Online monitoring of wind power plants using digital twin models

Edited by
Farid Khazaeli Moghadam, Zhen Gao, Arvind Keprate,
Valentin Chabaud and Mohammad Amin

Coordinated by
Sarah Barber

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

Online monitoring of wind power plants using digital twin models

Topic editors

Farid Khazaeli Moghadam — Norwegian University of Science and Technology, Norway

Zhen Gao — Norwegian University of Science and Technology, Norway

Arvind Keprate — Oslo Metropolitan University, Norway

Valentin Chabaud — SINTEF Energy Research, Norway

Mohammad Amin — EnchantedRock Management LLC, United States

Topic coordinator

Sarah Barber — OST Eastern Swiss University of Applied Sciences, Switzerland

Table of contents

Editorial: Online monitoring of wind power plants using digital twin models

Farid Khazaeli Moghadam[1]*, Arvind Keprate[2] and Zhen Gao[3]

[1]Department of Generation and Transmission, Nexans Norway AS, Oslo, Norway, [2]Department of Mechanical, Electronics and Chemical Engineering, Oslo Metropolitan University, Oslo, Norway, [3]Department of Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai, Shanghai Municipality, China

Editorial on the Research Topic
Online monitoring of wind power plants using digital twin models

In the fast-evolving renewable energy sector, offshore wind technology has made significant strides, evolving from bottom-fixed turbines to the integration of floating turbines. These developments have led to a critical focus on reducing Operational Expenditure (OPEX) and improving turbine reliability through predictive maintenance that uses real-time data to forecast and mitigate potential failures. Digital Twin models have acted as an enabling technology for performing predictive maintenance. This research topic brings together a set of nine studies dedicated to addressing the technical challenges and practical applications of digital twin models in wind energy, aiming to lay the groundwork for more sustainable, reliable, and cost-effective wind energy systems.

The featured articles explore various aspects of digital twin technology, from architectural development to fault detection, all with the goal of optimizing wind turbine maintenance. For instance, Marykovskiy et al. in *Architecting a Digital Twin for Wind Turbine Rotor Blade Aerodynamic Monitoring*, presents a novel digital twin system, Aerosense, designed for rotor blade monitoring using microelectromechanical sensors. Built on systems engineering principles, Aerosense uses microelectromechanical sensors to provide comprehensive monitoring and predictive modeling, moving the industry toward more adaptable and integrated digital twin frameworks.

Another significant contribution, *Fault Detection of a Wind Turbine Generator Bearing Using Interpretable Machine Learning* by Bindingsbø et al. addresses the challenge of detecting faults in generator bearings—a common issue that can compromise turbine performance. This study presents a machine learning-based approach for identifying temperature anomalies indicative of bearing faults, using Supervisory Control and Data Acquisition (SCADA) data. The emphasis on interpretable machine learning underscores the need for transparency and accountability in predictive maintenance, essential for broader adoption of digital twin models in wind energy.

This research topic also delves into powertrain degradation and control. In *Yaw Misalignment in Powertrain Degradation Modelling for Wind Farm Control in Curtailed Conditions*, Moghadam et al. examine how yaw misalignment affects powertrain

components and propose a damage-aware control framework for efficient power distribution within wind farms. This study highlights the integration of degradation metrics into farm-level control, demonstrating how digital twin models can simultaneously address efficiency and durability through advanced control strategies, and allow operators to make informed decisions about the power setpoints of individual turbines within large farms.

Another key study, *Optimal Operation Strategy for Distribution Network with High Penetration of Dispersed Wind Power* by Duan et al. addresses the integration of distributed wind power into local grids. This research presents an optimization model to enhance network performance while balancing high wind penetration and voltage stability. It highlights the potential of digital twins to manage complex, large-scale networks, providing essential decision-making support.

Control strategies for grid integration and stability are also central to *An MPC Based Active and Reactive Power Coordinated Control Strategy of PMSG Wind Turbines to Enhance the Support Capability* by Ma et al. The study advances model predictive control methods for regulating active and reactive power, underscoring the role of digital twins in managing grid support functions traditionally handled by conventional units. This contributes to a broader vision of the utilization of digital twin models in future grid where renewable sources play a more active role in maintaining grid stability.

Available Power Estimation of Wind Farms Based on Deep Spatio-temporal Neural Networks by Liu et al. illustrates the potential of artificial intelligence in digital twin applications. This study leverages deep learning techniques to forecast available power based on historical data, demonstrating how spatio-temporal neural networks can enhance prediction accuracy. The focus on turbine-level data highlights the potential of digital twins for fine-grained, turbine-specific insights, allowing operators to make informed decisions about individual turbines within large farms.

In addition to simulations, the research topic also includes an experimental study performed by Xiong et al. in *Experimental study of Dynamic Characteristics of an Ultra-Large Jacket Offshore Wind Turbine under Wind and Wave Loads Using Aero-Hydro-Structural Elastic Similarities*, which provides a physical testing framework for offshore turbine stability under extreme conditions. This study underscores the importance of empirical validation in digital twin models, ensuring their reliability across diverse operational conditions.

Power management and fault resilience are addressed by contributions such as *A Model Predictive Control Strategy for Enhancing Fault Ride Through in PMSG Wind Turbines Using SMES and Improved GSC Control* by Abdelkader et al. which introduces a novel strategy that combines superconducting magnetic energy storage (SMES) with model predictive control (MPC). This strategy boosts fault resilience by providing reactive power support during grid disruptions, showcasing how digital twins can predict and preempt faults, minimizing downtime.

Finally, *Study on Obtaining Real Power Curve of Wind Turbines by Using SCADA Data* by Dai et al. explores methods for deriving accurate power curves from SCADA data, an essential function of digital twins in assessing wind turbine performance. This study contributes to more precise performance assessments, directly supporting predictive maintenance and operational planning.

Together, these studies represent a comprehensive approach to tackling the challenges of digital twin modeling for predictive maintenance in wind energy. They collectively address the three core layers of digital twin realization—data acquisition, dynamic model processing, and decision support—each layer contributing to a robust framework for optimizing turbine performance. As digital twin technology continues to mature, these contributions highlight the technical challenges and the immense potential of predictive maintenance, steering the wind energy industry towards a more sustainable and efficient future.

Author contributions

FK: Writing–original draft, Writing–review and editing. AK: Writing–review and editing. ZG: Writing–review and editing.

Funding

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.

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

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.

Study on Obtaining Real Power Curve of Wind Turbines Using SCADA Data

Juchuan Dai[1]*, Huifan Zeng[1], Fan Zhang[1], Huanguo Chen[2] and Mimi Li[1]

[1]School of Mechanical Engineering, Hunan University of Science and Technology, Xiangtan, China, [2]Faculty of Mechanical Engineering and Automation, Zhejiang Sci-Tech University, Hangzhou, China

The key problem to be solved in the process of wind turbine (WT) operation and maintenance is to obtain the wind turbine performance accurately. The power curve is an important indicator to evaluate the performance of wind turbines. How to model and obtain the power curve of wind turbines has always been one of the hot topics in research. This paper proposes a novel idea to get the actual power curve of wind turbines. Firstly, the basic data preprocessing algorithm is designed to process the zero value and null value in the original supervisory control and data acquisition (SCADA) data. The moving average filtering (MAF) method is employed to deal with the wind speed, the purpose of which is to consider the comprehensive result of wind on the wind turbine power in a certain period. According to the momentum theory of the ideal wind turbine and combined with the characteristics of the anemometer installation position, the deviation between the measured wind speed and the actual wind speed is approximately corrected. Here, the influence of dynamic changes in air density is also considered. Then, the Gaussian fitting algorithm is used to fit the wind-power curve. The characteristics of the power curve before and after wind speed correction are compared and analyzed. At the same time, the influence of the parameter uncertainty on the reliability of the power curve is considered and investigated. Finally, the characteristics of the power curves of four wind turbines are compared and analyzed. The research results show that among these power curves, WT3 and WT4 are the closest, WT2 is the next, and WT1 has the farthest deviation from the others. The research work provides a valuable basis for on-site performance evaluation, overhaul, and maintenance of wind turbines.

Keywords: power curve, wind turbines, SCADA data, moving average filtering, wind speed correction

INTRODUCTION

The serious impact of environmental degradation has increased global interest in wind energy. In recent years, the wind power industry has developed rapidly (Dai et al., 2018a; Dawn et al., 2019). According to the wind power statistics revealed by the World Wind Energy Association (WWEA), the worldwide wind capacity has reached 744 gigawatts. In 2020, 93 gigawatts of new wind turbines were added, setting a new record. With batches of wind turbines in service, their operating efficiency

Abbreviations: ANN, artificial neural network; GP, Gaussian process; MAF, moving average filtering; MLE, maximum likelihood estimation; NSFM, non-symmetric fuzzy means; SCADA, supervisory control and data acquisition; SLF, standard logistic functions; TI, turbulence intensity; WT, wind turbine; WTPC, wind turbine power curve; WWEA, World Wind Energy Association.

(usually referred to as "power coefficient") has become the focus because this is directly related to the economic benefits of wind farm operation (Dai et al., 2016a; Dhunny et al., 2020; Sun et al., 2020; Bakir et al., 2021). However, due to the influence of many factors, the actual power characteristic of wind turbines operating in wind farms is often inconsistent with the designed power characteristic. So, modeling and obtaining the power curve of wind turbines has always been one of the hot topics in research (Rogers et al., 2020).

The power curve of wind turbines indicates the generated power versus wind speed (Ciulla et al., 2019). It is widely used for monitoring and evaluating wind turbine performance (Pandit et al., 2020; Astolfi et al., 2021a). The abundant SCADA data of wind farms provide a good database for wind power curve research. The different techniques used for wind turbine power curve (WTPC) modeling can be divided into parametric techniques and non-parametric techniques (Lydia et al., 2014). The power generation of wind turbines will vary with external environmental conditions. To investigate the influence of external conditions on wind speed and wind turbine power, Kim et al. analyze three atmospheric factors: atmospheric stability, turbulence intensity (TI), and wind shear (Kim et al., 2021). Various factors such as the age of the wind turbine, installation location, air density, and wind direction will cause inhomogeneity among the observation data, which usually affects the accuracy of the fitted power curve. To overcome this problem, the hybrid estimation method by Mehrjoo et al. is presented, which is based on weighted balanced loss functions (Mehrjoo et al., 2021). Saint-Drenan et al. develop an open-source model that can generate the power curve of any turbine to suit the specific conditions of any site (Saint-Drenan et al., 2020). Marčiukaitis et al. present a nonlinear regression model (three-parameter exponential model) for modeling power curve with application to the wind turbine of Seirijai wind farm (in Lithuania) (Marčiukaitis et al., 2017). To get highly accurate non-parametric power curve models, Karamichailidou et al. employ artificial neural network (ANN) belonging to the radial basis function architecture and train it using non-symmetric fuzzy means (NSFM) (Karamichailidou et al., 2021). Virgolino et al. introduce a semi-parametric method that combines Gaussian process (GP) regression, standard logistic functions (SLF), and probabilistic kernel-based machine learning models (Virgolino et al., 2020). Manobel et al. present a method based on GP data pre-filtering and ANN modeling of the power curve, where the prior filtering by GP modeling can improve the network performance (Manobel et al., 2018). Mehrjoo et al. propose two non-parametric techniques, which are based on the tilt method and the monotonic spline regression, to construct the WTPC that maintains monotonicity (Mehrjoo et al., 2020). Seo et al. construct a nonlinear parametric power curve model using the logistic function, and four parameters in the logistic function are obtained explicitly by the maximum likelihood estimation (MLE) method (Seo et al., 2019). The results provided by the

logistic functions are useful due to the continuity and adaptability. However, there are many types of logic functions, and how to choose is a question worth studying. From this scenario, the well-known logistic functions are employed and tested for modeling WTPC by Villanueva et al. (Villanueva and Feijóo, 2018). Yesilbudak et al. present a robust hybrid method for the power curve modeling of wind turbines, where Mahalanobis distance measure and the chi-square cumulative distribution are used for the power curve filtering (Yesilbudak, 2018). Usually, the SCADA-collected data are those averaged (typically with an averaging time of 10 min). Gonzalez et al. investigate the use of high-frequency SCADA data for wind turbine performance monitoring and propose a new framework based on multivariate non-parametric models (Gonzalez et al., 2017; Gonzalez et al., 2019).

However, the previous research mainly focused on the power curve algorithm itself, that is, how to improve the algorithm to improve the accuracy and reliability further, and the analysis of the impact caused by the physical properties of the data itself is insufficient. The knowledge gaps that need to be supplemented are mainly manifested in the following aspects, which are also the main contributions of this article.

- How to obtain the data needed for reliable power curve modeling from SCADA data. In SCADA data, the wind speed is provided by the anemometer installed on the nacelle, which is not the actual incoming wind speed (Dai et al., 2016b; Dai et al., 2019). Because the wind is blowing first on the wind rotor and then on the anemometer, some of the wind energy has already been absorbed by the wind rotor, the wind speed measured by the anemometer is smaller than the actual incoming wind speed. If the wind speed data in SCADA is used directly to obtain the power curve, the deviation must be significant (**Figure 1**). So, different from previous studies, this article first corrects the wind speed in SCADA and then performs power curve modeling.
- How to obtain the real mapping characteristics between wind speed and power. The wind speed recorded in the SCADA data is the instantaneous wind speed (once a second). Due to the inertia of the wind turbine, the output power cannot respond to the fast wind speed but is the result of a comprehensive response to the wind speed for a period. In other words, violent fluctuation in wind speed is common. However, the violent fluctuation of wind speed does not cause the violent fluctuation of generator power in a short period. The primary reason is that the wind rotor is a large inertial system. Therefore, the key question is how to find the true mapping relationship between the two. For this purpose, the moving average filtering (MAF) method will be employed to deal with the wind speed in the following sections. Then, a reasonable filter window that meets the characteristics of wind speed and power mapping will be further found.
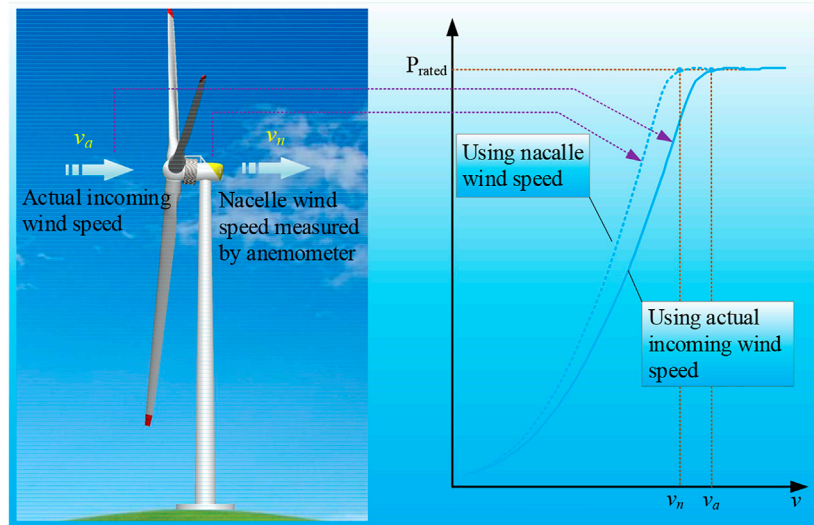
FIGURE 1 | Wind turbine and its power curve.
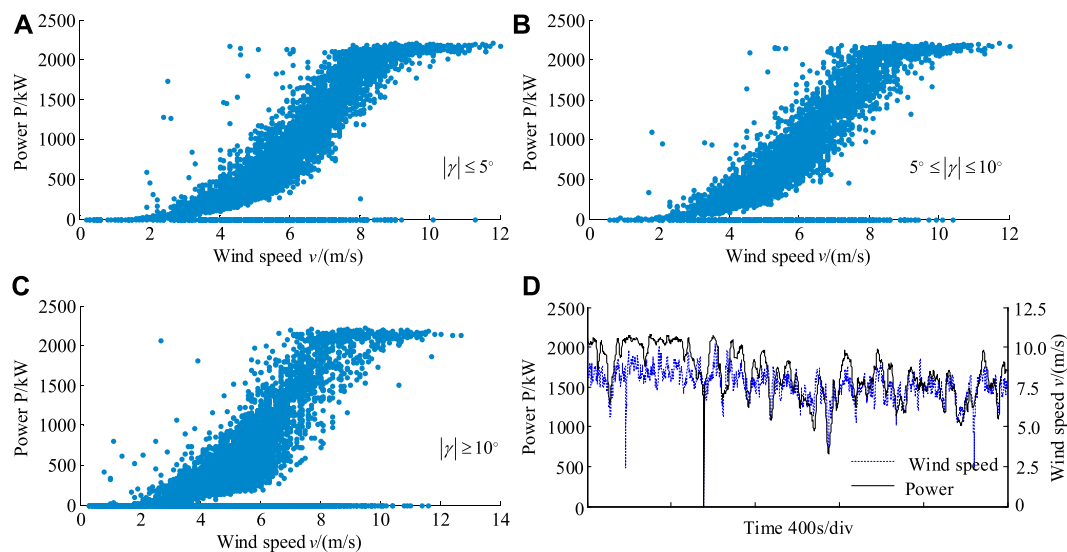
FIGURE 2 | Wind and power from raw SCADA data. **(A)** Wind speed /(m/s). **(B)** Wind speed /(m/s). **(C)** Wind speed /(m/s). **(D)** Time 400s/div.

- How to identify the reliability of the modeled power curve. Power curve modeling involves the processing of some parameters. If there is uncertainty (deviation) in the parameter value, it will affect the accuracy of power curve modeling. For example, the air temperature is used for the power curve modeling, which is recorded in the SCADA system. However, for mountain wind farms, significant differences in distance and altitude can cause temperature data bias. In another scenario, the effect of the sensor's error should also be considered. Specifically, the two direct parameters for power curve modeling are wind speed and power, both in the SCADA

system. The reliability of wind speed data and power data in the SCADA system should be judged. In this way, the actual power curve can be obtained better.

WIND AND POWER FROM RAW SCADA DATA

In the SCADA system of the investigated wind turbines, many operation parameters, such as wind speed, the rotational speed of the wind rotor, generator side power of the converter, and grid side power of the converter, are recorded. The investigated wind

turbines are 2 MW direct-driven type. For this type of wind turbines, the generator side power of the converter is the output power of the generator. If the loss of mechanical energy into electrical energy is ignored, the power output of the generator can also be approximately regarded as the mechanical energy output of the wind rotor. Therefore, the scatterplot between wind speed and generator power can be obtained by using wind speed data and generator power data, which is shown in **Figures 2A–C**. Here, about 57 h of SCADA data are used. Although the sampling interval is 1 s, a point is extracted every 10 s to reduce the size of the scatterplots. It should be pointed out that the wind speed data is measured by the anemometer mounted on the top of the nacelle. This means that the wind is blowing first on the wind rotor and then on the anemometer. As a result, the wind speed measured by the anemometer is smaller than the actual incoming wind speed because some of the wind energy has already been absorbed by the wind rotor. In addition, yaw misalignment tends to degrade wind turbine power production (Gao and Hong, 2021). Here, these SCADA data can be divided into three types according to the magnitude of the yaw angle: less than 5°, more than 5° but less than 10°, and more than 10°.

In **Figure 2A**, the corresponding yaw angle γ is less than 5° and there are 7,074 sets of data. In **Figure 2B**, the corresponding yaw angle γ is more than 5° but less than 10°, and there are 5,832 sets of data. In **Figure 2C**, the corresponding yaw angle γ is more than 10°, and there are 7,649 sets of data. By looking at the SCADA data, it can be found that the yaw angle is scattered in different angle ranges. This is determined by the various characteristics of natural wind direction and the yaw control strategy of the wind turbine. It also means that this distribution may be different in different periods or different wind turbines, with certain randomness. It is also not difficult to find that the variation of yaw angle may affect the wind-power characteristics by comparing the distribution characteristic of the scatterplots in the three subfigures. In addition, the distribution of some points is far away from their concentrated distribution area. For example, in **Figure 2A**, near a wind speed of 4 m/s, there are scattered points with a power of more than 2000 kW. This is not the normal performance of a wind turbine. There are two possible reasons for this phenomenon: there may be interference signals during data collecting, and the other is that there may be instantaneous fluctuations in wind speed. Violent fluctuation in wind speed is common. However, the violent fluctuation of wind speed does not cause the violent fluctuation of generator power. The primary reason is that the wind energy is absorbed by the wind rotor, a large inertial system. In other words, the instantaneous fluctuation of the wind does not cause the instantaneous fluctuation of the rotational speed of the wind rotor. **Figure 2D** shows the time history curves of wind speed and generator power over a period. The data sampling interval used here is 1 s. It can be seen from this subfigure that the change frequency of wind speed is greater than the change frequency of generator power. The changing trend of wind speed and the changing trend of generator power are not identical. Due to the inherent property of wind rotor inertia, it can be considered that the influence of wind speed on generator power has a certain lag effect, and the influence of wind speed on generator power is the result of the comprehensive impact of wind in a period. Under the condition of continuous wind speed, the generator power occasionally appears zero values, which is an error in the collection process and should be eliminated. Another thing to note is that both the wind speed and the generator power are measured by sensors; thus, the reliability of the data should be confirmed. In other words, the system error of the sensor should be eliminated as far as possible. The reason for proposing this issue is that the sensors may not be calibrated regularly. For example, an anemometer may not be calibrated during several years of operation.

Overall, the following points should be attentional for the investigation of the power curve of wind turbines.

- Compensation for the deviation between the wind speed measured by the anemometer and the actual incoming wind speed should be considered.
- The effect of wind direction (yaw angle) change on actual power performance should be considered.
- The lag effect between wind speed change and generator power change should be considered.
- The error data in the data collection process should be eliminated.
- The effect of wind speed on generator power should be regarded as the comprehensive result of wind in a certain period.
- The system error of the sensors related to wind and generator power should be eliminated.

WIND SPEED DATA CORRECTION

Since the effect of wind speed on generator power is the comprehensive result of wind in a certain period, it is more reasonable to model the power curve using the average of the wind speed rather than its instantaneous value. At the same time, the lag effect between wind speed variation and generator power variation is also considered. Based on these considerations, the moving average filtering (MAF) method is employed to deal with the wind speed. This method sequentially stores the data of n sampling points as a queue. When new data is collected, the first data in the queue is discarded and the arithmetic mean of the data is calculated. The discrete expression of the filter can be written as

$$y(n) = \frac{1}{N+1} \sum_{k=0}^{N} x(n-k) \tag{1}$$

where, $y(n)$ is the output of the filter, $x(n)$ is the input of the filter, and N is the window length of the MAF filter.

Another thing to note is that null values and zero values may occur during the sensor test. If a null value appears, it is filled with zero. If a zero value appears between two normally collected data, it can be repaired by averaging the two adjacent data shown as **Eq. 2**. Also, this processing should be done before filtering. If multiple zeros occur consecutively, the data should be rejected.
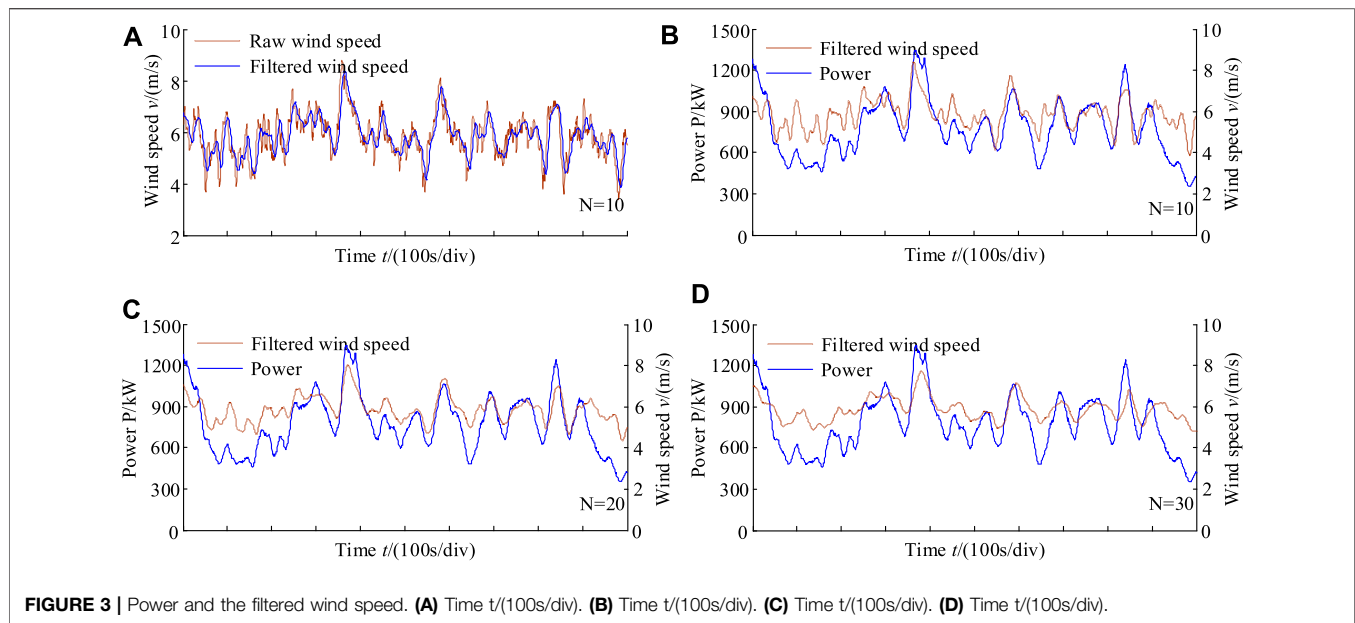
FIGURE 3 | Power and the filtered wind speed. **(A)** Time t/(100s/div). **(B)** Time t/(100s/div). **(C)** Time t/(100s/div). **(D)** Time t/(100s/div).

$$\begin{cases} x(n) = 0, \text{ if } x(n) \text{ is null} \\ x(n) = \dfrac{x(n-1) + x(n+1)}{2}, \text{ if } x(n) = 0, \ x(n-1) > 0 \& x(n+1) > 0 \quad (2) \\ x(n) = 0, \text{ if } x(n+1) = 0, \ x(n+2) = 0, \cdots \end{cases}$$

When using MAF filters, the critical problem is to determine the window length of the filter. Here, the basic idea to find the appropriate filter window length is to compare wind speed change trends and power change trends. Therefore, the filtered wind speed curve and the time-history power curve are given as shown in **Figure 3**.

In **Figure 3A**, the raw wind speed and the filtered wind speed are presented in a period. The window length N is set to be 10. The wind speed fluctuation is reduced after filtering. At the same time, the filtered wind speed curve has a slight time delay compared with the raw wind speed curve. **Figures 3B–D** show the wind speed curves for different filter window lengths and the corresponding time-history power curves. The longer the filter window is, the smoother the filtered wind speed curve is. Judging from the consistency of wind speed trend and power trend, it is not that the longer the filter window, the better. In contrast, a filter with a window length of 10 is better. Moreover, it should be noted that the value of the filter window length is only an approximate value because of the complexity of wind conditions.

On the other hand, the wind speed measured by the anemometer mounted on the top of the nacelle is not the actual wind speed. How to correct the wind speed has been an issue of great concern. For example, Malgaroli *et al.* propose a nacelle wind speed correction for evaluating wind turbine performance by estimating the wind speed entering the wind rotor (Astolfi et al., 2021b; Carullo et al., 2021). In this paper, the nacelle wind speed compensation is mainly based on the aerodynamic and energy flow characteristics of the wind turbines (Dai et al., 2016b). It should be noted here that the anemometers installed on the nacelle are considered in a normal working state. In particular, the anemometers are generally no longer calibrated regularly after service. Two issues require special consideration in the future. One is the deviation caused by the aging of the sensor itself, and the other is the deviation caused by the sensor failure.

According to the momentum theory, the power of a wind turbine to capture wind energy can be expressed as

$$P = \frac{1}{2}\rho S v_d \left(v_1^2 - v_2^2\right) \quad (3)$$

where, v_1 is the upstream wind speed of the wind rotor; v_d is the wind speed passing through the wind rotor; v_2 is the downstream wind speed of the wind rotor; ρ is the density of air; S is the area swept by the wind rotor.

The relationship between v_1, v_d, and v_2 can be written as (Hansen, 2008; Dai et al., 2016b)

$$v_2 = 2v_d - v_1 \quad (4)$$

Substituting **Eq. 4** into **Eq. 3**, there are

$$P = 2\rho S v_d^2 \left(v_1 - v_d\right) \quad (5)$$

Subsequently, **Eq. 5** can be transformed into (Dai et al., 2016b)

$$v_1 = \frac{P}{2\rho S v_d^2} + v_d \quad (6)$$

Since the anemometer is installed near the wind rotor, the wind speed measured by the anemometer can be considered as the speed flowing through the wind rotor, that is, v_d. From **Eq. 6**, the deviation between the measured wind speed and the actual incoming wind speed is $P/2\rho S v_d^2$. Using **Eq. 6**, the wind speed can be further corrected.

In **Eq. 6**, air density ρ is not a constant value, will change with ambient temperature, air pressure, and relative humidity. Furthermore, as the altitude changes, so does the atmospheric
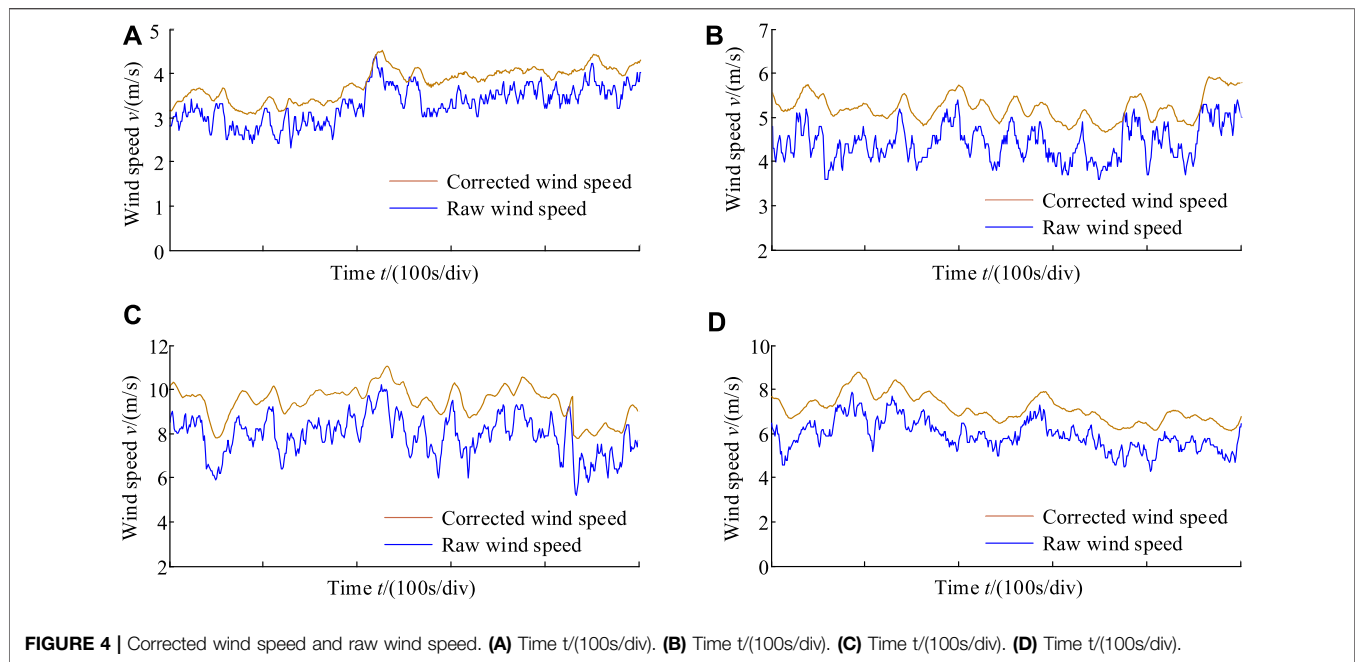
FIGURE 4 | Corrected wind speed and raw wind speed. **(A)** Time t/(100s/div). **(B)** Time t/(100s/div). **(C)** Time t/(100s/div). **(D)** Time t/(100s/div).

pressure. Here, the air density is calculated using the method shown in the "omni calculator" (https://www.omnicalculator.com). The calculation expression from the website is

$$
\begin{cases}
\rho = \dfrac{P_0 e^{\frac{-gM(h-h_0)}{R \cdot T_k}} - P_V}{R_d \cdot T_K} + \dfrac{P_V}{R_v \cdot T_K} \\[4mm]
P_V = 6.1078 \times 10^{\frac{7.5 T_C}{T_C + 237.3}} \cdot RH
\end{cases}
\tag{7}
$$

where, T_C is the measured temperature (degrees Celsius); T_K is the air temperature (Kelvins); RH is the relative humidity; P_V is the water vapor pressure (Pa); R_d is the gas constant for dry air (287.058 J/(kg·K)); R_v is the gas constant for water vapor (461.495 J/(kg·K)); P_0 is the pressure at the sea level; g is the gravitational acceleration; M is the molar mass of air (M = 0.0289,644 kg/mol); h is the altitude; R is the universal gas constant (R = 8.31432 N m/(mol·K)).

According to the above calculation method, the corrected wind speed can be calculated, that is, the wind speed filtering is carried out according to **Eq. 1**, and then the wind speed deviation is corrected according to **Eq. 6**. In **Figures 4A–D**, four wind speed correction curves under different wind speed conditions are given. Overall, the corrected wind speed curves have a delay in time, smoother curves, and an increase in numerical value compared with the raw wind speed curves.

POWER CURVE MODEL

According to the wind power theory, the power of a wind turbine is related to the inflow wind speed, rotational speed of the wind rotor, pitch angle, yaw angle, and so on. Usually, when the wind

turbine is studied, the power is written as a function of wind speed and power coefficient as shown in **Eq. 8**. Wind speed, wind direction, the rotational speed of the wind rotor, pitch angle and air density are all variables during wind turbine operation (Dai et al., 2018b).

$$
P = \frac{1}{2} \rho \pi R^2 C_P (v, \omega, \beta, \gamma) v^3
\tag{8}
$$

where, R is the radius of the wind rotor; C_P is the power coefficient; ω is the rotational speed of the wind rotor; β is the pitch angle; γ is the yaw angle.

When a wind turbine is in regular operation, the structure parameters and control parameters have been set. The performance of the wind turbine is mainly the ability to capture wind energy, which is reflected in the relationship between power and wind speed. In other words, from the user's point of view, the relationship between power and wind speed is their concern, which is also an important index for the on-site assessment of wind turbines. From this scenario, it is rather vital to obtain the actual wind-power curve. All the dynamic variations of the rotational speed of the wind rotor, pitch angle, and yaw angle are considered internal factors. This means that the power curve modeling is simplified to the form of **Eq. 9**.

$$
P = f(v)
\tag{9}
$$

However, only a scatterplot of wind and power can be obtained using SCADA data without an exact expression. A better description would be to get a relational expression and use a curve to describe it. Therefore, the key to solving this problem is to construct an effective nonlinear regression equation. Here, Gaussian fitting was used to process the
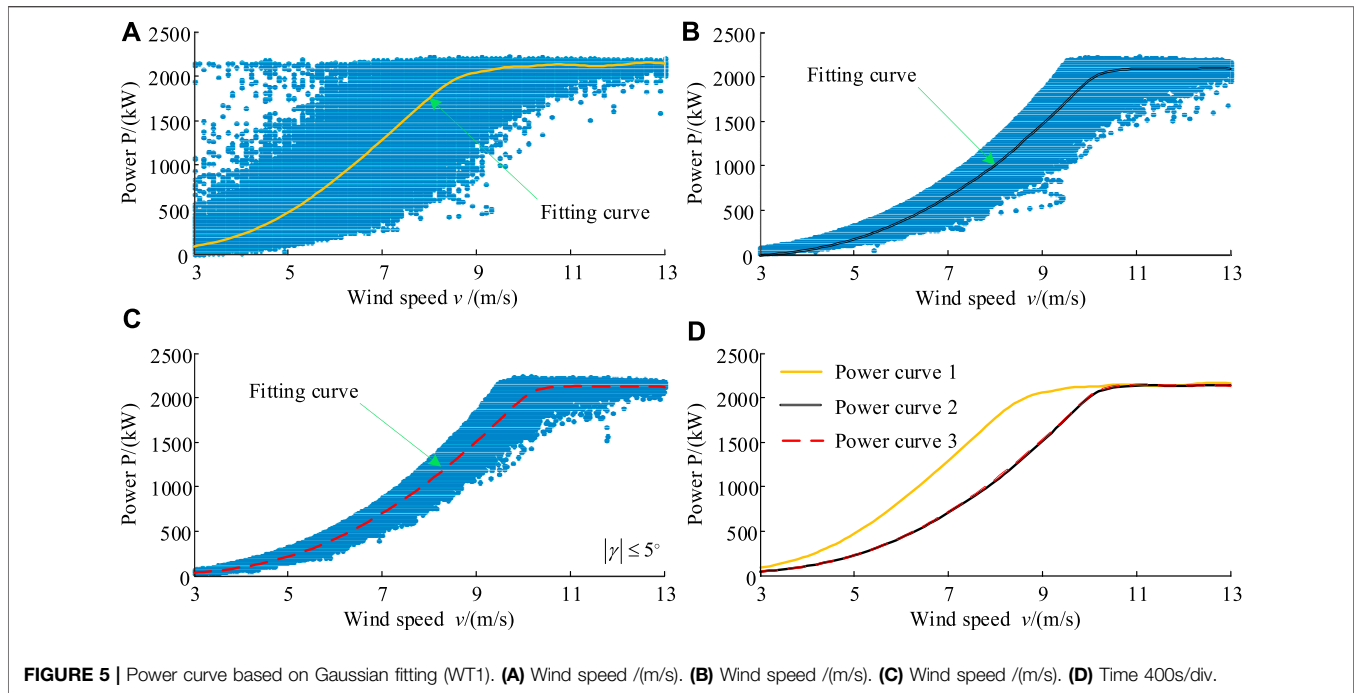
FIGURE 5 | Power curve based on Gaussian fitting (WT1). **(A)** Wind speed /(m/s). **(B)** Wind speed /(m/s). **(C)** Wind speed /(m/s). **(D)** Time 400s/div.

data through the comparison of various fitting forms, including Exponential fitting, Fourier fitting, Polynomial fitting, Power fitting, Sum of sin functions, etc. The expression of Gaussian fitting can be written as (Li et al., 2018; Xu et al., 2019)

$$P = f(v) = \sum_{i=1}^{n} \alpha_i \cdot e^{-\left[\frac{(v-\beta_i)}{\delta_i}\right]^2} \qquad (10)$$

where, v is wind speed; P is the power of the wind turbine; α_i, β_i and δ_i are the coefficients; n is the number of peaks to fit.

When using the Gaussian fitting method, the coefficient n in **Eq. 10** is set to be 8, the nonlinear least square fitting method is employed, and the algorithm is "trust region". The method has been integrated into the curve fitting toolbox of MATLAB. In the specific settings, the minimum change in coefficients for finite difference Jacobians is 1×10^{-8}, the maximum change in coefficients for finite difference Jacobians is 0.1, the maximum number of function (model) evaluations allowed is 600, the maximum number of fit iterations allowed is 400, the termination tolerance used on stopping conditions involving the function (model) value is 1×10^{-6}, the termination tolerance used on stopping conditions involving the coefficients is 1×10^{-6}. Also, the "center and scale" method is selected to use. When using this method, the abscissa of the data points to be fitted is changed by **Eq. 11** in MATLAB, that is

$$v' = \frac{v - \text{mean}(v)}{\text{std}(v)} \qquad (11)$$

Then, **Eq. 10** can be rewritten as

$$P = f\left(\frac{v - \text{mean}(v)}{\text{std}(v)}\right) = \sum_{i=1}^{n} \alpha_i \cdot e^{-\left[\frac{(x'-\beta_i)}{\delta_i}\right]^2} \qquad (12)$$

Figure 5 shows the power curve based on Gaussian fitting in three different scenarios. **Figure 5A** shows the wind-power scatters and the fitting curve using the raw wind speed data. Here, the selected wind speed ranges from 3 m/s to 13 m/s, and 0 values are excluded. The wind speed interval of 0.1 m/s is used to draw the fitting curve. Although the scatterplot covers a wide range, the trend of the fitting curve is consistent with the design law. It also illustrates that the curve fitting method is suitable. **Figure 5B** shows the wind-power scatters and the fitting curve using the corrected wind speed data in which all the yaw angles are contained. The range of wind-power scatterplot is significantly reduced after the wind speed correction. This shows that the wind speed correction is reasonable. It is also interesting to note that the upper contour of the wind-power scatterplot is relatively regular. In contrast, the lower contour of the wind-power scatterplot is still irregular.

Figure 5C shows the wind-power scatters and the fitting curve using the corrected wind speed data in which only yaw angles of less than 5° are included. For ease of comparison, the wind-power fitting curves for the three scenarios are put together in **Figure 5D**. Curve 1 denotes the wind-power fitting curve of scenario Ⅰ (**Figure 5A**), curve 2 denotes the wind-power fitting curve of scenario Ⅱ (**Figure 5B**), and curve 3 denotes the wind-power fitting curve of scenario Ⅲ (**Figure 5C**). **Table 1** shows the fitting coefficients and goodness of the power curve. In different scenarios, the fitting coefficient is different. For scenario Ⅰ, x is normalized by mean 6.258 and std 2.071; for scenario Ⅱ, x is

TABLE 1 | Fitting coefficients and goodness of the power curve (WT1).

Coefficients (with 95% Confidence Bounds)	Power Curve Using Raw SCADA Data	Power Curve Using Corrected SCADA Data (including all Yaw Angles)	Power Curve Using Corrected SCADA Data (including Yaw Angles of less than 5°)
α_1	27.41	214.9 (40.69, 389.1)	11.28 (2.124, 20.45)
β_1	1.648	1.201 (1.126, 1.276)	0.9859 (0.9808, 0.991)
δ_1	0.1773	0.4794 (0.3294, 0.6294)	0.00787 (0.00027, 0.01547)
α_2	2,570	17.18 (9.408, 24.96)	430.3 (254.5, 606.2)
β_2	1.681	0.5873 (0.565, 0.6096)	1.366 (1.185, 1.548)
δ_2	0.7932	0.1352 (0.0883, 0.1821)	0.6663 (0.5569, 0.7758)
α_3	0	−11.41 (−68.77, 45.96)	61.57 (−254.8, 377.9)
β_3	−8.822	2.026 (1.7, 2.351)	1.841 (1.69, 1.992)
δ_3	0.005298	0.2248 (−0.2693, 0.7189)	0.2534 (−0.07993, 0.5868)
α_4	−549.5(−1.672e4, 1.562e4)	32.94 (15.48, 50.39)	865.2 (673, 1,057)
β_4	1.646(0.843, 2.45)	1.184 (1.17, 1.198)	2.464 (2.191, 2.737)
δ_4	0.5395 (−1.427, 2.506)	0.1526 (0.1111, 0.1942)	0.557 (−0.07272, 1.187)
α_5	−2,198 (−1.028e6, 1.023e6)	5.622e4 (−2.523e8, 2.524e8)	8.968 (3.792, 14.14)
β_5	1.733 (−19.11, 22.57)	10.1 (−8,082, 8,102)	0.9242 (0.9107, 0.9378)
δ_5	0.8921 (−28.43, 30.22)	3.461 (−2,135, 2,142)	0.03104 (0.0094, 0.053)
α_6	−6.961 (−243.8, 229.8)	−14.44 (−29.38, 0.5091)	67.42 (53.08, 81.76)
β_6	1.358 (−0.32, 3.036)	0.8275 (0.825, 0.83)	1.122 (1.115, 1.129)
δ_6	0.2503 (−2.741, 3.242)	0.00302 (−0.00051, 0.0065)	0.1483 (0.1295, 0.1671)
α_7	1723 (−6,686, 1.013e4)	165.5 (−1.38e4, 1.413e4)	−24.9 (−42.53, −7.267)
β_7	3.523 (−4.671, 11.72)	0.3597 (−13.39, 14.11)	0.8373 (0.8359, 0.8388)
δ_7	1.557 (−34.62, 37.74)	1.393 (−11.7, 14.48)	0.00245 (0.00053, 0.0044)
α_8	1891 (−7,117, 1.09e4)	1860 (−7.483e4, 7.855e4)	1,685 (1,587, 1783)
β_8	1.383 (−2.771, 5.537)	1.883 (−16.19, 19.96)	1.39 (1.306, 1.473)
δ_8	1.686 (0.5422, 2.829)	1.826 (−16.85, 20.5)	1.719 (1.685, 1.754)
SSE	1.395e10	2.524e9	3.294e8
R-square	0.9022	0.983	0.9936
Adjusted R-square	0.9022	0.983	0.9936
RMSE	228.2	95.72	59.49

e in the table is the base of the exponential function with a value of 10.

TABLE 2 | Wind speed and the corresponding power based on fitting curve (WT1).

Wind Speed (m/s)	Power Curve Using Raw SCADA Data	Power Curve Using Corrected SCADA Data (including all Yaw Angles)	Power Curve Using Corrected SCADA Data (including Yaw Angles of less than 5°)
	Power (kW)	Power (kW)	Power (kW)
3	87	45	45
4	220	108	108
5	469	227	228
6	843	424	428
7	1,288	704	715
8	1765	1,057	1,079
9	2052	1,506	1,519
10	2,119	2001	2024
11	2,140	2,131	2,145
12	2,138	2,130	2,139
13	2,153	2,133	2,126

normalized by mean 7.319 and std 2.367; for scenario III, x is normalized by mean 7.487and std 2.384. In **Figure 5D**, curve 2 and curve 3 are basically coincident, and both are separated from curve 1. This shows that a wider yaw angle does not significantly affect wind-power fitting. This is because there are positive and negative yaw angles, and the effects of positive and negative yaw angles on the wind-power curve cancel each other out. In curve 1, the critical wind speed at which the wind turbine reaches the rated

output power is less than that in curve 2 (curve 3). What needs explanation is that taking into account the actual operating characteristics of wind turbines, in **Figure 5C**, data with power coefficients greater than 0.593 and less than 0.15 are excluded.

Table 2 gives wind speed and the corresponding power based on the fitting curve. In curve 1, when the wind speed is 9 m/s, the power of the generator is 2052 kW. Further, it can be found by calculating the curve fitting expression that when the wind speed is 8.7 m/s, the power of the wind turbine reaches 2000 kW. In curve 2, the wind speed is about 10 m/s which corresponds to the power of 2000 kW. Likewise, the wind speed is about 10 m/s which corresponds to the power of 2000 kW in curve 3. According to the wind turbine manufacturer, the designed rated wind speed is about 10.5 m/s. From this information, it is obvious that curve 2 and curve 3 fit better. This further shows that the wind speed correction is effective.

What needs to be further explained is that the fitted critical wind speed is 10 m/s, which is different from the 10.5 m/s designed by the manufacturer. There are several reasons to consider. The design value of a wind turbine is calculated based on a specific service condition, while the actual service conditions are variable, such as wind speed fluctuations, temperature changes, humidity changes, and so on. Because of the complexity of wind turbines, the physical model must be simplified in design, which leads to the difference between the physical model and the actual model. Because of the limitation of the level of the manufacturer, there are some differences between the manufactured turbine and the designed turbine. From the operating results, the constant power output of the wind turbine is not just 2000 kW, but more than 2000 kW. Generally, the wind-power curve after the wind speed correction is more in line with the actual situation. This is applicable to the performance evaluation of wind turbines in wind farms.

RELIABILITY ASSESSMENT OF POWER CURVE

Power curve modeling involves the processing of some parameters. If there is uncertainty (deviation) in the parameter value, it will affect the accuracy of power curve modeling. For example, the historical air relative humidity is used for the power curve modeling, which is not recorded directly in the SCADA system and can only be extracted by consulting the data from other websites or database and is the humidity data in the larger region, which is bound to have some differences with the real humidity data. Although the temperature is directly recorded in SCADA data, it is only the temperature data recorded near the wind tower in wind farms. For mountain wind farms, significant differences in distance and altitude can also cause temperature data bias. In **Eq. 4**, the parameter v_d denotes the wind speed in the wind rotor plane. However, the actual wind speed used is the measured result by the anemometer on the nacelle. Therefore, the deviation between the measured wind speed and the theoretical wind speed will also affect the accuracy of the power curve.

It should be noted that none of the above deviations are due to errors of the sensor itself. In other words, even if the sensors are very accurate, these deviations still exist. In another scenario, the effect of the sensor's error should also be considered. Specifically, the two direct parameters for power curve modeling are wind speed and power, both in the SCADA system. The reliability of wind speed data and power data in the SCADA system should be judged. In this way, the actual power curve can be obtained better.

· Uncertainty effect of air humidity, air temperature, air density, and wind speed

From **Eq. 7**, the deviation of relative humidity will affect the calculation of air density and then affect the calculation of wind speed in **Eq. 6**. If the air humidity has a δ_p% deviation, the effect on the air density can be written as

$$\frac{\hat{\rho}}{\rho} = \left[\frac{P_0 e^{\frac{-gM(h-h_0)}{R \cdot T_k}} - \hat{P}_V}{R_d \cdot T_K} + \frac{\hat{P}_V}{R_v \cdot T_K} \right] \Big/ \rho \qquad (13)$$

where, ρ is the actual air density; $\hat{\rho}$ is the air density with some deviation $P_V = 6.1078 \times 10^{\frac{7.5 T_C}{T_C + 237.3}} \cdot \text{RH}(1 + \delta_p\%)$.

If the air temperature has a deviation ΔT, the effect on the air density can be written as

$$\frac{\hat{\rho}}{\rho} = \left[\frac{P_0 e^{\frac{-gM(h-h_0)}{R \cdot (T_k + \Delta T)}} - \hat{P}_V}{R_d \cdot (T_K + \Delta T)} + \frac{\hat{P}_V}{R_v \cdot (T_K + \Delta T)} \right] \Big/ \rho \qquad (14)$$

where, $\hat{P}_V = 6.1078 \times 10^{\frac{7.5(T_C + \Delta T)}{T_C + 237.3 + \Delta T}} \cdot \text{RH}$.

From **Eq. 7**, if the air density has a deviation $\Delta\rho$, the effect on the wind speed calculation can be written as

$$\frac{\hat{v}_1}{v_1} = \left[\frac{P}{2(\rho + \Delta\rho)S v_d^2} + v_d \right] \Big/ v_1 \qquad (15)$$

If the measured wind speed v_d has a deviation Δv_d with the actual wind speed in the wind rotor plane, the effect on the wind speed calculation can be written as

$$\frac{\hat{v}_1}{v_1} = \left[\frac{P}{2\rho S(v_d + \Delta v_d)^2} + (v_d + \Delta v_d) \right] \Big/ v_1 \qquad (16)$$

Figure 6 shows the calculation results of the uncertainty effect of environmental parameters. In **Figure 6A**, the basic value of relative humidity is set to 0.8, and then given a deviation from -20 to 20%, the air density changes under different temperature conditions are shown. The higher the temperature is, the more significant the effect on air density is. Overall, deviations in relative humidity have little effect on air density. For example, at a temperature of 20°, a 20% deviation in humidity has only an effect of 0.15%. In **Figure 6B**, the temperature deviation is set in the range of -3°C to +3°C, and the fitting curves of air density change under three temperature base values are given. The three fitting curves basically coincide. Furthermore, the higher the temperature is, the lower the air density is. The numerical results show that the air density changes by 1.1% when the temperature deviation is
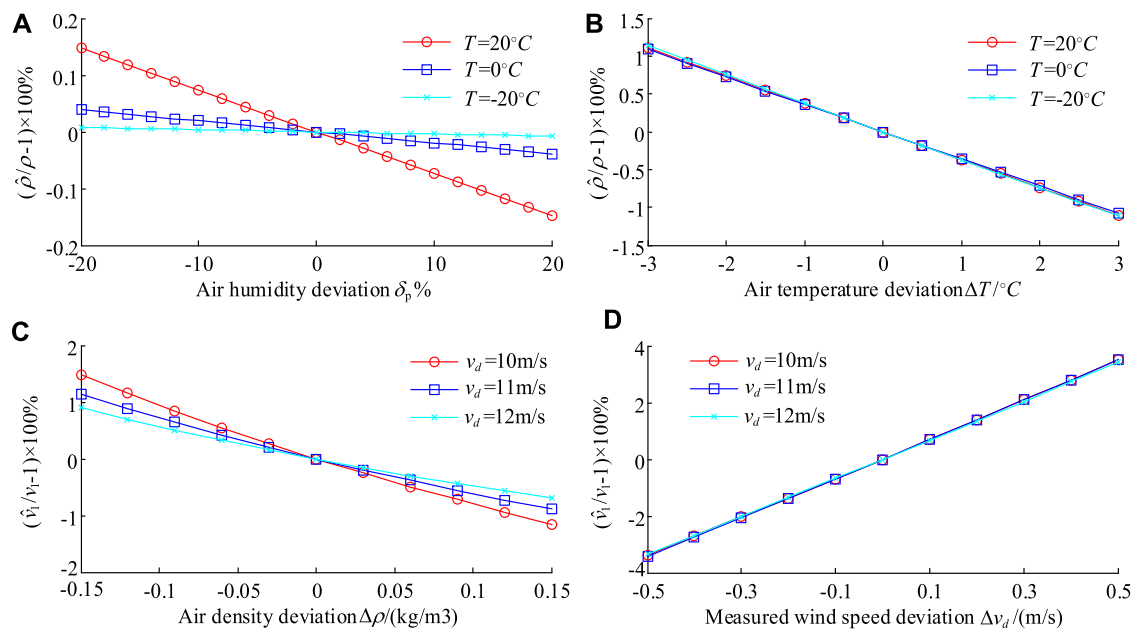
FIGURE 6 | Uncertainty effect of air humidity, air temperature, air density, and wind speed. **(A)** Air humidity deviation p%. **(B)** Air temperature deviation T /C. **(C)** Air density deviation /(kg/m3). **(D)** Measured wind speed deviation v d /(m/s).
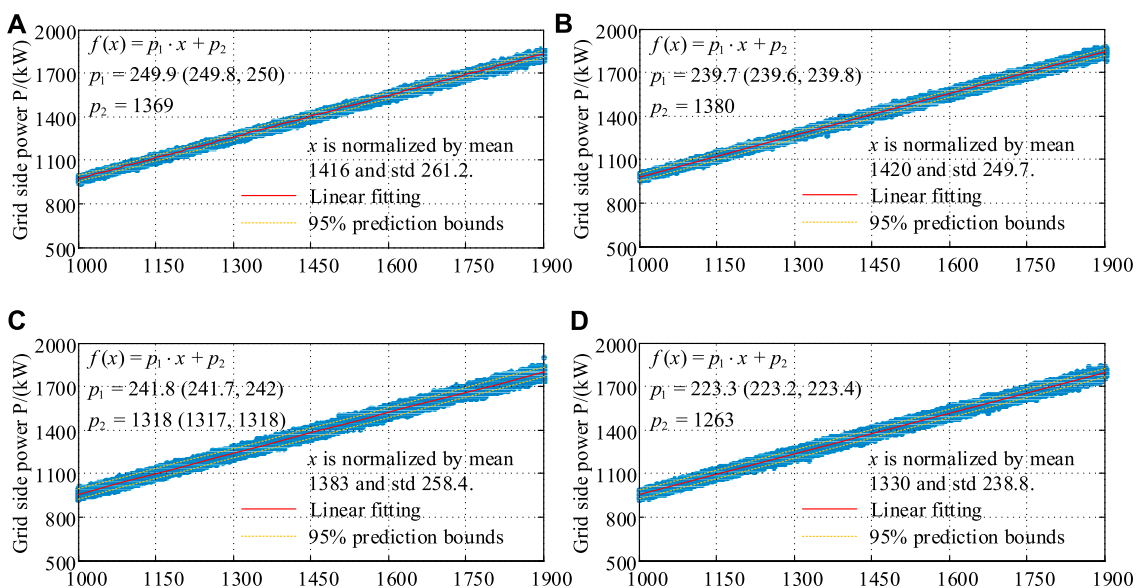
FIGURE 7 | Generator side power and grid side power. **(A)** WT1; **(B)** WT2; **(C)** WT3; **(D)** WT4.

3°C. In **Figure 6C**, the basic air density value is 1.2 kg/m³, the turbine power is set to be 2000 kW, and the given deviation ranges from -0.15 kg/m³ to 0.15 kg/m³. When the measured wind speed v_d is different, the effect of air density deviation on the corrected wind speed is slightly different. The smaller the measured wind speed v_d, the larger the corresponding effect. When v_d is 10 m/s and the air density deviation is 0.15 kg/m³, the effect on the corrected wind speed is 1.15%. In **Figure 6D**,

the given deviation of the measured wind speed ranges from -0.5 m/s to 0.5 m/s. Relatively, the deviation of the measured wind speed significantly influences the calculation result of the corrected wind speed. For instance, when the measured wind speed v_d is 10 m/s, the maximum effect on the corrected wind speed is 3.5%.

· Reliability assessment of wind speed and power data

TABLE 3 | Generator side power and grid side power based on fitting equation (kW).

P_{gen} P_{grid}	1,000	1,100	1,200	1,300	1,400	1,500	1,600	1700	1800	1900
WT1	971	1,067	1,162	1,258	1,354	1,449	1,545	1,641	1736	1832
WT2	977	1,073	1,169	1,264	1,361	1,457	1,553	1,649	1745	1841
WT3	960	1,053	1,147	1,240	1,334	1,427	1,521	1,615	1708	1802
WT4	954	1,048	1,141	1,235	1,328	1,422	1,515	1,609	1702	1796
$\frac{\max(P_{grid}) - \min(P_{grid})}{P_{gen}}$	0.023	0.023	0.023	0.022	0.024	0.023	0.024	0.024	0.024	0.024

FIGURE 8 | Four wind turbines in a mountain wind farm.

In the SCADA data, the generator side power (P_{gen}) and the grid side power (P_{grid}) of the converter are recorded simultaneously. It is believed that the two data will not have synchronously deviated. Therefore, the reliability can be judged by comparing the relationship between the two sets of power data. Specifically, the generator side power is given by the abscissa, and the grid side power is provided by the ordinate. The linear fitting curve and equation are provided by the scatter relationship between the two. In **Figure 7**, the generator-side power and grid-side power of the four wind turbines are shown. For the convenience of analysis, only the data of the generator side power between 1,000 kW and 1900 kW is selected. The relationship scatters between the two are very close and linearly related. In **Table 3**, the generator side power and grid side power based on the fitting equation are given, where the grid side power is calculated using the fitting equation. For a given generator side power, the calculated grid side power is different for different wind turbines. Overall, the power data of WT1 and WT2 are close to each other, while the power data of WT3 and WT4 are close to each other. To quantify the differences between different wind turbines, a relative difference $(\max(P_{grid}) - \min(P_{grid}))/P_{gen}$ is given in row 5 of **Table 2**. Under different power conditions, the values are basically the same, which shows that the relationship between generator side power and grid side power is stable. Moreover, the maximum grid side power always appears on WT2, and the minimum grid side power always appears on WT4.

To analyze the power curve of wind turbines more comprehensively, four wind turbines in a mountain wind farm in south China are investigated. The specific topography of the wind farm is shown in **Figure 8**. Their power curves are shown in **Figure 9**. Since **Figure 5** has given the power scatter of 1# wind turbine (WT1), **Figures 9A–C** only show the power scatters of 2# (WT2), 3# (WT3), and 4# (WT4). **Figure 9D** shows the power fitting curves of 4 wind turbines. The power curves of the four wind turbines do not entirely overlap, which seems to mean that although they are of the same type, the actual operating performance is always different. Of course, this difference cannot be ruled out due to the uncertainty of the data used. The benefit of obtaining these power curves is to provide a basis for further analysis of the performance of wind turbines. WT3 and WT4 are the closest among these power curves, WT2 is the next, and WT1 has the farthest deviation. In terms of the wind speed required to reach the designed rated power (2000 kW), WT1 is 10 m/s, WT2 is 10.9 m/s, WT3 is 11.4 m/s, and WT4 is 11.3 m/s. Since the design rated wind speed of the wind turbine is 10.5 m/s, WT1 is less than the rated wind speed, and the other three wind turbines are all greater than the designed rated wind speed. There may be several reasons for this phenomenon. 1) The power curve of WT1 deviates significantly from the other three wind turbines, which may be caused by the inaccurate anemometer of WT1. 2) If the measurement results of the anemometer are accurate, it is likely that the yaws of WT2, WT3, and WT4 have a large deviation, resulting in the need for greater wind speed to obtain the same power output. 3) If the
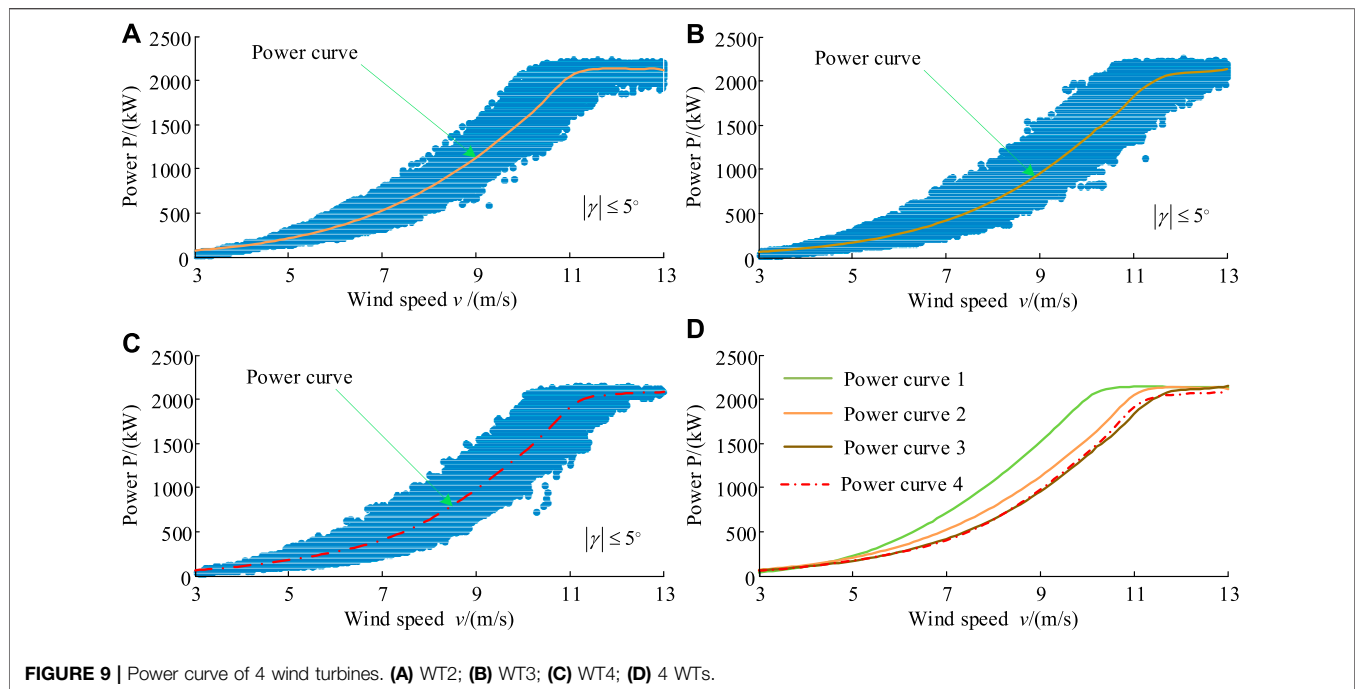
FIGURE 9 | Power curve of 4 wind turbines. **(A)** WT2; **(B)** WT3; **(C)** WT4; **(D)** 4 WTs.

previous two assumptions do not exist, then it should be caused by sensor deviation or the different performance characteristics of different wind turbines.

CONCLUSION

In this paper, a novel idea is proposed to obtain the actual power curve of wind turbines. A series of effective measures are taken to deal with the zero and null values in the original SCADA data, consider the comprehensive result of wind on the wind turbine power and correct the deviation between the measured wind speed and the real wind speed. The Gaussian fitting algorithm is used to fit the wind power curve; the power curve characteristics before and after the wind speed correction are compared and analyzed. Also, the characteristics of the power curves of four wind turbines are compared and analyzed. The results show that among these power curves, WT3 and WT4 are the closest, followed by WT2, and WT1 has the largest deviation from the other three wind turbines. The wind speed required for different turbines to reach the designed rated power is different from the actual power curves. Specifically, WT1 is 10 m/s, WT2 is 10.9 m/s, WT3 is 11.4 m/s, and WT4 is 11.3 m/s.

For the research topic of this paper, it is necessary to deepen it in the future further. For example, the research object of this paper is direct-drive wind turbines, and its algorithm can be transplanted to doubly-fed wind turbines or other types of wind turbines in the future. On the other hand, the power curve of wind turbines may evolve with the increase of service time, so historical

SCADA data can be used to observe its historical evolution trend further. More importantly, according to the characteristics of the power curve and the actual service conditions in wind farms, the operation strategy and maintenance strategy of wind turbines can be further studied.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

JD contributing to the conception and writing the manuscript. HZ data processing and analysis. FZ structure design and paragraph organization. HC writing assistance and English proofreading. ML data processing and analysis assistance.

FUNDING

REFERENCES

Astolfi, D., Castellani, F., Lombardi, A., and Terzi, L. (2021). Multivariate SCADA Data Analysis Methods for Real-World Wind Turbine Power Curve Monitoring. *Energies* 14 (4), 1105. doi:10.3390/en14041105

Astolfi, D., Malgaroli, G., Spertino, F., Amato, A., Lombardi, A., and Terzi, L. (2021). "Long Term Wind Turbine Performance Analysis through SCADA Data: A Case Study," in *2021 IEEE 6th International Forum on Research and Technology for Society and Industry (RTSI)* (IEEE), 7–12.

Bakir, I., Yildirim, M., and Ursavas, E. (2021). An Integrated Optimization Framework for Multi-Component Predictive Analytics in Wind Farm Operations & Maintenance. *Renew. Sustain. Energy Rev.* 138, 110639. doi:10.1016/j.rser.2020.110639

Carullo, A., Ciocia, A., Malgaroli, G., and Spertino, F. (2021). An Innovative Correction Method of Wind Speed for Efficiency Evaluation of Wind Turbines. *ACTA IMEKO* 10 (2), 46–53. doi:10.21014/acta_imeko.v10i2.1037

Ciulla, G., D'Amico, A., Di Dio, V., and Lo Brano, V. (2019). Modelling and Analysis of Real-World Wind Turbine Power Curves: Assessing Deviations from Nominal Curve by Neural Networks. *Renew. Energy* 140, 477–492. doi:10.1016/j.renene.2019.03.075

Dai, J., Cao, J., Liu, D., Wen, L., and Long, X. (2016). Power Fluctuation Evaluation of Large-scale Wind Turbines Based on SCADA Data. *IET Renew. Power Gener.* 11 (4), 395–402. doi:10.1049/iet-rpg.2016.0124

Dai, J., Liu, D., Wen, L., and Long, X. (2016). Research on Power Coefficient of Wind Turbines Based on SCADA Data. *Renew. Energy* 86, 206–215. doi:10.1016/j.renene.2015.08.023

Dai, J., Tan, Y., and Shen, X. (2019). Investigation of Energy Output in Mountain Wind Farm Using Multiple-Units SCADA Data. *Appl. Energy* 239, 225–238. doi:10.1016/j.apenergy.2019.01.207

Dai, J., Yang, X., Hu, W., Wen, L., and Tan, Y. (2018). Effect Investigation of Yaw on Wind Turbine Performance Based on SCADA Data. *Energy* 149, 684–696. doi:10.1016/j.energy.2018.02.059

Dai, J., Yang, X., and Wen, L. (2018). Development of Wind Power Industry in China: A Comprehensive Assessment. *Renew. Sustain. Energy Rev.* 97, 156–164. doi:10.1016/j.rser.2018.08.044

Dawn, S., Tiwari, P. K., Goswami, A. K., Singh, A. K., and Panda, R. (2019). Wind Power: Existing Status, Achievements and Government's Initiative towards Renewable Power Dominating India. *Energy Strategy Rev.* 23, 178–199. doi:10.1016/j.esr.2019.01.002

Dhunny, A. Z., Timmons, D. S., Allam, Z., Lollchund, M. R., and Cunden, T. S. M. (2020). An Economic Assessment of Near-Shore Wind Farm Development Using a Weather Research Forecast-Based Genetic Algorithm Model. *Energy* 201, 117541. doi:10.1016/j.energy.2020.117541

Gao, L., and Hong, J. (2021). Data-driven Yaw Misalignment Correction for Utility-Scale Wind Turbines. *J. Renew. Sustain. Energy* 13 (6), 063302. doi:10.1063/5.0056671

Gonzalez, E., Stephen, B., Infield, D., and Melero, J. J. (2017). On the Use of High-Frequency SCADA Data for Improved Wind Turbine Performance Monitoring. *J. Phys. Conf. Ser.* 926, 012009. doi:10.1088/1742-6596/926/1/012009

Gonzalez, E., Stephen, B., Infield, D., and Melero, J. J. (2019). Using High-Frequency SCADA Data for Wind Turbine Performance Monitoring: A Sensitivity Study. *Renew. Energy* 131, 841–853. doi:10.1016/j.renene.2018.07.068

Hansen, M. O. L. (2008). *Aerodynamics of Wind Turbines*. 2nd Edition. Eaethscan, UK and USA.

Karamichailidou, D., Kaloutsa, V., and Alexandridis, A. (2021). Wind Turbine Power Curve Modeling Using Radial Basis Function Neural Networks and Tabu Search. *Renew. Energy* 163, 2137–2152. doi:10.1016/j.renene.2020.10.020

Kim, D.-Y., Kim, Y.-H., and Kim, B.-S. (2021). Changes in Wind Turbine Power Characteristics and Annual Energy Production Due to Atmospheric Stability, Turbulence Intensity, and Wind Shear. *Energy* 214, 119051. doi:10.1016/j.energy.2020.119051

Li, L., Mu, X., Macfarlane, C., Song, W., Chen, J., Yan, K., et al. (2018). A Half-Gaussian Fitting Method for Estimating Fractional Vegetation Cover of Corn Crops Using Unmanned Aerial Vehicle Images. *Agric. For. Meteorology* 262, 379–390. doi:10.1016/j.agrformet.2018.07.028

Lydia, M., Kumar, S. S., Selvakumar, A. I., and Prem Kumar, G. E. (2014). A Comprehensive Review on Wind Turbine Power Curve Modeling Techniques. *Renew. Sustain. Energy Rev.* 30, 452–460. doi:10.1016/j.rser.2013.10.030

Manobel, B., Sehnke, F., Lazzús, J. A., Salfate, I., Felder, M., and Montecinos, S. (2018). Wind Turbine Power Curve Modeling Based on Gaussian Processes and Artificial Neural Networks. *Renew. Energy* 125, 1015–1020. doi:10.1016/j.renene.2018.02.081

Marčiukaitis, M., Žutautaitė, I., Martišauskas, L., Jokšas, B., Gecevičius, G., and Sfetsos, A. (2017). Non-linear Regression Model for Wind Turbine Power Curve. *Renew. Energy* 113, 732–741. doi:10.1016/j.renene.2017.06.039

Mehrjoo, M., Jafari Jozani, M., and Pawlak, M. (2021). Toward Hybrid Approaches for Wind Turbine Power Curve Modeling with Balanced Loss Functions and Local Weighting Schemes. *Energy* 218, 119478. doi:10.1016/j.energy.2020.119478

Mehrjoo, M., Jafari Jozani, M., and Pawlak, M. (2020). Wind Turbine Power Curve Modeling for Reliable Power Prediction Using Monotonic Regression. *Renew. Energy* 147, 214–222. doi:10.1016/j.renene.2019.08.060

Pandit, R. K., Infield, D., and Kolios, A. (2020). Gaussian Process Power Curve Models Incorporating Wind Turbine Operational Variables. *Energy Rep.* 6, 1658–1669. doi:10.1016/j.egyr.2020.06.018

Rogers, T. J., Gardner, P., Dervilis, N., Worden, K., Maguire, A. E., Papatheou, E., et al. (2020). Probabilistic Modelling of Wind Turbine Power Curves with Application of Heteroscedastic Gaussian Process Regression. *Renew. Energy* 148, 1124–1136. doi:10.1016/j.renene.2019.09.145

Saint-Drenan, Y.-M., Besseau, R., Jansen, M., Staffell, I., Troccoli, A., Dubus, L., et al. (2020). A Parametric Model for Wind Turbine Power Curves Incorporating Environmental Conditions. *Renew. Energy* 157, 754–768. doi:10.1016/j.renene.2020.04.123

Seo, S., Oh, S.-D., and Kwak, H.-Y. (2019). Wind Turbine Power Curve Modeling Using Maximum Likelihood Estimation Method. *Renew. Energy* 136, 1164–1169. doi:10.1016/j.renene.2018.09.087

Sun, H., Qiu, C., Lu, L., Gao, X., Chen, J., and Yang, H. (2020). Wind Turbine Power Modelling and Optimization Using Artificial Neural Network with Wind Field Experimental Data. *Appl. Energy* 280, 115880. doi:10.1016/j.apenergy.2020.115880

Villanueva, D., and Feijóo, A. (2018). Comparison of Logistic Functions for Modeling Wind Turbine Power Curves. *Electr. Power Syst. Res.* 155, 281–288. doi:10.1016/j.epsr.2017.10.028

Virgolino, G. C. M., Mattos, C. L. C., Magalhães, J. A. F., and Barreto, G. A. (2020). Gaussian Processes with Logistic Mean Function for Modeling Wind Turbine Power Curves. *Renew. Energy* 162, 458–465. doi:10.1016/j.renene.2020.06.021

Xu, N., Sheng, L., Chen, C., Li, Y., Su, T., Zhao, B., et al. (2019). Application of Gaussian Fitting to the Fast Search of Pulsar Periodic. *Optik* 198, 163253. doi:10.1016/j.ijleo.2019.163253

Yesilbudak, M. (2018). Implementation of Novel Hybrid Approaches for Power Curve Modeling of Wind Turbines. *Energy Convers. Manag.* 171, 156–169. doi:10.1016/j.enconman.2018.05.092

Experimental study of dynamic characteristics of an ultra-large jacket offshore wind turbine under wind and wave loads using aero-hydro-structural elastic similarities

Gen Xiong[1,2], Dongzhe Lu[3,4]*, Zuxing Pan[1,2,5], Wenhua Wang[3,4], Xin Li[3,4] and Qingquan Liu[1,2]

[1]Key Laboratory of Far-shore Wind Power Technology of Zhejiang Province, Hangzhou, China, [2]Renewable Energy Engineering Institute, Power China Huadong Engineering Corporation Limited, Hangzhou, China, [3]State Key Laboratory of Coastal and Offshore Engineering, Dalian University of Technology, Dalian, China, [4]Faculty of Infrastructure Engineering, Institute of Earthquake Engineering, Dalian University of Technology, Dalian, China, [5]School of Civil Engineering and Architecture, Guangxi University for Nationalities, Nanning, China

Owing to the difficulties in the scaled rotor-nacelle assembly (RNA) and support structure design, and alleviation of small scaling effects, the limited dynamic model tests are conducted for the jacket offshore wind turbines (OWTs), which are extensively constructed in the offshore wind farms located in the depth of 40−50 m. To address this limitation, an integrated test method based on aero-hydro-structural elastic similarities is proposed in this study. It comprises a performance-scaled RNA model and a scaled support structure model. A redesigned blade model is adopted in the scaled RNA model to ensure the similarities of aerodynamic thrust loads without modifications of the scaled test winds. Moreover, auxiliary scaled drivetrain and blade pitch control are designed to simulate the operational states of a practical OWT. The scaled model of the OWT support structure is fabricated based on the joint hydro-structural elastic similarity, and the small scaling effects are mitigated by introducing sectional bending stiffness similarities. Subsequently, the dynamic model tests of an ultra-large jacket OWT under wind-only, wave-only, and combined wind and wave conditions are carried out. The accuracy of the fabricated OWT test model is validated based on the recorded responses, and the influence of the dominant frequencies on the dynamic responses of the OWT model is quantitatively evaluated using the wavelet packet-based energy analysis method. Further, the coupling mechanisms of the scaled OWT model under typical wind and wave loads are investigated, and the interactions between the environmental loads and OWT motions are proved.

KEYWORDS

offshore wind turbine, scaled model design, basin model test, dynamic characteristics, wavelet packet-based energy analysis method, wind and wave loads

1 Introduction

Offshore wind energy is gradually becoming a mainstream renewable energy source, and the massive deployment of offshore wind turbines (OWTs) has accelerated this trend over the past decade (Ren et al., 2022). According to the Global Wind Energy Council (GWEC, 2022), the installed capacities of offshore wind farms reached 21.1 GW in 2021, and that in China exceeded 80%. Although superior wind fields enhance the output of OWTs compared to that of onshore wind fields (Park et al., 2021), the interactions among the environmental loads, rotor-nacelle assembly (RNA), and support system pose several problems, such as motion of the support structure, resulting in increased asymmetrical aerodynamic loading on the rotor blades (Wang et al., 2022), particularly for ultra-large OWTs. To accurately predict the structural responses of OWTs under complex environmental conditions, numerous fully coupled simulation tools have been developed, such as the widely used GH Bladed (Bossanyi, 2011), SIMA (DNV, 2018), FAST v8/ OpenFAST (Jonkman and Jonkman, 2016; Jonkman and Sprague, 2021), and HAWC2 (Larsen and Hansen, 2007). These coupled simulation tools employ blade element momentum theory and generalized dynamic wake model in aeroelastic analysis. Hydrodynamic loads on the highly slender and large structures are typically calculated *via* finite element analysis based on the dimensions of the support system using the Morison formula and potential flow theory, respectively. Meanwhile, the differences in the established RNA models in the above simulation tools should be pointed out. The flexibilities of rotor blades are generally simulated using the first two flapwise and the first edgewise modes in FAST v8 and OpenFAST, which is remarkably different from the corresponding finite element model using beam element in SIMA and HAWC2.

Presently, the extensively studies about the coupling mechanism of typical bottom-fixed and floating OWTs are carried out by researchers using the previously described simulation tools. For example, Kim et al. (2016) established the fully coupled numerical model of a monopile and jacket OWT in GH Bladed and investigated the differences in the structural responses between the monopile and jacket OWT. Ren et al. (2022) carried out the dynamic analysis of a multi-column tension leg platform floating OWT under operational and extreme environmental conditions in FAST v7 and investigated the tendon failure to examine the performance of OWT based on the accidental limit states specified in the design code DNV-RP-0286. Zhao et al. (2021) and Putri et al. (2020) studied the dynamic responses of a semi-submersible and spar-floating OWT using OpenFAST and SIMA, respectively, to elucidate the coupling effects between the RNA and mooring system. To verify the accuracy of numerical simulations, comparisons among the various

coupled simulation tools for OWTs are performed as part of the IEA OC4 project (Robertson et al., 2014). According to the studies of Larsen et al. (2014), discrepancies in the predicted OWT coupled motions using different numerical tools are attributed to the variability of the essential parameters in the different numerical models. Therefore, the results obtained by numerical simulation need to be further compared and analyzed, and the dynamic model test provides an effective approach (Zeng et al., 2022).

For OWTs under wind and wave loading conditions, besides the structural elastic similarity, the aerodynamic and hydrodynamic loads similarities should also be satisfied to ensure high quality and accuracy in the OWT dynamic tests. For aerodynamic load similarity, it means that the viscous and inertia forces should be appropriately scaled, and the Reynolds number similarity is priorly applied in this case (Çengel and Cimbala, 2006). Hydrodynamic inertia and gravity loads dominate the motions of offshore structures under wave excitation, and the associated Froude number similarity is generalized in the physical model tests of offshore structures under wave loads (Chakrabarti, 1994). Moreover, the differences in the required flow velocity scale ratios between the Reynolds and Froude number similarities should be considered. Therefore, appropriate coordination of the Froude and Reynolds number similarities is a critical issue in OWT dynamic tests.

Various experimental methods have been devised to ensure accurate modeling of aerodynamic and hydrodynamic loads in OWT tests. Wang et al. (2017) performed dynamic model tests of a pentapod OWT under earthquake, wind, wave, and current, and hydrodynamic loading conditions and the OWT model were scaled based on the Froude number and structural elastic similarities. The aerodynamic thrust loads were appropriately scaled using an equivalent circular disc, whereas the gyroscopic effects of the rotor system were neglected. Froude number similarity was applied in the semi-submersible floating OWT model design, and an approximately geometrically scaled RNA model was fabricated by Martin (2011). Subsequently, dynamic tests of the semi-submersible floating OWT model under winds and waves were performed, and the test winds were calibrated to ensure accurate scaling of aerodynamic loads. Ahn and Shin (2019) reported that aggerated wind loads occurred on the support structure model owing to the improved test winds used to accurately model the aerodynamic loads on the geometric-scaled RNA model, which impaired the accuracy of the OWT dynamic model tests. To address this, Du et al. (2016) and Ahn and Shin (2020) proposed a performance-scaled RNA model with updated blade airfoils for the OWT model design, and the accuracy was verified by comparing the results with the numerical results for the prototype RNA.

In addition, according to the recently released offshore wind energy development guidelines effective over the next 5 years in China, the exploration of offshore wind energy at water depths of

30–50 m and the application of ultra-large WTs are primary concerns. Therefore, the potential of commercial application of jacket substructures should be emphasized. Before commercialization, the dynamic characteristics and coupling mechanisms of ultra-large jacket OWTs should be adequately investigated. Hence, dynamic model tests of the ultra-large jacket OWT must be performed to validate numerical analyses and address their limitations. These observations indicate that the scaled RNA model design directly determines the quality of the tests, whereas the design method for ultra-large RNA still warrants further study and validation because of the oversized blade dimensions and inherent complexities of servo systems. Moreover, owing to significant discrepancies between the longitudinal principle and sectional dimensions, small scaling effects should be considered in the scaled model design of ultra-large support structures.

