faris et al., 2018) used adaptive operators to help the salp swarm algorithm break through the optimal local constraints in the process of follower position updating, so that the individual salp swarm has strong global convergence ability in the early stage, thus obtaining relatively accurate results in the later stage. Some researchers have also designed three new communication strategies, significantly improving the collaborative capability of SSA (Pan et al., 2021). Alternatively, starting from interval prediction, a novel prediction model based on wind speed distribution and multi-objective optimization is proposed by improving the SSA combination module (Wang and Cheng, 2021). The aforementioned studies by these scholars lay the foundation for the proposed multi-objective adaptive learning salp swarm algorithm (ASSA) in this paper.

In summary, this study proposes the VMD-Attention LSTM-ASSA (VMD-AtLSTM-ASSA) hybrid short-term wind speed prediction model containing decomposition, prediction, and optimization for short-term wind speed prediction. The variational mode decomposition (VMD), as a decomposition model in the hybrid model, decomposes the wind speed series data into a series of intrinsic mode functions (IMFs) that can adaptively update the optimal center frequency and bandwidth of each IMF component, which is helpful for the subsequent work of using the long short term memory networks (LSTM) prediction model to incorporate the attention mechanism effectively, which extracts the important slice information in each IMF component for high-precision prediction. Finally, the multi-objective adaptive learning rate salp swarm algorithm (ASSA) model is used to find the optimal weights for each IMF component, which is finally weighted to obtain the high-precision wind speed prediction value.

The main contributions and innovations of this paper are as follows.

(a) The use of long short term memory networks (LSTM) with the inclusion of an attention mechanism to individually predict the intrinsic mode functions (IMFs) obtained through variational mode decomposition (VMD). The Attention mechanism identifies the importance of slice information within each modal component, effectively improving the prediction accuracy and robustness of the LSTM network.
(b) On the basis of the salp swarm algorithm (SSA), improvements are made to address the problems of local optima trapping and premature convergence in the original salp swarm algorithm. This is achieved by proposing the adaptive learning operator and multi-objective operator in the multi-objective adaptive learning rate salp swarm optimization algorithm ASSA. Ultimately, this approach achieves global optimality and improves wind speed prediction accuracy.
(c) Through comprehensive comparisons with popular deep learning prediction models, decomposition models, and optimization models, this paper verifies the superiority of the proposed hybrid wind speed prediction model VMD-AtLSTM-ASSA in terms of individual components as well as overall predictive performance.

The structure of this paper is described as follows: Section 2 presents the algorithmic principles of the proposed model, including the model framework and execution process, and the model principles; Section 3 presents and discusses the case study; Section 4 gives the conclusions and future work.

2 Methodology

2.1 Overall framework and execution process of VMD-AtLSTM-ASSA

This section describes the framework structure of the proposed VMD-AtLSTM-ASSA combined model, and the specific flowchart is shown in Figure 1. The execution process of this study is in three phases, which are A. Wind speed sequence decomposition, B. Prediction of wind speed IMF components, and C. Weight optimization. In this study, the wind speed sequence decomposition stage utilizes the variational mode decomposition (VMD) model to decompose the complex original sequence into stable mode components, aiming to reduce the impact of non-stationarity and complexity of the original wind speed sequence on prediction accuracy. In the prediction stage, the LSTM model with attention mechanism (AtLSTM) is used to predict the wind speed from the decomposed IMF components. Since the predicted values of each IMF component are of differing importance to the actual values, the proposed multi-objective adaptive learning rate salp swarm algorithm (ASSA) algorithm is used to give the optimal weights to each component and then superimpose them to obtain the final highly accurate predicted values of wind speed.

2.2 Principle of VMD-AtLSTM-ASSA

2.2.1 Wind speed sequence decomposition—variational mode decomposition

The VMD is a signal decomposition method (Dragomiretskiy and Zosso, 2014), and the overall framework is a variational problem. That is, assuming that each "mode" is a finite bandwidth with different center frequencies, minimizing the sum of the estimated bandwidths of each mode becomes a problem. In order to solve this variational problem, the method adopts the alternating direction multiplier method, which constantly updates each mode and its center frequency, gradually demodulates each mode to the corresponding fundamental frequency band, and finally extracts each mode to the corresponding center frequency. Therefore, in this study, the VMD technique is employed to decompose the complex original wind speed sequence. The main objective is to decompose the original wind speed sequence, which exhibits nonlinearity and randomness, into a series of frequency-stable mode components, aiming to maximize the improvement in prediction accuracy. The specific process of VMD is as follows, and the results of the mode decomposition are shown in Figure 2.

Step 1: Assuming that each wind speed's intrinsic mode functions have a finite bandwidth with a center frequency, now find the decomposed wind speed modes such that the sum of the estimated bandwidths of each wind speed mode is minimized. The specific model is as follows:

$$\underset{\{u_k\}\{W_k\}}{min}\left\{\sum_k\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)^{*}u_k(t)\right]e^{-jW_kt}\right\|_2^2\right\} \tag{1}$$

$$s.t.\sum_k u_k(t)= f \tag{2}$$

Where, k is the number of modes to be decomposed (positive integer), u_k and w_k correspond to the k IMF and the center frequency of the decomposition, $\delta(t)$ is the Dirac function, and $*$ is the convolution operator.

Step 2: In order to solve the above model, introduce the penalty factor α (to reduce the effect of Gaussian noise) and Lagrange multiplier operator, transform the constrained problem into an unconstrained problem, and get the generalized Lagrange expression: the above equation constrained problem is equivalent to an unconstrained optimization problem through the generalized Lagrange function, and the mathematical formulas are as follows:

$$\begin{aligned}L(\{u_k\},\{W_k\},\lambda):\ &= a\sum_k\left\|\partial\left[\left(\delta(t)+\tfrac{j}{\pi t}\right)^{*}u_k(t)\right]e^{-jW_kt}\right\|_2^2\\&+\left\|f(t)-\sum_k u_k(t)\right\|_2^2+\left\langle\lambda(t),f(t)-\sum_k u_k(t)\right\rangle\end{aligned} \tag{3}$$

Step 3: Iteratively update the parameters, u_k w_k and λ by multiplier alternating direction method with the following equation.

$$\hat{u}_k^{n+1}(w)=\frac{\hat{f}(w)-\sum_{i\neq k}\hat{u}_i(w)+\frac{\hat{\lambda}(w)}{2}}{1+2\alpha(w-w_k)^2} \tag{4}$$

$$w_k^{n+1}=\frac{\int_0^{\infty}w\left|\hat{u}_k^{n+1}(w)\right|^2dw}{\int_0^{\infty}\left|\hat{u}_k^{n+1}(w)\right|^2dw} \tag{5}$$

$$\hat{\lambda}^{n+1}(w)=\hat{\lambda}^{n}(w)+\gamma\left(\hat{f}(w)-\sum_k\hat{u}_k^{n+1}(w)\right) \tag{6}$$

Where γ is the noise tolerance, which meets the fidelity requirement of signal decomposition; n is the number of iterations; $\hat{u}_k^{n+1}(w)$, $\hat{u}_i^{n+1}(w)$, $\hat{f}(w)$, $\hat{\lambda}(w)$ correspond to the Fourier transforms of $u_k^{n+1}(t)$, $u_i(t)^{n=1}$, $f(t)$, $\lambda(t)$, respectively.

Step 4: For a given precision ϵ $(\epsilon>0)$, if Eq. 7 is satisfied, the iteration stops, otherwise return to Step 3), and finally you can get the K, a decomposition of the IMF component denoted as IMFk.

$$\sum_k\frac{\left\|u_k^{n+1}-\hat{u}_k^{n}\right\|_2^2}{\left\|u_k^{n}\right\|_2^2}<\epsilon \tag{7}$$

The subsequent analysis focuses on the 15 intrinsic mode functions (IMFs) obtained through the variational mode decomposition (VMD), which are then utilized for short-term wind speed prediction using an Attention LSTM model. Additionally, the study investigates the optimization of weights associated with each IMF. The detailed process can be found in the flowchart depicted in Figure 1.

2.2.2 Prediction of wind speed IMF components—attention LSTM

The IMFs obtained by applying the variational mode decomposition (VMD) to the original wind speed sequence are individually predicted using an Attention LSTM model. LSTM network is a special type of recurrent neural network (RNN) (Hossain and Mahmood, 2020). Due to its special design, LSTM network memorizes long-term information by default, which can

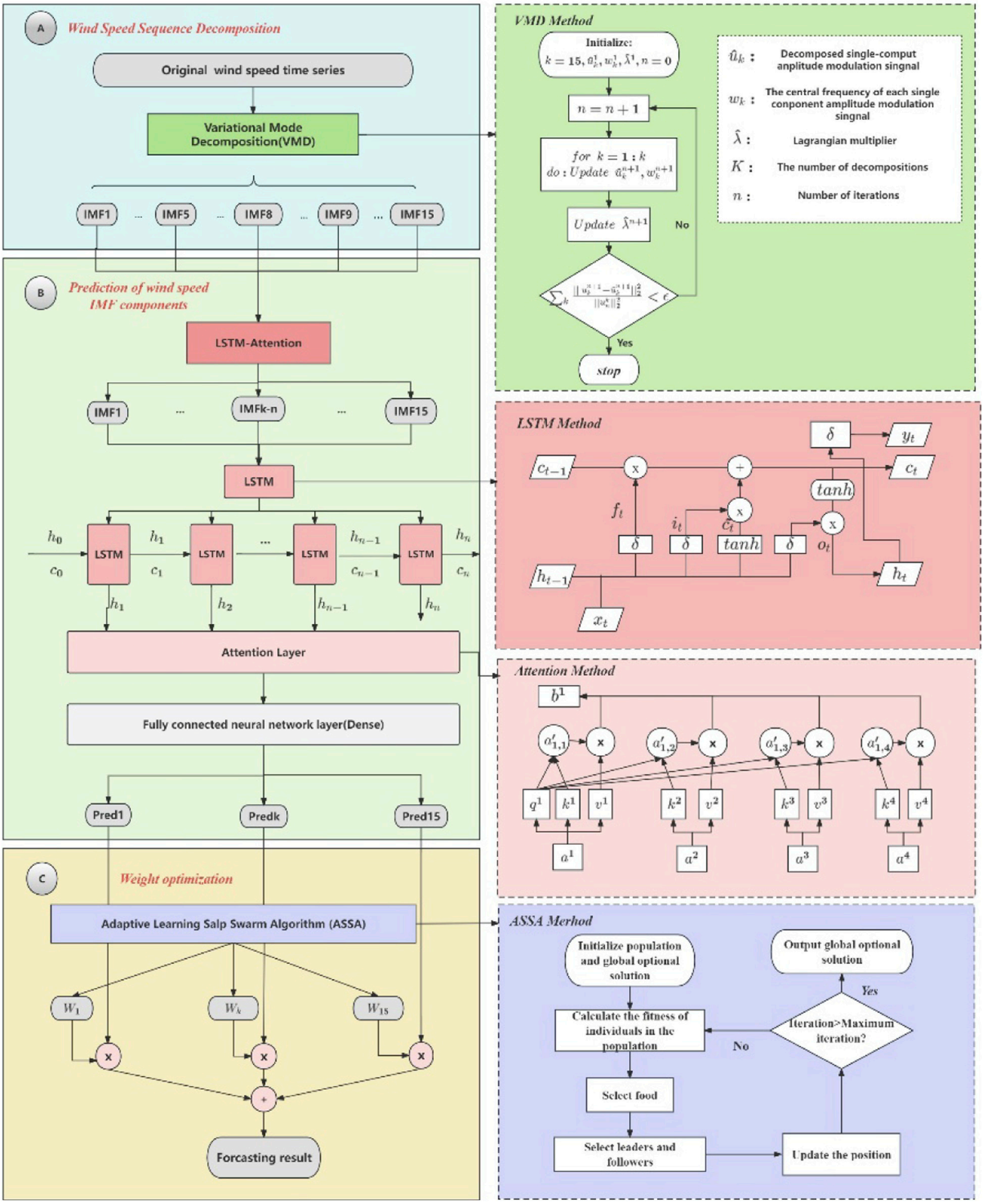


FIGURE 1
Framework and execution process of VMD-AtLSTM-ASSA model. This framework is divided into three steps, **(A)** is wind speed sequence decomposition, **(B)** is prediction of wind speed IMF conponents, and **(C)** is weight optimisation, and the three steps complete the prediction of the whole wind speed sequence in order.

effectively solve the long-term and short-term dependence problem when dealing with nonlinear sequence data. Compared to RNN networks, LSTM networks overcome the problems of gradient vanishing and gradient explosion as well as long-term memory (Hochreiter and Schmidhuber, 1997), because the core of the LSTM network is a memory cell state that replaces the hidden layer of

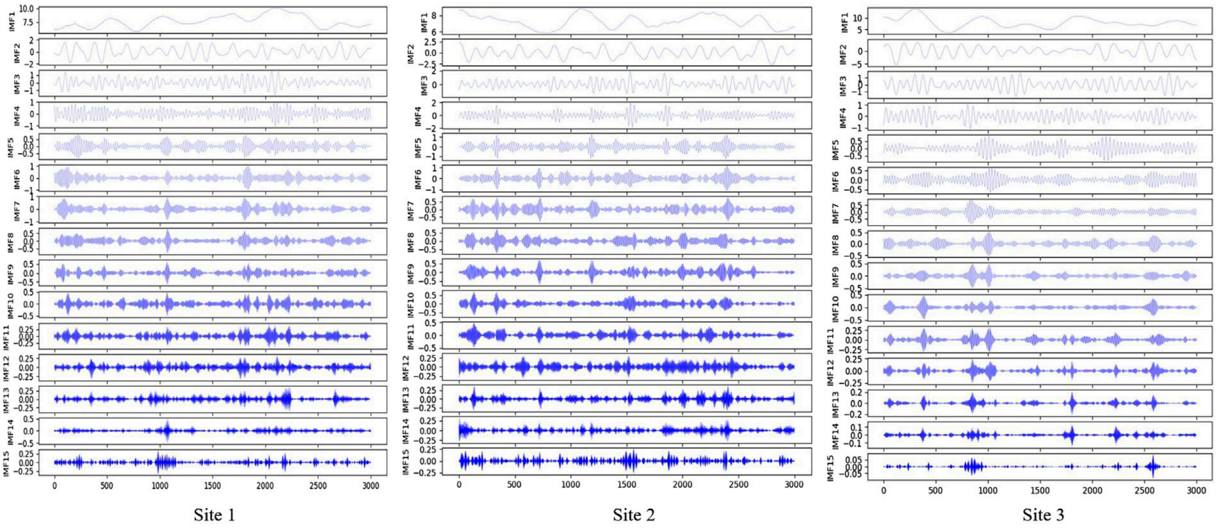


FIGURE 2
Results of variational mode decomposition (VMD) for three sites.

TABLE 1 Characteristics of the three-site wind speed datasets.

Dataset	Number	Statistical indicators			
		Mean (m/s)	Sd. (m/s)	Max (m/s)	Min (m/s)
Site1	3,000	7.6373	1.7722	14.4030	1.8014
Site2	3,000	7.2664	1.9386	15.8270	3.0015
Site3	3,000	8.0485	3.3500	18.1090	0.8450

TABLE 2 Three evaluation indicators for model evaluation.

Metrice	Definition	Equation
RMSE	Root mean square error	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{e}_1 - e_i)^2}$
MAE	Mean absolute error	$MSE = \frac{1}{N}\sum_{i=1}^{N}\lvert\hat{e}_1 - e_i\rvert$
MAPE	Mean Absolute percentage error	$MAPE = \frac{1}{N}\sum_{i=1}^{N}\lvert\frac{\hat{e}_1 - e_i}{e_i}\rvert \times 100$

traditional neurons C_t, which is similar to a conveyor belt so that the information is less likely to be forgotten, and therefore improves the accuracy of the short-term wind speed prediction. However, since in this study, each IMF component is predicted by LSTM model using rolling slice prediction method with a step size of 60, in fact, not every slice plays a key role in the prediction of the wind speed of each IMF component. Therefore, the attention mechanism (Potocnik et al., 2021) is introduced to effectively obtain the important feature relationships of the short-term wind speed slices of the component, so that different weights are assigned to each sample slice to improve the accuracy and also greatly improve the computational accuracy. Figure 1 describes the computation process of Attention LSTM. The principle of the Attention LSTM mechanism is explained as follows:

First, each IMF after decomposition is used as an input to the LSTM $\mathrm{x_t}$, and the flow of the LSTM network is as follows:

Step 1: Decide what information to discard from the memory cell state (calculate the “forget gate” state).

$$\mathrm{f_t = \sigma(W_f \cdot h_{t-1} + W_f \cdot x_t + b_f)} \tag{8}$$

$$\sigma(\mathrm{x}) = \frac{1}{1 + \mathrm{e^x}} \tag{9}$$

In the above equation, $\mathrm{h_{t-1}}$ represents the received output of the previous node, $\mathrm{x_t}$ is the input of IMFk, $\mathrm{W_f}$ is the corresponding weight matrix, $\mathrm{b_f}$ represents the deviation of the “forget gate”, and $\mathrm{f_t}$ represents the state of the “oblivious gate”.

Step 2: Decide which information is stored in the memory cell state (calculate the “input gate” state) and calculate the candidate values for the memory cell state.

$$\mathrm{i_t = \sigma(W_i \cdot h_{t-1} + W_i \cdot x_t + b_i)} \tag{10}$$

$$\mathrm{\tilde{C}_t = tanh(W_c \cdot h_{t-1} + W_c \cdot x_t + b_c)} \tag{11}$$

$$\mathrm{tanh(x)} = \frac{\mathrm{e^x - e^{-x}}}{\mathrm{e^x + e^{-x}}} \tag{12}$$

In the above equation, $\mathrm{h_{t-1}}$ and $\mathrm{x_t}$ are the same as above, $\mathrm{W_i}$ and $\mathrm{W_c}$ are the corresponding weight matrices, $\mathrm{b_i}$ and $\mathrm{b_c}$ represent the

TABLE 3 Related parameter settings.

Model	Parameter	Parameter value
SVR	Step size	60
	Kernal	linear
RNN	Step size	60
	Dropout ratio	0.1
	Epochs	150
	Batch size	64
BPNN	Step size	60
	Learning rate	1e-4
	Epochs	150
	Batch size	64
LSTM	Step size	60
	Dropout ratio	0.1
	Epochs	150
	Batch size	64
	Number of hidden neurons	64
GRU	Step size	60
	Dropout ratio	0.1
	Epochs	150
	Batch size	64
	Number of hidden neurons	64
AtLSTM	Step size	60
	Dropout ratio	0.1
	Epochs	150
	Batch size	64
	Number of hidden neurons	64
VMD	noise margin	0
	Alpha	7,000
	number of decomposition modes	15
ASSA	Population size	10
	Number of iterations	50

deviation of the "input gate" and the deviation of the candidate value of the memory cell state, i_t represents the input states, and $\tilde{C}_t$ represents the candidate value of the memory cell state.

Step 3: Update the current moment memory cell state with the "forget gate" state, the "input gate" state, the previous moment memory cell state, and the candidate value of memory cell state:

$$C_t = f_t{}^*C_{t=1} + i_t{}^*\tilde{C}_t \tag{13}$$

TABLE 4 Comparison of prediction errors of five single models with AtLSTM.

Dataset	Measurement model	Evaluation indicators		
		RMSE	MAE	MAPE
Site1	SVR	0.5831	0.4505	7.0699
	BPNN	0.6044	0.4657	7.2671
	RNN	0.5348	0.3997	6.2173
	GRU	0.5332	0.3967	6.1847
	LSTM	0.5309	0.3954	6.1427
	AtLSTM	**0.5067**	**0.3839**	**5.8560**
Site2	SVR	0.5409	0.4424	7.6596
	BPNN	0.5067	0.3988	6.6723
	RNN	0.5078	0.4075	7.1344
	GRU	0.4571	0.3545	5.7271
	LSTM	0.4562	0.3466	5.6348
	AtLSTM	**0.4418**	**0.3454**	**5.3974**
Site3	SVR	0.6314	0.5270	9.4655
	BPNN	0.3089	0.2229	3.9711
	RNN	0.2058	0.1386	2.5879
	GRU	0.2039	0.1333	2.5565
	LSTM	0.2024	0.1252	2.5397
	AtLSTM	**0.1998**	**0.1249**	**2.5198**

The best values for the evaluation indicators are bolded.

In the above equation C_t denotes the state of the memory cell at the current moment.

Step 4: Determine what information to output from the memory cell state (calculate the "output gate" state):

$$o_t = \sigma(W_o \cdot h_{t-1} + W_o \cdot x_t + b_o) \tag{14}$$

$$h_t = o_t{}^*\tanh(C_t) \tag{15}$$

In the above equation, o_t represents the state of "output gate", W_o is the corresponding weight matrix, b_o represents the deviation of "output gate", and h_t represents the output of current node.

When predicting each wind speed IMF component, it is obviously not rigorous enough to assign the same weight to all input slice information. While the Attention mechanism can capture the important features of wind speed, the Attention mechanism evaluates the importance of different input features, focuses the important information with high weights, ignores the less relevant information with low weights, and finally assigns different weights to them reasonably. Therefore, the Attention mechanism is introduced into the LSTM prediction of each IMF component, and the specific implementation steps of the mechanism are as follows: firstly, the weight coefficients are calculated, i.e., the attention distribution of the slices inside each IMF component is calculated; secondly, the

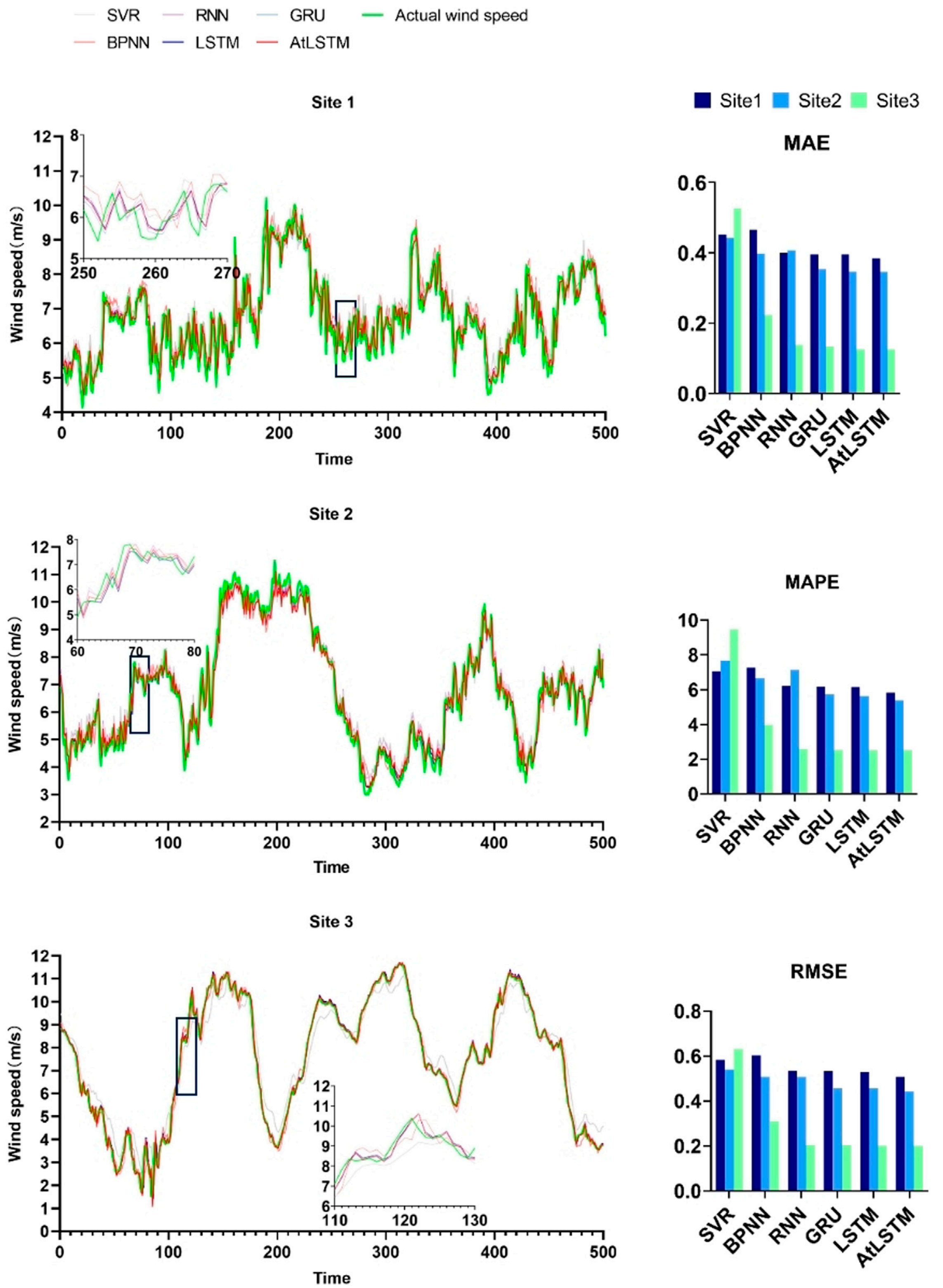


FIGURE 3
Bar charts of the fitting curves and metrics for 5 individual models and Attention LSTM.

weighted summation of the calculated weight coefficients is carried out, i.e., the weighted average of the slices of each IMF component is calculated, and the calculation process is as follows:

Step 1: Multiply the sliced samples a^i in each wind speed IMF component with the corresponding parameter matrix W^q, W^k, W^v to get the corresponding query (q^i), key (k^i), and value (v^i):

TABLE 5 Model error comparison of four decomposition methods combined with AtLSTM.

Dataset	Measurement model	Evaluation indicators		
		RMSE	MAE	MAPE
Site1	EMD-AtLSTM	0.3472	0.2673	4.1494
	EEMD-AtLSTM	0.1830	0.1445	2.2421
	CEEMDAN-AtLSTM	0.4289	0.3472	6.0182
	VMD-AtLSTM	**0.1813**	**0.1411**	**2.1213**
Site2	EMD-AtLSTM	0.2323	0.1743	2.7800
	EEMD-AtLSTM	0.2058	0.1712	2.8656
	CEEMDAN-AtLSTM	0.2296	0.1725	2.7366
	VMD-AtLSTM	**0.1678**	**0.1356**	**2.4595**
Site3	EMD-AtLSTM	0.1440	0.1052	1.9368
	EEMD-AtLSTM	0.1281	0.1037	1.7263
	CEEMDAN-AtLSTM	0.1169	0.1095	1.4721
	VMD-AtLSTM	**0.1072**	**0.0856**	**1.4199**

The best values for the evaluation indicators are bolded.

$$q^i = W^q \cdot a^i \, (i = 1, 2, 3 \ldots), \tag{16}$$

$$k^i = W^k \cdot a^i \, (i = 1, 2, 3 \ldots), \tag{17}$$

$$v^i = W^v \cdot a^i \, (i = 1, 2, 3 \ldots) \tag{18}$$

Step 2: query and key perform similarity calculation to get the weights $\alpha_{i,j}$:

$$\alpha_{i,j} = q^i \cdot k^j \, (i, j = 1, 2, 3 \ldots) \tag{19}$$

Step 3: The weights $\alpha_{i,j}$ are softmax normalized to get the normalized weights $\alpha'_{i,j}$:

$$\alpha'_{i,j} = \text{softmax}\left(\alpha_{i,j}\right), \tag{20}$$

$$\text{softmax}\left(x_{i,j}\right) = \frac{\exp\left(x_{i,j}\right)}{\sum_j \exp\left(x_{i,j}\right)} \tag{21}$$

Step 4: The normalized weights are weighted and summed with VALUE to get the final output of a certain IMF component prediction b^i:

$$b^i = \sum_j \left(\alpha'_{i,j} \cdot v^j\right) \tag{22}$$

2.2.3 Weight optimisation - multi-objective adaptive learning rate salp swarm algorithm

After the prediction of each IMF component sequence, the salp swarm algorithm (SSA) will find the optimal weights of each component, and finally weigh the superposition to get the final short-term wind speed prediction. The salp swarm algorithm (SSA) simulates the group behavior of salp swarm chains, which is a novel swarm intelligence optimization algorithm (Mirjalili et al., 2017). In this study, the sum of IMF components represents the salp swarm, while the individual intrinsic mode function (IMF) represents the individual salp. During the foraging process, the salp swarm will move towards the food in a chain behavior, and the salp at the head of the chain becomes the leader, and the subsequent ones become the followers. During the movement process, the leader carries out global exploration, while the followers fully carry out local exploration, and this search pattern greatly increases the precision of optimization. This foraging process is the process of finding the optimal weights for each wind speed IMF component in this study, where important information is given high weights and information of low relevance is given ground weights.

However, in the SSA, the salp swarm leader is eager to reach the local optimum from the beginning, which leads to insufficient searching and sometimes the algorithm has a low convergence accuracy. Therefore, this paper proposes multi-objective adaptive learning rate salp swarm algorithm (ASSA). Aiming to solve the problem of a lack of global awareness in population updating, we add two different learning operators in leader position updating and follower position updating respectively, which effectively solves the problem of the SSA easily falling into local extremes and improves the optimization accuracy of the algorithm. The flowchart of multi-objective adaptive learning rate salp swarm algorithm (ASSA) is shown in Figure 1, The optimization steps are as follows (Mirjalili et al., 2017):

Step 1: Population initialization. Let the search space be the Euclidean space of D × N, D represents the dimension of the space, and N represents the number of populations. The position of the salp swarm (IMF) is denoted by $X_n = [X_{n1}, X_{n2}, \ldots, X_{nD}]^T$ and the position of food (target weight) $F_n = [F_{n1}, F_{n2}, \ldots, F_{nD}]^T$ is denoted by $n = 1, 2, 3, \ldots, N$. The upper bound of the search space is $u_b = [u_{b1}, u_{b2}, \ldots, u_{bD}]$ and the lower bounds are $l_b = [l_{b1}, l_{b2}, \ldots, l_{bD}]$ and $j = 1, 2, 3, \ldots, N$. Leaders in the population are denoted by X_d^1 and followers by X_d^i, ; $i = 2, 3, 4, \ldots, N$ $d = 1, 2, 3, \ldots, D$

Step 2: Leader position update. During the movement and foraging process of the salp swarm chain, the position of the food source is the target position of all salp swarm individuals, so the leader's position update formula is expressed as:

$$X_d^1 = \begin{cases} F_d + c_1\left(\left(u_b - l_b\right)c_2 + l_b\right), c_3 \geq 0.5 \\ F_d + c_1\left(\left(u_b - l_b\right)c_2 + l_b\right), c_3 < 0.5 \end{cases} \tag{23}$$

Where: X_d^1 and F_d are the position of the first salp (leader) and the position of the food in the d dimension, respectively; u_b and l_b are the corresponding upper and lower bounds, respectively. Where c_1, c_2, c_3 are the control parameters. Eq. 23 shows that the update of the leader's position is only related to the position of the food, c_1 is the convergence factor in the optimization algorithm, which plays the role of balancing the global search and local exploitation, and the expression of c_1 is:

$$c_1 = 2e^{-\left(\frac{4l}{L}\right)^2} \tag{24}$$

Where: l is the current iteration number; L is the maximum iteration number. The convergence factor is a decreasing function

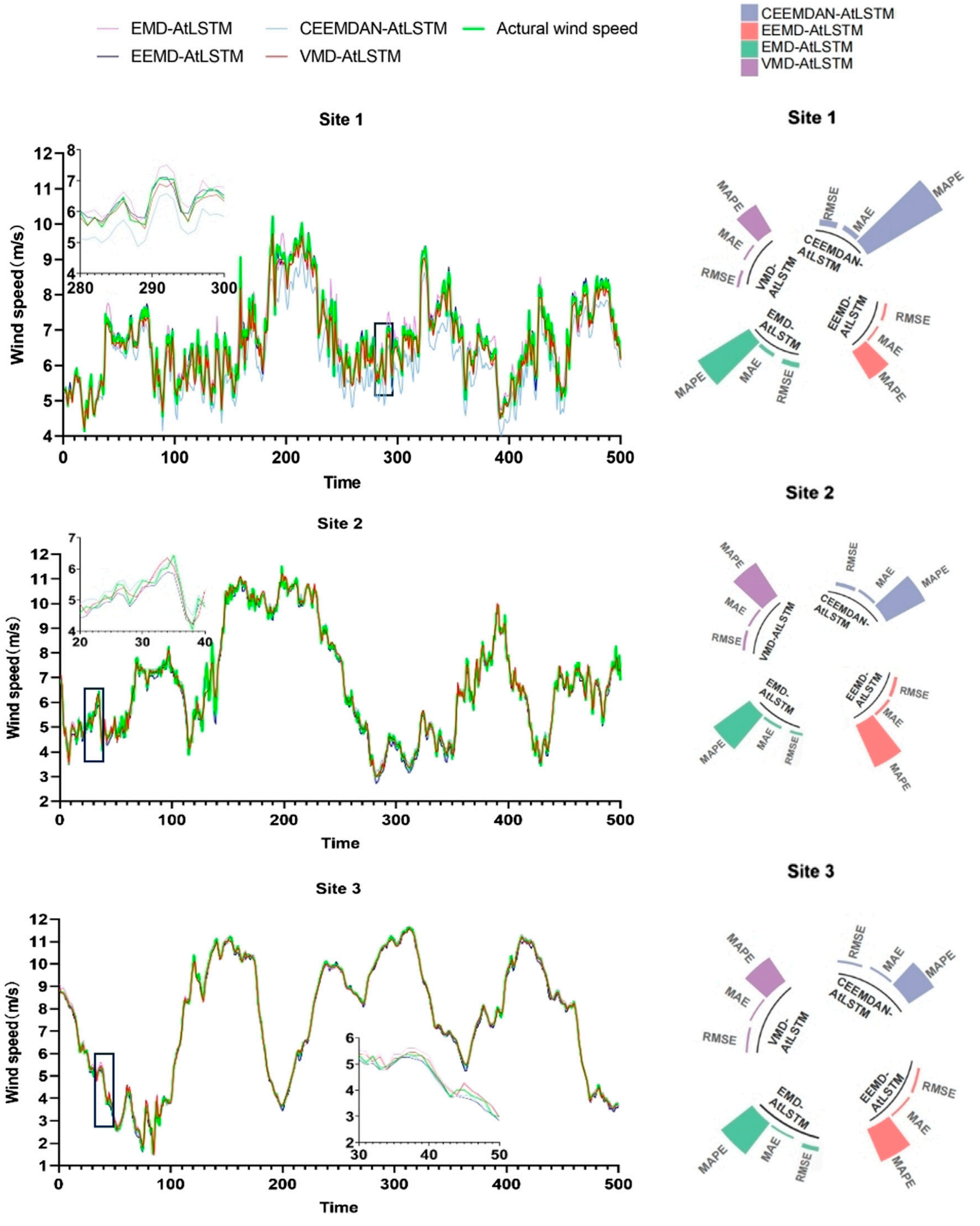


FIGURE 4
Circular bar charts comparing the fitting curves and error metrics of the four decomposition methods combined with AtLSTM.

from 2 to 0. The control parameters c_2, c_3 are random numbers between 0 and 1, which are used to enhance the randomness of X_d^1 to improve the global search ability and individual diversity of the chain cluster.

Step 3: Follower position update. During the movement and foraging process of the salp swarm chain, the followers move forward sequentially in a chain by influencing each other between the front and back individuals. Their displacements conform to Newton's laws of motion, and the equation for the follower's motion displacement is:

$$X = \frac{1}{2}at^2 - v_0 t \tag{25}$$

Where: t is the time; *a* is the acceleration, calculated as $a = (v_{\text{final}} - v_0)/t$; v_0 is the initial velocity, and $v_{\text{final}} = (X_d^i - X_d^{i-1})/t$. Considering that in the optimization

TABLE 6 Comparison of prediction errors based on VMD combined with various deep learning prediction models.

Dataset	Measurement model	Evaluation indicators			Running time(s)
		RMSE	MAE	MAPE	
Site1	VMD-SVR	0.3687	0.2902	4.4368	30.8748
	VMD-BPNN	0.1916	0.1497	2.3733	220.6907
	VMD-RNN	0.1867	0.1435	2.2144	850.7612
	VMD-GRU	0.1825	0.1409	2.1041	1990.3243
	VMD-LSTM	0.1810	0.1402	2.1066	2100.4321
	VMD-AtLSTM	**0.1803**	**0.1401**	**2.1213**	**2400.8764**
	VMD-AtLSTM-SSA	**0.1574**	**0.1212**	**2.0123**	**2405.5656**
	VMD-AtLSTM-ASSA	**0.1553**	**0.1204**	**1.8353**	**2410.0908**
Site2	VMD-SVR	0.5746	0.4599	8.1926	28.5463
	VMD-BPNN	0.2030	0.1581	2.8487	180.6700
	VMD-RNN	0.1790	0.1420	2.3702	780.3212
	VMD-GRU	0.1723	0.1375	2.3942	1897.5009
	VMD-LSTM	0.1707	0.1361	2.3410	1901.3221
	VMD-AtLSTM	**0.1678**	**0.1356**	**2.4595**	**2287.6543**
	VMD-AtLSTM-SSA	**0.1356**	**0.1091**	**1.7755**	**2293.1112**
	VMD-AtLSTM-ASSA	**0.1319**	**0.1003**	**1.6111**	**2296.9898**
Site3	VMD-SVR	0.5082	0.4252	7.1296	26.1276
	VMD-BPNN	0.1023	0.0734	1.3104	175.3435
	VMD-RNN	0.1004	0.0823	1.3261	809.9987
	VMD-GRU	0.1060	0.0851	1.4559	1799.3212
	VMD-LSTM	0.1090	0.0764	1.4377	1831.3221
	VMD-AtLSTM	**0.1062**	**0.0756**	**1.4309**	2108.8876
	VMD-AtLSTM-SSA	**0.0758**	**0.0540**	**1.0272**	**2118.7650**
	VMD-AtLSTM-ASSA	**0.0444**	**0.0273**	**1.0008**	**2123.5409**

Values of evaluation metrics for VMD-AtLSTM,VMD-AtLSTM-SSA,VMD-AtLSTM-ASSA are bolded.

algorithm, t is iterative, let $t = 1$ and $v_0 = 0$. Then the following equation can be obtained:

$$X = \frac{X_d^i - X_d^{i-1}}{2} \tag{26}$$

Where: $i \geq 2$; X_d^i, X_d^{i-1} are the positions of the two salps immediately connected to each other in the d dimension, respectively. Therefore, the position of the follower is denoted as:

$$X_d^{i'} = \frac{X_d^i - X_d^{i-1}}{2} \tag{27}$$

where $X_d^{i'}$ and X_d^{i-1} are the position of the updated follower and the position of the pre-updated follower in dimension d, respectively.

However, in the SSA algorithm, the salp swarm leader runs to the global optimum from the beginning of the iteration, which leads to insufficient global search, and an occasionally low convergence accuracy of the algorithm. To address this problem, this paper proposes the ASSA algorithm. For the problem of lack of global awareness in the population update, we add two different learning operators on the leader position update and follower position update respectively, which effectively solves the problem of the SSA algorithm easily falling into the local extreme value and improves the algorithm's optimization accuracy.

The learning operator for leader position update is added to make the population search more biased towards large-scale search in the early stage and focused towards the global optimal solution in the late stage of the search. The improved salp swarm leader position update process is:

$$x_{i,j} = F_d - k \cdot c_1 (\max - \min) c_2 + \min \tag{28}$$

$$k = \exp\left(-\frac{\text{count}}{\text{iter}}\right) \tag{29}$$

where count is the current iteration number in the range, [0, iter] *iter* is the maximum iteration number.

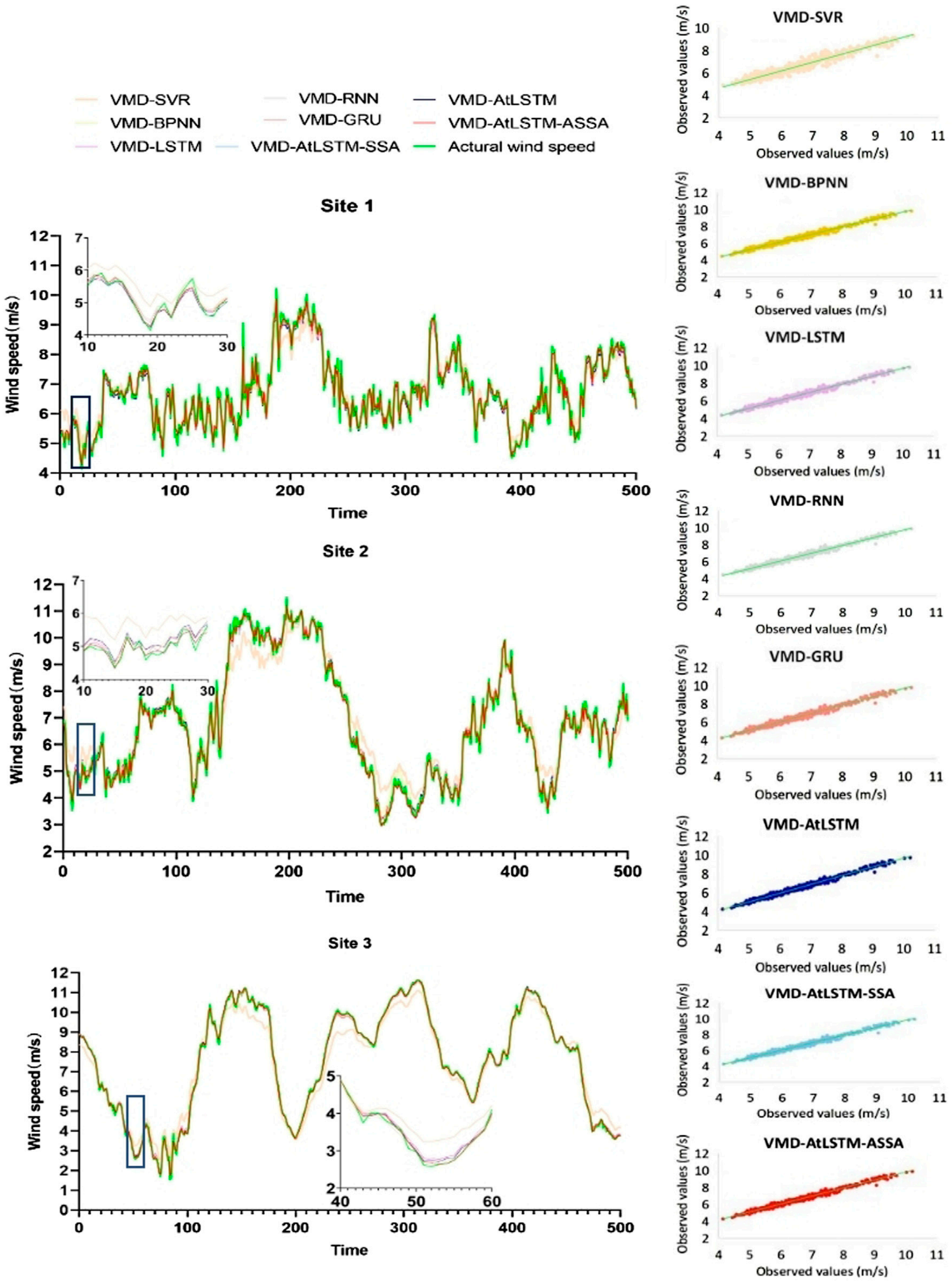
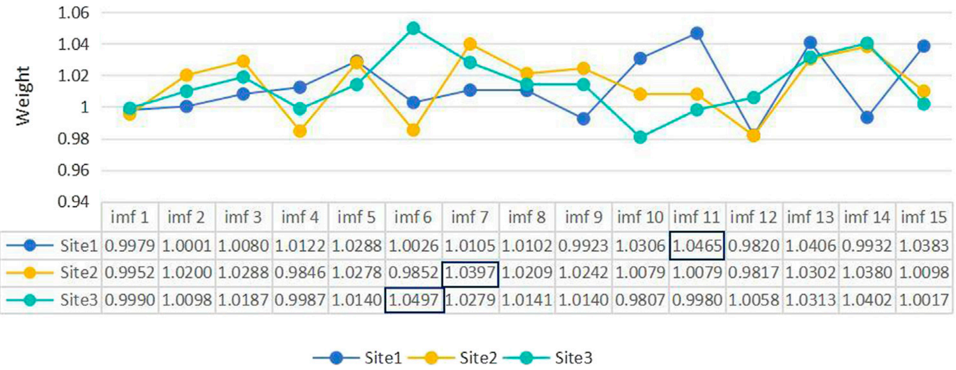


FIGURE 5
Curve fitting and regression fitting graphs of VMD combined with each model for prediction.

For the position update of a salp swarm follower, the individual position is always affected by the two individuals before and after it, and the fitness of the two individuals is unknown. Therefore, we propose that by calculating the fitness values of the two individuals and restricting the poorly adapted individual, we weaken the influence of the poorly adapted individual on the individual update at the current moment. The improved bottles sea squirt follower position update process is:



	imf 1	imf 2	imf 3	imf 4	imf 5	imf 6	imf 7	imf 8	imf 9	imf 10	imf 11	imf 12	imf 13	imf 14	imf 15
Site1	0.9979	1.0001	1.0080	1.0122	1.0288	1.0026	1.0105	1.0102	0.9923	1.0306	1.0465	0.9820	1.0406	0.9932	1.0383
Site2	0.9952	1.0200	1.0288	0.9846	1.0278	0.9852	1.0397	1.0209	1.0242	1.0079	1.0079	0.9817	1.0302	1.0380	1.0098
Site3	0.9990	1.0098	1.0187	0.9987	1.0140	1.0497	1.0279	1.0141	1.0140	0.9807	0.9980	1.0058	1.0313	1.0402	1.0017

FIGURE 6
Results of ASSA weight searching.

$$x_{i,j} = \begin{cases} \frac{1}{2}\left(x_{i,j} + k \cdot x_{i-1,j}\right) & f(x_i) < f(x_{i-1}) \\ \frac{1}{2}\left(k \cdot x_{i,j} + x_{i-1,j}\right) & f(x_i) > f(x_{i-1}) \end{cases} \tag{30}$$

Where count is the current number of iterations in the range, $[0, \text{iter}]$ *iter* is the maximum number of iterations, and $f(x_i)$ is the fitness value for each position.

This improved optimization algorithm has more than one objective function, thus the optimization problem is changed to a multi-objective optimization problem. We first de-measure the objective functions to ensure that the objective functions have the same measure; then average the objective functions, and then transform the multi-objective optimization problem into a simple single-objective optimization problem to solve the problem. As follows:

$$f(x_i) = \frac{(RMSE_i + MAE_i)}{2} \tag{31}$$

Step4: Judge whether the current iteration number count satisfies the maximum iteration number iter, if so, output the optimal weight results of each IMF component, otherwise return to Step2.

3 Case study

In this section, to verify the effectiveness of the proposed VMD-AtLSTM-ASSA model, we experimentally study the model using wind speed data collected from wind farms in three different regions. The VMD-AtLSTM-ASSA model is compared with popular models in the research field. All experiments are implemented under the deep learning framework under Python 3.7.3. The configuration of the emulated platform is Intel(R) Core(TM) i5-8250U CPU @ 1.60 GHz 1.80 GHz with 8 GB memory capacity.

3.1 Dataset

The study collected wind speed datasets from three sites on https://data.nrel.gov/search-page, each with 3,000 data points. Site1 came from the St. Thomas Wind Station in the Virgin Islands, the United States; the site 2 wind speed dataset came from the St. Croix Wind Station, the United States Virgin Islands, and the site 3 wind speed dataset came from the Woodburn Wind Station in the United States. At each of the three sites, the wind speed was collected. The last 500 data points were taken as the test set in all datasets, while the rest was taken as the training set. The characteristics of the dataset are shown in Table 1 below.

3.2 Experiments and evaluation indicators

To validate the effectiveness and high accuracy performance of the proposed hybrid model, three sets of comparative experiments were conducted. Experiment 1 compared the predictive performance of AtLSTM with currently popular single deep learning models, verifying the superior predictive performance of Attention LSTM. Experiment 2 compared the prediction results of different wind speed sequence decomposition methods combined with Attention LSTM, demonstrating the superiority of VMD followed by Attention LSTM prediction. Experiment 3 compared the prediction results of different deep learning models combined with VMD, as well as the performance of models incorporating the optimization models SSA and ASSA. This experiment validated the superiority of the VMD-Attention LSTM hybrid model and the excellent predictive performance and stability of the VMD-AtLSTM-ASSA model. The details of these three sets of comparative experiments will be presented in Sections 3.4–3.6.

In the experiments, three different evaluation indicators, root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE), were used to present and analyze the experimental results, and according to the value of the evaluation Indicators, the model's prediction performance was evaluated (Jiang et al., 2021). Their homologous expressions are shown in Table 2, it is worth noting that N represents the length of a predicted subsequence and e_i and $\hat{e}_i$ stand for the actual and predicted values, respectively.

3.3 Model parameter settings

In order to verify the validity of the proposed model, the model parameters used in this study are the same, eliminating the influence of model parameter settings on experimental results. The Attention LSTM is commonly referred to as AtLSTM in the experimental setting. Table 3 shows the model parameter settings used.

3.4 Experiment 1: validating the accuracy advantage of the AtLSTM model over a single model

In this experiment, AtLSTM was compared with SVR, BPNN, RNN, GRU, and LSTM models to validate the exceptional predictive performance of the proposed model. The evaluation metrics for model prediction performance are presented in Table 4, with bold font used to indicate the metrics of the AtLSTM model. Figure 3 provides a visual representation of the differences in predictive performance between the proposed model and the four deep learning models.

From Table 4; Figure 3, it is evident that there are variations in the experimental results across the three stations. Both LSTM and GRU demonstrate excellent predictive performance, with LSTM slightly outperforming GRU. By incorporating the Attention mechanism, AtLSTM exhibits a significant improvement in predictive performance compared to LSTM. As shown in Table 4, for different datasets and the five models considered, AtLSTM consistently achieves lower error values, indicating its superior predictive ability. Specifically, in the experiments conducted on the three stations, AtLSTM demonstrates a maximum reduction of 16.16% in RMSE, 17.56% in MAE, and 19.42% in MAPE when compared to other prediction models, namely, SVR, BPNN, RNN, GRU, and LSTM. Therefore, it can be reasonably concluded that AtLSTM possesses superiority in improving the accuracy of prediction results.

3.5 Experiment 2: Validating the decomposition advantages of the VMD model over other decomposition models

To demonstrate the superiority of AtLSTM based on the VMD decomposition model over other decomposition methods in improving wind speed prediction accuracy, we compared it with EMD-AtLSTM, EEMD-AtLSTM, and CEEMDAN-AtLSTM to validate the superior predictive performance of VMD-AtLSTM. The evaluation metrics for model prediction performance are presented in Table 5, with bold font used to indicate the metrics of the VMD-AtLSTM model. The fitting graph and circular bar chart in Figure 4 visually display the differences in predictive performance between VMD-AtLSTM and the other three decomposition models.

From Table 5, it can be observed that VMD-AtLSTM achieves the lowest error values across the three locations. Compared to other decomposition models, namely, EMD-AtLSTM, EEMD-AtLSTM, and CEEMDAN-AtLSTM, VMD-AtLSTM exhibits maximum reductions of 57.73%, 59.36%, and 64.75% in RMSE, MAE, and MAPE values, respectively. In conclusion, it can be reasonably argued that combining VMD with AtLSTM for wind speed prediction demonstrates superiority in enhancing short-term wind speed prediction accuracy compared to other signal decomposition techniques.

3.6 Experiment 3: validating the predictive performance advantages of VMD-AtLSTM and VMD-AtLSTM-ASSA

In order to validate the superior predictive performance of the proposed VMD-AtLSTM-ASSA model, we first compared the prediction errors of VMD-SVR, VMD-BPNN, VMD-RNN, VMD-GRU, and VMD-LSTM models, and then evaluated the superiority of VMD-AtLSTM. Furthermore, we verified the effectiveness of incorporating SSA in improving the accuracy of VMD-AtLSTM. Subsequently, a comparison of prediction errors was conducted between the VMD-AtLSTM-SSA and VMD-AtLSTM-ASSA models, ultimately confirming the significant positive impact of the proposed VMD-AtLSTM-ASSA on prediction accuracy. Table 6 displays the prediction error metrics and computational time for the eight hybrid models, while Figure 5 further illustrates the prediction results obtained by each model at the three stations.

The results from Table 6; Figure 5 indicate that among the VMD-based hybrid models, VMD-LSTM performs the best in terms of prediction accuracy, followed by VMD-RNN and VMD-GRU. The VMD-AtLSTM model exhibits improved accuracy compared to VMD-LSTM, suggesting that incorporating attention mechanism enhances the predictive accuracy of the LSTM model. Additionally, it can be observed that VMD-AtLSTM-SSA reduces the RMSE value by 12.7% compared to VMD-AtLSTM, while VMD-AtLSTM-ASSA further reduces the RMSE value by 1.33% based on VMD-AtLSTM-SSA. The proposed model achieves maximum reductions of 57.88%, 58.51%, and 58.63% in RMSE, MAE, and MAPE values, respectively, compared to other prediction models. In summary, compared to other VMD-based hybrid models, VMD-AtLSTM improves prediction accuracy by incorporating attention mechanism into LSTM and adding ASSA effectively optimizes prediction accuracy. Figure 6 displays the results of weight searching for 15 IMF components after applying the ASSA algorithm at the three stations. It can be observed that the importance of IMF components varies across different datasets. IMF11 in Site1, IMF7 in Site2, and IMF6 in Site3 are identified as the dominant modes influencing the prediction results, and therefore assigned higher weights.

Generally, the complexity of a model is related to its computational time. Table 6 presents the running time of each model, revealing that AtLSTM model takes slightly more time to execute compared to the LSTM model due to the attention mechanism requiring importance calculation for each slice of the prediction sequence. However, the inclusion of the optimized model ASSA only requires approximately 10 seconds. Overall, the VMD-AtLSTM-ASSA model demonstrates superior predictive performance.

4 Conclusion and future work

With the rapid development of China's economy, the consumption of traditional non-renewable resources (oil, coal, etc.) is huge, and wind energy, as a renewable and clean energy source, is becoming an important green power generation method for the modern power grid. However, due to the non-linear and non-stationary nature of wind speed, this trait seriously affects the safe and reliable operation of the power system, and finally leads to problems such as, difficult grid scheduling of wind farms. Therefore, the development of a high-precision and high-reliability short-term wind speed prediction model can, on the one hand, provide efficient and reliable planning for wind power, and on the other hand, stabilize the power grid and reduce the volatility. Numerous researchers have continuously invested in the study of wind speed prediction models, and a steady stream of wind speed prediction models have been proposed. Some examples of such models are, physical models based on meteorological data prediction; statistical models to establish the relationship with future wind speed function by calculating the historical wind speed; artificial intelligence prediction models based on training the model on training samples.

However, the above methods do not work well for fluctuating and complex data, so this paper proposes a short-term wind speed prediction model based on a mixture of the VMD model, the Attention LSTM prediction model, and an improved salp swarm algorithm (multi-objective adaptive learning rate salp swarm algorithm). In this study, the VMD model is employed to decompose the original wind speed sequence into multiple stable intrinsic mode functions (IMFs). Subsequently, the AtLSTM model is utilized to individually forecast each IMF component. Finally, the proposed ASSA algorithm is applied to assign weights to each IMF component, resulting in a weighted aggregation that yields highly accurate short-term wind speed predictions.

In this study, by simulating wind speed data from three wind farms and designing three aspects of comparison experiments, the experimental results illustrate that the data preprocessing strategy based on VMD technology can effectively reduce the volatility and complexity of the wind speed sequence, and significantly improve the accuracy of short-term high wind speed prediction. Furthermore, in the prediction module, the Attention LSTM (AtLSTM) with an incorporated attention mechanism is introduced. This attention mechanism enables the LSTM network to analyze the importance of each temporal slice of input data, assigning higher weight values to slices that have a significant impact on the prediction results. As a result, the predictive accuracy is enhanced. Finally, the multi-objective adaptive learning rate salp swarm algorithm (ASSA) proposed in the weight optimization part adds two operators on the basis of salp swarm algorithm (SSA) that effectively solve the problem of local optimal solution, which the original algorithm is prone to, so as to improve its accuracy in optimization searching. In summary, by setting up a large number of different comparison experiments, it has been verified that the hybrid short-term wind speed prediction model proposed in this paper based on the multi-objective adaptive learning rate salp swarm algorithm (ASSA), Attention LSTM, and VMD has fully demonstrated the accuracy advantage of the model.

In this study, a hybrid VMD-AtLSTM-ASSA short-term wind speed prediction model with decomposition algorithm and optimization algorithm is proposed to address the characteristics of short-term wind speed unsteadiness and nonlinearity and the lack of prediction accuracy of a single model for complex data. This proposed model shows excellent prediction performance. Nevertheless, this model still has more application scenarios and room for expansion. Firstly, this study mainly focuses on the processing and prediction of wind speed time series information, and other data inputs, such as, wind direction information, seasonal information, and spatial information between wind farms, can be considered to expand the model's environmental adaptability. Secondly, the K value of the variational modal decomposition algorithm used in this study is determined by judging whether the center frequency of each IMF is aliased or omitted, also, the α value is limited to 7,000, so the K value in this paper is selected for the experimental data in this paper, and it is not adaptive, so the introduction of optimization algorithms can be considered to achieve adaptive modal decomposition. Furthermore, within this research, we have observed that VMD-GRU demonstrates remarkable predictive accuracy and computational efficiency. Therefore, in future studies, we plan to introduce additional advanced models for comparative analysis. Additionally, we aim to conduct comprehensive optimizations addressing both the accuracy and model complexity limitations identified in these models during our research. In addition, the optimization algorithm for the machine learning algorithm in this study is the salp swarm algorithm (SSA). Considering the rapid progress in the research of swarm intelligence algorithms, more efficient swarm intelligence optimization algorithms can be added to the future research, and other optimization algorithms can be replaced to improve the prediction performance of the model. Finally, the hybrid VMD-AtLSTM-ASSA short-term wind speed prediction model proposed in this paper is also suitable for other datasets with complex data, high volatility, and high accuracy requirements, such as, crude oil prices and nuclear energy consumption.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

LW: Conceptualization, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing–original draft, Writing–review and editing. YL: Formal Analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing–original draft, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The work was supported by a grant from the National Natural Science Foundation of China (Grant No: 42171419), awarded to YL.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Altan, A., Karasu, S., and Zio, E. (2021). A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer. *Appl. Soft Comput.* 100, 106996. doi:10.1016/j.asoc.2020.106996

Duan, J., Zuo, H., Bai, Y., Duan, J., Chang, M., Chen, B., et al. (2021). Short-term wind speed forecasting using recurrent neural networks with error correction. *Energy* 217, 119397. doi:10.1016/j.energy.2020.119397

Dragomiretskiy, K., and Zosso, D. (2014). Variational mode decomposition. *Ieee Trans. Signal Process.* 62, 531–544. doi:10.1109/TSP.2013.2288675

Faris, H., Mafarja, M., Heidari, A., Aljarah, I., Al-Zoubi, A., Mirjalili, S., et al. (2018). An efficient binary Salp Swarm Algorithm with crossover scheme for feature selection problems. *Knowledge-Based Syst.* 154, 43–67. doi:10.1016/j.knosys.2018.05.009

Fu, W., Wang, K., Li, C., and Tan, J. (2019). Multi-step short-term wind speed forecasting approach based on multi-scale dominant ingredient chaotic analysis, improved hybrid GWO-SCA optimization and ELM. *Energy Convers. And Manag.* 187, 356–377. doi:10.1016/j.enconman.2019.02.086

Guliyev, F. (2020). Trump's "America first" energy policy, contingency and the reconfiguration of the global energy order. *Energy Policy* 140, 111435. doi:10.1016/j.enpol.2020.111435

Hansen, N., and Ostermeier, A. (2001). Completely derandomized self-adaptation in evolution strategies. *Evol. Comput.* 9, 159–195. doi:10.1162/106365601750190398

Hassan, R. U., Li, C., and Liu, Y. (2021). Online dynamic security assessment of wind integrated power system using SDAE with SVM ensemble boosting learner. *Int. J. Of Electr. Power and Energy Syst.* 125, 106429. doi:10.1016/j.ijepes.2020.106429

Higashiyama, K., Fujimoto, Y., and Hayashi, Y. (2018). *Feature extraction of NWP data for wind power forecasting using 3D-convolutional neural networks*. in, ed. P. Droege, 350–358. doi:10.1016/j.egypro.2018.11.043

Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* 9, 1735–1780. doi:10.1162/neco.1997.9.8.1735

Hossain, M. S., and Mahmood, H. (2020). "Short-term load forecasting using an LSTM neural network," in Proceeding of the 2020 IEEE Power and Energy Conference at Illinois (PECI), Champaign, IL, USA, February 2020 (IEEE), 1–6. doi:10.1109/PECI48348.2020.9064654

Hu, H., Wang, L., and Tao, R. (2021). Wind speed forecasting based on variational mode decomposition and improved echo state network. *Renew. Energy* 164, 729–751. doi:10.1016/j.renene.2020.09.109

Jaseena, K., and Kovoor, B. (2021). Decomposition-based hybrid wind speed forecasting model using deep bidirectional LSTM networks. *Energy Convers. And Manag.* 234, 113944. doi:10.1016/j.enconman.2021.113944

Jiang, P., Liu, Z., Niu, X., and Zhang, L. (2021). A combined forecasting system based on statistical method, artificial neural networks, and deep learning methods for short-term wind speed forecasting. *Energy* 217, 119361. doi:10.1016/j.energy.2020.119361

Kang, F., Li, J., and Dai, J. (2019). Prediction of long-term temperature effect in structural health monitoring of concrete dams using support vector machines with Jaya optimizer and salp swarm algorithms. *Adv. Eng. Softw.* 131, 60–76. doi:10.1016/j.advengsoft.2019.03.003

Kennedy, J., and Eberhart, R. (2011). Particle swarm optimization. *Proc. IEEE Int. Conf. Neural Netw.* 4, 1942–1948. Perth, Australia. doi:10.1109/ICNN.1995.488968

Khodayar, M., Kaynak, O., and Khodayar, M. (2017). Rough deep neural architecture for short-term wind speed forecasting. *Ieee Trans. Industrial Inf.* 13, 2770–2779. doi:10.1109/TII.2017.2730846

Lacal-Arantegui, R. (2019). Globalization in the wind energy industry: contribution and economic impact of European companies. *Renew. Energy* 134, 612–628. doi:10.1016/j.renene.2018.10.087

Liu, H., Tian, H., Chen, C., and Li, Y. (2010). A hybrid statistical method to predict wind speed and wind power. *Renew. Energy* 35, 1857–1861. doi:10.1016/j.renene.2009.12.011

Liu, Y., Qin, H., Zhang, Z., Pei, S., Jiang, Z., Feng, Z., et al. (2020a). Probabilistic spatiotemporal wind speed forecasting based on a variational Bayesian deep learning model. *Appl. Energy* 260, 114259. doi:10.1016/j.apenergy.2019.114259

Liu, Z., Jiang, P., Zhang, L., and Niu, X. (2020b). A combined forecasting model for time series: application to short-term wind speed forecasting. *Appl. Energy* 259, 114137. doi:10.1016/j.apenergy.2019.114137

Lowery, C., and O'Malley, M. (2012). Impact of wind forecast error statistics upon unit commitment. *Ieee Trans. Sustain. Energy* 3, 760–768. doi:10.1109/TSTE.2012.2210150

Ma, L., Luan, S., Jiang, C., Liu, H., and Zhang, Y. (2009). A review on the forecasting of wind speed and generated power. *Renew. Sustain. Energy Rev.* 13, 915–920. doi:10.1016/j.rser.2008.02.002

Mirjalili, S., Gandomi, A., Mirjalili, S., Saremi, S., Faris, H., and Mirjalili, S. (2017). Salp Swarm Algorithm: a bio-inspired optimizer for engineering design problems. *Adv. Eng. Softw.* 114, 163–191. doi:10.1016/j.advengsoft.2017.07.002

Mirjalili, S., and Lewis, A. (2016). The whale optimization algorithm. *Adv. Eng. Softw.* 95, 51–67. doi:10.1016/j.advengsoft.2016.01.008

Naik, J., Satapathy, P., and Dash, P. (2018). Short-term wind speed and wind power prediction using hybrid empirical mode decomposition and kernel ridge regression. *Appl. Soft Comput.* 70, 1167–1188. doi:10.1016/j.asoc.2017.12.010

Neshat, M., Majidi Nezhad, M., Mirjalili, S., Piras, G., and Garcia, D. A. (2022). Quaternion convolutional long short-term memory neural model with an adaptive decomposition method for wind speed forecasting: north aegean islands case studies. *Energy Convers. Manag.* 259, 115590. doi:10.1016/j.enconman.2022.115590

Neshat, M., Nezhad, M. M., Abbasnejad, E., Mirjalili, S., Tjernberg, L. B., Garcia, D. A., et al. (2021). A deep learning-based evolutionary model for short-term wind speed forecasting: a case study of the Lillgrund offshore wind farm. *Energy Convers. And Manag.* 236, 114002. doi:10.1016/j.enconman.2021.114002

Niu, Z., Yu, Z., Tang, W., Wu, Q., and Reformat, M. (2020). Wind power forecasting using attention-based gated recurrent unit network. *Energy* 196, 117081. doi:10.1016/j.energy.2020.117081

Paliwal, K., and Basu, A. (1987). "A speech enhancement method based on Kalman filtering," in ICASSP '87. IEEE International Conference on Acoustics, Speech, and Signal Processing, Dallas, TX, USA, April 1987 (IEEE), 177–180. doi:10.1109/ICASSP.1987.1169756

Pan, J., Shan, J., Zheng, S., Chu, S., and Chang, C. (2021). Wind power prediction based on neural network with optimization of adaptive multi-group salp swarm algorithm. *Clust. Computing-The J. Of Netw. Softw. Tools And Appl.* 24, 2083–2098. doi:10.1007/s10586-021-03247-x

Potocnik, P., Skerl, P., and Govekar, E. (2021). Machine-learning-based multi-step heat demand forecasting in a district heating system. *Energy And Build.* 233, 110673. doi:10.1016/j.enbuild.2020.110673

Ren, Y., Suganthan, P., and Srikanth, N. (2016). A novel empirical mode decomposition with support vector regression for wind speed forecasting. *Ieee Trans. Neural Netw. And Learn. Syst.* 27, 1793–1798. doi:10.1109/TNNLS.2014.2351391

Rodrigues Moreno, S., Gomes da Silva, R., Cocco Mariani, V., and dos Santos Coelho, L. (2020). Multi-step wind speed forecasting based on hybrid multi-stage decomposition model and long short-term memory neural network. *Energy Convers. Manag.* 213, 112869. doi:10.1016/j.enconman.2020.112869

Sahin, A., and Sen, Z. (2001). First-order Markov chain approach to wind speed modelling. *J. Of Wind Eng. And Industrial Aerodynamics* 89, 263–269. doi:10.1016/S0167-6105(00)00081-7

Scutaru, M., Vlase, S., Marin, M., and Modrea, A. (2020). New analytical method based on dynamic response of planar mechanical elastic systems. *Bound. Value Probl.* 2020, 104. doi:10.1186/s13661-020-01401-9

Shahid, F., Zameer, A., and Muneeb, M. (2021). A novel genetic LSTM model for wind power forecast. *Energy* 223, 120069. doi:10.1016/j.energy.2021.120069

Shao, Z., Han, J., Zhao, W., Zhou, K., and Yang, S. (2022). Hybrid model for short-term wind power forecasting based on singular spectrum analysis and a temporal convolutional attention network with an adaptive receptive field. *Energy Convers. And Manag.* 269, 116138. doi:10.1016/j.enconman.2022.116138

Storn, R., and Price, K. (1997). Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* 11, 341–359. doi:10.1023/A:1008202821328

Tanaka, I., and Ohmori, H. (2015). "Method selection in different regions for short-term wind speed prediction in Japan," in Proceeding of the 2015 54th Annual

Conference of the Society of Instrument and Control Engineers of Japan (SICE), Hangzhou, China, July 2015 (IEEE), 189–194.

