This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2020.3030165, IEEE Transactions on Instrumentation and Measurement

Critical Wind Turbine Components Prognostics: A Comprehensive Review

Milad Rezamand, Mojtaba Kordestani, *Senior Member, IEEE,* Rupp Carriveau, *Member, ASME,* David S-K. Ting, *Member, ASME,* Marcos E. Orchard, *Member, IEEE, PHMS Fellow,* and Mehrdad Saif, *Senior Member, IEEE.*

***Abstract*—As wind energy is becoming a significant utility source, minimizing the Operation and Maintenance (O&M) expenses has raised a crucial issue to make wind energy competitive to fossil fuels. Wind Turbines (WTs) are subject to unexpected failures due to operational and environmental conditions, aging, etc. An accurate estimation of time to failures assures reliable power production and lower maintenance costs. In recent years, a notable amount of research has been undertaken to propose prognosis techniques that can be employed to forecast the Remaining Useful Life (RUL) of wind farm assets. This paper provides a recent literature review on modeling developments for the remaining useful life prediction of critical WT components including physics-based, Artificial Intelligence (AI)-based, stochastic-based, and hybrid prognostics. In particular, the pros and cons of the prognosis models are investigated to assist researchers in selecting proper models for certain applications of WT RUL forecast. Our comprehensive review has revealed that hybrid methods are now the leading and most accurate tools for WT failure predictions over individual hybrid components. Strong candidates for future research include the consideration of variable operating environments, component interaction, physics-based prognostics, and Bayesian application in the development of WT prognosis methods.**

***Index Terms*—Wind Turbines, prognosis, bearings, blade, gearbox, generator.**

I. Introduction

WIND energy is playing an increasingly pivotal role in global energy systems. According to the Global Wind Energy Council (GWEC)'s report [1], the global installed wind power capacity reached 651 GW in 2019. It is excepted that over 355 GW of new capacity will be added between 2020 to 2024. The wind is also assuming a nascent and growing role in the expanding ancillary services market associated with evolving grids.

The authors would like to acknowledge that this work is part of the YR21 Investment Decision Support Program supported by progressive industrial partners, the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Ontario Centres of Excellence (OCE). Moreover, The work of M. E. Orchard was supported by FONDECYT Grant 1170044 and the Advanced Center for Electrical and Electronic Engineering, AC3E, Basal Project FB0008, ANID. (Corresponding author: Dr. Mojtaba Kordestani.)

Milad Rezamand, Rupp Carriveau and David S-K. Ting are with the Turbulence and Energy Laboratory, University of Windsor,401 Sunset Avenue, Windsor, O. N., N9B 3P4, Canada. E-mail: rezaman@ uwindsor.ca, rupp@ uwindsor.ca and dting@ uwindsor.ca

Mojtaba Kordestani and Mehrdad Saif are with the Department of Electrical and Computer Engineering, University of Windsor,401 Sunset Avenue, Windsor, O. N., N9B 3P4, Canada. E-mail: kordest@ uwindsor.ca, msaif@ uwindsor.ca

Marcos E. Orchard is with the Department of Electrical Engineering, University of Chile, Av. Tupper 2007, Suite 307, Santiago, Chile. E-mail: morchard@ing.uchile.cl

A. WT reliability

WTs are complex machines, assembled combinations of numerous technologies, functioning in challenging environmental and operating conditions including unpredictable loads due to gust wind, humidity, dustiness, corrosion, fatigue, wear, a wide range of temperatures, and air pressures. These severe environmental and operating conditions may result in increasing component defects and machine malfunctions. As an integrated system, some of the components are more critical than others. So, it is essential to identify components with the highest failure rate and downtime.

There have been some fundamental studies in recent decades on the reliability of wind farm components, as reviewed below. Based on 350 WTs operation over five years throughout Europe, Carroll et al. [2] revealed that the highest failure rates are related to generators, gearboxes, and blades. Shafiee et al. [3] showed that for onshore machines, the most frequent failures are related to the towers, gearboxes, and rotor blades, respectively; whereas in offshore settings, the gearboxes, rotor blades, generators, and towers have the highest failure rates. Hahn et al. [4] indicated that generators, gearboxes, and rotor blades have the most downtime according to 1467 WT (below 1 MW) data in the period from 1989 until the end of 2004. Tavner et al. [5] demonstrated that blades, generators, and gearboxes exhibit the highest downtime per failure based on a survey on 15000 WT-years. In Stenberg and Holttinen [6], a dataset from 72 operating wind turbines of Finland revealed that the gearboxes, hydraulic systems, brakes, and generators had the most maximum downtime over a period from 1996 to 2008. Reviews of these reliability summary studies showed that the gearbox and generator failure rates are distinctly high. The downtime for these failures is among the highest of all wind turbine components.

Numerous studies have been accomplished to obtain the distribution of failures by subassemblies in WTs [7–9]. They illustrated that bearings are by far the most liable subassemblies that are subject to failure, leading to increased WT downtime and maintenance costs. Therefore, it is essential to take the WT Prognostics and Health Management (PHM) into consideration for WT prognostics task. The WT PHM is thoroughly investigated in the following.

B. WT PHM

The unexpected failure of WT components and subcomponents can cause substantial economic losses, so, it may be prudent to employ WT PHM, which seeks to diminish the costly inspections and time-based maintenance through accurate monitoring, early fault detection, and impending failures forecast, i.e.,RUL estimation. This can provide wind farm owners with enhanced productivity, reduced unnecessary planned maintenance, extended operating intervals between maintenance, decreased downtime, reduced number, and severity of failures, especially unanticipated ones [10].

For this purpose, an appropriate monitoring system is essential to perform. Typical practice for WT health state assessment are Performance Monitoring (PM) and Condition Monitoring (CM). PM employs the Supervisory Control and Data Acquisition (SCADA) system to correlate multiple sets of variables such as wind speed and power to train models for normal operating states and utilize these models to detect abnormal abnormal behaviors and outliers. CM comprises inspecting WT components to identify changes in operation that can be indicative of a progressing fault by analysis of particular aspects and measurements of the operation. CM systems are capable of capturing high-frequency dynamics usually not achievable through the SCADA system. CM relies on analyses of particular aspects and measurements of the operation, e.g., vibration analysis. Table I summarizes the strengths and weaknesses of PM and CM systems [11–15].

Non-contact monitoring, an example of CM systems, contributes toward WT PHM using various techniques such as ultrasound and vibration analyses. Non-contact monitoring can lead to more efficient measurement, and, consequently, more accurate WT RUL prediction compared to the contact monitoring, since contact measurement can change the vibration characteristics of the object to be measured. Furthermore, non-contact monitoring requires a short time for measurement preparation and an easy operation to change the measurement locations and directions and prevents any interference with the operation of the wind turbine [16–22].

After data is provided at regular time intervals by sensors and measurement systems, Signal Processing is performed through signal transformations and features extraction, reduction, selection, etc. The next step of WT PHM is to determine the component state-of-health by detecting and localizing a fault based on pre-set operational limits. Consequently, the diagnostics module defines if the component condition has degraded. In the prognostic module, the future condition and the time to failure of the faulty components are predicted. As shown in Figure 1, RUL prediction made at t_0 utilizes a mathematical model of the failure fitted to online data. It is desirable to apply a robust technique to project an $RUL_{optimal}$ for the lifetime of a component based on a predefined failure criterion. Besides, uncertainties associated with prognostic techniques such as the deterioration process, and the unknown future operation condition must be considered in the RUL prediction. Thus, the confidence interval estimation is an essential component of this module, i.e., upper confidence bound (RUL_{max}) and lower confidence bound (RUL_{min}) estimation. Prognosis techniques consist of two main categories of physics-based and data-driven models [23–26]:

- Physics-based models: deal with the RUL prediction of critical components by employing mathematical or physical models of the degradation process, utilizing system-specific mechanistic knowledge, CM data, and damage evaluation formulas.
- Data-driven models: aim at transforming the data provided by CM into relevant models of the degradation's behavior rather than a physical understanding of the failure processes.

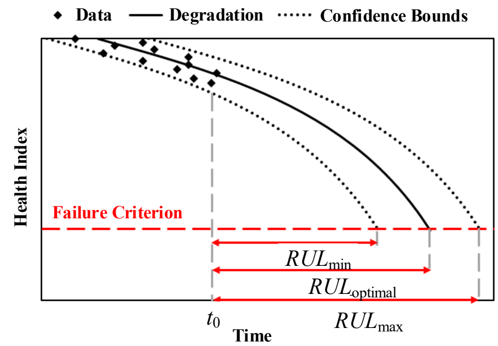


Fig. 1. A component health degradation curve

Finally, a proper maintenance approach can be scheduled based on the anticipated RUL. Different types of WT maintenance are corrective, preventive, and condition-based maintenance [27], [28]:

- Corrective maintenance is conducted after a failure occurs to restore the component to a condition to be able to perform its required purpose.
- Preventive maintenance is carried out preventively on a scheduled time interval without any prior knowledge of the time to failures.
- Condition-Based Maintenance (CBM) is performed before an occurrence of a failure, based on the condition of the component. As shown in Figure 2, the essential factors of practical CBM are data acquisition, data processing, and decision making, i.e., the recommendation of maintenance processes through diagnostics and prognostics [29].

C. Contributions

Although there have been seminal efforts that provide state-of-the-art reviews on WT CM, fault diagnostics, and maintenance scheduling [28–40], as well as a few reviews on individual WT component prognostic [41–43], the authors are not aware of any comprehensive reviews on WT prognostics. This paper presents a prognostic literature review of the most critical components, including gearboxes, generators, blades, and the most critical subcomponent, i.e., bearings. The main contributions of this work are summarized as follows:

- This article corresponds to the first comprehensive survey on WT prognostics that gathers, analyses, and separately classifies common failures for various parts of wind turbines. In each component, failures can be initiated by a combination of priory factors. These factors are further analysed to determine the roots of any failures.
- A review of the most recent literature on WT prognostics in component and subcomponent levels is presented. The

TABLE I
PERFORMANCE MONITORING VERSUS CONDITION MONITORING

Monitoring systems	Strengths	Weaknesses
PM system (SCADA)	Readily accessible, Capable of identifying abnormal behaviors and outliers	Incapable of fault isolation and identification, reporting many false alarms prompted by varying loads endured by WTs
CM system	Capable of fault isolation and identification, Facilitating condition-based maintenance and PHM, capable of capturing high-frequency dynamics usually not achievable through the SCADA system	Require resources on data analysis and result assessment, expensive due to instrumentation

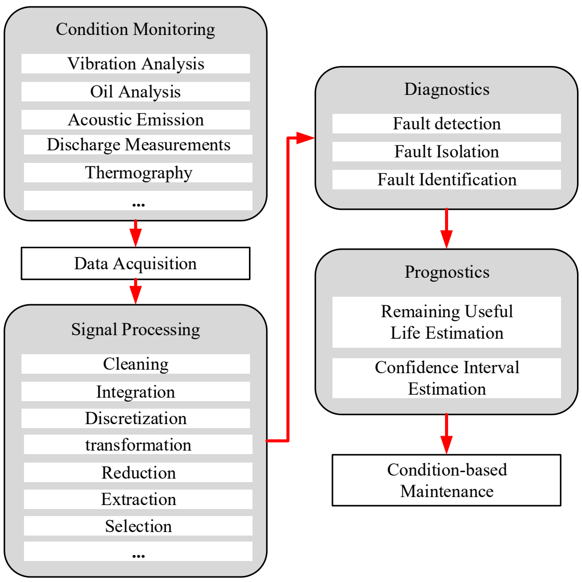


Fig. 2. WT PHM architecture

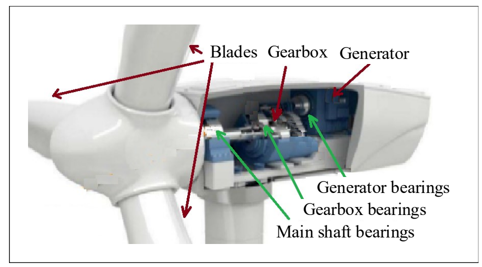


Fig. 3. The main components of wind turbines and some corresponding bearings

Fig. 4. Gearbox planetary tooth failure [47]

main approaches to WT prognostics associated with their pros and cons are provided. An effort to deepen the insight into weaknesses of the existing techniques is made, and possible future research directions for research are also discussed.

The work is organized as follows: Section II describes various failures in WT components. Section III illustrates the prognosis definition, and provides a thorough review of the WT assemblies' prognosis. In Section IV, a specific problem on prognostics of critical WT components is illustrated to give a brief idea of the practical applications related to WT prognostic field. Section V concludes the paper with an emphasis on future research challenges.

II. CRITICAL WT COMPONENTS FAILURES

Wind turbines are assembled integrations of several components, subjected to various failure modes due to harsh environmental and operating conditions. Figure 3 displays the main components of wind turbines and some corresponding bearings. In this research work, we review failure modes of gearboxes, generators, blades, and bearings.

A. Gearbox failures

Bearing damage, gear damage, leaking oil, broken shaft, and insufficient oil cooling are the typical defects observed in WT gearboxes. The WT gearboxes failures are due to a combination of several factors such as crack initiation and propagation, surface fatigue, surface wear, structural fatigue, and loss of lubrication [44]. It is worth noting that bearing failures are detected as the majority of the gearbox failures due to white structure flaking, scuffing, and micro-pitting. [45], [46]. Figure 4 exhibits tooth crack, detected on a WT gearbox planetary in a wind farm located in Southwestern Ontario in 2013 [47].

B. Generator failures

Common WT generator failures are as follows:

- Fluting (a typical bearing fault in electric drivetrain) [48], caused by the prolonged passage of relatively small electric current, usually due to current leakage.
- Stator Windings (SW) insulation breakdown [49]: caused by mechanical stress, contamination, electrical, and thermal stress. Figure 5 shows a visual inspection of an WT generator SW insulation fault, occurred in a wind farm located in Southwestern Ontario in 2015 [50].

Fig. 5. Generator SW insulation fault [50]

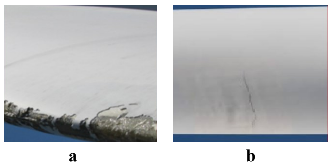


Fig. 6. Blade erosion (a) and crack (b) [56]

- Rotor electrical asymmetry [51]: caused by the rising resistance or open-circuit of the brush-gear circuits.

These failures may lead to prolonged torque pulsation, unbalances in the air-gap flux and phase currents, reduced average torque, excessive heating in the winding, increased losses, and diminished inefficiency [52].

C. Blade failures

WT blades regularly function in severe environmental conditions, including air salinity, wind gusts, water inclusions, air pollution, atmospheric oxidation, icing, and sand particle erosion [53]. These conditions can excite several damage types, including adhesive joint failure between skins along leading and/or trailing edges, adhesive debonding, fibre and laminate failure, delamination, blade deformation, and growth of cracks in the gel-coat [54]. Besides, there can be catastrophic blade-tower collisions that lead to the collapse of the whole turbine [55]. Figure 6 displays blade erosion and crack, investigated, analyzed, and detected by Rezamand et al. [56].

