Reliability Engineering and System Safety 260 (2025) 111004

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Review

A review of prognostics and health management techniques in wind energy

Jokin Cuesta [a,b,*], Urko Leturiondo [a], Yolanda Vidal [b,c], Francesc Pozo [b,c]

[a] Artificial Intelligence and Data Area. Ikerlan Technology Research Center, Basque Research and Technology Alliance (BRTA), Pº J. Mª. Arizmendiarrieta, 2, Arrasate/Mondragón, 20500, Spain

[b] Control, Data, and Artificial Intelligence (CoDAlab), Department of Mathematics, Escola d'Enginyeria de Barcelona Est (EEBE), Campus Diagonal-Besos (CDB), Universitat Politecnica de Catalunya (UPC), Eduard Maristany 16, Barcelona, 08019, Spain

[c] Institute of Mathematics (IMTech), Universitat Politecnica de Catalunya (UPC), Pau Gargallo 14, Barcelona, 08019, Spain

ARTICLE INFO

Keywords:
Prognostics and health management
Remaining useful life
Prognostics
Wind turbines
Wind energy
Predictive maintenance

ABSTRACT

This review aims to provide a holistic understanding of prognostics and health management (PHM) techniques in wind energy, particularly in the estimation of remaining useful life (RUL) of wind turbine (WT) components. The study begins with an introduction that discusses the principles of PHM and its critical role in the wind energy sector. This is followed by an overview of WT systems and the importance of accurate RUL predictions for specific failure modes. Then, various data sources, methods of feature extraction, and criteria for constructing health indices are explored, along with techniques for threshold determination. Degradation modeling techniques, essential for RUL prediction, are examined through three approaches: physics-based models, data-driven methods (including statistical and artificial intelligence techniques), and hybrid models. The performance of these models is evaluated using specific metrics which have been explored. Next, predictive maintenance strategies, optimized using RUL predictions, are presented to minimize downtime and maintenance costs. The paper concludes by identifying future research directions, emphasizing the need to manage uncertainty, integrate physical knowledge, address variable environmental and operational conditions, overcome data issues, and handle system complexity.

Contents

* Corresponding author at: Artificial Intelligence and Data Area. Ikerlan Technology Research Center, Basque Research and Technology Alliance (BRTA), Pº J. Mª. Arizmendiarrieta, 2, Arrasate/Mondragón, 20500, Spain.
E-mail address: jcuesta@ikerlan.es (J. Cuesta).

https://doi.org/10.1016/j.ress.2025.111004
Received 23 July 2024; Received in revised form 12 December 2024; Accepted 3 March 2025
Available online 12 March 2025

1. Introduction

The International Energy Agency (IEA) has set an ambitious goal of tripling global renewable power capacity by 2030, in line with their *Net Zero Emissions by 2050 Scenario* [1]. To achieve this, wind energy additions must double, underscoring the critical role of wind power in the energy transition. The expansion of wind energy, particularly offshore wind, is crucial for global renewable energy targets. However, it faces challenges due to variable macroeconomic conditions. As highlighted in the *Renewables 2023: Analysis and forecasts to 2028* report by the IEA [1], investment costs for offshore wind projects have surged, rising more than 20% than just a few years ago. This has led to project delays and cancellations, impacting 15 GW of offshore wind projects in the United States and the United Kingdom. These challenges emphasize the importance of optimizing operation and maintenance (O&M) costs, which typically account for 10%–30% of the levelized cost of energy (LCOE) for wind industry projects, according to the International Renewable Energy Agency [2].

The transition to a predictive maintenance strategy, facilitated by prognostics and health management (PHM) techniques, offers significant benefits in enhancing the efficiency of wind turbine (WT) maintenance. PHM is a computer-based analysis method that focuses on optimizing machinery and equipment management by reducing O&M costs and extending the useful life of a structure, system, or component [3]. This approach addresses three crucial tasks: fault detection, fault diagnostics, and prognostics. To establish a solid foundation for implementing these techniques, several standards provide frameworks on machinery PHM. For example, while International Organization for Standardization (ISO) 13381-1 [4] provides recommendations for the development and application of prognostic processes, Institute of Electrical and Electronics Engineers (IEEE) 1856-2017 [5] offers a standard framework for PHM of electronic systems. ISO 17359 [6] offers general guidelines for creating condition monitoring (CM) systems. Additionally, ISO 13373-9 [7] focuses on condition monitoring and diagnostics using vibration signals. Specific standards for WT CM also exist, such as VDI 3834 [8], which emphasizes the measurement and evaluation of mechanical vibrations in WT systems and their components. However, no standards were identified that provide specific guidelines for prognostics applied to WTs. In this context, a PHM framework for WTs is provided in Fig. 1, presented in [9], which provides a structured approach to achieve these goals that consists of observation, analysis, and action blocks. By effectively employing these blocks in the wind energy sector, it is possible to predict the remaining useful life (RUL) of critical WT components, enabling proactive maintenance planning and cost reduction.

To bridge the gap between theory and practical application, a comprehensive review of RUL prediction techniques is essential to enhance the application of PHM in the wind energy sector. This review will help identify future research directions and address challenges in this domain. For this purpose, the following subsections of the Introduction will present two crucial modules for RUL estimation. First, WT systems and their common failure modes will be covered. Second, prognostics with a particular emphasis on RUL estimation will be presented, both of which are integral to the effective implementation of PHM in WT systems.

1.1. Understanding the system: WTs

WTs are systems that convert the kinetic energy of the wind into mechanical energy, which is then used to generate electricity. To effectively implement PHM strategies for WTs, it is essential to explore their target components and their failure modes. The main components of a WT drivetrain include the main bearing, shafts, gearbox, generator, and power converter [10]. A schematic of the main critical WT components is shown in Fig. 2.

Once the operation of the system and its components is understood, critical components will be selected according to the failure rate, downtime, and economical repair costs. In onshore WTs, components with higher failure rates include electrical and control systems, blades, pitch systems, and the yaw system, as reported in some studies [12,13]. Similarly, offshore WTs exhibit the highest failure rates in the pitch system, followed by the generator [13]. Higher average failure rates can be observed in offshore WTs compared to onshore, which can be attributed to the severe environmental and operational conditions inherent to offshore environments, such as elevated mean wind speeds and exposure to corrosive saltwater. In relation to downtime, the gearbox, generator, and blades emerge as the most critical subsystems for both onshore and offshore WTs [13]. Furthermore, failures associated with the gearbox, rotor blades, yaw system, and generator represent a higher expenditure, in the specified order [14].

Following the identification of the critical components in WTs, the next crucial step is to identify potential failure modes. Every component within a WT has its unique set of failure modes, and the associated downtime contributions are carefully documented and analyzed. The main failure modes of these components are presented in Table 1 [12,15–23].

WT components are susceptible to various failure modes that can impact their performance. Conclusions drawn from Table 1 underscore the importance of detecting and understanding these specific failure modes. Each mode has distinct characteristics and implications, necessitating tailored prognostic approaches. The diversity of failure modes across WT components emphasizes the need for specialized data acquisition techniques, CM strategies, and predictive algorithms. A comprehensive study of these failures is essential for selecting the most suitable prognostic approach.

1.2. Prognosis: remaining useful life estimation

Understanding the critical components and their failure modes is crucial for prognostics, which, as an essential step of PHM, aims to predict their RUL. This is defined by the ISO as "*the remaining time before system health falls below a defined failure threshold*" [4]. It is a key element in prognostics, as its estimation leads to effective predictive maintenance plans and, subsequently, to optimization thereof. Traditionally, after the comprehensive study of the system and its failure modes is done, RUL estimation has consisted of five steps [24]:

(i) data acquisition,
(ii) data preprocessing,
(iii) feature extraction,
(iv) health index or health indicator (HI) construction,
(v) degradation modeling and RUL prediction.

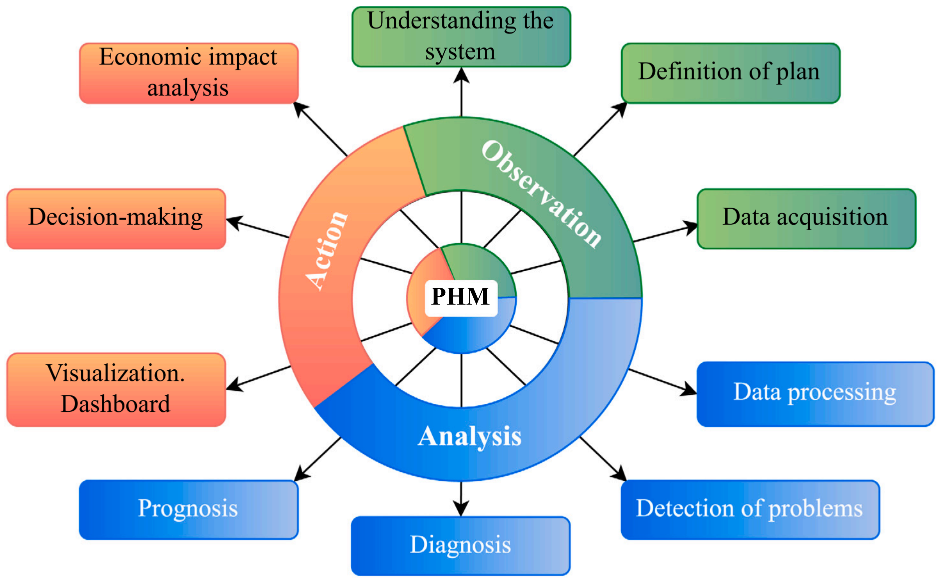


Fig. 1. Prognostics and health management framework [9].

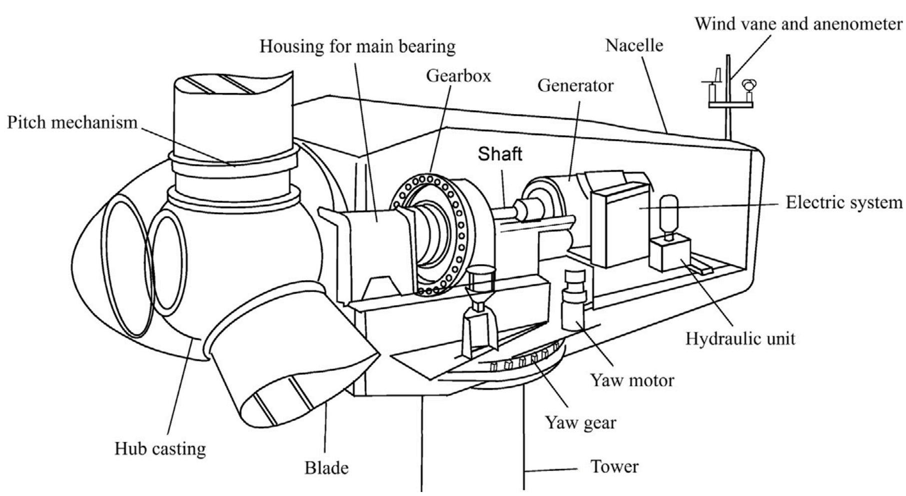


Fig. 2. Mechanical components of a traditional wind turbine [11].

Here, a HI of a component is a quantitative measure that provides an indication of its condition, performance, and proximity to failure [25]. Degradation models traditionally estimate RUL by predicting when the HI will cross a certain threshold [26]. However, as deep learning (DL) has raised, end-to-end predictions are feasible, which means that networks are trained directly with the raw input data, without any manual feature extraction. This minimizes or even eliminates the data pre-processing effort while maintaining the desired predictive performance [27,28]. Instead, traditional RUL estimation techniques (which can also use DL models to estimate the degradation trend) offer more interpretability, as each step in the process allows more transparency in understanding how the model arrives at its predictions.

The last step presented aims to model the evolution of system degradation and predict its RUL. Different approaches have been found in the literature for this, broadly categorized into three types: physics-based, data-driven models, and hybrid models. It should be highlighted that the selection of the model will depend on different factors, including the target component, its potential failure modes, the knowledge of its degradation physics, and the typology of the available data. Next, a brief description of each approach is presented.

1.2.1. Physics-based models

The first approach is known as physics-based. This group of prognostic methods aims to create mathematical models that explain how different failure modes occur due to the underlying physics of the system [29]. They describe the entire degradation process and the dynamic characteristics of the system, with the incorporation of degradation phenomena (mainly fatigue and discharge models) through mathematical equations [26], which are solved using numerical methods. The models are highly accurate as long as their underlying physics remain consistent across systems [30], and require less data than data-driven techniques. However, in the case of every engineered system, the development of a unique model and algorithm becomes challenging, especially with prior principles in real-world applications [30]. In such scenarios, data-driven models, which do not require a physical understanding of the system, seem more appropriate.

1.2.2. Data-driven models

Data-driven are commonly used when there is a lack of physical understanding or where the model exhibits a complex non-linear behavior. They only make use of historical data to predict the future state of a system and can be categorized into statistical and artificial

Table 1
WT component failure modes [9].

WT component	Failure mode	Description	References
Pitch system	Oil issues	Leaks, unscheduled oil changes, and unscheduled oil top ups	[16–18]
	Battery failure	Unexpected battery failure.	
	Motor failure	Wear or electrical issues	
	Motor converter failure	Electrical faults	
	Valve issues	Blockages or malfunctions	
	Accumulator problems	Leaks or pressure issues	
	Internal leakage	Internal leakage of proportional valve, internal leakage of solenoid valve, hydraulic cylinder leakage	
	Sensor failure	Inaccurate feedback for pitch control	
Frequency converter	Inverter failure	Generator-side or grid-side inverter failure	[17,19]
	Crowbar failure	Uncontrolled voltage spikes	
	Cooling Failure	Pump failures, blockages, or leaks in the cooling system	
	Control Board failure	Electric issues/component wear	
	Open/short circuit	Resistors, capacitors, and electronic switches	
Gearbox	Bearing failure	Axial cracking, spalling (contact fatigue), pitting (surface fatigue), and brinelling (fretting)	[20,21]
	Gear failure	Abrasive wear, pitting, cracking, scuffing	
	Oil leakage	Resulted from worn seals or damaged components	
Generator	Wearing	Wear and tear on generator components	[21,22]
	Electrical problems	Insulation failure and electrical imbalance	
	Winding damage	Due to electrical faults, overheating	
	Rotor asymmetries	Caused by imbalances	
	Bar break	Due to material fatigue or overloading	
	Overheating	Caused by electrical faults, inadequate cooling, or manufacturing defects	
Main shaft	Misalignment	Causing increased friction, vibration, and potential damage	[12]
	Crack	Caused by material fatigue, manufacturing defects, or excessive stress due to varying wind conditions	
	Corrosion	Caused by chemical reactions with the environment	
	Coupling failure	Caused by wear, misalignment, or overloading	
Yaw system	Abrasions of gear teeth	Caused by lack of greasing and bearing breakage	[15,23]
	Brake failure	Caused by brake pad wear, brake pad contamination and brake disc wear	
	Hydraulic failure	Hydraulic oil leaks and unstable braking forces	
	Electrical failure	Failures in the drive and drive motor, angular transducer, and abrasions in cables	
Blades	Crack fatigue	Fatigue cracks in the blades.	[23]
	Corrosion	Corrosion on the blades.	
	Aerodynamic imbalance/asymmetry	Due to manufacturing defects, non-uniform accumulation of ice, dirt, moisture or damage by lightning	

intelligence (AI)-based techniques. Statistical methods employ probabilistic techniques to fit data into stochastic process models, offering flexible RUL estimation [31]. However, they may struggle to align assumptions with the complexities of the real-world [32]. AI techniques, including machine learning (ML), DL, and ensemble methods, excel at learning from data. ML algorithms enable computers to learn instructions without explicit programming, while DL employs artificial neural networks (ANNs) to adaptively extract features for RUL prediction [33]. DL techniques minimize data preprocessing, but may face challenges with long-distance dependencies [34]. To address the issues raised from these techniques, ensemble methods are used, which combine predictions of more than one model for enhanced performance.

1.2.3. Hybrid models

Finally, hybrid models combine physics-based and data-driven approaches to obtain the strengths of each while minimizing their limitations. By incorporating domain-specific knowledge into machine learning (ML) frameworks, these models address issues such as the incompleteness of physics-based models and the limited representativeness of training datasets in data-driven models [35]. Notably, hybrid models offer enhanced interpretability and intuitiveness compared to purely data-driven approaches, which often lack transparency, earning the reputation of *black boxes* [36]. However, their efficacy in handling uncertainties increases algorithmic complexity and reliance on physical modeling, potentially demanding significant computational resources [26].

1.3. Discussion of related works on the field and contributions of this review

Several review papers in the field of PHM cover a wide range of mechanical systems, with WTs being mentioned only as specific examples among other applications. For instance, Liu et al. [37] review PHM techniques for electromechanical systems, but their focus is primarily on gear systems, not specifically WTs. Soleimani et al. [38] provide an in-depth analysis of diagnostics and prognostics for complex systems, again encompassing diverse applications beyond wind energy. Other reviews, like the one from Zonta et al. [39], focus on predictive maintenance strategies in Industry 4.0 environments without particular emphasis on WTs. Similarly, some reviews narrow their scope by focusing on specific prognostic techniques. For example, Xue et al. [40] examine similarity-based methods for RUL prediction, and Thoppil et al. [41] concentrate on DL algorithms. These reviews provide valuable insights but do not fully address the comprehensive challenges and methods for PHM specific to WT systems.

In contrast, other review papers focus more directly on WTs and their specific challenges. Rezamand et al. [42] offer a detailed analysis of WT components prognostics. However, there is minimal overlap between the studies discussed in their review and ours, as their work predates many recent advancements. While some common themes appear between the two reviews regarding general challenges and opportunities, their analysis, published in 2020, does not address more recent developments in data acquisition, processing, or feature extraction. Similar limitations are found in Zhang et al. [36] and Fox et al. [43], which are primarily concerned with prognostics techniques but do not cover the full pipeline required for accurate RUL estimation in WTs. Saidi et al. [26] review prognostics in renewable energy systems, but the scope is limited to prognostic techniques without exploring the broader context of health management. In comparison, Do et al. [44] present a more integrated approach, focusing on PHM control in utility-scale WTs. Our work builds upon and extends these reviews by not only covering prognostics techniques but also addressing a more comprehensive outlook on PHM in wind energy systems. Thus, this review offers a comprehensive analysis of the entire pipeline of PHM techniques for WT components. The key contributions of this work are summarized as follows:

1. This review covers the full spectrum of PHM processes for WTs, starting from system understanding, data acquisition, and feature extraction, to HI construction and maintenance optimization. Unlike previous reviews that focus on specific techniques or components, this work provides a holistic outlook on WTs PHM, addressing not only failure prediction but also the integration of these techniques into broader maintenance strategies.
2. This review also explores a wide range of degradation modeling techniques, such as physics-based, data-driven, and hybrid models, while evaluating their performance using specific metrics. This work deepens the understanding of the strengths and weaknesses of existing techniques, providing a critical evaluation that helps identify potential areas for improvement.
3. A detailed analysis of the challenges and opportunities specific to WT systems is provided. This includes handling uncertainty in RUL predictions, managing variable environmental and operational conditions, and addressing system complexity. The review highlights current research gaps and suggests future research directions for improving the robustness and applicability of PHM in wind energy sector.

Based on the need of PHM on WTs components, the aim of this review is to provide a comprehensive overview of the RUL estimation techniques in WTs blades and drivetrain components, considered as critical in Section 1.1, considering all the steps of the framework provided in Section 1.2: data acquisition, preprocessing, feature extraction, HI construction, degradation modeling, and RUL prediction models. Anomaly detection and diagnostics are beyond the scope of this review. The review presents in Section 3 eighty-seven works published from 2018 to June 21, 2024, thereby gathering recent advancements and trends in RUL prediction of critical WT components, including blades, gearboxes, bearings, and generators. The paper is structured as follows. Section 2 discusses the data sources, feature extraction and engineering, HI construction methods, and threshold determination; Section 3 explores the research carried out in the prediction of RUL classified into physics-based, data-driven and hybrid models; Section 4 introduces optimization strategies based on RUL prediction data; finally, Section 5 focuses on conclusions and key challenges for future research.

Table 2
Data used in the studied works.