Tascikaraoglu, A., and Uzunoglu, M. (2014). A review of combined approaches for prediction of short-term wind speed and power. *Renew. Sustain. Energy Rev.* 34, 243–254. doi:10.1016/j.rser.2014.03.033

Wang, J., and Cheng, Z. (2021). Wind speed interval prediction model based on variational mode decomposition and multi-objective optimization. *Appl. Soft Comput.* 113, 107848. doi:10.1016/j.asoc.2021.107848

Wang, J., and Li, W. (2018). Ultra-short-term forecasting of wind speed based on CEEMD and GWO. *Dianli Xit. Baohu yu Kongzhi/Power Syst. Prot. Control* 46, 69–74. doi:10.7667/PSPC170590

Wang, L., Zeng, Y., and Chen, T. (2015). Back propagation neural network with adaptive differential evolution algorithm for time series forecasting. *Expert Syst. Appl.* 42, 855–863. doi:10.1016/j.eswa.2014.08.018

Yu, C., Li, Y., Bao, Y., Tang, H., and Zhai, G. (2018). A novel framework for wind speed prediction based on recurrent neural networks and support vector machine. *Energy Convers. And Manag.* 178, 137–145. doi:10.1016/j.enconman.2018.10.008

Yuan, X., Tan, Q., Lei, X., Yuan, Y., and Wu, X. (2017). Wind power prediction using hybrid autoregressive fractionally integrated moving average and least square support vector machine. *Energy* 129, 122–137. doi:10.1016/j.energy.2017.04.094

Zaremba, W., Sutskever, I., and Vinyals, O. (2015). *Recurrent neural network regularization*. doi:10.48550/arXiv.1409.2329

frontiers | Frontiers in Energy Research

TYPE Brief Research Report
PUBLISHED 24 November 2023
DOI 10.3389/fenrg.2023.1305612

Interval model of a wind turbine power curve

Kai Zhou, Hao Han, Junfen Li, Yongjie Wang, Wei Tang, Fei Han, Yulei Li, Ruyu Bi, Haitao Zhao and Lingxiao Jiao*

State Grid Hubei Extra High Voltage Company, Wuhan, Hubei, China

Check for updates

OPEN ACCESS

EDITED BY
Xun Shen,
Osaka University, Japan

REVIEWED BY
Hardeep Singh,
University of Windsor, Canada
Shuang Zhao,
Hefei University of Technology, China
Yikui Liu,
Stevens Institute of Technology, United States

*CORRESPONDENCE
Lingxiao Jiao,
✉ 1062495017@qq.com

RECEIVED 02 October 2023
ACCEPTED 31 October 2023
PUBLISHED 24 November 2023

CITATION
Zhou K, Han H, Li J, Wang Y, Tang W, Han F, Li Y, Bi R, Zhao H and Jiao L (2023), Interval model of a wind turbine power curve.
Front. Energy Res. 11:1305612.
doi: 10.3389/fenrg.2023.1305612

COPYRIGHT
© 2023 Zhou, Han, Li, Wang, Tang, Han, Li, Bi, Zhao and Jiao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The wind turbine power curve model is critical to a wind turbine's power prediction and performance analysis. However, abnormal data in the training set decrease the prediction accuracy of trained models. This paper proposes a sample average approach-based method to construct an interval model of a wind turbine, which increases robustness against abnormal data and further improves the model accuracy. We compare our proposed methods with the traditional neural network-based and Bayesian neural network-based models in experimental data-based validations. Our model shows better performance in both accuracy and computational time.

KEYWORDS

abnormal detection, data cleaning, wind power prediction, prediction accuracy, stochastic optimization

1 Introduction

Wind power has become a significant renewable power source of global energy systems (Gilbert et al., 2020). To ensure operation safety and high efficiency, it is critical to monitor the operation conditions of wind turbines to predict wind power output and detect potential faults (Wang and Liu, 2021). The Supervisory Control and Data Acquisition (SCADA) system provides wind turbine data, for example, wind speed and power, for establishing models for wind turbine power prediction and fault detection (Y. Wang et al., 2019). A significant proportion of the SCADA data is abnormal due to communication failures, maintenance, and other reasons (Morrison et al., 2022). A model trained by a dataset with abnormal data is biased from a real model and suffers from reduced accuracy (Ye et al., 2021). It is necessary to consider the data cleaning-based method to obtain a model with improved robustness against abnormal data.

Prior results of the data cleaning method were based on clustering algorithms (Zheng et al., 2015; 2010; Yesilbudak, 2016). In clustering algorithm-based methods, k-means, manifold spectral clustering, and other algorithms are applied to separate the wind power curve into partitions and then identify the outliers based on distances to the cluster centers. An alternative method is to determine the upper and lower boundaries of the wind power curve by boundary models. For example, Shen et al. (2019) used change point and quantile to estimate a contour for normal data. However, the aforementioned methods fail to identify the outliers when there are many cluster centers. The setting of the algorithm parameters, for example, cluster number in the clustering algorithm-based method, is also unexplainable. These reasons make the clustering algorithm-based and existing boundary methods suffer from issues of misidentification.

Normal distribution model-based methods, proposed to overcome issues of misidentification, use the normal distribution to fit the power data's distribution and then calculate the probability contours. Data with low probability are regarded as abnormal data. Ouyang et al. (2017) calculated mean and standard deviation values based on exponential smoothing. Stephen et al. (2011) used bivariate joint distribution to fit the wind power data's distribution.

However, the aforementioned existing methods for abnormal detection toward improving the model accuracy have the following disadvantages:

- Many real outliers, especially those near the normal region, cannot be recognized.
- There are too many hyperparameters to be set. The performance highly depends on the hyperparameters, while the function of each hyperparameter is unexplainable.
- To ensure performance, the prior information on the normal points should be available, which is not practical in general cases.

Another practical way to quantify uncertainty is to directly use a confidence interval-aware model, known as the interval predictor model (Campi et al., 2009; Garattia et al., 2019). The scenario approach presented by Calafiore and Campi (2006); Campi and Garatti (2019); and (2011); Campi et al. (2015) can be used to establish the interval predictor model. However, the scenario approach cannot give an exact confidence bound since a small number of samples will give a bound with high risk, and a large number of samples will give a conservative bound. Luedtke and Ahmed (2008) proposed a sample average approach to obtain an approximate solution that exactly converges to the original as the sample number increases. In this paper, we propose the sample average approach-based interval models and extend them into extreme learning machines to provide a fast algorithm for training neural networks that can exactly give the desired confidence interval. The proposed interval model solves the abnormal detection and wind power curve regression problems together in a direct way. We implement experimental data-based validation to compare the proposed methods with several existing methods.

The rest of this paper is organized in the following way: Section 2 briefly introduces the wind power curve and then gives a formal problem statement. In Section 3, extreme learning machine and the theory of sample average approach for chance-constrained optimization are briefly reviewed; Section 4 presents the proposed interval models combining the extreme learning machine and sample average approach; and Section 5 presents the results and discussions of experimental data-based validation. Finally, Section 6 concludes the paper.

2 Problem description

As shown in Figure 1, a wind turbine power curve has the following three critical points, namely, A, B, and C, which divide the wind turbine power curve into four segments (Marvuglia and Messineo, 2012; Shokrzadeh et al., 2014). Point A is the cut-in wind speed from where the wind turbine starts to output the power. Point B represents the rated power before which the output power increases as the wind speed increases. Point C is the cut-out speed from where the output wind power decreases even with the increased wind speed.

The segment between points B and C gives the rated level of the wind power output. The segment between points A and B shows a non-linear correlation between wind speed and power. Let v be the wind speed and P be the wind power. The non-linear correlation can be described approximately by the following equation:

$$P = \frac{1}{2} C_p \rho \pi R^2 v^3, \tag{1}$$

where C_p is the coefficient of wind turbine power, ρ is the air density, and R is wind rotor's radius.

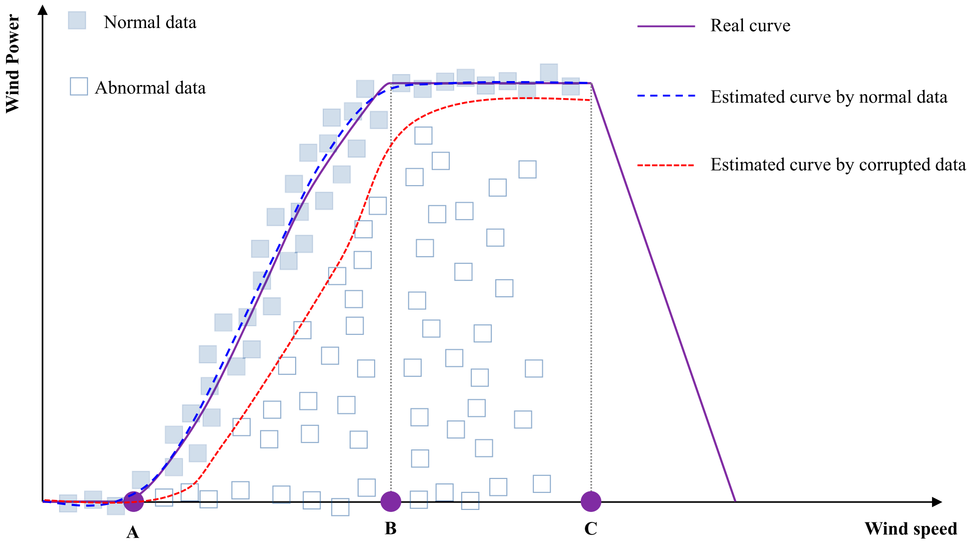


FIGURE 1
Brief illustration of the wind turbine power curve and estimation biases using abnormal data.

A dataset obtained by the SCADA system can be defined by the following equation:

$$\mathcal{D}_N^S = \{(v_i, P_i)\}_{i=1}^N, \tag{2}$$

where N is the number of the data samples. As shown in Figure 1, the SCADA system's data have normal and abnormal data. In other words, $\mathcal{D}_N^S$ is a corrupted dataset. Using the normal data, we can get an estimated curve (blue dashed line) close to the real curve (purple solid line). However, with the corrupted data, the estimated curve (red dashed line) is biased from the real curve.

Therefore, it is necessary to investigate a robust power curve estimation method for the abnormal data. In this paper, we address the following problem.

$$\begin{aligned} &\min_{\mathbf{f}} \sum_{i=1}^{N_c} |P_i - \mathbf{f}(v_i)|^2 \\ &s.t. \quad (v_i, P_i) \in \tilde{\mathcal{D}}_{N_r}. \end{aligned} \tag{3}$$

Here, the set $\tilde{\mathcal{D}}_{N_r}$ is a subset of $\mathcal{D}_N^S$, which is obtained by cleaning the abnormal data. Thus, we have to address the following two sub-problems:

- data cleaning problem to obtain a cleaned data $\tilde{\mathcal{D}}_{N_r}$;
- regression problem described by (3).

3 Preliminaries

This section briefly reviews the extreme learning machine and sample average approach as a preparation for introducing our proposed interval models.

3.1 Extreme learning machine

Extreme learning machine is a fast algorithm to train a single-layer neural network (Huang et al., 2006). A single-layer neural network has an input layer, a hidden layer, and an output layer. Let a positive integer L be the number of neurons. The hidden layer can be defined as a vector function by the following expression:

$$\mathbf{h}(\mathbf{x}) := [h_1(\mathbf{x}), \ldots, h_L(\mathbf{x})]^\top.$$

Each $h_i(\mathbf{x}), i = 1, \ldots, L$ is a neuron. Often, we choose the neuron as follows:

$$h_i(\mathbf{x}) := G(\mathbf{a}_i, \mathrm{b}_i, \mathbf{x}),$$

where $\mathbf{a}_i, \mathrm{b}_i$ are the hyperparameters in the $i-$ th neuron. The neuron can be a sigmoid function or a Gaussian function, etc. Let

$$\boldsymbol{\beta} := [\beta_1, \ldots, \beta_L]^\top$$

be the coefficient of the output layer. Then, we can write the single-layer network as follows:

$$\mathrm{y}(\mathbf{x}) = \boldsymbol{\beta}^\top \mathbf{h}(\mathbf{x}) = \sum_{i=1}^L \beta_i h_i(\mathbf{x}). \tag{4}$$

As a summary, the parameters that need to be trained are coefficient vector $\boldsymbol{\beta}$ and hyperparameters $(\mathbf{a}_i, \mathrm{b}_i)_{i=1}^L$. The extreme learning machine is used to train $\boldsymbol{\beta}$, $(\mathbf{a}_i, \mathrm{b}_i)_{i=1}^L$ with a giving set $\mathcal{D}_N$.

The algorithm of the extreme learning machine is summarized as follows:

- randomly generate hyperparameters $(\mathbf{a}_i, \mathrm{b}_i)_{i=1}^L$;
- estimate $\boldsymbol{\beta}$ by solving

$$\min_{\boldsymbol{\beta}} \sum_{t=1}^N \|\boldsymbol{\beta}^\top \mathbf{h}(\mathbf{x}_t) - \mathrm{y}_t\|^2, \tag{5}$$

which gives the solution as

$$\boldsymbol{\beta}_* := (\mathbf{H}^\top \mathbf{H})^{-1} \mathbf{H}^\top \mathbf{Y}_N, \tag{6}$$

where $\mathbf{H} \in \mathbb{R}^{N \times L}$ is defined as

$$\begin{bmatrix} G(\mathbf{a}_1, \mathrm{b}_1, \mathbf{x}_1) & \ldots & G(\mathbf{a}_L, \mathrm{b}_L, \mathbf{x}_1) \\ \ldots & \ldots & \ldots \\ G(\mathbf{a}_1, \mathrm{b}_1, \mathbf{x}_N) & \ldots & G(\mathbf{a}_L, \mathrm{b}_L, \mathbf{x}_N) \end{bmatrix} \tag{7}$$

and $\mathbf{Y}_N$ is defined as

$$\mathbf{Y}_N := \begin{bmatrix} \mathrm{y}_1 \\ \ldots \\ \mathrm{y}_N \end{bmatrix}. \tag{8}$$

Theorem 2.2 of Huang et al. (2006) provides the universal approximation property of an extreme learning machine-based single neural network regarding a dataset $\mathcal{D}_N$. We summarize it in Lemma 1.

Lemma 1: *For any given small ε and activation function $G(\cdot)$ which is infinitely differentiable in any interval, there exists $\bar{N}$ such that, for $N > \bar{N}$ arbitrary distinct samples $(\mathbf{x}_t, \mathrm{y}_t)$, for any $(\mathbf{a}_i, \mathrm{b}_i)_{i=1}^L$ randomly extracted according to any continuous probability distribution, with probability one,*

$$\sum_{t=1}^N \|\boldsymbol{\beta}_*^\top \mathbf{h}(\mathbf{x}_t) - \mathrm{y}_t\|^2 < \varepsilon. \tag{9}$$

Lemma 1 shows that we could use a single-layer neural network to approximate the wind power curve.

3.2 Sample average approach

Chance-constrained optimization seeks to optimize an objective under a stochastic constraint (Campi et al., 2015; Shen et al., 2020; 2021), which is written as follows:

$$\begin{aligned} &\min_{z \in \mathcal{Z}} J(z) \\ &s.t. \mathrm{Pr}_\xi \{G(z, \xi) \le 0\} \ge 1 - \alpha, \ \xi \in \Xi, \end{aligned} \tag{10}$$

where $z \in \mathcal{Z} \subset \mathbb{R}^{n_u}$ denotes the input variable with the compact feasible set $\mathcal{Z}$, uncertainty is represented by $\xi \in \Xi \subset \mathbb{R}^{n_\xi}$ defined on probability space $(\Xi, \mathcal{B}(\Xi), \mathrm{Pr}_\xi)$, $J(z)$: $\mathcal{Z} \to \mathbb{R}$ and $\forall \xi \in \Xi, G$: $\mathcal{Z} \times \Xi \to \mathbb{R}^l$ are continuous and differentiable objective function and constraint function in z, respectively. Problem (10) is a relaxation of robust optimization in which α is zero. The optimal objective function of (10) is defined by J_{α}^{*}. Let Z_{α}^{*} be the optimal solution set of (10).

The sample-based approximation is a practical way to solve chance-constrained optimization. This paper adopts the sample average approach presented by Luedtke and Ahmed (2008). In the sample average approach, samples are extracted from the sample space Ξ, and then, an approximate problem of the original chance-constrained optimization is established. Let $(\xi^{(1)},\ldots,\xi^{(N)})$ be an independent Monte Carlo sample set of the random variable ξ. After choosing $\epsilon \in [0,1)$ and $\eta > 0$, with sample set $(\xi^{(1)},\ldots,\xi^{(N)})$, the sample average approximation problem is defined as follows:

$$\begin{aligned} &\min_{z\in\mathcal{Z}} \; J(z) \\ &s.t. \; z \in \mathcal{Z}_{\epsilon,\gamma}^{N}, \end{aligned} \tag{11}$$

where $\mathcal{Z}_{\epsilon,\gamma}^{N}$ is defined as follows:

$$\mathcal{Z}_{\epsilon,\gamma}^{N} = \left\{ z \in \mathcal{Z} | \frac{1}{N}\sum_{i=1}^{N} \mathbb{I}\left(G\left(z,\xi^{(i)}\right)+\gamma\right) \geq 1-\epsilon \right\}, \tag{12}$$

where $\Lambda_{\eta}(\cdot)$ is defined by the following equation:

$$\mathbb{I}(t) = \begin{cases} 0, & t > 0, \\ 1, & \text{if } t \leq 0. \end{cases} \tag{13}$$

Let $\tilde{J}_{\epsilon,\gamma}^{N}$ be the optimal objective function of (10) and $\tilde{Z}_{\epsilon,\gamma}^{N}$ be the optimal solution. Note that both $\tilde{J}_{\epsilon,\gamma}^{N}$ and $\tilde{Z}_{\epsilon,\gamma}^{N}$ are decided by the sample set, and the sample set is randomly extracted. Thus, $\tilde{J}_{\epsilon,\gamma}^{N}$ is a random variable, and $\tilde{Z}_{\epsilon,\gamma}^{N}$ is a random set. The uniform convergence of $\tilde{J}_{\epsilon,\gamma}^{N}$ and $\tilde{Z}_{\epsilon,\gamma}^{N}$ should be addressed.

The following assumption on $G(\cdot)$ holds throughout this paper.

Assumption 1: *There exists $L > 0$ such that*

$$|G(z,\xi) - G(z',\xi)| \leq L\|z-z'\|_{\infty}, \; \forall z,z' \in \mathcal{Z} \text{ and } \forall \xi \in \Xi. \tag{14}$$

Assumption 1 is reasonable since we could choose an activation function that makes the neural networks satisfy it and also preserve the universal approximation.

The uniform convergence of $\tilde{J}_{\epsilon,\gamma}^{N}$ and $\tilde{Z}_{\epsilon,\gamma}^{N}$ is summarized from Luedtke and Ahmed (2008).

Lemma 2: *Suppose that Assumption 1 holds as $N \to \infty$, $\gamma \to 0$, and $\epsilon \to \alpha$, $\tilde{J}_{\epsilon,\gamma}^{N} \to J_{\alpha}^{*}$, and $\mathbb{D}(\tilde{Z}_{\epsilon,\gamma}^{N}, Z_{\alpha}^{*}) \to 0$ with probability 1.*

Lemma 2 shows that the approximate problem's solution converges to one in the solution set of the original problem if the number of samples increases to infinite. In addition, for a certain bounded value, we could use a large enough sample number to ensure that the approximate problem's solution is within that bound.

4 Proposed method

This section presents the proposed extreme learning machine with a confidence region. The convergence analysis is given. In addition, the proposed algorithm is presented.

4.1 Extreme learning machine with a confidence region

The previous extreme learning machine gives a single prediction value for a given input. In this paper, we investigate a computation method to give a confidence region for a given input with the center of the confidence region as the estimation of the curve and the normal data located in the confidence region with high probability. In this way, we can solve problem (3). The concept of the extreme learning machine with the confidence region is illustrated in Figure 2

We want to establish an interval model to give a power curve's interval for any given wind speed and require the correct wind power prediction to be within the interval at a given probability. This

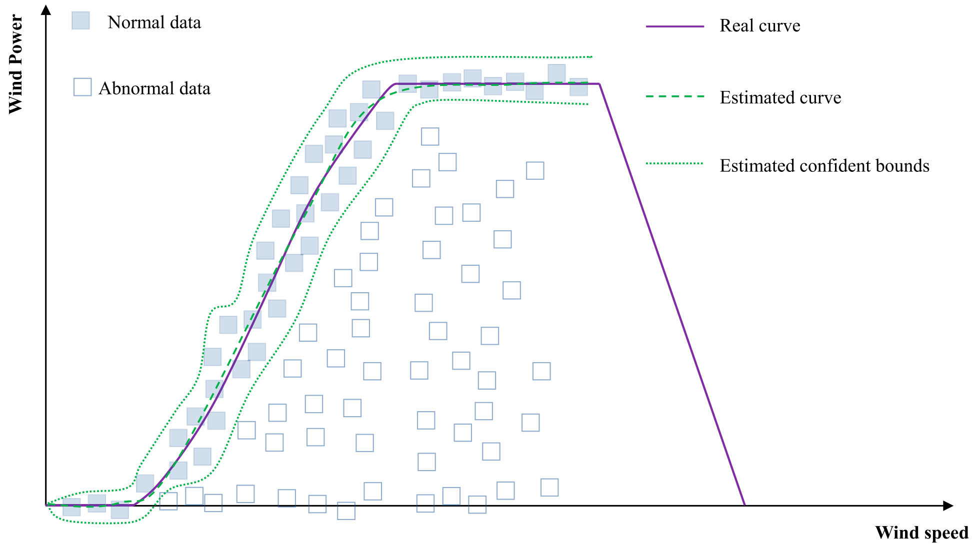


FIGURE 2
Basic concept of the extreme learning machine with a confidence region.

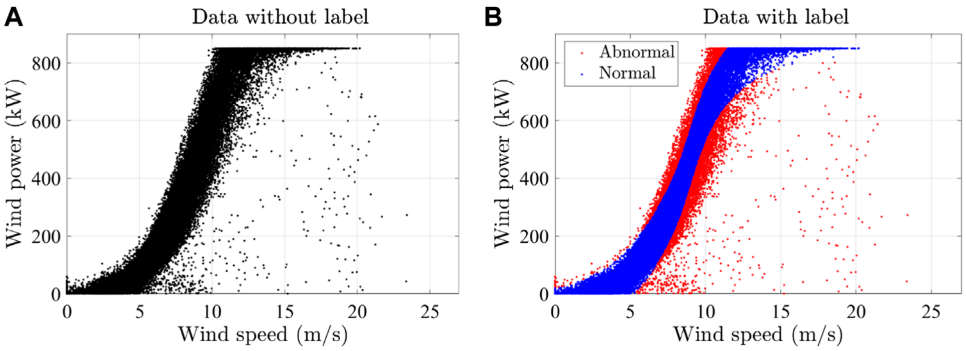


FIGURE 3
Experimental dataset: **(A)** data without labels and **(B)** data with labels.

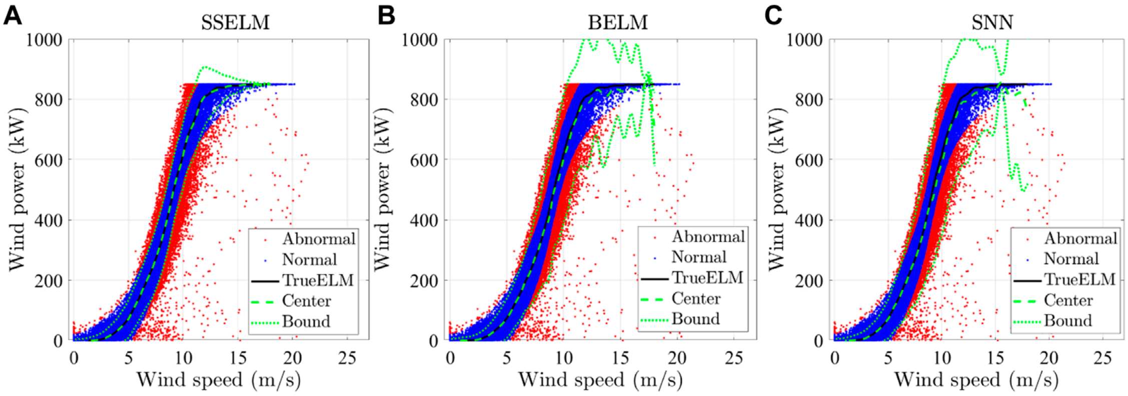


FIGURE 4
Examples of boundaries: **(A)** SSELM, **(B)** BELM, and **(C)** SNN.

interval should be the smallest since a large interval has the issue of being too conservative. An interval model based on a single-layer neural network can be defined by the following equation:

$$\mathcal{I}_{power} := \left\{\tilde{P} = \boldsymbol{\beta}^{\top}\mathbf{h}(v) + e, \boldsymbol{\beta} \in \mathcal{B} \subseteq \mathbb{R}^{L}, e \in [-\gamma, \gamma], \gamma \in \mathbb{R}^{+}\right\}. \tag{15}$$

Let

$$\boldsymbol{\theta} := (\boldsymbol{\beta}, e) \tag{16}$$

be the parameter vector that specifies the interval. We have a set of $\boldsymbol{\theta}$ as $\Theta = \mathcal{B} \times [-\gamma, \gamma]$. The wind power set for a given v and hyperparameters $(\mathbf{a}_i, \mathrm{b}_i)_{i=1}^{L}$ can be obtained by Θ.

In this paper, we use a ball set for $\mathcal{I}_{power}$. With a little manipulation of notation, $\mathcal{I}_{power}$ is also used for the ball interval of wind power. Therefore, the set $\mathcal{B}$ is a ball. Let $\boldsymbol{c}$ and r be the center of $\mathcal{B}$. Then, the set $\mathcal{B}$ can be specified by the following equation:

$$\mathcal{B} = \left\{\boldsymbol{\beta} \in \mathbb{R}^{L} : \|\boldsymbol{\beta} - \mathbf{c}\| \le r\right\}. \tag{17}$$

With $\mathcal{B}$ defined by (17), we can rewrite $\mathcal{I}_{\text{power}}$ in the following form:

$$\begin{aligned}\mathcal{I}_{\text{power}}\left(v, (\mathbf{a}_i, \mathrm{b}_i)_{i=1}^{L}, \boldsymbol{c}, r, \gamma\right) = [&\boldsymbol{c}^{\top}\mathbf{h}(v) - (r\|\mathbf{h}(v)\| + \gamma)\,,\\ &\boldsymbol{c}^{\top}\mathbf{h}(v) + (r\|\mathbf{h}(v)\| + \gamma)].\end{aligned} \tag{18}$$

TABLE 1 Abnormal data detection accuracy (%) of different methods.

	SSELM	BELM	SNN
Accuracy	91.3291	78.9714	63.5733

Then, the problem of solving the extreme learning machine with a confidence region is written as follows:

$$\begin{aligned}&\min_{(\mathbf{a}_i, \mathrm{b}_i)_{i=1}^{L}, \boldsymbol{c}, r, \gamma} \quad \eta r + \gamma\\ &s.t. \quad r, \gamma > 0,\\ &\qquad \mathbb{P}\left\{P(u) \in \mathcal{I}_{\text{power}}\left(v, (\mathbf{a}_i, \mathrm{b}_i)_{i=1}^{L}, \boldsymbol{c}, r, \gamma\right), \forall v\right\} \ge 1 - \varepsilon.\end{aligned} \tag{19}$$

Here, η is a positive number. By using the extreme learning machine algorithm (Lemma 1), we can obtain $(\mathbf{a}_i, \mathrm{b}_i)_{i=1}^{L}$ randomly. Then, after obtaining $(\mathbf{a}_i, \mathrm{b}_i)_{i=1}^{L}$ randomly, we simplify the problem to the following one:

$$\begin{aligned}&\min_{\boldsymbol{c}, r, \gamma} \; \eta r + \gamma\\ &s.t. \quad r, \gamma > 0,\\ &\qquad \mathbb{P}\left\{P(u) \in \mathcal{I}_{\text{power}}\left(v, (\mathbf{a}_i, \mathrm{b}_i)_{i=1}^{L}, \boldsymbol{c}, r, \gamma\right), \forall v\right\} \ge 1 - \varepsilon.\end{aligned} \tag{20}$$

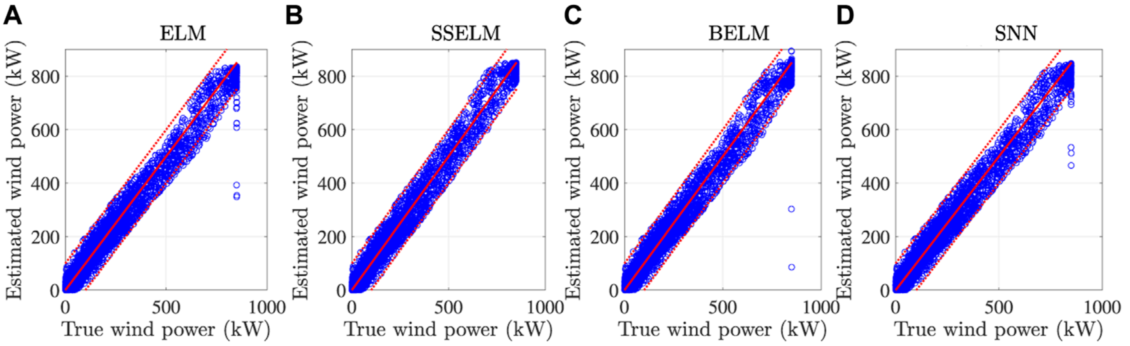


FIGURE 5
Results of wind power predictions on a test set by different methods: **(A)** ELM, **(B)** SSELM, **(C)** BELM, and **(D)** SNN.

TABLE 2 Error statistics of different methods.

	TrueELM	SSELM	ELM	BELM	SNN
Root of MSE	31.7651	31.8510	34.4626	32.9476	34.2956
MAE	21.3421	21.3935	23.3502	22.3977	22.8960

Let **z** and $\mathcal{Z}_\varepsilon$ be the decision variable and feasible region of (20), respectively. Let

$$\ell_\varepsilon^* := \min_{\mathbf{z}\in\mathcal{Z}_\varepsilon} \eta r + \gamma \tag{21}$$

be the optimal objective value of problem (20) and

$$\mathbf{Z}_\varepsilon^* := \left\{\mathbf{z} \in \mathcal{Z}_\varepsilon : \eta r + \gamma = \ell_\varepsilon^*\right\} \tag{22}$$

be the optimal solution set. Let $\mathbf{z}_\varepsilon^* \in \mathbf{Z}_\varepsilon^*$ be an optimal solution. With $\mathbf{z}_\varepsilon^*$, we can obtain $\mathcal{I}_{\text{power}}^*$ for any wind speed v. We can reject all data outside of $\mathcal{I}_{\text{power}}^*$ as abnormal data.

4.2 Sample-based approximation and proposed algorithm

Due to chance constraints, problem (20) is not intractable. With the sample set $\mathcal{D}_N^S$, we can obtain an approximate problem of problem (20) by the following expression:

$$\begin{aligned} \min_{c,r,\gamma} \quad & \eta r + \gamma \\ s.t. \quad & r, \gamma > 0, \\ & \sum_{t=1}^{N} \mathbb{I}\left\{P_t \in \mathcal{I}_{power}\left(v_t, (\mathbf{a}_i, \mathrm{b}_i)_{i=1}^{L}, \boldsymbol{c}, r, \gamma\right)\right\} \geq \left(1-\varepsilon'\right) N. \end{aligned} \tag{23}$$

Here, $\mathbb{I}\{\cdot\}$ is an indicator function defined by (13). Let $\tilde{\ell}_{\varepsilon'}^N$ and $\tilde{\mathbf{Z}}_{\varepsilon'}^N$ be the optimal objective value and optimal solution set of problem (23), respectively.

Theorem 1: *As* $N \to \infty$ *and* $\varepsilon' \to \varepsilon$, $\tilde{\ell}_{\varepsilon'}^N \to \ell_\varepsilon^*$, *and* $\mathbb{D}(\tilde{\mathbf{Z}}_{\varepsilon'}^N, \mathbf{Z}_\varepsilon^*) \to 0$ *with probability 1.*

```
Inputs: Dataset $\mathcal{D}_N^S$, $\varepsilon'$
  1: Solve Problem (23) to obtain $\tilde{\mathbf{z}}_{\varepsilon'}^N \in \tilde{\mathbf{Z}}_{\varepsilon'}^N$
  2: Abnormal data detection
       $\tilde{\mathcal{D}}_{N_r} = \{(v_t, P_t) \in \mathcal{D}_N^S : P_t \in \mathcal{I}_{power}(v_t, (\mathbf{a}_i, \mathrm{b}_i)_{i=1}^L, \tilde{\mathbf{c}}, \tilde{r}, \tilde{\gamma}).\}$   (24)
  3: Estimated curve
       $\tilde{P}_t = \tilde{\boldsymbol{\beta}}^{\top}\mathbf{h}(v_t),\ t = 1, \ldots, N.$   (25)
Output: $\tilde{\mathcal{D}}_{N_r}$, $\tilde{P}_t$, $t = 1, \ldots, N.$
```

Algorithm 1. Proposed algorithm of the extreme learning machine with a confidence region.

Proof. Theorem 1 can be proved by directly applying Lemma 2 since problem (20) satisfies Assumption 1.

The proposed algorithm of the extreme learning machine with a confidence region is presented as follows.

Notice that (25) gives the estimated curve. The confidence bound (upper and lower bounds) can be given using the interval obtained by solving Problem (23).

5 Experimental data-based validation

This section presents the results of experimental data-based validations. First, the experimental dataset is introduced. Then, the results given by the proposed method and several existing methods are compared.

5.1 Experimental data and settings

Figure 3 plots the data used in this validation. The dataset was collected from a wind farm in Hubei, China. Figure 3A shows all data, including abnormal data and normal data. Specialists were approached to give labels on the data set. The data with labels are plotted in Figure 3B.

In this paper, we compare the performance of the following methods:

- **TrueELM**: extreme learning machine using normal data;
- **SSELM**: the proposed extreme learning machine combined with the sample average approach;
- **ELM**: extreme learning machine without data cleaning;
- **BELM**: Bayesian extreme learning machine proposed by Soria-Olivas et al. (2011);
- **SNN**: Neural network trained by the scenario approach presented by Sadeghi et al. (2019).

In **BELM**, the parameter β is assumed to obey some predefined distribution. First, a prior distribution is set. With new data, the prior distribution is adjusted to a posterior distribution so that the data have maximum likelihood. Then, the corresponding output will also have a conditional probability, from which the confidence interval can be calculated. In addition, **SNN** can be regarded as a special case of **SSELM** with $\varepsilon' = 0$.

We evaluate the mean square error (MSE) and mean absolute error (MAE) regarding the normal data in the evaluations.

5.2 Results and discussions

Figure 4 gives the examples of boundaries estimated by **SSELM**, **BELM**, and **SNN**. For each method, 10,000 samples are used. For **SSELM**, the probability threshold is set as 0.09. Each method also gives a corresponding center point of the confidence region. The abnormal data detection performance of different methods is summarized in Table 1. **SSELM** shows a better performance than other methods. The reason that **SNN** shows a poor performance is that it includes more abnormal data since it essentially gives a completely robust interval.

The results of wind power predictions on a text set are plotted in Figure 5. Note that the proposed method, **SSELM**, gives a prediction that concentrates around the real value with a shorter distance. More comprehensive results of error statistics are summarized in Table 2, which shows that the proposed method, **SSELM**, performs very close to the results the method gave using normal data. This shows the effectiveness of the proposed method. The proposed method increases the robustness of the regression against the abnormal data since it can clean the abnormal data effectively and thus increases the accuracy of regression.

6 Conclusion

This paper proposes an interval model of wind turbine power curves to improve the accuracy of wind power prediction. The interval model combines an extreme learning machine and the sample average approach. Thus, the proposed interval model can give a confidence region and center point of the wind power prediction for a given wind speed. The confidence region can be used for abnormal data detection, and the center point can be used as the estimation of the wind turbine power curve point. Experimental data-based validations have been conducted to compare the proposed method with several existing methods. The results show that the proposed method improves the accuracy of both abnormal detection and wind power curve estimation.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

KZ: conceptualization, methodology, project administration, writing–original manuscript, and writing–review and editing. HH: data curation, formal analysis, software, supervision, validation, and writing–review and editing. JL: data curation, investigation, methodology, and writing–review and editing. YW: investigation, validation, visualization, and writing–review and editing. WT: data curation, formal analysis, resources, and writing–review and editing. FH: formal analysis, supervision, validation, and writing–review and editing. YL: methodology and writing–review and editing. RB: data curation, formal analysis, resources, and writing–review and editing. HZ: data curation, investigation, validation, and writing–review and editing. LJ: methodology, supervision, and writing–review and editing.

Funding

This work was supported by the Science and Technology Project of State Grid Hubei Electric Power Company (52152020003K).

Conflict of interest

Authors KZ, HH, JL, YW, WT, FH, YL, RB, HZ, and LJ were employed by State Grid Hubei Extra High Voltage Company.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Calafiore, G., and Campi, M. C. (2006). The scenario approach to robust control design. *IEEE Trans. Autom. Control* 51, 742–753. doi:10.1109/tac.2006.875041

Campi, M., Garatti, S., and Ramponi, F. (2015). A general scenario theory for nonconvex optimization and decision making. *IEEE Trans. Automatic Control* 63, 4067–4078. doi:10.1109/tac.2018.2808446

Campi, M. C., Calafiore, G., and Garatti, S. (2009). Interval predictor models: identification and reliability. *Automatica* 45, 382–392. doi:10.1016/j.automatica.2008.09.004

Campi, M. C., and Garatti, S. (2011). A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality. *J. Optim. Theory Appl.* 148, 257–280. doi:10.1007/s10957-010-9754-6

Campi, M. C., and Garatti, S. (2019). *Introduction to the scenario approach.* Philadelphia: MOS-SIAM Series on Optimization.

Garattia, S., Campib, M., and Care, A. (2019). On a class of interval predictor models with universal reliability. *Automatica* 110, 108542. doi:10.1016/j.automatica.2019.108542

Gilbert, C., Browell, J., and McMillan, D. (2020). Leveraging turbine-level data for improved probabilistic wind power forecasting. *IEEE Trans. Sustain. Energy* 11, 1152–1160. doi:10.1109/tste.2019.2920085

Huang, G., Zhu, Q. Y., and Siew, C. K. (2006). Extreme learning machine: theory and applications. *Neurocomputing* 70, 489–501. doi:10.1016/j.neucom.2005.12.126

Luedtke, J., and Ahmed, S. (2008). A sample approximation approach for optimization with probabilistic constraints. *SIAM J. Optim.* 19, 674–699. doi:10.1137/070702928

Marvuglia, A., and Messineo, A. (2012). Monitoring of wind farms' power curves using machine learning techniques. *Appl. Energy* 98, 574–583. doi:10.1016/j.apenergy.2012.04.037

Morrison, R., Liu, X., and Lin, Z. (2022). Anomaly detection in wind turbine scada data for power curve cleaning. *Renew. Energy* 184, 473–486. doi:10.1016/j.renene.2021.11.118

Ouyang, T., Kusiak, A., and He, Y. (2017). Modeling wind-turbine power curve: a data partitioning and mining approach. *Renew. Energy* 102, 1–8. doi:10.1016/j.renene.2016.10.032

Sadeghi, J., Angelis, M., and Patelli, E. (2019). Efficient training of interval neural networks for imprecise training data. *Neural Netw.* 118, 338–351. doi:10.1016/j.neunet.2019.07.005

Shen, X., Fu, X., and Zhou, C. (2019). A combined algorithm for cleaning abnormal data of wind turbine power curve based on change point grouping algorithm and quartile algorithm. *IEEE Trans. Sustain. Energy* 10, 46–54. doi:10.1109/tste.2018.2822682

Shen, X., Ouyang, T., Yang, N., and Zhuang, J. (2021). Sample-based neural approximation approach for probabilistic constrained programs. *IEEE Trans. Neural Netw. Learn. Syst.* 34, 1058–1065. doi:10.1109/tnnls.2021.3102323

Shen, X., Ouyang, T., Zhang, Y., and Zhang, X. (2020). Computing probabilistic bounds on state trajectories for uncertain systems. *IEEE Trans. Emerg. Top. Comput. Intell. early access* 7, 285–290. doi:10.1109/TETCI.2020.3019040

Shokrzadeh, S., Jafari Jozani, M., and Bibeau, E. (2014). Wind turbine power curve modeling using advanced parametric and nonparametric methods. *IEEE Trans. Sustain. Energy* 5, 1262–1269. doi:10.1109/tste.2014.2345059

Soria-Olivas, E., Gomez-Sanchis, J., Martin, J. D., Vila-Frances, J., Martinez, M., Magdalena, J. R., et al. (2011). Belm: Bayesian extreme learning machine. *IEEE Trans. Neural Netw.* 22, 505–509. doi:10.1109/tnn.2010.2103956

Stephen, B., Galloway, S. J., McMillan, D., Hill, D. C., and Infield, D. G. (2011). A copula model of wind turbine performance. *IEEE Trans. Power Syst.* 26, 965–966. doi:10.1109/tpwrs.2010.2073550

Wang, Y., Srinivasan, D., and Wang, Z. (2019). Wind power curve modeling and wind power forecasting with inconsistent data. *IEEE Trans. Sustain. Energy* 10, 16–25. doi:10.1109/tste.2018.2820198

Wang, Z., and Liu, C. (2021). Wind turbine condition monitoring based on a novel multivariate state estimation technique. *Measurement* 168, 108388. doi:10.1016/j.measurement.2020.108388

Ye, Z., Chen, Y., and Zheng, H. (2021). Understand the effect of bias in deep anomaly detection. *Proc. 30th Int. Conf. Artif. Intell.*, 3314–3320. doi:10.48550/arXiv.2105.07346

Yesilbudak, M. (2016). "Partitional clustering-based outlier detection for power curve optimization of wind turbines," in Proc. IEEE Int. Conf. Renewable Energy Res. Appl, Birmingham, UK, 20-23 November 2016, 1080–1084.

Zheng, L., Hu, W., and Min, Y. (2010). Short-horizon prediction of wind power: a data-driven approach. *IEEE Trans. Energy Convers* 25, 1112–1122. doi:10.1109/tec.2010.2043436

Zheng, L., Hu, W., and Min, Y. (2015). Raw wind data preprocessing: a data mining approach. *IEEE Trans. Sustain. Energy* 6, 11–19. doi:10.1109/tste.2014.2355837

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 29 December 2023
DOI 10.3389/fenrg.2023.1326248

Check for updates

OPEN ACCESS

EDITED BY
Xun Shen,
Osaka University, Japan

REVIEWED BY
Alan Wai Hou Lio,
Technical University of Denmark,
Denmark
Xue Jiang,
Zhejiang University, China

*CORRESPONDENCE
Wenting Chen,
✉ went_chen@ysu.edu.cn

RECEIVED 23 October 2023
ACCEPTED 20 November 2023
PUBLISHED 29 December 2023

CITATION
Wang Q, Du Z, Chen W, Zhang J, Lin Y and Liu H (2023), Incremental feedforward collective pitch control method for wind turbines.
Front. Energy Res. 11:1326248.
doi: 10.3389/fenrg.2023.1326248

COPYRIGHT
© 2023 Wang, Du, Chen, Zhang, Lin and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Incremental feedforward collective pitch control method for wind turbines

Qinwei Wang[1], Zeli Du[1], Wenting Chen[1,2,3]*, Jiarui Zhang[1,2], Yonggang Lin[2] and Hongwei Liu[2]

[1]State Key Laboratory of Crane Technology, Yanshan University, Qinhuangdao, Hebei, China, [2]State Key Laboratory of Fluid Power and Mechatronic Systems, Zhejiang University, Hangzhou, Zhejiang, China, [3]National-Local Joint Engineering Research Center for Advanced Manufacturing, Forming Technology and Equipment, Yanshan University, Qinhuangdao, China

In recent years, wind turbines are becoming larger, which will exacerbate the complexity of loads Complex load change affect the output power quality and wind turbine service life so that must be studied. Pitch control is usually used to reduce wind turbine load. In this paper, based on the Light Detection and Ranging (LiDAR) technology and incremental feedforward control theory, an incremental feedforward collective pitch controller is proposed. The controller can be directly superimposed on the traditional collective pitch controller so that the incremental pitch angle can fully compensate wind influence. The effectiveness of the controller is verified by multi-software platform joint simulation and hardware-in-the-loop experiment. The results show that the controller can effectively reduce the wind turbine power and load fluctuation when the variation trend of wind speed in the rotor plane estimate by LiDAR data is the same as the actual wind speed.

KEYWORDS

wind turbine, Light Detection and Ranging, collective pitch control, incremental feedforward control, predictive control, joint simulation, hardware-in-the-loop experiment

1 Introduction

With the deepening of wind power research, it has been found that the complex load changes of wind turbines caused by wind fluctuations and turbulence characteristics will adversely affect the output power quality and service life of wind turbines (Yuan et al., 2020). Currently, a wind turbine usually uses pitch control to reduce the load on the blades and ensure a smooth output of power. Therefore, in order to cope with these problems, it is necessary to optimize the control of the pitch angle and rotational speed of the wind turbine.

The application of Light Detection and Ranging (LiDAR) introduced wind speed information about reaching the plane to the wind turbine control system. Scholbrock et al. (2016) believed that by introducing this input quantity, the controller control performance could be optimized, and it might also reduce the load of the wind turbine and increase the power generation of the wind turbine. Khaniki et al. (2018) found that LiDAR-assisted feedforward control was better than pseudo-feedforward control. In summary, the collective pitch control technology based on the LiDAR wind measurement can be considered to reduce speed fluctuations and structural loads in the full-load area of wind turbines. It has been proposed that the LiDAR-predicted wind speed can be combined with pitch control to reduce the load of the wind turbine and extend the life of the wind turbine by 6–8 years by some researchers (Mikkelsen, 2014). However, some researchers have pointed out that the excessive pursuit of reducing load and increasing life faces the problem of reducing the net annual energy produced in wind energy (Mathur et al., 2017).

For the moment, the collective pitch control technology based on the LiDAR wind measurement currently has two mainstream research directions, one of which is to completely replace the traditional feedback control technology of the wind turbine. The collective pitch signal has been completely processed using the corresponding algorithm of the wind data measured by LiDAR, such as LiDAR-assisted control (David and Steffen, 2016), collective feedforward pitch control (Haizmann et al., 2015), model predictive control (Bottasso et al., 2014), radial basis function neural network control (Han et al., 2018a), and fuzzy control (Van et al., 2015).

Another research direction has been to keep the original controller unchanged and superimpose a feedforward pitch controller, such as the wind turbine feedforward-feedback H_2 optimal control based on the LiDAR wind measurement (Dunne and Pao, 2016), the LiDAR wind measurement feedforward pitch control combined with the cabin acceleration feedback (Yamaguchi et al., 2020), the joint control mode of the baseline feedback control enhanced by gain scheduling feedforward control based on the pseudo-LiDAR wind measurement (Bao et al., 2018), feedforward controller based on neural network catenary data weights (Han et al., 2018b), and the feedforward control carried out by looking up the static pitch angle curve table after using the LiDAR wind speed to estimate the effective wind speed of the rotor plane (Schlipf et al., 2015).

Through the experiments of different scholars, the positive effect of LiDAR on wind turbines has been proved. For example, the experimental results under different turbulence conditions show that the adaptive control method, that is, the radial basis neural network combined with the finite impulse response filter (RBFNNFIR) based on the LIDAR wind measurement (Jia et al., 2021), can effectively reduce the fluctuation of the generator speed and the maximum value of the load moment at the beginning of the blade. The optimal control of the prediction algorithm combining the linear wind turbine model and LiDAR-simulated wind disturbance can be applied to the non-linear wind turbine model, which can effectively improve the control performance, reduce the fluctuation of generator speed, and meet the pitch activity control limitation (Bao and Yue, 2022).

In this paper, the incremental feedforward collective pitch control method is used to calculate the increment of the pitch angle using the wind speed disturbance increment measured by LiDAR to reduce the impact of wind on the feedback given. This control method can realize the smooth transition of the pitch angle in the feedforward cut-out process. Furthermore, the wind turbine state change caused by the wind speed change is offset by the feedforward incremental pitch angle, which reduces the impact of the wind speed change on the wind turbine. The effectiveness of the proposed method is verified through joint simulation and hardware-in-the-loop experiment.

2 Incremental feedforward collective pitch controller

2.1 Design the incremental feedforward collective pitch control scheme

By using the LiDAR wind measurement technology, the wind speed time series of the rotor plane can be estimated according to the wind information measured by LiDAR, that is, the dynamic change of the wind can be measured. Therefore, the dynamic wind information measured by LiDAR can be used to calculate the feedforward pitch angle to compensate for the change in the load of the wind turbine caused by the wind change.

This paper proposes a wind turbine incremental feedforward collective pitch control method based on LiDAR prediction wind information. In this method, the calculated value is a given value, without artificially adjusting the proportion of the feedforward given in the total pitch angle. So, the controller designed using this control method can be directly added to the original collective pitch controller as a module. The specific steps are described as follows: first, the dynamic model of the wind turbine is established according to the dynamic characteristics. Then, based on the wind information measured by LiDAR, the wind disturbance increment signal about to reach the rotor plane is estimated. Finally, the wind signal in the wind turbine aerodynamic torque model is linearized to design the controller.

The collective pitch control of the wind turbine is shown in Figure 1. The input pitch angle $\Delta\theta$ is the superposition value of $\Delta\theta_{fb}$ and $\Delta\theta_{ff}$. $\Delta\theta_{fb}$ is the feedback pitch angle given by the baseline controller, and $\Delta\theta_{ff}$ is the feedforward pitch angle given by the incremental feedforward collective pitch controller.

Eq. 1 represents the overall equation of the pitch controller.

$$\Delta\boldsymbol{\theta} = \underbrace{\boldsymbol{K_P N_{Gear}\Delta\Omega} + \boldsymbol{K_I}\int_0^t \boldsymbol{N_{Gear}\Delta\Omega dt} + \boldsymbol{K_D N_{Gear}\Delta\dot{\Omega}}}_{\Delta\theta_{fb}} + \Delta\boldsymbol{\theta}_{ff}, \tag{1}$$

where K_P, K_I, and K_D are feedback control parameters of the collective pitch setting of the wind turbine; N_{Gear} is the gear ratio of the wind turbine; and $\Delta\Omega$ is the generator speed difference.

Considering the short sampling interval, the four-order of $\Delta\boldsymbol{v}$ can be regarded as the higher-order infinitesimal term. The incremental feedforward pitch angle $\Delta\theta_{ff}$ can be obtained as shown in Eq. 2.

$$\Delta\boldsymbol{\theta}_{ff} = -\frac{\left(\frac{\partial P}{\partial v}\right)}{\left(\frac{\partial P}{\partial \theta}\right)}\left(\frac{\mathbf{1}}{\mathbf{3!}}\Delta\boldsymbol{v}^3 + \frac{\mathbf{1}}{\mathbf{2!}}\Delta\boldsymbol{v}^2 + \Delta\boldsymbol{v}\right), \tag{2}$$

where Δv is the wind speed increments and $\Delta v = v_1 - v_2$ (where v_1 is the current time wind speed value of the rotor plane estimated from the wind speed measured by LiDAR; and v_2 is the wind speed value at the next sampling time, which is equivalent to linearization at time t_{v1}); $\partial P/\partial v$ is the sensitivity of power to wind speed; and $\partial P/\partial \beta$ is the sensitivity of power to changes in the pitch angle.

The following section will consider the FAST 5-MW wind turbine in FAST software as an example to provide the derivation method of this incremental feedforward collective pitch controller.

2.2 Theoretical derivation of incremental feedforward controllers

2.2.1 Onshore wind turbine baseline collective pitch controller

The baseline collective pitch controller is derived from the dynamic model of the wind turbine. Finally, the variable proportional gain PI feedback controller (Jonkman et al., 2009a)

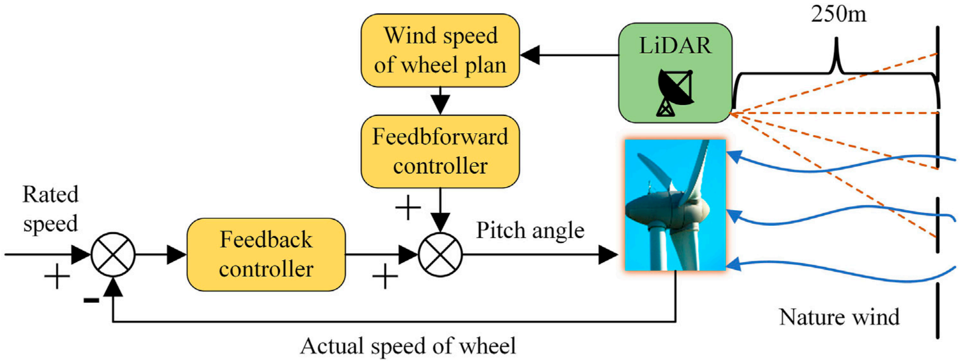


FIGURE 1
Incremental feedforward control assisted collective pitch baseline controller.

is obtained to control the constant speed of the generator. The dynamic model of the wind turbine wind turbine rotor can be expressed as

$$T_{Areo} - N_{Gear}T_{Gen} = \left(I_{Roter} + N_{Gear}^2 I_{Gen}\right)\frac{d}{dt}\left(\Omega_0 + \Delta\Omega\right) = I_{Drivertrain}\Delta\dot{\Omega}, \quad (3)$$

where T_{Areo} is the aerodynamic torque captured by the wind turbine rotor in the wind field; N_{gear} is the gear ratio of the wind turbine; T_{Gen} is the generator torque; I_{Rotor} is the equivalent moment of inertia of the wind turbine rotor; I_{Gen} is the equivalent moment of inertia of the generator; Ω_0 is the stable speed of the generator; $\Delta\Omega$ is the generator speed increment; and $I_{Drivertrain}$ is the equivalent moment of inertia of the drive chain.

Generator torque can be expressed as

$$T_{Gen}\left(N_{Gear}\Omega\right) = \frac{P_0}{N_{Gear}\Omega}, \quad (4)$$

where Ω is the generator speed and P_0 is the generator rated power.

The aerodynamic torque of the wind turbine can be expressed as

$$T_{Areo}\left(\theta\right) = \frac{P\left(\theta, \Omega_0\right)}{\Omega_0}, \quad (5)$$

where θ is the current pitch angle of the three blades of the wind turbine and the capture power of the wind turbine rotor $P(\theta, \Omega_0)$ is the function of its pitch angle and the generator speed.

The linearization equation of generator torque and the wind turbine rotor at the rated speed can be expressed as

$$T_{Gen} \approx \frac{P_0}{N_{Gear}\Omega_0} - \frac{P_0}{N_{Gear}\Omega_0^2}\Delta\Omega, \quad (6)$$

$$T_{Areo} \approx \frac{P_0}{\Omega_0} + \frac{1}{\Omega_0}\left(\frac{\partial P}{\partial \theta}\right)\Delta\theta_{fb}, \quad (7)$$

where $\Delta\theta_{fb}$ is the small disturbance in the pitch angle of the blades due to feedback control near the operating point.

For PID control, the feedback pitch angle $\Delta\theta_{fb}$ is related to the disturbance of the wind turbine rotor speed. So, the baseline collective pitch controller of the wind turbine can be expressed as

$$\Delta\theta_{fb} = K_p N_{Gear}\Delta\Omega + K_I\int_0^t N_{Gear}\Delta\Omega dt + K_D N_{Gear}\Delta\dot{\Omega}, \quad (8)$$

where K_P, K_I, and K_D are feedback control parameters for traditional collective pitch settings of the wind turbine. The specific parameter derivation process has been expounded by Jonkman et al. (2009b).

2.2.2 Theoretical derivation of the incremental feedforward collective pitch controller

The wind turbine rotor dynamic model of the FAST's 5-MW wind turbine is obtained under stable wind speed. The incremental feedforward collective pitch controller can be derived. First of all, wind speed in Eq. 5 is to be substituted by the effective rotor plane wind speed measured by LiDAR. Therefore, Eq. 9 can be obtained.

$$T_{Areo}\left(\theta\right) = \frac{P\left(\theta, \Omega_0, v\right)}{\Omega_0}. \quad (9)$$

The wind speed signal in Eq. 9 is linearized at the working point, and the high-order infinitesimal term of the wind speed change is ignored. Simulating Eqs 3, 7, 9, we can obtain the following equation:

$$\left(\frac{P_0}{\Omega_0} + \frac{1}{\Omega_0}\left(\frac{\partial P}{\partial \theta}\right)\Delta\theta_{fb} + \frac{1}{\Omega_0}\left(\frac{\partial P}{\partial v}\right)\frac{1}{3!}\Delta v^3 + \frac{1}{\Omega_0}\left(\frac{\partial P}{\partial v}\right)\frac{1}{2!}\Delta v^2 + \frac{1}{\Omega_0}\left(\frac{\partial P}{\partial v}\right)\Delta v\right) - N_{Gear}\left(\frac{P_0}{N_{Gear}\Omega_0} - \frac{P_0}{N_{Gear}\Omega_0^2}\Delta\Omega\right) = I_{Drivertrain}\Delta\dot{\Omega}. \quad (10)$$

The incremental feedforward collective pitch controller can be obtained using Eq. 10. Before that, some assumptions need to be made:

The speed of the wind turbine has been stabilized through the feedback pitch angle $\Delta\theta_{fb}$, that is, the speed increment of the wind turbine is zero, and the speed increment derivative is zero.

In this steady state, the change in the rotor speed caused by the wind speed disturbance is completely offset by the feedforward pitch angle $\Delta\theta_{ff}$.

(1) The LiDAR measurement sampling time is small enough, and the sampling time can be aligned with the sampling time of the wind turbine's own sampling system.

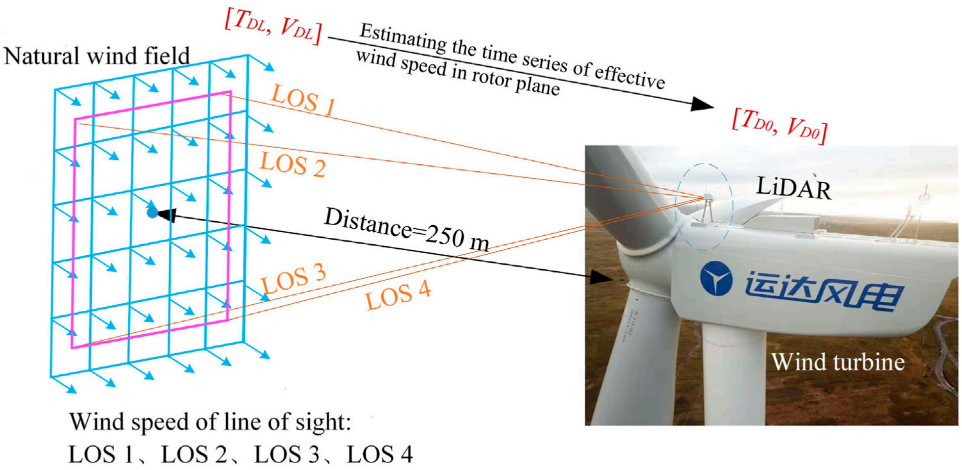


FIGURE 2
Diagram of LiDAR wind data preprocessing.

(2) The effective rotor plane wind speed estimated by the wind speed measured by LiDAR is exactly the same as the actual wind speed.
(3) The wind energy utilization coefficient model of the wind turbine is a collective and common mathematical simplification model.

On the basis of the assumption condition (1), that is, $\Delta\Omega = \Delta\dot{\Omega} = 0$, Eq. 11 can be obtained from Eq. 10.

$$\frac{1}{\Omega_0}\left(\frac{\partial P}{\partial \theta}\right)\Delta\theta_{ff} + \frac{1}{\Omega_0}\left(\frac{\partial P}{\partial v}\right)\frac{1}{3!}\Delta v^3 + \frac{1}{\Omega_0}\left(\frac{\partial P}{\partial v}\right)\frac{1}{2!}\Delta v^2 + \frac{1}{\Omega_0}\left(\frac{\partial P}{\partial v}\right)\Delta v = 0. \tag{11}$$

According to the assumptions, Eq. 11 can be used to obtain the incremental pitch angle (2). This controller can be added as a module to the traditional collective pitch controller to obtain Eq. 1, the incremental feedforward collective pitch controller. Furthermore, the calculated value is the given value, without artificially adjusting the proportion.

2.3 LiDAR wind measurement data pre-processing

LiDAR can only measure the wind information at a certain distance in front of the wind turbine rotor. However, for the wind turbine control, the best predicted wind information is the wind information that will arrive at the rotor plane exactly in the time required by the wind turbine pitch mechanism. Therefore, the LiDAR measurement needs to be processed accordingly (as shown in Figure 2). Then, the pitch control is made in advance using the wind information time series estimated in the rotor plane.

Under the condition that LiDAR has built in the line-of-sight wind speed and wind field reconstruction method, the main aim is to estimate the effective rotor plane wind speed by inverting the wind speed at the measurement point by processing the wind frequency, amplitude, and phase.

First, the filter is used to deal with the wind speed frequency measured by LiDAR so that the frequency unrelated to the wind speed of the rotor plane can be filtered (Schlipf, 2015). The filtering expression is shown in Eq. 12.

$$G_{filter}(z) = \frac{b_0 + b_1 z^{-1}}{a_0 + a_1 z^{-1}}. \tag{12}$$

Schlipf (2015) proposed a method for determining constants, b_0,b_1, a_0, and a_1.

Second, the wind speed amplitude of the rotor plane is estimated using the wind speed attenuation model of the induction area of the wind turbine (Zhang et al., 2021). Specific details are shown in Eq. 13.

$$\frac{v}{v_\infty} = \begin{cases} 1.004\left[1 - \frac{1}{3}\left(1 + k\left(1 + k^2\right)^{-\frac{1}{2}}\right)\right] k \le -1.4, \\ 1 - \frac{6.0}{v_\infty}\left(\frac{1}{3}\left(1 + k\left(1 + k^2\right)^{-\frac{1}{2}}\right)\right) - 1.4 < k \le 0, \end{cases} \tag{13}$$

where v is the effective wind speed of the rotor plane, v_∞ is the free-flow wind speed, κ is the relative distance, and $\kappa = x/R$ (x is the distance from the rotor plane, and x is positive for the upstream of the wind turbine rotor and negative for downstream; R is the radius of the rotor plane).

Finally, the wind speed phase estimation in the rotor plane mainly refers to the time when LiDAR measures the wind speed to reach the rotor plane. Its calculation method contains the use of Eq. 13 and differential equations for separable variables.

$$\int_{x_0}^{0} G'(x)dx = v_\infty t_{mov}, \tag{14}$$

where x_0 is the distance between the LiDAR measurement point and the rotor plane; $G(x)$ is the analytical solution of distance and time information function of the wind speed attenuation model,

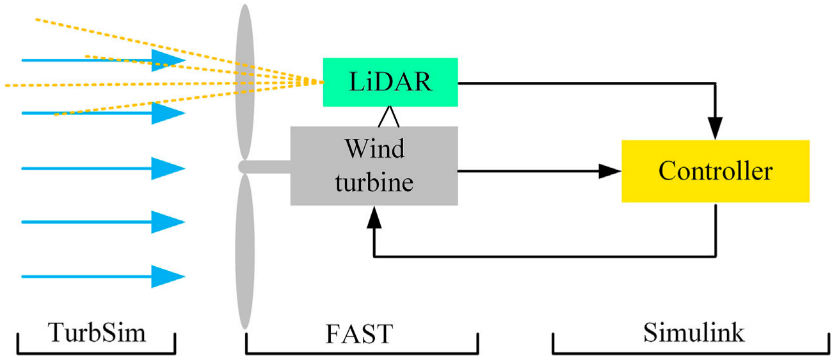


FIGURE 3
FAST/Simulink joint simulation diagram.

TABLE 1 Parameters of the 5-MW wind turbine.

Item	Description
Rated power	5 MW
Number of blades	3
Rotor direction	Upwind
Drive chain	Multistage gearbox
Transmission ratio	97
Diameter of the rotor	126 m
Diameter of the hub	3 m
Height of the engine room	90 m
Cut-in wind speed, rated wind speed, and cut-out wind speed	3 m/s, 11.4 m/s, and 25 m/s
Starting speed and rated speed of the rotor	6.9 r/min and 12.1 r/min
Collective pitch controller time constant τ	0.11 s

$G'(x) = v_\infty/v$; and t_{mov} is the time required for the wind speed at the LiDAR measurement point to move to the rotor plane.

The rotor plane effective wind time series is estimated using the above method (Zhang et al., 2022), as follows:

$$\left[\boldsymbol{t_{ori}} + \boldsymbol{t_{mov}} - \boldsymbol{t_{filter}},\ (\mathbf{1} - \boldsymbol{m})\boldsymbol{v}_\infty(\boldsymbol{t})\right], \tag{15}$$

where t_{ori} is the LiDAR measurement at the time point, t_{filter} is the time delay caused by the filter filtering process, and m is the wind speed attenuation coefficient calculated according to Eq. 13.

3 FAST/Simulink joint simulation

3.1 Simulation platform

As shown in Figure 3, FAST [an open-source software tool from the U.S. Department of Energy's National Renewable Energy Laboratory (NREL)] provides mathematical models and kinetic operations for a 5-MW onshore wind turbine, and TurbSim to simulate the wind farm, while FAST and Simulink are connected to exchange data via interfaces. After the control system built in Simulink takes the input signal from FAST, the control module calculates the control target which is fed into the FAST 5-MW wind turbine model. The wind turbine makes corresponding actions accordingly. The parameters of the 5-MW wind turbine are shown in Table 1.

3.2 Simulation results under ideal estimation of the rotor plane wind speed

During the simulation, it is assumed that the rotor plane wind speed estimated by LiDAR is the same as the actual wind speed of the rotor plane. In the application of feedforward controllers, the given signal increments are corrected so that the given paddle pitch angle signal of the feedforward controller has the least impact on the baseline feedback controller. Two correction methods are considered during the simulation. One is to limit the amplitude of the feedforward signal, and the other is to consider the correction coefficient method in the baseline collective pitch controller (Zhang et al., 2022).