To address these technical challenges associated with ultra-large jacket OWT dynamic model tests, an integrated test method based on aero-hydro-structural elastic similarities is proposed, comprising a performance-scaled RNA model and a scaled support structure model using joint hydro-structural elastic similarities. The remainder of this paper is organized as follows. The derivation of essential similarities for the scaled OWT model design is presented in Section 2. The redesigned scaled RNA model based on the released aerodynamic similarities and the scaled support structure model using hydro-structural elastic similarities are introduced. This section also describes the sensor arrangement and test facilities. The applied wavelet packet-based energy analysis method used for test data processing is introduced in Section 3. Furthermore, the validations of the fabricated OWT test model and the analysis of OWT coupling mechanisms based on the recorded data are presented in Section 4. Finally, based on the observed experimental phenomena, the conclusions regarding the OWT test model design method and coupling mechanisms under different environmental conditions and operation states are summarized in Section 5.

2 Physical dynamic test model design

2.1 Joint similarities for OWT model tests

For the OWT dynamic model tests, in addition to the structural elastic similarity expressed in Eq. 1, the aero and hydro similarities should be ensured, which are represented by the Reynolds and Froude number similarities defined in Eqs 2, 3, respectively. However, as defined in Eqs 2, 3, differences exist in the required velocity scale ratios between the Reynolds and Froude number similarities. Therefore, these similarities cannot be concurrently and strictly satisfied in the dynamic model tests. Considering that the objective of this study is to investigate the dynamic characteristics of an OWT under

typical wind and wave loads, strict Reynolds number similarity is substituted by performance-scaled similarities, which include similarities of aerodynamic thrust loads and essential operational parameters. Moreover, hydro-structural elastic similarity is formulated by combining Eqs 1–3, as expressed in Eq. 4. Based on this equation, in addition to the length scale ratio λ_L, an inertia radius scale ratio λ_r is used in this study to alleviate the small scaling effects in the design of the scaled OWT support structure model sectional geometries (Huan et al., 2022). To ensure the accuracy of the dynamic model tests, an appropriate scale ratio should be determined before the design of the scaled OWT model. Based on the geometries of the prototype OWT, water depth, and capacities of the experimental facilities, the length scale ratio is designated as 1/75. Subsequently, the similarities between the additional parameters can be obtained, as listed in Table 1.

$$\lambda_t^2 = \lambda_L^4 \cdot \lambda_\rho \cdot \lambda_E^{-1} \cdot \lambda_r^{-2} \tag{1}$$

$$\lambda_v = 1/\lambda_L \tag{2}$$

$$\lambda_v = \sqrt{\lambda_L} \tag{3}$$

$$\lambda_L^3 \cdot \lambda_r^{-2} = \lambda_E \tag{4}$$

where λ_t is the time scale ratio; λ_ρ, λ_E, and λ_v are the structural material density, elastic modulus, and velocity scale ratio, respectively.

2.2 Performance-scaled test model design

2.2.1 Prototype OWT concept

The prototype jacket OWT is illustrated in Figure 1A. The original OWT comprised an upwind WT system and jacket substructure. The RNA of the upper WT system was identical to that of the DTU 10 MW baseline WT (Bak et al., 2013). According to the recorded environmental conditions in the southeastern offshore regions in China, the distance from the top of the jacket substructure to the mean sea level needs to be 30.15 m to ensure structural safety, and the hub height of DTU 10 MW WT is 119.00 m, so the tower height should be 88.85 m. The original tower height of DTU 10 MW WT is 115.63 m. Therefore, the tower height is redesigned as 85.48 m to satisfy the requirements under the practical environments. The basic parameters of the RNA and the tower are listed in Table 2. As depicted in Figure 1A, the mud braces and four levels of X-braces are mounted along the four-legged jacket to ensure steady operation of the WT system and structural safety at a water depth of 40.00 m. The geometry and structure of the support system are also shown in this figure.

2.2.2 Scaled rotor-nacelle assembly model

Because of the length scale ratio and prototype RNA used in this study being consistent with the European Union

TABLE 1 Essential model parameter scale ratios defined by the ratio of model to prototype.

Parameters	Length	Density	Mass	Velocity	Acceleration	Rotor speed	Pitch angle	Time	Frequency	Force
Dimension	$[L]$	$[ML^3]$	$[M]$	$[LT^{-1}]$	$[LT^{-2}]$	$[\frac{1}{T}]$	—	$[T]$	$[\frac{1}{T}]$	$[MLT^{-2}]$
Similarity	λ_L	$\lambda_\rho = 1$	λ_L^3	$\sqrt{\lambda_L}$	$\lambda_g = 1$	λ_Ω	λ_θ	$\sqrt{\lambda_L}$	$\frac{1}{\sqrt{\lambda_L}}$	λ_L^3
Scale ratio	$\frac{1}{75}$	1.00	$\frac{1}{421\,875}$	0.11	1.00	8.66	1.00	0.11	8.66	$\frac{1}{421\,875}$

LIFES50 + project, the recommended scaled blade model design in the LIFES50 + project is adopted (Lifes50, 2021). The updated SD7032 airfoil is selected in the performance-scaled blade model design and a circular section is used at the blade root to connect the blade and hub models. The fabricated scaled blade model is shown in Figure 1B. Moreover, the sectional chord length is altered, and the maximum chord length could reach 0.115 m to withstand test wind conditions generated using a simplified wind generation system, as depicted in Figure 1C. The mass of each blade model is strictly limited to 200 g; therefore, high-strength and low-density carbon fibers are used in the fabrication. Meanwhile, to ensure the smoothness of the blade model and fabrication quality, the steel mold shown in Figure 1D is utilized in the fabrication process.

A scaled drivetrain system comprising an actuator, torque sensor, and shaft is designed to ensure similarity of essential operational parameters, such as rotor speed and thrust load. In the scaled drivetrain arrangement shown in Figure 1E, the components are connected using couplers and tightened in an aluminum nacelle model. Sequentially, the scaled model of the nacelle is manufactured based on the drivetrain system arrangement.

An additional scaled blade pitch system is designed to deploy blade pitch control strategies during the tests, as shown in Figures 1F,G. Three-pitch actuators mounted in the transition piece between the blade root and hub model are used to regulate the blade pitch angles. The material of the hub model with a diameter of 0.06 m is the same as that of the scaled nacelle model, and the total mass of the scaled RNA model is 2,860 g, as listed in Table 3.

2.2.3 Scaled support structure model

Based on previous mechanical tests, polymethyl methacrylate (PMMA) is selected as the material for the scaled OWT support model owing to its advantages such as stable properties and convenient fabrication. The measured elastic modulus of PMMA is 3.85 GPa, with a density of 1,198 kg/m³. Subsequently, the inertial radius scale ratio for the scaling of sectional geometries is calculated to be 1/88.8 according to the proposed hydro-structural elastic similarity and length scale ratio. The principal dimensions of the support structural model, such as the height and

sectional out-diameter of the tower and jacket model are determined by the length scale ratio, the member sectional thickness is scaled using the inertial radius scale ratio to alleviate small scaling effects and ensure fabrication accuracy. The scaled geometries of the tower and jacket models are shown in Figure 1H. Moreover, the density scale ratio is assumed to be 1.0; therefore, the additional weights are uniformly distributed along the support model to satisfy this assumption, except in the splash zone, as depicted in Figure 1I.

2.2.4 Scaled test cases and environmental conditions

Considering that the primary objective is to experimentally investigate the dynamic characteristics and coupling mechanisms of the jacket OWT, typical design load cases (DLCs) are selected according to the measured environmental conditions and offshore standard DNV GL-ST-0437 (DNV GL, 2016). Subsequently, the scale ratios listed in Table 1 are used to scale the recorded environmental conditions, as listed in the model scale column in Table 4. Three typical winds (DLCs 1–3) covering the cut-in to cut-out wind range are selected to study the coupling mechanisms of the dynamic response of OWT under different normal operation states. Additional extreme wind test case 4 is applied to investigate the dynamic behaviors of a parked OWT with the feathered rotor blades. Meanwhile, DLCs 5–8 are performed to investigate the dynamic responses of OWT under regular waves. To reveal the coupling mechanisms, subsequent dynamic model tests are conducted under combined conditions (DLCs 9–12). During the tests, the scaled winds and waves are aligned along the fore-aft (F-A) direction, as shown in Figure 2A. The rotational speed of the RNA model is regulated using the drivetrain system model, and the blade pitch and pitch-to-feather control strategies are deployed using the pitch control system model.

2.3 Test equipment and sensor arrangement

A joint wind and wave simulation system developed by the State Key Laboratory of Coastal and Offshore Engineering at Dalian University of Technology was employed to generate scaled wind and
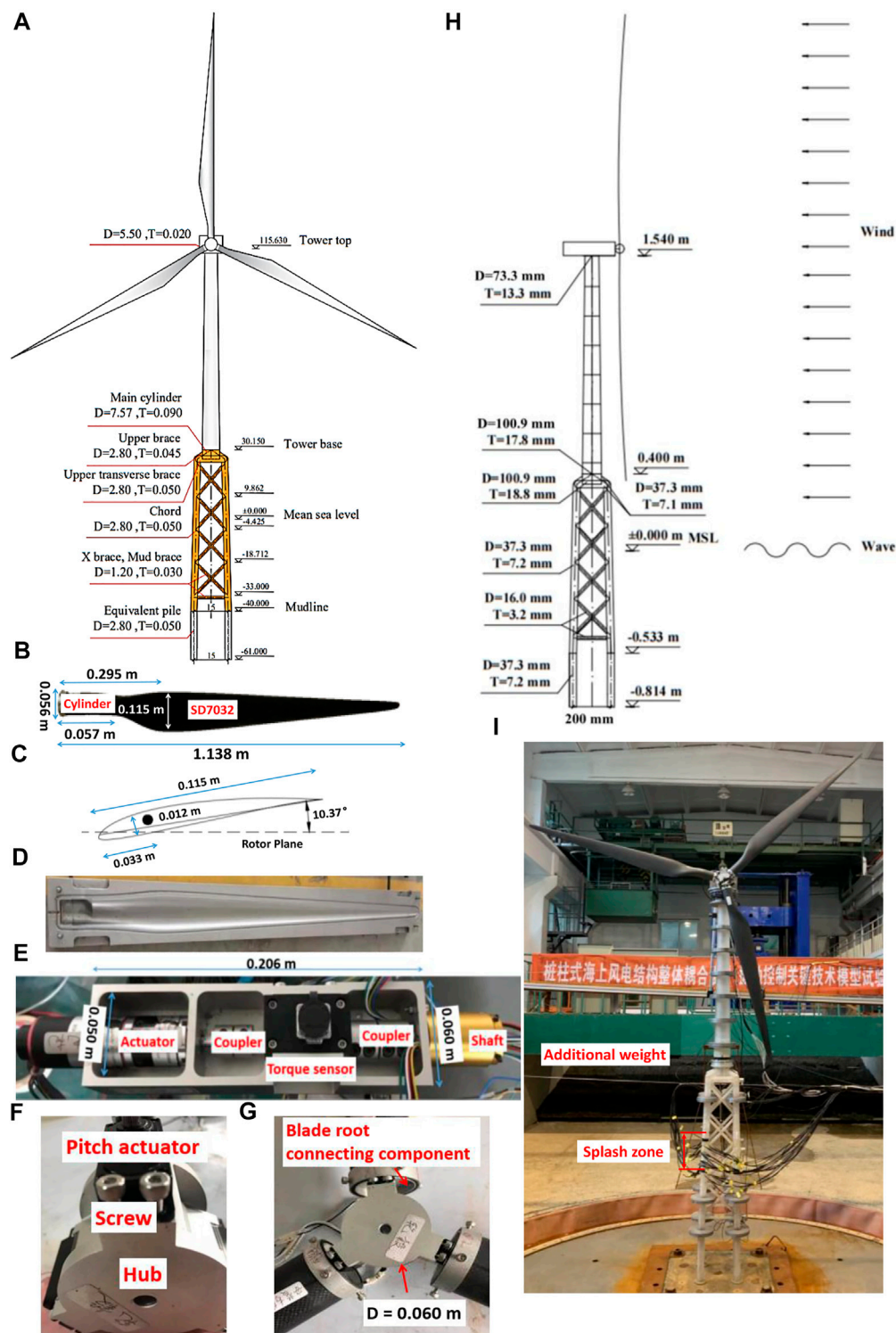
FIGURE 1
Design of performance-scaled test model. **(A)** Essential parameters of prototype jacket OWT (Unit: m); **(B)** carbon fiber blade model; **(C)** schematic of the blade model cross-section at maximum chord length; **(D)** steel mold; **(E)** arrangement of drivetrain system; **(F)** blade pitch system; **(G)** scaled hub model; **(H)** parameters of the support structure model; and **(I)** additional weights.

TABLE 2 Basic parameters of prototype OWT.

Characteristics	Value
Rating	10 MW
Rotor orientation and configuration	Upwind and 3 blades
Control	Variable speed, collective pitch
Single blade mass	41 732 kg
Hub mass	105 520 kg
Nacelle mass	446 036 kg
Hub height	119.00 m
Tower height	85.48 m
Tower top diameter and thickness	5.50 and 0.020 m
Tower base diameter and thickness	7.57 and 0.034 m
Rotor and hub diameter	178.30 and 5.60 m
Cut-in, rated, and cut-out wind speeds	4.0, 11.4, and 25.0 m/s
Cut-in and rated rotor speeds	6.0 and 9.6 rpm

wave conditions in the tests. Figure 2A shows the dynamic model tests of the jacket OWT model subjected to wind and wave loads. The model test equipment primarily included wave generation and wind generation systems, both of which are located on the left side of the basin, as shown in Figure 2B. Wave-absorber equipment is placed on the right side of the basin to reduce wave reflection. The dimensions of the basin are 22.0 × 5.4 × 1.2 m, and the maximum water depth is 1.0 m. The maximum height of generated waves is 0.33 m, and the allowable range of the wave period is 0.5–4 s. The usable area of the wind generation system is 2.54 × 2.54 m, and the maximum generated wind speed is 15 m/s.

The sensors used to record the input environmental conditions and dynamic responses of the OWT model are as follows. As depicted in Figure 1E, a torque sensor mounted in the nacelle was used to measure the torque of the designed drivetrain system, and the quality of the generated test wind fields was calibrated prior to testing using a wind sensor positioned at calibration points P1–P12, as shown in

TABLE 3 Masses of the major components of scaled RNA model.

Component	Nacelle	Actuator	Couplers	Torque sensor	Shaft	Hub	Pitch actuators	Connection components	Screws	Blades
Mass (g)	410	447	64	240	376	242	303	138	40	600

TABLE 4 Selected test cases.

Case no.	description	Full scale						Model scale (1:75)						
		V_{hub} (m/s)	H (m)	T (s)	Blade pitch (deg)	Rotor speed (rpm)	Water depth (m)	V_{hub} (m/s)	H (m)	T (s)	Blade pitch (deg)	Rotor speed (rpm)	Threefold blade passing frequency (Hz)	Water depth (m)
1	Steady wind	6.0			0.9	6.0	61.0	0.7			0.9	51.96	2.60	0.8
2		11.4			0.0	9.6		1.3			0.0	83.14	4.28	
3		18.0			15.2	9.6		2.1			15.2	83.14	4.28	
4		29.0			90.0	0.0		3.3			90.0	0.00		
5	Regular wave		1.50	8.7	0.9	0.00			0.02	1.0	0.9	0.00		
6			3.00	9.5	0.0	0.00			0.04	1.1	0.0	0.00		
7			4.50	10.4	15.2	0.00			0.06	1.2	15.2	0.00		
8			6.00	11.3	90.0	0.00			0.08	1.3	90.0	0.00		
9		6.0	1.50	8.7	0.9	6.0		0.7	0.02	1.0	0.9	51.96	2.60	
10	Steady wind and	11.4	3.00	9.5	0.0	9.6		1.3	0.04	1.1	0.0	83.14	4.28	
11	regular wave	18.0	4.50	10.4	15.2	9.6		2.1	0.06	1.2	15.2	83.14	4.28	
12		29.0	6.00	11.3	90.0	0.0		3.3	0.08	1.3	90.0	0.00		
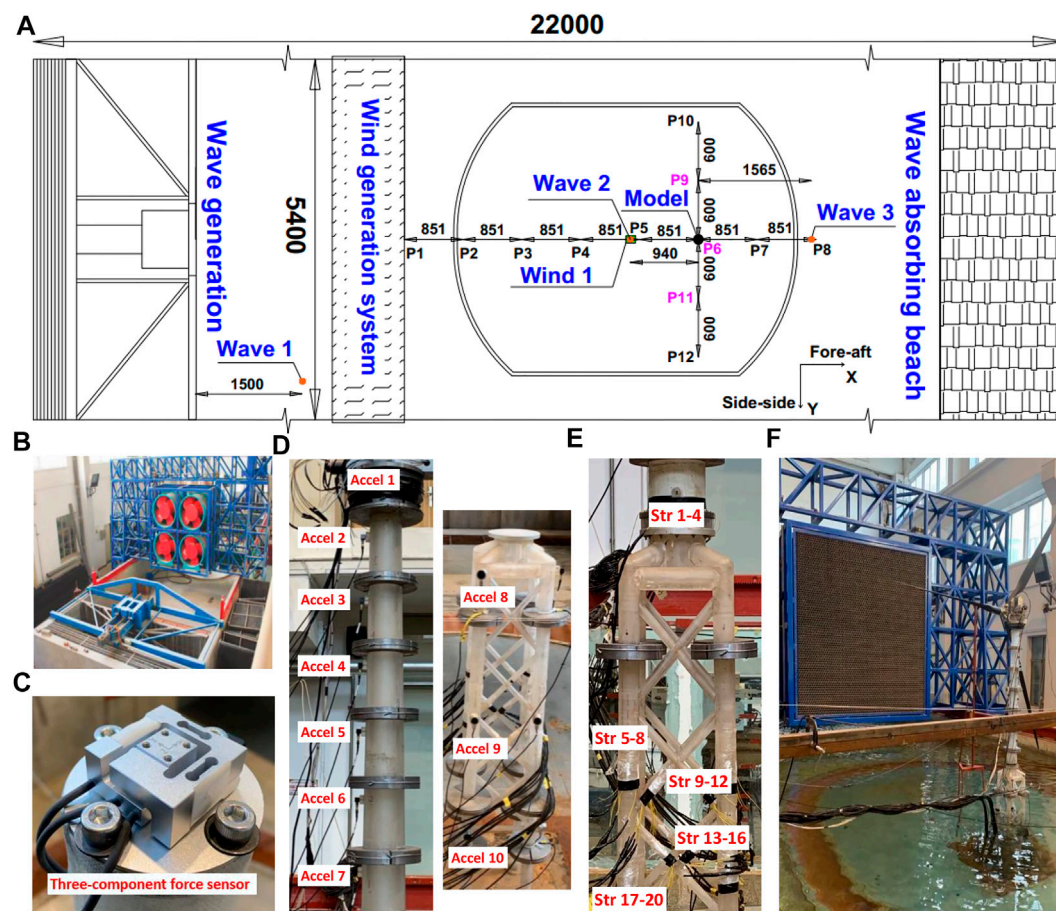
FIGURE 2
Primary sensor arrangement and test equipment. **(A)** Schematic of the model test basin (Unit: mm); **(B)** joint wave and wind load simulation system; **(C)** installed three-component load cell at the tower top; **(D)** acceleration sensor arrangement along model; **(E)** strain gauge arrangement; and **(F)** dynamic model testing of jacket OWT subjected to winds and waves.

Figure 2A. Moreover, the accuracy of the wave generator was measured by the wave gauge numbered "wave 1" positioned in front of the facility and wave gauge numbered "wave 2" placed around the OWT model. An additional wave gauge numbered "wave 3" was used in the tests to record the far field wave profiles. A force sensor with three translational degrees of freedom (DOFs) was installed at the top of the tower in the model to measure the thrust loads, as shown in Figure 2C. The measuring range and resolution of the force sensors on the DOFs are 100 and 0.1 N, respectively. Figure 2D shows ten acceleration sensors with a range of 100 m/s² uniformly arranged on the support structure in the F-A direction in the model to measure the motion of the structure under different loading conditions. Strain gauges were arranged at the tower base, and the local components of the jacket model measured the structural responses, as depicted in Figure 2E. The numbers and specifications of the sensors are listed in Table 5. In addition to sensor accuracy, the sensor mass was strictly

limited to eliminate the undesirable influence of installed sensors on local structural responses. For instance, the force gauge mass should be as low as 65 g to satisfy this requirement. The layout of the dynamic model tests for the jacket OWT is shown in Figure 2F.

2.4 Test wind field and wave calibrations

As depicted in Figure 2F, the wind generation system was pre-calibrated to ensure reliable generation of scaled wind fields during testing. Therefore, eight measurement points, P1–P8, were selected in the F-A direction, and the distance between pairs of points was set as 851 mm. The measured scaled wind speed of 2.1 m/s at the different points are listed in Table 6, and points P4–P6 are suitable for measurement compared with the theoretical values. Considering the fluctuations of the generated wind speed at these potential points, point P6 was selected as a

TABLE 5 Sensor specifications and arrangement in OWT dynamic tests.

No	Type	Number	Technical parameters
Wind 1	Wind sensor	1	Measuring range: 30 m/s
			Resolution: 0.001 m/s
Wave 1–3	Wave gauge	3	Measuring range: 30 cm
			Resolution: 0.03 cm
Force 1	Three-component load cell	1	Capacities: $F_x = F_y = F_z = 100$ N
			Resolution: 0.1 N
			Size: $4.0 \times 4.0 \times 2.0$ cm
			Mass: 65 g
Accel 1–10	(Single direction) Acceleration sensor	10	Capacity: 100 m/s^2
			Resolution: 0.004 m/s^2
			Size: $1.5 \times 1.0 \times 1.0$ cm
			Mass: 23 g
Str 1–20	Strain gauge	20	Capacity: 20000 με
			Resolution: 1 με

TABLE 6 Statistics of recorded wind speeds at different sites (Unit: m/s).

Item	Calibration site no.											
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
Mean	0.9	0.6	1.0	2.0	2.2	2.4	2.7	2.9	2.4	1.2	2.1	1.1
STD	0.03	0.13	0.14	0.16	0.14	0.11	0.11	0.12	0.42	0.31	0.23	0.25

TABLE 7 Comparison of wave parameters between measured and theoretical scaled wave around the OWT model installation point.

Wave condition	Regular wave 1		Regular wave 2		Regular wave 3		Regular wave 4	
Item and Unit	H (m)	T (s)	H (m)	T (s)	H (m)	T (s)	H (m)	T (s)
Measured in calibration test	0.018	0.99	0.041	1.09	0.061	1.17	0.076	1.30
Theoretical scaled	0.020	1.00	0.040	1.10	0.060	1.20	0.080	1.30

feasible location. Moreover, the qualities of the wind speed at point P6 were evaluated based on the measured statistics at evenly distributed points P9–P12 in the side-to-side (S-S) direction. The recorded mean values at points P9 and P11 were consistent with the recorded data at point P6 and the theoretical value; thus, the accuracy of the test wind field at point P6 was confirmed. However, the test wind decreased at points P10 and P12 and the fluctuations in the rotor plane increased, thereby affecting the steady state condition of the test wind fields.

The wave generator was calibrated using three wave gauges arranged in the basin, as shown in Figure 2A. The recorded wave parameters for the regular wave test cases at the selected OWT model installation point are listed in Table 7. Comparison with the scaled wave parameters indicated that the recorded regular wave heights and periods were consistent with the corresponding theoretical values. Subsequently, the accuracy of the experimental setup was measured based on the calibration. The developed setup satisfied the requirements of the OWT dynamic model tests.

3 Wavelet packet-based energy analysis method

In contrast to other signal processing methods, such as Fourier transform, fractal analysis, and Hilbert-Huang transform (Liao
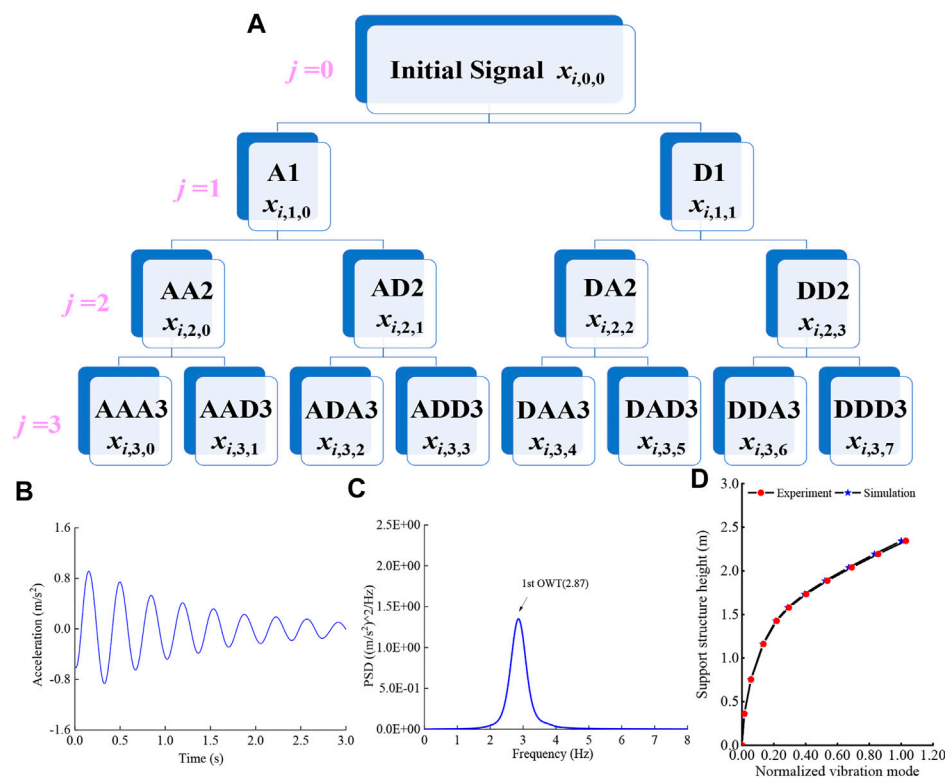
FIGURE 3
Schematic of the decomposed three-layer wavelet packet tree and free decay results of the OWT model under the parked state. **(A)** Schematic of the decomposed three-layer wavelet packet tree; **(B)** free decayed histories of tower top acceleration; **(C)** PSD of free decayed tower top accelerations; and **(D)** normalized first mode shapes of the scaled model and coupled numerical prototype OWT.

et al., 2017; Ling et al., 2019), wavelet packet analysis enables the decomposition of the input signal into low- and high-frequency bands (FBs) with the required resolution (Li et al., 2022), and the influence of the FBs of interest can be quantified using the energy ratio. Owing to these advantages, the wavelet packet-based energy analysis method (Zhang et al., 2021) was applied in subsequent data processing to identify the coupling mechanisms of the OWT model for the selected test cases.

As shown in Figure 3A, the initial signals $x_{i,0,0}$ is decomposed into two sub-frequency components A1 and D1 with respect to the low-frequency sub-signal $x_{i,1,0}^{*}$ expressed in Eq. 5 and the high-frequency sub-signal $x_{i,1,1}^{*}$ expressed in Eq. 6. As is evident from these equations, the wavelet basis functions should be selected prior to the decomposition. Meyer wavelet basis functions ($h(k)$ and $g(k)$) were used in this study, as expressed in Eqs 7, 8 (Zeng, 2016). The newly decomposed FBs can then be obtained after the second decomposition, resulting from the sub-signals in the previous layer. For example, the decomposed FBs AA2 and AD2 in the third layer are derived from FB A1 in the second layer. Based on the characteristics of the wavelet-packet-based energy analysis

method, 2^{j} FBs are obtained in the ($j+1$) layer after completing the j-th decomposition, and the general expressions of the decomposed sub-signals are expressed in Eqs 9, 10. Furthermore, the energy of each decomposed FB can be estimated using Eq. 11, the quantified influence of the decomposed FBs is evaluated based on the defined energy ratio $\phi_{j,\,m}$ in Eq. 12

$$x_{i,1,0}^{*} = \sum_{k} h(k-2i)x_{i,0,0} \tag{5}$$

$$x_{i,1,1}^{*} = \sum_{k} g(k-2i)x_{i,0,0} \tag{6}$$

$$h(k) = (2\pi)^{-1/2} e^{iw/2} \begin{cases} \sin\left(\dfrac{\pi}{2} v\left(\dfrac{3}{2\pi}|\omega|-1\right)\right) & \dfrac{2\pi}{3} \le |\omega| \le \dfrac{4\pi}{3} \\[2mm] \cos\left(\dfrac{\pi}{2} v\left(\dfrac{3}{4\pi}|\omega|-1\right)\right) & \dfrac{4\pi}{3} \le |\omega| \le \dfrac{8\pi}{3} \\[2mm] 0 & |\omega| \notin \left[\dfrac{2\pi}{3}, \dfrac{8\pi}{3}\right] \end{cases} \tag{7}$$

$$G(k) = (-1)^{k} h(-k+1) \tag{8}$$

$$x_{i,j+1,2m}^{*} = \sum_{k} h(k-2i)x_{i,j,m} \tag{9}$$

$$x_{i,j+1,2m+1}^{*} = \sum_{k} g(k-2i)x_{i,j,m} \tag{10}$$

where $x_{i,j+1,2m}^{*}$ and $x_{i,j+1,2m+1}^{*}$ denote the i-th decomposed low-frequency and high-frequency components of the m-th FB in the j-th layer, respectively; $h(k)$ and $g(k)$ represent the low-pass and high-pass filtered Meyer wavelet basis functions, respectively; k is a random number that varies from 1 to N; N denotes the number of discrete sampling points in the initial signal; ω is the angular frequency; ν is auxiliary function for constructing the Meyer wavelet; i is the serial number of discrete data in the original input signal; j is the number of decomposed layers; and m is the FB serial number in the jth layer.

$$E_{j,m} = \sum_{i=1}^{N} \left| x_{j,m}^{*}(i) \right|^2 \qquad (11)$$

$$\phi_{j,m} = \frac{E_{j,m}}{\sum_m E_{j,m}} \times 100\% \qquad (12)$$

4 Dynamic model tests of scaled jacket OWT

4.1 Estimation of OWT model first natural mode

Before conducting the dynamic tests under winds and waves, a free decay test was performed to identify the first natural mode and nondimensional damping ratio of the OWT model. An initial tower-top displacement in the F-A direction was applied to the scaled test model under the parked state in still water. The measured free decay histories of the tower top acceleration are presented in Figure 3B, a generally decreasing trend is observed owing to structural damping. According to the empirical formula based on structural dynamics (Clough and Penzien, 2006), expressed in Eq. 13, the nondimensional damping ratio can be estimated based on the structural free decayed motions, and the estimated damping ratio of the OWT model is 4.4%. The OWT model damping ratio is less than 5%; therefore, it can be assumed to be a small damping system, and the first natural modes of the system can be identified based on the narrow band peaks of the frequency response functions (Bendat and Piersol, 1980). The power spectral density (PSD) of the free decayed tower top acceleration is depicted in Figure 3C, and the first natural frequency of the OWT model is estimated as 2.87 Hz based on the small damping system assumption. Furthermore, the first natural mode of the OWT model can be identified based on the PSDs of the recorded free-decayed tower and jacket accelerations using the distributed acceleration sensors along the support system model. The normalized acceleration sensor is shown in Figure 3D.

Furthermore, a fully coupled numerical model of the prototype jacket OWT was established in the simulation tool FAST v8 (Jonkman and Jonkman, 2016), and the damping ratio of the support system model was set as 3.3%. Therefore, the

introduced small damping system assumptions are also applicable to the established coupled numerical model, and the first natural frequency of the prototype jacket OWT is 0.326 Hz based on the free decayed tower top accelerations. According to the frequency scale ratio defined in Table 1, the corresponding scaled theoretical value should be 2.82 Hz, which is consistent with the measure first natural frequency of the test model. The estimated and normalized first mode shapes of the coupled numerical model are shown in Figure 3D. Evidently, the first mode shape obtained from the numerical model is almost identical to the measured shape in the test model. This indicates that the first bending mode of the OWT support system is appropriately scaled in the tests using the proposed hydro-structural elastic similarity.

$$\xi = \frac{1}{2\pi} ln \left| \frac{\varphi_{Ap}}{\varphi_{Ap+1}} \right| \qquad (13)$$

where φ_{Ap} and φ_{Ap+1} denote the p-th and $p+1$-th peaks or valleys of the decay curve, respectively.

4.2 Dynamic response analysis of OWT model

According to the aero-hydro-elastic similarities, dynamic model tests of a scaled jacket OWT model were performed under the selected steady winds, regular waves, and combined test conditions. The results of steady wind test cases are compared to identify the variations of the structural responses under different operation states, and the regular wave test cases are conducted to investigate the effect of wave height and period. Then, the comparisons under combined wind and wave conditions are carried out to reveal the interactions among the wind, wave, and scaled OWT model responses. Each test case lasted for 330 s with a sampling frequency of 500 Hz, and the first 150 s of the recorded data in each test were discarded to eliminate the influence of transient effects on the results. The measured histories of the thrust and tower-top acceleration for the selected test cases listed in Table 4 are depicted in Figure 4, and the corresponding statistics are listed in Table 8.

Subsequently, the wavelet packet-based energy analysis method introduced in Section 3 was applied to analyze the measured thrust loads and tower top accelerations. To distinguish the frequency components of interest, such as the OWT model fundamental and rotor model rotational frequencies, the decomposed bandwidth in the last layer should be less than 0.25 Hz. This implies that the decomposed FBs should exceed 2^{10}, and the time cost is much higher if the original recorded data are directly applied. Considering this requirement, low-pass filtering and down-sampling procedures were performed based on the recommended analysis method by Zayed (2021) to reduce the sampling rate of the original data to
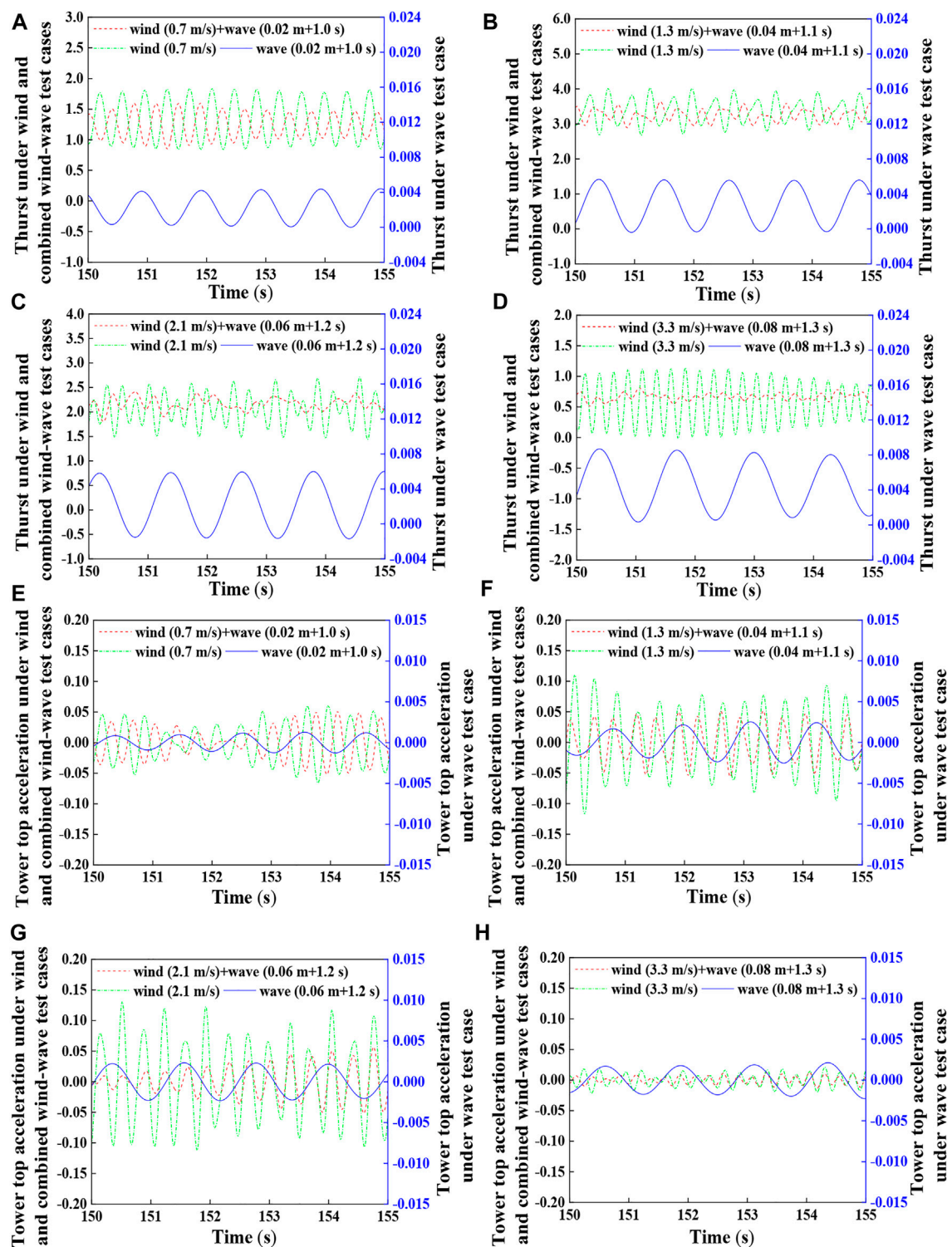
FIGURE 4

Time series of dynamic responses under different test winds and waves. **(A–D)** Time series of thrust under different test winds and waves; **(E–H)** time series of tower top acceleration under different test winds and waves.

TABLE 8 Statistical comparisons of dynamic responses under different test winds and waves.

Load cases	Thrust (N)				Tower top acceleration (m/s^2)		
	Mean	STD	95th MIN	95th MAX	STD	95th MIN	95th MAX
V_{hub} = 0.7 m/s	1.34	0.29	0.77	1.87	0.03	−0.06	0.06
H = 0.02 m, T = 1.0 s	1.8E-3	1.5E-3	3.4E-5	4.4E-3	7.2E-4	−1.6E-3	1.3E-3
V_{hub} = 0.7 m/s and H = 0.02 m, T = 1.0 s	1.24	0.20	0.84	1.59	0.02	−0.04	0.04
V_{hub} = 1.3 m/s	3.53	0.28	2.63	4.01	0.05	−0.07	0.07
H = 0.04 m, T = 1.1 s	2.6E-3	2.1E-3	5.9E-4	5.8E-3	1.2E-3	−2.3E-3	2.3E-3
V_{hub} = 1.3 m/s and H = 0.04 m, T = 1.1 s	3.25	0.17	2.86	3.62	0.03	−0.03	0.03
V_{hub} = 2.1 m/s	2.11	0.35	1.31	2.78	0.07	−0.10	0.10
H = 0.06 m, T = 1.2 s	3.6E-3	2.7E-3	1.6E-3	6.3E-3	1.4E-3	−2.6E-3	2.6E-3
V_{hub} = 2.1 m/s and H = 0.06 m, T = 1.2 s	2.06	0.12	1.78	2.40	0.03	−0.05	0.05
V_{hub} = 3.3 m/s	0.56	0.30	1.8E-3	1.12	9.0E-3	−0.02	0.02
H = 0.08 m, T = 1.3 s	4.6E-3	2.6E-3	1.9E-3	8.7E-3	1.6E-3	−2.8E-3	2.8E-3
V_{hub} = 3.3 m/s and H = 0.08 m, T = 1.3 s	0.66	0.15	0.31	1.01	6.2E-3	−0.01	0.01

9 Hz without frequency aliasing and maintain the precision and efficiency of data processing. The maximum identified frequency of the filtered data was 4.5 Hz. After five decompositions, 2^5 decomposed sub-signals were obtained in total. The bandwidth of each FB is 0.14 Hz, which can precisely identify the frequency components of interest. Moreover, the energy ratios of the OWT model responses calculated based on the time series of the decomposed sub-signals in the frequency range of 0–4.5 Hz under the scaled wind, wave, and combined test cases using Eq. 12 are presented in Figure 5.

4.2.1 Dynamic characteristics of thrust loads under different test winds and waves

- Steady wind test cases

As listed in Table 8, the measured mean thrust loads under the steady test wind speeds of 0.7, 1.3, 2.1, and 3.3 m/s are 1.34, 3.53, 2.11, and 0.56 N, respectively, whereas the corresponding theoretical scaled values of the prototype OWT are 1.18, 3.57, 1.74, and 0.61 N. It can be observed that the modeled mean thrust loads are in agreement with the scaled values. For example, the discrepancies between the measured and theoretical scaled values are only 0.04 N under the test wind speed of 1.3 m/s. The observed discrepancies could be due to the precision of the data acquisition system and the limitations of the performance-scaled RNA model. The latter can ensure the similarity of aerodynamic thrust load under the specified wind speeds without adjusting the scaled test winds using Froude number similarity, rather than the entire range of cut-in to cut-out wind speed. Moreover, an increasing trend of the measured mean thrust loads is observed as the test wind speeds increase to the scaled rated wind speed of 1.3 m/s. Under a normal

operation test wind of 2.1 m/s, pitch control strategies were deployed to alleviate the thrust loads on the rotor plane of the scaled OWT model. Minimal thrust loads are observed in the parked state under a test wind 3.3 m/s (rotor speed of 0 rpm and a blade pitch angle of 90°), and these are compared with the other test winds. The accuracy of the designed RNA test model was validated based on this comparison. The performance of the scaled rotor blades and the drivetrain and mechanical control systems, as well as the influence of test wind speeds and mechanical control strategies on the thrust loads were also determined.

The measured thrust load histories are shown in Figure 4. As indicated by the dotted green lines in Figures 4A–D, smoother curves are observed for test winds of 0.7 and 3.3 m/s, whereas the more complex variations are observed for other test winds. The reason for such differences in the thrust loads under different test winds can be explained based on the frequency domain shown in Figure 5. As depicted in Figures 5A,D, the unique dominant 3P and first blade collective flapwise FBs are observed under the test winds of 0.7 and 3.3 m/s, respectively. Consequently, approximately regular variations in thrust loads were observed under these test winds. Based on the definition of the energy ratio in Eq. 12, the quantified influences of the dominant FBs can reach 76% and 92% under each test wind, respectively.

For the remaining test winds, the influence of multiple FBs should be pointed out, as shown in Figures 5B,C. Particularly under the test wind of 2.1 m/s, the comparable influences of the scaled OWT model fundamental and 3P FBs are approximately 55% and 37%, respectively. Owing to the observed coupling effects between the scaled OWT model natural frequency and 3P under the test winds of 1.3 and 2.1 m/s, more complex variations in thrust loads under these test cases were observed.
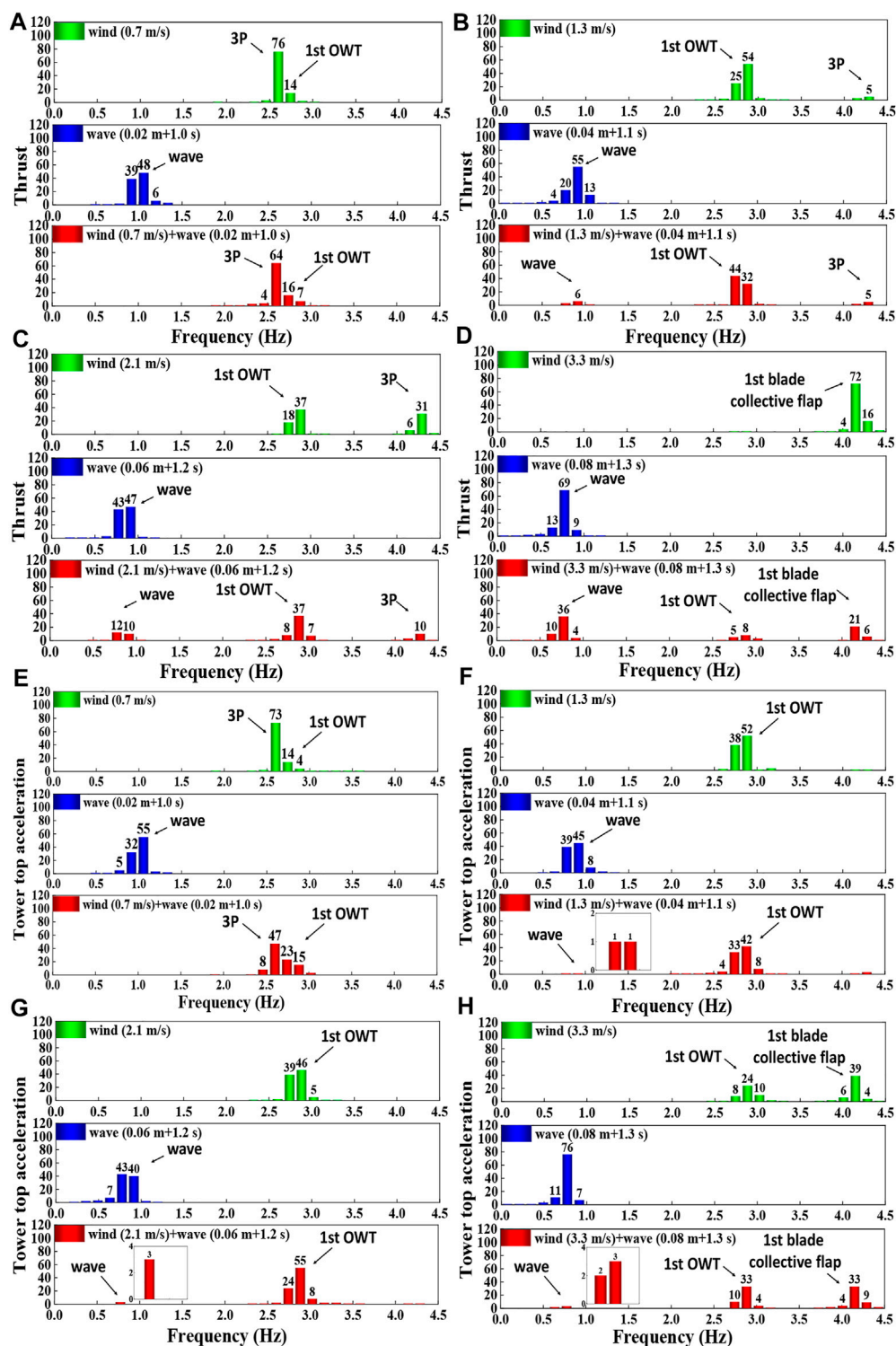
FIGURE 5

Wavelet packet-based energy ratios of dynamic responses under different test winds and waves (Unit: %). **(A–D)** Energy ratios of thrust loads under different test winds and waves; **(E–H)** energy ratios of tower top acceleration under different test winds and waves.

Therefore, it can be concluded that the coupling mechanisms of the measured thrust loads are sensitive to the inflow test winds and related control strategies. The observed coupling effects between the natural and rotational frequencies of the OWT model should be highlighted.

• Regular wave test cases

Although the OWT model was in a parked state under regular wave test cases, the internal forces of the transition piece between the nacelle and tower models were recorded using the mounted load gauge at the transition piece. The recorded histories of the internal shear force in the F-A direction of the scaled OWT model are plotted using solid blue lines in Figures 4A–D. Evidently, the internal shear forces under the scaled regular waves are much smaller than the thrust loads under the test winds, even for the parked OWT model. According to the energy ratios of the recorded shear forces shown in Figures 5A–D, the wave frequency dominates the response under regular wave excitations. For example, the energy ratio of the wave FB under the regular wave condition (H = 0.04 m, T = 1.1 s) can exceed 90%, as shown in Figure 5B.

• Combined wind and wave test cases

To investigate the variations in the structural responses of the scaled OWT model under the combined wind and wave test cases, the measured statistics of the thrust loads are compared, as listed in Table 8. As shown in the table, for the normal operation scaled OWT model, the measured mean thrust loads under the combined test cases are smaller than those under the steady wind test cases. By contrast, a slightly increased mean thrust load of the parked scaled OWT model was observed under the extreme combined test case compared with the measured thrust load under the extreme steady wind test case. Moreover, in comparison with the STDs and 95th MAX values of the thrust load under steady test winds, significant reductions in such statistics are observed under the related combined wind and wave test cases, in contrast to the increased 95th MIN values of the response listed in the table. Based on the compared thrust load statistics, the variations in the dynamic responses of the scaled OWT model under the combined test cases are quite complex was systematically analyzed using the wavelet packet-based energy analysis method.

The measured histories of the thrust loads under the combined test cases are presented in Figures 4A–D. Evidently, the inflow test winds are the dominant environmental loads under the combined test cases compared to the scaled regular waves. According to the recorded thrust loads, the fluctuations of the response are smaller than the measured data under the steady test winds owing to the interactions among the winds, waves, and scaled OWT model under the combined test cases.

Figure 5 presents the energy ratios of the measured thrust loads under the combined test cases, and the effects of the dominant FBs were quantified using the wavelet packet-based energy analysis method. For the below-scaled rated wind speed combined test case shown in Figure 5A, the rotor rotational frequency dominates the thrust load, which is consistent with the dynamic characteristics under the below-scaled rated test wind. The influence of the dominant rotational FB can exceed 60% in such a combined test case. Accordingly, more complex coupling mechanisms were observed under the scaled rated and above-scaled rated wind speed combined test cases, as depicted in Figures 5B,C. In addition to the rotational frequency, the scaled OWT model's first and wave frequencies should also be emphasized. For example, the influences of the OWT model structural, rotational, and wave frequencies under the combined test case (V_{hub} = 2.1 m/s, H = 0.06 m, T = 1.2 s) were approximately 52%, 10%, and 22%, respectively. Meanwhile, the proportion of the scaled OWT model's fundamental frequency exceeded 70% in the rated combined test case. Therefore, the influence of the scaled OWT model structural frequency is the most significant for the operation of the OWT model under these test cases. Moreover, the first blade collective flapwise frequency were observed for the parked OWT model under the extreme combined test cases and its proportion reached approximately 27%, in addition to the wave and OWT model natural frequencies, as shown in Figure 5D.

Furthermore, the PSDs of the thrust loads under steady test winds and the related combined test cases are shown in Figures 6A–D. As indicated in the figures, the remarkable hydrodynamic damping effects on the OWT model structural, rotor rotational, and first blade collective flapwise frequencies under the combined test cases are observed compared with the steady test winds. Therefore, the alleviated fluctuations in the recorded histories of thrust load under combined test cases are observed.

The influence of the interactions among the wind, wave, and scaled OWT models under the combined test cases was quantified in the above studies. For the operation OWT test model, the OWT model fundamental and wave frequencies are also observed for these scaled rated wind speed combined test cases, in addition to the rotor model rotational frequency. Moreover, for the parked OWT model, the first blade collective flapwise frequency should also be pointed out.

4.2.2 Dynamic characteristics of accelerations under different test winds and waves

• Steady wind test cases

The increased tower top accelerations are observed under the steady test wind of 1.3 m/s compared with the scaled test wind of 0.7 m/s, as listed in Table 8. As introduced in the previous section, the blades of the scaled RNA model were
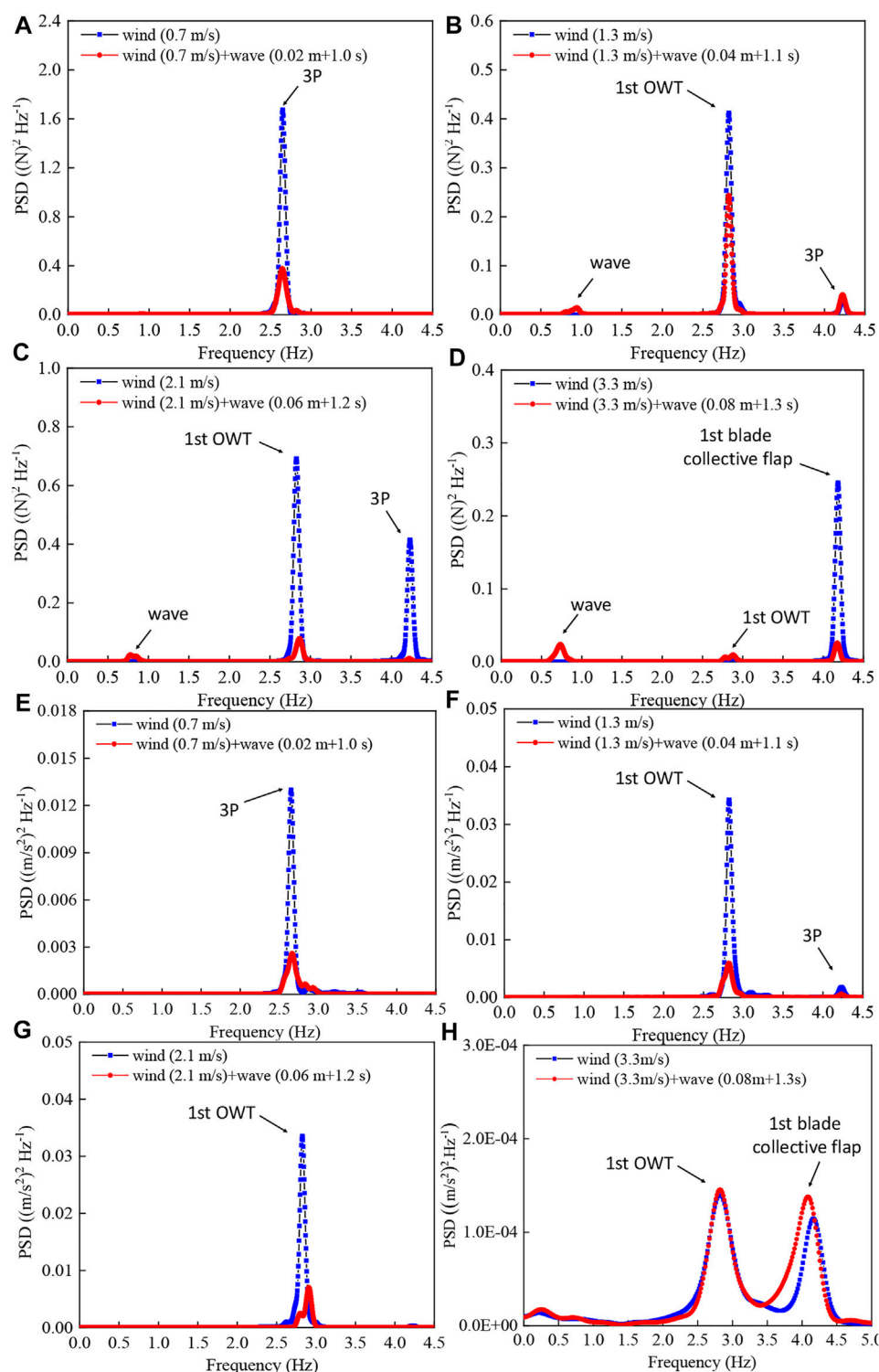
FIGURE 6
Power spectral densities of thrust and tower top acceleration under the steady wind and combined test cases. **(A)** Power spectral densities of thrust under DLC 1 and DLC 9; **(B)** power spectral densities of thrust under DLC 2 and DLC 10; **(C)** power spectral densities of thrust under DLC 3 and DLC 11; **(D)** power spectral densities of thrust under DLC 4 and DLC 12; **(E)** power spectral densities of tower top acceleration under DLC 1 and DLC 9; **(F)** power spectral densities of tower top acceleration under DLC 2 and DLC 10; **(G)** power spectral densities of tower top acceleration under DLC 3 and DLC 11; and **(H)** power spectral densities of tower top acceleration under DLC 4 and DLC 12.

pitched to 15.2°under a test wind of 2.1 m/s to alleviate the thrust loads, whereas the increased fluctuations in the STD of the thrust loads should be due to the non-uniformity of the generated test wind field. Accordingly, additional fluctuating tower top accelerations were observed under such a test wind. Meanwhile, by comparing with the normal operation test cases, the minimum tower top accelerations are observed for the parked OWT model with the blade pitch-to-feather control strategy under an extreme test wind of 3.3 m/s. The recorded histories of the tower top acceleration under various steady test winds are illustrated in Figures 4E–H.

As the dashed green lines in Figures 4E–G, the generally increasing trend can be seen with the inflow winds in the range of below-scaled to above-scaled rated test winds, and the most significant tower top accelerations are observed under the test wind of 2.1 m/s. Meanwhile, it can be seen that the deployed pitch-to-feather and shaft break procedures effectively mitigated the tower model motions of the parked OWT model under extreme test wind, in comparison with the normal operation test winds, as depicted in Figure 4H. Moreover, based on the wavelet packet-based energy analysis method, the dominant FBs of the tower top acceleration under each test case were identified and are presented in Figures 5E–H. For the normal operation OWT model, the influence of threefold blade passing FB should be highlighted under the test wind of 0.7 m/s, and the proportion of such FB can exceed 70%, as indicated in Figure 5E. Meanwhile, the fundamental FB of the OWT model dominates the response under the remaining operational test winds, as shown in Figures 5F,G. Further, the interactions between the OWT model first bending and RNA model first blade collective flapwise modes are proved under the extreme test wind, and the quantified comparable influence of each dominant FB is about 42% and 49%, respectively, as shown in Figure 5H.

According to the comparisons, it is approximately identical to the variations in the thrust load, and the first increasing and then decreasing trend of tower top accelerations under different test winds are revealed, which should be due to the deployed mechanical control strategies. Moreover, the coupling effects between RNA and OWT support structure models should be highlighted, even in the parked state.

- Regular wave test cases

Based on the statistics listed in Table 8, and the related histories shown in Figures 4E–H, for the presented regular wave test cases, the maximum tower top acceleration is observed under the test case with a wave height of 0.08 m; however, it is quite small with the related ones under the test winds. The energy distributions of the regular tower top accelerations under the regular wave test cases are depicted in Figures 5E–H, where it can

be observed that the wave FB dominates the response, and the quantified influence exceeds 90%.

Although the tower top accelerations under the selected regular waves were small, the influence of wave loads on the tower motions of the scaled OWT model under the combined test cases is discussed in the subsequent section.

- Combined wind and wave test cases

The statistics of the tower-top accelerations under the combined wind and wave test cases are listed in Table 8. As presented in the table, either the STDs or the absolute values of the extreme statistics under the combined test cases are smaller than the relevant ones under the related test winds. Meanwhile, the wind was proven to be the dominant load under the combined test cases, as shown in Figures 4E–H. In comparison with the steady test winds, the recorded histories of the tower top accelerations reduced remarkably under the combined test cases, particularly for the normal operation OWT model.

The identified dominant FBs of the response under the combined test cases are illustrated in Figures 5E–H. Owing to the prominent influence of the wind load, the dominant FBs of the response under the combined test cases are approximately identical to the data under steady test winds, for example, the scaled OWT model fundamental, threefold blade passing, and first blade collective flapwise FBs, as depicted in the figures. However, the proportions of the domain FBs are different because of the additional input regular waves in the combined test cases. For example, the proportions of the dominant OWT model fundamental and threefold blade passing FBs are 38% and 55% under the combined wind and wave case (V_{hub} = 0.7 m/s, H = 0.02 m, T = 1.0 s), while the relevant ones are 18% and 73% under the related test wind.

Moreover, as indicated in Figures 6E–H, significant reductions in the amplitudes of the dominant frequencies caused by hydrodynamic damping under the normal operation combined test cases can be distinctly observed compared with the steady test winds, in addition to the approximate variations under the extreme wind and combined tests. It can be observed that the tower accelerations are reduced, particularly for the motions located at the upper part of the tower, as shown in Figures 7A–C. Furthermore, it can be seen the remarkably decreased 95th MAX values of the tower accelerations under the combined test cases compared with the related scaled steady test winds, especially for the accelerations located at the upper part of tower model. For example, the 95th MAX value of tower top acceleration can reach 0.1 m/s² under the test wind of 2.1 m/s, while the relevant one is only 0.05 m/s² under the related combined test case. The above reductions of tower motions under the normal operation combined test cases should be owing to the hydrodynamic damping effects on the dominant OWT model first bending mode in the frequency domain, as shown in Figures 6E–G. Meanwhile, the reductions of the amplitudes of
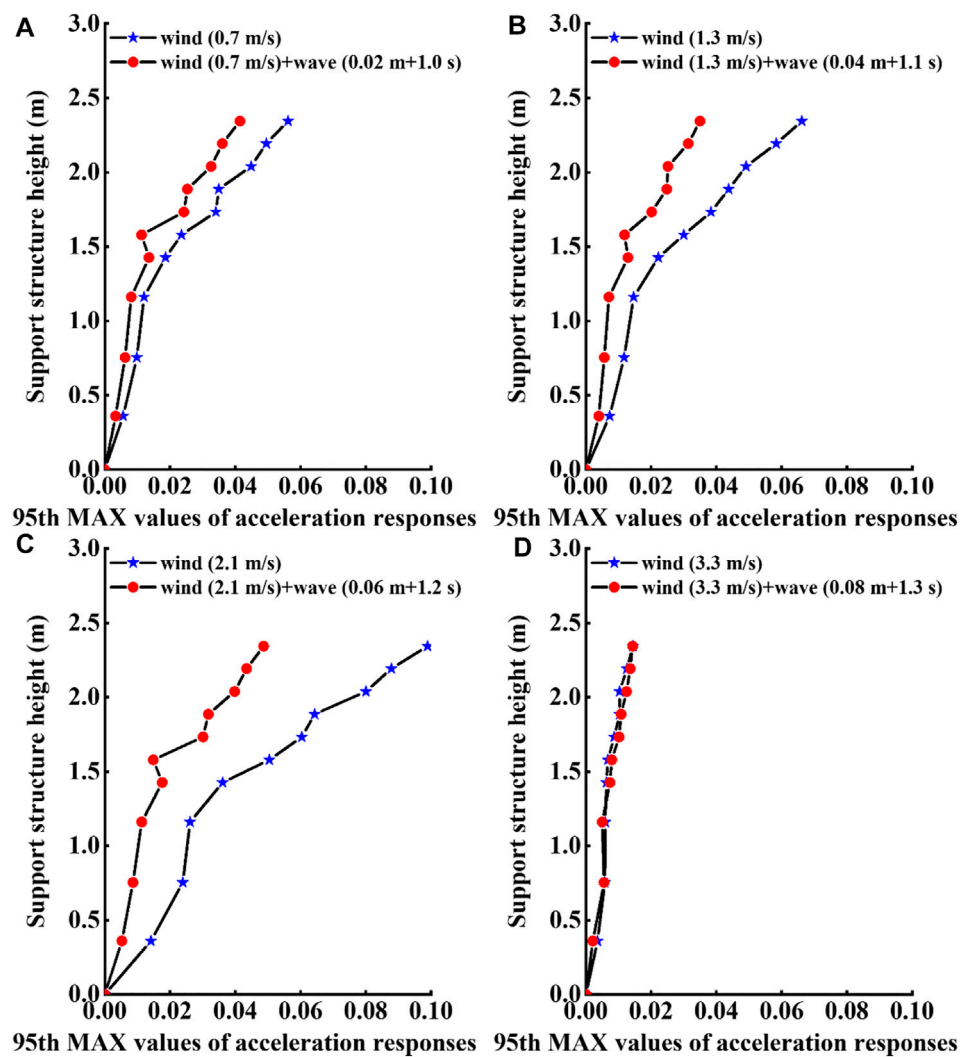
FIGURE 7
Schematic diagrams of the 95th MAX values of the acceleration along the model in the fore-aft direction under the steady wind and combined test cases. **(A)** Statistics of the measured accelerations under DLC 1 and DLC 9; **(B)** statistics of the measured accelerations under DLC 2 and DLC 10; **(C)** statistics of the measured accelerations under DLC 3 and DLC 11; and **(D)** statistics of the measured accelerations under DLC 4 and DLC 12.

dominant first bending and blade collective flapwise modes caused by hydrodynamic damping can not be observed under the extreme test case, and the tower accelerations under the combined case are comparable with the responses under the related extreme test wind, so it proves that the inflow wind dominates the tower motions under such extreme test case, as shown in Figures 6H, 7D. Moreover, owing to the activated OWT model first bending mode under the wind and combined test cases shown in Figures 6E–H, the maximum tower acceleration is located at the top of the scaled tower model.

The coupling mechanisms of the tower accelerations under the combined wind and wave test cases are studied, along with the identified dominant OWT model fundamental, RNA model rotational, and blade collective flapwise frequencies, and the distinctive hydrodynamic damping effects on the tower motions of the OWT model.

5 Conclusion

The dynamic characteristics and coupling mechanisms of an ultra-large jacket OWT were investigated based on dynamic model tests in this study. The scaled OWT model was designed based on aero-hydro-structural elastic similarities, and aerodynamic similarity was ensured using performance-scaled similarities consisting of thrust load and essential operational parameter similarities. Subsequently, a scaled blade model is designed, and additional drivetrain and

pitch control systems are fabricated. The principal geometries of the support model were scaled using hydro-structural elastic similarity, and the additional weights were generally uniformly distributed to ensure the density scale ratio. The three-component force sensor, acceleration sensors, and strain gauges mounted in the model were used to record the thrust loads, motions, and internal forces of the scaled OWT model. Based on the free decay, typical winds and waves, and combined test cases, the dynamic characteristics of the scaled OWT model were studied, and the following conclusions were drawn.

(1) The first natural mode of the scaled OWT model was estimated based on the free decay test and small damping system assumptions, and it was in good agreement with that of the prototype OWT. Thus, the applicability of the recommended hydro-structural elastic similarity was validated to a certain degree.

(2) The observed variations in the measured mean thrust loads under different test winds were consistent with the related scaled theoretical values. Hence, the fabrication accuracy of the scaled RNA model using performance-scaled similarities was proven. The fluctuations under some test winds are due to the asymmetry of the generated test wind fields using the simplified wind generation system and the scaling effects of the small-scale ratio OWT model tests.

(3) For the scaled steady wind test cases, the influence of the test wind speeds and mechanical control strategies on the thrust loads was proved. It is approximately identical to the variations in thrust loads; the first increasing and then decreasing trend of tower top accelerations under the different test winds are revealed, which should be due to the deployed blade pitch control strategies.

(4) According to the quantified influence of dominant FBs using the wave packet analysis method in the frequency domain, it can be concluded that in addition to the OWT model first natural frequency, the coupling effects between the RNA and support model should be emphasized, even for the OWT model under the parked state, such as the observed first blade collective flapwise frequency under the extreme test wind.

(5) The wave FBs dominate the motions of the OWT model under regular wave inputs, whereas the responses are relatively small compared with those under the steady test winds. The inflow test winds proved to be the dominant loads for the structural responses of the OWT model under the combined test cases; therefore, the coupling mechanisms are approximately identical to the scaled wind test cases, in addition to the observed wave FBs. However, owing to the influence of additional wave inputs under the combined test cases, the quantified influence of the dominant FBs differed from those under the scaled steady test winds.

(6) A significant decrease in the structural motions was observed under the normal operation test cases compared with the responses under steady test winds, which were caused by hydrodynamic damping under the combined test cases. Hence, although the motions excited by wave loads are smaller than those related to wind loads, the induced hydrodynamic damping effects should be pointed out, particularly for the normal operation OWT model.

(7) In future studies, a coupled aero-servo-hydro-elastic numerical model under winds and waves will be established, and fully coupled analyses under different limit states will be performed. Moreover, detailed comparisons of the measured test data with the analyzed results of the fully coupled numerical model will be performed to validate the observed coupling mechanisms and simulation accuracy of the coupled simulation tool.

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

Conceptualization, DL; methodology, DL and GX; investigation, GX and ZP; data analysis GX, DL, and ZP; writing—first draft preparation, writing, review and editing, GX, DL, WW, and XL; visualization and proofreading, GX, QL, and WW; supervision and funding acquisition, WW and XL; WW and XL set the objectives of the research. All authors contributed to the article and approved the submitted version.

Funding

Conflict of interest

GX, ZP, and QL were employed by Power China Huadong Engineering Corporation Limited.

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

Ahn, H., and Shin, H. (2020). Experimental and numerical analysis of a 10 MW floating offshore wind turbine in regular waves. *Energies (Basel)* 13, 2608–2617. doi:10.3390/en13102608

Ahn, H., and Shin, H. (2019). Model test and numerical simulation of OC3 spar type floating offshore wind turbine. *Int. J. Nav. Archit. Ocean Eng.* 11, 1–10. doi:10.1016/j.ijnaoe.2017.09.010

Bak, C., Zahle, F., Bitsche, R., Kim, T., Yde, A., Christian Henriksen, L., et al. (2013). *Department of wind energy I-report description of the DTU 10 MW reference wind turbine*. Copenhagen, Denmark: Department of Wind Energy.

Bendat, J., and Piersol, A. (1980). *Engineering applications of correlation and spectral analysis*. Hoboken, New Jersey, United States: John Wiley & Sons.

Bossanyi, E. (2011). *GH bladed user manual: Version 4.3*. Bristol, UK: GL Garrad Hassan.

Çengel, Y. A., and Cimbala, J. M. (2006). *Fluid mechanics: Fundamentals and applications - 1st ed.* McGraw-Hill.

Chakrabarti, S. K. (1994). *Offshore structure modeling*. Singapore: World Scientific Publishing.

Clough, R., and Penzien, J. (2006). *Dynamics of structures*. New York, United States: McGraw-Hill.

DNV GL (2016). *Loads and site conditions for wind turbines*. Oslo: DNV.

DNV (2018). *SIMA user guide*. Oslo, Norway: DNV.

Du, W., Zhao, Y., He, Y., and Liu, Y. (2016). Design, analysis and test of a model turbine blade for a wave basin test of floating wind turbines. *Renew. Energy* 97, 414–421. doi:10.1016/j.renene.2016.06.008

GWEC (2022). *Global offshore wind report 2022*. Brussels, Belgium: Department of Energy.

Huan, C., Lu, D., Zhao, S., Wang, W., Shang, J., Li, X., et al. (2022). Experimental study of ultra-large jacket offshore wind turbine under different operational states based on joint aero-hydro-structural elastic similarities. *Front. Mar. Sci.* 9, 915591. doi:10.3389/fmars.2022.915591

Jonkman, B., and Jonkman, J. M. (2016). *FAST user's guide: Version 8.16.00*. Golden, CO: National Renewable Energy Laboratory.

Jonkman, J. M., and Sprague, M. (2021). *OpenFAST documentation*. Golden, CO: National Renewable Energy Laboratory.

Kim, B., Jin, J., Bitkina, O., and Kang, K. (2016). Ultimate load characteristics of NREL 5-MW offshore wind turbines with different substructures. *Int. J. Energy Res.* 40, 639–650. doi:10.1002/er.3430

Larsen, T. J., and Hansen, A. M. (2007). *How 2 HAWC2, the user's manual*. Denmark: Roskilde.

Larsen, T. J., Yde, A., Verelst, D. R., Pedersen, M. M., Hansen, A. M., and Hansen, H. F. (2014). *IEA annex 30 offshore code collaboration continued (OC4) phase I+II.2014*. Denmark: Roskilde.

Li, Y., Ning, F., Jiang, X., and Yi, Y. (2022). Feature extraction of ship radiation signals based on wavelet packet decomposition and energy entropy. *Math. Problems Eng.* 2022, 1–12. doi:10.1155/2022/8092706

Liao, K., Lai, X., Zhou, Z., and Zhu, Q. (2017). Applying fractal analysis to detect spatio-temporal variability of soil moisture content on two contrasting land use hillslopes. *Catena (Amst)* 157, 163–172. doi:10.1016/j.catena.2017.05.022

Lifes50 (2021). *Qualification of innovative floating substructures for 10MW wind turbines and water depths greater than 50m.* Stuttgart: University of Stuttgart.

Ling, T., Zhang, L., Huang, F., Gu, D., Yu, B., and Zhang, S. (2019). OMHT method for weak signal processing of GPR and its application in identification of concrete micro-crack. *J. Cent. South Univ.* 26, 3057–3065. doi:10.1007/s11771-019-4236-y

Martin, H. R. (2011). *Development of a scale model wind turbine for testing of offshore floating wind turbine system*. Orono: University of Maine.

Park, M., Park, S., Seong, B., Choi, Y., and Jung, S. P. (2021). Current status and prospective of offshore wind power to achieve Korean renewable energy 3020 plan. *J. Korean Soc. Environ. Eng.* 43, 196–205. doi:10.4491/ksee.2021.43.3.196

Putri, R. M., Obhrai, C., Jakobsen, J. B., and Ong, M. C. (2020). Numerical analysis of the effect of offshore turbulent wind inflow on the response of a spar wind turbine. *Energies (Basel)* 13, 2506. doi:10.3390/en13102506

Ren, Y., Venugopal, V., and Shi, W. (2022). Dynamic analysis of a multi-column TLP floating offshore wind turbine with tendon failure scenarios. *Ocean. Eng.* 245, 110472. doi:10.1016/j.oceaneng.2021.110472

Robertson, A., Jonkman, J., Vorpahl, F., Popko, W., Qvist, J., Frøyd, L., et al. (2014). Offshore code comparison collaboration continuation within IEA wind task 30: Phase II results regarding a floating semisubmersible wind system. in Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering - OMAE 9B. San Francisco, California. June 8-13, 2014. U.S. Department of Energy. doi:10.13140/2.1.2822.9121

Wang, W., Gao, Z., Li, X., and Moan, T. (2017). Model test and numerical analysis of a multi-pile offshore wind turbine under seismic, wind, wave, and current loads. *J. Offshore Mech. Arct. Eng.* 139, 1–17. doi:10.1115/1.4035305

Wang, Y., Shi, W., Michailides, C., Wan, L., Kim, H., and Li, X. (2022). WEC shape effect on the motion response and power performance of a combined wind-wave energy converter. *Ocean. Eng.* 250, 111038. doi:10.1016/j.oceaneng.2022.111038

Zayed, A. I. (2021). Sampling theorem for two dimensional fractional Fourier transform. *Signal Process.* 181, 107902. doi:10.1016/j.sigpro.2020.107902

Zeng, J. (2016). Wavelet threshold image denoising algorithm based on MATLAB different wavelet bases. *Intelligent Comput. Appl.* 6, 6–8.

Zeng, Y., Shi, W., Michailides, C., Ren, Z., and Li, X. (2022). Turbulence model effects on the hydrodynamic response of an oscillating water column (OWC) with use of a computational fluid dynamics model. *Energy* 261, 124926. doi:10.1016/j.energy.2022.124926

Zhang, S., Zhang, L., Ling, T., Fu, G., and Guo, Y. (2021). Experimental research on evaluation of soil water content using ground penetrating radar and wavelet packet-based energy analysis. *Remote Sens.* 13, 5047. doi:10.3390/rs13245047

Zhao, Z., Shi, W., Wang, W., Qi, S., and Li, X. (2021). Dynamic analysis of a novel semi-submersible platform for a 10 MW wind turbine in intermediate water depth. *Ocean. Eng.* 237, 109688. doi:10.1016/j.oceaneng.2021.109688

*CORRESPONDENCE
Lupeng Chen,
✉ chenlp002@126.com

Optimal operation strategy for distribution network with high penetration of dispersed wind power

Jiandong Duan[1], Lupeng Chen[1]*, Wenchao Lu[1], Darui Zhu[1] and Qian Sun[2]

[1]School of Electrical Engineering, Xi'an University of Technology, Xi'an, Shaanxi, China, [2]Xi'an Xianyang International Airport, Xi'an, Shaanxi, China

As a significant approach to local utilization of clean renewable energy, the dispersed wind power (DWP) is becoming more and more widespread in engineering applications. With the fluctuating wind power widely and dispersedly integrated into distribution networks, it is urgent and pressing to effectively enhance the penetration rate of wind power based on the safe operation of distribution networks. This paper takes full advantage of the power regulation ability of wind turbines to actively participate in the operation of distribution networks, and then an optimal operation model of distribution networks is established with the optimization objectives of high penetration of DWP and maximum correntropy of node voltage. Aiming at the characteristics of the proposed model, the multi-objective brain storm optimization algorithm is adopted to solve the model to achieve the Pareto solution set, and the fuzzy membership function is used to determine the optimal operation scheme from the solution set. The simulation results on the expanded IEEE 33 bus system indicate the rationality of the proposed optimal operation strategy for distribution networks with high penetration of DWP. Meanwhile, the feasibility of the optimal operation scheme is verified through the case of a practical demonstration project.

KEYWORDS

dispersed wind power, distribution network, high penetration, regulation capability, maximum correntropy criterion, brain storm optimization algorithm

1 Introduction

The aggravated environmental pollution and global fossil energy shortage have been driving the rapid development of technologies in clean renewable energy based power generation. As one of the focal points in energy research, the wind power is prominent. China is committed to the development of dispersed wind power technology, especially in April 2018, the National Energy Administration of China promulgated the specific management rules for the projects of dispersed wind power (National Energy Administration, 2018), which intended to vigorously promote the nearby utilization capability of wind power. Due to the randomness and fluctuation of wind power output, however, the large-scale integration of DWP into distribution network will bring adverse effects on the safe and economic operation of the system (Ma et al., 2016). Therefore, the safe and efficient utilization of DWP has become a vital problem to be solved at present.

Many researches have been carried out on the optimal operation, energy management and active control of distribution network with diverse distributed generations in recent years. Two voltage control strategies for distribution power system were introduced in Reference (Joseph et al., 2020) which utilize distributed energy resources to contribute reactive power to support the grid voltage. Reference (Tang et al., 2017) presented an optimal economic dispatch model for active distribution network and developed a new optimization method to solve the model to manage multiple controllable resources, such as distributed generations, battery storage, and demand response and so on. In order to take advantage of technological superiority of distributed generations in the economic, environmental and energy aspects, Reference (Zhao et al., 2016) highlighted an optimal dispatch model for active distribution network aiming to protect environment and save energy, then a two-stage algorithm was presented to solve the model. In Reference (Li and Zheng, 2019), the consumption rate of renewable energy and the operation cost of hybrid AC/DC microgrid was regarded as the optimization objectives, the source-network-load coordination strategies were proposed to realize the efficient utilization of renewable energy and the optimal operation of microgrids. The optimization objectives were proposed in Reference (Tuladhar et al., 2016) which contains minimization of system power loss, voltage deviation and energy wastage from wind power and photovoltaic power. Based on that, an optimal dispatch method for network reconfiguration as well as reactive power optimization of distributed generations was presented to improve the network performances.

With the rapid development of wind turbines and their control technology, flexible regulation of dispersed wind power output has become a feasible technological means to achieve the optimal operation of distribution network (Zhang et al., 2017). Therefore, the researchers have paid great attention to the reactive power optimization, active-reactive power coordination optimization, and voltage control strategy for distribution network with dispersed wind power. In Reference (Zhang et al., 2017), the control model of maximum active power for wind power integration was established by making full use of the reactive power capability in doubly fed induction generator (DFIG). Furthermore, a control strategy and optimization process for the maximum active power of wind power integration was designed in the paper. The reactive power capability of wind turbines based on permanent magnet synchronous generators was considered in Reference (Li et al., 2020), a reactive power control strategy adopted by optimizing the reactive power of wind turbines was proposed to minimize the power loss and maximize the service life of wind turbines. Reference (Xiang et al., 2021), established an optimal operation model of reactive power for distributed wind generations integrated into distribution network, with the purpose of making the most of distributed wind generations and economic operation of the distribution network. Reference (Guo et al., 2019) presented a distributed coordinated active and reactive power control method to optimize the output of active power and reactive power of wind farms, and the effect of active power on voltages was considered to improve the voltage control. Reference (Londero et al., 2017) provided the influence mechanism of DFIG operational limits on long-term voltage stability of power system and the circumstances of

different wind speed was taken into account to explore possible simplified model. In Reference (Anilkumar et al., 2018), an intelligent method of optimal coordinated voltage control was proposed for distribution networks in the presence of distribution generations, which aimed to seek the optimal solution of different control variables.

Although most current researches on distributed renewable energy power focused on the optimal operation and energy management of distribution systems with a variety of controllable resources, it is not applicable to the engineering applications which only DWP is integrated into distribution networks. These researches took into account the coordination and collaboration of various controllable resources based on their respective characteristics. However, according to the application situation of practical projects, it is an important optimization goal to improve the penetration rate of wind power resources on the premise of ensuring the safe operation of power system. Besides, the power regulation capability of the wind turbines is not fully explored, and the regulation capability also needs to be verified by practical engineering. Therefore, it is necessary to conduct the research of the optimal operation strategy for DWP integrated into distribution networks.

Based on the analysis of power regulation capability of DFIG, this paper takes the high penetration of DWP and maximum correntropy of node voltage as optimization objectives to establish an optimal operation model for distribution networks, and then the multi-objective brain storm optimization algorithm is selected to solve the model. The expanded IEEE 33 bus system is adopted to test the validity of the proposed model and solution algorithm, moreover, the practical application effect of optimal operation strategy is also verified by a demonstration project of DWP integrated into the distribution network.

The rest of this paper is organized as follows: Section 2 analyzes the power regulation capability for wind turbine, which lays a theoretical foundation for the optimal operation strategy in this paper. In Section 3, the optimal operation model of distribution network with DWP is presented, and the solution method for the model is described in Section 4. Simulation results on the expanded IEEE 33 bus system are presented and discussed in Section 5, and the test on a demonstration project in China is shown in Section 6. Section 7 concludes the paper.

2 Analysis of power regulation capability for wind turbines

In general, all of the three main types of wind turbines, i.e., cage asynchronous generator, direct drive permanent magnet synchronous generator and doubly fed induction generator, have certain capability of power regulation. However, DFIG is the most popular type applied in dispersed wind farms in China. The power regulation capability of DFIG is the major foundation for the realization of DWP's active participation in the optimal operation of distribution network.