D. Bearing failures

Wind turbine bearings can be subject to defects induced by corrosive, high-speed, and high temperature operating conditions. The performance degradation of a bearing is a continuous irreversible process. Once the bearing is placed in its housing, there are certain expectations of long-term healthy service life. Eventually, minor early faults can arise that grow gradually at the initiation. Then, a major bearing failure in wind turbines can cause catastrophic downtime due to time-consuming reactive maintenance practices. Such lost production directly affects the wind farm bottom line [57]. Bearing defects can be categorized into two groups, including distributed and single-point defects. The distributed defect is characterized by degradation over large areas of the surface, which become rough, irregular, or deformed. A typical example is the overall surface roughness caused by contamination or lack of lubricant. This type of fault is difficult to identify by distinct frequencies. On the contrary, a single-point defect is localized and can be defined by specific frequencies that typically appear in the machine vibration. A typical example

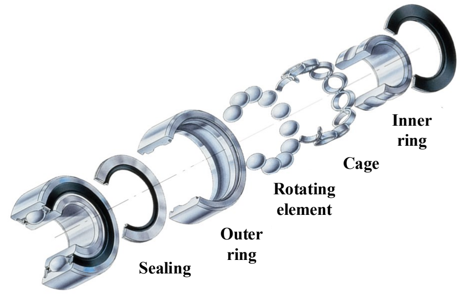


Fig. 7. Rolling bearing subcomponents

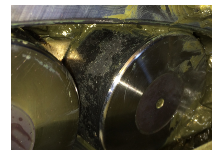


Fig. 8. Rolling element surface wear [60]

of a localized defect is a pit or spall. According to which component of the bearing affected, as shown in Figure 7, the single point defects can be categorized to inner raceway, outer raceway, rolling element, and cage defects [58], [59]. Fig. 8 indicates severe rolling element surface wear of a WT main-shaft bearing, caused by outer raceway failure, investigated by Rezamand et al. [60].

III. Review of WT prognosis

The main goal of prognosis is to evaluate how long a faulty component can work under reliable operating conditions, still achieving desired performance metrics [61–63]. Data-Driven Methods that transform historical data into relevant models of the degradation's behavior are widely used in WT prognosis due to the existence of historical wind farm data. However, a complete set of failure data based on all operating conditions is required to develop accurate data-driven prognosis methods [64].

WT prognosis methods are categorized into physics-based, AI-based, stochastic-based, and hybrid prognostic techniques, and a comprehensive review of the most recent literature on critical WT components and subcomponents is discussed for each group in the following.

A. Physics-based Prognostic methods

Physics-based prognostic methods attempt to construct mathematical models to describe failure modes physics, such as spall progression and crack growth. To do so, firstly, system and subsystem configuration, material specification, and after-treatment processes are defined. Then, potential failure modes and their causes in terms of the failure physics are identified at the individual component level, associated with operating and environmental conditions under which the failure is prone to occur. [65], [66]. Robust physics-based techniques comprise Paris' law crack growth modeling [67], Paris' law crack growth modeling with Finite Element Analysis (FEA) [68],

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2020.3030165, IEEE Transactions on Instrumentation and Measurement

Forman law crack growth modeling [69], fatigue spall initiation and progression model [70], contact analysis for bearing prognostics [71], and stiffness-based damage rule model [72].

Gray and Watson [73], firstly, identified different failure modes, their causes, and the damaging operating conditions of wind turbine gearboxes. Afterward, they proposed a prognostic approach based on a mathematical model, as shown in Algorithm 1, for WT gearbox damage calculation for a specific failure mode, bearing high cycle fatigue due to edge loading. The experimental study on six WTs experiencing severe gearbox failure among 160 WTs, recorded as heavy debris in lubricating oil, revealed the efficacy of the proposed method. Breteler et al. [74] proposed a generic physics-based diagnostics and prognostics for WT gearbox for a specific failure mode, helical gear tooth fault due to bending fatigue during misalignment. This study employs an FEA model to estimate bending stress based on the geat geometry and an averaged misalignment value, obtained through laser measurements of three-year WT operation. Next, the gear tooth damage is projected by employing the Palmgren-Miner rule (Eq. (5)) and degradation trend analysis. Although the proposed method was not able to detect misalignment continuously, it's robustness was shown, predicting a 20-year lead time to gear fault due to bending stresses.

Algorithm 1: Damage model for bearing high cycle fatigue due to edge loading [73]

1) Approximating the reaction of the contact forces at the gear teeth (F_{reac}) by assuming that mechanical and electrical efficiencies are constant:

$$F_{reac} = k_1 \frac{P}{w} \tag{1}$$

where k_1 is force scaling constant, P denotes generator electrical power, and w is shaft rotational speed.

2) Estimating the effective bearing load (F):

$$F = \frac{F_{reac}}{1-\delta} \tag{2}$$

where δ represents relative structure deformation.

3) Substituting $\delta = \frac{F_{reac}}{k_2}$ in Eq. 2:

$$F = \frac{F_{reac}}{1-\frac{F_{reac}}{k_2}} \tag{3}$$

where k_2 is the stiffness constant.

4) Estimating the expected reduction in the bearing life based on Lundberg-Palmgren rule:

$$N^{1/\alpha} = \frac{C}{F} \tag{4}$$

where N represents the number of shaft revolutions before the failure, C is the specific dynamic load capacity, and α equals 3 for ball bearings and $\frac{10}{3}$ for roller bearings.

5) Estimating the rate at which bearing damage accumulates before the failure based on Palmgren-Miner rule:

$$D = \sum_{i=1}^{K} \frac{n_i}{N_i} \tag{5}$$

where D represents high cycle fatigue damage, n_i is the number of cycles at load i, and K equals the number of load magnitudes in the total spectrum of loads to which the component is subjected. Note that $D \approx 1$ when failure is likely to occur.

Zhu et al. [75], [76] derived physical models including viscosity and dielectric constant, both as functions of temperature and particle contamination to determine the mathematical relationship between lubrication oil deterioration and particle contamination level. Then, a Particle Filter (PF) was implemented for lubrication oil RUL prediction. The experimental lab results confirmed the capability of the kinematic viscometer and dielectric sensor in lubricant deterioration monitoring and remaining useful life prediction. The studies provided discussions on the potential application of the given approach for WT gearbox health indication and RUL prediction. Grujicic et al. [77] employed FEA for fatigue life prediction of WT gearbox with a specific failure mode, helical-gear tooth-bending high-cycle fatigue under varying operating conditions. For this aim, two regimes of fatigue crack initiation (treated as a strain-controlled short cycle process) and growth (treated as a stress-controlled process) were discussed and modeled. The results indicated that gear misalignment, under a constant transferred-torque condition, can seriously reduce the service life of the gearbox helical gear.

Florian and Sorensen [78] constructed a generic crack propagation model based on a Paris' law approach for RUL prediction of a WT blade, subjected to variable loading over 20 years. Several series of 10-min simulations were conducted to establish the load cycle distribution as a function of the environment. It employed an aero-elastic simulator, covering all operational wind bins of the turbine. Finally, the crack growth in the bond line was determined based on the load cycles applied to the blade and the crack length, appropriating a log-normal distribution. The experimental results on a 5MW turbine with a rated wind speed of 11.4 m/s were leveraged for optimized maintenance scheduling.

Bechhoefer and Schlanbusch [79], firstly, applied bearing envelope analysis to generate condition indicators, related to bearing pass frequencies, and, then, condition indicators were mapped to Health Indicators (HIs). Finally, a simplification of Paris' law and Kalman filters were employed to forecast the WT bearing RUL. The HIs have been applied as inputs to compute bearing RUL. The results showed the efficacy of the proposed method. A mathematical model, Eq. (6), was employed by Butler et al. [80] to describe the progression of a WT main bearing degradation, using SCADA data including hydraulic brake temperature, blade pitch position, main shaft rotational speed, and hydraulic brake pressure. This study applied a particle filtering technique to address mathematical model and future load uncertainties to enhance WT main bearing RUL projection. Results revealed strong evidence of failure with a 30-day lead time.

$$x_j = x_{j-1} + \alpha_1 \frac{\exp(\frac{-\alpha_2}{t_j})}{t_j^2} + \alpha_3 \exp(\alpha_4 t_j) + w_j \quad (6)$$

where x_j is the degradation state at time t_j, the α_k values denote model parameters, tuned to fit the model to explain particular behavior, and w_j is a zero-mean Gaussian noise. Teng et al. [81] proposed a physics-based approach based on an improved Unscented PF to expect WT bearings RUL. To implement this improvement, first, the particles in the PF were replaced by the mean of the particles in unscented Kalman transform. Then, past known measurements were utilized to determine the likelihood function of the current step. Afterward, a modified resampling scheme using uniform distribution was developed to overcome the particle degeneracy. For RUL estimation, the degradation evolution of rolling bearings was described as an exponential model (Eq. 7). Experimental results on three life-cycle bearings from WT high-speed shaft indicated the effectiveness of the proposed method.

$$\begin{cases} x_j = \exp(b_j \cdot j) \cdot x_{j-1} \\ z_j = x_j + \sigma_z \end{cases} \quad (7)$$

where x_j describes the bearing state at the $(j)^{th}$ step, b_j denotes the model coefficient, z_j is the measurement, which can be the health indicator of rolling bearing. σ_z is the standard deviation of measurement noise. Table II summarizes the merits and limitations of physics-based prognostics.

This section provides strong evidence indicating that physics-based prognosis techniques are capable of providing accurate predictions with fewer data in comparison with data-driven techniques if the physics of models remain consistent across the component. However, physics-based models are defect-specific and complex to develop.

B. AI-based prognostic methods

AI-based prognostic methods such as Artificial Neural Networks (ANNs), Deep Learning (DL), and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) have been widely investigated in WTs.

ANNs estimate the RUL of a component using an input-output representative pattern, known as a black-box model, derived from observational data. ANNs provide a flexible tool for learning and recognizing system failures due to their ability to learn and generalize nonlinear relationships between input data and output data [82]. Networks consist of nodes connected in a layered format. A typical neural network is comprised of a single input layer, one or more hidden layers and an output layer, each including one or more nodes. Connections between nodes in adjacent layers are weighted. An activation function is associated with each node that determines how information is transferred to the following nodes. Estimated values of each node's function are then used as inputs to any subsequent nodes [83], [84].

ANNs are capable of handling noisy and incomplete data. Once they are trained, they can help with prediction and generalization at a high rate [85], [86]. ANNs are practical and efficient at modeling complex non-linear systems. However, they require a significant amount of data for training data that should be representative of the real data range and its variability [61]. Machine Learning (ML) is the procedure of constructing a model that learns from data to discover an underlying set of patterns to recognize relationships in data. Two main categories of ML consist of supervised learning methods, which project an output variable based on labeled inputs, and unsupervised learning techniques, which draw inferences from data based on unlabeled inputs. In supervised learning, an algorithm is applied to learn the mapping function from the input variable to the output variable. Classification and regression are described under the same category of supervised machine learning [15].

- Classification is the process of determining a mapping function to separate the input variables into multiple categorical classes, i.e., discrete values. For this aim, there are several steps to be taken including data preprocessing, classes equalization, feature extraction, feature selection, classification model fitting, cross-validation.
- Regression aims at the approximation of a mapping function from the input variables to numerical or continuous output variables. The correlation between features and outputs is obtained by fitting regression models when the system is in a healthy state. Whenever a new block of data becomes available, it is compared to the predicted healthy state, and if a deviation is observed for multiple consecutive time intervals, a detection alarm is raised.

Note that if the features are fed into the trained model with labels at future times, the solution is extended from diagnostics to prognostics. DL, which is a class of ML based on ANNs with representation learning, utilizes multiple layers to extract higher-level features from data. DL structures include Deep Neural Networks (DNNs), Deep Belief Networks (DBNs), Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs) [87–89]. ANFIS method, a combination of fuzzy logic and Neural Networks (NNs), constructs a hybrid intelligent system and benefits from the potentials of both techniques, including the simplicity and strength of NNs and the reasoning of fuzzy systems. ANFIS forms a series of fuzzy if-then rules with relevant membership functions to provide the specified input-output pairs. The result contributes to a robust framework for addressing practical classification problems [90–92]. Table II summarizes the advantages and disadvantages of aforementioned prognosis techniques.

1) AI-based prognostic methods for WT critical components: Hussain and Gabbar [93] compared a Nonlinear Autoregressive model with Exogenous inputs (NARX) to ANFIS for prognostics of wind turbine gearbox health conditions. For this aim, sun-spot activity data of the RWC Belgium World Data Center for years 1749-2012 and vibration data of the National Renewable Energy Laboratory from a planetary gearbox inside a wind turbine were practiced. Test results indicated that NARX outperforms ANFIS in anticipation of the WT gearbox prognostics. Chen et al. [94] introduced a Priori Knowledge (APK)-based ANFIS approach to predict WT pitch faults RUL based on SCADA data. The automated

TABLE II
A SUMMARY OF CRITICAL WT COMPONENTS PROGNOSTIC METHODS

Prognosis technique	Advantages	Limitations
AI-based methods (ANNs/ DL/ DBNs/ RNNs/ CNNs/ ANFIS)	Efficient and practical at modeling multi-dimensional, complex, and non-linear systems	Require a large amount of data as representative of actual data range and its variability for training, computationally intensive
Aggregate reliability functions	Capable of performing at all equipment hierarchy levels, particularly when a few failure modes exist, simple and well understood by reliability engineering community	Require a statistically meaningful sample size of each failure mode for reliable RUL predictions, require statistically independent and identically distributed failures
Bayesian networks	Can easily manage imprecise, noisy, or incomplete datasets	Computational challenging in determining a prior unknown network. Thus, a Bayesian network is only as beneficial as the previous knowledge is reliable
Gaussian Process Regression	Easy coding and implementation	Require a smooth and monotonic data trend
Markov/Semi-Markov	Well organized method, capable of modeling various system designs and failure modes, capable of managing incomplete data sets, providing confidence limits as part of the RUL prediction	Considering a single monotonic and non-temporal failure degradation trend, cannot model previously unanticipated faults and/or root causes
Hidden Markov/Semi-Markov	Capable of modeling varying stages of degradation. Therefore failure pattern does not require to be monotonic	Require a significant volume of data for training, proportional to the number of hidden states
Kalman Filter	Capable of accommodating incomplete and noisy measurements, Being computationally effective	computationally intensive of variants for non-linear systems, easily divergence of some variants
Particle Filter	Well-suited to perform the task of inference (state estimation) in nonlinear dynamic systems with non-Gaussian sources of uncertainty, and has become the de facto state of the art for real-time uncertainty characterization in the implementation of failure prognostic algorithms	Require a significant sample size to prevent degeneracy problem, can be more computationally intensive than basic Kalman filters
Physics-based prognostics	Require fewer data compared to data-driven methods, providing the most accurate estimates of all modeling options if the physics of models remain consistent across the component	Complexity in developing mathematical models, require detailed and complete knowledge of system behavior, being defect-specific

APK-ANFIS was able to accurately determine the WT pitch RUL within a prognostic horizon of up 21 days with optimal threshold and Window Size of 0.3 and 6, respectively. In another study, they [95] indicated that ANFIS outperforms other artificial intelligence techniques, such as K-means clustering, fuzzy inference system (FIS), ANN, and self-organizing map in WT pitch RUL forecast.

Pan et al. [96] proposed an Extreme Learning Machine (ELM) optimized by a Fruit Fly Optimization Algorithm (FOA) for wind turbine gearbox RUL forecast. FOA-ELM predicted model was trained on extracted health indicators from vibration signals. Then, the trained FOA-ELM predicted model was validated using an accelerated life test. Experimental results indicated that FOA-ELM is less time-consuming with higher accuracy compared to Particle Swarm Optimization (PSO)-ELM, Bat Algorithm (BA)-ELM, Genetic Algorithm (GA)-ELM, and Bacterial Foraging Optimization (BFO)-ELM.

2) AI-based prognostic methods for WT Bearings: Artificial intelligent methods are widely applied for condition monitoring of rotating machinery [97]. Malhi et al. [98] preprocessed vibration signals from a defect-seeded rolling bearing using a continuous wavelet transform. The preprocessed data were employed as candidate inputs to an RNN and, then, were clustered for effective representation into similar stages of bearing defect propagation. Analysis indicated that the proposed method is more accurate than the traditional incremental training technique in predicting bearing defect progressions. An approach to predict the RUL of bearings in wind turbine gearbox was proposed by Teng et al. [99]. They took an artificial neural network to train data-driven models and to predict short-term tendencies of feature series. By combining the predicted and training features, a polynomial curve reflecting the long-term degradation process of bearings was fitted. By determining the intersection between the fitted curve and the pre-defined threshold, the RUL was deduced. The results showed that the combination of the time and frequency features leads to more accurate prognostic results than those available from the individual features.