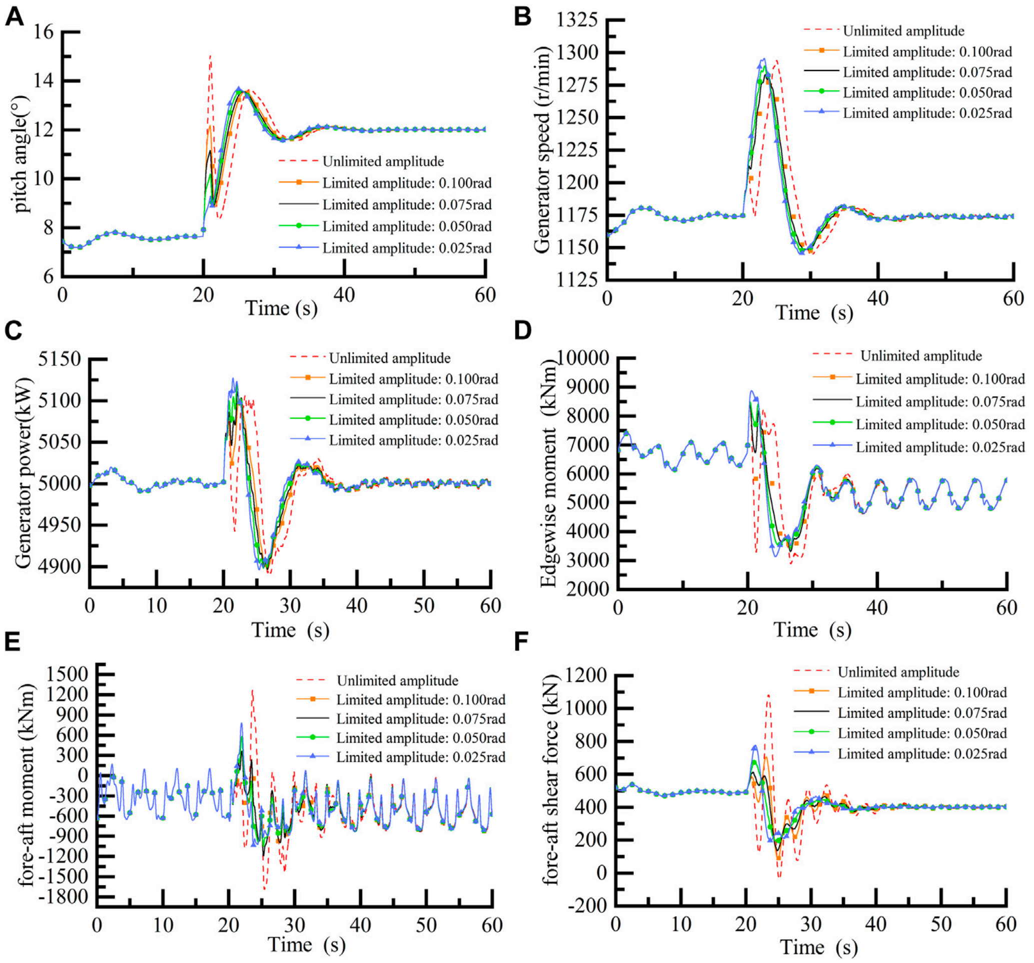


FIGURE 4
Wind turbine response with different limit values.

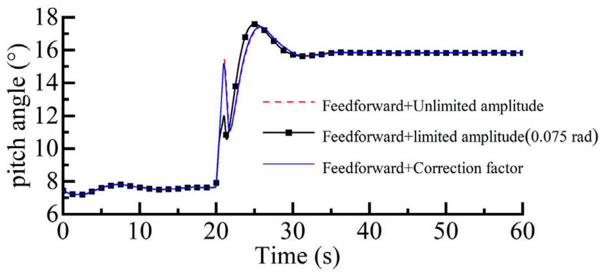


FIGURE 5
Wind turbine response under step wind speed.

The amplitude limiting method is used to increase the amplitude limitation for the feedforward signal and observe whether the control presetting of the baseline feedback controller has enough stiffness to quickly restore the stability after the feedforward control gives a small perturbation. When a step wind speed is given as 12 m/s–15 m/s, the incremental feedforward controller limitation is, respectively, set as 0.025 rad, 0.050 rad, 0.075 rad, and 0.1 rad, and the system response is shown in Figure 4. The overall control effect of the wind turbine with limiting is better than the state without limitation. The generator speed fluctuation performance is the best in amplitude limitation being 0.1 rad. The turbine power and edgewise moment are in good condition in a limitation of 0.050 rad. Fore-aft moment and fore-aft shear force have the most pronounced volatility improvement in a limitation of 0.075 rad. With comprehensive consideration, the limitation is determined to be 0.075 rad.

When the given step wind speed is 12–18 m/s, it can be seen from Figure 5 that the pitch angle has a relatively large fluctuation under feedforward control with the correction factor or unlimited amplitude. However, the pitch angle has a relatively gentle variation with the 0.075-rad amplitude limitation. Therefore, in the process of adding the incremental feedforward controller into the wind turbine baseline collective pitch, it needs to limit the given incremental feedforward pitch angle amplitude to the range of 0.075 rad.

Referring to the IEC Standard, the given gust signal and wind turbine response are shown in Figure 6. The LiDAR gust warning function is also added to the incremental feedforward controller,

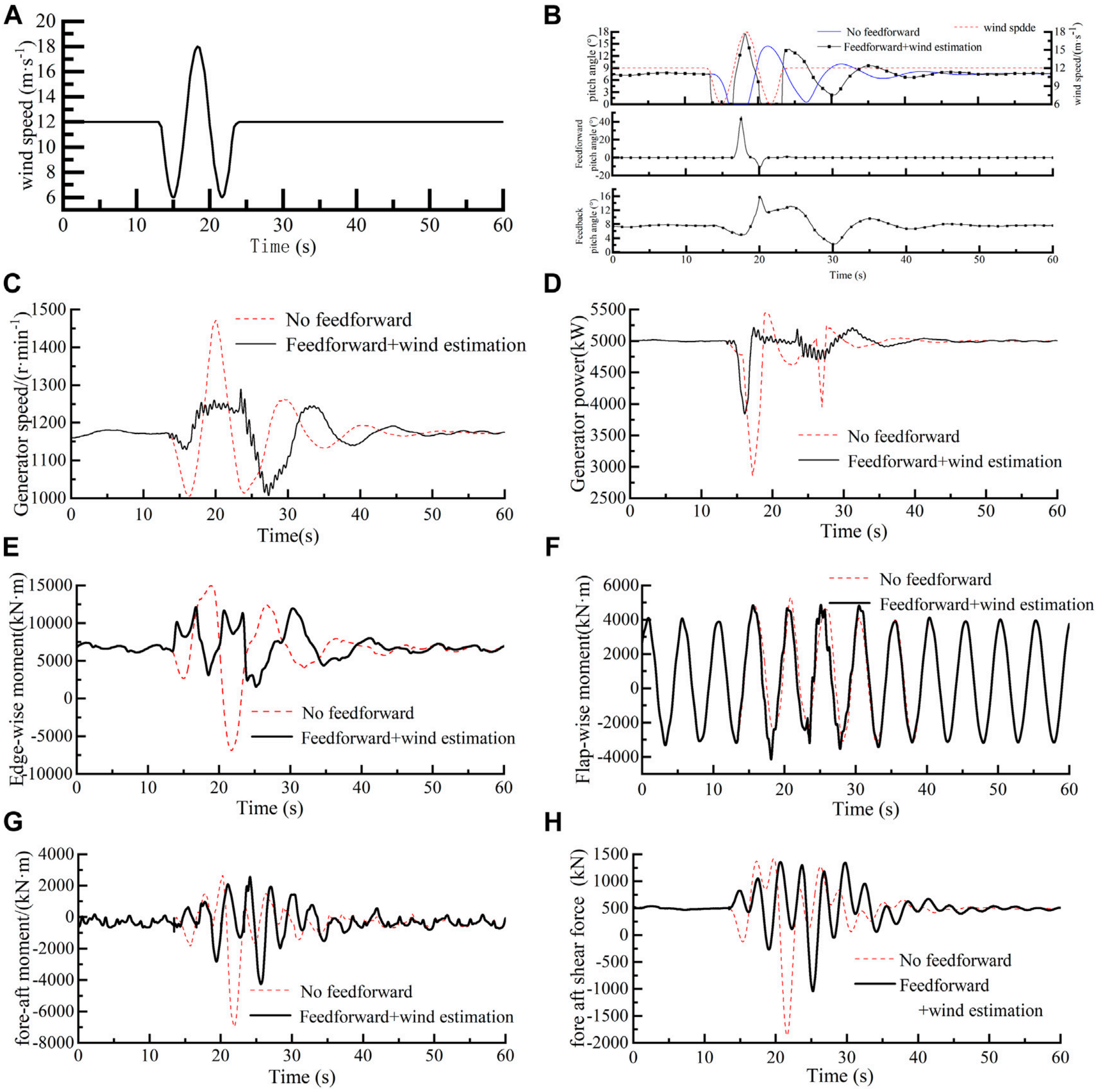


FIGURE 6
Wind turbine response under IEC gust input.

that is, when the wind speed measured by LiDAR is greater than the cut-in wind speed 3 m/s, or less than the rated wind speed 11.4 m/s, the pitch angle of the wind turbine is directly set to 0°, and when $v < 3$ m/s or $v \geq 25$ m/s, the pitch angle is directly set to 90°.

It can be seen from Figure 6 that the generator speed fluctuation decreases by 39.42%, and the peak decreases by 182.216 r/min. Generator power fluctuations are reduced by 46.82% and peaks reduced by 231.82 kW. The range of edgewise moment fluctuation decreases by 51.70%. However, the range of flap-wise moment fluctuation increases by 4.5%. The fore-aft moment fluctuation decreases by 51.91%, and the fore-aft shear force fluctuation range decreases by 56.59%.

In summary, even if it has a slight enhancement effect on the flap-wise moment fluctuation, the incremental feedforward and gust warning can effectively reduce the load of the wind turbine system, improving the performance of the wind turbine, diminishing the speed and power fluctuation of the generator, and reducing the moment and force fluctuation of the fore-aft load.

4 Experimental result analysis

The model of the wind turbine in the previous theoretical and simulation study uses NWTC's open-source simulation software FAST 5-MW onshore wind turbine. It can be called through an interface to be a module in Simulink. The Beckhoff PLC of the current new version supports Simulink programming and the real-time data interaction. The structure of the hardware-in-the-loop experiment system is shown in Figure 7. The LiDAR in the experimental system has no physical object, so it uses the real experimental data collected by the LiDAR of Molas NL in the wind farm. Since Molas NL can simultaneously measure the wind speed and wind direction information at

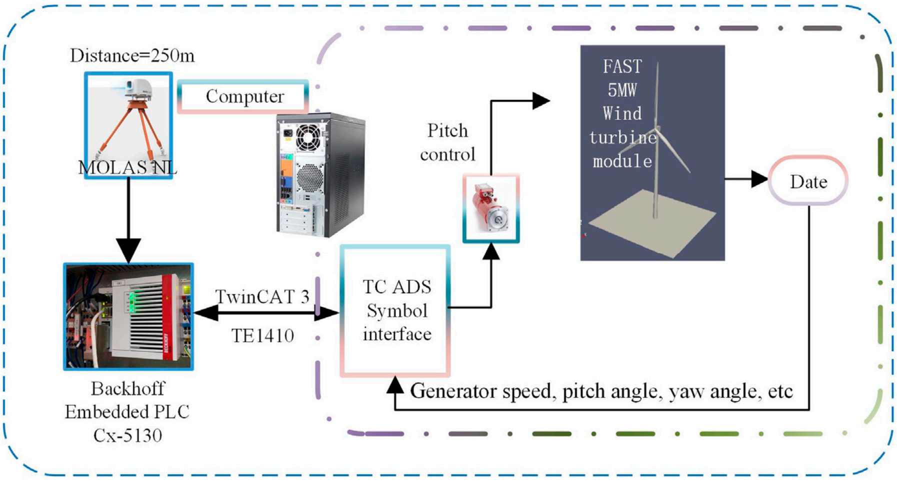


FIGURE 7
Frame of the hydraulic-mechanical transmission test rig.

TABLE 2 Semi-physical simulation experiment software and hardware system selection.

Model	Product description
CX5130-0125	Embedded PLC
CX2900-0038	Memory card (40G)
TC1220-0240	PLC/C++/MATLAB/Simulink, Performance Class 40
EL3008	Bus module
EL4008	Bus module
FASTv8.16.00.a-bjj	Wind turbine simulation software
MATLAB/Simulink R2016a	Visual simulation software

10 distances in front of the wind turbine rotor, LiDAR is used to measure the wind data at 250 m in front of the wind turbine rotor as the LiDAR prediction wind data to input Beckhoff Embedded PLC-CX5130, and the wind data at 50 m in front of the wind turbine rotor as the real data on the rotor plane is input into the rotor plane wind speed file of the wind turbine simulation model.

The selected software and hardware models are shown in Table 2. The sensor interface of the wind turbine in practical applications is not used during the hardware-in-the-loop experiment. During the application process, the Beckhoff Embedded PLC, which is a scalable PLC system, can add the corresponding interface to the sensor data, according to the sensor type to achieve control applications. For example, Molas NL wind-measuring LiDAR supports a variety of communication protocols, and the version of the early LiDAR communication protocol has PROFIBUS and CANopen, which can be ordered based on needs, while the communication protocol interface of the Beckhoff Embedded PLC master control system can also be expanded as needed so that the two can communicate normally and effectively. In summary, the final communication mode is determined according to the overall selection of the components of the wind turbine.

In the hardware-in-the-loop experiment, the real data measured by LiDAR are used to estimate the rotor plane wind speed for control research. This paper collected the data measured by LiDAR that is higher than 11.4 m/s from after 19 May 2019, 2:00 a.m. In this file, special step wind speeds, gusts, and continuously varying winds are selected to observe the system response with estimation errors.

As shown in Figure 8, the effective wind speed of the rotor plane is estimated using the wind speed measured by LiDAR at 250 m in front of the wind turbine rotor. In the time period of 16 s–25 s, the wind speed estimated by the step wind deviates greatly from the actual wind speed. At 20 s, the difference between the estimated wind speed and the actual wind speed is −1.34 m/s. At 24 s, the difference between the estimated wind speed and the actual wind speed is 1.35 m/s. Although the difference between the estimated wind speed and the actual wind speed is large, over this time period, the trend of the estimated wind speed remains basically the same as the actual wind. On the basis of wind speed estimation, comparing the effect of feedforward with no feedforward, the generator speed fluctuation is reduced by 17.13% and the peak by 21.52 r/min; generator power fluctuations decreased by 6.8%, and peaks, by 0.024 kW; and the swinging torque fluctuation decreased by 35.09% and the peak by 823.36 kN·m. However, the step wind speed estimated by the wind turbine at approximately 40 s has a significant lag, so the performance of the wind turbine does not improve due to wrong time estimation. However, at approximately 50 s, due to the large errors in estimating the wind speed trend, the wind turbine pitch angle is mishandled, which results in larger power fluctuations of the wind turbine. The above analysis shows that the incremental feedforward control works only when the wind speed estimated by LiDAR has the same change time and trend of the actual wind speed.

During the process, it is found that the wind turbine operation process had a great influence on the wind speed measurement at 50 m in front of the wind turbine rotor for the rotor movement. As shown in Figure 8A, at the moment of the wind step, the measurement trend of LiDAR measuring in

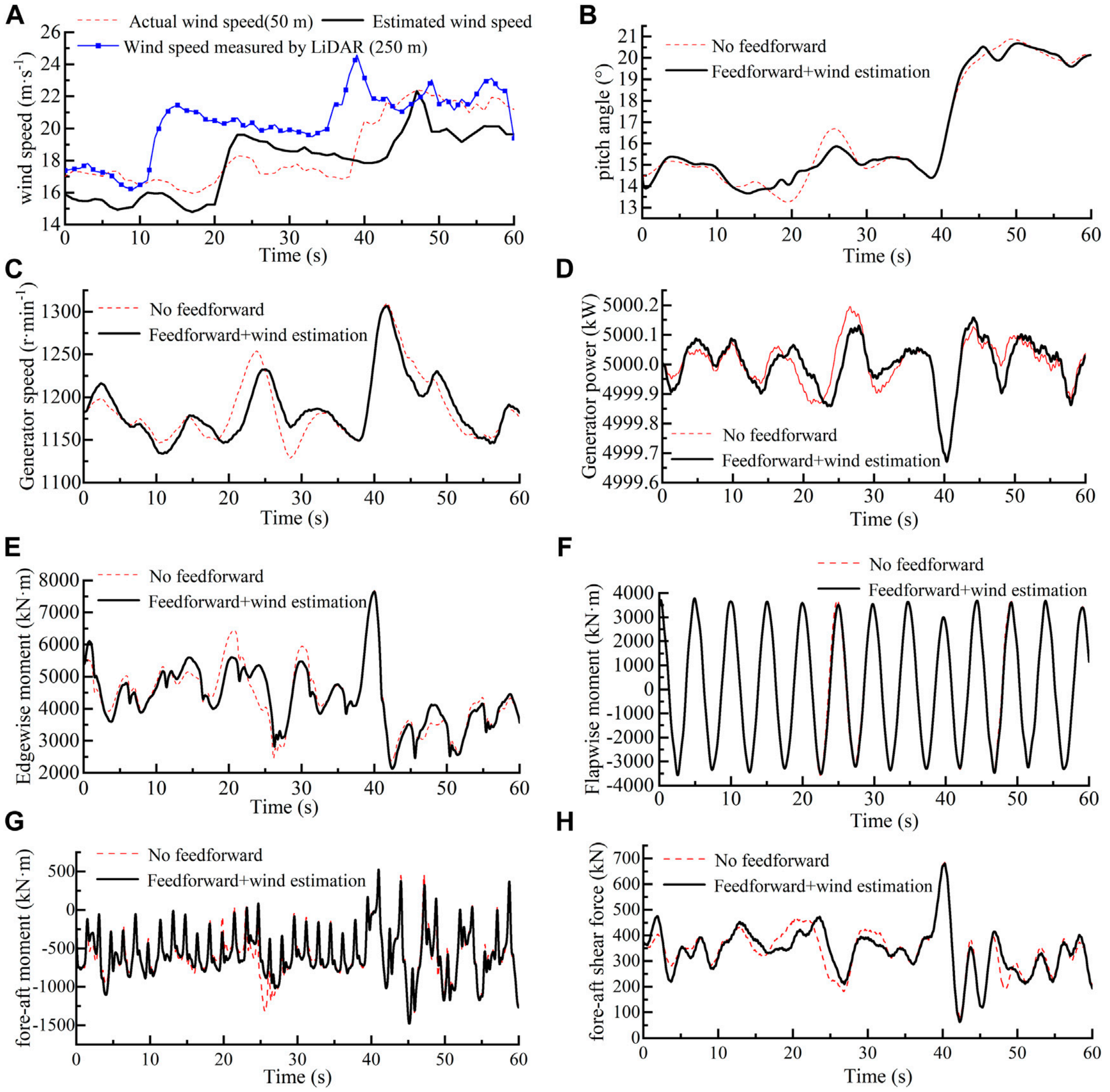


FIGURE 8
Experimental results of step wind speed.

250 m is consistent with 50 m, but the degree of attenuation is significantly different. On the one hand, the closer the wind speed point of the LiDAR measurement to the rotor plane, the greater the error. On the other hand, the control action of the wind turbine after the gust reaches the rotor plane affects the wind speed movement and attenuation form in front of the wind turbine rotor.

In the same day, this paper selects a period of continuously varying wind speeds with a better estimated trend. Since it is hard to directly obtain the rotor plane wind speed, the wind speed at a distance of 50 m upstream is considered the rotor plane actual speed. Furthermore, we used the wind speed of 250 m upstream to estimate the 50 m position wind speed at the same time. As shown in Figure 9A, at approximately 20 s, the wind speed measured by LiDAR at 250 m is inconsistent with the wind speed trend at 50 m, that is, the estimated wind speed curve is inconsistent with the actual wind speed curve trend of 50 m. It results in the incremental feedforward control misoperation. Moreover, it causes the rotational speed, power, and load in the incremental feedforward control have a greater fluctuation and peak than that in the non-incremental feedforward controller. Before 10 s and after 25 s, the trend of the estimated wind speed is consistent with the actual wind speed. It can be observed in Figure 9 that the generator speed fluctuation is reduced by 1.86%, the generator power fluctuation by 18.5%, the edge-wise moment fluctuation by 2.02%, and the fore-aft shear force fluctuation by 11.13%. The estimated wind speed of approximately 30 s is relatively close to the actual wind speed, so when the feedforward is active, the generator speed amplitude is reduced by 12.11 r/min, the generator power amplitude by 0.054 kW, the edge-wise moment amplitude by 417.85 kN·m, and the fore-aft shear force amplitude by 71.13 kN.

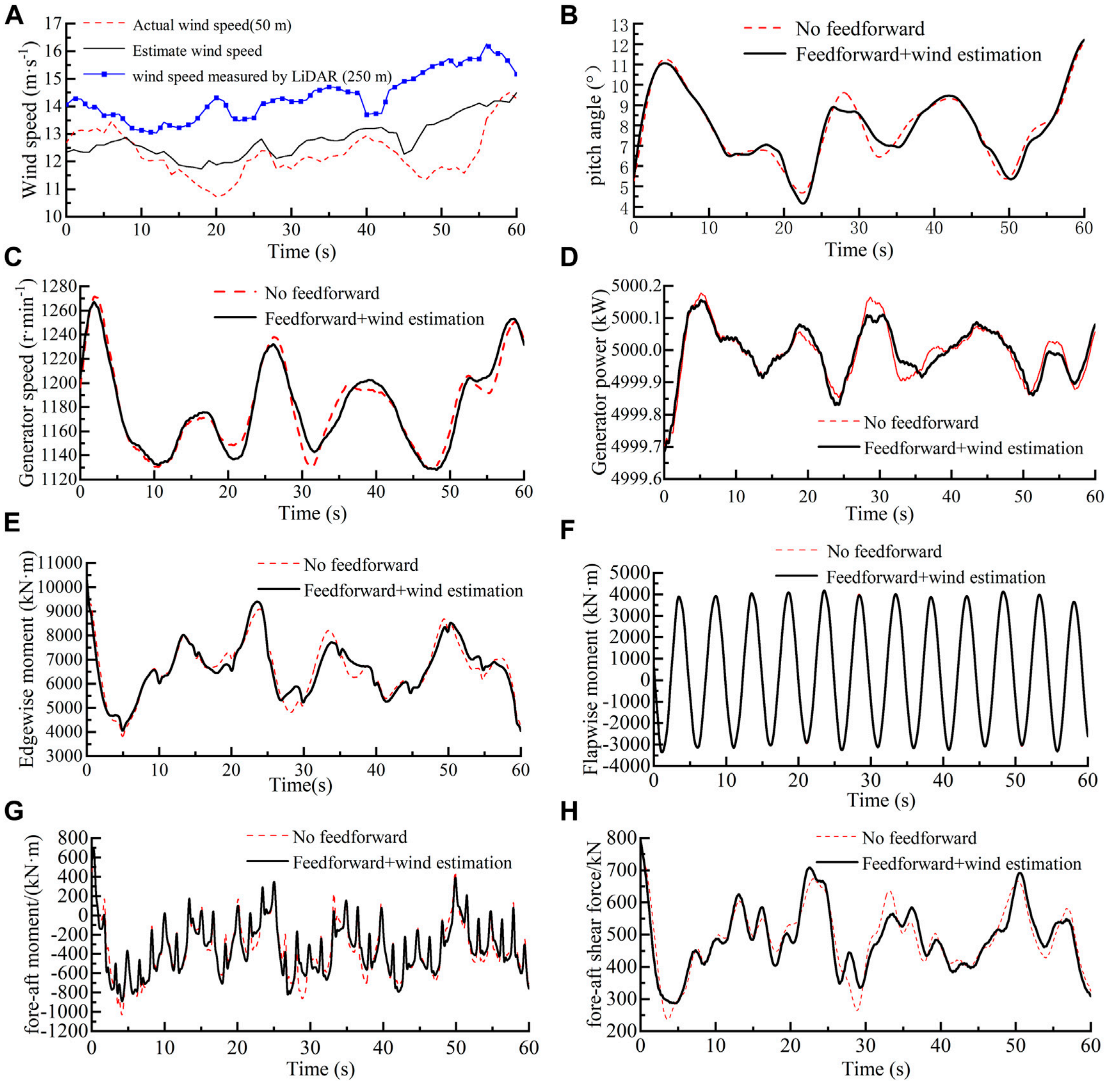


FIGURE 9
Experimental results of continuously varying wind speed.

From the above hardware-in-the-loop experiment, it can be seen that the incremental feedforward collective pitch controller based on LiDAR proposed in this paper can reduce the fluctuation range of the generator speed, power, and load. The precondition for the controller to play a role is that the estimated effective wind speed of the rotor plane should conform to the actual wind speed, and the specific values that complete consistency are not required; however, the trend should be consistent. This makes the specific requirements for wind speed estimation.

5 Conclusion

This paper derives the incremental feedforward collective pitch controller, which is superimposed on the baseline collective pitch controller to offset the impact of wind speed changes on the wind turbine. The joint simulation model of the FAST/Simulink 5-MW wind turbine is designed and established, and the incremental feedforward controller is simulated and verified under the input conditions of step wind and gust. The simulation results show that it can effectively reduce the unit load when the controller is with a limit range of 0.075 rad. For example, the load fluctuation range of gust conditions can be reduced by approximately 50%. The hardware-in-the-loop experiment results show that the incremental feedforward controller can achieve load reduction when the trend of the estimated wind speed is the same as the actual wind speed with no delay. The incremental feedforward controller is a pitch angle increment calculated by predicting wind disturbances, so it can be added as a module to the traditional collective pitch controller to reduce the system load. The modularity is suitable for commercial applications.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

QW: writing–original draft, methodology, and validation. ZD: writing–review and editing, data curation, and investigation. WC: funding acquisition, resources, writing–review and editing, methodology, project administration, and writing–original draft. JZ: writing–review and editing, investigation, software, and validation. YL: resources, writing–review and editing and project administration. HL: writing–review and editing, project administration, and validation.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Natural Science Foundation of China (U22A20178, 52205071), the Open Foundation of the State Key Laboratory of Fluid Power and Mechatronic Systems (GZKF-202110), the Hebei Science Foundation for Distinguished Young Scholars under grant number (E2021203020), the Central Government Guided Local Scientific and Technological Development Special Fund Project (ZYYD 2022C09), the Zhejiang Province Basic Public Welfare Research Program under grant number (LHZ21E090004).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Bao, J., and Yue, H. Design and assessment of a LIDAR-based model predictive wind turbine control[J]. *Energies*, 2022, 15(17): 6429, doi:10.3390/en15176429

Bao, J., Yue, H., Leithead, W. E. L., and Wang, J. Q. Feedforward control for wind turbine load reduction with pseudo-LiDAR measurement[J]. *Int. J. Automation Comput.*, 2018, 15 (2): 142–155. doi:10.1007/s11633-017-1103-x

Bottasso, C. L., Pizzinelli, P., Riboldi, C. E. D., and Tasca, L. (2014). LiDAR-enabled model predictive control of wind turbines with real-time capabilities[J]. *Renew. Energy*, 2014, 71: 442–452. doi:10.1016/j.renene.2014.05.041

David, S., and Steffen, R. (2016). *Turbulent extreme event simulations for LiDAR-assisted wind turbine control[C]//journal of physics*. Bristol: Conference Series.

Dunne, F., and Pao, L. Y. (2016). Optimal blade pitch control with realistic preview wind measurements. *J. Wind Energy* 19 (12), 2153–2169. doi:10.1002/we.1973

Haizmann, F., Schlipf, D., Raach, S., et al. (2015). *Optimization of a feed-forward controller using a CW- LiDAR system on the CART3[C]//American Control Conference*. Chicago.IEEE,

Han, B., Zhou, L., and Zhang, Z. (2018a). LiDAR-assisted radial basis function neural network optimization for wind turbines. *IEEJ Trans. Electr. Electron. Eng.* 13 (2), 195–200. doi:10.1002/tee.22514

Han, B., Zhou, L., and Zhang, Z. LIDAR-assisted radial basis function neural network optimization for wind turbines[J]. *IEEJ Trans. Electr. Electron. Eng.*, 2018b, 13(2): 195–200. doi:10.1002/tee.22514

Jia, C., Wang, L., Meng, E., Chen, L., Liu, Y., Jia, W., et al. Combining LIDAR and LADRC for intelligent pitch control of wind turbines[J]. *Renew. Energy*, 2021, 169: 1091–1105. doi:10.1016/j.renene.2021.01.065

Jonkman, J., Butterfield, S., Musial, W., et al. (2009a). *Definition of a 5 MW reference wind turbine for offshore system development[R]. Golden, CO (United States)*. USA, National Renewable Energy Lab NREL.

Jonkman, J., Butterfield, S., Musial, W., et al. (2009b). *Definition of a 5 MW reference wind turbine for offshore system development[R]*. Golden, CO (United States): National Renewable Energy Lab NREL.

Khaniki, M. S., Schlipf, D., and Cheng, P. (2018). *A comparison between LIDAR-based feedforward and DAC for control of wind turbines[C]//IEEE Conference on Control Technology and Applications (CCTA)*. Copenhagen, Denmark: CCTA.

Mathur, R. R., Rice, J. A., Swift, A., and Chapman, J. (2017). Economic analysis of LiDAR-based proactively controlled wind turbines. *Renew. Energy* 103, 156–170. doi:10.1016/j.renene.2016.10.069

Mikkelsen, T. (2014). Wind turbine improvements by wind-LiDAR-based preview and control. *Wash. Dep. Wind Energy Annu. Rep. 2013*, 18.

Schlipf, D. (2015). *LiDAR Assisted control concepts for wind turbines[D]*. Stuttgart: University of Stuttgart.

Schlipf, D., Fleming, P., Haizmann, F., et al. (2015). *Field testing of feedforward collective pitch control on the CART2 using a nacelle-based LiDAR scanner[C]//Journal of Physics*. Bristol, England: Conference Series.

Scholbrock, A., Fleming, P., Schlipf, D., et al. (2016). LiDAR-enhanced wind turbine control: past, present, and future[C]//American Control Conference. *Boston*.

Van, T. L., Nguyen, T. H., and Lee, D. C. Advanced pitch angle control based on fuzzy logic for variable-speed wind turbine systems[J]. *IEEE Trans. Energy Convers.*, 2015, 30(2): 578–587. doi:10.1109/tec.2014.2379293

Yamaguchi, A., Yousefi, I., and Ishihara, T. (2020). Reduction in the fluctuating load on wind turbines by using a combined nacelle acceleration feedback and LiDAR-based feedforward control. *Energies* 13 (17), 4558. doi:10.3390/en13174558

Yuan, Y., Chen, X., and Tang, J. (2020). Multivariable robust blade pitch control design to reject periodic loads on wind turbines. *Renew. Energy* 146, 329–341. doi:10.1016/j.renene.2019.06.136

Zhang, B. J., Lin, Y. G., Chen, W. T., et al. (2021). A novel rotor effective wind speed estimation method for LiDAR application: an experimental case study. *Proceedings of the Institution of Civil Engineers-Energy*, 1–41.

Zhang, B. J., Lin, Y. G., Chen, W. T., Liu, H., Li, W., Sun, Y., et al. (2022). Novel rotor effective wind speed estimation method for light detection and ranging application. *Proc. Institution Civ. Eng. -Energy* 175 (4), 177–194. doi:10.1680/jener.21.00086

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 05 January 2024
DOI 10.3389/fenrg.2023.1328899

Check for updates

OPEN ACCESS

EDITED BY
Unai Fernandez-Gamiz,
University of the Basque Country, Spain

REVIEWED BY
Koldo Portal-Porras,
University of the Basque Country, Spain
Mehdi Neshat,
University of South Australia, Australia

*CORRESPONDENCE
Theodoros Konstantinou,
✉ tkonstantinou@mail.ntua.gr

RECEIVED 27 October 2023
ACCEPTED 07 December 2023
PUBLISHED 05 January 2024

CITATION
Konstantinou T and Hatziargyriou N (2024), Complex terrains and wind power: enhancing forecasting accuracy through CNNs and DeepSHAP analysis. *Front. Energy Res.* 11:1328899.
doi: 10.3389/fenrg.2023.1328899

COPYRIGHT
© 2024 Konstantinou and Hatziargyriou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Complex terrains and wind power: enhancing forecasting accuracy through CNNs and DeepSHAP analysis

Theodoros Konstantinou[1]* and Nikos Hatziargyriou[1,2]

[1]School of Electrical and Computer Engineering, National Technical University of Athens (NTUA), Athens, Greece, [2]School of Technology and Innovations, University of Vaasa, Vaasa, Finland

Accurate prediction of wind power generation in regions characterised by complex terrain is a critical gap in renewable energy research. To address this challenge, the present study articulates a novel methodological framework using Convolutional Neural Networks (CNNs) to improve wind power forecasting in such geographically diverse areas. The core research question is to investigate the extent to which terrain complexity affects forecast accuracy. To this end, DeepSHAP—an advanced interpretability technique—is used to dissect the CNN model and identify the most significant features of the weather forecast grid that have the greatest impact on forecast accuracy. Our results show a clear correlation between certain topographical features and forecast accuracy, demonstrating that complex terrain features are an important part of the forecasting process. The study's findings support the hypothesis that a detailed understanding of terrain features, facilitated by model interpretability, is essential for improving wind energy forecasts. Consequently, this research addresses an important gap by clarifying the influence of complex terrain on wind energy forecasting and provides a strategic pathway for more efficient use of wind resources, thereby supporting the wider adoption of wind energy as a sustainable energy source, even in regions with complex terrain.

KEYWORDS

convolutional neural networks, DeepSHAP, terrain complexity, feature importance, wind power forecasting, Frontiers

1 Introduction

Wind energy is one of the most promising sources of renewable energy in the modern world. Its sustainability and low carbon footprint make it an attractive solution in the global effort to reduce greenhouse gas emissions and combat climate change (International Energy Agency IEA, 2022). As the penetration of wind energy increases, the ability to predict wind power generation becomes increasingly important for the operation of the electricity system. Accurate forecasting is essential not only to optimise energy production, but also to ensure grid stability and the successful integration of this variable energy source into power grids (Ahmed et al., 2020).

Forecasting wind power is a challenging task. The variability and unpredictability of wind, determined by many factors ranging from large-scale atmospheric dynamics to local geographical features, make it a complex phenomenon to predict. This challenge is even greater in regions with complex terrain. Mountains, valleys, coastlines and other topographic features add layers of complexity that can significantly affect wind patterns. For example, wind speeds can be amplified in mountain passes or become turbulent and erratic around steep cliffs and ridges. Predicting wind behaviour in such scenarios is critical, as these areas are often used to site wind farms due to their high wind potential. Traditional prediction models often fail to capture the nuanced interactions between wind and terrain (Bird et al., 2013; Hanifi et al., 2020).

As the demands on wind energy forecasting continue to grow, there is an urgent need for more advanced and accurate methods. While historical data and physical modelling have been the traditional basis for wind power prediction, the intricacies of wind behaviour in complex terrain require sophisticated computational techniques. In addition, to improve techniques, it is crucial to identify and understand the key factors that affect wind power forecasts. By identifying influential meteorological or geographical features, we can develop fine-tuned models that offer superior accuracy. This study uses a convolutional neural network (CNN) to predict wind power in areas with complex terrain. The aim is to address the unique challenges posed by these conditions and also to understand the factors that influence these predictions, in particular the relationship between terrain and wind dynamics.

2 Literature review

The study of wind power forecasting encompasses a wide range of methods, from classic time series analysis to cutting-edge machine learning strategies. Traditional techniques, in particular ARIMA, Exponential Smoothing and Vector Autoregression, have proven to be adept at adapting to the intricacies of complex terrain through their ability to capture the nuanced interplay between topography and wind flow. However, these methods have their limitations, particularly when it comes to accommodating a wide range of input variables and complex interdependencies between them (Chen et al., 2009).

Machine learning techniques have emerged in the field of wind energy forecasting and have been recognised for their ability to successfully deal with the complexity and non-linearities inherent in wind data (Wang et al., 2011; Giebel and Kariniotakis, 2017; Sideratos and Hatziargyriou, 2020; Tawn and Browell, 2022). From artificial neural networks to decision trees, support vector machines and advanced deep learning frameworks, these methods are redefining the benchmarks of forecast accuracy, especially in short-term forecast models. The advent of big data and cloud computing has further accelerated the adoption of advanced models, including convolutional and recurrent neural networks, leading to significant advances in regional wind power forecasting methodologies.

Several innovative techniques aimed at refining wind power forecasts have been presented in the literature. In (Ozkan and Karagoz, 2019), a data mining based strategy, known as the Regional Statistical Hybrid Wind Power Forecast Technique, is detailed for providing regional forecasts (Pinson et al., 2003). Presents a dynamic fuzzy neural network designed to improve forecast accuracy. In (Basu et al., 2020), a hybrid neural network model is developed that combines the capabilities of convolutional and multilayer perceptron networks for day-ahead forecasting. The study in (Dong et al., 2021) addresses the challenges of sparse data with a comprehensive approach, incorporating data correction and error analysis into a hybrid neural network to improve forecast accuracy. Furthermore, (Wood, 2022), presents a methodology that uses trend decomposition along with machine and deep learning for short-term wind capacity factor forecasting. Finally, (Yu et al., 2021), demonstrates the use of deep quantile regression for probabilistic forecasting, providing a robust method for dealing with forecast uncertainty. Deep learning has also been applied to wind speed forecasting, where the ability to predict and understand wind patterns is critical to the efficient operation of wind farms. In their seminal work, Wu et al. (2022a) presented an interpretable model for wind speed prediction using multivariate time series and temporal fusion transformers. This model is notable for its ability to process complex time-dependent data and provide insight into the temporal dynamics of wind speed, offering a significant advance over traditional methods. Similarly, Neshat et al. (2021) introduced a deep learning-based evolutionary model tailored for short-term wind speed forecasting at the Lillgrund offshore wind farm. Their approach combined the predictive power of deep neural networks with evolutionary algorithms to optimise the model's performance, demonstrating a case study where deep learning models significantly improved the accuracy of wind speed predictions. These studies are part of a growing body of literature confirming the superiority of deep learning methods in predicting wind speed, especially when compared to classical statistical models. For example, a study by Zhang et al. (2019) used a deep learning framework to analyse wind turbine data and achieved remarkable success in predicting wind speed, thereby optimising turbine performance. Furthermore, a study by Lei et al. (2009) explored the application of convolutional neural networks to predict wind speeds, which not only improved prediction accuracy but also provided a better understanding of the spatial features relevant to wind speed variations.

Despite their effectiveness, simple ANN-based forecasting methods can struggle in complex terrain (Castellani et al., 2016; Clifton et al., 2022). Recent studies have highlighted the potential of deep learning to address these challenges. Toumelin et al. (2023) presented "DEVINE," which uses CNNs to downscale weather forecasts with high-resolution topographic data, and demonstrated significant improvements in wind speed bias in complex terrain. Shin et al. (2023) emphasised the importance of spatio-temporal data for improving CNN forecasts, while Maldonado-Correa et al. (2021) and Eikeland et al. (2022) validated the effectiveness of hybrid models and the inclusion of historical weather data for probabilistic forecasting in difficult terrain. However, the use of ANNs has presented a paradoxical challenge. Although their performance exceeds that of traditional algorithms, the "black box" nature of their decision-making processes has attracted criticism (Montavon et al., 2017). The opacity of neural networks makes it difficult to discern the logic behind their accurate classifications and predictions, a significant problem in critical applications. To counter this, interpretive techniques such as DeepSHAP have emerged to provide a window into neural computation (Lundberg and Lee, 2017). DeepSHAP elucidates the influence of input features on model outputs, providing a level of transparency that enhances the interpretability of deep learning models (Doshi-Velez and Kim, 2017; Chen et al., 2018), thereby fostering confidence in their predictive capabilities.

3 Methodology

In this study, a methodology that evaluates terrain complexity metrics is developed for the region where wind power generation is expected. In conjunction with this, the DeepSHAP technique is applied to a CNN model to derive normalised importance values for the input features. These values are then compared to the terrain complexity matrices of the designated area. The primary goal is to integrate these methods to identify essential input features for wind power prediction and to discard redundant data from the input domain.

3.1 Convolutional neural networks

CNNs have reshaped the field of machine learning, particularly in tasks related to image and spatial data processing. Originally developed as a computational model for vision, CNNs are specifically designed to recognise and extract hierarchical patterns from structured, grid-like data (Alzubaidi et al., 2021). This makes them an ideal candidate for processing spatial data, such as images, where pixel relationships are essential, or, more relevant, weather grids, where spatial correlations between meteorological factors play a key role in forecasting. The cornerstone of CNNs lies in their ability to use convolutional layers to scan the input data with filters that detect local patterns. These patterns, initially simple in the early layers (such as edges or textures in images), become increasingly abstract and complex as the data progresses through deeper layers. This hierarchical pattern recognition is particularly useful for weather grids, where local interactions between variables such as temperature, pressure, and wind speed can lead to larger regional phenomena. In essence, CNNs can automatically and adaptively learn spatial hierarchies from the data, eliminating the need for manual feature engineering.

For the task of forecasting wind power generation, a simple CNN architecture is used that is suitable for handling the complexities of numerical weather predictions (Wang et al., 2022). The model consists of the following layers:

- Input layer: Accepts numerical weather predictions grids with dimensions representing spatial coordinates (latitude, longitude) and depth indicating various meteorological variables (e.g., wind speed, wind direction).
- Convolutional layers: Multiple layers are used, each with a set of filters to extract relevant features from the input data. The ReLU (Rectified Linear Unit) is used as the activation function to introduce nonlinearity.
- Pooling layers: Interspersed with the convolutional layers, these layers downsample the spatial dimensions, preserving essential information while reducing the computational burden. In this work, max-pooling was used, which retains the maximum value from each local region.
- Fully connected layers: Following the convolutional and pooling layers, one or more fully connected layers interpret the extracted features and drive the prediction mechanism.
- Output layer: Provides the wind power generation prediction for the region of interest.

Based on these characteristics, CNNs fundamentally revolve around a sequence of mathematical operations for processing spatial data, as presented in Eqs (1–9).

1. Convolution operation: Given an input matrix I (representing a small section of our spatial data) and a filter matrix F, the convolution operation is defined as follows:

$$(I \star F)(r,z) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} I(i,j) \cdot F(r-i,z-j) \tag{1}$$

For most applications, I and F are 2D matrices, and the convolution operates throughout the spatial extent of I.

2. Activation function: Post-convolution, an activation function is applied element-wise to introduce nonlinearity. One of the most popular is the Rectified Linear Unit (ReLU):

$$ReLU(x) = \max(0,x) \tag{2}$$

3. Pooling operation: Pooling layers reduce the spatial dimensions of the feature maps. For example, the max-pooling operation is defined as:

$$MaxPool(I)(r,z) = \max(I\{i,j\}\,\forall i,j \in [r,r+W],\; j \in [z,z+H]) \tag{3}$$

where W and H are the width and height of the pooling window, respectively.

4. Fully Connected Layers: In these layers, neurons are densely connected. Given an input vector X, weights A, and biases b, the output Y for a fully connected layer is:

$$Y = A \cdot X + b \tag{4}$$

Integrating these mathematical formulations, a CNN processes spatial data through convolution and pooling operations, introduces non-linearity through activation functions, and uses fully connected layers for final predictions, all while minimising a specified regression loss function. Batch normalisation and dropout techniques have also been incorporated into the architecture to ensure model stability and prevent overfitting. The general architecture of the proposed CNN model is shown in Figure 1. The convolution and pooling layers contained 16 filters with a kernel size of 3 in all dimensions, and the fully connected layers contained 100 nodes each. The model was trained using the Adam optimiser (learning rate = 0.01) with a mean squared error loss function, which is particularly suitable for regression tasks.

3.2 DeepSHAP

3.2.1 Explainable AI in general

As deep learning models become increasingly sophisticated, their predictions can often be hard to interpret, earning them the moniker “black-box” models. In critical applications, such as medical imaging, power system operation or finance, understanding the reasoning behind these predictions is crucial.

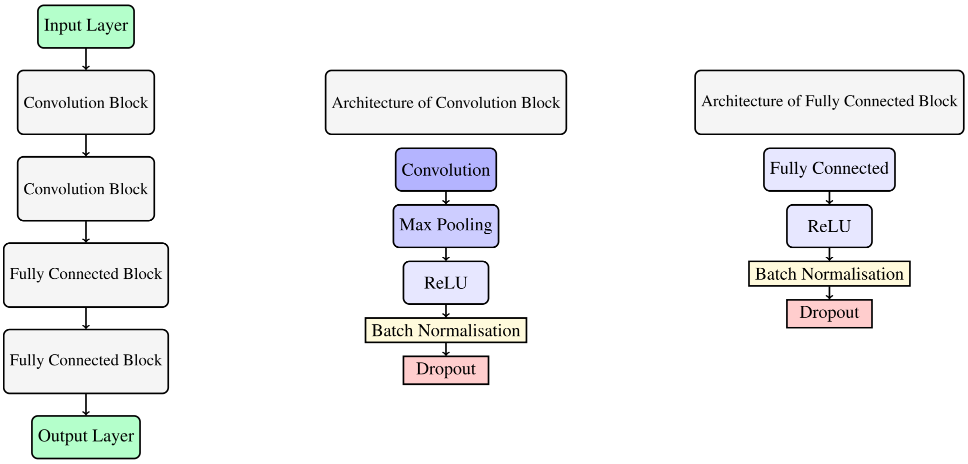


FIGURE 1
Architecture of the proposed CNN model.

This need for interpretability has led to Explainable AI (XAI), an interdisciplinary field that aims to make AI decision making transparent, interpretable and trustworthy (Arrieta et al., 2020). One prominent method in XAI is the concept of Shapley values, which originated from cooperative game theory. Imagine a group of workers working on a project. The Shapley value determines how much each worker contributed to the project success, considering all possible collaborations. In the context of machine learning, each "worker" is a feature and the "project's success" is the prediction. The Shapley value for a feature is then computed on the basis of its marginal contribution across all possible combinations of features. Mathematically, the Shapley value for a feature i is defined as

$$\phi_i(f) = \sum_{S\subseteq N\backslash i} \frac{|S|!\,(|N|-|S|-1)!}{|N|!}\,[f(S\cup i) - f(S)] \tag{5}$$

where f is the prediction function, N is the set of all features, and S is a subset of N without feature i. However, computing Shapley values can be computationally demanding, especially for DNN with numerous input features (Castro et al., 2009). Here DeepSHAP offers an efficient approximation by using a process analogous to backpropagation (Goodfellow et al., 2016).

3.2.2 DeepSHAP propagation in neural networks

DeepSHAP aims to approximate Shapley values for DNN, particularly feedforward neural networks. It does so by redistributing the Shapley values from the output through the network to the inputs (Lundberg and Lee, 2017). This backward pass redistributes the importance or contributions of the output rather than gradients. When attributing the contribution of neuron activations to their respective inputs, the activation of one neuron and the weight of the connection to the next must be accounted for. Mathematically:

$$\phi_{i\to j} = \sum_k \phi_{j\to k} \times \frac{a_i \times w_{i,k}}{\sum_l a_l \times w_{l,k}} \tag{6}$$

where $\phi_{i\to j}$ is the Shapley value of neuron i contributing to neuron j, $\phi_{j\to k}$ is the Shapley value of neuron j contributing to neuron k, a_i is the activation of neuron i, $w_{i,k}$ is the synaptic weight connecting neurons i and j, k an index referring to neurons that neuron j contributes to and l an index for summation, referring to all neurons that are inputs to neuron k.

Convolutional layers, prevalent in deep learning models for image processing, introduce an additional layer of complexity due to shared weights across spatial dimensions. Therefore, DeepSHAP must account for spatial relationships when redistributing contributions. For a specific convolutional filter applied across an input feature map, the contribution of a particular input pixel to an output pixel depends on the filter's weights and the relative position of the pixels. This relationship is described by:

$$\phi_{r,z}^{input} = \sum_{i,j} \phi_{i,j}^{output} \times \frac{I_{r,z} \times F_{i-r,j-z}}{\sum_{p,q} I_{p,q} \times F_{i-p,j-q}} \tag{7}$$

where $\phi_{r,z}^{input}$ is the Shapley value for the pixel at position (r,z) in the input feature map, $\phi_{i,j}^{output}$ is the Shapley value for the pixel at position (i,j) in the output feature map, $I_{r,z}$ is the pixel value at position (r,z) in the input feature map, $F_{i-r,j-z}$ is the weight of the convolutional filter at the relative position to the input pixel, i and j are indices referring to positions in the output feature map, p and q are indices for summation, referring to all positions in the input feature map that contribute to a specific output position.

DeepSHAP's treatment of convolutional layers provides a detailed perspective into which patterns or regions in input feature maps are pivotal for the model's decision, considering not just the importance of a feature but its spatial context in the decision-making process.

3.3 Terrain complexity metrics

The complexity of a terrain can significantly influence the environmental and atmospheric dynamics, especially wind patterns.

Several metrics have been developed to quantify different aspects of this complexity. Understanding these metrics is crucial when integrating them with advanced machine learning techniques, such as DeepSHAP, to decipher the intricate interplay between terrain and wind dynamics. The importance of these terrain metrics in various environmental processes has been highlighted by several studies (Stock and Dietrich, 2006; Wu et al., 2022b).

- Topographic Ruggedness Index (TRI): This index measures roughness based on elevation variances between a cell and its neighboring cells (Riley et al., 1999). Mathematically, it's expressed as:

$$TRI = \sqrt{\sum_{i=1}^{n} (a_i - a_{mean})^2} \tag{8}$$

where a_i denotes the elevation of cell i, and a_{mean} represents the mean elevation of all adjacent cells. The value of n corresponds to the number of cells considered.

- Standard Deviation of Elevation (SDE): A rudimentary metric, it calculates the standard deviation of elevation values within a specified area (Jenny and Hurni, 2011), symbolised as:

$$SDE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (a_i - \mu)^2} \tag{9}$$

where a_i is each elevation value, N is the total number of values and μ is the mean elevation.

In the context of wind power prediction using CNN and DeepSHAP, these terrain complexity metrics play a key role. DeepSHAP determines the importance of each feature by calculating the Shapley values from the output to the input layer. For spatial datasets, such as numerical weather predictions, this reveals which regions or patterns are critical to the model's decision. Comparing DeepSHAP's feature importance values with terrain complexity indices can be revealing. For example, areas identified as high importance by DeepSHAP, when overlaid with regions with high TRI or SDE values, could indicate the importance of rugged terrain in influencing wind power predictions. In essence, if a complex terrain metric closely matches DeepSHAP importance values in a region, it suggests that terrain complexity is a dominant factor in model decisions in that area. Such an investigation provides an empirical way to understand how terrain undulations and complexity affect wind predictability and variability. As a result, prediction models can be refined to ensure that they are better suited to the unique challenges posed by different terrains.

4 Case studies

Exploring the complexities of predicting wind power generation requires an in-depth understanding of the complex interaction between atmospheric conditions and different terrain features. In this context, the selection of Greece, Bulgaria and Romania as our case studies provides a unique opportunity. These countries, each with their own topographical characteristics, provide a diverse landscape for our investigation. Greece's landscape is a mixture of rugged mainland terrain, numerous islands and extensive coastlines. Bulgaria, on the other hand, offers a mix of mountainous regions and flat plains, while Romania's topography is characterised by the Carpathian Mountains, rolling hills and vast plains. This diversity in the geography of these countries allows for a more comprehensive analysis and helps us to understand regional differences in wind power generation.

Recognising that topographic complexity is shaped by a range of factors beyond simply elevation, a comprehensive set of metrics is employed. These include elevation maps, which capture the variation in elevation from coastlines to mountain peaks in all three countries. In addition, metrics such as the TRI and SDE are used to quantify the ruggedness and heterogeneity of each terrain. Shifting focus, the second analysis evaluates the capabilities of a CNN trained on numerical weather predictions and regional wind power generation measurements for Greece, Bulgaria and Romania, as shown in Table 1. The primary objective of this training is to accurately predict regional wind power generation. The input features for each case study consist of wind speed and direction forecasts at a height of 10 m, obtained from the Global Forecast System (GFS). These forecasts are structured in a 3D grid format, where the first two

TABLE 1 Case study information.

Case	Greece	Bulgaria	Romania
Input features	Wind speed (10 m), Wind direction (10 m)	Wind speed (10 m), Wind direction (10 m)	Wind speed (10 m), Wind direction (10 m)
Forecasting horizon (hours)	24	24	24
Time step (hours)	1	1	1
Dataset	ENTSO-e/GFS	ENTSO-e/GFS	ENTSO-e/GFS
Training period	01/01/2019–31/12/2019	01/01/2019–31/12/2019	01/01/2019–31/12/2019
Testing period	01/01/2020–31/12/2020	01/01/2020–31/12/2020	01/01/2020–31/12/2020

dimensions represent the geographical coordinates (latitude and longitude) covering the respective region, and the third dimension contains the wind speed and direction forecasts. This 3D grid is essentially an image-like array that the CNN interprets in a similar way to a visual image. During the training phase, this 3D grid is fed into the CNN, allowing the model to learn the spatial and temporal patterns of wind behaviour in the different terrains of the three countries. The model is trained to recognise how these patterns correlate with actual wind power generation, a crucial step in making accurate predictions. The output data for the model comes from the ENTSO-e platform, which provides actual measurements of the wind power generated in each region. This output is normalised by the installed capacity in each area to standardise the data and ensure that the model's predictions are proportionate and comparable across different regions with different capacities.

However, achieving high prediction accuracy is only one aspect of the objective; it is equally important to understand which input features the model considers critical for its predictions. To this end, we use DeepSHAP to generate Feature Importance Factors (FIV). This technique provides insight into which aspects of numerical weather prediction have the most influence on the model's prediction process. The third analysis attempts to combine the results of the previous two analyses. The feature importance matrices produced by DeepSHAP are compared with the terrain complexity matrices for each country. This comparison will highlight the extent to which terrain complexity affects the importance of different input variables in the prediction model. Such an integrative approach allows us to draw more holistic conclusions about the interaction between terrain complexity and wind energy production in different geographical landscapes.

5 Results

5.1 Terrain complexity

The analysis of terrain complexity in Greece, Bulgaria and Romania is visually summarised in Figure 2 (Greece), 3 (Bulgaria) and 4 (Romania). The left sub-figures display the elevations map of each country. The middle sub-figures display the TRI, highlighting areas of significant topographic variability. Finally, the right sub-figures display the SDE of each country, which provides a quantitative perspective on elevation variability within each region shown. These visual representations serve as a foundation for understanding the complicated relationship between terrain complexity and wind energy prediction in these different geographical areas.

In Greece (Figure 2), the elevation map shows a high contrast between high mountain peaks and sea level, indicative of the mountainous regions of the country and the extensive coastline. The TRI highlights the regions of Greece that are particularly variable in topography, which is likely to have a significant impact on wind flow patterns. The SDE further quantifies these variations, painting a picture of the ruggedness of the terrain. Moving to Bulgaria (Figure 3), the elevation map shows a mixture of flat plains and mountain areas, the TRI highlighting the variability of the Balkan Mountains. The SDE for Bulgaria reflects a more uniform landscape in the plains, with pockets of complexity in the mountainous areas, which could indicate localised areas of more unpredictable wind behaviour. Finally, Romania's landscape (Figure 4) is captured by an elevation map that outlines the extensive Carpathian mountain range as well as the lower elevation regions. The TRI highlights the complexity of the Carpathians, which may correlate with areas of complex wind patterns. The SDE map confirms this complexity, with

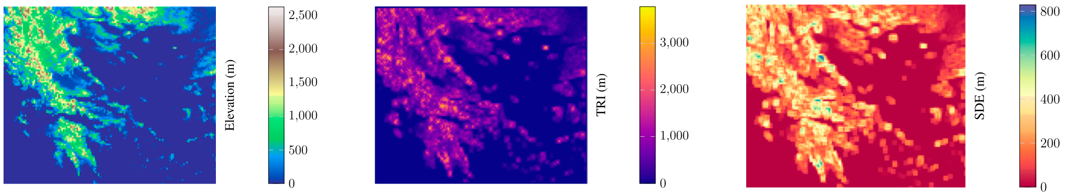


FIGURE 2
Elevation and terrain complexity metrics over the extended area of Greece.

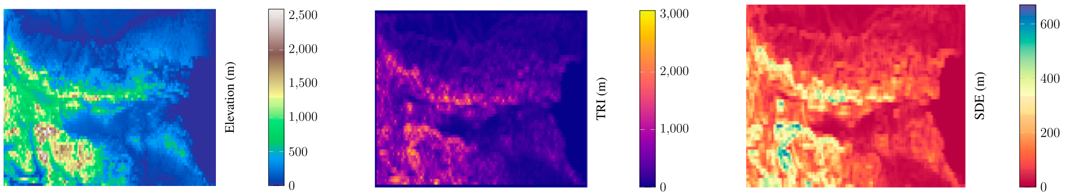


FIGURE 3
Elevation and terrain complexity metrics over the extended area of Bulgaria.

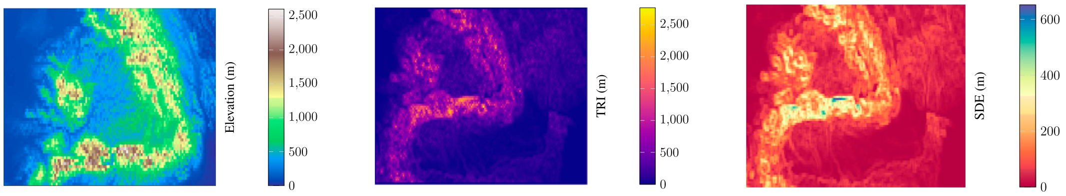


FIGURE 4
Elevation and terrain complexity metrics over the extended area of Romania.

variability in elevation that can affect both micro- and macro-scale wind flows.

5.2 Feature importance analysis

In this analysis, we examine the influence of terrain complexity on wind power prediction by analysing the normalised FIV for Greece, Bulgaria and Romania. Figures 5, 6 show a comparison of these values for wind speed and wind direction predictions in relation to the topographic metrics of each country. In the case of Greece (Figure 5, left), the FIV of the elevation map shows a higher importance in coastal areas and a lower importance in the mountainous regions. This suggests that while the highlands contribute to wind variability, it is the coastal areas where consistent wind patterns dominate the model's focus, possibly due to the unobstructed flow of sea breezes that are crucial for wind power generation. The TRI visualisation further supports this by showing less importance in regions with high topographic variability, suggesting that CNN may find it difficult to predict wind patterns where the terrain is most rugged. For wind direction (Figure 6, left), the FIV is particularly significant along the sea coast, highlighting the importance of offshore influences on the wind pattern for both the mainland and the islands. In the case of Bulgaria (Figure 5, middle), the FIV for wind speed suggests that the model assigns different degrees of importance across the country, reflecting Bulgaria's combination of flat terrain and mountainous areas. Areas of significant FIV align with regions of lower topographic complexity, suggesting that in Bulgaria, unlike Greece, the simpler terrain of the interior may provide more reliable wind conditions for power generation. The values of the importance of the wind direction (Figure 6, middle) show a scattered pattern, suggesting that the impact of wind direction on power forecasting is influenced by the combination of the Balkan Mountains and the surrounding plains. The Romanian analysis (Figure 5, right) shows a clear distribution of FIV across the Carpathians and the vast plains. The model places less emphasis on wind speed predictions in the highly complex Carpathian region, possibly due to the unpredictability of wind behaviour in such terrain. On the contrary, the plains, with their more predictable wind patterns, receive higher FIV scores. For wind direction (Figure 6, right), the FIV is noticeably concentrated in areas that serve as natural wind corridors, suggesting that certain flat and valley regions are key to the prediction process of the prediction model. It is clear that while complex terrain can introduce forecast variability, consistent and predictable wind patterns, particularly in maritime regions, are critical in shaping the focus of the forecast model. This underscores the importance of considering both land- and sea-based influences in the development of accurate wind power forecasts.

5.3 Forecasting performance

Using the knowledge from the feature importance and terrain complexity analysis, the proposed work is focussing on the refinement of the input data by emphasising areas of significance while filtering out potential noise can significantly enhance a model's performance. In our efforts to optimise the input for a CNN model, we systematically investigated three approaches. Each method contains its own unique philosophy, based on computational insights derived from the model or observations of the landscape. The basic goal remained the same: to mask out inputs that help the model deliver accurate wind power forecasts. The following sections clarify these three correspondences and the rationale behind their design.

5.3.1 Approach 1: feature selection based on FIV using DeepSHAP

To improve the predictive accuracy of our CNN model for wind power forecasting, our first approach exploits the strategic use of FIV as determined by DeepSHAP. This method is based on the premise that not all regions within the input data contribute equally to the model's predictions. In particular, regions with low FIVs, as identified by DeepSHAP, are considered to have a minimal or even detrimental effect on prediction accuracy. These regions could represent noise or irrelevant information that could potentially bias the model performance (Lundberg and Lee, 2017; Molnar, 2020). To implement this approach, we applied a selective filtering process to the training data on the key weather variables: wind speed and wind direction. For each of these variables, we examined the normalised FIV values across the input grid. Areas where the FIV was below a threshold of 0.2 were considered to be of low importance. To mitigate their influence, we set the values in these areas to a placeholder or dummy value of −1. This value acts as a signal to the model, effectively "masking" these regions during the training process. The motivation for this decision is twofold. Firstly, by reducing the influence of less important regions, we reduce the likelihood of the model being misled by noise or irrelevant data points. Secondly, and more importantly, this approach sharpens

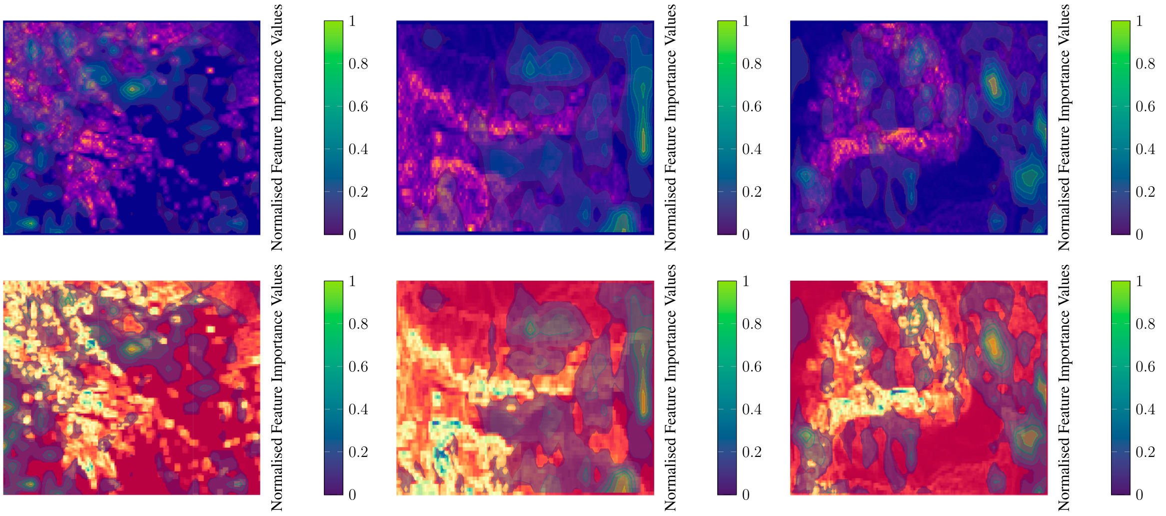


FIGURE 5
Display of Normalised feature importance values of wind speed predictions over the ruggedness metrics' maps of all case studies (Left: Greece, Middle: Bulgaria, Right: Romania).

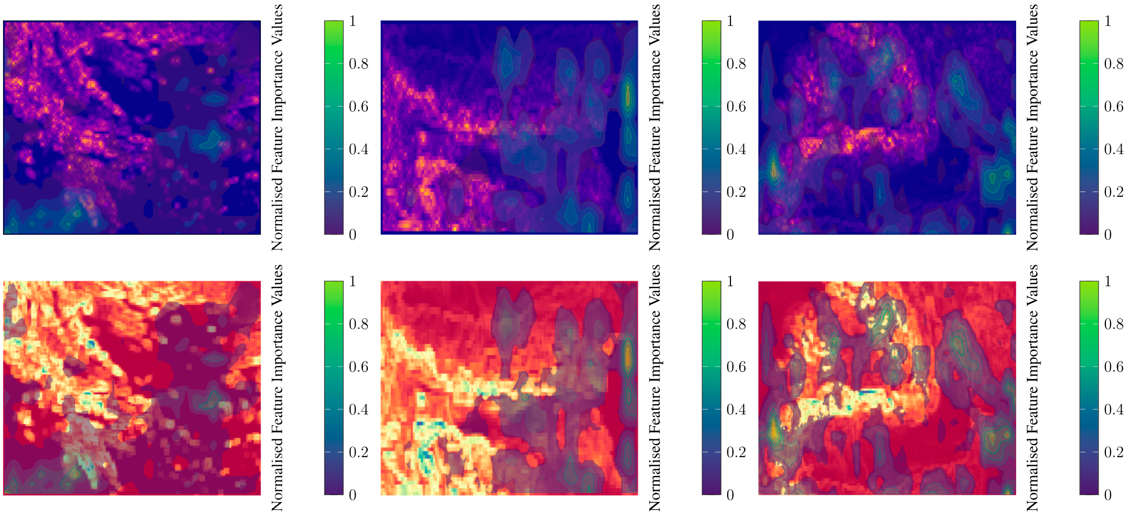


FIGURE 6
Display of Normalised feature importance values of wind direction predictions over the ruggedness metrics' maps of all case studies (Left: Greece, Middle: Bulgaria, Right: Romania).

the model's focus on higher FIV regions, which are theoretically more important in determining accurate wind power forecasts. This method is consistent with the strategies adopted in recent studies where researchers have successfully used feature selection techniques based on importance values to streamline model training and improve overall accuracy.

5.3.2 Approach 2: data filtering based on terrain complexity

The second approach focuses on the dynamic relationship between terrain complexity and wind behaviour, an aspect less emphasised in traditional models. Instead of relying exclusively on FIV, this method integrates SDE as a key metric to assess terrain complexity. This approach is based on the hypothesis that regions with less rugged terrain, as indicated by a lower SDE, are likely to have more predictable and consistent wind profiles. In contrast, areas with a higher SDE, indicating greater ruggedness, may contribute to the unpredictability of wind patterns. To incorporate this terrain-based information into our CNN model, we manipulated the input training data for both wind speed and wind direction. Specifically, regions with an SDE value greater than 400 m were assigned a dummy value of −1. This threshold of 400 m, determined based on the average SDE in each region under study as shown in Figure 2, serves as an arbitrary yet strategic boundary to differentiate between areas of low and high terrain complexity. By applying this filter, we aim to sharpen the focus of the model, allowing it to focus on regions where terrain complexity is less likely to distort wind patterns. By selectively masking regions with high SDE values, we potentially enhance the ability of the model to recognise and adapt to the varying effects of terrain complexity on wind dynamics.

5.3.3 Approach 3: integrating FIV and terrain complexity for improved data filtering

Approach 3 represents a synergistic integration of the first two methods, merging the model-driven insights derived from FIV with the empirical understanding of terrain complexity as indicated by SDE. This approach is based on the premise that a more robust and accurate forecast model can be achieved through a more sophisticated data filtering process that takes into account both the learned patterns of the model and the physical characteristics of the terrain. In practice, this integrated approach involves a two-step filtering mechanism applied to the input training data. First, for each weather variable—wind speed and wind direction—the regions where the normalised FIV falls below the threshold of 0.2 are identified. The values in these regions are then set to −1, effectively "masking" them in the training data set. This step is based on the principle that regions with low FIV contribute less to the model's predictive accuracy and may even act as noise, affecting the model's performance. Following the initial FIV-based filtering, the approach further incorporates considerations of terrain complexity. Areas where the SDE exceeds a predefined threshold of 400 m are also assigned a value of −1. This threshold was chosen to distinguish regions with significant terrain variation from those with more uniform topographic features. The choice of 400 m as a threshold is strategic, as it aims to filter out regions where complex terrain could introduce unpredictability in wind patterns, potentially complicating the forecasting task. By combining these two filtering criteria, Approach 3 creates a training dataset that is both selective and strategic. It emphasises regions that are not only considered important by the model (as per high FIV), but also those with less complex terrain (as per low SDE), and thus potentially more predictable in terms of wind behaviour. This refined dataset is expected to guide the CNN model to focus on the most relevant and reliable features for wind power prediction, thereby improving its overall prediction accuracy.

5.3.4 Evaluation results

To objectively assess the efficacy of the three data preprocessing approaches, a set of reliable evaluation metrics was used. Normalised mean absolute error (NMAE) and normalised mean squared error (NMSE) were used to gain an understanding of the average magnitude of errors and the model prediction accuracy (Willmott and Matsuura, 2005; Chai and Draxler, 2014). The NMAE indicates the average absolute discrepancy, while the NMSE magnifies the effect of larger errors, thus providing an indication of the model's forecast reliability (Hyndman and Koehler, 2006). Additionally, the standard deviation was calculated to measure the variability or spread of prediction errors and to assess the consistency of the model's forecasting ability. The bias was also calculated to identify any systematic overprediction or underprediction tendencies in the model. To compare the performance of the three approaches, metrics were calculated and compared to a baseline scenario, where the input training data was not masked. Through this comparative analysis, our objective is to determine the added value, if any, of the data preprocessing steps.