Type	Source
Public dataset	First Industrial Big Data Competition [45], IEEE PHM Data Challenge 2012 for FEMTO [46], Harvard dataset [47], EDP Open data [48]
Simulation	Blades [49–51], gearbox [52], generator [53–55], others [56]
Accelerated degradation testbeds	Blades [57–60], gearbox [61–66], generator [67]
Operational proprietary CMS data	Gearbox [34,68–81], generator [29,82–96], others [97–105]
Operational proprietary SCADA data	Blades [106,107], gearbox [66,70,108–111], generator [112–115], others [103,116–118]

2. Data acquisition, preprocessing and health index construction

Developing accurate and reliable RUL prediction models for WT components begins with data acquisition, preprocessing, and HI construction. This section aims to provide a comprehensive overview of the data sources, feature extraction techniques, and feature engineering for HI construction approaches employed in the reviewed literature, as well as failure threshold determination methods used.

2.1. Data sources

The availability and quality of data are crucial factors in developing effective RUL estimation algorithms for WTs. Various data sources, including public datasets, simulations, accelerated degradation tests, and proprietary operational condition monitoring system (CMS) and supervisory control and data acquisition (SCADA) data, are used in the studied works. The data sources used in them are gathered in Table 2. It is important to clarify that public datasets can provide, among others, valuable real-world operational data, but are categorized separately due to their accessibility and distinction from private or proprietary datasets.

Access to public datasets offers several benefits to the advancement of RUL algorithms, as it allows researchers to develop, train, and validate models, as well as the evaluation and comparison between the results. However, the availability of such datasets is often limited, posing a significant constraint [26,119]. A great example of this is that in some of the studied works, the first step for WTs bearing RUL estimation is training models with commercial modular aero-propulsion system simulation (C-MAPSS) [120] data provided by NASA (examples in [63,77–81,105]), intelligent maintenance system (IMS) bearing data by University of Cincinnati [121] (examples in [75,122,123]), and PRONOSTIA [124] (examples in [34,46,94,96,101]) datasets, before making the leap to WT operational data. These are industrial datasets, but not directly data of bearings from WTs. The implementation and approval of a benchmark database would enable the evaluation of various PHM algorithms based on their effectiveness, interpretability, scalability, and dependability [125]. Recently, three public datasets containing information of WTs have been found in the literature. For instance, in [45], SCADA data from the first Industrial Big Data Innovation Competition was used, which contains SCADA data of 4 months of 2 WTs. In [47,126], SCADA data from a 1.5 MW WT with a hub height of 80 meters are used to estimate the RUL of main bearings, available in Harvard Dataverse [127]. In [48], a public dataset offered by Energias de Portugal (EDP) [128] is used, which comprises measurements SCADA measurements, meteorological recordings, and the logs of the component failure from five offshore WTs, collected at ten-minute intervals over a span of two years, resulting in a total of 87.208 samples for each turbine. The fact that only these have been used shows the limited availability of such datasets. Addressing this limitation requires the establishment of benchmark databases addressed to WTs to facilitate the evaluation of RUL algorithms across various metrics [125].

Generating synthetic data may be found as a solution to the problem, either using mathematical models, or simulation tools. FAST v8 [49,50], openFAST v2.3.0 [51], SIMPACK [52] and MATLAB/SIMULINK [53,54] tools have been used to model WTs. However, these are not always capable of generalizing effectively real-world conditions [129]. Moreover, the simulation models might not sufficiently capture the wide spectrum of WT configurations, environmental conditions, and operational scenarios.

Another option for scale-up testing is the use of accelerated degradation testbeds. Various of them have been employed to predict the RUL of WT components, such as blades, gearboxes, and generators. For blades, experimental setups include a helical gearbox and generator system to emulate WT dynamics [57], and diverse sensor arrays on composite blades to monitor damage progression under controlled vibrations [58]. Scale-down models, such as a 3-kW blade with embedded electrodes for condition monitoring, have also been used [59]. Additionally, pull and release tests on large-scale blades have provided data on structural response to stress [60]. On the other hand, gearbox testing includes custom planetary gearboxes to emulate operational conditions and detect faults using vibration data [61,62,65,66]. For generators, slewing bearings test rigs with hydraulic systems to apply forces and accelerometer data acquisition systems have been implemented to validate monitoring techniques [67]. These testbeds enable detailed analysis and prediction of RUL by simulating real-world stressors and collecting high-frequency data.

The final validation steps of RUL prediction involve applying the models to operational data, as dealing with data with multiple characteristics is one of the main challenges [38]. Regarding the proprietary operational data monitored in real WTs, SCADA data and CMS data are the primary sources [129–131]. Based on the study by Wang et al. [132], parameters obtained from a wind farm SCADA system can be classified into wind parameters (e.g., wind speed, deviation), performance parameters (e.g., power output, rotor speed, or blade pitch angle), vibration parameters (e.g., tower acceleration or drivetrain vibration), and temperature parameters (e.g., bearing temperature) which are very useful in this context because of their slow inertia [116]. Even if SCADA data are appropriate for predictive maintenance, their low sampling frequency (typically 10-minute averages) limits the usage of many conventional CM techniques (e.g., spectral analysis) [133]. However, after recognizing the importance of SCADA-based detection systems, companies are now starting to store data at higher frequencies, such as 1 Hz or higher, instead of averaging it over ten minutes [134]. Another primary challenge is the absence of standardized protocols for WTs SCADA data. The lack of uniformity in data formats, parameters, and naming conventions across different WT systems makes difficult the integration and analysis of information, limiting the development of cohesive PHM solutions [133,135]. The lack of uniformity in data formats makes it challenging to create algorithms and models that can work seamlessly across diverse WT systems, as well as robust analytics tools for performance monitoring and optimization. Data quality is another key factor to consider [135]. This negatively affects the PHM capabilities, possibly hiding short-lived events [136]. The current data enhancement methods for incomplete data cannot directly generate balanced and usable high-quality training datasets, and the highly imbalanced condition between healthy and unhealthy data is still one of the problems in the research of equipment life prediction [137]. Normal behavior modeling (NBM) models have emerged to solve this problem [116], where only healthy data are needed.

On the other hand, CMS data involve continuous monitoring of parameters such as strains, vibrations, acoustic emissions, oil debris counts, oil condition measurements, electric currents, and voltages to detect potential faults and predict maintenance needs [138]. It has been shown that CMS data are efficient for early fault diagnosis, but their real-world exploitation has some limitations due to the difficulty in isolating faulty vibration signatures acquired in such complex systems as WTs, considering their highly predominant non-stationary conditions [139]. One of the main types of signals used in predictive maintenance is vibration signatures, specially in rotating machinery such as drivetrains and pitch bearings [140].

In summary, given the challenges associated with diverse data, ensuring data quality and uniformity is crucial. Standardized protocols and higher-frequency SCADA data collection would greatly improve predictive maintenance for WTs. As the field progresses, combining SCADA and CMS data with better data enhancement methods would be key to developing accurate RUL estimation models. This comprehensive approach would lead to more efficient maintenance strategies, enhancing the performance and maintenance costs of WTs. These data would then be exploited with effective feature extraction techniques, a variety of statistical, spectral, and cross-domain approaches that are essential for transforming raw sensor data into meaningful WT HIs.

2.2. Feature extraction

Feature extraction in RUL estimation involves isolating key characteristics from raw data that significantly influence predictive accuracy. The complex structure and non-stationary operating conditions of WTs introduce noise and interference in data, which can adversely affect RUL prediction accuracy [94]. Therefore, it is necessary to obtain effective features that represent the degradation of the system. In practical applications, the extraction of features for monitoring processes or analyzing signals depends on various factors, such as the nature of the process being monitored or the characteristics of the measured signals. To mitigate the fluctuations in individual features resulting from variable amplitude signals, both time-domain and frequency-domain features are integrated. This combination yields more precise prognostic outcomes than those derived from each type of feature alone [141]. Ultimately, the selection of features will depend on the target component addressed and the type of signals used for the features [58] as shown in Table 3.

- **Time domain**: Time domain features are directly extracted from time-series samples of the measured signal and offer a balance between simplicity, interpretability, and effectiveness. These features are calculated based on the behavior of the signal over time, primarily using traditional statistical methods. They are particularly useful for detecting changes in the signal's characteristics, especially when a fault develops. Commonly employed functions used for this purpose include statistical moments (see Table 3). In WT components, for instance, temporal features are reliable for the planetary gearbox degradation [66], thus time-domain vibration features are selected, such as the range, variance, and root mean square (RMS), which represents the energy change of the signals [87,97]. Additional time-domain features often reported in the literature include skewness, crest factor, standard deviation, kurtosis, peak to peak, shape factor, energy, impulse factor, and margin factor.
- **Frequency domain**: By providing insights into the spectral characteristics of signals, these methods can identify the dominant frequencies associated with different components [123] and highlight the presence and locations of transients in the frequency spectrum [95]. These features are especially beneficial when faults manifest as distinct frequency patterns. The fast Fourier transform is a widely used technique for such frequency domain analysis. Examples of features in this domain include mean

Table 3
Feature extraction methods used in the studied works.

Feature domain	Methods	Signals
Time-domain	Statistics: mean, variance, kurtosis, skewness etc.	Vibration [56,69,80,82,89,94,97,123]
	Waveform parameters extraction: peak to peak, crest factor, impulse factor, energy, etc.	Vibration [63,83,95–97,106]
Frequency-domain	Power spectral density	Current, vibration [57,142]
Time-frequency domain	Hilbert-Huang transform	Vibration [58]
	Spectral kurtosis	Vibration [89,90,95,113,123]
	Spectral shape factor (SSF)	Vibration [92,95]
	Energy entropy	Vibration [84]
Cross-domain	Convolutional neural network (CNN)	Vibration [45], generated power, temperature, wind speed [110]
	Spatially multidifferential CNN (SMCNN)	Vibration [78]
	Deep belief network-self organizing map (DBN-SOM)	Vibration [65]
	Multi-branch 1D involution neural network (MINN)	Vibration [80]
	Multi-cellular long short-term memory (MCLSTM) units	Vibration [81]
	Gated graph convolutional layers	Vibration [102]
	Bayesian large-kernel attention network (BLKAN) layers	Vibration [104]
	Six-layer dual-view graph transformer (DVGTformer)	Vibration [105]

power spectral density (PSD) [142], maximum PSD, and mean frequency. Spectral kurtosis (SK) is considered a powerful tool for prognostics that can indicate and pinpoint nonstationary or non-Gaussian behavior and detect impulsive signatures since they could be masked by other sources of vibration [63]. It has been proved to be able to highlight local fault induced impulses from the noise background by identifying the corresponding resonance frequency band [62].

- **Time-frequency domain**: In this approach, features are extracted from the time-evolving frequency representation of the measured signal. Wavelet transform [62] and Hilbert–Huang transform [58] are commonly employed methods to capture the time–frequency characteristics of signals. The Hilbert–Huang transform uses the empirical mode decomposition (EMD) method to decompose a signal into intrinsic mode functions (IMFs), then the Hilbert spectral analysis method is applied to the IMFs to obtain instantaneous frequency data, as used in [58] focusing on WT blades. Another time-frequency domain feature used was energy entropy (EE), which represents the uncertainty of the energy distribution in different frequency bands. In WT bearings, this EE value decreases first and then increases as the fault severity increases [84].
- **Cross-domain**: as discussed in Section 1.2, DL can be used to develop an end-to-end model, or other applications such as feature extraction. These features extracted do not necessary belong to time-domain or frequency-domain features, thus they are classified as cross-domain in this review. One notable approach is the utilization of CNNs, which are used to automatically extract local features from raw sensor data. For instance, in [45], the effectiveness of CNNs was demonstrated in extracting local features from WT blades condition data, as localized features were extracted from sensor data, optimizing for spatial and temporal variations. Similarly, in [110], a two-dimensional CNN was employed to capture spatial information from the WT main bearing, thereby capturing location-based relationships within the data. Moreover, spatially multidifferential CNNs (SMCNNs) have been proposed in [78] as a means to extract features from various sensor signals of WT gearbox bearings. SMCNNs used defined stage division units to adaptively extract features, enhancing their capability to handle diverse data types.
In addition to CNN-based methods, the introduction of multi-branch 1D involution neural network (MINN) is found in [80], which addressed the need for feature extraction from WT gearboxes multi-source data. Recurrent neural networks (RNNs) also play a significant role in feature extraction for RUL prediction tasks. For instance, in [81], the authors integrated multiple multi-cellular long short-term memory (MCLSTM) units into their feature extraction module, enabling the extraction of diverse sub-cell outputs. Furthermore, DL methodologies extend beyond traditional neural network (NN) architectures to include graph convolutional networks (GCNs). In [102], gated graph convolutional layers were introduced to accurately extract degradation features from WT bearings multi-sensor signals. Moreover, a six-layer dual-view graph transformer (DVGTformer) proposed in [105] was used for spatio-temporal graph representations of multi-sensor signals, facilitating the extraction of degradation information of bearings.

In summary, effective feature extraction is essential for accurate RUL estimation in WTs. By using diverse methodologies, from traditional statistical measures to advanced DL techniques, researchers can enhance the predictive power of their models. The integration of time-domain, frequency-domain, and time-frequency domain features, along with cross-domain approaches, provides a robust framework for capturing the complex and non-stationary nature of WT data. These extracted features will be the first step for developing reliable HIs with feature engineering techniques.

2.3. Feature engineering for health index indicator construction

Feature engineering is a process of creating new input features from raw data using domain knowledge [25]. Once features that describe the system's degradation evolution are extracted, as seen in the previous subsection, feature engineering must be applied to construct the HI which will be used to predict the RUL. The techniques found in the reviewed papers have been classified into three categories: dimensionality reduction, data fusion, and feature selection, gathered in Table 4.

Dimensionality reduction or data compression are unsupervised ML tasks that reduce the number of features or variables in a dataset while preserving critical information, and, therefore, the complexity of the model, decreasing the chances of overfitting and poor model

Table 4
Feature engineering techniques for health indicator construction used for WTs.

Feature engineering	Method	References
Dimensionality reduction	Principal component analysis (PCA)	[46,61,67,88,96,123]
	t-distributed stochastic neighbor embedding (t-SNE)	[104]
	Fast independent component analysis (fastICA)	[114]
Data fusion	Ordered weighted averaging (OWA) operator	[29,93]
	Autoencoder (AE)	[76,79]
Feature selection	Particle swarm optimization (PSO)	[48,53,54]
	Minimum quantization error (MQE)	[65]
	Monotonicity, trendability, and prognosability study	[66,76,89,90,92,143]
	Filter-based selection	[93]
	Feature degradation ratio	[101]

performance [114]. In the studied papers, principal component analysis (PCA) [46,61,67,88,96,123], t-distributed stochastic neighbor embedding (t-SNE) [104], and fast independent component analysis (fastICA) [114] methods have been used to reduce dimensionality. PCA is the most prevalent technique, identifying directions of greatest variance in input data to capture essential information, with initial components often retaining significant defect-sensitive information [61]. PCA-derived HIs are instrumental in RUL prediction, which effectively predict when the HI will surpass a predefined failure threshold to enable timely predictions [96]. Another technique, t-SNE, reduces dimensionality by transforming high-dimensional data into a lower-dimensional space while preserving relative spatial structures, facilitating the extraction of features critical for RUL prediction [104]. As last approach, independent component analysis (ICA) decomposes multivariate signals into statistically independent subcomponents, with fastICA demonstrating effectiveness in identifying defect-sensitive features for RUL prediction [114,144]. Despite their utility, these techniques must be applied with caution, considering the specific characteristics of high-dimensional spaces and the nature of the data, as no single model universally outperforms others across all tasks [129]. For instance, PCA can impact interpretability and may not capture non-linear degradation patterns effectively [76,145]. Moreover, as it searches for the direction of maximum variance, it almost always separates regions of operation, rather than failures. Then, it can only be used when restricted to specific regions of operation. Otherwise, rather than detecting faults, it will detect different regions of operation [146]. Therefore, it must be highlighted the need for problem-specific model selection and careful application within defined operating regions to ensure effective features.

Data fusion techniques represent an alternative approach to feature engineering, aiming to combine information from multiple sources to construct the HI, enhancing decision-making and analysis [147]. In recent advancements, various data fusion techniques have been explored to improve the prediction of the RUL of WT components. One notable method is the ordered weighted averaging (OWA) operator, which combines attributes through a weighted summation, with the effectiveness of the OWA operator hinging on the careful tuning of its weighting factors, typically optimized using gradient descent [29,93]. Parallel to this, data-driven approaches, such as autoencoders (AEs), have shown significant promise. An AE is a type of unsupervised neural network that learns to encode the underlying representations of data in a high-dimensional space. It compresses input data into a lower-dimensional latent space using hidden layers, and then reconstructs the output to match the input as closely as possible, thereby enhancing RUL prediction. Building on this, a novel technique named the supervised multi-head self-attention autoencoder (SMSAE) has been introduced in [79], which incorporates a multi-head self-attention module within the AE framework to extract HIs directly from raw vibration signals.

As part of the last feature engineering technique found in the literature, feature selection has been done to identify meaningful features to construct the HI of WT components. Two ways have been used for this: algorithm-based and knowledge-based. In the first approach, particle swarm optimization (PSO) has been employed to optimize the selection of features by moving a population of candidate solutions through the search space to find the optimal solution [29,93]. Given the irreversible degradation of components, such as bearings, metrics including monotonicity, trendability, and prognosability are applied to determine whether a feature can represent the degradation process effectively [63]. Monotonicity measures the underlying increasing or decreasing trend of a feature, trendability reflects its correlation with time, and prognosability assesses the spread of failure values relative to the parameter pathway [92]. A suitable indicator should combine high values in these metrics to accurately represent degradation, overcoming the limitations of relying on single or uncorrelated features [123]. However, selecting the best indicator for RUL estimation can be complex, as high monotonicity might not coincide with high trendability and prognosability. To address this, a combined fitness function incorporating all three previous metrics has been proposed [89,92,148]. An example of a signal with high metrics in WTs is the RMS of bearings, which has been identified as a favorable candidate for health prognostics due to its high monotonicity and trendability [68,97]. Finally, other feature selection methods found in the literature include the feature degradation ratio, which compares the current feature to historical values considering vibration energy [101], filter-based methods that rank features using statistical metrics, selecting the highest-ranked ones for analysis [93], and minimum quantization error (MQE), which can effectively track gearbox degradation performance and provide an effective evaluation model for RUL prediction [65].

In other works, extracted features discussed in Section 2.2, which have the capacity to represent the degradation evolution, have been selected directly as HIs based on previous knowledge. Vibration signals are the most frequently chosen for assessing the health condition of WT components due to their reliability and informativeness. For instance, features sensitive to global damage, such as RMS, are commonly used. Despite its susceptibility to operational variations, RMS remains effective due to its explicit threshold in vibration criteria as stated in Germany industrial standard VDI 3834 [8]. Although RMS serves as an accurate HI, relying only on it has limitations in capturing the complexity of the input data, leading to reduced model performance [123]. Hence, other features like kurtosis, which indicates the degree of impact as sensitive to local damage, and EE, complemented RMS in bearing health assessment in [84,87]. In particular, kurtosis exhibits higher monotonicity in bearing vibration signals, making it a valuable feature in bearing degradation prediction [100,148]. Additionally, non-dimensional indicators such as waveform factor and crest factor, which are independent of working conditions, have been chosen for their sensitivity to damage and faults [96]. Furthermore, the shape factor, calculated through time-frequency decomposition, offers insight into signal shape, aiding in fault prognostics [92].

In conclusion, feature engineering is vital to construct HIs and predict the RUL of WT components. Dimensionality reduction methods such as PCA, *t*-SNE, and ICA reduce model complexity while retaining crucial information. Data fusion techniques, including the OWA operator and AEs, combine multiple data sources to enhance RUL prediction. Feature selection methods, such as PSO and metrics like monotonicity, trendability, and prognosability ensure the use of relevant features. A combined problem-specific approach using these techniques can effectively address the challenges of RUL prediction, as each of these techniques has its strengths and limitations. As highlighted above, the feature selection technique as HI will depend on specific characteristics of the component and the type of degradation.

2.4. Threshold determination

Once the HI is defined, and before constructing predictive models for RUL estimation, a crucial step is the determination of a threshold that delimits the end of the useful life for WT components. This threshold serves as a benchmark, enabling the prediction of RUL as the time interval between the current operational state and the projected point at which the performance of the component is expected to exceed the designated threshold. In the reviewed papers, three main strategies have been observed for threshold determination: expert judgment, statistical analysis, and risk management considerations.