Xie and Zhang [100] developed a prognosis scheme employing an Echo State Network (ESN) and Recurrent Multilayer Perceptron (RMLP), based on the vibration signal of rotating machinery. Both ESN and RMLP are functional forms of a recurrent neural network. The experimental tests on faulty bearings demonstrated that these prognostic methods enhance the bearing performance forecast within a relatively short time interval and even with limited data availability. Guo et al. [101] examined six related-similarity features and eight time-frequency features to create an original feature set that exhibited rich degradation signatures of bearings. Correlation metrics were then employed to choose the most appropriate fault features. Finally, the selected features were combined into RNN based HI (RNN-HI) for RUL prediction of bearings. The performance of the RNN-HI was validated through two experimental bearing data sets. The results indicated that the ability of RNN-HI to obtain better performance than a self-organization map-based method.

A study of wavelet neural network classifier bearing fault diagnosis was presented by Karim et al. [102]. In this work, the statistical features of vibration signals such as standard deviation, kurtosis, and wavelet energy were employed as input to an ANN classifier. The results showed that these parameters could be applied as an operational status indicator to distinguish between a safe operational mode and a defective one. Kramti et al. [103] developed an Elman Neural Network (ENN) architecture for RUL prediction of a High-Speed Shaft Bearing (HSSB) utilizing real data provided by the Green Power Monitoring Systems in the USA. For this aim, prognosability, and monotonicity characteristics were employed to select the best features as inputs to ENN. The proposed method indicated precise forecast ability even with noisy signals and severe environmental conditions. A two-stage approach using DNN was proposed in Xia et al. [104] to estimate the RUL of bearings. A denoising auto encoder-based DNN was employed to classify the acquired signals into different degradation states. Then, regression models based on shallow neural networks were constructed for each health state. The proposed approach obtained satisfactory prediction performance on a real bearing degradation dataset with different working conditions.

Li et al. [105] proposed an intelligent RUL prediction method based on DL. Multi-scale feature extraction was executed using the CNN method. It was shown that related network structures without multi-scale feature extraction such as Single Scale-Low and Single Scale-High, are less able to capture the degrading behavior, especially in the late stage near to failure. Despite the promising prognostic experimental results achieved by the proposed method on an accelerated aging platform PRONOSTIA of rolling bearings, it was noted that adequate labeled data was demanded to initialize and train the proposed deep neural network. Deutsch and He [89] developed a DL-based method through the combination of an DBN and a Feedforward Neural Network (FNN) algorithm for RUL forecasting of rotating equipment. The proposed DBN FNN algorithm benefits from the feature learning ability of the DBN and the prediction power of the FNN. The test result indicated the promising RUL prediction performance of the DL-based DBN FNN. An HI based Hierarchical Gated Recurrent Unit Network (HGRUN) was proposed by Li et al. [106] for rolling bearing health prognosis. The HGRUN was formed by stacking various hidden layers. An open experimental bearing data was practiced to validate the capacity of the proposed approach. The results proved that HGRUN outperforms the other techniques, including Back-Propagation (BP) neural network, Support Vector Machine (SVM), and basic DBN.

A machine condition prognosis approach based on ANFIS was proposed by Chen et al. [107] to model a fault propagation trend. The high-order particle filtering was then employed to carry out the prediction. The results of experimental data from a faulty bearing demonstrated a higher prediction accuracy compared to RNNs. Cheng et al. [108] introduced a case-based data-driven prognostic framework using the ANFIS. First, large historical data was processed to build an ANFIS model-case library. Then, the prognosis of a new machinery system was implemented by applying the suitable ANFIS model extracted from the model-case library. In simulation tests, it was shown that the prognostic framework has better accuracy compared to the traditional data-driven systems. Soualhi et al. [109] proposed a time series forecasting model, neo-fuzzy neuron to predict the degradation of bearings. The neo-fuzzy neuron (NFN) is an intelligent tool that contributes to modeling complex systems by the simplicity of its structure. The Root Mean Square (RMS) extracted from vibration signals was employed as an input of the NFN in order to determine the growth of the bearing's degradation in time. A comparative study between the NFN and ANFIS was conducted to evaluate their prediction capabilities.

This section confirms the AI's robustness in the prognostic modeling of WTs. However, a large amount of data over a wide range of operating conditions is required to train the prognostic model to achieve reasonable prediction accuracy, which, often in practice, is limited, especially for complex systems. Moreover, The review on AI-based techniques indicates the highest accuracy in RUL prediction of WT components compared to conventional prognosis methods, including NARX, HGRUN, FOA-ELM, ANFIS, NFN, ESN, and RNN. The details are shown in Table III.

C. *Stochastic-based prognostic techniques*

In this section, stochastic prognostic techniques are introduced, and various studies on stochastic prognosis of critical WT components and bearings are reviewed.

Aggregate Reliability Functions (ARFs) employ a Probability Density Function (PDF) to determine the times to failure of a population of machine component/failure modes [61]. Bayesian networks are a type of probabilistic open-chain graphical model for estimating probabilities [110]. A

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2020.3030165, IEEE Transactions on Instrumentation and Measurement

TABLE III
A SUMMARY OF THE LITERATURE REVIEW ON AI-BASED PROGNOSTIC METHODS

Reference	Proposed Architecture	Application	Comparison&Accuracy
Hussain and Gabbar [93]	NARX	Gearbox prognosis	NARX outperforms ANFIS, Mean Square Error (MSE) $= 3.63 \times 10^{-6}$ (Vibration data) and $= 2.14 \times 10^{-5}$ (Sunspot data)
Chen et al. [94], [95]	APK-ANFIS	Blade pitch prognosis	ANFIS outperforms k-means, FIS, ANN, and self-organizing map-based method, Error Rate (ER) $= 14.1\%$
Pan et al. [96]	FOA-ELM	Gearbox prognosis	FOA-ELM is less time-consuming with higher accuracy compared to PSO-ELM, BA-ELM, GA-ELM, and BFO-ELM, Root Mean Square Error (RMSE) $= 0.82$
Malhi et al. [98]	RNN	Bearing prognosis	RNN is more accurate than the traditional incremental training technique, MSE $= 0.04$
Teng et al. [99]	ANN	Bearing prognosis	Error $= 12.78\%$
Xie and Zhang [100]	ESN and RMLP	Bearing prognosis	RMSE $= 0.0136$ (ESN) and $= 0.0262$ (RMLP)
Guo et al. [101]	RNN-HI	Bearing prognosis	RNN-HI obtains better performance than a self-organization map-based method, Mean of error $= 23.24\%$
Karim et al. [102]	ANN	Bearing prognosis	MSE $= 10^{-5}$
Kramti et al. [103]	ENN	Bearing prognosis	MSE $= 0.0023$
Li et al. [105]	DNN	Bearing prognosis	Mean Absolute Error (MAE) $= 30.4\%$
Li et al. [106]	HGRUN	Bearing prognosis	HGRUN outperforms BP neural network, SVM, and basic DBN, Maximum of Absolute Error (MaxAE) $= 18.79\%$
Chen et al. [107]	ANFIS	Bearing prognosis	ANFIS outperforms RNN, RMSE $= 0.0812$
Cheng et al. [108]	ANFIS	Bearing prognosis	Average RMSE $= 0.0503$
Soualhi et al. [109]	NFN	Bearing prognosis	NFN outperforms ANFIS, RMSE $= 0.000428$

Bayesian network is comprised of nodes, which correspond to random variables that can take on distinct states. These are connected by directional arcs representing conditional dependencies between nodes [111]. A Bayesian network can be utilized to assess the likelihood of different scenarios being the root cause of an event, or in the case of time series modeling, determine probabilities associated with a particular future event. The most common Bayesian techniques used in engineering prognostics consist of Markov models, Kalman Filters (KFs), and Particle Filters (PFs).

Markov models aim at estimating probabilities of future failure by determining probabilities associated with each state and probabilities associated with transitioning from one state to another. A primary characteristic of all Markov models is that future states are only dependent on the immediately prior state. For Markov prognostic purposes, the following assumptions are considered. [61].

- Transition probabilities are independent of time (i.e., a constant failure rate).
- The waiting time in a distinct state has an exponential trend.
- The sum of all transition probabilities for leaving one state and entering different states must be equal to one.

On the other hand, Semi-Markov models assume that the time spent in a particular state can be attributed to any distribution. This implies that the sum of probabilities for each state transitioning into other different states can be less than one. Thus, they are more advantageous for predicting RUL than traditional Markov chains. Despite Markov and Semi-Markov models' explicit flexibility in modeling a number of various system designs and failure scenarios, the primary drawback is the underlying assumption of a constant failure rate, which is quite idealistic [112], [113]. This can be addressed by employing the hidden and semi-hidden Markov variants.

Hidden Markov Model (HMM) and Semi-Hidden Markov Model (SHMM) are an extension of Markov chains in which not all states are directly observable. Thus, corresponding transition probabilities are not directly assignable. An HMM is characterized by the number of model states, the number of distinct observation symbols per state, a state transition probability distribution, an observation symbol probability distribution, and an initial state distribution [61]. The stochastic model is trained with failure data to overcome the lack of transition information to and from hidden states. The main benefit of HMM is its capability in the modeling of both spatial and temporal phenomena, so time-series data can be analyzed without a physical understanding of the failure, so long as enough data is available for training. A weakness of all forms of the Markov model is that it is computationally expensive, even for the simplest models with few states. The number of calculations to evaluate how well the model fits the observation data set is proportional to the number of states squared [114].

KFs are recursive processing methods applied to determine the unknown state of a dynamic system from a set of noisy measurements based on mean squared error minimization. The KF accomplishes this goal through linear projections. These are based on the assumption that process noise and measurement noise are Gaussian, white, independent of each other, and additive [61]. PFs are alternatives to KF for determining the posterior distribution. These are not restricted by linearity or Gaussian noise assumptions. They are especially helpful with conditions where the posterior distribution is multivariate and non-standard. Whereas KFs determine the posterior PDF by extrapolating from the previous state, PFs use Sequential Importance Sampling (SIS) to predict the entire next state in every iteration of the filter [61]. Table II summarizes the advantages and disadvantages of aforementioned prognosis techniques.

1) Stochastic-based prognostic techniques for WT Critical components: Rezamand et al. [115] employed Weibull life data analysis to predict the reliability of a population of generators and ALTA life data analysis to indicate how electrical loads may affect turbine generator reliability based on truncated wind farm data records. The naive prediction interval procedure was also applied to provide an approximate interval for the remaining life of individual generators. The experimental results indicated that Non-Linear Rank Regression (NLRR) outperforms Maximum Likelihood Estimation (MLE) in parameter estimation of generators failure distribution models. This study provided efficient insight into the reliability of WT generators. Fan et al. [116] introduced a framework based on PF for the WT gearbox RUL forecast. The framework determined the posterior probability distribution, i.e., the evolution of the system model and the state vector. The experimental result on a 1.5MW WT using gearbox vibration data confirmed the efficacy of the PF method in the RUL prediction of gearboxes.

2) Stochastic-based prognostic techniques for WT Bearings: Hong and Zhou [117] proposed a robust Bayesian machine learning method called Gaussian Process Regression (GPR) for bearing degradation evaluation. From the test results, it was shown that the GPR model application in bearing prognosis could achieve higher performance compared to the Wavelet Neural Network (WNN). Since covariance function is the key factor GPR properties controls, three covariance functions, including Squared Exponential (SE), Rational Quadratic (RQ), and composite, were also compared in this study. It was illustrated that the composite covariance function outperformed SE and RQ covariance functions.

Kundu et al. [118] proposed a Weibull Accelerated Failure Time Regression (WAFTR) model for the RUL projection of WT bearing under the effect of the multiple operating conditions. The study confirmed the efficacy of the WAFTR model when including operating condition data. A fault diagnosis using an HMM method was developed for rolling bearings in Zhang and Kang [119]. Afterward, prognosis was further implemented based on a Hierarchical Hidden Markov Model (HHMM). Their research work indicated that the accuracy of the method depended on the sample size of historical data. In Chen et al. [120], a multi-sensor Hidden Semi-Markov Model (HSMM) was proposed, which is an extension of classical hidden semi-Markov models. Experimental results revealed that the prognostic method was promising to achieve more reliable performance than classical HSMMs.

Singleton et al. [121] applied an Extended Kalman Filter (EKF) for anticipating the RUL of bearings. For this purpose, an affine function that best approximates the fault degradation is utilized to learn the parameters of the EKF. Then, the learned EKF is examined to forecast the RUL of bearing faults under different operating conditions. Bearing vibration data from the ”PRONOSTIA platform”, an experimental platform for bearings accelerated degradation tests, was applied to the proposed algorithm. It showed the convergence of the algorithm in different conditions. Lim and Mba [122] introduced Switching Kalman Filter (SKF) for fault diagnosis and prognosis of a gearbox bearing. For this purpose, it was presumed that the degradation trend would grow through time, and the various deterioration processes were modeled using an KF each. The SKF would then practice various models. Then, the most probable one would be selected from the CM data through the employment of Bayesian estimation for the RUL forecast. The experimental results showed that the developed approach was a promising tool to improve maintenance decision-making.

Mohammad et al. [123] presented a statistical method to predict WT bearing states by using Bayesian inference of WT bearing temperature residuals and Gaussian processes. Evaluated on a limited set of time series, it was confirmed that the approach was capable of bearing failure prediction one month in advance. A stochastic modeling method, based on interacting multiple model technique and PF for RUL prediction of bearing, was proposed by Wang et al. [124]. For this aim, a set of particle filter modules run in parallel with generating several models for different fault modes. Due to the close representation of each model to system behavior under every fault mode, the process noise for each model was highly reduced, and, thus, the prediction accuracy was improved. Experiments were carried out on a customized bearing test rig to illustrate the effectiveness of the proposed method compared to traditional PF. Chen et al. [125] presented a generic PF-based framework with application in bearing spalling fault diagnosis and prognosis. The results suggested that the system was capable of meeting performance requirements.

Information provided in this section confirms the stochastic-based prognostic potentials in the RUL prediction of WT components due to their capability in modeling the uncertainty inherent in the prediction horizon. The review on stochastic-based techniques confirms the highest accuracy in RUL prediction of WT bearings compared with conventional prognosis methods, including GPR, HHMM, Multi-sensor HSMM, EKF, SKF, PF, and Generic PF. The details are presented in Table IV.

TABLE IV
A SUMMARY OF THE LITERATURE REVIEW ON STOCHASTIC-BASED PROGNOSTIC METHODS

Reference	Proposed Architecture	Application	Comparison&Accuracy
Rezamand et al. [115]	ARF	Generator prognosis	NLRR outperforms MLE in parameter estimation of generator failure distribution models, RMSE $= 1.15$
Hong and Zhou [117]	GPR	Bearing prognosis	GPR model application in bearing prognosis achieves higher performance compared to WNN, Relative Error (RE) $= 6.32\%$
Zhang and Kang [119]	HHMM	Bearing prognosis	Error $= 13.64\%$
Chen et al. [120]	Multi-sensor HSMM	Bearing prognosis	A multi-sensor HSMM outperforms the classical HSMMs
Singleton et al. [121]	EKF	Bearing prognosis	Mean error of OPs $= 32.8\%, 73.2\%, 44\%$
Lim and Mba [122]	SKF	Bearing prognosis	Error $= 13.3\%$
Wang et al. [124]	PF	Bearing prognosis	Error $= 3.0\%$
Chen et al. [125]	Generic PF	Bearing prognosis	low MSE $= 10^{-2}$

D. Hybrid prognostic techniques

Hybrid prognosis methods are constructed using a combination of various prognostic approaches [126].