6 Discussion

The comparative analysis of data processing approaches in Greece, Bulgaria and Romania, as shown in Tables 2–4, provides a detailed evaluation of their impact on CNN-based wind power forecasting.

In Greece (Table 2), the baseline approach sets the standard for comparison, with an NMAE of 4.26% and an NMSE of 0.18%. The bias and standard deviation provide information on the average prediction error of the model and its variability. After implementing Approach 1, which incorporates FIV-based data filtering, a reduction in all metrics is observed, indicating improved accuracy and model stability. Approach 2, which focusses on terrain complexity, yields improvements but falls short of the gains made by Approach 1, suggesting the dominance of FIV-driven regions in influencing wind pattern predictions. However, Approach 3, which combines both FIV and terrain complexity considerations, outperforms the individual approaches, achieving the lowest NMAE, NMSE, bias and standard deviation, thereby demonstrating superior forecast performance and reliability. For Bulgaria (Table 3), the baseline metrics are higher compared to Greece, indicating a greater initial error in the predictions. The adaptation of Approach 1 again proves to be beneficial, as evidenced by the lower NMAE and NMSE. Interestingly, Approach 2 leads to an increase in NMSE despite a reduction in other metrics, suggesting a complex interaction between the features and the terrain. However, Approach 3 emerges as the most effective, significantly reducing all metrics, highlighting the value of a hybrid approach that uses both model-driven and empirical data insights. The results for Romania (Table 4) show the highest baseline NMAE and NMSE of the three countries, highlighting initial challenges in the forecasting model due to possibly more complex wind patterns or varied terrain. Approach 1 and Approach 2 show improvements, but Approach 2 shows a negative NMSE value that may require further investigation to understand anomalous model behaviour. Approach 3 demonstrates its robustness by significantly improving the accuracy and consistency of the forecast, as indicated by significant reductions in all metrics.

The results of the analysis from Greece, Bulgaria and Romania clearly indicate that a hybrid approach combining both FIV and terrain complexity metrics consistently improves CNN-based wind power prediction. This combined strategy, as shown by the data in Tables 2–4, consistently outperforms the individual use of either FIV or terrain complexity metrics alone. The singular use of FIV-based data filtering (Approach 1) while beneficial in reducing error metrics such as NMAE and NMSE, may not fully capture the diverse influence of complex terrain on wind patterns. Similarly, Approach 2 by focusing exclusively on terrain complexity provides a limited view and occasionally leads to inconsistent results, such as the unexpected increase in NMSE for Bulgaria. It is the fusion of both approaches that provides a comprehensive understanding, integrating the data-driven insights of FIV with the empirical knowledge of terrain effects. This dual strategy exploits the strengths of both approaches:

TABLE 2 Case study: Greece—evaluation results for the three approaches.

Approach	NMAE (%)	NMSE (%)	Bias (%)	Standard deviation (%)
Baseline (No Masking)	4.26	0.18	0.38	2.16
Approach 1 (FIV-based)	3.86	0.15	0.32	1.91
Approach 2 (Terrain Complexity)	3.94	0.16	0.35	1.84
Approach 3 (Combined)	3.65	0.14	0.28	1.63

TABLE 3 Case study: Bulgaria—evaluation results for the three approaches.

Approach	NMAE (%)	NMSE (%)	Bias (%)	Standard deviation (%)
Baseline (No Masking)	5.31	0.23	−0.15	2.46
Approach 1 (FIV-based)	4.67	0.18	0.23	2.20
Approach 2 (Terrain Complexity)	4.74	0.30	0.35	2.24
Approach 3 (Combined)	4.02	0.16	0.16	1.85

TABLE 4 Case study: Romania—evaluation results for the three approaches.

Approach	NMAE (%)	NMSE (%)	Bias (%)	Standard deviation (%)
Baseline (No Masking)	7.65	0.34	0.34	2.83
Approach 1 (FIV-based)	6.89	0.28	0.31	2.67
Approach 2 (Terrain Complexity)	6.56	−0.25	0.35	2.54
Approach 3 (Combined)	5.74	0.20	−0.16	2.12

FIV's ability to identify predictive regions within the data, and the complexity of the terrain, which reflects the geographical influence on wind behaviour. The superior performance of Approach 3 in all three countries underlines the synergy achieved by combining these methods. It fine-tunes the forecast model to account for the unique geographical characteristics of each region, resulting in more accurate, reliable and interpretable wind power forecasts. The consistent improvement across all metrics with this combined approach confirms its effectiveness and demonstrates the value of integrating different data processing methods to improve forecasting capability in complex, real-world applications.

7 Conclusion

This research conducted a comprehensive study of the interaction between terrain complexity and feature importance values derived from deep learning models, with a particular focus on their collective impact on wind power predictions. Convolutional Neural Networks using numerical weather prediction were used to extract the intricate correlations influencing wind power generation in Greece, Bulgaria and Romania. The research used metrics such as Standard Deviation of Elevation and Terrain Ruggedness Index, which showed a discernible effect on wind behaviour across the diverse landscapes of these countries. The feature importance analysis, facilitated by the DeepSHAP methodology, identified critical areas within each country that had a significant impact on the forecasting process. A consistent pattern emerged from the analysis; regions with pronounced rugged terrain, particularly inland, generally showed reduced importance. In contrast, maritime regions emerged as a significant contributor to wind dynamics, underlying the importance of coastal and marine areas in the forecast models. The study tested three data filtering approaches to improve forecast accuracy: one based on FIV, another based on terrain complexity, and a third that combined both sets of approaches. Across all case studies, the combined method proved superior, consistently outperforming the others by providing the most accurate forecasts, minimising errors and reducing variability of results. This method effectively combines the data-driven focus of FIV with the empirical knowledge of the field, providing a robust framework for forecasting. In general, this research highlights the value of integrating terrain characteristics with deep learning-derived algorithmic predictions. By adopting such an integrated approach, the potential for optimising wind energy forecasting is greatly enhanced, offering a way to improve the sustainability of energy resources in regions characterised by complex terrain.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://www.ncei.noaa.gov/products/weather-climate-models/global-forecast, https://transparency.entsoe.eu/generation/r2/actualGEnerationPerProductionTYpe/show?name=&defaultValue=true&viewType=TABLE&areaType=BZN&atch=false&datepicker-day-offset-select-dv-date-from_input=D&dateTIme.dateTime=24.10.2023+00:00|CET|DAYTIMERANGE&dateTIme.endDAteTime=24.10.2023+00:00|CET|DAYTIMERANGE&area.values=CTY|10YGR-HTSO-----Y!BZN|10YGR-HTSO-----Y&productionTYpe.values=B01&productionTYpe.values=B02&productionTYpe.values=B03&productionTYpe.values=B04&productionTYpe.values=B05&productionTYpe.values=B06&productionTYpe.values=B07&productionTYpe.values=B08&productionTYpe.values=B09&productionTYpe.values=B10&productionTYpe.values=B11&productionTYpe.values=B12&productionTYpe.values=B13&productionTYpe.values=B14&productionTYpe.values=B20&productionTYpe.values=B15&productionTYpe.values=B16&productionTYpe.values=B17&productionTYpe.values=B18&productionTYpe.values=B19&dateTIme.timezone=CET_CEST&dateTime.timezone_input=CET+(UTC+1)+/+CEST+(UTC+2), https://developers.google.com/maps/documentation/elevation/overview.

Author contributions

TK: Writing–original draft. NH: Writing–review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ahmed, S. D., Al-Ismail, F. S. M., Shafiullah, M., Al-Sulaiman, F. A., and El-Amin, I. M. (2020). Grid integration challenges of wind energy: a review. *IEEE Access* 8, 10857–10878. doi:10.1109/access.2020.2964896

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A. Q., Duan, Y., Al-Shamma, O., et al. (2021). Review of deep learning: concepts, cnn architectures, challenges, applications, future directions. *J. Big Data* 8, 53. doi:10.1186/s40537-021-00444-8

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., et al. (2020). Explainable artificial intelligence (xai): concepts, taxonomies, opportunities and challenges toward responsible ai. *Inf. Fusion* 58, 82–115. doi:10.1016/j.inffus.2019.12.012

Basu, S., Watson, S., Lacoa Arends, E., and Cheneka, B. (2020). *Day-ahead wind power predictions at regional scales: Post-processing operational weather forecasts with a hybrid neural network.*

Bird, L., Milligan, M., and Lew, D. (2013). *Integrating variable renewable energy: challenges and solutions.*

Castellani, F., Astolfi, D., Mana, M., Burlando, M., Meißner, C., and Piccioni, E. (2016). "Wind power forecasting techniques in complex terrain: ann vs. ann-cfd hybrid approach," in Journal of Physics: Conference Series, 753, 082002. doi:10.1088/1742-6596/753/8/082002

Castro, J., Gomez, D., and Tejada, J. (2009). Polynomial calculation of the shapley value based on sampling. *Comput. Intell. Games* 36, 1726–1730. doi:10.1016/j.cor.2008.04.004

Chai, T., and Draxler, R. R. (2014). Root mean square error (rmse) or mean absolute error (mae)? arguments against avoiding rmse in the literature. *Geosci. Model Dev.* 7, 1247–1250. doi:10.5194/gmd-7-1247-2014

Chen, J., Song, L., Wainwright, M. J., and Jordan, M. I. (2018). *Learning to explain: an information-theoretic perspective on model interpretation*, 07814. CoRR abs/1802.

Chen, P., Pedersen, T., Bak-Jensen, B., and Chen, Z. (2009). Arima-based time series model of stochastic wind power generation. *IEEE Trans. power Syst.* 25, 667–676. doi:10.1109/tpwrs.2009.2033277

Clifton, A., Barber, S., Stökl, A., Frank, H., and Karlsson, T. (2022). Research challenges and needs for the deployment of wind energy in hilly and mountainous regions. *Wind Energy Sci.* 7, 2231–2254. doi:10.5194/wes-7-2231-2022

Dong, W., Sun, H., Tan, J., Li, Z., Zhang, J., and Zhao, Y. Y. (2021). Short-term regional wind power forecasting for small datasets with input data correction, hybrid neural network, and error analysis. *Energy Rep.* 7, 7675–7692. doi:10.1016/j.egyr.2021.11.021

Doshi-Velez, F., and Kim, B. (2017). *Towards a rigorous science of interpretable machine learning.*

Eikeland, O. F., Hovem, F. D., Olsen, T. E., Chiesa, M., and Bianchi, F. M. (2022). Probabilistic forecasts of wind power generation in regions with complex topography using deep learning methods: an arctic case. *Energy Convers. Manag. X* 15, 100239. doi:10.1016/j.ecmx.2022.100239

Giebel, G., and Kariniotakis, G. (2017). Wind power forecasting—a review of the state of the art. *Renewable Energy Forecasting.* Woodhead Publishing, 59–109. doi:10.1016/B978-0-08-100504-0.00003-2

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning.* MIT press.

Hanifi, S., Liu, X., Lin, Z., and Lotfian, S. (2020). A critical review of wind power forecasting methods—past, present and future. *Energies* 13, 3764. doi:10.3390/en13153764

Hyndman, R. J., and Koehler, A. B. (2006). Another look at measures of forecast accuracy. *Int. J. Forecast.* 22, 679–688. doi:10.1016/j.ijforecast.2006.03.001

International Energy Agency (IEA) (2022). *Renewables 2022.* Paris: IEA. License: CC BY 4.0.

Jenny, B., and Hurni, L. (2011). Studying cartographic heritage: analysis and visualization of geometric distortions. *Computers and Graphics* 35 (2), 402–411. doi:10.1016/j.cag.2011.01.005

Lei, M., Shiyan, L., Chuanwen, J., Hongling, L., and Yan, Z. (2009). A review on the forecasting of wind speed and generated power. *Renew. Sustain. Energy Rev.* 13, 915–920. doi:10.1016/j.rser.2008.02.002

Lundberg, S., and Lee, S. I. (2017). *A unified approach to interpreting model predictions.*

Maldonado-Correa, J., Valdiviezo-Condolo, M., Viñan-Ludeña, M. S., Samaniego-Ojeda, C., and Rojas-Moncayo, M. (2021). Wind power forecasting for the villonaco wind farm. *Wind Eng.* 45, 1145–1159. doi:10.1177/0309524x20968817

Molnar, C. (2020). *Interpretable machine learning.* Lulu.com.

Montavon, G., Samek, W., and Müller, K. (2017). Methods for interpreting and understanding deep neural networks. *Digit. Signal Process.* 7, 1–15. doi:10.1016/j.dsp.2017.10.011

Neshat, M., Nezhad, M. M., Abbasnejad, E., Mirjalili, S., Tjernberg, L. B., Astiaso Garcia, D., et al. (2021). A deep learning-based evolutionary model for short-term wind speed forecasting: a case study of the lillgrund offshore wind farm. *Energy Convers. Manag.* 236, 114002. doi:10.1016/j.enconman.2021.114002

Ozkan, M. B., and Karagoz, P. (2019). Data mining-based upscaling approach for regional wind power forecasting: regional statistical hybrid wind power forecast technique (regionalshwip). *IEEE Access* 7, 171790–171800. doi:10.1109/access.2019.2956203

Pinson, P., Siebert, N., and Kariniotakis, G. (2003). "Forecasting of regional wind generation by a dynamic fuzzy-neural networks based upscaling approach," in Proceedings EWEC 2003 (European Wind energy and conference).

Riley, S. J., DeGloria, S. D., and Elliot, R. (1999). A terrain ruggedness index that quantifies topographic heterogeneity. *Intermt. J. Sci.* 5, 23–27.

Shin, H., Rüttgers, M., and Lee, S. (2023). Effects of spatiotemporal correlations in wind data on neural network-based wind predictions. *Energy* 279, 128068. doi:10.1016/j.energy.2023.128068

Sideratos, G., and Hatziargyriou, N. D. (2020). A distributed memory rbf-based model for variable generation forecasting. *Int. J. Electr. Power and Energy Syst.* 120, 106041. doi:10.1016/j.ijepes.2020.106041

Stock, J. D., and Dietrich, W. E. (2006). Erosion processes in steep terrain—truths, myths, and uncertainties related to forest management in the pacific northwest. *For. Ecol. Manag.* 224 (1), 199–225. doi:10.1016/j.foreco.2005.12.019

Tawn, R., and Browell, J. (2022). A review of very short-term wind and solar power forecasting. *Renew. Sustain. Energy Rev.* 153, 111758. doi:10.1016/j.rser.2021.111758

Toumelin, L. L., Gouttevin, I., Helbig, N., Galiez, C., Roux, M., and Karbou, F. (2023). Emulating the adaptation of wind fields to complex terrain with deep learning. *Artif. Intell. Earth Syst.* 2, e220034. doi:10.1175/aies-d-22-0034.1

Wang, H. K., Song, K., and Cheng, Y. (2022). A hybrid forecasting model based on cnn and informer for short-term wind power. *Front. Energy Res.* 9. doi:10.3389/fenrg.2021.788320

Wang, X., Guo, P., and Huang, X. (2011). "A review of wind power forecasting models," in The Proceedings of International Conference on Smart Grid and Clean Energy Technologies (ICSGCE 2011, 770–778. Energy Procedia, Energy Procedia12. doi:10.1016/j.egypro.2011.10.103

Willmott, C. J., and Matsuura, K. (2005). Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Clim. Res.* 30, 79–82. doi:10.3354/cr030079

Wood, D. A. (2022). Trend decomposition aids short-term countrywide wind capacity factor forecasting with machine and deep learning methods. *Energy Convers. Manag.* 253, 115189. doi:10.1016/j.enconman.2021.115189

Wu, B., Wang, L., and Zeng, Y. R. (2022a). Interpretable wind speed prediction with multivariate time series and temporal fusion transformers. *Energy* 252, 123990. doi:10.1016/j.energy.2022.123990

Wu, J., Wang, G., Chen, W., Pan, S., and Zeng, J. (2022b). Terrain gradient variations in the ecosystem services value of the qinghai-tibet plateau, China. *Glob. Ecol. Conservation* 34, e02008. doi:10.1016/j.gecco.2022.e02008

Yu, Y., Yang, M., Han, X., Zhang, Y., and Ye, P. (2021). A regional wind power probabilistic forecast method based on deep quantile regression. *IEEE Trans. Industry Appl.* 57, 4420–4427. doi:10.1109/tia.2021.3086077

Zhang, S., Zhang, S., Wang, B., and Habetler, T. G. (2020). Deep learning algorithms for bearing fault diagnostics – a comprehensive review. *IEEE Access* 8, 29857–29881. doi:10.1109/ACCESS.2020.2972859

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 23 January 2024
DOI 10.3389/fenrg.2023.1334588

The energy management strategy of a loop microgrid with wind energy prediction and energy storage system day-ahead optimization

Bin Xu[1,2], Feng Zhang[1], Rui Bai[3], Hui Sun[1]* and Shichuan Ding[2]

[1]Electric Power Research Institute, State Grid Anhui Electric Power Co., Ltd., Hefei, Anhui, China, [2]School of Electrical Engineering and Automation, Anhui University, Hefei, Anhui, China, [3]School of Electrical Engineering and Automation, Hefei University of Technology, Hefei, Anhui, China

Check for updates

OPEN ACCESS

EDITED BY
Xun Shen,
Osaka University, Japan

REVIEWED BY
Boyu Wang,
Pensacola State College, United States
Yuheng Wu,
John Deere, United States
Xingchen Zhao,
Virginia Tech, Blacksburg, United States in collaboration with reviewer [YW]

*CORRESPONDENCE
Hui Sun,
✉ sunhui1990@gmail.com

RECEIVED 07 November 2023
ACCEPTED 26 December 2023
PUBLISHED 23 January 2024

CITATION
Xu B, Zhang F, Bai R, Sun H and Ding S (2024), The energy management strategy of a loop microgrid with wind energy prediction and energy storage system day-ahead optimization.
Front. Energy Res. 11:1334588.
doi: 10.3389/fenrg.2023.1334588

COPYRIGHT
© 2024 Xu, Zhang, Bai, Sun and Ding. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Microgrid has been extensively applied in the modern power system as a supplementary mode for the distributed energy resources. The microgrid with wind energy is usually vulnerable to the intermittence and uncertainty of the wind energy. To increase the robustness of the microgrid, the energy storage system (ESS) is necessary to compensate the power imbalance between the power supply and the load. To further maximize the economic efficiency of the system, the system level control for the microgrid is desired to be optimized when it is integrated with the utility grid. Aiming at the aforementioned problem, this paper comprehensively analyzes the power flow of a typical loop microgrid. A transformer-based wind power prediction (WPP) algorithm is proposed and compared with recurrent neural networks algorithm. With the historical weather data, it can accurately predict the 24 h average wind energy. Based on the predicted wind energy and the time-of-use (TOU) electricity price, a day-ahead daily cycling profile of the ESS with particle swarm optimization algorithm is introduced. It comprehensively considers the system capacity constraints and the battery degree of health. The functionality of the proposed energy management strategy is validated from three levels. First, WPP is conducted with the proposed algorithm and the true historical weather data. It has validated the accuracy of the transformer algorithm in prediction of the hourly level wind energy. Second, with the predicted wind energy, a case study is given to validate the day-ahead daily cycling profile. A typical 1 MVA microgrid is utilized as the simulation model to validate performance of the daily cycling optimization algorithm. The case study results show that the ESS daily cycling can effectively reduce the daily energy expense and help to shave the peak power demand in the grid.

KEYWORDS

wind power prediction, optimization, microgrid, energy storage system, time-of-use price

1 Introduction

The renewable energy such as photovoltaic and wind energy has multiple superiorities including zero carbon emission and reduced generation cost over conventional power Byers and Botterud (2020). However, they also pose challenges to the modern power system due to their distributed characteristics. The modern power system is based on concentrated load and generationCady et al. (2015). The increasing proportion of the distributed generations (DGs) is challenging for the power dispatching and operating efficiency improvement. Thus, microgrid is an emerging supplementary power generation mode to address the challenge posed by the increasing penetration of distributed energy resources in the modern power system Zhao et al. (2017).

Among all renewable energy resources, wind energy has the advantage of cost-effectiveness which makes it appropriate for areas with rich wind energy resources such as off shore areas Tessarolo et al. (2017). However, its intrinsic intermittence features make it undesired for power grid operation. This problem is particularly serious in weak grid conditions such as microgrid which has lower capacity than the large power grid. The mismatch between the power generation and load can lead to wind curtailment which is detrimental to increasing the operating efficiency Zhao et al. (2016). In the worst case, it raises multiple power quality concerns such as frequency deviation or voltage flickering and finally operating failures.

To address the problems posed by wind energy in the microgrid, considerable research effort has gone into the power industry which includes the following steps. The development of battery technology enables energy storage systems (ESS) in the microgrid to reach higher capacity with reduced cost. There are various types of ESS such as the lithium-ion (Li-ion) battery, flywheel, lead-acid batter, *etc.* Prakash et al. (2022). Li-ion battery has been extensively applied in the EVs for the sake of its large energy density and relatively mature manufacturing technology Deng (2015). Therefore, some microgrids also utilize recycled Li-ion batteries as the ESS unit. A large ESS with a fast dynamic response speed can greatly shave the peak generating power and fill the load demand. The energy storage system for microgrid generally includes battery packs and power conversion systems. Via adjusting the output power of the PCS, the power in the grid can be changed and the bus voltage is maintained. However, the application of ESS introduces extra cost for the microgrid and its cycling can lead to aging of the battery. Using ESS daily cycling to minimize the power expense is emerging recently Zhao et al. (2019). The Time-of-Use (TOU) price of the electricity varies hour by hour because of the load changes hourly. Accordingly, when the microgrid is tied to the utility grid, the ESS charges when the TOU price is low and discharges when TOU price is high. For the microgrid, due to the capacity limitation, the daily optimization algorithm of the ESS can be different.

Another effort is developing more accurate wind power prediction (WPP) algorithms. The application of emerging artificial intelligence algorithms enables wind energy models to be more complex and accurate with data-driven methods Zhao et al. (2017). An accurate wind energy prediction strategy is essential for the day-ahead optimization of the microgrid operation plan. Based on the optimization operation plan, the energy management system of the microgrid can effectively adjust the output power of the ESS and DG to minimize the generated power curtailmentYang et al. (2014). The existing WPP algorithms are usually based on recurrent neural networks (RNN) models or the improved RNN models such as short-term memory and gated recurrent units. However, RNN-based models process data sequentially instead of in parallel which makes it struggle with long-range dependencies, and under-perform with non-stationary time-series data. Transformers, which is renowned in natural language processing and image tasks, are now capturing the interest of time series researchers. They show great performance at handling global information while predefined cutoff lengths is not needed. It shows great promise in capturing long-range dependencies, particularly in recent time series prediction tasks Sun et al. (2021). As an emerging algorithm, the industry application of transformer-based model still has large space for further development.

Aiming at addressing the aforementioned problems, this paper introduces a kind of daily optimization method of the microgrid with wind energy generation and ESS. Considering the TOU price, the ESS operates in daily cycling mode to maximize the saving for the users. A machine-learning-based WPP algorithm is introduced. It employs the local historical wind energy data to train the algorithm. Compared with the widely used time-series WPP methods, the proposed energy management strategy utilizes transformer-based algorithm to process the weather data. Compared with the conventional RNN-based methods, its parallel data processing feature makes it more appropriate for the day-ahead hourly-level WPP since the operating efficiency is increased. Also, With the prediction results, a day-ahead operation plan of the ESS cycling which considers the battery aging is introduced with Particle Swarm Optimization (PSO). Simulation results experimental results have validated the functionality of the proposed forecast algorithm and the energy management strategy.

The other sections of this paper are organized in the following way: Section 1 summarizes the state-of-the-art WPP algorithm and the energy management strategy for the microgrid. Section 2 analyzes the microgrid power flow, introduces the proposed WPP method, and the TOU price profile. Based on the predicted wind energy profile and local TOU price profile, Section 4 introduces the day-ahead optimization algorithm of the microgrid and the ESS with PSO. Section 5 introduces a case study to validate the functionality of the proposed energy management strategy. Conclusions are drawn in Section 6.

2 Related work

Microgrid is proposed for various applications to connect the distributed generation resources, loads and ESSs. Since it was first introduced in Hatziargyriou et al. (2007), both industry and academia have dedicated much effort to improving the control and structure of the microgrid to increase its robustness under various operating conditions. The structure of the microgrid, such as the terminal can be different based on the application scenarios. Generally, the energy management of the microgrid can be categorized into three types: hierarchical control Zhao et al. (2016), autonomous control Pogaku et al. (2007), and master-slave control Xie et al. (2021). The power flow control of the microgrid is conducted via changing the operating modes of the power converters. Tayab et al. (2017) introduces the droop control for the microgrid. Rosso et al. (2021) introduce the grid-forming converter for the DGs inside the microgrid. The targets of microgrid energy management can be concluded into two aspects: maximize the economic efficiency and improve the system operating stability. Improving the economic efficiency can be implemented via shaving the peak power demand and filling the valley demand. Therefore, the TOU price is employed by the utility companies to assist in the power dispatching. Yang et al. (2012) employ game-theory to optimize the TOU pricing and maximize the benefit for the utility grid.

Compared with the conventional large grid, microgrid is usually more vulnerable to the uncertainties of the DGs due to the limited system capacity Shuai et al. (2016). Among all distributed energy resources, wind and solar energy occupy the dominant penetration in the current renewable energy structure. Compared with solar energy, wind energy is more efficient. However, wind energy has the intrinsic characteristics of uncertainties and intermittence which are detrimental to the robust operation of the microgrid Wood (2020). As summarized in Xue et al. (2014), there are two major reasons for wind energy uncertainties: average wind speed in the long term and fluctuating wind speed in the short term.

The integration of wind energy brings about multiple concerns for the microgrid such as frequency fluctuation and voltage flickering. To minimize its impact on the microgrid frequency and the voltage, Gautam et al. (2010) propose a blade pitch angle compensation control method and Qi and Tsuji (2023) introduce a time-series coordinated frequency control strategy for the wind farm. It can also result in low-frequency oscillation in the doubly fed induction generators. Therefore, Yang et al. (2011) propose a model to explain the mechanism and figures out the potential solutions. Apart from the microgrid robustness, it can also affect the carbon emission Yao et al. (2012), electricity market Ghadikolaei et al. (2012) *etc.*

To minimize the downsides brought by wind energy integration, various methods are proposed to control and optimize the system's energy management. The extensively applied methods include wind power prediction (WPP) Rodríguez et al. (2020), wind farm system-level control Andersson et al. (2021), fault ride-through Zhang et al. (2020), and energy storage. Among the aforementioned four method, using ESS is the only hardware-based method. The capacity of ESS is highly associated with the robustness of the microgrid. Liu et al. (2018) utilize particle swarm optimization (PSO) algorithm to determine the capacity of a hybrid ESS.

Compared with the other methods, WPP can be assisted in making power dispatching plan to optimize the system operation in advance. With the increasing penetration of wind energy in the current energy structure, WPP becomes more significant since it provides theoretical basis to the comprehensive grid operating optimization. Both academia and industry have dedicated numerical effort into the research of WPP. As a critical part of the numerical weather forecast (NWF), WPP utilizes the wind speed and wind direction predicted with NWP as the input data. However, due to the different locations of each wind turbine, the NWP results cannot be directly employed to replace WPP. Also, the mutual impact between wind turbines can also lead to prediction errors. For instance, the wind turbine at the front side can generate a wake flow which will affect the operation of the following wind turbines []. Therefore, an accurate WPP model should incorporate the physical model of the wind turbines.

Due to the intermittency and uncertainty of wind, WPP is always challenging particularly for a single wind turbine. Generally, WPP can be categorized into two types: short-term prediction which aims at the time-scale shorter than 24-h Sanchez (2006), and long-term prediction which aims at the time-scale longer than 24-h Ahmadi et al. (2020). For the short-term prediction, the wind speed variance and disturbance increase as the prediction period increases due to the inertial of wind power fluctuation. The error of short-term prediction varies due to the intermittence and inertia of the wind turbine which is highly associated with location, environment, and the predicted time. Due to the prediction errors, the post-processing of the prediction results is usually necessary Zhao et al. (2022b). Various methodologies are employed for the short-term WPP such as the Markov chain, Kalman filter, *etc.* Li et al. (2001) propose a first-order artificial neural network model to estimate the very short-term wind turbine power curve. Blonbou (2011) introduces an artificial neural network model with adaptive Bayesian earning and Gaussian process approximation for short-term prediction. Senjyu et al. (2006) introduce a recurrent neural network algorithm to conduct the long-term-ahead wind power generation profile.

According to the aforementioned literature review investigation, this paper will introduce a comprehensive optimization algorithm for the loop microgrid with wind energy and ESS. The WPP is conducted with transformer which shows superiority over the conventional algorithms such as long short-term memory (LSTM) and recurrent neural networks (RNNs). Also, particle swarm optimization algorithm is utilized to make the daily cycling plan for the battery which comprehensively considers the constraints of the system capacity.

3 Microgrid system topology introduction

3.1 Three-terminal microgrid system introduction

A typical three-terminal loop microgrid is given in Figure 1. It includes a wind energy generation terminal, a battery ESS, a load bank and utility grid. The battery ESS is usually located close to the load to compensate for the load demand. The ESS consists of a battery bank and a bidirectional power converter. An transformer is employed to connect the microgrid and the utility grid. Different terminals are connected via transmission line. The bus voltage is set to be 220 V ac. The transmission cable is three-phase four-line. Generally, the transmission cables have circuit breakers on the line.

It can be operated under islanding mode and grid-connected mode. Islanding mode occurs when the utility grid is no longer accessible. The top-priority target of the islanding mode is to minimize the power outage time and maximize the system's robustness. In contrast, when the utility grid is available, the energy management system will aim at maximizing the operating efficiency.

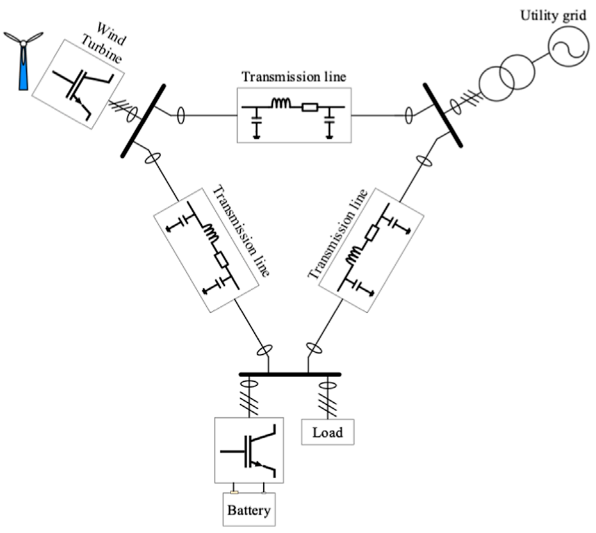


FIGURE 1
The topology of a three-terminal loop microgrid.

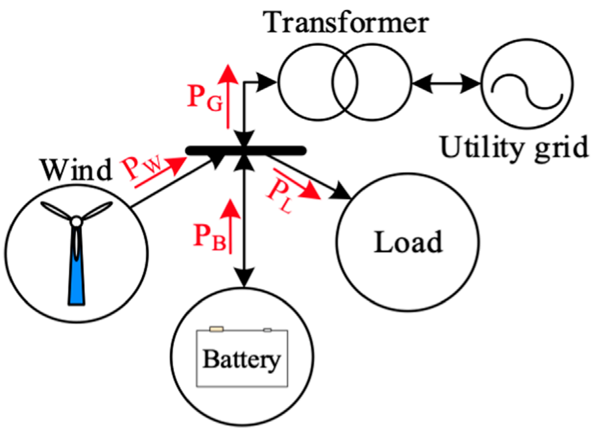


FIGURE 2
The power flow of all terminals in the three-terminal loop microgrid.

Generally, due to the generation intermittence of DG and the missing of a large grid, only a wind energy terminal as the power source is not enough for a microgrid in islanding mode. To maintain the power balance, as many types of DGs as possible such as photovoltaic energy, bioenergy, and tidal energy should be employed for a microgrid. Also, the energy storage system can operate in grid-forming mode to maintain the bus voltage.

In the grid-connected mode, the robustness of the microgrid can be supported by the utility grid and it is no longer the major concern. Thus, the operating efficiency is the top priority target for the microgrid. It can be conducted by minimizing the wind energy curtailment and daily cycling of the ESS to benefit from the TOU price. This paper focuses on the grid-connected mode of the microgrid. The control of the microgrid in grid-connected mode can be found in Figure 2.

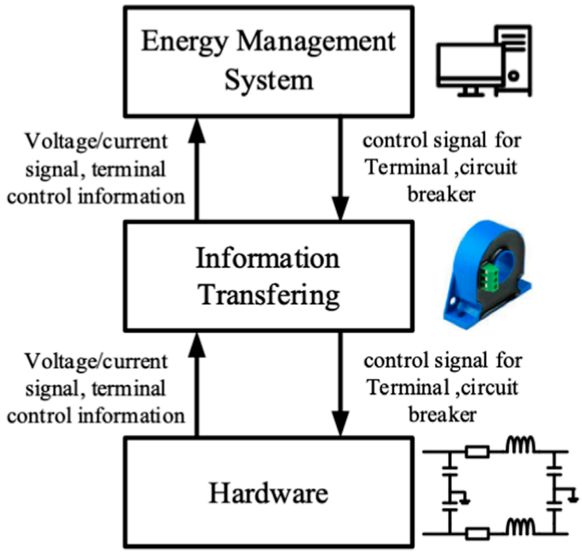


FIGURE 3
The structure of the energy management strategy.

The power flow analysis is essential for operation optimization. Figure 2 shows the power flow of the microgrid. The wind energy generation system and the load are all unidirectional while the transformer and the battery are bidirectional. PB, PW, PL and PG denote the output power of the battery, wind energy generation terminal, load, and the grid. The reference direction of each variable is marked in Figure 2.

The energy management strategy of the microgrid is hierarchical control which includes three layers as shown in Figure 3. The hardware layer include all physical components such as the transmission line, the converter, the connectors, *etc.* The local control of the power converter is also in the hardware layer. The information transferring layer plays a key role in the information exchange between the system layer and the hardware layer. It delivers the local information such as the power flow, voltage or current to the energy management system while the control signals are sent to the hardware layer. The energy management system controls the whole microgrid and dispatches the power flow in the grid. It should be noted that the day-ahead optimization is performed in the energy management system.

3.2 Prediction of the wind power

In 2019, wind energy played a substantial role in advancing renewable energy, contributing to one-third of the overall growth Wood (2020). While the evolution of wind power technology brings evident economic and environmental benefits, the integration of large-scale wind power into the grid faces challenges due to its intermittent nature. WPP emerges as a pivotal solution to tackle this issue.

As demonstrated in Section 2, WPP can be categorized based on the time frame of predictions: ultra-short-term (0–4 h) for managing intraday operations, short-term (typically 1 day or a few

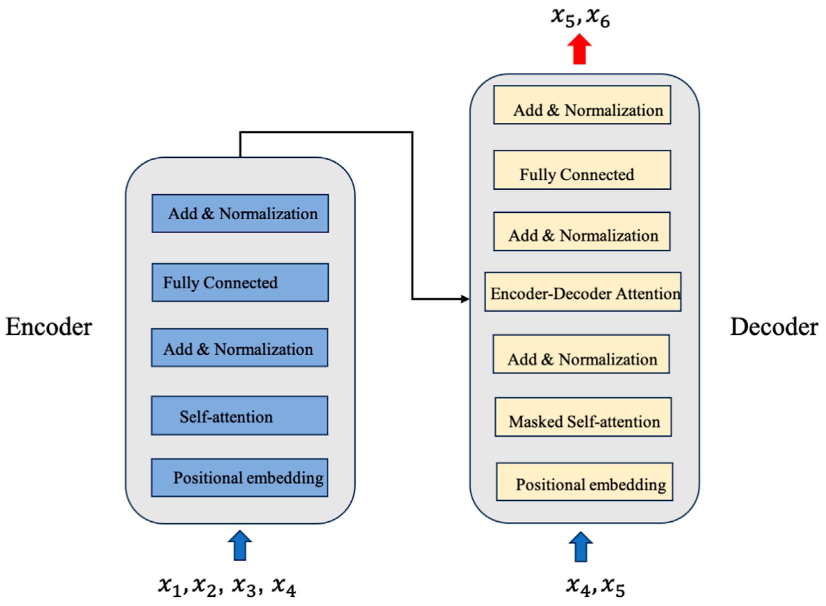


FIGURE 4
The encoder-decoder structure.

days in advance) for day-ahead planning and unit commitment, and medium/long-term forecasting (spanning weekly, seasonal, annual, or more) for wind resource assessment and selection of wind farm sites.

WPP methodologies can be broadly classified into four primary categories: physical models, statistical approaches, intelligent systems, and hybrid methods as shown in Figure 4. Predictions can also be categorized based on the types of input features they use, such as multi-variable predictions and historical predictions. This research, in particular, focuses on intelligent, short-term historical predictions. For the wind prediction task, the prediction process is defined in Eq. 1.

$$\widehat{x}_{t+1} = f\left(x_t, \ldots, x_{t-d+1}\right) + e \tag{1}$$

where $\widehat{x}_{t+1}$ is the predicted wind power, $f(.)$ is the prediction model, and e is the prediction error.

In this study, a transformer model is employed for the prediction task. The encoder component takes the historical time series as input, while the decoder predicts future values in an auto-regressive manner. To establish a connection between the encoder and the decoder, an attention mechanism is employed. This allows the decoder to learn how to prioritize the most pertinent historical values from the time series before making a prediction. Furthermore, the decoder utilizes masked self-attention to prevent the model from gaining an unfair advantage during training by peeking ahead and using future data to predict past values. Note that the prediction can inevitably introduce the prediction errors. Therefore, it is recommended to incorporate the prediction error correction algorithm in prior to performing optimization for the system. Some post-processing algorithms for wind energy prediction can be found by Jiang and Huang (2017) and Zhao et al. (2022b).

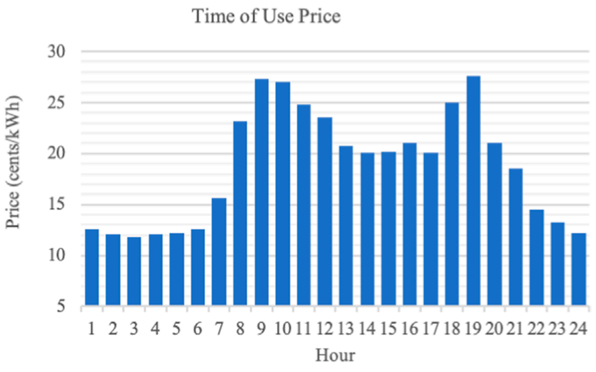


FIGURE 5
A typical TOU price profile (Own representation of data from www.epexspot.com).

3.3 Time-of-use price

The utility company employs the TOU price profile to assist in shaving the peak demand the filling the valley. The price of the energy changes hourly based on the demand. A typical TOU pricing profile is given in Figure 5. Generally, the price is higher at the peak demand time while it decreases at the valley demand time. The customers will adjust their usage of electricity accordingly to minimize their daily energy expense. For instance, the charging of electric vehicles can be delayed to midnight when the TOU price is low. In this case, the power of midnight in the utility grid can be slightly increased and the operation efficiency is improved. Also, the utility company does not need to invest more money into the construction of backup capacity.

4 Day-ahead ESS cycling plan

4.1 Object function and optimization constraint

The optimization object which is improved on the basis of the proposed algorithm in Zhao et al. (2019) aims to minimize the total energy expense of the microgrid which is usually a community network. As mentioned earlier, via charging the battery when TOU is low and discharging the battery when TOU is high, it can maximize the operation efficiency for the customers. Therefore, the object function as shown in 2 is the part the changed expense caused by the ESS daily cycling. It should be noted that in this paper, the power loss on the transmission line can be ignored due to the small size of microgrid.

$$M = \mathbf{min}\left(\sum_{t=1}^{24} P_B(N)\cdot \mathbf{TOU}(\mathbf{N})\right) \tag{2}$$

Herein, $P_B(N)$ and $TOU(N)$ denote the output power of the ESS and the Time-of-Use price at the hour-N. Since the load is determined by the users and it is independent of the system-level control, the only considered variable is ESS output power. The optimization should follow multiple constraints such as the SoC of the battery, the maximum power limit of each terminal and the SoC daily cycling constraints which can be expressed in 3 - 5.

$$P_{B_{ch}} \le P_B(t) \le P_{B_{disch}} \tag{3}$$

$$SoC_{low} \le SoC(t) \le SoC_{high} \tag{4}$$

$$SoC(0) = SoC(24) = SoC(initial) \tag{5}$$

In 3, the output power of the battery is limited to the maximum power rating for charging and discharging. In 4, there is a limitation for the battery SoC. When SoC is too high, the battery is over-charged and it can damage the hardware. In contrast, SoC is also not desired to be very low since deep discharging can lead to a shortened lifetime of the battery. Therefore, it is desired to limit the battery SoC to be within an appropriate range to prevent the over charge/discharge damaging the battery. Usually, SoC_{low} which is the lower limit is between 20% and 40% and SoC_{high} which is the upper limit should be lower than 95%.

$SoC(N)$ denotes the SoC of the battery at the end of hour-N. To finish a completed charging cycle, it is needed to make the SoC same level after completing a daily cycle. In other words, after the charging/discharging for a day, the SoC should return to the initial point, i.e., SoC ($initial$) in the equation. The SoC of the battery at the end of each 30 min can be calculated with 6. Herein, the loss on the power converter is neglected since it is usually very low. Currently, the efficiency of a power converter with wide bandgap semiconductor devices can go to 98% Zhao et al. (2022a). To make the SoC prediction more accurate, it is usually necessary to consider the power loss of the power converter. In this case, the output power of the converter should be multiplied by the efficiency.

$$SoC(N) = SoC(0) + \sum_{i=1}^{t} P_B(N)\cdot \frac{3.6}{C_B\cdot DoH\cdot V_B} \tag{6}$$

In 6, $SoC(N)$ is the SoC at the end of hour-N in a day. SoC (0) is the initial SoC of a day. The SoC (0) can be assumed to be 60% at 0:00 a.m. since 60% is the recommended SoC to maintain the battery lifetime. Note that, due to the battery aging in the daily cycling, the capacity fading should be considered in this case. DoH denotes the battery degree of health which should be measured daily to avoid SoC estimation errors. It can be conducted using an offline method with a coulombmeter as given in Yang et al. (2018). C_B is the nominal battery capacity and V_B denote the battery output voltage. Also, due to the limitation of the transformer capacity, the wind power, load and the battery output power should follow the constraint as given in 7. It can avoid the curtailment of the wind energy caused by the extra-large ESS output power.

$$P_{G\_min} < P_B + P_W + P_L < P_{G\_max} \tag{7}$$

4.2 The optimization algorithm process of ESS output power

The optimization flow is shown in Figure 6. Usually, the optimization process should be performed a day ahead.

All information including the predicted wind power, ESS specs and status, TOU profile from the utility company, and the microgrid parameters is input to the optimization process as the original data. Based on the given data, the fitness function as given in Eq. 2 and the constraints as given in Eqs 3–7 can be determined. Then the optimization can be performed. The selection of various optimization algorithms such genetic algorithm, PSO, simulated annealing algorithm, *etc.*, depends on the application scenario. To avoid the calculation results being trapped into the local optimal point, the initial conditions and maximum iteration generation should be carefully selected. After iteration, the optimized ESS output power can be generated and it will be utilized as the reference value for the ESS converter P/Q control scheme.

5 Experimental study

5.1 Wind power prediction results

In this segment, we utilize actual wind power data to assess the efficacy of forecasting strategies based on transformer models. The data, generated by the Weather Research and Forecasting (WRF) model, consists of hourly samples, amassing a total of 8,640 data points. For our analysis, we consider a historical sequence of 10 samples to predict future wind power output. The dataset is partitioned into a training set comprising 70% of the data, and a testing set making up the remaining 30%. We employ LSTM, GRU, and transformer models to conduct both single-step and multi-step predictions. In single-step forecasting, the immediate next data

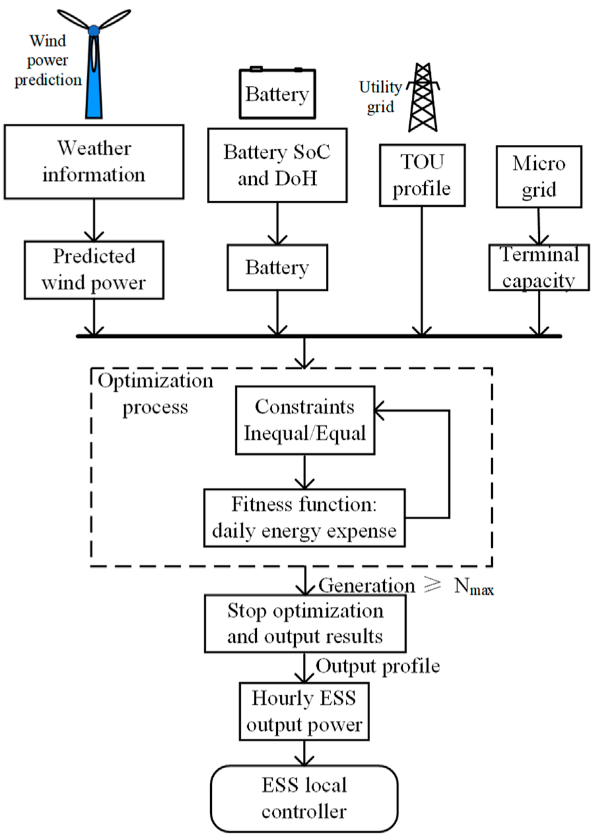


FIGURE 6
The optimization flow chart of the proposed energy management strategy.

TABLE 1 The specs of the microgrid for case study.

Terminal	Parameter	Value
Transmission line	Voltage	$380V3\phi$
Transformer	Capacity	$1000kVA$
Wind turbine	Capacity	$500kVA$
Battery	Voltage	$800V$
	Maximum discharging power	$400kW$
	Maximum charging power	$350kW$
	Capacity	$600kWh$
	Original SoC	60%
	SoC limitation	25%–95%
Load	Capacity	$800kVA$

point is predicted, while in multi-step forecasting, the prediction is made for the third data point in the sequence.

The experiments are conducted on a Linux PC equipped with an AMD Ryzen 5 3550H processor, clocked at 2.1 GHz, along with 16 GB of RAM. The computational environment includes Python 3.9 and TensorFlow 2.8.0. Our aim is to validate the effectiveness of the proposed forecasting approach. TensorFlow is employed due to its ability to handle large datasets and perform complex computations efficiently. It is very appropriate for the evaluation of model accuracy and its deployment for practical, real-time forecasting application. The performance of each model is assessed using two metrics: Mean Squared Error (MSE) and Mean Absolute Error (MAE). MSE quantifies the average squared deviation between the predicted and actual values, offering insight into the model's precision. Conversely, MAE calculates the average absolute discrepancy between the predicted and actual values, showcasing the extent of deviation in the model's predictions.

Upon conducting a thorough analysis of the experimental results in Figures Figure 7, Figure 8, Figure 9, Figure 10, it is observed that the transformer model exhibits superior performance in comparison to the LSTM and GRU models in both single-step and multi-step forecasting tasks. The transformer model demonstrates comparable proficiency to the LSTM model, and maintaining an advantage over the GRU model. These empirical findings underscore the transformer model's potential as an efficacious approach for time series forecasting in wind power prediction endeavors.

5.2 Optimization results: a case study

A case study is given in this paper to validate the functionality of the proposed algorithm. A 500 kVA microgrid system is utilized for the case study. The specs of the microgrid are listed in Table 1.

The hour-level predicted wind power from the original data and the load profile can be plotted in Figure 11. After all data such as the wind energy generation, TOU price, load profile, and the constraints are determined, the optimization process can be performed. In this paper, PSO is employed. The optimization results of the battery output power are plotted in Figures 12, 13.

From Figures 12, 13, the battery output power matches the TOU price. When the TOU price is high, the battery output power increases. When the TOU price is low, the battery charges. From Figure 7, the battery SoC is also within the set constraint 25% and 95%.

The daily expense of the microgrid is plotted in Figure 14. From Figure 14, without the wind energy and the battery daily cycling, the daily energy expense is 6,138 dollar. The ESS daily cycling can reduce the energy expense by 188 dollar. If a larger battery bank is utilized, the cost reduced by daily cycling can be larger. With wind energy and the battery, the total daily cost of energy can be reduced to 3,200 dollar. Therefore, the DG terminal and ESS can effectively reduce the user energy cost.

6 Discussion

To maximize the advantages of microgrid using wind power and battery ESS, this paper proposes a kind of system-level control strategy and the ESS daily cycling algorithm. Using WPP results and

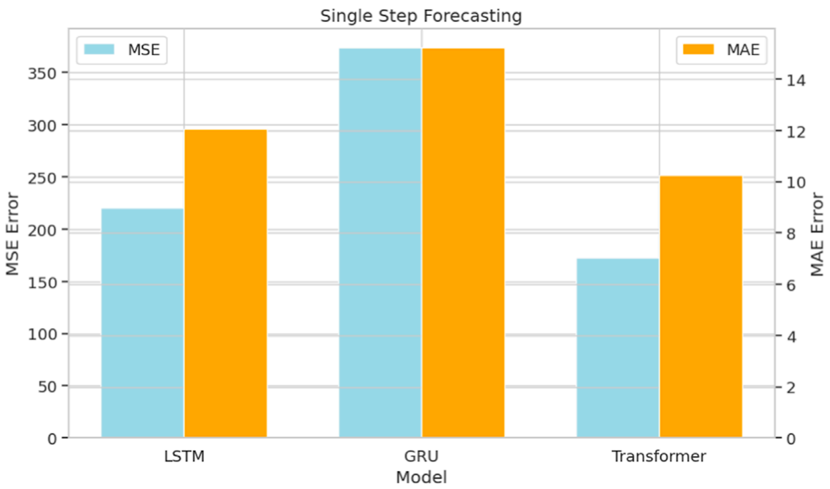


FIGURE 7
Model performance evaluation for the single-step forecast task.

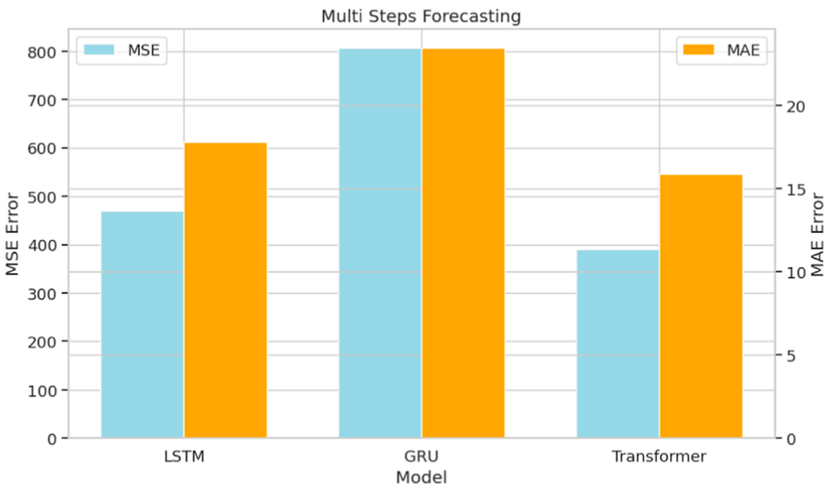


FIGURE 8
Model performance evaluation for the multi-step forecast task.

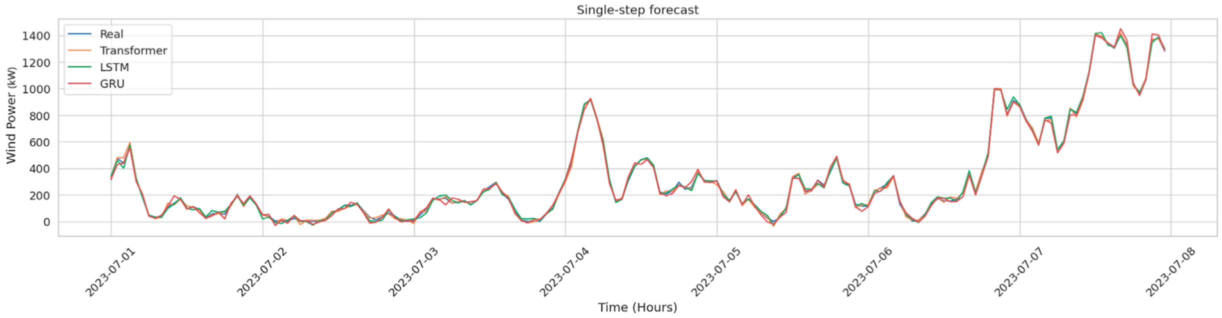


FIGURE 9
Single-step wind power forecasting.

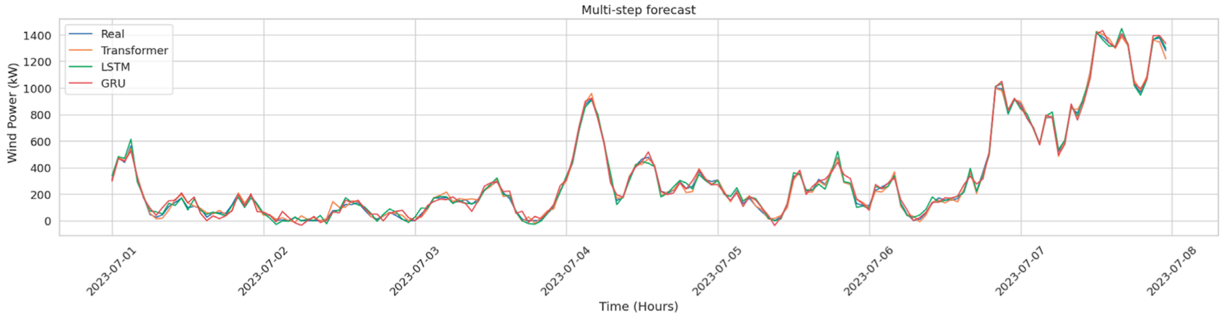


FIGURE 10
Multi-step wind power forecasting.

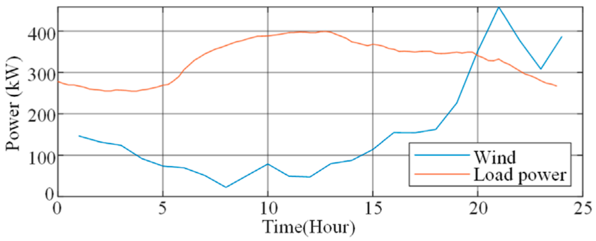


FIGURE 11
The predicted results of the load power and generated wind power.

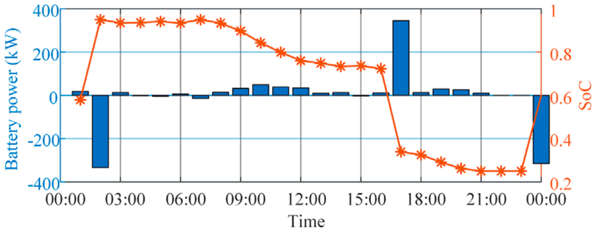


FIGURE 13
The ESS SoC and the output power.

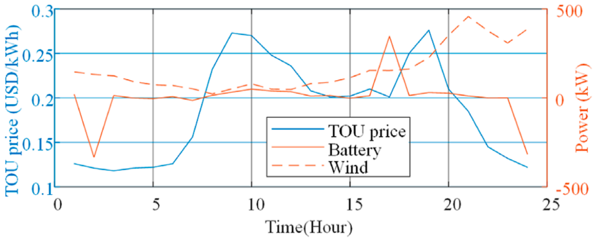


FIGURE 12
The TOU price, the predicted wind energy, and the ESS battery output power.

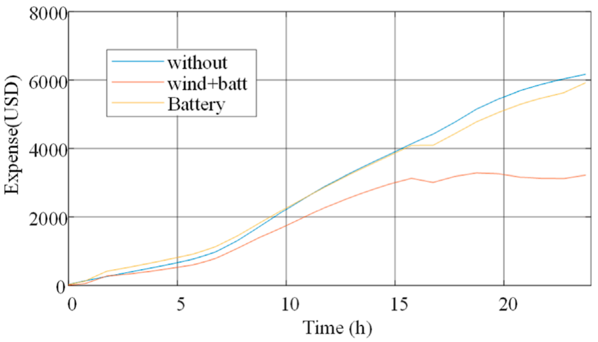


FIGURE 14
The daily electricity expense.

the TOU price, it can effectively reduce the daily energy expense and help the utility grid to dispatch energy by shaving the peak demand. The experimental study is conducted in two levels and the results are discussed as below.

6.1 WPP results

A transformer-based WPP algorithm is proposed and compared with the state-of-the-art machine learning algorithms. From the comparison results as given in Figure 7 - 8, all machine learning methods can predict the wind power at hourly level. However, among the listed algorithms, the transformer algorithm shows superiority over the LSTM and GRU model. The MSE error of the transformer algorithm in single-step prediction is much better than LSTM and GRU model. For multi-step prediction, the GRU model still have the largest MSE error. The MSE error of transformer algorithm is slightly lower than LSTM. Therefore, for 24 h WPP, transformer can be a suitable algorithm.

6.2 ESS daily cycling optimization

From the simulation study in last section, the daily cycling of the ESS can effectively reduce the electricity expense of the users. The charging time focuses at 1:00-3:00 when the TOU is the lowest during a day. The discharging time is usually between 15:00-18:00 when the TOU is highest. It should be noted that it also charges at 23:00 since the battery SoC should return to 60% at the beginning of a daily operating cycle. The simulation results also reveal that the majority of saved expense is still from the wind energy which has no cost. The daily cycling has lower benefit per day. However, for 2 years, the daily cycling of ESS can cover the cost of the battery lifetime aging. Also, it should be noted that in the simulation study, PSO is employed. To achieve better optimization results, other optimization algorithms such as GA, grey wolf optimizer, *etc.*, can be tried.

7 Conclusion

In this paper, a system level control strategy is proposed for the loop microgrid with wind energy generation terminal and the ESS. As the theoretical basis for the system-level operation strategy optimization, a transformer-based WPP algorithm is proposed and compared with LSTM. The comparison results reveal that transformer-based model has better prediction results over the conventional time-series algorithm. With the predicted wind power profile, the ESS daily cycling plan is made based on the optimization algorithm with TOU price. PSO is utilized for the ESS optimization. The case study shows that for a 1MVA microgrid with 800 kVA regular load, the daily cycling of a 400 kVA 600 kWh ESS can reduce the daily energy expense by 3%. Thus, the proposed energy management strategy can effectively improve the system operating efficiency.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

BX: Conceptualization, Writing–original draft, Writing–review and editing. FZ: Software, Validation, Writing–review and editing. RB: Data curation, Software, Validation, Writing–review and editing. HS: Data curation, Formal Analysis, Resources, Software, Writing–original draft, Writing–review and editing. SD: Project administration, Resources, Supervision, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was funded partially by the State Grid Anhui Electric Power Co., Ltd. Scientific and Technology Project (No. 5212G0230002).

Conflict of interest

Authors BX, FZ, and HS were employed by State Grid Anhui Electric Power Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ahmadi, A., Nabipour, M., Mohammadi-Ivatloo, B., Amani, A. M., Rho, S., and Piran, M. J. (2020). Long-term wind power forecasting using tree-based learning algorithms. *IEEE Access* 8, 151511–151522. doi:10.1109/access.2020.3017442

Andersson, L. E., Anaya-Lara, O., Tande, J. O., Merz, K. O., and Imsland, L. (2021). Wind farm control-part i: a review on control system concepts and structures. *IET Renew. power Gener.* 15, 2085–2108. doi:10.1049/rpg2.12160

Blonbou, R. (2011). Very short-term wind power forecasting with neural networks and adaptive bayesian learning. *Renew. Energy* 36, 1118–1124. doi:10.1016/j.renene.2010.08.026

Byers, C., and Botterud, A. (2020). Additional capacity value from synergy of variable renewable energy and energy storage. *IEEE Trans. Sustain. Energy* 11, 1106–1109. doi:10.1109/TSTE.2019.2940421

Cady, S. T., Domínguez-García, A. D., and Hadjicostis, C. N. (2015). A distributed generation control architecture for islanded ac microgrids. *IEEE Trans. Control Syst. Technol.* 23, 1717–1735. doi:10.1109/TCST.2014.2381601

Deng, D. (2015). Li-ion batteries: basics, progress, and challenges. *Energy Sci. Eng.* 3, 385–418. doi:10.1002/ese3.95

Gautam, D., Goel, L., Ayyanar, R., Vittal, V., and Harbour, T. (2010). Control strategy to mitigate the impact of reduced inertia due to doubly fed induction generators on large power systems. *IEEE Trans. power Syst.* 26, 214–224. doi:10.1109/tpwrs.2010.2051690

Ghadikolaei, H. M., Ahmadi, A., Aghaei, J., and Najafi, M. (2012). Risk constrained self-scheduling of hydro/wind units for short term electricity markets considering intermittency and uncertainty. *Renew. Sustain. Energy Rev.* 16, 4734–4743. doi:10.1016/j.rser.2012.04.019

Hatziargyriou, N., Asano, H., Iravani, R., and Marnay, C. (2007). Microgrids. *IEEE Power Energy Mag.* 5, 78–94. doi:10.1109/MPAE.2007.376583

Jiang, Y., and Huang, G. (2017). Short-term wind speed prediction: hybrid of ensemble empirical mode decomposition, feature selection and error correction. *Energy Convers. Manag.* 144, 340–350. doi:10.1016/j.enconman.2017.04.064

Li, S., Wunsch, D. C., O'Hair, E., and Giesselmann, M. G. (2001). Comparative analysis of regression and artificial neural network models for wind turbine power curve estimation. *J. Sol. Energy Eng.* 123, 327–332. doi:10.1115/1.1413216

Liu, Z., Chen, Y., Zhuo, R., and Jia, H. (2018). Energy storage capacity optimization for autonomy microgrid considering chp and ev scheduling. *Appl. Energy* 210, 1113–1125. doi:10.1016/j.apenergy.2017.07.002

Pogaku, N., Prodanovic, M., and Green, T. C. (2007). Modeling, analysis and testing of autonomous operation of an inverter-based microgrid. *IEEE Trans. Power Electron.* 22, 613–625. doi:10.1109/TPEL.2006.890003

Prakash, K., Ali, M., Siddique, M. N. I., Chand, A. A., Kumar, N. M., Dong, D., et al. (2022). A review of battery energy storage systems for ancillary services in distribution grids: current status, challenges and future directions. *Front. Energy Res.* 10, 4831–4842. doi:10.3389/fenrg.2022.971704

Qi, J., and Tsuji, T. (2023). A coordinated frequency control in microgrid based on asymmetric synthetic inertia considering rotational speed of wind turbine. *IEEJ Trans. Electr. Electron. Eng.* 18, 525–536. doi:10.1002/tee.23763

Rodríguez, F., Florez-Tapia, A. M., Fontán, L., and Galarza, A. (2020). Very short-term wind power density forecasting through artificial neural networks for microgrid control. *Renew. energy* 145, 1517–1527. doi:10.1016/j.renene.2019.07.067

Rosso, R., Wang, X., Liserre, M., Lu, X., and Engelken, S. (2021). Grid-forming converters: control approaches, grid-synchronization, and future trends—a review. *IEEE Open J. Industry Appl.* 2, 93–109. doi:10.1109/ojia.2021.3074028

Sanchez, I. (2006). Short-term prediction of wind energy production. *Int. J. Forecast.* 22, 43–56. doi:10.1016/j.ijforecast.2005.05.003

Senjyu, T., Yona, A., Urasaki, N., and Funabashi, T. (2006). "Application of recurrent neural network to long-term-ahead generating power forecasting for wind power generator," in IEEE PES power systems conference and exposition, USA, JUNE 27, 2022 (IEEE), 1260. –1265.

Shuai, Z., Sun, Y., Shen, Z. J., Tian, W., Tu, C., Li, Y., et al. (2016). Microgrid stability: classification and a review. *Renew. Sustain. Energy Rev.* 58, 167–179. doi:10.1016/j.rser.2015.12.201

Sun, Z., Cao, S., Yang, Y., and Kitani, K. M. (2021). "Rethinking transformer-based set prediction for object detection," in Proceedings of the IEEE/CVF international conference on computer vision, China, 11-17 Oct. 2021 (IEEE), 3611–3620.