In some studies, the threshold selection is based on experimental evidence, previous knowledge, or expert judgment [49,56,58,61–63,66,68–70,72,94,96,102,109,149]. For instance, in a study focusing on vibration signals of WT gearbox bearings, the RMS value of acceleration exhibited an amplitude jump time, regarded as the failure time, with the corresponding amplitude serving as the threshold for bearing health assessment [34,75]. Moreover, thresholds derived from operational parameters such as maximum speed and temperature [113], from proportional relationship between temperature and rotor speed [112], or based on the relationship between WT operation points and fuzzy rules, have been proposed [54].

Alternatively, statistical time-varying analysis criteria such as the 3σ and 6σ rules have been proposed for threshold determination [29,65,104,116,123,148]. In a normal or Gaussian distribution, approximately 99.7% of the data falls within three standard deviations (plus or minus) from the mean, which serves as a basis for the 3σ method. This method helps to identify significant deviations from normal operating conditions. The 6σ method further extends this concept. In this context, coefficients 3 and 6 are sometimes referred to as *complete failure criteria coefficients*, indicating their use in defining failure thresholds. Additionally, thresholds may be calculated as the average value of the noise to signal ratio (NSR) or through other statistical parameters [57,110,142].

Lastly, risk management considerations such as anomaly operation index (AOI) failure threshold determination involves selecting the critical AOI value at the point of failure, given by the ratio between the number of faulty data points and the total number of data points within a defined time window [114]. Furthermore, specialized techniques such as kernel space threshold transform (KSST) offer a data-driven approach by identifying common failure patterns from a set of indicators [101].

Upon completing the crucial steps for an effective HI construction and threshold determination, predictive models that accurately estimate the degradation of WT components over time and predict their RUL can be built.

3. Degradation modeling and remaining useful life prediction

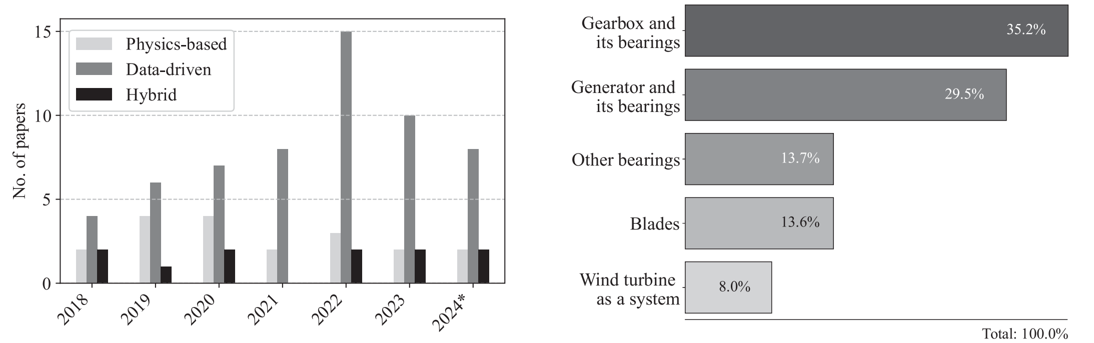
After the data preprocessing steps, the development of degradation models and RUL predictions are next. This section breaks down the approaches mentioned in Section 1.2 into specific algorithms for RUL prediction in WT components. Physics-based, data-driven and hybrid models will be discussed, providing the applications of each algorithm, as well as its advantages and disadvantages. The applications align with the most critical components in terms of downtime and repair costs, classified into blades, gearbox, generator, other bearings (which include predictions of RUL of bearings whose location is not specified) and those that consider WT as a system. It is important to note that most of the works found focus on bearing prediction, many of them located on the high-speed shaft (HSS). These can be gearbox high-speed bearings, gearbox intermediate-speed bearings, and generator bearings [150]. When the paper specifies the location of these, they are included in the gearbox/generator subsection. If not, they are included in other bearings. The distribution of eighty-seven papers among years and components can be found in Fig. 3. It can be seen that data-driven approaches are the most common ones to predict the RUL of WT components (Fig. 3(a)). Furthermore, the most studied components have been the gearbox and the generator, respectively (Fig. 3(b)). Fig. 4 gathers all the methods found in the literature, classified by type and component.

3.1. Physics-based models

Physics-based models have been widely used in estimating the RUL of WT components. These models require a fundamental understanding of physical processes and dynamic behaviors to predict failure modes [29]. By constructing kinematic or structural models, these methods can accurately describe phenomena such as crack growth and fatigue.

The application of physics-based models in WT prognostics is well-documented, with various studies demonstrating their effectiveness. These approaches go from bond graphs to fatigue damage accumulation, each aiming to incorporate detailed physical phenomena into RUL prognostics effectively [51,64,72,98,103]. For instance, physics-based models predict critical metrics like bearing loads [72] and internal shaft forces [98], employing fatigue analysis methods such as Miner's rule to estimate component RUL in dynamic wind conditions. Studies also explore comprehensive WT submodels, using aeroelastic simulations [51] and durability and damage tolerance analysis (DADTA) by determining the number of cycles until failure using material S-N curves or Goodman diagrams to generate holistic RUL predictions [151]. In one such approach, the modified life rating as defined in ISO 281:2007 [152] was applied to predict RUL, allowing for a more standardized estimation of bearing life based on load and material factors [64]. However, despite their precision, these models face key limitations. Accurate model development demands a thorough understanding of WT dynamics, and simplifying assumptions, such as linear load-damage relationships in Miner's rule, may not fully capture non-linear interactions under variable operational environments, potentially reducing model accuracy [35,103]. Additionally, reliance on specific data inputs, like strain time histories or frictional energy metrics, can constrain applicability in large-scale settings, where consistent data collection is challenging [72,98]. Computational demands further limit real-time monitoring feasibility, as the complex equations required may require extensive resources in certain operational contexts [35]. To enhance these approaches, adaptable methodologies that can accommodate non-linear behaviors and uncertainties are needed, particularly given the dynamic and varied operational conditions in wind energy systems.

To address these limitations and enhance the adaptability of RUL predictions under uncertain and variable operational conditions, Bayesian updating methods, including the Kalman filter (KF) and particle filter (PF), have gained attention in WT components RUL estimation due to their capacity for state estimation with uncertainty and noise. The KF and PF use recursive algorithms for updating state estimates in real-time, but they differ fundamentally in handling non-linearities and non-Gaussian noise. While the KF is optimal for linear systems with Gaussian noise, the PF is more versatile, allowing for accurate state estimation in non-linear systems with non-Gaussian noise, making



(a) Number of papers found in this review, categorized by modeling approach (physics-based, data-driven, and hybrid), over the years.

(b) Percentage distribution of RUL prediction techniques found across WT components.

Fig. 3. Classification of papers in this review: (a) by year and approach, and (b) by component. Note*: data for 2024 is only up to 21st June.

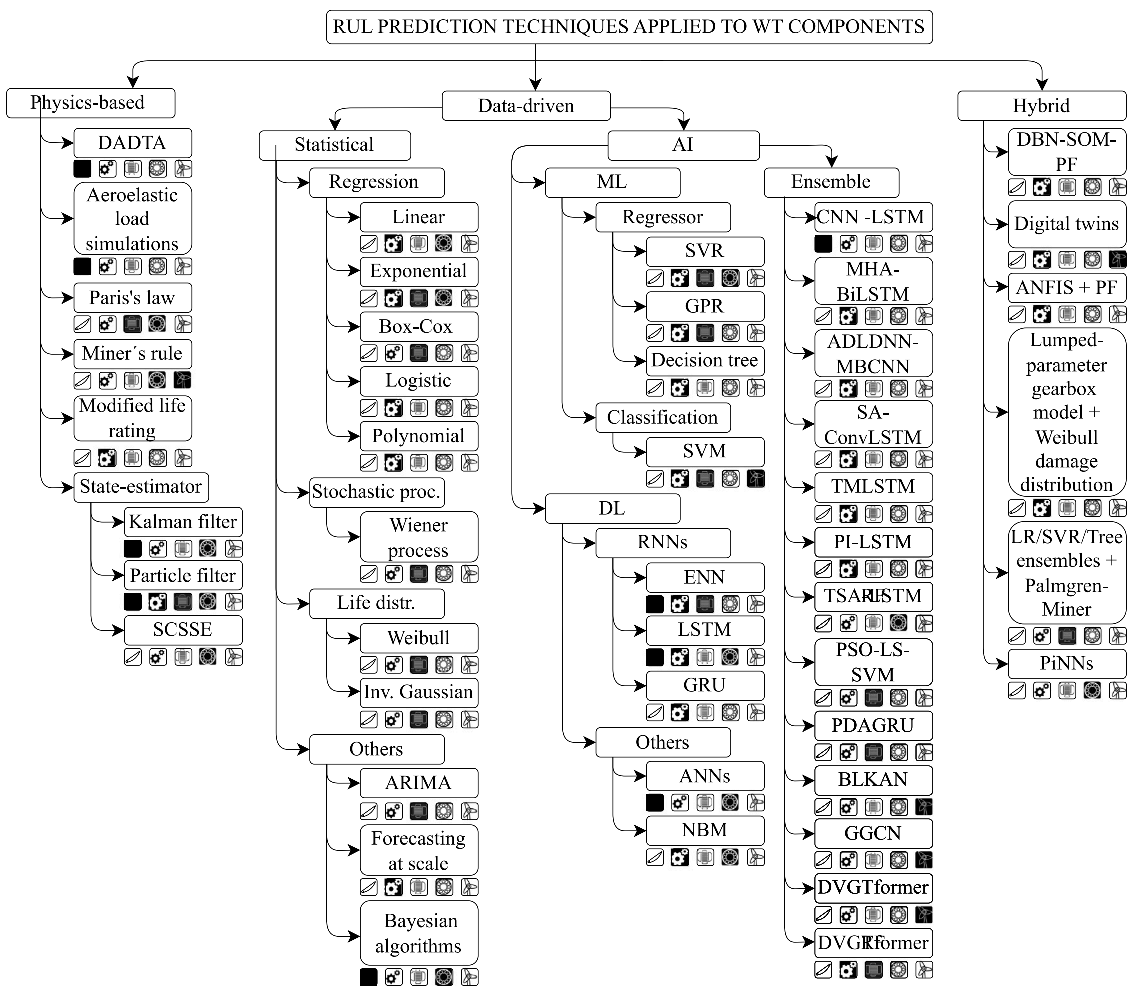


Fig. 4. Techniques found in papers to predict RUL of WT components. The five categorical boxes correspond to, in order: blades, gearbox and its bearings, generator and its bearings, other bearings, and the WT as a system. The components highlighted with a black background represent those studied with the corresponding technique.

it a particularly valuable tool in WT prognostics, where operational conditions can be unpredictable [63,153].

On the one hand, the KF is a recursive algorithm suited for linear, Gaussian state-space models, estimating the state of a dynamic system by minimizing the mean squared error [153]. By integrating classical crack growth equations like Paris's law with the KF, studies have demonstrated improvements in RUL estimation accuracy, particularly in cases involving phase delays in vibration data from high-speed shaft bearings in WTs [82]. Extensions of this approach, such as the Kalman smoother, further enhanced degradation tracking by considering state transitions over time, obtaining more accurate predictions of wear and failure trends across multiple states [100].

Additionally, innovations like the application of extended Kalman filter (EKF) with the maximum correntropy criterion enabled better performance of time-varying degradation under non-Gaussian noise conditions and provided robust confidence interval computation for uncertainty management [154]. While the KF approach is advantageous due to its straightforward assumptions on state-space models and its minimal data requirements compared to data-driven techniques [153], it faces limitations, as the linearity and Gaussian noise assumptions, foundational to the KF, limit its applicability in more complex, real-world WT scenarios, where non-linear dynamics and high measurement uncertainty are present [154].

On the other hand, PF, also known as the sequential Monte Carlo technique, is widely used for state estimation in non-linear and non-Gaussian state-space systems based on given measurements [101]. It estimates an output given the uncertainty of both the state and the measurements to solve the problem of characterizing the posterior probability density function (PDF) [58]. PFs are particularly useful when the posterior distribution is non-standard or multivariate. This versatility makes them particularly well-suited for forecasting a non-linear process, such as crack growth [49,50]. In the context of RUL estimation of WT components, studies conducted in [49,50,58,59,63] demonstrated the efficacy of PF in accurately predicting RUL under varying conditions. This is a key concept, as many studies make overly simplistic assumptions about stable work environments. A great example of considering varying operating conditions with PF is found in [49], by comparing RUL estimations for blades under variable versus constant average wind speeds. Using simulated blade responses in FAST v8, fatigue crack growth was modeled through Walker's equation to generate propagation data under fluctuating wind speeds (2 to 22 m/s) and a constant 12 m/s baseline. The PF method then estimated RUL by projecting crack growth over time, accounting for the nonlinear and stochastic nature of fatigue damage. Results showed that variable wind speed produced more conservative and realistic RUL estimates. Similarly, in [58], varying operational conditions were integrated into the prognostic stage through a dynamic degradation model paired with Bayesian estimation techniques. By using natural frequency shifts as damage indicators, the model naturally adapted to different operational scenarios. A Bayesian framework then continuously updated damage estimations as new data was collected, allowing for robust handling of uncertainties. PFs further enhance the system's adaptability by filtering out noise and managing stochastic variability. Observation equations account for diverse load and stress conditions, while validation against experimental data, inclusive of real-world noise, ensured reliability and accuracy under fluctuating conditions. In [62], due to varying speed operating conditions, vibration and speed data were measured at the sampling frequency of 97 kHz in a run-to-failure test. However, a limitation of these conventional PFs is their reliance on predetermined model parameters, which may lead to prediction errors due to the stochastic nature of machinery defect propagation under these varying conditions. To address this, an integrated state prediction and parameter estimation framework has been proposed [68], combining PFs with the expectation–maximization algorithm. Using hidden degradation state and in-process measurements, model parameters were adaptively estimated, improving the prediction accuracy to predict bearing defects from vibration signals in a real-world WT gearbox.

Another problem of PFs is the particle impoverishment problem, which arises when the number of particles with large weights significantly outweighs those with small weights, leading to a reduction in the diversity of resampled particles. As a consequence, the resampled particles may only represent a limited range of distinct state values, resulting in an inaccurate approximation of the posterior PDF of the state variable. To overcome the particle impoverishment issue, an enhanced particle filter (EPF) algorithm has been used [68], tailored for bearing RUL prediction in a 2.5 MW WT gearbox. The EPF algorithm overcame particle impoverishment issues, demonstrating superior performance compared to traditional PF methods. The work in [97] emphasized a model-based approach using an improved unscented PF to study bearings located in the gearbox high-speed shaft, generator drive end, and generator non-drive end. Moreover, built upon this foundation, in [101], degraded feature fusion models were integrated, along with threshold determination techniques, and self-constraint state-space estimator (SCSSE) to further enhance RUL prediction accuracy. These steps have been crucial towards more reliable predictive maintenance strategies for WTs, yet there is the need to improve PF approaches. While effective, EPFs are even more computationally demanding than traditional PFs, which could limit their practical use in large-scale predictive maintenance systems. Furthermore, additional feature fusion models and complex techniques may increase the system's dependence on high-quality data, which could limit their deployment in varying operational conditions. Future efforts should focus on balancing model complexity with real-time performance, as well as addressing the data requirements necessary for accurate and reliable RUL predictions.

Overall, through various innovative applications, these physics-based models have demonstrated their capability to enhance the reliability and accuracy of WT prognostics, ensuring better maintenance strategies and operational efficiency. Table 5 resumes the application of each technique, as well as its advantages and drawbacks.

In general, physics-based models for RUL estimation in WT components offer a robust method based on physical laws and dynamic behavior. If the physical or mechanical model of generator exists, the best choice of prognosis is a physics-based method [53]. One of the primary advantages of physics-based models is their precision. Since these models are grounded in the fundamental physics of material behavior and failure modes, they can provide highly accurate predictions. Physics-based models typically require less data compared to data-driven models, which makes them valuable in scenarios where data are scarce or expensive to obtain. Physics-based approaches are particularly adept at handling specific failure modes, such as crack propagation, where understanding material fatigue and stress is crucial [36]. However, physics-based models have limitations. They require a comprehensive understanding of the physical behavior of the system, and when the dynamics of the system are not fully known or highly complex, developing an accurate model is challenging [24]. Furthermore, these models often involve the solution of complex mathematical equations, which can be computationally intensive and time consuming [35].

Most of the current research in the field of RUL estimation through physics-based models focuses on blades due to their relatively straightforward modeling. In contrast, more complex components such as gearboxes and generators often require a data-driven approach. These systems exhibit intricate dynamics and failure modes that are not easily captured through purely physics-based models. Instead, data-driven techniques use extensive operational data to infer patterns and predict RUL more accurately. These data-driven models will be explored in detail in the next subsection.

3.2. Data-driven models

Data-driven models, which use historical data (e.g., sensor readings, operational history, and SCADA), are commonly used when there is a lack of physical understanding or where the model exhibits complex non-linear behavior, making the physics-based approaches discussed in the previous section impractical. These data-driven models have been classified into two main types: statistical models and AI models. Both approaches make use of large datasets to detect patterns and predict the RUL of WT components.

3.2.1. Statistical methods

Statistical model-based approaches estimate the RUL by constructing statistical models based on empirical knowledge. RUL prediction results are often presented as conditional PDFs based on the observed data [31]. Instead of relying on physical laws or principles, these

Table 5
Advantages and disadvantages of physics-based models found in the literature.

Method	Application	Advantages	Disadvantages
Durability and damage tolerance analysis	Blades prognosis [151]	Provides insights into long-term structural integrity, transferable techniques from aerospace.	Highly dependent on accurate material properties and loading conditions, complexity in modeling real-time damage growth.
Aeroelastic load simulations	Blades prognosis [51]	Captures dynamic interactions between turbine and environment, suitable for fatigue analysis.	Computational complexity may limit real-time implementation, and requires detailed environmental data.
Paris's law	High-speed shaft bearing (HSSB) prognosis [82]	Simple, widely applicable for fatigue crack growth analysis, integrates well with linear damage models.	Limited applicability to non-linear crack growth and complex failure modes, overlooks varying load conditions.
Miner's rule	Blade prognosis [149]	Easy to implement, intuitive for cumulative damage analysis.	Oversimplifies damage accumulation, limited in capturing non-linear or dynamic failure processes.
Modified life rating	Gearbox prognosis [64]	Incorporates detailed understanding of bearing dynamics and material fatigue, provides holistic predictions.	Requires extensive calibration and validation, lacks generality across different turbine models and environments.
Kalman filter	Bearings prognosis [154], HSSB prognosis [100]	Effective at estimating system state and predicting future behavior, requires minimal data.	Assumes linear system dynamics, struggles with non-linear effects and large uncertainties, data quality sensitive.
Particle filter	Blades prognosis [49,50,58,59], gearbox bearings prognosis [63,68], HSSB prognosis [97]	Handles non-linear and non-Gaussian processes well, robust in varying conditions, suitable for real-time updating of predictions.	Computationally intensive, requires accurate system modeling, suffers from particle impoverishment, complex to implement in real-time.
Self-constrained state-space estimator	Bearings prognosis [101]	Can deal with noisy and incomplete data, offers enhanced robustness in dynamic systems.	Requires precise system dynamic models, risk of convergence issues, computationally heavy for large-scale applications.

models are built by fitting observed data to stochastic process models or random coefficient models using probabilistic techniques. This section contains the most commonly used methods for RUL estimation of WT components, including regression, Wiener processes, and life distribution techniques.

First, regression has served as a fundamental statistical tool for estimating RUL by examining the relationship between independent variables (features) and a dependent variable (target). The main advantages of this approach are the simplicity and easy implementation, as well as the high interpretability of the results. Nevertheless, these models may not capture complex dependencies between the variables and the RUL. An example of linear regression can be found in [47], where the authors developed a RUL prediction method for WT bearings, achieving an accuracy rate of 99.18%. Although it may be seen as an impressive result, the authors do not explain the use of variables such as viscosity damage or fatigue damage to fit the linear regression model, which are obtained by other work [126] and not available in real cases. Therefore, the current work only concludes that, for this case, the linear regression model fits better to the available variables for predicting the RUL compared to RNNs.