1) Hybrid prognostic techniques for WT critical components: Djeziri et al. [127] presented a hybrid method for investigation of prognosis and RUL prediction of WTs subjected to multiple faults. First, the geolocation principal was employed to project the WTs RUL. Then, Euclidean distances between the normal operation and faulty operation clusters were calculated, and, finally, the RUL is forecast as the ratio of the Euclidean position and the moving speed of the degradation. The experimental results on a 750 KW WT showed the effectiveness of the proposed method in the RUL prediction of critical WT components. Zhao et al. [128] introduced a new health index, Anomaly Operation Index (AOI), to estimate wind turbine performance deterioration, and to predict wind turbine generator RUL. In this regard, Zhao et al. employed a Density-Based Spatial Clustering of Applications with Noise to recognize abnormal data and normal data from unlabeled historical SCADA data, an SVM to categorize anomaly AOI and normal AOI in runtime, and an ARIMA to analyze real-time AOIs for wind turbine generator RUL forecast. The experimental results illustrated that the presented scheme could provide the time to failure of wind turbine generators for maintenance scheduling.

Cheng et al. [129] proposed a combination of ANFIS and PF to anticipate RUL of WT gearboxes utilizing current signals. The proposed approach employed an ANFIS to learn the state transition function of the extracted fault feature and a PF to predict the gearbox RUL based on the trained state transition function. Results illustrated that the ANFIS outperforms the RNN in learning the state transition function of the fault feature in the PF algorithm. Ding et al. [130] expected WT gearbox fatigue crack propagation and remaining life by explicitly examining varying external load. The proposed approach integrated the physical gear model utilizing Finite Element Stress Analysis (FESA) in modeling and accessible health state data. Finally, RUL prediction was enhanced by updating the distribution of the uncertain material parameter modeled in the crack degradation process via Bayesian inference. The case studies demonstrated the efficacy of the introduced varying load approach and its benefits compared to the constant load approximation method.

2) Hybrid prognostic techniques for WT Bearings: Rezamand et al. [60] proposed an integrated prognosis method based on signal processing and an adaptive Bayesian algorithm to predict the RUL of WT bearings, as indicated in Algorithm 2. Here, to improve the accuracy of RUL estimation, OWA operator, which combined the RULs obtained from various features, was employed. Two experimental case studies confirmed that the proposed real-time fusion prognosis approach achieved higher Average Relaive Accuracy (ARA) in RUL prediction compared to the Choquet integral fusion approach and the Bayesian algorithm obtained from single-feature driven methods.

In Caesarendra et al. [131], a combination of Vector Machine (VM), Logistic Regression (LR), and Autoregressive Moving Average (ARMA)/Generalized Autoregressive Conditional Heteroscedastic (GARCH) models was proposed to assess bearing failure degradation as shown in Figure 9. The results confirmed the ability of the proposed method for bearing failure degradation assessment.

Sun et al. [132] proposed an SVM-based model for bearing prognosis. In this model, Principal Component Analysis (PCA) was employed for feature extraction from a vibration signal, and the SVM parameters were optimized using PSO. The expected result based on bearing run-to-failure experimental data confirmed that the proposed model was more accurate L_{10} life formula, shown in Eq. 12. Note that L_{10} is basic rating life, n denotes shaft speed, C is basic dynamic load rating, P is equivalent dynamic bearing load, and p expresses exponent of the life equation.

$$L_{10} = \frac{10^6}{60n}(\frac{C}{P})^p \tag{13}$$

Dong and Luo [133] developed an approach to determine bearing degradation based on a combination of PCA and an optimized Least-Squares Support Vector Machine (LSSVM)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2020.3030165, IEEE Transactions on Instrumentation and Measurement

Algorithm 2: Bayesian-based RUL prediction [60]

1) A healthy dataset is set: $x^{(1)}, x^{(2)}, ..., x^{(w-1)}$
2) A sliding window data as a training set with the degradation trend is selected: $x^{(w)}, x^{(w+1)}, ..., x^{(z)}$
3) A failure criterion, FC is set using the mean (μ), the standard deviation (σ) of healthy dataset, and the coefficient of complete failure criteria (λ_q)

$$FC = \mu + \lambda\sigma \tag{7}$$

4) An optimal affine function of discrete-time t on the training set, presented in step 2), is identified over a sliding window with a length of $z - w$:

$$y_t = \hat{m}t + \hat{n} + e_t \tag{8}$$

Where e_t is a Gaussian white noise error with zero mean and variance σ^2, and $\hat{m}$ and $\hat{n}$ are the optimal affine function estimated parameters.

5) The probability of failure $p(F_{t_0+k})$ is estimated at time $t_0 + k$:

$$p(F_{t_0+k}) = p(F_{t_0+k}|H_{t_0:t_0+k-1})p(H_{t_0:t_0+k-1}) \tag{9}$$

where $p(H_{t_0:t_0+k-1})$ is the probability of staying healthy until $t_0 + k - 1$:

$$p(F_{t_0+k}|H_{t_0:t_0+k-1}) = Q(\frac{FC - y_{t_0+k}}{\sigma\sqrt{k+1}}) \tag{10}$$

Probability $p(F_{t_0+k}|H_{t_0:t_0+k-1})$ is known as likelihood function, and can be estimated by using the standard probability Gaussian distribution function (Q):

$$p(H_{t_0:t_0+k-1}) = [1 - p(F_{t_0+1}|H_{t_0})]\times \\ \cdots \times [1 - p(F_{t_0+k-1}|H_{t_0:t_0+k-2})] \tag{11}$$

6) RUL is calculated using the time of a complete failure described as the component deficiency to accomplish its tasks ($t_{Failure}$) and the time at which prediction is made (t_0):

$$RUL = t_{Failure} - t_0 \tag{12}$$

Note that based on Gaussian distribution theory, the failure probability sequences, $p(F_{t_0+k})$ of the prediction horizon k exhibit monotonic growth. The $t_{Failure}$ of the system is determined at a time k where the probability of failure is at its peak.

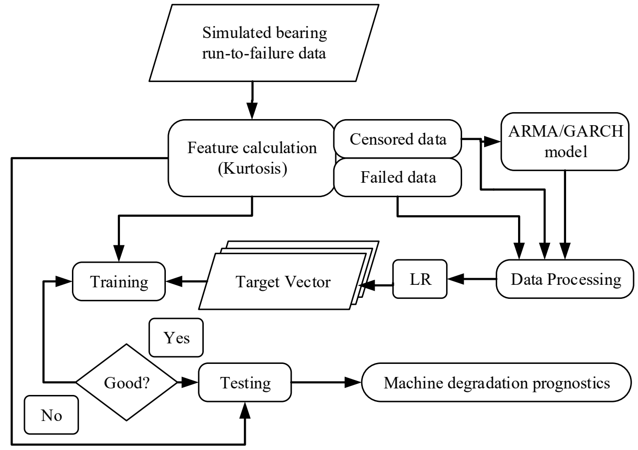


Fig. 9. A combination of VM, LR and ARMA GARCH for bearings RUL estimation [131]

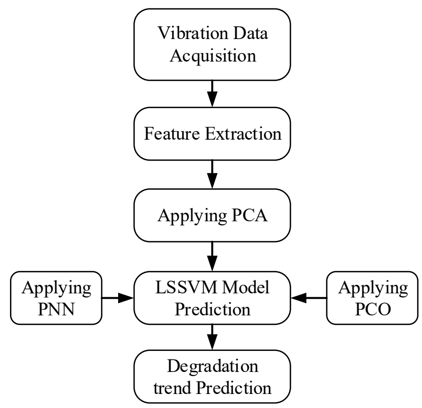


Fig. 10. A combination of PCA and LSSVM for bearings RUL estimation [133]

method, as shown in Figure 10. Firstly, PCA was employed to decrease the dimension of the extracted features. Then, the LSSVM model was formed and trained based on the extracted features for bearing degradation trend estimation. The Pseudo Nearest Neighbor (PNN) and the PSO were applied for the input number of the model estimation and the LSSVM parameter selection, respectively. The experimental results confirmed the effectiveness of the methodology. A hybrid approach for prognostics based on the Least Squares Support Vector Regression (LSSVR), and the HMM was proposed by Liu et al. [134]. Features extracted from vibration signals were utilized for training HMMs. The LSSVR algorithm was employed to predict feature trends. The predicted features probabilities for each HMM were estimated using forward or backward algorithms. Then, these probabilities helped with determining future health states and anticipating the RUL. Test results illustrated that the LSSVR/HMM approach predicted faults before their occurrence.

Hong et al. [135] proposed a combination of Wavelet Packet Decomposition (WPD), Empirical Mode Decomposition (EMD) and Self-Organizing Map (SOM) neural network techniques, as shown in Figure 11 for assessing the state of the bearing's degradation and estimating the RUL. A health indicator named Confidence Value (CV) was derived from the SOM network. The results indicated that the CV could effectively identify the degradation stage and help to estimate the RUL accurately. Later, the CV change rate was used to classify degradation stages into normal, slight degradation, severe degradation, and failure stages. Then, the corresponding prognosis models are chosen to determine the health trend

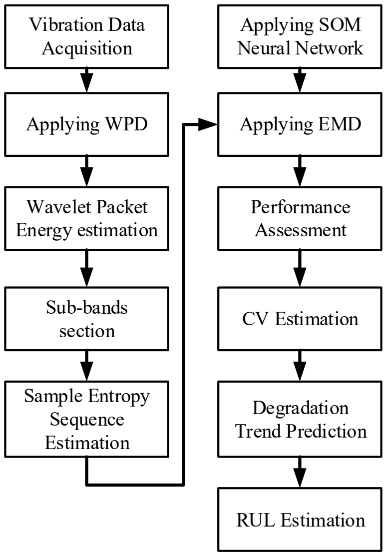


Fig. 11. A combination of WPD, EMD and SOM neural network techniques for bearings RUL estimation [135]

and RUL. The proposed hybrid approach enhanced accuracy when entering the severely degraded stage compared to the traditional single method, such as WNN [136].

Soualhi et al. [137] presented a prognostic methodology that combines HMM, the multistep time series prediction, and the ANFIS for providing the imminence of the next degradation state and estimating the remaining time before the next degradation state. The experimental results showed the proposed methodology potential for the detection, diagnosis, and prognosis in roller bearings. A combination of Simplified Fuzzy Adaptive Resonance Theory Map (SFAM) neural network and Weibull Distribution (WD) was developed by Ali et al. [138] for bearing prognosis. Experimental results showed that the capability of the proposed method to estimate the RUL of rolling element bearings based on vibration signals.

Soualhi et al. [139] proposed a hybrid approach that combined the Hilbert Huang Transform (HHT) to extract feature indexes from raw vibration signals, an SVM to detect the degradation states, and the Support Vector Regression (SVR) to estimate the RUL of ball bearings. The experimental results confirmed that the use of the HHT, the SVM, and the SVR is a suitable strategy to enhance the detection, diagnosis, and prognosis of bearing degradation. Their proposed prognostic approach is shown in Algorithm 3. Wang et al. [140] proposed a two-stage strategy prognosis including, first, estimation of degradation by determining the deviation of extracted features from a known healthy state and, then, estimating the RUL of the bearing using an enhanced KF and an Expectation–Maximization (EM) algorithm. The results confirmed that their proposed approach could provide higher estimation accuracy and narrower PDFs in comparison with Gebraeel's model [141] and Si's model [142].

Zhao et al. [143] presented a feature extraction system for vibration-based bearing prognosis using Time-Frequency Representation (TFR) and supervised dimensionality reduction.

Algorithm 3: SVR-based RUL using HHT HI [139]

1) Setting HI up to time j using HHT, and a time series of observation extracted from this indicator: $X_j = (x_j, x_{j-q}, x_{j-2q}, x_{j-3q}, \ldots, x_{j-(n-1)q})$

Note that q denotes the interval of the measure and $(n-1)$ is the length of the series.

2) Determining the health indicator, x_{j+p} at time $j+p$, where p is the horizon of prediction.

$$\begin{aligned}\hat{x}_{j+p} &= f(x_j, x_{j-q}, x_{j-2q}, x_{j-3q}, \ldots, x_{j-(n-1)q}) \\ &= f(X_j)\end{aligned} \tag{14}$$

Note that $f(X_j)$ denotes the prediction model of the time series X_j.

3) Choosing the prediction model $f(X_j)$ based on SVR

$$f(X_t) = \Omega = w^T.\Phi(X_t) + b \tag{15}$$

where

$$w = \sum_{i=1}^{n-1} (\alpha_i^+ - \alpha_i^-)\Phi(x_i) \tag{16}$$

$$b = \Omega_s - \varepsilon - \sum\nolimits_{m\in s} (\alpha_m^+ - \alpha_m^-)\Phi(x_m)\cdot\Phi(x_s) \tag{17}$$

where s is the set of indices of the support vectors, Ω_s presents the class that represents the degradation state of the bearing, $\Phi : x \to \Phi(x)$ denotes a nonlinear function that projects the observation x into a higher dimensional space, ε is a margin of tolerance, set to tolerate the deviation of the regression from the real values, and variables α_i^+ and α_i^- are obtained by applying a quadratic programming solver on Lagrange multipliers presented in [139].

4) Estimating the RUL as the smallest of RULs deduced by several x_{j+p}:

$$\underset{\text{estimated}}{\text{RUL}} \leftarrow \min\left(\text{RUL}_1, \text{RUL}_2, \text{RUL}_3, \ldots\right) \tag{18}$$

A combination of TFR, Gaussian pyramid (GP), and Local Binary Pattern (LBP) was used to evaluate lifetime information represented by highly dimensional features. The RULs are determined by employing simple Multiple Linear Regressions (MLRs). The experimental results demonstrated that the proposed method outperforms techniques employing traditional statistical features and PCA. In Jin et al. [144], a health index was proposed to detect bearing health states. A nonlinear form was developed to track the bearings' degradation process, and an EKF was employed for the RUL prediction. Test results showed that the advance warning of bearing failure could be obtained, and ongoing maintenance can be scheduled by identifying the anomaly successfully.

Jiang et al. [145] proposed an evaluation approach for bear-

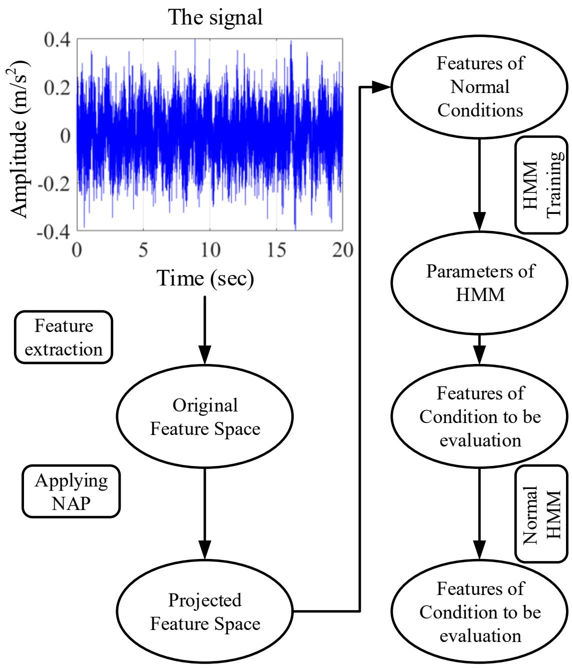


Fig. 12. A combination of HMM and NAP bearing performance descending evaluation [145]

ing performance degradation using a combination of HMM and Nuisance Attribute Projection (NAP), as shown in Figure 12. It was illustrated that the NAP could remove the impact of nuisance attributes, and the new feature space calculated by the NAP was barely affected by other interference occurring during operation. The experimental results showed that their approach improved the accuracy of the bearing performance assessment system.

A prognostic method based on vibration signals, including health monitoring methodology for WT HSSBs, was introduced by Saidi et al. [146] using a Spectral Kurtosis (SK) data-driven approach. Based on the historical run-to-failure vibration data analysis and SK indices, SVR was employed to implement RUL prediction. Although it was shown that SK-derived features could provide an early warning for bearing defects and helped with the evaluation of bearing degradation, the predictions initially overestimated the RUL. Aye and Heyns [147] proposed an optimal GPR, an integration of mean and covariance functions, for capturing the bearing degradation trend. The GPR also captured the irregularities within the data and, subsequently, improved the RUL estimation for slow speed bearings. The experimental outcomes indicated that their model demonstrated improvement over simpler GPR models.