Tayab, U. B., Roslan, M. A. B., Hwai, L. J., and Kashif, M. (2017). A review of droop control techniques for microgrid. *Renew. Sustain. Energy Rev.* 76, 717–727. doi:10.1016/j.rser.2017.03.028

Tessarolo, A., Luise, F., Pieri, S., Benedetti, A., Bortolozzi, M., and De Martin, M. (2017). Design for manufacturability of an off-shore direct-drive wind generator: an insight into additional loss prediction and mitigation. *IEEE Trans. Industry Appl.* 53, 4831–4842. doi:10.1109/TIA.2017.2700285

Wood, D. (2020). Grand challenges in wind energy research. *Front. Energy Res.* 8, 624646. doi:10.3389/fenrg.2020.624646

Xie, X., Quan, X., Wu, Z., Cao, X., Dou, X., and Hu, Q. (2021). Adaptive master-slave control strategy for medium voltage dc distribution systems based on a novel nonlinear droop controller. *IEEE Trans. Smart Grid* 12, 4765–4777. doi:10.1109/tsg.2021.3104413

Xue, Y., Lei, X., Xue, F., Yu, C., Dong, Z., Wen, F., et al. (2014). A review on impacts of wind power uncertainties on power systems. *Proc. CSEE* 34, 5029–5040. doi:10.13334/j.0258-8013.pcsee.2014.29.004

Yang, H.-T., Huang, C.-M., Huang, Y.-C., and Pai, Y.-S. (2014). A weather-based hybrid method for 1-day ahead hourly forecasting of pv power output. *IEEE Trans. Sustain. Energy* 5, 917–926. doi:10.1109/TSTE.2014.2313600

Yang, L., Xu, Z., Østergaard, J., Dong, Z. Y., Wong, K. P., and Ma, X. (2011). Oscillatory stability and eigenvalue sensitivity analysis of a dfig wind turbine system. *IEEE Trans. Energy Convers.* 26, 328–339. doi:10.1109/tec.2010.2091130

Yang, P., Tang, G., and Nehorai, A. (2012). A game-theoretic approach for optimal time-of-use electricity pricing. *IEEE Trans. Power Syst.* 28, 884–892. doi:10.1109/tpwrs.2012.2207134

Yang, Z., Patil, D., and Fahimi, B. (2018). Online estimation of capacity fade and power fade of lithium-ion batteries based on input–output response technique. *IEEE Trans. Transp. Electrification* 4, 147–156. doi:10.1109/TTE.2017.2775801

Yao, F., Dong, Z. Y., Meng, K., Xu, Z., Iu, H. H.-C., and Wong, K. P. (2012). Quantum-inspired particle swarm optimization for power system operations considering wind power uncertainty and carbon tax in Australia. *IEEE Trans. industrial Inf.* 8, 880–888. doi:10.1109/tii.2012.2210431

Zhang, T., Yao, J., Sun, P., Pei, J., Zhang, H., Liu, K., et al. (2020). Improved continuous fault ride through control strategy of dfig-based wind turbine during commutation failure in the lcc-hvdc transmission system. *IEEE Trans. Power Electron.* 36, 459–473. doi:10.1109/tpel.2020.3000515

Zhao, S., Kempitiya, A., Chou, W. T., Palija, V., and Bonfiglio, C. (2022a). Variable dc-link voltage llc resonant dc/dc converter with wide bandgap power devices. *IEEE Trans. Industry Appl.* 58, 2965–2977. doi:10.1109/tia.2022.3151867

Zhao, S., Umuhoza, J., Zhang, Y., Moquin, J., Farnell, C., and Mantooth, H. A. (2017). "Analysis and optimization of a high-efficiency residential energy harvesting system with dual half-bridge converter," in IEEE Applied Power Electronics Conference and Exposition (APEC), USA, 15 Jul 2021 (APEC), 2838. –2844. doi:10.1109/APEC.2017.7931100

Zhao, S., Zhang, Y., Moquin, J., and Mantooth, H. A. (2016). "The hierarchical energy management control for residential energy harvesting system," in IEEE Energy Conversion Congress and Exposition (ECCE), China, October 29 2011 (IEEE). 1–7. doi:10.1109/ECCE.2016.7855501

Zhao, S., Zhao, X., Farnell, C., Mantooth, H. A., Umuhoza, J., and Zhang, Y. (2019). "A daily optimization method for a pv-battery microgrid considering the battery lifetime and time-of-use pricing," in IEEE Applied Power Electronics Conference and Exposition (APEC), USA, February 25-29, 2024 (APEC), 3243. –3250. doi:10.1109/APEC.2019.8722088

Zhao, X., Sun, Q., Tang, W., Yu, S., and Wang, B. (2022b). A comprehensive wind speed forecast correction strategy with an artificial intelligence algorithm. *Front. Environ. Sci.* 10, 1034536. doi:10.3389/fenvs.2022.1034536

frontiers | Frontiers in Energy Research

TYPE Brief Research Report
PUBLISHED 25 January 2024
DOI 10.3389/fenrg.2023.1321917

Meta-reservoir computing for learning a time series predictive model of wind power

Li Zhang*, Han-Xiao Ai, Ya-Xin Li, Li-Xin Xiao and Cao Dong

Xiangyang Electric Power Supply Company, State Grid, Xiangyang, Hubei, China

Check for updates

OPEN ACCESS

EDITED BY
Xun Shen,
Osaka University, Japan

REVIEWED BY
Yikui Liu,
Stevens Institute of Technology,
United States
Hao Zhong,
China Three Gorges University, China

*CORRESPONDENCE
Li Zhang,
✉ zhang.li336@gmail.com

RECEIVED 15 October 2023
ACCEPTED 24 November 2023
PUBLISHED 25 January 2024

CITATION
Zhang L, Ai H-X, Li Y-X, Xiao L-X and Dong C (2024), Meta-reservoir computing for learning a time series predictive model of wind power.
Front. Energy Res. 11:1321917.
doi: 10.3389/fenrg.2023.1321917

COPYRIGHT
© 2024 Zhang, Ai, Li, Xiao and Dong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Wind energy has become an essential part of the energy power source of current power systems since it is eco-friendly and sustainable. To optimize the operations of wind farms with the constraint of satisfying the power demand, it is critical to provide accurate predictions of wind power generated in the future. Although deep learning models have greatly improved prediction accuracy, the overfitting issue limits the application of deep learning models trained under one condition to another. A huge number of data are required to train a new deep learning model for another environment, which is sometimes not practical in some urgent situations with only very little training data on wind power. In this paper, we propose a novel learning method, named meta-reservoir computing (MRC), to address the above issue, combining reservoir computing and meta-learning. The reservoir computing method improves the computational efficiency of training a deep neural network for time series data. On the other hand, meta-learning is used to improve the initial point and other hyperparameters of reservoir computing. The proposed MRC method is validated using an experimental dataset of wind power compared with the traditional training method. The results show that the MRC method can obtain an accurate predictive model of wind power with only a few shots of data.

KEYWORDS

meta-learning, deep learning model, wind power prediction accuracy, time series data, reservoir computing

1 Introduction

The utilization of wind power has dramatically improved in the last decade. Wind power generation is random due to the uncertain property of wind speed. The uncertainty of wind power generation brings challenges to the power system dispatch with safety constraints and operational stability (Ummels et al., 2007). Thus, accurate wind turbine power generation prediction is critical for improving the safety and efficiency of utilizing wind energy in power systems (Lange, 2005). Nowadays, wind turbines are often equipped with Supervisory Control and Data Acquisition (SCADA) systems that record the real-time data on wind turbine operations. The data from the SCADA system can be applied to monitor the status of the wind turbines. On the other hand, we can also use the data to build predictive models for wind turbine power.

The research on wind power prediction has been mainly focused on providing time series predictions based on time series data (Burke and O'Malley, 2011). Deep learning models have been applied to improve the accuracy of wind power prediction. One mainstream method is to use a long short-term memory (LSTM) neural network to model the time series wind power model. For example, Chen et al. (2019) proposed a two-layer method, which combines extreme learning machine and LSTM to address the nonlinear property of the wind power model and overcome the weakness of linear combination by using only one layer. In addition, Ko et al. (2021) proposed a deep residual network that integrates long and short bidirectional LSTM to improve accuracy and training efficiency further. Recently, a probabilistic prediction of wind power has also been addressed. Zhang et al. (2021) designed a multi-source and temporal attention network to improve prediction performance by introducing three specific designed sources. Furthermore, Safari et al. (2018) used ensemble empirical mode decomposition to divide the wind power time series into different components with different time–frequency characteristics. Then, the authors used chaotic time series analysis to discover the components with chaotic properties. Subsequently, the predictive model provides the predictions for the chaotic and nonchaotic parts separately, which improves the prediction accuracy. Zhao et al. (2021) proposed an integrated probabilistic forecasting and decision framework to optimize the prediction interval of wind power and quantify the probabilistic reserve simultaneously. An extreme learning machine is applied to reduce the time efficiency of establishing the prediction interval. In addition, a novel closed-form prediction for wind speed and wind power is presented by Wang et al. (2021). Liu et al. (2018)integrated wavelet packet decomposition, gray wolf optimizer, adaptive boosting.MRT, and robust extreme learning machine to increase the accuracy of multi-step prediction for wind power.

Recent research has discovered that the wind speed dynamical model and the wind turbine power curve depend on the environment, such as atmospheric conditions and temperature (Cascianelli et al., 2022; Pandit et al., 2023). None of the above research on wind power predictive models has considered environmental changes. Wu et al. (2023) presented a heuristic result that considers the atmospheric model in wind power prediction. However, it does not give hints on building a more general model. Deep models encounter overfitting issues (Duffy et al., 2023). As the environment changes, the prediction by deep models deviates from the real value and needs to be modified by using data from the new environment. The traditional training methods for deep models need a sufficiently large number to train the model, which is computationally complex for real-time modification. In addition, it may not be practical to quickly obtain many new data.

The reservoir computing method is a computationally efficient method to train neural network models (Hamedani et al., 2018; Nokkala et al., 2022), including recurrent neural networks (RNNs) and LSTM neural networks. Although using the reservoir computing method for deep models can significantly reduce the computational complexity for training, the issue of not having enough data quickly is still unresolved. Meta-learning has been validated to adapt the deep model to a new situation with only a few shots of data (Li and Hu, 2021; Tian et al., 2022). This paper combines the advantages of reservoir computing and meta-learning and proposes a novel wind power predictive model, named the meta-reservoir computing method. Meta-learning optimizes part of the hyperparameters of the reservoir computing algorithm based on a multiple-task dataset. Then, the enhanced reservoir computing algorithm can efficiently adapt the predictive model of wind power to a new task with a few data samples. We conducted experimental data-based validations to evaluate the proposed meta-reservoir computing method. The main contributions of this paper are summarized as follows:

- This is the first study to consider the problem of adapting a deep learning wind power predictive model with small samples.
- Meta-learning is combined with reservoir computing for the first time to improve the training efficiency of deep learning models for wind power prediction with the constraint of small samples.

The remainder of this paper is organized as follows: Section 2 presents the addressed problem after formulating the model, integrating the environment factors; Section 3 explains the proposed meta-reservoir computing method for wind power predictive modeling; Section 4 presents the validation results of applying the proposed meta-reservoir computing method to an experimental dataset; and Section 5 presents the conclusions of this paper.

2 Addressed problem: fast model learning for wind power prediction

Let the time index be $k = 0, 1, 2, \ldots, T, \ldots$. At every time k, the wind power is defined by p_k. Wind power is generated from the wind turbine and depends on the wind speed at the current time index k. Let s_k be the wind speed at time step k. A nonlinear map called the wind turbine power curve (Luo et al., 2022) describes the correlation between wind speed and wind power output, which is expressed as follows:

$$p_k = h(s_k, w_k), \tag{1}$$

where w_k defines the uncertainty related to the measurements and the model bias.

On the other hand, the mechanism of generating wind speed s_k is essentially a Markov process defined by

$$s_{k+1} = f(s_k, v_k), \tag{2}$$

where v_k is the system noise and $f(\cdot)$ is the function that describes the state transition with randomness. Note that the randomness is addressed by the system noise v_k. Then, we can equivalently regard the wind power itself following a Markov process defined by

$$p_{k+1} = g(p_k, \delta_k). \tag{3}$$

In practice, $g(\cdot)$ is not available. One basic solution is to use the time series dataset of wind power to estimate $g(\cdot)$, which essentially follows a data-driven fashion. Let

$$\mathcal{D}_T := \{p_k\}_{k=0}^{T} \tag{4}$$

be the available dataset. The traditional problem is to solve

$$\min_{\tilde{g}} \sum_{k=0}^{T-1} (p_{k+1} - \tilde{g}(p_k))^2. \tag{5}$$

However, recent research reveals that the wind speed dynamical model and the wind turbine power curve vary as the environment changes (Cascianelli et al., 2022; Pandit et al., 2023). Namely, instead of using (Eqs 1, 2), the following model should be used:

$$p_k = h(s_k, \theta, w_k), \tag{6}$$

$$s_{k+1} = f(s_k, \theta, v_k), \tag{7}$$

where θ defines an unknown variable to represent the influence of the environment change. Then, the dynamical model of the wind power is written by

$$p_{k+1} = g(p_k, \theta, \delta_k). \tag{8}$$

Suppose the dataset $\mathcal{D}_T$ includes the data collected from different environments specified by labels from $\{1, \ldots, K\}$. If we train the model $\tilde{g}$ by solving Eq. 5 directly based on $\mathcal{D}_T$, the solution will not fit any model conditioned on the label $i, i = 1, \ldots, K$. On the other hand, when the new dataset $\mathcal{D}_{T'}^{K+1}$ conditioned on a new environment specified by θ_{K+1} comes with very few samples (namely, T' is very small), there is no effective way to adapt the solution to dataset $\mathcal{D}_T$ to the new environment model.

This paper addresses the problem of providing a solution $\tilde{g}^*$ robust to each environment parameter $\theta_i, i = 1, \ldots, K$. On the other hand, $\tilde{g}^*$ should also have a property that it can be adapted to the solution of a new set, $\tilde{g}_{K+1}$, very efficiently, with only very few data obtained on the new environment.

3 Meta-reservoir computing for wind power prediction

3.1 Reservoir computing

Reservoir computing is a computational approach for time series data processing based on neural networks. Reservoir computing was first proposed by Jaeger and Haas (2004) for optimizing RNN models for given training data. Since RNN is widely used for time series data modeling, reservoir computing can also be generalized to the applications of time series data processing (Tanaka et al., 2019).

In reservoir computing, the time series data are supposed to be generated from unknown dynamical models driven by sequences of inputs, and the system outputs sequences of outputs. It can also be applied to autonomous systems by setting the input at each time as zero. In this paper, since we do not have input for the dynamics of wind speed, the input is not considered. The *reservoir* in reservoir computing is essentially the state variable of the established dynamical model for predicting the output, and it does not have to represent the underlying state of the actual physical systems (Tanaka et al., 2019). Let $\mathbf{r}_k$ be the reservoir at time step k. The measured output at time step k is defined by $\mathbf{y}_k$. The reservoir at time step $k+1$, $\mathbf{r}_{k+1}$, is a function of $\mathbf{r}_k$ and $\mathbf{y}_k$, written by

$$\mathbf{r}_{k+1} = \mathbf{f}_{\mathrm{rc}}(\mathbf{W}_{\mathrm{rc}}\mathbf{r}_k + \mathbf{W}_{\mathrm{back}}\mathbf{y}_k), \tag{9}$$

where $\mathbf{f}_{\mathrm{rc}}$ is a neural network and $\mathbf{W}_{\mathrm{rc}}$ and $\mathbf{W}_{\mathrm{back}}$ are the weight matrices for reservoir–reservoir connections and output–reservoir connections, respectively. The output at time index $k+1$ is predicted by

$$\mathbf{y}_{k+1} = \mathbf{W}_{\mathrm{out}}\mathbf{r}_{k+1}. \tag{10}$$

The computational complexity is immense if we want to train $\mathbf{W}_{\mathrm{out}}$, $\mathbf{W}_{\mathrm{back}}$, and $\mathbf{W}_{\mathrm{rc}}$ together. Note that the model capacity is substantial if there are enough reservoirs and neurons. The model can achieve high accuracy even though $\mathbf{W}_{\mathrm{back}}$ and $\mathbf{W}_{\mathrm{rc}}$ are randomly given and only $\mathbf{W}_{\mathrm{out}}$ is trained. The algorithm of implementing reservoir computing with a dataset $\mathcal{D}_T^{\mathrm{rc}} := \{\mathbf{y}_k\}_{k=0}^T$ is summarized in Algorithm 1.

Note that λ is a parameter for regularization. Using a large λ confers the method a higher robustness but may lose some accuracy. With a small λ, the obtained model will have better accuracy but may encounter the overfitting issue. The choice of λ should be made according to the problem and the user demands.

3.2 Meta-learning

The meta-learning discussion first focused on learning in a multiple-task scenario. To specify the training process of meta-learning, it is formulated as a bi-level optimization problem. We will clearly explain how the bi-level optimization framework of meta-learning fits our problem.

As introduced in Section 2, the dataset $\mathcal{D}_T$ includes the data obtained from different environments specified by the task labels $\{1, \ldots, K\}$. Then, the dataset can be divided into K different tasks. Each task has a corresponding dataset $\mathcal{D}_{T_i}^i$, $i = 1, \ldots, K$. Instead of only considering the parameter vector in the model to be learned in meta-learning, another important variable, ω, which specifies the algorithm about how to learn the parameter, is also optimized. The variable ω can include the initial point of the parameter, the hyperparameters for the gradient-descent method, the choice of cost function, and the selected model.

The dataset of each task is separated into a training set $\mathcal{D}_{T_i}^{i,\mathrm{tr}}$ and test set $\mathcal{D}_{T_i}^{i,\mathrm{te}}$. Note that the parameter obtained by each dataset depends not only on the training dataset but also on the learning variable ω. The loss function depends on the trained parameter

```
Inputs: dataset $\mathcal{D}_T^{\mathrm{rc}} = \{\mathbf{y}_k\}_{k=0,\ldots,T}$
  Select the model $\mathbf{f}_{\mathrm{rc}}$ and reservoir $\mathbf{r}_k$
  Generate weight matrices $\mathbf{W}_{\mathrm{back}}$ and $\mathbf{W}_{\mathrm{rc}}$ randomly
  Generate initial reservoir $\mathbf{r}_0$ randomly
  Obtain the weight matrix $\mathbf{W}_{\mathrm{out}}$ by solving the following problem
      $\min \sum_{k=0}^{T} \|\mathbf{W}_{\mathrm{out}}\mathbf{r}_k - \mathbf{y}_k\|^2 + \lambda\|\mathbf{W}_{\mathrm{out}}\|_2^2$    (11)
Output: Initial reservoir $\mathbf{r}_0$, Weight matrices $\mathbf{W}_{\mathrm{out}}$, $\mathbf{W}_{\mathrm{back}}$, and $\mathbf{W}_{\mathrm{rc}}$
```

Algorithm 1. Implementation of reservoir computing for a time series dataset $\mathcal{D}_T^{\mathrm{rc}}$.

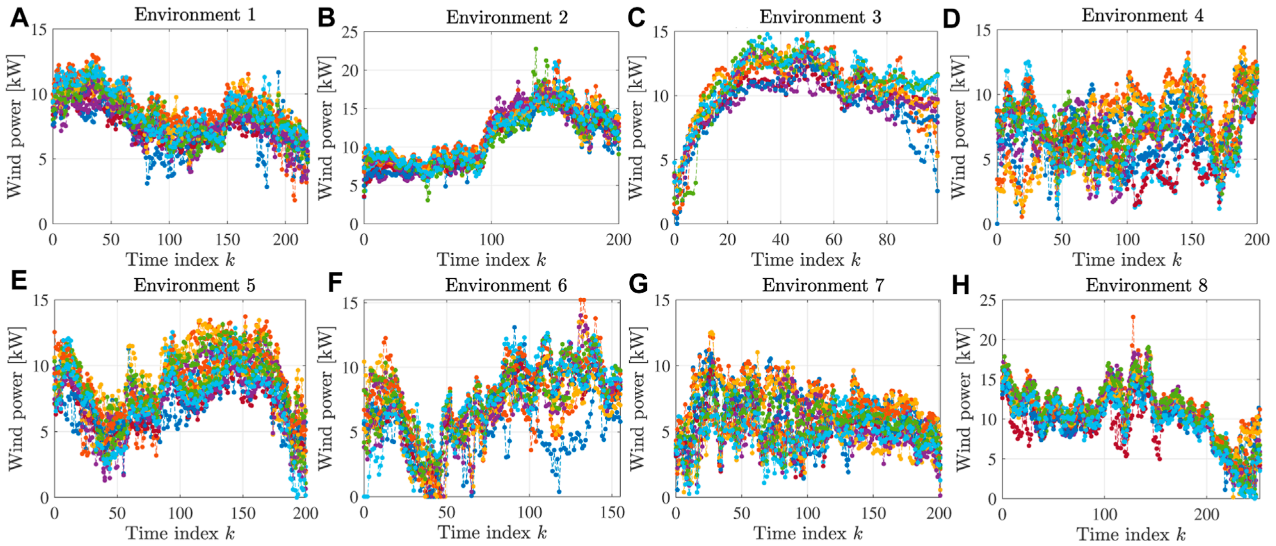


FIGURE 1
Experimental data used in this validation. The dataset includes data from eight different environments, plotted in **(A–H)**. For each environment, there are 13 different profiles.

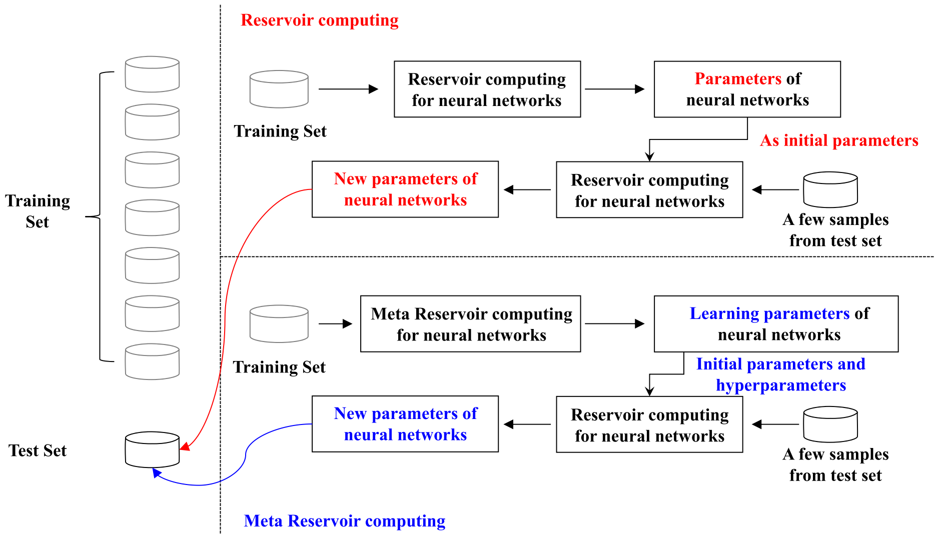


FIGURE 2
Implementations of the meta-reservoir computing and reservoir computing methods.

vector. Let $\mathcal{L}^{\text{train}}(\cdot)$ and $\mathcal{L}^{\text{test}}(\cdot)$ be the loss function for training and testing, respectively. Then, the training process in meta-learning can be formulated as

$$\min_{\omega} \sum_{i=1}^{K} \mathcal{L}^{\text{test}}\left(\mathcal{D}_{T_i}^{i,\text{te}}, \theta_*^{(i)}(\omega), \omega\right), \tag{12}$$

$$\text{s.t.} \quad \theta_*^{(i)}(\omega) = \arg\min_{\theta} \mathcal{L}^{\text{train}}\left(\mathcal{D}_{T_i}^{i,\text{tr}}, \theta, \omega\right), \; i = 1, \ldots, K. \tag{13}$$

Let ω^* be the solution to the bi-level optimization problem. In every iteration, the learning parameter ω is optimized for a given θ, and finally, it converges to the optimal value for learning a task. The optimality here refers to the given training dataset. Even for a newly given task, the learning parameter ω^* can provide better efficiency to find the optimal parameter for the newly given task.

3.3 Algorithm for meta-reservoir computing

This paper proposes a novel wind power predictive model learning algorithm that combines reservoir computing and meta-learning. As introduced in Section 2, we have the dataset $\mathcal{D}_T$

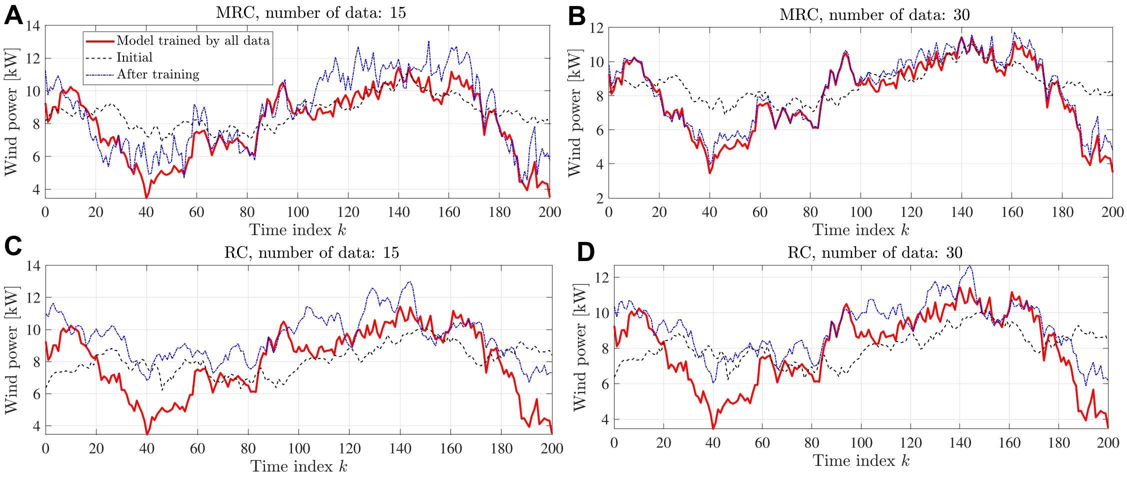


FIGURE 3
Few-shot adaptation for environment 5. **(A)** MRC with 15 samples; **(B)** MRC with 30 samples; **(C)** RC with 15 samples; and **(D)** RC with 30 samples.

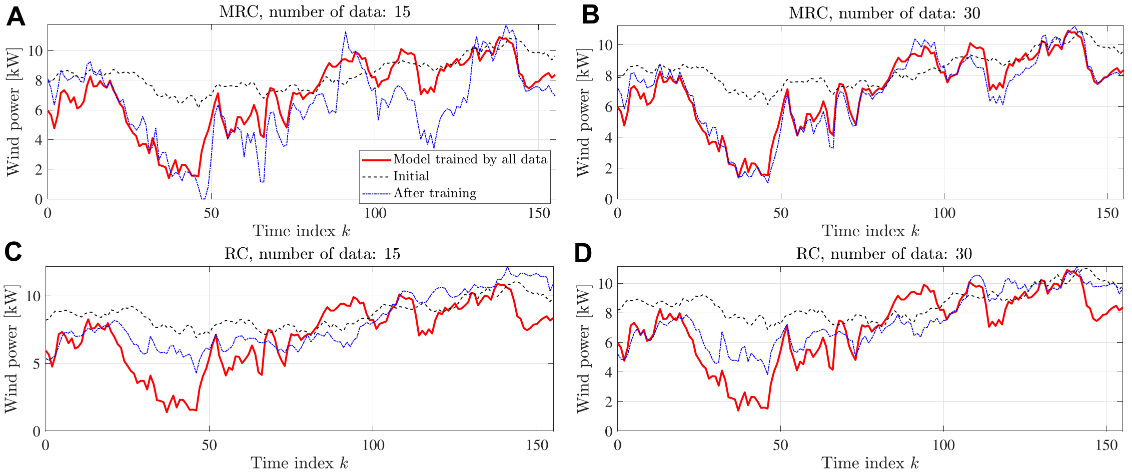


FIGURE 4
Few-shot adaptation for environment 6. **(A)** MRC with 15 samples; **(B)** MRC with 30 samples; **(C)** RC with 15 samples; and **(D)** RC with 30 samples.

obtained from multiple environments. Regarding each environment as a task, we separate the dataset into

$$\mathcal{D}^i := \left\{p_k^{(i)}\right\}_{k=0}^{T_i}, \quad i = 1,\ldots,K. \tag{14}$$

Note that we have

$$\bigcup_{i=1}^{K} \mathcal{D}^i = \mathcal{D}_T \tag{15}$$

and

$$\mathcal{D}^i \bigcap \mathcal{D}^j = \emptyset, \ \text{if } i \neq j. \tag{16}$$

For each task, we further separate the dataset into data for training and data for testing as follows:

$$\mathcal{D}^i_{\text{train}} := \left\{p_k^{(i),\text{train}}\right\}_{k=0}^{T_{i,\text{train}}}, \ \mathcal{D}^i_{\text{test}} := \left\{p_k^{(i),\text{test}}\right\}_{k=0}^{T_{i,\text{test}}}. \tag{17}$$

Note that we have

$$\mathcal{D}^i_{\text{train}} \bigcap \overset{i}{\mathcal{D}}_{\text{test}} = \emptyset, \tag{18}$$

$$\mathcal{D}^i_{\text{train}} \bigcup \overset{i}{\mathcal{D}}_{\text{test}} = \mathcal{D}^i, \tag{19}$$

for every task $i = 1, \ldots, K$.

We use reservoir computing to train a temporal prediction model of wind power. Thus, the parameter to be trained is $\mathbf{W}^{(i)}_{\text{out}}$. There are a lot of hyperparameters to be optimized, such as the initial point of reservoirs, the initial point of the solution of the optimization problem for obtaining $\mathbf{W}^{(i)}_{\text{out}}$, the gradient-descent rate, and the ratio λ for regularization. This study adopts the initial point

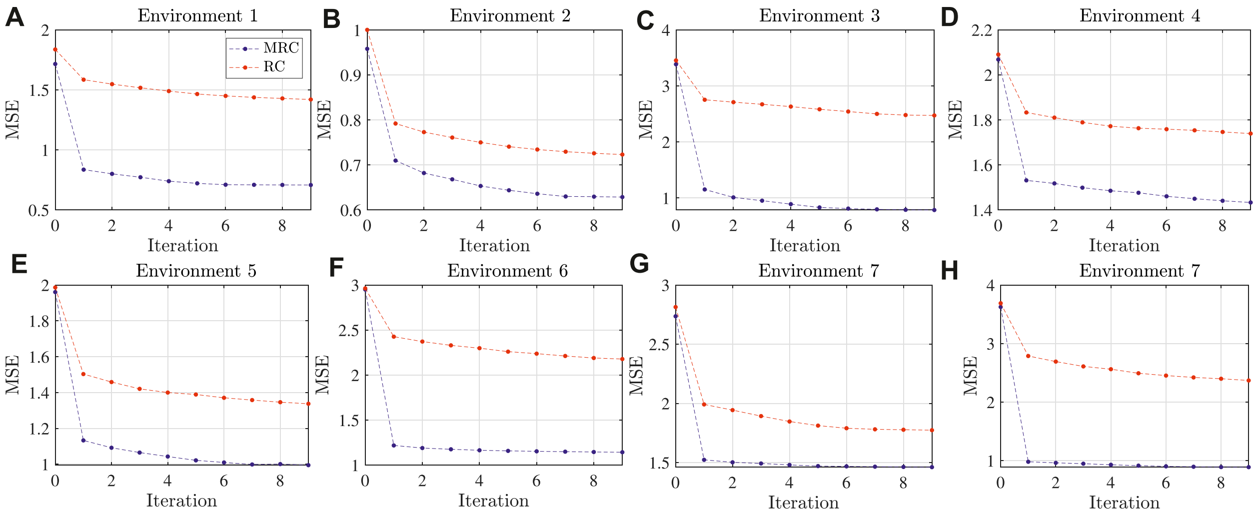


FIGURE 5
Quantitative prediction results showing the MSE curve along with the training iteration (30 samples): **(A)** Environment 1; **(B)** Environment 2; **(C)** Environment 3; **(D)** Environment 4; **(E)** Environment 5; **(F)** Environment 6; **(G)** Environment 7; **(G)** Environment 8.

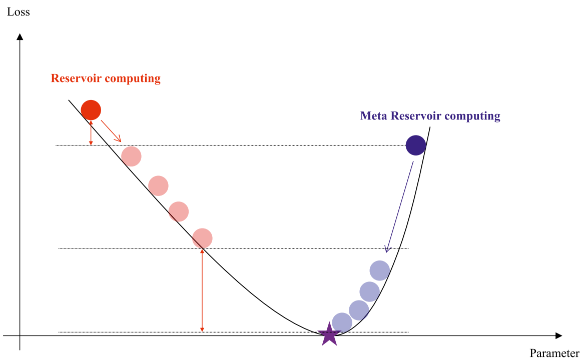


FIGURE 6
Intuitive explanation of the quantitative prediction results.

```
Inputs: dataset $\mathcal{D}_T = \{p_k\}_{k=0,\ldots,T} = \{\mathcal{D}^1_{\mathrm{train}}, \mathcal{D}^1_{\mathrm{test}}, \ldots\ldots, \mathcal{D}^K_{\mathrm{train}}, \mathcal{D}^K_{\mathrm{test}}\}$, the regularization ratio $\lambda$, a new dataset $\mathcal{D}^{K+1}_{T'}$
  1: Select the model $\mathbf{f}_{\mathrm{rc}}$ and reservoir $\mathbf{r}_k$
  2: Generate weight matrices $\mathbf{W}_{\mathrm{back}}$ and $\mathbf{W}_{\mathrm{rc}}$ randomly
  3: Solve the problem described by Eqs 23, 24 with dataset $\mathcal{D}_T$ and obtain $\omega*$
  4: Obtain the weight matrix $\mathbf{W}^{(K+1)}_{\mathrm{out}}$ by solving the following problem described by Eq. 11 with dataset $\mathcal{D}^{K+1}_{T'}$, parameters $\mathbf{f}_{\mathrm{rc}}$, $\mathbf{r}_k$, $\omega*$, $\mathbf{W}_{\mathrm{back}}$, and $\mathbf{W}_{\mathrm{rc}}$
Output: $\omega = (\mathbf{r}_0, \mathbf{W}_{\mathrm{out}}(0))$, weight matrices $\mathbf{W}^{K+1}_{\mathrm{out}}$, $\mathbf{W}_{\mathrm{back}}$, and $\mathbf{W}_{\mathrm{rc}}$
```

Algorithm 2. Implementation of meta-reservoir computing for learning a wind power predictive model.

of the solution to the optimization problem, $\mathbf{W}_{\mathrm{out}}(0)$, and the initial point of reservoirs $\mathbf{r}_0$ as the learning parameter ω^*. Namely, we have

$$\omega^* = \left(\mathbf{W}_{\mathrm{out}}(0), \mathbf{r}_0\right). \tag{20}$$

Note that the learning parameter ω^* is common for each task, and the parameter differs in each task. For the loss function, we adopt the loss function of Eq. 11 for both training and testing processes. It is written as

$$\mathcal{L}^{\mathrm{train}}\left(\mathcal{D}^i_{\mathrm{train}}, \mathbf{W}^{(i)}_{\mathrm{out}}, \omega\right) = \sum_{k=0}^{T_{i,\mathrm{test}}} \left\| W^{(i)}_{\mathrm{out}} \mathbf{r}^{(i)}_k - p^{(i),\mathrm{train}}_k \right\|^2 + \lambda \|\mathbf{W}^{(i)}_{\mathrm{out}}\|^2_2, \tag{21}$$

$$\mathcal{L}^{\mathrm{test}}\left(\mathcal{D}^i_{\mathrm{train}}, \mathbf{W}^{(i),*}_{\mathrm{out}}, \omega\right) = \sum_{k=0}^{T_{i,\mathrm{train}}} \left\| W^{(i),*}_{\mathrm{out}} \mathbf{r}^{(i)}_k - p^{(i),\mathrm{test}}_k \right\|^2 + \lambda \|\mathbf{W}^{(i),*}_{\mathrm{out}}\|^2_2. \tag{22}$$

Then, the training process in meta-reservoir computing is written as

$$\min_{\omega} \sum_{i=1}^{K} \mathcal{L}^{\mathrm{test}}\left(\mathcal{D}^i_{\mathrm{train}}, \mathbf{W}^{(i),*}_{\mathrm{out}}, \omega\right), \tag{23}$$

$$\text{s.t.} \quad \mathbf{W}^{(i),*}_{\mathrm{out}} = \arg\min_{\mathbf{W}^{(i)}_{\mathrm{out}}} \mathcal{L}^{\mathrm{train}}\left(\mathcal{D}^i_{\mathrm{train}}, \mathbf{W}^{(i)}_{\mathrm{out}}, \omega\right), \; i = 1, \ldots, K. \tag{24}$$

According to the above discussions, we summarize the meta-reservoir computing algorithm for learning a wind power predictive model in Algorithm 2.

4 Experimental validation

In this section, we first introduce the experimental dataset and several settings for validation. The validation results are then presented with detailed discussions.

4.1 Dataset for validation

This validation uses the experimental dataset shown in Figure 1. This dataset includes time series data obtained from eight different environments, as shown in Figures 1A–H, respectively. In addition, for each environment, we have 13 different profiles. The same environment means that the data were collected in the same period, places, and weather.

For meta-learning, seven profiles in each environment are used as training data and the rest as test data. For a fair comparison, we compare our MRC with normal reservoir computing (RC) without meta-learning. In each validation, we use data from seven environments to train a model and then use the rest for validation.

The comparison of the implementations of the MRC and RC methods is shown in Figure 2. For RC methods, the training set is used to train a recurrent neural network. We obtain only a parameter vector for the recurrent neural network. When a new task comes as a test set, only a few shots in the test set are used for training a new recurrent neural network. The trained parameter vector can be used as the initial value of reservoir computing for updating the recurrent neural network for the new task. For MRC methods, the training set is used for meta-learning. Except for one parameter vector for the training set, the learning parameter, including a good initial point, is also obtained. When a new task comes, the learning parameter is used to learn a new parameter vector for the new task.

4.2 Validation results

The performance of the MRC and RC methods is evaluated by checking the accuracy of the model learned by each method with a fine-tuning process based on a sample number of 15 or 30 from a new task. During fine-tuning, each gradient-descent step is computed with the same data points. Figures 3, 4 provide the qualitative results for using environments 5 and 6 for the test. The red solid line is the model trained by using all the data in the test set, which can be regarded as a perfect model. The results show that the MRC method can provide a model very close to perfection, even with a few shots of data. Note that both MRC and RC methods do not have good initial points. However, the MRC method can adapt the model very quickly. The RC method fails to adapt the model to a proper model with the limited data number.

Figure 5 provides more quantitative prediction results. The mean square errors (MSEs) of the model at each iteration are plotted for each case with a different environment as the test set. It is obvious that the proposed MRC method can adapt the model to a given environment even though the initial MSE is almost the same as that of the RC method. The reason is that the MRC method optimizes the initial value of the reservoir, which may provide some information to find a better gradient to reduce the loss. The intuitive explanation is given in Figure 6.

5 Conclusion

A wind power prediction model must be able to be adapted to a new environment, with a few samples of data from the new environment. The traditional deep learning methods encounter the overfitting issue and are hard to be adapted to a new environment. A huge dataset is still needed. This paper proposes a novel learning method for a wind power prediction model. The reservoir computing algorithm is combined with meta-learning to efficiently adapt the wind power prediction model to a new environment with only a few samples. The algorithmic structure of reservoir computing significantly reduces the computational complexity of learning a deep model. On the other hand, the initial points and other hyperparameters of reservoir computing are optimized by meta-learning based on the historical dataset. Experimental datasets have validated the proposed meta-reservoir computing method for learning the wind power prediction model. The validation results show that the proposed meta-reservoir computing can find a good model for the new environment in a very small number of iterations with a few shots of new data.

The proposed method opens a new avenue for training wind power predictive models for different environments. Instead of giving the best point for each environment, it is better to find a good learning parameter to be ready for new tasks. In future work, we will investigate comparing the proposed method with more existing deep models.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

LZ: conceptualization, data curation, formal analysis, investigation, methodology, project administration, supervision, validation, writing–original draft, and writing–review and editing. H-XA: formal analysis, funding acquisition, investigation, methodology, project administration, resources, validation, and writing–review and editing. Y-XL: data curation, funding acquisition, investigation, resources, validation, and writing–review and editing. L-XX: data curation, formal analysis, methodology, project administration, supervision, validation, and writing–review and editing. CD: formal analysis, project administration, resources, validation, and writing–review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

Authors LZ, H-XA, Y-XL, L-XX, and CD were employed by Xiangyang Electric Power Supply Company, State Grid.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Burke, D., and O'Malley, J. (2011). A study of principal component analysis applied to spatially distributed wind power. *IEEE Trans. Power Syst.* 26, 2084–2092. doi:10.1109/TPWRS.2011.2120632

Cascianelli, S., Astolfi, D., Castellani, F., Cucchiara, R., and Fravolini, M. L. (2022). Wind turbine power curve monitoring based on environmental and operational data. *IEEE Trans. Industrial Inf.* 18, 5209–5218. doi:10.1109/TII.2021.3128205

Chen, M., Zeng, G. Q., Lu, K. D., and Weng, J. (2019). A two-layer nonlinear combination method for short-term wind speed prediction based on elm, enn, and lstm. *IEEE Internet Things J.* 6, 6997–7010. doi:10.1109/JIOT.2019.2913176

Duffy, K., Vandal, T. J., Wang, W., Nemani, R. R., and Ganguly, A. R. (2023). A framework for deep learning emulation of numerical models with a case study in satellite remote sensing. *IEEE Trans. Neural Netw. Learn. Syst.* 34, 3345–3356. doi:10.1109/TNNLS.2022.3169958

Hamedani, K., Liu, L., Atat, R., Wu, J., and Yi, Y. (2018). Reservoir computing meets smart grids: attack detection using delayed feedback networks. *IEEE Trans. Industrial Inf.* 14, 734–743. doi:10.1109/TII.2017.2769106

Jaeger, H., and Haas, H. (2004). Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. *Science* 304, 78–80. doi:10.1126/science.1091277

Ko, M., Lee, K., Kim, J. K., Hong, C. W., Dong, Z. Y., and Hur, K. (2021). Deep concatenated residual network with bidirectional lstm for one-hour-ahead wind power forecasting. *IEEE Trans. Sustain. Energy* 12, 1321–1335. doi:10.1109/TSTE.2020.3043884

Lange, M. (2005). On the uncertainty of wind power predictions—analysis of the forecast accuracy and statistical distribution of errors. *J. Sol. Energy Eng.* 127, 177–184. doi:10.1115/1.1862266

Li, J., and Hu, M. (2021). Continuous model adaptation using online meta-learning for smart grid application. *IEEE Trans. Neural Netw. Learn. Syst.* 32, 3633–3642. doi:10.1109/TNNLS.2020.3015858

Liu, H., Duan, Z., Li, Y., and Lu, H. (2018). A novel ensemble model of different mother wavelets for wind speed multi-step forecasting. *Appl. Energy* 228, 1783–1800. doi:10.1016/j.apenergy.2018.07.050

Luo, Z., Fang, C., Liu, C., and Liu, S. (2022). Method for cleaning abnormal data of wind turbine power curve based on density clustering and boundary extraction. *IEEE Trans. Sustain. Energy* 13, 1147–1159. doi:10.1109/TSTE.2021.3138757

Nokkala, J., Martinez-Pena, R., Zambrini, R., and Soriano, M. C. (2022). High-performance reservoir computing with fluctuations in linear networks. *IEEE Trans. Neural Netw. Learn. Syst.* 33, 2664–2675. doi:10.1109/TNNLS.2021.3105695

Pandit, R., Infield, D., and Santos, M. (2023). Accounting for environmental conditions in data-driven wind turbine power models. *IEEE Trans. Sustain. Energy* 14, 168–177. doi:10.1109/TSTE.2022.3204453

Safari, N., Chung, C. Y., and Price, G. C. D. (2018). Novel multi-step short-term wind power prediction framework based on chaotic time series analysis and singular spectrum analysis. *IEEE Trans. Power Syst.* 33, 590–601. doi:10.1109/TPWRS.2017.2694705

Tanaka, G., Yamane, T., Héroux, J. B., Nakane, R., Kanazawa, N., Takeda, S., et al. (2019). Recent advances in physical reservoir computing: a review. *Neural Netw.* 115, 100–123. doi:10.1016/j.neunet.2019.03.005

Tian, P., Li, W., and Gao, Y. (2022). Consistent meta-regularization for better meta-knowledge in few-shot learning. *IEEE Trans. Neural Netw. Learn. Syst.* 33, 7277–7288. doi:10.1109/TNNLS.2021.3084733

Ummels, B., Gibescu, M., Pelgrum, E., Kling, W. L., and Brand, A. J. (2007). Impacts of wind power on thermal generation unit commitment and dispatch. *IEEE Trans. Energy Convers.* 22, 44–51. doi:10.1109/TEC.2006.889616

Wang, J., AlShelahi, A., You, M., Byon, E., and Saigal, R. (2021). Integrative density forecast and uncertainty quantification of wind power generation. *IEEE Trans. Sustain. Energy* 12, 1864–1875. doi:10.1109/TSTE.2021.3069111

Wu, Y., Huang, C. L., Wu, S. H., Hong, J. S., and Chang, H. L. (2023). Deterministic and probabilistic wind power forecasts by considering various atmospheric models and feature engineering approaches. *IEEE Trans. Industrial Appl.* 59, 192–206. doi:10.1109/TIA.2022.3217099

Zhang, H., Yan, J., Liu, Y., Gao, Y., Han, S., and Li, L. (2021). Multi-source and temporal attention network for probabilistic wind power prediction. *IEEE Trans. Sustain. Energy* 12, 2205–2218. doi:10.1109/TSTE.2021.3086851

Zhao, C., Wan, C., and Song, Y. (2021). Operating reserve quantification using prediction intervals of wind power: an integrated probabilistic forecasting and decision methodology. *IEEE Trans. Power Syst.* 36, 3701–3714. doi:10.1109/TPWRS.2021.3053847

Frontiers in Energy Research

TYPE Brief Research Report
PUBLISHED 20 February 2024
DOI 10.3389/fenrg.2023.1239973

Check for updates

OPEN ACCESS

EDITED BY
Xun Shen,
Osaka University, Japan

REVIEWED BY
Shuang Zhao,
Hefei University of Technology, China
Yahui Zhang,
Yanshan University, China
Hardeep Singh,
University of Windsor, Canada

*CORRESPONDENCE
Meng-Yi Wu,
✉ wmy@itsc.cn

RECEIVED 14 June 2023
ACCEPTED 25 July 2023
PUBLISHED 20 February 2024

CITATION
Gao L-D, Li Z-H, Wu M-Y, Fan Q-L, Xu L,
Zhang Z-M, Zhang Y-P and Liu Y-Y
(2024), Interval reservoir computing:
theory and case studies.
Front. Energy Res. 11:1239973.
doi: 10.3389/fenrg.2023.1239973

COPYRIGHT
© 2024 Gao, Li, Wu, Fan, Xu, Zhang, Zhang and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Interval reservoir computing: theory and case studies

Lan-Da Gao[1,2], Zhen-Hua Li[1,2], Meng-Yi Wu[1,2]*, Qing-Lan Fan[1,2], Ling Xu[1,2], Zhuo-Min Zhang[1,2], Yi-Peng Zhang[1,2] and Yan-Yue Liu[1,2]

[1]Research Institute of Highway, Ministry of Transport, Beijing, China, [2]Key Laboratory of Technology on Intelligent Transportation Systems, Ministry of Transport, Beijing, China

The time series data in many applications, for example, wind power and vehicle trajectory, show significant uncertainty. Using a single prediction value of wind power as feedback information for wind turbine control or unit commitment is not enough since the uncertainty of the prediction is not described. This paper addresses the uncertainty issue in time series data forecasting by proposing the novel interval reservoir computing method. The proposed interval reservoir computing can capture the underlying evolution of the stochastic dynamical system for time series data using the recurrent neural network (RNN). On the other hand, by formulating a chance-constrained optimization problem, interval reservoir computing outputs a set of parameters in the RNN, which maps to an interval of prediction values. The capacity of the interval is the smallest one satisfying the condition that the probability of having a prediction inside the interval is lower than the required level. The scenario approach solves the formulated chance-constrained optimization problem. We implemented an experimental data-based validation to evaluate the proposed method. The validation results show that the proposed interval reservoir computing can give a tight interval of time series data forecasting values for wind power and traffic trajectory. In addition, the confidence probability over the feasibility goes to 1 very quickly as the sample number increases.

KEYWORDS

uncertain dynamical systems, probabilistic prediction, time series data, wind power forecasting, vehicle trajectory

1 Introduction

Time series data prediction is vital in many applications for pursuing better control or decision-making performance toward achieving a better society or quality of life. For example, to accomplish the net-zero carbon goal, it is vital to establish a reliable power system with renewable energy for energy supplement instead of high-carbon power generation (Evans et al., 2021). Wind energy is one of the best choices among different kinds of renewable energy resources. However, wind power has a stochastic nature, which makes it challenging to realize the optimal wind power supplementation with high reliability (Zhao et al., 2018; Ge et al., 2022). It is necessary to provide a reliable wind power prediction for the security-constrained unit commitment (SCUC) problem to improve the optimality and reliability of wind power supplementation (Hu and Wu, 2016). Instead of using one single wind power prediction, the SCUC problem involved with wind power often considers the uncertain nature of wind power. It is formulated as a stochastic program (Chen et al., 2015). The random variables, such as wind power, are assumed to be within a bounded set in the formulated stochastic program (Hu et al., 2014; Dai et al., 2016). Only with a reliable set for the random variable will the solution of the stochastic program be faithful. Another critical application scene is safety control in complex traffic environments. It is crucial to give reliable sets for the trajectories of traffic participants surrounding the self-vehicle (Liu et al, 2022; Shen and Raksincharoensak, 2022). For example, Liu et al. (2022) proposed a dynamic lane-changing trajectory generator based on the uncertain evaluation of other vehicle trajectories. Yu et al. (2022) proposed a robust and safe trajectory planning method, considering a bounded uncertainty for the other vehicle trajectories. Lyu et al. (2022) improved the vehicle trajectory prediction accuracy using the information from the connected environment. Thus, a reliable set for the random variable's prediction is vital for robust control and decision-making.

However, giving a reliable set for the random variable's prediction is challenging due to the computational complexity issue. Conformal prediction is a method to provide scores on confidence in the prediction value and then gives a confidence interval (Wang et al., 2021). It can also be applied to deep neural networks (Wen et al., 2021). However, it suffers from the "curse of dimensionality." The computational complexity becomes impractical for applications as the dimension of parameters in the model increases. The Bayesian neural network is an alternative way to provide the confidence interval of the predictions (Neal, 2012; Chen et al., 2021; Xue et al., 2022). The uncertainty is represented by giving the weights on the parameters of the neural network. However, the Bayesian neural network needs many assumptions for practical implementation. A neural network that maps an input to an interval of the predictions is called an interval neural network (INN), first proposed in Ishibuki et al. (1993). Compared to the Bayesian neural network, the INN requires fewer assumptions and can provide probabilistic guarantees on the reliability of the obtained neural network (Ak et al., 2015). Recently, a machine learning-based method, called interval predictor models, was proposed in Campi et al. (2009) and Garattia et al. (2019). The problem of constructing an interval predictor model can be formulated as a chance-constrained optimization problem (Shen et al., 2023). The above methods do not consider the neural networks for dynamic systems. In this paper, we extend the above method to recurrent neural networks combining reservoir computing (Jaeger and Haas, 2004) to address the uncertain quantification problem for predictions in dynamic systems. We call the proposed method "interval reservoir computing." The proposed interval reservoir computing can capture the underlying evolution of the stochastic dynamical system for time series data using the recurrent neural network (RNN). On the other hand, by formulating a chance-constrained optimization problem, interval reservoir computing outputs a set of parameters in the RNN, which maps to an interval of wind power prediction values. The capacity of the interval is the smallest one satisfying the condition that the probability of having a prediction inside the interval is lower than the required level. The scenario approach solves the formulated chance-constrained optimization problem. We implemented experimental data-based validation to evaluate the proposed method.

The rest of this paper is organized as follows: Section 2 gives a general problem formulation of interval prediction in dynamical systems; Section 3 presents the proposed interval reservoir after briefly introducing reservoir computing and the scenario approach; Section 4 gives the experimental data-based validation; Section 5 concludes the whole paper and discusses future work.

2 Problem formulation: prediction in dynamical systems

In wind power or vehicle trajectory forecasting applications, time series data are generated by an underlying stochastic dynamical system. The stochastic dynamical system has system inputs, hidden states which cannot be observed, and observations that sensors can measure. A graphical model of the addressed stochastic dynamical system is illustrated by Figure 1. Let $t \in \mathbb{Z}$ be the time index. The hidden state is denoted by $x_t \in \mathbb{R}^k$. The system input is represented by $u_t \in \mathbb{R}^c$. The observation is $y_t \in \mathbb{R}^d$. Note that x_t is not available, and only the data on y_t and u_t can be obtained from the sensors. The observation y_t depends on u_t and x_t. However, for given values of u_t and x_t, the observation y_t is not deterministic but with uncertainty. The observation y_t is a random variable with a conditional probability distribution $p_t(y|x_t, u_t)$. An alternative way is to use a function involved with random variables. Let $g\colon \mathbb{R}^k \times \mathbb{R}^c \times \mathbb{R}^m \to \mathbb{R}^d$ be the function that gives the observation from state and input in the following way:

$$y_t = g\left(x_t, u_t, w_t\right), \tag{1}$$

where $w_t \in \mathbb{R}^m$ denotes the m-dimension observation noise with the probability density function $r(w)$. On the other hand, the system transition is also involved with uncertainty. Let $f\colon \mathbb{R}^k \times \mathbb{R}^c \times \mathbb{R}^l \to \mathbb{R}^k$ be the function that gives the state of the next time index from the state and input in the following way:

$$x_{t+1} = f\left(x_t, u_t, v_t\right), \tag{2}$$

where $v_t \in \mathbb{R}^l$ is the l-dimension system noise with the probability density function $q(v)$. The initial state vector x_0 is distributed according to the probability density $p_0(x_0)$.

The information on $f(\cdot)$, $g(\cdot)$, $r(w)$, and $q(v)$ is unavailable. In this study, the available information is the dataset $U_T = \{u_0, u_1, \ldots, u_T\}$ of system inputs and the dataset $Y_T = \{y_0, y_1, \ldots, y_T\}$ of observations. We want to learn models $\tilde{f}(\cdot)$, $\tilde{g}(\cdot)$, $\tilde{r}(w)$, and $\tilde{q}(v)$. The traditional view is to learn $\tilde{f}(\cdot)$, $\tilde{g}(\cdot)$, $\tilde{r}(w)$, and $\tilde{q}(v)$ for the sake of improving the performance of the root mean square (RMS) of predictions or maximizing the likelihood of the dataset. In this paper, we obtain a novel prediction model that gives a predictive interval of the observation. We define an interval of y_t as follows.

Definition 1: *Let $\mathcal{F}_Y$ be the Borel set of $\mathbb{R}^d$. An interval $I_t \in \mathcal{F}_Y$ is a set of y_t. For a given probability level $\alpha \in (0, 1)$, if $I_t \in \mathcal{F}_Y$ satisfies*

$$\Pr_t\{y_t \in I_t \in \mathcal{F}_Y\} \geq 1 - \alpha, \tag{3}$$

where $\Pr_t\{\cdot\}$ is the underlying probability measure defined on $\mathcal{F}_Y$ at time index t, we call I_t as a α-reliable interval. The set of all α– confidence intervals is defined as $\mathcal{I}_{t,\alpha}$. For a given probability level $\alpha \in (0, 1)$, an optimal interval I_t^* satisfies

$$\mathbb{C}\left(I_t^*\right) \leq \mathbb{C}\left(I_t\right), \quad \forall I_t \in \mathcal{I}_\alpha, \tag{4}$$

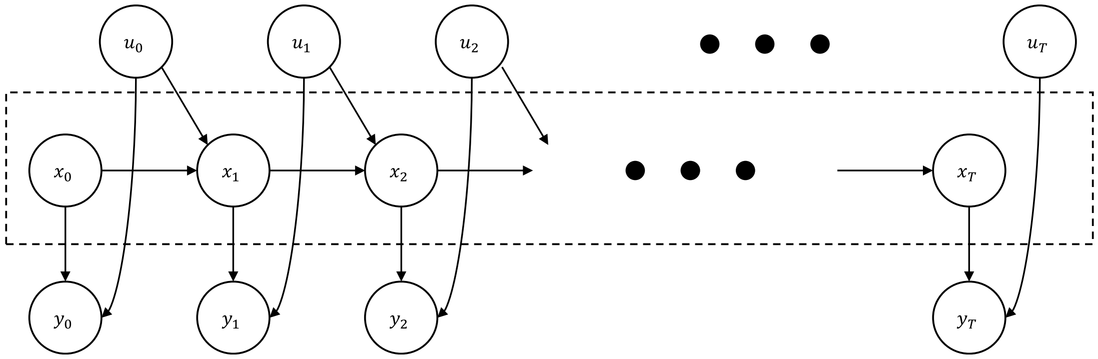


FIGURE 1
Probabilistic graphical model for stochastic dynamical systems with hidden states x_t, inputs u_t, and observations y_t.

where $\mathbb{C}(\cdot)$ *denotes the capacity of a set.*

The problem is formally summarized in Problem 1.

Problem 1: *Given the system input set* $U_T = \{u_0, u_1, \ldots, u_T\}$ *and observation set* $U_T = \{y_0, y_1, \ldots, y_T\}$, *to obtain* $\tilde{f}^*(\cdot)$, $\tilde{g}^*(\cdot)$, $\tilde{r}^*(w)$, *and* $\tilde{q}^*(v)$ *by solving*

$$\begin{aligned} &\min_{\tilde{f},\tilde{g},\tilde{r},\tilde{q}} \mathbb{C}\left(I_t\left(\tilde{f},\tilde{g},\tilde{r},\tilde{q}\right)\right) \\ &\qquad \Pr_t\left\{y_t \in I_t\left(\tilde{f},\tilde{g},\tilde{r},\tilde{q}\right) \in \mathcal{F}_Y\right\} \geq 1-\alpha, \ \forall t = 1,\ldots,T. \end{aligned} \tag{5}$$

3 Proposed method

In this section, we briefly review the reservoir computing and scenario approach. Then, we give the concept of interval reservoir computing, combining reservoir computing and a scenario approach. The probabilistic reliability of interval reservoir computing to Problem 1 is also given.

3.1 Brief introduction to reservoir computing

Reservoir computing is a novel algorithm to train a RNN. This study uses the echo state network (ESN) method presented in Jaeger and Haas (2004) to construct RNN. Let $x_{\mathrm{act},t}$ be the activation state of RNN at time index t. The terminology “echo” implies that $x_{\mathrm{act},t}$ is a function of all the input history $u_{t-1}, \ldots$ related to the network. The ESN consists of multiple sigmoidal units in discrete time, the so-called reservoir or dynamic reservoir. A general ESN has a discrete-time neural network with internal network units (for state $x_{\mathrm{sta},t}$), input units (for input u_t), and output units (for observation y_t). The internal units are updated as follows.

$$x_{\mathrm{act},t+1} = f_{\mathrm{act}}\left(W^{\mathrm{in}}u_t + Wx_{\mathrm{act},t} + W^{\mathrm{back}}y_t\right), \tag{6}$$

where f_{act} is the vector function of the internal unit written as $f_{\mathrm{act}} = \left[f_{\mathrm{act}}^1, \ldots, f_{\mathrm{act}}^{n_{\mathrm{act}}}\right]^\top$.

On the other hand, the output is computed as

$$y_t = W^{\mathrm{out}}x_{\mathrm{act},t}, \tag{7}$$

where W^{out} is the output weight. Reservoir computing is to train W^{in}, W, W^{back}, and W^{out}, and the algorithm is summarized as follows:

- Design of a reservoir vector: a reservoir vector $x_{\mathrm{act},t}$ and the internal unit, as shown in Eq. 6, are established.
- Randomly generating W^{in}, W, and W^{back}, which comprise a sparse random matrix with the maximal eigenvalue controlled.
- Determining the output layer by

$$\min_{W^{\mathrm{out}}} \sum_{t=1}^{T} \|W^{\mathrm{out}}x_t - y_t\|^2 + \beta \mathrm{Trace}\left(W^{\mathrm{out}}W^{\mathrm{out},\top}\right). \tag{8}$$

Figure 2 illustrates the reservoir computing concept.

3.2 Scenario approach

The scenario approach has been applied to obtain the probabilistic boundary for a given nonlinear state space model (Shen et al., 2020a). The theory of the scenario approach has been presented in Calafiore and Campi (2005) for solving robust optimization with the convex objective function and constraint functions. The result has been extended to non-convex cases in Campi et al. (2015). This paper reviews the method of Campi et al. (2015).

The decision variable is $\theta \in \Theta \subseteq \mathbb{R}^{n_\theta}$. Let $J{:}\Theta \to \mathbb{R}$ be the objective function. The uncertain variable is denoted by $\delta \in \Delta \subseteq \mathbb{R}^{n_\delta}$. For every instance $\delta \in \Delta$, a subset of Θ is defined by

$$\Theta_\delta = \{\theta \in \Theta : h(\theta, \delta) \leq 0\},$$

where $h{:}\,\Theta \times \Delta \to \mathbb{R}^m$ is a constraint function. Then, a robust optimization problem can be written as

$$\begin{aligned} &\min_{\theta \in \Theta} \ J(\theta) \\ &s.t. \quad \theta \in \Theta_\delta, \ \forall \delta \in \Delta. \end{aligned} \tag{9}$$

Problem Eq. 9 is NP-hard and cannot be solved by any algorithms for a general optimization problem. An approximate problem of

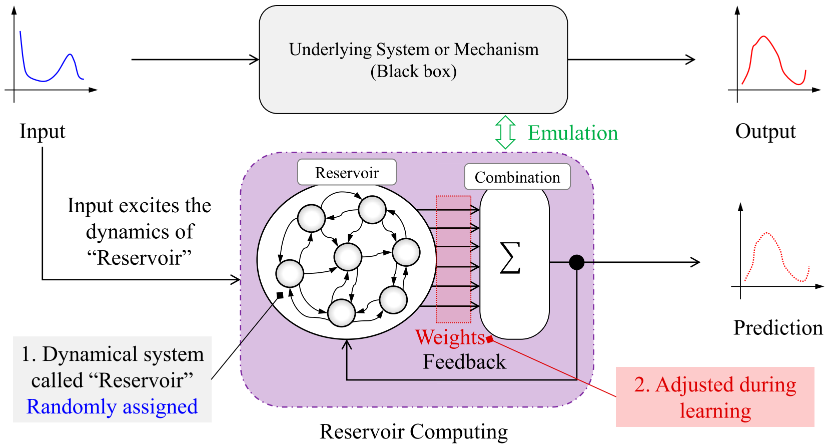


FIGURE 2
Intuitive introduction for reservoir computing.

that in Eq. 9 is obtained by sampling a dataset $\delta^{(1)}, \ldots, \delta^{(N)}$, which is written by

$$\begin{aligned} &\min_{\theta \in \Theta} \; J(\theta) \\ &s.t. \quad \theta \in \Theta_{\delta^{(i)}}, \; \forall i = 1, \ldots, N. \end{aligned} \tag{10}$$

Let θ_N^* be an optimal solution of the problem in Eq. 10 and $\mathcal{A}_N$ be an algorithm that is able to get θ_N^* for a given dataset $(\delta^{(1)}, \ldots, \delta^{(N)}) \in \Delta^N$. Then, we have

$$\theta_N^* = \mathcal{A}_N\left(\delta^{(1)}, \ldots, \delta^{(N)}\right). \tag{11}$$

It is natural to doubt whether θ_N^* is a feasible solution of the problem in Eq. 9 since θ_N^* does not consider constraints for all $\delta \in \Delta$. Here, since the sampling process of the dataset $(\delta^{(1)}, \ldots, \delta^{(N)}) \in \Delta^N$ is random, we consider the feasibility of θ_N^* in a probabilistic sense. We define the violation probability herein.

Definition 2: *The violation probability of any decision* $\theta \in \Theta$ *is written as*

$$\mathbb{V}(\theta) \coloneqq \Pr_\delta \{\delta \in \Delta : \theta \notin \Theta_\delta\},$$

where $\Pr_\delta\{\cdot\}$ *defines the probability measure defined on the* σ*-algebra of* Δ.

For a given probability level $\varepsilon \in (0,1)$ and a given confidence bound $1-\beta \in (0,1)$, we want to get a bound of sample number $\bar{N}(\beta, \varepsilon)$ such that

$$\Pr_\delta\left\{\mathbb{V}\left(\theta_N^*\right) \leq \varepsilon\right\} \geq 1-\beta, \quad \forall N \geq \bar{N}(\beta,\varepsilon).$$

Theorem 1 of Campi et al. (2015) is stated as follows:

Lemma 1: *Let* $\varepsilon : \{0, \ldots, N\} \to [0,1]$ *be a function such that*

$$\varepsilon(N) = 1 \tag{12}$$

and

$$\sum_{i=0}^{N-1} \binom{N}{i} (1-\varepsilon(i))^{N-i} = \beta. \tag{13}$$

It holds that

$$\Pr^N\left\{\mathbb{V}\left(\theta_N^*\right) > \varepsilon\left(m_N^*\right)\right\} \leq \beta, \tag{14}$$

where m_N^* *is the number of irreducible subsamples of* $(\delta^{(1)}, \ldots, \delta^{(N)})$.

3.3 Interval reservoir computing

In this study, compared to obtain W^{out} by solving Eq. 8, we intend to find an interval of W^{out} for every given u_t, $x_{\text{act},t}$ that can finally give an α-confident interval of y_t, as defined in Definition 1. In other words, the output will locate in an interval with a probability larger than the given level $\alpha \in (0,1)$. In addition, the interval is expected to be optimal with the smallest area. Here, a sub-optimal interval is targeted as the approximation of I_t^*. For RNN, the interval is written as

$$I_{\text{RNN}} \coloneqq \left\{y = W^{\text{Out}} x_{\text{act},t} + e, W^{\text{out}} \in \mathcal{W} \subseteq \mathbb{R}^{d \times n_{\text{act}}}, |e| \leq \gamma \in \mathbb{R}^+\right\}. \tag{15}$$

Note that the set I_{RNN} is obtained by varying the values of W^{out}, e in $\mathcal{W}$, and $\mathbb{R}^+$. A possible choice for the set Ω is a ball with center c and radius $r > 0$:

$$\Omega = \mathcal{B}_{c,r} = \left\{W^{\text{out}} \in \mathcal{W} : \|\omega - c\|_2 \leq r\right\}. \tag{16}$$

The interval output of the RNN obtained via Eq. 15 is explicitly written as

$$I_{\text{RNN},\mathcal{B}_{c,r}}(x_{\text{act},t}, \gamma) = \left[c x_{\text{act},t} - \left(r\|x_{\text{act},t}\| + \gamma\right), c_{\text{act},t} + \left(r\|x_{\text{act},t}\| + \gamma\right)\right]. \tag{17}$$

Then, the problem of obtaining a spherical INN is written as

$$\begin{aligned} &\min_{c,r,\gamma} \; \eta r + \gamma \\ &s.t. \quad r, \gamma > 0, \\ &\qquad \Pr\left\{y_t \in I_{\text{RNN},\mathcal{B}_{c,r}}(x_{\text{act},t}, \gamma), \forall t\right\} \geq 1-\alpha, \qquad (\mathcal{P}_{\text{B},\alpha}) \\ &\qquad y_t \in \mathbb{R}^d, \end{aligned}$$

where η is a positive number. Let θ_{B} be the decision variable of Problem $\mathcal{P}_{\text{B},\alpha}$ including c, r, and γ. Let $\Theta_{\text{B},\alpha}$ be the feasible region of θ_{B} of Problem $\mathcal{P}_{\text{B},\alpha}$. Defining the optimal objective function of Problem $\mathcal{P}_{\text{B},\alpha}$ by

$$J_{\text{B},\alpha}^* \coloneqq \min_{\theta_{\text{B}} \in \Theta_{\text{B},\alpha}} \eta r + \gamma. \tag{18}$$

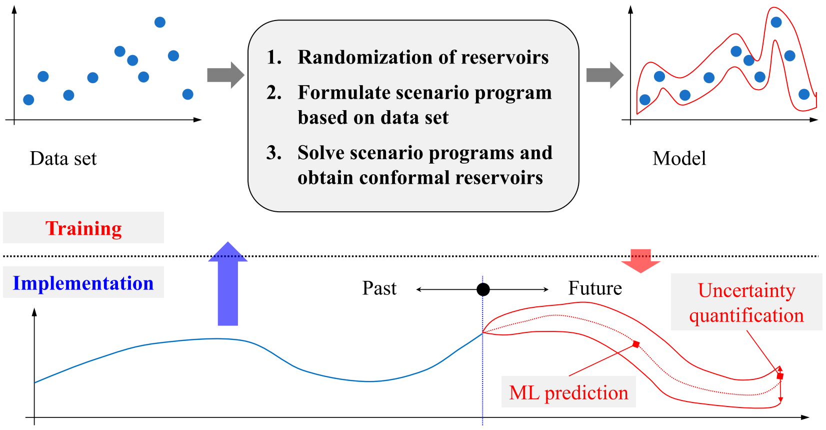


FIGURE 3
Proposed framework of the interval reservoir computing method.

Defining the optimal solution set of Problem $\mathcal{P}_{\mathrm{B},\alpha}$ by

$$\mathcal{A}^*_{\mathrm{B},\alpha} := \left\{\theta_{\mathrm{B}} \in \Theta_{\mathrm{B},\alpha} : \eta r + \gamma = J^*_{\mathrm{B},\alpha}\cdot\right\}. \tag{19}$$

Suppose that the dataset $\mathcal{D}_T = \{u(t), y(t)\}_{t=1,\dots,T}$ is available. Then, we can formulate the scenario program of Problem $\mathcal{P}_{\mathrm{B},\alpha}$ as follows:

$$\begin{aligned}&\min_{c,r,\gamma}\ \eta r+\gamma\\ &s.t.\quad r,\gamma>0,\\ &\qquad y_t\in I_{\mathrm{RNN},\mathcal{B}_{c,r}}(x_{\mathrm{act},t},\gamma),\quad \forall t=1,\dots,T. \qquad \left(\tilde{\mathcal{P}}^T_{\mathrm{B},\alpha}\right)\end{aligned}$$

Let $\tilde{\Theta}^T_{\mathrm{B},\alpha}$ be the feasible region of θ_{B} of Problem $\tilde{\mathcal{P}}^T_{\mathrm{B},\alpha}$. Defining the optimal objective function of Problem $\tilde{\mathcal{P}}^T_{\mathrm{B},\alpha}$ by

$$\tilde{J}^T_{\mathrm{B},\alpha} := \min_{\theta_{\mathrm{B}} \in \tilde{\Theta}^T_{\mathrm{B},\alpha}} \eta r + \gamma. \tag{20}$$

Defining the optimal solution set of Problem $\tilde{\mathcal{P}}^T_{\mathrm{B},\alpha}$ by

$$\tilde{\mathcal{A}}^T_{\mathrm{B},\alpha} := \left\{\theta_{\mathrm{B}} \in \tilde{\Theta}^T_{\mathrm{B},\alpha} : \eta r + \gamma = \tilde{J}^T_{\mathrm{B},\alpha}\cdot\right\}. \tag{21}$$

By adapting Lemma 1, we have the following theorem on the probabilistic feasibility of $\tilde{\theta}^T_{\mathrm{B},\alpha} \in \tilde{\mathcal{A}}^T_{\mathrm{B},\alpha}$.