Nonetheless, non-linear regression models are preferable, as the relationship between bending moments and dynamic drivetrain responses appears to be non-linear. This non-linearity is particularly evident when modeling the relationship between these variables [115]. The non-linear regression models include exponential, polynomial and Cox regressions. Several studies have successfully implemented exponential models in the RUL estimation framework [61,69,88,155]. However, a common limitation across these works is the necessity of incorporating additional steps, typically needed to refine input features or to adjust for varying degradation patterns, which exponential models alone may not capture effectively. Some examples of this dependency are pre-processing stages (e.g., HI construction through PCA [61,88,155] or a multi-stage degradation identification [156]) and post-modeling integration methods with polynomial models and NNs [69], which highlight that their effectiveness relies heavily on these complementary stages to achieve reliable and accurate RUL predictions. Other regression models, like logistic regression or Box–Cox regression, have been used in recent studies. While the logistic approach did not reach promising results in [114], a study that incorporated the Box–Cox transformation with the regime-switching model obtained favorable outcomes [157]. This adaptation allowed the model to adjust the dataset to meet the requirements of a switching linear system while also estimating the points of regime transition. Consequently, the model could operate effectively in environments characterized by uncertainty, nonlinear behavior, and phase transitions (or regime switches) between healthy and degradation states, where the forward degradation rate is typically low in the healthy phase and significantly increases in the fault phase [96]. Collectively, these studies establish non-linear regression models as an alternative for capturing diverse degradation trends of WTs components.

Secondly, another approach documented in the literature for RUL is the Wiener process, particularly suitable for modeling non-strict monotonic degradation [158]. Through applications of inverse Gaussian distributions and data fusion, these models enhance reliability even with sparse failure data and complex operational loads [67, 112,113,122]. Nonlinear adaptations, such as multi-stage monitoring, improve accuracy by capturing detailed degradation trends under varying conditions [113,122]. Each study contributes specific advancements: for instance, the use of Bayesian analysis with temperature data fusion for improved reliability [67], and real-time monitoring using a non-linear Wiener model tailored for complex bearing parameters [122]. However, some models are validated on datasets like IMS [122], which are not WT-specific as discussed in Section 2.1, or test rigs [67], limiting transferability to real systems, while assumptions like Brownian motion for noise and numerical approximations introduce uncertainties [158]. Despite these limitations, the Wiener process framework remains valuable for WT maintenance, enabling adaptive and data-informed decisions on component health.

Third, lifetime distributions play a crucial role in expressing performance degradation and quantifying reliability variation, offering a solution to model degradation processes and predict RUL [79]. These distributions, characterized by their PDFs, provide a statistical framework to analyze the longevity and failure characteristics of components. Among the various lifetime distributions, the Weibull distribution stands out as a widely adopted model in reliability assessment, which forecasts the future performance of the population by modeling a statistical distribution based on life data [85]. In practical applications, estimating RUL involves identifying the point at which the cumulative distribution function (CDF) of the Weibull distribution reaches a specific threshold [85]. Using the CDF, RUL estimation becomes feasible, providing valuable information on the RUL of the components. This also illustrates a valuable example of how to estimate component RUL based on limited data records. It is the case of the study in [85], where reliability metrics were explored using truncated WT generator data, employing Weibull and accelerated life testing analysis to identify best-fitted distribution models and propose predictive PDFs and hazard functions for the generator group. However, despite its utility, limitations remain in fully capturing operational uncertainty, which did not account for statistical uncertainty, an essential factor given WT variability. Future models integrating Monte Carlo simulations and bootstrap methods could address this gap by refining uncertainty quantification, improving both RUL accuracy and the robustness of lifetime models across other critical WT components.

Fourth, other statistical approaches such as auto regressive integrated moving average (ARIMA), forecasting-at-scale, and Bayesian algorithms have advanced RUL prediction for WT components by effectively handling time series data and adapting to the complexities of degradation using real-world SCADA data. ARIMA models, incorporating data preprocessing and anomaly detection, allow for accurate short-term forecasting based on SCADA signals, providing an extended lead time for maintenance planning [114]. However, they require meticulous management of data stationarity and validation to sustain reliability. Forecasting-at-scale techniques, which are particularly suited to gearbox health index generation using SCADA data, further enhance RUL prediction accuracy and demonstrate adaptability to operational variability, although they may struggle with highly irregular turbine behavior [70]. Bayesian methods offer a sophisticated approach to RUL estimation by capturing non-linear failure trends and providing real-time updates, proving effective for complex degradation processes such as those seen in blades and generator bearings [93,106]. While Bayesian algorithms show robust performance, their reliance on substantial computational resources and extensive data poses a limitation, though they exceed traditional methods in addressing uncertainty and dynamic failure characteristics specific to WT systems.

In conclusion, statistical methods estimate the RUL by constructing models based on empirical knowledge, offering information on the health prognostics of the components and facilitating maintenance decision-making. From regression analysis to Wiener processes, and lifetime distributions, a diverse array of techniques has been explored to capture the degradation patterns of WT components. Regression analysis, particularly non-linear models, has proven effective in capturing complex dependencies, while Wiener processes offer valuable insights into non-strict monotone degradation modeling. Furthermore, lifetime distributions such as the Weibull distribution provide a statistical framework for reliability assessment and RUL estimation. Other statistical approaches, such as ARIMA models and forecasting-at-scale algorithms, showcase the integration of data preprocessing, anomaly detection, and time series forecasting to improve prognostic capabilities. However, the degradation processes of key components in practice are often influenced by nonlinearity, switching behaviors, and the stochasticity of environmental conditions, resulting in complex and variable degradation characteristics. Classical statistical models often follow uniform model assumptions throughout the degradation process. As a result, traditional statistically driven methods often face limitations when modeling flexibility in complex operating conditions [96]. Together, these statistical methods contribute to a comprehensive toolkit for predictive maintenance strategies in the wind energy sector, enabling efficient operations and enhancing the RUL of WT components. Table 6 shows the applications of each technique, as well as its advantages and drawbacks.

3.2.2. Artificial intelligence-based methods

The exponential growth in the usage of AI has presented promising opportunities to progress from traditional statistical methods to AI-based techniques in the wind energy sector, enabling the full exploitation of SCADA and CMS data. These algorithms, including ML and DL models, have demonstrated significant potential in analyzing large amounts of operational data from WTs to accurately predict the RUL of critical components. When comparing both approaches, one significant advantage of DL over ML methods is that DL can process raw data directly without necessarily using hand-crafted features, thus enabling end-to-end learning [28]. Additionally, transfer learning in DL allows pre-trained networks to be reused through a fine-tuning process to solve new problems, making it increasingly popular in PHM applications. In contrast, ML techniques rely heavily on domain expertise for feature extraction and are generally limited to what they are explicitly designed for, whereas DL reduces the reliance on feature engineering. This subsection will explore recent advancement of ML, DL, and ensemble models used for RUL estimation of WT components.

Machine learning

ML techniques have become powerful tools for predicting the RUL of WT components. While ML can handle several tasks without tagged data through unsupervised learning, RUL prediction typically relies on supervised learning tasks, such as classification and regression, due to the necessity of labeled faulty data. In the context of classification models for RUL prediction, multi-class methods have been used, where labeled data are grouped in time windows. For instance, Carroll et al. [108] used labeled SCADA data from 1 month, 2 months, 3 months, and 6 months before failure, as well as data in healthy conditions (greater than 1 year), so the model could accurately predict gearbox failures with vibration data. The main advantage of this method is that it does not require a degradation model or trend, yet it still achieves good accuracy [56]. However, it only provides only a narrow range of RUL time predictions.

Regressor models, on the other hand, predict outputs where the target is a numerical variable, making them ideal for RUL estimation. In this work, the term *regressor* is used to denote a machine/deep learning model that predicts a continuous target variable (numerical) based on input features. Various supervised algorithms have been employed in this context, including support vector regression (SVR) and Gaussian process regression (GPR).

SVR is a popular regressor that extends the principles of support vector machine (SVM) to regression tasks, proposed in [159]. It aims to find a function that deviates from the actual data points by a value no greater than a specified margin, while simultaneously minimizing model complexity. It is particularly effective in handling high-dimensional data [145]; however, SVR models alone do not always demonstrate superior performance for RUL estimation. This conclusion is supported by four studies. First, a comparative analysis concluded that random forest (RF) are more accurate than SVR methods for bearing prognosis [46]. Second, the enhanced exponential exhibited superior accuracy than SVR in [73]. Third, the prediction error of SVR was significantly higher than that of their proposed model, the interval whitenization Gaussian process (IWGP) in [84]. Finally, the study in [160] concluded that SVR models were conservative because they underestimated the RUL. Therefore, to enhance SVR models, a combination of SVR with EMD and indicators such as RMS and kurtosis to denoise and extract fault signals from vibration data were introduced [87]. Validated with real datasets from two real-world WTs, this method proved improved results in comparison to SVR alone. Another improvement for SVRs is found in [95], where three innovative models integrate sparrow search algorithm (SSA) with SVR, RF regression, and GPR to accurately forecasted the RUL of HSSBs [95]. Their models, driven by vibration signal analysis and feature selection based on monotonicity, showcased exceptional performance when using SVR, successfully validating their approach using real-world data from a 2

Table 6
Advantages and disadvantages of statistic-based models found in the literature.

Method	Application	Advantages	Disadvantages
Linear regression	Bearings prognosis [47]	Simple and easy to implement, fast computation, well-understood and interpretable.	Assumes a linear relationship, which may not be realistic, sensitive to outliers.
Exponential regression	Gearbox prognosis [61], HSSB prognosis [69], other bearings prognosis [88,155].	Captures exponential growth/decay trends well, useful for modeling time-to-failure data.	Dependent of pre/post-stages.
Box–Cox regression	Bearings prognosis [96]	Can handle non-linearity through transformation, enhances normality of residuals.	Choosing the correct transformation parameter can be complex, interpretation of results can be less intuitive.
Logistics regression	Bearing prognosis [108]	Effective for binary classification problems, provides probabilities for failure events.	Assumes a logistic distribution of the outcome.
Polynomial regression	HSSB prognosis [69]	Can model non-linear relationships, flexible in fitting a wide range of data shapes.	Prone to overfitting with high-degree polynomials, computationally more intensive than linear regression.
Wiener process	Generator bearings prognosis [67,112,113,122]	Suitable for modeling continuous degradation, can handle random shocks and wear.	Requires complex parameter estimation, may need extensive historical data.
Weibull life distribution	Generator bearings prognosis [85,112]	Effective for reliability analysis, can model different failure rates over time.	Assumes a specific failure distribution, parameter estimation can be complex.
ARIMA	Generator bearings prognosis [114]	Effective for time series forecasting, can model various types of temporal data.	Requires stationarity of data, model selection can be challenging.
Forecasting-at-scale	Gearbox bearings prognosis [70]	Scalable to large datasets, can handle complex time series patterns.	May require significant computational resources, implementation can be complex.

MW commercial WT. These works highlight the need for enhancing SVR methods in RUL prediction applications.

The second ML regressor identified for predicting the RUL of WT components is GPR, a non-parametric, Bayesian approach to regression that provides a probabilistic prediction model. It is particularly useful for RUL estimation because it not only predicts the mean value of the RUL but also quantifies the uncertainty associated with the prediction. GPR has shown significant potential to model the complex relationships inherent in WT operational data. An example of it is found in [71], where results illustrated that the proposed GPR outperforms others in predicting RUL of real-world WT gearboxes with the lowest error among various ML techniques including decision trees, SVMs, ANNs, and mixture discriminant analysis. It is worth noting that, for the prognostic analysis of this dataset, twelve statistical features were initially extracted, followed by the application of PCA for dimensionality reduction. No alternative preprocessing techniques were explored in this study, leading to the conclusion that the GPR method is the most effective for this specific case. In another study, a data-driven method that combined the interval whitenization method with a Gaussian process (GP) algorithm for RUL prediction of WT generator bearings, the IWGP [84]. The main contribution of this work is to introduce the interval whitenization method for the first time to reduce the fluctuation of HIs caused by non-stationary operation conditions. More several prognostic techniques that address non-linear and non-stationary processes are discussed in this work, including match matrix, PF, EKF, and order tracking analysis. Each method has its limitations: match matrix is time-intensive and requires extensive historical data; PF struggles with high computation and storage demands; EKF performs poorly with approximately non-Gaussian processes; and order tracking analysis depends on accurate rotational speed and high time-frequency resolution. In contrast, the presented method offers faster computation, independence from rotational speed, and reduced sensitivity to time-frequency resolution. It also performs well with smaller datasets, making it a suitable choice for handling non-stationary conditions in this study. Therefore, GPR is a useful ML regressor in this context.

Deep learning

With recent advancements in DL, the use of NNs in the prognostics of WTs has shown promising results in recent studies. Several works have used ANNs to predict the RUL of WT components, demonstrating the effectiveness of these models in various contexts, mainly using vibrations signals: HSSB prognosis [69], gearbox prognosis [108] or WT system prognosis [48,90]. ANN outperformed other techniques with remarkable accuracy, specially using vibration data [69,108]. These approach also avoid various problems such as vanishing gradients and short memories in NN models while automatically feeding back the reflection of the system's past with each current observation [48]. This gives the opportunity to enrich the RUL model with knowledge achieved from degradation dynamics. Nonetheless, one of the main drawbacks of these works is that in most cases, only healthy SCADA data or too few faulty data are available, as the highly imbalanced condition between healthy and unhealthy data is still one of the problems to be solved urgently in the research of equipment life prediction [137]. To overcome this, as a particular approach of ANNs, NBMs have been used, capable of predicting RUL based solely on healthy SCADA. This approach can be applied to any wind farm, even when no faulty data have been recorded [116]. As an example of this, an extensive exploration of the use of high-frequency SCADA data was done in [74], employing advanced techniques to address imbalanced operational regimes and enhance detection capabilities in WT gearbox failure prediction, and using ANN-based NBM and one-class SVM. Another development of ANN-based NBM prognostic approach was done in [116], which relied solely on SCADA data to predict main bearing failure, enabling strategic maintenance scheduling several months in advance. Although these works were presented as prognostic studies, they do not provide specific values for RUL predictions. Instead, the authors suggested only that fault detection could occur months in advance, raising questions as to whether these studies are truly prognostic or merely fault detection in long time in advance.

Among ANNs, RNNs are particularly effective for time-series forecasting. RNNs are designed for sequential data processing; unlike traditional ANNs, RNNs can retain a memory of an arbitrarily long context window, thanks to the loops within their network architecture [29]. In addition, RNNs are capable of learning the variable-length sequences and sharing features learned in the subsequent neural nodes. They can effectively adapt to complex mapping methods due to their non-linear dynamics [41].

One of the fundamental RNN architectures is the Elman neural network (ENN), which is a local-feedback recursive RNN, categorized under non-linear state space models [92]. The ENN operates synchronously, maintains fixed recurrent weights, and its training is based on the backpropagation algorithm [83]. In the context of RUL estimation, ENNs have been mainly used for RUL estimation of HSSBs [83,89,92]. These applications have explored different input approaches, i.e., statistical time-domain features derived from vibration signals estimation [83]. Based on this, the authors enhanced their prognostic model by incorporating metrics such as monotonicity, trendability, and prognosability to increase robustness [89]. Additionally, a novel HI derived from the spectral shape factor entropy and the Teager energy operator was introduced in [92]. In another approach, the accuracy of icing failure prediction in WT blades was improved through a novel balancing algorithm based on boundary division synthetic minority oversampling technology (BD-SMOTE) and a multistep prediction process using multiple ENNs [157]. In conclusion, these fundamental RNNs can capture long-term and transient dependencies from time-series and sequential data, but they have several drawbacks, such as the exploding gradient problem. The issue occurs because the gradient is multiplied repeatedly as it propagates through time steps, leading to an exponential increase in the gradient's magnitude [161]. This can hinder the model's learning effectiveness. To overcome this issue, new versions of RNNs were introduced: long short-term memory (LSTM) and gated recurrent unit (GRU).

Traditional LSTM networks have widely been used for time series forecasting, as this approach tends to outperform statistics time-series forecasting methods, such as ARIMA, when it comes to long-term dependencies [145]. They are extensions of traditional RNNs and are specifically designed to overcome the vanishing gradient problem, a common problem in conventional RNNs. LSTM have a more complex structure that includes three gates (input, forget, and output gates) that control the flow of information in and out of the cell state, allowing the network to store and access data over a more extended period [162]. Nevertheless, they show limitations in RUL prediction, such as challenges posed by their inability to effectively capture global trends over time and to use backward and forward connections within time series data [76]. Several recent works have aimed to address these issues, as well as the challenges found on the way, as the effective feature extraction, addressing non-stationary operating conditions or the availability of limited samples. These have been used for blades icing prognosis [45], gearbox prognosis [76–78,81,143], and bearings prognosis [34,118]. Overall, one of the main limitations of traditional LSTMs is the heavy dependency on the massive degradation data, due to the nature of data-driven methods, which is not often available in real-world cases [143]. Because of this, it is essential to further work on techniques that enhance the training phase of the models from the limited samples. A way to do this has been the introduction of CNNs layers to the LSTM network for automatic feature extraction, which has been applied to SCADA data to predict blades icing [45]. Here, the dataset used has blade icing labels, not available in many cases, so it seems difficult to transfer to other environmental conditions. In another approach, the tree seed algorithm optimized long short-term memory (TSA-LSTM) was introduced [118], which uses tree seed algorithm (TSA) to search for the hyperparameters of the LSTM model globally to obtain the best combination of hyperparameters. Moreover, data augmentation techniques have been used [143], where a pre-interaction LSTM was designed to enhance the capture of sequential features in time-series data with limited samples, particularly during periods of interrupted continuous features. The prior knowledge of an empirical model for data augmentation was used based on the raw limited samples and then using the deep neural network (DNN) to learn from the augmented data. Another big limitation is that traditional LSTM do not highlight important degradation information in the prediction process, as they assume that all input data make equal contributions to the output [34]. Namely, they cannot learn non-stationary degradation characteristics by using different updating modes based on the time importance of information [78]. For instance, in the case of bearings, the vibration they produce changes as they wear down. As these parts get closer to failure, their vibration patterns become more intense [76]. In this context, the attention module is a tool that helps the network focus on the parts of these time-based vibration patterns that are really related to the degradation or wear, to address the varying input contributions over time [77]. An example of it was found in [34], where the self-attention ConvLSTM (SA-ConvLSTM) was introduced, which combined ConvLSTM architecture with a self-attention mechanism to selectively focus on important information. Based on the same idea, the multi-head attention bidirectional-long-short-term-memory (MHA-BiLSTM) was introduced [76], which incorporated a multi-head attention mechanism to assign weights to each item of the degraded data according to the similarity between the generated prediction data and the incoming degraded information. In the aim of addressing the same issue, a combination of a temporally and spatially multidifferential LSTM with the multitrend division unit and multicellular unit was proposed in [78]. By combining both, a spatiotemporally multi-differential deep neural network was developed for predicting the RUL, which enhanced the ability of feature extraction from the spatiotemporal perspective by using the multi-trend and multistage information. In a similar approach, the automatic multi-differential learning deep neural network (ADLDNN) was introduced. ADLDNN first groups data into levels of characteristic information, then uses a special multi-branch CNN to extract features for each level separately. Next, a bidirectional LSTM is used to identify trends in both forward and backward directions. Finally, a fully connected layer and regression layer predict the machine's RUL based on these features. In a final approach, motivated by the accuracy reduction problem caused by the time-varying characteristics of life-cycle data in the cross domain WT RUL prediction scenario, the MCLSTM was proposed [81] to obtain multiple differentiated distributions of monitoring data. There, domain adversarial and active screen mechanisms were used for transfer learning to predict the RUL of new equipment. These advancements highlight the continuous evolution of LSTM models to estimate the RUL in varying conditions, aiming to address the inherent challenges and improve prediction accuracy with limited faulty data.

Compared to LSTM networks, GRU networks have advantages of faster convergence and comparable prediction performance, since they combine the input gate with the forget gate into a single update gate [102]. Nonetheless, authors have found challenges in improving the accuracy, quantifying uncertainty, and adequately considering temporal and spatial dependencies in multi-sensor signals, which are critical for WTs. Specifically, the WT degradation process can be classified into normal, gradually degrading, and seriously degrading states, requiring adaptive data processing in various states for accurate predictions [80]. Existing models, however, often rely on fixed feature extraction patterns that do not adapt to the intrinsic characteristics of varying degradation states, which limits RUL prediction accuracy [80]. To address this issue, a concise self-adapting deep learning network (CSDLN) was developed [80]. It featured a MINN, and an embedded trend recognition unit that integrated a modified GRU to further reduce parameter count. Fully connected layers were then used for RUL regression. Based on the same idea, the parallel gated recurrent unit with dual-stage attention mechanism (PDAGRU) prediction model was presented in [94], enhancing the model with a non-parametric uncertainty quantification method based on the kernel density estimation and Monte Carlo dropout. This required less prior knowledge, while the parallel structure improved prediction accuracy. This last study also proposes the probabilistic RUL prediction based on transfer learning as future research, to get a higher generalization, robustness, and higher prediction accuracy with varying operational conditions. The study in [102] also worked on the two main challenges mentioned above: how to fuse multi-sensor signals and manage uncertainty, introducing a novel approach named gated graph convolutional network

(GGCN). The structure of GGCN involved a novel approach using a GRU as the basic block, where the traditional linear connections were replaced with graph connections. This facilitated the encoding of spatial and temporal information embedded in multi-sensor signals related to degradation states. The three studies show the advances in overcoming challenges that arise from the use of RNN, yet it is still an open research line.

These advancements highlight the growing applicability and effectiveness of DL-based methods in enhancing prognostic capabilities and operational efficiency in WT maintenance. Moreover, in an attempt to improve the performance of these models, ensemble methods have been proposed that combine predictions from multiple models, which are discussed in the next subsection.

Ensemble or composite data-driven methods

As a final approach to AI models, ensemble methods, also known as composite methods, have gained significant attention in the field of RUL estimation due to their ability to improve predictive accuracy by combining multiple models [163]. Some of these have been presented in the previous subsection, such as traditional LSTM and GRU enhancing models.

Among the diverse ensemble techniques, RF has been studied for its application in RUL estimation. This ensemble learning method constructs multiple decision trees and merges them, which makes it extremely great at taking care of tabular datasets with numerical features or categorical features with a limited number of classes. In contrast to linear models, RF can catch non-linear activities among the features and the objective [46]. The study in [46] conducted an investigation on the application of RF for RUL prediction in WT bearings. RF and SVR models were used, combined with PCA for feature selection. The results highlighted the superior accuracy of RF regression over SVR. Although these results were well-documented, additional comparisons with other techniques, such as those introduced in the present work, could provide further insights into the strengths and limitations of RF. It is the case of the work in [95] where SSA was integrated with different models to forecast the RUL of HSSB: RF, SVR, and GPR. Using real-world data from a commercial WT, their approach, enhanced by vibration signal analysis and monotonicity-based feature selection, showed that SVR ultimately outperformed RF in predictive accuracy. Therefore, RF is not the most suitable technique in every case. Lastly, RF has also been used as a multiclass classifier in [117], achieving an accuracy of over 80%. While it was easy to implement and cost-effective with limited data from WT rotors, only a few classes were categorized: less than seven days, between seven and thirty days, between thirty and sixty days, and over sixty days. However, the authors justified these intervals by explaining that they were defined in collaboration with the maintenance team, who confirmed them as suitable time frames for effective preventive actions.

In another approach, an innovative framework for prognostics was presented [53], focusing on the failure behavior of the doubly fed induction generator due to rotor electrical asymmetries. A PSO-least squares (LS)-SVM method was used, with parameter tuning for LS-SVM optimization and a radial basis function (RBF) kernel. Expanding their work, the authors [54] introduced a novel method for fault prognostics related to generator rotor winding, using feature level fusion and adaptive thresholds based on the fuzzy rules and WT operation point. The model was able to adapt to real-world cases such as variable dynamics, rotor speed, and simulated breakdown scenarios, and comparisons with SVM- and NN-based approaches, showcased superior performance. However, a key limitation of both studies lies in their exclusive reliance on MATLAB Simulink simulations rather than real-world SCADA data, which may limit their robustness under the diverse, unpredictable conditions found in actual WT environments.

Lastly, ensemble methods between DNNs and other methods, including optimization, are attracting attention nowadays [28]. These approaches leverage stacked architectures of neural networks to combine the strengths of different methods, improving prediction accuracy and robustness. Nevertheless, while these architectures benefit from enhanced performance, they face several challenges. The computational complexity of large-scale neural networks remains a significant limitation, though recent advancements in GPUs and cloud computing have alleviated some of these issues by accelerating training times. Moreover, while models such as the Bayesian large-kernel attention network (BLKAN) offer improved RUL predictions and uncertainty quantification, they have yet to fully address the uncertainties inherent in real-world data, including measurement errors, fluctuating working conditions, and the imprecision of RUL labels. Furthermore, models like the DVGTformer showed promise in handling multi-modal sensor data for better capturing degradation patterns but often struggle with the interpretability of results and the need for large labeled datasets. These limitations suggest that further advancements are necessary to integrate expert knowledge, handle diverse uncertainty sources, and refine multi-modal data fusion techniques for more reliable and interpretable RUL estimation.

Based on the reviewed works, it can be concluded that ML- and DL-based approaches are increasingly adopting ensemble or composite methods, often coupled with optimization algorithms to enhance performance, reflecting the evolving technological advances addressing the needs of predictive maintenance in complex systems such as WTs. All applications of each technique, as well as its advantages and drawbacks, are presented in Table 7. The LSTM and GRU improving ensemble methods have been included in each model (LSTM and GRU, respectively).

These approaches offer advantages in requiring less physical knowledge, effectively quantifying prognosis uncertainty, and processing high-dimensional data. However, they may suffer from interpretability issues and struggle with generalization when data availability is limited, a common challenge in many WT datasets. Explainable AI (XAI) methods have been developed to address the interpretability issues by providing insights into the decision-making process of AI models, making it easier to understand and trust their predictions. Examples of this can be found in other applications, such as C-MAPSS datasets, where SHapley Additive exPlanations (SHAP) and local interpretable model-agnostic explanations (LIME) approaches have been applied to RUL estimation [165,166], yet it is needed to be considered in wind energy application. Moreover, to further mitigate the challenges of limited data availability and enhance model robustness, recent research has introduced hybrid methods.

3.3. Hybrid models

The preceding sections have examined the limitations of purely physics-based and data-driven approaches, highlighting the interest for a hybrid approach to overcome these drawbacks. Specifically, challenges include developing robust models for complex systems with physics-based methods and addressing issues of interpretability and generalization when dealing with sparse or low-quality data in data-driven models. Thus, WT hybrid prognosis techniques, which combine physical degradation models and data-driven approaches, have been used in critical WT components due to their higher accuracy over individual prognosis methods [29]. They allow for accurate modeling of uncertainties and improvements in prediction capabilities. As the approaches of these hybrid models are very varied, they have been ordered according to their application rather than their nature.

Hybrid models for gearbox health monitoring have seen significant advancements, particularly with the integration of SCADA data and physics-based models to enhance fault detection capabilities for WT gearboxes. A great limitation of SCADA systems is that, although they provide valuable insights into the overall condition of WTs, they often lack the specificity needed for monitoring individual gearbox components like bearings. To address this, models that enriched bearing

Table 7
Advantages and disadvantages of AI-based models found in the literature.

Method	Application	Advantages	Disadvantages
Support vector regression	Gearbox bearings prognosis [46,84,160], HSSB prognosis [73]	Effective in high-dimensional spaces, robust to overfitting in high-dimensional feature space, works well with small datasets.	Can be computationally expensive, choice of kernel can significantly impact performance, sensitive to the choice of hyperparameters.
Gaussian process regression	HSSB prognosis [95], gears prognosis [71], bearings prognosis [84]	Provides probabilistic predictions with uncertainty estimates, works well with small datasets, flexible in modeling complex relationships.	Computationally expensive for large datasets, requires careful selection of kernel functions, sensitive to the choice of hyperparameters.
Decision tree	Gears prognosis [71]	Simple to understand and interpret, requires little data preprocessing, can handle both numerical and categorical data.	Prone to overfitting, especially with deep trees, unstable as small variations in data can lead to a completely different tree, biased towards dominant classes.
Support vector machine	Gearbox prognosis [108]	Effective in high-dimensional spaces, robust to overfitting, works well for time-windowed classification tasks.	Computationally intensive, choice of kernel and hyperparameters is crucial, not well suited for very large datasets.
Traditional ANNs	HSSB prognosis [69], bearings prognosis [104,105,108], gears prognosis [71], generator prognosis [90,164], WT prognosis [48,116]	Can model complex non-linear relationships, highly adaptable to different types of data, good at handling large and diverse datasets.	Requires extensive training data, prone to overfitting without proper regularization, computationally intensive.
Elman neural network	HSSB prognosis [83,89,92]	Can capture complex non-linear relationships, adaptable to different problem domains, good at handling noisy data.	Requires a large amount of training data, computationally intensive, can be prone to overfitting without proper regularization.
Long short-term memory	Blades prognosis [45], gearbox prognosis [76–78,81,143], bearings prognosis [34,118]	Effective at capturing long-term dependencies in time series data, robust to varying sequence lengths, good at handling sequential data with complex patterns.	Requires large datasets for training, computationally expensive, difficult to interpret.
Gated recurrent unit	Generator prognosis [94], gearbox bearings prognosis [80], WT prognosis [102]	Simplified structure compared to LSTM, effective at capturing dependencies in sequential data, faster training time than LSTM.	May not capture as complex patterns as LSTM, requires large datasets, still computationally intensive.
Random forest	Gearbox bearings prognosis [46], HSSB prognosis [95], WT prognosis [117]	Robust to overfitting by averaging multiple decision trees, can handle large datasets and high-dimensional spaces, good at capturing non-linear relationships.	Can be less interpretable than single decision trees, requires more computational resources for training, sensitive to the number of trees and other hyperparameters.

fault signatures by incorporating physics-based simulations based on gearbox design parameters were proposed [109]. This approach has led to the emergence of WT digital twins (DTs), which aim to offer more precise, component-level monitoring [52,111,167]. DT implementations, such as those in [52,111], combine CMS and SCADA data with physics-based models. However, these models still face limitations; for instance, though effective for RUL estimation through state estimation and fatigue analysis methods were carried out, the model was validated only under fixed load conditions, raising concerns about its adaptability to real operational variability [52]. The DT proposed in [167] addressed some of these concerns by testing their model under varying wind conditions, improving the applicability to fluctuating loads. Yet, even in more refined models [111], which use high-fidelity dynamic load data to calculate cumulative damage in real-time, there are still gaps. Specifically, while this model estimates accumulated damage accurately, it stops short of a full prognostic output by not converting this information into direct RUL predictions.

Expanding on gearbox applications, in recent years, predictive models for RUL in WT gearboxes have predominantly been validated on test benches under controlled, accelerated life testing conditions. While these studies employing hybrid methodologies like deep belief network (DBN)-self-organizing feature map (SOM)-PF [65] and ANFIS-based PFs [57,68] demonstrate high accuracy in degradation prediction and uncertainty reduction, their focus remains largely on idealized setups or single, specific failure modes. This controlled environment, though valuable for initial validation, limits the generalizability of findings to real-world applications, where WTs operate under variable speeds and complex environmental conditions. These models, while innovative, rely on signal resampling techniques and noise representations tailored for test conditions, raising questions about their robustness in operational turbines exposed to diverse stressors. Furthermore, while some models incorporate field data from SCADA systems and failure records, such as the multi-stage RUL models in [72] analysis across multiple WTs (considering non-stationary conditions, wind speeds, directions, and even gearboxes provided by multiple suppliers) this practice remains rare. Thus, there is a pressing need for further research focused on real-world validation, where models are tested across a broad spectrum of operational conditions, turbine types, and failure modes to ensure their practical applicability and reliability in predicting RUL in active wind farms.

In an innovative application of hybrid models for WTs generators, a DT framework to predict the RUL of generator-side bearings in the high-speed shaft was developed in [115]. This framework estimated aerodynamic hub loads and monitored accumulated fatigue damage using various regression techniques, such as linear regression, SVR, and tree ensembles, complemented by low-fidelity physics-based models. Fatigue damage and RUL calculations adhere to ISO 281:2007 guidelines [152] and the Palmgren-Miner model. While low-fidelity, quasi-static models proved effective for virtual sensing with low error rates in bearing damage prediction, the study highlights the challenge of capturing the intricate dynamics of the drivetrain, which are more complex than those of the tower and blades. Given that most operators lack sufficient drivetrain data, further research is needed to assess the uncertainties associated with quasi-static models and to improve model fidelity for real-world applications.

Another application of hybrid models used for WT main bearing can be found in [99], where the authors tackled the complex issue of predicting the RUL by accounting for the critical, yet uncertain, role of grease condition in fatigue life. Their hybrid model, using a physics-informed neural network (PiNN) embedded within an RNN cell, integrates reduced-order physics models for bearing fatigue accumulation, ISO 281:2007 life formulas [152], and NN representations for grease degradation. This approach aims to reduce inaccuracies in bearing life predictions caused by variable grease quality, which traditionally poses significant challenges due to degradation mechanism uncertainties and batch quality variations. The model was validated with synthetic data, derived from a National Renewable Energy Laboratory database, which was augmented to emulate SCADA data resolution and extended for long-term projections. Their model highlights the practicality of integrating SCADA-like data, such as wind speed and bearing temperature, although real-world data might require additional preprocessing to address missing values and variability. They acknowledge, however, that the current load models in use assume standard operational conditions, omitting extreme loads from events like startups, emergency brakes, or yaw misalignment. The study suggests that further robustness could be achieved by incorporating field data from proximity sensors, allowing the model to better capture load variations. Additionally, they propose that the framework's uncertainty quantification for grease quality could be expanded to account for other sources of uncertainty, such as noise from inspections or environmental factors across multiple sites. While this study provides valuable insights at the site level, future extensions to broader environmental contexts across multiple wind farms could enhance its generalizability and predictive accuracy in diverse conditions.

As a final application of the hybrid models, studies have considered WT as a holistic system of components [56,168]. The study in [168] dealed with fault prognosis ofWT in presence of multiple faults combining physical modeling, data clustering, and a geolocation-based degradation estimation. The model compensated the lack of failure information of early-stage data by using simulated physical data to represent both normal and faulty states, addressing gaps in data. Similarly, in their related study [56], the proposed method for RUL estimation is based on the similarity measurement between the reference attributes of the failure operation, identified offline, and the current attributes calculated continuously online at each sampling time. The main advantage of this geometric approach based on Euclidean distance calculation is the lack of requirement of prior knowledge about the profile of the degradation process. However, structured and unstructured uncertainties were managed through spherical clustering thresholds, though further real-world data is needed for refinement. Additionally, PCA was used to handle variable dependencies, though overlapping fault clusters and noise in degradation rates suggest areas for future improvement [168].

Collectively, these studies highlight the efficacy of hybrid approaches in addressing the inherent limitations of physics-based and data-driven methodologies, thus advancing the accuracy and interpretability of RUL estimation for WT components. However, the development of hybrid models remains challenging, particularly in terms of model selection and uncertainty characterization. The maturity of these algorithms has not yet been realized. Still, there is a growing interest towards adopting these hybrid models, such as the PiNNs, marking a direction away from purely physics-based or data-driven approaches in future research. There is no single definitive approach within hybrid models as seen in this subsection, but each of them offers unique advantages and limitations.

3.4. Model performance evaluation metrics

Once the models have been developed, evaluation metrics are used to validate them. This subsection presents the metrics that have been used in WT applications, although the models are not properly comparable, as they have not been validated with the same datasets. Accuracy is a commonly used metric to evaluate the performance of regression models in WT RUL prediction. Yue et al. [45] used accuracy as one of the primary metrics to evaluate their CNN-based model for blade fault detection. Similarly, Guo et al. [72] used accuracy to measure the performance of their model for gearbox monitoring. Additionally, root-mean-square error (RMSE) has been widely adopted as the reference evaluation metric. For instance, Lazaro et al. [70] used RMSE to evaluate their SCADA-based model for gearbox health monitoring, while Pan et al. [65] employed RMSE to assess the accuracy of their DL model in predicting gearbox RUL. Other studies, such as those by Xiang et al. [77] and Pagitsch et al. [64], also used RMSE to validate their prognosis models, highlighting its applicability in this field. Mean absolute error (MAE) and mean absolute percentage error (MAPE) are alternative metrics to RMSE, offering different perspectives on prediction errors. He et al. [75] and Li et al. [34] used both MAE and MAPE to evaluate their vibration-based prognosis models for gearbox bearings. These metrics provide insights into the average absolute deviation and percentage error of predictions, respectively, allowing for a more intuitive understanding of model performance.

The diverse range of metrics used in WT health monitoring and fault prognostics reflects the complexity of this field. However, standardized model performance evaluation metrics are needed to facilitate meaningful comparisons between models. Establishing common benchmarks and methodologies for evaluating WT health monitoring and fault prognostics models would enhance transparency and facilitate advancements in predictive maintenance strategies.

4. Maintenance optimization based on remaining useful life

PHM with RUL predictions enables prescriptive maintenance for WTs prior to failure, thus minimizing corrective and preventive maintenance that may be expensive and cause long downtimes. Built upon the predictions of RUL, suggestions about specific actions to optimize maintenance schedules and mitigate failure risks are made [39]. Prescriptive maintenance not only anticipates when failures may occur, but also recommends the most cost-effective, efficient, and timely interventions. However, despite its recognized potential, prescriptive maintenance has seen limited development over the last decade. Although it has been highlighted as a promising evolution of predictive maintenance [169], there remains a notable gap in the implementation of robust methodologies that translate its theoretical benefits into practical applications, particularly in the wind energy sector. Consequently, prescriptive maintenance for WTs remains at an early stage of maturity. Nevertheless, its potential for reducing costs, enhancing efficiency, and improving the resilience of wind energy systems is undeniable.

As this prescriptive approach needs to evolve, maintenance decision-making emerges as a critical module in the PHM framework (Fig. 1). It is defined as the process of selecting the logically best choice from a list of available options [170]. In the context of PHM, it involves the use of information obtained from RUL predictions to inform timely and optimized choices regarding maintenance actions. This insight, in turn, facilitates strategic decision-making, allowing for the prioritization of maintenance tasks based on the urgency and severity of potential failures. Thus, RUL prediction provides a powerful basis for decision-makers to plan predictive maintenance based on future conditions [171].

In the conventional approach, maintenance optimization relies on an analysis of degradation thresholds, identification of maintenance opportunities, definition of maintenance workflows, spare parts availability, and other elements directly associated with maintenance tasks, among others. Nevertheless, as engineering systems such as WTs become more complex and integrated, the number of factors and parameters to be considered for maintenance optimization is much larger than for a single independent system [135]. This process can be viewed as an optimization problem, which can mathematically be defined as a multi-criteria optimization problem [135,170,172]. In such problems, multiple objectives must be achieved simultaneously. For example, maintenance optimization might involve minimizing costs, maximizing reliability, and ensuring system availability, all while adhering to various constraints. These constraints could include operational limits, resource availability, and safety requirements. The goal is to find the best compromise among these competing objectives. Based on a study by Syan et al. [173], the most common criterion for general maintenance optimization is cost, considered in 61.3% of all applications. This is followed by availability at 28% and reliability at 26%. In offshore WTs, cost minimization and reliability maximization are the most widely used objectives [172].

An example of the most representative mathematical model known as multi-criteria optimization is found in [174]. In this paper, a maintenance strategy optimization framework is built for offshore wind farms, considering the uncertainty with Monte Carlo methods. The proposed maintenance model for offshore wind farms operates by making use of RUL predictions to optimize maintenance strategies, assuming that all WTs in the farm are of the same type and represented as a series of critical components, each degrading over time until failure. This degradation is modeled using a Weibull distribution. Then, three types of maintenance opportunity are considered: failure-based (when an offshore WT breaks down due to degradation failure of a component), incident-based (when an offshore WT suffers from a sudden critical incident), and aging-based (when no failure occurs in the farm, but a certain percentage of components reach a specific degree of degradation). Thus, if no degradation failure or incident occurs, the aging-based opportunity is determined based on component condition, which is defined by the RUL. Depending on the condition of the components, four types of maintenance actions are considered: failure replacement, preventive replacement, major repair, and basic maintenance. These are triggered based on two kinds of output: maintenance-related cost, which includes the cost of materials for repair, mobilization, vessels, and technicians; and production losses during turbine downtime, evaluated based on wind speed data and the design parameters of the WTs. These costs are assessed to ensure that the maintenance strategy effectively balances cost-efficiency and operational reliability. The proposed approach was implemented on a 150 MW offshore wind farm located in the North Sea, comprising 50 WTs, each with a capacity of 3 MW. Each WT consists of five critical components: the gearbox, generator, rotor and blades, main bearing, and pitch system, which reflects a high correlation between these components and those identified in this review. The paper concludes that enhancing RUL prediction accuracy of WT components is crucial to planning an effective maintenance strategy that balances cost-efficiency with operational reliability.

Another method to determine the optimal predictive maintenance opportunity for a WT with RUL predictions is the real option analysis (ROA) [175]. It is a financial modeling technique used to evaluate the value of investment decisions under uncertainty. It treats investment opportunities as real options, similar to financial options, giving the decision-maker the right, but not the obligation, to take certain actions in the future. An example of this is found in the work of Lei and Sandborn [176], where a simulation-based ROA model is used to maximize the value of predictive maintenance within the context of a power purchase agreement (PPA), the contractual agreement between energy buyers and sellers. This framework enables maintenance scheduling by evaluating the balance between maintenance and potential revenue loss. The model considers the operational state of multiple WTs, energy delivery targets, pricing, and penalties for under-delivery defined in the PPA. Here, the impact of RUL on the maintenance model is significant. The model simulates various future scenarios to determine the optimal maintenance timing, factoring in uncertainties in RUL predictions and wind speed. The inclusion of RUL allows the model to adjust maintenance schedules to ensure optimal decision-making, considering the economic implications of different maintenance strategies.

In a different approach, Mazidi et al. [177] presented a hybrid method combining NNs and the proportional hazard model to monitor WT behavior for PHM. This approach evaluates the efficiency of previously implemented maintenance plans [170]. Initially, a NN is constructed to model the normal behavior of the WT using data collected from a SCADA system. By comparing the real-time data with the predictions of NN, a deviation signal is generated, indicating the state of stress and the health status of WT. This signal is then used to assess the effectiveness of past maintenance actions and provide recommendations for future maintenance planning. Notably, this work considers the impact of maintenance decisions on the overall system. RUL was incorporated into the maintenance schedules and the cost function [170].

In conclusion, accurate prediction of RUL for WT components is essential in the effective decision-making of maintenance plans. By using RUL predictions within a predictive maintenance framework, operators can proactively schedule maintenance activities based on the anticipated health deterioration of critical components. This approach not only minimizes the risk of unexpected failures and associated downtime but also optimizes maintenance costs by enabling timely interventions before components reach critical states, thus maximizing the operational reliability and longevity of wind farms. Therefore, enhancing the accuracy of RUL predictions is fundamental for achieving an optimal balance between cost-efficiency and operational reliability in WT maintenance strategies.

5. Conclusions and future research lines

This review illustrates the dynamic field of RUL estimation for WT components, showcasing the evolution and diversity of methodologies and their respective challenges. The comprehensive analysis of eighty-seven papers published since 2018 on RUL prediction models reveals a clear alignment with the identified critical subsystems and failure modes within WT systems in Section 1. The predominant focus on predicting the RUL of gearboxes (35.2%), generators (29.5%), and blades (13.5%) aligns with the findings from the downtime analysis and failure costs in the introduction, where these components emerged as the most critical.

In the reviewed papers, all the RUL prediction framework steps have been studied, from data acquisition to data preprocessing, feature extraction, HI construction, and degradation modeling and RUL prediction (although novel DL models offer end-to-end approaches, skipping all the intermediate steps). As the first step, diverse data sources have been identified, including public datasets, simulated data, accelerated degradation testbeds, and proprietary operational data. The establishment of a benchmark database would facilitate the initial validation of PHM algorithms by evaluating their effectiveness, interpretability, scalability, and reliability. Another significant challenge identified is the absence of standardized protocols for SCADA data in WTs. The lack of uniformity in data formats, parameters, and naming conventions across different WT systems makes the integration and analysis of information difficult, limiting the development of cohesive PHM solutions.

In terms of preprocessing, feature extraction techniques have encompassed various domains such as time-domain, frequency-domain, time-frequency domain, and cross-domain. However, a definitive guide is needed to determine the appropriate domain for each characteristic related to specific components and failure modes. Feature engineering has explored a range of dimensionality reduction, data fusion, and feature selection algorithms. Yet, further research is essential to enhance the accuracy of HIs. Particularly, the selection and development of dynamic thresholds needs additional investigation to improve prognostic capabilities.

Degradation modeling and RUL prediction involves physics-based, data-driven, and hybrid models, all gathered in Fig. 4. When the underlying physics of the system is known, physics-based methods offer interpretable results and accuracy without the need for large amounts of data. However, it is difficult to obtain robust physics-based models of complex systems, such as gearboxes and generators. Data-driven

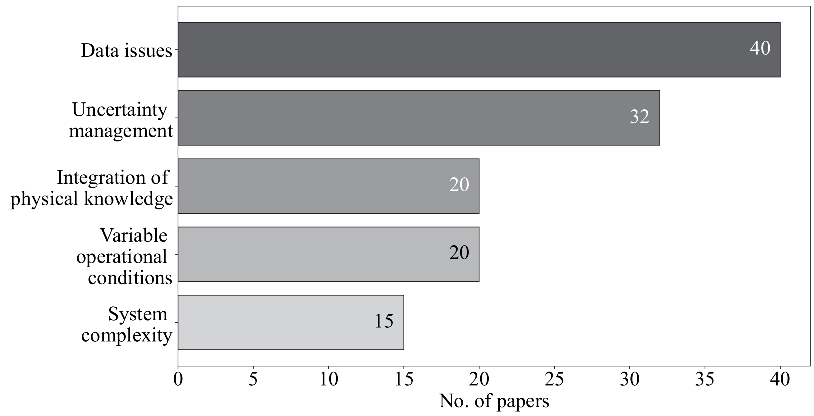


Fig. 5. Number of papers addressing challenges in WT component RUL prediction, from 2018 to June 21, 2024.

methods, while requiring less physical knowledge, can effectively quantify prognosis uncertainty and process high-dimensional data, though they may lack interpretability and generalize poorly with limited data. This is a common occurrence in many WT datasets, where data availability is often sparse or low quality. Hybrid methods combine the advantages of various approaches but may face challenges in model selection and characterization of uncertainty. Fig. 3(a) illustrates that data-driven methodologies are predominant in predicting the RUL of WT components. Further development of hybrid models is expected in the following years, as their interest is evident from the maintenance optimization point of view, as concluded in Section 4.

While the primary objective of these models is to improve the accuracy and robustness of their predecessors, significant barriers remain. The challenges outlined in the review on RUL estimation methods for WT components were derived from an extensive analysis of existing literature and research findings. Through a systematic review process, recurring obstacles were identified and categorized into five distinct groups: inherent uncertainty management, integration of physical knowledge, consideration of varying operational conditions, data issues, and complex system dynamics. This classification was based on the underlying nature of the challenges and their impact on prognostic accuracy and reliability, as shown in Fig. 5. The distribution of papers that address these issues is also depicted in this figure. It is important to clarify that a systematic count was conducted for the studies included in the reviewed literature. Additionally, since a single paper may address multiple challenges, so the count reflects the total number of instances where each issue is discussed, not a unique paper count. These are discussed below.

1. **Uncertainty management.** Uncertainty management is a challenge in all prognostic applications, but WTs face unique sources of uncertainty due to their dynamic operating environment and the complexity of their systems. Unlike many other systems that operate under more controlled conditions, WT are subject to highly varying environmental conditions and loads, which introduces significant uncertainty into degradation models. Accurate RUL prediction depends on the effective estimation and management of this uncertainty arising from various sources, which can be broadly categorized into data-related, model-related, and environmental factors [28,104]. First, the inherent uncertainty in CM data, such as sensor noise or measurement errors, can weaken the reliability of RUL predictions. These issues arise from environmental noise caused by other components in the system, as well as from inaccuracies in sensor readings themselves [104]. In the context of WTs, the operating conditions are further complicated by fluctuations in environmental factors like wind speed and temperature, which can introduce additional noise into the CM data and affect the accuracy of the RUL labels [28]. Such fluctuations create a disconnection between the sensor data and the true degradation states, as the latter are difficult to observe directly, particularly during operation. This uncertainty in labeling is a significant challenge, as RUL labels are typically based on the time interval between the current time and predicted end-of-life, a method that may not always reflect the true condition of the equipment [104]. Model-related uncertainties also contribute to the overall uncertainty in RUL prediction. Modeling errors, including fitting inaccuracies, mean that even the best-trained models are not perfect solutions but rather approximations of reality. The final trained model will always contain some degree of error, further aggravating the challenge of RUL prediction [104]. Additionally, DL models, while flexible, introduce another layer of uncertainty due to their highly adaptable nature. The optimization of network parameters in DL models can lead to overconfidence in point predictions, making the resulting RUL estimates less trustworthy [104]. Moreover, the inherent variability of working conditions in WT complicates the relationship between CM data and RUL labels. For instance, wind gusts or extreme weather events can have a substantial impact on component degradation, making it harder to reliably map sensor readings to actual RUL values [31]. These working condition fluctuations, combined with sensor noise, create a layer of uncertainty that can significantly affect model accuracy. Therefore, developing specialized uncertainty quantification methods that can better handle seasonal effects, wind gusts, or extreme weather events on WT components would have a considerable impact on the effective and trusty decision-making based on RUL prediction [31,104]. One of the most promising approaches to overcome these difficulties is the application of Bayesian methods, which use probabilistic inference to account for the inherent uncertainty in degradation models [178]. Bayesian neural networks have gained traction in this regard, as they can quantify uncertainty through the use of probability distributions rather than single-point estimates. This allows for a more comprehensive view of the possible degradation state of WT components, as demonstrated in studies [104,178]. Additionally, Monte Carlo methods, specifically Monte Carlo dropout, have been employed to approximate inference by introducing dropout layers during the testing phase of neural network models [179]. This method simulates parameter uncertainty by randomly omitting neurons, thereby generating multiple possible outcomes and providing a probabilistic estimation of RUL [178]. Monte Carlo dropout has been successfully applied to various fields, including infrastructure damage prediction, and shows potential for WT RUL prediction, as explored by Deng et al. [180].

Beyond Bayesian methods and Monte Carlo techniques, other fields have successfully implemented a range of advanced approaches to uncertainty management, which could be further investigated for use in WT prognostics. These include variational inference, Markov chain Monte Carlo, Bayesian neural network weights, Bayesian active learning, Bayes by backpropagation, variational autoencoders, Laplacian approximations, and reinforcement learning, which could be studied as a further application of uncertainty management in RUL prediction [179]. These models could allow WT systems to dynamically adjust operational parameters in response to real-time RUL predictions under uncertainty, optimizing both turbine performance and component lifespan.

2. **Integration of physical knowledge**. A significant challenge in the development of reliable RUL prediction models for WT components is the effective integration of physical knowledge into data-driven approaches. WTs are highly complex systems, and purely data-driven models often struggle to capture the intricate physical interactions between different components, especially in the context where failure data records are limited. While data-driven methods, such as ML models, have shown promise in processing large datasets and handling high-dimensional data, they often lack interpretability and the ability to generalize across different operational conditions. To address these shortcomings, incorporating physical knowledge of turbine systems, such as underlying physics and degradation laws, is essential [31].
Physics-based models, which rely on mathematical formulations to describe the behavior of components under stress or fatigue, offer valuable insights into the degradation processes of critical WT components, as seen in Section 1.2.1. When the degradation process is well understood, these models can produce accurate RUL predictions with a high level of interpretability [104]. However, the challenge with WTs lies in developing robust physics-based models for highly complex systems like gearboxes and generators, where numerous factors such as variable load conditions, material properties, and component interactions make it difficult to capture the full degradation process accurately. To overcome this limitation, as well as the limited samples of faulty data, recent research has focused on hybrid models that combine physics-based approaches with data-driven methods. These models take advantage of the strengths of both paradigms: the interpretability and physical understanding of the physics-based models, and the flexibility and adaptability of the data-driven models. A great example of these techniques is PiNNs, which have been proposed to enhance data-driven models by incorporating physical laws directly into the architecture of NNs, ensure that the predictions align with known physical constraints [178]. In the context of WT components, this could involve integrating fatigue damage models into the network structure, allowing the model to learn both from historical data and the physical behavior of the system. Such an approach could significantly enhance the accuracy and reliability of RUL predictions by grounding the model's learning process in real-world physics [31]. For instance, by embedding physics-based cumulative damage laws within a data-driven model, researchers have been able to predict the degradation of bearings more accurately. These models allow for the real-time adjustment of RUL estimates based on operational data, while also providing physically interpretable insights into how and why a component is degrading [35]. An example of this is found in [181], where the authors propose a methodology for predicting the damage level in WT bearings by modeling grease degradation using PiNNs. Their approach highlights the advantage of integrating hybrid physics-data models, which allow the prediction of grease degradation increments based on current measurements and various operational parameters. This method enables more interpretable predictions compared to black-box models, as it incorporates physical insights that are not directly observable from data alone.
However, there remain several challenges to fully integrating physical knowledge into RUL prediction models for WTs. One major setback is the modeling complexity. The interactions between different components in a WT, such as between the drivetrain, rotor, and electrical systems, introduce additional layers of complexity that are difficult to encapsulate within a single model. Current hybrid approaches often rely on simplifications that may not fully capture the multi-physics interactions present in real-world turbine operations [31]. Additionally, there are challenges related to data compatibility. The integration of physical models into data-driven frameworks requires compatible data sources, and it is often difficult to collect high-quality physical data from operational WTs. For example, while SCADA data is widely available, it is typically low-frequency and may not contain the detailed physical measurements needed to feed physics-based models [133]. Accelerated degradation testing and digital twin simulations offer potential solutions, as they allow for controlled experiments where both physical and operational data can be collected. This combined data could then be used to train hybrid models [35]. Finally, the model selection process in hybrid approaches remains a challenge. With multiple ways to combine physics-based and data-driven methods, selecting the right model structure, training strategy, and optimization algorithm requires careful consideration. Research is ongoing to address these challenges, with recent studies focusing on optimizing hybrid model architectures for specific WT components, such as gearboxes and bearings [104].

3. **Varying operational conditions**. WTs operate in environments with far greater variability than many other mechanical or industrial systems. WTs are typically installed in locations with high wind resources, where conditions are highly dynamic. Factors such as wind speed, direction, temperature, humidity, and, in offshore locations, saltwater exposure, all contribute to the degradation of critical components [42]. Moreover, events like startups, emergency brakes, or yaw misalignment are frequent, resulting in extreme loads [99]. The impact of these conditions on the reliability and lifetime of WT components is profound, and this dynamic, fluctuating environment introduces specific challenges that are not as present in other industrial prognostic applications. Moreover, due to time-varying environmental loads, offshore WTs typically operate in dynamic conditions, resulting in vibration signals displaying robust non-stationarity [182]. In non-stationary conditions, fault characteristic frequencies may fluctuate with speed variations, causing blurred frequency lines in spectra. Moreover, in the case of offshore WTs face unique environmental challenges, such as the harsh marine environment, which accelerates corrosion and material fatigue. The constant exposure to saltwater and the more extreme weather conditions offshore result in different degradation patterns compared to onshore wind farms. Additionally, wind speeds in offshore environments are typically higher and more consistent, which alters the mechanical loading on WTs. These factors show the need for the development of offshore-specific prognostic models that can account for the effects of marine environmental factors.
Among the structural components of the WTs, rotor blades are subjected to the most dynamically varying loads [50]. The loads due to start and stoppage of WT, sudden gusts, and variation of wind speeds between the cut-in and cut-out speed of the operation are ignored for simplicity and time conservation during simulations [50]. WT blades regularly function in severe environmental conditions, including air salinity, wind gusts, water inclusions, air pollution, atmospheric oxidation, icing, and sand particle erosion [42]. The WT planetary gearbox always works

under time-varying speed because the corresponding wind speed varies a lot with time [62]. Compared to the turbine generator systems spinning in the traditional thermal and hydropower plants, WTs generators have relatively higher failure rates due to the highly varying operating conditions and harsher environments, and therefore, require more frequent inspection and maintenance [57]. Moreover, the fault characteristic frequencies in current signals are time-varying due to nonstationary shaft rotating speeds of the WT drivetrains. [57]. Most studies make overly simplistic assumptions about the work environment, such as constant load or stable conditions, which do not reflect the reality of the WT working environment, particularly for offshore WTs where maintenance costs are high.

In the studied works, many techniques to consider variable conditions have been used. For instance, in physics-based models, Bayesian updating approaches have been used, KF [82,100,154] and PF [49,50,58,59]. Here, the physics behind the degradation process and fault related features representing the degradation status are modeled in a Bayesian framework, and KF and PF are employed to estimate the model parameter online and predict the degradation status with uncertainty quantification [62]. These methods work well for tracking component degradation over time, as they continuously update predictions based on incoming sensor data, which reflects current operational conditions, and are often used with sensor-fusion techniques to handle multiple data streams. The need for continuous online learning underscores the urgency for novel algorithms [28]. In this context, transfer learning and domain adaptation present an opportunity to adapt models to varying operating conditions, although challenges remain to ensure prediction accuracy across different equipment [125]. Therefore, the transferability assessment of different domains continues to pose a substantial challenge. In the studied works, transfer learning has been used [45, 81], paving the way for a promising line of research. Transfer learning has been proposed for this [81], constructing a RUL prediction model capable of aligning various subspaces of both the source and target domains. This adaptive alignment allows for more accurate prognostics by extracting domain-invariant features over time, even as operating conditions shift.

Future research needs to focus on integrating these environmental stressors into RUL prediction models. For instance, the development of climate-adaptive prognostic models could take into account seasonal variations and extreme weather events like storms, which significantly impact WT reliability, particularly in offshore environments. Moreover, WTs experience highly varying operational conditions due to changing wind speeds and loads. Unlike other mechanical systems, which often operate under more stable and controlled conditions, WTs must endure constant fluctuations. These fluctuations create non-linear degradation patterns, which are difficult to capture using traditional prognostic models. Future research should explore dynamic, non-linear degradation models that can adjust to changes in operational conditions in real-time. Techniques like adaptive prognostics, where the model evolves based on real-time operational data, could provide more accurate RUL predictions for components subjected to variable loads, such as blades and gearboxes.

4. **Data issues**. Data quality, availability, and compatibility pose some of the most significant challenges in predicting the RUL of WT components. These challenges mainly arise from SCADA systems as primary data sources. At present, the data are mainly collected and stored by the SCADA to trigger alarms and shut down the WTs by comparing the real power with the theoretical power when there is a certain deviation [45]. The particularities of WT data differ greatly from other industrial settings, creating unique difficulties for developing accurate and generalizable prognostic models for WTs. One of the primary issues is SCADA data frequency, which are typically collected at 10-minute intervals, offering low-frequency snapshots of turbine operations [133]. While this frequency is sufficient for monitoring general performance trends, it lacks the resolution required to capture early signs of component degradation. This negatively affects the PHM capabilities, possibly hiding short-lived events [136]. To enhance the detection of fast-evolving faults, researchers must explore data enrichment techniques. One potential approach is the use of advanced data fusion techniques, where SCADA data is combined with higher-frequency data from CMS, such as vibration or acoustic emissions [31], which have been explored in Section 2.3.

   Another critical challenge is the unbalanced nature of SCADA datasets, where the majority of the data represents healthy operation, with only a small fraction containing fault or failure information [125]. This imbalance complicates the training of AI models, leading to biased predictions favoring healthy states and poor performance when identifying faults. For instance, WT bearings have a large volume, slow operating speed, and high maintenance cost, and the degradation process can be easily affected by the external environment. Therefore, it is difficult to obtain a large amount of destructive data. To provide a more accurate analysis of the RUL, multiple factors need to be considered, including modeling for the degradation process as accurately as possible, fusing multiple sensor data, more accurate feature extraction methods, and developing fused degradation models [67]. As discussed in Section 2.1, generating synthetic data has been proposed as a solution to address this issue by creating artificial fault data to balance the datasets [28]. However, synthetic data is not always representative of real-world failure modes, particularly in WTs, where operational environments and component configurations vary widely across wind farms [38]. To overcome the obstacles of unbalanced datasets with most of the data categorized as healthy, models such as NBM [74,116] have been developed to train with purely healthy SCADA data.

   Moreover, the inconsistent and non-standardized formats of SCADA data across different WT models and manufacturers create a significant obstacle in developing scalable RUL models [135]. Variability in parameter naming conventions, data acquisition protocols, and feature availability prevents the creation of generalizable models that can work across multiple turbine systems [133]. This is particularly challenging for multi-turbine wind farms, where different turbines may produce data that is incompatible with one another, limiting the applicability of predictive maintenance strategies.

   In addition to data format issues, the scarcity of publicly available WT datasets present difficulties when developing robust RUL models. Public datasets are essential for training, testing, and validating AI models, yet many wind energy operators are reluctant to share their operational data due to confidentiality concerns. This scarcity of accessible data has led to calls for the establishment of benchmark datasets for WTs, which would enable the fair evaluation of prognostic algorithms across different research groups [26,119]. In the absence of such datasets, researchers have turned to federated learning, an approach that allows AI models to be trained across multiple decentralized datasets without sharing raw data [183]. This privacy-preserving approach could revolutionize RUL modeling by enabling collaboration across organizations while maintaining data security [184].

   In conclusion, addressing the data challenges unique to WT RUL prediction requires a multifaceted approach that combines advanced machine learning techniques with physical modeling, data enrichment, and standardization efforts. The wind energy industry must push towards improved data collection practices, increased data-sharing cooperation, and the development of hybrid models that integrate physics-based insights with data-driven methods. These steps will enable more reliable, generalizable, and interpretable RUL predictions, ultimately contributing to the broader goal of optimizing WT maintenance and reducing downtime.

5. **System complexity**. WTs are complex systems, with multiple subsystems interacting in ways that influence the degradation processes of individual components. Unlike other industrial systems, where components may operate more independently, the performance of one component in a WT can directly affect others. For example, wear and tear on the blades can alter the load on the gearbox, which in turn can accelerate wear in the bearings and generator. This systemic interaction is unique to WTs and poses significant challenges for RUL prediction.
   Most existing RUL models focus on individual components in isolation, without accounting for the interdependencies between components. This approach overlooks the fact that faults in one component can propagate through the system, leading to secondary failures in other parts of the turbine. To address this, future research should explore system-level prognostic models that capture the interactions between multiple components. These models would consider how the degradation of one component affects the entire turbine system, enabling more comprehensive and accurate RUL predictions.
   Moreover, the complexity of WT systems is further exacerbated by the fact that multiple faults can occur simultaneously or sequentially. For instance, a minor fault in the pitch control system could increase the stress on the blades, leading to accelerated degradation. Traditional prognostic models are often not equipped to handle multi-fault scenarios, and future research should focus on developing multi-fault detection and prognostic frameworks. These frameworks could use fault propagation models to predict how an initial fault in one subsystem might trigger failures in other subsystems.
   Additionally, holistic degradation models that account for the entire turbine as an integrated system are needed. These models could be developed using digital twin technology, where a virtual replica of the turbine is continuously updated with real-time data, allowing for the simulation of degradation processes across the entire system. Such an approach would provide a more accurate and system-wide view of the turbine's health, helping operators to better prioritize maintenance actions.

As research progresses, it is essential to address these challenges systematically, paying particular attention to the critical components within WT components. The implementation of PHM will facilitate the development of a robust predictive maintenance plan, which will contribute to the reduction of LCOE associated with O&M costs, thus aligning with the objectives of achieving green transition goals.

CRediT authorship contribution statement

Jokin Cuesta: Writing – original draft, Visualization, Validation, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Urko Leturiondo:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization. **Yolanda Vidal:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization. **Francesc Pozo:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Funding acquisition, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used AI-assisted technologies only to improve the readability and language of the work. After using these technologies, the authors reviewed and edited as needed and take full responsibility for the content of the publication.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Jokin Cuesta reports financial support was provided by Spanish agencia estatal de investigación - Ministerio de Ciencia e Innovación. Urko Leturiondo reports financial support was provided by Spanish agencia estatal de investigación - Ministerio de Ciencia e Innovación. Yolanda Vidal reports financial support was provided by Spanish agencia estatal de investigación - Ministerio de Ciencia e Innovación. Francesc Pozo reports financial support was provided by Spanish agencia estatal de investigación - Ministerio de Ciencia e Innovación. Yolanda Vidal reports financial support was provided by Generalitat de Catalunya. Francesc Pozo reports financial support was provided by Generalitat de Catalunya. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was partially funded by the Spanish Agencia Estatal de Investigación (AEI) — Ministerio de Ciencia e Innovación, the Fondo Europeo de Desarrollo Regional (FEDER) through the research projects PID2021-122132OB-C21, PID 2021-122132OB-C22, and TED2021-129512B-I00; and by the Generalitat de Catalunya, Spain through the research project 2021-SGR-01044.

Data availability

No data was used for the research described in the article.

References

[1] International Energy Agency (IEA). Renewables 2023: Analysis and forecasts to 2028. 2024, International Energy Agency.
[2] International Renewable Energy Agency (IREA). Renewable power generation costs 2022. 2023, International Renewable Energy Agency.
[3] Polverino L, Abbate R, Manco P, Perfetto D, Caputo F, Macchiaroli R, et al. Machine learning for prognostics and health management of industrial mechanical systems and equipment: A systematic literature review. Int J Eng Bus Manag 2023;15.
[4] International Organization for Standardization (ISO). 13381-1: Condition monitoring and diagnostics of machines-prognostics-part 1. 2015.
[5] IEEE standard framework for prognostics and health management of electronic systems. 2017, p. 1–31, IEEE Std 1856- 2017.
[6] International Organization for Standardization (ISO). Condition monitoring and diagnostics of machines — General guidelines. Tech. rep. 17359:2018, International Organization for Standardization; 2018.
[7] International Organization for Standardization (ISO). Condition monitoring and diagnostics of machines — Vibration condition monitoring. Tech. rep. 13373-9:2017, International Organization for Standardization; 2017.
[8] Verein Deutscher Ingenieure (VDI). 3834: Measurement and evaluation of the mechanical vibration of wind turbines and their components - Wind turbines with gearbox. Tech. rep., Verein Deutscher Ingenieure; 2015.
[9] Cuesta J, Leturiondo U, Vidal Y, Pozo F. Challenges on prognostics and health management for wind turbine components. J Physics: Conf Ser 2024;2745:012003.
[10] Nejad AR, Keller J, Guo Y, Sheng S, Polinder H, Watson S, et al. Wind turbine drivetrains: State-of-the-art technologies and future development trends. Wind Energy Sci 2022;7:387–411.
[11] Jiang Z, Hu W, Dong W, Gao Z, Ren Z. Structural reliability analysis of wind turbines: A review. Energies 2017;10:2099.

[12] Gao Z, Liu X. An overview on fault diagnosis, prognosis and resilient control for wind turbine systems. Processes 2021;9:1–19.
[13] Dao C, Kazemtabrizi B, Crabtree C. Wind turbine reliability data review and impacts on levelised cost of energy. Wind Energy 2019;22:1848–71.
[14] Tazi N, Châtelet E, Bouzidi Y. Using a hybrid cost-FMEA analysis for wind turbine reliability analysis. Energies 2017;10.
[15] Santelo TN, de Oliveira CMR, Maciel CD, Roberto JRBJ. Wind turbine failures review and trends. J Control Autom Electr Syst 2022;33:505–21.
[16] Carroll J, McDonald A, McMillan D. Failure rate, repair time and unscheduled O&M cost analysis of offshore wind turbines. Wind Energy 2016;19:1107–19.
[17] Gayo JB. Final publishable summary of results of project RELIAWIND. 2011.
[18] Walgern J, Fischer K, Hentschel P, Kolios A. Reliability of electrical and hydraulic pitch systems in wind turbines based on field-data analysis. Energy Rep 2023;9:3273–81.
[19] Liang J, Zhang K, Al-Durra A, Muyeen SM, Zhou D. A state-of-the-art review on wind power converter fault diagnosis. Energy Rep 2022;8:5341–69.
[20] Owolabi OI, Madushele N, Adedeji PA, Olatunji OO. FEM and ANN approaches to wind turbine gearbox monitoring and diagnosis: a mini review. J Reliab Intell Environ 2023;9:399–419.
[21] Olabi AG, Wilberforce T, Elsaid K, Sayed ET, Salameh T, Abdelkareem MA, et al. A review on failure modes of wind turbine components. Energies 2021;14.
[22] Lydia M, Kumar GEP. Condition monitoring in wind turbines: a review. In: Non-destructive testing and condition monitoring techniques in wind energy. Elsevier; 2023, p. 229–47.
[23] Catelani M, Ciani L, Galar D, Patrizi G. Optimizing maintenance policies for a yaw system using reliability-centered maintenance and data-driven condition monitoring. IEEE Trans Instrum Meas 2020;69:6241–9.
[24] Li H, Zhang Z, Li T, Si X. A review on physics-informed data-driven remaining useful life prediction: Challenges and opportunities. Mech Syst Signal Process 2024;209:111120.
[25] Sim J, Kim S, Park HJ, Choi JH. A tutorial for feature engineering in the prognostics and health management of gears and bearings. Appl Sci (Switzerland) 2020;10:5639.
[26] Saidi L, Benbouzid M. Prognostics and health management of renewable energy systems: State of the art review, challenges, and trends. Electron (Switzerland) 2021;10:2732.
[27] Zhou K, Tang J. A wavelet neural network informed by time-domain signal preprocessing for bearing remaining useful life prediction. Appl Math Model 2023;122:220–41.
[28] Elattar HM, El-Brawany MA, Elminir HK, Ibrahim DA, Ramadan E. Artificial intelligence-based data-driven prognostics in industry: A survey. Comput & Ind Eng 2023;109605.
[29] Rezamand M, Kordestani M, Carriveau R, Ting DS, Saif M. An integrated feature-based failure prognosis method for wind turbine bearings. IEEE/ASME Trans Mechatronics 2020;25:1468–78.
[30] Ahmadzadeh F, Lundberg J. Remaining useful life estimation: review. Int J Syst Assur Eng Manag 2014;5:461–74.
[31] Lei J, Zhang W, Jiang Z, Gao Z. A review: Prediction method for the remaining useful life of the mechanical system. J Fail Anal Prev 2022;22:2119–37.
[32] Gordan M, Sabbagh-Yazdi SR, Ismail Z, Ghaedi K, Carroll P, McCrum D, et al. State-of-the-art review on advancements of data mining in structural health monitoring. Measurement: J Int Meas Confed 2022;193:110939.
[33] Zuo T, Zhang K, Zheng Q, Li X, Li Z, Ding G, et al. A hybrid attention-based multi-wavelet coefficient fusion method in RUL prognosis of rolling bearings. Reliab Eng Syst Saf 2023;237:109337.
[34] Li B, Tang B, Deng L, Zhao M. Self-attention ConvLSTM and its application in RUL prediction of rolling bearings. IEEE Trans Instrum Meas 2021;70:1–11.
[35] Xu Y, Kohtz S, Boakye J, Gardoni P, Wang P. Physics-informed machine learning for reliability and systems safety applications: State of the art and challenges. Reliab Eng Syst Saf 2023;230:108900.
[36] Zhang W, Vatn J, Rasheed A. A review of failure prognostics for predictive maintenance of offshore wind turbines. J Physics: Conf Ser 2022;2362:012043.
[37] Liu D, Shi J, Liao Z, Guo H. Prognostics and health management for electromechanical system: A review. J Adv Manuf Sci Technol 2022;2. This article presents a review on PHM methods for gears, presenting four PHM methods (Physical-Based Method, Knowledge model-based method, data driven model-based methods and hybrid methods..
[38] Soleimani M, Campean F, Neagu D. Diagnostics and prognostics for complex systems: A review of methods and challenges. Qual Reliab Eng Int 2021;37:3746–78.
[39] Zonta T, da Costa CA, da Rosa Righi R, de Lima MJ, da Trindade ES, Li GP. Predictive maintenance in the industry 4.0: A systematic literature review. Comput Ind Eng 2020;150.
[40] Xue B, Xu H, Huang X, Zhu K, Xu Z, Pei H. Similarity-based prediction method for machinery remaining useful life: A review. Int J Adv Manuf Technol 2022;121:1501–31.
[41] Thoppil NM, Vasu V, Rao CS. Deep learning algorithms for machinery health prognostics using time-series data: A review. J Vib Eng Technol 2021;9:1123–45.
[42] Rezamand M, Kordestani M, Carriveau R, Ting DS, Orchard ME, Saif M. Critical wind turbine components prognostics: A comprehensive review. IEEE Trans Instrum Meas 2020;69:9306–28.
[43] Fox H, Pillai A, Friedrich D, Collu M, Dawood T, Johanning L. A review of predictive and prescriptive offshore wind farm operation and maintenance. Energies 2022;15:504.
[44] Do MH, Söffker D. State-of-the-art in integrated prognostics and health management control for utility-scale wind turbines. Renew Sustain Energy Rev 2021;145.
[45] Yue G, Ping G, Lanxin L. An end-to-end model based on CNN-LSTM for industrial fault diagnosis and prognosis. In: 2018 international conference on network infrastructure and digital content. 2018, p. 274–8.
[46] Tayade A, Patil S, Phalle V, Kazi F, Powar S. Remaining useful life (RUL) prediction of bearing by using regression model and Principal Component Analysis (PCA) technique. In: Vibroengineering procedia, vol. 23, EXTRICA; 2019, p. 30–6.
[47] Jellali A, Maatallah H, Ouni K. Predicting remaining useful life of wind turbine bearing using linear regression. In: 2022 5th international conference on advanced systems and emergent technologies. 2022, p. 357–62.
[48] Rajaoarisoa L, Randrianandraina R, Sayed-Mouchaweh M. Predictive maintenance model-based on multi-stage neural network systems for wind turbines. In: International conference on artificial intelligence, computer, data sciences, and applications. Institute of Electrical and Electronics Engineers Inc.; 2024, p. 1–7.
[49] Valeti B, Pakzad SN. Remaining useful life estimation of wind turbine blades under variable wind speed conditions using particle filters. In: Annual conference of the PHM society, vol. 10, 2018, p. 1–10.
[50] Valeti B, Pakzad SN. Estimation of remaining useful life of a fatigue damaged wind turbine blade with particle filters. In: Pakzad S, editor. In: Dynamics of civil structures, vol. 2, Cham: Springer International Publishing; 2019, p. 319–28.
[51] Saathoff M, Rosemeier M, Kleinselbeck T, Rathmann B. Effect of individual blade pitch angle misalignment on the remaining useful life of wind turbines. Wind Energy Sci 2021;6(5):1079–87.
[52] Mehlan FC, Nejad AR, Gao Z. Digital twin based virtual sensor for online fatigue damage monitoring in offshore wind turbine drivetrains. J Offshore Mech Arct Eng 2022;144(6):060901.
[53] Dameshghi A, Refan MH. Combination of condition monitoring and prognosis systems based on current measurement and PSO-LS-SVM method for wind turbine DFIGs with rotor electrical asymmetry. Energy Syst 2021;12(1):203–32.
[54] Kamarzarrin M, Refan MH, Amiri P, Dameshghi A. A new intelligent fault diagnosis and prognosis method for wind turbine doubly-fed induction generator. Wind Eng 2022;46(1):308–40.
[55] Farhat MH, Chaari F, Chiementin X, Bolaers F, Haddar M. Dynamic remaining useful life estimation for a shaft bearings system. In: Smart monitoring of rotating machinery for industry 4.0. Cham: Springer International Publishing; 2022, p. 169–78.
[56] Benmoussa S, Djeziri MA, Sanchez R. Support vector machine classification of current data for fault diagnosis and similarity-based approach for failure prognosis in wind turbine systems. In: Artificial intelligence techniques for a scalable energy transition: advanced methods, digital technologies, decision support tools, and applications. Cham: Springer International Publishing; 2020, p. 157–82.
[57] Qiao W, Qu L. Prognostic condition monitoring for wind turbine drivetrains via generator current analysis. Chin J Electr Eng 2018;4(3):80–9.
[58] Jaramillo F, Gutiérrez JM, Orchard M, Guarini M, Astroza R. A Bayesian approach for fatigue damage diagnosis and prognosis of wind turbine blades. Mech Syst Signal Process 2022;174:109067.
[59] Lee IY, Roh HD, Park YB. Prognostics and health management of composite structures under multiple impacts through electromechanical behavior and a particle filter. Mater Des 2022;223:111143.
[60] Vettori S, Di Lorenzo E, Peeters B, Luczak MM, Chatzi E. An adaptive-noise augmented Kalman filter approach for input-state estimation in structural dynamics. Mech Syst Signal Process 2023;184:109654.
[61] Praveen HM, Shah D, Pandey KD, Vamsi I, Sabareesh GR. PCA based health indicator for remaining useful life prediction of wind turbine gearbox. In: Vibroengineering procedia, vol. 29, EXTRICA; 2019, p. 31–6.
[62] Wang J, Gao RX, Yuan Z, Fan Z, Zhang L. A joint particle filter and expectation maximization approach to machine condition prognosis. J Intell Manuf 2019;30(2):605–21.
[63] Wang J, Liang Y, Zheng Y, Gao RX, Zhang F. An integrated fault diagnosis and prognosis approach for predictive maintenance of wind turbine bearing with limited samples. Renew Energy 2020;145:642–50.
[64] Pagitsch M, Jacobs G, Bosse D. Remaining useful life determination for wind turbines. J Phys: Conf Ser 2020;1452(1):012052.
[65] Pan Y, Hong R, Chen J, Wu W. A hybrid DBN-SOM-PF-based prognostic approach of remaining useful life for wind turbine gearbox. Renew Energy 2020;152:138–54.
[66] Zheng H, Deng W, Song W, Cheng W, Cattani P, Villecco F. Remaining useful life prediction of a planetary gearbox based on meta representation learning and adaptive fractional generalized Pareto motion. Fractal Fract 2024;8(1):14.

[67] Liu M, Dong Z, Shi H. Multi-sensor information fusion and multi-model fusion-based remaining useful life prediction of fan slewing bearings with the nonlinear Wiener process. Sustain (Switzerland) 2023;15(15):12010.
[68] Cheng F, Qu L, Qiao W, Hao L. Enhanced particle filtering for bearing remaining useful life prediction of wind turbine drivetrain gearboxes. IEEE Trans Ind Electron 2019;66(6):4738–48.
[69] Elasha F, Shanbr S, Li X, Mba D. Prognosis of a wind turbine gearbox bearing using supervised machine learning. Sensors 2019;19(14):3092.
[70] Lázaro R, Yürüen NY, Melero JJ. Determining remaining lifetime of wind turbine gearbox using a health status indicator signal. J Phys: Conf Ser 2020;1618(2):022037.
[71] Elforjani M. Diagnosis and prognosis of real world wind turbine gears. Renew Energy 2020;147(1):1676–93.
[72] Guo Y, Sheng S, Phillips C, Keller J, Veers P, Williams L. A methodology for reliability assessment and prognosis of bearing axial cracking in wind turbine gearboxes. Renew Sustain Energy Rev 2020;127:109888.
[73] Li Z, Zhang X, Kari T, Hu W. Health assessment and remaining useful life prediction of wind turbine high-speed shaft bearings. Energies 2021;14(15):4612.
[74] Verma A, Zappalá D, Sheng S, Watson SJ. Wind turbine gearbox fault prognosis using high-frequency SCADA data. J Phys: Conf Ser 2022;2265(3):032067.
[75] He K, Su Z, Tian X, Yu H, Luo M. RUL prediction of wind turbine gearbox bearings based on self-calibration temporal convolutional network. IEEE Trans Instrum Meas 2022;71:1–12.
[76] Shen Y, Tang B, Li B, Tan Q, Wu Y. Remaining useful life prediction of rolling bearing based on multi-head attention embedded bi-LSTM network. Measurement: J Int Meas Confed 2022;202:111803.
[77] Xiang S, Qin Y, Liu F, Gryllias K. Automatic multi-differential deep learning and its application to machine remaining useful life prediction. Reliab Eng Syst Saf 2022;223:108531.
[78] Xiang S, Qin Y, Luo J, Pu H. Spatiotemporally multidifferential processing deep neural network and its application to equipment remaining useful life prediction. IEEE Trans Ind Inform 2022;18(10):7230–9.
[79] Qin Y, Yang J, Zhou J, Pu H, Mao Y. A new supervised multi-head self-attention autoencoder for health indicator construction and similarity-based machinery RUL prediction. Adv Eng Informatics 2023;56:101973.
[80] Xiang S, Qin Y, Luo J, Wu F, Gryllias K. A concise self-adapting deep learning network for machine remaining useful life prediction. Mech Syst Signal Process 2023;191:110187.
[81] Xiang S, Li P, Luo J, Qin Y. Micro transfer learning mechanism for cross-domain equipment RUL prediction. IEEE Trans Autom Sci Eng 2024;1–11.
[82] Saidi L, Ali JB, Benbouzid M, Bechhofer E. An integrated wind turbine failures prognostic approach implementing Kalman smoother with confidence bounds. Appl Acoust 2018;138:199–208.
[83] Kramti SE, Ben Ali J, Saidi L, Sayadi M, Bechhoefer E. Direct wind turbine drivetrain prognosis approach using elman neural network. In: 2018 5th international conference on control, decision and information technologies. 2018, p. 859–64.
[84] Cao L, Qian Z, Zareipour H, Wood D, Mollasalehi E, Tian S, et al. Prediction of remaining useful life of wind turbine bearings under non-stationary operating conditions. Energies 2018;11(12):3318.
[85] Rezamand M, Carriveau R, Ting DS, Davison M, Davis JJ. Aggregate reliability analysis of wind turbine generators. IET Renew Power Gener 2019;13(11):1902–10.
[86] Kramti SE, Saidi L, Ali JB, Sayadi M, Bechhoeferer E. Particle filter based approach for wind turbine high-speed shaft bearing health prognosis. In: 2019 international conference on signal, control and communication. 2019, p. 46–50.
[87] Cao L, Qian Z, Pei Y. Remaining useful life prediction of wind turbine generator bearing based on EMD with an indicator. In: Proceedings of prognostics and system health management conference, PHM-chongqing. IEEE; 2019, p. 375–9.
[88] Wang P, Long Z, Wang G. A hybrid prognostics approach for estimating remaining useful life of wind turbine bearings. Energy Rep 2020;6(9):173–82.
[89] Kramti SE, Ali JB, Saidi L, Sayadi M, Bouchouicha M, Bechhoefer E. A neural network approach for improved bearing prognostics of wind turbine generators. EPJ Appl Phys 2021;93(2):20901.
[90] Hayder AN, Saidi L. Applications of artificial neural networks with input and output degradation data for renewable energy systems fault prognosis. In: 2021 12th international renewable energy congress. 2021, p. 1–6.
[91] Rezamand M, Kordestani M, Orchard ME, Carriveau R, Ting DS, Saif M. Improved remaining useful life estimation of wind turbine drivetrain bearings under varying operating conditions. IEEE Trans Ind Inform 2021;17:1742–52.
[92] Merainani B, Laddada S, Bechhoefer E, Chikh MAA, Benazzouz D. An integrated methodology for estimating the remaining useful life of high-speed wind turbine shaft bearings with limited samples. Renew Energy 2022;182:1141–51.
[93] Kordestani M, Rezamand M, Orchard ME, Carriveau R, Ting DS, Rueda L, et al. New condition-based monitoring and fusion approaches with a bounded uncertainty for bearing lifetime prediction. IEEE Sensors J 2022;22(9):9078–86.
[94] Cao L, Zhang H, Meng Z, Wang X. A parallel GRU with dual-stage attention mechanism model integrating uncertainty quantification for probabilistic RUL prediction of wind turbine bearings. Reliab Eng Syst Saf 2023;235:109197.
[95] Pandit R, Xie W. Data-driven models for predicting remaining useful life of high-speed shaft bearings in wind turbines using vibration signal analysis and sparrow search algorithm. Energy Sci Eng 2023;11(1):4557–69.
[96] Peng Y, Bi R, Wang Y. Regime-switching model with adaptive adjustments for degradation prognosis. IEEE Trans Instrum Meas 2023;72(6):1–11.
[97] Teng W, Han C, Hu Y, Cheng X, Song L, Liu Y. A robust model-based approach for bearing remaining useful life prognosis in wind turbines. IEEE Access 2020;8:47133–43.
[98] Moynihan B, Liberatore S, Moaveni B, Khan U. Fatigue life analysis of main shaft bearings in wind turbines using strain measurements collected on blades. In: Conference proceedings of the society for experimental mechanics series. Springer; 2021, p. 185–92.
[99] Yucesan YA, Viana FA. A hybrid physics-informed neural network for main bearing fatigue prognosis under grease quality variation. Mech Syst Signal Process 2022;171:108875.
[100] Bousebsi M, Medoued A, Saidi L. Data-driven and physics-based approaches for wind turbine high-speed shaft bearing prognostics. In: Proceedings of 2023 international conference on control, automation and diagnosis. IEEE; 2023, p. 1–6.
[101] Li X, Teng W, Peng D, Ma T, Wu X, Liu Y. Feature fusion model based health indicator construction and self-constraint state-space estimator for remaining useful life prediction of bearings in wind turbines. Reliab Eng Syst Saf 2023;233:109124.
[102] Wang L, Cao H, Xu H, Liu H. A gated graph convolutional network with multi-sensor signals for remaining useful life prediction. Knowl-Based Syst 2022;252:109340.
[103] Binsbergen DV, Soares MN, Pedersen E, Nejad AR. A physics-, SCADA-based remaining useful life calculation approach for wind turbine drivetrains. J Phys: Conf Ser 2022;2265(3):032079.
[104] Wang L, Cao H, Ye Z, Xu H. Bayesian large-kernel attention network for bearing remaining useful life prediction and uncertainty quantification. Reliab Eng Syst Saf 2023;238:109421.
[105] Wang L, Cao H, Ye Z, Xu H, Yan J. DVGTformer: A dual-view graph transformer to fuse multi-sensor signals for remaining useful life prediction. Mech Syst Signal Process 2024;207:110935.
[106] Rezamand M, Kordestani M, Orchard M, Carriveau R, Ting D, Saif M. Condition monitoring and failure prognostic of wind turbine blades. In: 2021 IEEE international conference on systems, man, and cybernetics. 2021, p. 1711–8.
[107] Muto K, Namura N, Ukei Y, Takeda N. Model-based load estimation for wind turbine blade with Kalman filter. In: 2019 8th international conference on renewable energy research and applications. 2019, p. 191–9.
[108] Carroll J, Koukoura S, McDonald A, Charalambous A, Weiss S, McArthur S. Wind turbine gearbox failure and remaining useful life prediction using machine learning techniques. Wind Energy 2019;22(3):360–75.
[109] Desai A, Guo Y, Sheng S, Phillips C, Williams L. Prognosis of wind turbine gearbox bearing failures using SCADA and modeled data. In: Annual conference of the PHM society, vol. 12, 2020, p. 10.
[110] Bermúdez K, Ortiz-Holguin E, Tutivén C, Vidal Y, Benalcázar-Parra C. Wind turbine main bearing failure prediction using a hybrid neural network. J Phys: Conf Ser 2022;2265(3):032090.
[111] Zhou Y, Zhou J, Cui Q, Wen J, Fei X. Digital twin-driven online intelligent assessment of wind turbine gearbox. Wind Energy 2024;1–19.
[112] Hu Y, Li H, Shi P, Chai Z, Wang K, Xie X, et al. A prediction method for the real-time remaining useful life of wind turbine bearings based on the Wiener process. Renew Energy 2018;127:452–60.
[113] Song M, Youliang S, Kai J, Cheng L, Tao W. Remaining life prediction of wind turbine bearing based on Wiener process. IOP Conf Ser: Mater Sci Eng 2020;788(1):012089.
[114] Peter R, Zappalá D, Schamboeck V, Watson SJ. Wind turbine generator prognostics using field SCADA data. J Phys: Conf Ser 2022;2265(3):032111.
[115] Mehlan FC, Keller J, Nejad AR. Virtual sensing of wind turbine hub loads and drivetrain fatigue damage. Forsch Im Ingenieurwesen/Eng Res 2023;87(1):207–18.
[116] Encalada-Dávila Á, Puruncajas B, Tutivén C, Vidal Y. Wind turbine main bearing fault prognosis based solely on SCADA data. Sensors 2021;21(6):2228.
[117] Gomes GdP, Lopes SMd, Araujo DCP, Flauzino RA, Pinto MM, Alves MEG. Wind turbine remaining useful life prediction using small dataset and machine learning techniques. J Control Autom Electr Syst 2024;35:337–45.
[118] Li L, Jian Q. Remaining useful life prediction of wind turbine main-bearing based on LSTM optimized network. IEEE Sensors J 2024;21143–56.
[119] Maldonado-Correa J, Martín-Martínez S, Artigao E, Gómez-Lázaro E. Using SCADA data for wind turbine condition monitoring: A systematic literature review. Energies 2020;13(12).
[120] National Aeronautics and Space Administration (NASA). Commercial modular aero-propulsion system simulation (C-MAPSS) datasets. 2008, https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/#turbofan.
[121] Lee J, Qiu H, Yu G, Lin J, Services RT. Bearing Data Set. 2007, NASA Ames Research Center, Moffett Field, CA, http://ti.arc.nasa.gov/project/prognostic-data-repository.

[122] Lan X, Liang D, Liu W, Wang Y, Han Y, Yu Z. Remaining useful life prediction of wind turbine generator bearings based on nonlinear Wiener process. In: Proceedings - 2023 3rd power system and green energy conference. Institute of Electrical and Electronics Engineers Inc.; 2023, p. 1124–9.

[123] Le T-T, Lee S-J, Dinh M-C, Park M. Design of an improved remaining useful life prediction model based on vibration signals of wind turbine rotating components. Energies 2024;17(1):19.

[124] Nectoux P, Gouriveau R, Medjaher K, Ramasso E, Chebel-Morello B, Martin T, et al. PRONOSTIA: An experimental platform for bearings accelerated degradation tests. In: IEEE international conference on prognostics and health management. IEEE; 2012, p. 1–8.

[125] Ramezani SB, Cummins L, Killen B, Carley R, Amirlatifi A, Rahimi S, et al. Scalability, explainability and performance of data-driven algorithms in predicting the remaining useful life: A comprehensive review. IEEE Access 2023;11(4):41741–69.

[126] Yucesan YA, Viana FAC. Wind turbine main bearing fatigue life estimation with physics-informed neural networks. Annu Conf PHM Soc 2019;11(1).

[127] Yucesan Y. Wind turbine main bearing fatigue life prediction with PINN. 2019, Harvard Dataverse, https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/ENNXLZ.

[128] Energias de Portugal (EDP). Dataset wind turbines. 2023, URL https://www.edp.com/en/innovation/open-data/data.

[129] Stetco A, Dinmohammadi F, Zhao X, Robu V, Flynn D, Barnes M, et al. Machine learning methods for wind turbine condition monitoring: A review. Renew Energy 2019;133:620–35.

[130] Márquez FPG, Gonzalo AP. A comprehensive review of artificial intelligence and wind energy. Arch Comput Methods Eng 2022;29:2935–58.

[131] Maldonado-Correa J, Martín-Martínez S, Artigao E, Gómez-Lázaro E. Using SCADA data for wind turbine condition monitoring: A systematic literature review. Energies 2020;13:3132.

[132] Wang KS, Sharma VS, Zhang ZY. SCADA data based condition monitoring of wind turbines. Adv Manuf 2014;2:61–9.

[133] Leahy K, Gallagher C, O'Donovan P, O'Sullivan DT. Issues with data quality for wind turbine condition monitoring and reliability analyses. Energies 2019;12.

[134] Remigius WD, Natarajan A. Identification of wind turbine main-shaft torsional loads from high-frequency SCADA (Supervisory Control and Data Acquisition) measurements using an inverse-problem approach. Wind Energy Sci 2021;6:1401–12.

[135] Hu Y, Miao X, Si Y, Pan E, Zio E. Prognostics and health management: A review from the perspectives of design, development and decision. Reliab Eng Syst Saf 2022;217:108063.

[136] Vidal Y, Pozo F, Tutivén C. Wind turbine multi-fault detection and classification based on SCADA data. Energies 2018;11.

[137] Chen J, Huang R, Chen Z, Mao W, Li W. Transfer learning algorithms for bearing remaining useful life prediction: A comprehensive review from an industrial application perspective. Mech Syst Signal Process 2023;193.

[138] Sheng S. Prognostics and health management of wind turbines—Current status and future opportunities. In: Probabilistic prognostics and health management of energy systems. Cham: Springer International Publishing; 2017, p. 33–47.

[139] Liu X, Du J, Ye Z-S. A condition monitoring and fault isolation system for wind turbine based on SCADA data. IEEE Trans Ind Inform 2022;18(2):986–95.

[140] Sandoval D, Leturiondo U, Pozo F, Vidal Y, Salgado O. Revision of the condition based monitoring techniques of the pitch bearings from wind turbines. Dyna 2019;94(6):636–42.

[141] Teng W, Zhang X, Liu Y, Kusiak A, Ma Z. Prognosis of the remaining useful life of bearings in a wind turbine gearbox. Energies 2017;10:32.

[142] Cheng F, Qu L, Qiao W. Fault prognosis and remaining useful life prediction of wind turbine gearboxes using current signal analysis. IEEE Trans Sustain Energy 2018;9(1):157–67.

[143] Wang Z, Gao P, Chu X. Remaining useful life prediction of wind turbine gearbox bearings with limited samples based on prior knowledge and PI-LSTM. Sustain (Switzerland) 2022;14(19):12094.

[144] Hyvärinen A, Oja E. Independent component analysis: Algorithms and applications. Neural Netw 2000;13:411–30.

[145] Taşcı B, Omar A, Ayvaz S. Remaining useful lifetime prediction for predictive maintenance in manufacturing. Comput Ind Eng 2023;184:109566.

[146] Roberts C, Cava DG, Avendaño-Valencia LD. Addressing practicalities in multivariate nonlinear regression for mitigating environmental and operational variations. Struct Heal Monit 2023;22:1237–55.

[147] Huang L, Pan X, Liu Y, Gong L. An unsupervised machine learning approach for monitoring data fusion and health indicator construction. Sensors 2023;23:7239.

[148] Du X, Jia W, Yu P, Shi Y, Gong B. RUL prediction based on GAM–CNN for rotating machinery. J Braz Soc Mech Sci Eng 2023;45(3):142.

[149] Boutrous K, Puig V, Nejjari F. A set-based prognostics approach for wind turbine blade health monitoring. IFAC- Pap 2022;55(6):402–7.

[150] Liu Z, Zhang L. A review of failure modes, condition monitoring and fault diagnosis methods for large-scale wind turbine bearings. Measurement: J Int Meas Confed 2020;149:107002.

[151] Moroney PD, Verma AS. Durability and damage tolerance analysis approaches for wind turbine blade trailing edge life prediction: A technical review. Energies 2023;16(24):7934.

[152] International Organization for Standardization (ISO). Rolling bearings — Dynamic load ratings and rating life. Tech. rep. 281:2007, International Organization for Standardization; 2007.

[153] Li G, Wei J, He J, Yang H, Meng F. Implicit Kalman filtering method for remaining useful life prediction of rolling bearing with adaptive detection of degradation stage transition point. Reliab Eng Syst Saf 2023;235:109269.

[154] Shiri H, Zimroz P, Wyłomańska A, Zimroz R. Estimation of machinery's remaining useful life in the presence of non-Gaussian noise by using a robust extended Kalman filter. Measurement 2024;114882.

[155] Le T-T, Lee S-J, Dinh M-C, Park M. Design of an improved remaining useful life prediction model based on vibration signals of wind turbine rotating components. Energies 2024;17:19.

[156] Wang H, Ding X, Huang Z, Zhen H, Yu W, Huang W. Multi-stage exponential model based on subspace clustering distribution for bearing remaining useful life prediction. Meas Sci Technol 2024;35(7):076129.

[157] Peng C, Chen Q, Zhou X, Wang S, Tang Z. Wind turbine blades icing failure prognosis based on balanced data and improved entropy. Int J Sen Netw 2020;34(2):126–35.

[158] Guan Q, Wei X. The statistical data-driven remaining useful life prediction—A review on the Wiener process-based method. In: 2023 prognostics and health management conference. Institute of Electrical and Electronics Engineers (IEEE); 2023, p. 64–8.

[159] Vapnik VN, Golowich SE, Smola A. Support vector method for function approximation, regression estimation and signal processing. In: Neural information processing systems. 1996, p. 1–7.

[160] Vieira JLd, Farias FC, Ochoa AAV, de Menezes FD, da Costa ACA, da Costa JAP, et al. Remaining useful life estimation framework for the main bearing of wind turbines operating in real time. Energies 2024;17:1430.

[161] Rashid NA, Abdul Aziz I, Hasan MHB. Machine failure prediction technique using recurrent neural network long short-term memory-particle swarm optimization algorithm. In: Silhavy R, editor. Artificial intelligence methods in intelligent algorithms. Cham: Springer International Publishing; 2019, p. 243–52.

[162] Zhang X, Zhong C, Zhang J, Wang T, Ng WW. Robust recurrent neural networks for time series forecasting. Neurocomputing 2023;526:143–57.

[163] Rezaeianjouybari B, Shang Y. Deep learning for prognostics and health management: State of the art, challenges, and opportunities. Measurement: J Int Meas Confed 2020;163:107929.

[164] Mammadov E, Farrokhabadi M, Canizares CA. AI-enabled predictive maintenance of wind generators. In: Proceedings of 2021 IEEE PES innovative smart grid technologies Europe: smart grids: toward a carbon-free future. Institute of Electrical and Electronics Engineers Inc.; 2021, p. 1–5.

[165] Youness G, Aalah A. An explainable artificial intelligence approach for remaining useful life prediction. Aerospace 2023;10:474.

[166] Khan T, Ahmad K, Khan J, Khan I, Ahmad N. An explainable regression framework for predicting remaining useful life of machines. In: 2022 27th international conference on automation and computing: smart systems and manufacturing. Institute of Electrical and Electronics Engineers Inc.; 2022, p. 1–6.

[167] Pacheco-Blazquez R, Garcia-Espinosa J, Di Capua D, Sanchez AP. A digital twin for assessing the remaining useful life of offshore wind turbine structures. J Mar Sci Eng 2024;12:573.

[168] Djeziri MA, Benmoussa S, Sanchez R. Hybrid method for remaining useful life prediction in wind turbine systems. Renew Energy 2018;116:173–87.

[169] Santiago RAdF, Barbosa NB, Mergulhão HG, Carvalho TFd, Santos AAB, Medrado RC, et al. Data-driven models applied to predictive and prescriptive maintenance of wind turbine: A systematic review of approaches based on failure detection, diagnosis, and prognosis. Energies 2024;17(5).

[170] Bougacha O, Varnier C, Zerhouni N. A review of post-prognostics decision-making in prognostics and health management. Int J Progn Heal Manag 2020;11:31.

[171] Mosallam A, Medjaher K, Zerhouni N. Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction. J Intell Manuf 2016;27:1037–48.

[172] Ren Z, Verma AS, Li Y, Teuwen JJ, Jiang Z. Offshore wind turbine operations and maintenance: A state-of-the-art review. Renew Sustain Energy Rev 2021;144:110886.

[173] Syan CS, Ramsoobag G. Maintenance applications of multi-criteria optimization: A review. Reliab Eng Syst Saf 2019;190:106520.

[174] Li M, Jiang X, Carroll J, Negenborn RR. A multi-objective maintenance strategy optimization framework for offshore wind farms considering uncertainty. Appl Energy 2022;321:119284.

[175] Lei X, Sandborn P, Sandborn PA. PHM-based wind turbine maintenance optimization using real options. Int J Progn Heal Manag 2016;2153–648.

[176] Lei X, Sandborn PA. Maintenance scheduling based on remaining useful life predictions for wind farms managed using power purchase agreements. Renew Energy 2018;116:188–98.

[177] Mazidi P, Tjernberg LB, Bobi MAS. Wind turbine prognostics and maintenance management based on a hybrid approach of neural networks and a proportional hazards model. Proc Inst Mech Eng Part O: J Risk Reliab 2017;231(2):121–9.
[178] Zhu R, Chen Y, Peng W, Ye ZS. Bayesian deep-learning for RUL prediction: An active learning perspective. Reliab Eng Syst Saf 2022;228:108758.
[179] Abdar M, Pourpanah F, Hussain S, Rezazadegan D, Liu L, Ghavamzadeh M, et al. A review of uncertainty quantification in deep learning: Techniques, applications and challenges. Inf Fusion 2021;76:243–97.
[180] Deng W, Nguyen KT, Medjaher K, Gogu C, Morio J. Physics-informed machine learning in prognostics and health management: State of the art and challenges. Appl Math Model 2023;124:325–52.
[181] Yucesan Y, Viana F. A physics-informed neural network for wind turbine main bearing fatigue. Int J Progn Heal Manag 2020;11:17.
[182] Gao Z, Odgaard P. Real-time monitoring, fault prediction and health management for offshore wind turbine systems. Renew Energy 2023;218.
[183] Bagheri B, Rezapoor M, Lee J. A unified data security framework for federated prognostics and health management in smart manufacturing. Manuf Lett 2020;24:136–9.
[184] Cheng X, Shi F, Liu Y, Liu X, Huang L. Wind turbine blade icing detection: a federated learning approach. Energy 2022;254:124441.