Lu et al. [148] proposed a prognostic algorithm applying a combination of the Variable Forgetting Factor Recursive Least-Square (VFF-RLS), an ARMA model. To demonstrate the capability of the proposed methodology, the accuracy of the prediction of the proposed model is examined utilizing bearing experimental data compared to an Auto-Regressive Integrated Moving Average (ARIMA) model without adaptation. Results confirmed accurate predictions of the hybrid prognostic method over the ARIMA model. Elforjani et al. [149] proposed Signal Intensity Estimator (SIE) as a new indicator to detect individual types of early fault in real-world wind turbine bearings. This study indicated the ability of the proposed indicator to accurately estimate the RUL for wind turbine bearings in a combination of Regression Trees (RT) and multilayer ANN models. The experimental results demonstrated that SIE has an advantage over the other fault indicators, such as Crest Factor (CF) and Kurtosis, if sufficient data are provided.

Ahmad et al. [150] presented a hybrid method that employed regression-based adaptive predictive techniques to learn the degradation trend to project the RUL of a bearing. The approach applied a gradient-based method to determine the Time to Start Prediction (TSP) accurately using linear regression analysis, which contributes to relatively more accurate RUL predictions. A deep feature optimization fusion method was proposed by Zhao et al. [151] to extract centrifugal pump bearing degradation features from large amounts of vibration data. It benefited from the capability of DNN in extracting highly abstracted features that correlate well with bearing degradation. The detailed experiments on real datasets showed that the developed method has an advantage over other methods and creates degradation trajectories with potential predictive capabilities, therefore enhancing the accuracy of RUL prediction.

Wang et al. [152] predicted WT bearing RUL by employing a combination of physical knowledge and statistical model in a Bayesian framework. For this aim, first, an empirical model for the spalling evolution based on Paris' formula, shown in Eq. 19, was developed. Then, PF was developed as a recursive numerical approach based on the sequential Monte Carlo sampling technique to estimate the posterior probability density function of the state, as indicated in Eq. 20. The experimental results confirmed that the proposed method was capable of inferring the hidden defect state of the bearing from noisy measurement based on Bayesian inference and quantifying the uncertainty of the RUL prognosis in a probabilistic manner.

$$x_{t+1} = [x_t^{(1-m)} + c(1-m)]^{\frac{1}{1-m}} + u_t \tag{19}$$

$$\begin{cases} Prediction : p(x_{t+1}|z_t) = \sum_{j=1}^{M} w_t^j p(x_{t+1}|x_t^j) \\ Update : w_{t+1}^j = w_t^j p(z_{t+1}|x_{t+1}^j) \end{cases} \tag{20}$$

where x_{t+1} represents the spalling area at time index $t+1$, u_t denotes the noise in the state evolving process, and the model parameters c and m are initialized as unknown variables. $p(x_{t+1}|z_t)$ represents probability distribution at time $t+1$, i is the index of a particle, M is the total number of particles, and w_{t+1}^j denotes the weight of particle i at time $t+1$.

Elforjani and Shanbr [153] employed the combination of SVM regression, multilayer ANNs models, and GPR to estimate the RUL of slow speed bearings by correlating features with the corresponding natural wear throughout a series of laboratory experiments. It was concluded that the neural networks model with a backpropagation learning algorithm

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2020.3030165, IEEE Transactions on Instrumentation and Measurement

outperformed the other models in predicting the RUL for slow speed bearings. This was true when the appropriate network structure was chosen, and enough data was provided. Qiu et al. [154] presented a prognostic procedure by combining a HI and PF to determine the bearing RUL. The process included applying the Structural Information of the Spectrum (SIOS) algorithm to build the HI called SIOS-based Indicator (SIOSI) for bearing deterioration monitoring. Then, they assessed the Initial Degradation Point through an index calculated with a self-zero space observer and predicted the bearing RUL using the SIOSI and a PF-based algorithm that was aided by a degradation model. Experimental results have shown that the bearing RUL could be acceptably anticipated by the proposed method, and its performance was superior to conventional prognostic methods.

Rezamand et al. [155] developed a hybrid approach for RUL prediction of WT bearings under varying operating conditions. For this aim, first, SCADA measurements are categorized into two states of normal and aggressive using Kernel Fuzzy C-Means. Then, HMM and Viterbi were utilized to determine the most likely switching between the states. Next, a Damage Progression Model (DPM) is determined for each state based on extracted features form vibration signals. Finally, different paths for future transitions in the HMM were generated to define the most likely state in which WT will operate in the future. For each path, the RUL was forecast via adaptive Bayesian Algorithm on defined DPMs conditional to the projected state. RUL estimates of generated paths were averaged to achieve an accurate RUL. Experimental results on two bearings with outer and inner raceways failure indicated the efficacy of the proposed approach.

Rai et al. [156] introduced a data-driven prognosis approach based on an NARX neural network model that utilized a wavelet-filter technique for bearing RUL estimation. In time-domain modeling, an NARX is a nonlinear autoregressive model that has exogenous inputs. This implies that the model links the current value of a time series to past values of the same series and current and past values of the driving (exogenous) series [157]. The proposed approach was comprised of several steps, as shown in Figure 13 .

- In order to boost the impulsive aspects of bearing signals and enhance the quality of fault feature extraction, the vibration signals provided by an experimental test rig were preprocessed with the proposed wavelet-filter.
- To address the highly non-monotonic behavior of the extracted features due to the bearing degradation, an HI based on Mahalanobis Distance (MD) criterion [158], [159] and Cumulative Sum (CUMSUM) chart [160] was introduced.
- The NARX neural network was developed as a Time Delay Neural Network (TDNN) model, which was trained by the introduced HI and bearing age as inputs and bearing life percentage as output for bearing RUL estimation.

The results confirmed that the proposed method could accurately predict the RUL of bearings and outperformed the application of the self-organizing map-based indicator. Hu et al. [161] presented a real-time performance degradation model

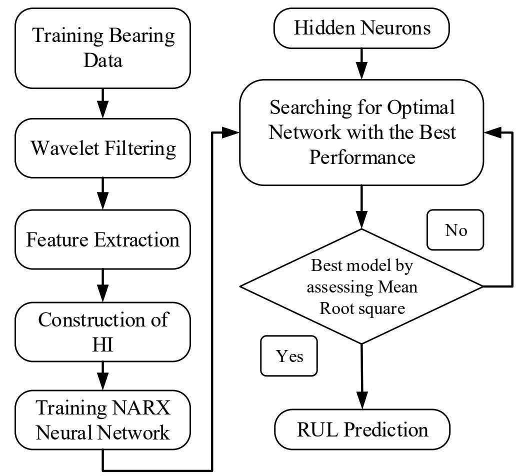


Fig. 13. Flowchart for prognosis approach based on a NARX neural network model in association with a wavelet-filter technique for bearing RUL estimation [156]

based on temperature characteristic parameters for prognosis of wind turbine bearings. Here a combination of the Wiener process for establishing the performance degradation model, the maximum likelihood estimation method for obtaining the parameters of the developed model, and an inverse Gaussian distribution approach for RUL prediction was employed to achieve this. The comparison of the predicted RUL and actual RUL revealed that the hybrid prediction method was correct and effective.

Hemmer et al. [162] suggested a framework based on three fault classifiers of CNN, SVM, and Sparse Autoencoder (SA)-based SVM utilizing transfer learning. The effectiveness of the proposed technique was examined employing vibration and acoustic emission signal datasets from roller bearings with artificial damage. The survey showed the ability of the combination of a trained CNN and SVM for extracting features and classification, respectively, in detecting faults in roller bearings based on robustness, easy implementation, and computational weight. However, the combination of a trained CNN and sparse autoencoder for extracting features and, then, decreasing dimensions of extracted features increased the computational weight and complexity as well as reducing the accuracy.

This section reveals how a combination of various WT prognosis techniques can lead to higher accuracy compared to individually employed prognosis methods. For instance, ANFIS, a combination of fuzzy logic and NNs benefits from the simplicity and strength of NNs and the reasoning of fuzzy systems, outperforms NNs in WT bearing RUL projection. The details are shown in Table V.

IV. The Problem of Prognostics in Critical WT Components

To provide a brief idea of the practical applications related to WT prognostic field, the authors propose a hybrid approach based on the idea of [155] using both SCADA and CM measurements for critical WT components prognostics in the following steps.

TABLE V
A SUMMARY OF THE LITERATURE REVIEW ON HYBRID PROGNOSTIC METHODS

Reference	Proposed Architecture	Application	Comparison&Accuracy
Rezamand et al. [60]	Bayesian+OWA	Bearing prognosis	OWA operator achieves higher Average Relative Accuracy (ARA) in bearing RUL prediction compared to the Choquet integral fusion approach and the Bayesian algorithm obtained from single-feature driven methods, ARA $= 76\%$ (Outer raceway failure), ARA $= 91\%$ (Inner raceway failure)
Zhao et al. [128]	SVM+ARIMA	Generator prognosis	Mean relative error $= 27\%$
Cheng et al. [129]	ANFIS+PF	Gearbox prognosis	ANFIS performs better than the RNN to learn the state transition function of the fault feature in the PF algorithm and results in higher RUL estimation accuracy, RMSE $= 0.048$
Ding et al. [130]	FESA+Bayesian	Gearbox prognosis	Improved accuracy $= 83\%$
Caesarendra et al. [131]	VM+LR+ ARMA/GARCH	Bearing prognosis	ARMA/GARCH model outperforms Dempster-Shafer regression
Sun et al. [132]	SVM+PCA+PSO	Bearing prognosis	SVM+PCA+PSO is more accurate than L_{10} life formula, Error $= 3.2\%$
Dong and Luo [133]	PCA+LSSVM+PNN+ PSO	Bearing prognosis	RMSE $= 0.000118$
Liu et al. [134]	LSSVR+HMM	Bearing prognosis	MaxAE $= 43.75\%$
Hong et al. [135]	WPD+EMD+SOM	Bearing prognosis	WPD+EMD+SOM enhances prediction accuracy when entering the severely degraded stage compared to the traditional single method, such as WNN, MaxAE $= 51.8\%$
Ali et al. [138]	SFAM-NN+WD	Bearing prognosis	Error $= 2.23\%$
Soualhi et al. [139]	HHT+SVM+SVR	Bearing prognosis	MaxAE $= 1.25\%$
Wang et al. [140]	Enhanced KF+EM algorithm	Bearing prognosis	Enhanced KF+EM algorithm provides higher estimation accuracy and narrower PDFs in comparison with Gebraeel's model and Si's model, Error $= 12.09\%$
Zhao et al. [143]	TFR+GP+LBP+MLR	Bearing prognosis	TFR+GP+LBP+MLR outperforms techniques employing traditional statistical features and PCA
Jin et al. [144]	Autoregressive model+EKF	Bearing prognosis	RMSE $= 0.865$
Aye and Heyns [147]	GPR	Bearing prognosis	RMSE $= 0.0069$
Lu et al. [148]	VFF-RLS+ARMA	Bearing prognosis	VFF-RLS+ARMA outperforms ARIMA model, MSE $= 0.0164$
Elforjani et al. [149]	SIE+RT+multilayer ANN	Bearing prognosis	SIE has an advantage over the other fault indicators, such as CF and Kurtosis
Ahmad et al. [150]	Regression-based hybrid method	Bearing prognosis	MaxAE $= 32.1\%$
Wang et al. [152]	Physical model and statistical model	Bearing prognosis	The proposed method outperforms ARMA in quantifying the prediction uncertainty.
Qiu et al. [154]	SIOS+PF	Bearing prognosis	SIOS+PF is more accurate than L_{10} life formula, MaxAE $= 13.42\%$
Rai et al. [156]	NARX	Bearing prognosis	MSE $= 0.0059$
Hemmer et al. [162]	CNN+SVR+SA-based SVM	Bearing prognosis	MaxAE $= 27\%$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2020.3030165, IEEE Transactions on Instrumentation and Measurement

- First, SCADA data is utilized to determine the role of operating and environmental conditions in which a WT operates. To do so, a combination of a clustering method, HMM, and Viterbi can be utilized to categorize "n" Different Operating States (DOS) (e.g., normal, mild, and aggressive for n=3) and determine the most likely switching between the categorized states.
  Algorithm 4 indicates the proposed HMM and Viterbi approach [163], [164].

Algorithm 4: HMM and Viterbi

1) Let $q_t \in [s_1, s_2, ..., s_n]$ as the value of the hidden state at time t
where n is the number of hidden states,
$o_t \in [v_1, v_2, ..., v_m]$ is set as the observed state based on measured sensor values, and m expresses the possible number of the observed values corresponding to each state
2) Probability distribution over states initialization:

$$\pi = (\pi_1, \pi_2, ..., \pi_n) \tag{20}$$

3) Setting a transition probability matrix, $A = (a_{ij})$:

$$a_{ij} = P(q_t = s_i | q_{t-1} = s_j), 1 \leq i, j \leq n \tag{21}$$

This indicates the probability that the state is s_j at time $t-1$, conditional to, the state s_i at time t.
4) Setting a confusion matrix, $B = (b_{jk})$:

$$b_{jk} = P(o_t = v_k | q_t = s_j), 1 \leq j \leq n, 1 \leq k \leq m \tag{22}$$

This indicates the probability that the hidden state is s_j with v_k as the observed state.
5) Finding the most likely sequence of hidden states $Q = [q_1, q_2, ..., q_t]$ using the system parameter $\lambda = (\pi, A, B)$ and the observation sequence $O = [o_1, o_2, ..., o_t]$ as follows:
– Initialization:

$$\delta_1(i) = \pi_i b_{io_1} \psi_1(i) = 0, 1 \leq i \leq n \tag{23}$$

– Recursion:

$$\begin{aligned} &\delta_{t+1}(j) = \max_{1 \leq i \leq n} [\delta_t(i) a_{ij}] b_{jo_{t+1}} \\ &\psi_{t+1}(j) = \arg \max_{1 \leq i \leq n} [\delta_t(i) a_{ij}] b_{jo_{t+1}}, 1 \leq t \leq T, 1 \leq j \leq n \end{aligned} \tag{24}$$

– Computing States Sequences (with retrospect):

$$P(Q, O | \lambda) = \max_{1 \leq i \leq n} \delta_T(i), Q_{t-1} = \psi_t(Q_t), T \geq t \geq 1 \tag{25}$$

- Then, different prognostic techniques such as physics-based, regression-based methods, etc. can be employed to identify an DPM for each state ("n" DPMs for "n" DOSs) utilizing CM measurements (e.g., vibration signals for bearings or temperature measurements for generators). Kalman-based or particle-filtering-based algorithms need to switch between different DPMs as necessary when estimating the current damage condition in the system component.
- Finally, for long term prediction horizons, different paths for future transitions in the HMM can be generated to define the most likely state in which WT will operate in the future. For each path, the RUL can be forecast via prognostic techniques such as an adaptive Bayesian Algorithm, shown in 2, on "n" defined DPMs conditional to the projected state. Finally, RUL estimates of generated paths are averaged to achieve an accurate RUL.

V. Conclusions and Research Directions

This study provided a review of recent modeling developments for critical WT prognostics. Moreover, basic definitions and elemental reliability concepts were discussed. The pros and cons of each prognosis method were also highlighted. Our review has revealed the following key findings:

- WT hybrid prognosis techniques are now the leading tools for critical WT components failure prediction because of their higher accuracy over individual prognosis methods.
- WT physics-based prognostics provide the most accurate predictions with fewer data compared to data-driven techniques if the physics of models remain consistent across the component. However, physics-based models are defect-specific and complex to develop.
- WT AI-based prognostics are capable of modeling complex and nonlinear systems. However, a large amount of data over a wide range of operating conditions is required to train the prognostic model to achieve reasonable prediction accuracy, which, often in practice, is limited, especially for complex systems.
- WT stochastic-based prognostic are robust in the RUL prediction of WT components owning to their capability in modeling the uncertainty inherent in the prediction horizon of WT components.
- AI-based techniques reach a higher RUL prediction accuracy than conventional prognosis methods, including NARX, HGRUN, FOA-ELM, ANFIS, NFN, ESN, and RNN.
- Stochastic-based prognosis techniques are more accurate than conventional RUL estimation methods, including GPR, HHMM, Multi-sensor HSMM, EKF, SKF, PF, and Generic PF.