Theorem 1: *Let $\tilde{\theta}^T_{\mathrm{B},\alpha} \in \tilde{\mathcal{A}}^T_{\mathrm{B},\alpha}$ be the solution of $\tilde{\mathcal{P}}^T_{\mathrm{B},\alpha}$. The interval at t associated with $\tilde{\theta}^T_{\mathrm{B},\alpha}$ is denoted by $\tilde{I}^\alpha_{\mathrm{RNN},\mathcal{B}_{c,r}}(x_{\mathrm{act},t},\gamma)$. Then, the following holds:*

$$\mathrm{Pr}^T\left\{\mathrm{Pr}\left\{y_t \notin \tilde{I}^\alpha_{\mathrm{RNN},\mathcal{B}_{c,r}}(x_{\mathrm{act},t},\gamma)\right\} > \varepsilon\left(m^*_T\right)\right\} \le \beta,\ \forall t, \tag{22}$$

where m^*_T is the number of irreducible subsamples of $((u_1,y_1),\dots,(u_T,y_T))$ and ε satisfies

$$\varepsilon(T) = 1 \tag{23}$$

and

$$\sum_{i=0}^{T-1}\binom{T}{i}(1-\varepsilon(i))^{T-i} = \beta. \tag{24}$$

```
Inputs: data set $\mathcal{D}_T = \{u(t), y(t)\}_{t=1,\dots,T}$
  design of reservoir vector and function according to Eqs 6, 7
  randomly generate $W^{\mathrm{in}}$, $W$, and $W^{\mathrm{back}}$
  solve Problem $\tilde{\mathcal{P}}^T_{\mathrm{B},\alpha}$ and obtain $\tilde{\theta}^T_{\mathrm{B},\alpha}$
Output: $\tilde{\theta}^T_{\mathrm{B},\alpha}$
```

Algorithm 1. Algorithm for interval reservoir computing.

Proof. Since $\tilde{\theta}^T_{\mathrm{B},\alpha}$ is a feasible solution of $\tilde{\mathcal{P}}^T_{\mathrm{B},\alpha}$, by Lemma 1, we have

$$\mathrm{Pr}^T\left\{\mathbb{V}\left(\tilde{\theta}^T_{\mathrm{B},\alpha}\right) > \varepsilon\left(m^*_T\right)\right\} \le \beta, \tag{25}$$

where $\mathbb{V}(\tilde{\theta}^T_{\mathrm{B},\alpha}) = \mathrm{Pr}\left\{y_t \notin \tilde{I}^\alpha_{\mathrm{RNN},\mathcal{B}_{c,r}}(x_{\mathrm{act},t},\gamma)\right\}$. Thus, Eq. 22 holds.

By Theorem 1, we know that it can adjust the sample number T to regulate the violation probability. Using the scenario approach directly, it cannot regulate the violation probability to the desired one. We leave this issue for future work. Based on the theoretical analysis, the algorithm for interval reservoir computing is designed, and the pseudo-code is written in Algorithm 1.

Figure 3 illustrates the proposed framework for implementing the interval reservoir computing method. It follows the general framework widely used to validate the time series model (Shen et al., 2020b). The online obtained history data range from the blue line (not the whole line). Then, the data are used to give the future maximum likelihood prediction (the red dotted line) and the confidence region (the red line) by the model trained by the training dataset.

4 Validations

4.1 Wind power prediction

Let $x_s(t)$ be the wind speed at time index t and $y_p(t)$ be the wind power at time index t. The mechanism behind the evolution of wind

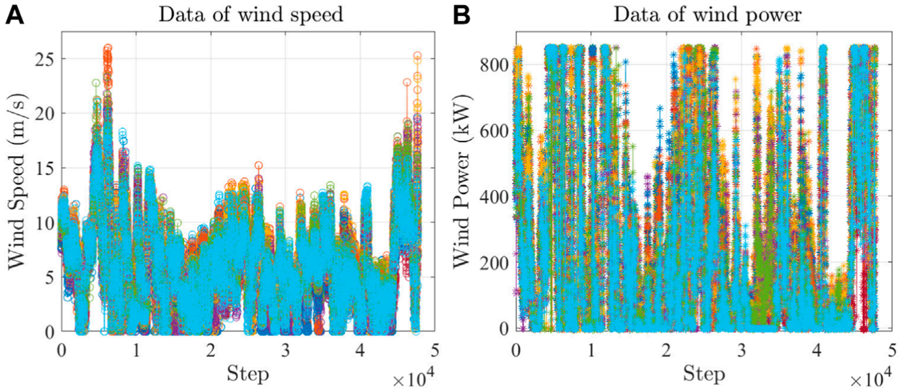


FIGURE 4
Experimental dataset in this study.

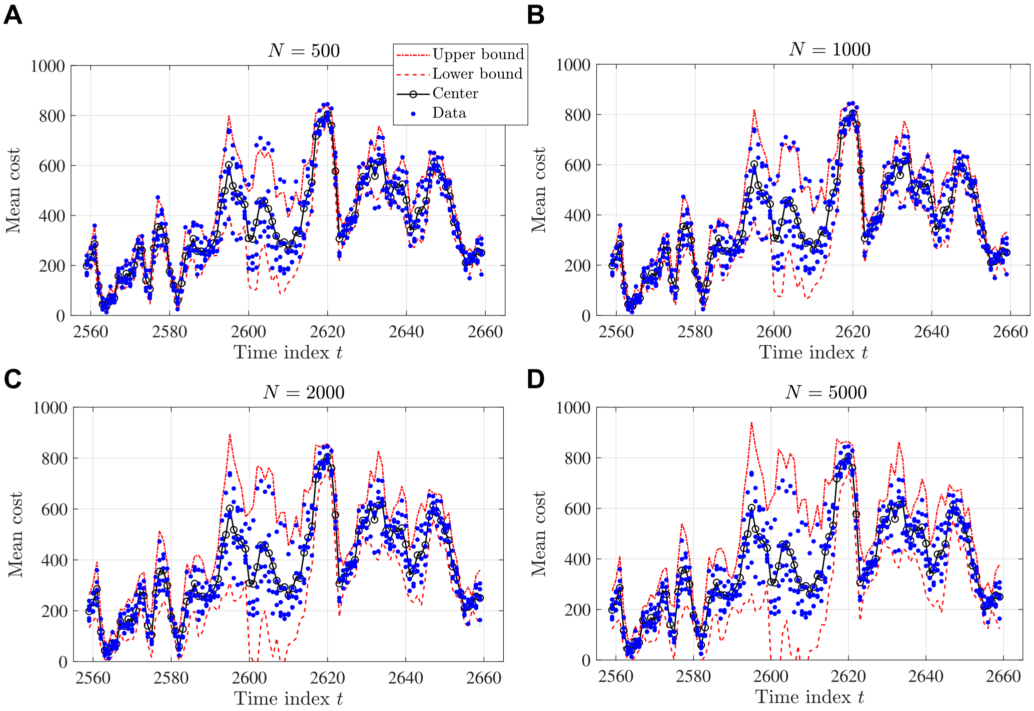


FIGURE 5
Results of a one-step-ahead prediction by the proposed interval reservoir computing (from $t = 2560$ to $t = 2660$): **(A)** $N = 500$, **(B)** $N = 1000$, **(C)** $N = 2000$, and **(D)** $N = 5000$.

speed and wind power can be described by

$$x_s(t+1) = f_{\text{wind}}(x_s(t), w_s(t)), \tag{26}$$

$$y_p(t) = g_{\text{wind}}(x_s(t), v_s(t)). \tag{27}$$

Here, both f_{wind} and g_{wind} are unknown.

The experimental dataset shown in Figure 4 is used in this validation. Figure 4A plots the time series data on wind speed, and Figure 4B plots the time series data on the wind power at the same time. There are a total of 13 groups of data. Eight groups are used as training datasets; the other groups are used as test datasets.

In this validation, we set the threshold for violation probability as $\alpha = 0.05$. The number of samples, N, is from $\{100, 500, 1000, 2000, 5000, 10,000\}$. Figures 5, 6 show two examples of the one-step-ahead prediction by the proposed interval reservoir computing. The parts (a), (b), (c), and (d) of each figure provide the results with $N = 500$, $N = 1000$, $N = 2000$, and $N = 5000$, respectively. As the sample number N increases, the size of the interval also increases, while the center of the interval does not change significantly. In particular, as N surpluses 2,000, the probability of having the data inside the interval is less than the required value $\alpha = 0.05$, implying that the proposed method gives a more conservative interval than we expect.

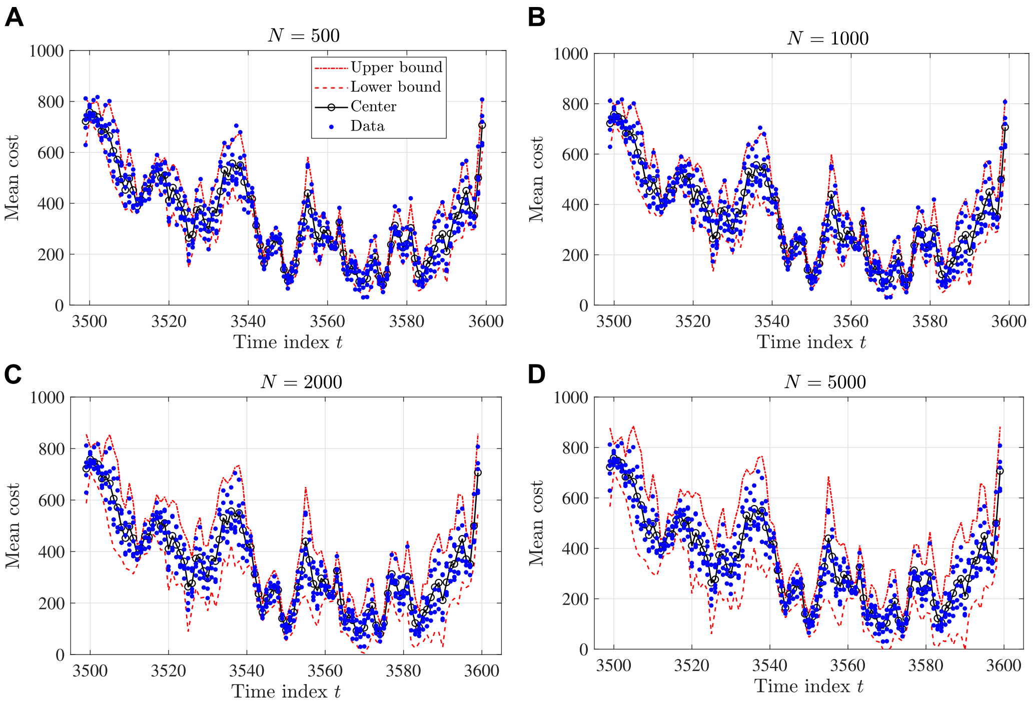


FIGURE 6
Results of a one-step-ahead prediction by the proposed interval reservoir computing (from $t = 2560$ to $t = 2660$): **(A)** $N = 500$, **(B)** $N = 1000$, **(C)** $N = 2000$, and **(D)** $N = 5000$.

TABLE 1 Statistical performance of the proposed method with different sample numbers *N*. Mean values of 5,000 Monte Carlo trials are presented.

Method/*N*	100	500	1,000	2,000	5,000	10,000
$\Pr^N\{\mathbb{V}(\theta_N^*) > 0.05\}$	0.293	0.138	0.002	0.000	0.000	0.000
CPU time (s)	0.076	0.154	0.197	0.231	0.277	0.359

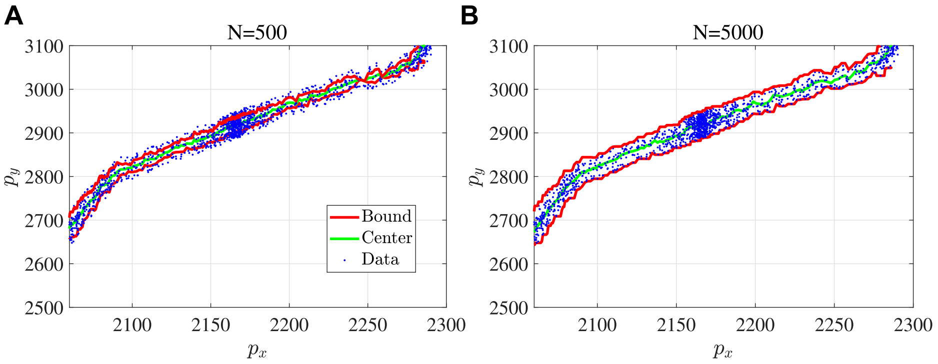


FIGURE 7
One example of a one-step-ahead prediction by the proposed interval reservoir computing for vehicle trajectory prediction: **(A)** $N = 500$ and **(B)** $N = 5000$.

A statistical analysis has been conducted to check the performance of the proposed interval reservoir computing. Monte Carlo tests have been repeated 5,000 times for each choice of sample number $N = 100, 500, 1000, 2000, 5000, 10,000$. We check the violation probability and CPU time in this Monte Carlo simulation. The metric for checking the performance of the violation probability is $\Pr^N\{\mathbb{V}(\theta_N^*) > 0.05\}$, the chance that the violation probability is larger than 0.05. As N increases, $\Pr^N\{\mathbb{V}(\theta_N^*) > 0.05\}$ decreases to zero quickly, as shown in Table 1. On the other hand, the computation time increases as N increases while it is still at an acceptable level.

TABLE 2 **Statistical performance of the proposed method with different sample numbers *N*. Mean values of 5,000 Monte Carlo trials are presented.**

Method/N	100	500	1,000	2,000	5,000	10,000
$\Pr^N\{\mathbb{V}(\theta_N^*) > 0.05\}$	0.127	0.005	0.000	0.000	0.000	0.000
CPU time (s)	0.102	0.218	0.409	0.531	0.591	0.677

4.2 Vehicle trajectory prediction

In another example, we have applied the proposed interval reservoir computing to vehicle trajectory prediction. The experimental data are the public dataset "US Highway 101 Dataset." As in the example of wind power prediction, the number of samples is chosen from $\{100, 500, 1000, 2000, 5000, 10,000\}$, and the violation probability threshold is $\alpha = 0.05$. Figure 7 shows one example of a one-step-ahead prediction for the vehicle trajectory prediction with $N = 500$ and $N = 5000$ results.

Statistical analysis has also been conducted in this example of vehicle trajectory prediction. The settings of Monte Carlo tests are the same as the example of wind power prediction. As shown in Table 2, the results are consistent with the results of wind power prediction.

4.3 Discussion

In this validation, we mainly check the performance of the proposed method with different sample numbers. Indeed, it is necessary to compare the proposed method with other uncertainty quantification methods, such as conformal prediction and Bayesian neural networks. We will further research on this as future work.

One drawback of the proposed interval reservoir computing is that it cannot give an exact interval for a given violation probability α. This drawback comes from using a scenario approach to solve the problem $\mathcal{P}_{B,\alpha}$. The scenario approach ensures the approximate solution's feasibility while not considering the convergence of the approximate solution's optimality. Using sample discarding presented in (Campi and Garatti, 2011) seems to be an excellent choice to make a trade-off between optimality and feasibility. However, sample discarding will dramatically increase the computational complexity of solving $\mathcal{P}_{B,\alpha}$. We leave the issue of optimality for future work.

5 Conclusion and future work

This paper proposes an improved version of interval reservoir computing for time series data forecasting, for example, wind power forecasting and vehicle trajectory forecasting. More than giving a maximum likelihood prediction value of wind power or vehicle trajectory, interval reservoir computing provides an interval of the prediction. The future data will be located inside the interval with a probability larger than the required value. To obtain the interval, a chance-constrained optimization has to be solved for obtaining the interval of the parameters in an RNN. We apply a scenario approach to solve the chance-constrained optimization problem. Experimental data-based validations have been conducted to evaluate the proposed interval reservoir computing. Although the results show that the proposed interval reservoir computing can give a tight interval for wind power forecasting and vehicle trajectory forecasting, the following issues remain to be resolved in future work.

- It is necessary to compare the proposed method with other uncertainty quantification methods, such as the conformal prediction and Bayesian neural networks.
- The scenario approach for solving chance-constrained optimization cannot ensure the convergence of the optimality of the approximate solution. Thus, it is necessary to develop a method that ensures the convergence of the optimality to solve the chance-constrained optimization.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material; further inquiries can be directed to the corresponding author.

Author contributions

L-DG: methodology, software, validation, formal analysis, data curation, writing–original draft, and writing–review and editing. Z-HL: conceptualization, methodology, writing–original draft, writing–review and editing, supervision, and funding acquisition. M-YW: conceptualization, methodology, writing–original draft, and writing–review and editing. Q-LF: methodology, writing–original draft, and writing–review and editing. LX: data curation. Z-MZ: data curation. Y-YL: data curation. Y-PZ: supervision and funding acquisition. All authors contributed to the article and approved the submitted version.

Acknowledgments

This work is supported by the Opening Project of Key Laboratory of Technology on Intelligent Transportation Systems, Ministry of Transport, Beijing, China.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ak, R., Vitelli, V., and Zio, E. (2015). An interval-valued neural network approach for uncertainty quantification in short-term wind speed prediction. *IEEE Trans. Neural Netw. Learn. Syst.* 26, 2787–2800. doi:10.1109/tnnls.2015.2396933

Calafiore, G., and Campi, M. C. (2005). Uncertain convex programs: randomized solutions and confidence levels. *Math. Program.* 102, 25–46. doi:10.1007/s10107-003-0499-y

Campi, M., and Garatti, S. (2011). A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality. *J. Optim. Theory Appl.* 148, 257–280. doi:10.1007/s10957-010-9754-6

Campi, M., Garatti, S., and Ramponi, F. (2015). A general scenario theory for nonconvex optimization and decision making. *IEEE Trans. Automatic Control* 63, 4067–4078. doi:10.1109/tac.2018.2808446

Campi, M. C., Calafiore, G., and Garatti, S. (2009). Interval predictor models: identification and reliability. *Automatica* 45, 382–392. doi:10.1016/j.automatica.2008.09.004

Chen, C., Lu, C., Wang, B., Trigoni, N., and Markham, A. (2021). Dynanet: neural kalman dynamical model for motion estimation and prediction. *IEEE Trans. Neural Netw. Learn. Syst.* 32, 5479–5491. doi:10.1109/tnnls.2021.3112460

Chen, Z., Wu, L., and Shahidehpour, M. (2015). Effective load carrying capability evaluation of renewable energy via stochastic long-term hourly based scuc. *IEEE Trans. Sustain. Energy* 6, 188–197. doi:10.1109/tste.2014.2362291

Dai, C., Wu, L., and Wu, H. (2016). A multi-band uncertainty set based robust scuc with spatial and temporal budget constraints. *IEEE Trans. Power Syst.* 31, 4988–5000. doi:10.1109/tpwrs.2016.2525009

Evans, M., Bono, C., and Wang, Y. (2021). Toward net-zero electricity in europe: what are the challenges for the power system? *IEEE Power Energy Mag.* 20, 44–54. doi:10.1109/mpe.2022.3167575

Garattia, S., Campib, M., and Care, A. (2019). On a class of interval predictor models with universal reliability. *Automatica* 110, 108542. doi:10.1016/j.automatica.2019.108542

Ge, X., Qian, J., Fu, Y., Lee, W., and Mi, Y. (2022). Transient stability evaluation criterion of multi-wind farms integrated power system. *IEEE Trans. Power Syst.* 37, 3137–3140. doi:10.1109/tpwrs.2022.3156430

Hu, B., and Wu, L. (2016). Robust scuc with multi-band nodal load uncertainty set. *IEEE Trans. Power Syst.* 31, 2491–2492. doi:10.1109/tpwrs.2015.2449764

Hu, B., Wu, L., and Marwali, M. (2014). On the robust solution to scuc with load and wind uncertainty correlations. *IEEE Trans. Power Syst.* 29, 2952–2964. doi:10.1109/tpwrs.2014.2308637

Ishibuki, H., Tanaka, H., and Okada, H. (1993). An architecture of neural networks with interval weights and its application to fuzzy regression analysis. *Fuzzy Sets Syst.* 57, 27–39. doi:10.1016/0165-0114(93)90118-2

Jaeger, H., and Haas, H. (2004). Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. *Science* 304, 78–80. doi:10.1126/science.1091277

Liu, H., Flores, C. E., Spring, J., Shladover, S. E., and Lu, X. Y. (2022a). Field assessment of intersection performance enhanced by traffic signal optimization and vehicle trajectory planning. *IEEE Trans. Intelligent Transp. Syst.* 23, 11549–11561. doi:10.1109/tits.2021.3105329

Liu, Y., Zhou, B., Wang, X., Li, L., Cheng, S., Chen, Z., et al. (2022b). Dynamic lane-changing trajectory planning for autonomous vehicles based on discrete global trajectory. *IEEE Trans. Intelligent Transp. Syst.* 23, 8513–8527. doi:10.1109/tits.2021.3083541

Lyu, N., Wen, J., Duan, Z., and Wu, C. (2022). Vehicle trajectory prediction and cut-in collision warning model in a connected vehicle environment. *IEEE Trans. Intelligent Transp. Syst.* 23, 966–981. doi:10.1109/tits.2020.3019050

Neal, R. (2012). *Bayesian learning for neural networks*, 118. Springer Science and Business Media.

Shen, X., Ouyang, T., Yang, N., and Zhuang, J. (2023). Sample-based neural approximation approach for probabilistic constrained programs. *IEEE Trans. Neural Netw. Learn. Syst.* 34, 1058–1065. doi:10.1109/tnnls.2021.3102323

Shen, X., Ouyang, T., Zhang, Y., and Zhang, X. (2020a). Computing probabilistic bounds on state trajectories for uncertain systems. *IEEE Trans. Emerg. Top. Comput. Intell. early access* 7, 285–290. doi:10.1109/TETCI.2020.3019040

Shen, X., and Raksincharoensak, R. (2022). Pedestrian-aware statistical risk assessment. *IEEE Trans. Intelligent Transp. Syst.* 23, 7910–7918. doi:10.1109/tits.2021.3074522

Shen, X., Zhang, Y., Sata, K., and Shen, T. (2020b). Gaussian mixture model clustering-based knock threshold learning in automotive engines. *IEEE/ASME Trans. Mechatronics* 25, 2981–2991. doi:10.1109/tmech.2020.3000732

Wang, P., Wang, P., Wang, D., and Xue, B. (2021). A conformal regressor with random forests for tropical cyclone intensity estimation. *IEEE Trans. Geoscience Remote Sens.* 60, 1–14. doi:10.1109/tgrs.2021.3139930

Wen, H., Gu, J., Ma, J., Yuan, L., and Jin, Z. (2021). Probabilistic load forecasting via neural basis expansion model based prediction intervals. *IEEE Trans. Smart Grid* 12, 3648–3660. doi:10.1109/tsg.2021.3066567

Xue, B., Hu, S., Xu, J., Geng, M., Liu, X., and Meng, H. (2022). Bayesian neural network language modeling for speech recognition. *IEEE/ACM Trans. Audio, Speech, Lang. Process.* 30, 2900–2917. doi:10.1109/taslp.2022.3203891

Yu, Y., Shan, D., Benderius, O., Berger, C., and Kang, Y. (2022). Formally robust and safe trajectory planning and tracking for autonomous vehicles. *IEEE Trans. Intelligent Transp. Syst.* 23, 22971–22987. doi:10.1109/tits.2022.3196623

Zhao, Y., Ye, L., Wang, W., Sun, H., Ju, Y., and Tang, Y. (2018). Data-driven correction approach to refine power curve of wind farm under wind curtailment. *IEEE Trans. Sustain. Energy* 9, 95–105. doi:10.1109/tste.2017.2717021

Frontiers in Energy Research

TYPE Original Research
PUBLISHED 24 May 2024
DOI 10.3389/fenrg.2024.1379863

Check for updates

OPEN ACCESS

EDITED BY
Xun Shen,
Osaka University, Japan

REVIEWED BY
Giuseppe Piras,
Sapienza University of Rome, Italy
Alan Wai Hou Lio,
Technical University of Denmark, Denmark

*CORRESPONDENCE
Fanny Kristianti,
fanny.kristianti@epfl.ch

RECEIVED 31 January 2024
ACCEPTED 25 April 2024
PUBLISHED 24 May 2024

CITATION
Kristianti F, Gerber F, Gonzàlez-Herrero S, Dujardin J, Huwald H, Hoch SW and Lehning M (2024), Influence of air flow features on alpine wind energy potential. *Front. Energy Res.* 12:1379863. doi: 10.3389/fenrg.2024.1379863

COPYRIGHT
© 2024 Kristianti, Gerber, Gonzàlez-Herrero, Dujardin, Huwald, Hoch and Lehning. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Influence of air flow features on alpine wind energy potential

Fanny Kristianti[1]*, Franziska Gerber[1,2], Sergi Gonzàlez-Herrero[2], Jérôme Dujardin[1,2], Hendrik Huwald[1,2], Sebastian W. Hoch[3] and Michael Lehning[1,2]

[1]Environmental Engineering Institute, Laboratory of Cryospheric Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL) Valais/Wallis, Sion, Switzerland, [2]WSL Institute for Snow and Avalanche Research SLF, Davos, Switzerland, [3]Department of Atmospheric Sciences, University of Utah, Salt Lake City, UT, United States

Wind energy is one of the potential options to fill the gap in renewable energy production in Switzerland during the winter season when the energy demand exceeds local production capacities. With likely further rising energy consumption in the future, the winter energy deficit may further increase. However, a reliable assessment of wind energy potential in complex terrain remains challenging. To obtain such information, numerical simulations are performed using a combination of the “Consortium for Small-scale Modeling” and “Weather Research and Forecasting” (COSMO-WRF) models initialized and driven by COSMO-1E model, which allows us to simulate the influence of topography at a horizontal resolution of 300 m. Two LiDAR measurement campaigns were conducted in the regions of Lukmanier Pass and Les Diablerets, Switzerland. Observational LiDAR data and measurements from nearby wind sensor networks are used to validate the COSMO-WRF simulations. The simulations show an improved representation of wind speed and direction near the ground compared to COSMO-1E. However, with increasing height and less effect of the terrain, COSMO-WRF tends to overestimate the wind speeds, following the bias that is already present in COSMO-1E. We investigate two characteristic mountain–terrain flow features, namely waves and Foehn. The effect of mountain-induced waves of the flow is investigated through an event that occurred in the area of Diablerets. One-year analysis for the frequency of conditions that are favorable for mountain wave formation is estimated. The Foehn impact on wind was observed in the Lukmanier domain. We attempt quantification of the probability of occurrence using the Foehnix model. The result shows a high probability of Foehn occurrence during the winter and early spring seasons. Our study highlights the importance of incorporating complex terrain-related meteorological events into the wind energy assessment. Furthermore, for an accurate assessment of wind speed in complex terrain, our study suggests the necessity to have a better representation of the topography compared to COSMO-1E.

KEYWORDS

wind energy, mountain wave, leewave, complex terrain, Foehn, lidar

1 Introduction

Switzerland has stated the objective to entirely transition to renewable energy resources as formulated in the Energy Strategy 2050 (Swiss Federal Office of Energy, 2023b). It is therefore crucial to explore renewable energy options, to reduce and replace the use of unsustainable fossil fuels and thus reduce greenhouse gas emissions (Sims, 2004; Olabi and Abdelkareem, 2022). Currently, the largest part of Switzerland's renewable energy production is based on hydropower (Swiss Federal Office of Energy, 2023a). However, the mismatch of over-production during summer and a demand exceeding production during winter constitutes a significant challenge, i.e., seasonal peaks of production and demand are not aligned. During summer, hydro-power production increases due to the seasonal snow melt and precipitation dynamics (Bavay et al., 2009), while in winter the energy demand increases (Dujardin et al., 2017). Dujardin et al. (2021) suggests that adding wind as a renewable energy resource could be part of the solution to alleviate this energy mismatch. A general increase in mean wintertime wind speed in North and Central Europe has been reported, fueling the motivation to explore the wind as potential energy resource (Archer and Jacobson, 2013; Graabak and Korpås, 2016; Clark et al., 2017; Grams et al., 2017).

The Alps cover two-thirds of Switzerland's area (Federal Office of Topography, 2023). These complex terrain characteristics add to the difficulty of assessing the wind energy potential of the country (Alfredsson and Segalini, 2017; Lange et al., 2017; Mann et al., 2017). The characteristic spatial and temporal patterns of the wind vary substantially according to the specific local topography of the complex terrain. This makes it difficult to accurately observe and quantify near-surface flow using standard measurement techniques and instrumentation (Kruyt et al., 2018), in particular due to sparse spatial distribution of instruments. Most of the wind measurements are taken near the surface at 10 m above ground (MeteoSwiss, 2022; WSL SLF, 2022), while most of the operational wind turbines feature a hub height of 100 m. A logarithmic wind profile is often utilized to extrapolate the wind speed measured near the ground to the turbine hub height. However, the vertical wind profile in complex terrain generally does not follow a logarithmic shape (Dar et al., 2019; Elgendi et al., 2023). Despite the additional problems and challenges, it has been found that terrain complexity can also provide benefits to the local wind power potential (Clifton et al., 2014). If this mechanism is well understood, the interplay between wind and complex terrain could be an untapped potential for wind energy resources. To estimate local wind speed in complex terrain at typical turbine hub heights where measurements are unavailable, several simulation techniques, such as computational fluid dynamics (CFD) (Dhunny et al., 2017; Tabas et al., 2019) and numerical weather models (NWP) Kruyt (2019), have been used while available measurements serve for validating the accuracy of the simulations.

Simulating airflow over complex terrain requires the ability of a model to combine the synoptic flow field and regional scale topography (Lehner and Rotach, 2018). However, the high computational demand of CFD and NWP models and limited computer resources make it difficult to simulate with a fine grid and sufficient domain size for a sufficiently long period to represent the synoptic scale processes. On the other hand, running the model in a too-coarse resolution often results in incorrect, unrealistic representation of the local topography, making it inadequate for resolving the complex terrain processes (Toumelin et al., 2023). The accurate representation is essential for reliable wind energy assessment in complex terrain where the spatial and temporal variability of wind speed is high (Pickering et al., 2020; Clifton et al., 2022; Dujardin and Lehning, 2022). Currently, assessments of large-scale wind resources are often based on reanalyses with a typically very coarse horizontal resolution of 50–100 km and a coarse time step of 1–3 h (Archer and Jacobson, 2005; Archer and Jacobson, 2013; Tobin et al., 2015; Grams et al., 2017). In such a framework, significant wind energy potential in complex terrain likely remains undiscovered due to insufficient spatial resolution for capturing local topographic effects.

In this paper, we propose to simulate wind in complex terrain at a spatial resolution of 300 m, i.e., a resolution within the so-called gray zone (Chow et al., 2019) (also referred to as "terra incognita" in the context of turbulence modeling (Wyngaard, 2004)). The gray zone is a range of resolutions for which certain physical processes start to be explicitly resolved (approximately 100 m to 1 km, Kealy et al. (2019)). When modeling turbulence, it is defined when the turbulence length scale is comparable to the filter length scale (Wyngaard, 2004). In the context of complex terrain, the gray zone challenges include the correct representation of topography, turbulence, and convective processes (Chow et al., 2019). For the scale of the Swiss Alps complex terrain, running simulations at gray zone resolution is cheaper in computational resources compared to classical micro-scale simulations. Another benefit of simulations in the gray zone is to have more insight into the interplay between the meso-scale motion and the smaller-scale motion that occurs at the higher resolution scale because it allows for larger domains to be covered compared to microscale simulations. The representation of the interplay of flow at the two different scales is crucial for gaining accurate information on the wind energy potential in complex terrain (Koletsis et al., 2009; Koletsis et al., 2010).

A previous study by Gerber et al. (2018) implemented the Weather Research and Forecasting (WRF) model (Skamarock et al., 2008), initialized by a 2.2 km resolution Consortium for Small-scale Modeling (COSMO) analysis. This model, hereafter called COSMO-WRF, was used to study wind and terrain-controlled distribution of snow in the complex terrain of Dischma Valley near Davos, Switzerland. The results of the simulation were validated and discussed against operational weather radar measurements acquired at the nearby Weissfluh summit provided by the Federal Office of Meteorology and Climatology (Meteoswiss). Kruyt (2019) have also used COSMO-WRF simulations to investigate the wind speed at a 450 m horizontal grid resolution in the Swiss Alps. That resolution resulted in a significant improvement of the representation of wind speed at the hub height of wind turbines and in the prediction of resulting power production, compared to the results of simulations using COSMO-1 alone, which has a spatial resolution of 0.01°. This improvement is attributed to the better terrain representation in the model. Results from these studies motivate us to further explore the utilization of numerical weather models for the study of wind speed in complex terrain areas. We combine WRF and an ensemble of 11 forecasts with a spatial resolution of 1.1 km called COSMO-1E (Federal Office of Meteorology and Climatology MeteoSwiss, 2023a) to study the flow in complex terrain.

Understanding the impact of typical wind features in complex terrain on potential wind energy production is crucial. Examples of complex terrain phenomenon that affect wind power production are mountain waves (Draxl et al., 2021; Xia et al., 2021) and Foehn wind (Pickering et al., 2020). Mountain waves tend to occur when a stably stratified air mass ascends a mountain barrier and triggers buoyancy perturbations when it descends on the lee side of the barrier. The wave oscillation on the lee side can result in a disturbance of the airflow, which can be propagated downward to a level of 90 m above ground, as shown in the METCRAX II field experiment (Lehner et al., 2016). It has also been reported that mountain wave fluctuations can change the total power output of a wind farm in the region of the Columbia River (United States) up to 11% (Draxl et al., 2021). Foehn wind, on the other hand, is a strong, warm, and dry down-slope wind (Chow et al., 2013). A statistical mixture model named Foehnix (Plavcan et al., 2014) can be used to separate the Foehn and non-Foehn events. This model was tested in the Wipp Valley, Austria, using wind data of a station situated on the crest (Sattelberg station, 11.47889°*E*/47.01083°*N*, 2107 m a.s.l.) and a station in the valley (Ellbögen station, 11.42889°*E*/47.18694°*N*, 1,080 m a.s.l.). The complex terrain phenomenon mentioned above have typically not been included in the wind energy assessment process. Understanding these events can lead to a better selection for wind turbine infrastructure and to a more accurate forecast of energy production. Therefore in this paper, we study the effect of mountain waves and Foehn on the wind in the Alpine area, as two examples of complex terrain effects.

Data and results from a measurement campaign in the Swiss Alps using a Light Detection and Ranging (LiDAR) instrument, combined with simulation results from COSMO-WRF are used to investigate the impact of complex terrain phenomenon on wind speed. Wind LiDAR instruments were used in previous experiments to measure mountain waves (Lehner et al., 2016; Udina et al., 2020) and Foehn events (Beffrey et al., 2006). A filtering technique to derive wind speed in complex terrain using LiDAR measurements is provided in Kristianti et al. (2023). This technique was tested for data obtained at the Diablerets and Lukmanier areas, both Swiss Alps (Figure 1), during the winter season of 2020/21.

This paper aims to propose a method for studying the spatial variability of wind speed in complex terrain and quantifying the effect of complex terrain phenomenon on wind speed. We propose the use of COSMO-WRF simulations in the so-called gray zone of spatial resolution to obtain information on wind interplay between the synoptic and local scales. Two complex terrain phenomenon, namely waves and Foehn, are analyzed to study the impact on wind speed in Swiss Alps area. The focus of the study is on the wind assessment process, for which many aspects also need to be considered (i.e. potential of social impact, etc), however, we limit the scope of the study to wind potential only. Section 2 describes the methods used for the assessment process. Details about the wind measurement network and the LiDAR measurement campaign are described in Sections 2.1.1, 2.1.2, respectively. We also provide a short description of the models used in this study, namely COSMO-1E (Section 2.2.1), COSMO-WRF (Section 2.2.2), and Foehnix (Section 2.2.3). The validation of wind simulations in the gray zone resolution is discussed in Section 3.1. The analysis of the complex terrain wind features is presented in Sections 3.2, 3.3 for mountain waves and Foehn, respectively. The main findings and conclusion are summarized in Section 4.

2 Data and methods

2.1 Measurements

2.1.1 Wind measurement network

Data from two wind measurement networks are used to validate the COSMO-WRF simulation: (a) the Inter-Cantonal Measurement and Information System, IMIS (WSL SLF, 2022), and (b) the Meteo Swiss SwissMetNet, SMN (MeteoSwiss, 2022). By 2021, the IMIS network counts 186 measuring stations which are scattered over the Swiss Alps. The IMIS stations are situated at high-elevation locations to provide data for operational avalanche forecasts and warnings. IMIS wind speed data are measured by R.M.Young wind sensors (model 05103) at approximately 7.5 m a.g.l. (Lehning et al., 2000). SMN stations are distributed at middle and low altitudes of Switzerland. SMN measures wind speed using Lambrecht L14512 cup anemometers and Thies 2D ultrasonic anemometers at 10 m a.g.l. (Federal Office of Meteorology and Climatology MeteoSwiss, 2023b).

The stations used for validation of the COSMO-WRF simulation and as input for the Foehnix model are shown in Figure 1 (star symbols) and station details are given in Table 1. Stations Nara/Motto Crostel (SLFNAR) and Tujetsch/Crispalt (SLFTUJ) are used for COSMO-WRF validation of the Lukmanier domain and stations Evionnaz (EVI) and Färmel/Färmelberg (SLFFA2) are used for COSMO-WRF validation of the Diablerets domain. SLFNAR is located on the peak of Motto Crostel, in the middle of Valle Leventina and Valle di Blenio, Canton Ticino, Switzerland. SLFTUJ is located on the Crispalt ridgeline, on the northwest side of Oberalpass, Canton Glarus, Switzerland. EVI is located in Evionnaz city on the west side of the Rhone Valley in Canton Valais and SLFANV is located in Färmelberg, Canton Bern. The stations Altdorf (ALT), Biasca (BIA), and Gütsch (GUE) are used as input data for the Foehnix model. Both ALT and BIA stations are used as valley stations for the Foehnix model. ALT is used to represent the southerly Foehn. BIA is used to represent the northerly Foehn. GUE is used as crest station input of the Foehnix model for both northerly and southerly Foehn.

2.1.2 LiDAR measurements

Two measurement campaigns were conducted deploying a Halo Photonics Streamline XR Scanning Doppler wind LiDAR. The first field campaign was conducted on the west slope of Piz Scopi, Lukmanier, and the second campaign was conducted at Cabane station, Diablerets (dots in Figure 1). The Lukmanier and Diablerets campaigns were conducted from 20/10/2020 to 16/12/2020, and from 20/02/2021 to 02/05/2021, respectively. The coordinates of the LiDAR location were 46.58409°*N*, 8.81890°*E* (blue dot, Figure 1) at 2519 m a.s.l. at Lukmanier, and 46.33995°*N*, 7.21491°*E* (red dot, Figure 1) at 2523 m a.s.l. at Diablerets. The LiDAR configuration was the same for both campaigns except for the elevation angle of 45° at Lukmanier and 70° at Diablerets. The gate overlapping mode was used to collect LiDAR data, with a range gate length of 30 m. This resulted in radial velocity retrieval from 30 m

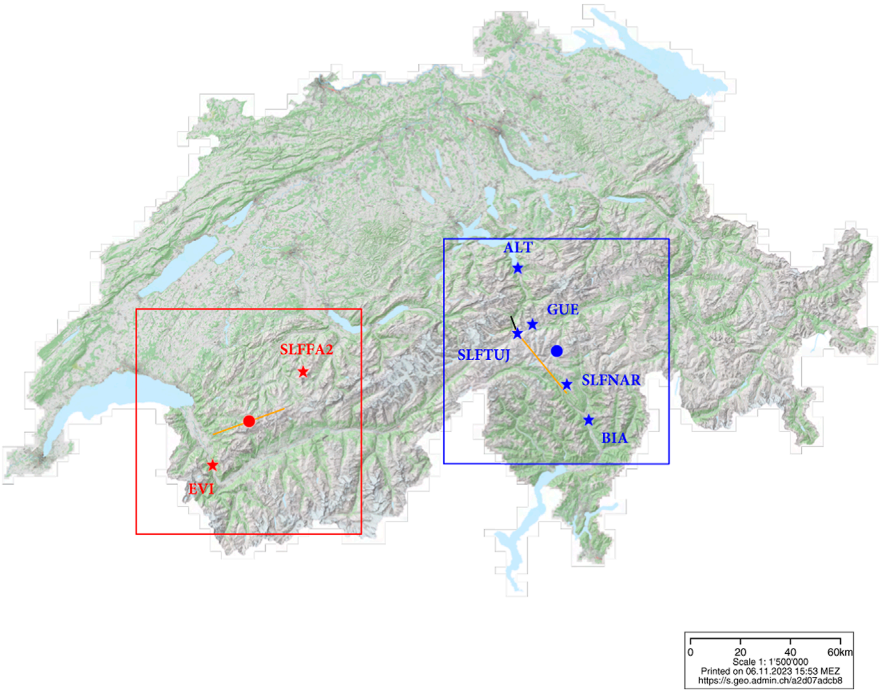


FIGURE 1
Simulation domains of Diablerets (red box) and Lukmanier (blue box). The LiDAR locations in Diablerets and Lukmanier are shown by red and blue dots, respectively. Star symbols represent the SMN and IMIS wind measurement stations. The orange line inside the Diablerets domain represents the vertical cross-section shown in Figure 10. The black and orange line inside the Lukmanier domain represents the vertical cross-section shown in Figure 14. Map source: (Federal Office of Topography, 2023).

TABLE 1 Details of the SMN and IMIS wind measurement stations.

Station	Region	Lat(N)/Lon(E)	Station elevation (m a.s.l.)	COSMO-WRF elevation (m a.s.l.)	COSMO-1E elevation (m a.s.l.)	Network
ALT	Altdorf	46°53′13”/8°37′19″	437	444.83	452.75	SMN
BIA	Biasca	46°20′10”/8°58′41″	278	429.40	620.88	SMN
EVI	Evionnaz	46°10′43”/7°06′47″	482	493.01	589.72	SMN
GUE	Gütsch, Andermatt	46°39′09”/8°36′56″	2286	1999.96	2018.88	SMN
SLFFA2	Färmel/Färmelberg	46°10′59”/7°01′36″	1970	1899.23	1994.16	IMIS
SLFNAR	Nara/Motto Crostel	46°27′55”/8°52′01″	2302	2022.74	1866.16	IMIS
SLFTUJ	Tujetsch/Crispalt	46°41′01”/8°41′40″	3028	2791.66	2429.81	IMIS

gates with a 3-m spacing. Under ideal conditions, the use of this setting enables us to observe wind velocity up to a radial distance of 2.1 km. We used 6-point and 12-point step-stare Plan Position Indicator (PPI) scans at an elevation angle of 45° and 70° for Lukmanier and Diablerets sites, respectively. The scan sequence was repeated at 5- and 10-min intervals for the Lukmanier and Diablerets sites, respectively. Post-processing from the radial velocity to the u, v, w wind speed components followed the procedure described in Kristianti et al. (2023).

2.2 Models

2.2.1 COSMO-1E

COSMO-1E is a numerical weather forecasting model run over Switzerland at a horizontal resolution of 1.1 km (Federal Office of Meteorology and Climatology MeteoSwiss, 2023a). COSMO-1E includes an ensemble of 11 forecasts computed eight times per day. From a single forecast, several iterations are produced to predict the probability of weather events. Therefore, the

TABLE 2 Simulation details.

Simulation domain	LiDAR coordinates	Simulated days	Main wind direction	100 m a.g.l. Mean wind speed at the LiDAR location (m/s)		
				LiDAR	COSMO-1E	COSMO-WRF
Lukmanier	46°35′03″N/8°49′08″E	23/10/2020	Southerly	6.13	12.78	13.75
		09/12/2020	Northerly	4.79	10.01	8.26
Diablerets	46°20′24″N/7°12′54″E	27/02/2021	Easterly	5.71	9.50	9.74
		11/03/2021	Southwesterly	14.22	15.55	19.09

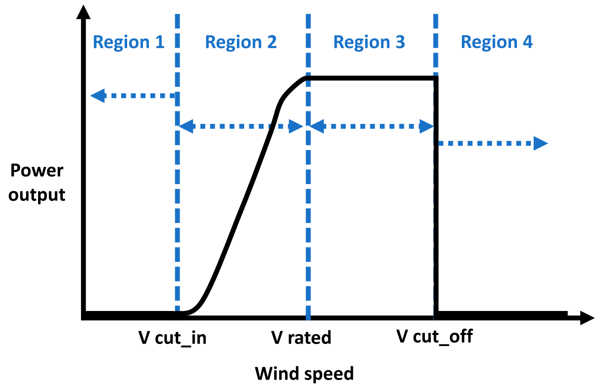


FIGURE 2
Schematic diagram of typical power curve of wind turbine.

reliability of the forecast is improved and so is the quality of short to medium-range forecasts for extreme or highly localized weather, compared to the deterministic forecast (Schraff et al., 2016).

2.2.2 COSMO-WRF

Numerical modeling is used to investigate the spatial variations of wind speed in complex terrain and its effect on the wind at typical turbine hub heights of 100 m above ground level. We use the Weather Research and Forecasting (WRF) model, Version 4.4.5 (Skamarock et al., 2021) initialized and forced with COSMO-1E (Federal Office of Meteorology and Climatology MeteoSwiss, 2023a) data provided by Meteoswiss (COSMO-WRF, hereafter). COSMO-WRF is used to simulate representative cases of flow events over the complex terrain at the Diablerets and Lukmanier sites based on observational LiDAR data. The topography input is based on the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) global digital elevation model (DEM) v003 (NASA/METI/AIST/Japan Space systems and U.S./Japan ASTER Science Team, 2019) and the land use is taken from the Coordination of Information on the Environment (CORINE) dataset (European Environmental Agency, 2006) as provided by Gerber and Lehning (2021). More technical details about COSMO-WRF can be found in (Gerber and Sharma, 2018).

Owing to the steep slopes in the selected complex terrain domains, which may lead to numerical instabilities in the simulations, a pre-processing of the topographic data is needed. Pre-processing of WRF is performed by using the WRF Pre-processing System (WPS) Version 4.4 (Skamarock et al., 2021). We use three cycles of the 1-2-1 smoothing algorithm to reduce steep slopes over 45° (Gerber and Sharma, 2018). After the smoothing process, the maximum slope angle in the simulation domain is 46.1° and 43.6° for the Lukmanier and Diablerets sites, respectively. The simulation domains are set to 90 × 90 km, with the LiDAR position located at the center (Figure 1). A single domain with no nesting is used for the simulations, following the gray zone recommendation of Chow et al. (2019). The horizontal grid resolution is 300 m resulting in a domain composed of 301 × 301 grid points.

The model applies *eta*-level coordinates with 60 vertical levels. The simulation is run with a time step of 0.5 s. The barometric pressure at the top of the domain is set to 15′000 Pa. The planetary boundary layer uses the Shin-Hong Scale scheme (Shin and Hong, 2015). The Morrison 2-moment scheme is selected (Morrison et al., 2009) to parameterize the cloud microphysics. Longwave and shortwave radiation use the rrtmg parameterization (Mlawer et al., 1997). For the surface layer a Monin-Obukhov Similarity scheme is implemented (Dyer and Hicks, 1970; Paulson, 1970; Webb, 1970; Zhang and Anthes, 1982; Beljaars, 1995). Land surface processes are parameterized by the Noah-MP scheme (Niu et al., 2011; Yang et al., 2011). No cumulus option is used when running WRF. The w-Rayleigh damping option (Klemp et al., 2008) is activated in WRF. The namelist used to prescribe the simulation can be found on (Kristianti et al., 2024). At the grid cells where the wind measurement stations and LiDAR are located, model output is saved at every time step using the tslist options of the WRF model.

Two representative flow situations are simulated for each field site, resulting in four simulation cases (two cases for Lukmanier and two cases for Diablerets). For the Lukmanier site, observations of 23/10/2020 and 09/12/2020 are used to represent the southerly and northerly flow regimes, the two principal wind directions during the campaign duration. For Diablerets, observations of 27/02/2021 and 11/03/2021 are used to represent the easterly and southwesterly flow regimes, the two principal wind directions during the campaign. A more detailed wind direction analysis can be found in Kristianti et al. (2023). Details of the simulations are summarized in Table 2.

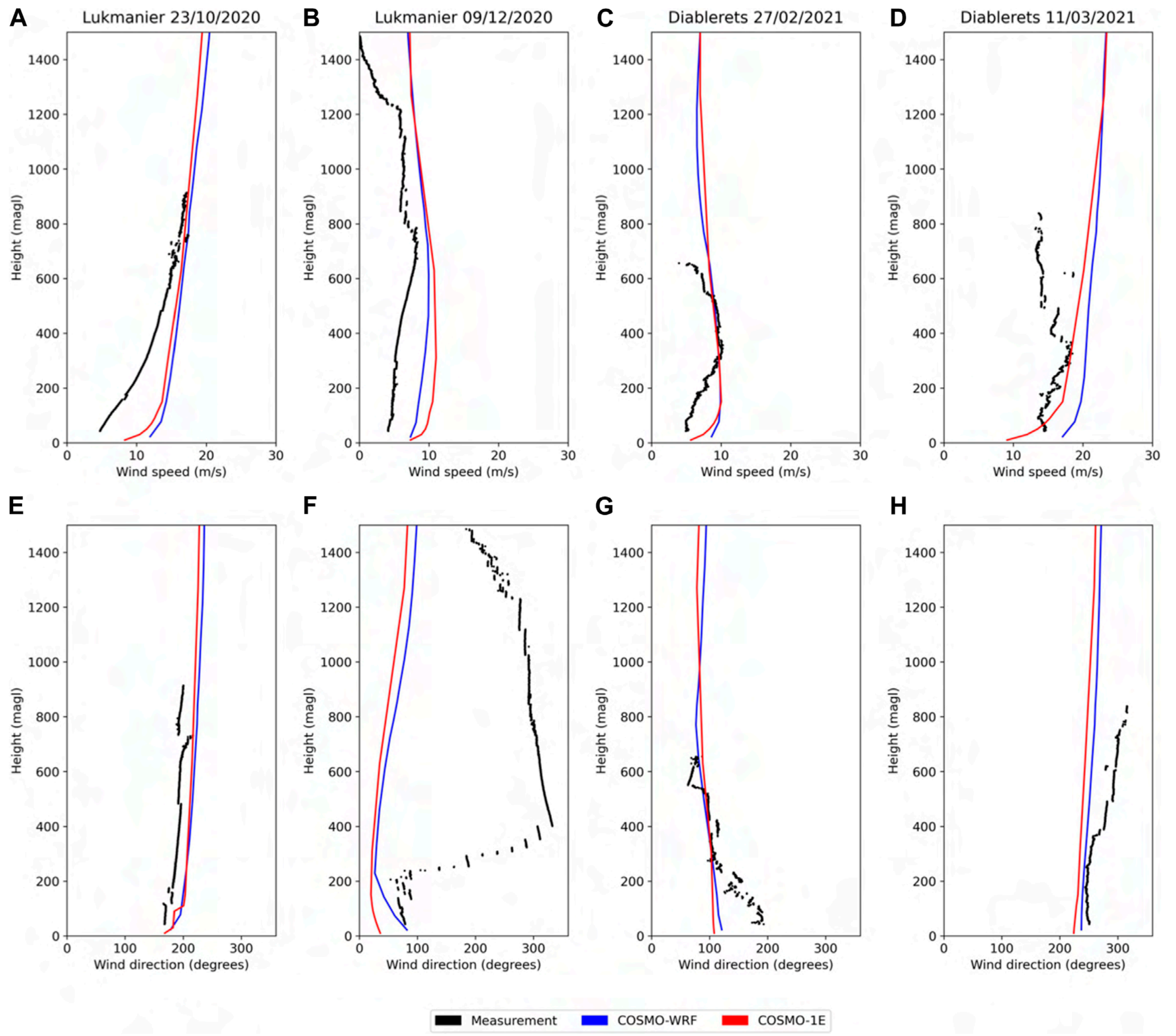


FIGURE 3
Daily average wind speed (top) and wind direction (bottom) profile at the field sites measured by the LiDAR (black), and simulated by COSMO-WRF (blue), and COSMO-1E (red) during 23/10/2020 **(A,E)** and 09/12/2020 **(B,F)** in the Lukmanier domain, and 27/02/2021 **(C,G)** and 11/03/2021 **(D,H)** in the Diablerets domain.

2.2.3 Foehnix

Foehn events are investigated and validated using a statistical mixture model named Foehnix (Plavcan et al., 2014). The model can distinguish between Foehn and no Foehn wind using wind speed, wind direction, relative humidity, and temperature differences as indicators. First, a wind direction filter is applied. Then, the temperature difference between the two stations is selected as the dominant variable, while wind speed and relative humidity are used as concomitant variables. The mixture model uses the wind speed distribution and divides it into downslope wind and Foehn. The Foehn phenomenon has a strong seasonal cycle, therefore to capture this cycle, a minimum data set comprising at least 1 year is required as model input.

2.3 Power curve of wind turbine

The wind turbine power curve can be used as a tool to estimate the power extractions from the incoming wind speed. A typical wind turbine power curve consists of four regions of wind speed (Figure 2). The first region represents the area where wind speed is less than the minimum wind speed for power production (v_{cut_in}), therefore it does not produce any power. The second region represents the area between the v_{cut_in} and the rated wind speed (v_{rated}). In this region, the power rises rapidly until the wind speed reaches the v_{rated}. The third region produces a constant power where the wind speed is between the v_{rated} and the maximum operational wind speed (v_{cut_off}). If the wind speed goes higher than the (v_{cut_off}), the wind turbine does not operate to protect its

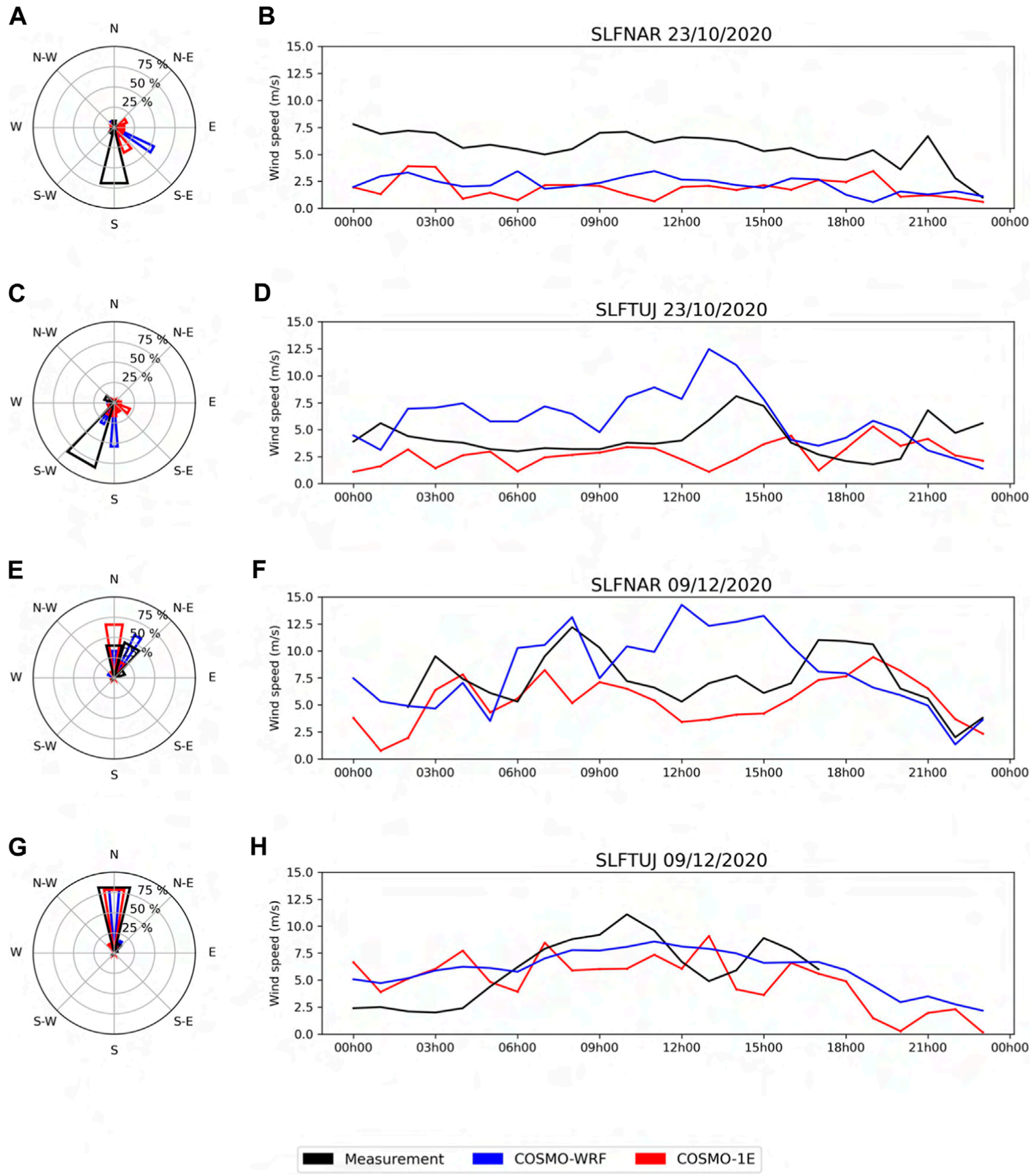


FIGURE 4
Windrose (left) and wind speed time series (right) from the Lukmanier simulation domain at SLFNAR **(A,B,E,F)** and SLFTUJ **(C,D,G,H)** stations on 23/10/2020 **(A,B,C,D)** and 09/12/2020 **(E,F,G,H)**. The black color shows the result of the wind measurement station, the blue color shows the COSMO-WRF result and the red color shows the COSMO-1E. On 09/12/2020 **(H)**, some observational data is missing during the night for the SLFTUJ station.

components from possible damage due to high wind. The wind speed above the (v_{cut_off}) is represented in the fourth region, which produces no power, similar to the first region.

3 Results and discussion

3.1 Validation of wind simulation

This section compares the wind simulation results from COSMO-WRF to the LiDAR and wind station measurements at the two field sites. Three aspects of wind (direction, vertical profiles, and time series of wind speed) are utilized for this purpose. In addition, a comparison with COSMO-1E data is provided. The comparison with wind station data represents the conditions near the ground level and the comparison with LiDAR data represents the conditions at a higher elevation level.

The comparison of the observed daily averaged LiDAR wind speed and wind direction profiles, the COSMO-1E, and the COSMO-WRF simulations for four different periods is presented in Figures 3A-H. The wind speed profiles in the top row show that both COSMO-1E and COSMO-WRF overestimate the wind speed compared to the LiDAR measurements, especially at low elevations (Figures 3A, C). A relatively good agreement of the wind direction profile between measurement, COSMO-1E, and COSMO-WRF can be seen for the profiles with little change of wind direction with height (Figures 3E, H). However, the marked wind direction change with height (veering or backing) (Figures 3F, G), was neither captured in COSMO-1E nor COSMO-WRF. The slightly improved representation of the COSMO-WRF wind direction profile especially near the ground compared to COSMO-1E (Figure 3F), may be explained by the better terrain representation in COSMO-WRF compared to COSMO-1E. However, with increasing height and lower influence of the terrain, the wind direction follows COSMO-1E as this is used as forcing. The wind direction change from 200° (southerly) to 100° (easterly) in the lowest 300 m above ground level (Figure 3G) is also not represented by both models. The wind direction profile from COSMO-1E and COSMO-WRF is rather constant with height (approximately 100° and agrees well with observations higher than 300 m a.g.l.

Overall, COSMO-1E and COSMO-WRF show wind speed overestimation in all wind speed profiles and some disagreements in wind direction, especially for profiles with veering winds. This disagreement illustrates the complexity of the wind in mountain regions and the difficulty of simulating it. The overestimation of wind speed by COSMO-WRF can be explained by the overestimated input data from COSMO-1E. Table 2 presents the 100 m a.g.l. mean wind speed at the LiDAR location in COSMO-1E and COSMO-WRF, representing a wind turbine hub height. Compared to the LiDAR measurements, the simulated wind speed is overestimated by approximately 5 m/s by both models.

Further analysis compares the time series of wind speed and the wind roses from COSMO-WRF and COSMO-1E to wind measurement stations. Figures 4, 5 show wind roses and time series of wind speed at wind measurement stations situated in the Lukmanier and Diablerets domains, respectively. Details of these measurement stations used for validation can be seen in Table 1. For the case study of 23/10/2020 (Figures 4A, C), the wind roses from COSMO-1E and COSMO-WRF at SLFNAR and SLFTUJ show more distributed directions compared to the observations, which are mainly clustered around the southerly (SLFNAR) and southwesterly (SLFTUJ) sectors. The observed difference in wind direction is explained by the COSMO-1E wind direction input data, while COSMO-WRF shows a slight deviation from its initial direction in COSMO-1E. For the case study of 09/12/2020 (Figures 4E, G), on the other hand, the wind rose shows better agreement between the measurements, COSMO-1E, and COSMO-WRF.

Figures 4B, D, F, H show the wind speed time series at the SLFNAR and SLFTUJ station, respectively. The time series of wind speed shows an underestimation by COSMO-1E and COSMO-WRF compared to the measurements, except in the COSMO-WRF case of SLFTUJ on the 23/10/2020 and SLFNAR on the 09/12/2020. This contradicts the comparison between LiDAR and simulation profiles, where COSMO-1E and COSMO-WRF overestimate the wind speed (Figure 3). This might be the result of a lower elevation represented in COSMO-1E and COSMO-WRF, compared to the real elevation of the SLFNAR and SLFTUJ stations (Table 1). During the smoothing process in WPS, the steepness of the slope in COSMO-WRF is reduced. After smoothing, to reach a maximum steepness of approximately 45°, topographic peaks get “shaved” and valleys are “filled.” The elevation difference is more significant for the near ground level wind speed, where the comparison with wind measurement stations is performed. At higher elevations, as we can see from the comparison with LiDAR, COSMO-WRF tends to be closer to COSMO-1E as the terrain influence diminishes. A slight improvement in the near-ground wind speed comparison from COSMO-WRF (right column, Figure 4) can be the result of a smaller elevation difference between COSMO-WRF and the stations compared to COSMO-1E and the stations (Table 1) since with 300 m horizontal grid resolution COSMO-WRF has a better terrain representation than the 1 km of COSMO-1E.

Figure 5 shows the comparison of wind direction and wind speed time series for the Diablerets domain. For 27/02/2021 at the SLFFA2 station (Figure 5A), observed winds are from all directions, while COSMO-WRF, and COSMO-1E show dominant NE-E directions. For the case study of 11/03/2021 at SLFFA2 station (Figure 5E), measurements show a significant difference in wind direction compared to the models, which could be due to the local terrain sheltering of the station still insufficiently resolved. The time series of wind speed (Figures 5B, F), show a slight overestimation by COSMO-1E and COSMO-WRF at the SLFFA2 location which may again be due to an overestimation of the SLFFA2 station elevation in COSMO-1E. This would result in an overestimation of wind speed from COSMO-1E compared to the measurement. On 27/02/2021 at EVI station (Figure 5C), the wind directions from models and measurements are well aligned. On 11/03/2021 at EVI station (Figure 5G), the modeled wind direction (SE) is well aligned with part of the measured wind, however, the other part is opposite to the modeled wind. At the EVI station, the measured and modeled wind speed time series show a good agreement (Figures 5D, H). For wind speed near ground level, COSMO-WRF seems to perform well, if the model terrain elevation in COSMO-WRF is similar to the real terrain elevation. This explanation is consistent with a better agreement of both elevation and near-ground time series of wind speed at the SLFFA2 and EVI stations, compared to the SLFNAR and SLFTUJ stations. At higher elevations of the atmosphere, however, the influence of the input and boundary conditions from COSMO-1 in the COSMO-WRF model becomes stronger and might result in overestimation.

In conclusion, COSMO-WRF shows improved simulation results near the ground compared to COSMO-1E, as a result of the better terrain representation in COSMO-WRF. However, model performance is limited by the input data used (COSMO-1E), which tends to overestimate the wind speed at the height, where the wind turbines are located (cf. Table 2). Therefore, existing biases in the forcing data aloft cannot be completely rectified with improved surface representation. The COSMO-WRF simulations are improving surface representation but remain limited due to the (still) coarse horizontal resolution of 300 m and the maximum allowed slope angle of approximately 45°. These two limitations bear the risk of compromising the terrain’s full influence. As the simplification of the terrain leads to an overestimation or underestimation of wind speed depending on the location, it is important to use multiple sites for a robust and representative comparison between models and measurements. The further use of the COSMO-WRF model in this study is to study and quantify the effect of complex terrain on wind power potential, thus the model needs to capture the event mechanism. This will contribute to the

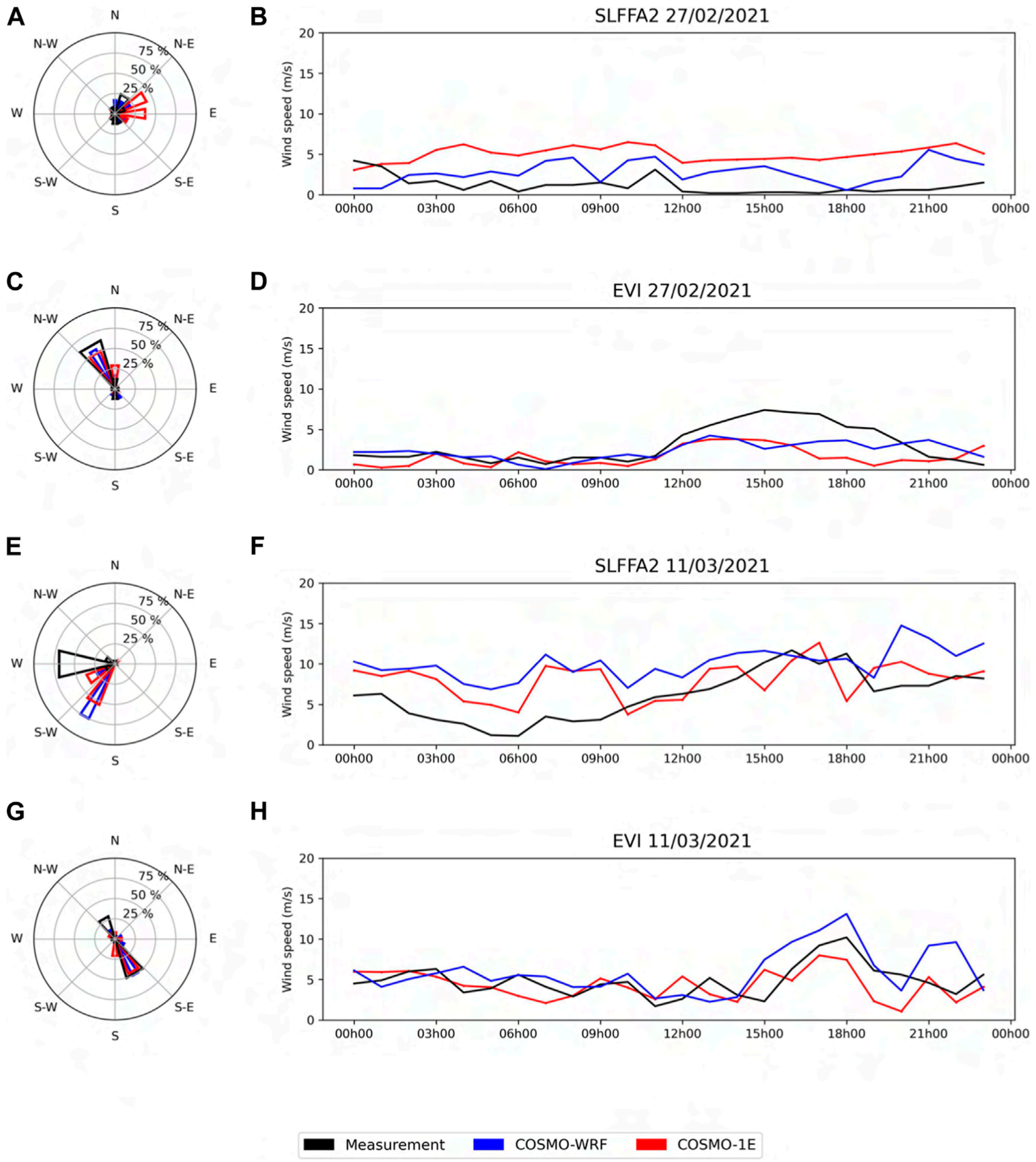


FIGURE 5
Windrose (left) and wind speed time series (right) from the Diablerets simulation domain at EVI **(A,B,E,F)** and SLFFA2 **(C,D,G,H)** stations on 27/02/2021 **(A,B,C,D)** and 11/03/2021 **(E,F,G,H)**. The black color shows the result of the wind measurement station, the blue color shows the COSMO-WRF result and the red color shows the COSMO-1E.

main goal of our study, namely for a better understanding of the possible mechanism of Foehn and mountain waves that influence wind variability in complex terrain.