Taking the DFIG as an example, the analysis of power characteristics shows that the stator side and the grid-side converter are the main source of reactive power (Zhang et al., 2017; Xiang et al., 2021). Current research indicates that the
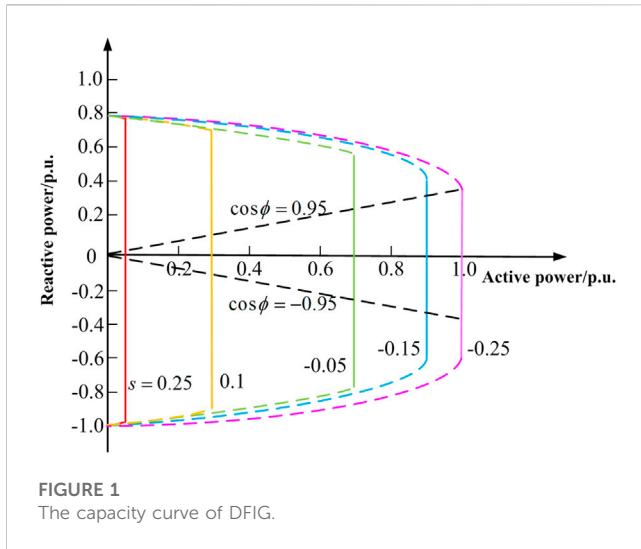
FIGURE 1
The capacity curve of DFIG.

reactive power regulation capability of DFIG is greatly affected by the rotor current. The rotor current is restricted by the thermal current limit of the rotor winding and the maximum current of the converter, so it is considered that the maximum rotor current is a key factor affecting the reactive power of DFIG. Taking into account the current constraints of rotor side, the reactive power capability from the stator side can be expressed as follows:

$$\begin{cases} Q_s^{\max} = -\dfrac{3U_s^2}{2\omega_s L_s} + \sqrt{\left(\dfrac{3}{2}\dfrac{L_m}{L_s}U_s I_{\max}\right)^2 - P_s^2} \\ Q_s^{\min} = -\dfrac{3U_s^2}{2\omega_s L_s} - \sqrt{\left(\dfrac{3}{2}\dfrac{L_m}{L_s}U_s I_{\max}\right)^2 - P_s^2}, \end{cases} \quad (1)$$

where $Q_s^{\max}$ and $Q_s^{\min}$ denote the maximum reactive power which issued and absorbed by the stator side respectively; U_s is the stator voltage; $I_{\max}$ is the maximum permissible current of rotor; P_s denotes the active power output of the stator side; ω_s denotes the synchronous rotation angular velocity; L_s denotes the equivalent inductance of stator; L_m denotes the excitation inductance.

In addition, the grid-side converter can also provide reactive power regulation capability for the wind turbines, which is restricted by the maximum capacity of grid-side converter. The reactive power capability can be expressed as:

$$\begin{cases} Q_g^{\max} = \sqrt{S_{c\,\max}^2 - s^2 P_s^2} \\ Q_g^{\min} = -\sqrt{S_{c\,\max}^2 - s^2 P_s^2}, \end{cases} \quad (2)$$

where $Q_g^{\max}$ and $Q_g^{\min}$ denote the maximum reactive power which issued and absorbed by the grid-side converter; $S_{c\,\max}$ is the capacity of the grid-side converter; s is the slip ratio.

Therefore, the reactive power capability of a single DFIG unit, i.e., the adjustable range of reactive power, can be expressed as Eq. 3:

$$\begin{cases} Q^{\max} = Q_s^{\max} + Q_g^{\max} \\ Q^{\min} = Q_s^{\min} + Q_g^{\min}, \end{cases} \quad (3)$$

where $Q^{\max}$ and $Q^{\min}$ denote the maximum reactive power which issued and absorbed by DFIG unit respectively.

The power capacity curve of DFIG is shown in Figure 1, each curve in the figure is corresponding to the regulation range of reactive power in different slip ratio. It can be seen that the reactive power of wind turbines can be adjusted in a wide range and gradually decreases with the increase of active power. Therefore, the reactive power capability of wind turbine makes it have sufficient conditions to actively participate in the voltage regulation of distribution networks.

3 Optimal operation model of distribution network with dispersed wind power

The power regulation capability of wind turbines makes it possible for dispersed wind power to participate in the operation and regulation of distribution network. In order to give full play to the advantages of dispersed wind power in the economic development and environmental protection, this paper presents an optimal operation model of distribution network with DWP, which focuses on the high penetration of DWP and maximum correntropy of node voltage. The purpose of optimal operation model is not only the full utilization of clean renewable energy, but also the improvement of voltage quality in distribution networks.

3.1 Objective function

3.1.1 High penetration of dispersed wind power integration

Making full use of clean renewable energy based power generation is one of the effective ways to solve the environmental pollution and energy shortage crisis. Moreover, the operation cost of wind power generation does not need fuel expenses. On the premise of satisfying the demand of power load, the higher wind power consumption, the lesser the electricity purchased from the bulk power system which is dominated by thermal power, and the better the economical operation of the system. Therefore, in the perspective of environmental and economic factors, realizing the high penetration of wind power integration is the primary objective to achieve the optimal operation of distribution network with DWP.

The concept of energy penetration ratio has been recognized by many researchers in judging the utilization level of clean and renewable energy generation in distribution network. It mainly refers to the percentage of electricity supplied by distributed generations to load consumption of distribution network (Karimi et al., 2016; Yarahmadi and Shakarami, 2018). Therefore, based on the energy penetration ratio of DWP in an operation cycle, the objective function for pursuing high penetration of dispersed wind power integration is established as Eq. 4:

$$E_p = \max \sum_{t=1}^{K} \left(\frac{\sum\limits_{w \in W} P_w(t) \cdot \Delta t}{P_{load}(t) \cdot \Delta t} \right) \times 100\%, \quad (4)$$

where E_p denotes the energy penetration ratio of DWP; K is the number of stages in a operation cycle; W denotes the number of

dispersed wind farms which integrated into distribution network; $P_w(t)$ denotes the active power output of wth dispersed wind farm at time t; $P_{load}(t)$ denotes the load power in distribution area at time t; Δt denotes the length of each unit operation stage.

Eq. 4 refers to the percentage of the wind power consumption to the load of distribution network. It is intended to ensure that the active power generated by dispersed wind farms can be consumed to the greatest extent on the premise of safe operation of the system, so as to achieve the objective of high penetration and full utilization of DWP by taking full advantages of the power regulation capability of wind turbines.

3.1.2 Maximum correntropy of node voltage

Node voltage can reflect both the operation status and the power quality of distribution network. The improvement of voltage quality for distribution network by DWP integration has been verified in theoretical research and practical engineering application (Londero et al., 2017). The node voltage is usually regarded as the constraint condition in most of optimal operation models, but it will make the optimized node voltage very close to its boundary value. So the node voltage is extremely liable to exceed the specified limit, resulting in the large-scale disconnection of dispersed wind farms from the distribution network, which will cause the loss of consumers. Therefore, in order to avoid the dispersed wind farms quitting operation under voltage fluctuation as far as possible and maintain a certain voltage margin, it is necessary to consider the optimization objective of node voltage quality improvement.

Because the node voltage in the distribution network actually presents the characteristics of non-Gaussian and nonlinear, the maximum correntropy criterion (MCC) (Liu et al., 2007) is applied as a cost function to establish the objective function of voltage quality improvement in distribution networks, and its general expression can be defined as Eq. 5:

$$\hat{V}_{N,\sigma}(e) = \max \frac{1}{N} \sum_{i=1}^{N} G_\sigma(e_i) = \max \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(e_i)^2}{2\sigma^2}}, \quad (5)$$

where N denotes the number of sample points; $G_\sigma(\bullet)$ denotes the kernel function of MCC; σ denotes the kernel width of the kernel function, which is usually taken as constant 1; e_i denotes the difference between the output and expectation of the ith vector to be solved.

The MCC maximizes the similarity between the expected value and predicted output value in the sense of "entropy". The criterion focuses on the high-order statistics of the data, it is robust when dealing with the non-Gaussian and nonlinear data (Bessa et al., 2009). As a result, the MCC is suitable for processing of the data with nonstationary and time-varying characteristics, such as node voltage of distribution networks. According to Eq. 5, the objective function of voltage quality improvement in distribution networks based on the MCC can be expressed as Eq. 6:

$$\Delta U_{MCC} = \max \sum_{t=1}^{K} \left(\frac{1}{N_d} \sum_{i=1}^{N_d} G_\sigma \left(\phi \left(|\Delta U_i(t)| - \delta U \right) - 0 \right) \right)$$
$$= \max \sum_{t=1}^{K} \left(\frac{1}{N_d} \sum_{i=1}^{N_d} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\phi(|\Delta U_i(t)| - \delta U))^2}{2\sigma^2}} \right), \quad (6)$$

where ΔU_{MCC} denotes the voltage quality index based on MCC; N_d is the number of nodes in the distribution network; $\Delta U_i(t)$ denotes the voltage deviation of the ith node at time t, and it is a state variable in the optimization process, which satisfies the power flow function corresponding to the control variable of the active power and reactive power output, i.e., $\Delta U = f(P_w, Q_w)$; δU denotes the maximum allowable deviation of node voltage; function ϕ can be calculated according to Eq. 7:

$$\phi(x) = \begin{cases} 0, & x \le 0 \\ x, & x > 0, \end{cases} \quad (7)$$

Eq. 6 is the objective function of node voltage, which is constructed based on MCC shown in Eq. 5. The deviation of node voltage is the optimization objective which refers to the difference between the output and expectation in Eq. 5. This objective function means that the voltage of all nodes in the distribution network should be controlled within a safe and high-quality range to the fullest extent under the circumstances of load or DWP output fluctuation, so as to solve the disconnection problem of dispersed wind farms due to overvoltage. This problem has arisen many times in the trial operation stage of the first demonstration project of DWP integrated into 10 kV distribution network in China (Zhang et al., 2017).

3.2 Constrains

3.2.1 Power flow constraints

The power flow constraints of the distribution network mainly include equality constraints of power balance and inequality constraints of node voltage and branch current, which can be expressed as Eqs. 8, 9.

$$\begin{cases} P_i = U_i \sum_{j \in i} U_j \left(G_{ij} \cos \theta_{ij} + B_{ij} \sin \theta_{ij} \right) \\ Q_i = U_i \sum_{j \in i} U_j \left(G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij} \right), \end{cases} \quad (8)$$

where P_i and Q_i denote the active power and reactive power which feed into ith node respectively; U_i and U_j denote the voltage amplitude of node i and j respectively; G_{ij} and B_{ij} denote the real and imaginary part of bus admittance matrix respectively; θ_{ij} denotes the phase angle difference between node i and j.

$$\begin{cases} U_i^{min} \le U_i \le U_i^{max} \\ I_i \le I_i^{max}, \end{cases} \quad (9)$$

where U_i denotes the voltage amplitude of node i; U_i^{min} and U_i^{max} denote the maximum and minimum voltage value of node i respectively; I_i denotes the current amplitude of branch i; I_i^{max} denote the maximum limit of branch current.

3.2.2 The power output constraints

The power capacity of DFIGs in dispersed wind farms is finite. Moreover, its reactive power is subject to the value of active power and power factor, which is shown in Figure 1. The active power and reactive power output of dispersed wind farms can be expressed as inequality constraints, which are shown in Eq. 10.

$$\begin{cases} P_w^{min} \le P_w(t) \le P_w^{max} \\ Q_w^{min} \le Q_w(t) \le Q_w^{max}, \end{cases} \quad (10)$$

where $P_w(t)$ and $Q_w(t)$ denote the active power and reactive power output of wth dispersed wind farm at time t respectively; $P_w^{\min}$ and $P_w^{\max}$ denote the active power capacity bound of wth dispersed wind farm respectively; $Q_w^{\min}$ and $Q_w^{\max}$ denote the reactive power bound value respectively, which are corresponding to the active power output $P_w(t)$ according to Figure 1, i.e., $[Q_w^{\min}, Q_w^{\max}] = f(P_w(t))$.

4 Solution method for the optimal operation model

The above Eqs. 4–10 constitute the optimal operation model of the distribution network with DWP. With high dimensional control variables, multiple optimization objectives, multiple optimization periods and nonlinear constraints, the optimal operation model is essentially a complex nonlinear programming problem. Traditional optimization algorithms, such as interior point method and nonlinear programming method, have shown some difficulties in solving the power optimization problems of the distribution network with DWP. Therefore, more and more researchers are focusing on artificial intelligence algorithms which have more powerful global searching ability. Aiming at the characteristics of the proposed optimal operation model, it is necessary to choose an intelligent algorithm with simple principle and wider adaptability to solve the model. Brain storm optimization (BSO) algorithm is a novel swarm intelligence search algorithm, which has the advantages of clear concept, excellent global optimization, wide adaptability and so on (Shi et al., 2013; Xiong and Shi, 2018). It has been applied to various optimization problems such as economic dispatch of power system. Considering the characteristics of the model and the advantages of BSO algorithm, the multi-objective brain storm optimization (MBSO) algorithm is introduced to solve the optimal operation model to obtain optimal operation scheme in this paper.

4.1 Description of BSO algorithm

BSO algorithm was derived from the process of brainstorming meeting when people take the approach of group decision making to deal with the problems. It is a mathematical solution algorithm abstracted from this process. In the discussion process, participants can be divided into three categories, i.e., a moderator, discussants and a few of problem originators. The mission of the moderator is to coordinate discussion process and promote participants to generate more good "ideas". To avoid prejudice, the moderator should have no relevant professional background for unresolved problems. The discussants are the main force in the discussion process, they discuss the problems and provide a variety of "ideas". They should have different academic knowledge and open mind as much as possible, which will help them to put forward much more different "ideas" where the best one is generated. After each round of brainstorming, the problem originators should make a choice from the "ideas" proposed by the discussants. They all have their own expertise and experience for the problems, so they will choose good idea according to different standards.

The iterative process of BSO algorithm corresponds to the discussion process of brainstorming meeting. Solutions in the algorithm can be regarded as "ideas" generated during brainstorming discussions. In addition, the fitness functions of the algorithm play the same role as problem originators in evaluation and selection. Similar to other swarm intelligence search algorithms, BSO algorithm also contains three core parts: generation of initial solutions, iterative update of solutions, and evaluation and selection of solutions. In each iteration process, the algorithm mainly includes two modules: clustering and updating. The clustering module divides all the solutions into k clusters based on K-means clustering method. The cluster center reflects the characteristics of these solutions and improves the efficiency of subsequent iteration. The updating module promotes local search through parallel optimization of all kinds of solutions, it also promotes global search by means of inter-cluster collaboration and Gauss mutation, which can ensure the convergence performance of the algorithm and the diversity of solutions in the process of optimization.

4.2 Solution method using MBSO algorithm

Because the presented optimal operation model has two objective functions, the multi-objective brain storm optimization algorithm is used to solve the model to obtain the Pareto solution set and select the optimal operation scheme. The algorithm flow is shown below.

(1) The parameters of the distribution network and DWP are initialized firstly, and the initial parameters of the MBSO algorithm should be set reasonably, which include the amount of solutions N_s, maximum iterations $KI_{\max}$, four probability parameters P_1, P_2, P_3, P_4, the amount of solutions M in the archive set (AS), and the amount of clusters k;

(2) N_s solution vectors are generated randomly and evaluated by the objective functions, on the basis of analysis of these solution vectors, the non-dominated solutions can be singled out and be put into the AS;

(3) K-means clustering method is used to divide N_s vectors into k clusters in the two-dimensional target space composed of two objective functions. For the k clusters, the clusters containing non-dominated solutions can be regarded as elite clusters and the others are ordinary clusters;

(4) The process of selecting solution vectors is shown in Figure 2 to generate new solutions. In order to generate new solution vectors, the solution vectors that need to be mutated are selected according to the selection flow in Figure 2. Then, the Gauss mutation method shown in Eq. 11 can be used to update the elements in the selected solution vectors.

$$x_{new}^d = x_{selected}^d + \xi \cdot N_r(\mu_r, \sigma_r^2), \quad (11)$$

$$\xi = \log sig\left(\frac{KI_{\max}/2 - t_c}{K_s}\right) \cdot random() = \frac{random()}{1 + \exp\left(\frac{t_c - \frac{KI_{\max}}{2}}{K_s}\right)}, \quad (12)$$

where $x_{selected}^d$ is the dth element of selected vector; x_{new}^d is the dth element of new vector; $N_r(\mu_r, \sigma_r^2)$ is a random function subjected to Gauss distribution with the parameters of μ_r and σ_r; ξ is a weight
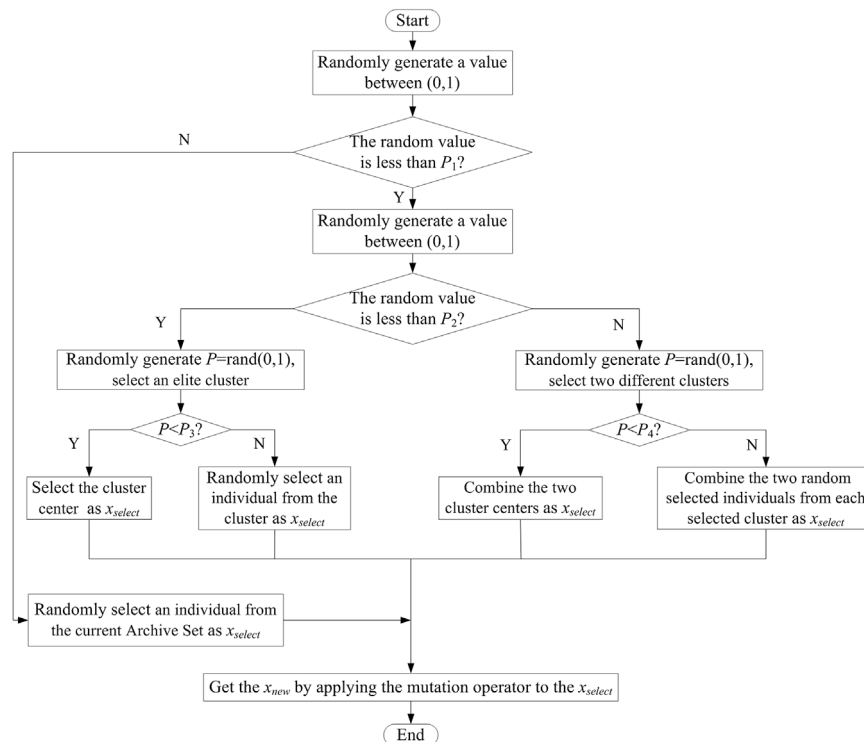
FIGURE 2
Update operation flow of BSO algorithm.

coefficient shown in Eq. 12, where $KI_{\max}$ and t_c denote the maximum and the current iterations respectively; K_s denotes the slope of function $\log sig(x)$.

The old and new solution vectors can be evaluated according to the two objective functions and the non-dominated vector should be kept for the next iterations (If they do not dominate each other, either of them could be kept randomly).

(5) A new non-dominated solution is selected and compared with all the solutions in the AS. If it dominates a number of solutions in the AS, these solutions will be eliminated and this new solution will be stored into the AS. However, if the new non-dominated solution is dominated by a solution in the AS, this new solution will be eliminated and the next new non-dominated solution is checked. When all the non-dominated solutions are compared, the number of solutions in the AS is counted clearly. If the number of solutions exceeds the prescribed capacity of the AS, the crowding-distance of each solution in the AS should be calculated and the redundant solutions should be eliminated. The method of calculating crowding-distance is shown in Eq. 13.

$$D(i) = \sum_{n=1}^{N_{obj}} \left(\frac{|f_n(i+1) - f_n(i-1)|}{2} + \min\left[|f_n(i+1) - f_n(i)|, |f_n(i) - f_n(i-1)|\right]\right), \quad (13)$$

where N_{obj} denotes the number of objective functions and $f_n(i)$ denotes the nth objective function value of ith solution.

(6) The optimal solutions could be searched according to the above steps. When the preset convergence accuracy or maximum iteration number are reached, i.e., the termination condition is satisfied, the search stops and the current AS should be output, which is the Pareto optimal solution set.

In actual operation of the distribution network, the operator must choose one of them from Pareto set, i.e., an optimal operation scheme that can balance the voltage quality and the penetration of wind power. In this paper, the fuzzy membership function shown in Eq. 14 is used to evaluate the satisfactory degree of each objective function corresponding to each solution in Pareto set. According to Eq. 15, the overall satisfactory degree of each solution can be calculated. The operator can choose the solution with the largest overall satisfactory degree as the optimal operation scheme.

$$\mu_i = \begin{cases} 0 & f_i \leq f_{i_{\min}} \\ \dfrac{f_i - f_{i_{\min}}}{f_{i_{\max}} - f_{i_{\min}}} & f_{i_{\min}} < f_i < f_{i_{\max}} \\ 1 & f_i \geq f_{i_{\max}} \end{cases} \quad (14)$$

$$\mu = \frac{1}{N_{obj}} \sum_{i=1}^{N_{obj}} \mu_i, \quad (15)$$

where μ_i is the satisfactory degree of the ith objective function; f_i is the value of the ith objective function; $f_{i_{\max}}$ and $f_{i_{\min}}$ form the scope of the ith objective function; μ is the overall satisfactory degree.

To sum up, the solution method for optimal operation model is shown in Figure 3. The method can be used to obtain the Pareto solution set, from which the optimal operation scheme is selected.

FIGURE 3
Flowchart of solution method for optimal operation model.
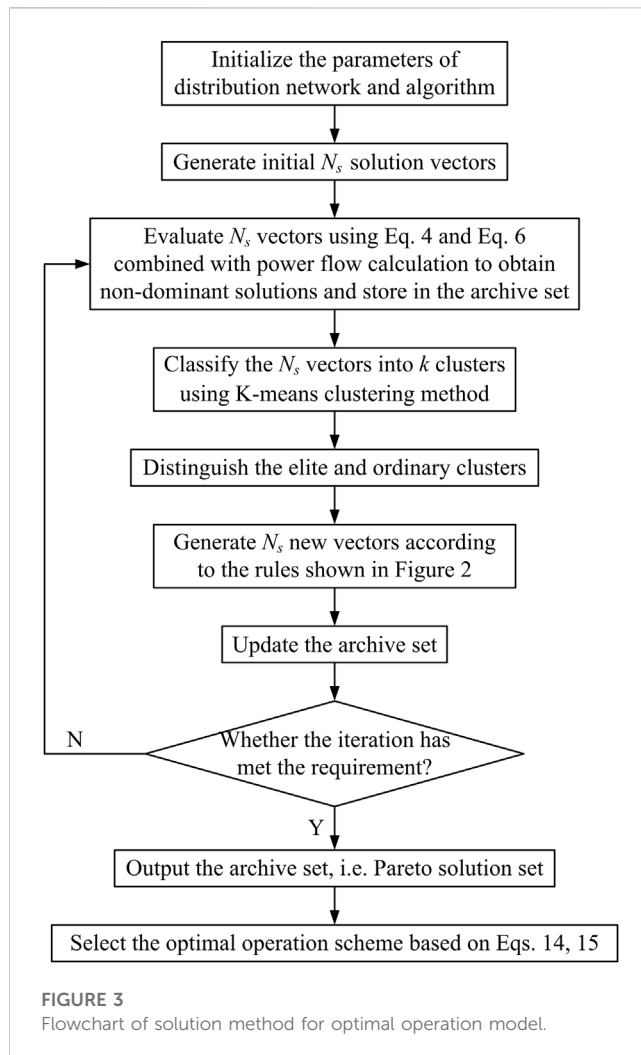
5 Case study

5.1 Case system and parameters

In order to test the effectiveness of the proposed optimal operation model and the solution method, an extended IEEE 33 bus distribution system is used as an example for simulation analysis. The topology structure of the case system is presented

FIGURE 4
Diagram of dispersed wind farms integrated into IEEE33 bus distribution system.

TABLE 1 Parameters of dispersed wind farms.

Number	Node	Capacity/MW
WT1	15	0.6
WT2	18	1.2
WT3	32	2.7

in Figure 4, and the details of each dispersed wind farm are shown in Table 1.

The regulation range of reactive power is confined to the actual situation of active power output for each dispersed wind farm, which is shown in Figure 1. According to Chinese national standard, the allowable deviation of voltage for 10 kV distribution systems is limited to ±7% 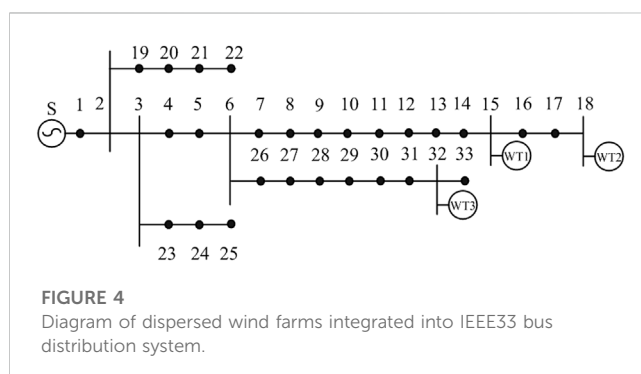of nominal voltage. In order to prevent voltage disqualification caused by the fluctuation of wind power and load, the maximum allowable deviation of node voltage (δU in Eq. 6) is set as 0.05 p.u.

As shown in Figure 5, the load curve and forecasting active power output curves of dispersed wind farms in a typical day are used to simulate the proposed method. This typical day contains various operating conditions, which can be simulated and tested in different scenarios. The operating cycle is 24 h and the time interval is 15 min.

5.2 Simulation results and analysis

The MBSO algorithm is utilized to solve the optimal operation model to obtain the operation scheme. The proper configuration of algorithm parameters should be set up firstly, i.e., the initial number of individuals $N_s = 50$, while the capacity of archive set $M = 20$, the maximum number of iterations $KI_{max} = 500$, the number of clusters $k = 4$, and the four probability parameters are $P_1 = 0.9$, $P_2 = 0.3$, $P_3 = P_4 = 0.5$.

The simulation results are shown in Figures 6, 7, i.e., the optimal operation scheme obtained by the proposed method of high penetration of DWP integration. Figure 6 shows the reactive power output of each dispersed wind farm after optimization. Figures 7A–C show the active power output of each dispersed wind farm after optimization during a typical time period. According to the simulation results, we can conclude that the wind turbines can make full use of its own capability of power regulation to satisfy the integration conditions of DWP and maintain the safe operation of the system.

5.2.1 Comparison and analysis of wind power utilization

The utilization indices of wind power about the constant power factor operation mode ($\cos\varphi = 1$) and the optimal high penetration operation scheme are calculated respectively as shown in Table 2. The indices are the utilization ratio of wind power (i.e., the percentage of actual active power output to the forecasting maximum active power of dispersed wind farms) and the energy penetration ratio of DWP (i.e., Eq. 4). The constant

FIGURE 5
Diagram of case data in a typical day.
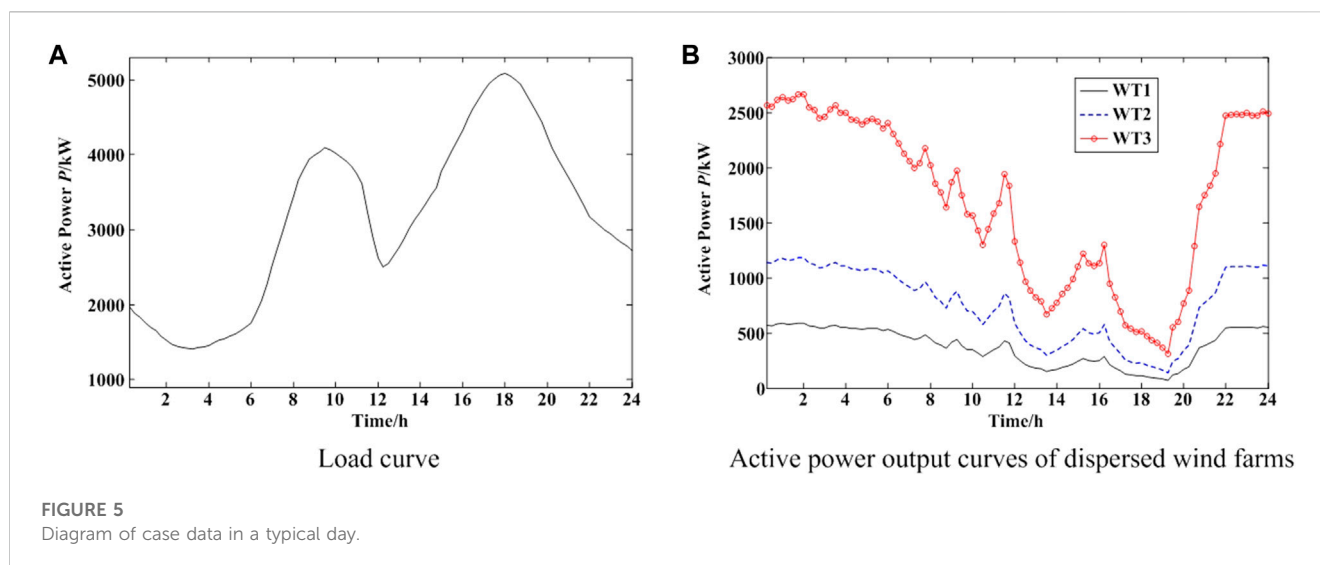
power factor operation mode represents a kind of operation method of distribution network with distributed generations. It did not explore the power regulation capability of the wind turbines, and the active power output of DWP was adjusted passively according to the load demand of distribution network. In order to ensure the normal operation of the system, in cases when the node voltage of the distribution network exceeds the specified limits in the constant power factor operation mode, the measures of wind power curtailment and electricity restriction are adopted to reduce the active power output to maintain the node voltage to be qualified.

As shown in Table 2, the utilization ratio and the penetration ratio of wind power are increased by 20.59% and 18.56% respectively based on the high penetration operation scheme. Taking the dispersed wind farm (WT2) which integrated into Node 18 as an example, Figure 8 shows the comparison on the active power output of wind power in different operation modes. During 0 a.m.–8 a.m. and 10 p.m. to 0 a.m. on the next day, the

FIGURE 6
Reactive power output curves of dispersed wind farms.

active power output of wind power based on the high penetration operation scheme is increased significantly compared with the constant power factor operation mode. It indicates that the proposed optimal operation strategy achieves high penetration and efficient utilization of wind power resources.

5.2.2 Comparison and analysis of voltage quality

The voltage curves of typical nodes about the constant power factor operation mode and the optimal high penetration operation scheme are shown in Figure 9 (Wind turbines only output active power in the constant power factor operation mode). In order to avoid redundancy in image display, three typical nodes are selected in Figure 9 to demonstrate the effect of the optimal operation strategy. Node 18 and Node 32 are both the points of common coupling (PCC) of DWP in different feeder sections of the distribution network, the voltage of these two nodes is most affected by the wind power output. Moreover, according to the results of simulation calculation, the value of node voltage at Node 18 is the upper voltage limit in the test system. Therefore, these two nodes are typical examples of their respective feeder sections to show the optimization effect of node voltage. The location of Node 10 is between the 10 kV bus and the PCC of wind power, it can reflect the voltage of some nodes which are far away from the PCC of DWP to a certain extent. The comparison between the Figures 9A,B shows that the node voltage will exceed the specified limits during part of time period in constant power factor operation mode, while the optimized node voltage can meet the requirements for safe operation of the distribution network.

Combined with the Figures 5–7, we can clearly see that: 1) During 0 a.m.–6 a.m., in cases when the distribution network is weakly loaded while the output of wind power is close to full capacity, the dispersed wind farms will be separated from the distribution network because of overvoltage in the event of untimely adjustment, which is shown in Figure 9A. But when the high penetration operation scheme of wind power is adopted, the reactive power capability of the wind turbines is made full use to regulate the node voltage of the distribution network. As shown in
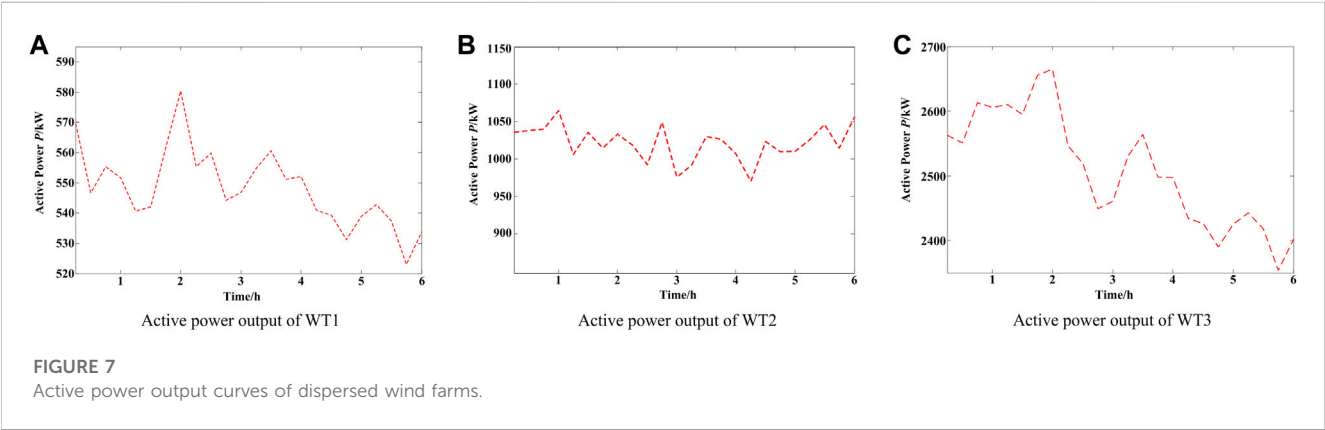
FIGURE 7
Active power output curves of dispersed wind farms.
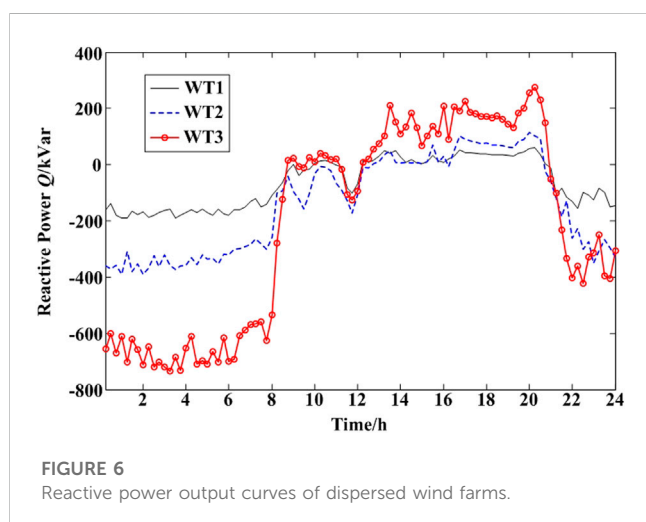
TABLE 2 Comparison on two operation modes of wind power.

Index	Utilization ratio (%)	Penetration ratio (%)
Constant power factor operation mode	77.46	69.87
High penetration operation scheme	98.05	88.43

FIGURE 8
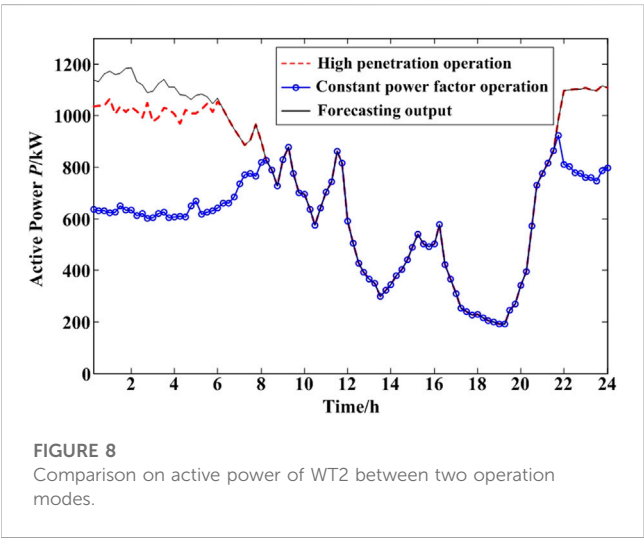Comparison on active power of WT2 between two operation modes.

Figure 9B, the capacity of reactive power is limited, so the active power output of each wind farm should be reduced appropriately to adjust the voltage to the defined scope. 2) During 5 p.m.–8 p.m., in cases when the distribution network is heavily loaded while the output of wind power is at a low level, the active power output of the dispersed wind farms cannot meet the load requirement of the system, as shown in Figure 9A, the voltage of some nodes will be less than the lower boundary in the constant power factor operation mode. At this time, the quality of the node voltage is failed to reach the required standard of safe operation. If it continues to operate like that, there may even appear quite serious consequences such as voltage instability. During this period, as shown in Figure 9B, the

high penetration operation scheme can adjust the reactive power output of wind turbines to regulate the node voltage actively and keep them above 0.95 p. u. Generally speaking, the optimal high penetration operation scheme can keep the voltage of each node within the safe and high-quality range of 0.95–1.05p.u., the voltage quality is obviously improved.

In summary, the simulation test on the IEEE 33 bus system demonstrates that the proposed optimal operation method can realize the high penetration and efficient utilization of wind power on the premise of ensuring high quality of voltage. When the problem of overvoltage cannot be solved only by the reactive power capacity of dispersed wind farms, the active power output of wind farms should be reduced to adjust the node voltage to a qualified range.

6 Case analysis of demonstration project

The optimal operation model and solution method proposed in this paper have been applied to a demonstration project in Northern Shaanxi province of China. It is one of the first demonstration engineering applications for dispersed wind farms integrated into the 10 kV distribution networks in China. The local distribution network structure of the demonstration project is presented in Figure 10, and the details of the dispersed wind farm are shown in Table 3.

As shown in Figure 11, the load curve and active power output curves of dispersed wind farms in a typical day are used to simulation. The operating cycle is 24 h and the time interval is 15 min. The algorithm parameters are set up as: the initial number of individuals $N_s = 50$, while the capacity of archive
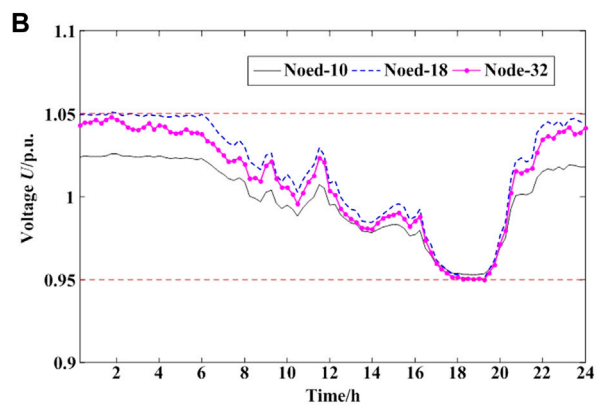
FIGURE 9
Voltage curves of typical nodes.
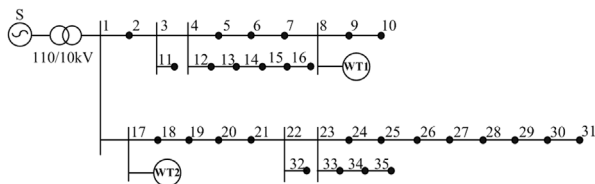
FIGURE 10
Diagram of dispersed wind farm integrated into the local distribution network.

TABLE 3 The parameters of the dispersed wind farm.

Number	Node	Capacity/MW
WT1	8	4.2
WT2	17	6.3

set $M = 20$, the maximum number of iterations $KI_{max} = 500$, the number of clusters $k = 4$, and the four probability parameters are $P_1 = 0.9$, $P_2 = 0.3$, $P_3 = P_4 = 0.5$.

The MBSO algorithm is used to solve the Pareto solution set and select the optimal operation scheme. The simulation results are shown as follows. Figure 12A shows the reactive power output of dispersed wind farms after optimization. The voltage curves of typical nodes about the constant power factor operation mode and the optimal high penetration operation scheme are shown in Figures 12B,C.

Combined with Figures 11, 12, the optimized results illustrate that:1) During 0 a.m.–8 a.m., that is the time of valley load of the distribution network and almost full active power output of wind power, the voltage of the point of common coupling exceeds the upper boundary (1.07p.u.) in the constant power factor operation mode, as shown in Figure 12B. While the optimal operation scheme of wind power can adjust the node voltage to the scope of safe operation shown in Figure 12C, which indicates that the reactive power regulation capability of wind turbines is sufficient
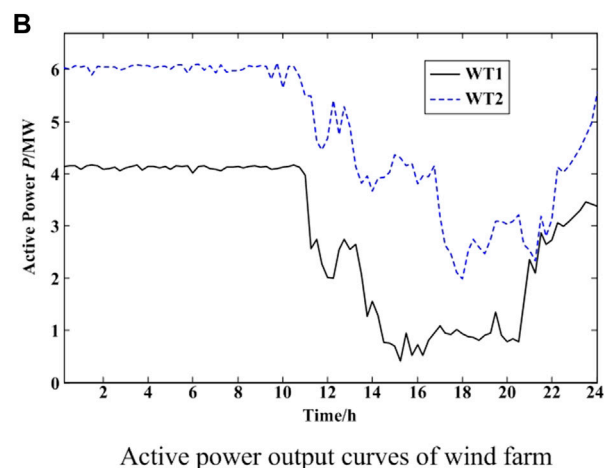
FIGURE 11
Diagram of engineering case data.

FIGURE 12
Diagram of optimization results.
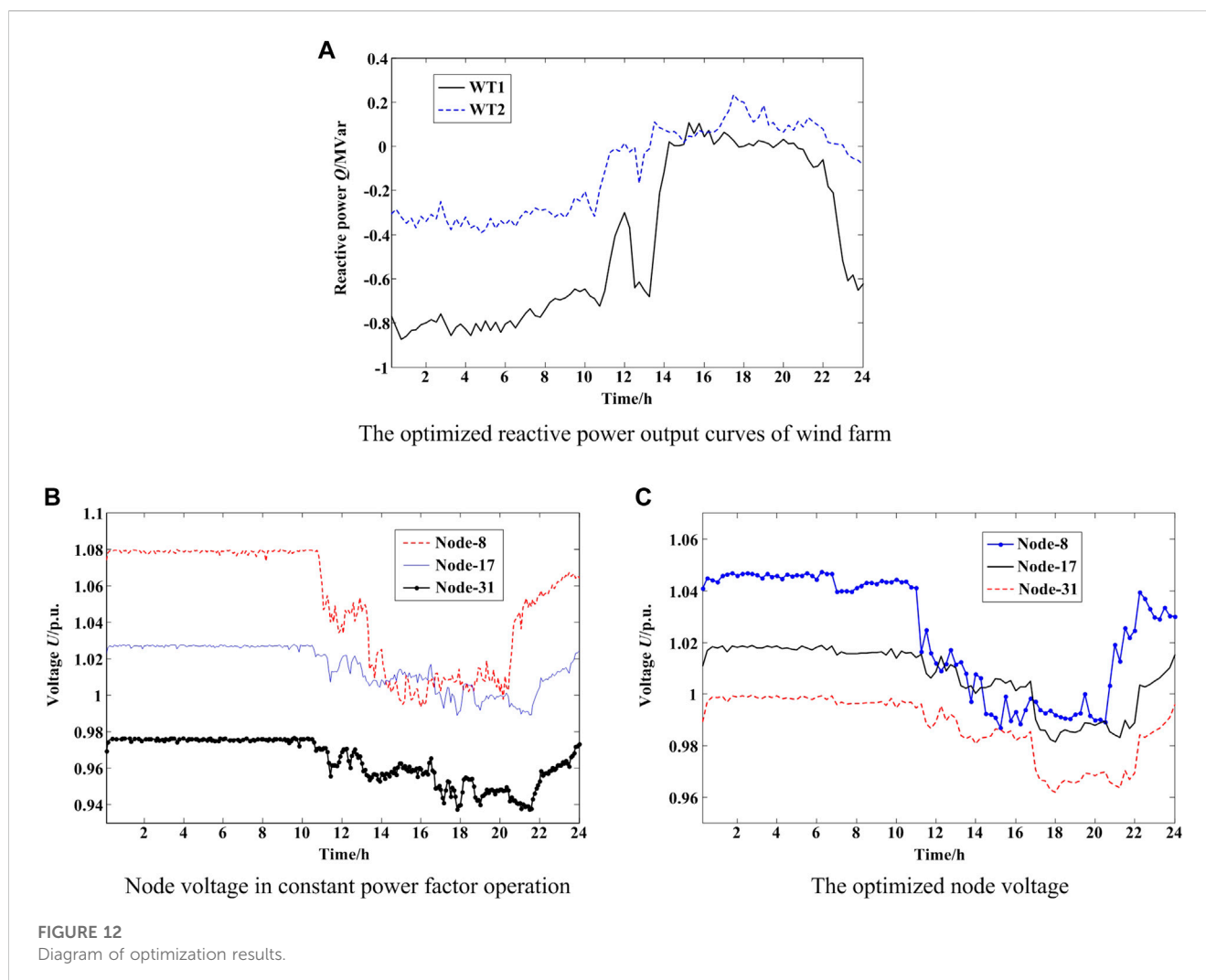
to meet the voltage requirements and realize the efficient utilization of wind power. 2) During 6 p.m.–10 p.m., that is the time of peak load of the distribution network and low active power output of wind power, the voltage is slightly low at the terminal node of feeder lines in the constant power factor operation mode, which is shown in Figure 12B. When the load or wind power output fluctuates, there is a high probability that the node voltage will exceed the lower boundary. The reactive power output of wind turbines can improve the quality of node voltage based on the optimal operation scheme, as shown in Figure 12C, which indicates that the reactive power of dispersed wind farms can participate in the voltage regulation of distribution network effectively.

In this practical case, the high penetration operation scheme has no restriction on active power output, it means that the active power output of wind farms has been fully utilized, i.e., the penetration ratio of wind power reaches the maximum. The above simulation results are basically consistent with the actual operation of the dispersed wind farm. According to statistics, the annual operation time of the demonstration project is more than 3,000 h. The utilization ratio of active power of DWP maintains at a relatively high level, which is the genuine implementation of the efficient utilization of wind power.

7 Conclusion

This paper proposes an optimal operation strategy of distribution network with high penetration of dispersed wind power, in which the power regulation capability of wind turbines is utilized sufficiently. A novel optimal operation model of distribution network is established, which pursues high penetration integration of DWP and maximum correntropy of node voltage. Considering the characteristics of the model, the MBSO algorithm is adopted to obtain the optimal operation scheme for the distribution network. The simulation results on expanded IEEE 33 bus system verify that the proposed optimal operation strategy can effectively ensure the high quality of node voltage and realize the high penetration integration of wind power. Moreover, a case of engineering application for dispersed wind farms also indicates the feasibility of the

optimal operation strategy, which shows great application prospect and popularization value.

Data availability statement

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

Author contributions

Conceptualization, JD and LC; methodology, LC; software, WL; validation, DZ and QS.

Funding

Acknowledgments

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

Anilkumar, R., Devriese, G., and Srivastava, A. K. (2018). Voltage and reactive power control to maximize the energy savings in power distribution system with wind energy. *IEEE Trans. Industry Appl.* 54 (1), 656–664. doi:10.1109/TIA.2017.2740850

Bessa, R. J., Miranda, V., and Gama, J. (2009). Entropy and correntropy against minimum square error in offline and online three-day ahead wind power forecasting. *IEEE Trans. Power Syst.* 24 (4), 1657–1666. doi:10.1109/tpwrs.2009.2030291

Guo, Y., Gao, H., Wu, Q., Østergaard, J., Yu, D., and Shahidehpour, M. (2019). Distributed coordinated active and reactive power control of wind farms based on model predictive control. *Int. J. Electr. Power and Energy Syst.* 104, 78–88. doi:10.1016/j.ijepes.2018.06.043

Joseph, A., Smedley, K., and Mehraeen, S. (2020). Secure high DER penetration power distribution via autonomously coordinated Volt/VAR control. *IEEE Trans. Power Deliv.* 35 (5), 2272–2284. doi:10.1109/tpwrd.2020.2965107

Karimi, M., Mokhlis, H., Naidu, K., Uddin, S., and Bakar, A. H. A. (2016). Photovoltaic penetration issues and impacts in distribution network–A review. *Renew. Sustain. Energy Rev.* 53, 594–605. doi:10.1016/j.rser.2015.08.042

Li, J., Wang, N., Zhou, D., Hu, W., Huang, Q., Chen, Z., et al. (2020). Optimal reactive power dispatch of permanent magnet synchronous generator-based wind farm considering levelised production cost minimisation. *Renew. Energy* 145, 1–12. doi:10.1016/j.renene.2019.06.014

Li, P., and Zheng, M. (2019). Multi-objective optimal operation of hybrid AC/DC microgrid considering source-network-load coordination. *J. Mod. Power Syst. Clean Energy* 7 (5), 1229–1240. doi:10.1007/s40565-019-0536-3

Liu, W., Pokharel, P. P., and Principe, J. C. (2007). Correntropy: Properties and applications in non-Gaussian signal processing. *IEEE Trans. Signal Process.* 55 (11), 5286–5298. doi:10.1109/tsp.2007.896065

Londero, R. R., Affonso, C. D. M., and Vieira, J. P. A. (2017). Effects of operational limits of DFIG wind turbines on long-term voltage stability studies. *Electr. Power Syst. Res.* 142, 134–140. doi:10.1016/j.epsr.2016.09.007

Ma, Z., An, T., and Shang, Y. (2016). State of the art and development trends of power distribution technologies. *Proc. CSEE* 36 (6), 1552–1567. (in Chinese). doi:10.13334/j.0258-8013.pcsee.2016.06.011

National Energy Administration (2018). Interim management measures for development and construction of dispersed wind power projects. http://zfxxgk.nea.gov.cn/auto87/201804/t20180416_3150.htm [Accessed April 3, 2018]. (in Chinese).

Shi, Y., Xue, J., and Wu, Y. (2013). Multi-objective optimization based on brain storm optimization algorithm. *Int. J. Swarm Intell. Res.* 4 (3), 1–21. doi:10.4018/ijsir.2013070101

Tang, J., Wang, D., Wang, X., Jia, H., Wang, C., Huang, R., et al. (2017). Study on day-ahead optimal economic operation of active distribution networks based on Kriging model assisted particle swarm optimization with constraint handling techniques. *Appl. Energy* 204, 143–162. doi:10.1016/j.apenergy.2017.06.053

Tuladhar, S. R., Singh, J. G., and Ongsakul, W. (2016). Multi-objective approach for distribution network reconfiguration with optimal DG power factor using NSPSO. *IET Generation Transm. Distribution* 10 (12), 2842–2851. doi:10.1049/iet-gtd.2015.0587

Xiang, Y., Zhou, L., Huang, Y., Zhang, X., Liu, Y., and Liu, J. (2021). Reactive coordinated optimal operation of distributed wind generation. *Energy* 218, 119417. doi:10.1016/j.energy.2020.119417

Xiong, G., and Shi, D. (2018). Hybrid biogeography-based optimization with brain storm optimization for non-convex dynamic economic dispatch with valve-point effects. *Energy* 157, 424–435. doi:10.1016/j.energy.2018.05.180

Yarahmadi, M., and Shakarami, M. R. (2018). An analytical and probabilistic method to determine wind distributed generators penetration for distribution networks based on time-dependent loads. *Int. J. Electr. Power and Energy Syst.* 103, 404–413. doi:10.1016/j.ijepes.2018.06.025

Zhang, X., Duan, J., Shi, X., Wang, K., Dong, Y., Mu, T., et al. (2017). Research of maximum control strategy with dispersed wind power considering reactive capability of DFIG. *Proc. CSEE* 37 (7), 2001–2008. (in Chinese). doi:10.13334/j.0258-8013.pcsee.152819

Zhao, F., Si, J., and Wang, J. (2016). Research on optimal schedule strategy for active distribution network using particle swarm optimization combined with bacterial foraging algorithm. *Int. J. Electr. Power and Energy Syst.* 78, 637–646. doi:10.1016/j.ijepes.2015.11.112

frontiers | Frontiers in Energy Research

*CORRESPONDENCE
Ping Yang,
✉ eppyang@scut.edu.cn

An MPC based active and reactive power coordinated control strategy of PMSG wind turbines to enhance the support capability

Xiyuan Ma[1], Jingyi Yu[1], Ping Yang[2]*, Pengyu Wang[1] and Peng Zhang[2]

[1]Digital Grid Research Institute of China Southern Power Grid, Guangzhou, China, [2]Guangdong Key Laboratory of Green Energy Technology, South China University of Technology, Guangzhou, China

As wind turbines are constantly replacing traditional units, it is becoming a consensus that wind turbines should participate in the grid support that was only responsible by traditional units in the past. In order to enhance the grid support capabilities (including active power support and reactive power support) of permanent magnet synchronous generator (PMSG) based wind turbines, this paper constructs an active and reactive power coordinated control strategy. Compared with the current active and reactive power coordinated control strategy of PMSG wind turbines, the method of the proposed one innovatively considers the climbing coordinated restriction between active and reactive power, flexible prioritization arrangement between active and reactive power, the accurate amplitude and climbing constraints of grid-side converters' output voltage, and the model predictive control (MPC) technique. The simulation results verify that the proposed power control strategy can make PMSG wind turbines achieve excellent power output performance and thus better meet the requirements of power grid support.

KEYWORDS

grid support, permanent magnet synchronous generator (PMSG) wind turbine, active and reactive power coordinated control strategy, model predictive control (MPC), amplitude and climbing

1 Introduction

Traditionally, the frequency and voltage of power systems are mainly controlled by thermal units and hydropower units. Wind and photovoltaic units account for a small generating capacity, and their power decouples with the grid frequency and voltage, standing at maximum power points and fixed power factors (Gaied et al., 2022; Zeng et al., 2022). However, with the implementation of the "double carbon" strategy in recent years, the installed capacity of wind and photovoltaic energy in China is increasing significantly, which reached approximate 706 GW by the end of September 2022 (National Energy Administration, 2023). The traditional units responsible for frequency and voltage control are gradually being replaced by the wind and photovoltaic units, which is gradually weakening the frequency and voltage control capability of power systems. Hence, it is a growing consensus that wind and photovoltaic units should provide frequency and voltage support services to power systems (i.e., grid support), so as to maintain the security operation of power systems (Hansen et al., 2006; Feltes et al., 2009).

Permanent magnet synchronous generator (PMSG) wind turbine, one of the mainstream wind turbines, is receiving more and more popularity nowadays (Guo et al., 2021). The strengths of a PMSG wind turbine include high energy efficiency and low maintenance costs (Li et al., 2012; Musarrat et al., 2019). In addition, a PMSG wind turbine contains a back-to-back full-scale converter for connecting to power systems so that power system faults and abnormal conditions will not directly affect the power output of the generator. This means that it holds inherent advantages in fault ride through and grid support (Tan et al., 2017; Sheng et al., 2021).

To provide grid support services, it is necessary for a PMSG wind turbine to adjust its active and reactive power to respond to the changes of grid frequency and voltage. Topics about power control strategies of a PMSG wind turbine/wind farm to provide grid support could be found in a great deal of literature (Wu et al., 2017; Peng et al., 2021; Zhong et al., 2021; Li et al., 2022; Okedu, 2022). For instance, a control strategy is proposed in literature (Wu et al., 2017) for a PMSG wind turbine coordinating with a battery system to provide frequency support, which is realized by instantly raising its active power to a predefined level once grid frequency disturbance occurs. In (Peng et al., 2021), the authors design a reactive voltage support method for a wind farm with static synchronous compensators considering remaining capacities and voltage unbalanced factors for different PMSG wind turbines. Yet, the power control strategies in literature (Wu et al., 2017; Peng et al., 2021; Zhong et al., 2021; Li et al., 2022; Okedu, 2022) do not fall under the category of coordinated control between active and reactive power.

Unlike loose regulatory codes for daily power generation, in the periods of providing grid support, the active and reactive power of a PMSG wind turbine is required to meet exact amplitude and response rate requirements, which are calculated based on the deviation or slope of frequency and voltage (Mohseni and Islam, 2012; Liu et al., 2015; You et al., 2015). In this context, active and reactive power coordinated control which involves active and reactive power coordinated restriction and priority decisions becomes an unavoidable and meaningful problem. On the one hand, active and reactive power coordinated restriction can improve the power control performance of a PMSG wind turbine in order to better provide grid support. On the other hand, when the apparent power of a PMSG wind turbine is greater than its rated apparent power, the priority of active and reactive power must be judged and then reduce the party with lower priority in order to ensure that the party with higher priority meets the power grid support requirements.

As shown in Figure 1, the shaded part represents the amplitude range of active and reactive power that PMSG wind turbines can output, which picture forms a semicircle or rectangle when active and reactive power is restricted coordinately or separately. To obey the given apparent power amplitude constraint, the semicircle's radius is equal to maximum apparent power S, and the rectangle is contained in a semicircle with a radius of maximum apparent power S. Obviously, under the given apparent power amplitude constraint, active and reactive power amplitude coordinated restriction enables PMSG wind turbines to output a larger range of active and reactive power amplitude. By the same token, under the given apparent

FIGURE 1
Amplitude range of active and reactive power that PMSG wind turbines can output. **(A)** coordinated restriction. **(B)** separated restriction.
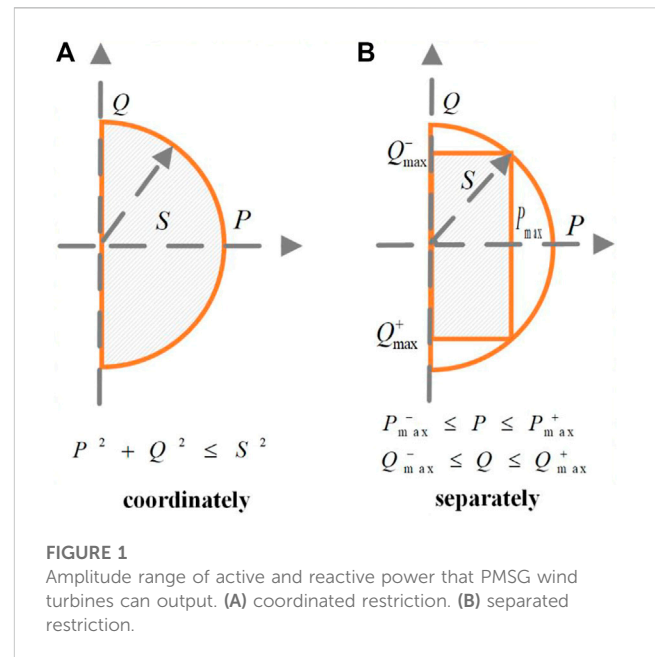
power climbing constraint, active and reactive power climbing coordinated restriction allows PMSG wind turbines to perform a larger range of active and reactive power climbing. Therefore, active and reactive power coordinated restriction is contributing to improving the PMSG wind turbines' power control performance.

For this reason, some literature pays attention to designing the coordinated control strategies of active and reactive power for PMSG wind turbines providing grid support (Nguyen et al., 2013; Moghadasi and Sarwat, 2015; Yan et al., 2016; Zhang et al., 2016; Tripathi et al., 2019). Ref (Dong et al., 2012). presents a coordinated control strategy to enhance the low voltage ride through and grid support capability of PMSG wind turbines. In (Khazaei et al., 2020), a consensus-based control strategy is proposed to regulate the output of PMSG wind turbines and distributed batteries in a wind farm to deliver active and reactive power to the load. These strategies only focus on amplitude coordinated restriction of active and reactive power, which will become more perfect if the climbing coordinated restriction of active and reactive power can be taken into account at the same time.

In summary, most previous power control strategies for PMSG wind turbines providing grid support either do not explore the coordinated control between active and reactive power, or only focus on amplitude coordinated restriction and ignore the climbing coordinated restriction between active and reactive power. In addition, the priority decisions of active and reactive power in previous power coordinated control strategies are not flexible enough to apply to multiple grid support scenarios. Model predictive control (MPC) (Rodriguez et al., 2009; Mayne, 2014), which can handle complex constraints and achieve multiple optimization objectives, is suitable to structure flexible power coordinated control strategies. MPC is numerous applicated in industrial process control (Qin and Badgwell, 2003; Venkat et al., 2008; Vazquez et al., 2014) and power control of PMSG wind turbines (Maaoui-Ben Hassine

FIGURE 2
The composition of the PMSG wind turbine studied.
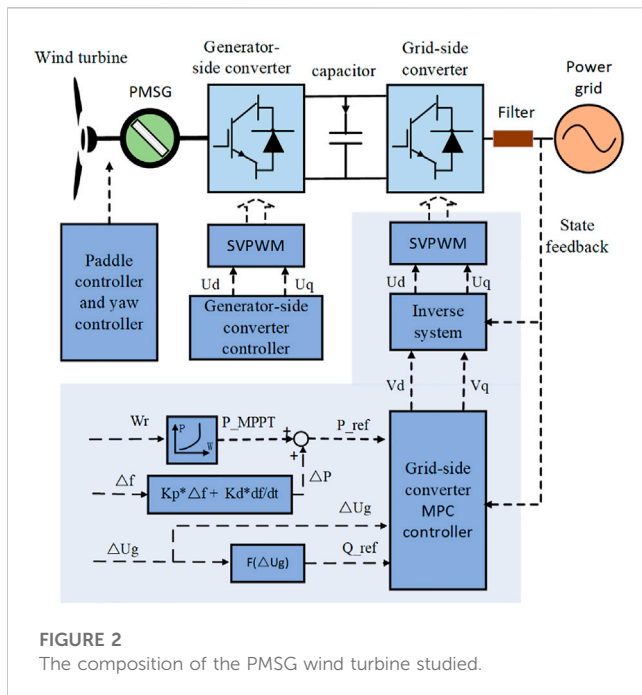
et al., 2016; Shehata, 2017; Mishra and Saha, 2020), while has not been found to be applied in PMSG wind turbines' active and reactive power coordinated control. Finally, the current MPC models on the PMSG wind turbines lack the accurate amplitude and climbing constraints of grid-side converters' output voltage.

Based on the mentioned issues, this paper proposes an MPC based active and reactive power coordinated control strategy of PMSG wind turbines to enhance the grid support capability. The contributions of this article are as follows.

(1) Take the lead in adopting MPC to build an active and reactive power coordinated control strategy of PMSG wind turbines that can be flexibly applied in different grid support scenarios due to the MPC can flexibly set the priority of active and reactive power.

(2) Not only the existing amplitude coordinated restriction between active and reactive power but also the innovative climbing coordinated restriction between active and reactive power is considered in the proposed control strategy, which enables PMSG wind turbines to perform a wider range of active and reactive power amplitude and climbing. So it enhances the power control performance and the grid support capability of PMSG wind turbines.

(3) The amplitude and climbing constraint model of grid-side converter voltage is structured in the proposed MPC based control strategy in order to protect the grid-side converters and dignify the output voltage waveform.

The rest of this paper is organized as follows. Section 2 describes the overview of PMSG wind turbines. Section 3 displays the mathematical model of PMSG wind turbines. The MPC based active and reactive power coordinated control strategy is introduced in Section 4 and simulated in Section 5. Finally, Section 6 summarizes this paper.

2 The system description

The composition of the PMSG wind turbine studied is shown in Figure 2, which mainly consists of a wind turbine, a PMSG, a back-to-back converter, and multiple controllers. In this paper, the generator-side converter is employed to maintain capacitor voltage and the grid-side converter controls active power to realize maximum power point tracking and provide grid frequency support. Meanwhile, the grid-side converter is also used to control the reactive power exchange to the power grid, so as to provide grid voltage support. Therefore, the key to providing frequency and voltage support for the power grid is to control the active and reactive power of the grid-side converter.

In order to obtain the wider range of active and reactive power amplitude and climbing under the given apparent power amplitude and climbing constraints, in this paper, both the amplitude coordinated restriction and climbing coordinated restriction of the grid-side converter's active and reactive power are considered. In view of the obvious advantages of MPC in dealing with multi-input and multi-output control problems with complex constraints and specific objectives, this paper applies MPC to struct the active and reactive power coordinated control strategy of the grid-side converter.

The MPC based active and reactive power coordinated control system of the grid-side converter is presented in Figure 2. The MPC controller uses the predictive model of the controlled system to predict the behavior of the controlled system under different control actions and selects the optimal control action which minimizes the objective function. According to the situations of the power grid and the PMSG wind turbine, the objective function is automatically adjusted so that the proposed control strategy can adapt to different grid support scenarios.

Firstly, the mathematical model of the grid-side converter is established. Secondly, an inverse system is designed in order to change the grid-side converter into a pseudolinear composite system. Thirdly, the prediction model and constraint conditions of the MPC controller are built according to the state space equation and operation boundary of the pseudolinear composite system, and the objective function of the MPC controller is built according to the reference value and priority of active and reactive power. Finally, the optimal control action is obtained by solving the optimization problem.

3 The system model

3.1 Mathematical model of grid side converter

By analyzing the three-phase circuit between the grid-side converter and the power system as shown in Figure 2, the three-phase mathematical model of the grid-side converter system can be written as follows:

$$U_{abc} - E_{abc} = I_{abc}R + L\frac{dI_{abc}}{dt} \tag{1}$$

where U_{abc} and E_{abc} represent the three-phase voltage of the grid-side converter and the power system respectively. R and L are the

resistor and inductor of the filter. I_{abc} represents the three-phase grid current flowing into the power grid from the grid-side converter.

The three-phase mathematical model (1) of the grid-side converter can be transformed to a mathematical model in dq synchronously rotating reference frame, which can be expressed as follows.

$$L\frac{di_d}{dt} = -Ri_d + \omega Li_q + u_d - e_d$$

$$L\frac{di_q}{dt} = -Ri_q - \omega Li_d + u_q - e_q \tag{2}$$

$$\begin{aligned} e_d &= e \\ e_q &= 0 \end{aligned} \tag{3}$$

where i_d and i_q, u_d and u_q, e_d and e_q are the grid current, converter voltage, grid voltage in dq-axis. e and ω are the amplitude and angular frequency of the grid voltage, which are uncertain and varying parameters. The active and reactive power flowing into the power grid can be given by

$$P = \frac{3}{2}ei_d$$

$$Q = \frac{3}{2}ei_q \tag{4}$$

3.2 Grid Side Converter Model Linearization

It can be seen that the nonlinearities exist in the grid-side converter's mathematical model (2) since it contains the nonlinear terms e and ω. In this paper, an inverse system is designed and connected in series to the control loop to cope with the nonlinearity existing in the grid-side converter's mathematical model. The inverse system uses real time state feedback to compensate the control variables v_d and v_q, which the compensation law can be expressed as follows.

It is worth mentioning that, due to the need to solve the optimization model, the sampling interval of an MPC controller cannot be set as a small value, while the sampling interval of an inverse system can be set to very small, so the inverse system can sample the state variables (i_d and i_q), amplitude e and angular frequency ω in real time to compensate the control variables v_d and v_q. The real time system states (i_d and i_q) can be obtained by measuring the real time three-phase current and making a Park's Transformation.

$$\begin{aligned} u_d &= Lv_d + Ri_d - \omega Li_q + e \\ u_q &= Lv_q + Ri_q + \omega Li_d \end{aligned} \tag{5}$$

As shown in Figure 2, v_d and v_q are the inputs of the inverse system. u_d and u_q are the outputs of the inverse system and as converter control voltages to input in the grid-side converter system. The mathematical model of the composite system composed of the inverse system and the grid-side converter system can be obtained by substituting (5) into (2)–(4).

$$\frac{di_d}{dt} = v_d$$

$$\frac{di_q}{dt} = v_q \tag{6}$$

$$P = \frac{3}{2}ei_d$$

$$Q = \frac{3}{2}ei_q \tag{7}$$

Therefore, a grid-side converter composite system is structed which inputs are v_d and v_q, and outputs are the active power P and reactive power Q.

The state variables and output variables of the composite system are the same as those of the grid-side converter system. Therefore, to control the composite system is essential to control the converter system grid-side converter system.

The state representation Eq. 2 of the grid-side converter system contains the nonlinear terms e and ω. If the amplitude e and angular frequency ω of the grid voltage change, the MPC controller may not achieve the expected effect of predicting and controlling the state variables. In contrast, the state representation Eq. 6 of the composite system is linear and definite, so that the state variables can be accurately predicted and controlled by the MPC controller whether the amplitude and angular frequency of the grid voltage change or not. Therefore, compared with the direct control of the grid-side converter system, the control of the composite system can obtain a more reliable control effect.

The grid-side converter composite system model (6)–(7) can be written as a normal linear state space model.

$$\begin{aligned} \dot{x} &= Ax + Bv \\ y &= Cx \\ x &= [i_d, i_q]^T, \ y = [P, Q]^T \\ A &= \begin{bmatrix} 0,0 \\ 0,0 \end{bmatrix}, \ B = \begin{bmatrix} 1,0 \\ 0,1 \end{bmatrix}, \ C = \begin{bmatrix} 1.5^*e, \ 0 \\ 0, \ 1.5^*e \end{bmatrix} \end{aligned} \tag{8}$$

If the amplitude of the grid voltage is stable, the system model can accurately describe the output characteristic of the system. So that the MPC controller can accurately predict and control the outputs of the system. The system outputs will gradually reach the reference values under the control of the MPC controller.

In contrast, if the amplitude of the grid voltage changes after being measured, this means that the output characteristic of the system has changed and does not match the currently established system model during this sampling interval. The MPC controller may not achieve the expected effect of predicting and controlling the system outputs during this sampling interval. However, the system model will be corrected because the MPC controller will re-measure the grid voltage amplitude at the beginning of the next sampling interval. Based on the correct system model, the system outputs will gradually reach the reference values under the control of the MPC controller. Therefore, the change of the grid voltage amplitude will not cause the system to lose stability, but make the outputs of the system fail to reach the expected value within the current sampling interval.

The MPC algorithm is not executed continuously in the controller, but at regular interval, which is called sampling interval or control period and is denoted by T. Discretization (8) can result in

$$\begin{aligned} x(k+1) &= A_d x(k) + B_d v(k) \\ y(k) &= C_d x(k) \\ x(k) &= [i_d(k), i_q(k)]^T, \ v(k) = [v_d(k), v_q(k)]^T \\ y(k) &= [P(k), Q(k)]^T \\ A_d &= e^{AT}, \ B_d = \int_0^T e^{At}dt \cdot B, \ C_d = C \end{aligned} \tag{9}$$

where $x(k)$, $v(k)$, $y(k)$ are the states, inputs and outputs of the grid-side converter composite system in discrete time. Next, the MPC controller is designed to control the grid-side converter composite system, based on the discrete state-space model (9).

4 The MPC based control strategy

The prediction horizon refers to the time range from the current time point to a certain time point in the future. The MPC algorithm needs to predict the state variables of the composite system in this time range. The time length of the prediction horizon is an integral multiple of the MPC sampling interval/control interval, so it can be expressed as $N_P{*}T$. Similarly, The control horizon refers to the time range from the current time point to a certain time point in the future. The MPC algorithm needs to compute the optimal control variables of the composite system in this time range. The time length of the control horizon is an integral multiple of the MPC sampling interval/control interval, so it can be expressed as $N_C{*}T$.

In this paper, $N_P = 3$ and $N_C = 2$ when we introduce the formulas of the optimization problem. This means that the MPC algorithm needs to predict the state variables of the system from the time $t = (k+1){*}T$ to the time $t = (k+3){*}T$, meanwhile, the MPC algorithm needs to compute the optimal control variables of the system in the period from the time $t = kT$ to the future time $t = (k+2){*}T$. In fact, they can take other values. With the increase in the N_P and N_C, the computing time and controller performance of the MPC algorithm will increase. When setting the values of N_P and N_C, it is necessary to ensure that the computing time of the MPC algorithm cannot exceed the MPC sampling interval.

In each sampling interval, the MPC controller solves the optimization problem (8–23) and obtains the optimal control variables $v(t) = [v_d(t), v_q(t)], t = kT, (k+1)T, \ldots (k+N_C)T$ of future N_C sampling interval, and then only the first set of optimal control variable $[v_d(kT), v_q(kT)]$ will be provided for the composite system.

4.1 Constraint conditions of power amplitude

Generalized power amplitude constraint conditions include the power amplitude constraints and current amplitude constraints. Based on the discrete state-space model (9), the grid current amplitudes in the future triple control periods can be predicted as follows.

$$
\begin{aligned}
x(k+1) &= A_d x(k) + B_d v(k) \\
x(k+2) &= A_d x(k+1) + B_d v(k+1) = A_d^2 x(k) + A_d B_d v(k) \\
&\quad + B_d v(k+1) \\
x(k+3) &= A_d x(k+2) + B_d v(k+2) = A_d^3 x(k) + A_d^2 B_d v(k) \\
&\quad + A_d B_d v(k+1) + B_d v(k+2)
\end{aligned}
$$

(10)

The prediction model (10) of grid current amplitudes can be rewritten in matrix form.

$$
\begin{aligned}
X &= A_D x(k) + B_D V \\
X &= [x(k+1); x(k+2); x(k+3)] \\
V &= [v(k); v(k+1); v(k+2)] \\
A_D &= [A_d; A_d^2; A_d^3] \\
B_D &= \begin{bmatrix} B_d, & 0_{(2\times2)}, & 0_{(2\times2)} \\ A_d B_d, & B_d, & 0_{(2\times2)} \\ A_d^2 B_d, & A_d B_d, & B_d \end{bmatrix}
\end{aligned}
$$

(11)

The future grid current amplitudes should not be greater than of the rated current amplitude of the grid-side converter. Then, we have.

$$
|x(j)|^2 = i_d(j)^2 + i_q(j)^2 \le i_N^2, \, j = k+1, \ldots, k+3 \quad (12)
$$

Compared to separately restricting each active current $i_d(j)$ and reactive current $i_q(j)$ amplitude such as $i_d{}^{\min} \le i_d(j) \le i_d{}^{\max}, i_q{}^{\min} \le i_q(j) \le i_q{}^{\max}$, coordinately restricting active current i_d and reactive current $i_q(j)$ amplitudes as (12) makes the amplitude range of active current and reactive current larger.

Similarly, based on discrete state-space model (9), the power amplitudes in the future triple control periods can be predicted.

$$
\begin{aligned}
y(k+1) &= C_d A_d x(k) + C_d B_d v(k) \\
y(k+2) &= C_d A_d x(k+1) + C_d B_d v(k+1) \\
&= C_d A_d^2 x(k) + C_d A_d B_d v(k) + C_d B_d v(k+1) \\
y(k+3) &= C_d A_d x(k+2) + C_d B_d v(k+2) \\
&= C_d A_d^3 x(k) + C_d A_d^2 B_d v(k) + C_d A_d B_d v(k+1) \\
&\quad + C_d B_d v(k+2)
\end{aligned}
$$

(13)

The prediction model (12) of the power amplitudes can be rewritten in matrix form.

$$
\begin{aligned}
Y &= C_D x(k) + D_D V \\
Y &= [y(k+1); y(k+2); y(k+3)] \\
V &= [v(k); v(k+1); v(k+2)] \\
C_D &= [(C_d A_d); (C_d A_d^2); (C_d A_d^3)] \\
D_D &= \begin{bmatrix} C_d B_d, & 0_{(2\times2)}, & 0_{(2\times2)} \\ C_d A_d B_d, & C_d B_d, & 0_{(2\times2)} \\ C_d A_d^2 B_d, & C_d A_d B_d, & C_d B_d \end{bmatrix}
\end{aligned}
$$

(14)

The future apparent power amplitudes of the grid-side converter should not be greater than its rated apparent power amplitude. Then, we have.

$$
|y(j)|^2 = P(j)^2 + Q(j)^2 \le S_N^2, \, j = k+1, \ldots, k+3 \quad (15)
$$

Compared to separately restricting each active power $P(j)$ and reactive power $Q(j)$ amplitude such as $P^{\min} \le P(j) \le P^{\max}, Q^{\min} \le Q(j) \le Q^{\max}$, coordinately restricting active power $P(j)$ and reactive power $Q(j)$ amplitudes as (15) makes the amplitude range of active power and reactive power larger.

4.2 Constraint conditions of power climbing and voltage

Based on the state space model (6)–(9), we can see that the climbing of active power and reactive power depends on the control variables v_d and v_q. To restrict the maximum climbing of apparent

FIGURE 3
The variation characteristic of the control variables, current and converter voltage.
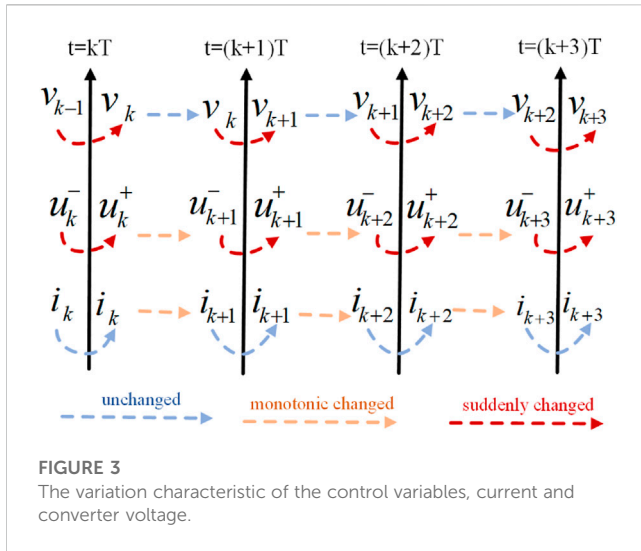
power, the coordinately restricting climbing of active and reactive power can be set as

$$v_d(j)^2 + v_q(j)^2 \leq v_{max}^2, \; j = k, ..., (k+2) \tag{16}$$

Compared to separately restricting each control input such as $v_d^{\,min} \leq v_d(j) \leq v_d^{\,max}$, $v_q^{\,min} \leq v_q(j) \leq v_q^{\,max}$, coordinately restricting v_d and v_q as (16) makes the climbing range of active and reactive power larger.

In order to protect the grid-side converter and dignify the output voltage waveform, the amplitude coordinated restriction and climbing coordinated restriction between grid-side converter voltages u_d and u_q are structed in this sub-section. The variation characteristic of the converter voltage u_d and u_q can be analyzed as follow and is shown as Figure 3.

(1) The variation characteristic of the control variables v_d and v_q

As mentioned above, The MPC algorithm is not executed continuously, but at regular control period T. At the starting point of each control period T, the MPC controller solves the MPC algorithm and outputs a new round of control variables v_d and v_q. Then, the outputs of the MPC controller will remain unchanged until the next control period.

Therefore, during a control period T, the control variables v_d and v_q remain unchanged. At the moments of two control period junctures, the MPC controller will give a new round of control variables v_d and v_q, so that the control variables v_d and v_q will suddenly change at these moments. The variation characteristic of the control variables v_d and v_q can be seen in Figure 3.

(2) The variation characteristic of the current i_d and i_q

According to mathematical model (6), it can be seen that the change rates of the current i_d and i_q are the control variables v_d and v_q, respectively. During a control period T, the control variables v_d and v_q remain unchanged. Therefore, the current i_d and i_q will monotonically increase or decrease. The variation characteristic of

the current i_d and i_q can be seen in Figure 3. The variation characteristics of the power P and Q are consistent with those of the current i_d and i_q.

(3) The variation characteristic of the converter voltage u_d and u_q.

Based on inverse system model (5), the derivative of the converter voltage u_d and u_q can be calculated as $du_d/dt = Rv_d - \omega Lv_q$ and $du_q/dt = Rv_q + \omega Lv_d$. It can be seen that the derivatives of the converter voltage u_d and u_q are constant, so they vary monotonically during a control period T.

At the moments of two control period junctures (such as t = $kT...$t=$(k+3)T$), the MPC controller will give a new round of control variables v_d; v_q. According to mathematical model (5), the converter voltage u_d and u_q will suddenly change at these moments due to the sudden change of the control variables v_d and v_q. The variation characteristic of the converter voltage u_d and u_q can be seen in Figure 3.

Based on the inverse system model (5), the future converter voltage $u(j)^+$, $j = k...k+2$ in Figure 3 can be calculated as

$$\begin{aligned} u(k)^+ &= Lv(k) + R_j x(k) + e_j \\ u(k+1)^+ &== Lv(k+1) + R_j x(k+1) + e_j \\ u(k+2)^+ &= Lv(k+2) + R_j x(k+2) + e_j \\ u(k)^+ &= \begin{bmatrix} u_d(k)^+ \\ u_q(k)^+ \end{bmatrix}, \quad R_j = \begin{pmatrix} R, -L\omega_0 \\ L\omega_0, R \end{pmatrix}, \quad e_j = \begin{bmatrix} e \\ 0 \end{bmatrix} \end{aligned} \tag{17}$$

The prediction model (17) of converter voltage can be rewritten in matrix form.

$$\begin{aligned} U^+ &= LV + R_J W + e_J \\ U^+ &= [u(k)^+; u(k+1)^+; u(k+2)^+] \\ W &= [x(k); x(k+1); x(k+2)] \\ V &= [v(k); v(k+1); v(k+2)] \\ e_J &= [e_j; e_j; e_j] \\ R_J &= diag(R_j, R_j, R_j) \end{aligned} \tag{18}$$

Based on the inverse system model (5), the future converter voltage $u(j)^-$, $j = k+1...k+3$ in Figure 3 can be calculated as

$$\begin{aligned} u(k+1)^- &== Lv(k) + R_j x(k+1) + e_j \\ u(k+2)^- &= Lv(k+1) + R_j x(k+2) + e_j \\ u(k+3)^- &= Lv(k+2) + R_j x(k+3) + e_j \\ u(k)^- &= \begin{bmatrix} u_d(k)^- \\ u_q(k)^- \end{bmatrix}, \; R_j = \begin{pmatrix} R, -L\omega_0 \\ L\omega_0, \; R \end{pmatrix}, \; e_j = \begin{bmatrix} e \\ 0 \end{bmatrix} \end{aligned} \tag{19}$$

The prediction model (19) of converter voltage can be rewritten in matrix form.

$$\begin{aligned} U^- &= LV + R_J X + e_J \\ U^- &= [u(k+1)^-; u(k+2)^-; u(k+3)^-] \\ X &= [x(k+1); x(k+2); x(k+3)] \\ V &= [v(k); v(k+1); v(k+2)] \\ e_J &= [e_j; e_j; e_j] \\ R_J &= diag(R_j, R_j, R_j) \end{aligned} \tag{20}$$

The future voltage amplitudes of the grid-side converter should not be greater than its rated voltage amplitude. Then, we have.

$$\begin{aligned} |u(j)^+|^2 &= |u_d(j)^+|^2 + |u_q(j)^+|^2 \leq u_N^2, \quad j = k, ..., k+2 \\ |u(j)^-|^2 &= |u_d(j)^-|^2 + |u_q(j)^-|^2 \leq u_N^2, \quad j = k+1, ..., k+3 \end{aligned} \tag{21}$$

As shown in Figure 3, at the moments of two control period T junctures (such as $t = kT \ldots t=(k+3)T$), the output converter voltage $u = [u_d, u_q]^T$ will suddenly change due to the suddenly change of control input $v = [v_d, v_q]^T$. The climbing of the converter voltage depends on the change of the control input which is denoted by $\Delta v = [\Delta v_d, \Delta v_q]^T$. Δv_d and Δv_d are coordinated restricted in order to avoid excessive transient climbing of converter voltage.

$$
\begin{aligned}
&|\Delta v(j)|^2 = \Delta v_d(j)^2 + \Delta v_q(j)^2 \leq \Delta v^2_{max}, \; j = k, \ldots, (k+2) \\
&\Delta v(k) = v(k) - v(k-1) \\
&\Delta v(k+1) = v(k+1) - v(k) \\
&\Delta v(k+2) = v(k+2) - v(k+1)
\end{aligned} \tag{22}
$$

The square of the current amplitude is $|i|^2 = i_d^2 + i_q^2$ and the square of the converter voltage amplitude is $|u|^2 = u_d^2 + u_q^2$. Based on $di_d/dt = v_d$ and $di_q/dt = v_q$, the second derivative of $|i|^2$ can be calculated as $d^2|i|^2/dt^2 = i_d v_d + 2i_q v_q$. Similarly, based on $du_d/dt = Rv_d - \omega L v_q$ and $du_q/dt = Rv_q + \omega L v_d$, the second derivative of $|u|^2$ can be calculated as $d^2|u|^2/dt^2 = 2(Rv_d - \omega L v_q)^2 + 2(Rv_q + \omega L v_d)^2$.

It can be seen that the second derivatives of both $|i|^2$ and $|u|^2$ are greater than or equal to 0. This means that during a control period T, the curves of both $|i|^2$ and $|u|^2$ are concave downward. The maximum values of $|i|^2$ and $|u|^2$ during a control period T will appear at the beginning and end instants of the control period. The square of the apparent power amplitude $|S|^2$ and the square of the current amplitude $|i|^2$ have the same characteristics. Therefore, it is only necessary to restrict the amplitude square of converter voltage, current and power at the beginning and end instants of each control period T such as constraint conditions (12), (15), (21). Even without the above reasons, since the MPC controller is discrete, it can only carry out discrete constraints on voltage, current and power.

4.3 Objective function

The optimization objective is to minimize the deviations between the output values and reference values of the grid-side converter composite system's active and reactive power.

$$
\begin{aligned}
&\min J = \Delta Y^T R \Delta Y \\
&\Delta Y = Y_{ref} - Y \\
&Y = [y(k+1); y(k+2); y(k+3)] \\
&Y_{ref} = \left[y(k+1)_{ref}; y(k+2)_{ref}; y(k+3)_{ref} \right] \\
&y(k)_{ref} = \left[P(k)_{ref}, Q(k)_{ref} \right]^T, \; y(k) = [P(k), Q(k)]^T \\
&r = \begin{bmatrix} r_P, 0 \\ 0, r_Q \end{bmatrix}, \quad R = diag(r, r, r)
\end{aligned} \tag{23}
$$

When the apparent power reference of the grid-side converter composite system lesser than its rated apparent power, the output values of active and reactive power can reach their reference values. If the apparent power reference is greater than rated apparent power, the priority of active and reactive power must be set and the party with lower priority cannot reach its reference.

This paper sets the priority of active and reactive power by setting the deviation coefficients r_P and r_Q. When r_P is greater than r_Q, it means that the active priority is higher. On the contrary, when r_P is less than r_Q, reactive power priority is higher. The party with higher priority can achieve greater amplitude and climbing.

According to the National standard of China "Technical Regulations for Wind Farm Access to Electric Power System, GB/T 19963.1-2021", when the grid voltage change is in the domain of $e < 0.9e_N$ or $e > 1.1e_N$, the wind farm should provide reactive power support for the grid voltage recovery, and when the grid voltage change is in the domain of $0.9e_N < e < 1.1e_N$, the wind farm should withdraw reactive power support.

Therefore, in the proposed control strategy, according to the situation of the power grid and the PMSG wind turbine, the deviation coefficients r_P and r_Q are automatically adjusted so that the proposed control strategy can adapt to different grid support scenarios. For example, when the grid voltage change is in the domain of $0.9e_N < e < 1.1e_N$, the active power priority is set higher ($r_P > r_Q$), so that the PMSG wind turbine can carry out maximum wind energy capture and frequency support. When the grid voltage change is in the domain of $e < 0.9e_N$ or $e > 1.1e_N$, the priority task of the PMSG wind turbine is to provide reactive power support for the grid voltage recovery, and of cause the reactive power priority is set higher ($r_P < r_Q$). The priority of active and reactive power can also be set according to other conditions, such as grid frequency or rotor speeds of the PMSG wind turbine.

4.4 Stability analysis

The optimization problem of the MPC controller can be sort out as follows

$$
\begin{aligned}
&\min J(\{V\}) = \Delta Y^T R \Delta Y \\
&s.t. \quad (8)-(23) \begin{cases} c_i(\{V\}) \leq 0, \; i = 1, \cdots, m \\ h_j(\{V\}) = 0, \; j = 1, \cdots, n \end{cases}
\end{aligned} \tag{24}
$$

where $c_i(\{V\})$ and $h_j(\{V\})$ represent the inequality constraints and equality constraints in (8–23). The optimization variables $\{V\}$ of the optimization problem (24) are $V = [v(k); \cdots; \quad v(k+N_C)]$.

4.4.1 The feasibility of the optimization process

To avoid the feasible set determined by the constraints becoming empty, the optimization problem (24) can be modified as follows.

$$
\begin{aligned}
&\min J(\{V, \varepsilon\}) = \Delta Y^T R \Delta Y + \sum_{i=1}^{m} \rho_i |\varepsilon_i|^2 \\
&s.t. \quad \begin{cases} c_i(\{V, \varepsilon\}) \leq \varepsilon_i, \; i = 1, \cdots, m \\ \varepsilon_i \geq 0, \; i = 1, \cdots, m \\ h_j(\{V\}) = 0, \; j = 1, \cdots, n \end{cases}
\end{aligned} \tag{25}
$$

The optimization variables $\{V, \varepsilon\}$ of the optimization problem (25) include original optimization variables $V = [v(k); \cdots; \quad v(k+N_C)]$ and slack optimization variables $\varepsilon = [\varepsilon_1; \cdots; \quad \varepsilon_m]$. The values of the slack optimization variables $\varepsilon = [\varepsilon_1; \cdots; \quad \varepsilon_m]$ are flexible, thus any values of the original optimization variables $V = [v(k); \cdots; \quad v(k+N_C)]$ can satisfy the constraints. In other word, the feasible set determined by the constraints is always non-empty so that the optimization problems (25) always exists the optimal solution during the optimization process.

The slack optimization variables $\varepsilon = [\varepsilon_1; \cdots; \quad \varepsilon_m]$ are contained in the objective function, so that all the slack optimization variables $\varepsilon = [\varepsilon_1; \cdots; \quad \varepsilon_m]$ will eventually approach 0 during the

optimization process. In the end, the optimization problem (25) is the same as the optimization problem (24).

4.4.2 The stability analysis of the control system.

The control system including the MPC controller and the grid-side converter composite system. Stability analysis is to construct an energy function for the control system and determine whether the energy function meets the stability conditions. The energy function chosen here is $E_F(x(k)) = y(k)^T r\ \ y(k)$, where $y(k) = C_d x(k)$. Next, we will prove that it satisfies the stability conditions.

1) $E_F(x(k))$ is a positive definite function.

If $x(k) \neq \vec{0}$, we have $y(k) = C_d x(k) \neq \vec{0}$, then $E_F(x(k)) = y(k)^T r\ \ y(k) > 0$.
If $x(k) = \vec{0}$, we have $y(k) = C_d x(k) = \vec{0}$, then $E_F(x(k)) = y(k)^T r\ \ y(k) = 0$.
Therefore, $E_F(x(k))$ is a positive definite function.

2) when $|x(k)| \to \infty$, the $E_F(x(k)) \to \infty$.

Firstly, $E_F = y(k)^T r\ \ y(k) = \alpha |y(k)|^2$, α is a constant.
If $|x(k)| \to \infty$, we have $|y(k)| = |C_d x(k)| \to \infty$, then $E_F(x(k)) \to \infty$.

3) If $x(k) \neq \vec{0}$, $E_F(x(k))$ is going to go down over time.

When judging the change trend of the energy function $E_F(x(k))$, the system input vector is 0 vector ($Y_{ref} = \vec{0}$) but the initial system state vector is not 0 vector ($x(k) = [i_d(k), i_q(k)]^T \neq \vec{0}$). Given that $\Delta Y = Y_{ref} - Y$ and $Y_{ref} = \vec{0}$, the optimization problem (25) can be rewritten as:

$$\min J(\{V, \varepsilon\}) = Y^T RY + \sum_{i=1}^{m} \rho_i |\varepsilon_i|^2 = \sum_{l=(k+1)}^{k+N_P} y(l)^T r\ \ y(l) + \sum_{i=1}^{m} \rho_i |\varepsilon_i|^2$$

$$s.t. \begin{cases} c_i(\{V, \varepsilon\}) \le \varepsilon_i, i = 1, \cdots, m \\ \varepsilon_i \ge 0, i = 1, \cdots, m \\ h_j(\{V\}) = 0, j = 1, \cdots, n \end{cases}$$

(26)

When initial state variables are $x(k) = [i_d(k), i_q(k)]^T \neq \vec{0}$, the energy function is $E_F(x(k)) = y(k)^T r\ \ y(k)$. Solving the optimization problem (26), the optimal variables obtained are denoted by $V^k = [v(k); \cdots;\ \ v(k + N_C)]$, $\varepsilon^k = [\varepsilon_1^k; \cdots;\ \ \varepsilon_m^k]$, and the optimal state sequence is denoted by $X^k = [x(k + 1); \cdots;\ \ x(k + N_P)]$, and the optimal output sequence is denoted by $Y^k = [y(k + 1); \cdots;\ \ y(k + N_P)]$. Then, the objective function is $J(\{V, \varepsilon\})^k = \sum_{l=(k+1)}^{k+N_P} y(l)^T r\ \ y(l) + \sum_{i=1}^{m} \rho_i |\varepsilon_i^k|^2$. By analyzing the optimization problem (26), the optimization results have the following two hypothesis.

If the value taking of the original optimization variables $V^k = [v(k); \cdots;\ \ v(k + N_C)]$ makes the state variables and outputs constantly approach 0 (that is $|x(k)| > |x(k + 1)| \ge \cdots \ge |x(k + N_P)| \ge 0$ and $|y(k)| > |y(k + 1)| \ge \cdots \ge |y(k + N_P)| \ge 0$), the value taking of the slack optimization variables $\varepsilon^k = [\varepsilon_1^k; \cdots;\ \ \varepsilon_m^k]$ can be smaller or remain 0. so that the result of

the objective function $J(\{V, \varepsilon\})^k = \sum_{l=(k+1)}^{k+N_P} y(l)^T r\ \ y(l) + \sum_{i=1}^{m} \rho_i |\varepsilon_i^k|^2$ will be smaller.

On the contrary, if the value taking of the original optimization variables $V^k = [v(k); \cdots;\ \ v(k + N_C)]$ make the state variables and outputs constantly away from 0 (that is $|x(k)| < |x(k + 1)| < \cdots < |x(k + N_P)|$ and $|y(k)| < |y(k + 1)| < \cdots < |y(k + N_P)|$), the slack optimization variables $\varepsilon^k = [\varepsilon_1^k; \cdots;\ \ \varepsilon_m^k]$ will be forced to take larger values to ensure that the inequality constraint can be satisfied. In this way, the result of the objective function $J(\{V, \varepsilon\})^k = \sum_{l=(k+1)}^{k+N_P} y(l)^T r\ \ y(l) + \sum_{i=1}^{m} \rho_i |\varepsilon_i|^2$ will be larger.

Given that the optimal variables obtained $V^k = [v(k); \cdots;\ \ v(k + N_C)]$, $\varepsilon^k = [\varepsilon_1^k; \cdots;\ \ \varepsilon_m^k]$ always minimize the objective function $J(\{V, \varepsilon\}) = \sum_{l=(k+1)}^{k+N_P} y(l)^T r\ \ y(l) + \sum_{i=1}^{m} \rho_i |\varepsilon_i|^2$. Therefore, the optimization result will satisfy the first hypothesis. That is to say, the optimal state variables and optimal outputs obtained from solving the optimization problem (26) will constantly approach 0 ($|x(k)| > |x(k + 1)| \ge \cdots \ge |x(k + N_P)| \ge 0$ and $|y(k)| > |y(k + 1)| \ge \cdots \ge |y(k + N_P)| \ge 0$).

Under the control of the MPC controller, in the next sampling interval, the system state variables changes to $x(k + 1) = [i_d(k + 1), i_q(k + 1)]^T$, the energy function is $E_F(x(k + 1)) = y(k + 1)^T r\ \ y(k + 1)$. Given that $|y(k)| > |y(k + 1)|$, we have $E_F(x(k)) > E_F(x(k + 1))$. The energy function $E_F(x(k))$ declines over time.

If $x(k + 1) \neq \vec{0}$, solving the optimization problem (26), the optimal variables obtained are denoted by $V^{k+1} = [v(k + 1); \cdots;\ \ v(k + 1 + N_C)]$, $\varepsilon^{k+1} = [\varepsilon_1^{k+1}; \cdots;\ \ \varepsilon_m^{k+1}]$, and the optimal state sequence is denoted by $X^{k+1} = [x(k + 2); \cdots;\ \ x(k + 1 + N_P)]$, and the optimal output sequence is denoted by $Y^{k+1} = [y(k + 2); \cdots;\ \ y(k + 1 + N_P)]$. Similarly, the optimization result will satisfy that $|x(k + 1)| > |x(k + 2)| \ge \cdots \ge |x(k + 1 + N_P)| \ge 0$ and $|y(k + 1)| > |y(k + 2)| \ge \cdots \ge |y(k + 1 + N_P)| \ge 0$). Under the control of the MPC controller, the system state variables will changes to $x(k + 2)$ in the next sampling interval, and we have $E_F(x(k + 1)) > E_F(x(k + 2))$.

By analogy, if the state vector is not 0 vector, the energy function $E_F(x(k))$ is going to go down over time under the control of the MPC controller.

4.5 Optimization solving program

The optimization problem (25) is a convex optimization problem with a convex quadratic objective function and multiple convex quadratic constraints which can be solved in many ways. Specifically, a convex optimization problem can be sort out as the following general form.

$$\min f(\{x_k\})$$
$$s.t. \quad g_i(\{x_k\}) \le 0, i = 1, \cdots, m \qquad (27)$$
$$h_j(\{x_k\}) = 0, j = 1, \cdots, n$$

where $\{x_k\}$ is the set of optimization variables, and $f(\{x_k\})$ is the objective function of the convex optimization problem. $g_i(\{x_k\})$ and $h_j(\{x_k\})$ represent inequality constraints and equality constraints of

FIGURE 4
The flow of the MPC algorithm.

FIGURE 5
The active and reactive power step references of the PMSG wind turbine.
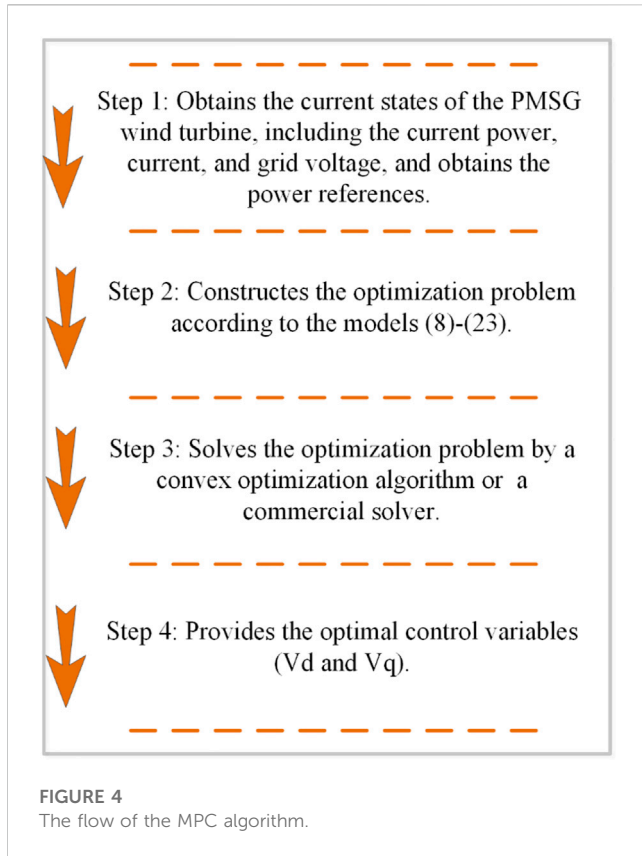
the convex optimization problem, respectively. The Lagrange function is defined as:

$$L\left(\{x_k\}, \{\lambda_i\}, \{\mu_j\}\right) = f\left(\{x_k\}\right) + \sum_{i=1}^{m} \lambda_i g_i\left(\{x_k\}\right) + \sum_{j=1}^{n} \mu_j h_j\left(\{x_k\}\right) \quad (28)$$

where $\{\lambda_i\}$ and $\{\mu_j\}$ are called Lagrange multipliers. The optimal solution of convex optimization problem (27) can be obtained by solving the KKT conditions (Xu et al., 2001). as follows.

$$\frac{\partial L}{\partial x_k} = 0, \ k = 1, \cdots, K$$

$$\frac{\partial L}{\partial \mu_j} = 0, \ j = 1, \cdots, J$$

$$\lambda_i g_i\left(\{x_k\}\right) = 0, \ i = 1, \cdots, m \qquad (29)$$

$$\lambda_i \geq 0, \ i = 1, \cdots, m$$

$$g_i\left(\{x_k\}\right) \leq 0, \ i = 1, \cdots, m$$

In (29), equality constraints are used to solve the optimal solution, and inequality constraints are used to verify the optimal solution. It is worth mentioning that in addition to the above method, there are many solving methods (An, 2000) and mature commercial solvers for convex optimization problems. The commercial solvers (Gurobi optimization, 2023; IBM CPLEX Optimizer, 2023) can be used to solve convex optimization problems conveniently and quickly, without the need for users to write optimization programs.

In each sampling interval T, the actions of the MPC controller include measuring the system states, constructing and solving the
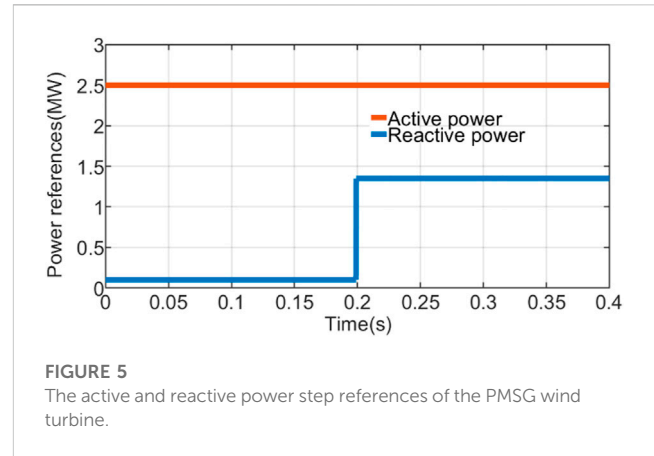
optimization problem (8–23), and providing the optimal control variables (V_d and V_q) to the system. The flow of the MPC algorithm is shown in Figure 4.

5 Case study

5.1 Simulation description

The performance of the proposed active and reactive power coordinated control strategy is evaluated in this section by testing the tracking performance of the PMSG wind turbine to the active and reactive power step references. The references of active and reactive power are shown in Figure 5. In order to capture maximum wind energy, the active power reference remains at 2.5 MW unchanged. The grid voltage stays in the rated value at 0–0.2s and the reactive power reference is set as 0.1 MW at 0–0.2 s. While the grid voltage drops down to 50% of the rated value at 0.2 s and the duration is 0.6 s. According to the reactive current support requirements in grid codes [21], the reactive power reference is stepped from 0.1 MW to 1.35 MW at 0.2 s.

Three different scenarios are simulated in this section using MATLAB/Simulink.

S1: Active and reactive power are controlled using the MPC controller without coordinated restrictions and priority arrangement
S2: Active and reactive power are coordinately controlled using the MPC controller, which considers amplitude coordinated restrictions and priority arrangement of active and reactive power.
S3: Active and reactive power are coordinately controlled using the MPC controller, which considers amplitude coordinated restrictions and priority arrangement of active and reactive power. Meanwhile, the climbing coordinated restrictions of active and reactive power is considered.
S4: Active and reactive power are coordinately controlled using the PI controller, which considers amplitude coordinated restrictions and priority arrangement of active and reactive power.

TABLE 1 Parameters of the grid-side converter system.

Parameters	Description	Value
R (Ω)	The resistor and inductor of the filter	0.027
L (mH)	The inductor of the filter	1.65
e_N (V)	Rated phase voltage of the power grid	1732
S_N (MW)	Rated apparent power of the grid side converter	3
i_N (A)	Rated current of the grid side converter	577.35
u_N (V)	Rated phase voltage of the grid side converter	1803
T (ms)	Control period of the MPC controller	10

The other conditions and parameters in the three simulated scenarios are the same. The parameters of the grid-side converter system are indicated in Table 1 (Yaramasu and Wu, 2014). The prediction horizon is set as five times control period T and the control horizon is set as four times control period T when simulating.

5.2 Results and analysis

Figures 6–8 depict the power reference tracking results of the PMSG wind turbine in the three different scenarios. Due to the grid voltage droping down to 50% of the rated value at 0.2 s, the maximum apparent power of the PMSG wind turbine change from 3MW to 1.5 MW at 0.2 s. In Scenario S1, the active and reactive power coordinated restrictions and priority decisions are not considered. Regardless of the power grid status, the amplitude range and climbing range of active and reactive power are fixed and do not interfere with each other. Therefore, the PMSG wind turbine cannot guarantee reactive power output by reducing the active power output. Therefore, when the grid voltage drop, the reactive power of the PMSG wind turbine cannot meet the 1.35 MW reference which is the reactive power support requirement.

Amplitude coordinated restrictions and priority decisions of active and reactive power are considered in Scenario S2. When the grid voltage drops down to 50% of the rated value at 0.2s, the priority task of the PMSG wind turbine is to provide reactive power support for the grid voltage recovery, and of cause the reactive power priority is set higher. As shown in Figure 6, by reducing the amplitude of active power, the PMSG wind turbine can output 1.35 MW reactive power to meet the reactive power support requirements of the grid. That is why we say it is very necessary and meaningful to coordinately control the active and reactive power amplitude of the wind turbine under the background of grid support.

The proposed control strategy in this paper is simulated in Scenario 3 which considers amplitude and climbing coordinated restrictions and priority decisions between active and reactive power. Similarly to Scenario 2, the PMSG wind turbine can meet the reactive power support requirements of the grid when the grid voltage drops. The difference between Scenario 2 and Scenario 3 is that, the PMSG wind turbine shows a faster response speed in Scenario 3, which the dynamic time in Scenario 3 is 20 ms and it is 40 ms in Scenario 2. This demonstrates the superiority of the control strategy proposed in this paper.

Figures 9, 10 show the current climbing of the PMSG wind turbine in Scenario 2 and Scenario 3. It can be found that the active and reactive current climbing limit value of the PMSG wind turbine in Scenario 3 is larger than that in Scenario 2. The greater ability to climb allows the PMSG wind turbine to show faster response speed in response to a step change in the power reference value. The results explain why the PMSG wind turbine shows a faster response speed in Scenario 3. The simulation results verify that under the given apparent power/current climbing constraint, active and reactive power/current climbing coordinated restrictions allow PMSG wind turbines to perform a larger range of active and reactive power/current climbing.

The voltage of the grid-side converter is shown in Figure 11 when the PMSG wind turbine executes the proposed control strategy
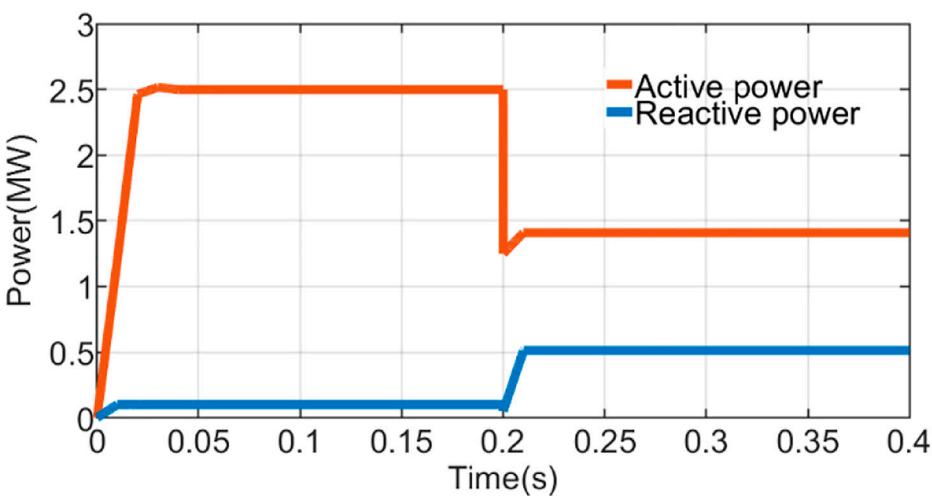

FIGURE 6
The power references tracking results of the PMSG wind turbine in the S1 scenario.
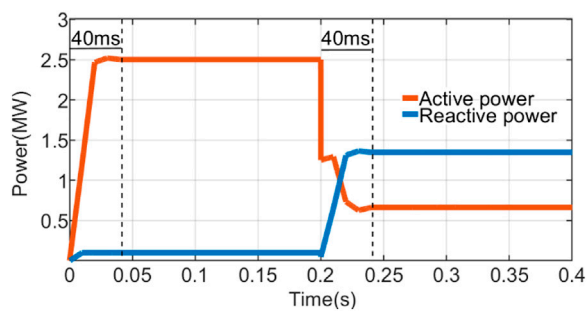
FIGURE 7
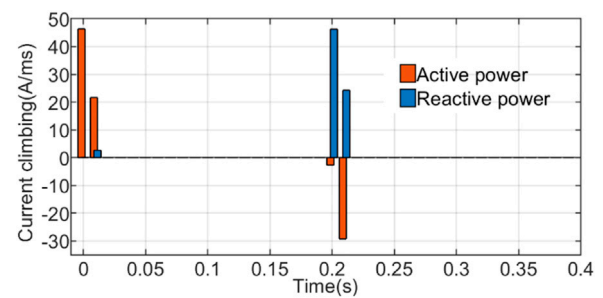The power references tracking results of the PMSG wind turbine in the S2 scenario.

FIGURE 10
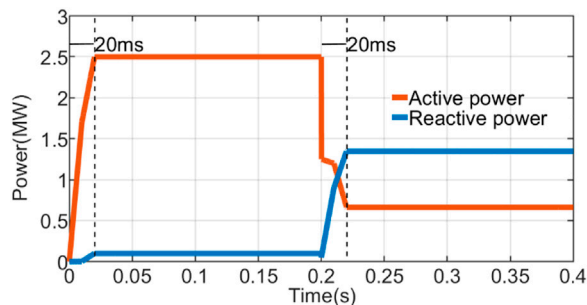The power climbing of the PMSG wind turbine in Scenario 3.

FIGURE 8
The power references tracking results of the PMSG wind turbine in the S3 scenario.
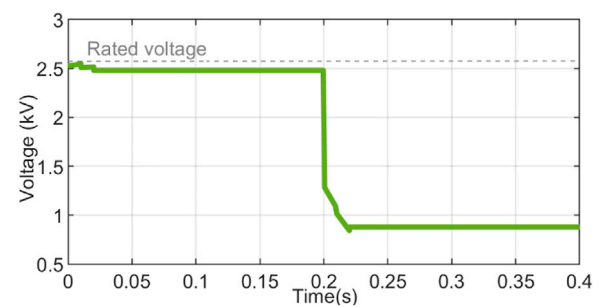

FIGURE 11
The voltage of the grid-side converter in Scenario 3.
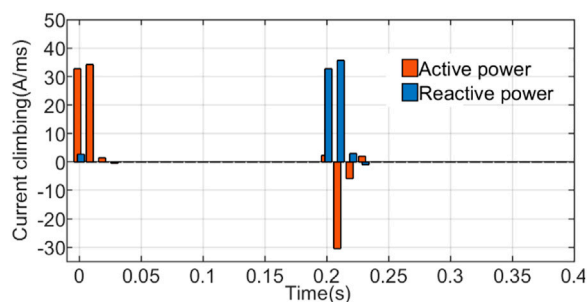
FIGURE 9
The power climbing of the PMSG wind turbine in Scenario 2.
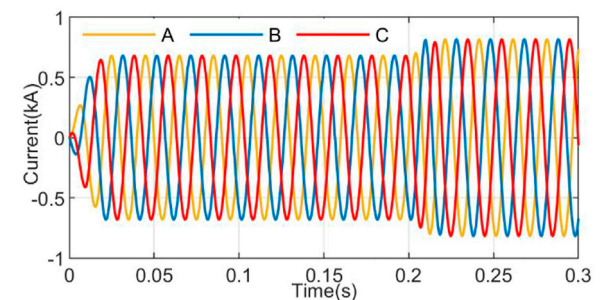
FIGURE 12
The three-phase current of the grid-side converter in Scenario 3.

in Scenario 3. It is easy to see that the voltage amplitude does not exceed its maximum allowed value in the whole process. In addition, there are no sharp climb and spike in the voltage waveform, which can protect the grid-side converters and improve the electricity quality of the PMSG wind turbine. Hence, The voltage amplitude and climbing constraint model proposed in this paper is effective.

The three-phase current of the grid-side converter is shown in Figure 12 when the PMSG wind turbine executes the proposed

control strategy in Scenario 3. The rated peak value of single-phase current is 0.816 kA. It is easy to see that the amplitudes of the three-phase current do not exceed the rated peak value 0.816 kA in the whole process. In addition, there are no sharp climb and spike in the three-phase current waveforms, which can protect the grid-side converters and improve the electricity quality of the PMSG wind turbine. Hence, The current amplitude and climbing constraint model proposed in this paper is effective.

Figure 13 shows the result of the PMSG wind turbine tracking a 3 MW step active power reference, which is simulated with different

FIGURE 13
The active power references tracking results of the PMSG wind turbine under different prediction horizon.

FIGURE 14
The power references tracking results of the PMSG wind turbine in the S4 scenario.

TABLE 2 The steady-state amplitude of active and reactive power with different deviation coefficients.
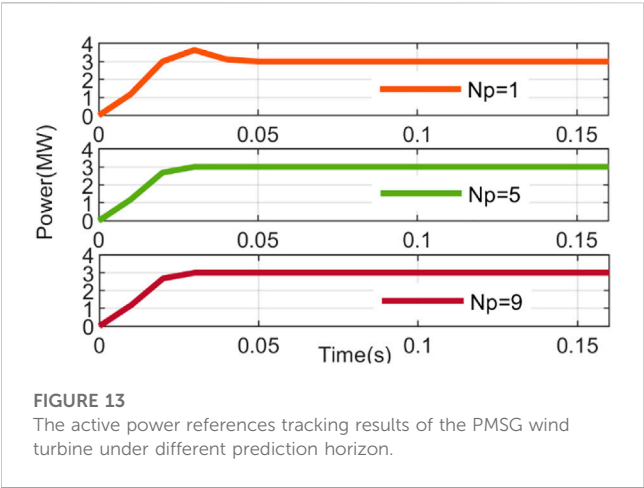
Deviation coefficients	Power references	Power outputs
$r_P = 1$ $r_Q = 100000$	Pref = 2.5 MW	p = 0.65 MW
	Qref = 1.35 MW	Q = 1.35 MW
$r_P = 100000$ $r_Q = 1$	Pref = 2.5 MW	p = 1.5 MW
	Qref = 1.35 MW	Q = 0 MW
$r_P = 1$ $r_Q = 1$	Pref = 2.5 MW	p = 1.32 MW
	Qref = 1.35 MW	Q = 0.713 MW
$r_P = 1$ $r_Q = 10$	Pref = 2.5 MW	p = 0.95 MW
	Qref = 1.35 MW	Q = 1.161 MW

FIGURE 15
The voltage of the grid-side converter in Scenario 4.

prediction horizons. As can be seen from Figure 13, when the prediction horizon is short (Np = 1), the output of the PMSG wind turbine exists the overshoot. This is because the MPC model in this paper contains constraints on the power climbing change. The active power output of the PMSG wind turbine has a large climbing in the early stage, and the climbing cannot be changed quickly after approaching the active power reference value, so overdrive occurs. When the prediction horizon is long, the MPC controller can predict the corresponding future multi-step power output under different control variables, and select the optimal control variable to avoid overshoot. It is worth mentioning that the longer the prediction horizon, the better the control effect. In the scenario, the control effect is exactly the same when Np = 5 as when Np = 9. It is worth mentioning that as the prediction horizon increases, the control effect will not get better. In the current scenario, the control effect is exactly the same when Np = 5 and Np = 9.

In the proposed MPC based control strategy, the priority of active and reactive power can be flexibly arranged by setting the deviation coefficients r_P and r_Q in the objective function. The steady-state amplitude of active and reactive power at 0.2–0.4s in Scenario 3 are displayed in Table 2, which is simulated with different deviation coefficients r_P and r_Q. Due to the grid voltage drops of
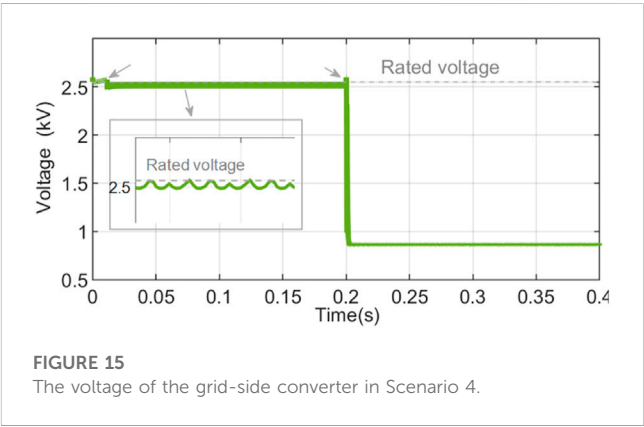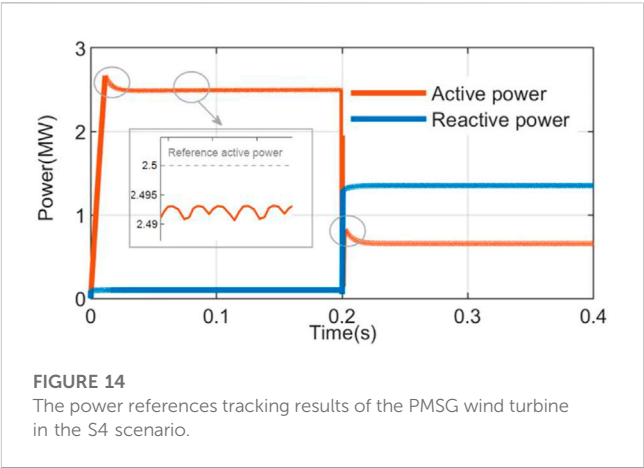
50% at 0.2 s, the maximum apparent power of the PMSG wind turbine changes from 3MW to 1.5 MW. Under the given 1.5 MW apparent power constraint, different priority settings will result in different active and reactive power outputs. So that the PMSG wind turbine can operate in different modes such as giving priority to the active power support/output, giving priority to the reactive power support/output, and giving the same/close priority to the active power support and the reactive power support. The MPC controller can flexibly set the priority of active and reactive power, so that the proposed control strategy can be flexibly applicated in different grid support scenarios.

Figures 14, 15 depict the power reference tracking results and the grid-side converter voltage of the PMSG wind turbine when the PMSG wind turbine is controlled by a PI controller in Scenario 4. It can be seen in Figure 14 that there are overshoots and steady-state errors existing when the PMSG wind turbine tracks the active power references. It is worth mentioning that increasing the integral parameter in the PI controller can eliminate steady-state errors, but can lead to larger overshoots. As can be seen in Figure 15, the grid-side converter voltage may exceed its rated value during the dynamic process, which may cause damage to the grid-side converter and degrade the power quality. In addition, the amplitudes of output power and the grid-side converter voltage continue to fluctuate during the steady state period. This means that the power quality of the PMSG wind turbine output is lower.

As shown in Figures 8–13, the MPC based active and reactive power coordinated control strategy proposed in this paper can enable one of active power and reactive power (the one with higher priority) to track its references perfectly with zero steady-state error, zero overshoot and transient dynamic process. The steady-state power curve and the steady-state voltage curve are smooth straight lines. The power amplitude constraint, power climbing constraint, power priority arrangement, voltage amplitude constraint and voltage ramp constraint set in the proposed control strategy have achieved the expected results. That is because the MPC controller is good at dealing with multi-input and multi-output control problems with complex constraints and specific objectives. The above complex constraints and specific objectives are difficult to implement by PI controllers.

6 Conclusion

In this paper, an MPC based active and reactive power coordinated control strategy is proposed to enhance the power control performance and grid support capability of PMSG wind turbines. Firstly, the constraint conditions in the MPC controller are constructed which include amplitude and climbing coordinated restrictions between active and reactive power, voltage amplitude and climbing constraints of grid-side converters. These constraint conditions enable PMSG wind turbines to perform a wider range of active and reactive power amplitude and climbing, and a smoother output voltage waveform. Then, an objective function is constructed to minimize the deviation of active and reactive power from their reference values. By setting the deviation coefficients in the objective function, the priority of active power output and reactive power output can be flexibly arranged. So that PMSG wind turbines can meet the power output requirements in different grid support scenarios. In the end, the simulation results show that the proposed control strategy can make PMSG wind turbines achieve excellent power control performance and thus better meet the requirements of power grid support. In the future, the influences of sampling delay and MPC calculation delay on the proposed control strategy will be analyzed and the proposed control strategy will be improved by using delay compensation technology.

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

XM proposed the idea and carried out the model establishment, simulation experiment, result analysis, meanwhile he wrote the first draft. JY assisted in establishment of model and improvement of the idea. PY was in charge of knowledge coaching and error correction. PW is responsible for data collection, analysis, and processing and assisted in result analysis. PZ helped the revision of the manuscript and typeset the manuscript.

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

An, L. T. H. (2000). An efficient algorithm for globally minimizing a quadratic function under convex quadratic constraints. *Math. Program. Ser. B* 87, 401–426. doi:10.1007/s101070050003

Dong, S., Wang, Y., and Li, H. (2012). Coordinated control for active and reactive power of PMSG-based wind turbine to enhance the LVRT capability. *2012 15th int. Conf. Electr. Mach. Syst., (ICEMS)*, 33–36.

Feltes, C., Engelhardt, S., Kretschmann, J., Fortmann, J., Koch, F., and Erlich, I. (2009). Comparison of the grid support capability of DFIG-based wind farms and conventional power plants with synchronous generators. *2009 IEEE power energy soc. Gen. Meet.*, 1–7. doi:10.1109/PES.2009.5275441

Gaied, H., Naoui, M., Kraiem, H., Goud, B. S., Flah, A., Alghaythi, M. L., et al. (2022). Comparative analysis of MPPT techniques for enhancing a wind energy conversion system. *Front. Energy Res.* 10, 1–15. doi:10.3389/fenrg.2022.975134

Guo, L., Ren, Y., Wang, Z., Zhu, X., Wang, X., Li, X., et al. (2021). Double-layer feedback control method for synchronized frequency regulation of PMSG-based wind farm. *IEEE Trans. Sustain. Energy* 12, 2423–2435. doi:10.1109/TSTE.2021.3096724

Gurobi optimization (2023). *Gurobi*. Available at: http://www.gurobi.com (Accessed Feb, 2023).

Hansen, A. D., Sørensen, P., Iov, F., and Blaabjerg, F. (2006). Grid support of a wind farm with active stall wind turbines and AC grid connection. *Wind Energy* 9, 341–359. doi:10.1002/we.176

IBM CPLEX Optimizer (2023). *High-performance mathematical programming solver for linear programming, mixed integer programming, and quadratic programming*. Available at: https://www.ibm.com/cn-zh/analytics/cplex-optimizer (Accessed Feb, 2023).

Khazaei, J., Nguyen, D. H., and Asrari, A. (2020). Consensus-based demand response of PMSG wind turbines with distributed energy storage considering capability curves. *IEEE Trans. Sustain. Energy* 11, 2315–2325. doi:10.1109/TSTE.2019.2954796

Li, Q., Ren, B., Li, Q., Wang, D., Tang, W., Meng, J., et al. (2022). Virtual inertial control strategy based on fuzzy logic algorithm for PMSG wind turbines to enhance frequency stability. *Front. energy Res.* 10, 1–9. doi:10.3389/fenrg.2022.907770

Li, S., Haskew, T. A., Swatloski, R. P., and Gathings, W. (2012). Optimal and direct-current vector control of direct-driven PMSG wind turbines. *IEEE Trans. Power Electron.* 27, 2325–2337. doi:10.1109/TPEL.2011.2174254

Liu, Y., Gracia, J. R., King, T. J., and Liu, Y. (2015). Frequency regulation and oscillation damping contributions of variable-speed wind generators in the U.S. Eastern interconnection (EI). *IEEE Trans. Sustain. Energy* 6, 951–958. doi:10.1109/TSTE.2014.2318591

Maaoui-Ben Hassine, I., Naouar, M. W., and Mrabet-Bellaaj, N. (2016). Predictive control strategies for wind turbine system based on permanent magnet synchronous generator. *ISA Trans.* 62, 73–80. doi:10.1016/j.isatra.2015.12.002

Mayne, D. Q. (2014). Model predictive control: Recent developments and future promise. *Automatica* 50, 2967–2986. doi:10.1016/j.automatica.2014.10.128

Mishra, R., and Saha, T. K. (2020). Performance analysis of model predictive technique based combined control for PMSG-Based distributed generation unit. *IEEE Trans. Ind. Electron.* 67, 8991–9000. doi:10.1109/TIE.2020.2992970

Moghadasi, A., and Sarwat, A. I. (2015). Optimal analysis of resistive superconducting fault current limiters applied to a variable speed wind turbine system. *Conf. Proc. - IEEE S.*, 1–7. doi:10.1109/SECON.2015.7132944

Mohseni, M., and Islam, S. M. (2012). Review of international grid codes for wind power integration: Diversity, technology and a case for global standard. *Renew. Sustain. Energy Rev.* 16, 3876–3890. doi:10.1016/j.rser.2012.03.039

Musarrat, N., Islam, R., Member, S., Muttaqi, K. M., and Member, S. (2019). Enhanced frequency support from a PMSG-based wind energy conversion system integrated with a high temperature SMES in standalone power supply systems. *IEEE Trans. Appl. Supercond.* 29, 1–6. doi:10.1109/tasc.2018.2882429

National Energy Administration (2023). *Transcript of the online press conference of the National Energy Administration in the fourth quarter of 2022.* Available at: http://www.nea.gov.cn/2022-11/14/c_1310676392.htm (Accessed Feb, 2023).

Nguyen, T. H., Lee, D. C., Van, T. L., and Kang, J. H. (2013). Coordinated control of reactive power between STATCOMs and wind farms for PCC voltage regulation. *J. Power Electron.* 13, 909–918. doi:10.6113/JPE.2013.13.5.909

Okedu, K. E. (2022). Improving the performance of PMSG wind turbines during grid fault considering different strategies of fault current limiters. *Front. Energy Res.* 10, 1–12. doi:10.3389/fenrg.2022.909044

Peng, Y., Li, Y., Lee, K. Y., Tan, Y., Cao, Y., Wen, M., et al. (2021). Coordinated strategy of pmsg and cascaded H-bridge STATCOM in dispersed wind farm for suppressing unbalanced grid voltage. *IEEE Trans. Sustain. Energy* 12, 349–359. doi:10.1109/TSTE.2020.2995457

Qin, S. J., and Badgwell, T. A. (2003). A survey of industrial model predictive control technology. *Control Eng. Pract.* 11, 733–764. doi:10.1016/S0967-0661(02)00186-7

Rodriguez, J., Cortes, P., Kennel, R., and Kazmierkowski, M. P. (2009). Model predictive control - a simple and powerful method to control power converters. *IEEE Trans. Ind. Electron.* 56, 41–49. doi:10.1109/IPEMC.2009.5289335

Shehata, E. G. (2017). A comparative study of current control schemes for a direct-driven PMSG wind energy generation system. *Electr. Power Syst. Res.* 143, 197–205. doi:10.1016/j.epsr.2016.10.039

Sheng, Y., Li, C., Jia, H., Liu, B., Li, B., and Coombs, T. A. (2021). Investigation on FRT capability of PMSG-based offshore wind farm using the SFCL. *IEEE Trans. Appl. Supercond.* 31, 2021–2024. doi:10.1109/TASC.2021.3091054

Tan, Y., Member, S., Muttaqi, K. M., Member, S., Ciufo, P., Member, S., et al. (2017). Enhanced frequency response strategy for a PMSG-based wind energy conversion system using ultracapacitor in remote area power supply systems. *IEEE Trans. Ind. Appl.* 53, 549–558. doi:10.1109/tia.2016.2613074

Tripathi, S. M., Tiwari, A. N., and Singh, D. (2019). Low-voltage ride-through enhancement with the ω and T controls of PMSG in a grid-integrated wind generation system. *IET Gener. Transm. Distrib.* 13, 1979–1988. doi:10.1049/iet-gtd.2018.6275

Vazquez, S., Leon, J., Franquelo, L., Rodriguez, J., Young, H., Marquez, A., et al. (2014). Model predictive control. *IEEE Ind. Electron. Mag.* 8, 91–94. doi:10.1007/978-3-319-91707-8_12

Venkat, A. N., Hiskens, I. A., Rawlings, J. B., and Wright, S. J. (2008). Distributed MPC strategies with application to power system automatic generation control. *IEEE Trans. Control Syst. Technol.* 16, 1192–1206. doi:10.1109/TCST.2008.919414

Wu, Z., Gao, D. W., Member, S., Zhang, H., Yan, S., and Wang, X. (2017). Coordinated control strategy of battery energy storage system and PMSG-WTG to enhance system frequency regulation capability. *IEEE Trans. Sustain. Energy* 8, 1330–1343. doi:10.1109/tste.2017.2679716

Xu, Y., Wang, W., and Gao, Z. (2001). The algorithm of sequential KKT equations by nonmonotone search for arbitrary initial point. *Comput. Optim. Appl.* 18, 221–232. doi:10.1023/A:1011281102385

Yan, Z., Chen, M., Zhen, X., Li, X., and Wubin, W. (2016). An experimental system for LVRT of direct-drive PMSG wind generation system. *2016 IEEE 8th int. Power electron. Motion control conf.* Hefei: IPEMC-ECCE Asia 2016, 1452–1456. doi:10.1109/IPEMC.2016.7512505

Yaramasu, V., and Wu, B. (2014). Predictive control of a three-level boost converter and an NPC inverter for high-power PMSG-based medium voltage wind energy conversion systems. *IEEE Trans. Power Electron.* 29, 5308–5322. doi:10.1109/TPEL.2013.2292068

You, R., Barahona, B., Chai, J., and Cutululis, N. A. (2015). Frequency support capability of variable speed wind turbine based on electromagnetic coupler. *Renew. Energy* 74, 681–688. doi:10.1016/j.renene.2014.08.072

Zeng, C., Yang, B., Cao, P., Li, Q., Deng, J., and Tian, S. (2022). Current status, challenges, and trends of maximum power point tracking for PV systems. *Front. Energy Res.* 10, 1–5. doi:10.3389/fenrg.2022.901035

Zhang, B., Hu, W., Hou, P., Soltani, M., and Chen, Z. (2016). "Coordinated power dispatch of a PMSG based wind farm for output power maximizing considering the wake effect and losses," in *2016 IEEE power energy soc. Gen. Meet. (PESGM)*, 1–5. doi:10.1109/PESGM.2016.7741738

Zhong, C., Lv, Y., Zhou, Y., and Li, H. (2021). An equivalent rotor speed compensation control of PMSG-based wind turbines for frequency support in islanded microgrids. *Front. energy Res.* 9, 1–13. doi:10.3389/fenrg.2021.717327

Available power estimation of wind farms based on deep spatio-temporal neural networks

Yu Liu[1], Kunpeng Huang[2], Jincheng Liu[2], Pei Zhang[2] and Zhao Liu[2]*

[1]North China Branch of State Grid Corporation of China, Beijing, China, [2]Beijing Jiaotong University, Beijing, China

With the development of advanced digital infrastructure in new wind power plants in China, the individual wind-turbine level data are available to power operators and can potentially provide more accurate available wind power estimations. In this paper, considering the state of the wind turbine and the loss in the station, a four-layer spatio-temporal neural network is proposed to compute the available power of wind farms. Specifically, the long short-term memory (LSTM) network is built for each wind turbine to extract the time-series correlations in historical data. In addition, the graph convolution network (GCN) is employed to extract the spatial relationship between neighboring wind turbines based on the topology and patterns of historical data. The case studies are performed using actual data from a wind farm in northern China. The study results indicate that the computation error using the proposed model is lower than that using the conventional physics-based methods and is also lower than that using other artificial intelligence methods.

KEYWORDS

wind farm available power, deep spatio-temporal network, long short-term memory network, wind power, artificial intelligence

1 Introduction

In recent years, the installed capacity of wind power has gradually increased, and the proportion of new energy power generation has gradually increased. In July 2022, the Global Wind Energy Council released the "Global Wind Report 2022." This report shows that the new installed capacity of global wind power is 93.6 GW. By the end of 2021, the cumulative installed capacity of global wind power reached 837 GW, a year-on-year increase of 12.4%. However, due to the random, fluctuating, and intermittent characteristics of wind power, it brings significant challenges to real-time dispatching of power grids (Li et al., 2019a). In this work, we focus on the available power estimation of wind farms, which refers to the theoretical power subtracting the power output of wind turbines under losses in wind farms (State Grid Corporation of China, 2018). Accurate available power is important for system operators to determine the optimal dispatch on the wind farm and other types of generators. It could also serve as an important input for many real-time monitoring and control systems.

At present, the theoretical power computation and prediction of wind farms can be divided into physics-based methods, statistics-based methods, artificial intelligence-based methods, and hybrid methods. The physics-based method mainly uses the numerical weather forecast model to calculate the future wind speed. Then, the predicted wind speed is brought into the relevant wind farm power curve (Tascikaraoglu et al., 2014) (usually provided by the wind farm manufacturer) to predict the wind farm power generations. The physical methods include the prototype machine method (Ding et al., 2016), the wind measurement tower extrapolation method (Guo et al., 2019), and the nacelle wind speed method (Jiang et al., 2014). When using the prototype method, using the prototype data to represent other wind turbine data of the same

model will result in large errors. Since the wind tower extrapolation method needs to consider several factors, such as the topography, humidity, and pressure of the wind farm, the macro- and micro-climate model is very complicated. In addition, the estimation accuracy is acceptable for larger regional levels. However, for the wind farm level, it suffers from poor performance.

The statistical method is based on the statistical analysis of the correlation between wind farm power generations, wind speed, and wind direction data and establishes the mapping relationship between wind speed and wind direction and the output power of the wind farm. In the literature (Rajagopalan and Santoso, 2009), according to the actual measurement data of the wind farm, the model is based on the autoregressive moving average model (ARMA). Since ARMA can only handle stationary time series, the researchers applied the autoregressive integral moving average model (ARIMA), which can handle non-stationary time series, to wind power forecasting. However, statistical methods only analyze the superficial relationship between variables in time series. It is difficult to deal with complex and non-linear relationships.

At present, artificial intelligence-based methods are generally used for wind power forecasting and available power estimations. The artificial intelligence method uses historical power data, NWP data, etc., as input information to establish a non-linear mapping relationship between the output and multi-variables. Compared with statistical methods, the adaptability and self-learning ability of artificial intelligence methods have been significantly improved. Li et al. (2019b) used the Spearman rank correlation coefficient method to determine the hyperparameters of the long short-term memory (LSTM) network prediction model, which can effectively determine the initial step size range. Compared with the BP neural network, the wind power prediction accuracy based on the LSTM model is higher. Kisvari et al. (2021) combined the grid search method to adjust the hyperparameters of the gated recurrent unit (GRU) neural network. The proposed method achieves high accuracy with low computational cost. It shows robustness and low sensitivity to noise. The aforementioned wind power prediction methods only consider historical time-series features and do not consider complex spatial relationships. Therefore, the convolutional neural network (CNN) is used to extract the spatial features of wind farms to improve the accuracy of wind power prediction. Bai et al. (2018) redesigned the structure of CNN and proposed a temporal convolutional network (TCN). The research object of the aforementioned literature is only a single wind farm, and in most cases, regional prediction of wind power is required. Therefore, the temporal and spatial correlation between multiple wind farms needs to be further considered. Wang et al. (2022b) considered the dynamic spatio-temporal correlation between adjacent wind farms and calculated the spatio-temporal correlation matrix, modeling a graph structure with dynamic spatio-temporal correlation information as a graph convolution network (GCN) input. Experiments prove that the prediction accuracy of theoretical power generation of wind farms has been improved.

A combined model is a combination of two or more models, which eliminates the limitations of the individual models by combining their advantages in order to maximize the advantages of each method and improve the accuracy of wind farm power prediction. In a study by He et al. (2022), the weather is divided into different types according to the meteorological characteristics, and the IOWA operator is applied to assign different weight coefficients to the CNN and LSTM. The final power prediction is obtained by weighting the outputs of the two

models. In a study by Liang et al. (2021), a method of CNN combined with LSTM is proposed to obtain spatial distribution characteristics of the long-term wind speed and short-term time-series characteristics. Since the CNN can only be used to process regularly arranged images, the GCN is proposed to enable feature extraction for non-Euclidean structured data. Kan and Liu (2019) used the LSTM model to extract temporal features from historical data and used the graph convolution technique to extract spatial features from multiple PV plant data in the same region. Liao et al. (2022) combined the GCN and LSTM network, adopted the GCN, captured the complex spatial correlation between adjacent wind farms through the adjacency matrix, and learned the dynamic change of the wind power curve based on LSTM. Compared with the single wind power prediction method, the combined method predicts the wind farm power with higher accuracy (Chen et al., 2021), which can retain the advantages of each model. Therefore, wind farm power prediction models based on combined methods have received extensive attention from researchers. In recent years, some scholars have adopted more advanced hybrid models, including the correlation-constrained and sparsity-controlled vector autoregressive models for spatio-temporal wind power forecasting (Zhao et al., 2018), feature extraction of meteorological factors for wind power prediction based on the variable weight-combined method (Lu et al., 2021), spatial model-based short-term wind power prediction (Ye et al., 2017), and ultra-short-term combined prediction approach based on kernel function with a specially designed switch mechanism (Peng et al., 2021).

Most of the existing wind farm power prediction methods used numerical weather forecast data or measurement data from weather towers in wind farms as model inputs and the theoretical power of the whole wind farm as the model output. Since the state of the wind turbines and the loss in the field are not taken into account, they can only be used for theoretical power generation calculation and theoretical prediction of wind farms. Few studies have been conducted to calculate the available power of wind farms.

The previous theoretical power estimation is aimed at using numerical weather forecasts to estimate the total power generated by the entire wind farm, which cannot be accurate to each wind turbine, and the error is large. The emergence of a stand-alone information system can provide the nacelle wind speed of each wind turbine so that the grid dispatching department can be based on stand-alone actual measurement data to more accurately estimate the wind turbine and wind farm available power generation. Second, in most areas of the country, the only data available to the grid dispatching department are the weather forecast data and the active power of the grid connection point, which does not include the status of waiting for wind, planned shutdown, unit failure shutdown, and other wind turbine operating conditions. However, the active power emitted by the wind turbine in the normal power generation operation state reflects the real power generation capacity of the wind turbine.

Regional control centers in China enhanced the existing SCADA system to collect each wind turbine's operating status, wind speed, and wind direction from every wind farm, which provides an opportunity for accurate computation of available power of each wind farm. In this paper, we propose a deep spatio-temporal neural network for calculating the available power generation in wind farms based on the data to integrate wind farm's SCADA into the control center's SCADA. Long short-term memory (LSTM) is used to extract temporal information, and the graph convolutional network (GCN) is used to

describe the topology between wind turbines in wind farms and then solve the problem of spatial correlation and station loss.

In this paper, in order to improve the prediction accuracy, based on the real-time meteorological information of each wind turbine provided by the stand-alone information system, an ultra-short-term available power calculation method is proposed. This method combines the GCN and LSTM and considers the spatio-temporal correlation of wind farms and station losses. The key contributions are as follows:

1) Stand-alone information systems are used in wind power forecasting. They can provide real-time meteorological information such as the wind speed of each wind turbine and the information on the operating status of the wind turbine.
2) A novel graph neural network-based hybrid approach is proposed for ultra-short-term power prediction. LSTM is used to excavate the temporal characteristics of the wind speed of the wind turbine. The spatial position relationship of the wind turbine constitutes graph data, which is used as an input to the GCN to capture spatial dependence.
3) Based on the construction of the electrical topology connection diagram, the station loss problem is solved in combination with the GCN.

2 Problem formulation of available power estimation of wind farms

The available power of a wind farm refers to the maximum power theoretically available from the wind energy subtracting the power output of a wind turbine under maintenance and line losses in the wind farm. Additionally, instead of the actual power output of the wind farm, it may also be subject to curtailment and dispatch signals. To estimate the maximum available power of a wind farm, the wind speed information is used as the input. In this work, we used the historical wind speed data measured at each turbine and the turbine output power data under the maximum power point tracking working condition. Therefore, the summation of power generations of all the turbines, subtracting the total losses of the wind farm, equals the maximum available power of the wind farm.

$$P_{avai,i} = G\left(s_{windspeed,i}\right), \tag{1}$$

$$P_{farm} = \sum_1^n P_{avai,i} - P_{loss}, \tag{2}$$

where $s_{windspeed,i}$ is the wind speed at the ith wind turbine and $G(\bullet)$ is the estimation method which takes the wind speed as the input and outputs the maximum power generation of the turbine. $P_{avai,i}$ is the available power of the ith turbine. n is the total number of turbines, and P_{loss} is the loss within the wind farm. P_{farm} is the estimated maximum available power of the wind farm.

The aforementioned available power calculation does not take into account the in-station losses. Since the sources of loss in the station are diverse and difficult to estimate, therefore, it is proposed to construct an electrical connection diagram of wind farm wind turbines based on the GCN using the wind farm electrical main wiring diagram, and the characteristics of the loss in the station are included in the wind turbine electrical connection diagram.

There is a correlation between the available power generation of wind turbines and the historical nacelle wind speed. The LSTM neural

network layer can be constructed to effectively mine the time-series correlation information of wind turbine wind speed data. The complex topography and wake effects in the wind farm space have an impact on the power generated by the wind turbines. There is a correlation between adjacent wind turbines. Due to the different spatial distribution of WTGs and their own operating conditions, even the available power generation of WTGs of the same model varies. Therefore, based on the GCN, the fan relationship connection diagram is considered to extract spatial features.

## 3 The spatio-temporal NN-based algorithm for available power estimation

In terms of the structure of the proposed algorithm, in this work, a two-stage deep neural network-based algorithm is proposed, which mainly consists of four major layers. The first stage contains the first two layers, namely, the LSTM layer and the GCN layer. LSTM is a type of recurrent neural network that is proven to be very effective in terms of handling the input data with temporal relationships. The time-series wind speed data are used as the input of the LSTM layer. The GCN layer is implemented to extract the spatial topology of the wind dynamics in the wind farm to help the estimation of the available power. In the second stage, a third GCN layer is used to calculate the losses in the wind farm, where the topology of the wind farm is also used. The last layer is a fully connected layer for the final output. It should be noted that the layer here refers to a section of small network, which contains multiple sub-layers, like single GCN layers, activation function, and pooling layer. The first stage and the second stage can be pre-trained separately under the supervised learning scheme with SGD and then combined together for fine-tuning the parameters. Figure 1 shows the general framework of the proposed model, which demonstrated the structure of the proposed model. The details of the model are presented in the following sub-sections.

3.1 The correlation of time sequence extraction by the LSTM algorithm

There is a correlation between the available power generation of wind turbines and the historical nacelle wind speed (Wang et al., 2021). As a typical time series, forecasting the current wind power is not only related to forecasting the current input but also to the previous input and output. In order to fully exploit the information contained in the historical wind speed, this paper uses a long short-term memory network to extract the potential time-series information of wind turbine wind speed data for the estimation of available power generation of wind turbines.

The long short-term memory network is a variant of the recurrent neural network (RNN), which solves the problem of RNN gradient explosion and gradient disappearance. Based on the RNN, LSTM changes the single neural network layer in the RNN to one with four neural networks.

By changing the structure of neurons, LSTM introduces the gate mechanism and removes and adds information in the neurons through the gate mechanism (Hochreiter and Schmidhuber, 1997). The LSTM network can consist of one or more LSTM units that represent the data independent of each other in the time series and the data of multiple consecutive moments on the data of the current
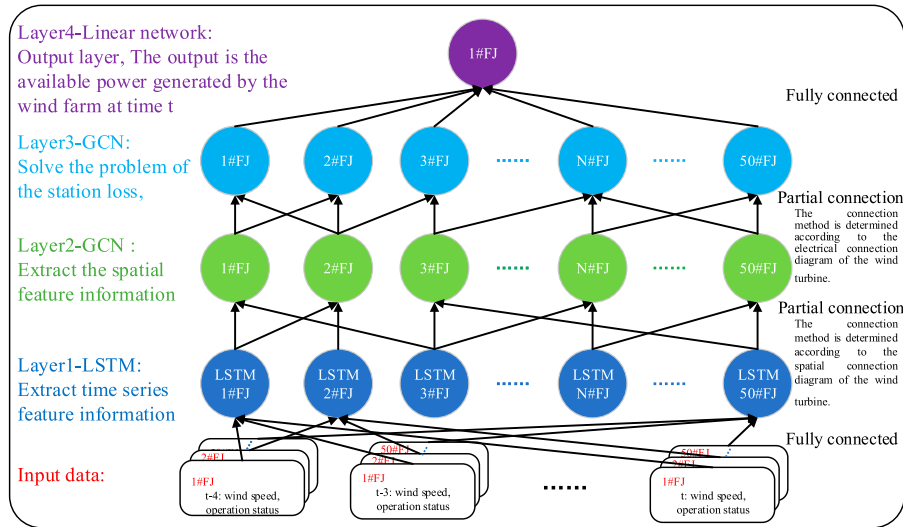
FIGURE 1
Model general frame diagram.

FIGURE 2
LSTM internal structure diagram.
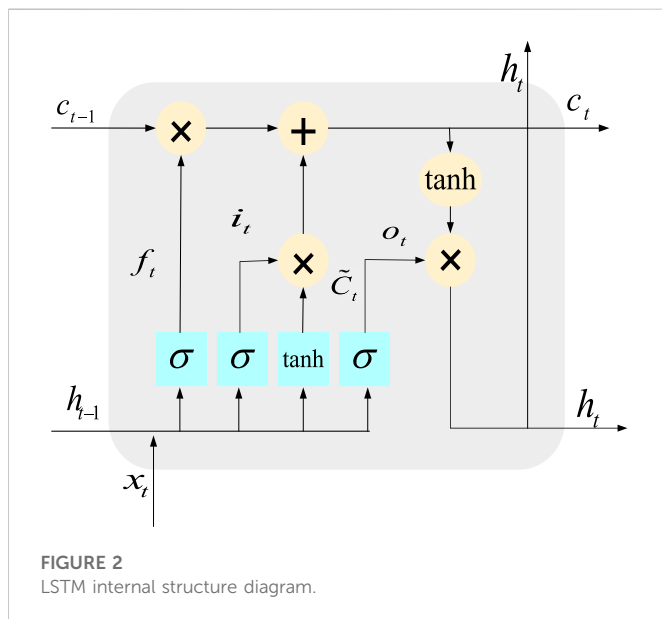
moment. The structure principle is shown in Figure 2. In Figure 2, x_t represents the nacelle wind speed and the individual operating state at moment t; h_t represents the power generated by the WTGs at moment t; and c_t represents the effect of the power generated by the WTGs at moments t-4, t-3, t-2, and t-1 on the power generated by the WTGs at moment t.

The LSTM unit consists of four parts: information storage chain, forgetting gate, memory gate, and output gate. The information storage chain runs through all the LSTM units before and after the LSTM network and is responsible for the storage and transmission of wind speed and operation status information of wind turbines at historical moments, and the information in the information storage chain of each LSTM unit is updated. ω and b denote the weight vector

and offset value in each gate mechanism, respectively; σ is the sigmoid activation function.

The role of the forgetting gate is to selectively forget the wind speed and operating status information components of some historical moments and to avoid too much information from the historical moments to affect the neural network's processing of the wind speed and operating status inputs of the wind turbine at the current moment. By the effect of the forgetting gate, the information in the historical moments of WTGs that are not strongly correlated with the estimated moments can be eliminated. The mathematical expression of the forgetting gate is shown in Eq. 3.

$$f_t = \sigma\big(W_f.[h_{t-1}, x_t] + b_f\big). \tag{3}$$

The memory gate is the control unit used to control whether the wind speed and operating status data of the WTGs at time t (now) are incorporated into the network cell state. First, the $tanh$ function layer is used to extract the valid information from the present vector, and the output is $\tilde{C}_t$; then, the sigmoid function is used to control "how much" of this memory is to be put into the unit state, and the output is i_t. By using the memory gate, the strong correlation between the WTG historical moments and the estimated moments can be retained in the network cell state and passed to the next moment. The mathematical representation of the memory gate is as follows:

$$\tilde{C}_t = \tanh(W_c.[h_{t-1}, x_t] + b_c), \tag{4}$$
$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i). \tag{5}$$

The output gate integrates the output data processed by the forgetting gate and the memory gate as the output of the power generated by the wind turbine at moment t. The mathematical expression of the output gate is as follows:

$$o_t = \sigma(W_o.[h_{t-1}, x_t] + b_0), \tag{6}$$
$$h_t = o_t * \tanh(C_t). \tag{7}$$

The network unit status is used to store information about the current WTG power, wind speed, and operating status and pass it to the next moment, which affects the output of the WTG power at the next moment. The update equation of the network unit status is shown in Eq. 8.

$$C_t = f_t {}^\star C_{t-1} + i_t {}^\star \tilde{C}_{t.}$$ (8)

The LSTM neural network layer can be constructed to effectively mine the time-series correlation information of wind turbine wind speed data. The key historical moments of wind speed are first screened out, and then the wind speed and operation status of the key historical moments are passed into the LSTM network as input vectors, and the valid information in the key historical moments is selected through memory gates and forgetting gates to update the network unit states, thus making the LSTM network fully consider the temporal correlation when estimating the available power generation of wind turbines and thus improving the available power estimation accuracy.

3.2 The temporal correlation by the GCN algorithm

The complex topography and wake effects in the wind farm space have an impact on the power generated by the wind turbines. The wind turbines are not distributed in isolation in the wind farm space, and the data of neighboring turbines have a large contribution to the estimation of the available power generation of wind turbines. From the perspective of wind speed in wind turbine nacelle, the nacelle wind speed of a WTG at moment t in a certain wind direction has a strong correlation with the nacelle wind speed of an upstream WTG at a moment t before, and when the wind direction changes, the upstream unit of that turbine will also change. In addition, due to the different spatial distribution of WTGs and their own operating conditions, even the available power generation of WTGs of the same model varies. Therefore, in this paper, we will analyze the spatial correlation of WTGs based on the Pearson correlation coefficient of the nacelle wind speed of each WTG and construct a WTG connection relationship diagram. Then, we use a GCN to extract the spatial information within the wind farm based on the output information of the LSTM network in the previous section and establish a differentiated WTG available power generation estimation model for the wind farm.

Traditional convolutional neural networks are limited to modeling Euclidean spatial data only, while graph convolutional networks can process non-Euclidean spatial data using graph representation, which makes them more suitable for modeling all wind turbines in a wind farm. In the wind farm available power estimation, the graph data are first constructed based on the WTG connection relationship graph and the nacelle wind speed and operation status data of WTGs, and then the extraction of spatial features among WTGs is completed based on the graph convolutional neural network.

Before extracting the spatial features of wind farms using the GCN, the wind turbine connection relationship graph should be constructed based on the correlation of all wind turbines in the wind farm. Each wind turbine in the wind farm is abstracted into nodes, and the correlation of nacelle wind speed between all wind turbines in the wind farm is analyzed by the Pearson correlation coefficient, and the units with the correlation coefficient greater than the set threshold can be judged as having a correlation, and the units with a correlation are

connected to form edges. Taking a wind farm with six wind turbines as an example, the connection relationship diagram of wind turbines in the wind farm is constructed, as shown in Figure 3.

The GCN is a neural network that performs feature extraction on graphs. A graph consists of a set of vertices and edges connecting the vertices [27]; the vertices are the objects studied, and the edges are specific relationships between two objects. In the wind power plant turbine connectivity graph, turbines are abstracted as nodes, and units with a strong correlation with each other constitute edges. The fan connection relationship graph notation can be expressed as $= (V, E)$, where V is the set of nodes (turbines) and E is the set of edges. The connections between turbines can be represented by the adjacency matrix $\tilde{A}$, and the number of edges directly connected by a given turbine is represented by the degree matrix $\tilde{D}$. The adjacency and degree matrices for the aforementioned example are constructed and shown in Figure 4.

The graph convolutional neural network uses the structural information of the edge–vertex connections of the WTG connectivity graph $\tilde{A}$ and the input data H of each WTG to perform feature extraction of the implicit graph information. The message propagation between layers is given by Eq. 9, and the feature aggregation between turbines is given by Eq. 10.

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)} \right),$$ (9)

$$\begin{aligned} \left(\tilde{D}^{0.5} \tilde{A} \tilde{D}^{0.5} H \right)_i &= \left(\tilde{D}^{0.5} \tilde{A} \right)_i \tilde{D}^{0.5} H \\ &= \left(\sum_k \tilde{D}_{ik}^{0.5} \tilde{A}_i \right) \tilde{D}^{0.5} H \\ &= \tilde{D}_{ii}^{0.5} \sum_j \tilde{A}_{ij} \sum_k \tilde{D}_{jk}^{0.5} H_j \\ &= \tilde{D}_{ii}^{-0.5} \sum_j \tilde{A}_{ij} \tilde{D}_{jj}^{-0.5} H_j \\ &= \sum_j \frac{1}{\sqrt{\tilde{D}_{ii} \tilde{D}_{jj}}} \tilde{A}_{ij} H_j \end{aligned}$$ (10)

In Eq. 9, $H^{(l)}$ represents the input data of WTGs in the layer l network after l-1 network aggregation, $H^{(l+1)}$ represents the output data of the layer l network, n is the total number of turbines in the turbine connectivity graph, and each turbine is represented by a d-dimensional feature vector; $\tilde{A} = A + I_N$, where if A is multiplied directly with H, only the neighboring turbine features are considered, so it is common to add the self-loop, i.e., the unit matrix, to take into account the turbines' own $\tilde{D}$, which is the degree matrix of the undirected graph, $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$; $W^{(l)}$ is the training parameter needed for the model and h is the output dimension; σ is the activation function, usually ReLU or Sigmoid. After aggregation and multiplication by the weight parameter matrix, the input of the next layer of the network is obtained after the non-linear activation function $H^{(l+1)}$.

The spatial correlation problem between wind turbines can be effectively solved by using the GCN, and the schematic diagram of extracting spatial features using the GCN is shown in Figure 5. In the wind farm wind turbine connectivity graph, the initial attribute of each node, i.e., the input, is the temporal feature vector of each wind turbine extracted after LSTM. After the first layer of the GCN, the wind speed and operation status information of each wind turbine's neighboring units are fully aggregated, and the feature vector of each node is
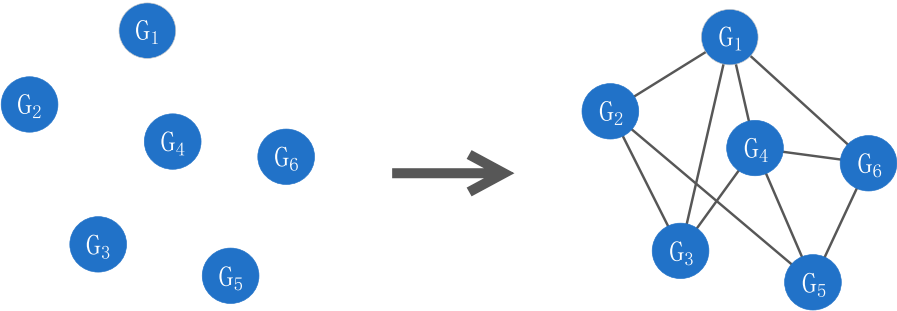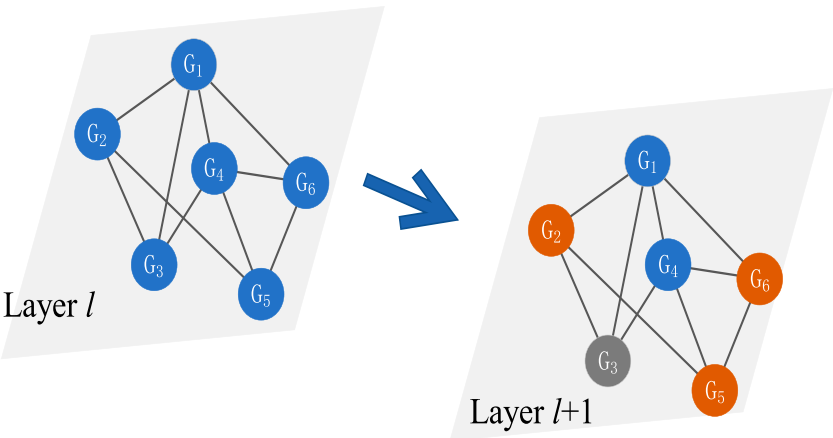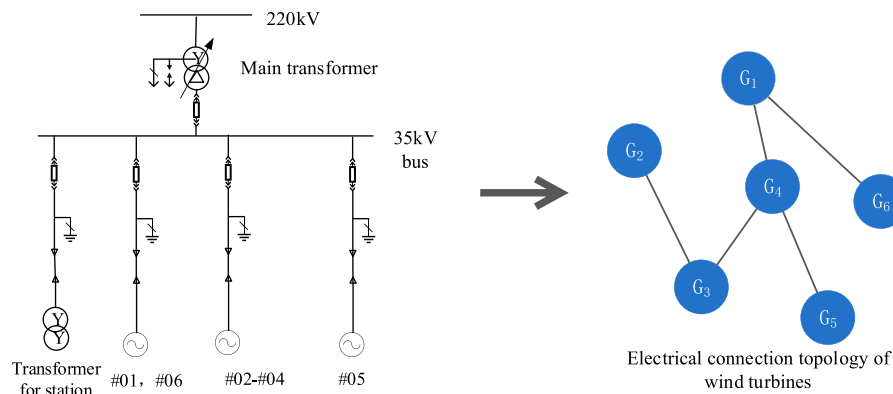
FIGURE 3
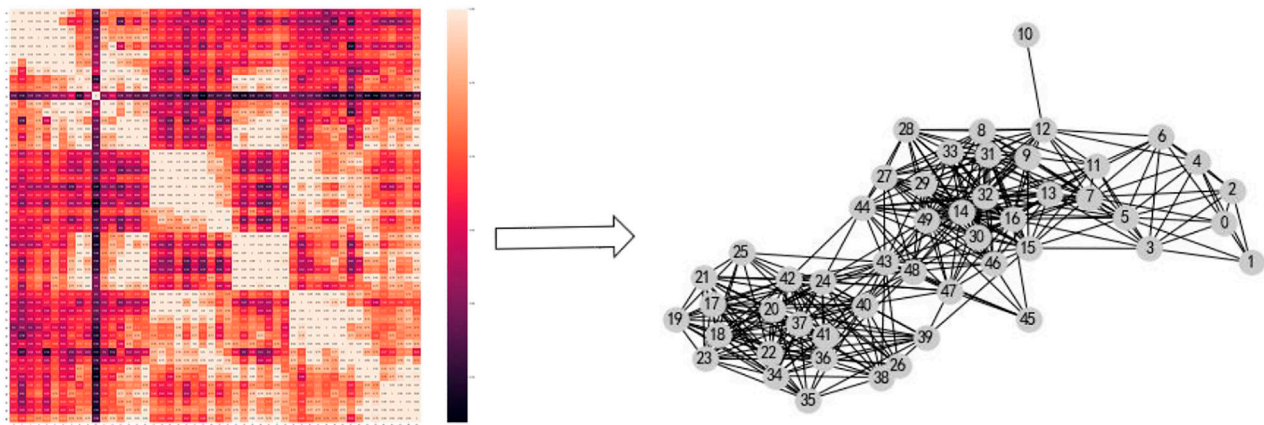Schematic diagram of building a connection diagram of wind turbines.

FIGURE 4
Adjacency matrix and degree matrix.

FIGURE 5
Schematic diagram of extracting spatial features by the GCN.

69

FIGURE 6
Electrical connection topology of the wind turbines.

FIGURE 7
Construction of the wind farm wind turbine connection diagram.

transformed into one dimension, outputting N*1 dimensional data, so that the model considers both the temporal characteristics of the wind speed and the spatial characteristics within the wind farm, and the output of this layer is the available power generated by the wind turbine.

3.3 The station loss problem based on the layer 2 GCN algorithm

Station losses also need to be taken into account when making available power estimates. Wind turbines in the wind farm through the transformer step-up, and the power lines are connected together, and they finally go through the wind farm main transformer step-up voltage to the grid. This process will inevitably produce line losses, box transformer losses, main transformer losses, etc., which are collectively referred to as station losses. Because the station loss is not part of the power sent through the grid connection point, the station loss does not belong to the wind farm available power generation. The calculation of the wind farm available power generation needs to subtract this part of the power loss. However,

station losses come from a variety of sources and are difficult to estimate, making it difficult to eliminate the impact of direct calculation on their accuracy.

Using the graph convolutional neural network, the wind turbine electrical connection relationship diagram can be constructed according to the main wind farm electrical wiring diagram, and the characteristics of the station losses are implied in the wind turbine electrical connection relationship diagram so as to indirectly complete the calculation of the losses of the convergence line, the main transformer in the wind farm. The schematic diagram of the wind turbine electrical connection relationship diagram is shown in Figure 6.

The wind turbines are abstracted as nodes in the second layer of the GCN graph, and the power transmission paths are abstracted as edges between the nodes. The propagation process of the nodes' own attributes in the GCN well-simulates the losses occurring in the power transmission process of the WTGs' power generation. The input of the second GCN layer is the output matrix of the first GCN layer $H_1 = R^{(N*l)}$, and the output features are the available power generation of each WTG in the wind farm, $H_2 = R^{(N*l)}$. Finally, after the fully connected layer, the node characteristics of N wind
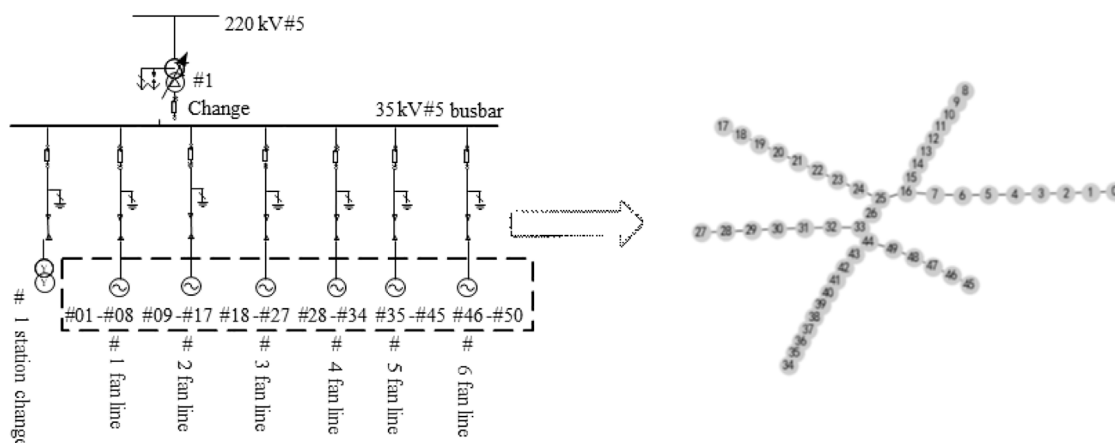
FIGURE 8
Construction of the electrical connection diagram of the wind turbine.
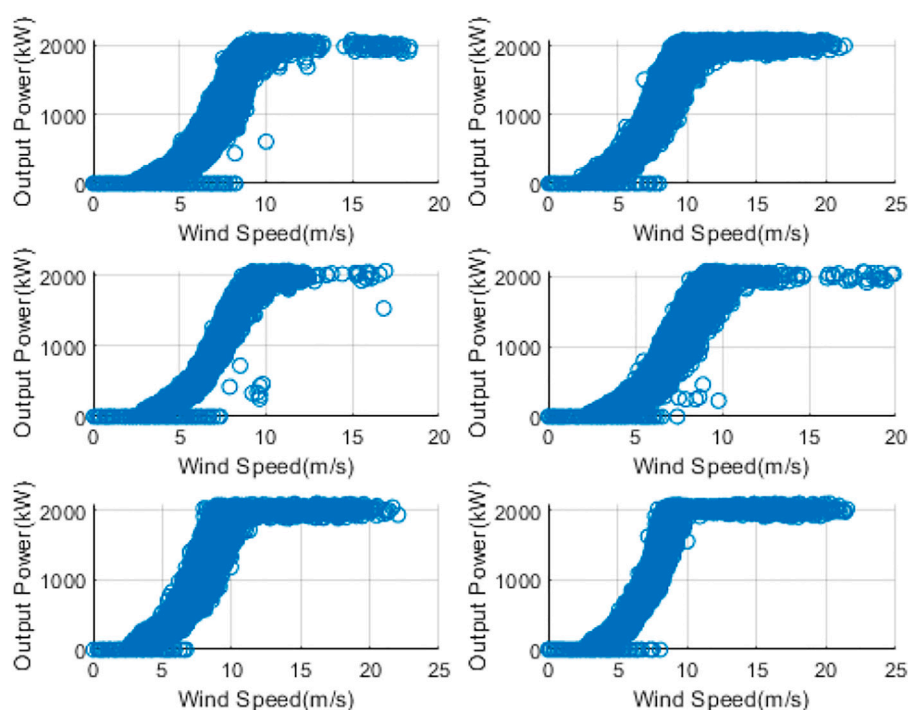
FIGURE 9
Plot of wind speed versus power output for six turbines.

turbines H_2 are used as the input to output the available power generated by the whole wind farm.

4 Case study

The actual operation data of a wind farm are used as an example to test the model of this paper. The installed capacity of the wind farm is 100 MW, including 50 turbines with a single capacity of 2 MW, one main transformer with a rated capacity of 120 MW, 50 box

transformers with a rated capacity of 2,300 kW, and six convergence lines; the turbine models are consistent. The wind farm data include the actual power of wind turbines, wind speed, and the actual power of wind farm grid connection points, and the data sampling interval is 1 December 2020–31 December 2020, with a sampling frequency of 1 min/point and a total of 44,640 samples. To use the data more efficiently for training, testing, and validation, the 10-fold cross validation is used.

The proposed deep neural network is trained with the stochastic gradient descent (SGD) optimization algorithm and

TABLE 1 Errors in the test set of the theoretical power generation estimation model for wind turbines.

Model	Root mean square (RMS) error (%)	Mean absolute error (MAE) (%)
Proposed method (model_1)	2.005	1.679
Prototype machine method (Ding et al., 2016) (model_2)	9.486	7.269
LSTM-based method (Li et al., 2019b) (model_3)	3.265	2.169
Power curve method (Tascikaraoglu et al., 2014) (model_4)	5.286	4.568
LSTM + GCN method (model_5)	2.925	1.354

FIGURE 10
Estimated power output versus the real power outputs of the six turbines.
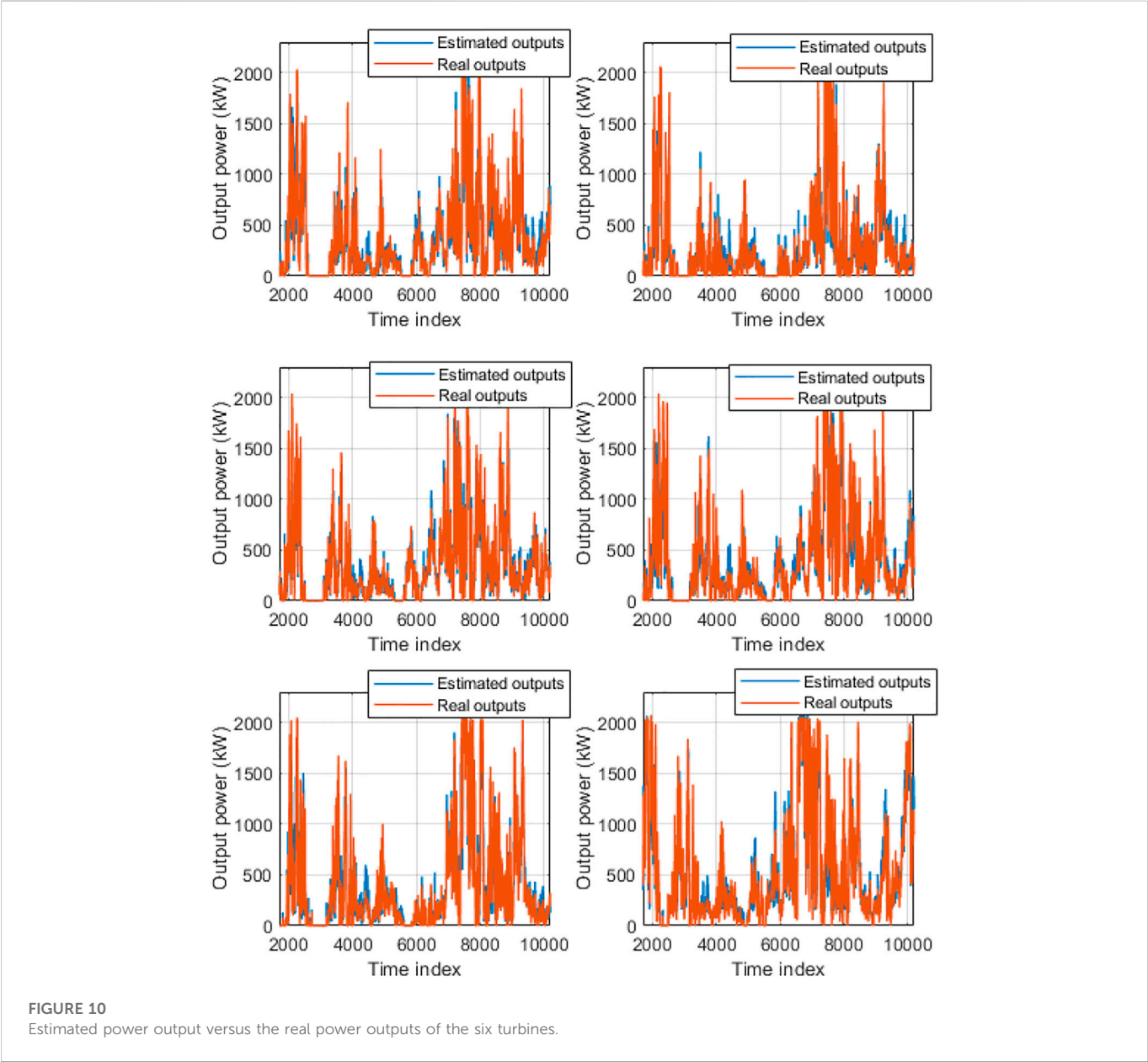
the cross-entropy loss functions. The cross-entropy loss increases as the predicted probability diverges from the actual label. A perfect model has a loss of 0. However, balancing the over and under fitting is important for the training process. Therefore, the loss will not be 0 in reality. The SGD is an iterative method for optimizing an objective function with smoothness properties. It can be considered a stochastic approximation of the conventional gradient decent optimization, where an estimated gradient is used as the replacement of the actual gradient.

FIGURE 11
Comparison of estimated power values and real power values of a part of the test set.
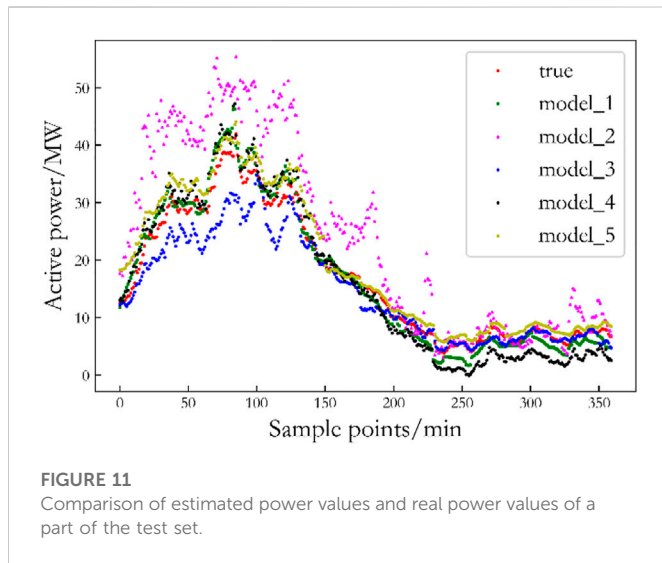
In this paper, four models are compared with the proposed method. The proposed model is marked as model_1, which establishes a wind farm available output estimation model based on LSTM and GCN considering the spatial and temporal correlation within the wind farm and the loss within the wind farm; model 2 is the prototype machine method (Ding et al., 2016), and model 4 is the power curve method (Tascikaraoglu et al., 2014); model 3 considers the temporal sequence within the wind farm, based on LSTM to complete the estimation. The estimation of available power generation of wind farms is done based on LSTM (Li et al., 2019b). Model_5 is an estimation model based on LSTM and the single-layer GCN without considering in-station losses.

After the adjustment of the model structure and parameters, the parameters of the proposed model are as follows:

(1) LSTM layer: Since the input variables of the model are only the nacelle wind speed and operation status, the dimension of the input layer is 2; the number of key historical moments is determined to be 4 based on correlation analysis, so the number of time steps of the input layer is 5; for the estimation of available power generation in wind farms, this paper only considers a single-layer LSTM estimation model, so the number of hidden layer parameters is 1; the dimension of the hidden layer is generally chosen to be four times the number of input variables Therefore, the number of hidden layer dimensions of the model is 8.

(2) GCN layer: The model in this paper uses two layers of the GCN. The graph topology of the first layer network is shown in Figure 7, which is determined by the spatial correlation of WTGs, with 50 nodes and 736 edges, and the input layer dimension of each node is determined by the hidden layer dimension of LSTM, which is 8, and the output layer dimension is 1; the graph topology of the second layer network is shown in Figure 8, which is determined by the wind farm electrical connection diagram with 50 nodes and 98 edges, each of which has an input layer dimension of 1 and an output layer dimension of 1.

(3) Fully connected layer: The fully connected layer is selected as the output layer of the spatio-temporal neural network. The input

layer dimension of this layer is 50, which represents the available power generation of 50 wind turbines; the output layer dimension is 1, which represents the estimated value of the available power generation of wind farms.

Based on the data of pre-processed samples completed the aforementioned four wind farm available power generation estimation models, the test set in each model estimation results is shown in Figure 9; the root mean square error and the average absolute error of the estimated value of each model and the real value are shown in Table 1.

In Figure 9, the original data in terms of wind speed versus output power are plotted for six different turbines. We have the data for all of the turbines in the wind farm. Due to the space limitation, only six of them are presented here. From the pattern in the figure, we can see that the maximum power of each turbine is around 2 MW. The output power of the turbines saturated around 2 MW. When the wind speed is low, in addition to the dead zone of the turbines, there are also chances that the wind turbines do not produce any power due to operation dispatches. Furthermore, the data contain outlier points due to maintenance or other reasons.

In Figure 10, the estimated power output results of the proposed method are shown for the six wind turbines. From the result, it can be seen that the proposed method can provide relatively accurate results, comparing with the real outputs.

From Figure 11, the estimation error of model_2 is unstable and increases significantly when the wind farm is at high power generation, and the estimation error of some samples is significantly larger than that of others at low power generation, compared with the model proposed in this paper (model_1), the LSTM model (model_3), and the power curve method (model_4), which are more stable. The estimated value of the LSTM model (model_3) is lower than the real value when the wind farm is at high power generation and slightly higher than the real value when it is at low power generation. The estimated values of the proposed model (model_1) and the power curve method (model_4) are slightly higher than the true values when the wind farm is at high power generation and slightly lower than the true values when it is at low power generation. As can be seen from Figure 9, the difference between the estimated values and the true values for all the samples tested in the proposed model (model_1) is not large when the wind farm is at low power, while the difference between the estimated values and the true values for some samples of the prototype method (model_2) is small, but there are also some samples that differ significantly from the true values.

From Table 1, it can be seen that the estimation error of the wind farm available power estimation model (model_1) proposed in this paper is 2.005%, and it is improved by 7.481% compared with the model_2 and by 1.26%, 3.281%, and 0.92% compared with model 3, model 4, and model_5, respectively. Model_5 adds a layer of GCN relative to model_3 to consider the spatial correlation within the wind field, reducing its root mean square error by 0.34%.

To sum up, based on the long short-term memory network and the graphical convolutional neural network, this paper establishes a model (model_1) that considers the temporal order of the wind speed of wind turbines and the spatial correlation of wind turbines inside wind farms and the station loss and has a more stable performance than that of model 2, model 3, model 4, and model_5, and its estimation error is lower and closer to the real value; by comparing model_3 and model_5 with the models proposed in this paper, it can be seen that the two-

layer GCN in the model can effectively extract the spatial information inside the wind farm and solve the problem of loss in the station.

5 Conclusion

We investigate the problem of estimating the available power generation of wind farms and propose a model for calculating the available power generation of wind farms based on long short-term memory networks and graph convolutional networks.

When comparing the proposed method with the four available power generation estimation methods, namely, the sample machine method, the LSTM model, the power curve method, and the available generation power estimation method that considers only the spatio-temporal correlation method in the wind farm, the root mean square error of the proposed method is 2.138%, which is 77.5%, 34.5%, 59.6%, and 26.9% higher than that of the sample machine method, the LSTM model, the power curve method, and the method that considers only the spatio-temporal correlation within the wind farm, respectively, proving the feasibility and superiority of the proposed method.

The improvement of the accuracy of the estimation of the available power generation of wind farms will facilitate the online dispatching and optimization of the strategy of the direct wind power AGC (automatic generation control) system by the dispatching department of the State Grid Corporation, promote the consumption of new energy, and help the country achieve the reduction of carbon emissions.

In terms of the findings, limitations, and recommendations, the authors would like to share the following:

1) Due to the aerodynamics, the output power of wind farms has strong spatial correlations.
2) With the help of the data from individual turbines, more accurate estimation results in terms of available power can be generated because the state of individual turbines can be considered, including the rotating maintenance of turbines within the wind farm.
3) The performance of the machine learning algorithms is limited by the quality of the input/training dataset. Therefore, it is very important to wash and clean the dataset and find the outliers, resulting in improved estimation accuracy.

The limitation of the proposed method is also the data availability. It is fortunate that we have the data of individual wind turbines in terms of wind speed, output power, and operation states. Therefore, the model can be developed in this work, and the available power can be estimated based on individual turbine outputs and losses. Otherwise, the conventional wind forecasting is the best way to predict the output power of a wind farm.

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

YL and PZ proposed the algorithm, KH and JL wrote the manuscript, and ZL provided valuable comments.

Funding

Conflict of interest

Author YL was employed by North China Branch of the State Grid Corporation of China.

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

References

Bai, S., Kolter, J. Z., and Vjapa, Koltun (2018). An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv:1803 https://arxiv.org/abs/1803.01271.

Chen, H., Birkelund, Y., and Yuan, F. J. E. R. (2021). Examination of turbulence impacts on ultra-short-term wind power and speed forecasts with machine learning. Energy Rep. 7, 332–338. doi:10.1016/j.egyr.2021.08.040

Ding, Kun, Qingquan, Lü, and Xu, Cai (2016). Statistical analysis of measured data on the calculation of wind farm curtailed wind power by model machine method. Renew. energy 1, 56–63.

Focken, U., Lange, M., and Waldl, H. P. (2001). "Previento-a wind power prediction system with an innovative upscaling algorithm," in Proceedings of the European Wind Energy Conference, Copenhagen, Denmark, 276.

Fu, Yang, Ren, Zixu, Shurong, Wei, Wang, Yang, Huang, Lingling, and Jia, Feng (2022). Ultra-short-term power prediction of offshore wind power based on improved LSTM-TCN model. Chin. J. Electr. Eng. 12. doi:10.13334/j.0258-8013.pcsee.210724

Guo, Haisi, He, Hui, and Bao, Daen (2019). Theoretical power generation calculation and comparative optimization of wind turbines. Electr. Power Energy 40 (3), 339–343.

He, B., Ye, L., Pei, M., Lu, P., Dai, B., and Li, Z. (2022). A combined model for short-term wind power forecasting based on the analysis of numerical weather prediction data. Energy Rep. 8, 929–939. doi:10.1016/j.egyr.2021.10.102

Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. Neural Comput. 9 (8), 1735–1780. doi:10.1162/neco.1997.9.8.1735

Jiang, Wenling, Feng, Shuanglei, and Sun, Yong (2014). Research on the calculation method of abandoned wind power in wind farms based on wind speed data of engine room. Power Grid Technol. 38 (3), 647–652. doi:10.13335/j.1000-3673.pst.2014.03.016

Kan, Bowen, and Liu, Guangyi (2019). Prediction of distributed photovoltaic power generation based on graph machine learning. Power Supply 36 (11), 20–27. doi:10.19421/j.cnki.1006-6357.2019.11.003

Kisvari, A., Lin, Z., and Liu, X. J. R. E. (2021). Wind power forecasting–A data-driven method along with gated recurrent neural network. *Renew. Energy* 163, 1895–1909. doi:10.1016/j.renene.2020.10.119

Li, J., Geng, D., Zhang, P., Meng, X., Liang, Z., and Fan, G. "Ultra-short term wind power forecasting based on LSTM neural network," in Proceedings of the 2019 IEEE 3rd International Electrical and Energy Conference (CIEEC), Beijing, China, September 2019 (IEEE).

Li, K., Zhang, P., Li, G., Wang, F., Mi, Z., and Chen, H. J. I. A. (2019). Day-ahead optimal joint scheduling model of electric and natural gas appliances for home integrated energy management. *IEEE Access* 7, 133628–133640. doi:10.1109/access.2019.2941238

Liang, Chao, Liu, Yongqian, Zhou, Jiakang, Yan, Jie, and Lu, Zongxiang (2021). Multi-point wind speed prediction method in wind farm based on convolutional recurrent neural network. *Power Grid Technol.* 45 (2), 534–542. doi:10.13335/j.10003673.pst.2020.0767

Liao, W., Bak-Jensen, B., Pillai, J. R., Yang, Z., and Liu, K. J. E. P. S. R. (2022). Short-term power prediction for renewable energy using hybrid graph convolutional network and long short-term memory approach. *Electr. Power Syst. Res.* 211, 108614. doi:10.1016/j.epsr.2022.108614

Liu, Zhongyu, Li, Yanlin, and Zhou, Yang (2020). *Graph neural networks: Analysis of GNN principles*. Beijing, China: Machinery Industry Press.

Lu, P., Ye, L., Zhao, Y., Dai, B., Pei, M., and Li, Z. (2021). Feature extraction of meteorological factors for wind power prediction based on variable weight combined method. *Renew. Energy* 179, 1925–1939. doi:10.1016/j.renene.2021.08.007

Peng, L. A., Lin, Y. A., Yong, T. B., Yz, A., Wz, B., and Ying, Q. C. (2021). Ultra-short-term combined prediction approach based on kernel function switch mechanism. *Renew. Energy* 164, 842–866. doi:10.1016/j.renene.2020.09.110

State Grid Corporation of China (2018). *Calculation method of wind power theoretical power and blocked power: Q/Gdw 11900-2018*. Beijing: State Grid Corporation of China

Rajagopalan, S., and Santoso, S. "Wind power forecasting and error analysis using the autoregressive moving average modeling," in Proceedings of the IEEE Power & Energy Society General Meeting, Calgary, AB, Canada, July 2009.

Tascikaraoglu, A., Uzunoglu, M. J. R., and Reviews, S. E. (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, Y., Zou, R., Liu, F., Zhang, L., and Liu, Q. (2021). A review of wind speed and wind power forecasting with deep neural networks. *Applied Energy* 304, 117766. doi:10.1016/j.apenergy.2021.117766

Wang, F., Chen, P., Zhen, Z., Yin, R., Cao, C., and Zhang, Y. (2022). Dynamic spatio-temporal correlation and hierarchical directed graph structure based ultra-short-term wind farm cluster power forecasting method. *Appl. Energy* 323, 119579. doi:10.1016/j.apenergy.2022.119579

Wang, L., He, Y., Liu, X., Li, L., and Shao, K. J. E. R. (2022). M2TNet: Multi-modal multi-task Transformer network for ultra-short-term wind power multi-step forecasting. *Energy Rep.* 8, 7628–7642. doi:10.1016/j.egyr.2022.05.290

Ye, L., Zhao, Y., Zeng, C., and Zhang, C. (2017). Short-term wind power prediction based on spatial model. *Renew. energy* 101 (4), 1067–1074. doi:10.1016/j.renene.2016.09.069

Zeiler, A., Brooks, D., Blau, G., and Pekny, J. (2011). Elsevier, Improved wind power forecasting with ARIMA models*Comput. Aided Chem. Eng.*, 29 1789–1793.

Zhao, Y., Ye, L., Pinson, P., Tang, Y., and Lu, P. (2018). Correlation-constrained and sparsity-controlled vector autoregressive model for spatio-temporal wind power forecasting. *IEEE Trans. Power Syst.* 33, 5029–5040. doi:10.1109/tpwrs.2018.2794450

A model predictive control strategy for enhancing fault ride through in PMSG wind turbines using SMES and improved GSC control

Sobhy M. Abdelkader[1,2], Ernest F. Morgan[1], Tamer F. Megahed[1,2], Wesam Rohouma[3]* and Omar Abdel-Rahim[1,4]*

[1]Electrical Power Engineering, Egypt-Japan University of Science and Technology (E-JUST), New Borg El-Arab City, Egypt, [2]Electrical Engineering Department, Faculty of Engineering, Mansoura University, El-Mansoura, Egypt, [3]Electrical Power and Renewable Energy, University of Doha for Science and Technology, Doha, Qatar, [4]Electrical Engineering Department, Faculty of Engineering, Aswan University, Aswan, Egypt

Wind energy has emerged as a prominent player in the realm of renewable energy sources, both in terms of capacity and technological adaptability. Among the various renewable energy technologies, wind turbine generators stand out as the most widely employed. Recently, gearless permanent magnet synchronous generators have gained traction in the wind energy sector due to their appealing features, such as reduced maintenance costs and the elimination of gearboxes. Nevertheless, challenges remain, particularly concerning the grid-friendly integration of wind turbines, specifically with regard to high voltage ride-through (HVRT) and low voltage ride-through (LVRT) improvements. These challenges pose a threat to grid stability, impede Wind Turbine Generator performance, and may lead to significant damage to wind turbines. To address these concerns, this research proposes an integrated strategy that combines a model predictive control (MPC) superconducting magnetic energy storage (SMES) device with a modified WTG grid-side converter control. By coupling SMES devices to the dc-link of Permanent Magnet Synchronous Generator WTGs, the proposed approach aims to achieve an overvoltage suppression effect during grid disturbances and provide support for grid reactive power. Through various test scenarios, the feasibility and practicality of this suggested technique are demonstrated.

KEYWORDS

fault ride-through, grid faults, model predictive control, permanent magnet synchronous generators, wind energy

1 Introduction

Governments worldwide have shown significant interest in renewable energy production, leading to substantial investments and development efforts. The wind sector experienced remarkable growth. Globally, 77.6 GW of new wind power capacity was connected to power grids in 2022, bringing total installed wind capacity to 906 GW, a year-on-year (YoY) growth of 9%. Projections indicate that the total installed wind capacity will reach 2 TW by 2030 (Global Wind Energy Council, 2023). This surge in wind energy has

prompted network operators and researchers to focus on enhancing the efficiency of wind power generators integrated into electrical grids. Consequently, the integration of wind power facilities using permanent magnet synchronous generators (PMSGs) into utility networks is on the rise (Jiang et al., 2021; Abdel-Rahim et al., 2022; Morgan et al., 2022a; Raouf et al., 2023).

PMSG wind turbine generators (WTGs) offer numerous advantages, such as excellent operational efficiency, self-excitation capability, gearbox elimination, high power factor, and reliability (Abdel-Rahim et al., 2014). However, the impact of wind energy on power networks necessitates addressing issues related to security, stability, and operation. Controlling voltage, frequency, and power while avoiding network disruptions is crucial. In the event of a malfunction, it was previously acceptable to isolate wind energy conversion systems (WECs) from the electric grid. However, to prevent power outages, strict grid codes have been established by transmission line operators worldwide, mandating wind turbines to remain connected to the grid even under adverse conditions. Failure to provide adequate protection can lead to disconnection of WTGs or damage to turbine and converter switches (Lyu et al., 2020).

To address these challenges, fault ride-through (FRT) methods have been developed. Notably, direct-drive PMSG wind turbines (WTs) have the advantage of contributing minimally to fault currents, allowing them to meet FRT criteria to some extent. However, they lack the ability to stabilize system voltage during grid variations. To maintain grid voltage, WTGs must contribute reactive power (López et al., 2009; Sarkar et al., 2018). As a result, FRT is crucial for WTs to withstand voltage fluctuations without disconnection from the grid. The main focus of FRT in direct-drive PMSG WTGs is to manage excess energy stored in the WTG's internal dc-link while supporting the grid with reactive power during disturbances (Hu et al., 2017).

Various FRT methods incorporate hardware schemes, such as braking choppers, dynamic voltage restorers (DVRs), energy storage systems (ESS), and STATCOM, as well as soft schemes like modified control for back-to-back converters (Benali et al., 2018; Djagarov et al., 2019; Huang et al., 2020a; Kim and Kim, 2021; Nasiri and Arzani, 2022). To enhance grid connection efficiency, using ESS to offset erratic active power supply during grid faults has been considered favorable (Makhad et al., 2022). High-capacity energy storage devices play a crucial role in quick dynamic power adjustment, which improves transient stability and guarantees consistent electricity output (Abhinav and Pindoriya, 2016). However, batteries used for reactive power assistance result in frequent charging and discharging cycles, reducing battery lifespan (Das et al., 2018). Hence, superconducting magnetic energy storage (SMES) is a more suitable option for FRT compared to other high-capacity energy storage devices like battery energy storage (BES). SMES devices offer advantages such as high energy storage efficiency, minimal self-discharge rate (when coupled with a superconducting switch), extended lifespan, and minimal environmental impact (Bar et al., 2021). Research has shown that SMES outperforms BES in terms of faster response and better dc voltage stabilization in dc-dc converters (Nikolaidis and Poullikkas, 2018).

SMES devices are widely used in WTG applications, either integrated into the WTG's dc-link or linked to the point of common coupling (PCC) outside the WTG. At the dc-link, SMES controls its dc-dc chopper to reduce excess power and compensate for reactive power from the WTG to the grid. When connected at PCC, SMES supports voltage and compensates for delivered wind power from PCC to the grid.

However, some SMES-based schemes that involve adding a dc-ac inverter and control adjustments in the external circuit can lead to increased capital costs and control complexity (Zheng et al., 2017; Morgan et al., 2022b). Additionally, these schemes may not be cost-effective since they require additional devices like SMES to enhance FRT in PMSG WTGs. A more viable approach is to incorporate a modified control scheme for the WTG to maintain the DC link voltage close to a constant. Various control methods have been employed to achieve this objective, including fuzzy logic controllers (FLCs) and artificial neural network (ANN) controllers. However, these methods have drawbacks, such as complex architecture and lengthy training times (Mukherjee and Rao, 2019; Ahsan and Mufti, 2020).

In contrast, proportional-integral (PI) controllers have proven effective in modern manufacturing processes due to their resilience, broad stability margin, simplicity, and low cost (Jannati et al., 2016). Despite these advantages, PI controllers may encounter challenges in heavy nonlinear systems, particularly when uncertainties are involved. Various optimization techniques, such as the genetic algorithm, continuous mixed p-norm (CMPN) algorithm, grey wolf optimizer, water cycle algorithm, and whale optimization algorithm, have been suggested for fine-tuning PI controllers for industrial applications (Dahiya et al., 2019; Qais et al., 2019; Qais et al., 2020a; Soliman et al., 2020; Joseph et al., 2022). While these algorithms excel in finding the best global optimum solution, they are time-consuming and require intricate flowcharts. To address these issues, the model predictive control (MPC) method is adopted in this study. MPC offers rapid dynamic response, adaptability to multiple variables and inequalities, improved resilience, and stability. MPC is extensively used in industry, has a straightforward calculating procedure, and has undergone rigorous long-term practice tests (Qais et al., 2020b; Chen et al., 2020).

Previous research has explored the coordination of SMES devices and modified converter control for other wind turbine types. However, these techniques cannot be directly applied to PMSG WTGs, necessitating further research into cooperative schemes involving enhanced WTG control and SMES to improve the stable grid connection of future wind farms.

To address these challenges, this paper presents a comprehensive control strategy that combines SMES with auxiliary reactive power support. A modified WTG converter control and a model predictive control SMES system are employed to achieve the objectives. The paper's main contributions are as follows:

1) A cooperative control strategy is proposed to enhance FRT using SMES.
2) A prediction model of SMES is established, and a corresponding model predictive control method is proposed to track power instructions for storing, charging, and discharging. Simulation results compare the performance of MPC with other control algorithms.
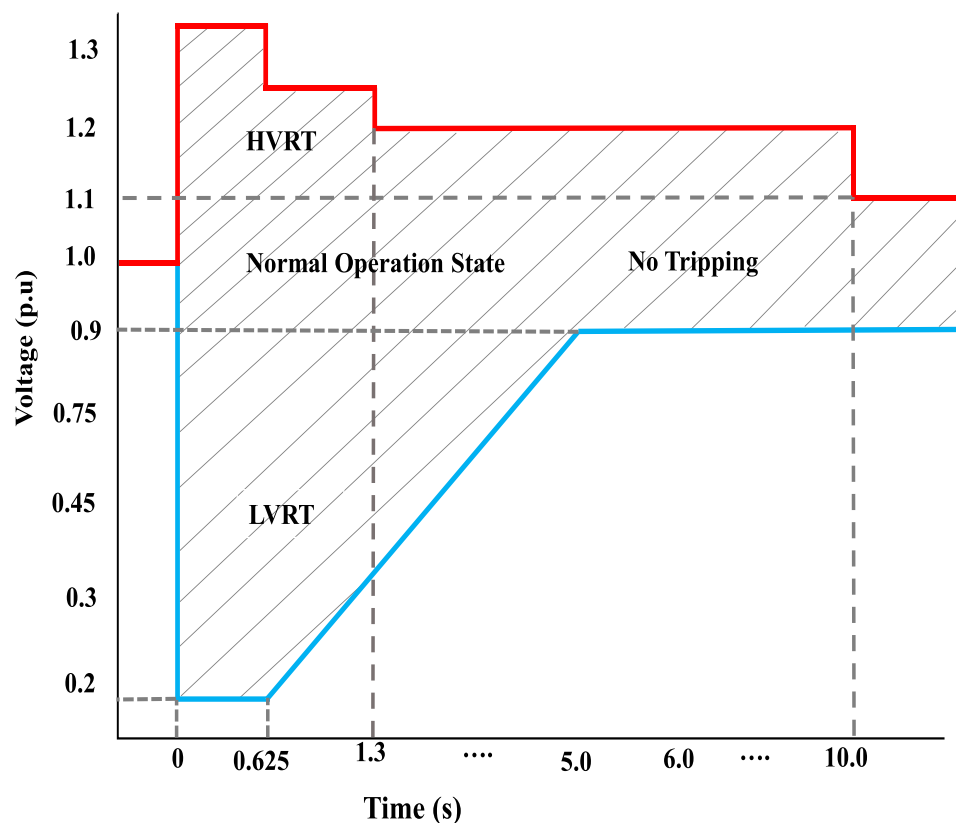
FIGURE 1
A typical grid code for fault ride-through (FRT).

2 FRT grid code requirement for PMSG wind turbines

Various countries have pursued diverse approaches to enhance their wind power generation capacities. In regions where the wind energy industry is well-established, grid codes have been implemented to accommodate the specific requirements of grid integration. One crucial aspect of wind power integration is the ability to withstand faults and disturbances in the grid, which is known as fault ride-through (FRT) capability (Abdel-Rahim and Funato, 2014; Huang et al., 2020b).

Utility grid standards impose requirements on wind turbine generators (WTGs) to possess sufficient High Voltage Ride-Through (HVRT) and Low Voltage Ride-Through (LVRT) capabilities. Figure 1 provides an illustrative example of the HVRT and LVRT criteria. According to these standards, the WTG should remain connected to the grid when the grid voltage falls within the shaded gray region. On the other hand, if the grid voltage exceeds this region, the WTG can safely trip without any adverse consequences. Moreover, when the terminal voltage ranges from 0.9 per unit (p.u.) to 1.1 p.u., the wind turbines should continue operating, but under different voltage conditions, a brief disconnection might be necessary (Luo et al., 2018; Abdel Aleem et al., 2017).

3 Mathematical modeling of the proposed hybrid WECS-SMES

The comprehensive system under study is depicted in Figure 2, encompassing components such as a wind turbine (WT) model, a Permanent Magnet Synchronous Generator (PMSG), power converters, a Superconducting Magnetic Energy Storage (SMES) device, and a grid model. The proposed method entails the parallel connection of the SMES device to the direct current (dc) link of the 1.5 MW integrated WTG. The essential characteristics of the integrated 1.5 MW WTG are summarized in Table 1.

3.1 Mathematical modeling of WECS

The mechanical power produced from the wind by the PMSG-WT is expressed as (Okedu, 2022):

$$P_w = \frac{1}{2}\rho A V_W^3 C_P(\lambda, \beta) \qquad (1)$$

In Eq. 1, P_w represents the wind power harnessed, measured in watts (W). The parameter C_P denotes the power coefficient, while ρ denotes the air density in kilograms per cubic meter (Kg/m^3). A represents the swept area in square meters (m^2), and V_w stands for

FIGURE 2
PMSG-SMES based wind turbine connected to the Grid.

TABLE 1 Key features of the PMSG WTG system.

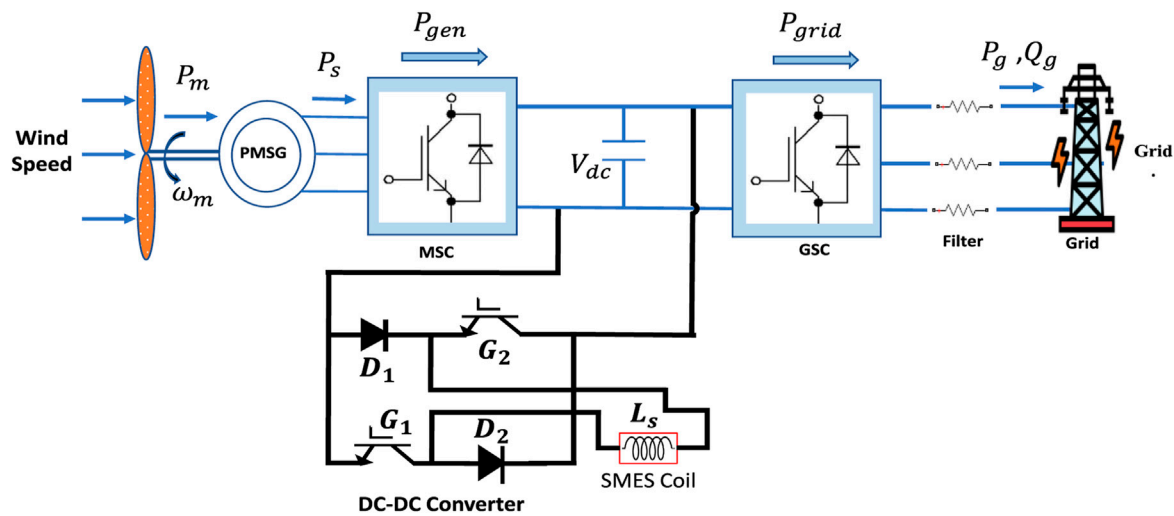| Parameter	Value	Parameter
Rated Power	1.5	MW
Rated wind speed	12	m/s
Rated Voltage	595	V
Number of poles pairs	40	
DC-link Voltage	1150	V

the wind velocity without any rotor interference, expressed in meters per second (m/s). The power coefficient C_P of the wind generator is determined by the ratio of the tip speed (λ) to the pitch angle (β), and this relationship is expressed in Eq. 2 (Okedu, 2022).

$$C_P\left(\lambda, \beta\right) = c_1\left(\frac{c_2}{\lambda_i} - c_3\beta - c_4\right)e^{\frac{-c_5}{\lambda_i}} + c_6\lambda \tag{2}$$

Where

$$\frac{1}{\lambda_i} = \frac{1}{\lambda - 0.08\beta} - \frac{0.035}{\beta^3 + 1} \tag{3}$$

Eq. 2 incorporates coefficients c_1 to c_6, which represent the characteristics of the wind turbine (WT). In the context of Permanent Magnet Synchronous Generator (PMSG) wind turbines, the maximum power point tracking (MPPT) relies on the rotor speed, enabling the attainment of maximum power output (Wang et al., 2018).

$$P_{MAX} = \frac{1}{2}\rho A\left(\frac{\omega_r R}{\lambda_{opt}}\right)^3 c_{popt} \tag{4}$$

The variables in question are as follows: λ_{opt} represents the optimal tip speed, c_{popt} denotes the optimal power coefficient, and ω_r indicates the rotor speed of the wind generator.

3.2 Mathematical modeling of PMSG

The dynamic model of the PMSG wind turbine in the d-q reference rotating frame is presented as shown in reference (Wang et al., 2018).

$$\frac{d\psi_{sd}}{dt} = -V_{sd} - R_s I_{sd} - \omega_e \psi_{sq} \tag{5}$$

$$\frac{d\psi_{sq}}{dt} = -V_{sq} - R_s I_{sq} - \omega_e \psi_{sd} \tag{6}$$

From Eqs 5, 6

$$\psi_{sd} = (L_{sd} + L_{md})I_{sd} + \psi_m \tag{7}$$

$$\psi_{sq} = (L_{sq} + L_{mq})I_{sq} \tag{8}$$

In the dynamic model of the PMSG wind turbine, the following variables are involved: V_{sd} and V_{sq} represent the stator circuit's voltage, R_s denotes the stator winding resistance, and I_{sd} and I_{sq} are the currents in the d and q reference frames, respectively. Furthermore, ω_e signifies the rotational speed of the wind generator, while ψ_{sd} and ψ_{sq} stand for the flux linkages of the stator circuit. Additionally, L_{sd} and L_{sq} represent the stator winding leakage inductances, and L_{md} and L_{mq} denote the magnetizing inductances. Finally, ψ_m signifies the flux linkage of the machine's permanent magnet. By substituting Eqs 7, 8 into Eqs 5, 6, the resulting differential equation can be obtained.
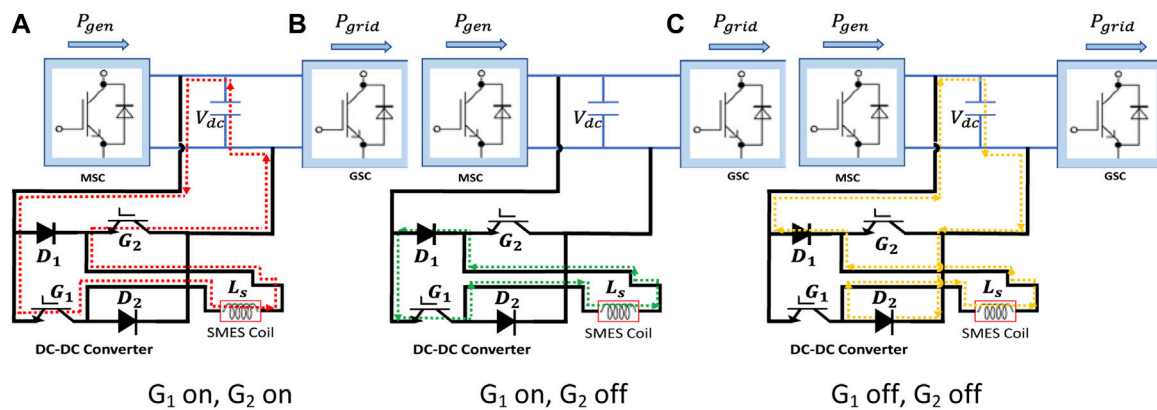
$$L_d\frac{dI_{sd}}{dt} = -V_{sd} - R_s I_{sd} - \omega_e I_q I_{sq} \tag{9}$$

$$L_q\frac{dI_{sq}}{dt} = -V_{sq} - R_s I_{sq} + \omega_e L_d I_{sd} + \omega_e \psi_m \tag{10}$$

$$L_d = L_{sd} + L_{md} \tag{11}$$

$$L_q = L_{sq} + L_{mq} \tag{12}$$

The active and reactive power of the Permanent Magnet Synchronous Generator (PMSG) are expressed as follows:

FIGURE 3
Operation mode of SMES **(A)** Charging mode **(B)** Standby mode **(C)** Discharging mode.

$$P_g = \frac{3}{2}\left(V_{sd}I_{sd} + V_{sq}I_{sq}\right) \tag{13}$$

$$Q_g = \frac{3}{2}\left(V_{sq}I_{sd} - V_{sd}I_{sq}\right) \tag{14}$$

The electro-torque of the wind generator with pole pairs (p) is represented as follows:

$$T_e = 0.5p\left(\psi_m I_{sq} + (L_d - L_q)I_{sd}I_{sq}\right) \tag{15}$$

In the case of a surface-seated Permanent Magnet Synchronous Generator (PMSG), it is reasonable to assume that $L_d = L_q$. Consequently, the expression for T_e can be formulated as follows:

$$T_e = 0.5p\left(\psi_m i_{sq}\right) \tag{16}$$

3.3 SMES Modeling

Among the various commercial energy storage devices (ESD) available, superconducting magnetic energy storage (SMES) stands out due to its rapid response speed, high power density, and extended lifespan (Gouda et al., 2020). In this study, SMES is emulated as an inductor with negligible resistance. The DC-DC side of the converter comprises two Insulated Gate Bipolar Transistors (IGBT) and two diodes, enabling a bidirectional, two-quadrant operation of the SMES magnet. This allows the SMES to charge, store, and discharge energy while the current is positive. The switching pattern of the transistors is adjusted to produce either positive or negative voltage at the magnet's terminals, turning it into an intrinsic zero-voltage device. The fundamental electrical circuit for the SMES chopper is depicted in Figure 3.

In the charging mode, when the controller commands the SMES to absorb energy, the current path is illustrated in Figure 3A. During this mode, IGBT G1 and IGBT G2 are activated to charge the SMES from the DC bus. The voltage V(t) across the inductor can be expressed as follows:

$$V(t) = L_s\frac{di(t)}{dt} \tag{17}$$

The expression for the energy stored in the SMES can be obtained as follows, taking into account the DC bus voltage (V), the coil inductance (L_s), and the charging current (i(t)) as a function of time (t). The DC current flowing in the coil is represented by I_s, while R_1 denotes the coil resistance, with the inner resistance considered negligible ($R_1 \approx 0$). Therefore, the energy stored in the SMES can be calculated using the given parameters:

$$E(t) = \frac{1}{2}L_s I_s^2(t) \tag{18}$$

In the discharge mode, IGBT G1 and G2 are turned off, and the controller manages the duty cycle of IGBT G2 to achieve the desired output current I_{ref}. For the discharge process, assuming the initial current in the SMES is I_0, and the duty ratio for IGBT G2 is denoted by D, the current in the SMES at time t can be expressed as follows:

$$I(t) = I_0 \exp{-\frac{R*t*D}{L_s}} \tag{19}$$

When the SMES does not require power exchange with the power grid, the chopper operates in standby mode to maintain the stored energy in the SMES. In this standby state, the current path is depicted in Figure 3C, where IGBT G1 remains continuously turned on, and IGBT G2 remains continuously turned off. The current flowing in the superconductor during the storage process can be represented as follows:

$$I(t) = I_0 \exp\left(-\frac{R_1*t}{L_s}\right) \tag{20}$$

If the superconductor material exhibits zero resistance ($(R_1 \approx 0)$, both the storing current and the stored energy remain constant. The control diagram of the chopper will be discussed in the subsequent section. For a summary of the essential characteristics of the SMES device, please refer to Table 2.

4 Principle of operation

The dual operation modes of the protection circuit exist, as described below:

TABLE 2 Key features of SMES Tokyo Model (Xu et al., 2019).

Parameter	Value	Parameter
Critical coil current	3375	A
Self-inductance	0.1	H

4.1 During voltage sag and swell

Eq. 21 illustrates the behavior of the DC link, where C and V_c represent the capacitance and voltage of the dc-link, respectively. P_f and P_r correspond to the grid filter and rotor power, respectively. To maintain dc-link voltage stability, it is necessary for the power transferred from the generator side (P_r) to equal the power transferred from the dc-link capacitor to the grid filter (P_f) (Xu et al., 2019). However, during voltage dips or spikes, a power imbalance occurs, leading to variations in the values of P_r and P_f Consequently, a rise in voltage occurs across the dc-link capacitor.

During High Voltage Ride-Through (HVRT), if the grid voltage increases, the output of the MSC (P_f) will not change significantly. Nevertheless, an overvoltage condition may cause energy to flow back toward the WTG, resulting in an unexpected spike in the capacitor voltage. Such an occurrence poses a threat to the normal operation of the WTG (De Siqueira and Peng, 2021).

To address this issue, the excess energy received by the dc-link capacitor during voltage fluctuations needs to be dissipated, or alternatively, the total capacitance of the DC-link (C_Total) should be increased, as depicted in Eq. 22, to maintain the dc-link voltage within a safe range (De Siqueira and Peng, 2021).

$$CV_C \frac{dV_c}{dt} = P_f - P_r \qquad (21)$$

$$V_c = \sqrt{\frac{2E_c}{C_{Total}}} \qquad (22)$$

4.2 Converter and SMES control

The primary aim of this study is to enhance the Fault Ride-Through (FRT) capability in the PMSG-based Wind Turbine Generator (WTG). To achieve this objective, the methods presented in Nielsen et al. (2010), which pertain to Machine and Grid Side Converter control, are adopted. Consequently, these methods will not be reiterated in this work. Instead, this section focuses on a modified control strategy for the grid-side converter that incorporates SMES to address the FRT issue, as illustrated in Figure 4.

4.3 Model predictive control strategy for SMES

The dc-dc chopper operates with binary on and off settings. Consequently, it has a set of switching states that encompass all possible permutations. However, certain combinations leading to short-circuits in the DC link are avoided. To determine the optimal switching state of the dc-dc chopper corresponding to the dc-link voltage, an optimization process is carried out. Subsequently, the optimal switching states are applied. Unlike traditional PID control, where pulse width modulation (PWM) is used to generate switching signals for the converters, the MPC-based SMES system directly produces the required switching signals for the chopper. Figure 5 illustrates the predictive control method for the dc-dc chopper. For calculating the expected voltage across the dc link, the predicted current flowing through the SMES coil at the sampling time T_s is utilized. Below is a discrete-time model representing the dynamic current in the SMES coil at the sampling time T_s.

The expected SMES current at time k is given by:

$$i_L^p(k+1) = \frac{i_L(k)}{T_s} + \left(\frac{T_s}{L_s}\right)(V_{dc}(k)) \qquad (23)$$

Where, L_s represents the inductance of the SMES coil, $i_L(k)$ represents the current of the SMES coil at time k, Dc link Voltage at time k is denoted by, $V_{dc}(k)$.

With its accompanying switching function as:

$$s = \begin{cases} 0, & \text{while } = G_1 \text{ is off and } G_2 \text{ is on} \\ 1, & \text{while } = G_1 \text{ is on and } G_2 \text{ is of} \end{cases} \qquad (24)$$

Table 3 illustrates the complete switching sequence for the DC-DC chopper. Where G_1 and G_2 represents the dc-dc chopper's switches.

At time k, the capacitor's expected current is:

$$i_c(k) = i_s(k) - i_l(k) - i_g(k) \qquad (25)$$

But since the SMES coil is considered already charged, the next available state is the discharge state, hence $i_c(k)$ is represented as;

$$i_c(k) = i_s(k) + i_l(k) - i_g(k) \qquad (26)$$

The relationship between voltage and capacitor current is as follows:

$$i_C = \frac{CdV_{dc}}{dt} \qquad (27)$$

$$i_c(k+1) = \frac{C[V_{dC}(K+1) - V_{dC}(K)]}{T_s} \qquad (28)$$

The derivative capacitor voltage $\frac{dV_C}{dt}$, is also replaced by the forward Euler approximation. Thus, the capacitor's predicted dc voltage at time k+1 can be derived as follows:

$$V_{dc}^p(k+1) = i_c(k) * \frac{T_S}{c} + V_{dc}(k) \qquad (29)$$

where, $V_{dc}^p(k+1)$ represents the capacitor's predicted voltage at time k + 1.

The cost functions g_1, thus the difference between the reference and forecast signal levels is defined as Eq. 30.

$$g_1 = \left| V_{dc}^p(k+1) - V_{dc}^*(K) \right| \qquad (30)$$

Where V_{dc}^* represents the reference value of the dc-link voltage.

Figure 6 depicts the control algorithm used by the MPC-based SMES system. The dc voltage forecast relies on Eq. 29.

FIGURE 4
Schematic representation of the MPC-based SMES control system.
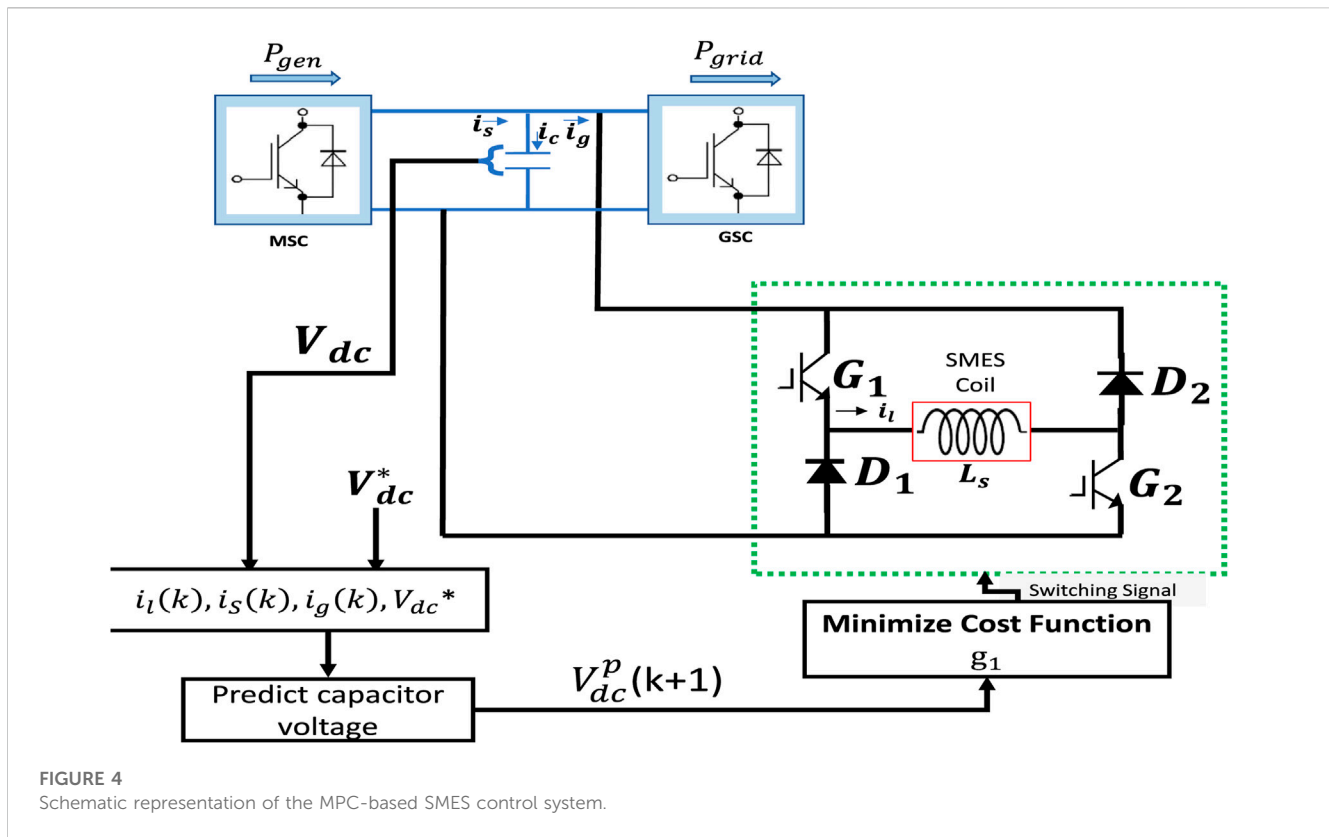
4.4 Modified Grid Side Converter control

For grid voltage angle detection, a phase-locked loop (PLL) is utilized (Yuan et al., 2020). During regular operation, iqref is set to 0, enabling the Grid-Side Converter (GSC) to operate in unity power factor mode. However, when the WTG is in Fault Ride-Through (FRT) mode, the dynamic reactive current reference is determined based on the magnitude of the voltage dips, as expressed in Eq. 31.

$$i_{gqref} = k \times I_N \times (V_N - V) \tag{31}$$

In Eq. 32, K represents the coefficient, i_gqref denotes the reactive current reference, and V refers to the instantaneous terminal voltage. V_N and I_N represent the nominal voltage and current of the WTG, respectively. During operation, if the terminal voltage is below the nominal voltage, a capacitive current reference will be transmitted to the converter. Conversely, if the terminal voltage exceeds the nominal voltage, an inductive current reference will be sent to the converter. It is essential to ensure that the active current reference is governed by Eq. 32 to restrict its value within the IGBT's current threshold during Fault Ride-Through (FRT) mode.

$$i_{gdref} \le \sqrt{I_{max}^2 - I_{gqref}^2} \tag{32}$$

The threshold current of the IGBT is denoted as I_{max}, and i_gqref represents the active current reference. For the purpose of this study, the maximum allowable current is set to 1.8 pu. As depicted in Figure 4, the active and reactive references remain the same as in conventional vector control during normal operating conditions. However, when the Fault Ride-Through (FRT) mode is activated, the i_gqref changes to Eq. 31 to provide reactive power assistance either in a capacitive or inductive manner, while ensuring that the i_gqref stays within the safe operating current limit of the IGBT.

During High Voltage Ride-Through (HVRT) and Low Voltage Ride-Through (LVRT) events, the Grid-Side Converter (GSC) is configured to operate in the Q-priority mode. This means that the converter prioritizes the contribution of reactive current over the contribution of active current.

5 Case studies

In Matlab\Simulink, a 1.5 MW PMSG-based WECS with a model predictive control SMES model is developed. The following are the conditions for the test case studies:

Case 1: 80% balanced voltage sag.

Case 2: 20% balanced voltage swell.

5.1 Case 1: Behaviours under 80% symmetrical voltage sag

In this experimental study, a balanced voltage drop of 0.80 pu is applied to the grid at $t = 3$s and persists for 150 m. Figure 6A illustrates
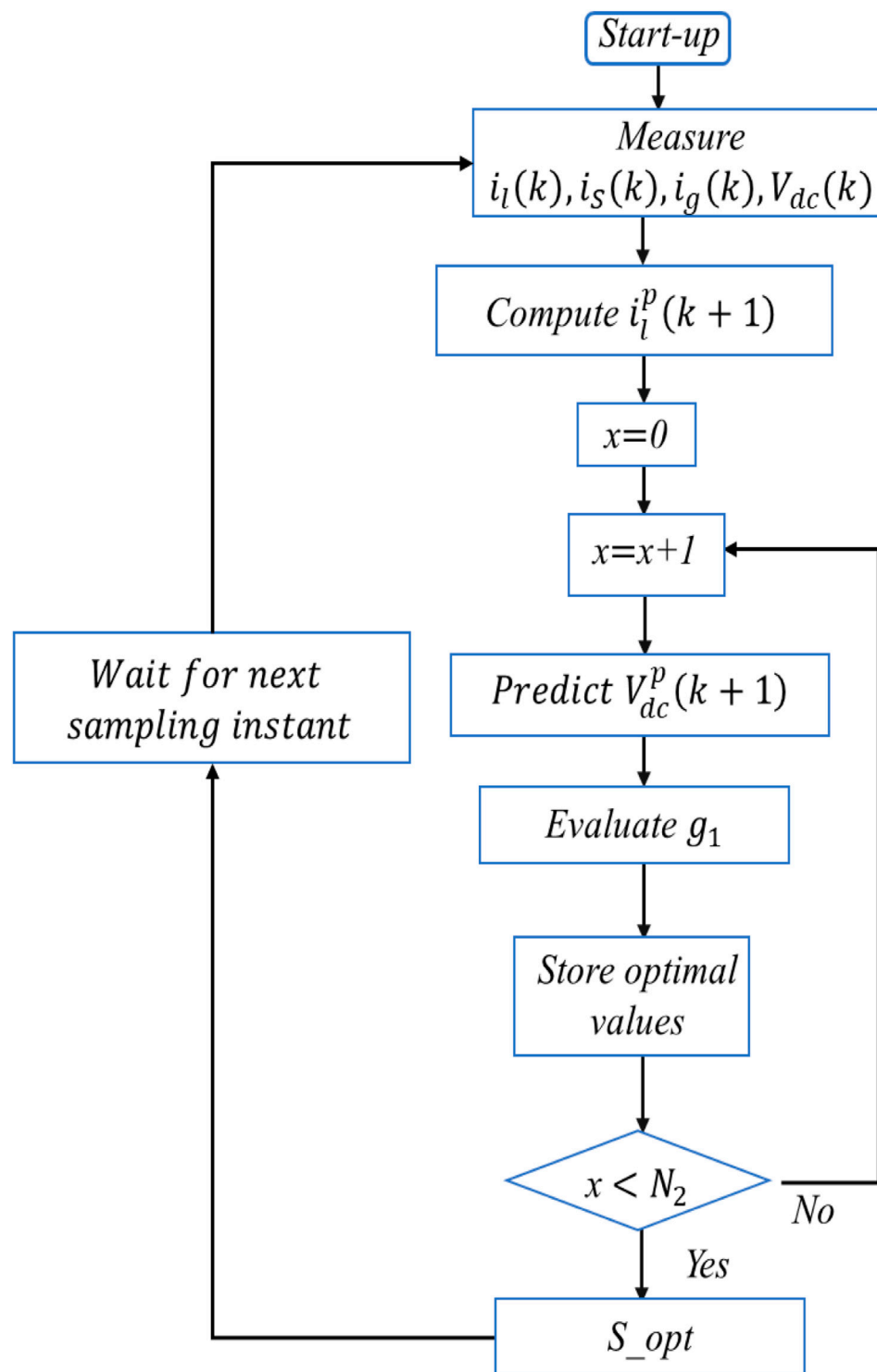
FIGURE 5
Algorithm for controlling the systems as implemented.

the 0.80 pu voltage drop at the grid-connected point during the fault. The root mean square value of the voltage drop at the grid-connected point is depicted in Figure 6B. Without any Q support, the voltage at the grid-connected point peaks at 0.2 pu and gradually increases to 0.25 pu when the SMES is controlled by MPC (Model Predictive Control).

Figure 6C presents the behavior of the DC-link voltage. Without Q support, the DC-link voltage exhibits uncontrolled growth, exceeding the fault period, resulting in a 5,450 V increase over a time period of 3.5s, surpassing the fault duration. However, when MPC-controlled SMES is activated, the DC-link voltage is effectively

TABLE 3 DC-DC chopper Switching States.

G_1	G_2	Decisions
0	0	Discharge (V_{dc})
0	1	Freewheeling (0)
1	0	Freewheeling (0)
1	1	Charge ($-V_{dc}$)

regulated to 1,200 V, enabling the wind turbine to ride through the fault within the fault period.

Figure 6D depicts the active power management profile under two scenarios: without Q support and with SMES + MPC schemes. Without Q support, the grid's active power decreases from its nominal value of 1.5 MW to approximately 0.365 MW until t = 3.47s, before returning to its nominal value. When SMES regulated by MPC is engaged, the grid-side active power is cushioned at around 0.736 MW until t = 3.15s, before returning to its nominal value.

Similarly, Figure 6E shows the behavior of the grid's reactive power. Without any Q support mechanism, there is no grid reactive power support during the fault. However, with the MPC + SMES

technique, the reactive power reaches 0.35 MVar to facilitate PCC (Point of Common Coupling) voltage recovery. As depicted in Figure 6F, the SMES's active current is injected into the grid during this recovery period. Following the elimination of the grid fault, the SMES current, controlled by MPC, is increased to 1,750 A, ensuring that these operations remain below the SMES critical current threshold.

Table 4 provides a summary of the advantages of MPC in the proposed technique, as detailed in reference (Aimene et al., 2022). The comparison includes the percentage change in the DC-link voltage (Δ V_dc) and the control settling time (s). Table 4 demonstrates that the MPC + SMES technique, owing to its capability to handle control uncertainties, facilitates a reasonable system recovery.

5.2 Case 2: Behaviours under 20% symmetrical voltage swell

In this experimental study, a balanced voltage swell of 0.20 pu is applied to the grid at t = 3s and lasts for 150 m. Figure 7A illustrates the 0.20 pu voltage swell at the grid-connected point during the fault. The root mean square value of the voltage swell at the grid-connected point is depicted in Figure 7B.
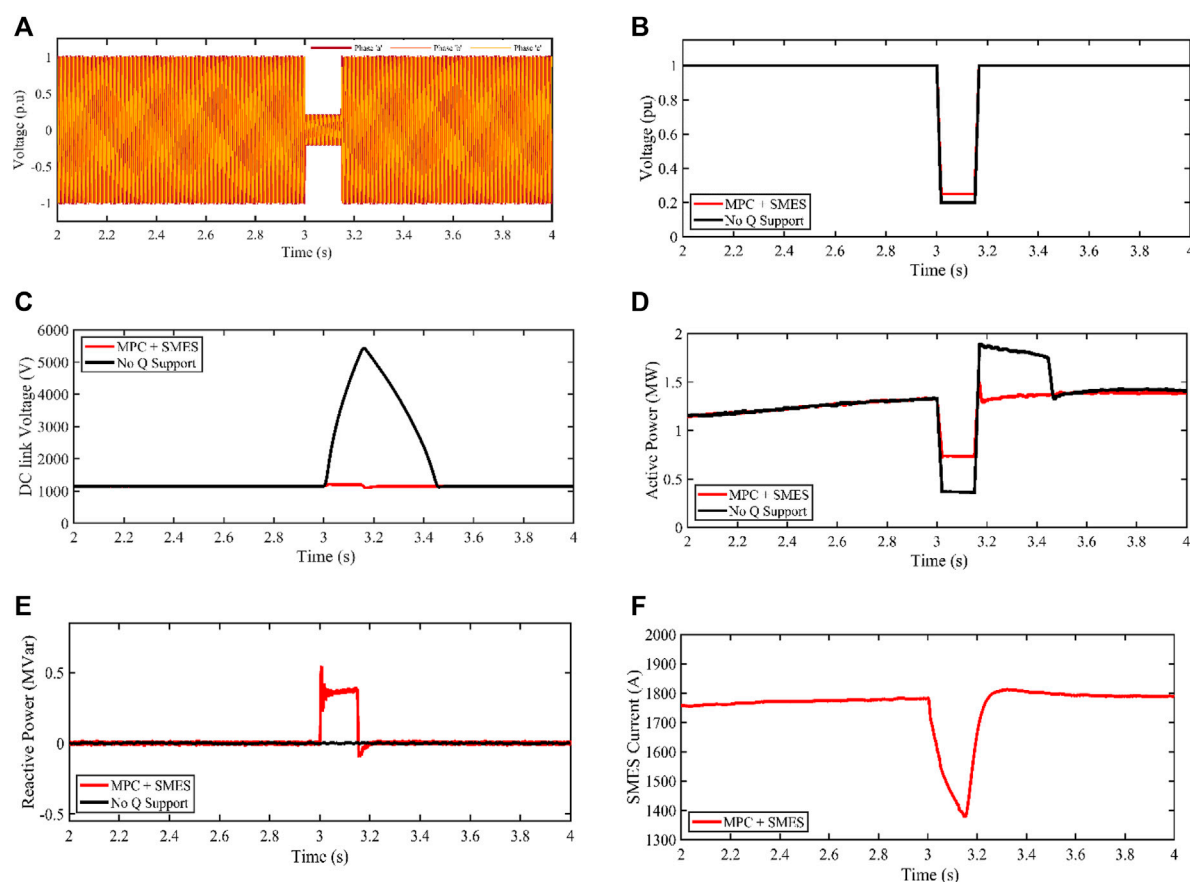

FIGURE 6
PMA wind generator response for 0.80pu balanced voltage sag **(A)** Phase abc of PCC voltage profile **(B)** RMS of Terminal voltage **(C)** DC-link voltage **(D)** Active power **(E)** Reactive power **(F)** SMES current profile.

METHODS	SETTLING TIME (S)	
	ΔVDC%	VDC
MPC + SMES	**4.348**	**3.2**
MAHM OUD et al. (2020); Aimene et al. (2022)	13.043	3.6

Bold is our achieved results.

Methods	Settling time (s)	
	ΔVdc%	Vdc
MPC + SMES	**8.695**[a]	**3.4**
Yuan et al. (2020); Mahmoud et al. (2022)	15.25	3.52

[a]Bold signifies best values.
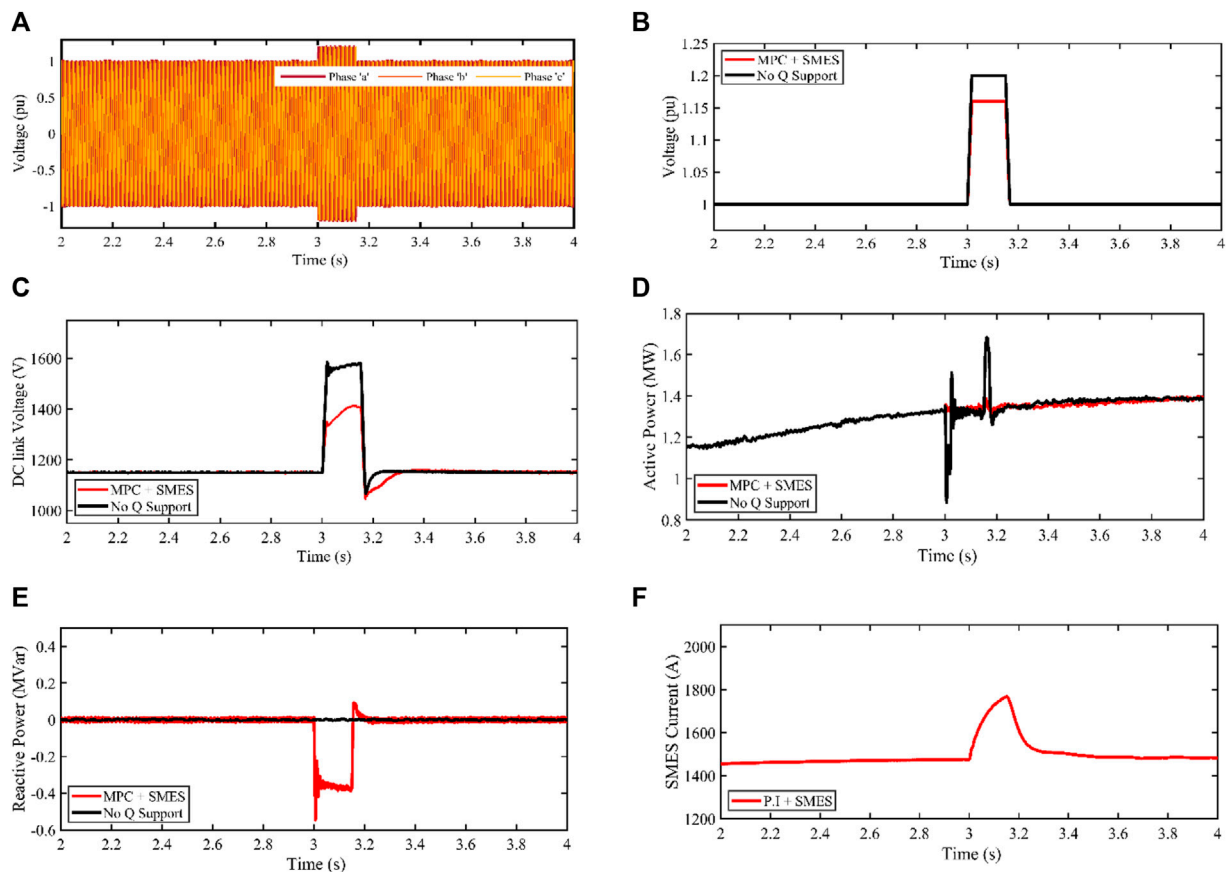
FIGURE 7
PMA wind generator response for 0.20 pu balanced voltage swell **(A)** Phase abc of PCC voltage profile **(B)** RMS Terminal voltage **(C)** DC-link voltage **(D)** Active power **(E)** Reactive power **(F)** SMES current profile.

Without any Q support, the voltage at the grid-connected point peaks at 1.2 pu and gradually decreases to 1.16 pu when the SMES is controlled by MPC (Model Predictive Control). The DC-link behavior is shown in Figure 7C. If no Q support mechanism is introduced, the DC-link voltage would increase uncontrollably, resulting in a 1,600 V rise in the DC-link voltage. However, when MPC-controlled SMES is activated, the DC-link voltage decreases to about 1,250 V, enabling the wind turbine to ride through the fault within the fault period.

Figure 7D depicts the active power management profile under two scenarios: without Q support and with MPC + SMES schemes. Without Q support, the grid's active power decreases from its nominal value of 1.5 MW to around 1 MW until t = 3.2s, before returning to its nominal value. When SMES regulated by MPC is engaged, the grid-side active power is cushioned at about 1.3 MW for a duration of t = 3.15s before returning to its nominal value.

Similarly, Figure 7E shows the behavior of the grid's reactive power. Without any Q support mechanism, there is no grid reactive power support during faults. However, with the MPC-SMES technique, the reactive power reaches a noticeable rise of −0.4 MVar to facilitate PCC (Point of Common Coupling) voltage recovery.

As shown in Figure 7F, during the voltage swell at t = 3s, the SMES coil increases, leading to a positive slope with MPC. Energy is transferred from the grid to the SMES coil during this recovery period. Following the elimination of the grid fault, the SMES current returns steadily to 1,500 A, which is well below its critical current threshold.

Table 5 provides a summary of MPC's advantages over other techniques proposed in the literature. The comparison includes the percentage change in the DC-link voltage ((ΔV_{dc}) and the control settling time (s). From Table 5, it can be seen that the MPC-SMES technique, due to its capability to handle control uncertainties, facilitates a reasonable system recovery.

6 Conclusion

In this paper, a grid-tied PMA wind turbine with SMES connected at the dc-link. The model incorporates a wind turbine with a capacity of 1.5 MW, an SMES device, and a machine and grid-side converter.

The WECS performance was evaluated under balanced voltage sag and swell scenarios.

The 80% voltage dip experiences an overvoltage of up to 5450 V on the DC-link, while the 0.20% swell reaches 1000 V from the reference value, indicating how detrimental sags are to the dc-link capacitor.

This means that PMA wind turbines are at a higher risk from balanced voltage sags compared to balanced voltage swells. Reactive power is managed via a strategy that regulates its transfer between the grid and the converter. Reactive power flows from the grid to the converter during a swell and the other way around during a sag. This contributes to the reliability of the power system.

There is clear evidence that the suggested ride through technique improves the performance of PMA-WTG in the face of grid voltage fluctuations.

Data availability statement

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

Author contributions

Funding

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

References

Abdel Aleem, S. H. E., Abdelaziz, A. Y., and Zobaa, A. F. (2017). "Egyptian grid code of wind farms and power quality," in *Handbook of distributed generation: electric power technologies, economics and environmental impacts* (Berlin, Germany: Springer), 227–245. doi:10.1007/978-3-319-51343-0_7

Abdel-Rahim, O., Chub, A., Vinnikov, D., and Blinov, A. (2022). DC integration of residential photovoltaic systems: a survey. *IEEE Access* 10, 66974–66991. doi:10.1109/ACCESS.2022.3185788

Abdel-Rahim, O., and Funato, H. (2014). "A novel model predictive control for high gain switched inductor power conditioning system for photovoltaic applications," in Proceedings of the 2014 IEEE Innovative Smart Grid Technologies - Asia (ISGT ASIA), Kuala Lumpur, Malaysia, May, 2014, 170–174. doi:10.1109/ISGT-Asia.2014.6873784

Abdel-Rahim, O., Funato, H., Abu-Rub, H., and Ellabban, O. (2014). "Multiphase Wind Energy generation with direct matrix converter," in Proceedings of the 2014 IEEE International Conference on Industrial Technology (ICIT), Busan, Korea (South), February, 2014, 519–523. doi:10.1109/ICIT.2014.6894994

Abhinav, R., and Pindoriya, N. M. (2016). "Grid integration of wind turbine and battery energy storage system: review and key challenges," in Proceedings of the 2016 IEEE 6th International Conference on Power Systems, ICPS, New Delhi, India, March. 2016. doi:10.1109/ICPES.2016.7583998

Ahsan, H., and Mufti, M. ud D. (2020). Systematic development and application of a fuzzy logic equipped generic energy storage system for dynamic stability reinforcement. *Int. J. Energy Res.* 44 (11), 8974–8987. doi:10.1002/ER.5606

Aimene, M., Payman, A., and Dakyo, B. (2022). Comparative study between flatness-based and field-oriented control methods of a grid-connected wind energy conversion system. *Process. 2022* 10 (2), 378. doi:10.3390/PR10020378

Barra, P. H. A., de Carvalho, W. C., Menezes, T. S., Fernandes, R. A. S., and Coury, D. v. (2021). A review on wind power smoothing using high-power energy storage systems. *Renew. Sustain. Energy Rev.* 137, 110455. doi:10.1016/J.RSER.2020.110455

Benali, A., Khiat, M., Allaoui, T., and Denaï, M. (2018). Power quality improvement and low voltage ride through capability in hybrid wind-pv farms grid-connected using dynamic voltage restorer. *IEEE Access* 6, 68634–68648. doi:10.1109/ACCESS.2018.2878493

Chen, T., Abdel-Rahim, O., Peng, F., and Wang, H. (2020). An improved finite control set-MPC-based power sharing control strategy for islanded AC microgrids. *IEEE Access* 8, 52676–52686. doi:10.1109/ACCESS.2020.2980860

Dahiya, P., Sharma, V., and Naresh, R. (2019). Optimal sliding mode control for frequency regulation in deregulated power systems with DFIG-based wind turbine and TCSC–SMES. *Neural Comput. Appl.* 31 (7), 3039–3056. doi:10.1007/s00521-017-3250-y

Das, C. K., Bass, O., Kothapalli, G., Mahmoud, T. S., and Habibi, D. (2018). Overview of energy storage systems in distribution networks: placement, sizing, operation, and power quality. *Renew. Sustain. Energy Rev.* 91, 1205–1230. doi:10.1016/J.RSER.2018.03.068

De Siqueira, L. M. S., and Peng, W. (2021). Control strategy to smooth wind power output using battery energy storage system: a review. *J. Energy Storage* 35, 102252. doi:10.1016/J.EST.2021.102252

Djagarov, N., Grozde, Z., Enchev, G., and Djagarova, J. (2019). "Mathematical model for study of low-voltage ride through of wind permanent magnet synchronous generator by means of STATCOM," in Proceedings of the 2019 20th International Scientific Conference on Electric Power Engineering, EPE 2019, Kouty nad Desnou, Czech Republic, May 2019. doi:10.1109/EPE.2019.8778121

Global Wind Energy Council, (2023). Global wind report 2023. Available: https://gwec.net/wp-content/uploads/2023/03/GWR-2023_interactive.pdf.

Gouda, E. A., Abd-Alaziz, A., and El-Saadawi, M. (2020). Design modeling, and control of multi-stage SMES integrated with PV system. *J. Energy Storage* 29, 101399. doi:10.1016/j.est.2020.101399

Hu, Y. L., Wu, Y. K., Chen, C. K., Wang, C. H., Chen, W. T., and Il Cho, L. (2017). A review of the low-voltage ride-through capability of wind power generators. *Energy Procedia* 141, 378–382. doi:10.1016/J.EGYPRO.2017.11.046

Huang, C., Zheng, Z., Xiao, X., and Chen, X. (2020b). Cooperative strategy of SMES device and modified control for cost-effective fault ride-through enhancement and power smoothing of 10 MW class superconducting wind turbine. *J. Renew. Sustain. Energy* 12 (3), 033302. doi:10.1063/1.5143565

Huang, C., Zheng, Z., Xiao, X., and Chen, X. (2020a). Enhancing low-voltage ride-through capability of PMSG based on cost-effective fault current limiter and modified WTG control. *Electr. Power Syst. Res.* 185, 106358. doi:10.1016/J.EPSR.2020.106358

Jannati, M., Hosseinian, S. H., Vahidi, B., and jie Li, G. (2016). ADALINE (ADAptive Linear NEuron)-based coordinated control for wind power fluctuations smoothing with reduced BESS (battery energy storage system) capacity. *Energy* 101, 1–8. doi:10.1016/j.energy.2016.01.100

Jiang, P., van Fan, Y., and Klemeš, J. J. (2021). Impacts of COVID-19 on energy demand and consumption: challenges, lessons and emerging opportunities. *Appl. Energy* 285, 116441. doi:10.1016/J.APENERGY.2021.116441

Joseph, S. B., Dada, E. G., Abidemi, A., Oyewola, D. O., and Khammas, B. M. (2022). Metaheuristic algorithms for PID controller parameters tuning: review, approaches and open problems. *Heliyon* 8 (5), e09399. doi:10.1016/J.HELIYON.2022.E09399

Kim, C., and Kim, W. (2021). Low-voltage ride-through coordinated control for pmsg wind turbines using de-loaded operation. *IEEE Access* 9, 66599–66606. doi:10.1109/ACCESS.2021.3076787

López, J., Gubía, E., Olea, E., Ruiz, J., and Marroyo, L. (2009). Ride through of wind turbines with doubly fed induction generator under symmetrical voltage dips. *IEEE Trans. Industrial Electron.* 56 (10), 4246–4254. doi:10.1109/TIE.2009.2028447

Luo, X., Wang, J., Wojcik, J., Wang, J., and Draganescu, M., (2018). Review of voltage and frequency grid code specifications for electrical energy storage applications. *Energies* 11 (5), 1070. doi:10.3390/EN11051070

Lyu, J., Ma, B., Yan, H., Ji, Z., and Ding, J. (2020). A modified finite control set model predictive control for 3L−NPC Grid−Connected inverters using virtual voltage vectors. *J. Electr. Eng. Technol.* 15 (1), 121–133. doi:10.1007/s42835-019-00305-8

Mahmoud, M. M., Aly, M. M., Salama, H. S., and Abdel-Rahim, A.-M. M. (2022). An internal parallel capacitor control strategy for DC-link voltage stabilization of PMSG-based wind turbine under various fault conditions. *Wind Eng.* 46 (3), 983–992. doi:10.1177/0309524X211060684

Makhad, M., Zazi, K., Zazi, M., and Loulijat, A. (2022). Adaptive super-twisting terminal sliding mode control and LVRT capability for switched reluctance generator based wind energy conversion system. *Int. J. Electr. Power & Energy Syst.* 141, 108142. doi:10.1016/J.IJEPES.2022.108142

Morgan, E. F., Abdel-Rahim, O., Megahed, T. F., Suehiro, J., and Abdelkader, S. M. (2022a). Fault ride-through techniques for permanent magnet synchronous generator wind turbines (PMSG-WTGs): a systematic literature review. *Energies* 15 (23), 9116. doi:10.3390/EN15239116

Morgan, E. F., Megahed, T. F., Suehiro, J., and Abdelkader, S. M. (2022b). A Fault Ride-Through technique for PMSG wind turbines using superconducting magnetic energy storage (SMES) under grid voltage sag conditions. *Renew. Energy Power Qual. J.* 20, 79–83. doi:10.24084/REPQJ20.223

Mukherjee, P., and Rao, V. v. (2019). Superconducting magnetic energy storage for stabilizing grid integrated with wind power generation systems. *J. Mod. Power Syst. Clean Energy* 7 (2), 400–411. doi:10.1007/s40565-018-0460-y

Nasiri, M., and Arzani, A. (2022). Robust control scheme for the braking chopper of PMSG-based wind turbines–A comparative assessment. *Int. J. Electr. Power & Energy Syst.* 134, 107322. doi:10.1016/J.IJEPES.2021.107322

Nielsen, K. E., and Molinas, M. (2010). "Superconducting Magnetic Energy Storage (SMES) in power systems with renewable energy sources," in Proceedings of the 2010 IEEE International Symposium on Industrial Electronics, Bari, Italy, July, 2010, 2487–2492. doi:10.1109/ISIE.2010.5637892

Nikolaidis, P., and Poullikkas, A. (2018). Cost metrics of electrical energy storage technologies in potential power system operations. *Sustain. Energy Technol. Assessments* 25, 43–59. doi:10.1016/J.SETA.2017.12.001

Okedu, K. E. (2022). Improving the performance of pmsg wind turbines during grid fault considering different strategies of fault current limiters. *Front. Energy Res.* 10. doi:10.3389/fenrg.2022.909044

Qais, M. H., Hasanien, H. M., Alghuwainem, S., and Elgendy, M. A. (2019). "Output power smoothing of grid-tied pmsg-based variable speed wind turbine using optimal controlled SMES," in Proceedings of the 2019 54th International Universities Power Engineering Conference, Bucharest, Romania, September. 2019. doi:10.1109/UPEC.2019.8893530

Qais, M. H., Hasanien, H. M., and Alghuwainem, S. (2020a). Output power smoothing of wind power plants using self-tuned controlled SMES units. *Electr. Power Syst. Res.* 178, 106056. doi:10.1016/J.EPSR.2019.106056

Qais, M. H., Hasanien, H. M., and Alghuwainem, S. (2020b). Whale optimization algorithm-based Sugeno fuzzy logic controller for fault ride-through improvement of grid-connected variable speed wind generators. *Eng. Appl. Artif. Intell.* 87, 103328. doi:10.1016/j.engappai.2019.103328

Raouf, A., Tawfiq, K. B., Tag Eldin, E., Youssef, H., and E El-Kholy, E. (2023). Wind energy conversion systems based on a synchronous generator: comparative review of control methods and performance. *Energies* 16 (5), 2147. doi:10.3390/en16052147

Sarkar, M. N. I., Meegahapola, L. G., and Datta, M. (2018). Reactive power management in renewable rich power grids: a review of grid-codes, renewable generators, support devices, control strategies and optimization Algorithms. *IEEE Access* 6, 41458–41489. doi:10.1109/ACCESS.2018.2838563

Soliman, M. A., Hasanien, H. M., Al-Durra, A., and Alsaidan, I. (2020). A novel adaptive control method for performance enhancement of grid-connected variable-speed wind generators. *IEEE Access* 8, 82617–82629. doi:10.1109/ACCESS.2020.2991689

Wang, D., Gao, X., Meng, K., Qiu, J., Lai, L. L., and Gao, S. (2018). Utilisation of kinetic energy from wind turbine for grid connections: a review paper. *IET Renew. Power Gener.* 12 (6), 615–624. doi:10.1049/iet-rpg.2017.0590

Xu, L., Lin, R., Ding, L., Zhang, H., Li, S., and Huang, C. (2019). A new frt method of pmsg under grid faults by using improved msc control and smes device. *IOP Conf. Ser. Mater Sci. Eng.* 490 (7), 072032. doi:10.1088/1757-899X/490/7/072032

Yuan, L., Meng, K., Huang, J., Dong, Z. Y., Zhang, W., and Xie, X. (2020). Development of HVRT and LVRT control strategy for PMSG-based wind turbine generators. *Energies* 13, 5442–5520. doi:10.3390/en13205442

Zheng, Z. X., Chen, X. Y., Xiao, X. Y., and Huang, C. J. (2017). Design and evaluation of a mini-size smes magnet for hybrid energy storage application in a kw-class dynamic voltage restorer. *IEEE Trans. Appl. Supercond.* 27 (7), 1–11. doi:10.1109/TASC.2017.2748954

Fault detection of a wind turbine generator bearing using interpretable machine learning

Oliver Trygve Bindingsbø[1], Maneesh Singh[1]*, Knut Øvsthus[1] and Arvind Keprate[2]

[1]Department of Mechanical and Marine Engineering, Western Norway University of Applied Sciences, Bergen, Norway, [2]Department of Mechanical, Electrical and Chemical Engineering, Oslo Metropolitan University, Oslo, Norway

Introduction: During its operational lifetime, a wind turbine is subjected to a number of degradation mechanisms. If left unattended, the degradation of components will result in its suboptimal performance and eventual failure. Hence, to mitigate the risk of failures, it is imperative that the wind turbine be regularly monitored, inspected, and optimally maintained. Offshore wind turbines are normally inspected and maintained at fixed intervals (generally 6-month intervals) and the program (list of tasks) is prepared using experience or risk-reliability analysis, like Risk-based inspection (RBI) and Reliability-centered maintenance (RCM). This time-based maintenance program can be improved upon by incorporating results from condition monitoring involving data collection using sensors and fault detection using data analytics. In order to properly carry out condition assessment, it is important to assure quality & quantity of data and to use correct procedures for interpretation of data for fault detection. This paper discusses the work carried out to develop a machine learning based methodology for detecting faults in a wind turbine generator bearing. Explanation of the working of the machine learning model has also been discussed in detail.

Methods: The methodology includes application of machine learning model using SCADA data for predicting operating temperature of a healthy bearing; and then comparing the predicted bearing temperature against the actual bearing temperature.

Results: Consistent abnormal differences between predicted and actual temperatures may be attributed to the degradation and presence of a fault in the bearing.

Discussion: This fault detection can then be used for rescheduling the maintenance tasks. The working of this methodology is discussed in detail using a case study.

KEYWORDS

bearing, condition monitoring, fault detection, machine learning, offshore wind turbine, SCADA, SHAP

1 Introduction

In order to meet the increasing demand for energy and yet reduce dependency on conventional fossil fuels, there has been a spurt in growth of wind farms (IEA, 2021). These wind farms are comprised of arrays of wind turbines (typically horizontal), installed either onshore or offshore, to produce electricity from the wind. However, despite recent advances

in the design, manufacturing, operation and maintenance of wind turbines, their acceptance has been muted due to a number of reasons, including difficulties and high costs associated with their operation and maintenance.

When compared to the onshore wind turbines, the offshore counterparts offer more reliable power generation due to higher mean wind speeds and more steady wind supply. Unfortunately, the operation and maintenance difficulties and costs are also higher due to multiple reasons, including faster degradation of equipment by harsh marine conditions, difficulties in accessing the site from distant shores, rough weather conditions, scarcity of skilled personnel and need for specialized vessels. Thus, the operation and maintenance costs account for approximately a third of the Levelized Cost of Energy (LCOE) (Wiggelinkhuizen et al., 2007; Stehly et al., 2020).

During their operational lifetime, various components of a wind turbine are subjected to a number of environmental & operational attacks resulting in their degradation. This degradation results in deterioration in performance and at times failure. Failure of a component takes place when the applied load is greater than the maximum safe working load of the component. The applied load and maximum safe working load of the component vary with time. The applied load can vary due to the changes in the operating conditions, environmental conditions or accident; and the maximum safe working load may change with time due to degradation caused to the component by different types of degradation mechanisms. Hence, it becomes difficult to predict when the failure will take place (Arabian-Hoseynabadi, et al., 2010; Kahrobaee and Asgarpoor 2011; Shafiee and Dinmohammadi 2014; Luengo and Kolios 2015; Zhang et al., 2016).

To help in predicting the time of failure, detailed failure analysis involving the following stages needs to be carried out (Kandukuri et al., 2016):

- *Fault Detection*—detection of abnormal changes in the structure or behavior of a component that can help to identify faulty condition
- *Fault Diagnosis*—analysis of the abnormal changes in the structure or behavior to identify cause or mechanism of the degradation that would cause the failure
- *Fault Quantification*—analysis of the behavior or performance to quantify the degree of degradation and fault (partial or complete)
- *Fault Prognosis*—analysis of the time-based changes to predict the outcome of further degradation or prognosis of fault

Failure (or fault) analysis can be used to develop detailed failure profiles (failure causes, failure mechanisms, etc.), which can subsequently be used for developing an appropriate maintenance schedule to prevent or manage the failure. In a maintenance schedule, the maintenance activities can be either preventive or corrective depending on whether the task is carried out before or after failure. These maintenance activities involve detailed inspection (visual, auditory, NDT), testing, service (lubrication, cleaning, repair, etc.), repair and replacement tasks.

The preventive maintenance programs are often time-based, for example, preventive maintenance activities of wind turbines are normally planned at 6-month intervals (Nilsson and Bertling, 2007). Since these time-based inspection and maintenance programs are expensive, there have been efforts to develop

methodologies for preparing more efficient and effective maintenance programs. This involves development of maintenance schedules based on formalized risk/reliability analysis (e.g., Risk Based Inspection and Maintenance or Reliability Centered Maintenance).

In order to improve the technical asset integrity management of wind farms there is an increasing move towards condition-based maintenance as opposed to scheduled or reactive maintenance to reduce downtime and lost production. This is achieved by a) continuous monitoring using sensors; b) data analytics; and c) developing condition-based maintenance plans.

To continuously monitor, all modern wind turbines come with a Supervisory Control and Data Acquisition (SCADA) system. This system is comprised of a multitude of sensors that constantly monitor various parameters regarding environment, process, operation, and condition of components (equipment or structure). The data from the sensors is transmitted and stored in SCADA supervisory computers. At the control office the data is interpreted, and the information gained is then used to control the process or operation. The same data can be used to develop optimized condition-based maintenance schedules.

While the collection, transmission and storage of data has become relatively easy in recent years, the challenge is to identify and extract relevant information from the available data. Thus, sensible data collection requires understanding the system, making decisions related to collection and rationalization of data to make it suitable for further analysis, and finally, to use the preprocessed data to extract useful information, like, fault detection and identification, so that necessary decisions can be taken. There are a number of approaches by which the data analysis can be carried out, to include machine learning, fuzzy logic, artificial neural networks, and deep learning.

Machine learning techniques have been widely explored for analyzing data from offshore wind turbines and these have been found to be suitable for detecting anomalies and assisting in decision-making (Stetco, et al., 2019). However, while machine learning models may have high prediction accuracy, they often lack interpretability. This is because models often act as black-boxes, thereby making their results challenging to understand and interpret, and users may not have knowledge of the underlying decisions in the predicting process (Ekanayake et al., 2022).

Interpretable machine learning tools can be applied to gain insight into the working of machine learning models. Thus, it is easier to understand the factors that drive their predictions and increase confidence in their predictions. This understanding may be used to justify the use of the model and to further improve its working (Adadi and Berrada, 2018). Interpretable machine learning is currently at a stage where it is sufficiently developed and mature, but there are still some challenges that need to be addressed (Mahesh, 2020; Vilone and Longo, 2020).

In recent years, the research community has become more interested in Shapley additive explanations (SHAP) method, which proposes a model agnostic representation of feature importance estimated by Shapley values in a computationally efficient manner. Shapley values are a solution concept from collaborative game theory. The SHAP method is an additive feature attribution method that considers the features as "the players", combinations of different features as "the coalitions," and the prediction as "the total payout." The average marginal contribution for feature i over all possible coalitions is the Shapley

value ϕ_i, hence it explains each feature's contribution to a prediction (Lundberg and Lee, 2017; Lundberg et al., 2018).

Besides SHAP there are other methods for interpreting machine learning results such as Individual Conditional Expectation (ICE) plots (Goldstein et al., 2015) and Local interpretable model-agnostic explanations (LIME) (Ribeiro et al., 2016). ICE plots visualize the dependence of model predictions on a feature for each instance separately. By varying the values of a feature for a particular instance while keeping the values of all other features fixed, it shows the relationship between the feature and the model's predictions across a range of values by repeating this process. Each line in the ICE plot represents the predicted outcome for a different instance, allowing us to see the individual effects of a feature on the model's predictions. LIME works by approximating the machine learning model locally around a specific instance, using a simpler, interpretable model. It perturbs the instance, creates a dataset, fits an interpretable model on the perturbed instances, and generates explanations based on the model's feature weights. These explanations help us understand why a particular prediction was made on a local level.

While ICE plots and LIME focus on local explanations for individual predictions, SHAP provides both model-agnostic and global explanations. SHAP values capture the contribution of each feature to a prediction across the entire dataset, allowing for a more comprehensive understanding of feature importance. Additionally, SHAP is applicable to a wide range of models and is able to handle feature interactions, thus providing a more nuanced understanding of how features interact to influence predictions. Based on these advantages, SHAP is selected as the best fitting interpretable machine learning method.

After the SCADA data has been analyzed using appropriate models, the results from the model have to be used to decide maintenance activities. These activities are triggered when some condition indicator crosses a preset limit. This guides the maintenance activities to take place based on the actual condition, as against faulty condition in corrective maintenance and perceived condition in preventive maintenance. Hence, condition-based maintenance strategy offers advantages that are associated with (Koukoura et al., 2021):

- maintenance activities carried out only when required, e.g., reduced human errors in maintenance
- not conducting unnecessary scheduled replacement of parts before their end of useful life, e.g., cost saving
- advanced planning of maintenance activities, e.g., better planning

In spite of these advantages, use of a condition-based maintenance approach is still restricted and needs further research and development. This is because of the difficulties associated with the:

- quality and quantity of collected data
- handling of imperfect (spurious, inconsistent, inaccurate, uncertain, or irrational) data collected from faulty sensors
- interpretation of data to information regarding failure profile
- reasoning of information into knowledge about the existing status of the equipment
- converting knowledge to decision regarding maintenance scheduling

- handling of unreliable analysis that may trigger false alarm (false positive) or failure to respond (false negative)

Hence, a solution that integrates the traditional (corrective and preventive) maintenance methods with condition-based maintenance methods may provide a solution that is robust, effective, and efficient. In this integrated method:

- the failure analysis is carried out in the traditional manner, and then the results of failure profile is used judiciously to develop a maintenance strategy;
- the time for inspection and maintenance of a component is adjusted based upon condition monitoring.

This paper discusses the work carried out to develop methodology for identifying faults in a wind turbine generator bearing using interpretable machine learning models and using the results for rescheduling of its maintenance time. The methodology includes preprocessing of data to remove outlier data, use of machine learning models to predict bearing temperature, identification of deviation between predicted and actual temperatures, critical analysis of results, and recommendations for rescheduling of maintenance tasks.

2 Proposed fault detection methodology

2.1 Description of the process

In order to develop an effective and efficient asset management program for a component, it is important to understand the process in terms of the structure, environment, and operation.

A wind-turbine contains 20 to 25 bearings, all of which must be considered in a system-level reliability calculation of life expectancy [wind power engineering]. A typical roller bearing consists of four components: a) inner ring, b) outer ring, c) cage, and d) rollers. During an operation, these components are subjected to different levels of dynamic and static loads, which can be in axial, radial or combination direction under constant or alternating conditions. These loads cause degradation of the material because of wear (contact wear—peeling, scoring, smearing, etc.), fatigue (contact fatigue—flaking, spalling, etc.), corrosion, electrical erosion, plastic deformation, and fracture and cracking (ISO, 2017), thereby resulting in the deterioration of the components and ultimately failure (Sankar et al., 2012). As the degradation progresses, it also results in changes in the behavior patterns of parameters like temperature, vibration, noise, rotational speed, etc. By monitoring these parameters using appropriate sensors, it may be possible to diagnose the health of the bearings. Commonly used parameters for identifying fault in a bearing include temperature, vibration, and noise.

2.2 Feature selection

As discussed in the previous section, temperature is a commonly measured parameter to monitor the health of a bearing, because it is
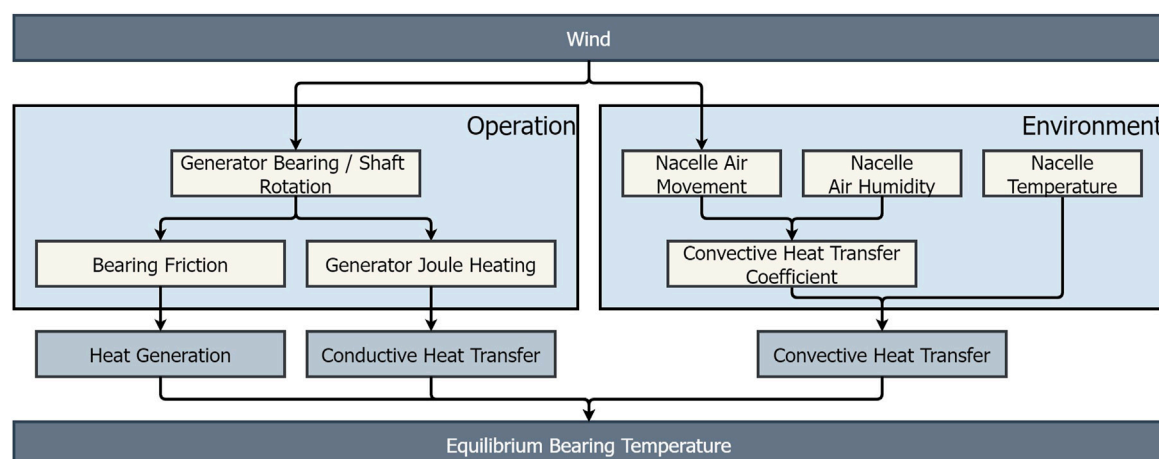
FIGURE 1
Flowchart showing the heat transfers taking place in bearings.

easy to continuously monitor and analyze in order to identify any abnormal behavior.

Figure 1 shows the simplified flowchart of heat transfers taking place in a bearing. A bearing is at a thermal equilibrium when it reaches a steady temperature. At this temperature, there is a balance between:

1. Heat generation due to bearing friction (rolling, sliding, etc.) and seal friction—During an operation, the friction among the components of a bearing results in generation of heat, the amount of which is dependent upon a number of factors, including the rotational speed, type of bearing, bearing geometry, elastic deformation under load of the rolling elements and raceways, type of lubricant and its application, and sliding friction between the components. The friction also results in its wear as a result of which there is an increase in bearing surface imperfections (deformation, pitting, craters, depressions, surface irregularities, spalling, cracking, etc.). The formation of surface imperfections leads to an increase in friction resulting in an increase in heat generation. Thus, an increase in friction due to structural imperfections or deterioration in lubrication increases the temperature of bearings.
2. Conductive heat transfer from or to the adjacent parts—Temperature of a bearing depends upon the heat input from or heat output to the adjacent parts. One piece of equipment that can significantly affect the bearing temperature is the generator itself. When the generator shaft rotates, heat is generated due to electrical resistance in the windings, resulting in heating of the generator. Since the temperature of the generator is higher than the temperature of the bearing, there is thus a heat transfer from generator to bearing. By measuring the temperature of the generator in stator windings, it may be possible to estimate the effect of the generator temperature on the temperature of the bearing.
3. Convective heat dissipation to environment—Temperature of a bearing in operation is generally above the environmental temperature, hence the bearing continuously dissipates heat to

the environment. The rate of convective heat transfer is a function of:

- Convective heat transfer coefficient—The convective heat transfer coefficient depends upon a number of parameters, including the air velocity over the solid surface and the specific heat capacity of humid air. The specific heat capacity of humid air is approximately proportional to the absolute humidity of air. Thus, as the humidity increases the value of convective heat transfer coefficient increases, resulting in an increase in heat loss (Boukhriss et al., 2013). Thus, the temperature of a bearing depends upon the speed of air circulation around it and the relative humidity of air.
- Temperature difference between the bearing and the environment—The rate of heat loss is proportional to the difference in the temperatures of the solid (bearing) and the environment. Thus, the temperature of the bearing depends upon the ambient temperature.

Based on the understanding of the heat transfers, five variables have been selected to predict the bearing temperature. These are:

1. Generator Shaft/Bearing Rotational Speed—This is the rotational speed of the high-speed shaft connected to the generator. The shaft is supported by the generator bearings, and thus rotation of the shaft leads to rotation of the bearing resulting in generation of heat in the bearings due to friction.
2. Generator Temperature—This measures the temperature of the generator stator windings. When the generator shaft rotates, heat is generated by electrical resistance in the windings. The windings are located close to the generator bearings and heat is transferred from the windings to the bearings.
3. Wind Speed—In a wind turbine, wind turns its rotor which in-turn rotates the shaft of the generator. Thus, wind speed determines the rotational speed of the generator shaft and bearing. Additionally, since the nacelle is not airtight, the wind speed impacts air movement inside the nacelle, which in turn influences the convective heat transfer rate.
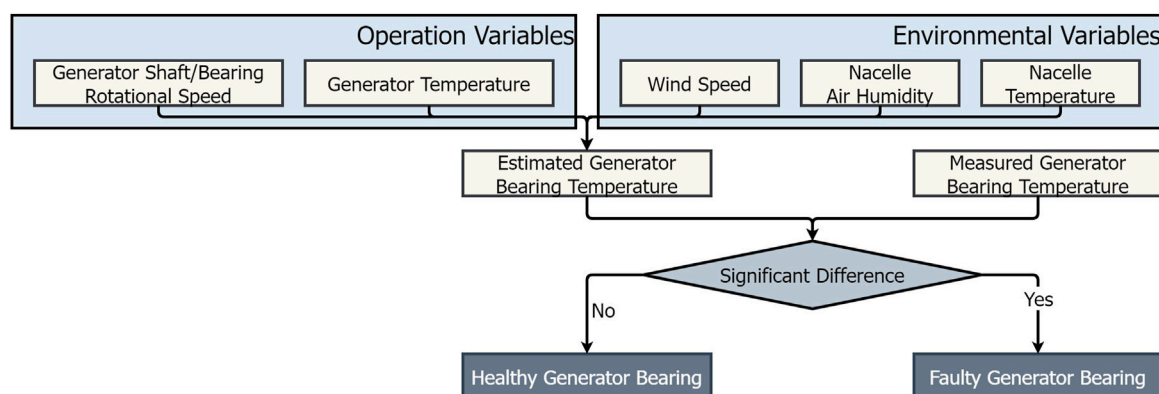
FIGURE 2
Flowchart showing the proposed fault detection methodology.
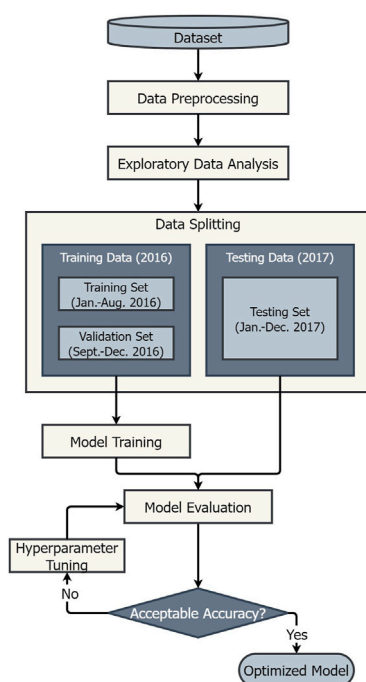
FIGURE 3
Flowchart for developing the proposed interpretable machine learning model.

4. Nacelle Air Humidity—This is the relative humidity of air inside the nacelle.
5. Nacelle Temperature—This is the temperature measured in the confined space housing the wind turbine drivetrain. The generator is located at the back of the nacelle and is therefore affected by the ambient temperature in the nacelle.

Figure 2 shows the flowchart of the methodology employed for detecting fault in a bearing. Using the five parameters, it may be possible to estimate temperature of a healthy bearing and if the measured temperature is above the predicted value, then there is a possibility that the higher temperature is the result of increased friction due to degradations in the bearing or lubrication.

2.3 Proposed model for predicting bearing temperature

As discussed in the previous section, the first step is to predict the bearing temperature using the five input variables. Figure 3 shows the flowchart of proposed methodology for predicting bearing temperature using machine learning algorithms.

2.3.1 Selection of regression algorithms

In this project a number of machine learning algorithms have been considered for developing a predictive model. These included:

- Linear Models—Linear Regression (LR), Lasso, Ridge, and Bayesian Ridge Regression
- Tree-based Models—Decision Trees, Random Forest (RF)
- Boosting Models—AdaBoost, XGBoost and LGBoost
- Support Vector Regression (SVR)

The short-listing of suitable algorithms have been carried out based on two key criteria.

- Firstly, the algorithms that demonstrate high compatibility with interpretable machine learning tools (example, SHAP) have been prioritized. This consideration is crucial as it ensures that the developed models are not just black boxes, rather their decision-making processes can be understood and explained. This aspect is particularly important for applications where transparency and trust in the model's predictions are paramount.
- Secondly, one representative algorithm from each of the aforementioned categories—linear models, tree-based models, boosting models, and support vector machines have been deliberately selected. This enables comparison regarding their behavior and strengths.

These selection criteria help to identify the most effective algorithm that not only delivers high accuracy but also aligns

with the interpretability and applicability requirements of our project. Thus, out of the above mentioned algorithms, four algorithms—Linear Regression (LR), Random Forest (RF), Support Vector Regression (SVR) and XGBoost—have been shortlisted for further testing.

2.3.2 Data preprocessing

Data preprocessing is an important step of any machine learning model. This is because raw data is typically created, processed, and stored by a mix of humans and business processes, often resulting in imperfections like vague, inconsistent, irrational, duplicate or missing values. These imperfections need to be corrected for the algorithms to work properly. Hence, an important step in preprocessing is to identify and handle (often remove) outliers. The outliers are removed only from the training and evaluation data so that the models can be trained and evaluated on healthy turbine operation data. This improves the models' capability to detect anomalies in the test data.

2.3.3 Exploratory data analysis (EDA)

Exploratory data analysis is used to analyze and investigate the data set and summarize the main characteristics by employing data visualization methods. Common methods include the use of Pearson, Kendall, or Spearman correlation metrics. These metrics depict the correlation between all the possible pairs of values and is a powerful tool to identify and visualize patterns in data.

2.3.4 Data splitting—training, validation and testing data

In supervised machine learning tasks, best practice is to split data into three independent data sets: training set, validation set and test set.

2.3.5 Model training

Model training is the process of teaching a machine learning model to make predictions or perform a specific task by exposing it to a labeled data set. The goal of model training is to enable the model to learn patterns, relationships, and rules from the training data so that it can generalize its knowledge to make accurate predictions on unseen or future data.

2.3.6 Model evaluation

In order to select the best performing algorithm out of the four, some criteria for evaluation need to be applied. These criteria should be able to judge a model's performance regarding a) accuracy of prediction, b) compatibility with interpretable machine learning tools, c) time usage for carrying out the calculations, and d) simplicity. The selection of the best model is based on an overall assessment of all the criteria.

To evaluate the accuracy of prediction, Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), and Coefficient of Determination (R^2) have been used.

2.3.7 Hyperparameter tuning

Many machine learning algorithms require hyperparameters that need to be defined before running them. First-level model parameters are decided during training, but the second-level tuning

parameters need to be tuned to optimize the performance. Typically, this is done by performing cross-validation or evaluating predictions on a separate test set (Probst et al., 2019).

In this analysis, hyperparameter tuning is performed using grid search (Bergstra and Bengio 2012) and hyperparameter values suggested by Probst, Boulesteix et al. (2019). This method runs through all possible combinations of the parameters within their search ranges forming a grid. It is performed using the scikit-learn library for python programming language. The grid search finally ranks all the combinations by their mean RMSE score across the same cross-validation folds used for model evaluation. Results from the grid search are used to select the optimal values for the hyperparameters.

Besides grid search there are additional hyperparameter tuning methods such as random search and Bayesian optimization. Grid search is selected due to its transparency and reproducibility, as well as its robustness against local optima. By evaluating all possible combinations, it reduces the risk of getting stuck in suboptimal regions of the hyperparameter space, and hence it increases the likelihood of finding the best set of hyperparameters for a given problem.

2.4 Model interpretation using SHAP

Once the model has been tuned using optimal hyperparameters, it is ready to be interpreted. SHAP has been used to interpret outputs of the best performing machine learning model and quantifying impact of each features to predictions. A negative SHAP value indicates a negative impact that decreases the value of the model output, whereas a positive SHAP value indicates a positive impact that increases the value of the model output. Although a SHAP analysis does not explicitly imply causalities, it helps in interpreting how each feature contributes to the model output and helps to identify importance of a feature in a model prediction.

3 Illustrative case study

3.1 SCADA data

To demonstrate the feasibility of the proposed methodology, SCADA data made available by the energy company EDP (2017) from four horizontal axis wind turbines located off the western coast of Africa has been used. The data has been recorded over a period of 2 years (2016 and 2017) at a 10-min averaging interval. The datasets contain values of 76 parameters. Besides this, associated datasets about meteorological conditions have also been provided for the same time instances. Failure logs containing timestamp, damaged component and associated remarks are also available. For this work, Turbine Number 7 ("T07") has been selected because its failure log has recorded generator bearing failure. For Turbine Number 7, the total number of instances are 52,445 and 52,294 for 2016 and 2017, respectively. Table 1 shows the selected features and target used for developing the model.

The generator uses two bearings, one on the drive-end and one on the driven end. The failure log records damage of generator bearings on 20 August 2017, at 08:08:00, and damage of generator shortly afterwards on 21 August 2017, at 16:47:00 (Table 2). The downtime caused by the generator failures is highlighted in green in
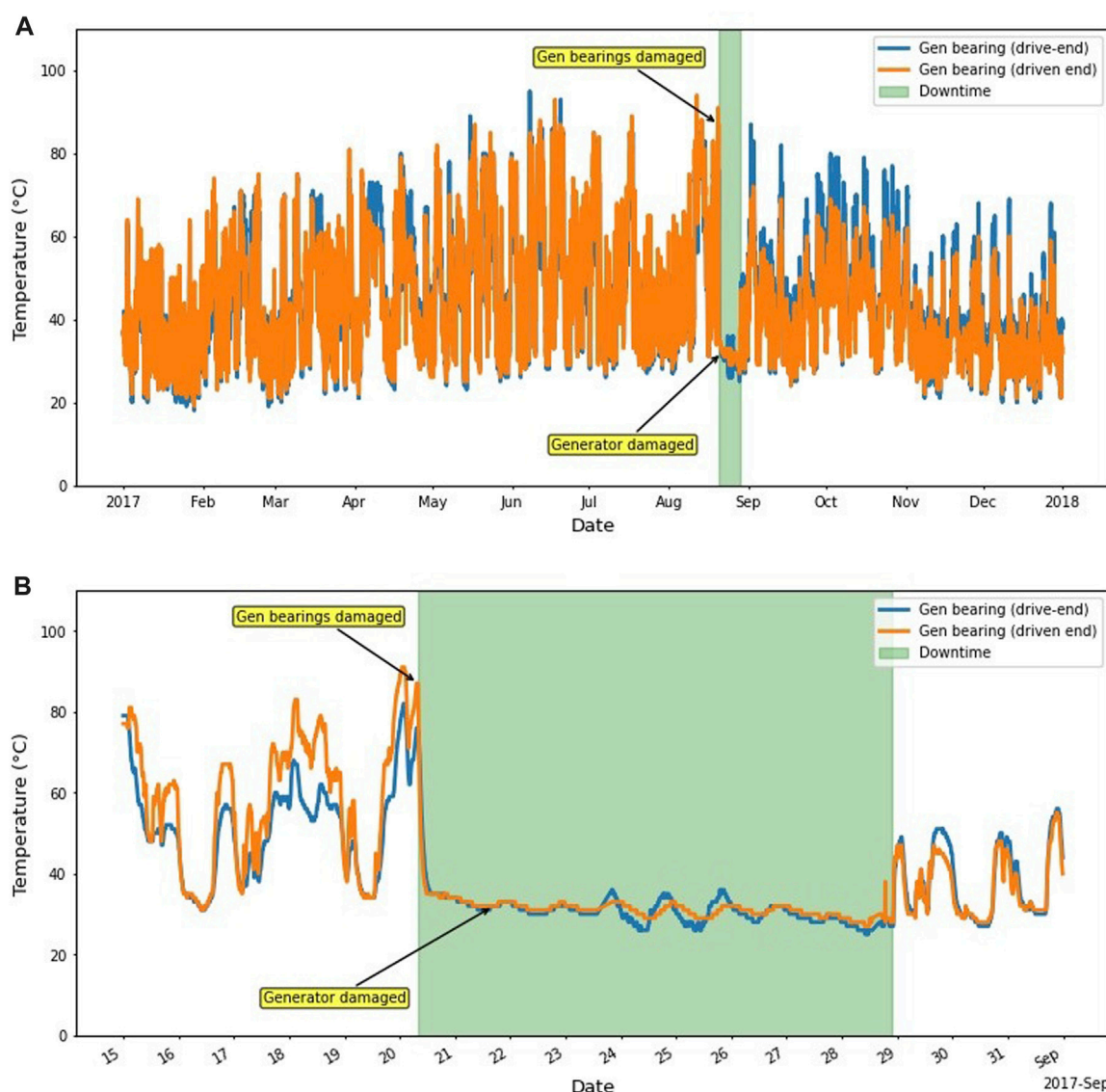
FIGURE 4
Bearings temperature during the bearing and generator failures in **(A)** 2017 and **(B)** August 2017.

Figure 4 and lasts from 20 August 2017, at 08:10:00 until 28 August 2017, at 21:50:00. The model shall attempt to predict these failures.

3.2 Data preprocessing

3.2.1 Identification of data outliers

Quite often SCADA data contains outliers that arise due to imperfections in the SCADA system and do not reflect the actual condition of process, environment, or component. For the development of a predictive model, it is important to remove these outliers because their presence can lead to biases in the model.

One common reason for outliers in the data is the inputs from faulty sensors. Since health prognosis of a bearing relies heavily on the data collected by the sensors, the reliability of analysis thus depends upon the reliability of the collected data. Hence, the

reliability of results from the proposed methodology also depends upon the quality of data used for the analysis.

Figure 5 shows plots of the temperature data versus selected periods of the two bearings. Sudden spike in the recoded temperatures can only be due to errors in the data collection, possibly arising due to the faulty sensor. This is justified by the record showing that the sensor was replaced on 30 April 2016 12:40 after recording *High temperature in generator bearing 1*. Outliers like those shown in the figure need to be handled during the data preprocessing.

In this model outliers have been identified by the use of box plots, shown in Figure 6. In a box plot, the lower limit of the whisker marks the minimum value, excluding outliers, whereas the upper limit of the whisker marks the maximum value, excluding outliers. The lower limit of the box is the first quartile (Q1 or the 25th percentile), whereas the upper point of the box is the third quartile (Q3 or the 75th percentile). All values within the box between

FIGURE 5
Effect of faulty sensors on recorded temperature of bearings.
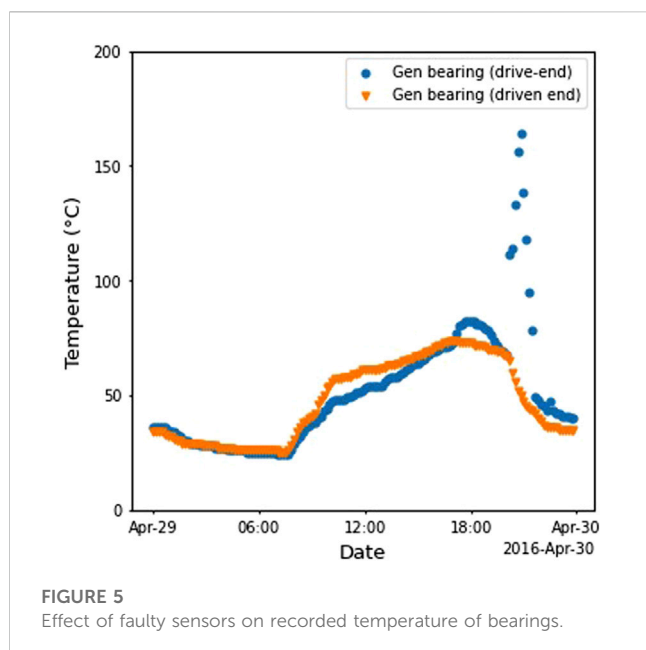
Q1 and Q3, also called the interquartile range (IQR), are calculated using Eq. 1. The horizontal red line in the box is the median value. An outlier in this case is defined as a value outside 1.5 times the IQR above Q3 or below Q1.

$$IQR = Q3 - Q1 \qquad (1)$$

where: IQR = Interquartile range $Q1$ = the first quartile, or the 25th percentile $Q3$ = the third quartile, or the 75th percentile

3.2.2 Data cleaning

Depending upon the characteristics of specific variables, rules for identification and handling of outliers have also been adopted.

For example, a threshold of 100°C has been set for the generator bearing temperature and all values higher than this have been removed. Similarly, relative humidity values are missing in the period 3 January 2017 to 6 May 2017, and this gap has been filled with values from the previous year.

Further cleaning has been performed using DBSCAN (Ester et al., 1996). DBSCAN is a density-based clustering algorithm that works on the assumption that clusters are dense regions in space separated by regions of lower density. "Densely clustered" data points are gathered into a single cluster.

The results before and after cleaning are shown in Figure 7. Figure 7A shows the presence of a significant number of outliers which indicate that either the turbine is not operating despite the blowing wind, or the sensors are not working properly. Additionally, there are many instances of the turbine not operating at its maximum potential. Figure 7B shows the plot after the removal of the most significant outliers and the remaining data points sufficiently fit the theoretical power curve.

3.3 Exploratory data analysis (EDA)

Figure 8 shows the Pearson correlation matrix of the input features and target. Some signals are highly correlated, for example a) wind speed and generator rotational speed, b) wind speed and generator phase temperature, and c) generator phase temperature and bearing temperature. The matrix shows that the selected features are significantly relevant to the target variable.

To further understand the correlation between the features and target, pairwise relationships between them in the training set have been plotted (Figure 9). The marginal histograms have been prepared by dividing signal values into 25 bins.

FIGURE 6
Box plot of SCADA signals.
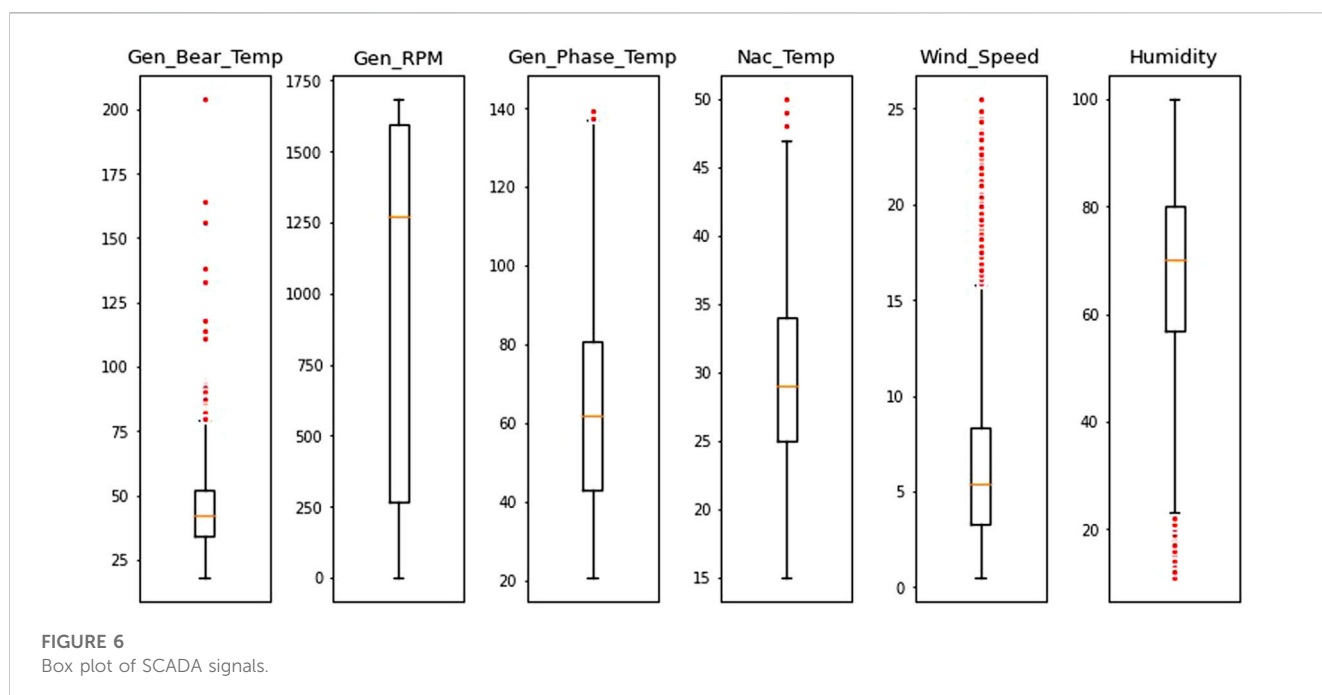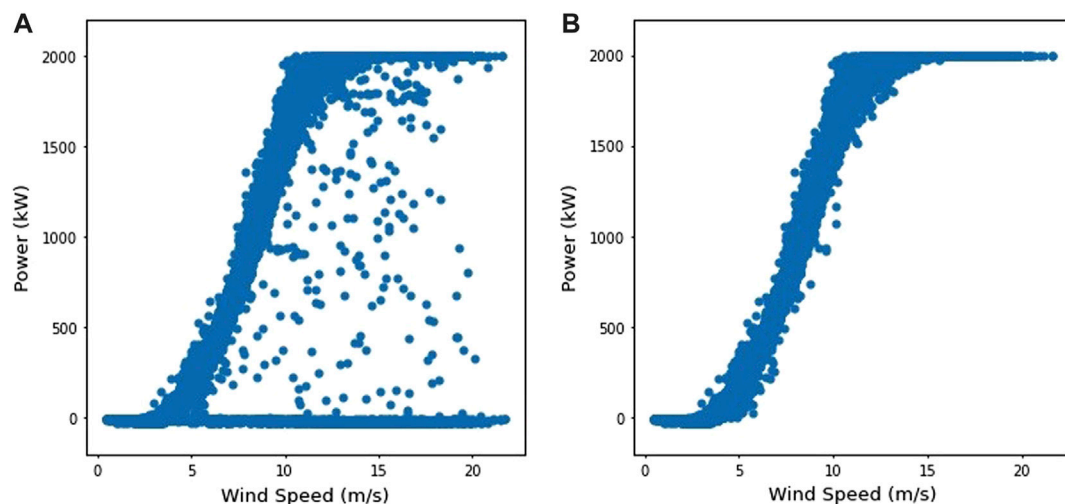
FIGURE 7
Plot of power generated versus wind speed using data of training period **(A)** Using raw. **(B)** Using data after cleaning outliers.
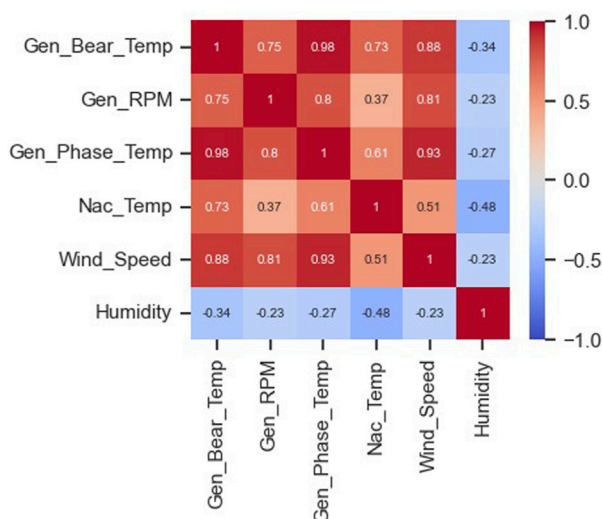
FIGURE 8
Pearson correlation matrix of the input features.

3.3.1 Effect of generator shaft/bearing rotational speed on bearing temperature

The time averaged wear rate of a bearing can be given as (Gupta, 2013):

$$W(T) = \frac{1}{T}\frac{K}{H}\int_o^T Q(t)u(t)dt \qquad (2)$$

where: W = Time-averaged wear rate over the time interval (T) K = Wear coefficient H = Hardness of the material being subjected to wear Q = The time-dependent load at a given interaction u = Sliding velocity as a function of time

The equation shows the dependence of wear on the parameters Q and u, which in turn are dependent upon the rotational speed.

Thus, the wear rate increases with an increase in the rotational speed. Corresponding to the increase in wear, the heat generated due to friction also increases with the increase in the rotational speed. This increase in heat generation manifests itself as an increase in the temperature.

Figure 9 shows the bearing temperature (*Gen_Bear_Temp*) is a function of the rotational speed of generator shaft/bearing (*Gen_RPM*).

3.3.2 Effect of generator temperature on bearing temperature

In a generator, heat is produced in the windings of the stators due to the passage of electricity through the electric wiring (Joule Heating). This heat is dissipated to the surrounding through conduction and convection. A part of dissipated heat also increases the temperature of the generator bearings.

Figure 9 shows the approximately linear relationship between the generator temperature (*Gen_Phase_Temp*) and the bearing temperature (*Gen_Bear_Temp*).

3.3.3 Effect of wind speed on bearing temperature

Wind speed has two opposing effects on the bearing temperature. On the one hand, an increase in wind speed increases the rotational speed of bearing resulting in increase in temperature due to friction. On the other hand, wind speed also increases air circulation within the nacelle, thereby increasing the convective heat transfer coefficient and subsequently heat loss from the bearing.

Figure 9 shows that there is a net increase in bearing temperature (*Gen_Bear_Temp*) with an increase in wind speed (*Wind_Speed*).

3.3.4 Effect of nacelle air humidity on bearing temperature

Since the specific heat capacity of humid air increases with an increase in the relative humidity of air, expectedly an increase in
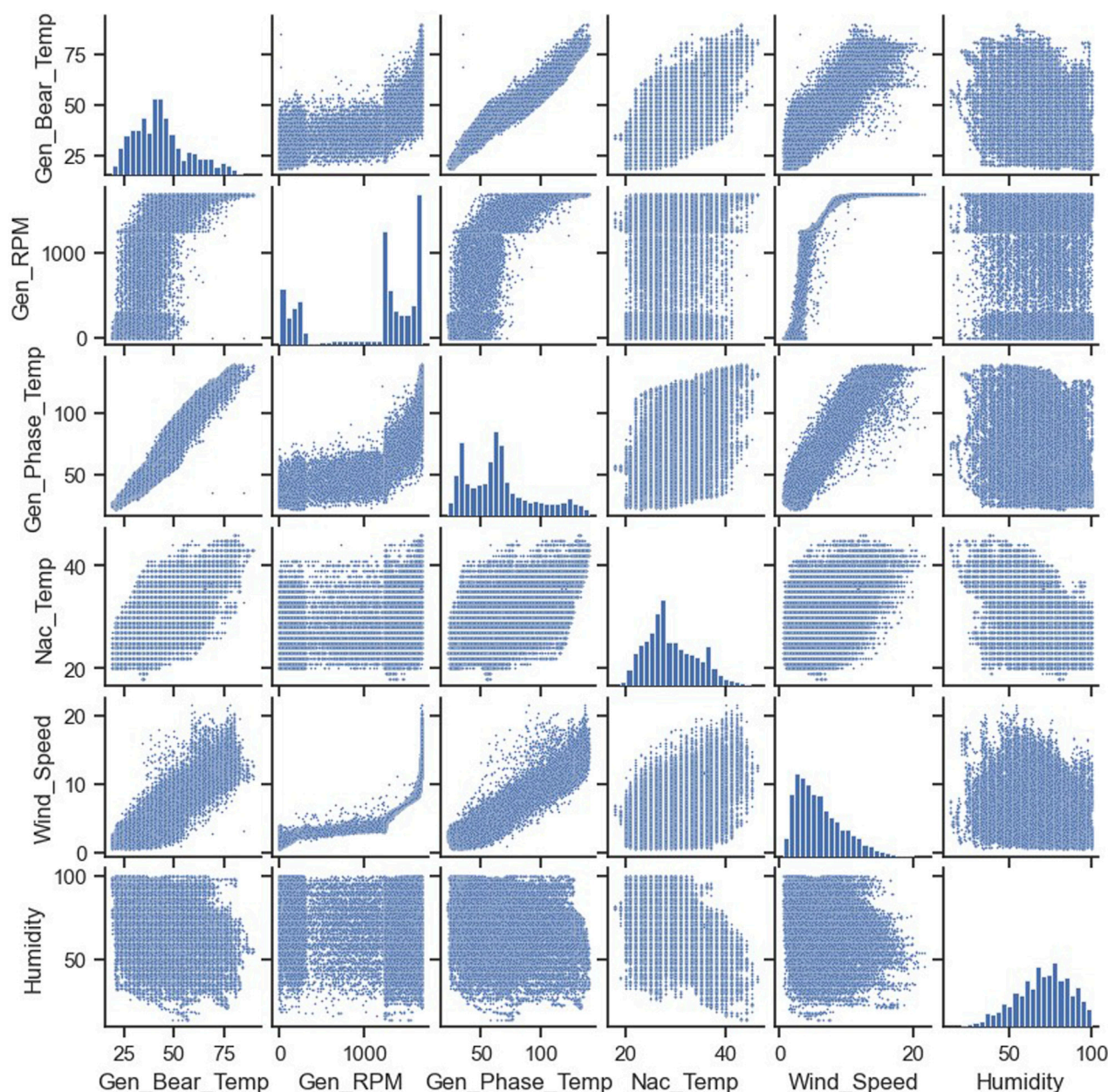
FIGURE 9
Pairwise relationships between input features.

relative humidity increases the convective heat transfer coefficient and subsequently increases heat loss from the bearing.

Figure 9 shows a weak correlation between the relative humidity of air (*Humidity*) and the bearing temperature (*Gen_Bear_Temp*).

3.3.5 Effect of nacelle temperature on bearing temperature

The ambient temperature in the nacelle follows an annual cycle, whereby the temperature is lower during winters and higher during summers. Since the convective heat transfer is proportional to the temperature difference between a bearing's surface temperature and the ambient temperature, this variation in the ambient temperature

has an effect on the heat dissipation from bearing to the environment.

Figure 9 shows an increase in the bearing temperature (*Gen_Bear_Temp*) with an increase in ambient temperature inside nacelle (*Nac_Temp*).

3.4 Data splitting—training, validation and test data

The data from 2016, after the removal of outliers, has been used for training the model in two steps. In the first step, the clean

TABLE 1 Selected features and target for developing the model.

Variable	Description	Units
Timestamp	10-min resolution	
Features		
Gen_RPM	Generator shaft/bearing rotational speed	rpm
Gen_Phase_Temp	SCADA dataset gives the average temperature inside generator in stator windings Phase 1, 2 and 3. Since the temperatures are nearly the same, Gen_Phase_Temp is an average temperature of the three temperatures	˚C
Wind_Speed	Ambient wind speed	m/s
Humidity	Relative nacelle air humidity	%
Nac_Temp	Nacelle temperature	˚C
Target		
Gen_Bear_Temp	Temperature in generator bearing 1 (Driven End)	˚C

TABLE 2 Failure log for Turbine Number 7 ("T07").

Timestamp	Component	Remarks
20 August 2017, 08:08:00	Generator bearing	Generator bearings damaged
21 August 2017, 16:47:00	Generator	Generator damaged

TABLE 3 Cross validation RMSE scores.

Model	Fold 0	Fold 1	Fold 2	Fold 3	Mean
LR	1.61	1.74	1.62	1.57	1.64
RF	1.53	1.68	1.57	1.58	1.59
SVR	1.48	1.55	1.46	1.31	1.45
XGBoost	1.48	1.74	1.48	1.51	1.55

2016 data is split into two parts—training data and validation data. The data from the first 8 months is used to train the algorithms, while the data from the last 4 months is used to evaluate (validate) the algorithms. Four month-long validation data can be considered sufficient to cover different parts of the time series such as trends and seasonality patterns. The validation data has been divided into four folds, each lasting for nearly a month. The initial part of the validation set is correlated with the last part of the training set. In order to increase independence between training and validation, a gap of 24 h is removed from the end of the training set close to the validation set.

In the second step, the best performing model has been trained on all data in 2016 in order to capture any seasonal variations.

Thus, the complete dataset has been split into training data (33%), validation data (17%) and test data (50%). The dataset contains over 100,000 timestamps, and hence using only 33% (in the first step) and 50% (in the second step) of the data for training is sufficient. Holding out 17% of the data for validation is in the recommended range (Belyadi and Haghighat 2021).

3.5 Model training

The four shortlisted Ralgorithms—Linear Regression (LR), Random Forest (RF), Support Vector Regression (SVR) and XGBoost—are trained using the training data set. For the algorithms to be evaluated on equal terms, all algorithm parameters are set to their default values during initial training.

3.6 Model evaluation

In the first step, performance of the four algorithms—Linear Regression (LR), Random Forest (RF), Support Vector Regression (SVR) and XGBoost—have been evaluated. Table 3 presents the RMSE scores for the four algorithm from the cross validation. The table shows that Support Vector Regression (SVR) has the best RMSE mean score whereas Linear Regression (LR) has the worst. The existence of almost equal RMSE values across different folds signifies that the data is evenly distributed over the time period.

Table 4 presents the results of the evaluation of the four models on the whole 1-year test set (2017). There is a noticeable difference in the RMSE scores when the models predict a whole year compared to only the folds in the cross validation. This is due to the test set containing faulty turbine operational data whereas the cross validation set consists of only healthy turbine operational data similar to the training set used to learn the model. The evaluation results suggest that:

- Linear Regression (LR)—This has a decent score and shortest fit and prediction time.
- Random Forest (RF)—This has a good score but somewhat long fit time.
- Support Vector Regression (SVR)—This goes from top performing algorithm on the validation data to worst performing on the test data in almost all parameters, highest RMSE and longest fit and predict time.
- XGBoost–This scores on top while having an acceptable fit and predict time.

TABLE 4 Performance of models with default parameters.

Model	MAE	MAPE	MSE	RMSE	R^2	Fit time [s]	Predict time [s]
LR	1.569	0.039	4.436	2.106	0.980	0.011	0.005
RF	1.479	0.035	3.888	1.972	0.982	18.104	0.889
SVR	1.521	0.037	4.887	2.211	0.978	90.701	188.590
XGBoost	1.436	0.034	3.824	1.955	0.983	1.266	0.019

TABLE 5 Hyperparameter search range.

Hyperparameter	Search range	Optimal value
n_estimators	[200, 400, 600, 800, 1,000]	1,000
max_depth	[3, 4, 5, 6, 7, 8, 9]	4
learning_rate	[0.1, 0.05, 0.01]	0.05

To visualize the performance of the algorithms, plots of the predicted temperatures versus observed temperatures are shown in Figure 10.

- Linear Regression (LR)—This tends to predict rather low values
- Random Forest (RF)—Along with XGBoost this appears to give the best fit
- Support Vector Regression (SVR)—This predicts high values for some low bearing temperatures and low values for some high bearing temperatures.
- XGBoost–This appears to be the most accurate model, even though at times it predicts high values for some low bearing temperatures

While the SVR shows good performance in scoring metrics, it is important to note that the algorithm demands significantly more time for model fitting and prediction compared to XGBoost. This increased computational time, especially while dealing with large datasets or in real-time analysis, often makes SVR unsuitable. In contrast, XGBoost with its efficient handling of large data and faster execution emerges as a more practical choice.

Upon detailed evaluation, XGBoost has been identified as the most suitable algorithm because it strikes an optimal balance between accuracy and computational efficiency. Furthermore, this algorithm can be fine-tuned using hyperparameter tuning techniques, thereby, enhancing its performance. This process involved systematically adjusting the algorithm's parameters to find the combination that yields the best results in terms of prediction accuracy and processing speed. The fine-tuned XGBoost model demonstrates a marked improvement in performance, confirming its suitability for the required predictive modeling tasks.

3.7 Hyperparameter tuning

As described in the previous section, the XGBoost model has been selected as the most suitable model for further analysis. An important part of machine learning optimization is the tweaking and tuning of hyperparameters. Hyperparameter tuning is performed in the XGBoost model to enhance the model's accuracy before trying it on the test data set. The selected hyperparameters and their suggested ranges (Probst et al., 2019) for tuning are presented in Table 5. In addition to the parameters in Table 5, the parameters colsample_bytree and colsample_bylevel have been set to 0.6. In order to determine the optimal combination of hyperparameters grid search with cross validation strategy has been performed.

Results from the grid search are displayed in Figure 11. The figure shows that as compared to max_depth, learning_rate and n_estimators have more effect on performance of the algorithm in terms of RMSE, MAE and R^2. The optimal values of these parameters are given in Table 5.

Table 6 shows the performance of XGBoost algorithm after hyperparameter tuning using the optimized parameter values given in Table 5. As shown, there is an improvement in the performance of the algorithm after hyperparameter tuning.

3.8 Prediction of generator bearing temperature

The optimized XGBoost algorithm-based model (Figure 3) has been used to predict bearing temperature using the Testing Data (2017).

Figure 12 shows the plots of the actual and predicted values for the period 1 January to 15 January 2017, the curves of which are for:

- actual temperature
- predicted temperature
- predicted plus/minus 2 standard deviation temperature

The figure shows that the actual temperature remains within the (predicted ±2 standard deviation or approximately 3.5°C) temperature range.

3.9 Sources of error

Inaccuracies in the output results may arise due to:

- The high correlations between feature and target variables may impact how the machine learning model learns. This risk is partly mitigated by using hyperparameters colsample_bytree and colsample_bylevel.
- Faulty sensors
- Wrong calibration or drift in calibration of sensors
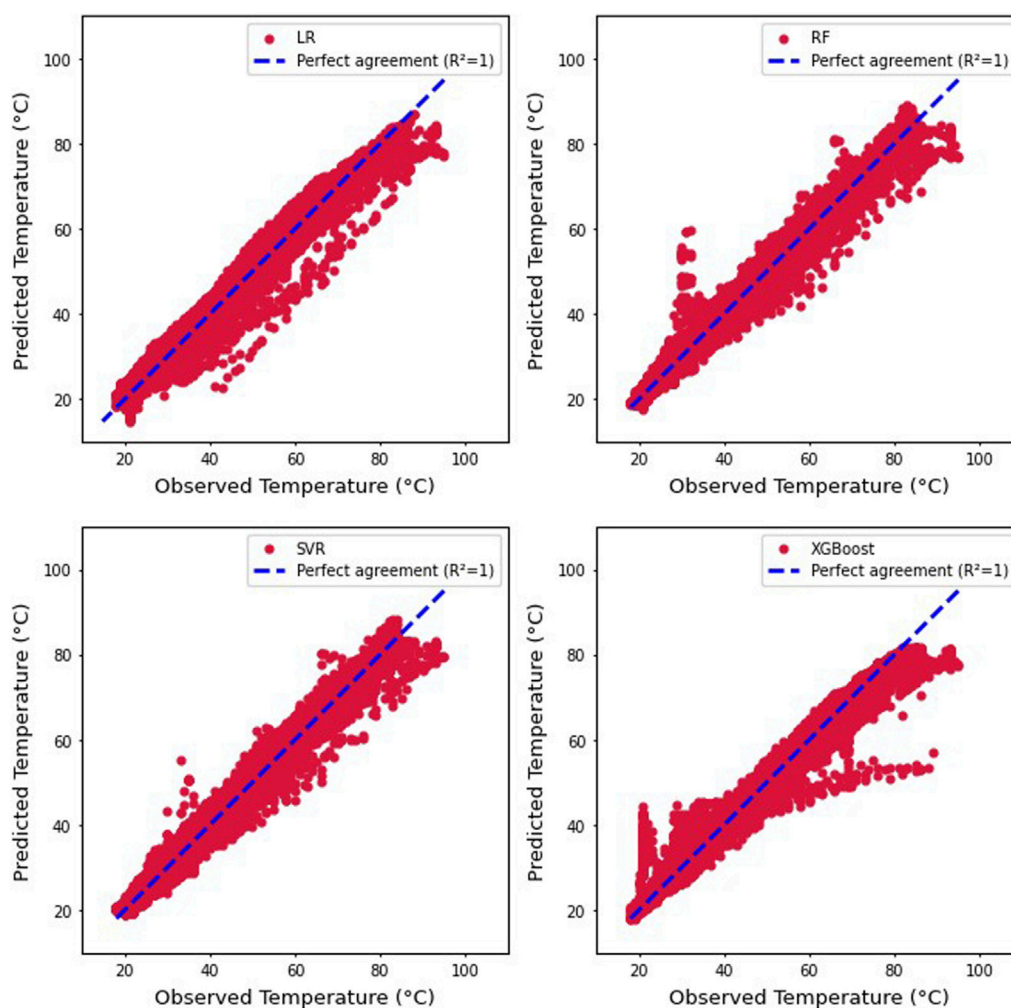
FIGURE 10
Predicted and observed temperatures for all models.
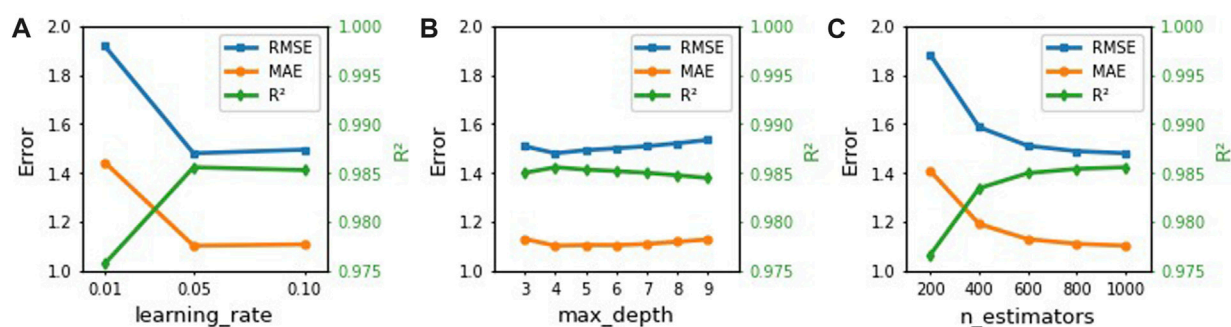
FIGURE 11
Model impact changing **(A)** *learning_rate*, **(B)** *max_depth* and **(C)** *n_estimators*.

In the case study there may be additional sources of errors, including:

- Replacing the missing humidity data with the values from the previous year.

3.10 Fault detection and recommendation for rescheduling maintenance plan

Figure 13 shows the plots of the actual and predicted values for the period from 7 June to 23 June 2017. During this period there are

TABLE 6 Optimized XGBoost performance on test data and validation data.

Test data performance					
Model	MAE	MAPE	MSE	RMSE	R^2
XGBoost	1.436	0.034	3.824	1.955	0.983
Optimized XGBoost	1.389	0.033	3.354	1.832	0.985
Change [%]	3.272	2.941	12.291	6.292	0.203
Validation Data Performance [RMSE]					
Model	Fold 0	Fold 1	Fold 2	Fold 3	Mean
XGBoost	1.48	1.74	1.48	1.51	1.55
Optimized XGBoost	1.41	1.65	1.44	1.40	1.48
Change [%]	4.73	5.17	2.70	7.29	4.52

times when the actual bearing temperature exceeds the predicted value by more than two standard deviations (approximately 3.5°C) over significantly long periods, and this is highlighted in green. For example, on 7 June 2017, the actual value reaches 95°C whereas the model prediction is 76°C, a difference of 19°C.

After 7 June 2017, there is a tendency for the actual bearing temperature to be higher than the predicted bearing temperature. At times it often crosses the two standard deviation limit. This indicates two possibilities:

- Malfunctioning of the bearing sensor.
- Possibility that the bearing is getting hotter than expected perhaps due to increased friction. The increased friction could be either because of increased wear or improper lubrication.

Both of these possibilities warrant special inspection and monitoring activity.

Based on the detection of faulty bearing, recommendation may be made for scheduling maintenance activities at the earliest opportunity. This recommendation is justified by the fact that the bearing breaks down 2 months later on 20 August 2017.

4 Model interpretation using SHAP

The XGBoost algorithm-based model used for the case study gives reasonably good predictions for the temperature of a generator bearing. The model needs to be further evaluated to interpret it is working. Since XGBoost is a tree-based model, the Tree SHAP algorithm proposed by Lundberg et al. (2018) for tree ensembles can be used to calculate the SHAP values that could be used for the interpretation of the working.

4.1 Global explanations

Figure 14A shows the mean absolute SHAP values for the used features. The figure shows that:

- The generator phase temperature has by far the highest impact on the model predictions. This is reasonable due to the adjacent location of the bearing and generator.
- Nacelle temperature and wind speed have moderate average impact on the model predictions, which should be expected since the convective heat loss from bearing is directly proportional to the difference in temperature between the
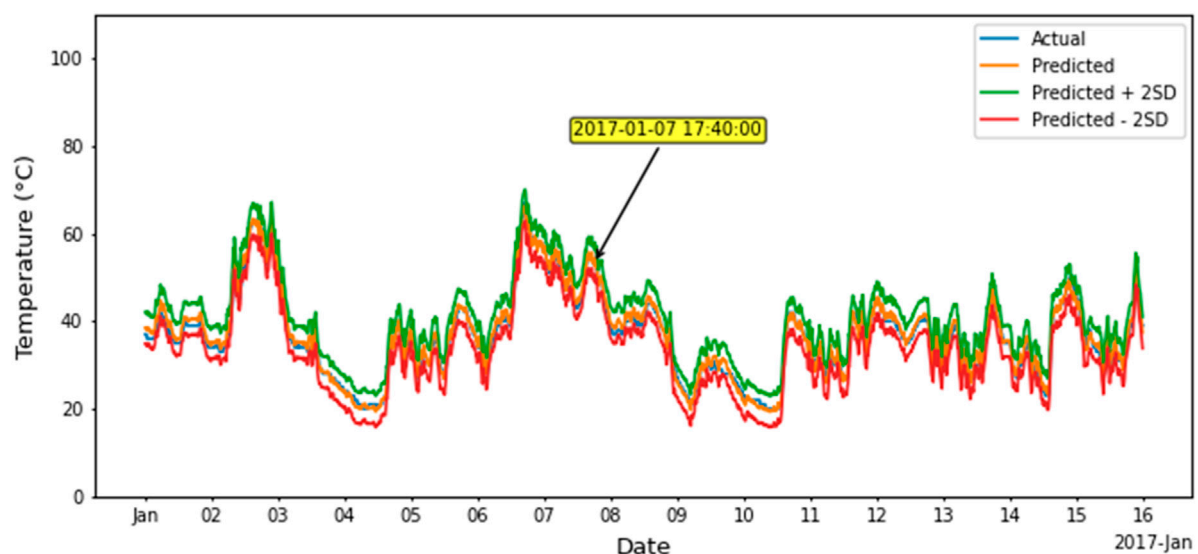
FIGURE 12
Actual and predicted temperatures of generator bearing for the period January 1 to 15 January 2017.
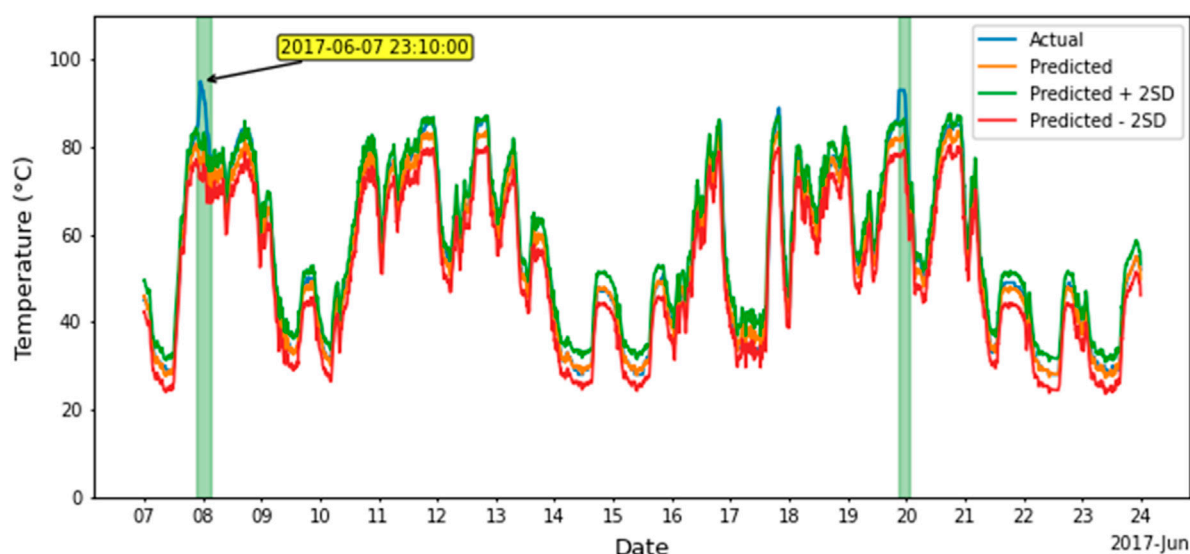
FIGURE 13
Actual and predicted temperatures of generator bearing for the period 7 June to 23 June 2017.

bearing and the nacelle temperature. Wind speed affects not only the rotational speed but also the convective heat loss.

- Generator or bearing rotational speed and relative humidity have low impact.

Figure 14B shows the changes in the SHAP value for changes in the feature value. For all features except the humidity, a higher feature value has a positive impact on the model prediction, and a low feature value has a negative impact on the model output. As is to be expected, the humidity has the opposite impact for its feature values, because increase in humidity increases the specific heat capacity of air resulting in higher convective heat loss from the bearing and a decrease in temperature.

SHAP treats each feature as a "player," hence there are interaction effects between features. The SHAP main effect plots in Figure 15 remove all interaction effects between features and thus display the raw impact of each feature.

- Generator Shaft/Bearing Rotational Speed—Generator rotational speed has a low impact with a small positive spike near its max rotation speed.
- Generator Temperature—The generator phase temperature has a dominant and nearly linear impact on the model output.
- Wind Speed—At the cut-in wind speed of 4 m/s, there is a marked increase in the impact of wind speed. It increases up until the rated wind speed of 12 m/s and from there on stays constant.
- Nacelle Air Humidity—The impact of humidity is rather weak and decreases slowly across its range.
- Nacelle Temperature—Nacelle temperature has an increased positive impact in the temperature range 20°C–45°C.

4.2 Local explanations

SHAP waterfall plots are used for explaining individual predictions. Starting from the expected value of the model output

(the average prediction of the model on the training data) at the bottom of the waterfall plot, each row shows the contribution of each feature to the model output for a prediction. A positive (red) contribution moves the initial output value higher whereas a negative (blue) contribution moves the initial output value lower.

4.2.1 Explanation of prediction for 7 January 2017

Figure 12 shows the plots of the actual and predicted values for the period of 1 January to 15 January 2017. During this period all predicted values are within two standard deviations of the actual value, indicating a possibility that the bearing is operating normally. From this period, an instance (7 January 2017, 17:40:00) has been randomly selected for local explanation.

According to Figure 15, the temperature of bearing is influenced most by the generator temperature because of its high temperature and proximity to the bearing. This is followed by the nacelle temperature and wind speed. The generator rotational speed and humidity have relatively minor effect.

On 7 January 2017, at 17:40:00 the actual generator bearing temperature is 53°C. The SHAP waterfall plot in Figure 16 explains how the XGBoost model arrived at a prediction of 54°C.

- Generator Shaft/Bearing Rotational Speed—Rotational speed has minor effect on the predicted temperature value, hence the net heating effect on the predicted bearing temperature (+0.52°C) is relatively small.
- Generator Temperature—The high generator phase temperature (89.3°C) has by far the most significant positive influence (+8.52°C) on the bearing temperature.
- Wind Speed—Wind speed makes relatively small positive effect (+2.02°C) on the predicted value. Wind speed has two opposing effects—increase in temperature due to increased friction and decrease in temperature due to increased convective heat loss. In this case the rotational speed has small effect (+0.52°C) and hence a greater
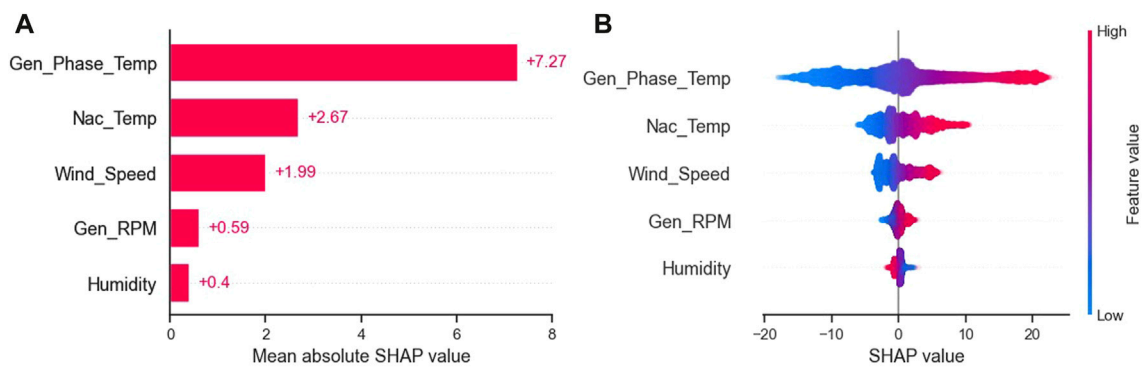
FIGURE 14
(A) Mean absolute SHAP value per feature. **(B)** Matrix plot of SHAP values for different features.
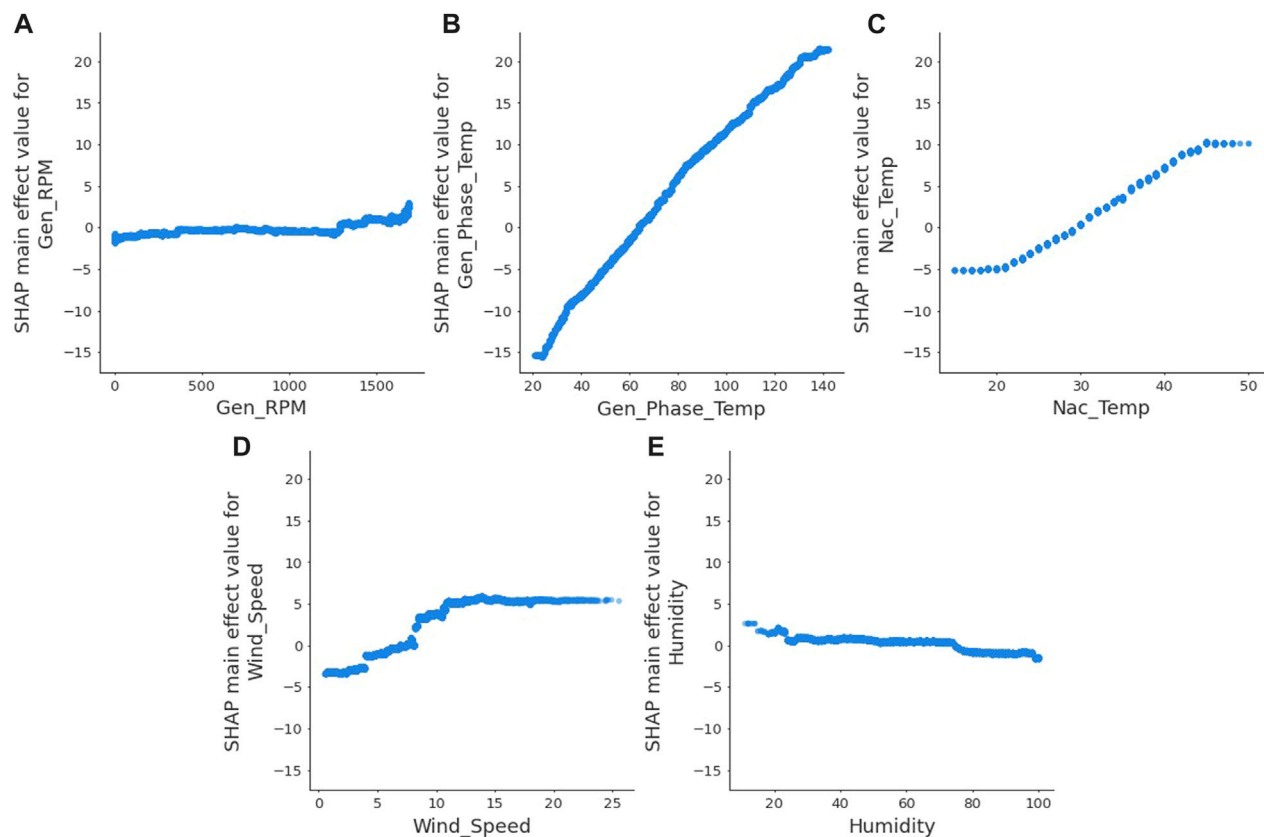
FIGURE 15
SHAP main effects plot for **(A)** generator rpm, **(B)** generator phase temperature, **(C)** nacelle temperature, **(D)** wind speed and **(E)** humidity.

positive effect may be due to the interaction between the wind speed, the generator temperature and the bearing temperature.

- Nacelle Air Humidity—The high relative humidity (78%) also does not significantly (−0.52°C) affect the predicted temperature value, because relative humidity itself does not have any significant role.
- Nacelle Temperature—The nacelle temperature (30°C) is close to the average annual temperature, ranging between 15°C and

50°C, and hence does not play a significant role (−0.01°C) in the fall of temperature on predicted value.

4.2.2 Explanation of prediction for 7 June 2017

Figure 13 shows the plots of the actual and predicted values for the period 7 June to 23 June 2017. On 7 June 2017 (Summer), the environmental and operating temperatures are quite different
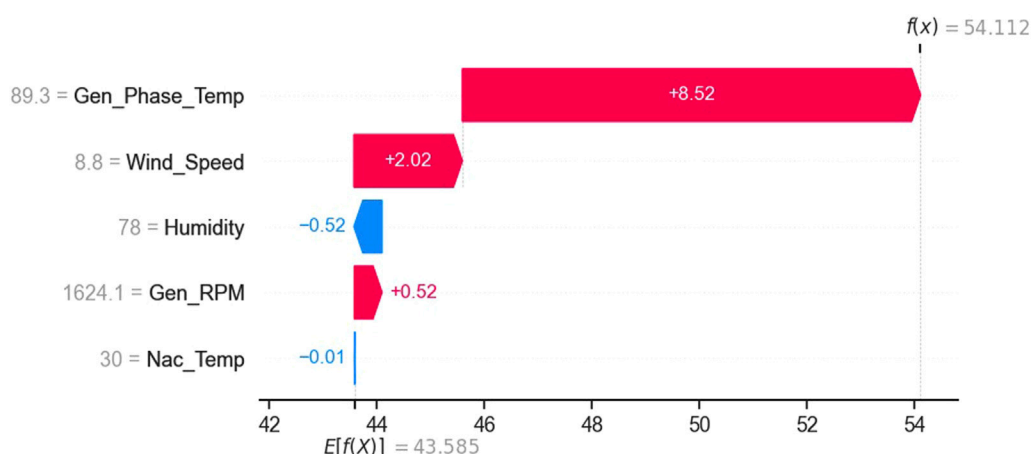
FIGURE 16
Local explanation on 7 January 2017, 17:40:00 by waterfall plot.
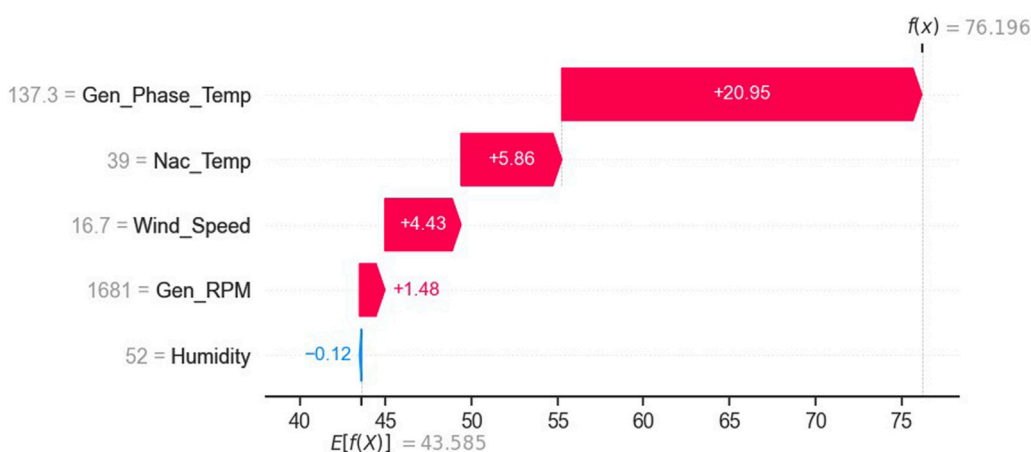
FIGURE 17
Local explanation on 7 June 2017, 23:10:00 by waterfall plot.

from those of 7 January 2017 (Winter). Based on the SHAP waterfall plot (Figure 17), an attempt is made to explain the working of the model.

- Generator Shaft/Bearing Rotational Speed—As in the previous case (7 January 2017), the rotational speed has a minor effect on the predicted temperature value, and hence the net heating effect on the predicted bearing temperature (+1.48°C) is relatively small. The small increase could be due to the small positive spike that appears near its max rotation speed (Figure 14A).
- Generator Temperature—The generator temperature is very high (137.3°C) and this significantly (+20.95°C) raises the temperature of the bearing.
- Wind Speed—Compared to the previous case, wind speed gives relatively higher positive effect (+4.43°C) on the

predicted value. This may be because of higher interaction between the wind speed, the generator temperature, and the bearing temperature.
- Nacelle Air Humidity—As in the previous case, nacelle relative humidity has negligible (−0.12°C) effect on the predicted temperature value.
- Nacelle Temperature—Compared to the previous case, the nacelle temperature (39°C) is 9°C higher than the previous case, and hence there is significantly (+5.86°C) higher effect on the predicted temperature.

The analysis provides a reasonable explanation for the predicted bearing temperature. A high generator temperature (137°C) increases the predicted bearing temperature significantly (+20.95°C) and the remaining features also contribute to bringing the predicted bearing temperature to 76.2°C.

5 Conclusion

This paper presents a simple and robust methodology for making a machine learning based model for detecting faults in wind turbine generator bearing. In this model, the predicted bearing temperature is compared against the actual bearing temperature and a significant difference between the two indicates a possibility of fault(s) in the bearing or its lubrication. Either of these may result in failure. As a case study, the idea has been demonstrated on a generator bearing, using real-life SCADA data. The results show that it is possible to detect potential failure well in advance. This knowledge can be used for planning maintenance.

Four different machine learning algorithms, Linear Regression (LR), Random Forest (RF), Support Vector Regression (SVR) and XGBoost, have been evaluated and XGBoost has been found to be the most suitable algorithm for the task.

The paper also examines the role of five features, generator shaft/bearing rotational speed, generator temperature, wind speed, nacelle air humidity, and nacelle temperature, on the predicted bearing temperature. Out of these, the generator temperature has been found to play the major role, followed by the wind speed and nacelle temperature. Bearing rotational speed and relative humidity of nacelle air play minor roles.

To take the research work further, the following tasks have been identified:

(a) analysis of data from different wind turbines,
(b) testing of other machine learning/artificial intelligence algorithms, like artificial neural networks,
(c) consideration of the impact of more features,
(d) use of other interpretable machine learning tools such as Individual Conditional Expectation (ICE) plots (Goldstein et al., 2015) and LIME (Local interpretable model-agnostic explanations (LIME) (Ribeiro et al., 2016),
(e) expanding the scope from component to system level.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: For training the link is: https://www.edp.com/en/wind-turbine-scada-signals-2016, For testing the link is: https://www.edp.com/en/innovation/open-data/wind-turbine-scada-signals-2017.

Author contributions

OB: Writing–original draft. MS: Writing–review and editing. KØ: Writing–review and editing. AK: Writing–review and editing.

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

Adadi, A., and Berrada, M. (2018). Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). *IEEE Access* 6, 52138–52160. doi:10.1109/ACCESS.2018.2870052

Arabian-Hoseynabadi, H., Oraee, H., and Tavner, P. J. (2010). Failure modes and effects analysis (FMEA) for wind turbines. *Int. J. Electr. Power Energy Syst.* 32 (7), 817–824. doi:10.1016/j.ijepes.2010.01.019

Belyadi, H., and Haghighat, A. (2021). *Machine learning guide for oil and gas using Python: a step-by-step breakdown with data, algorithms, codes, and applications*. Houston: Gulf Professional Publishing.

Bergstra, J., and Bengio, Y. (2012). Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* 13 (2). Available at: https://www.jmlr.org/papers/volume13/bergstra12a/bergstra12a.pdf.

Boukhriss, M., Khalifa, Z., and Ghribi, R. (2013). Study of thermophysical properties of a solar desalination system using solar energy. *Desalination Water Treat.* 51, 1290–1295. doi:10.1080/19443994.2012.714925

EDP (2017). Data. Available at: https://www.edp.com/en/innovation/open-data/data (Accessed June 14, 2023).

Ekanayake, I., Meddage, D., and Rathnayake, U. (2022). A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP). *Case Stud. Constr. Mater.* 16, e01059. doi:10.1016/j.cscm.2022.e01059

Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). "A density-based algorithm for discovering clusters in large spatial databases with noise," in Second International Conference on Knowledge Discovery and Data Mining (KDD'96). Proceedings of a conference held, August 2-4, 226–231.

Goldstein, A., Kapelner, A., Bleich, J., and Pitkin, E. (2015). Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation. *J. Comput. Graph. Statistics* 24 (1), 44–65. doi:10.1080/10618600.2014.907095

Gupta, P. K. (2013). "Analytical modeling of rolling bearings," in *Encyclopedia of tribology*. Editors Q. J. Wang and Y. W. Chung (Boston, MA: Springer). doi:10.1007/978-0-387-92897-5_741

IEA (2021). *Net zero by 2050*. Paris. Available at: https://www.iea.org/reports/net-zero-by-2050 (Accessed August 26, 2022).

ISO (2017). ISO 15243:2017 Rolling bearings — damage and failures — terms, characteristics and causes. Available at: https://www.iso.org/standard/59619.html (Accessed June 13, 2022).

Kahrobaee, S., and Asgarpoor, S. (2011). "Risk-based failure mode and effect analysis for wind turbines (RB-FMEA)," in 2011 North American Power Symposium, Boston, MA, USA, 04-06 August 2011.

Kandukuri, S. T., Klausen, A., Karimi, H. R., and Robbersmyr, K. G. (2016). A review of diagnostics and prognostics of low-speed machinery towards wind turbine farm-level health management. *Renew. Sustain. Energy Rev.* 53, 697–708. doi:10.1016/j.rser.2015.08.061

Koukoura, S., Scheu, M. N., and Kolios, A. (2021). Influence of extended potential-to-functional failure intervals through condition monitoring systems on offshore wind turbine availability. *Reliab. Eng. Syst. Saf.* 208, 107404. doi:10.1016/j.ress.2020.107404

Luengo, M. M., and Kolios, A. (2015). Failure mode identification and end of life scenarios of offshore wind turbines: a review. *Energies* 8, 8339–8354. doi:10.3390/en8088339

Lundberg, S. M., Erion, G. G., and Lee, S.-I. (2018). Consistent individualized feature attribution for tree ensembles. arXiv e-prints. Available at: https://doi.org/10.48550/arXiv.1802.03888.

Lundberg, S. M., and Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Adv. neural Inf. Process. Syst.* 30. doi:10.48550/arXiv.1705.07874

Mahesh, B. (2020). Machine learning algorithms - a review. *Int. J. Sci. Res. (IJSR)* 9 (1), 381–386. Available at: https://www.ijsr.net/archive/v9i1/ART20203995.pdf.

Molnar, C., Casalicchio, G., and Bischl, B. (2020). "Interpretable machine learning – a brief history, state-of-the-art and challenges," in *ECML PKDD 2020 workshops*. Editor I. Koprinska (Cham: Springer International Publishing), 417–431.

Nilsson, J., and Bertling, L. (2007). Maintenance management of wind power systems using condition monitoring systems—life cycle cost analysis for two case studies. *IEEE Trans. energy Convers.* 22 (1), 223–229. doi:10.1109/tec.2006.889623

Probst, P., Boulesteix, A.-L., and Bischl, B. (2019). Tunability: importance of hyperparameters of machine learning algorithms. *J. Mach. Learn. Res.* 20 (1), 1934–1965. doi:10.48550/arXiv.1802.09596

Ribeiro, M. T., Singh, S., and Guestrin, C. (2016). "Why should i trust you?," in *Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. Editor

A. f. C. Machinery (San Francisco, CA (United States): Association for Computing Machinery), 1135–1144.

Sankar, S., Nataraj, M., and Prabhu Raja, V. (2012). Failure analysis of bearing in wind turbine generator gearbox. *J. Inf. Syst. Commun.* 3 (1), 302–309.

Shafiee, M., and Dinmohammadi, F. (2014). An FMEA-based risk assessment approach for wind turbine systems: a comparative study of onshore and offshore. *Energies* 7, 619–642. doi:10.3390/en7020619

Stehly, T., Beiter, P., and Duffy, P. (2020). *2019 cost of wind energy review*. NREL/TP-5000-78471. Golden, CO (United States): National Renewable Energy Laboratory. Available at: https://www.nrel.gov/docs/fy21osti/78471.pdf.

Stetco, A., Dinmohammadi, F., Zhao, X., Robu, V., Flynn, D., Barnes, M., et al. (2019). Machine learning methods for wind turbine condition monitoring: a review. *Renew. Energy* 133, 620–635. doi:10.1016/j.renene.2018.10.047

Vilone, G., and Longo, L. (2020). Explainable artificial intelligence: a systematic review. arXiv preprint. Available at: https://doi.org/10.48550/arXiv.2006.00093.

Wiggelinkhuizen, E., Rademakers, L., Verbruggen, T., Watson, S., Xiang, J., Giebel, G., et al. (2007). *Conmow final report*. Netherlands: Energy research Centre of the Netherlands.

Zhang, X., Sun, L., Sun, H., Guo, Q., and Bai, X. (2016). Floating offshore wind turbine reliability analysis based on system grading and dynamic FTA. *J. Wind Eng. Industrial Aerodynamics* 154, 15421–15433. doi:10.1016/j.jweia.2016.04.005

Yaw misalignment in powertrain degradation modeling for wind farm control in curtailed conditions

Farid Khazaeli Moghadam[1]*, Zhen Gao[2], Valentin Chabaud[3] and Spyridon Chapaloglou[3]

[1]Department of Marine Technology, Norwegian University of Science and Technology, Trondheim, Norway, [2]School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai, China, [3]Energy Systems Department, SINTEF Energy Research, Trondheim, Norway

A framework characterizing the degradation of wind turbines for use in multiple-input damage-aware farm control is suggested. The focus is on the fatigue damage of the powertrain (drivetrain + generator) system, but the methodology may be extended to other components. A database of steady-state damage analyses for different operating conditions (average wind speeds, turbulence levels, power demands, and yaw misalignment angles) using aero-hydro-servo-elastic simulations is first generated. Then, a weighted damage index based on probabilistic long-term fatigue damage analysis of the powertrain system components is suggested and used to represent degradation at the farm level for control purposes. The focus is on curtailed conditions where the farm controller dispatches power commands to individual turbines in order to track a demanded power reference (rather than seeking to maximize power) at the farm level. As a secondary objective, the controller seeks to mitigate degradation through a smart combination of power commands and yaw offset angles, making use of the weighted degradation index. The potential of the proposed approach is demonstrated through a case study on the TotalControl Reference Wind Power Plant in a FLORIS-based simulation framework. The proposed farm controller is compared with the conventional one without damage mitigation feature and with damage mitigation but without yaw angle as the control input. It is found that combining yawing and downregulation effectively slows down degradation on the main bearing and powertrain as a whole.

KEYWORDS

database, fatigue damage, powertrain, farm control, power tracking, curtailment, yaw misalignment

1 Introduction

Enhancing the cost-effectiveness of offshore wind is unanimously cited as essential to solve the energy crisis, and the importance of optimizing the operation and maintenance of wind farms is underscored. Regarding operation, an aspect that takes increasing space in wind farm operators' economics is the provision of ancillary services to grid operators. In particular, farm curtailment services for secondary (directly asked by the grid operator) or tertiary (through intraday bidding in electricity markets) frequency support is expected to increase dramatically as the penetration of intermittent energy sources (wind and solar)

into the grid increases (Yasuda et al., 2022; Wiser et al., 2023). In this context, wind farms are asked to provide a desired power below the maximum available in the wind, leaving some freedom in how individual turbine power outputs should be dispatched. Maintenance, on the other hand, refers more to the degradation of wind energy systems over time, associated with repair and downtime costs, which is all the more acute as wind farms move further from the shore. Recent literature has developed digital twin models based on simulation or operational data for monitoring the health condition and optimizing the operation of the various electrical, mechanical, and structural systems of wind turbines and power plants (Xia and Zou, 2023; Moghadam and Nejad, 2022; van Dinter et al., 2022). A powertrain system consisting of the gearbox, shafts, main bearings, back-to-back (BTB) frequency converter, generator, and rotor is on average responsible for approximately 50% of wind turbine total failures and downtime (Pfaffel et al., 2017). Among these, powertrain components such as bearings and gearbox are prone to fatigue damage, which greatly depends on how much power the turbine is asked to produce (Moghadam et al., 2023). This lays the motivation behind this study, as a way to reduce the levelized cost of energy (LCOE) of wind power, combining wind farm curtailment with degradation mitigation by smart power dispatch to influence the remaining useful lifetime of powertrain components, in order to reduce maintenance costs while making profit from ancillary service provision.

However, this requires a multidisciplinary and multiscale simulation framework for development and analysis. In particular, complexity has to be reduced by using a simple indicator for degradation, avoiding the requirement for running turbine- and component-scale models at the farm level. To this end, Moghadam et al. (2023) suggested a mapping between the damage index and operating conditions (wind speed, turbulence intensity, and power set point) based on turbine-level simulations and a quasi-static degradation model. High-fidelity physics-based (state-space models of varying degrees of complexity with constant/time-variant lumped parameters (Moghadam et al., 2021; Zhang et al., 2021; Moghadam and Desch, 2023), multi-body (Peeters et al., 2006) and finite element (Hart et al., 2020), and data-driven (random forest; Azzam et al., 2022) models and artificial neural network (Azzam et al., 2021) have been used in the literature to estimate loads on powertrain components, but the overall complexity sets a limit to their applicability. Higher fidelity would also be beneficial regarding the effect of wake on powertrain, especially when wake flow impacts only part of the downstream turbine's rotor as has been identified by van Binsbergen et al. (2020) using FAST.Farm simulations featuring the dynamic wake meandering model were carried out by Madsen et al. (2010). Again, with complexity as the limiting factor and with farm-level use in mind, Moghadam et al. (2023) modeled wake-added turbulence as a simple increase in turbulence intensity in the ambient wind spectrum as recommended by the IEC standard IEC 61400-1:2019 (2019). This way, wake-added turbulence is readily encompassed by adjusting turbulence intensity as an input parameter when reading in the database.

Another multidisciplinary bottleneck lies in modeling local wind variations between turbines, since they drive the local available power that is a key input to power dispatch—setting an upper bound to the individual set points. Variations in mean wind speed due to wake-induced velocity deficit are typically readily included in available farm simulation tools. Farm-wide wind fluctuations due to turbulence are, however, not trivial to include, although their effect on power dispatch is tremendous. The mid-fidelity approach used in Moghadam et al. (2023) is based on models for synthetic turbulence generation that have been specifically developed for farm-wide applications (Chabaud, 2023, based on Sørensen et al., 2008).

Damage mitigation in curtailed operations is receiving increasing attention in the literature (Knudsen et al., 2015; Stock et al., 2020; Merz et al., 2021; Sood et al., 2023). However, the powertrain has not been the focus as modeling fatigue damage on its components requires multidisciplinary collaboration and a change of metric from damage-equivalent load to fatigue damage (Moghadam et al., 2023; Sood et al., 2023). Also, modeling wind fluctuations has been either simplistic (based on frozen turbulence assumption, only valid at the turbine level) or exceedingly costly (based on large eddy simulations, LES), limiting the studies' validity or flexibility, respectively.

Yaw angle offset control has always been used in the context of power maximization: the basic use is to follow wind direction and reduce misalignment, and the advanced use is wake steering to reduce wake losses on downstream turbines. In this context, the influence of yaw misalignment on drivetrain loads has been studied by Cardaun et al. (2019) and van Binsbergen et al. (2020), where the authors showed that yaw misalignment does not necessarily have a negative impact; however, a combination with the farm power curtailment controller has not been studied. This work adds a new dimension to the previous study by Moghadam et al. (2023), showing how yaw angle offset and power-tracking active controllers can cooperate at the farm level to reduce fatigue damage and spread it among turbines. To this end, a powertrain fatigue damage analysis is performed considering different power set points and yaw offset angles in addition to wind conditions. The resulting damage database is utilized for tuning the farm controller gains. The potential of the proposed control scheme is then demonstrated in case studies.

The main contributions of this article are

- developing a powertrain system fatigue damage database that encompasses yaw misalignment angle variations in addition to wind field and demand variations,
- designing a power tracking and yaw angle offset coordinated controller by using the developed database to enhance the load mitigation feature of the farm controller, and
- demonstrating the use of the database within the proposed farm controller and its potential to steer farm-wide degradation.

The article is organized as follows: the database approach, the farm controller, and the utilization of the database are described in Section 2; the results related to database generation and demonstration are discussed in Section 3; and finally, the article is concluded in Section 4.

2 Methodology

2.1 Database generation

The methodology described in Moghadam et al. (2023) with yaw misalignment added is shown in Figure 1. The powertrain system damage database as a function of wind speed, turbulence intensity, power demand, and yaw misalignment angle is created by turbine-level aero-hydro-servo-elastic simulations in NREL's OpenFAST—the *servo* part uses a custom version of the DTU wind energy controller (Meng et al., 2020; Hansen and Henriksen, 2013) featuring active power derating; the *hydro* part is not included at this stage and left as further work. The steps are

- generating input turbulent wind field by NREL's TurbSim with desired mean wind speed and turbulence intensity,
- running the aero-hydro-servo-elastic simulation with the desired power command and yaw misalignment angle to obtain global powertrain loads,
- propagating these loads to powertrain subcomponents using a quasi-static approach,
- calculating the load and stress of each subcomponent, and
- calculating the fatigue damage of each subcomponent and the weighted damage index.

The last step is based on a powertrain degradation model, further detailed in the following:

- The physics-based quasi-static model proposed by Moghadam et al. (2023) is employed.
- The load elements of the bearings and gears of the powertrain, then the equivalent load of the bearings and maximum stress of the gears are calculated.
- The load calculation is carried out based on the types of components, basic geometrical parameters, powertrain input loads, and safety factors from ISO 281 and ISO 6336 standards.
- Rainflow cycle counting and Miner's rule are applied to calculate the fatigue damage of the gears.
- Load–duration–distribution and Miner's rule are applied to calculate the fatigue damage of the bearings.

The database created by turbine-level simulations establishes the damage index (an index that represents the long-term fatigue damage of all the drivetrain components) as a function of environmental (average wind speed and turbulence intensity) and operational (generated power and yaw misalignment angle) variables. The wake effect for individual turbines is considered by the reduced mean wind speed and the increased wake-induced turbulence intensity. When used in farm control, it is desirable to characterize degradation via a scalar quantity, here called the damage index. Depending on the control objective, the damage index may be specific to a single powertrain component or aggregated over all components. In the latter case, a weighted damage index may be used based on the vulnerability of each component. In this article, a simple weighting is suggested based on probabilistic long-term fatigue damage analysis:
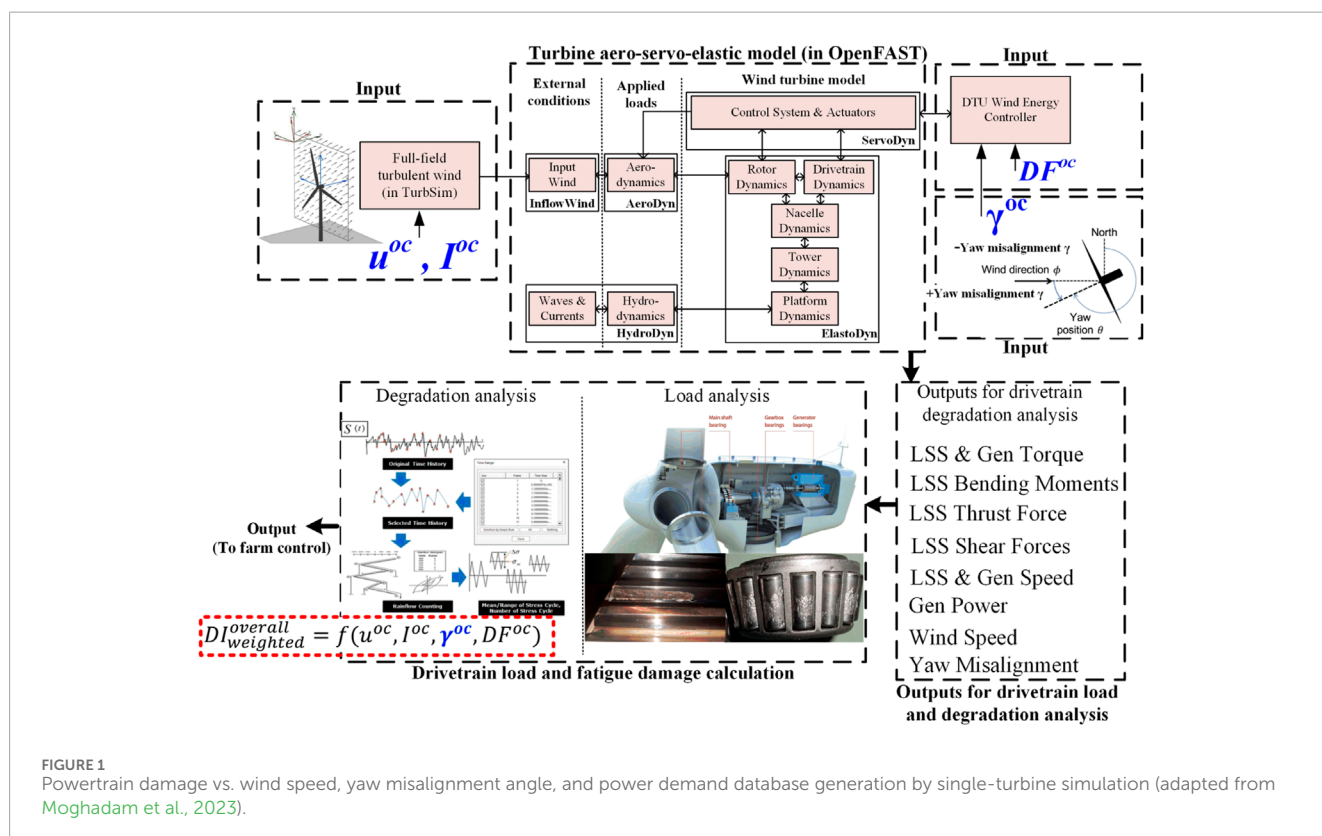
Powertrain damage vs. wind speed, yaw misalignment angle, and power demand database generation by single-turbine simulation (adapted from Moghadam et al., 2023).

$$DI^{overall}_{weighted}(u^{oc}, I^{oc}, \gamma^{oc}, DF^{oc}) = \frac{1}{M}\sum_{i=1}^{M}\left(\alpha_i D^{t,oc}_i(u^{oc}, I^{oc}, \gamma^{oc}, DF^{oc})\right),$$

$$s.t. \sum_{i=1}^{M}\alpha_i = 1,$$

$$\alpha_i = \frac{\bar{D}_i}{\max(\bar{D}_i)},$$

$$\bar{D}_i = \int\int\int\int D_i(u, I, \gamma, DF).f(u, I, \gamma, DF)$$

$$\times du\, dI\, d\gamma\, dDF \quad for \quad i = 1, \dots, M,$$

(1a)

(1b)

where u^{oc}, I^{oc}, and γ^{oc} are, respectively, the mean value of wind speed, the turbulence intensity, and the yaw misalignment angle during the operating condition ^{oc} during the time interval t. $D^{t,oc}_i$ is the absolute value of accumulated damage for the i-th subcomponent during this time interval. DF is the derating factor equal to the ratio of generated power over available power—saturated to rated power at higher wind speeds. M is the number of powertrain subcomponents considered in the degradation analysis. α_i is a weight factor representing the normalized damage for the i-th subcomponent. $f(u, I, \gamma, DF)$ is the joint probability distribution of the operating conditions, with the integral equal to 1.

2.2 Farm control and database utilization

The wind farm controller coordinates the operation of the turbine's active power—through pitch angle and generator torque—and yaw controllers within the farm to meet the power demand while steering degradation. The proposed farm power tracking controller is an adapted version from Moghadam et al. (2023); Merz et al. (2021) now augmented with a yaw angle offset component. The integration of the powertrain degradation database to the farm controller is shown in Figure 2, featuring the yaw angle offset as an additional control input for the damage mitigation function.

The wind farm control layout is presented in Figure 3. A distributed architecture based on individual PI regulators and time-adaptive gains is adopted. The wind farm controller is responsible for dispatching the turbines' power, $P_{c,i}(t)$, such that the reference

wind farm power, at the point of common coupling (PCC), $P_{pcc}(t)$, is tracked. Each individual turbine controller (TC) consists of a feedforward term and a feedback loop, as shown in Figure 3. The former defines the bulk part of the PCC power reference requested from the i-th turbine through the gain $\lambda_{i1}(t)$. The latter uses the farm output power error at the PCC, $\epsilon_{pcc}(t)$, by first dispatching between turbine regulators through the gains $\lambda_{i2}(t)$ and then using it to compensate for the tracking errors through the proportional and integral feedback gains, K_P and K_I, respectively. It is worth mentioning that the power-sharing quality is directly affected by the locally available power, $P_{a,i}(t)$, and smart anti-windup mechanisms are necessary to prevent sudden power injections due to changing inflow conditions from weather, turbulence, and wakes ($u_i(t)$). The participation of each turbine in tracking the farm power output is adjusted based on the turbine's accumulated damage and through time-dependent gains $\lambda_{ij}(t)$ (see Section 2), which are also communicated among the turbines and encapsulate information from the database of the weighted damage index (see Section 2.1). The additional decision feature that is included, when compared to Moghadam et al. (2023), allows for setting the yaw angle offset of the damaged turbine(s), $\gamma_i(t)$, to the value corresponding to the lowest possible damage accumulation, under specified environmental and operational conditions (power dispatch). In this sense, the updated damage database, which includes the extra dimension of the yaw angle offset, is used instead, further enhancing the damage mitigation capabilities for the damaged turbines. The adaptive gains $\lambda_{ij}(t)$ ($j \in \{1, 2\}$) are defined as

$$\lambda_{ij} = \frac{g_{ij}(DI(u, I, \gamma, DF))}{\sum_{turbines} g_{ij}(DI(u, I, \gamma, DF))}, \quad s.t. \quad \sum_{i=1}^{N}\lambda_{ij} = 1. \quad (2)$$

where N is the number of turbines of the farm, and g_{ij} is a tunable mapping from $DI_i = DI(u, I, \gamma, DF)$ to λ_{ij}, for turbine i, as in Moghadam et al. (2023). In case of no fatigue damage mitigation in the farm controller, the gains λ_{i1} and λ_{i2} will take equal values for all the turbines of the farm.

The yaw angle controller is an open-loop add-on. Optimal yaw misalignment values are first found, then translated to yaw angle
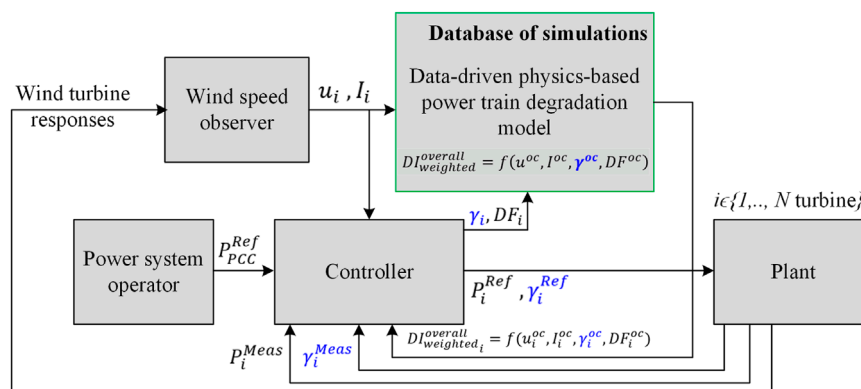
FIGURE 2
Utilization of powertrain degradation database with the additional wake steering dimension for the farm power tracking controller design (adapted from Moghadam et al., 2023).
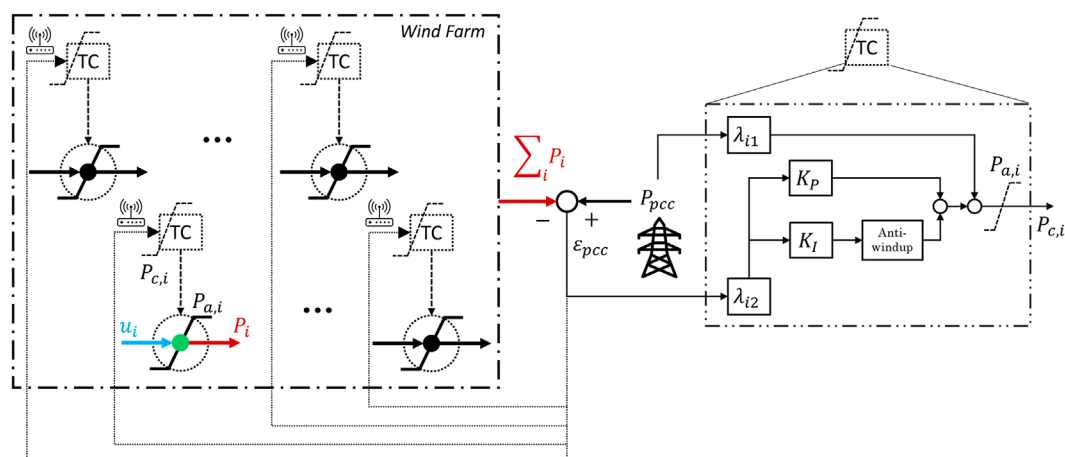
FIGURE 3
Wind farm control configuration and schematic of the distributed controller.

offset commands considering wind direction[1]. Losses induced by yaw misalignment are compensated through the PI feedback of the power dispatch control loop. Note that yaw angle control is not meant to increase power production on downstream turbines through wake steering in this context. Wake deflection is not the focus and is included in the modeling framework only as a side effect. The wake effects are modeled through the simulation framework presented in Moghadam et al. (2023), employing a modified, quasi-steady version of NREL's FLORIS farm flow model, which is capable of capturing the effect of the derating commands to the turbines.

3 Results and discussion

3.1 Case study

The conditions of the case study simulated to generate the database are listed as follows:

- The turbine model is DTU 10 MW reference wind turbine (Bak et al., 2013) fixed-bottom installed on a reference monopile as shown in Figure 4. The 10 MW powertrain system design, configuration, and selection of the subcomponents for the fatigue damage analysis are based on Moghadam and Nejad (2020), Wang et al. (2020), and Moghadam et al. (2023);
- Design Load Case (DLC) 1.2 from IEC 61400-1 IEC 61400-1:2019 (2019) (fatigue damage in normal operation) with wind field parameters adjusted to offshore conditions using IEC 61400-3 IEC 61400-3-1:2019 (2019);
- In the simulations, the average wind speed changes from 4 to 25 m/s in steps of 3 m/s with the resolution of 1 m/s between

---

1 In this study, the wind direction was kept constant in farm-level simulations, so the yaw angle offset commands and yaw misalignment input in the database are equivalent.

FIGURE 4
DTU 10 MW reference wind turbine installed on a reference monopile.

7 and 13 m/s, which is the region that is rated with a high chance of occurrence. Effective turbulence intensity or the turbulence intensity by taking into consideration the wake-added turbulence changes from 0% to 32% in steps of 8%. This range is selected based on the results reported by Frandsen

(2007). The derating factor DF changes from 0% to 100% in steps of 25%. Yaw misalignment angle changes from $-20°$ to $+20°$ in steps of $10°$.

- For each case, six independent 1-h wind field realizations, with the first 10 min removed, are run. In total, 9,000 independent simulations have been run.
- The weighted damage index is calculated using a simple weighted integral along the wind speed only (other dimensions are fixed to their input value) using a standard probability distribution from IEC 61400-1:2019 (2019), wind turbine class I.
- Standard settings in FLORIS are used, *i.e.*, the Gauss and GCH wake model with default parameters (Laboratory, 2023).

3.2 Controller performance

The performance of the DTU controller featuring derating functionality is shown in Figure 5, which shows how the rotor speed and pitch regulators cooperate at different wind speeds and yaw misalignment angles to achieve different values of power demand. In this figure, ω_{rot} is the average rotor speed and β is the average blade pitch angle. As seen, the yaw misalignment angle has a negligible influence on the controller's operating state.

3.3 Powertrain degradation analysis

Figure 6 shows the selected time series of dynamic equivalent loads of the bearings (the gearbox-side main bearing, the gearbox high-speed shaft drive end bearing, and the generator non–drive end beari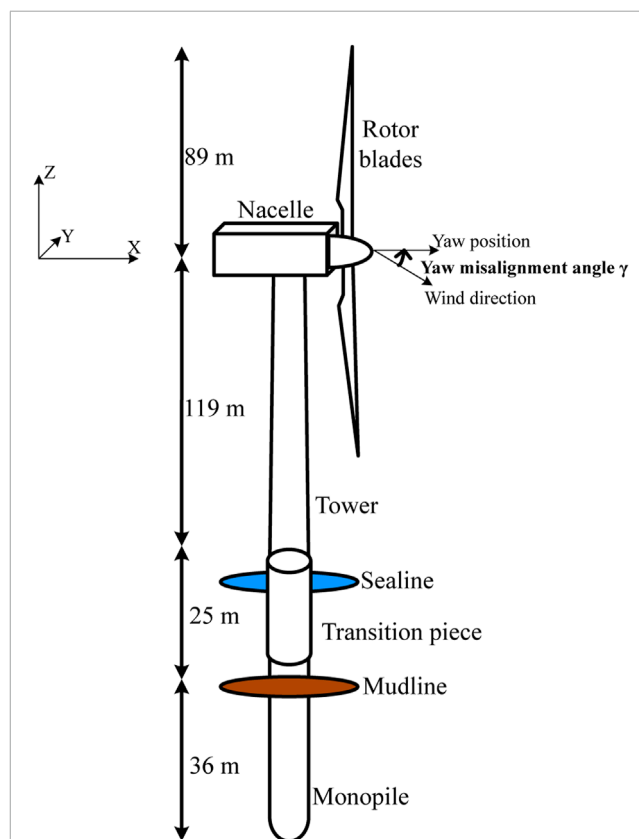ng) and contact and root bending stresses of the gears (the planet gears of the first planetary gear stage) as a function of the average wind speed, turbulence, yaw misalignment angle, and power demand variations. It can be seen that yaw misalignment has the largest influence on the main bearing load in terms of both mean value and oscillation amplitude. Further investigation showed that

the main shaft bending moment around the *z*-axis (yaw moment) is the main contributor.

The results of the long-term fatigue damage analysis of the powertrain system are shown in Figure 7. This sets the base for selecting the weight factors α_i in Section 2.1 to calculate the weighted average damage index. In practice, the main bearings dominate damage in most operating conditions. At very high wind speeds, the planet gears take over. The contributions from the other components to the damage index are minor.

3.4 Turbine-level database

A global overview of the degradation database mapping the damage index to operating conditions is shown in Figure 8, which maps the weighted damage index of the powertrain under consideration to mean wind speed, turbulence intensity, derating factor, and yaw misalignment angle. The damage ratio in this figure is defined as the ratio of the weighted damage index at a non–zero yaw misalignment over its zero yaw misalignment value. It shows the conditions when yawing is beneficial or detrimental, strongly depending on the yawing direction. The rotational direction of the rotor is the same for positive and negative yaw misalignment angles. As a result, the angle of attack for each blade cross-section due to the combined effect of the inflow velocity and the rotational velocity will be different for positive and negative yaw angles. This gives different induced velocities and the resulting aerodynamic forces on the blade cross-sections. As a result, the drivetrain fatigue damages are also different for positive and negative yaw angles, as shown in Figure 9.

Figure 9 shows the normal operation (not downregulated) case in more detail, where it is clear that yawing in any direction would mostly increase damage. However, by crossing observations with Figure 8, it is seen that in highly downregulated conditions, yawing may prove beneficial. This may in turn be exploited by the farm controller to reduce damage whenever the demand is
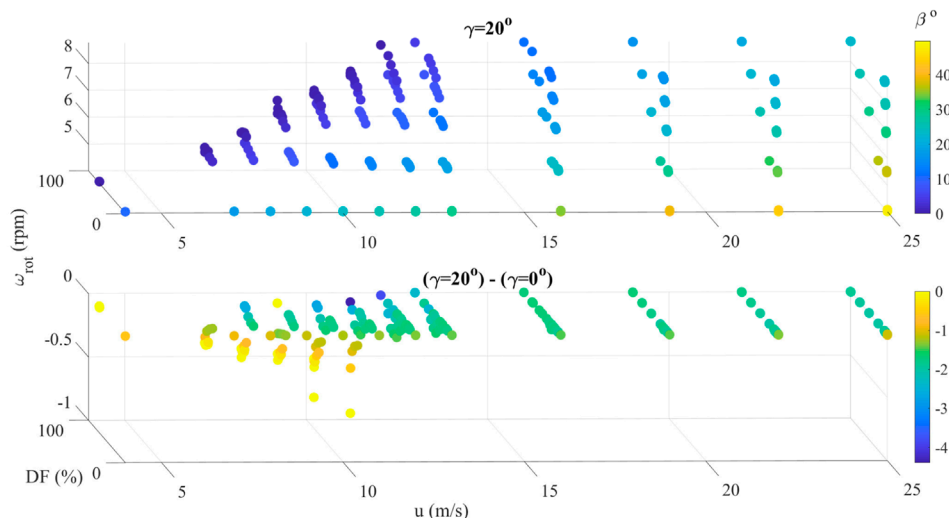
FIGURE 5
DTU controller performance in different yaw misalignment angles.
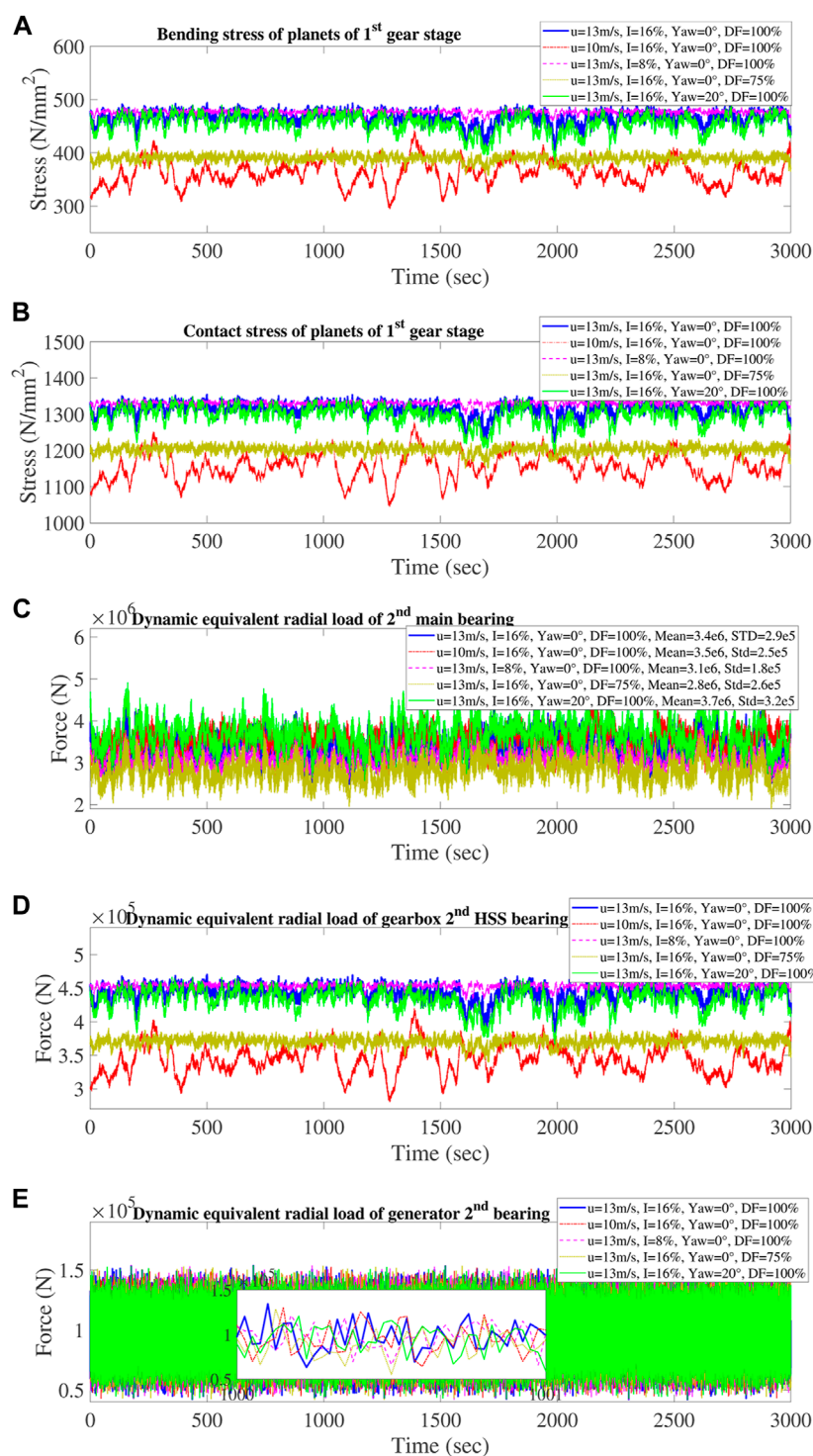
FIGURE 6
Estimated powertrain components load and stress for different load conditions. **(A)** Bending stress for the planet gears—stage 1. **(B)** Contact stress for the planet gears—stage 1. **(C)** Main bearing MB-B equivalent radial dynamic load. **(D)** Gearbox bearing HSS-DE equivalent radial dynamic loads. **(E)** Generator bearing GEN-NDE equivalent radial dynamic load.

met, as demonstrated in Section 3.5. Multiple damage indices for a given wind speed and yaw angle in Figure 9 are associated with the different values of power demand. As the farm is operating in the power curtailment mode of operation, the power set point and

therefore the generated power of each turbine can be different from the available power.

To gain an understanding of how each parameter participates in the damage, a sensitivity analysis is performed, with results as

FIGURE 7
Vulnerability map of the powertrain while taking into consideration the wake steering controller.

FIGURE 8
Powertrain weighted damage index vs. mean wind speed, turbulence intensity, power demand, and yaw misalignment conditions. Plain/hollow markers show conditions when yawing is detrimental/beneficial, respectively, regarding the overall powertrain fatigue damage.

presented in Figure 10. Figure 10B shows that damage increases as turbulence intensity increases, as expected. Figure 10A shows that the damage index increases as the mean wind speed increases up to the rated mean wind speed 11.4 m/s, due to the increased mean thrust force and its effect on main bearing damage. After the rated wind speed is reached, there is a mitigating trend in the overall damage attributed to the reduction of thrust force (both mean value and fluctuations) through pitch control. For high wind speeds and/or turbulence intensities, gears dominate damage—the damage of gears is mainly driven by mean torque and torque oscillations, not thrust—showing a different damage pattern increasing linearly with demanded power. More interesting is the relationship between damage and yaw misalignment angle shown in Figure 10C, depicting in more detail the preliminary observations on a possible reduction of damage through yawing in downregulated conditions. It shows that the combined effect of downregulating and

yawing is nontrivial and cannot be generalized in a simple control law; case-by-case lookup in the database should be used in the farm controller.

3.5 Applications in farm control

The same setup as used in Moghadam et al. (2023) has been chosen for comparability:

- Turbine layout: 32 turbines, staggered, 5-diameter spacing (Andersen et al., 2018).
- Wind speed: 12.4 m/s—damage is largely dominated by the main bearing.
- Wind direction: North, i.e., perpendicular to dominant (featuring maximum wake effect, eight rows with 5D spacing, starting from turbines T29 to T32).
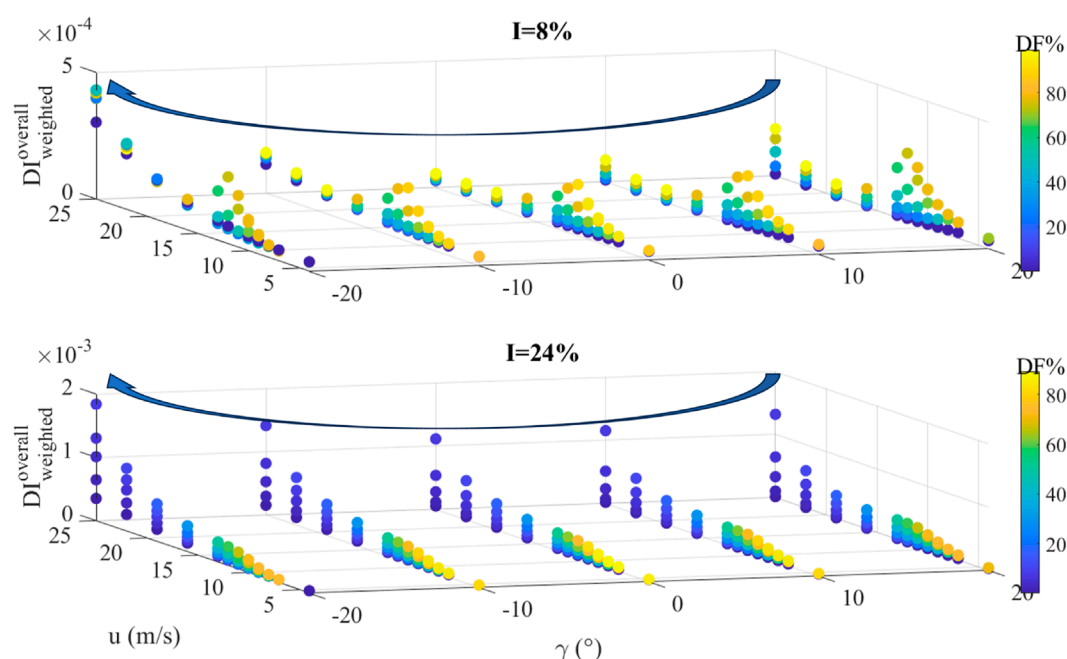
FIGURE 9
3D sensitivity analysis to see the influence of yaw misalignment angle on the powertrain damage for a wide range of wind speed variations.

- Wind fluctuations: obtained from the model of Sørensen et al. (2008) and Vigueras-Rodríguez et al. (2010, 2012).
- Tunable mapping functions g_i: same for both $i \in \{1, 2\}$ readings.

$$g_i \left(DI_i \right) = 1 - 0.5 \left(\frac{DI_i - \inf\limits_{i \in N} DI_i}{\sup\limits_{i \in N} DI_i - \inf\limits_{i \in N} DI_i} \right). \tag{3}$$

It introduces an affine re-scaling of the damage index DI_i, which ensures that turbines with low DI_i values will contribute more to the power tracking task (g_i takes values closer to 1), while turbines with higher DI_i values will contribute less (g_i takes values closer to 0.5) and with a minimum contribution that is lower bound by half of the turbine's naive contribution (without any damage information). This tuning is not cost-optimized and only devised for qualitative demonstration purposes. The *sup* and *inf* operators represent the supremum and infimum for the set DI_i.

The goal is to show how the yaw angle can be used by the controller as an additional control input to improve the fatigue mitigation feature. In total, 18 simulations are run corresponding to two control objective scenarios, three control methods, and three curtailment profiles. The two control objective scenarios are

- Scenario 1: All turbines are equally damaged. The objective is to reduce the overall damage and spread it evenly over the turbines. This case has a high cost-saving potential but would require a global cost–benefit analysis that includes all turbine components.
- Scenario 2: One particular turbine (turbine 1) is damaged. The objective is to minimize further damage on this turbine only. This case is particularly relevant as powertrain components typically show early signs of failure, motivating the use of

damage mitigating control to influence the remaining lifetime and coordinate with maintenance actions, thus reducing downtime.

The three control methods are

- Method 1: No fatigue damage mitigation (equalLambdas),
- Method 2: Fatigue damage mitigation without yaw angle control as in Moghadam et al. (2023) (loadMitigation),
- Method 3: Fatigue damage mitigation with yaw angle control (loadMitigation yaw).

The three curtailment profiles correspond to constant power references equal to 0.2, 0.5, and 0.8 per unit (pu) of the installed farm power (320 MW).

Power dispatch commands follow the control scheme in Figure 3 and are updated every minute. Yaw angle commands are selected to yield minimum damage on the turbine(s) of interest (all in scenario 1 and only turbine 1 in scenario 2) and updated every 10 min, following the quasi-steady update rate of the waked flow field, by the simulation framework described in Section 2.2.

Results for scenario 1 are shown in Figure 11 with one subfigure per curtailment profile. It can be seen that power dispatch only can even out damage between upstream and downstream turbines but does not significantly reduce overall damage. Adding the yaw degree of freedom, on the other hand, aims at decreasing damage and is particularly efficient in heavily curtailed conditions, which is consistent with the database simulation results.

Results for scenario 2 are shown in Figures 12, 13, where the power dispatch (moving average) and the corresponding time evolution of the damage accumulation for the damaged turbine
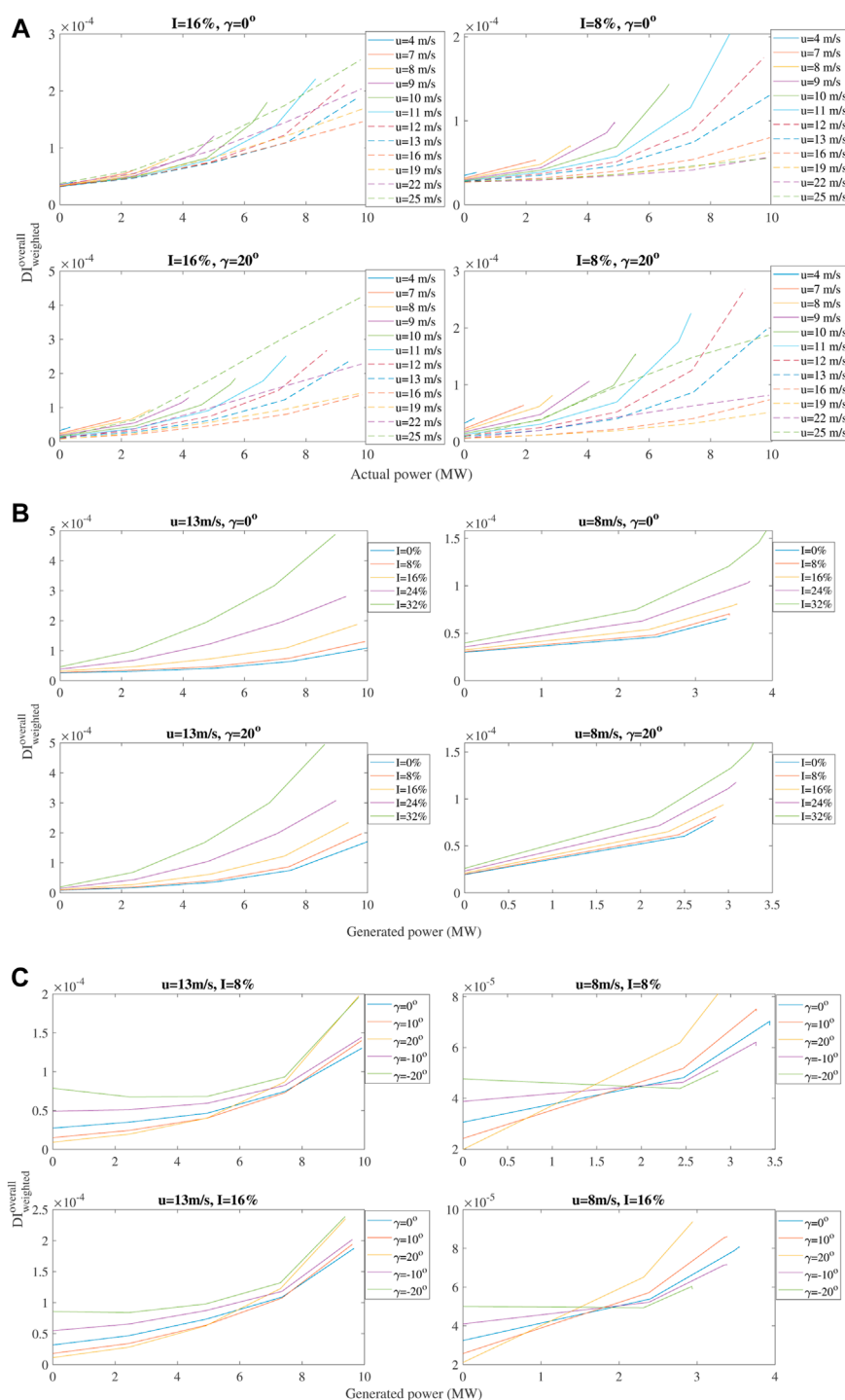
FIGURE 10
Powertrain fatigue damage vs. turbine-generated power sensitivity analysis. **(A)** Damage index vs. generated power at different wind speeds. **(B)** Damage index vs. generated power at different turbulence intensities. **(C)** Damage index vs. generated power at different yaw misalignment angles.

can be observed. Initially, P_{pcc} is naively distributed among the turbines. After the first control update based on the database at 600 s (because it requires information from the last 10 min), the farm controller has realized that it can track P_{pcc} with less contribution from the damaged turbine, and turbine 1 is re-dispatched to lower power levels. Both versions of the farm controller calculate the same

powerdispatch plan for turbine 1, as observed from the moving average of the power dispatch in Figure 12A. However, since the proposed, enhanced version can also set the yaw angle of turbine 1 to a proper value, at the end of the simulation period, turbine 1 has accumulated less damage (solid purple line in Figure 12) than the case without a yaw angle offset (dashed green line in Figure 12). On
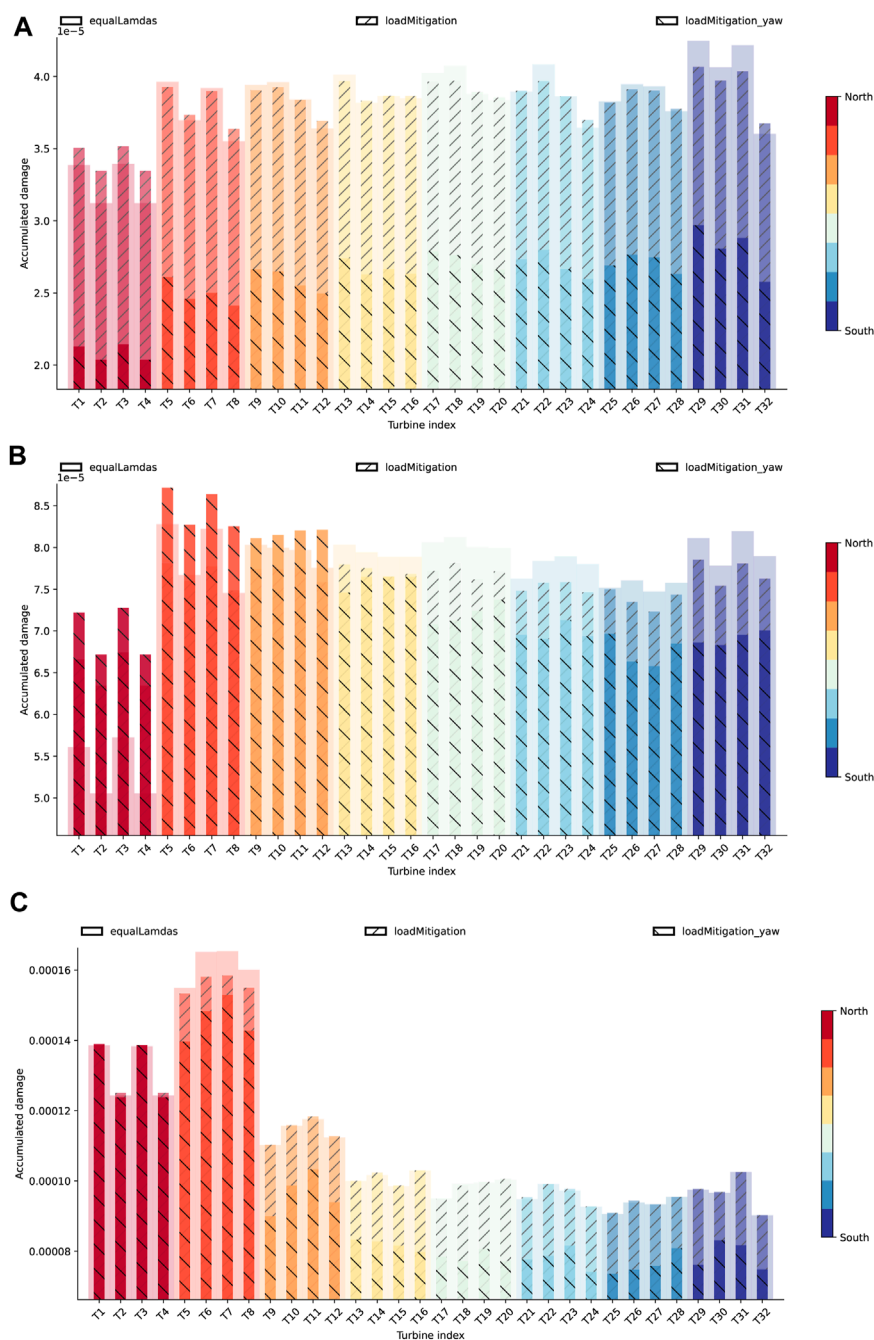
FIGURE 11
Distribution of accumulated damage over the farm turbines under different operational strategies. **(A)** Farm reference power level 0.2 pu (north wind, 12.4 m/s). **(B)** Farm reference power level 0.5 pu (north wind, 12.4 m/s). **(C)** Farm reference power level 0.8 pu (north wind, 12.4 m/s).

the contrary, the naive approach that does not consider any damage information (solid red lines in Figure 12) instructs a different power dispatch plant for turbine 1, which would result in even higher damage accumulation by the end of the simulation. The simulated case's corresponding optimal yaw angle offset sequence (adopted only by the controller version with yaw information) is presented in Figure 12B.

Similar results were also observed for the case of farm reference power level 0.2 pu, and in general, for relatively lower values of

farm reference power (equivalently higher levels of curtailment). For relatively higher values of farm reference power, however, the effectiveness of the proposed control scheme is lower. To illustrate this, we present the wind farm simulation results for the case of farm reference power 0.8 pu, shown in Figure 13, and following the same format as for the case of 0.5 pu (Figure 12). As can be observed, in this case, the power dispatch for the damaged turbine (turbine 1) is the same for both versions of the controller (with and without yaw), and for most of the simulation period, they are
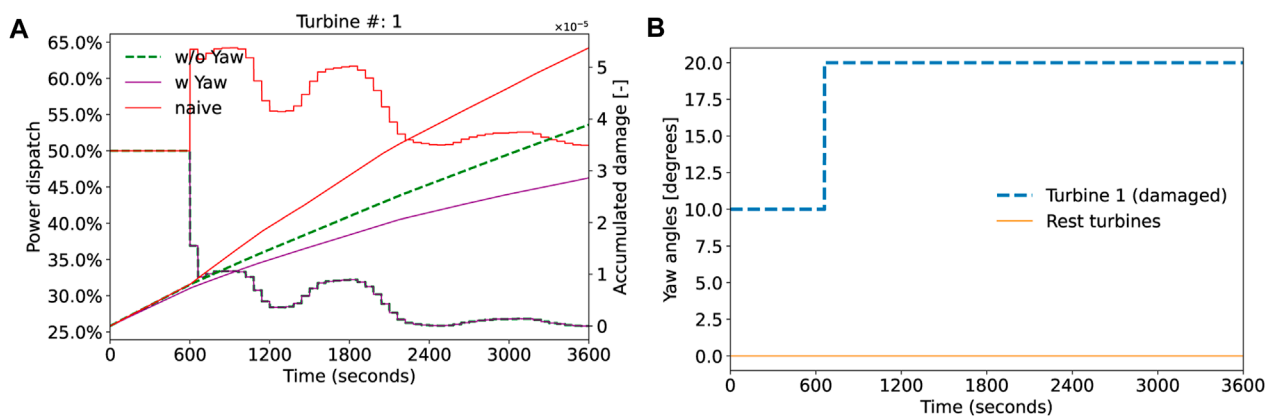
FIGURE 12
Wind farm simulation and effects of the wind farm controller on the damaged turbine (farm reference level 0.5 pu). **(A)** Power dispatch (moving average) and damage accumulation. **(B)** Yaw angle offset (case with Yaw).
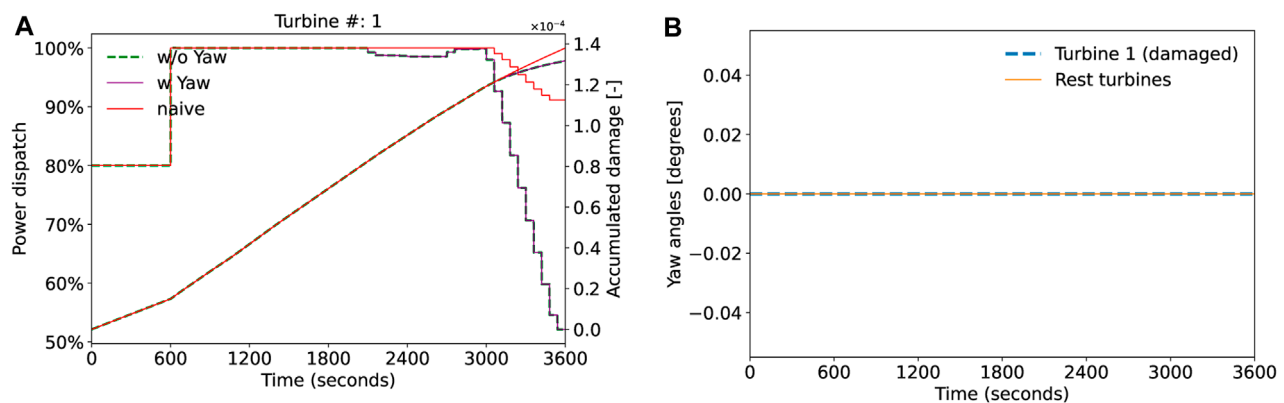
FIGURE 13
Wind farm simulation and effects of the wind farm controller on the damaged turbine (farm reference level 0.8 pu). **(A)** Power dispatch (moving average) and damage accumulation. **(B)** Yaw angle offset (case with Yaw).

even the same with the naive approach. This is primarily because the requested farm power is relatively high when compared to the available, for this case, considering the pessimistic wind direction (north wind) that is associated with the largest impact from the wake effect. It is only at the last part of the simulation period that the farm controller can re-dispatch turbine 1 since the farm output can be tracked with less turbine 1 contribution. This is shown in Figure 13A, where the power decreases to lower values after 3,000 s. However, the decrease for the controller versions that take into account the damage is much more significant when compared to the naive approach, and as a consequence of this, the accumulated damage for the turbine is slightly lower than the naive approach by the end of the simulation. Again, the corresponding optimal yaw angle offset sequence is presented in Figure 13B. In this case, we observe that no offset yaw angle is decided, meaning that both versions of the controller (with and without yaw) are identical, which also justifies the identical damage accumulation for both versions, as depicted in Figure 13A. This observation is in line with the updated results of the

damage database when including the yaw misalignment capability, where for higher turbine-generated power (lower derating levels), the effect of the yaw on turbine damage (based on the weighted damage index) becomes less significant.

4 Conclusion

A methodology characterizing powertrain fatigue damage for use in farm control is suggested. This study focuses on the effect of yaw misalignment, enabling yaw angle control at the farm level for damage mitigation purposes. It complements previous work on powertrain degradation modeling for multi-objective farm control in curtailed conditions, where it was shown how the damage on individual turbines may be steered while tracking a below-maximum power reference from the grid operator by acting on the power set points sent to individual turbines (the so-called power dispatch). The farm controller makes use of a database

mapping operating conditions (wind speed, turbulence intensity, power set point, and yaw misalignment) to an overall indicator of powertrain degradation. It is built from numerous aero-servo-elastic turbine simulations feeding a quasi-static degradation model of each powertrain component—bearings and gears—following design standards, whose results are assembled to give a single metric for degradation. It is found that yawing-downregulated—asked to produce less power than available—turbines may further reduce damage, especially for the main bearing. This finding is then exploited in farm-level case studies with various levels of curtailment, where the farm controller reacts to local wind variations from turbulence and wakes, with power tracking as a primary objective and damage mitigation as a secondary objective. This demonstration of damage-aware farm control shows the benefit of adding the yaw degree of freedom to the load mitigation functionality.

These results are readily valuable when the objective is to reduce downtime on a damaged upstream turbine showing early signs of failure in a powertrain component. Future work will be devoted to quantifying uncertainty in the database by using high-fidelity models for wake effects and including the fatigue damage of other turbine systems, namely, blades, pitch bearings, and support structure, focusing particularly on combined yawing and downregulation. This will enable the extension of the proposed farm control methodology to holistic farm-wide lifetime extension and optimal maintenance planning purposes.

Data availability statement

To access the raw data of this research please contact the corresponding author at farid.k.moghadam@ntnu.no.

References

Andersen, S. J., Madariaga, A., Merz, K., Meyers, J., Munters, W., and Rodriguez, C. (2018). *Reference wind power plant d1. 03.*

Azzam, B., Schelenz, R., and Jacobs, G. (2022). Sensor screening methodology for virtually sensing transmission input loads of a wind turbine using machine learning techniques and drivetrain simulations. *Sensors* 22, 3659. doi:10.3390/s22103659

Azzam, B., Schelenz, R., Roscher, B., Baseer, A., and Jacobs, G. (2021). Development of a wind turbine gearbox virtual load sensor using multibody simulation and artificial neural networks. *Forsch. Ing.* 85, 241–250. doi:10.1007/s10010-021-00460-3

Bak, C., Zahle, F., Bitsche, R., Kim, T., Yde, A., Henriksen, L., et al. (2013). *The dtu 10-mw reference wind turbine.* DTU Library, Denmark: Danish Wind Power Research. 2013; Conference date: 27-05-2013 Through 28-05-2013.

Cardaun, M., Roscher, B., Schelenz, R., and Jacobs, G. (2019). Analysis of wind-turbine main bearing loads due to constant yaw misalignments over a 20 years timespan. *Energies* 12, 1768. doi:10.3390/en12091768

Chabaud, V. (2023). Synthetic turbulence modelling for offshore wind farm engineering models using coherence aggregation. *Wind Energy.* doi:10.1002/we.2875

Frandsen, S. (2007). *Turbulence and turbulence-generated structural loading in wind turbine clusters*

Hansen, M. H., and Henriksen, L. C. (2013). *Basic DTU wind energy controller. Tech. Rep. DTU Wind Energy No.0028.*

Hart, E., Clarke, B., Nicholas, G., Kazemi Amiri, A., Stirling, J., Carroll, J., et al. (2020). A review of wind turbine main bearings: design, operation, modelling, damage mechanisms and fault detection. *Wind Energy Sci.* 5, 105–124. doi:10.5194/wes-5-105-2020

IEC 61400-1:2019 (2019). *Iec 61400-1:2019: Wind energy generation systems, Part 1: Design requirements.* Geneva, Switzerland: Standard, International Electrotechnical Commission.

IEC 61400-3-1:2019 (2019). *Iec 61400-3-1:2019: Wind energy generation systems, Part 3-1: Design requirements for fixed offshore wind turbines. Standard.* Geneva, Switzerland: International Electrotechnical Commission.

Knudsen, T., Bak, T., and Svenstrup, M. (2015). Survey of wind farm control—power and fatigue optimization. *Wind Energy* 18, 1333–1351. doi:10.1002/we.1760

Laboratory, N. R. E. (2023). *Wake models.*

Madsen, H. A., Larsen, G. C., Larsen, T. J., Troldborg, N., and Mikkelsen, R. (2010). *Calibration and validation of the dynamic wake meandering model for implementation in an aeroelastic code.*

Meng, F., Lio, W., and Barlas, T. (2020). DTUWEC: an open-source DTU wind energy controller with advanced industrial features. *J. Phys. Conf. Ser. (Online)* 1618, 022009. doi:10.1088/1742-6596/1618/2/022009

Merz, K., Chabaud, V., Garcia-Rosa, P. B., and Kolle, K. (2021). "A hierarchical supervisory wind power plant controller," in Journal of Physics: Conference Series (IOP Publishing), vol. 2018, Ancona, Italy, 20–22 June 2018.012026

Moghadam, F. K., Chabaud, V., Gao, Z., and Chapaloglou, S. (2023). Power train degradation modelling for multi-objective active power control of wind farms. *Forsch. Im. Ingenieurwes.* 87, 13–30. doi:10.1007/s10010-023-00617-2

Moghadam, F. K., and Desch, N. (2023). Life cycle assessment of various pmsg-based drivetrain concepts for 15 mw offshore wind turbines applications. *Energies* 16, 1499. doi:10.3390/en16031499

Moghadam, F. K., and Nejad, A. R. (2020). Evaluation of pmsg-based drivetrain technologies for 10-mw floating offshore wind turbines: pros and cons in a life cycle perspective. *Wind Energy* 23, 1542–1563. doi:10.1002/we.2499

Author contributions

FK: Writing–original draft. ZG: Writing–review and editing. VC: Writing–original draft. SC: Writing–original draft.

Funding

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.

Moghadam, F. K., and Nejad, A. R. (2022). Online condition monitoring of floating wind turbines drivetrain by means of digital twin. *Mech. Syst. Signal Process.* 162, 108087. doi:10.1016/j.ymssp.2021.108087

Moghadam, F. K., Rebouças, G. F. d. S., and Nejad, A. R. (2021). Digital twin modeling for predictive maintenance of gearboxes in floating offshore wind turbine drivetrains. *Forsch. Im. Ingenieurwes.* 85, 273–286. doi:10.1007/s10010-021-00468-9

Peeters, J. L., Vandepitte, D., and Sas, P. (2006). Analysis of internal drive train dynamics in a wind turbine. *Wind Energy An Int. J. Prog. Appl. Wind Power Convers. Technol.* 9, 141–161. doi:10.1002/we.173

Pfaffel, S., Faulstich, S., and Rohrig, K. (2017). Performance and reliability of wind turbines: a review. *energies* 10, 1904. doi:10.3390/en10111904

Sood, I., del Fosse et d'Espierres, C., and Meyers, J. (2023). Quasi-static closed-loop wind-farm control for combined power and fatigue optimization. Available at: https://arxiv.org/pdf/2305.11710.pdf.

Sørensen, P., Cutululis, N. A., Vigueras-Rodríguez, A., Madsen, H., Pinson, P., Jensen, L. E., et al. (2008). Modelling of power fluctuations from large offshore wind farms. *Wind Energy An Int. J. Prog. Appl. Wind Power Convers. Technol.* 11, 29–43. doi:10.1002/we.246

Stock, A., Cole, M., Leithead, W., and Amos, L. (2020). "Distributed control of wind farm power set points to minimise fatigue loads," in 2020 American Control Conference (ACC), Denver, Colorado, USA, 1-3 July 2020 (IEEE), 4843–4848.

van Binsbergen, D. W., Wang, S., and Nejad, A. R. (2020). Effects of induction and wake steering control on power and drivetrain responses for 10 mw floating wind turbines in a wind farm. In Journal of Physics: Conference Series (IOP Publishing), vol. 1618, 022044.

van Dinter, R., Tekinerdogan, B., and Catal, C. (2022). Predictive maintenance using digital twins: a systematic literature review. *Inf. Softw. Technol.* 151, 107008. doi:10.1016/j.infsof.2022.107008

Vigueras-Rodríguez, A., Sørensen, P., Cutululis, N. A., Viedma, A., and Donovan, M. (2010). Wind model for low frequency power fluctuations in offshore wind farms. *Wind Energy* 13, 471–482. doi:10.1002/we.368

Vigueras-Rodríguez, A., Sørensen, P., Viedma, A., Donovan, M. H., and Lázaro, E. G. (2012). Spectral coherence model for power fluctuations in a wind farm. *J. Wind Eng. Industrial Aerodynamics* 102, 14–21. doi:10.1016/j.jweia.2011.12.006

Wang, S., Nejad, A. R., and Moan, T. (2020). On design, modelling, and analysis of a 10-mw medium-speed drivetrain for offshore wind turbines. *Wind Energy* 23, 1099–1117. doi:10.1002/we.2476

Wiser, R., Bolinger, M., Hoen, B., Millstein, D., Rand, J., Barbose, G., et al. (2023). *Land-based wind market report: 2023 edition. Tech. Rep.* Berkeley, CA (United States): Lawrence Berkeley National Laboratory LBNL.

Xia, J., and Zou, G. (2023). Operation and maintenance optimization of offshore wind farms based on digital twin: a review. *Ocean. Eng.* 268, 113322. doi:10.1016/j.oceaneng.2022.113322

Yasuda, Y., Bird, L., Carlini, E. M., Eriksen, P. B., Estanqueiro, A., Flynn, D., et al. (2022). Ce (curtailment–energy share) map: an objective and quantitative measure to evaluate wind and solar curtailment. *Renew. Sustain. Energy Rev.* 160, 112212. doi:10.1016/j.rser.2022.112212

Zhang, S., Wei, J., Xu, Z., Tang, B., and Niu, R. (2021). Research on the influence of system parameters on the electromechanical dynamics of a large wind turbine drivetrain. *Energy Rep.* 7, 7835–7851. doi:10.1016/j.egyr.2021.11.020

Architecting a digital twin for wind turbine rotor blade aerodynamic monitoring

Yuriy Marykovskiy[1,2]*, Thomas Clark[3], Julien Deparday[1], Eleni Chatzi[2] and Sarah Barber[1]*

[1]Institute of Energy Technology, Eastern Switzerland University of Applied Sciences, Rapperswil, Switzerland, [2]Chair of Structural Mechanics and Monitoring, ETH Zürich, Zurich, Switzerland, [3]Octue Ltd., Cambridge, United Kingdom

Digital twins play an ever-increasing role in maximising the value of measurement and synthetic data by providing real-time monitoring of physical systems, integrating predictive models and creating actionable insights. This paper presents the development and implementation of the Aerosense digital twin for aerodynamic monitoring of wind turbine rotor blades. Employing low-cost, easy-to-install microelectromechanical (MEMS) sensors, the Aerosense system collects aerodynamic and acoustic data from rotor blades. This data is analysed through a cloud-based system that enables real-time analytics and predictive modelling. Our methodological approach frames digital twin development as a systems engineering problem and utilises design patterns, design thinking, and a co-design framework from applied category theory to aid in the development process. The paper details the architecture, deployment, and validation of a 'Digital Shadow'-type twin with simulation/prediction functionalities. The solution pattern is discussed in terms of its implementation challenges and broader applicability. By providing a practical solution to integrating all the digital twin components into a holistic system, we aim to help wind energy specialists learn how to transform a conceptual idea of a digital twin into a functional implementation for any application.

1 Introduction

1.1 Rotor blade aerodynamic monitoring

Due to the increasing size and flexibility of wind turbine rotor blades, it is becoming more and more important to measure and monitor their aerodynamic and acoustic behaviour operation (Schepers and Schreck, 2019). This can help wind turbine manufacturers (OEMs) to improve their aerodynamic models and blade designs, owner/operators to optimise operation and researchers to understand complex aerodynamic phenomena. The complexity of the installation and use of such measurement systems means that there have not yet been a large number of publications on the topic, despite the increasing demand. For example, the DanAERO project from 2013–2016 investigated the aerodynamic and acoustic properties of wind turbine blades in wind tunnel and field

tests (Madsen et al., 2016; Troldborg et al., 2013). The field tests included instrumenting a 2 MW wind turbine rotor blade with 50 flush-mounted microphones. It was shown that such aero-acoustic field measurements have the potential to provide a high added value to the wind industry through furthered understanding of three-dimensional effects. Furthermore, the international collaboration IEA Wind Task 47 aims to cooperate and share experiences in aerodynamic measurements on Megawatt-scale wind turbines[1]. A more detailed study of previous literature can be found in the introduction of a recent paper by the current authors (Barber et al., 2022). It can be summarised that there is a large unmet need for easy-to-install and affordable rotor blade aerodynamic monitoring systems in the sector.

1.2 Digital twins

Digital twins are a promising technology for creating value from wind turbine monitoring, such as rotor blade aerodynamic monitoring, particularly when a scale-up of a single measurement campaign is desired. A "digital twin" is a top-level conceptualisation based on two fundamental principles: "duality" and "strong similarity" (Grieves, 2022). The digital twin paradigm thus spans such a broad range of applications that two particular implementations may share few - or no - technological solutions between them. Recently, the current authors proposed a digital twin classification system based on a Simple Knowledge Organisation System (SKOS) (W3C, 2009) data model, to act as a starting point for the development of digital twins by allowing comparison or solution re-use between implementations (Marykovskiy et al., 2023a). An excerpt from the classification system is presented in Figure 1, referencing some digital twin types that have previously been described in literature, such as `DigitalTwinPrototype` and `DigitalTwinInstance` described by Grieves and Vickers (2017), `DigitalModel`, `DigitalShadow`, `DigitalTwin` described by Kritzinger et al. (2018), and various types categorised by their functionalities as described by Wagg et al. (2020).

When classified based on their functional capabilities (Wagg et al., 2020), digital twins can range from 'Supervisory', in which data from measurements is simply ingested and stored, to 'Operational', in which analysis of the operational data is undertaken, to 'Simulation/Prediction', in which models, simulations, validation, and verification and uncertainty quantification enhance the measurements, to 'Intelligent/Learning', which includes Decision Support Systems (DSS), and finally to 'Autonomous/Management', in which autonomous asset control is implemented. According to Wagg et al. (2020), the main transformative aspect of a digital twin is to improve the predictive capability of a system by augmenting computational models with data to create a virtual prediction tool that can evolve over time. 'Intelligent/Learning' digital twins have recently been shown to allow accelerated and informed decision-making related to physical systems by representing them virtually and including a continuous feedback loop between the virtual representation and a physical

```
DigitalTwinConceptualModel
|___ DigitalTwinSystem
|       |___ PhysicalSystem
|       |___ DigitalSystem
|       |___ ConnectionSystem
|___ DigitalTwinType
        |___ PhysicalSystemLifetimeStageType
        |       |___ DigitalTwinPrototype
        |       |___ DigtialTwinInstance
        |___ ConnectionSystemAutomationType
        |       |___ DigitalModel
        |               |___ DigitalShadow
        |                       |___ DigitalTwin
        |___ SystemFunctionalityType
                |___ SupervisoryDigitalTwin
                        |___ OperationalDigitalTwin
                                |___ SimulationPredictionDigitalTwin
                                        |___ IntellegentLearningDigitalTwin
                                                |___ AutonomousManagementDigitalTwin
```

FIGURE 1
Excerpt from Digital Twin Conceptual Model (DTCM) SKOS taxonomy (Marykovskiy et al., 2023a).

system (Arista et al., 2023; D'Amico et al., 2022; Grieves, 2022; Wagg et al., 2020; Zheng et al., 2021).

1.3 Developing digital twins in the wind energy context

Recently, publications related to digital twins and DSS in wind energy were reviewed by Marykovskiy et al. (2024) in the broader context of artificial intelligence systems and domain semantics. Most digital twin implementations in wind energy were found to belong to the functional levels 'Supervisory' (26 out of 111), 'Operational' (22) or 'Simulation-Prediction' (60). Only three papers belong to the functional levels 'Intelligent-Learning' (2) and 'Autonomous-Management' (1). For wind energy Operations and Management (O&M), previous 'Supervisory' or 'Operational' digital twins included continuous structural monitoring of a wind farm (Hines et al., 2023). 'Simulation/Prediction' digital twins included an augmented Kalman filter with a reduced mechanical model to estimate tower wind turbine loads (Branlard et al., 2020), integration of degradation processes in a strategic offshore wind farm O&M simulation model using a Markov process for blade degradation (Welte et al., 2017), and modelling the probabilistic characteristics of mooring line fatigue stresses for the purpose of risk-based inspection (Lone et al., 2022). 'Intelligent/Learning'

digital twins include a probabilistic framework for updating the structural reliability of offshore wind turbine substructures based on digital twin information (Augustyn et al., 2021).

This distinction based on functional capabilities, however, is not the only possible classification of digital twin implementations. As we can see from Section 1.2, it is also possible to distinguish and group digital twins based on the manner in which the connection between the physical object and its digital representation is achieved or based on the fidelity levels of simulations employed. Here, the specific technological solutions used in composing the system into the one whole will highly depend on the use case and its requirements.

The sheer variety of different types and applications of digital twins has left many wind energy domain specialists struggling to clarify development processes that suit their specific needs: the passage from the conceptual idea of a digital twin to a functional implementation is often unclear. There exist a myriad of technology stacks, modelling and simulation tools, algorithms and system integration requirements, the selection of which is nuanced and made more complex by technical and specialised jargon. This is not helped by the focus of previous literature on the physical models, rather than on the system architecture.

In its core, the digital twin conceptual model does not necessarily imply an introduction or development of new modelling techniques or simulation applications. However, practical solutions to integrate all the digital twin components into a holistic system, with all of its constituents interconnected and valorised, remain obscure. For instance, one possibility to enable system orchestration is through the use of semantic artefacts. The term 'semantic artefact' is used to denote conceptualisations with various degree of expressiveness, such as controlled vocabularies, taxonomies, schemas and ontologies (Le Franc et al., 2020). However, in the aforementioned review by Marykovskiy et al. (2024), a lack of adoption of semantic artefacts in the research of digital twin and DSS was found, reflected by the low number of publications that refer to them (35 out of 181). This can be attributed to multitude of factors that plague multi- and interdisciplinary developments such as a natural tendency towards knowledge siloing within organisations and communities (wind energy from information technology professionals, industry from academia, etc.) as well as to the overall digitalisation challenges in the areas of data, culture and coopetition (Clifton et al., 2023). There is therefore a high value for the wind energy community to present developed digital twin instances from system architecture and technology implementation points of view.

1.4 The Aerosense system

The Aerosense system was developed to address the high demand for easy-to-use and cost effective rotor blade aerodynamic monitoring systems combined with a high potential of digital twin applications in this field, as mentioned in the previous two sections. Aerosense is a cost-effective microelectromechanical systems (MEMS)-based aerodynamic and acoustic wireless measurement system that is thin, non-intrusive, easy to install, low power, and self-sustaining, which was previously introduced by the authors of

this present paper (Barber et al., 2022). The hardware is composed of sensor nodes installed on the blade and a base station receiving and sending the data to the cloud (Figure 2A). Figure 2B shows a sensor node of the Aerosense measurement system installed on a wind turbine blade.

Previous publications related to this work have focused strongly on the hardware development, showing that the sensors are capable of delivering relevant results continuously in the wind tunnel (Barber et al., 2022; Polonelli et al., 2023a). Additionally, various methods for using the measurements to provide added value to the wind energy industry have been introduced, including Leading Edge Erosion (LEE) detection and classification (Duthé et al., 2021), inferring angle of attack and wind speed (Marykovskiy et al., 2023c), detecting structural damage (Abdallah et al., 2022) and flow-field reconstruction (Duthé et al., 2023). The overall design of the digital twin, including software integration and the cloud data storage design, has not yet been discussed.

1.5 This contribution

In this paper, we present and demonstrate the top-level system design of a digital twin for wind turbine rotor blade aerodynamic monitoring, which was developed as part of the Aerosense project. By providing a practical solution to integrating all the digital twin components into a holistic system, we aim to help wind energy specialists learn how to transform a conceptual idea of a digital twin into a functional implementation for any application. In Section 2 we present the system architecture of the Aerosense digital twin from a conceptual point of view. Then, we discuss the system design in Section 3, with a focus on the cloud data storage solution and the software integration. In Section 4 we present the results of a field test case, including the test set-up, the measurement results, and the demonstration of added value. In Section 5 we discuss its wider application, and in Section 5.3 we present the conclusions.

2 Architecture of the Aerosense digital twin

A multitude of design methodologies, decision support tools, and optimisation algorithms exist for facilitating design and architecting processes in general. Here we used several well-established methodologies including design thinking (Pearce, 2020), design patterns (Gamma et al., 1994; Tekinerdogan and Verdouw, 2020), decision trees and applied category theory (Censi, 2016; Zardini et al., 2021). According to the design patterns approach, before developing a concrete realisation of a digital twin, it is opportune to establish the desired digital twin type. Type selection is guided by the context in which the development of the digital twin is occurring. Adopting the design thinking methodology, the use case for the Aerosense digital twin is therefore first presented (Section 2.1), followed by digital twin type selection (Section 2.2), which served as a starting point in establishing the overall system architecture (Section 2.3).
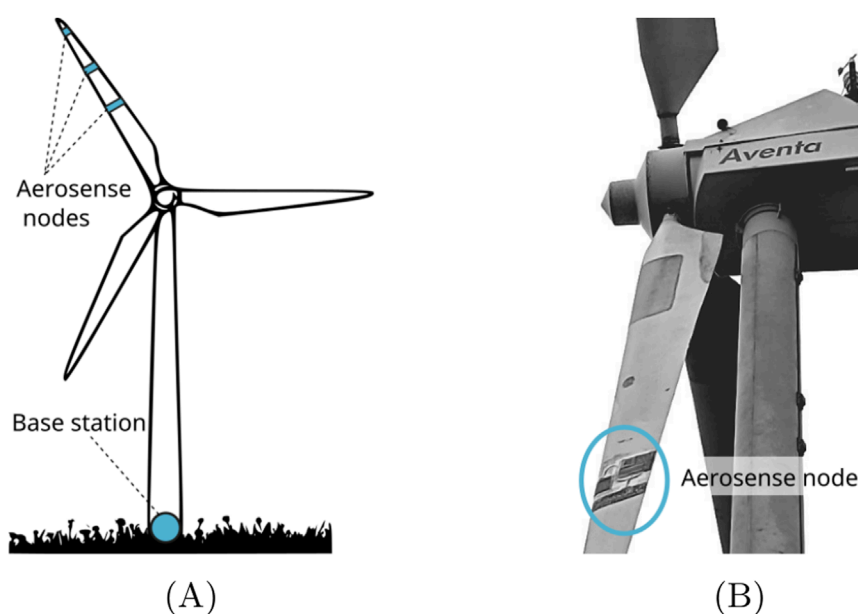
FIGURE 2
Aerosense system sensor hardware and its placement on a wind turbine blade. **(A)** General concept of the Aerosense system. **(B)** View of an Aerosense measurement system installed on a 6 m long wind turbine blade.

2.1 The use case

In order to define the priority use case for the Aerosense project, a design thinking strategy was applied. Design thinking is an iterative methodology for framing problems and co-creating implementable solutions using visual thinking and prototyping (Pearce, 2020). It consists of the phases "Empathise", "Define", "Ideate", "Prototype", "Test" and "Implement". For the "Empathise" phase, extensive "user story" interviews were carried out with potential customers from both industry and academia at the beginning of the project, in which several imagined but realistic "user stories" were presented and discussed. The results were used to define and prioritise the most important use cases for the "Define" phase. Through this process, we discovered that the Aerosense system has a high potential to provide OEMs, owner/operators and researchers with added value, including to improve aero-elastic models, detect and classify surface damage, and even detect structural damage. For the remaining phases, we applied a design pattern methodology, as discussed in the next sections.

Analysing a variety of use cases revealed one *foundational* application. Since wind turbines have grown larger and more flexible in recent years, established 2D assumptions used for aerodynamic tools have become less likely to hold valid (Bangga et al., 2017). Thus, one of the use cases, "improving aerodynamic models", was seen as most important, underpinning further analysis or damage detection methods. The beneficiaries, value statements and required outputs of this use case are given in Table 1. This information was used as a design basis for the system, and will be revisited in Section 4.

The required outputs from Table 1 can be summarised as functionality requirements for interactive dashboards (required outputs indicated by letter 'a') and Colab notebooks, which is a cloud-hosted Jupyter notebooks service (required outputs indicated by letters 'b' through 'e'). For dashboards, these functionalities include visualisation, exploration, and inspection of sensors' time series data and pressure coefficient distributions through interactive plots. Colab notebooks, on the other hand, allow for more flexible and custom uses, more accommodating of the defined user stories. Here, the main functionality to ensure is access to the sensor data, data processing algorithms, and simulations for further data transformation and analysis. The data processed in Colab notebooks can also be visualised, explored, and inspected through interactive plots.

2.2 Type classification

Comparing the digital twin classifications of Figure 1 to the required outputs from the use case exercise in Table 1, the `DigitalTwinType` of the Aerosense digital twin was classified as follows:

- `PhysicalSystemLifetimeStageType:`
  `DigitalTwinInstance`
  (because the intention is to work with an existing instance of a wind turbine, not, say, a prototype of a turbine not yet in existence)
- `ConnectionSystemAutomationType:`
  `Digital Shadow`
  (because there is a one-way automated connection from physical to digital system, as opposed to two-way, which would enable control or other adaptive behaviour)
- `SystemFunctionalityType:`
  `SimulationPredictionDigitalTwin`

TABLE 1 Description of the use case "improving aerodynamic models" for this work.

Beneficiary	Value added	Required output
1. OEM measurement and data science teams	1a. Allow initial data inspection and download of measurement data for further analysis or comparison by the customer	1a. Corrected and calibrated time series data for each sensor, classified based on operating conditions, available on a dashboard
	1b. Allow detailed analyses such as extracting data and plotting pressure distributions to be carried out based on the time series and averaged data according to the needs of the customer or partner, without having to write a new code that works with the downloaded data	1b. Colab notebook for analysis of time series and averaged data
2. OEM aerodynamic modelling teams	2a. Enable data exploration for improved understanding of the aerodynamic behaviour	2a. Pressure coefficient distribution plots for specific time instances available on a dashboard
	2b. Enable pressure distributions to be plotted, examined and compared	2b. Installed sensor placement obtained via photogrammetry available through an API and accessible via a Colab notebook
	2c. Enable direct comparisons with data from the customer, e.g., measured or simulated 2D pressure coefficient distributions	2c. Phase-averaged pressure and pressure coefficient distributions at different operating conditions available in a Colab notebook
	2d. Enable inference of the angle of attack at the sensor location, which allows measured pressure distributions to be compared with measured or simulated 2D pressure distributions at different angles of attack	2d. Aerofoil inflow inference model with uncertainty quantification available in a Colab notebook
	2e. Allow direct comparison of observed pressure distributions with 2D simulation results, leading to an improved understanding of the aerodynamic behaviour of the wind turbine in the field, validation of the fundamental assumptions adopted during modelling, as well as to recommendations for the improvement of aerodynamic models	2e. Simulated vs. measured phase-averaged pressure distributions at different angles of attack available in a Colab notebook

(because the extent of the use case outputs include simulation and prediction applications, incorporating operational and supervisory aspects like visualisation of system state)

Each of the aforementioned types is accompanied by a specific design pattern (Tekinerdogan and Verdouw, 2020) reflected in the overall digital twin architecture and system hardware implementations as discussed in the following section.

2.3 Aerosense digital twin conceptual model and related hardware

Generally, a digital twin system can be conceptually divided into three main sub-systems: physical system, digital system, and connection system. Sensors, in general, are considered to be a part of the physical system (Singh et al., 2021; Tao et al., 2018) or its interface. However, this conceptual division may not always coincide with the boundaries of the actual physical hardware (a more convenient division) requiring some pragmatism in classifying system components.[2] A conceptual diagram of the Aerosense digital twin system and its hardware is shown in Figure 3. It comprises

sensor node, base station, and cloud infrastructure sub-systems, which are classified and described below.

2.3.1 Physical system: wind turbine and sensors

As a `DigitalTwinInstance` the Aerosense system provides a digital twin for a wide range of generic turbines - from small test platforms to massive, multi-Megawatt scale devices. The latter impose demanding design requirements, especially in terms of wireless transmission ranges. Aerosense prototypes were tested on the Aventa AV-7 wind turbine[3], a small 6 kW device, located in Taggenberg (CH), with a rotor diameter of 12.8 m: this is the physical instance that is "twinned" here. However, the design specifications enable use with much larger devices.

The Aerosense sensor node pictured in Figure 2 contains a suite of sensors to provide the measurements necessary for the outputs defined by the "improving aerodynamics models" use case. The sensors included in the suite are: absolute pressure senors, differential pressure sensors, acoustic sensors, a 9 Degrees of Freedom (DOF) Inertial Measurement Unit (IMU), and microphones. These sensors are controlled by an in-house data processing and transfer unit equipped with a Bluetooth Low Emission (BLE) wireless interface for data transmission (Polonelli et al., 2023a). Up to five sensor nodes can potentially be installed to allow for measurements at different locations on the

---

2 For example, each of the subsystems mentioned contains elements that conceptually would be considered a connection system, but a coarser classification based on the purpose of each subsystem results in a clearer discussion.
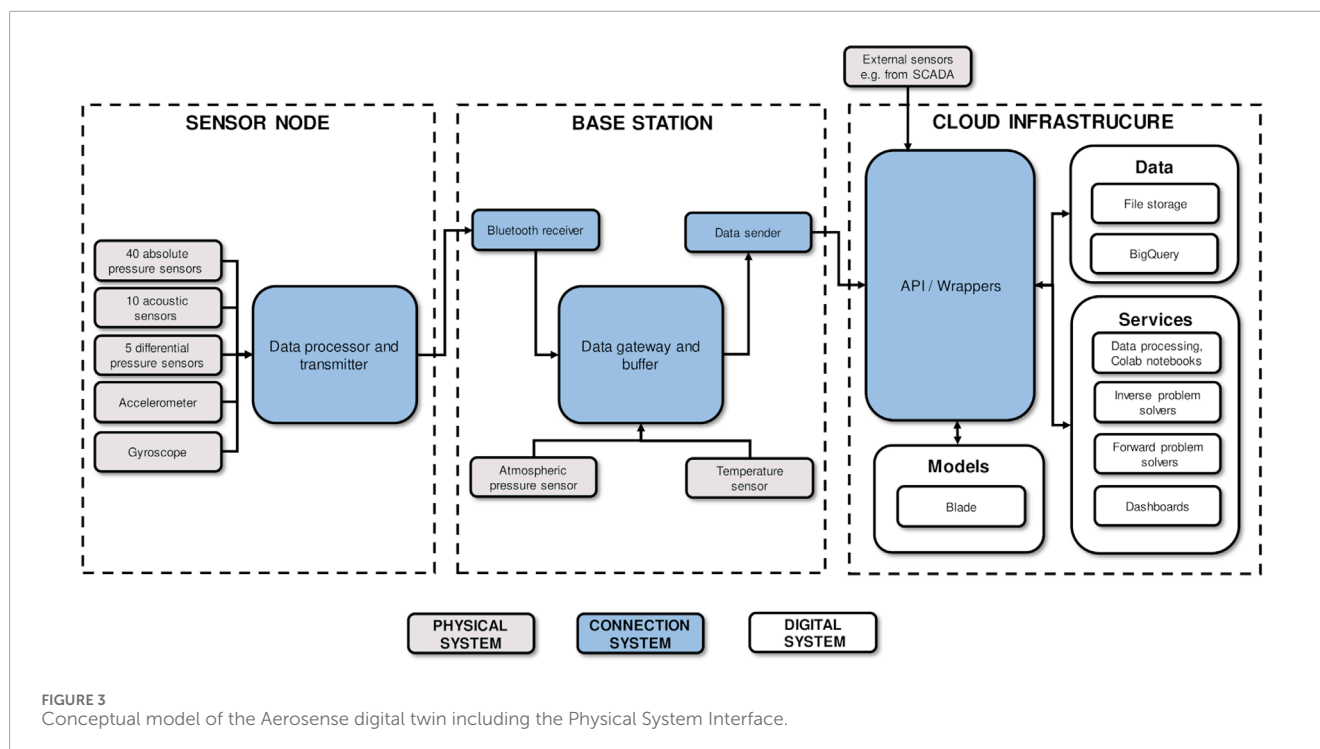
---

3 More information about the Aventa AV-7 wind turbine can be found at https://doi.org/10.5281/zenodo.8192149

FIGURE 3
Conceptual model of the Aerosense digital twin including the Physical System Interface.

rotor (Polonelli et al., 2023b). Section 3.2 discusses the design of the sensor nodes. External data sources such as Supervisory Control and Data Acquisition (SCADA) system data from turbines, weather forecasts, etc., could also be considered part of the physical system interface.

2.3.2 Connection system

The connection system[4] forms the infrastructure for data retrieval from the sensor(s) through to the cloud infrastructure. The data flow from physical to digital system in a `DigitalShadow` is typically unidirectional, as can be seen from Figure 3. The main physical element of this connection system is the base station (see Figure 2), which acts as a gateway and a buffer for sensor data on its way from the node(s) to the data ingress of the cloud infrastructure. It orchestrates sessions of sampling and data download from the node(s), and allows sessions to be controlled remotely (see Section 3.3). The Application Programming Interface (API) serves as a "connecting tissue" between different applications within the digital system. Additionally it serves as an entry point for the sensor data arriving from the gateway (running on the base station) as well as from external sensors including SCADA and other data describing physical system quasi-static properties (e.g., geometry). For the Aerosense system, a project-specific API

was developed. The design of the API and software integration is discussed in detail in Section 3.4.

2.3.3 Digital system (data, services and models)

As can be seen in Figure 3, the digital system can be conceptually divided into three sub-systems: Data (for data storage and retrieval), Models (which are the virtual entities representing the physical system) and Services (which run Models for analysis, provide data transformation and support applications like dashboards or other monitoring tools). This classification maps ontologically to the "five-dimension" digital twin model proposed by Tao et al. (2019) and used in the development of a prognostics and health management wind turbine digital twin (Tao et al., 2018). For the specific implementation of this `SimulationPredictionDigitalTwin`, the Services include forward solvers to provide the simulation capabilities, inverse solvers to infer non-measured quantities, data processing algorithms and Colab notebooks to perform data transformation and analysis, and dashboards for immediate data visualisation, exploration and inspection. The Models include Computer Aided Design (CAD) geometries of the blades, as well as sectional models for the aforementioned forward solvers. The Data sub-system provides data storage through two modalities: file storage (for long term data persistence) and BigQuery tables (for when the data needs to be queried by a user or a service). The design of the digital system is discussed in detail in Section 3.4.

3 Hardware and digital system design

A concrete realisation of the chosen digital twin type requires an implementation of hardware and digital system solutions,

---

4 According to Tao et al. (2019), connections within the digital system - such as messaging queues and data pipelines within the cloud infrastructure - would be part of the "Connection System". However, since these are such fundamental components of the data system and services architecture, their design is discussed as a part of the cloud infrastructure and digital system for the purposes of this work.

which provide the end users with a desired set of functionalities (e.g., supervisory through to intelligent learning functions, fidelity levels, twin synchronisation times, etc.) within the bounds of available resources (e.g., production and operational costs, etc.). Furthermore, when creating complex multi-scale systems such as wind turbine digital twins, it is common for the development process to occur in different teams. In this case, a team may work on a system component that has certain functionalities, which in turn satisfy resource requirements for the other team. For example, an electronics team working on a measurement system development chooses sensors that capture certain data with certain accuracy and precision. This data is later used by a cloud services development team as an input to their solvers. The changes adopted by one team will affect the performance of other, yet the final digital twin should still satisfy the user-defined constraints. This can be seen as an applied category theory collaborative design (or co-design) problem as defined by Censi (2016) and Fong and Spivak (2018). The co-design problem in general, and specifically for the entire Aerosense digital twin system, is presented in Section 3.1, followed by a system-level overview of sensor(s) design in Section 3.2. The base station design is touched upon in Section 3.3, and the cloud infrastructure and digital system implementation is discussed in detail in Section 3.4.

3.1 Co-design problem

Before formally defining a co-design problem, it is necessary to formalise a single design problem with implementation (DPI):

$$DPI = (\mathcal{F}, \mathcal{R}, \mathcal{I}, \mathbf{prov}, \mathbf{req})$$

where:

- $\mathcal{F}, \mathcal{R}, \mathcal{I}$ are posets, called Functionality, Resources, and Implementation spaces respectively;
- $\mathbf{prov}: \mathcal{I} \rightarrow \mathcal{F}$ is a mapping from an implementation to the functionality it provides;
- $\mathbf{req}: \mathcal{I} \rightarrow \mathcal{R}$ is a mapping from an implementation to the resources it requires;

A co-design problem, then, is defined as a multigraph of design problems. This allows to treat an overall design of the system in a compositional manner (i.e., divide the system into its components) and to introduce different levels of abstraction.

In the case of the Aerosense digital twin, the constraints on the Functionality and Requirement spaces are presented in Table 2. These are specific quantitative (when possible) and qualitative top-level constraints resulting from the use-story studies, described previously in Section 2.1. The overall system co-design problem can be visually represented using the graphical language as in Figure 4, with an abstraction on the hardware components and digital system levels. In these type of figures, the co-design graph is presented, allowing for an immediate overview of various interdependences in the system. Each labeled node represents an Implementation of a component or an assembly, while the edges can be of either Functionality or Recourse type. To which degree an assembly should be split to sub-assemblies and sub-sub-assemblies is arbitrary, enabling various levels of abstraction. For example, it is possible to

consider the senor node(s) as a whole or, as an assembly of sensors, power, housing, compute, and transmission assemblies. In the next three sections, the design of the three main components, senor node(s), base station, and cloud infrastructure and digital system is presented.

3.2 Sensor node design

The development of the sensor node is the most complex part of the Aerosense system design. From a system point of view, it requires a close collaboration between teams of diverse backgrounds and expertise such as development of integrated circuit boards and relative firmware (Center for Project Based Learning at ETH Zurich), experimental and computational fluid dynamics (Institute for Energy Technology at OST), structural health monitoring and machine learning (Structural Mechanics and Monitoring at ETH Zurich), data engineering (Octue), and additive manufacturing (Institute of Materials Engineering and Plastics Processing at OST). Hence, here we describe the design constraints and implementation characteristics on a system level. A detailed description of the sensor node design from an electronics point of view is available in the relevant preceding publications (Polonelli et al., 2023a).

The functionality and resources graph defined by each team during sensor node design is visualised in Figure 5. In addition to the overall design constraints already defined in Table 2 it illustrates the interdependence between various components within the sensor node design problem. For instance, a change to the desired measurement data characteristics inevitably updates the constraints on compute, power, and transmission systems. This, in turn, may influence the housing design, for example, by requiring it to provide more useable volume for a bigger battery. In terms of actual implementations, which provide the desired functionalities within the bounds of available required resources, the sensor node components have the characteristics described hereinafter.

Measurement data is the key functionality of the sensor node component, as it also constitutes a required resource for the digital system and its services. The types of sensors utilised, measurement characteristics (precision, accuracy, sampling frequency), and measurement session periods are all ultimately driven by necessity to capture the physical system state in sufficient resolution to describe the underlying phenomena. This process is at the core of the digital twin concept in that of the physical system being twinned to its digital representation. The fidelity and the resolution of this digital representation, in the end, should provide the functionalities and the added value desired by the digital twin users. The reader may refer directly to Section 3.4.3, which discusses digital system services, for more information on the intended use of the measurement data.

In terms of the sensor suite, the hardware implementation is the following. An array of 40 MEMS absolute pressure sensors (ST LPS27HHW) are distributed along the chord of the blade, sampling at 100 Hz. Following thorough calibration, an absolute accuracy of 11 Pa is achieved. Given the expected dynamic pressure of 1,000 Pa on a 5 MW wind turbine, (Deparday et al., 2022), it suggests that a precision of 1% can be reached in pressure measurements. Five differential pressure sensors measure differences of pressure around the leading-edge. The sensors have
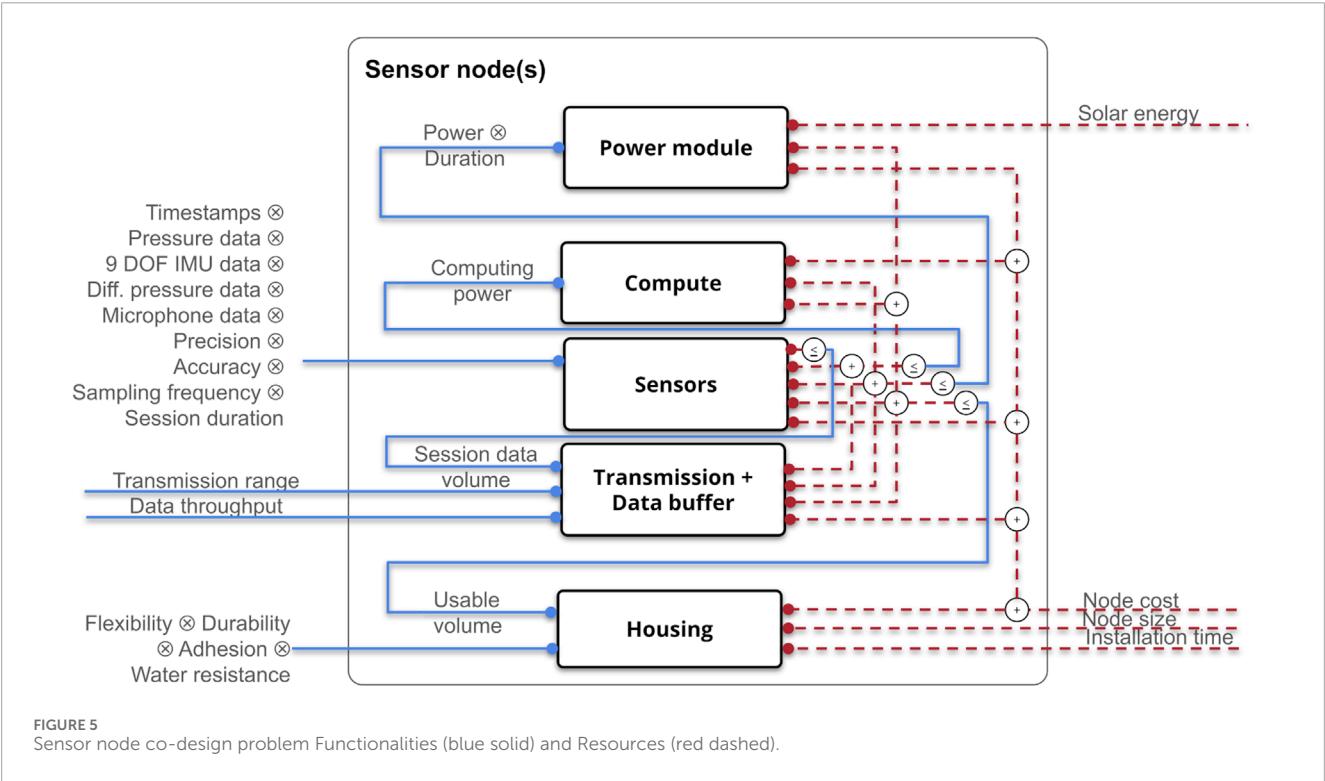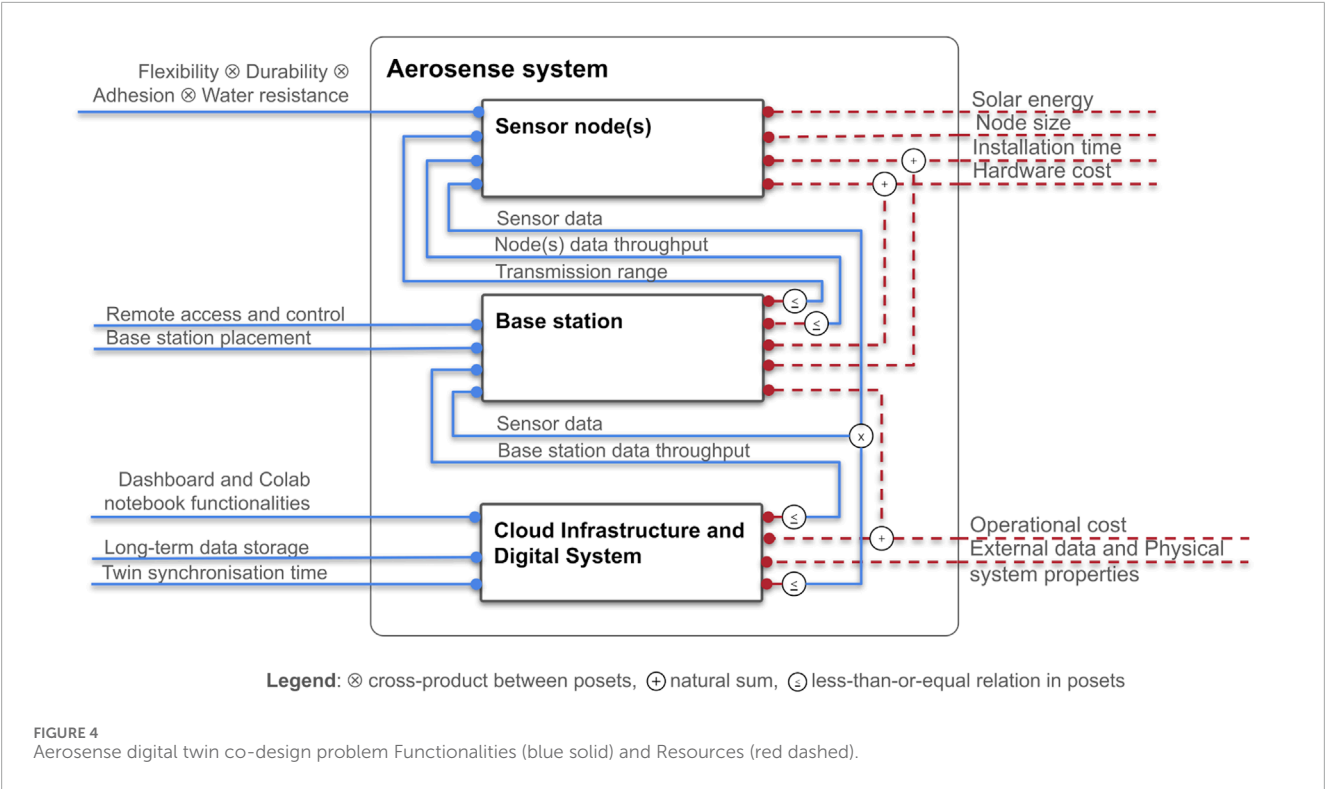
TABLE 2 Functionality and requirements constraints for the Aerosense digital twin system.

Digital twin component	Functionality/Resources	Constraint
Sensor Node	**Functionality**	
	Pressure data ⊗ IMU data ⊗ Differential pressure data ⊗Microphone data ⊗ Precision ⊗Accuracy ⊗ Sampling frequency ⊗Session duration	≥resources required for the Services
	Housing flexibility	≥minimum curvature radius of 10 mm
	Housing durability	≥withstand weather conditions defined in IEC 61400–1
	Housing adhesion	≥withstand peeling and not cause blade degradation
	Housing water resistance	≥IP55
	Resources	
	Node cost	≤5′000€ to 10′000€
	Node size	housing thickness ≤4 mm
	Installation time	≤4 h
Base Station	**Functionality**	
	Connectivity	= mobile network
	Placement	= tower base
	Resources	
	Installation time	≤1 h
	Base station cost	≤500€ to 1′000€
	Operational cost	≤100€ per month
Cloud Infrastructure and Digital System	**Functionalities**	
	Dashboard and Colab notebook functionalities	as per use case required output
	Long-term data storage	≥wind turbine life-time
	Twin synchronisation time	≤1 h
	Resources	
	Physical system geometry	~1 mm accuracy
	Operational cost	≤500€ per month

an accuracy of 0.25% Full Scale, +1 Last Significant Bit, at 25°C and a sampling frequency of 1.2 kHz, sufficient to resolve fast dynamics of the turbulent inflow. Ten acoustic sensors (Vesper VM2020) sampling at 16 kHz are installed at the trailing edge. An Inertial Measurement Unit (IMU) is included, comprising an accelerometer, a gyroscope and a magnetometer (Bosch BMX160). The IMU data is sampled at 100 Hz.

On-board compute, sensor controls, and data transmission are provided by a CC2652P microcontroller by Texas Instruments. It embeds a 48 MHz ARM Cortex-M4 processor, and a Bluetooth Low Emission (BLE) wireless interface to capture data from the individual sensors and communicate with the base station. This solution provides a low power consumption for sensor readout, and a long-range transmission with a range up to 400 m at a

FIGURE 4
Aerosense digital twin co-design problem Functionalities (blue solid) and Resources (red dashed).

FIGURE 5
Sensor node co-design problem Functionalities (blue solid) and Resources (red dashed).

rate up to 2 Mbps (Fischer et al., 2021). This allows for a flexible base station placement even on a large-scale wind turbines. The implementation of the BLE is a result of the power consumption requirement. However, this implementation results in a significant data throughput limitation. This design problem does not have a

data streaming solution with the current technology. Instead, a batch processing approach was adopted, in which sampling periods (i.e., measurement sessions) are intermittent with data transfer periods. The manner in which this conditioned the development of the Aerosense system is further described in Section 3.3 and Section 3.4.

The housing for the sensors, integrated circuits, and power module is implemented with a custom-made PolyJet 3D-printed sleeve, which is flexible enough to bend around airfoil section where the system may be potentially installed. To provide necessary adhesion characteristics, the sleeve is fixed onto the blade with the same type of adhesion tape that is used for leading edge protection of wind turbine blades. This solution also ensures an easy installation by a technician even on mounted blades, and the possibility of the system removal without damaging the blade.

3.3 Base station design

The design of the base station is less complex compared to the other parts of the Aerosense system. The constraints imposed (see Figure 4) are also less restrictive, with multiple possible solutions in terms of hardware and software implementations. Hence, in the case of the Aerosense digital twin, there is no necessity to formalise the design of individual components of the base station as a co-design problem, maintaining a higher abstraction level. The base station hardware comprises a BLE transceiver, a local computing unit running the gateway software (on a Linux distribution), and a mobile network modem which provides a connection to cloud resources. The open-source gateway software[5] was implemented as a Command Line Interface (CLI) in python, with a multi-threaded implementation (to stream packets from the node whilst simultaneously caching, batching and uploading their data contents). Software was deployed using balena.io, which allowed automated update across multiple prototypes as well as facilitating remote connection via Secure Shell Protocol (SSH). The gateway uploads data files (containing batches of sensor values) to the Data Ingress area (see Figure 7), and interfaces with the API to retrieve and update installation configurations (containing information related to equipped sensors, geolocation of the site, sensor geometry and so on) (Clark and Lugg, 2022).

3.4 Cloud infrastructure and digital system design

The cloud infrastructure provides the necessary resources for the digital system implementation (see Figure 6). From the top level point of view, the required functionalities of the cloud infrastructure include data storage, management and querying for the data system. At the same time, cloud infrastructure provides necessary compute and orchestration capabilities for digital system services. Lastly, for models, there is a requirement of management solutions. In terms of resources required by the cloud infrastructure itself, the main limitation is imposed by the operational costs, as modern cloud solutions are capable of managing Big Data type datasets and providing high performance computing (HPC).

For the Aerosense digital twin, Figure 7 shows a more detailed view of the cloud infrastructure supporting the implementation of the digital system, highlighting the data storage, retrieval and data

processing services that comprise the digital twin. This architecture was determined from a bottom-up analysis of the data requirements discussed in Section 3.4.1, and is described in more details in the following sub-sections.

3.4.1 Data

The design of an efficient Data sub-system is fundamental for a sustainable and scalable digital twin solution. The design decisions include data ingress organisation (such as communication protocols, endpoints, and APIs), data management and storage solutions (such as database types and databases management systems (DBMS) selection), data modelling and querying implementations (such as data conceptual models and database schemas), as discussed below.

3.4.1.1 Data ingress

The data ingress area of Figure 7 represents the final step in connection of the physical to the digital system. Data ingress has two aspects:

- Gateway API. A very limited set of endpoints[6] is exposed, allowing the gateway CLI to register new installations and update node configuration data. Because the set of endpoints is so limited and tightly scoped, serverless Cloud Functions are used to avoid the creation and maintenance of server-related infrastructure.
- Gateway Batch Ingress. A write-only cloud storage bucket is configured to accept authenticated uploads of files containing raw sensor data. A serverless Cloud Function is triggered on upload, its sole purpose being to read batched data from the files and stream values into long term storage (Tier 1 in Figure 7). In addition to the above advantages, using serverless functions in this case facilitates massive scalability: with data rates being extremely substantial when multiple nodes are downloading, but intermittent for much of the time, maintaining statically-resourcing servers presents either a choke-point on data ingress or a high cost for over-provisioned capacity most of the time.

The Cloud Functions for data ingress are developed in the same repository as the rest of the data gateway code, and deployed in the same Continuous Deployment process. This ensures that edge gateway code running on the base stations is always compatible with its counterpart cloud-side.

3.4.1.2 Data management and storage

The architecture of the tiered data lakehouse shown in Figure 7 was not developed in a top-down approach, but the opposite: its design emerged from a bottom-up consideration of 1) *what* data sources would have to be stored/retrieved, 2) *why* end users (researchers) would access them and 3) *how* they would do that. To start this process, a decision tree was built, not considering Aerosense in particular but governing an entirely general problem of what kinds of data storage are suitable for what kinds of data. This is shown in Figure 8. Next, each different

---

6 The endpoints are designed with a RESTful pattern of GET, POST, PUT requests.

Cloud Infrastructure and Digital System

FIGURE 6
Cloud infrastructure and digital system co-design problem Functionalities (blue solid) and Resources (red dashed).

FIGURE 7
Cloud infrastructure for the digital twin and its underlying data lakehouse.
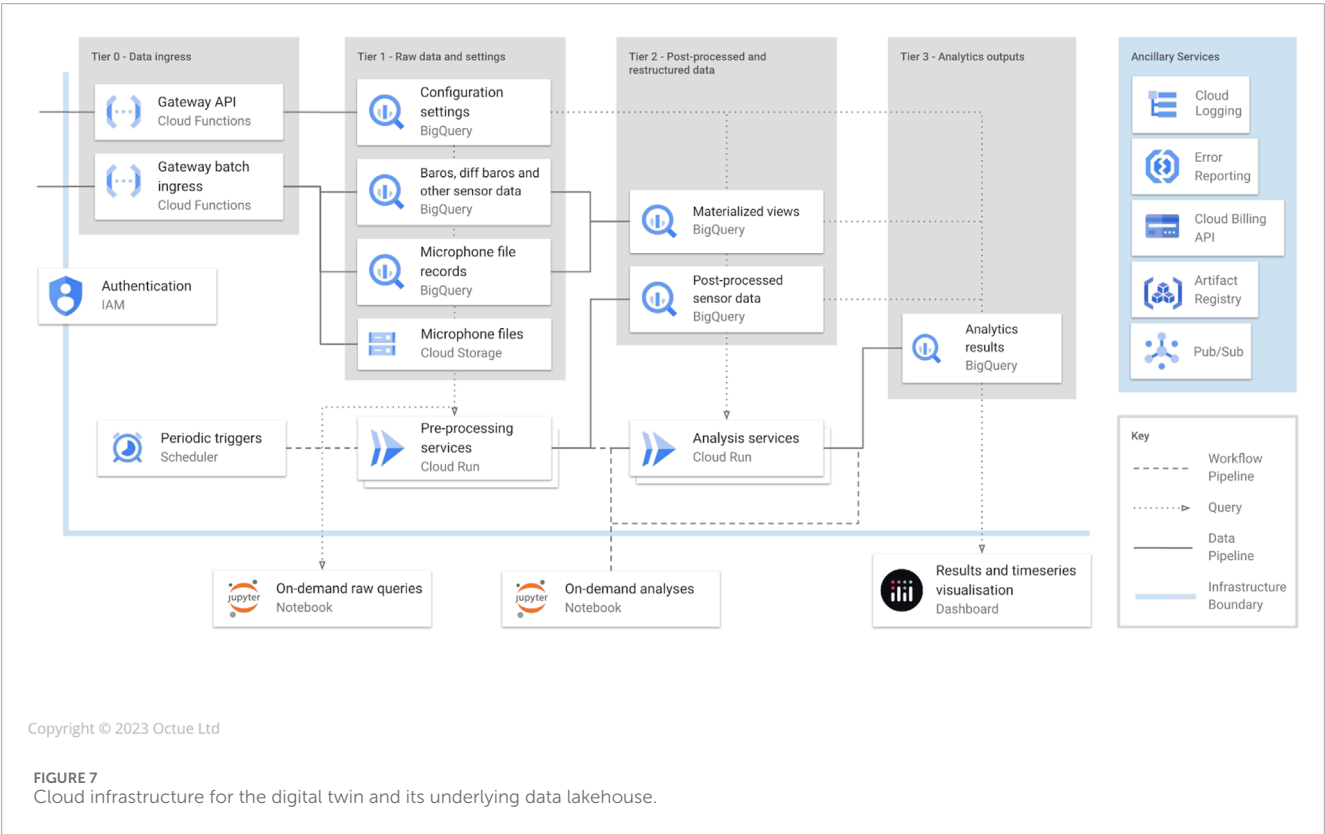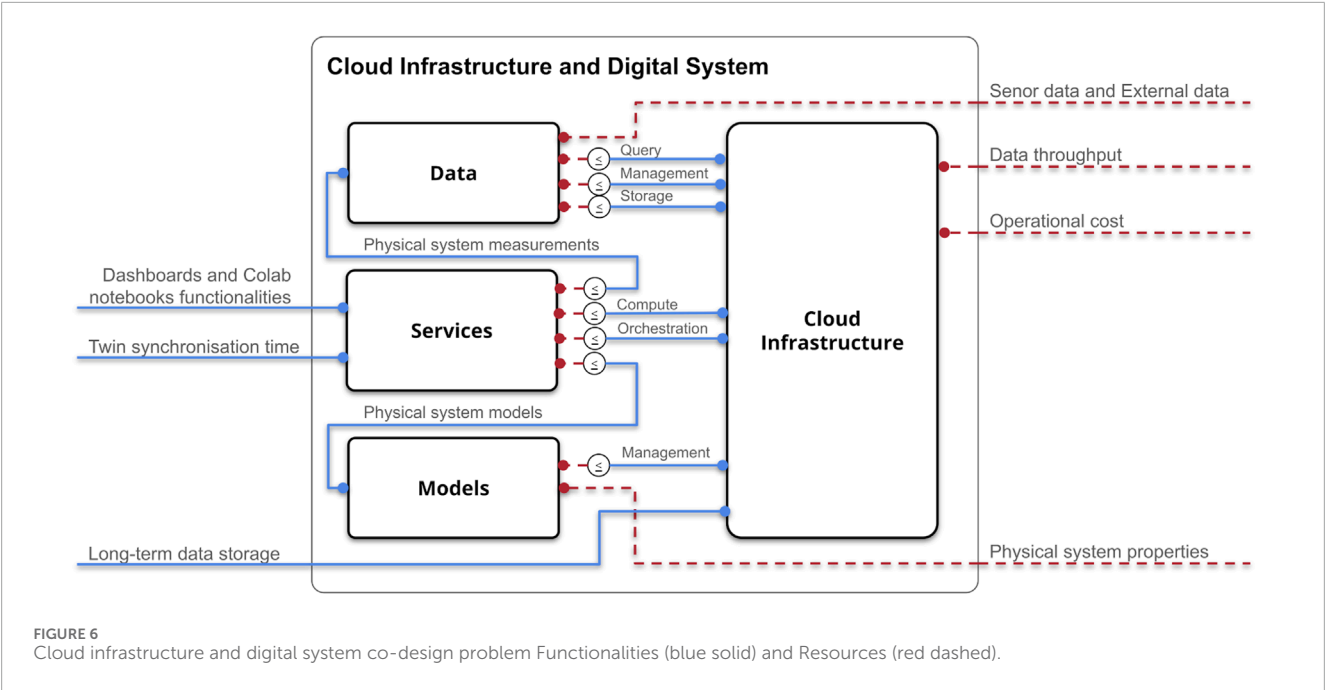
kind of data that the Aerosense project would produce was listed and the volume of that data kind was estimated.[7] Drawing

on the user profiles and journeys discussed in Section 2.1, a process was followed for each data kind to choose the ultimate storage decision. One example for the pressure sensor data (the kind requiring the most sophisticated approach) is shown in the Figure 9.

The recommended solution was to use a data warehouse for the following data sources:
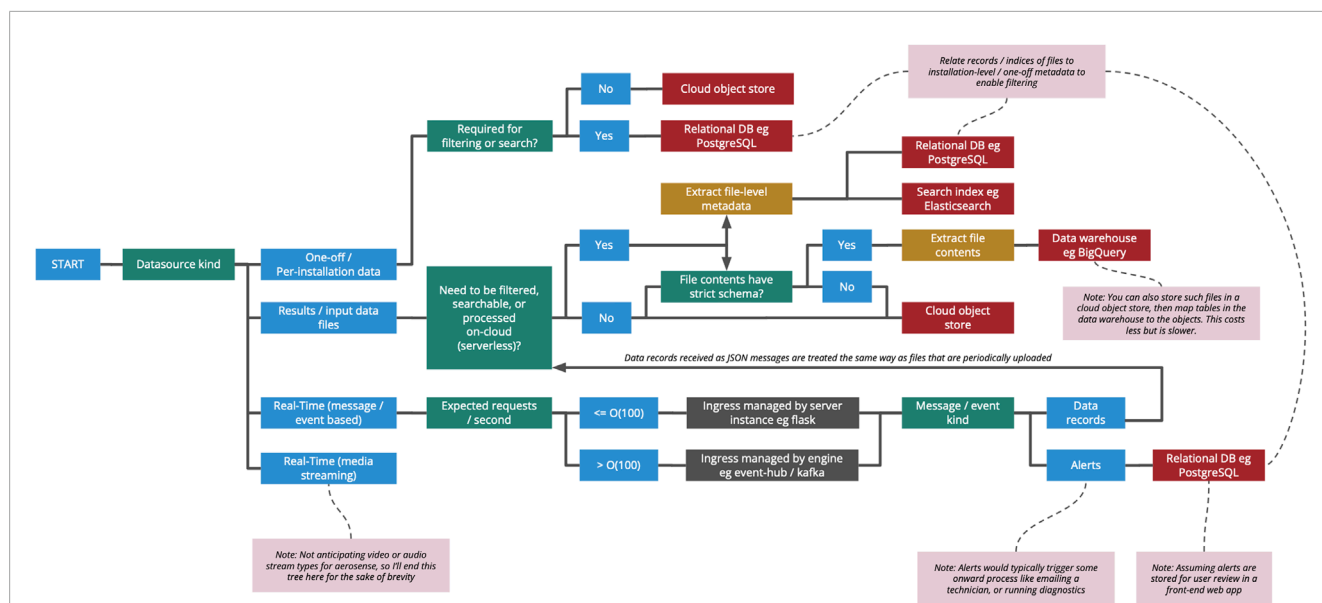
_____

7 Whilst exact specifications of sensors and related equipment were not known *a priori*, the stated initial objectives of the project allowed good estimation of the type and volume of data early in the project.

Decision tree for determining cloud data storage options based on data type and volume.
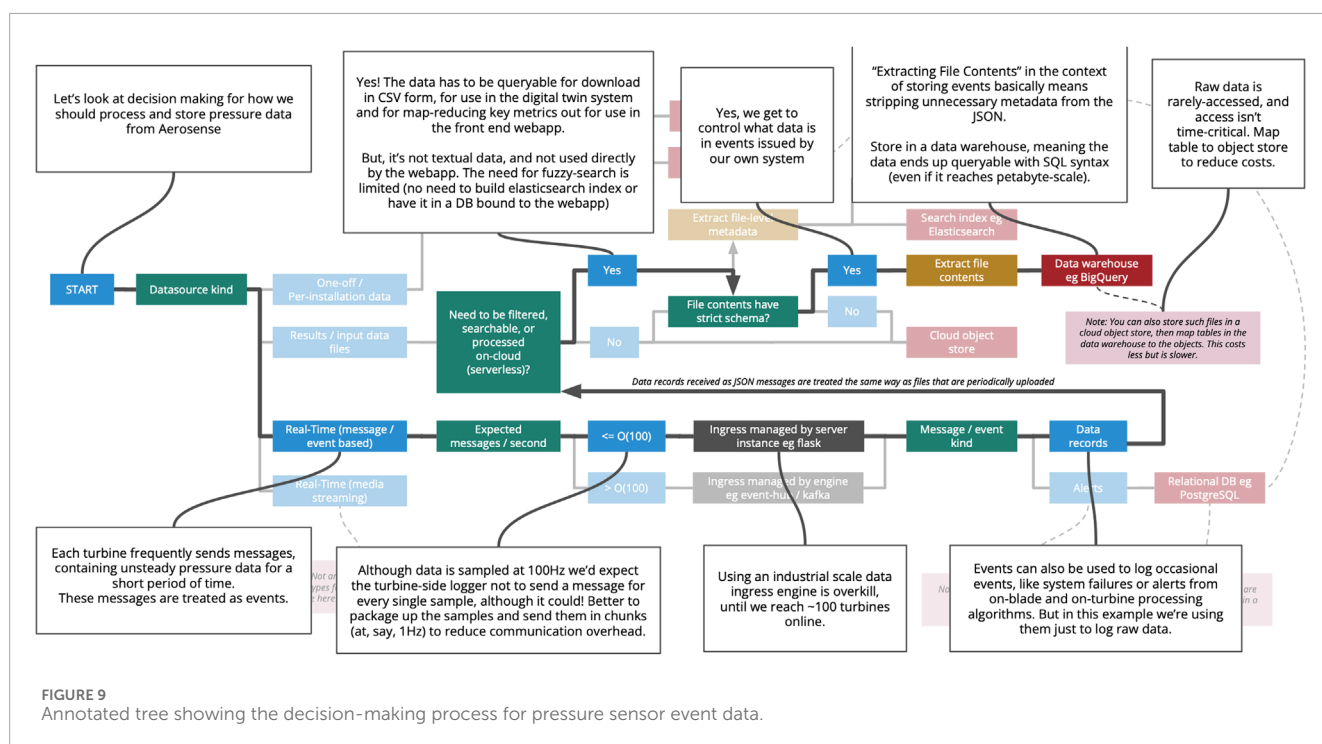
Annotated tree showing the decision-making process for pressure sensor event data.

- Unsteady pressure and accelerometer measurements, with timeseries of individual data points batched into a stream of events.
- Intermittent events stored on the same time basis, allowing efficient and easy extraction of data records corresponding to, for example, system alerts or commands issued.
- Fetched data from third party systems (e.g., wind speed, weather metrics etc.), fetched and cleaned by one or more digital twins, resolved onto the time basis of the warehouse.

- 'Materialised views' of same (in which raw data in a root master table, or derived/cleaned representations of the same, is recorded in a table having a more efficient access pattern (working like a cache for fast fetching and reduced query cost).
- Records of all file-like object entries (see below), time-synced and labeled where appropriate, enabling user to query for a manifest of the file objects relevant to a given period of time or experiment.

Columns in the warehouse tables were defined to include associated system metadata and timestamps, allowing filtering of results by experimental session, by time, or by other tag values. The recommended solution was to store the following as file-like objects ('blobs') in a store:

- Microphone recordings (being high bandwidth and efficiently compressible, treating their sensor data as a series of audio files was most appropriate).
- Trained Models used for classification, feature detection, etc. These data blobs (typically binary) never need to be queried internally so simple blob storage is ideal.
- Geometry Files such as aerofoil shapes.
- Legacy input/output files for simulations. For example, software for aerodynamic simulation such as XFOIL and OpenFAST (see Section 3.4.3) require particular format text/ASCII files for definition of simulation variables, geometry, etc., which are best retained in their original forms for the service layer.

The outstanding data kind is system metadata, which included:

- Geometrical details of the installed sensor locations.
- Installation records of the particular combinations of hardware installed on each turbine
- Session records of the sequences of commands issued in sequence
- Low-level configuration metadata (e.g., buffering and cache settings, communication port configuration, etc.

This metadata would ideally be stored in a relational database, to facilitate development of improved workflows and more interactive update of the data (e.g., via a web application). However, since metadata is tightly coupled to the data itself and the volume is extremely small, it was stored in tables in the warehouse 1) to support a simpler querying and permissions system, 2) to reduce cost by avoiding the need for maintaining a highly available relational database instance to store only a few kilobytes of data and 3) to avoid distribution of tightly coupled data into two different stores.

Positioning of the sensors (in a frame of reference local to the installed strip) is included in configuration data for each installation, which is registered (along with other metadata like session details) using the Gateway API (see Section 3.4.1). This type of metadata is serialised as JavaScript Object Notation (JSON), validated against a schema[8] and stored for future uses such as plotting pressure distributions. Additional measurements (the radius of the node from the rotation axis, the geolocation of the turbine and the shape of the blade section at that radius) are included to enable later conversion of coordinates to blade-local, turbine-local and world coordinate systems as required.

3.4.1.3 Data querying

A python based client was developed, along with a process to supply the user with credentials (a 'service account') sufficient for querying the warehouse and object store. The python client encapsulates the more challenging SQL statements required for

querying tables in the warehouse - this step is important, because different query implementations can have significant cost implications.[9] The python client can be installed on researchers' personal laptops, within a Colab notebook, and in applications like dashboards to facilitate a universal access to data. To enable working with raw file-like objects, the python client leveraged the file manifesting capability of the Octue SDK (Octue, 2022), enabling the warehouse to be queried for a list of relevant files (e.g., for a time period or experiment session) which can then be opened or downloaded directly (with the mechanics of managing cloud file storage abstracted away from the user).

3.4.2 Models

Physical system models is one of the required resources for the digital system services (see Figure 6). To satisfy the required output 2e, in which measured pressure distributions are compared to the ones obtained from the simulations, the 2D aerofoil sectional models are needed by the forward solvers (see Section 3.4.3). These models can be created by utilising data from several sources external to the Aerosense digital twin such as wind turbine technical documentation, drawings or CAD models. In practice, there are several obstacles for model creation:

- Data is not available for discontinued and legacy equipment or due to legal limitations.
- Variations from the original design specifications during manufacturing process or due to in-operation degradation.
- Difficulties in precise collocation of the sensors during the node installation on the blade.

To overcome these obstacles in the Aerosense digital twin, a photogrammetry process has been developed to evaluate the position of the sensors and the shape of the blade. The process involves taking videos and photos of the sensor nodes and wind turbine blade from different positions, and reconstructing the 3D shape of the wind turbine blade and obtaining the position of the sensors using triangulation. Detailed drawings or patterns on the housings and some additional speckled tapes make this photogrammetry process more accurate. The requirement on the accuracy was evaluated through the uncertainty quantification procedure described by Marykovskiy et al. (2023c), specifically for the inflow inference problem.

The reconstructed blade shape is further processed by extracting a point-cloud relative to the section of interest, approximating the aerofoil geometry with Bèzier curves, for smoothing and re-sampling. The resulting ordered lists of aerofoil section coordinates can be used directly as inputs to panel-code type forward solvers. As for finite volume method solvers, a Construct 2D Meshing utility was integrated into the model creation pipeline. This software creates structured, high-quality 2D aerofoil meshes. The modified version of the software developed by Fraunhofer IWES[10] was

---

FIGURE 10
Overview of the services design problem, in terms of data resources.

wrapped with Octue SDK and implemented as a child-process for the OpenFOAM service.

Additionally to blade sectional models, full wind turbine models for the solvers based on Blade Element Momentum (BEM) method were created during the Aerosense project. The integration of these models and solvers in the Aerosense digital twin was not required by the use case presented in this work. However, these developments were considered for further digital twin developments and improvements to satisfy the use cases discovered during the user-story interviews.

3.4.3 Services

As described in Section 2.3.3, Services, run Models for analysis, provide intermediate data transformation or processing (e.g., from Tier 1 to Tier 2 in Figure 7), or provide applications like dashboards or other monitoring tools. In essence, the services convert ingressed and stored raw data (Section 3.4.1) into the requisite outputs of the use case (Table 1).

Figure 10 presents this process as a co-design problem where for each service its inputs are viewed as resources and its outputs as functionalities. For the sake of clarity, the diagram omits other design parameters such as computational cost (resource), parallelisation possibilities (functionality), and fidelity (functionality). Nevertheless, these aspects should be considered when approaching digital twin design, especially for a digital twin with functionality level of `SimulationPrediction` and above (see Section 2.2). These considerations in the context of wind turbine twinning often necessitate introduction of surrogate and reduced order models, multi-fidelity and hybrid modeling techniques (Li et al., 2022; Quick et al., 2019; Quick et al., 2022; Renganathan et al., 2020).

3.4.3.1 Service wrappers

To implement a model or a data processor in a digital twin, the code or application must be "wrapped", enabling it to be deployed to cloud infrastructure and invoked as part of a data pipeline. Commonly, legacy software applications or libraries must be either re-implemented entirely or adapted to meet these requirements. The Octue SDK (Octue, 2022), which embodies Octue's 'twined' framework, was developed for this purpose (with significant work on the framework inspired by the needs of the Aerosense project).

The premise of the framework is as follows:

- Any new or legacy scientific analysis app/code is possible to wrap for use in an automated data processing pipeline.
- A system of communication called a question is the basis of the wrapper. Services can be asked a question and should answer with a series of updates culminating in a result.
- Any service can ask one or more questions of any other service.
- A service is bounded by its 'Data API':
  - inputs (data that will change on a per-analysis basis),
  - outputs (data returned as a result),
  - configuration (input settings, constants or static data that change on a per-service basis),
  - logs (semi-structured textual data reporting progress, warnings and errors)
  - monitors (structured numeric data for reporting progress, such as residual values in a CFD calculation or an estimated time remaining)
- The inputs, outputs and configuration may be supplied in the form of 'files' (a manifest of file-like objects in cloud storage with associated metadata) or 'values' (raw JSON data).

• A service has a set of JSONSchema[11] defining the expected inputs, configuration, outputs and monitors.
• All services have an identical 'Calling API' (the mechanics of asking a question as an http request and receiving an event stream of responses). This is encapsulated meaning that researchers developing services need no expertise of web APIs, servers, message queues or deployment infrastructure.

This system was implemented for each of the models discussed in Section 3.4.2, with an automated deployment process set up so that each subsequent release of code resulted in a new service revision (with corresponding version tag, allowing questions to be routed to specific versions).

3.4.3.2 Service implementations

The digital system Services required by the Aerosense use case for the realisation of the digital twin are the following:

• Data processing algorithms compute derived quantities such as static pressure, pressure coefficient distributions or blade azimuth angle. As shown in Figure 10, to compute pressure coefficient distribution on the aerofoils, measurements from absolute pressure sensors are corrected for sensor drift. Additionally it is necessary to account for the contribution of the atmospheric pressure and hydrostatic pressure variations (Deparday et al., 2023). While atmospheric pressure at the base of the wind turbine is provided by the sensors installed on the base station, the hydrostatic pressure component varies with height. IMU data is proceeded by in-house fusion algorithm (Trummer et al., 2023), to compute the largest deflections of the blade as well as its azimuthal position when rotating, enabling the estimation of hydrostatic pressure.

• Inverse problem solvers infer quantities including the angle of attack and farfield wind speed. Differential pressure measurements, once processed, are used by the inference service as input for the hybrid model, based on the inviscid flow theory (Marykovskiy et al., 2023c). Farfield wind speed is used to estimate dynamic pressure contribution in pressure coefficient calculations and along with angle of attack is used as an input to forward solvers.

• Forward problem solvers allow for comparisons between measurements and simulations, and can predict non-measured quantities such as the structural deformation of the blade.
  – OpenFOAM[12] and XFOIL[13] produce simulated data for the purpose of comparison with measured quantities, as required by the use case. To integrate these solvers with Octue SDK, xfoil-python[14] and pyFoam[15] python-based software wrapers were used. Additionally these solvers,

along with their automated workflow and data pipelines can be dual-purposed to also generate large synthetic data sets. These are then used to train machine learning algorithms and perform uncertainty quantification. This provides a bridgehead for further digital twin system augmentation with new algorithms and analysis tools.

  – OpenFAST[16] (with TurbSim) software was used to generate an inflow data and perform aeroelastic simulations with AVENTA AV-7 wind turbine model. A python package `openfast_toolbox`[17] was used to provide an intermediate integration layer between OpenFAST and OctueSDK.

• The Aerosense dashboard allows to explore the Aerosense data lake in a visual manner according to the selected metadata defined in our data model. It allows for filtering based on different measurement campaigns, installations, sensors, time periods and other metadata such as wind turbine rotor speed or relative statistical information (min, max, mean, standard deviation etc.) for chosen variables of interest. The data can be explored via the interactive plotly functions such as data inspection, view controls (such as zooming and panning), and selecting individual data sets.
• Colab notebooks providing additional post processing capabilities are available to team members as well as invited external researchers. The "Aerosense data playground" is a set of Colab notebooks that can be shared with chosen collaborators. The resulting code is continuously refactored into a python library "aerosense-tools" hosted on GitHub (Lugg et al., 2023). This library is used in dashboard development. This allows different partners to work together, develop code and get insights into the data.

4 Application and results

4.1 Test set-up

After some initial tests of the robustness of the housings and the power consumption of the Aerosense system (Polonelli et al., 2023a), a field test campaign was carried out with the final Aerosense prototype (Figure 11). As mentioned in Section 2.3.1, the design specifications enable the use of the system on multi-Megawatt scale wind turbines. However, for practical and cost considerations, these initial field tests were performed on a 6 kW Aventa wind turbine. The turbine delivers 10-min averaged SCADA data including the active power, the nacelle wind speed and wind direction, the blade pitch angle and the generator temperature. The sensor node was installed at a radial position of 74% of the blade length (6.7 m from the centre of rotation) on one blade. It is shown attached to the blade in Figure 11A. Several measurement campaigns were carried out between July 2022 and

---

11 https://json-schema.org/last access: 01 March 2024.

12 https://openfoam.org/, last access: 01 March 2024.

13 https://web.mit.edu/drela/Public/web/xfoil/, last access: 01 March 2024.

14 https://github.com/DARcorporation/xfoil-python, last access: 01 March 2024.

15 http://openfoamwiki.net/index.php/Contrib_PyFoam, last access: 01 March 2024.

16 https://www.nrel.gov/wind/nwtc/openfast.html, last access: 01 March 2024.

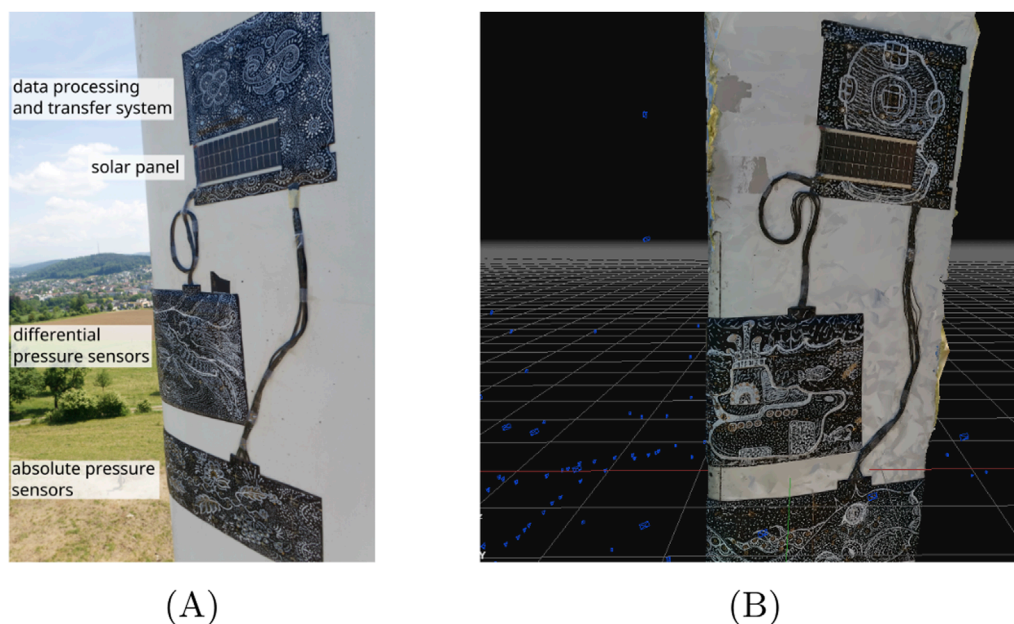17 https://github.com/OpenFAST/openfast_toolbox, last access: 01 March 2024.

FIGURE 11
View of the sensor node installed in the field. **(A)** The Aerosense node installed on the Aventa turbine, **(B)** 3D reconstruction of the installed sensor node.

April 2023; however, with several improvements being made to the hardware and firmware. This results in several weeks of useable barometer, differential pressure sensor, microphone, and IMU data during this time, which is available in Marykovskiy et al. (2023b). All the samples are timestamped with UTC time. Wind turbine, sensor location and sensor properties metadata is provided as JSON files. The wind turbine metadata follows the WindIO schema and the pressure sensor locations metadata follows the Aerosense sensor geometry schema. The pressure sensor locations were calculated as described in Section 3.4.2 with a photogrammetry process[18], reconstructing the sensor node and the blade shape, as demonstrated in Figure 11B. An accuracy in the order of 1 mm was achieved, enabling accurate positioning of the aforementioned sensors and the use of this information in post-processing or transformation of measurement data.

4.2 Digital twin demonstration

The field test allowed us to demonstrate the value of the use case introduced in Section 2.1 according to the outputs in Table 1. As described in the aforementioned section, the two main modes to use and analyse the data are the dashboard and the Colab notebooks. Here, we limit the demonstration to these two functionalities of the digital twin, while the detailed analysis of the data itself merits a separate study to thoroughly explore the insights gained from the field tests.

----

18   https://sketchfab.com/3d-models/aventa-blade-
2ebed0e05e0e4c3c95a7308af3a494d3, last access: 01 March 2024.

4.2.1 Dashboard

The dashboard displays times series for each sensor, classified in different installations of the Aerosense measurement system, as shown in Figure 12 for pressure data from the barometers. Specific sensor types and measurements periods can be selected. No data needs to be downloaded or specific code to be computed to obtain an initial data inspection of measurement data. This fulfils the output 1a of Table 1. The dashboard can also display pressure coefficient distribution plots for an immediate visualisation relative to specific time instances. This functionality enables the verification of the physical plausibility of pressure data recorded by the measurement system and a general understanding of the aerodynamic behaviour at specific time instants. The synchronisation time of the dashboard plots to the on-site measurements is under 1 hour, allowing for quasi-real time monitoring. This fulfils the output 2a and 2b of Table 1.

4.2.2 Colab notebooks

The Colab notebooks developed within this work allow users to perform detailed analyses, such as extracting data and plotting pressure distributions. This is achieved without the need to write a new code that works with the downloaded data. The versatility of the Colab notebook allows more complex analyses, for example, based on conditional averaging with specific weather conditions in a large time period or for specific azimuthal position of the blade. The Colab notebook enables, for example, the comparison of data from multiple wind turbine installations. This fulfils the output 1b of Table 1.

Phase-averaged pressure and pressure coefficient distributions at different operating conditions can be computed and analysed in the Colab notebook. This allows, for example, a more detailed analysis of the pressure distribution depending on the azimuth
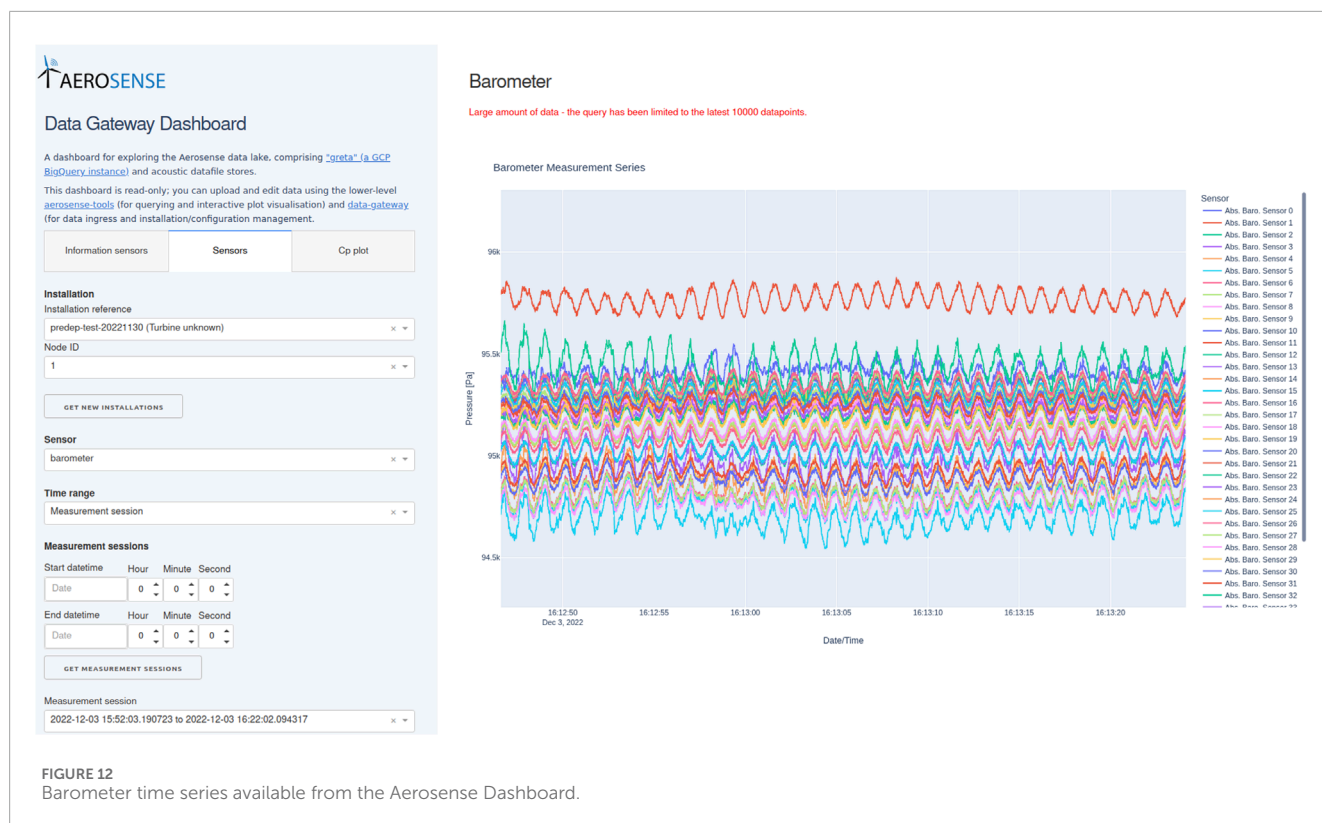
FIGURE 12
Barometer time series available from the Aerosense Dashboard.

position of the rotor blade for different operating conditions (output 2b from Table 1) It also enables direct comparisons with data from the customer, for instance measured or simulated 2D pressure coefficient distributions (output 2c). Furthermore, phase-averaged pressure distribution can also be directly compared to pressure distributions obtained from XFOIL simulations for given inflow conditions, as illustrated in Figure 13 (output 2e). This figure depicts phase-averaged pressure distributions at different rotor blade azimuth positions (indicated by the colour of the points) for a 1-min time interval when wind directions and wind speeds were considered stable. They are compared to XFOIL simulations for two inflow conditions. The inflow conditions were computed using the Aerofoil inference model (Marykovskiy et al., 2023c) available in the Python package used by the Colab notebook (output 2d). Figure 13 illustrates that the local wind speed and corresponding angle of attack differ at different positions on the blade when rotating. This may be due to non-uniformity of the wind or yaw misalignment of the wind turbine. These findings can facilitate a more comprehensive understanding of the aerodynamic behaviour of the wind turbine in operational conditions, as well as to recommendations for the improvement of aerodynamic models.

5 Discussion

In this work, we have seen that the development of digital twins in the wind energy context primarily represents a characteristic system engineering problem. In this discussion, we summarise the main challenges we experienced in this project, followed

by an evaluation of the methodologies used to overcome these challenges. Finally, we discuss how the results can be useful for wider applications. This presents domain experts with questions, which can often lie outside their primary expertise. The main challenges encountered when architecting the Aerosense digital twins were found to be:

- Establishing priority use cases for the digital twin system, and preventing a "scope creep" from introducing ever increasing requirements to the system.
- Collaboration and management of teams from diverse domains.
- Selecting and adopting appropriate supporting digital technologies for sustainable and scalable results.

The authors believe these challenges are not unique to the Aerosense project, and the methods presented in this work can also be applied in the development of digital twins with different applications and different twinned physical systems, including the multi-Megawatt scale wind turbines. This methodological approach to digital twin design, in fact, becomes even more crucial for the successful technology implementation as the complexity of twinned systems increases, and the use cases require integration of signals from increased number of sensors and analyses resulting from multi-scale and multi physics models. In the next sections, we evaluate the success of our use of the design thinking, design patterns, decision trees and applied category theory methodologies to overcome these challenges.
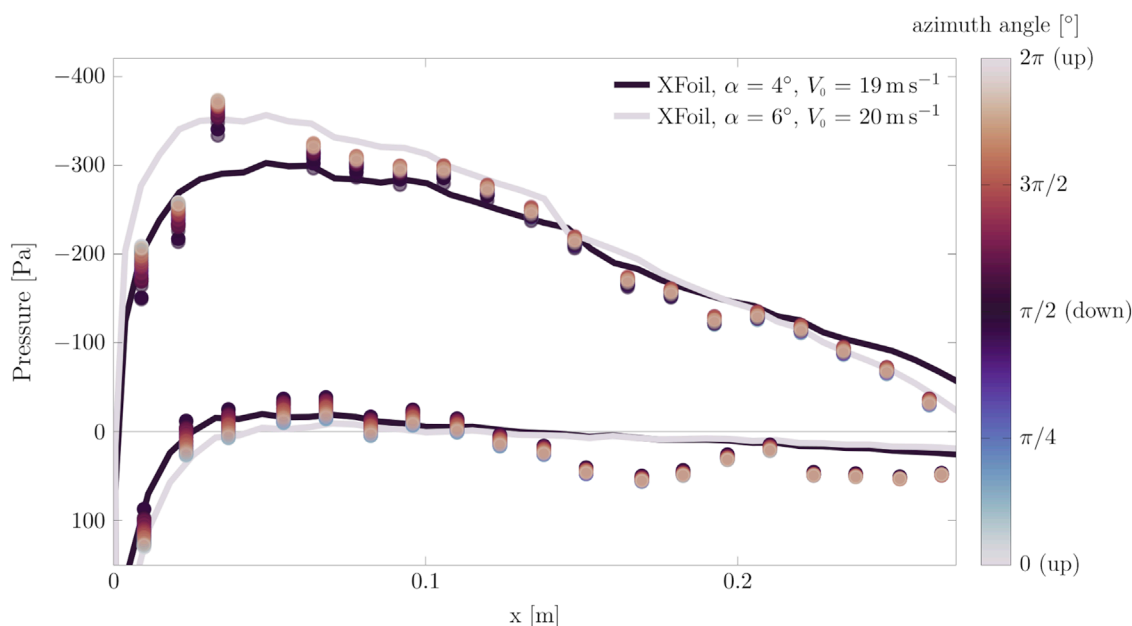
FIGURE 13
Aerosense Colab notebooks produced plot. Phase-averaged pressure distributions at different rotor blade azimuth position for a 1-min time interval when wind directions and wind speed were stable. Comparison with XFOIL simulation results.

5.1 Design thinking for digital twins

The main focus of the design thinking methodology in this work was on the "Empathise" phase, in which "user stories" were explored in order to then define the foundational use case that was used as a basis for the design of the Aerosense digital twin. It proved as essential tool for providing the required amount of focus to then apply the design patterns methodology, as discussed below. Furthermore, a range of other use cases were defined, which could be used to further develop the present solution.

5.2 Design patterns for digital twins

In the Aerosense project, when approaching the development of the digital twin from a systems architecture perspective, the authors adopted well-established twin types as well as common conceptualisation schemes such as the 5-dimension digital twin model. These conceptual models have provided an initial indication on the overall design patterns, to serve as a blueprint for further development. Currently, "the catalogue" of solution patterns is rather limited and only a few digital twin type definitions are commonly accepted and recognised across engineering domains. A more fine-grained and widely accepted digital twin classification can facilitate more streamlined digital twin development by applying proven methodologies. Additionally, this approach opens avenues for collaborative innovation. By utilising a classification system that is acknowledged across industries, future digital twin projects can increase interoperability and knowledge exchange, thus gaining access to otherwise untapped resources.

Furthermore, the use of a standardised classification schemes aids in documenting and communicating the functionality and scope of digital twins more effectively, both within and across industries. This not only enhances the visibility of research outcomes but also aids in selection and adoption of technical solutions, making them more accessible to domain experts from different sectors, and facilitating cross-domain knowledge transfer. An example of such knowledge transfer is the adoption of digital twin technology innovations in the manufacturing sector by the wind energy domain, where the principles of operational efficiency and predictive maintenance are equally valuable.

5.3 Applied category theory and co-design framework for digital twins

During the design of individual system components, the discussion centered around the component assemblies and their physical boundaries, rather than conceptual division into physical, digital and connection systems. This type of compositional thinking finds its fundamental basis in category theory. Co-design formalisation originally proposed by Zardini et al. (2021) in the context of robotics and autonomous systems also lends itself to the digital twin development. The theoretical grounding of this framework serves as means of knowledge representation for functional and resource requirements for a given digital twin, its components or component assemblies. Additionally, this formal basis allows for multiple solution searches and optimisations as a twin evolves to include new uses cases or technology implementations.

6 Conclusion

In this paper the authors approached the design, development, and deployment of digital twin from a system engineering perspective, to architect the Aerosense system for aerodynamic monitoring of wind turbine blades.

The Aerosense system use case and requirements (Section 2.1) were adopted in Section 2.2, leading to characterisation of the particular digital twin with a "Simulation/Prediction" functionality, "Digital Shadow" connectivity, and being at "Instance" physical system lifetime stage. By undertaking the same classification process for prospective digital twins, readers can understand whether the technology solutions used in this work are appropriate for reuse to accelerate their own project(s).

Casting the design of the digital twin into a co-design formalism of applied category theory (Section 3.1), proved instrumental for obtaining desired twin functionalities and affronting interdisciplinary challenges.

The technology solutions form a fully-tested, production-ready data system for turbine blade data collection, cloud ingress, data lakehouse storage and access/use. Both software and hardware solutions have been developed and published as open source (together with documentation) to provide a complete implementation example.

The Aerosense digital twin brings this all together to provide an out-of-the-box solution to monitoring wind turbines in the field, collecting blade sensor data for research purposes and augmenting that data with simulations to form a digital twin of the type classification described above.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://doi.org/10.34808/ypae-8684.

Author contributions

YM: Conceptualization, Data curation, Investigation, Methodology, Software, Visualization, Writing–original draft, Writing–review and editing. TC: Conceptualization, Investigation, Methodology, Software, Visualization, Writing–original draft, Writing–review and editing. JD: Investigation, Software, Visualization, Writing–original draft, Writing–review and editing. EC: Resources, Supervision, Writing–review and editing. SB: Conceptualization, Funding acquisition, Investigation, Methodology, Resources, Supervision, Writing–original draft, Writing–review and editing.

Conflict of interest

Author TC was employed by Octue 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.

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

Abdallah, I., Duthé, G., Barber, S., and Chatzi, E. (2022). Identifying evolving leading edge erosion by tracking clusters of lift coefficients. *J. Phys. Conf. Ser.* 2265, 032 089. doi:10.1088/1742-6596/2265/3/032089

Arista, R., Zheng, X., Lu, J., and Mas, F. (2023). An Ontology-based Engineering system to support aircraft manufacturing system design. *J. Manuf. Syst.* 68, 270–288. doi:10.1016/j.jmsy.2023.02.012

Augustyn, D., Ulriksen, M. D., and Sørensen, J. D. (2021). Reliability updating of offshore wind substructures by use of digital twin information. *Energies* 14, 5859. doi:10.3390/en14185859

Bangga, G., Weihing, P., Lutz, T., and Krämer, E. (2017). Effect of computational grid on accurate prediction of a wind turbine rotor using delayed detached-eddy simulations. *J. Mech. Sci. Technol.* 31, 2359–2364. doi:10.1007/s12206-017-0432-6

Barber, S., Deparday, J., Marykovskiy, Y., Chatzi, E., Abdallah, I., Duthé, G., et al. (2022). Development of a wireless, non-intrusive, MEMS-based pressure and acoustic measurement system for large-scale operating wind turbine blades. *Wind Energy Sci.* 7, 1383–1398. doi:10.5194/wes-7-1383-2022

Branlard, E., Giardina, D., and Brown, C. S. D. (2020). Augmented Kalman filter with a reduced mechanical model to estimate tower loads on a land-based wind turbine: a step towards digital-twin simulations. *Wind Energy Sci.* 5, 1155–1167. doi:10.5194/wes-5-1155-2020

Censi, A. (2016). A mathematical theory of Co-design. doi:10.48550/ARXIV.1512.08055

Clark, T. H., and Lugg, M. (2022). From blade to BigQuery: a case study, doi:10.5281/zenodo.10925800

Clifton, A., Barber, S., Bray, A., Enevoldsen, P., Fields, J., Sempreviva, A. M., et al. (2023). Grand challenges in the digitalisation of wind energy. *Wind Energy Sci.* 8, 947–974. doi:10.5194/wes-8-947-2023

D'Amico, R. D., Erkoyuncu, J. A., Addepalli, S., and Penver, S. (2022). Cognitive digital twin: an approach to improve the maintenance management. *CIRP J. Manuf. Sci. Technol.* 38, 613–630. doi:10.1016/j.cirpj.2022.06.004

Deparday, J., Marikovskiy, Y., Polonelli, T., Clark, T., and Barber, S. (2023). How to analyse blade aerodynamics on an operating wind turbine with low-cost pressure sensors?. doi:10.5281/zenodo.7974881

Deparday, J., Müller, H., Polonelli, T., and Barber, S. (2022). An experimental system to acquire aeroacoustic properties on wind turbine blades. *J. Phys. Conf. Ser.* 2265, 022 089. doi:10.1088/1742-6596/2265/2/022089

Duthé, G., Abdallah, I., Barber, S., and Chatzi, E. (2021). Modeling and monitoring erosion of the leading edge of wind turbine blades. *Energies* 14, 7262. doi:10.3390/en14217262

Duthé, G., Abdallah, I., Barber, S., and Chatzi, E. (2023). Graph neural networks for aerodynamic flow reconstruction from sparse sensing. *arXiv preprint arXiv:2301.03228*. doi:10.48550/arXiv.2301.03228

Fischer, R., Mueller, H., Polonelli, T., Benini, L., and Magno, M. (2021). "WindNode: a long-lasting and long-range Bluetooth wireless sensor node for pressure and acoustic monitoring on wind turbines," in 2021 4th IEEE international conference on industrial cyber-physical systems (ICPS) (IEEE), 393–399.

Fong, B., and Spivak, D. I. (2018). Seven sketches in compositionality: an invitation to applied category theory, doi:10.48550/ARXIV.1803.05316

Gamma, E., Helm, R., Johnson, R., and Vlissides, J. (1994). "Design patterns: elements of reusable object-oriented software," in *Addison-wesley professional computing series*. Reading, MA: Pearson Education.

Grieves, M. (2022). Intelligent digital twins and the development and management of complex systems. *Digit. Twin* 2, 8. doi:10.12688/digitaltwin.17574.1

Grieves, M., and Vickers, J. (2017). "Digital twin: mitigating unpredictable," in *Undesirable emergent behavior in complex systems*. Cham: Springer International Publishing, 85–113. doi:10.1007/978-3-319-38756-7_4

Hines, E. M., Baxter, C. D., Ciochetto, D., Song, M., Sparrevik, P., Meland, H. J., et al. (2023). Structural instrumentation and monitoring of the block island offshore wind farm. *Renew. Energy* 202, 1032–1045. doi:10.1016/j.renene.2022.11.115

Kritzinger, W., Karner, M., Traar, G., Henjes, J., and Sihn, W. (2018). Digital Twin in manufacturing: a categorical literature review and classification. *IFAC-PapersOnLine* 51, 1016–1022. 16th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2018. doi:10.1016/j.ifacol.2018.08.474

Le Franc, Y., Parland-von Essen, J., Bonino, L., Lehväslaiho, H., Coen, G., and Staiger, C. (2020). D2.2 FAIR semantics: first recommendations, doi:10.5281/zenodo.3707985

Li, K., Kou, J., and Zhang, W. (2022). Deep learning for multifidelity aerodynamic distribution modeling from experimental and simulation data. *AIAA J.* 60, 4413–4427. doi:10.2514/1.J061330

Lone, E. N., Sauder, T., Larsen, K., and Leira, B. J. (2022). Probabilistic fatigue model for design and life extension of mooring chains, including mean load and corrosion effects. *Ocean. Eng.* 245, 110396–396. doi:10.1016/j.oceaneng.2021.110396

Lugg, M., Marykovskiy, Y., and Clark, T. (2023). Aerosense-ai/aerosense-tools: AeroSense operational twin alpha. doi:10.5281/zenodo.8084257

Madsen, H. A., Bertagnolio, F., Fischer, A., Bak, C., and Paulsen, U. S. (2016). A novel full scale experimental characterization of wind turbine aero-acoustic noise sources.

Marykovskiy, Y., Abdallah, I., Barber, S., and Chatzi, E. (2023a). Extended taxonomy of digital twins. doi:10.5281/zenodo.8021787

Marykovskiy, Y., Clark, T., Day, J., Wiens, M., Henderson, C., Quick, J., et al. (2024). Knowledge engineering for wind energy. *Wind Energ. Sci.* 9, 883–917. doi:10.5194/wes-9-883-2024

Marykovskiy, Y., Deparday, J., Abdallah, I., and Barber, S. (2023b). AeroSense measurements: Aventa AV-7 Taggenberg winter 2022-2023 campaign. doi:10.34808/ypae-8684

Marykovskiy, Y., Deparday, J., Abdallah, I., Duthé, G., Barber, S., and Chatzi, E. (2023c). Hybrid model for inflow conditions inference on airfoils under uncertainty. *AIAA J.* 61, 4913–4925. doi:10.2514/1.J063108

Octue (2022). Octue SDK. doi:10.5281/zenodo.10961974

Pearce, B. (2020). Design thinking. doi:10.5281/zenodo.3717021

Polonelli, T., Moallemi, A., Kong, W., Müller, H., Deparday, J., Magno, M., et al. (2023b). A self-sustainable and micro-second time synchronised multi-node wireless system for aerodynamic monitoring on wind turbines. *IEEE Access*, 1–18. doi:10.1109/ACCESS.2023.3327422

Polonelli, T., Müller, H., Kong, W., Fischer, R., Benini, L., and Magno, M. (2023a). Aerosense: a self-sustainable and long-range Bluetooth wireless sensor node for aerodynamic and aeroacoustic monitoring on wind turbines. *IEEE Sensors J.* 23, 715–723. doi:10.1109/JSEN.2022.3224307

Quick, J., Hamlington, P. E., King, R., and Sprague, M. A. (2019). Multifidelity uncertainty quantification with applications in wind turbine aerodynamics, doi:10.2514/6.2019-0542

Quick, J., King, R. N., Barter, G., and Hamlington, P. E. (2022). Multifidelity multiobjective optimization for wake-steering strategies. *Wind Energy Sci.* 7, 1941–1955. doi:10.5194/wes-7-1941-2022

Renganathan, S. A., Harada, K., and Mavris, D. N. (2020). Aerodynamic data fusion toward the digital twin paradigm. *AIAA J.* 58, 3902–3918. doi:10.2514/1.J059203

Schepers, J. G., and Schreck, S. J. (2019). Aerodynamic measurements on wind turbines. *Wiley Interdiscip. Rev. Energy Environ.* 8, e320. doi:10.1002/wene.320

Singh, S., Shehab, E., Higgins, N., Fowler, K., Reynolds, D., Erkoyuncu, J. A., et al. (2021). Data management for developing digital twin ontology model. *Proc. Institution Mech. Eng. Part B J. Eng. Manuf.* 235, 2323–2337. doi:10.1177/0954405420978117

Tao, F., Zhang, M., Liu, Y., and Nee, A. (2018). Digital twin driven prognostics and health management for complex equipment. *CIRP Ann.* 67, 169–172. doi:10.1016/j.cirp.2018.04.055

Tao, F., Zhang, M., and Nee, A. (2019). "Chapter 3 - five-dimension digital twin modeling and its key technologies," in *Digital twin driven smart manufacturing*. Editors F. Tao, M. Zhang, and A. Nee (Academic Press), 63–81. doi:10.1016/B978-0-12-817630-6.00003-5

Tekinerdogan, B., and Verdouw, C. (2020). Systems architecture design pattern catalog for developing digital twins. *Sensors* 20, 5103. doi:10.3390/s20185103

Troldborg, N., Bak, C., Madsen, H. A., and Skrzypinski, W. (2013). *DANAERO MW: Final Report*. E No. 0027(EN). Roskilde, Denmark: DTU Wind Energy.

Trummer, P., Polonelli, T., Deparday, J., Abdallah, I., and Magno, M. (2023). "Blade position and motion estimation on the surface of a rotating wind turbine through a single MEMS IMU," in *2023 9th international workshop on advances in sensors and interfaces (IWASI)*, 40–45. doi:10.1109/IWASI58316.2023.10164363

W3C (2009). SKOS simple knowledge organization system reference: W3C reccomendation. Available at: https://www.w3.org/TR/skos-reference/ (Accessed February 01, 2024).

Wagg, D. J., Worden, K., Barthorpe, R. J., and Gardner, P. (2020). Digital twins: state-of-the-art and future directions for modeling and simulation in engineering dynamics applications. *ASCE-ASME J Risk Uncert Engrg Sys Part B Mech Engrg* 6, 030901. doi:10.1115/1.4046739

Welte, T. M., Sperstad, I. B., Sørum, E. H., and Kolstad, M. L. (2017). Integration of degradation processes in a strategic offshore wind farm O&M simulation model. *Energies* 10, 925. doi:10.3390/en10070925

Zardini, G., Milojevic, D., Censi, A., and Frazzoli, E. (2021). "Co-Design of embodied intelligence: a structured approach," in IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Prague, Czech Republic, 27 September 2021 - 01 October 2021, 7536–7543. doi:10.1109/IROS51168.2021.9636513

Zheng, X., Lu, J., and Kiritsis, D. (2021). The emergence of cognitive digital twin: vision, challenges and opportunities. *Int. J. Prod. Res.* 0, 7610–7632. doi:10.1080/00207543.2021.2014591

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

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact

frontiers

Frontiers in
Energy Research

frontiers | Research Topics