Beyond this, there are a number of challenges that merit further study. We summarize them as follows, and also include the authors recommendations to address these challenges and a specific problem on the practical applications related to prognostics of critical WT components.

A. Challenges

1) **Considering operating conditions in monitoring methods.**

Prognostic studies have largely been executed over constant environmental (operating) conditions, and the prognostic techniques are developed using monitoring methods such as vibration analysis [96], [98], [99], [105], [139], [143], [146], [150]. However, it is vital to consider varying operating conditions, which can include environmental variables like wind speed and ambient temperature. It should be recognized that damage progression can be a function of the stress and loading applied to components which subsequently affect the RUL estimation, i.e., classification of operating conditions based on the severity of an environmental condition. To consider varying operating conditions, the authors recommend to categorize different operating states in which a WT component undergos. For this aim, SCADA data can be clustered using a robust clustering technique. Then, a probablistic-based technique can be utilized to determine transitions between the categorized states, and predict the most likely state in which WT will function. Recently, the authors attempted to consider this by introducing a hybrid prognostic approach for RUL prediction of WT bearings under varying operating conditions [155] based on SCADA and CM measurements.

2) **Investigating component interactions for the prognosis task.**
Almost all studies reviewed were for individual bearings [98–100], [131–133]. However, component interactions should also be considered in the degradation process (for instance, the interaction between bearings and gears in a gearbox). More signal processing procedures could also be applied to the machine degradation process to differentiate bearing fault signals from other component signals. The authors recommend to focus future research efforts on the implementation of Bayesian processors, such as sequential Monte Carlo methods, and novel model architectures for uncertainty characterization and probabilistic characterization of stress (either structural, mechanical, thermal, and electrical) propagation between different components, such as the stress-based model architecture constructed in [165].

3) **Applying hybrid methods for the prognosis task.**
Hybrid methods use a combination of various prognosis techniques, which can lead to higher accuracy compared to individually employed prognosis methods [127–133]. Hence, it is beneficial to construct more hybrid methods to continue to achieve improved accuracies.

4) **Application of Bayesian methods.**
Bayesian approaches intrinsically consider probability theory, which may be more appropriate for RUL prediction owing to the probabilistic characteristics of the RUL task [60], [117], [122], [123], [130], [152], [155]. Therefore, additional emphasis should be applied to the Bayesian style of analysis.

5) **Physics-based methods merits to WT prognostics.**
There have been a few studies on the application of physics-based methods for the WT prognosis task [73–81]. However, the approach is robust in prognostics because of resulting in higher accuracy, if the physics of models remain consistent across the component and requiring fewer data, compared to data-driven methods. Hence, it is beneficial to apply additional emphasis on developing a mathematical model for each WT component failure mode.

VI. Appendix

The nomenclature of short abbreviations used in the system are listed in Tables VI and VII.

References

[1] J. Lee and F. Zhao, "GLOBAL WIND REPORT 2019," Global Wind Energy Council, Tech. Rep., 03 2020.

[2] J. Carroll, A. McDonald, and D. McMillan, "Failure rate, repair time and unscheduled o&m cost analysis of offshore wind turbines," *Wind Energy*, vol. 19, no. 6, pp. 1107–1119, 2016.

[3] M. Shafiee and F. Dinmohammadi, "An fmea-based risk assessment approach for wind turbine systems: a comparative study of onshore and offshore," *Energies*, vol. 7, no. 2, pp. 619–642, 2014.

[4] B. Hahn, M. Durstewitz, and K. Rohrig, "Reliability of wind turbine–experiences of 15 years with 1500 wts," *Institut für Solare Energieversorgungstechnik (ISET): Hesse, Germany*, 2005.

[5] P. Tavner, D. Greenwood, M. Whittle, R. Gindele, S. Faulstich, and B. Hahn, "Study of weather and location effects on wind turbine failure rates," *Wind energy*, vol. 16, no. 2, pp. 175–187, 2013.

[6] A. Stenberg and H. Holttinen, "Analysing failure statistics of wind turbines in finland," in *European Wind Energy Conference, April*, 2010, pp. 20–23.

[7] B. Nivedh, "Major failures in the wind turbine components and the importance of periodic inspections," *Wind Insid*, 2014.

[8] H. D. M. de Azevedo, A. M. Araújo, and N. Bouchonneau, "A review of wind turbine bearing condition monitoring: State of the art and challenges," *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 368–379, 2016.

[9] N. Chen, R. Yu, Y. Chen, and H. Xie, "Hierarchical method for wind turbine prognosis using scada data," *IET Renewable Power Generation*, vol. 11, no. 4, pp. 403–410, 2016.

[10] N. M. Vichare and M. G. Pecht, "Prognostics and health management of electronics," *IEEE transactions on components and packaging technologies*, vol. 29, no. 1, pp. 222–229, 2006.

[11] H. Alian, S. Konforty, U. Ben-Simon, R. Klein, M. Tur, and J. Bortman, "Bearing fault detection and fault size estimation using fiber-optic sensors," *Mechanical Systems and Signal Processing*, vol. 120, pp. 392–407, 2019.

[12] S. Sheng and W. Yang, "Wind turbine drivetrain condition monitoring-an overview (presentation)," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2013.

[13] W. Yang, P. J. Tavner, C. J. Crabtree, Y. Feng, and Y. Qiu, "Wind turbine condition monitoring: technical and commercial challenges," *Wind Energy*, vol. 17, no. 5, pp. 673–693, 2014.

[14] S. Sheng, "Prognostics and health management of wind turbines—current status and future opportunities," in *Probabilistic prognostics and health management of energy systems*. Springer, 2017, pp. 33–47.

[15] A. Stetco, F. Dinmohammadi, X. Zhao, V. Robu, D. Flynn, M. Barnes, J. Keane, and G. Nenadic, "Machine learning methods for wind turbine condition monitoring: A review," *Renewable energy*, vol. 133, pp. 620–635, 2019.

[16] C. Warren, C. Niezrecki, and P. Avitabile, "Optical non-contacting vibration measurement of rotating turbine blades ii," in *Structural Dynamics and Renewable Energy, Volume 1*. Springer, 2011, pp. 39–44.

[17] J. Baqersad, P. Poozesh, C. Niezrecki, and P. Avitabile, "A noncontacting approach for full-field strain monitoring of rotating structures," *Journal of Vibration and Acoustics*, vol. 138, no. 3, 2016.

[18] F. X. Ochieng, C. M. Hancock, G. W. Roberts, and J. Le Kernec, "A review of ground-based radar as a noncontact sensor for structural health monitoring of in-field wind turbines blades," *Wind Energy*, vol. 21, no. 12, pp. 1435–1449, 2018.

[19] R. Wu, D. Zhang, Q. Yu, Y. Jiang, and D. Arola, "Health monitoring of wind turbine blades in operation using three-dimensional digital image correlation," *Mechanical Systems and Signal Processing*, vol. 130, pp. 470–483, 2019.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2020.3030165, IEEE Transactions on Instrumentation and Measurement

TABLE VI
TABLE OF ACRONYMS

Abbreviation	Definition
AE	Absolute Error
AI	Artificial Intelligence
ANFIS	Adaptive Neuro-Fuzzy Inference Systems
ANN	Artificial Neural Network
AOI	Anomaly Operation Index
ARA	Average Relaive Accuracy
ARF	Aggregate Reliability Function
ARMA	Autoregressive Moving Average
ARIMA	Auto-Regressive Integrated Moving Average
BA	Bat Algorithm
BFO	Bacterial Foraging Optimization
BP	Back Propagation
CF	Crest Factor
CM	Condition Monitoring
CMS	Condition Monitoring Systems
CNN	Convolutional Neural Network
CUMSUM	Cumulative Sum
CV	Confidence Value
DBN	Deep Belief Network
DL	Deep Learning
DNN	Deep Neural Network
DOS	Different Operating States
DPM	Damage Progression Model
EKF	Extended Kalman Filter
ELM	Extreme Learning Machine
EM	Expectation–Maximization
EMD	Empirical Mode Decomposition
ENN	Elman Neural Network
ESN	Echo State Network
FDP	Fault Diagnosis and Prognosis
FEA	Finite Element Analysis
FESA	Finite Element Stress Analysis
FL	Fuzzy Logic
FNN	Feedforward Neural Network
FOA	Fruit Fly Optimization Algorithm
GA	Genetic Algorithm
GP	Gaussian Pyramid
GPR	Gaussian Process Regression
GWEC	Global Wind Energy Council
HGRUN	Hierarchical Gated Recurrent Unit Network
HHHM	Hierarchical Hidden Markov model
HHT	Hilbert Huang Transform
HI	Health Indicator
HMM	Hidden Markov Model
HSMM	Hidden Semi-Markov Model
HSSB	High-Speed Shaft Bearing
KF	Kalman Filter
LBP	Local Binary Pattern
LFGRU	Local Feature-based Gated Recurrent Unit
LR	Logistic Regression
LSSVM	Least Squares Support Vector Machine
LSSVR	Least Squares Support Vector Regression
MaxAE	Maximum of Absolute Error

TABLE VII
TABLE OF ACRONYMS (CONTINUE)

Abbreviation	Definition
MAE	Mean Absolute Error
MD	Mahalanobis distance
ML	Machine Learning
MLE	Maximum Likelihood Estimation
MLP	Multilayer Perceptron
MLR	Multiple Linear Regression
MSE	Mean Square Error
NARX	Nonlinear Autoregressive Exogenous
NAP	Nuisance Attribute Projection
NFN	Neo-Fuzzy Neuron
NLRR	Non-Linear Rank Regression
NN	Neural Network
O&M	Operation and Maintenance
PCA	Principal Component Analysis
PCO	Pseudo Nearest Neighbor
PDF	Probability Density Function
PF	Particle Filter
PHM	Prognostics and Health Management
PM	Performance Monitoring
PNN	Pseudo Nearest Neighbor
PSO	Particle Swarm Optimization
RLS	Recursive Least-Square
RMLP	Recurrent Multilayer Perceptron
RMS	Root Mean Square
RNN	Recurrent Neural Network
RMSE	Root Mean Square Error
RQ	Rational Quadratic
RT	Regression Trees
RUL	Remaining Useful Lifetime
SA	Sparse Autoencoder
SCADA	Supervisory Control and Data Acquisition
SE	Squared Exponential
SKF	Switching Kalman Filter
SK	Spectral Kurtosis
SHMM	Semi-Hidden Markov Model
SIE	Signal Intensity Estimator
SIOS	Structural Information Of the Spectrum
SIOSI	SIOS-based Indicator
SOM	Self-Organizing Map
SVM	Support Vector Machine
SVR	Support vector regression
SW	Stator Windings
TDNN	Time Delay Neural Network
TFR	Time-Frequency Representation
TSP	Time to Start Prediction
VFF	Variable Forgetting Factor
VM	Vector Machine
WAFTR	Weibull Accelerated Failure Time Regression
WD	Weibull Distribution
WPD	Wavelet Packet Decomposition
WNN	Wavelet Neural Network
WT	Wind Turbine

[20] W. Luo, J. Li, X. Ma, and W. Wei, "A novel static deformation measurement and visualization method for wind turbine blades using home-made lidar and processing program," *Optics and Lasers in Engineering*, vol. 134, p. 106206, 2020.

[21] A. Khadka, B. Fick, A. Afshar, M. Tavakoli, and J. Baqersad, "Non-contact vibration monitoring of rotating wind turbines using a semi-autonomous uav," *Mechanical Systems and Signal Processing*, vol. 138, p. 106446, 2020.

[22] Y. Du, S. Zhou, X. Jing, Y. Peng, H. Wu, and N. Kwok, "Damage detection techniques for wind turbine blades: A review," *Mechanical Systems and Signal Processing*, vol. 141, p. 106445, 2020.

[23] A. Heng, S. Zhang, A. C. Tan, and J. Mathew, "Rotating machinery prognostics: State of the art, challenges and opportunities," *Mechanical systems and signal processing*, vol. 23, no. 3, pp. 724–739, 2009.

[24] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, and N. Zerhouni, "Prognostics and health management of pemfc–state of the art and remaining challenges," *International Journal of Hydrogen Energy*, vol. 38, no. 35, pp. 15 307–15 317, 2013.

[25] P. Tchakoua, R. Wamkeue, M. Ouhrouche, F. Slaoui-Hasnaoui, T. A. Tameghe, and G. Ekemb, "Wind turbine condition monitoring: State-of-the-art review, new trends, and future challenges," *Energies*, vol. 7, no. 4, pp. 2595–2630, 2014.

[26] F. P. G. Márquez, A. M. Tobias, J. M. P. Pérez, and M. Papaelias, "Condition monitoring of wind turbines: Techniques and methods," *Renewable Energy*, vol. 46, pp. 169–178, 2012.

[27] G. d. N. P. Leite, A. M. Araújo, and P. A. C. Rosas, "Prognostic techniques applied to maintenance of wind turbines: a concise and specific review," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1917–1925, 2018.

[28] Y. Merizalde, L. Hernández-Callejo, O. Duque-Perez, and V. Alonso-Gómez, "Maintenance models applied to wind turbines. a comprehensive overview," *Energies*, vol. 12, no. 2, p. 225, 2019.

[29] E. Byon and Y. Ding, "Season-dependent condition-based maintenance for a wind turbine using a partially observed markov decision process," *IEEE Transactions on Power Systems*, vol. 25, no. 4, pp. 1823–1834, 2010.

[30] Y. Amirat, M. E. Benbouzid, B. Bensaker, and R. Wamkeue, "Condition monitoring and ault diagnosis in wind energy conversion systems: a review," in *2007 IEEE International Electric Machines & Drives Conference*, vol. 2. IEEE, 2007, pp. 1434–1439.

[31] Z. Daneshi-Far, G.-A. Capolino, and H. Henao, "Review of failures and condition monitoring in wind turbine generators," in *The XIX International Conference on Electrical Machines-ICEM 2010*. IEEE, 2010, pp. 1–6.

[32] H. Badihi, Y. Zhang, and H. Hong, "A review on application of monitoring, diagnosis, and fault-tolerant control to wind turbines," in *2013 Conference on Control and Fault-Tolerant Systems (SysTol)*. IEEE, 2013, pp. 365–370.

[33] M. Kabir, A. M. Oo, and M. Rabbani, "A brief review on offshore wind turbine fault detection and recent development in condition monitoring based maintenance system," in *2015 Australasian Universities Power Engineering Conference (AUPEC)*. IEEE, 2015, pp. 1–7.

[34] W. Qiao and D. Lu, "A survey on wind turbine condition monitoring and fault diagnosis—part i: Components and subsystems," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10, pp. 6536–6545, 2015.

[35] ——, "A survey on wind turbine condition monitoring and fault diagnosis—part ii: Signals and signal processing methods," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10, pp. 6546–6557, 2015.

[36] J. Tautz-Weinert and S. J. Watson, "Using scada data for wind turbine condition monitoring–a review," *IET Renewable Power Generation*, vol. 11, no. 4, pp. 382–394, 2016.

[37] W. Caesarendra and T. Tjahjowidodo, "A review of feature extraction methods in vibration-based condition monitoring and its application for degradation trend estimation of low-speed slew bearing," *Machines*, vol. 5, no. 4, p. 21, 2017.

[38] C. Li, J. V. De Oliveira, M. Cerrada, D. Cabrera, R. V. Sánchez, and G. Zurita, "A systematic review of fuzzy formalisms for bearing fault diagnosis," *IEEE Transactions on Fuzzy Systems*, vol. 27, no. 7, pp. 1362–1382, 2018.

[39] H. Seyr and M. Muskulus, "Decision support models for operations and maintenance for offshore wind farms: a review," *Applied Sciences*, vol. 9, no. 2, p. 278, 2019.

[40] S. Zhang, S. Zhang, B. Wang, and T. G. Habetler, "Deep learning algorithms for bearing fault diagnosticsx—a comprehensive review," *IEEE Access*, vol. 8, pp. 29 857–29 881, 2020.

[41] N. Jammu and P. Kankar, "A review on prognosis of rolling element bearings," *International Journal of Engineering Science and Technology*, vol. 3, no. 10, pp. 7497–7503, 2011.

[42] J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao, and D. Siegel, "Prognostics and health management design for rotary machinery systems—reviews, methodology and applications," *Mechanical systems and signal processing*, vol. 42, no. 1-2, pp. 314–334, 2014.

[43] S. Baltazar, C. Li, H. Daniel, and J. V. de Oliveira, "A review on neurocomputing based wind turbines fault diagnosis and prognosis," in *2018 Prognostics and System Health Management Conference (PHM-Chongqing)*. IEEE, 2018, pp. 437–443.

[44] B. Lu, Y. Li, X. Wu, and Z. Yang, "A review of recent advances in wind turbine condition monitoring and fault diagnosis," in *2009 IEEE Power Electronics and Machines in Wind Applications*. IEEE, 2009, pp. 1–7.

[45] X. Y. Zhong, L. C. Zeng, C. H. Zhao, J. Zhang, and S. Q. Wan, "Research of condition monitoring and fault diagnosis techniques for wind turbine gearbox," in *Applied Mechanics and Materials*, vol. 197. Trans Tech Publ, 2012, pp. 206–210.

[46] Y. Feng, Y. Qiu, C. J. Crabtree, H. Long, and P. J. Tavner, "Monitoring wind turbine gearboxes," *Wind Energy*, vol. 16, no. 5, pp. 728–740, 2013.

[47] I. Zuccolin, "KEC T78 GEARBOX PLANETARY FAILURE REPORT," Kruger Energy Inc., Ontario, Tech. Rep., 12 2013.

[48] J. Cibulka, M. K. Ebbesen, G. Hovland, K. G. Robbersmyr, and M. R. Hansen, "A review on approaches for condition based maintenance in applications with induction machines located offshore," 2012.

[49] G. Stone and J. Kapler, "Stator winding monitoring," *IEEE Industry Applications Magazine*, vol. 4, no. 5, pp. 15–20, 1998.

[50] "MSD Quality System Documentation, MSD Initial Repair Report 042214," Kruger Energy Inc., Ontario, Tech. Rep., 05 2015.

[51] M. Zaggout, P. Tavner, C. Crabtree, and L. Ran, "Detection of rotor electrical asymmetry in wind turbine doubly-fed induction generators," *IET Renewable Power Generation*, vol. 8, no. 8, pp. 878–886, 2014.

[52] L. M. Popa, B.-B. Jensen, E. Ritchie, and I. Boldea, "Condition monitoring of wind generators," in *38th IAS Annual Meeting on Conference Record of the Industry Applications Conference, 2003.*, vol. 3. IEEE, 2003, pp. 1839–1846.

[53] L. Cheng, J. Lin, Y.-Z. Sun, C. Singh, W.-Z. Gao, and X.-M. Qin, "A model for assessing the power variation of a wind farm considering the outages of wind turbines," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 3, pp. 432–444, 2012.

[54] M. Sundaresan, M. Schulz, and A. Ghoshal, "Structural health monitoring static test of a wind turbine blade," 2002.

[55] C. C. Ciang, J.-R. Lee, and H.-J. Bang, "Structural health monitoring for a wind turbine system: a review of damage detection methods," *Measurement science and technology*, vol. 19, no. 12, p. 122001, 2008.

[56] M. Rezamand, M. Kordestani, R. Carriveau, D. S. Ting, and M. Saif, "A new hybrid fault detection method for wind turbine blades using recursive pca and wavelet-based pdf," *IEEE Sensors Journal*, 2019.

[57] S. Sankar, M. Nataraj, and P. V. Raja, "Failure analysis of bearing in wind turbine generator gearbox," *Journal of Information Systems and Communication*, vol. 3, no. 1, p. 302, 2012.

[58] M. Blodt, P. Granjon, B. Raison, and G. Rostaing, "Models for bearing damage detection in induction motors using stator current monitoring," *IEEE transactions on industrial electronics*, vol. 55, no. 4, pp. 1813–1822, 2008.

[59] J. R. Stack, T. G. Habetler, and R. G. Harley, "Fault classification and fault signature production for rolling element bearings in electric machines," *IEEE Transactions on Industry Applications*, vol. 40, no. 3, pp. 735–739, 2004.

[60] M. Rezamand, M. Kordestani, R. Carriveau, D. S.-K. Ting, and M. Saif, "An integrated feature-based failure prognosis method for wind turbine bearings," *IEEE/ASME Transactions on Mechatronics*, 2020.

[61] J. Sikorska, M. Hodkiewicz, and L. Ma, "Prognostic modelling options for remaining useful life estimation by industry," *Mechanical systems and signal processing*, vol. 25, no. 5, pp. 1803–1836, 2011.

[62] M. Kordestani, M. F. Samadi, M. Saif, and K. Khorasani, "A new fault diagnosis of multifunctional spoiler system using integrated artificial neural network and discrete wavelet transform methods," *IEEE Sensors Journal*, vol. 18, no. 12, pp. 4990–5001, 2018.

[63] M. Kordestani, A. Zanj, M. E. Orchard, and M. Saif, "A modular fault diagnosis and prognosis method for hydro-control valve system based on redundancy in multisensor data information," *IEEE Transactions on Reliability*, no. 99, pp. 1–12, 2018.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2020.3030165, IEEE Transactions on Instrumentation and Measurement

[64] M. Kordestani, M. Saif, M. E. Orchard, R. Razavi-Far, and K. Khorasani, "Failure prognosis and applications—a survey of recent literature," *IEEE transactions on reliability*, 2019.

[65] A. Fatemi and L. Yang, "Cumulative fatigue damage and life prediction theories: a survey of the state of the art for homogeneous materials," *International journal of fatigue*, vol. 20, no. 1, pp. 9–34, 1998.

[66] C. A. A. Magniez, "Combining information flow and physics-of-failure in mechatronic products," 2007.

[67] Y. Li, T. Kurfess, and S. Liang, "Stochastic prognostics for rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 14, no. 5, pp. 747–762, 2000.

[68] C. J. Li and H. Lee, "Gear fatigue crack prognosis using embedded model, gear dynamic model and fracture mechanics," *Mechanical systems and signal processing*, vol. 19, no. 4, pp. 836–846, 2005.

[69] C. H. Oppenheimer and K. A. Loparo, "Physically based diagnosis and prognosis of cracked rotor shafts," in *Component and Systems Diagnostics, Prognostics, and Health Management II*, vol. 4733. International Society for Optics and Photonics, 2002, pp. 122–132.

[70] R. F. Orsagh, J. Sheldon, and C. J. Klenke, "Prognostics/diagnostics for gas turbine engine bearings," in *ASME Turbo Expo 2003, collocated with the 2003 International Joint Power Generation Conference*. American Society of Mechanical Engineers Digital Collection, 2003, pp. 159–167.

[71] S. Marble and B. P. Morton, "Predicting the remaining life of propulsion system bearings," in *2006 IEEE Aerospace Conference*. IEEE, 2006, pp. 8–pp.

[72] J. Qiu, B. B. Seth, S. Y. Liang, and C. Zhang, "Damage mechanics approach for bearing lifetime prognostics," *Mechanical systems and signal processing*, vol. 16, no. 5, pp. 817–829, 2002.

[73] C. S. Gray and S. J. Watson, "Physics of failure approach to wind turbine condition based maintenance," *Wind Energy*, vol. 13, no. 5, pp. 395–405, 2010.

[74] D. Breteler, C. Kaidis, T. Tinga, and R. Loendersloot, "Physics based methodology for wind turbine failure detection, diagnostics & prognostics," *EWEA 2015 Annual Event*, 2015.

[75] J. Zhu, J. Yoon, D. He, B. Qiu, and E. Bechhoefer, "Online condition monitoring and remaining useful life prediction of particle contaminated lubrication oil," in *2013 IEEE Conference on Prognostics and Health Management (PHM)*. IEEE, 2013, pp. 1–14.

[76] J. Zhu, J. M. Yoon, D. He, and E. Bechhoefer, "Online particle-contaminated lubrication oil condition monitoring and remaining useful life prediction for wind turbines," *Wind Energy*, vol. 18, no. 6, pp. 1131–1149, 2015.

[77] M. Grujicic, R. Galgalikar, S. Ramaswami, J. Snipes, V. Chenna, and R. Yavari, "Finite-element analysis of horizontal-axis wind-turbine gearbox failure via tooth-bending fatigue," *International Journal of Material and Mechanical Engineering*, vol. 3, no. 1, pp. 6–15, 2014.

[78] M. Florian and J. D. Sørensen, "Wind turbine blade life-time assessment model for preventive planning of operation and maintenance," *Journal of Marine Science and Engineering*, vol. 3, no. 3, pp. 1027–1040, 2015.

[79] E. Bechhoefer and R. Schlanbusch, "Generalized prognostics algorithm using kalman smoother," *IFAC-PapersOnLine*, vol. 48, no. 21, pp. 97–104, 2015.

[80] S. Butler, F. O'Connor, D. Farren, and J. V. Ringwood, "A feasibility study into prognostics for the main bearing of a wind turbine," in *2012 IEEE International Conference on Control Applications*. IEEE, 2012, pp. 1092–1097.

[81] W. Teng, C. Han, Y. Hu, X. Cheng, L. Song, and Y. Liu, "A robust model-based approach for bearing remaining useful life prognosis in wind turbines," *IEEE Access*, vol. 8, pp. 47 133–47 143, 2020.

[82] Y. Lei, *Intelligent fault diagnosis and remaining useful life prediction of rotating machinery*. Butterworth-Heinemann, 2016.

[83] M. Morshedizadeh, M. Kordestani, R. Carriveau, D. S. Ting, and M. Saif, "Improved power curve monitoring of wind turbines," *Wind engineering*, vol. 41, no. 4, pp. 260–271, 2017.

[84] M. Morshedizadeh, M. Kordestani, R. Carriveau, D. S.-K. Ting, and M. Saif, "Power production prediction of wind turbines using a fusion of mlp and anfis networks," *IET Renewable Power Generation*, vol. 12, no. 9, pp. 1025–1033, 2018.

[85] A. Zameer, J. Arshad, A. Khan, and M. A. Z. Raja, "Intelligent and robust prediction of short term wind power using genetic programming based ensemble of neural networks," *Energy conversion and management*, vol. 134, pp. 361–372, 2017.

[86] M. C. Mabel and E. Fernandez, "Analysis of wind power generation and prediction using ann: A case study," *Renewable energy*, vol. 33, no. 5, pp. 986–992, 2008.

[87] J. S. L. Senanayaka, H. Van Khang, and K. G. Robbersmyr, "Autoencoders and recurrent neural networks based algorithm for prognosis of bearing life," in *2018 21st International Conference on Electrical Machines and Systems (ICEMS)*. IEEE, 2018, pp. 537–542.

[88] Q. Cui, Z. Li, J. Yang, and B. Liang, "Rolling bearing fault prognosis using recurrent neural network," in *2017 29th Chinese Control And Decision Conference (CCDC)*. IEEE, 2017, pp. 1196–1201.

[89] J. Deutsch and D. He, "Using deep learning-based approach to predict remaining useful life of rotating components," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 1, pp. 11–20, 2018.

[90] M. Morshedizadeh, M. Kordestani, R. Carriveau, D. S.-K. Ting, and M. Saif, "Application of imputation techniques and adaptive neuro-fuzzy inference system to predict wind turbine power production," *Energy*, vol. 138, pp. 394–404, 2017.

[91] M. Mostafaei, H. Javadikia, and L. Naderloo, "Modeling the effects of ultrasound power and reactor dimension on the biodiesel production yield: Comparison of prediction abilities between response surface methodology (rsm) and adaptive neuro-fuzzy inference system (anfis)," *Energy*, vol. 115, pp. 626–636, 2016.

[92] D. Zurita, J. A. Carino, M. Delgado, and J. A. Ortega, "Distributed neuro-fuzzy feature forecasting approach for condition monitoring," in *Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA)*. IEEE, 2014, pp. 1–8.

[93] S. Hussain and H. A. Gabbar, "Vibration analysis and time series prediction for wind turbine gearbox prognostics," *International Journal of Prognostics and Health Management*, vol. 4, pp. 69–79, 2013.

[94] B. Chen, P. C. Matthews, and P. J. Tavner, "Wind turbine pitch faults prognosis using a-priori knowledge-based anfis," *Expert Systems with Applications*, vol. 40, no. 17, pp. 6863–6876, 2013.

[95] P. C. Matthews, B. Chen, and P. J. Tavner, "Automated on-line fault prognosis for wind turbine pitch systems using supervisory control and data acquisition," *IET Renewable Power Generation*, vol. 9, no. 5, pp. 503–513, 2015.

[96] Y. Pan, R. Hong, J. Chen, J. Singh, and X. Jia, "Performance degradation assessment of a wind turbine gearbox based on multi-sensor data fusion," *Mechanism and Machine Theory*, vol. 137, pp. 509–526, 2019.

[97] R. Liu, B. Yang, E. Zio, and X. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: A review," *Mechanical Systems and Signal Processing*, vol. 108, pp. 33–47, 2018.

[98] A. Malhi, R. Yan, and R. X. Gao, "Prognosis of defect propagation based on recurrent neural networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 3, pp. 703–711, 2011.

[99] W. Teng, X. Zhang, Y. Liu, A. Kusiak, and Z. Ma, "Prognosis of the remaining useful life of bearings in a wind turbine gearbox," *Energies*, vol. 10, no. 1, p. 32, 2016.

[100] Y. Xie and T. Zhang, "The application of echo state network and recurrent multilayer perceptron in rotating machinery fault prognosis," in *2016 IEEE Chinese Guidance, Navigation and Control Conference (CGNCC)*. IEEE, 2016, pp. 2286–2291.

[101] L. Guo, N. Li, F. Jia, Y. Lei, and J. Lin, "A recurrent neural network based health indicator for remaining useful life prediction of bearings," *Neurocomputing*, vol. 240, pp. 98–109, 2017.

[102] B. Karim and L. Abderrazak, "A statistical parameters and artificial neural networks application for rolling element bearing fault diagnosis using wavelet transform preprocessing." ICEE-B, 10 2017, pp. 1–6.

[103] S. E. Kramti, J. B. Ali, L. Saidi, M. Sayadi, and E. Bechhoefer, "Direct wind turbine drivetrain prognosis approach using elman neural network," in *2018 5th International Conference on Control, Decision and Information Technologies (CoDIT)*. IEEE, 2018, pp. 859–864.

[104] M. Xia, T. Li, T. Shu, J. Wan, Z. Wang *et al.*, "A two-stage approach for the remaining useful life prediction of bearings using deep neural networks," *IEEE Transactions on Industrial Informatics*, 2018.

[105] X. Li, W. Zhang, and Q. Ding, "Deep learning-based remaining useful life estimation of bearings using multi-scale feature extraction," *Reliability Engineering & System Safety*, vol. 182, pp. 208–218, 2019.

[106] X. Li, H. Jiang, X. Xiong, and H. Shao, "Rolling bearing health prognosis using a modified health index based hierarchical gated recurrent unit network," *Mechanism and Machine Theory*, vol. 133, pp. 229–249, 2019.

[107] C. Chen, B. Zhang, G. Vachtsevanos, and M. Orchard, "Machine condition prediction based on adaptive neuro–fuzzy and high-order particle filtering," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 9, pp. 4353–4364, 2011.

[108] F. Cheng, L. Qu, and W. Qiao, "A case-based data-driven prediction framework for machine fault prognostics," in *2015 IEEE Energy*

Conversion Congress and Exposition (ECCE). IEEE, 2015, pp. 3957–3963.

[109] A. Soualhi, G. Clerc, H. Razik, and F. Rivas, "Long-term prediction of bearing condition by the neo-fuzzy neuron," in *2013 9th IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives (SDEMPED)*. IEEE, 2013, pp. 586–591.

[110] M. Kordestani, M. F. Samadi, M. Saif, and K. Khorasani, "A new fault prognosis of mfs system using integrated extended kalman filter and bayesian method," *IEEE Transactions on Industrial Informatics*, 2018.

[111] S. Dey and J. Stori, "A bayesian network approach to root cause diagnosis of process variations," *International Journal of Machine Tools and Manufacture*, vol. 45, no. 1, pp. 75–91, 2005.

[112] M. Rausand and A. Høyland, *System reliability theory: models, statistical methods, and applications*. John Wiley & Sons, 2004, vol. 396.

[113] P. O'Connor and A. Kleyner, *Practical reliability engineering*. John Wiley & Sons, 2012.

[114] Q. Miao and V. Makis, "Condition monitoring and classification of rotating machinery using wavelets and hidden markov models," *Mechanical systems and signal processing*, vol. 21, no. 2, pp. 840–855, 2007.

[115] M. Rezamand, R. Carriveau, D. S.-K. Ting, M. Davison, and J. J. Davis, "Aggregate reliability analysis of wind turbine generators," *IET Renewable Power Generation*, vol. 13, no. 11, pp. 1902–1910, 2019.

[116] X. Fan, X. Yang, X. Li, and J. Wang, "A particle-filtering approach for remaining useful life estimation of wind turbine gearbox," in *International Conference on Chemical, Material and Food Engineering*. Atlantis Press, 2015.

[117] S. Hong and Z. Zhou, "Application of gaussian process regression for bearing degradation assessment," in *2012 6th International Conference on New Trends in Information Science, Service Science and Data Mining (ISSDM2012)*. IEEE, 2012, pp. 644–648.

[118] P. Kundu, A. K. Darpe, and M. S. Kulkarni, "Weibull accelerated failure time regression model for remaining useful life prediction of bearing working under multiple operating conditions," *Mechanical Systems and Signal Processing*, vol. 134, p. 106302, 2019.

[119] X.-H. Zhang and J.-S. Kang, "Hidden markov models in bearing fault diagnosis and prognosis," in *Computational Intelligence and Natural Computing Proceedings (CINC), 2010 Second International Conference on*, vol. 2, 2010, p. 364.

[120] Z. Chen, Y. Yang, Z. Hu, and Q. Zeng, "Fault prognosis of complex mechanical systems based on multi-sensor mixtured hidden semi-markov models," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 227, no. 8, pp. 1853–1863, 2013.

[121] R. K. Singleton, E. G. Strangas, and S. Aviyente, "Extended kalman filtering for remaining-useful-life estimation of bearings," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 3, pp. 1781–1790, 2015.

[122] C. K. R. Lim and D. Mba, "Switching kalman filter for failure prognostic," *Mechanical Systems and Signal Processing*, vol. 52, pp. 426–435, 2015.

[123] H. Mohammad, L. Niels, S. Esmaeil *et al.*, "Bayesian state prediction of wind turbine bearing failure," *Renewable energy*, 2017.

[124] J. Wang and R. X. Gao, "Multiple model particle filtering for bearing life prognosis," in *2013 IEEE Conference on Prognostics and Health Management (PHM)*. IEEE, 2013, pp. 1–6.

[125] C. Chen, D. Brown, C. Sconyers, G. Vachtsevanos, B. Zhang, and M. E. Orchard, "A. net framework for an integrated fault diagnosis and failure prognosis architecture," in *2010 IEEE AUTOTESTCON*. IEEE, 2010, pp. 1–6.

[126] L. Liao and F. Köttig, "Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction," *IEEE Transactions on Reliability*, vol. 63, no. 1, pp. 191–207, 2014.

[127] M. Djeziri, S. Benmoussa, and R. Sanchez, "Hybrid method for remaining useful life prediction in wind turbine systems," *Renewable Energy*, vol. 116, pp. 173–187, 2018.

[128] Y. Zhao, D. Li, A. Dong, J. Lin, D. Kang, and L. Shang, "Fault prognosis of wind turbine generator using scada data," in *2016 North American Power Symposium (NAPS)*. IEEE, 2016, pp. 1–6.

[129] F. Cheng, L. Qu, and W. Qiao, "Fault prognosis and remaining useful life prediction of wind turbine gearboxes using current signal analysis," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 1, pp. 157–167, 2017.

[130] F. Ding, Z. Tian, F. Zhao, and H. Xu, "An integrated approach for wind turbine gearbox fatigue life prediction considering instantaneously varying load conditions," *Renewable energy*, vol. 129, pp. 260–270, 2018.

[131] W. Caesarendra, A. Widodo, P. H. Thom, B.-S. Yang, and J. D. Setiawan, "Combined probability approach and indirect data-driven method for bearing degradation prognostics," *IEEE Transactions on Reliability*, vol. 60, no. 1, pp. 14–20, 2011.

[132] C. Sun, Z. Zhang, and Z. He, "Research on bearing life prediction based on support vector machine and its application," in *Journal of Physics: Conference Series*, vol. 305, no. 1. IOP Publishing, 2011, p. 012028.

[133] S. Dong and T. Luo, "Bearing degradation process prediction based on the pca and optimized ls-svm model," *Measurement*, vol. 46, no. 9, pp. 3143–3152, 2013.

[134] Z. Liu, Q. Li, X. Liu, and C. Mu, "A hybrid lssvr/hmm-based prognostic approach," *Sensors*, vol. 13, no. 5, pp. 5542–5560, 2013.

[135] S. Hong, Z. Zhou, E. Zio, and K. Hong, "Condition assessment for the performance degradation of bearing based on a combinatorial feature extraction method," *Digital Signal Processing*, vol. 27, pp. 159–166, 2014.

[136] S. Hong, Z. Zhou, E. Zio, and W. Wang, "An adaptive method for health trend prediction of rotating bearings," *Digital Signal Processing*, vol. 35, pp. 117–123, 2014.

[137] A. Soualhi, H. Razik, G. Clerc, and D. D. Doan, "Prognosis of bearing failures using hidden markov models and the adaptive neuro-fuzzy inference system," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 6, pp. 2864–2874, 2014.

[138] J. B. Ali, B. Chebel-Morello, L. Saidi, S. Malinowski, and F. Fnaiech, "Accurate bearing remaining useful life prediction based on weibull distribution and artificial neural network," *Mechanical Systems and Signal Processing*, vol. 56, pp. 150–172, 2015.

[139] A. Soualhi, K. Medjaher, and N. Zerhouni, "Bearing health monitoring based on hilbert–huang transform, support vector machine, and regression," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 1, pp. 52–62, 2015.

[140] Y. Wang, Y. Peng, Y. Zi, X. Jin, and K.-L. Tsui, "A two-stage data-driven-based prognostic approach for bearing degradation problem," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 3, pp. 924–932, 2016.

[141] N. Z. Gebraeel, M. A. Lawley, R. Li, and J. K. Ryan, "Residual-life distributions from component degradation signals: A bayesian approach," *IiE Transactions*, vol. 37, no. 6, pp. 543–557, 2005.

[142] X.-S. Si, W. Wang, C.-H. Hu, M.-Y. Chen, and D.-H. Zhou, "A wiener-process-based degradation model with a recursive filter algorithm for remaining useful life estimation," *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 219–237, 2013.

[143] M. Zhao, B. Tang, and Q. Tan, "Bearing remaining useful life estimation based on time–frequency representation and supervised dimensionality reduction," *Measurement*, vol. 86, pp. 41–55, 2016.

[144] X. Jin, Y. Sun, Z. Que, Y. Wang, and T. W. Chow, "Anomaly detection and fault prognosis for bearings," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 9, pp. 2046–2054, 2016.

[145] H. Jiang, J. Chen, and G. Dong, "Hidden markov model and nuisance attribute projection based bearing performance degradation assessment," *Mechanical Systems and Signal Processing*, vol. 72, pp. 184–205, 2016.

[146] L. Saidi, J. B. Ali, E. Bechhoefer, and M. Benbouzid, "Wind turbine high-speed shaft bearings health prognosis through a spectral kurtosis-derived indices and svr," *Applied Acoustics*, vol. 120, pp. 1–8, 2017.

[147] S. Aye and P. Heyns, "An integrated gaussian process regression for prediction of remaining useful life of slow speed bearings based on acoustic emission," *Mechanical Systems and Signal Processing*, vol. 84, pp. 485–498, 2017.

[148] Y. Lu, Q. Li, Z. Pan, and S. Y. Liang, "Prognosis of bearing degradation using gradient variable forgetting factor rls combined with time series model," *IEEE Access*, vol. 6, pp. 10 986–10 995, 2018.

[149] M. Elforjani, S. Shanbr, and E. Bechhoefer, "Detection of faulty high speed wind turbine bearing using signal intensity estimator technique," *Wind Energy*, vol. 21, no. 1, pp. 53–69, 2018.

[150] W. Ahmad, S. A. Khan, and J.-M. Kim, "A hybrid prognostics technique for rolling element bearings using adaptive predictive models," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 2, pp. 1577–1584, 2018.

[151] L. Zhao and X. Wang, "A deep feature optimization fusion method for extracting bearing degradation features," *IEEE Access*, vol. 6, pp. 19 640–19 653, 2018.

[152] J. Wang, Y. Liang, Y. Zheng, R. X. Gao, and F. Zhang, "An integrated fault diagnosis and prognosis approach for predictive maintenance of

wind turbine bearing with limited samples," *Renewable Energy*, vol. 145, pp. 642–650, 2020.

[153] M. Elforjani and S. Shanbr, "Prognosis of bearing acoustic emission signals using supervised machine learning," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5864–5871, 2018.

[154] M. Qiu, W. Li, F. Jiang, and Z. Zhu, "Remaining useful life estimation for rolling bearing with sios-based indicator and particle filtering," *IEEE Access*, vol. 6, pp. 24 521–24 532, 2018.

[155] M. Rezamand, M. Kordestani, M. E. Orchard, R. Carriveau, D. S. Ting, and M. Saif, "Improved remaining useful life estimation of wind turbine drivetrain bearings under varying operating conditions (voc)," *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2020.

[156] A. Rai and S. Upadhyay, "The use of md-cumsum and narx neural network for anticipating the remaining useful life of bearings," *Measurement*, vol. 111, pp. 397–410, 2017.

[157] S. A. Billings, *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. John Wiley & Sons, 2013.

[158] R. De Maesschalck, D. Jouan-Rimbaud, and D. L. Massart, "The mahalanobis distance," *Chemometrics and intelligent laboratory systems*, vol. 50, no. 1, pp. 1–18, 2000.

[159] J. Lin and Q. Chen, "Fault diagnosis of rolling bearings based on multifractal detrended fluctuation analysis and mahalanobis distance criterion," *Mechanical Systems and Signal Processing*, vol. 38, no. 2, pp. 515–533, 2013.

[160] D. M. Hawkins and D. H. Olwell, *Cumulative sum charts and charting for quality improvement*. Springer Science & Business Media, 2012.

[161] Y. Hu, H. Li, P. Shi, Z. Chai, K. Wang, X. Xie, and Z. Chen, "A prediction method for the real-time remaining useful life of wind turbine bearings based on the wiener process," *Renewable energy*, vol. 127, pp. 452–460, 2018.

[162] M. Hemmer, H. Van Khang, K. Robbersmyr, T. Waag, and T. Meyer, "Fault classification of axial and radial roller bearings using transfer learning through a pretrained convolutional neural network," *Designs*, vol. 2, no. 4, p. 56, 2018.

[163] A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE transactions on Information Theory*, vol. 13, no. 2, pp. 260–269, 1967.

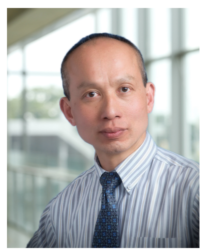
[164] D. A. Pola, H. F. Navarrete, M. E. Orchard, R. S. Rabié, M. A. Cerda, B. E. Olivares, J. F. Silva, P. A. Espinoza, and A. Pérez, "Particle-filtering-based discharge time prognosis for lithium-ion batteries with a statistical characterization of use profiles," *IEEE Transactions on Reliability*, vol. 64, no. 2, pp. 710–720, 2015.

[165] F. Tamssaouet, K. T. Nguyen, K. Medjaher, and M. E. Orchard, "Degradation modeling and uncertainty quantification for system-level prognostics," *IEEE Systems Journal*, 2020.

Milad Rezamand received his Bachelor's degree in Mechanical Engineering from Bahonar University of Kerman, Iran and Master's degree in Aerospace engineering from K.N.T University of Technology, Tehran, Iran in 2008 and 2012, respectively. His research interests include fault diagnostics, prognosis, and predicting remaining useful life. He received his Ph.D. degree in Mechanical Engineering from University of Windsor, Canada in 2019. He is currently a post-doctoral fellow in University of Windsor, Canada. His ongoing research includes estimating remaining useful life of wind turbine components.

Mojtaba Kordestani (S'09-GS'15-SM'19) received his Bachelor's degree in electronic engineering and Master's degree in control engineering from Malek Ashtar University and Tehran Azad University, Iran, in 2002 and 2008, respectively. He received his Ph.D. in electrical engineering from University of Windsor, Canada in 2018. He is a Postdoctoral fellow in University of Windsor. His research interests are in control, estimation, fault diagnostics, and failure prognosis. He has published about 40 refereed journal and conference papers in these areas.

Rupp Carriveau received the bachelor's degree in Civil Structural Engineering from the University of Windsor. He studied Fluids Engineering for both his MASc and PhD degrees at the Western University. He is a Director in the Turbulence and Energy Laboratory and the Centre for Energy and Water Advancement at the University of Windsor. Dr. Carriveau serves on the Editorial Boards of Wind Engineering, Advances in Energy Research, and the International Journal of Sustainable Energy. He is the current President of the Underwater Energy Storage Society.



David S-K. Ting obtained his BASc degree from the University of Manitoba. He studied Turbulent Combustion for both his MASc and PhD degrees at the University of Alberta. Dr. Ting explored Fluid-Structure Interactions at McGill University prior to joining the University of Windsor as a faculty member in the Department of Mechanical, Automotive and Materials Engineering. Professor Ting is the founder of the Turbulence & Energy Laboratory. He has cosupervised over seventy-five graduate students, coauthored more than one hundred and forty journal papers, authored four textbooks, and coedited eight books.

Marcos E. Orchard Marcos E. Orchard is Professor with the Department of Electrical Engineering at Universidad de Chile and Associate Researcher at the Advanced Center for Electrical and Electronic Engineering (UTFSM). He received the Ph.D. and M.S. degrees in Electrical and Computer Engineering from The Georgia Institute of Technology, Atlanta, GA, USA, in 2005 and 2007 respectively. He has authored and co-authored more than 100 papers. His current research interests include failure prognosis, electromobility, mining industry, and finance. Dr. Orchard is a Fellow of the Prognostic and Health Management Society.

Mehrdad Saif (S'84-M'84-SM'05) received B.S. in 1982, M.S. in 1984, and DEng in 1987, all in Electrical Engineering from Cleveland State University, Ohio, United States. He has been the Dean of the Faculty of Engineering at the University of Windsor since July 2011. Dr. Saif's research interests are in systems and control; estimation and observer theory; model based fault diagnostics; condition monitoring, diagnostics & prognostic; and application of these areas to automotive, power, autonomous systems, and other complex engineering systems. He has published over 350 refereed journal and conference papers plus an edited book in these areas.