3.2 Influence of mountain waves on wind speed at turbine hub height

This section aims to investigate how events of mountain waves influence the wind at the hub height of a potential turbine at the Diablerets site and to quantify the conditions favoring the generation of mountain waves. By utilizing satellite images, the observed mountain wave event is validated and analyzed using COSMO-WRF simulations. Afterward, we describe the effect of mountain waves at turbine height levels and underline the importance of including this aspect in wind energy assessment in complex terrain.

For finding a mountain wave, the high wind speed period measured at Diablerets on 11/03/2021 is selected (Figure 6, blue dots). The area within the white box in Figure 7A is selected to check the corrected reflectance from Moderate Resolution Imaging Spectro-Radiometer (MODIS) satellite images on 11 and 12 March 2021 (Figures 7B, C). On 12/03/2023 (Figure 7C), we see cloud bands, perpendicular to the wind over the Alps, as seen from the

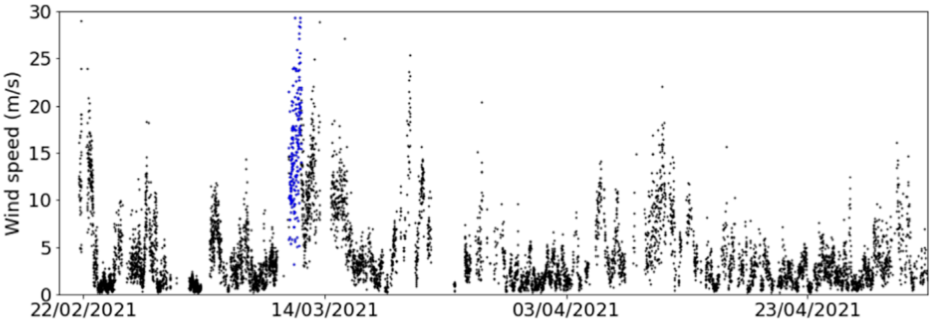


FIGURE 6
LiDAR wind speed measurements at Diablerets from 20/02/2021 to 02/05/2021 (black dots) and the simulation period on 11/03/2021 (blue dots) at 100 m a.g.l.

LiDAR's wind rose (Figure 9A). This cloud pattern is associated with mountain waves that often occur downwind of a mountain when the atmosphere is stably stratified, especially during the winter season. A similar pattern can also be seen in the blue-circled area on 11/03/2021 and 12/03/2021, suggesting the event lasted for 2 days. Due to the limited availability of LiDAR data on 12/03/2021, only data from 11/03/2021 are analyzed. Over Switzerland, COSMO-1E at approximately averaged 2,400 m a.g.l. of terrain following coordinates show a pattern of alternating positive and negative vertical wind velocities (Figure 8). The horizontal wind direction is consistent with the LiDAR's wind rose and the cloud pattern shown in the satellite images. The simulated wave pattern is strongest in the southwestern Alpine region (tallest mountains) and weakest over the (Rhone River) valley. However, the alternating vertical velocity pattern covers most of Switzerland, apart from the Ticino region in the South. This shows the wide impact area of the wave event and the importance of understanding the event to accurately assess its influence.

To study the impact of mountain waves at the turbine height, the event was simulated with COSMO-WRF in the Diablerets domain depicted as a red box in Figure 1. The result of COSMO-WRF is used to analyze the mountain wave event. Following the main wind direction obtained from LiDAR measurements and the COSMO-WRF simulation (Figures 9A, B), which shows a very good agreement between both, a cross-section at 70°–250° is plotted (Figure 1, orange line in Diablerets domain). The vertical cross-section is plotted at 08h00 11/03/2021 for approximately 15 km radius distance from the LiDAR location (black dot, Figures 10A, B). Vertical cross-sections of horizontal (Figure 10A) and vertical wind speed (Figure 10B) from COSMO-WRF are utilized to visualize the mountain wave event. We can see an undulating pattern of potential temperature contours and horizontal wind speed, indicating the presence of the mountain wave (Figure 10A). The pattern of potential temperature (Figure 10A) agrees very well with the alternating upward-downward vertical velocity (Figure 10B). This simulated oscillating pattern shows the model's ability to simulate the mountain wave event. The potential temperature profile at the LiDAR site increases with height, indicating a stable atmosphere, and favoring the formation of mountain waves.

Further, the Hovmoller diagram is used to present the evolution of the horizontal and vertical wind speed cross-section (x-axis) in time (y-axis) (Figures 10C, D). The Hovmoller diagram of vertical wind speed (Figure 10D) shows relatively stationary positive and negative velocity patterns, especially at the east side of the LiDAR site (black vertical line). This quasi-constant vertical velocity pattern indicates stationary mountain waves also visible in Figure 10B. The Hovmoller diagram of the horizontal wind speed shows an increase in wind speed in several areas which is interpreted as another propagating mountain wave from west to east (blue dashed line, Figure 10C). To better understand the timing of the event, another pair of Hovmoller diagrams of horizontal (Figure 10E) and vertical (Figure 10F) wind speed is provided at the location of the LiDAR. The horizontal wind speed increases at 10h00, which coincides with the downdraft wind inferred from the vertical velocity pattern, as evident in Figure 10F. This increase in wind speed might be caused by a downdraft by the mountain wave event, as described in Lehner et al. (2016).

Figure 11 shows the time series of wind speed at the Diablerets LiDAR site measured by the wind LiDAR, and simulated with COSMO-WRF and COSMO-1E. All results are hourly averaged. We notice a prominent oscillation in the LiDAR data at the wind turbine hub height, especially after 10h00. This indicates a correlation with the downdraft from the propagated wave seen from the Hovmoller diagram. The timing and amplitude of the oscillations simulated by COSMO-1E and COSMO-WRF correspond well, suggesting that COSMO-1E is suitable for providing initial data for a mountain wave simulation case. The oscillations observed during 11/03/2021 at the Diablerets LiDAR site were approximately 10 m/s, as can be seen from the COSMO-WRF simulation. The study of (Draxl et al., 2021) (Cascade Range region USA) showed that wind speed oscillations on the order of 5 m/s can already create oscillations in wind turbine power output in their case study. Therefore, we may anticipate oscillations of wind speed if a turbine is placed in a region influenced by mountain waves, such as the Diablerets region.

During the measurement period, the wind rose of the Diablerets site shows two predominant wind directions, i.e., southwesterly and northeasterly [as shown in Kristianti et al. (2023)]. These main wind directions are also found as the annual average using data of the Wind Atlas Switzerland (Koller and Humar, 2016). The LiDAR measurement and the COSMO-WRF simulation during the 11/03/2021 event represent a case study of a situation with a main wind direction from the Southwest (Figures 9A, B), and

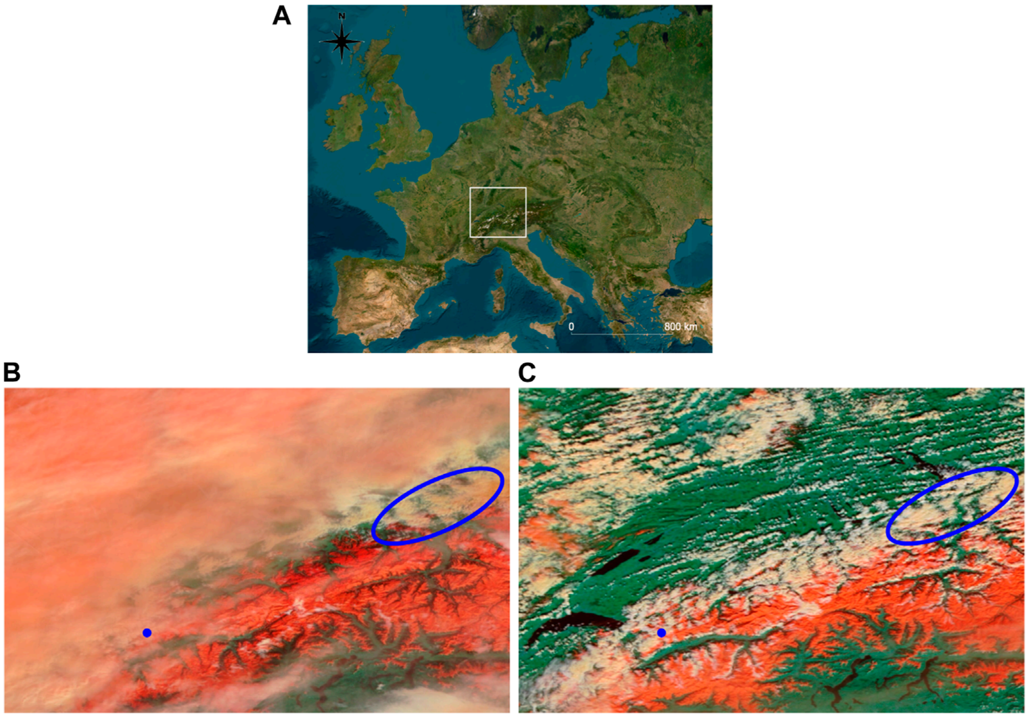


FIGURE 7
(A) Selected area (white rectangle) of the MODIS satellite images [map source (ESRI, 2023)]. EOSDIS satellite image [source: NASA EARTHDATA (2023)] on **(B)** 11/03/2021 and **(C)** 12/03/2021 of the region indicated in **(A)** at 60-m resolution. Snow and ice on the surface are shown in red color. Green and white represent land and clouds, respectively. The blue dot indicates the Diablerets site.

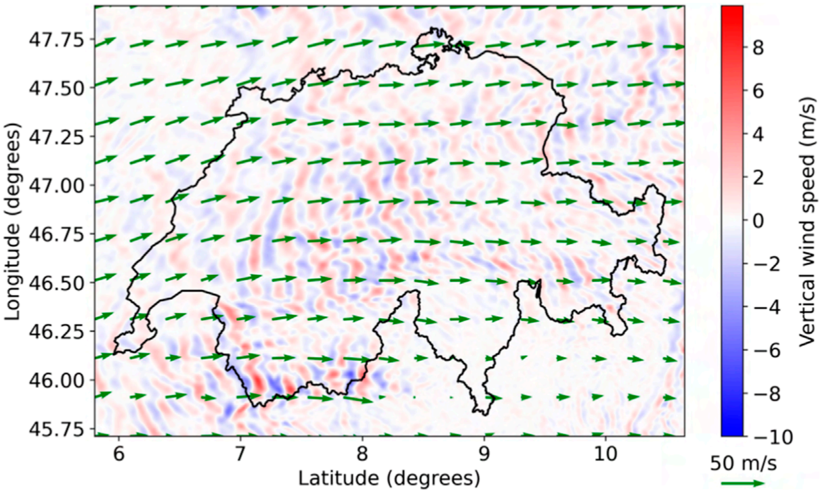


FIGURE 8
Vertical velocity from the terrain following coordinate of COSMO-1E with average height of 2,400 m a.g.l. on 11/03/2021, 12h00. Green arrows represent the horizontal wind speed and direction. The red and blue colors represent upward and downward vertical velocity, respectively. Country border is provided by GADM (2022).

the Wind Atlas data indicate that this case is not an isolated but rather frequently recurring event in the Diablerets area. Cloud lines perpendicular to the southwesterly wind direction can be seen from the satellite images (Figure 7), indicating the same wind direction as seen from the LiDAR measurement and simulation result (Figures 9A, B).

COSMO-1E 2019 data is used to find the relative frequency of atmospheric conditions favoring the formation of mountain waves. We adopted the wind speed threshold and static stability used in Díaz-Fernández et al. (2022). Note that this threshold is mainly based on the atmospheric condition neglecting the topographic characteristics. The purpose of focusing only on the atmospheric

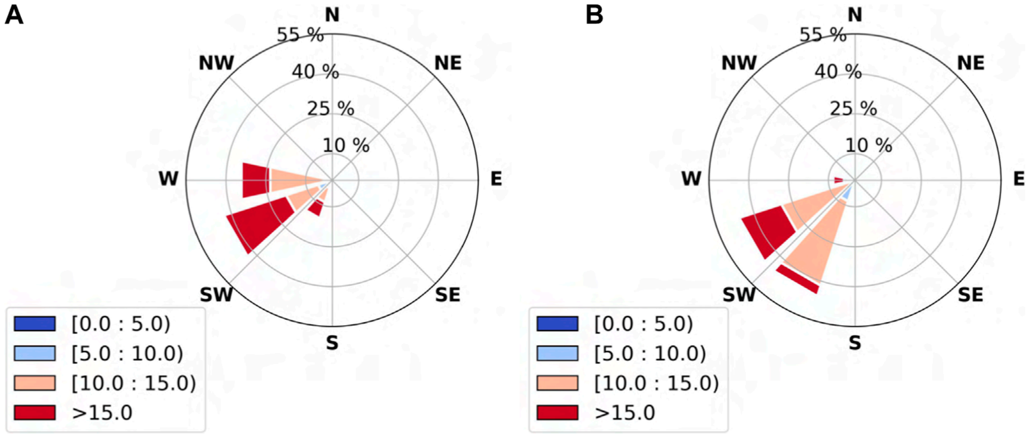


FIGURE 9
Windrose on 11/03/2021 at 100 m a.g.l. at the Diablerets LiDAR site from **(A)** LiDAR measurements and **(B)** the COSMO-WRF simulation. Colors represent wind speed in m/s unit.

conditions is to create a universal threshold that can be quickly adapted to any location in the Alpine region. As shown in Figure 7B almost the entire Swiss Alpine area may be affected by mountain waves, therefore we focus on atmospheric conditions that allow generation of mountain waves and assume spatial coverage at the scale of the Alpine region. We acknowledged that this crude estimation may lead to significant overestimation but serves here as an order of magnitude characterization to get an idea of the significance of mountain wave events.

For a mountain wave to develop, wind speed needs to be high enough to traverse the mountain ridge, otherwise flow separation or other topographic wind flow will occur instead of a wave. Following the recommendation of Reichmann (1978) and the study of Draxl et al. (2021), a wind speed aloft larger than 8 m/s would be sufficient for a mountain wave to be formed. Therefore, the first threshold is a wind speed at 1,164 m a.g.l. equal or larger than 8 m/s. The second threshold is the static stability (ST, Eq. 1) number following the equation and threshold used by Díaz-Fernández et al. (2022). For a wave to form, the static stability (ST) number is required to be between 0.0002 and 0.0014 K/Pa. T is temperature, θ is potential temperature, and p is the barometric pressure. Variable $d(\theta)$ and dp are calculated using potential temperature and pressure difference from 1,164 m a.g.l. and 10 m a.g.l. from COSMO-1E data.

$$ST = \frac{-\left(\frac{T}{\theta}\right) * d(\theta)}{dp} \tag{1}$$

The percentage of atmospheric conditions favorable for the potential formation of mountain waves is shown in Figure 12 for the Diablerets and Lukmanier LiDAR sites. Both sites show a higher percentage of favorable conditions during the winter time, covering up to 80% of the time. This result demonstrates the importance of considering mountain waves when assessing wind energy, especially, when wind energy is designed to respond to the increased winter energy demand. Both sites also show a slightly higher percentage during the first half of wintertime in October to December compared to January to March. The percentage is lower during the spring and summer seasons from April to September. However, there is still a significant number of days with the potential of mountain wave formation of approximately 40%. This might be due to the relatively high wind speed at the Diablerets and Lukmanier LiDAR sites and the stable atmospheric conditions during the night time, both leading to the defined threshold being met.

The fluctuation in wind speed related to the mountain wave event has the potential to influence wind energy production, depending on its location in the region of the power curve (Figure 2). In this paper, we are solely focusing on the impact of wind speed fluctuations and we exclude other variables related to mountain wave event that might also influence the power output (i.e. turbulence, etc). For a mountain wave to occur, it requires high wind speed, as we defined in the threshold above. Depending on the type of wind turbine, it is less likely that the fluctuation will occur in the first region, below the v_{cut_in}. If the fluctuation occurs within the second region (between v_{cut_in} and v_{rated}), depending on the scale of the fluctuation amplitude, we can expect a high impact on the power output production. For stable output of power production and minimum impact of mountain wave on power production, the range of fluctuation ideally occurs within the third region (between v_{rated} and v_{cut_off}) but not reaching the fourth region, which in this case could lead to wind turbines not operating. The region distribution of the power curve varies between the wind turbine infrastructures. The goal of introducing the possibility of mountain wave occurrence during the planning and wind assessment process is to help the process of infrastructure selection to maximize the potential power production in the area.

Mountain wave events in the Diablerets region have been shown to influence the wind speed at the typical height of wind turbine hubs. For regions prone to a high occurrence of mountain waves, we suggest consideration of mountain waves potentially propagating down to the level of wind turbine hubs when assessing wind potential in complex Alpine terrain. In Switzerland, wintertime energy demand increases; at the same time, the stable atmospheric conditions during winter create a favorable situation for mountain

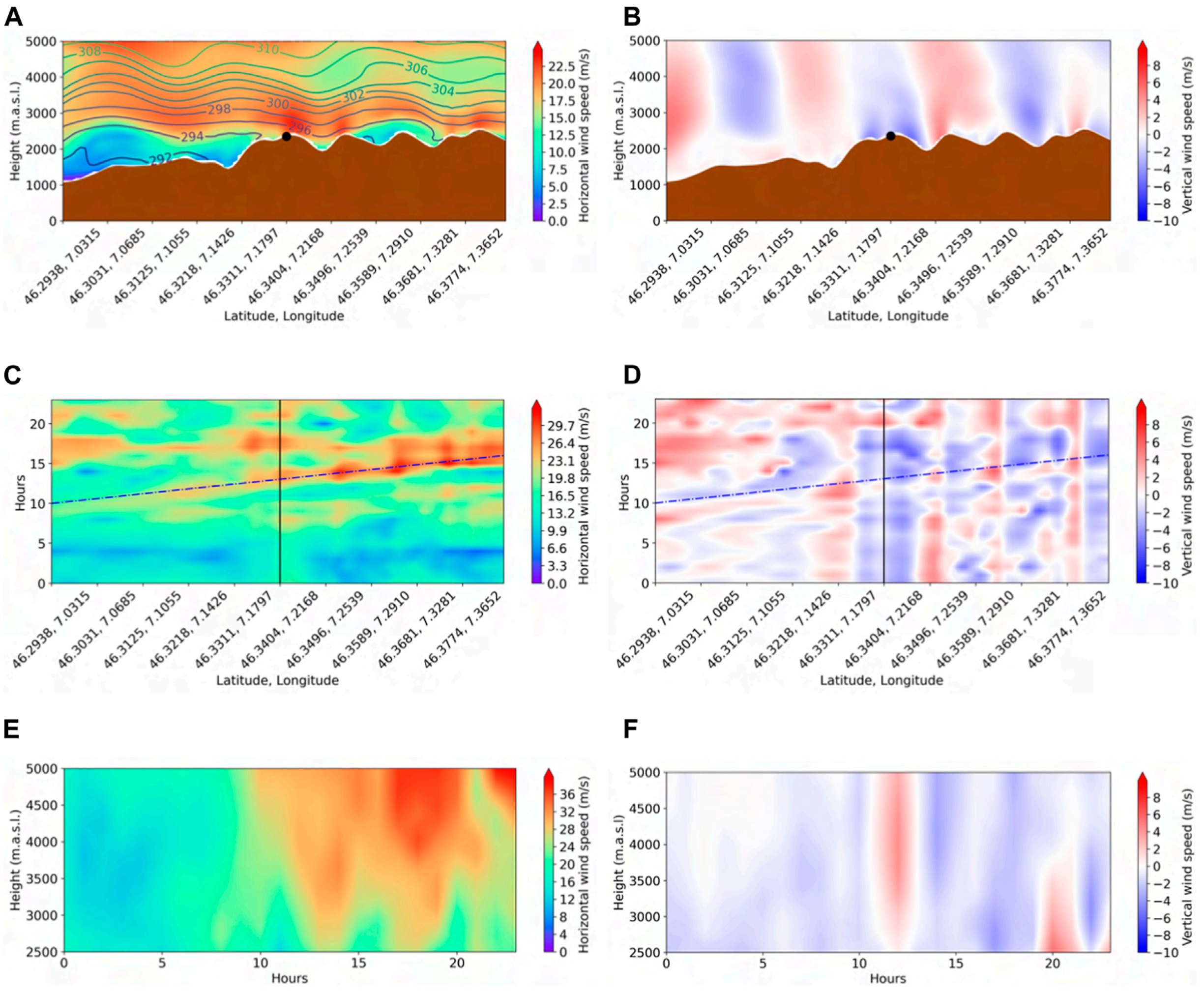


FIGURE 10
Vertical cross-section from COSMO-WRF on 11/03/2021, 08h00, following the orange line in Figure 1 of **(A)** horizontal wind speed and **(B)** vertical wind speed. The black dot indicates the LiDAR site. The contour lines in **(A)** show the potential temperature. The horizontal axis represents the latitude and longitude from west to east. The color bars in (ace) and (bdf) represent horizontal and vertical wind speeds, respectively. Red in the **(B,D,F)** color bar indicates upward vertical wind. Hovmoller diagram of **(C)** horizontal and **(D)** vertical wind speed at the height of 2564.90 m a.s.l. following the orange line as above. Hovmoller diagram of **(E)** horizontal wind speed and **(F)** vertical wind speed at the LiDAR location. The black vertical line in **(C,D)** represents the LiDAR location. The blue dashed line in **(C,D)** represents the propagating wave. The vertical axis of (abef) represents the height above sea level and the vertical axis of (cd) represents the hours.

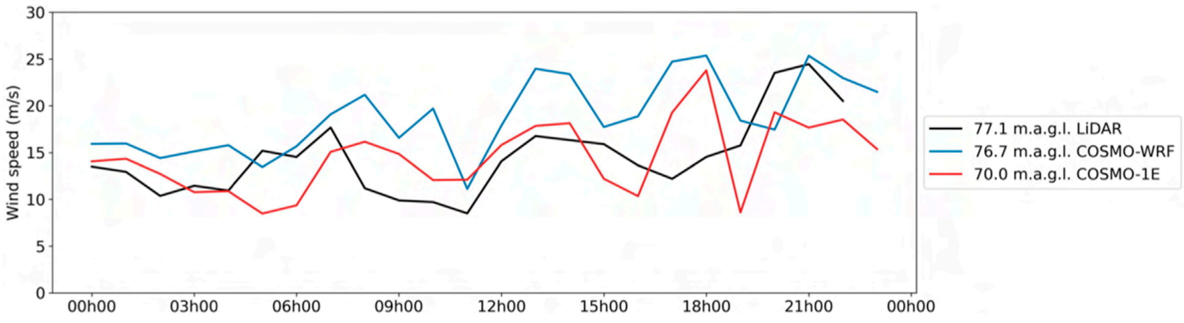


FIGURE 11
Hourly averaged time series of wind speed from LiDAR, COSMO-WRF, and COSMO-1E at the Diablerets LiDAR site on 11/03/2021. The color legend represents the height above ground level.

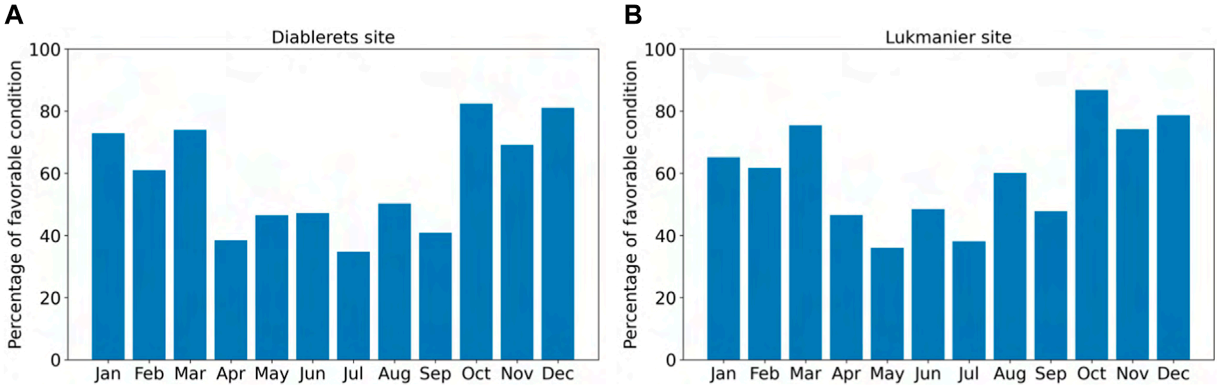


FIGURE 12
Time fraction (%) of conditions favorable for the occurrence of mountain waves based on the COSMO-1E 2019 data at the **(A)** Diablerets and **(B)** Lukmanier sites.

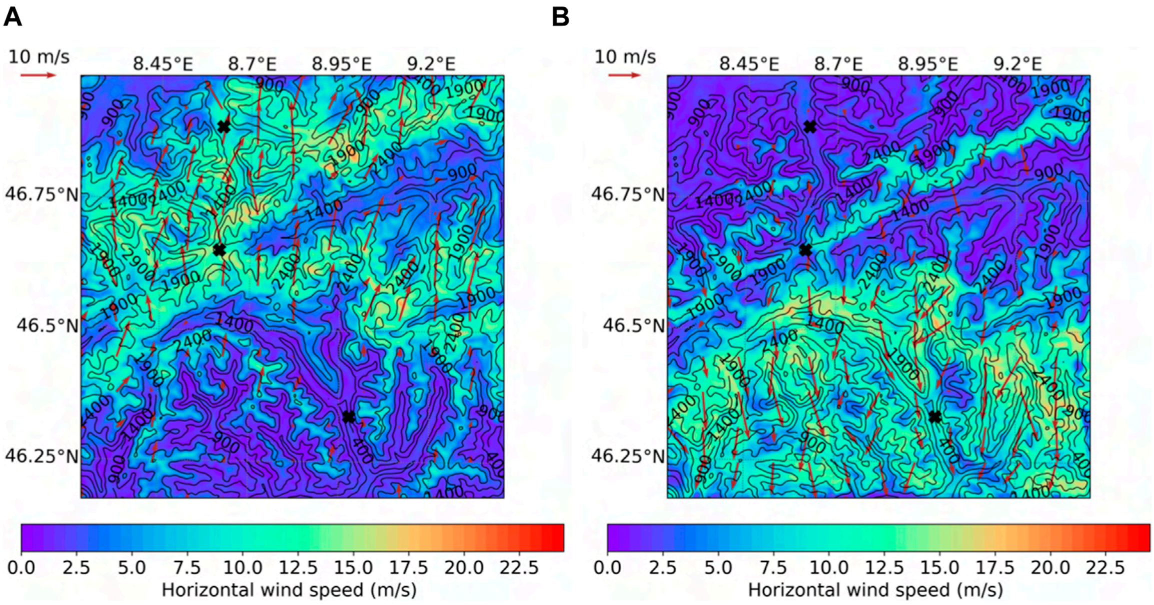


FIGURE 13
Daily mean wind speed from COSMO-WRF at 100 m a.g.l. in the Lukmanier region on 23/10/2020 **(A)** and 09/12/2020 **(B)**. The arrows indicate the wind direction. The contour lines show the elevation above sea level in the COSMO-WRF model. The black crosses show the locations of the wind measurement stations used as input for the Foehnix model.

waves to occur. We have shown that the percentage of favorable atmospheric conditions for wave formation is higher during the winter time up to 80% at the Lukmanier and Diablerets LiDAR sites. Further investigation is still needed to study how the downward propagation of mountain waves exactly influences wind energy production.

3.3 Influence of Foehn wind at turbine hub height

In this section, the influence of Foehn on potential wind is explored and a statistical estimation of event occurrence is provided, to support an accurate wind assessment in complex terrain. First, we provide analyses of two Foehn events in the Lukmanier domain (Figure 1, blue box) using two simulations with a southerly (23/10/2020) and northerly (09/12/2020) wind direction, respectively. Then the statistical estimation of how often Foehn events occur is provided based on meteorological measurements in the surroundings of the Lukmanier area in 2022 using the Foehnix model (Plavcan et al., 2014) (Sect. 2.2.3). Figure 13A shows the mean wind speed from COSMO-WRF on 23/10/2020 at 100 m a.g.l. in the Lukmanier domain during southerly Foehn. The simulation shows high wind speed over the lee slopes of the main mountain ridges, corresponding to the northern part of the domain. Figure 13B shows the simulated mean wind speed

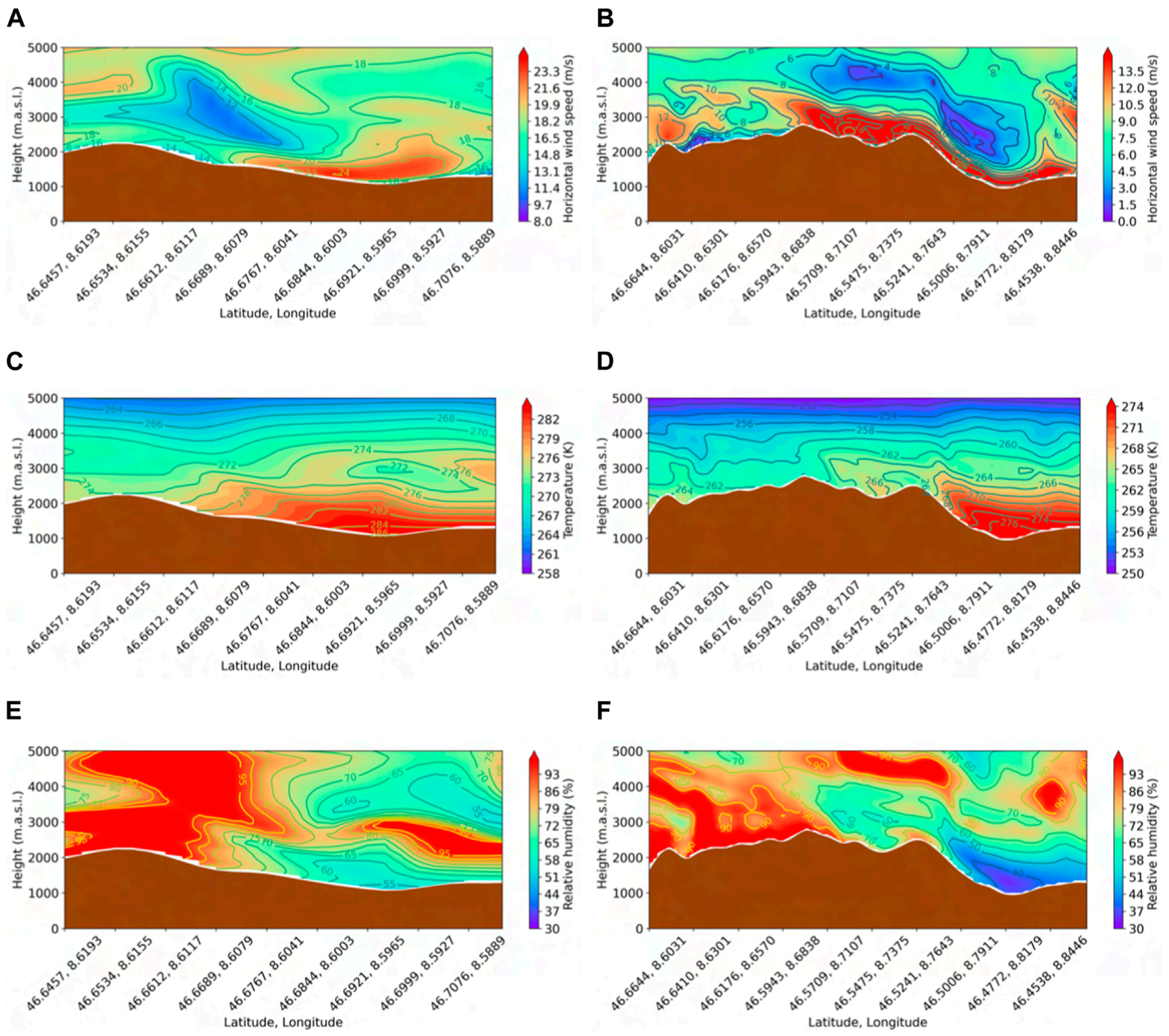


FIGURE 14
Vertical cross-section from COSMO-WRF on 23/10/2020, 13h00 for southerly Foehn (left) and on 09/12/2020, 09h00 for northerly Foehn (right) for **(A,B)** horizontal wind speed (m/s), **(C,D)** temperature (K), and **(E,F)** relative humidity (%). For the southerly Foehn event, the vertical cross-section starts from 46.64560°*N*, 8.61934°*E* (south of GUE crest station) to 46.71181°*N*, 8.58495°*E* (east side of Göschenen valley) as shown by the black line in the Lukmanier domain, Figure 1. For the northerly Foehn event, the vertical cross-section starts from 46.66363°*N*, 8.60292°*E* (northwest of GUE crest station) to 46.43380°*N*, 8.86705°*E* (along Valle Leventina) as shown by the orange line in the Lukmanier domain, Figure 1.

on 09/12/2020 during northerly Foehn. In this case, areas of high wind speed over the lee slopes are located in the southern part of the domain.

For further analysis, a vertical cross-section is shown in Figure 14 with the wind direction left to right and including the crest and valley in the direction of Foehn wind. For both, the northerly and southerly case, we see an increase in wind speed at the lee side of the mountain (Figures 14A, B) confirming the situation shown in Figure 13. The increase in horizontal wind speed is accompanied by an increase in temperature (Figures 14C, D) and a decrease in relative humidity (Figures 14E, F), all typical characteristics of a warm dry Foehn wind. As we see in Figure 14, the altitude affected by the Foehn event involves the height where wind turbines operate. Therefore, for wind assessment purposes in areas known to be affected by Foehn, we recommend including the frequency of Foehn occurrence for a more accurate wind assessment.

Applying the Foehnix model, we select the GUE crest as the central indicator station for both the northerly and southerly Foehn. BIA and ALT are selected as downwind indicator valley stations for northerly and southerly Foehn, respectively. More details of these stations can be found in Table 1 and their location is shown in Figure 1. The wind direction filter was chosen based on the wind roses of the three stations and the topography situation. Figure 15 shows the wind roses of the stations ALT, GUE, and BIA for 2022. A total number of 8,760 hourly wind speed records were utilized. ALT station shows a major wind direction from the south-southeast, while BIA presents a dominant wind direction from the north-northwest. Both stations show a secondary sector almost

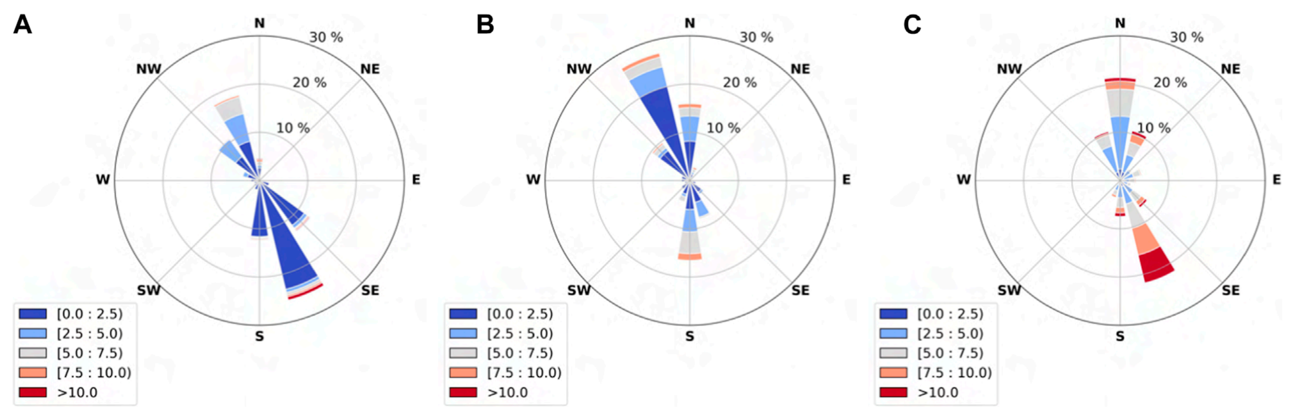


FIGURE 15
Wind roses for the measurement stations of **(A)** Altdorf (ALT), **(B)** Biasca (BIA), and **(C)** Gütsch (GUE) for the year 2022. The color bar represents wind speed (m/s).

opposite to the respective dominant sector. GUE station, located on the crest, shows prevailing winds from both directions, north and south. Taking the local topography situation and wind direction into consideration, we pick the southeasterly as the main axis for the wind direction filter of southerly Foehn. We also pick northerly for the wind direction filter for the northerly Foehn. Hence, we applied a wind direction filter of 45°–225° and 270°–90° for southerly and northerly Foehn, respectively. The wind direction filters span within a sector of 180° following a recommendation by Plavcan et al. (2014) and is set equally for both crest and valley stations. The defined wind direction filter cluster the measured wind in the specified sectors. Then, based on air temperature difference, wind speed, and relative humidity, the Foehnix model identifies the probability of occurrence of Foehn events.

When the estimated probability of Foehn exceeds 50%, it is assumed that a Foehn event occurs. The threshold of 50% follows the classification threshold used by Plavcan et al. (2014) The Foehnix model provides the number of hours with favorable conditions for Foehn generation (Figures 16A, B). The northerly Foehn has the highest frequency during the winter months of January and February. A similarly high frequency of northerly Foehn during winter has been reported by Meteoswiss for the Poschiavo station in eastern Switzerland from 2008 until 2020 (MeteoSwiss, 2023). This comparison is made to show the representativity of the seasonal patterns for a larger region and for a longer period of time. Poschiavo and Altdorf are some of the most representative stations for the northerly and southerly Foehn according to MeteoSwiss (2023) (Figures 16C, D). Poschiavo station, in particular, has been known for its record-high Foehn activity MeteoSwiss (2023). This can also be seen from the high hours of favorable conditions produced by the Foehnix model for the northerly Foehn (Figure 16A). The frequency of northerly Foehn from the Foehnix model varies strongly from spring to autumn 2022, while the Foehn hours from Poschiavo station are lowest from August to November. Following the Foehnix results, the southerly Foehn has the highest occurrence in March 2022. The long-term data from Altdorf shows the highest frequency in April from 1991 to 2020 (MeteoSwiss, 2023). Both show the lowest frequency during July and August. The difference between the result of the Foehnix model from measurement stations in 2022 and the long-term records from Meteoswiss can be attributed to the inter-annual variability and different locations of the stations.

The Foehn event increases the wind speed significantly in the valley region and has a high probability to be higher than the v_{cut_off} (fourth region, Figure 2). Therefore, wind assessment in the area with a high probability of Foehn occurrence should be done thoroughly. The benefit of including the Foehn event in the wind assessment is not only for a better selection of appropriate infrastructure but also to give us a better picture of the future potential production of wind turbines. With the right infrastructure, a wind turbine could handle the high wind speed event of Foehn and reduce the number of nonoperational wind turbines. Even when the Foehn event still results in a non-operational wind turbine, including it in the assessment process will improve the accuracy of the production forecast. A high number of hours of Foehn as seen in Figure 16 should be taken into consideration, especially if the goal is to fulfill the energy demand during the wintertime. The warm dry high wind speed produced by the Foehn event can also reduce the icing issue for the wind turbine during the wintertime. For the installation of wind turbines in a remote complex terrain area, this would mean less maintenance needed.

Foehn has been shown to increase the wind speed on the lee slopes in the Lukmanier domain. Results from an analysis of historical Foehn data from Meteoswiss, together with model predictions based on data measured in 2022 show a significantly higher frequency of Foehn events especially during the winter (northerly Foehn) and spring season (southerly Foehn), strengthening the motivation to include Foehn analysis in wind power assessments with the objective to reduce the energy production gap during the winter season. More research is needed to quantify the impact of Foehn on turbine power yields, the impact of turbulence, and to develop better forecasts for accurate wind speed assessment.

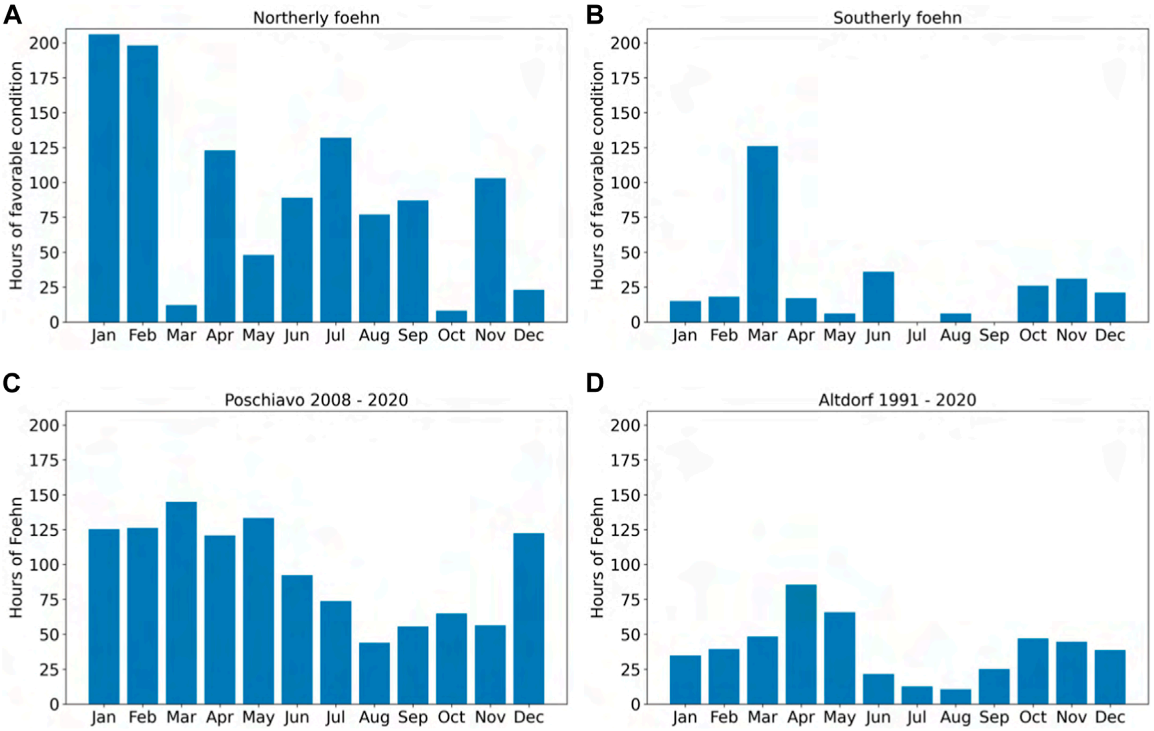


FIGURE 16
Number of hours of favorable conditions for **(A)** northerly Foehn (BIA and GUE stations) and **(B)** southerly Foehn (ALT and GUE stations) to occur based on the Foehnix model using measurements of 2022. Number of hours of **(C)** northerly Foehn at Poschiavo from 2008 to 2020, and **(D)** southerly Foehn at Altdorf from 1991 to 2020 (MeteoSwiss, 2023).

4 Conclusion

For providing the most advanced assessment of wind potential in complex mountain terrain, it is necessary to improve our understanding and the modeling of topography-induced effects on wind. This paper presents examples of terrain effects on wind in the Swiss Alps, namely mountain waves and Foehn, and estimates the occurrence of these phenomena. Two field measurement campaigns were conducted in the Lukmanier and Diablerets areas of the Swiss Alps collecting wind data using a Doppler wind LiDAR instrument. The COSMO-WRF model is used to investigate meteorological events that led to stronger wind at the mountain tops and the typical height of wind turbine hubs during the field measurements. The numerical simulations with COSMO-WRF using wind data from measurement stations located in the model domain show an improvement in wind speed representation near ground level in complex mountain terrain. However, modeled high-resolution wind further aloft is mainly driven by the input data, COSMO-1E, which shows an overestimation compared to the LiDAR measurement data. More realistic input data is needed for a more accurate simulation of wind at higher altitudes.

Mountain waves and Foehn are investigated as examples of meteorological phenomena that happen in Alpine complex terrain with significant impact for the wind energy production, especially during winter. Mountain waves occur during stable atmospheric conditions which are more prevalent in winter, when the demand of energy is high. Our study shows that the wind speed fluctuations associated with mountain waves can propagate downward to the height above ground where the wind turbines typically operate, i.e., 100 m. Further study on the frequency of events and the downward propagation process is still needed for an accurate assessment of wind speed in complex terrain.

The simulation of Foehn events in the Lukmanier area shows that such winds may have a strong effect at the level where the wind turbines operate. Using the Foehnix model, we estimated the probability of Foehn occurrence, which was found higher during winter and early springtime than during summer and fall. This information is useful for establishing a more accurate assessment of wind power potential in complex terrain. This finding adds the significance of including the Foehn assessment for accurate wind prediction in the Alpine complex terrain. The performed measurements, simulations, and analyses enable an improved accuracy of wind assessment especially during winter by considering prominent characteristic flow features and the meteorological conditions favoring their genesis and occurrence. This underlines the dominant influence of the local terrain and topography on the wind speed and wind direction and thus on potential wind power production by turbines deployed at selected favorable sites for that purpose.

This study underlines the need for sufficiently detailed assessment, including the near surface effects such as through Foehn and mountain waves, to assess their quantitative impact on the potential power production of wind turbines in the Swiss mountains. The future study will include a more in-depth analysis of the complex terrain phenomenon mechanism in the Swiss Alps region, such as analysis for the area of impact and a more solid recommendation for wind energy community in the area with a high probability of mountain wave and Foehn cases. Due to the limited data availability during our measurement campaign, we have a limited number of cases to investigate. A longer and more thorough campaign will help to add more cases to investigate and create a more generalized conclusion. It might also discover more examples of complex terrain phenomenon which have impacts on wind energy, other than mountain waves and Foehn. Integrating the information on the impact of complex terrain phenomenon with machine learning [i.e. Dujardin and Lehning (2022)] and Digital Twin and its integration to Geographical Information System (GIS) (Agostinelli et al., 2022; Yousef et al., 2023; Piras et al., 2024) will also increase the accuracy of monitoring and performance prediction of wind turbines. The present research acts as a step forward in accurately estimating potential wind power and optimally using it for renewable energy production, particularly during periods when it is most needed, i.e., the winter season when power demand is high and other renewable sources are limited.

Data availability statement

The data presented in the study are deposited in the EnviDat repository and can be found at https://www.doi.org/10.16904/envidat.501.

Author contributions

FK: Writing–original draft, Conceptualization, Investigation, Software, Writing–review and editing. FG: Software, Investigation, Writing–review and editing. SG-H: Validation, Visualization, Writing–review and editing. JD: Writing–review and editing. HH: Investigation, Writing–review and editing. SH: Writing–review and editing. ML: Supervision, Project Administration, Funding acquisition, Conceptualization, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The research published in this publication was carried out with the support of the Swiss Federal Office of Energy as part of the SWEET consortium EDGE, Swiss National Science Foundation (SNSF): Grant 179130, and the Swiss National Supercomputing Centre (CSCS) projects s938, s1115 and s1242. The authors bear sole responsibility for the conclusions and the results presented in this publication.

Acknowledgments

We acknowledge the use of imagery from the Worldview Snapshots application (https://wvs.earthdata.nasa.gov), part of the Earth Observing System Data and Information System (EOSDIS). We thank Glacier3000, Planair, MeteoSwiss, Swiss Federal Department of Defence, Civil Protection and Sport, and Halo Photonics Online support for data provisions, collaborations, and logistical support of the field campaigns.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Agostinelli, S., Cumo, F., Nezhad, M. M., Orsini, G., and Piras, G. (2022). Renewable energy system controlled by open-source tools and digital twin model: zero energy port area in Italy. *Energies* 15, 1817. doi:10.3390/en15051817

Alfredsson, P.-H., and Segalini, A. (2017). Introduction wind farms in complex terrains: an introduction. *Philosophical Trans. R. Soc. A Math. Phys. Eng. Sci.* 375, 20160096. doi:10.1098/rsta.2016.0096

Archer, C. L., and Jacobson, M. Z. (2005). Evaluation of global wind power. *J. Geophys. Res. Atmos.* 110. doi:10.1029/2004JD005462

Archer, C. L., and Jacobson, M. Z. (2013). Geographical and seasonal variability of the global "practical" wind resources. *Appl. Geogr.* 45, 119–130. doi:10.1016/j.apgeog.2013.07.006

Bavay, M., Lehning, M., Jonas, T., and Löwe, H. (2009). Simulations of future snow cover and discharge in alpine headwater catchments. *Hydrol. Process.* 23, 95–108. doi:10.1002/hyp.7195

Beffrey, G., Jaubert, G., and Dabas, A. (2006). Föhn flow and stable air mass in the rhine valley: the beginning of a map event. *Q. J. R. Meteorological Soc.* 130, 541–560. doi:10.1256/qj.02.228

Beljaars, A. C. (1995). The parametrization of surface fluxes in large-scale models under free convection. *Q. J. R. Meteorological Soc.* 121, 255–270. doi:10.1256/smsqj.52202

Chow, F. K., Schär, C., Ban, N., Lundquist, K. A., Schlemmer, L., and Shi, X. (2019). Crossing multiple gray zones in the transition from mesoscale to microscale simulation over complex terrain. *Atmosphere* 10, 274. doi:10.3390/atmos10050274

Chow, F. K., Snyder, B. J., and Wekker, S. F. J. D. (2013). Mountain weather research and forecasting: recent progress and current challenges. *Mt. Weather Res. Forecast.* doi:10.1007/978-94-007-4098-3

Clark, R. T., Bett, P. E., Thornton, H. E., and Scaife, A. A. (2017). Skilful seasonal predictions for the european energy industry. *Environ. Res. Lett.* 12, 024002. doi:10.1088/1748-9326/aa57ab

Clifton, A., Barber, S., Stökl, A., Frank, H., and Karlsson, T. (2022). Research challenges and needs for the deployment of wind energy in hilly and mountainous regions. *Wind Energy Sci.* 7, 2231–2254. doi:10.5194/wes-7-2231-2022

Clifton, A., Daniels, M., and Lehning, M. (2014). Effect of winds in a mountain pass on turbine performance. *Wind Energy* 17, 1543–1562. doi:10.1002/we.1650

Dar, A. S., Berg, J., Troldborg, N., and Patton, E. G. (2019). On the self-similarity of wind turbine wakes in a complex terrain using large eddy simulation. *Wind Energy Sci.* 4, 633–644. doi:10.5194/wes-4-633-2019

Dhunny, A., Lollchund, M., and Rughooputh, S. (2017). Wind energy evaluation for a highly complex terrain using computational fluid dynamics (cfd). *Renew. Energy* 101, 1–9. doi:10.1016/j.renene.2016.08.032

Díaz-Fernández, J., Bolgiani, P., Santos-Muñoz, D., Quitián-Hernández, L., Sastre, M., Valero, F., et al. (2022). Comparison of the wrf and harmonie models ability for mountain wave warnings. *Atmos. Res.* 265, 105890. doi:10.1016/j.atmosres.2021.105890

Draxl, C., Worsnop, R. P., Xia, G., Pichugina, Y., Chand, D., Lundquist, J. K., et al. (2021). Mountain waves can impact wind power generation. *Wind Energy Sci.* 6, 45–60. doi:10.5194/wes-6-45-2021

Dujardin, J., Kahl, A., Kruyt, B., Bartlett, S., and Lehning, M. (2017). Interplay between photovoltaic, wind energy and storage hydropower in a fully renewable Switzerland. *Energy* 135, 513–525. doi:10.1016/j.energy.2017.06.092

Dujardin, J., Kahl, A., and Lehning, M. (2021). Synergistic optimization of renewable energy installations through evolution strategy. *Environ. Res. Lett.* 16, 064016. doi:10.1088/1748-9326/abfc75

Dujardin, J., and Lehning, M. (2022). Wind-topo: downscaling near-surface wind fields to high-resolution topography in highly complex terrain with deep learning. *Q. J. R. Meteorological Soc.* 148, 1368–1388. doi:10.1002/qj.4265

Dyer, A., and Hicks, B. (1970). Flux-gradient relationships in the constant flux layer. *Q. J. R. Meteorological Soc.* 96, 715–721. doi:10.1002/qj.49709641012

Elgendi, M., AlMallahi, M., Abdelkhalig, A., and Selim, M. Y. (2023). A review of wind turbines in complex terrain. *Int. J. Thermofluids* 17, 100289. doi:10.1016/j.ijft.2023.100289

ESRI (2023). World imagery. Available at: https://www.arcgis.com/home/item.html?id=10df2279f9684e4a9f6a7f08febac2a9 (Accessed November 13, 2023).

European Environmental Agency (2006). CORINE land cover (CLC) 2006 raster data,Version 13. Available at: https://www.eea.europa.eu/dataand-maps/data/clc-2006-raster (Accessed May 28, 2021).

Federal Office of Meteorology and Climatology MeteoSwiss (2023a). Cosmo forecasting system. Available at: https://www.meteoswiss.admin.ch/weather/warning-and-forecasting-systems/cosmo-forecasting-system.html (Accessed January 4, 2023).

Federal Office of Meteorology and Climatology MeteoSwiss (2023b) Measurement instruments. Available at: https://www.meteoswiss.admin.ch/weather/measurement-systems/land-based-stations/automatic-measurement-network/measurement-instruments.html (Accessed February 9, 2023).

Federal Office of Topography (2023). Swiss map raster. Available at: https://map.geo.admin.ch/ (Accessed May 23, 2023).

GADM (2022). GADM version 4.1. Available at: https://geodata.ucdavis.edu/gadm/ (Accessed November 21, 2023).

Gerber, F., Besic, N., Sharma, V., Mott, R., Daniels, M., Gabella, M., et al. (2018). Spatial variability in snow precipitation and accumulation in cosmo–wrf simulations and radar estimations over complex terrain. *Cryosphere* 12, 3137–3160. doi:10.5194/tc-12-3137-2018

Gerber, F., and Lehning, M. (2021). High resolution static data for WRF over Switzerland. Tech. rep.

Gerber, F., and Sharma, V. (2018) *Running cosmo-wrf on very-high resolution over complex terrain.* doi:10.16904/envidat.35

Graabak, I., and Korpås, M. (2016). Variability characteristics of european wind and solar power resources—a review. *Energies* 9, 449. doi:10.3390/en9060449

Grams, C. M., Beerli, R., Pfenninger, S., Staffell, I., and Wernli, H. (2017). Balancing europe's wind-power output through spatial deployment informed by weather regimes. *Nat. Clim. Change* 7, 557–562. doi:10.1038/nclimate3338

Kealy, J. C., Efstathiou, G. A., and Beare, R. J. (2019). The onset of resolved boundary-layer turbulence at grey-zone resolutions. *Boundary-Layer Meteorol.* 171, 31–52. doi:10.1007/s10546-018-0420-0

Klemp, J. B., Dudhia, J., and Hassiotis, A. D. (2008). An upper gravity-wave absorbing layer for nwp applications. *Mon. Weather Rev.* 136, 3987–4004. doi:10.1175/2008MWR2596.1

Koletsis, I., Lagouvardos, K., Kotroni, V., and Bartzokas, A. (2009). The interaction of northern wind flow with the complex topography of crete island – part 1: observational study. *Nat. Hazards Earth Syst. Sci.* 9, 1845–1855. doi:10.5194/nhess-9-1845-2009

Koletsis, I., Lagouvardos, K., Kotroni, V., and Bartzokas, A. (2010). The interaction of northern wind flow with the complex topography of crete island – part 2: numerical study. *Nat. Hazards Earth Syst. Sci.* 10, 1115–1127. doi:10.5194/nhess-10-1115-2010

Koller, S., and Humar, T. (2016) *Windpotentialanalyse für windatlas.ch: Jahresmittelwerte der modellierten windgeschwindigkeit und windrichtung.*

Kristianti, F., Dujardin, J., Gerber, F., Huwald, H., Hoch, S. W., and Lehning, M. (2023). Combining weather station data and short-term lidar deployment to estimate wind energy potential with machine learning: a case study from the swiss alps. *Boundary-Layer Meteorol.* 188, 185–208. doi:10.1007/s10546-023-00808-y

Kristianti, F., Gerber, F., and Lehning, M. (2024) *Cosmo-wrf dataset for swiss alps simulations in gray-zone resolution.* doi:10.16904/envidat.501

Kruyt, A. C. (2019). Potential and uncertainty of wind energy in the swiss alps. *EPFL thesis*, 126. doi:10.5075/epfl-thesis-9350

Kruyt, B., Dujardin, J., and Lehning, M. (2018). Improvement of wind power assessment in complex terrain: the case of cosmo-1 in the swiss alps. *Front. Energy Res.* 6, 102. doi:10.3389/fenrg.2018.00102

Lange, J., Mann, J., Berg, J., Parvu, D., Kilpatrick, R., Costache, A., et al. (2017). For wind turbines in complex terrain, the devil is in the detail. *Environ. Res. Lett.* 12, 094020. doi:10.1088/1748-9326/aa81db

Lehner, M., and Rotach, M. W. (2018). Current challenges in understanding and predicting transport and exchange in the atmosphere over mountainous terrain. *Atmosphere* 9, 276. doi:10.3390/atmos9070276

Lehner, M., Whiteman, C. D., Hoch, S. W., Crosman, E. T., Jeglum, M. E., Cherukuru, N. W., et al. (2016). The metcrax ii field experiment: a study of downslope windstorm-type flows in Arizona's meteor crater. *Bull. Am. Meteorological Soc.* 97, 217–235. doi:10.1175/BAMS-D-14-00238.1

Lehning, M., Doorschot, J., and Bartelt, P. (2000). A snowdrift index based on snowpack model calculations. *Ann. Glaciol.* 31, 382–386. doi:10.3189/172756400781819770

Mann, J., Angelou, N., Arnqvist, J., Callies, D., Cantero, E., Arroyo, R. C., et al. (2017). Complex terrain experiments in the new european wind atlas. *Philosophical Trans. Ser. A, Math. Phys. Eng. Sci.* 375, 20160101. doi:10.1098/rsta.2016.0101

MeteoSwiss (2022) Meteoswiss idaweb. Available at: https://gate.meteoswiss.ch/idaweb/ (Last access: 2022.

MeteoSwiss (2023). Foehn frequency. Available at: https://www.meteoswiss.admin.ch/weather/weather-and-climate-from-a-to-z/foehn-frequency.html (Accessed November 6, 2023).

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A. (1997). Radiative transfer for inhomogeneous atmospheres: rrtm, a validated correlated-k model for the longwave. *J. Geophys. Res. Atmos.* 102, 16663–16682. doi:10.1029/97jd00237

Morrison, H., Thompson, G., and Tatarskii, V. (2009). Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: comparison of one-and two-moment schemes. *Mon. weather Rev.* 137, 991–1007. doi:10.1175/2008mwr2556.1

NASA EARTHDATA (2023). EOSDIS Worldview. Available at: https://worldview.earthdata.nasa.gov/ (Accessed November 13, 2023).

NASA/METI/AIST/Japan Space systems and U.S./Japan ASTER Science Team (2019). ASTER global digital elevation model V003, *NASA EOSDIS Land Process. DAAC, Distrib. by NASA EOSDIS Land Process. DAAC.* doi:10.5067/ASTER/ASTGTM.003

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community noah land surface model with multiparameterization options (noah-mp): 1. model description and evaluation with local-scale measurements. *J. Geophys. Res. Atmos.* 116, D12109. doi:10.1029/2010jd015139

Olabi, A., and Abdelkareem, M. A. (2022). Renewable energy and climate change. *Renew. Sustain. Energy Rev.* 158, 112111. doi:10.1016/j.rser.2022.112111

Paulson, C. A. (1970). The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. *J. Appl. Meteorology Climatol.* 9, 857–861. doi:10.1175/1520-0450(1970)009<0857:tmrows>2.0.co;2

Pickering, B., Grams, C. M., and Pfenninger, S. (2020). Sub-national variability of wind power generation in complex terrain and its correlation with large-scale meteorology. *Environ. Res. Lett.* 15, 044025. doi:10.1088/1748-9326/ab70bd

Piras, G., Agostinelli, S., and Muzi, F. (2024). Digital twin framework for built environment: a review of key enablers. *Energies* 17, 436. doi:10.3390/en17020436

Plavcan, D., Mayr, G. J., and Zeileis, A. (2014). Automatic and probabilistic foehn diagnosis with a statistical mixture model. *J. Appl. Meteorology Climatol.* 53, 652–659. doi:10.1175/jamc-d-13-0267.1

Reichmann, H. (1978) *Cross-country soaring*. Santa Monica, Ca, United States: Thomson Publications.

Schraff, C., Reich, H., Rhodin, A., Schomburg, A., Stephan, K., Periáñez, A., et al. (2016). Kilometre-scale ensemble data assimilation for the cosmo model (kenda). *Q. J. R. Meteorological Soc.* 142, 1453–1472. doi:10.1002/qj.2748

Shin, H. H., and Hong, S.-Y. (2015). Representation of the subgrid-scale turbulent transport in convective boundary layers at gray-zone resolutions. *Mon. Weather Rev.* 143, 250–271. doi:10.1175/mwr-d-14-00116.1

Sims, R. (2004). Renewable energy: a response to climate change. *Sol. Energy* 76, 9–17. Solar World Congress 2001. doi:10.1016/S0038-092X(03)00101-4

Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Wang, W., et al. (2008) *A description of the advanced research wrf version 3 27*, 3–27.

Skamarock, W., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., et al. (2021) *A description of the advanced research wrf model version 4.3*. doi:10.5065/1dfh-6p97

Swiss Federal Office of Energy (2023a). Energy consumption in Switzerland 2021. Available at: https://www.bfe.admin.ch/bfe/en/home/supply/statistics-and-geodata/energy-statistics/overall-energy-statistics.html/ (Accessed May 30, 2023).

Swiss Federal Office of Energy (2023b) Energy perspectives 2050+. Available at: https://www.bfe.admin.ch/bfe/en/home/policy/energy-perspectives-2050-plus.html (Accessed May 30, 2023).

Tabas, D., Fang, J., and Porté-Agel, F. (2019). Wind energy prediction in highly complex terrain by computational fluid dynamics. *Energies* 12, 1311. doi:10.3390/en12071311

Tobin, I., Vautard, R., Balog, I., Bréon, F.-M., Jerez, S., Ruti, P., et al. (2015). Assessing climate change impacts on European wind energy from ENSEMBLES high-resolution climate projections. *Clim. Change* 128, 99–112. doi:10.1007/s10584-014-1291-0

Toumelin, L. L., Gouttevin, I., Helbig, N., Galiez, C., Roux, M., and Karbou, F. (2023). Emulating the adaptation of wind fields to complex terrain with deep learning. *Artif. Intell. Earth Syst.* 2, e220034. doi:10.1175/aies-d-22-0034.1

Udina, M., Bech, J., Gonzalez, S., Soler, M. R., Paci, A., Miró, J. R., et al. (2020). Multi-sensor observations of an elevated rotor during a mountain wave event in the eastern pyrenees. *Atmos. Res.* 234, 104698. doi:10.1016/j.atmosres.2019.104698

Webb, E. K. (1970). Profile relationships: the log-linear range, and extension to strong stability. *Q. J. R. Meteorological Soc.* 96, 67–90. doi:10.1002/qj.49709640708

WSL SLF (2022) Wsl slf institute for snow and avalanche research. Available at: https://www.slf.ch/en/avalanche-bulletin-and-snow-situation/measured-values/description-of-automated-stations.html (Last access: 2022.

Wyngaard, J. C. (2004). Toward numerical modeling in the "*terra incognita*". *J. Atmos. Sci.* 61, 1816–1826. doi:10.1175/1520-0469(2004)061⟨1816:TNMITT⟩2.0.CO;2

Xia, G., Draxl, C., Raghavendra, A., and Lundquist, J. K. (2021). Validating simulated mountain wave impacts on hub-height wind speed using sodar observations. *Renew. Energy* 163, 2220–2230. doi:10.1016/j.renene.2020.10.127

Yang, Z.-L., Niu, G.-Y., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., et al. (2011). The community noah land surface model with multiparameterization options (noah-mp): 2. evaluation over global river basins. *J. Geophys. Res. Atmos.* 116, D12110. doi:10.1029/2010jd015140

Yousef, L. A., Yousef, H., and Rocha-Meneses, L. (2023). Artificial intelligence for management of variable renewable energy systems: a review of current status and future directions. *Energies* 16, 8057. doi:10.3390/en16248057

Zhang, D., and Anthes, R. A. (1982). A high-resolution model of the planetary boundary layer—sensitivity tests and comparisons with sesame-79 data. *J. Appl. Meteorology* 21 (1962-1982), 1594–1609. doi:10.1175/1520-0450(1982)021<1594:ahrmot>2.0.co;2

Frontiers in Energy Research

TYPE Original Research
PUBLISHED 01 July 2024
DOI 10.3389/fenrg.2024.1346369

Check for updates

OPEN ACCESS

EDITED BY
Lorenzo Ferrari,
University of Pisa, Italy

REVIEWED BY
Cristian Paul Chioncel,
Babeş-Bolyai University, Romania
Seçkin Karasu,
Bülent Ecevit University, Türkiye

*CORRESPONDENCE
Yongfeng Qiu,
✉ qiu19851219@126.com

RECEIVED 29 November 2023
ACCEPTED 30 May 2024
PUBLISHED 01 July 2024

CITATION
Man J, Xu K, Wang D, Liu Y, Zhan J and Qiu Y (2024), Multi-device wind turbine power generation forecasting based on hidden feature embedding.
Front. Energy Res. 12:1346369.
doi: 10.3389/fenrg.2024.1346369

COPYRIGHT
© 2024 Man, Xu, Wang, Liu, Zhan and Qiu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Multi-device wind turbine power generation forecasting based on hidden feature embedding

Junfeng Man[1,2,3], Ke Xu[1], Dian Wang[4], Yong Liu[4], Jun Zhan[2,3] and Yongfeng Qiu[5,6]*

[1]School of Computer, Hunan University of Technology, Zhuzhou, China, [2]School of Intelligent Manufacturing, Hunan First Normal University, Changsha, China, [3]Key Laboratory of Industrial Equipment Intelligent Perception and Maintenance, College of Hunan Province, Hunan First Normal University, Changsha, China, [4]CRRC Zhuzhou Electric Locomotive Research Institute Co., Ltd., Zhuzhou, China, [5]Guiyang Aluminum Magnesium Design and Research Institute Co., Ltd., Guiyang, China, [6]Hunan Tianqiao Jiacheng Intelligent Technology Co., Ltd., Zhuzhou, China

In recent years, the global installed capacity of wind power has grown rapidly. Wind power forecasting, as a key technology in wind turbine systems, has received widespread attention and extensive research. However, existing studies typically focus on the power prediction of individual devices. In the context of multi-turbine scenarios, employing individual models for each device may introduce challenges, encompassing data dilution and a substantial number of model parameters in power generation forecasting tasks. In this paper, a single-model method suitable for multi-device wind power forecasting is proposed. Firstly, this method allocates multi-dimensional random vectors to each device. Then, it utilizes space embedding techniques to iteratively evolve the random vectors into representative vectors corresponding to each device. Finally, the temporal features are concatenated with the corresponding representative vectors and inputted into the model, enabling the single model to accomplish multi-device wind power forecasting task based on device discrimination. Experimental results demonstrate that our method not only solves the data dilution issue and significantly reduces the number of model parameters but also maintains better predictive performance. Future research could focus on using more interpretable space embedding techniques to observe representation vectors of wind turbine equipment and further explore their semantic features.

KEYWORDS

wind power generation, time series forecasting, space embedding, hidden feature, long short-term memory

1 Introduction

Since the Industrial Revolution in the 18th century, with the advancement of technology and social progress, the demand for energy has grown rapidly (Wang et al., 2019). Conventional energy sources such as oil, coal, and natural gas not only have limited reserves but also contribute to environmental pollution and global warming (Wang et al., 2019). Wind energy, as a clean and widely distributed renewable energy, has gained global attention in recent years (Liu and Chen, 2019; Wang et al., 2021; Yang et al., 2021). However, the fluctuation of wind energy leads to the instability of power output in wind farms, which imposes additional burdens on energy storage devices and potentially affects the reliability of power supply (Parsons et al., 2004).

Incorporating efficient wind power forecasting methods into power control systems can effectively reduce operational costs and significantly enhance the reliability of wind power systems (Contaxis and Kabouris, 1991; Kariniotakis et al., 1996). Existing wind power forecasting methods mainly focus on individual devices. However, in practical applications, multiple wind turbines often operate in parallel within a wind power system. Assigning independent forecasting models to each device would result in two problems: firstly, dividing the dataset based on devices would lead to limited training data for each model, causing data dilution; secondly, each device having an independent model would result in a large number of total parameters, making accurate forecasting of wind turbine power generation increasingly challenging. In this paper, we propose a training method for prediction models applicable to multi-device scenarios, aiming to address the challenges of data dilution and excessive parameters.

2 Relevant work

The existing time series forecasting methods can be mainly divided into two categories: one consists of classical statistical methods with high interpretability and theoretical foundations, while the other category comprises more efficient methods based on artificial neural networks. The method proposed in this paper combines deep neural networks with space embedding technology from the field of natural language processing, aiming to effectively address the issue of multi-device wind turbine power generation forecasting.

2.1 Traditional wind power forecasting methods

The traditional wind power forecasting methods include physics-based models and statistic-based methods. The physics-based models play an crucial role in traditional forecasting methods, which consider meteorological factors (such as pressure, humidity, and temperature) from numerical weather prediction (NWP) and local topography for forecasting (Jung and Broadwater, 2014; Fang and Chiang, 2016; Hu et al., 2020). In terms of short-term forecasting capability, these methods generally perform moderately well, and their results are more suitable as a reference for long-term forecasting (Hu et al., 2020; Wang et al., 2021). The Autoregressive Integrated Moving Average (ARIMA) model, which is based on the theory of differencing, transforms non-stationary processes into stationary ones to address prediction problems (Ariyo et al., 2014). However, this method can only model individual attributes and fails to consider the correlations among multiple attributes at different time steps. In addition, ARIMA has high computational costs and is rarely applied to modeling and forecasting tasks involving long sequences.

2.2 Time-series forecasting methods based on deep learning

The commonly used wind power forecasting methods based on deep learning include two methods: Recurrent Neural Networks (RNN) and Transformer models. In comparison to the classical RNN model, its variants, such as Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber, 1997)and Gated Recurrent Unit (GRU) (Cho et al., 2014), are more prevalent. LSTM was proposed to address the problem of vanishing gradients caused by long sequence backpropagation. GRU, compared to LSTM, reduces the number of gate units and parameters, making it easier to train it to convergence. Additionally, it exhibits similar performance to LSTM in multiple tasks (Chung et al., 2014). Lai et al. advocated for forecasting models that encompass the impacts of both long-term patterns (such as day-night and season) and short-term patterns (like cloud fluctuations and wind direction). Building upon this concept, they introduced LSTNet, a variant of Convolutional Recurrent Neural Network (CRNN) (Lai et al., 2018). RNN possesses inherent capability in modeling time series data. However, the issue of gradient explosion has not been entirely solved yet. Moreover, their auto-regressive output mode not only extends the output time for long sequence forecasting tasks, but also increases training time due to challenges in parallel training.

The Transformer architecture was initially proposed for machine translation tasks (Vaswani et al., 2017). Although the Transformer model exhibits excellent performance in the field of NLP, its drawbacks are also evident: the model structure is complex, it has a large number of parameters, and it requires a relatively long time to produce outputs. Informer (Zhou et al., 2021) is a variant of the Transformer model designed for time series forecasting tasks. It incorporates the ProbSparse attention mechanism to reduce sampling time and introduces a generative decoder that can output the entire prediction sequence in a single step, significantly reducing the time complexity of the forecasting task. AutoFormer (Wu et al., 2021) introduces a novel attention mechanism called Auto-Correlation, which has stronger information aggregation capability, enabling it to achieve superior forecasting performance compared to variants such as Informer. However, the main advantage of the Transformer architecture lies in its multi-head attention mechanism, which exhibits permutation invariance. Even with the addition of positional encoding in the data, the application of attention mechanisms inevitably results in the loss of temporal information. In the field of natural language processing, semantics and word order are not entirely bound, but in the domain of time series forecasting, the output results are highly correlated with the temporal order. Zeng et al., 2023 have demonstrated that in some time series prediction tasks, a single-layer linear neural network outperforms Transformer-based networks and offers significant advantages.

2.3 Feature engineering

Feature engineering is the process of transforming raw data into features that better represent the essence of the problem. Effective feature engineering can consistently enhance the forecasting accuracy of the model. Two-dimensional Discrete Wavelet Transform (2D-DWT) and the 2D Fast Discrete Orthonormal Stockwell Transform (2D-FDOST) method are used to extract new effective dynamic features from dynamic electrical signals (Karasu and Sarac, 2019; Karasu and Saraç 2022). Compared to the Fourier transform, these methods exhibit stronger adaptability and noise resistance, allowing for localized analysis in different frequency domains and thereby capturing detailed features more effectively. However, these methods have a high computational complexity and are not suitable for scenarios requiring real-time processing. The Multi-Objective Grey Wolf Optimizer (MOGWO) is commonly used to extract a small set of useful features from a large volume of dynamic electrical signals, improving data quality and reducing computational overhead (Karasu and Saraç, 2020). The Grey Wolf algorithm has fewer parameters, is easy to implement, and requires less computational time, but the solutions found may not always be optimal. A feature selection method based on Sequential Floating Forward Selection (SFFS) has been used to reduce the historical operating data of lots of wind turbines in a wind farm environment to 660 effective features (Peng et al., 2021). This not only reduces the computational overhead of the forecasting model but also enhances its forecasting accuracy. However, the hyper-parameters of this algorithm are not adaptive, making it highly dependent on empirical expertise.

Methods for extracting dynamic features from historical power data are commonly used to assist neural network models in forecasting. However, the static features of wind turbine equipment have received less attention from researchers. Static factors, such as the type of wind turbine components, geographic environmental conditions, equipment layout, and equipment failure status, can also have a long-term impact on the power generation patterns of wind turbines, making them not entirely dependent on measurable internal conditions and weather factors.

Mikolov et al. (2013a), Mikolov et al. (2013b) both schemes infer the properties of words based on the distributional order of words in sentences. In natural language, there exist semantic and syntactic correlations between words, and deep learning models need to discover the semantic features hidden beneath the distributional order of words to accurately predict their sequence. However, apart from geographical location, there is no obvious distributional correlation among wind turbine devices. Therefore, CBOW and Skip-gram schemes are not suitable for the task of feature representation for wind turbines. The graph embedding technique (Grover and Leskovec, 2016) requires the model to predict the connectivity structure between nodes in the graph. Then, it utilizes gradient descent algorithm to infer high-dimensional vector representations of nodes or the entire graph. However, in a distributed wind farm, the geographical positions of wind turbines do not conform to the structure of a graph because there is no explicit connection between the nodes representing the wind turbines. Therefore, graph embedding techniques cannot be directly applied to the representation of wind turbine devices, nor can they directly uncover the hidden static features that influence the device's own power generation patterns. Position Embedding (Vaswani et al., 2017), which is a manually specified method for encoding sequence order, utilizes a fixed calculation approach without neural networks or gradient descent algorithms. This method is applied in the position encoding of Transformer models. However, the hidden features of wind turbines are more complex than sequential order, and representation vectors calculated using manually specified algorithms based on device identifiers are unable to effectively reflect the characteristics of wind turbine devices.

Considering the complexity of static factors that influence wind turbine power generation patterns and the implicit correlations among turbines, this paper sets the task of training device representation vectors as power generation forecasting. To achieve this, the gradient descent algorithm is employed to evolve randomly initialized data into vectors that represent the static factors of the devices. This method directly uncovers hidden static factors that impact device power generation patterns. Different from traditional space embedding methods:

(1) Traditional space embedding methods are commonly used to generate generic representation vectors that are not specific to particular business scenarios. As a result, they fail to capture representation information in specific task scenarios. The proposed space embedding technique in this paper specifically addresses the task scenario of distributed wind turbine power forecasting, generating representation vectors that are exclusively applicable to this task scenario.
(2) In traditional space embedding methods, the tasks of generating entity representation vectors and the subsequent tasks of using these vectors often differ. The proposed space embedding method in this paper, however, aligns the task of generating vectors with the subsequent task, both of which is power generation forecasting.

2.4 Work presented in this paper

Classical statistical models, recurrent neural networks, and deep neural networks based on the Transformer architecture are all forecasting models designed for single devices. However, in multi-device scenarios, allocating independent prediction models for each device would result in data fragmentation, significantly reducing the available dataset for each model, while substantially increasing the total number of parameters. This paper presents an innovative method for wind power forecasting: instead of splitting the dataset according to devices or providing independent models for each device, a single model is trained to predict the power generation for each individual wind turbine device. The static characteristics of individual devices have an impact on the power generation patterns. However, a single prediction model cannot differentiate between different devices or take into account the differences in device operating patterns, leading to a loss in forecasting accuracy. To address this issue, this paper utilizes space embedding technology to infer the hidden features of each device and applies it to represent the wind turbines. The essence of space embedding technology is the same as that of neural networks, both of which are derived from causal effects and use gradient descent algorithms to calculate the static attributes that effectively affect the target task. Therefore, this article aims to propose a method that does not rely on expert knowledge and complex modeling processes to obtain the static properties of wind turbine equipment (including inherent equipment features and some long-term climate characteristics that do not change). During the power generation forecasting process, the representation vectors are concatenated with the temporal data and inputted into the neural network model. This approach enables the model to consider both the dynamic historical data and the inherent static characteristics of the devices. Experimental validation shows that the proposed method achieves a superior forecasting performance while reducing the parameter quantity to only 0.74% of the comparative method.

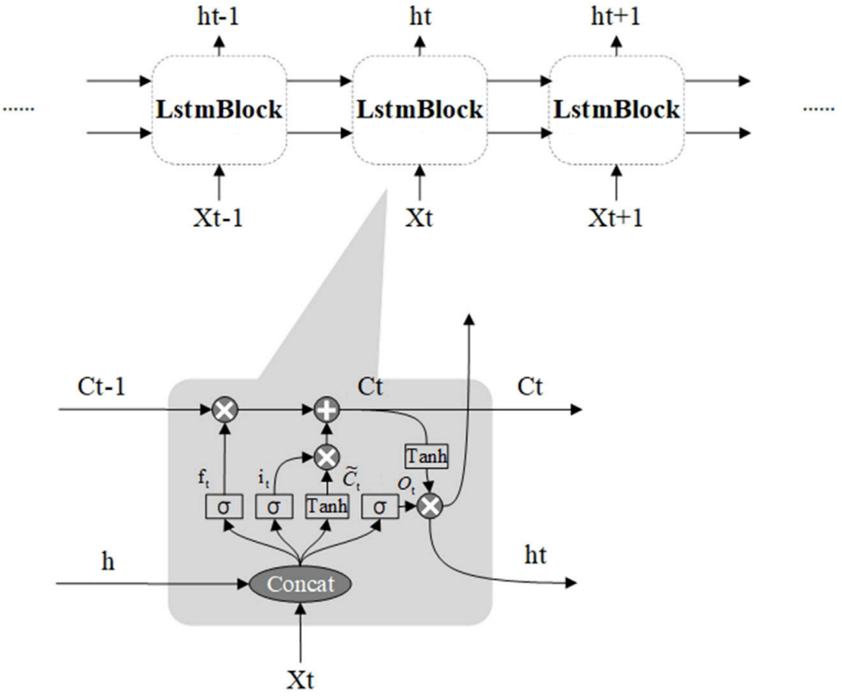


FIGURE 1
Illustration of LSTM layer structure.

The main contributions of this paper are as follows:

(1) Proposed a method that utilizes the complete dataset for model training and employs a single model for wind power forecasting across multiple wind turbines. This method addresses the issue of data dilution and significantly reduces the number of model parameters.
(2) Introduced a space embedding technique specifically designed for wind turbines. This technique is used to represent the impact of hidden static features of the devices on power generation patterns, addressing the issue of predictive performance loss caused by an individual model's inability to differentiate between devices.
(3) The experiments demonstrate that the single-model method using the complete dataset not only significantly reduces the number of parameters but also improves predictive performance. Building upon this foundation, the utilization of wind turbine embedding technology further enhances prediction accuracy. This paper verifies a positive correlation between the dimension of representation vectors and the accuracy of power generation forecasting. However, there is limited improvement in performance when the dimension becomes excessively large.

3 Theoretical background

3.1 Long Short-Term Memory

Long Short-Term Memory (LSTM) is a special type of Recurrent Neural Network (RNN). Compared to traditional RNN, LSTM alleviates the issues of vanishing and exploding gradients in modeling long sequences. When receiving input from the upper layers of the network, the LSTM layer needs to unfold itself horizontally to match the shape of the input data. The data flow mechanism of LSTM makes it naturally suitable for modeling sequential data, but also hinders parallel computation. The diagram below illustrates the data propagation and internal structure of LSTM during horizontal unfolding.

In Figure 1, the LSTM layer consists of multiple blocks, where each block shares the same parameters, and data propagation occurs strictly in linear order. Each LSTM block includes a forget gate f_t, an input gate i_t, and an output gate o_t. The forget gate f_t controls whether historical information in the memory cell should be forgotten. The input gate i_t determines whether the input data X_t should be written into the memory cell. The output gate o_t decides the extent to which information from the previous time step is transmitted to the next LSTM block. The formulas for these three gate units and the memory cell are as follows:

$$f_t = \sigma\left(W_f \cdot [X_t, h_{t-1}] + b_f\right) \tag{1}$$

$$i_t = \sigma(W_i \cdot [X_t, h_{t-1}] + b_i) \tag{2}$$

$$o_t = \sigma(W_o \cdot [X_t, h_{t-1}] + b_o) \tag{3}$$

$$\tilde{C}_t = tanh(W_c \cdot [X_t, h_{t-1}] + b_c) \tag{4}$$

In the formulas, W and b represent the learnable parameters and bias terms for each gate unit. σ stands for applying the *sigmoid* activation function after performing matrix multiplication between the input data and the network parameter matrix. The formula for the *sigmoid* activation function is as follows:

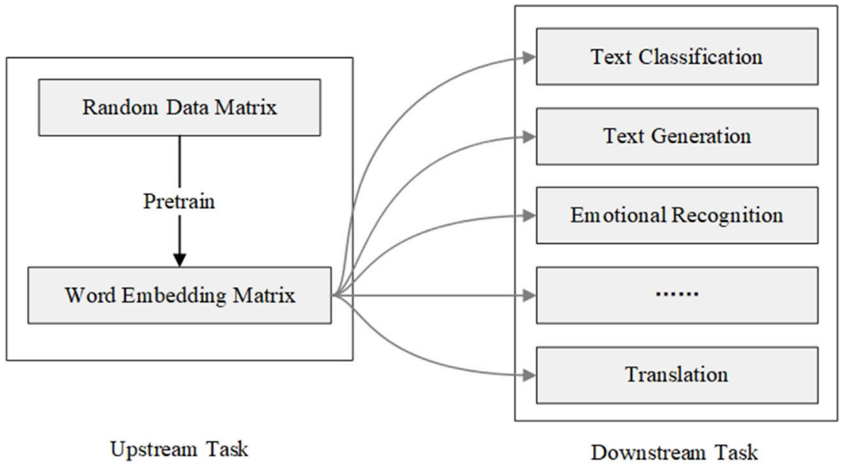


FIGURE 2
Diagram illustrating the association between upstream and downstream tasks.

$$sigmod(x) = \frac{1}{1+e^{-x}} \tag{5}$$

The activation function will map the input data to a value between 0 and 1. The closer the value is to 0, the smaller the influence of the mapped data will be in the subsequent matrix multiplication. The formulas for the cell state C_t and hidden state h_t calculations are as follows:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{6}$$

$$h_t = o_t + \ tanh(C_t) \tag{7}$$

The derivative of C_t with respect to C_{t-1} is represented as $\frac{\partial C_t}{\partial C_{t-1}} = f_t + C_{t-1}\frac{\partial f_t}{\partial C_{t-1}} + \frac{\partial i_t}{\partial C_{t-1}}\tilde{C}_t + i_t\frac{\partial \tilde{C}_t}{\partial C_{t-1}}$, where f_t falls between 0 and 1. Therefore, the internal structure of LSTM can effectively prevent the occurrence of gradient explosion or gradient vanishing.

3.2 Space embedding technology

Space embedding technology is a technique that computes the continuous vector representation of entities in a high-dimensional space. It originated from word embedding in the field of Natural Language Processing (NLP). Typically, space embedding technology evolves the vector representation based on the distribution phenomena or behavioral patterns of entities in specific tasks, evolving random data into high-dimensional vectors with representational capabilities. In the field of NLP, performing space embedding computation is an upstream task. This task is not specific to particular business scenario but rather aims to convert abstract natural language into a more easily processable data format. Conversely, downstream tasks are tailored to specific business scenarios and rely on the representations vector generated by upstream tasks. Figure 2 illustrates the relationship between upstream and downstream tasks.

Regarding the word embedding technology, predicting word distribution tasks are considered upstream tasks, while using evolved word vectors for tasks such as machine translation, sentiment analysis, or named entity recognition is referred to as downstream tasks. The distributional hypothesis proposed by Harris in 1954 serves as the theoretical foundation of word embedding technology. This hypothesis posits that words with similar contexts also have similar meanings and should correspond to similar high-dimensional continuous representation vectors (Harris, 1954). Word embedding technology derives high-dimensional continuous vector representations based on the phenomenon of word distribution. Typically, researchers train a deep neural network to predict word distributions and employ the gradient descent algorithm to update network parameters and word vector matrices simultaneously. After training the neural network until convergence, the high-dimensional vector representations corresponding to words have evolved from their random initial states to appropriate states. These representations can be used to describe the hidden features associated with each word. Representative models of this technology include word2vec, Elmo (Sarzynska-Wawer et al., 2021), Bert (Devlin et al., 2018).

After word embedding technology, embedding techniques have further developed into graph embedding for graph structures (Grover and Leskovec, 2016), position embedding for sequential order (Vaswani et al., 2017), and data embedding architecture known as data2vec for multi-modal data (Baevski et al., 2022), among other techniques or approaches.

It is crucial to recognize that power fluctuation patterns are influenced by both dynamic factors, such as changes in internal turbine states and meteorological conditions, and the static attributes of the turbine equipment. The design of turbine blades and the control strategy significantly affect energy capture and conversion efficiency, while the geographical and climatic context of the equipment directly impacts power generation fluctuations. Additionally, the static attributes of the turbine have a lasting impact on its power generation patterns, making its power output not entirely dependent on real-time internal and meteorological data. Even turbines of the same types may exhibit differences in their power generation patterns due to variations in environment, layout, and maintenance conditions.

However, characterizing the static features of turbines faces three challenges:

1. Although common SCADA datasets include a wealth of dynamic data from the operational phase, hey do not record details on common static features such as blade shape, wind adaptability, altitude, climate, and geographical environment.
2. It is challenging to analyze the correlation between a single static feature and power fluctuations, which limits our ability to discern and rank the relationship between static features and power fluctuations.
3. Analyzing the correlation between a single static feature and power fluctuations is challenging, which limits researchers' ability to distinguish and rank the extent of correlation between static features and power fluctuations.

In the absence of effective features, researchers can predict the distribution of entities by training models to ensure the models capture the static hidden features of the entities. The Word2vec model infers the hidden semantic features of words based on the order of word distribution, and the graph embedding model node2vec infers the hidden features of nodes based on their connectivity structure in the graph. To overcome the challenges mentioned above and effectively capture the static features of turbine equipment, it is feasible to consider inferring potential static features based on the turbine power patterns.

We considered training models to infer hidden turbine features during the process of forecasting wind speed or direction. However, the static feature vectors generated in this way only reflect the climatic characteristics and do not represent the equipment characteristics (such as conversion efficiency, wind adaptability, or fault conditions). Constructing correlation graphs based on the similarity of power fluctuation patterns between turbines and using graph embedding techniques can also capture the static features of turbine nodes, though the static features obtained this way tend to represent inter-device correlations more. To comprehensively characterize the static factors affecting the equipment's power pattern, we set both upstream and downstream tasks as the same, namely, the turbine power forecasting task. The representation vector generated based on this describes the static factors that affect the power fluctuation pattern. The vector semantics are not limited to climatic factors, equipment models, operational strategies, etc., but may also include other related factors that have not been researched but have a tangible correlation.

This paper presents an embedding technique that does not rely on entity distribution correlation. Specifically, when performing wind power forecasting tasks using neural networks, this paper utilizes the gradient descent algorithm to iteratively evolve randomly initialized vectors into high-dimensional representations of the wind turbine's hidden static factors. Traditional space embedding techniques rely on predicting the distribution patterns of entities. However, in our method, we generate representation vectors by predicting the target attribute, i.e., Active power, directly. This method is not only applicable for generating representation vectors of entities without specific distribution phenomena, such as turbine generators, but also directly discovers hidden features that are highly correlated with the target attributes. The representation vectors generated by embedding technology are derived from longer segments of the training dataset, which allows them to encompass features from a wider time span. In contrast, the wind power forecasting task only accesses data from a limited number of historical time steps. The representation vectors provide additional evidence for the forecasting task. Subsequent experiments evaluated the impact of different dimensional representation vectors on the forecasting model, and verified the capability of embedding techniques to enhance the performance of the forecasting models.

4 Methodology

4.1 Task description

Wind power forecasting tasks fall into the category of time series forecasting tasks, which require models to predict future time steps based on historical time step data. Each time step corresponds to a sampling for factors such as turbine power, wind attributes, and internal device states. Typically, such tasks involve input data $X \in R^{K \times T}$, where K and T represent the number of historical time steps and the number of features per time step, respectively. After being processed by the forecasting model, the model's output is denoted as $Y \in R^{N \times T}$, where N is the specified number of forecasting steps for the task. Specifically, in the scenario of single-property forecasting tasks, the model's output is $Y \in R^{N \times 1}$. In time series forecasting tasks, we aim for minimal discrepancy between the model's predictions and actual measurement values.

In the proposed method, each wind turbine contains a vector $h_i \in R^M$ to represent its own characteristics, where M represents the number of attributes contained in the representation vector. The representation vector of the wind turbine will be concatenated with the dynamic temporal data and sent as part of the input data into the model. In this scenario, the input data of the model is $X \in R^{K \times (T+M)}$, and the model output is $Y \in R^{N \times 1}$.

4.2 Overview of method

In the proposed method of this paper, the wind power forecasting task is decomposed into two tasks: an upstream embedding task to infer the hidden features of wind turbines, and a downstream task to forecast power generation based on the hidden features.

The upstream and downstream tasks are not completely independent. As shown in Figure 3, the same data processing method and model architecture are used for model training in both the upstream and downstream tasks. The hidden static features generated by the upstream task are used as additional features, which will be concatenated with the historical time steps in the downstream task, and inputted into the LSTM model. In the proposed method, after the input data pass through the LSTM layer, Dropout layer, and linear projection layer, only the data representing future time steps is used as output, while the content representing historical time steps is discarded. The discarded portion does not contribute to the calculation of the loss function and does not have a positive effect on the optimization of the model parameters.

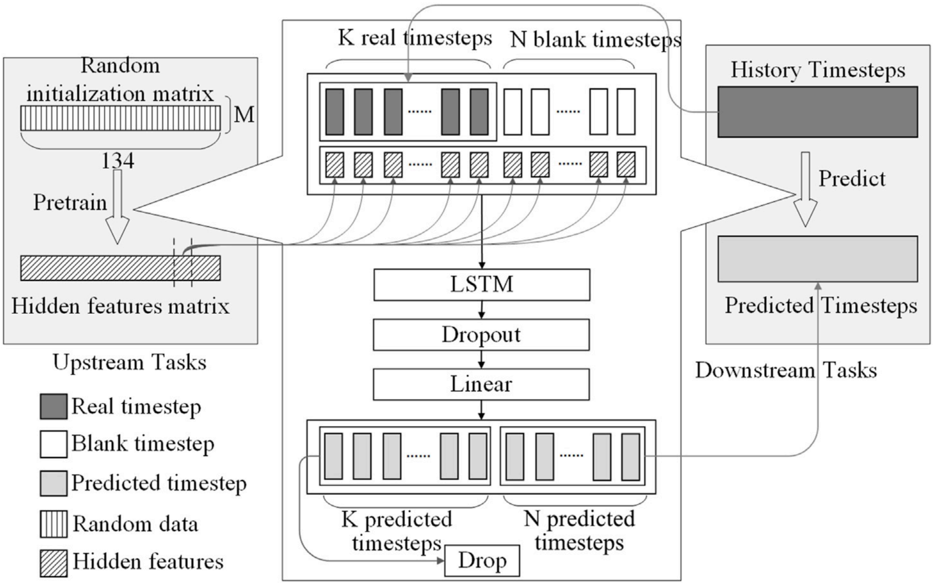


FIGURE 3
Overview of the method architecture.

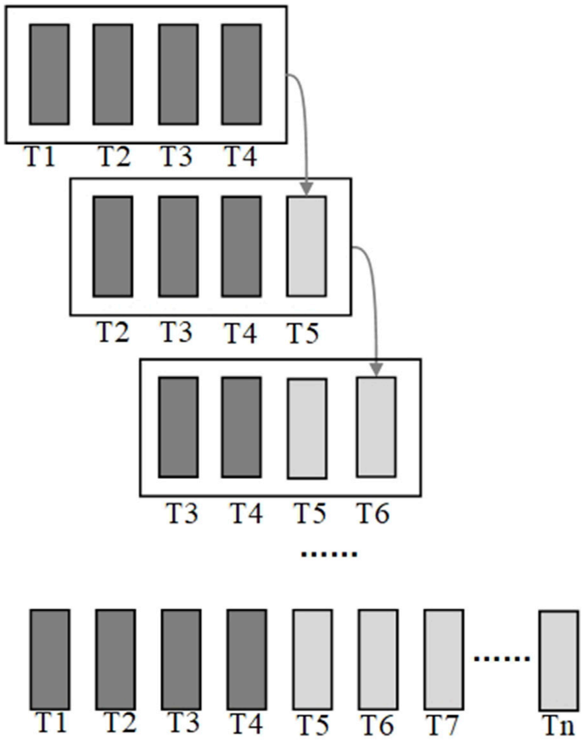


FIGURE 4
Diagram of autoregressive prediction method.

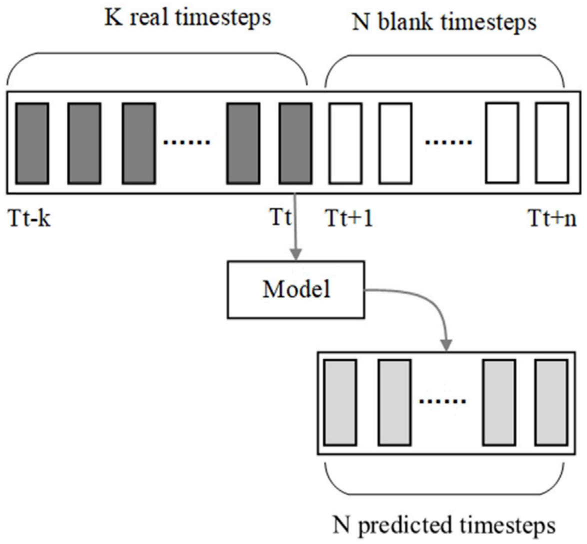


FIGURE 5
Diagram of the novel prediction method.

4.2.1 Concatenation of blank time steps

In time series forecasting tasks, historical data is inputted into a neural network model, which then generates future data as output. In classical scenarios, researchers commonly utilize a sliding window approach to predict a limited number of time steps. Taking Figure 4 as an example, 4 previous time steps are used to forecast 1 subsequent time step, resulting in the generation of an entire time-series through multiple autoregressive iterations.

This paper argues that the method of the sliding window results in wastage of computational resources and time, as it requires inputting K historical time steps into the model for each prediction, and a long sequence needs multiple iterations to be fully generated. Therefore, this paper concatenates additional blank data with the historical time step to align the output format with the expected format. As shown in Figure 5, the prediction model needs to output data for N

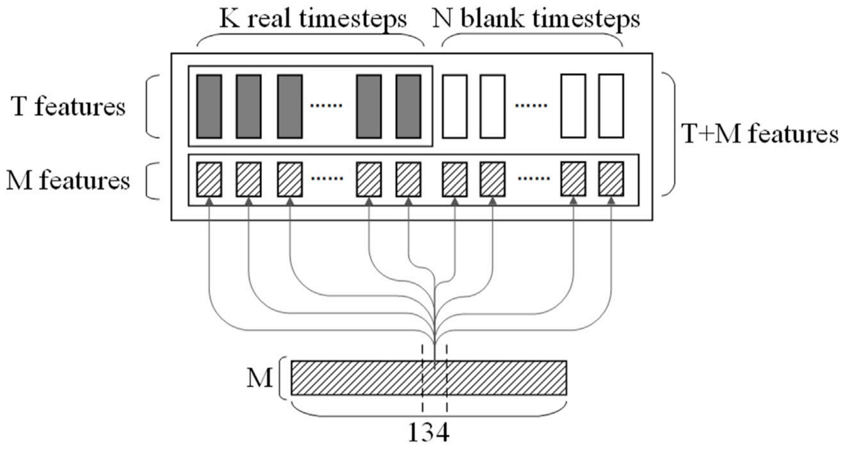


FIGURE 6
Replication and Concatenation process of hidden feature vectors.

future time steps, with the number of blank time steps matching the desired output.

Due to the fact that LSTM can propagate data through the cell state C_t and hidden state h_t when unfolded, inputting blank time steps into the LSTM model does not cause interruption in information transmission.

4.2.2 Replication and concatenation of hidden features

The proposed hidden feature space embedding method aims to extract the hidden static features of the wind turbine devices thereby uncovering the latent factors that influence the operation patterns of the wind turbines.

As shown in Figure 6, in the process of pre-training, the hidden feature vector $h_i \in R^M$ corresponding to wind turbine i will be replicated $K + N$ times. It is concatenated directly with historical time steps $X_i^h \in R^{K\times T}$ and blank time steps $X^b \in R^{N\times T}$, forming an input matrix $Input^{(K+N)\times(T+M)}$ with a shape of $(K + N) \times (T + M)$, which is then inputted into the LSTM model. This paper adopts the unique spatial dynamic wind power forecasting dataset, SDWPF (Zhou et al., 2022), provided by Longyuan Power Group Co., Ltd. This dataset contains a total of 134 wind turbines, so the hidden feature embedding matrix $H \in R^{134\times M}$ contains the representation vectors of 134 turbines. It is essential to ensure that the hidden feature vector h_i and the historical time step X_i^h are from the same turbine.

The hidden features, as input data, participate in computation and obtain corresponding gradients through a backward propagation process. Subsequently, multiple rounds of iteration are performed using the gradient descent algorithm. The randomly initialized vectors gradually evolves into representation vectors that capture the hidden static features of the wind turbine devices. The algorithmic procedure is illustrated in Algorithm 1.

4.3 Evaluation metrics

This paper employs four performance evaluation metrics for forecasting models: Mean squared error (*MSE*), Mean Absolute Error (*MAE*), Pearson correlation coefficient (*Corr*), and coefficient of determination (R^2). The formulas for calculating these metrics are as follows:

$$MSE = \frac{1}{n}\sum_{1}^{n}\left(\hat{y}_i - y_i\right)^2 \tag{8}$$

$$MAE = \frac{1}{n}\sum_{1}^{n}\left|\hat{y}_i - y_i\right| \tag{9}$$

$$Corr = \frac{Cov(\hat{y}, y)}{\sqrt{Var(\hat{y}) \times Var(y)}} = \frac{\sum_{i}^{n}\left(\hat{y}_i - \bar{\hat{y}}\right) \times \left(y_i - \bar{y}\right)}{\sqrt{\sum_{1}^{n}\left(\hat{y}_i - \bar{\hat{y}}\right)^2 \times \sum_{1}^{n}\left(y_i - \bar{y}\right)^2}} \tag{10}$$

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{1}^{n}\left(\hat{y}_i - \bar{y}\right)^2}{\sum_{1}^{n}\left(y_i - \bar{y}\right)^2} \tag{11}$$

In the above equations, ŷ represents the predicted values generated by the model, while y represents the measured values. The metrics of Mean Squared Error (*MSE*) and Mean Absolute Error (*MAE*) measure the discrepancy between the predicted values and the measured values, where smaller values indicate better performance.The Correlation (*Corr*) metric describes the degree of correlation between the predicted sequence and the actual sequence, with its value ranging from −1 to 1. A larger value indicates a stronger positive correlation, while a smaller value indicates a stronger negative correlation. The R-squared (R^2) metric quantifies the fitting degree of the predicted values to the actual values, with a value ranging between 0 and 1. Ideally, it should approach 1.

The forecasting model is prone to generating a straight line at the mean of the actual values as the prediction result, which exhibits a poor correlation with the actual values. Although the *MSE* and *MAE* metrics have small values in this case, the R-squared (R^2) metric approaches zero, indicating the model's limited ability to capture volatility.

Furthermore, this paper introduces a custom comprehensive evaluation metric called Mean Standardized Score (*MSS*). It is calculated using the following formula:

$$Score(x) = Corr(x) + R^2(x) - MSE(x) - MAE(x) \tag{12}$$

TABLE 1 Overview of dataset contents.

Index	0	1	2	……
TurbID	1	1	1	……
Day	1	1	1	……
Tmstamp	0:00	0:10	0:20	……
Wspd	12.23	11.58	11.21	……
Wdir	−0.83	−3.32	−1.38	……
Etmp	29.08	29.01	29.17	……
Itmp	41.9	42.01	42.24	……
Ndir	−23.73	−23.7	−28.84	……
Pab1	1.07	1.06	1.04	……
Pab2	1.07	1.06	1.04	……
Pab3	1.07	1.06	1.04	……
Prtv	−0.21	−0.25	−0.25	……
Patv	1549.53	1549.71	1534.77	……

$$MSS = Mean\left(\frac{Score(x) - Mean(Score(x),0)}{Std(Score(x),0)}, 1\right) \tag{13}$$

The calculation of the *MSS* metric consists of three steps:

(1) First, the original four evaluation metrics are summed according to the principle of adding positive gains and subtracting negative gains. This yields the total score of each model across all criteria.
(2) Then, the scores for each task are normalized. Due to the differences in task difficulty, the dimensions of the scores are inconsistent, resulting in tasks with higher scores having a greater impact on the final evaluation. Through the normalization operation, we ensure that the scores of each task have the same dimension, eliminating the influence of task difficulty on the final evaluation. Here, “$Mean(Score(x),0)$” and “$Std(Score(x),0)$”"mean taking the mean and standard deviation of the scores for different models under the same task.
(3) Finally, the scores of each task under the same model are averaged to determine the overall score of the model.

5 Experiment and analysis

5.1 Data preprocessing

This paper utilizes the unique spatial dynamic wind power forecasting dataset, SDWPF (Zhou et al., 2022), provided by Longyuan Power Group Co., Ltd. The dataset spans a period of 184 days and includes sampled data from 134 wind turbines. The SCADA system compiles the collected data at 10-min intervals, with each wind turbine accounting for 184 (days) × 24 (hours) × 6 (intervals), resulting in a total of 26,496 time steps. The entire dataset contains 26,496 × 134 (units), summing up to 3,550,464 time steps. Each time step is associated with 13 dynamic features, including data from internal features of the wind turbine equipment as well as climate-related data.

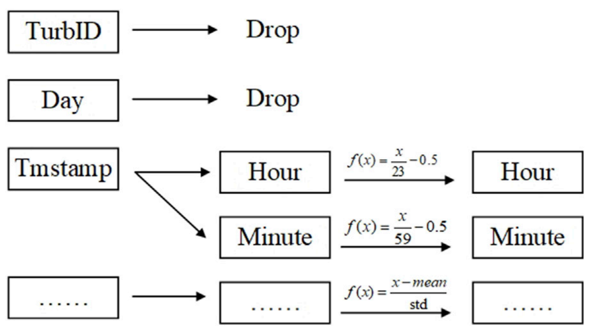


FIGURE 7
Diagram of data feature processing and normalization scheme.

The content and format of the dataset are presented in the Table 1.

During the data preprocessing stage, the following steps were conducted on the dataset in this paper:

1. The feature of turbine ID was discarded. This paper employed space embedding technique to obtain a multi-dimensional vector representation of the hidden features of turbines. This method can provide richer turbine feature information for the model, whereas the turbine ID does not contain descriptive information about the static features of the device.
2. The feature of operating days was discarded. This feature is used to identify the sequential relationship between data. However, recurrent neural networks have the inherent ability to model time series data. Additionally, the data in the test set and validation set belong to future data, and this feature differs from the training set in terms of mean and variance, which can affect the model’s judgment. Therefore, this paper chooses to remove this feature.
3. Recoding the time feature. The format of this feature is “hour: minute,” and its content is not numerical, making it unsuitable for direct input into the model. In this paper, the timestamp was split to create two new dimensions. We hope the model can recognize the pattern of the relationship between power generation and the time variation within a day.

Figure 7 shows the data preprocessing process in a more intuitive way after preprocessing, the attributes of the dataset and their descriptions are shown in the Table 2.

This paper explores the correlation between multidimensional features and active power (Prtv), as shown in the heat map.

As depicted in Figure 8, a strong correlation is evident between active power (feature 10) and wind speed (feature 1), located at coordinates (1, 10). Additionally, there is an insignificant correlation between the active power (feature 10) and the temperature inside the turbine nacelle (feature 4), corresponding to coordinates (4, 10). Furthermore, the heatmap exhibits a strong negative correlation between the active power (feature 4) and the pitch angle of the three blades (feature 6, 7, and 8), corresponding to coordinates (7, 10), (7,

TABLE 2 Display of dataset features and descriptions.

Column	Column name	Description
1	Wspd (m/s)	The wind speed recorded by the anemometer
2	Wdir (°)	The angle between the wind direction and the position of turbine nacelle
3	Etmp (℃)	Temperature of the surrounding environment
4	Itmp (℃)	Temperature inside the turbine nacelle
5	Ndir (°)	Nacelle direction, i.e., the yaw angle of the nacelle
6	Pab1 (°)	Pitch angle of blade 1
7	Pab2 (°)	Pitch angle of blade 2
8	Pab3 (°)	Pitch angle of blade 3
9	Prtv (kW)	Reactive power
10	Patv (°)	Active power (target variable)
11	Hour	Hour of the record
12	Minute	Minute of the record

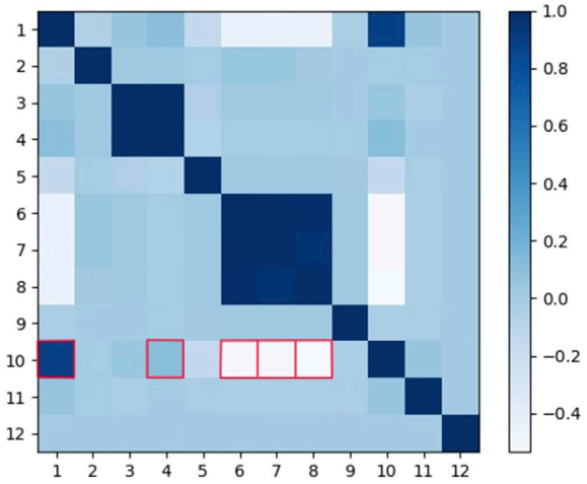


FIGURE 8
Heat map for attribute correlation analysis of the dataset.

10) and (8, 10) respectively. The time feature (feature 10 and 11) shows few correlation with other features.

5.2 Comparison between the improved method and traditional methods

In this research, 80% of the dataset is used as the training set, while the remaining 20% is allocated for the validation and test sets. Apart from the time feature, each dimension of the features is normalized using mean and standard deviation. The proposed method in this paper for multi-device power generation forecasting is not dependent on a specific neural network model, but can complement the improved methods of neural network models. Choosing an appropriate neural network model can improve the accuracy of power generation forecasting tasks in specific application scenarios. Conventional neural network models include but are not limited to Transformer and its variants, as well as recurrent neural network models such as LSTM or GRU. We adopts the LSTM as the experimental object and compares the performance difference between the traditional method and the single-model method that integrates space embedding technology. Table 3 displays the best results obtained from three experiments under the same conditions. Bold in the table is used to highlight the best results under the same experimental conditions, and the following table is the same.

The "Multi LSTM" method in the table does not utilize the space embedding technique to obtain device representation vectors. Instead, it assigns a independent LSTM model to each turbine device for power generation forecasting. Since each device has an independent model, the dataset is also divided by devices. In the "Single LSTM & 8 Hidden Features" method, we use a single model and an undivided training set to forecasting the power generation of 134 turbine devices. At the same time, we introduce an 8-dimensional vector to represent the hidden static features of the turbine devices. The "ARIMA" scheme employs the classical statistical model ARIMA for power generation forecasting. The data in the table represents the average performance of all turbines' predictions.

Based on the table results, it can be observed that the forecasting model using hidden features has fewer model parameters and demonstrates significant advantages across all four metrics. Particularly noteworthy is the 23.6% improvement in the MSE metric for ultra-short-term forecasting (1 h, 6 time steps). Compared to traditional approaches that merely input historical power data into neural network models, the method presented in this paper utilizes hidden features to represent the impact of wind turbine static attributes on their power generation patterns, enabling the forecasting model to make more accurate predictions based on the inherent properties of the equipment. Additionally, in traditional methods, since neural network models cannot distinguish which

TABLE 3 Performance comparison between improved method and traditional method.

Model	Params (Million)	Metrics	Horizon						Count
			6	12	24	48	96	144	
ARIMA	-	MSE	0.8245	0.8252	0.8262	0.8283	0.8339	0.8398	0
		MAE	0.7296	0.7295	0.7298	0.7300	0.7313	0.7332	
		Corr	0.0003	-1e-5	−0.0010	−0.0021	−0.0027	−0.0033	
		R^2	0.0392	0.0556	0.0778	0.1137	0.1410	0.1470	
Multi LSTM	6.8	MSE	0.2149	0.2983	0.4646	**0.5730**	0.6811	0.7489	2
		MAE	0.3140	0.3753	0.4566	0.5513	0.6260	0.6421	
		Corr	0.8601	0.8000	0.6912	**0.5196**	0.3174	0.2688	
		R^2	0.6382	0.5413	0.4314	0.2734	0.1556	0.1172	
Single LSTM & 8 Hidden Features	**0.104**	MSE	**0.1584**	**0.2572**	**0.4033**	0.5953	**0.6596**	**0.6703**	**23**
		MAE	**0.2552**	**0.3313**	**0.4376**	**0.5428**	**0.5730**	**0.5981**	
		Corr	**0.8957**	**0.8284**	**0.7085**	0.5120	**0.3990**	**0.3322**	
		R^2	**0.7819**	**0.6262**	**0.5834**	**0.3956**	**0.2708**	**0.2063**	

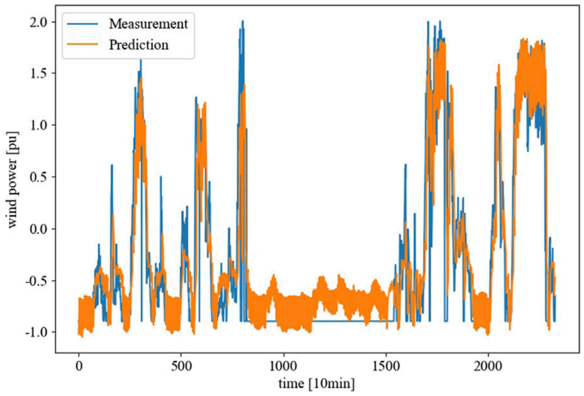


FIGURE 9
Forecasting performance of traditional method on wind turbine #1.

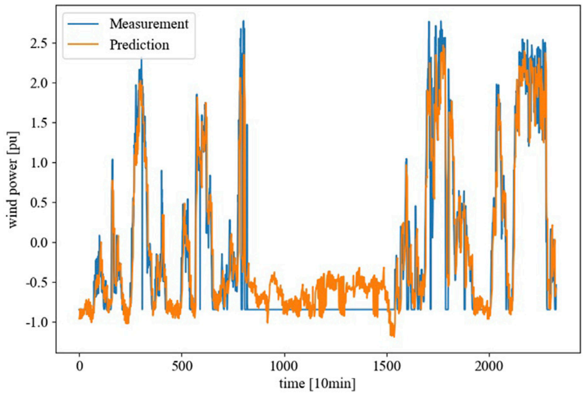


FIGURE 10
Forecasting performance of the Single Model method with 8 Hidden Features on wind turbine #1.

device the time series data originates from, separate forecasting models are assigned to each device. The dataset consists of 134 turbines, and the parameter size of the traditional method would reach 134 times that of the proposed method. This leads to wastage of computational resources without yielding significant performance improvements. However, the approach in this paper distinguishes devices based on hidden features which represent the differences between turbines, allowing the entire dataset to be used for model training. The augmentation of data also supports enhancements in forecasting accuracy. Additionally, due to the high randomness of the data, the performance of the statistic-based model ARIMA is not satisfactory.

We selected Wind Turbine NO. 1 and uses a 1-h ahead prediction to assess the short-term effectiveness of the forecasting model. By comparing Figure 9 and Figure 10, two advantages of the model trained with hidden features and the complete dataset can be observed:

- Higher accuracy: The results shown in Figure 10 demonstrate a stronger correlation between the orange line and the blue line. This observation aligns with the model's superior performance over the traditional models in terms of correlation (*Corr*) and determination coefficient (R^2) indicators.
- Higher certainty: Compared to Figure 9, the predicted values in Figure 10 exhibit smaller short-term fluctuations.

It should be noted that, the measured values from time step 700 to 1600 in the graph are displayed as 0, which is actually a result of data set incompleteness and filled with 0 instead of real measurements.

TABLE 4 Comparative results of models with different dimensions of hidden features on four criteria.

Dimension	Metrics	Horizon						Count
		6	12	24	48	96	144	
0	MSE	0.1641	0.2586	**0.3899**	0.6156	0.6622	0.7032	3
	MAE	0.2581	0.3396	**0.4253**	0.5724	0.5906	0.6272	
	Corr	0.8935	0.8234	**0.7175**	0.4906	0.3794	0.2774	
	R^2	0.7154	0.6597	0.5897	0.3892	0.2259	0.1770	
4	MSE	0.1614	0.2652	0.3974	**0.5935**	0.7026	0.7091	2
	MAE	0.2547	0.3390	0.4454	0.5534	0.5991	0.6178	
	Corr	0.8938	0.8186	0.7121	0.5166	0.3416	0.2631	
	R^2	0.7622	**0.7140**	0.4774	0.4146	0.2737	0.1576	
8	MSE	**0.1584**	**0.2572**	0.4033	0.5953	**0.6596**	**0.6703**	**14**
	MAE	**0.2552**	**0.3313**	0.4376	**0.5428**	**0.5730**	**0.5981**	
	Corr	**0.8957**	**0.8284**	0.7085	0.5120	0.3990	**0.3322**	
	R^2	**0.7819**	0.6262	**0.5834**	0.3996	0.2708	0.2063	
12	MSE	0.1627	0.2645	0.3994	0.6017	0.6549	0.7102	5
	MAE	0.2554	0.3430	0.4388	0.5437	0.5749	0.6162	
	Corr	0.8934	0.8188	0.7111	**0.5300**	**0.4057**	0.3016	
	R^2	0.7380	0.6781	0.5473	**0.4266**	**0.2783**	**0.2661**	

5.3 Performance evaluation on hidden feature dimensions

The above experiment compared the performance between the traditional method and the improved method that utilizes 8-dimensional hidden features. We will further examine the impact of additional dimensions of hidden features on the efficacy of the forecasting model in this research.

There are similarities in the power generation patterns among different wind turbines. Therefore, training the model using data from other wind turbines can enhance its emphasis on the conversion pattern between climate factors and power generation, thus reducing the risk of overfitting. The experiment demonstrates that even when the dimension of device representation vectors is 0, the forecasting performance of a single model is still superior to the traditional method of assigning independent models to each device. This implies that the negative impact caused by the inability of the model to differentiate between devices is smaller than the positive gain achieved through dataset augmentation. This phenomenon verifies the similarity in power generation patterns among wind turbine devices. In addition, the advantages of the forecasting model are more significant when the dimension of the device representation vector is higher. This phenomenon confirms the existence of differences in the power generation patterns among different devices. The information contained in the device representation vector provides additional features to the forecasting model, enabling more accurate predictions.

From Table 4, it can be observed that different dimensions of hidden features exhibit varying gain effects on the model. Among them, the 8-dimensional hidden features contribute the highest gain to the model. It is worth noting that Table 4 only explicitly compares the optimal performance under different conditions, without considering the negative impact of non-optimal attributes (non-bolded fields) on the performance of the method. Therefore, in order to compare the relative differences in model performance under different scenarios, we adopts the comprehensive scoring criterion *MSS*, aiming to comprehensively evaluate the methods.

TABLE 5 Score table of models with different dimensions of hidden features under the *MSS* criterion.

Model	*MSS*
Multi LSTM	−1.69729642
Single LSTM & 0 Hidden Features	0.09489346
Single LSTM & 4 Hidden Features	0.14731688
Single LSTM & 8 Hidden Features	**0.72910813**
Single LSTM & 12 Hidden Features	0.72597795

As shown in Table 5, it can be observed that although the 12-dimensional hidden feature model does not perform as well as the 8-dimensional model in terms of the number of optimal score quantity, its negative impact on non-optimal scores is less severe compared to the 8-dimensional hidden feature model. This results in a small difference in overall scores between the two models.

The above performance differs from the space embedding task in natural language processing (NLP) tasks. In NLP tasks, word vectors usually have higher dimensions (512–1024 dimensions), while the dimensions of the hidden features of wind turbines are

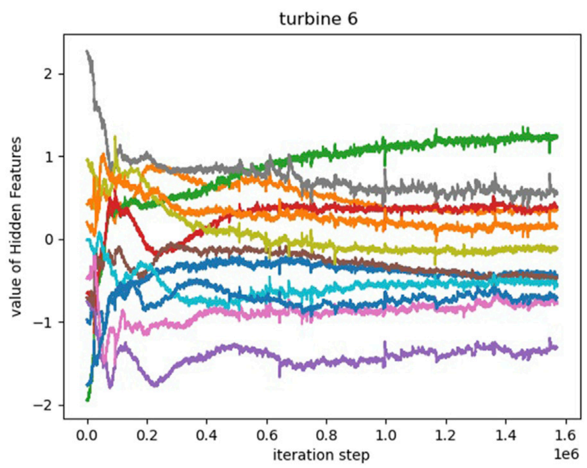


FIGURE 11
The 12-dimensional hidden features of Wind Turbine 6 converge to 8 points.

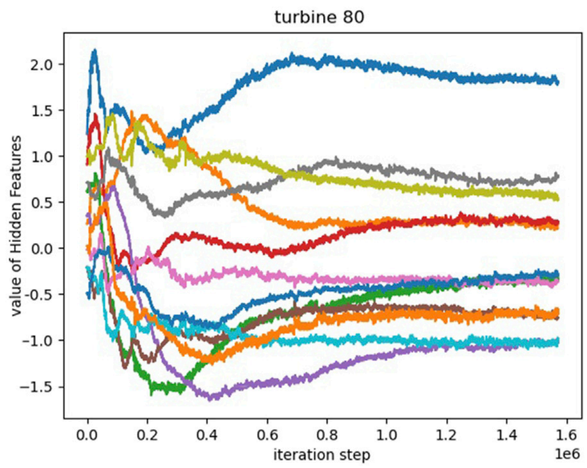


FIGURE 12
The 12-dimensional hidden features of Wind Turbine 80 converge to 7 points.

much lower than this range. 8-bit binary numbers can encode 256 entities, while the representation vectors generated by space embedding technique exhibit excellent representational capacity. Given that there are only 134 wind turbines, the number of hidden features should not differ significantly from 8 ($[\log_2^{134}] + 1$). The gain of the model exhibited a turning point when the dimension of the representation vector was 8. This indicates that there is a certain degree of coupling between the static factors among wind turbines and their power generation patterns, and the number of independent influential factors is not substantial. The following figure demonstrates the iterative process of inferring the 12-dimensional representation vectors for wind turbines.

From Figure 11 and Figure 12, it can be observed that after being initialized with a normal distribution, some representation vectors exhibit the convergence of multiple features to the same point. This phenomenon confirms the existence of certain coupling between hidden static feature representations. Therefore, it can be concluded that a 12-dimensional hidden feature is not the most compact embedding representation for wind turbines. Excessive hidden features not only increase computational burden, but may also lead to overfitting of the prediction model. This paper argues that in this task scenario of distributed wind farm power generation forecasting, the number of hidden features should not be excessive. The results of the experiment demonstrate the effectiveness of static features in assisting forecasting, indicating that detailed features that affect the target task can be inferred to a certain extent without relying on specific expert knowledge and on-site detail modeling. However, the correspondence and representation effect between hidden features and real features in the on-site environment still need further research.

The hyperparameters used in the forecasting model for the experiment will be displayed in Table 6.

6 Conclusion

This paper investigates the problem of multi-device power generation forecasting in distributed power grid scenarios and proposes a forecasting method that combines space embedding techniques from the field of natural language processing. This method utilizes space embedding techniques to uncover hidden static features of each power generation device and uses these features as device identifiers. This allows a single model to distinguish between devices and accurately predict the power generation of multiple devices. The proposed method is independent of experimental models and does not rely on specific neural network architectures. It complements the improvements made in neural network algorithms. The experiments have shown that the proposed forecasting method, which integrates space embedding technology, not only significantly reduces the number of model parameters but also achieves higher prediction accuracy. The experimental results also indicate that the gain of representation vectors varies across different dimensions. The gain utility becomes less apparent, When the dimension of device representation vector is excessively large in the scenario described in this paper.

The proposed method in this paper focuses on using a single model to perform forecasting tasks for devices within the entire distributed power grid. However, there are several aspects that can be improved in the future:

(1) This paper confirms the compatibility of LSTM and space embedding technology. The subsequent investigation should involve considering the use of variants of the Transformer architecture to replace the classical LSTM model and verify the compatibility of space embedding technology with Transformer models in time series forecasting tasks.
(2) The representation vector of the wind turbine device is static data and does not vary with the time series, which is different from the temporal data. Currently, we concatenate the representation vector with the time steps data directly. In the future, we will consider using a more robust approach to integrate the device representation vector with the temporal data.

TABLE 6 Display of experimental parameter.

Parameter	Value	Explanation
Batch size	32	
Learning rate	1e-4	
Activation function	leakyRelu	Used between LSTM and Dropout layers
negative_slope	0.01	leakyRelu parameter
LSTM hidden_size	32	
Dropout	0.1	
Linear input size	32	
Early stopper patience	3	The number of times the valid error is allowed to increase

(3) The proposed prediction method integrates space embedding technology without relying on specific neural network architectures. Combining space embedding technology with models specific to the business scenario may lead to even better performance. In future work, we consider incorporating convolutional operations and attention mechanisms into the neural network to further enhance the accuracy of forecasting models.

```
Input:
$H = \{h_1, h_2, ..., h_{134}\}$, hidden features Matrix, includes hidden features of 134 turbines.
$X = \{x_1^1, x_2^1, ..., x_K^1, x_1^2, x_2^2, ..., x_K^2, ..., x_1^{134}, x_2^{134}, ..., x_K^{134}\}$, Historical data of 134 wind turbines, each sequence containing K time steps.
$Y = \{y_1^1, y_2^1, ..., y_N^1, y_1^2, y_2^2, ..., y_N^2, ..., y_1^{134}, y_2^{134}, ..., y_N^{134}\}$, Future data of 134 wind turbines, each sequence containing N time steps.
$lr$, the learning rate
Output: hidden features Matrix $H$.
for epoch in range (1, 10), do:
    for i in range (1, 134), do:
        $input = concat(repeat(h^i, K, dim = 0), X^i, dim = 1)$
        $\widehat{y^i}$ = LSTM (input)
        loss = MSELoss($\widehat{y^i}, y^i$)
        loss.backward
        LSTM.parameters = LSTM.parameters − LSTM.parameters.grads × $lr$
        $h_i = h_i - h_i.grad \times lr$
    end for
end for
return $H$
```

Algorithm 1. Turbine Embedding.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://aistudio.baidu.com/aistudio/competition/detail/152/0/datasets.

Author contributions

JM: Conceptualization, Methodology, Writing–review and editing. KX: Software, Writing–original draft. DW: Data curation, Investigation, Writing–review and editing. YL: Investigation, Methodology, Writing–review and editing. JZ: Investigation, Writing–review and editing. YQ: Methodology, Resources, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was funded by Key Technologies for Intelligent Monitoring and Analysis of Equipment Health Status of High of Science and Technology Innovation Team in College of Hunan Province, grant numbers 2023-233 (Xiang Jiao Tong), Natural Science Foundation of Hunan Province, grant numbers 2022JJ50002 and 2021JJ50049, Science and Technology Innovation Leading Plan Project of Hunan High Tech Industry, grant number 2021GK4008.

Acknowledgments

In my upcoming paper, I would like to take this opportunity to express my deepest gratitude to those individuals and institutions who have provided crucial support and assistance throughout my research journey. First and foremost, I want to sincerely thank my advisor. Thank you for providing valuable guidance, advice, and encouragement, helping me find the right direction in my research and offering unwavering support during times of difficulty. Furthermore, I would like to express my gratitude to the university for providing an excellent learning and research environment. Here, I had access to laboratory equipment and resources that were essential for my research work. Lastly, I would like to thank all my friends, classmates, and family members who have supported and assisted me throughout my research journey.

Conflict of interest

Authors DW and YL were employed by the CRRC Zhuzhou Electric Locomotive Research Institute Co., Ltd. Authors YQ were employed by the Guiyang Aluminum Magnesium Design and Research Institute Co., Ltd., Hunan Tianqiao Jiacheng Intelligent Technology Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fenrg.2024.1346369/full#supplementary-material

References

Ariyo, A. A., Adewumi, A. O., and Ayo, C. K. (2014). "Stock price prediction using the ARIMA model," in 2014 UKSim-AMSS 16th international conference on computer modelling and simulation, Cambridge, UK, 26-28 March 2014 (IEEE), 106–112.

Baevski, A., Hsu, W.-N., Xu, Q., Babu, A., Gu, J., and Auli, M. (2022). "Data2vec: a general framework for self-supervised learning in speech, vision and language," in *International conference on machine learning* (PMLR), 1298–1312.

Cho, K., van Merriënboer, B., Bahdanau, D., and Bengio, Y. (2014). "On the properties of neural machine translation: encoder–decoder approaches," in *Proceedings of SSST-8, eighth workshop on syntax, semantics and structure in statistical translation*, 103–111.

Chung, J., Gulcehre, C., Cho, K. H., and Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv Prepr. arXiv:1412.3555.*

Contaxis, G. C., and Kabouris, J. (1991). Short term scheduling in a wind/diesel autonomous energy system. *IEEE Trans. Power Syst.* 6 (3), 1161–1167. doi:10.1109/59.119261

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: pre-training of deep bidirectional transformers for language understanding. *arXiv Prepr. arXiv:1810.04805.*

Fang, S., and Chiang, H.-D. (2016). A high-accuracy wind power forecasting model. *IEEE Trans. Power Syst.* 32 (2), 1–1590. doi:10.1109/tpwrs.2016.2574700

Grover, A., and Leskovec, J. (2016). "node2vec: scalable feature learning for networks," in *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, 855–864.

Harris, Z. S. (1954). Distributional structure. *Word* 10 (2-3), 146–162. doi:10.1080/00437956.1954.11659520

Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* 9 (8), 1735–1780. doi:10.1162/neco.1997.9.8.1735

Hu, J., Heng, J., Wen, J., and Zhao, W. (2020). Deterministic and probabilistic wind speed forecasting with de-noising-reconstruction strategy and quantile regression based algorithm. *Renew. Energy* 162, 1208–1226. doi:10.1016/j.renene.2020.08.077

Jung, J., and Broadwater, R. P. (2014). Current status and future advances for wind speed and power forecasting. *Renew. Sustain. Energy Rev.* 31, 762–777. doi:10.1016/j.rser.2013.12.054

Karasu, S., and Sarac, Z. (2019). Investigation of power quality disturbances by using 2D discrete orthonormal S-transform, machine learning and multi-objective evolutionary algorithms. *Swarm Evol. Comput.* 44, 1060–1072. doi:10.1016/j.swevo.2018.11.002

Karasu, S., and Saraç, Z. (2020). Classification of power quality disturbances by 2D-Riesz Transform, multi-objective grey wolf optimizer and machine learning methods. *Digit. signal Process.* 101, 102711. doi:10.1016/j.dsp.2020.102711

Karasu, S., and Saraç, Z. (2022). The effects on classifier performance of 2D discrete wavelet transform analysis and whale optimization algorithm for recognition of power quality disturbances. *Cognitive Syst. Res.* 75, 1–15. doi:10.1016/j.cogsys.2022.05.001

Kariniotakis, G. N., Stavrakakis, G. S., and Nogaret, E. F. (1996). Wind power forecasting using advanced neural networks models. *IEEE Trans. Energy Convers.* 11 (4), 762–767. doi:10.1109/60.556376

Lai, G., Chang, W.-C., Yang, Y., and Liu, H. (2018). "Modeling long-and short-term temporal patterns with deep neural networks," in *The 41st international ACM SIGIR conference on research & development in information retrieval*, 95–104.

Liu, H., and Chen, C. (2019). Data processing strategies in wind energy forecasting models and applications: a comprehensive review. *Appl. Energy* 249, 392–408. doi:10.1016/j.apenergy.2019.04.188

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013a) *Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.*

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013b). Distributed representations of words and phrases and their compositionality. *Adv. neural Inf. Process. Syst.* 26.

Parsons, B., Milligan, M., Zavadil, B., Brooks, D., Kirby, B., Dragoon, K., et al. (2004). Grid impacts of wind power: a summary of recent studies in the United States. *Wind Energy An Int. J. Prog. Appl. Wind Power Convers. Technol.* 7 (2), 87–108. doi:10.1002/we.111

Peng, X., Cheng, K., Lang, J., Zhang, Z., Cai, T., and Duan, S. (2021). Short-term wind power prediction for wind farm clusters based on SFFS feature selection and BLSTM deep learning. *Energies* 14 (7), 1894. doi:10.3390/en14071894

Sarzynska-Wawer, J., Wawer, A., Pawlak, A., Szymanowska, J., Stefaniak, I., Jarkiewicz, M., et al. (2021). Detecting formal thought disorder by deep contextualized word representations. *Psychiatry Res.* 304, 114135. doi:10.1016/j.psychres.2021.114135

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. *Adv. neural Inf. Process. Syst.* 30.

Wang, Y., Hu, Q., Li, L., Foley, A. M., and Srinivasan, D. (2019). Approaches to wind power curve modeling: a review and discussion. *Renew. Sustain. Energy Rev.* 116, 109422. doi:10.1016/j.rser.2019.109422

Wang, Y., Zou, R., Liu, F., Zhang, L., and Liu, Q. (2021). A review of wind speed and wind power forecasting with deep neural networks. *Appl. Energy* 304, 117766. doi:10.1016/j.apenergy.2021.117766

Wu, H., Xu, J., Wang, J., and Long, M. (2021). Autoformer: decomposition transformers with auto-correlation for long-term series forecasting. *Adv. Neural Inf. Process. Syst.* 34, 22419–22430.

Yang, J., Fang, L., Song, D., Su, M., Yang, X., Huang, L., et al. (2021). Review of control strategy of large horizontal-axis wind turbines yaw system. *Wind Energy* 24 (2), 97–115. doi:10.1002/we.2564

Zeng, A., Chen, M., Zhang, L., and Xu, Q. (2023). Are transformers effective for time series forecasting? *Proc. AAAI Conf. Artif. Intell.* 37 (9), 11121–11128. doi:10.1609/aaai.v37i9.26317

Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., et al. (2021). Informer: beyond efficient transformer for long sequence time-series forecasting. *Proc. AAAI Conf. Artif. Intell.* 35 (12), 11106–11115. doi:10.1609/aaai.v35i12.17325

Zhou, J., Lu, X., Xiao, Y., Su, J., Lyu, J., Ma, Y., et al. (2022). Sdwpf: a dataset for spatial dynamic wind power forecasting challenge at kdd cup 2022. *arXiv Prepr. arXiv:2208.04360.*

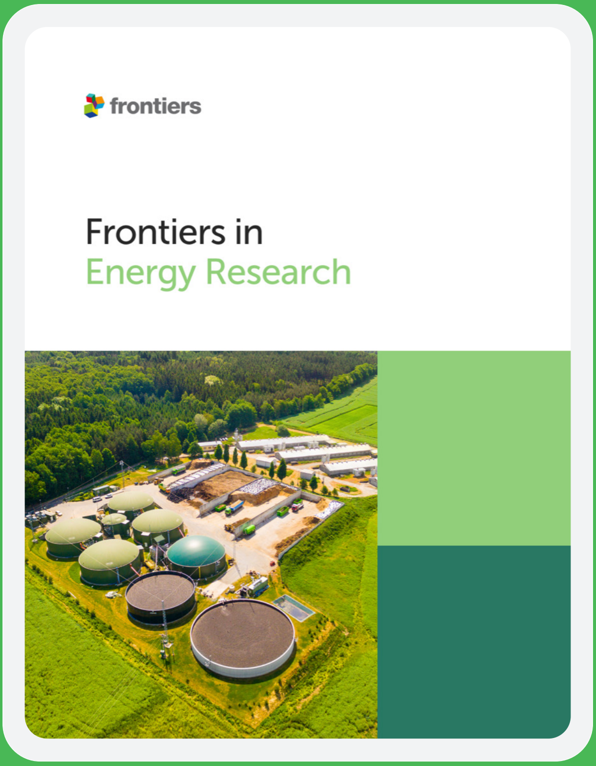
Frontiers in
Energy Research

Advances and innovation in sustainable, reliable and affordable energy

Explores sustainable and environmental developments in energy. It focuses on technological advances supporting Sustainable Development Goal 7: access to affordable, reliable, sustainable and modern energy for all.

Discover the latest Research Topics

See more →



frontiers

Frontiers in
Energy Research

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact