Critical Wind Turbine Components Prognostics: A Comprehensive Review

Milad Rezamand[ID], Mojtaba Kordestani[ID], *Senior Member, IEEE*, Rupp Carriveau[ID], *Member, IEEE*,
David S.-K. Ting[ID], Marcos E. Orchard[ID], *Member, IEEE*, and Mehrdad Saif[ID], *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, and so on. 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 article provides a recent literature review on modeling developments for the RUL 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 the Bayesian application in the development of WT prognosis methods.

Index Terms—Bearings, blade, gearbox, generator, prognosis, wind turbines (WTs).

NOMENCLATURE

AE	Absolute error.
AI	Artificial intelligence.
ANFIS	Adaptive neurofuzzy inference systems.
ANN	Artificial neural network.

Milad Rezamand, Rupp Carriveau, and David S.-K. Ting are with the Turbulence and Energy Laboratory, University of Windsor, Windsor, ON N9B 3P4, Canada (e-mail: rezaman@uwindsor.ca; rupp@uwindsor.ca; dting@uwindsor.ca).

Mojtaba Kordestani and Mehrdad Saif are with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B 3P4, Canada (e-mail: kordest@uwindsor.ca; msaif@uwindsor.ca).

Marcos E. Orchard is with the Department of Electrical Engineering, University of Chile, Santiago 1025000, Chile (e-mail: morchard@ing.uchile.cl).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIM.2020.3030165

AOI	Anomaly operation index.
ARA	Average relative accuracy.
ARF	Aggregate reliability function.
ARMA	Autoregressive moving average.
ARIMA	Autoregressive integrated moving average.
BA	Bat algorithm.
BFO	Bacterial foraging optimization.
BP	Backpropagation.
CF	Crest Factor.
CM	Condition monitoring.
CMS	CM system.
CNN	Convolutional neural network.
CUMSUM	Cumulative sum.
CV	Confidence value.
DBN	Deep belief network.
DL	Deep learning.
DNN	Deep neural network.
DOSs	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 AE.
MAE	Mean AE.
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	Nonlinear rank regression.
NN	Neural network.
O&M	Operation and maintenance.
PCA	Principal component analysis.
pdf	Probability density function.
PF	Particle filter.
PHM	Prognostics and health management.
PM	Performance monitoring.
PNN	Pseudonearest neighbor.
PSO	Particle swarm optimization.
RLS	Recursive least-square.
RMLP	Recurrent MLP.
rms	Root mean square.
RNN	Recurrent NN.
RMSE	Root mse.
RQ	Rational quadratic.
RT	Regression trees.
RUL	Remaining useful life.
SA	Sparse autoencoder.
SCADA	Supervisory control and data acquisition.
SE	Squared exponential.
SKF	Switching KF.
SK	Spectral kurtosis.
SHMM	Semihidden 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 winding.
TDNN	Time delay NN.
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 NN.
WT	Wind turbine.

I. INTRODUCTION

WIND energy is playing an increasingly pivotal role in global energy systems. According to the 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 and 2024. The wind is also assuming a nascent and growing role in the expanding ancillary services market associated with evolving grids.

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. Thus, 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 will be reviewed in the following. Based on 350-WT operation over five years throughout Europe, Carroll *et al.* [2] revealed that the highest failure rates are related to generators, gearboxes, and blades. Shafiee and Dinmohammadi [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 15 000 WT years. In [6], a data set from 72 operating WTs 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 WT 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 PHM into consideration for the 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 practices for WT health state assessment are PM and CM. PM employs the 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 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. CMSs 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 CMSs [11]–[15].

Noncontact monitoring, an example of CMSs, contributes toward WT PHM using various techniques, such as ultrasound and vibration analyses. Noncontact monitoring can lead to more efficient measurement and, consequently, more accurate WT RUL prediction compared with the contact monitoring since contact measurement can change the vibration characteristics of the object to be measured. Furthermore, noncontact 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 WT [16]–[22].

After data are provided at regular time intervals by sensors and measurement systems, signal processing is performed through signal transformations and features extraction, reduction, selection, and so on. The next step of WT PHM is to determine the component state-of-health by detecting and localizing a fault based on preset 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 Fig. 1, RUL prediction made at t_0 utilizes a mathematical model of the failure fit to the 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].

1) *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.
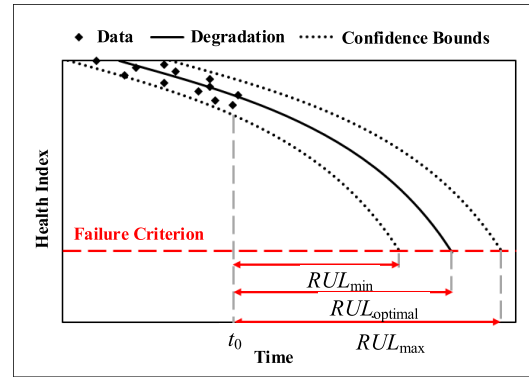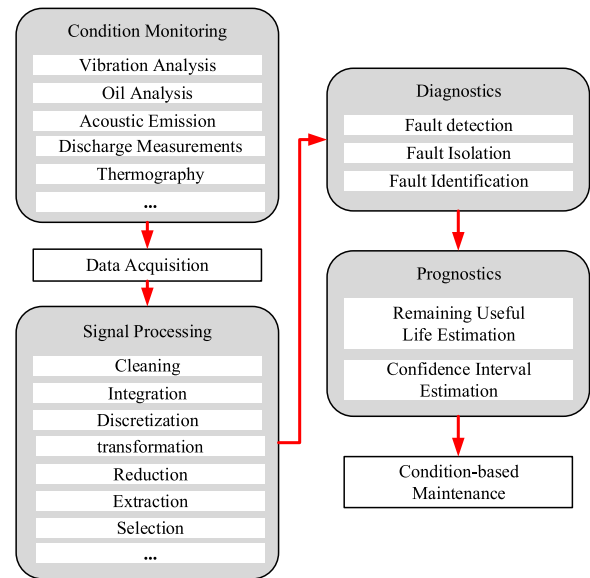

Fig. 1. Component health degradation curve.

Fig. 2. WT PHM architecture.

2) *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.

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 (CBM) [27], [28]:

1) Corrective maintenance is conducted after a failure occurs to restore the component to a condition to be able to perform its required purpose.
2) Preventive maintenance is carried out preventively on a scheduled time interval without any prior knowledge of the time to failures.
3) CBM is performed before an occurrence of a failure based on the condition of the component. As shown in Fig. 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].

TABLE I
PM VERSUS CM

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

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 article 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.

1) This article corresponds to the first comprehensive survey on WT prognostics that gathers, analyses, and separately classifies common failures for various parts of WTs. In each component, failures can be initiated by a combination of priory factors. These factors are further analyzed to determine the roots of any failures.

2) A review of the most recent literature on WT prognostics in component and subcomponent levels is presented. The 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 the WT prognostic field. Section V concludes this article with an emphasis on future research challenges.

II. CRITICAL WT COMPONENTS FAILURES

WTs have assembled integrations of several components, subjected to various failure modes due to harsh environmental and operating conditions. Fig. 3 displays the main components of WTs and some corresponding bearings. In this research work, we review the 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
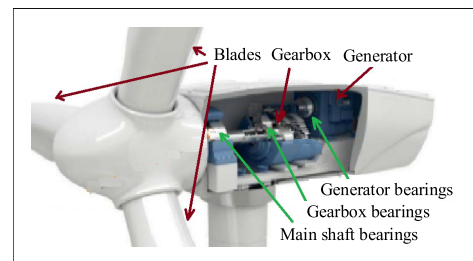

Fig. 3. Main components of WTs and some corresponding bearings.

Fig. 4. Gearbox planetary tooth failure [47].

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 micropitting [45], [46]. Fig. 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:

1) *Fluting (a Typical Bearing Fault in Electric Drivetrain) [48]:* It is caused by the prolonged passage of relatively small electric current, usually due to current leakage.

2) *SW Insulation Breakdown [49]:* It is caused by mechanical stress, contamination, electrical, and thermal stress. Fig. 5 shows a visual inspection of a WT generator SW insulation fault, occurred in a wind farm located in Southwestern Ontario in 2015 [50].

3) *Rotor Electrical Asymmetry [51]:* It is caused by the rising resistance or open circuit of the brush-gear circuits.
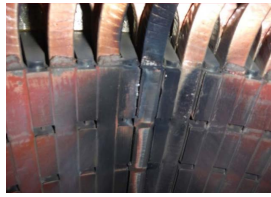
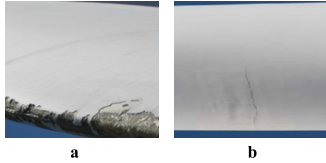Fig. 5.   Generator SW insulation fault [50].

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

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, fiber 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]. Fig. 6 displays blade erosion and crack, investigated, analyzed, and detected by Rezamand *et al.* [56].

D. Bearing Failures

WT 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 WTs 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 becomes 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 of a localized defect is a pit or spall. According to which component of
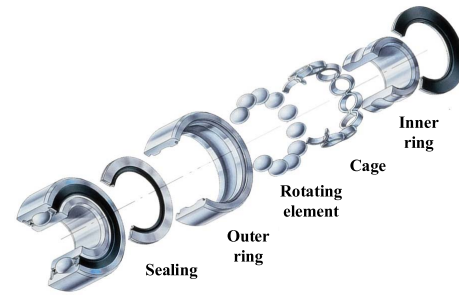

Fig. 7.   Rolling bearing subcomponents.

Fig. 8.   Rolling element surface wear [60].

the bearing affected, as shown in Fig. 7, the single point defects can be categorized into the 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, first, system and subsystem configurations, 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 FEA [68], the 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], first, identified different failure modes, their causes, and the damaging operating conditions of WT 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 gear 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 [see (5)] and degradation trend analysis. Although the proposed method was not able to detect misalignment continuously, its robustness was shown, predicting a 20-year lead time to gear fault due to bending stresses.

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 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 RUL 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 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 aeroelastic 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 5-MW turbine with a rated wind speed of 11.4 m/s were leveraged for optimized maintenance scheduling.

Bechhoefer and Schlanbusch [79], first, applied bearing envelope analysis to generate condition indicators, related to bearing pass frequencies, and then, the condition indicators

---

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.

---

were mapped to HIs. Finally, a simplification of Paris' law and KFs was employed to forecast the WT bearing RUL. HIs have been applied as inputs to compute bearing RUL. The results showed the efficacy of the proposed method. A mathematical model, (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 the enhance WT main bearing RUL projection. Results revealed strong evidence of failure with a 30-day lead time

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

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 the 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 ((7)). Experimental results on three life-cycle bearings from WT high-speed shaft indicated the effectiveness of the proposed method

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

where x_j describes the bearing state at the (j)th step, b_j denotes the model coefficient, and z_j is the measurement, which can be the HI 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 ANNs, DL, and 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 NN is comprised of a single input layer, one or more hidden layers, and an output layer, each including one or more nodes. Connections between the 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 nonlinear 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]. 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 that project an output

variable based on labeled inputs and unsupervised learning techniques that 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 ML [15].

1) 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, and cross-validation.

2) 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 with the predicted healthy state, and if a deviation is observed for multiple consecutive time intervals, then 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 DNNs, DBNs, RNNs, and CNNs [87]–[89]. The ANFIS method, a combination of FL and 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 WT 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 WT 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 the SCADA data. The automated APK-ANFIS was able to accurately determine the WT pitch RUL within a prognostic horizon of up 21 days with an optimal threshold and a window size of 0.3 and 6, respectively. In another study, Matthews *et al.* [95] indicated that ANFIS outperforms other AI techniques, such as K-means clustering, fuzzy inference system (FIS), ANN, and SOM in WT pitch RUL forecast.

Pan *et al.* [96] proposed an ELM optimized by an FOA for WT gearbox RUL forecast. FOA-ELM predicted model

TABLE II
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

was trained on extracted HIs 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 with PSO-ELM, BA-ELM, GA-ELM, and BFO-ELM.

2) AI-Based Prognostic Methods for WT Bearings: Artificial intelligent methods are widely applied for CM 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 WT gearbox was proposed

by Teng *et al.* [99]. They took an ANN to train data-driven models and 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 fit. By determining the intersection between the fit curve and the predefined 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 ESN and RMLP, based on the vibration signal of rotating machinery. Both ESN and RMLP are functional forms of an RNN. 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 the RUL prediction of bearings. The performance of the RNN-HI was validated through two experimental bearing data sets. The results indicated the ability of RNN-HI to obtain better performance than a self-organization map-based method.

A study of the WNN classifier bearing fault diagnosis was presented by Karim and Abderrazak [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 ENN architecture for RUL prediction of an 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 autoencoder-based DNN was employed to classify the acquired signals into different degradation states. Then, regression models based on shallow NNs were constructed for each health state. The proposed approach obtained satisfactory prediction performance on a real bearing degradation data set with different working conditions.

Li et al. [105] proposed an intelligent RUL prediction method based on DL. Multiscale feature extraction was executed using the CNN method. It was shown that related network structures without multiscale 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 were demanded to initialize and train the proposed DNN. Deutsch and He [89] developed a DL-based method through the combination of a DBN and an 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 HGRUN was proposed by Li et al. [106] for rolling bearing health prognosis. The HGRUN was formed by stacking various hidden layers. The open experimental bearing data were practiced to validate the capacity of the proposed approach. The results proved that HGRUN outperforms the other techniques, including BP NN, 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 with RNNs. Cheng et al. [108] introduced a case-based data-driven prognostic framework using the ANFIS. First, large historical data were 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 with the traditional data-driven systems. Soualhi et al. [109] proposed a time-series forecasting model, NFN, to predict the degradation of bearings. The NFN is an intelligent tool that contributes to modeling complex systems by the simplicity of its structure. The 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 with 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 the stochastic prognosis of critical WT components and bearings are reviewed.

ARFs employ a pdf to determine the times to failure of a population of machine component/failure modes [61]. The Bayesian networks are a type of probabilistic open-chain graphical model for estimating probabilities [110]. The Bayesian network is comprised of nodes that correspond to random variables that can take on distinct states. These are connected by directional arcs representing conditional dependencies between nodes [111]. The 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 the Markov models, KFs, and PFs.

The 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].

1) Transition probabilities are independent of time (i.e., a constant failure rate).

TABLE III
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 $= 0.000428$

2) The waiting time in a distinct state has an exponential trend.
3) The sum of all transition probabilities for leaving one state and entering different states must be equal to one.

On the other hand, the 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 semihidden Markov variants.

HMM and 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 the time-series data can be analyzed without a physical understanding of the failure, so long as enough data are 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 nonstandard. While 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 the 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 NLRR outperforms 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 1.5-MW 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 ML method called 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 with the WNN. Since covariance function is the key factor GPR properties controls, three covariance functions, including SE, 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 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, the prognosis was further implemented based on an HHMM. Their research work indicated that the accuracy of the method depended on the sample size of historical data. In [120], a multisensor HSMM was proposed, which is an extension of classical HSMMs. Experimental results revealed that the prognostic method was promising to achieve more reliable performance than classical HSMMs.

Singleton *et al.* [121] applied an 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, were applied to the proposed algorithm. It showed the convergence of the algorithm in different conditions. Lim and Mba [122] introduced SKF for the FDP 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

a 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 the 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 the 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 techniques and PF for RUL prediction of bearing, was proposed by Wang and Gao [124]. For this aim, a set of PF 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 with traditional PF. Chen *et al.* [125] presented a generic PF-based framework with application in bearing spalling FDP. 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, Multisensor HSMM, EKF, SKF, PF, and Generic PF. The details are presented in Table IV.

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, the 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, AOI, to estimate WT performance deterioration and predict WT generator RUL. In this regard, Zhao *et al.* [128] 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 WT generator RUL forecast. The experimental results illustrated that the presented

TABLE IV
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}

scheme could provide the time to failure of WT 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 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 the Bayesian inference. The case studies demonstrated the efficacy of the introduced varying load approach and its benefits compared with 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 ARA in RUL prediction compared with the Choquet integral fusion approach and the Bayesian algorithm obtained from single-feature driven methods.

In [131], a combination of VM, LR, and ARMA/generalized autoregressive conditional heteroscedastic (GARCH) models was proposed to assess bearing failure degradation, as shown in Fig. 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, PCA was employed for feature extraction from a vibration signal, and the SVM parameters
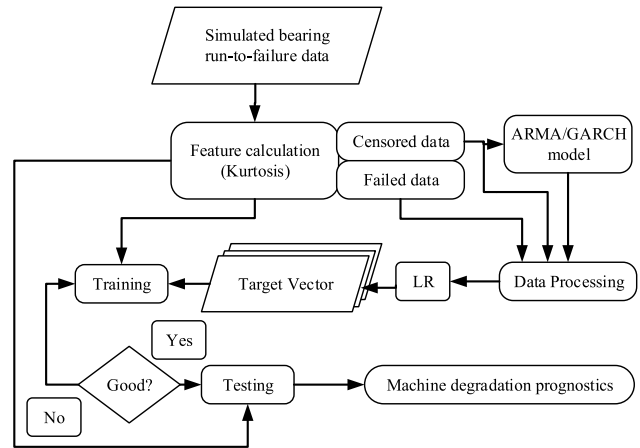

Fig. 9. Combination of VM, LR and ARMA GARCH for bearings RUL estimation [131].

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 (14). 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

$$L_{10} = \frac{10^6}{60n} \left(\frac{C}{P} \right)^p. \tag{14}$$

Dong and Luo [133] developed an approach to determine bearing degradation based on a combination of PCA and an optimized LSSVM method, as shown in Fig. 10. First, 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 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 LSSVR, and the HMM was proposed

---

Algorithm 2 Bayesian-Based RUL Prediction [60]

1) A healthy data set is set: $x^{(1)}, x^{(2)}, \ldots, x^{(w-1)}$

2) A sliding window data as a training set with the degradation trend is selected: $x^{(w)}, x^{(w+1)}, \ldots, x^{(z)}$

3) A failure criterion, FC is set using the mean (μ), the standard deviation (σ) of healthy data set, and the coefficient of complete failure criteria (λ_q)

$$FC = \mu + \lambda \sigma \qquad (8)$$

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 $z - w$:

$$y_t = \hat{m}t + \hat{n} + e_t \qquad (9)$$

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 $t_0 + k$:

$$p(F_{t_0+k}) = p(F_{t_0+k}|H_{t_0:t_0+k-1})p(H_{t_0:t_0+k-1}) \qquad (10)$$

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

$$p(F_{t_0+k}|H_{t_0:t_0+k-1}) = Q(\frac{FC - y_{t_0+k}}{\sigma\sqrt{k+1}}) \qquad (11)$$

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):

$$p(H_{t_0:t_0+k-1}) = [1 - p(F_{t_0+1}|H_{t_0})]$$
$$\times \cdots \times [1 - p(F_{t_0+k-1}|H_{t_0:t_0+k-2})] \qquad (12)$$

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):

$$RUL = t_{Failure} - t_0 \qquad (13)$$

Note that based on the 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.
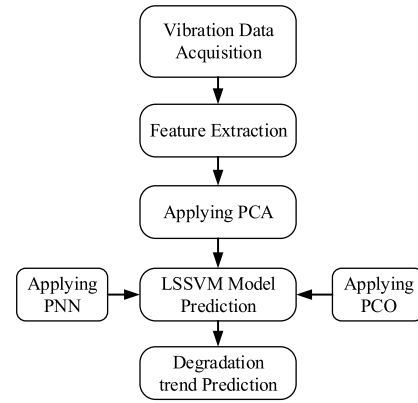

Fig. 10.   Combination of PCA and LSSVM for bearings RUL estimation [133].
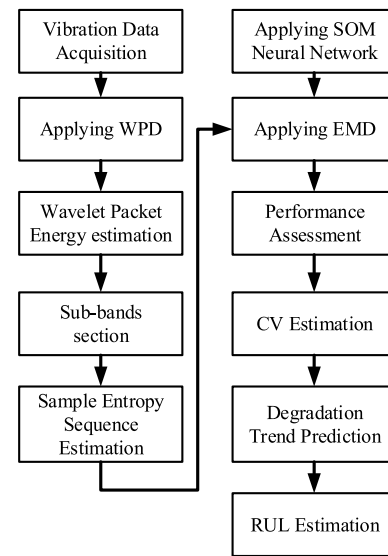
Fig. 11.   Combination of WPD, EMD and SOM NN techniques for bearings RUL estimation [135].

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 WPD, EMD, and SOM NN techniques, as shown in Fig. 11 for assessing

the state of the bearing's degradation and estimating the RUL. An HI named 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 and RUL. The proposed hybrid approach enhanced accuracy when entering the severely degraded stage compared with 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) NN and WD

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.

$$\hat{x}_{j+p} = f(x_j, x_{j-q}, x_{j-2q}, x_{j-3q}, \ldots, x_{j-(n-1)q})$$
$$= f(X_j) \tag{15}$$

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

$$f(X_t) = \Omega = w^T . \Phi(X_t) + b \tag{16}$$

where

$$w = \sum_{i=1}^{n-1} (\alpha_i^+ - \alpha_i^-)\Phi(x_i) \tag{17}$$

$$b = \Omega_s - \varepsilon - \sum_{m \in s} (\alpha_m^+ - \alpha_m^-)\Phi(x_m) \cdot \Phi(x_s) \tag{18}$$

where s is the set of indices of the support vectors, Ω_s presents the class that represents the degradation state of the bearing, $\Phi : x \rightarrow \Phi(x)$ 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}

$$\underset{\text{estimated}}{\text{RUL}} \leftarrow \min(\text{RUL}_1, \text{RUL}_2, \text{RUL}_3, \ldots) \tag{19}$$

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

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 HHT to extract feature indexes from raw vibration signals, an SVM to detect the degradation states, and the SVR to estimate the RUL of ball bearings. The experimental results confirmed that the use of the HHT, SVM, and 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 EM algorithm. The results confirmed that their proposed approach could provide
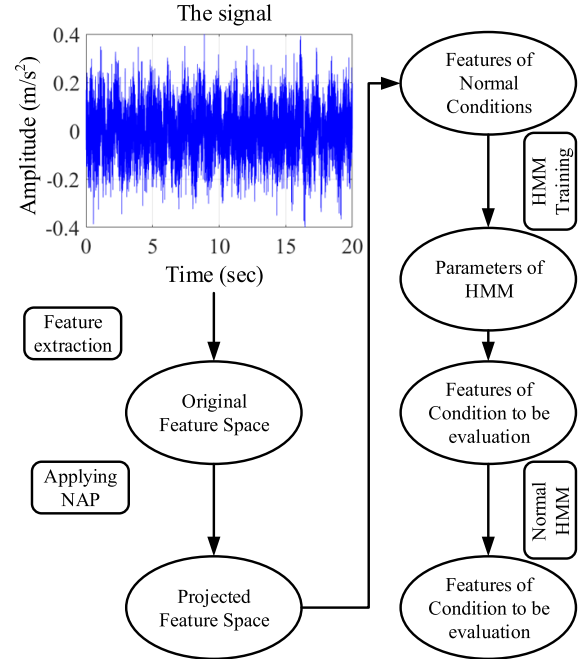
higher estimation accuracy and narrower PDFs in comparison with Gebraeel's model [141] and Si's model [142].

Zhao *et al.* [143] presented a feature extraction system for vibration-based bearing prognosis using TFR and supervised dimensionality reduction. A combination of TFR, GP, and LBP was used to evaluate lifetime information represented by highly dimensional features. The RULs are determined by employing simple MLRs. The experimental results demonstrated that the proposed method outperforms techniques employing traditional statistical features and PCA. In [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 bearing performance degradation using a combination of HMM and NAP, as shown in Fig. 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 an 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 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 with an ARIMA model without adaptation. Results confirmed accurate predictions of the hybrid prognostic method over the ARIMA model. Elforjani et al. [149] proposed SIE as a new indicator to detect individual types of early fault in real-world WT bearings. This study indicated the ability of the proposed indicator to accurately estimate the RUL for WT bearings in a combination of RT and multilayer ANN models. The experimental results demonstrated that SIE has an advantage over the other fault indicators, such as 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 TSP accurately using linear regression analysis, which contributes to relatively more accurate RUL predictions. A deep feature optimization fusion method was proposed by Zhao and Wang [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 data sets 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 the Bayesian framework. For this aim, first, an empirical model for the spalling evolution based on Paris' formula, as shown in (20), was developed. Then, PF was developed as a recursive numerical approach based on the sequential Monte Carlo sampling technique to estimate the posterior pdf of the state, as indicated in (21). The experimental results confirmed that the proposed method was capable of inferring the hidden defect state of the bearing from noisy measurement based on the Bayesian inference and quantifying the uncertainty of the RUL prognosis in a probabilistic manner

$$x_{t+1} = \left[x_t^{(1-m)} + c(1-m)\right]^{\frac{1}{1-m}} + u_t \tag{20}$$

$$\begin{cases} \text{Prediction: } p(x_{t+1}|z_t) = \sum_{j=1}^{M} w_t^j \, p\left(x_{t+1}\big|x_t^j\right) \\ \text{Update: } w_{t+1}^j = w_t^j \, p\left(z_{t+1}\big|x_{t+1}^j\right) \end{cases} \tag{21}$$

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 the 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 NNs model with a BP learning algorithm 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 were provided. Qiu et al. [154] presented a prognostic procedure by combining an HI and PF to determine the bearing RUL. The process included applying the SIOS algorithm to build the HI called 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 DPM is determined for each state based on extracted features from 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 and Upadhyay [156] introduced a data-driven prognosis approach based on an NARX NN 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 Fig. 13.

1) 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.
2) To address the highly nonmonotonic behavior of the extracted features due to the bearing degradation, an HI based on MD criterion [158], [159] and CUMSUM chart [160] was introduced.

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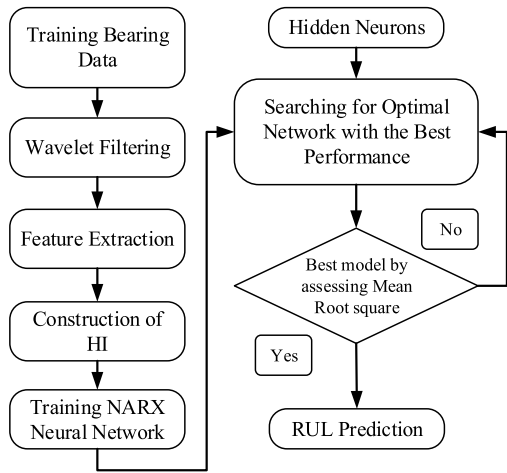
Fig. 13. Flowchart for prognosis approach based on a NARX NN model in association with a wavelet-filter technique for bearing RUL estimation [156].

3) The NARX NN was developed as a 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 SOM-based indicator. Hu *et al.* [161] presented a real-time performance degradation model based on temperature characteristic parameters for prognosis of WT bearings. Here, a combination of the Wiener process for establishing the performance degradation model, the MLE 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 SA-based SVM utilizing transfer learning. The effectiveness of the proposed technique was examined employing vibration and acoustic emission signal data sets 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 SA 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 with individually employed prognosis methods. For instance, ANFIS, a combination of FL and NNs, which 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. PROBLEM OF PROGNOSTICS IN CRITICAL WT COMPONENTS

To provide a brief idea of the practical applications related to the 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.

1) First, SCADA data are 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" DOSs (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].

2) Then, different prognostic techniques, such as physics- and regression-based methods, can be employed to identify a DPM for each state ("n" DPMs for "n" DOSs) utilizing CM measurements (e.g., vibration signals for bearings or temperature measurements for generators). The 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.

3) 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, as 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. CONCLUSION 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.

1) WT hybrid prognosis techniques are now the leading tools for critical WT components failure prediction because of their higher accuracy over individual prognosis methods.

2) WT physics-based prognostics provide the most accurate predictions with fewer data compared 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.

3) 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.

Algorithm 4 HMM and Viterbi

---

1) Let $q_t \in [s_1, s_2, \ldots, s_n]$ as the value of the hidden state at time t
where n is the number of hidden states,
$o_t \in [v_1, v_2, \ldots, v_m]$ 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:

$$\pi = (\pi_1, \pi_2, \ldots, \pi_n) \tag{22}$$

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

$$a_{ij} = P(q_t = s_i | q_{t-1} = s_j), \quad 1 \le i, \ j \le n \tag{23}$$

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, $B = (b_{jk})$:

$$b_{jk} = P(o_t = v_k | q_t = s_j), \quad 1 \le j \le n, \ 1 \le k \le m \tag{24}$$

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 $Q = [q_1, q_2, \ldots, q_t]$ using the system parameter $\lambda = (\pi, A, B)$ and the observation sequence $O = [o_1, o_2, \ldots, o_t]$ as follows:
– Initialization:

$$\delta_1(i) = \pi_i b_{io_1} \psi_1(i) = 0, \quad 1 \le i \le n \tag{25}$$

– Recursion:

$$\delta_{t+1}(j) = \max_{1 \le i \le n} [\delta_t(i) a_{ij}] b_{jo_{t+1}}$$
$$\psi_{t+1}(j) = \arg \max_{1 \le i \le n} [\delta_t(i) a_{ij}] b_{jo_{t+1}}$$
$$1 \le t \le T, \quad 1 \le j \le n \tag{26}$$

– Computing States Sequences (with retrospect):

$$P(Q, O|\lambda) = \max_{1 \le i \le n} \delta_T(i), \quad Q_{t-1} = \psi_t(Q_t), \ T \ge t \ge 1 \tag{27}$$

---

4) WT stochastic-based prognostics are robust in the RUL prediction of WT components due to their capability in modeling the uncertainty inherent in the prediction horizon of WT components.

5) AI-based techniques reach a higher RUL prediction accuracy than conventional prognosis methods, including NARX, HGRUN, FOA-ELM, ANFIS, NFN, ESN, and RNN.

6) Stochastic-based prognosis techniques are more accurate than conventional RUL estimation methods, including GPR, HHMM, Multisensor 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, such as wind speed and ambient temperature. It should be recognized that damage progression can be a function of the stress and loading applied to components that 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 categorizing DOSs in which a WT component undergoes. For this aim, SCADA data can be clustered using a robust clustering technique. Then, a probabilistic-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 focusing future research efforts on the implementation of the 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 with 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:* The Bayesian approaches intrinsically consider probability theory, which may be more appropriate for RUL prediction due 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 with data-driven methods. Hence, it is beneficial to apply additional emphasis on developing a mathematical model for each WT component failure mode.

ACKNOWLEDGMENT

REFERENCES

[1] J. Lee and F. Zhao, "Global wind report 2019," Global Wind Energy Council, Brussels, Belgium, Tech. Rep. 3, 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, Jun. 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, Feb. 2014.

[4] B. Hahn, M. Durstewitz, and K. Rohrig, "Reliability of wind turbine– experiences of 15 years with 1500 WTs," Institut für Solare Energiev-ersorgungstechnik (ISET), Hesse, Germany, Tech. Rep., 2005.

[5] P. J. Tavner, D. M. Greenwood, M. W. G. 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, Mar. 2013.

[6] A. Stenberg and H. Holttinen, "Analysing failure statistics of wind turbines in Finland," in *Proc. Eur. Wind Energy Conf.*, Apr. 2010, pp. 20–23.

[7] B. S. Nivedh, "Major failures in the wind turbine components and the importance of periodic inspections," Wind Services, UL India PVT LTD., Bengaluru, India, Tech. Rep., 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," *Renew. Sustain. Energy Rev.*, vol. 56, pp. 368–379, Apr. 2016.

[9] N. Chen, R. Yu, Y. Chen, and H. Xie, "Hierarchical method for wind turbine prognosis using SCADA data," *IET Renew. Power Gener.*, vol. 11, no. 4, pp. 403–410, Mar. 2017.

[10] N. M. Vichare and M. G. Pecht, "Prognostics and health management of electronics," *IEEE Trans. Compon. Packag. Technol.*, vol. 29, no. 1, pp. 222–229, Mar. 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," *Mech. Syst. Signal Process.*, vol. 120, pp. 392–407, Apr. 2019.

[12] S. Sheng and W. Yang, "Wind turbine drivetrain condition monitoring-an overview (presentation)," Nat. Renew. Energy Lab., Golden, CO, USA, Tech. Rep. NREL/PR-5000-58774, 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, May 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*. Cham, Switzerland: Springer, 2017, pp. 33–47.

[15] A. Stetco *et al.*, "Machine learning methods for wind turbine con- dition monitoring: A review," *Renew. Energy*, vol. 133, pp. 620–635, Apr. 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*, vol. 1. Cham, Switzerland: Springer, 2011, pp. 39–44.

[17] J. Baqersad, P. Poozesh, C. Niezrecki, and P. Avitabile, "A noncon-tacting approach for full-field strain monitoring of rotating structures," *J. Vib. Acoust.*, vol. 138, no. 3, Jun. 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, Dec. 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," *Mech. Syst. Signal Process.*, vol. 130, pp. 470–483, Sep. 2019.

[20] W. Luo, J. Li, X. Ma, and W. Wei, "A novel static deformation mea-surement and visualization method for wind turbine blades using home-made LiDAR and processing program," *Opt. Lasers Eng.*, vol. 134, Nov. 2020, Art. no. 106206.

[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," *Mech. Syst. Signal Process.*, vol. 138, Apr. 2020, Art. no. 106446.

[22] Y. Du, S. Zhou, X. Jing, Y. Peng, H. Wu, and N. Kwok, "Damage detection techniques for wind turbine blades: A review," *Mech. Syst. Signal Process.*, vol. 141, Jul. 2020, Art. no. 106445.

[23] A. Heng, S. Zhang, A. C. C. Tan, and J. Mathew, "Rotating machinery prognostics: State of the art, challenges and opportunities," *Mech. Syst. Signal Process.*, vol. 23, no. 3, pp. 724–739, Apr. 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," *Int. J. Hydrogen Energy*, vol. 38, no. 35, pp. 15307–15317, 2013.

[25] P. Tchakoua, R. Wamkeue, M. Ouhrouche, F. Slaoui-Hasnaoui, T. 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, Apr. 2014.

[26] F. P. G. Márquez, A. M. Tobias, J. M. P. Pérez, and M. Papaelias, "Condition monitoring of wind turbines: Techniques and methods," *Renew. Energy*, vol. 46, pp. 169–178, Oct. 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," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 1917–1925, Jan. 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, Jan. 2019.

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

[30] Y. Amirat, M. E. H. Benbouzid, B. Bensaker, and R. Wamkeue, "Condition monitoring and Ault diagnosis in wind energy conversion systems: A review," in *Proc. IEEE Int. Electr. Mach. Drives Conf.*, vol. 2, May 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 *Proc. 19th Int. Conf. Electr. Mach. (ICEM)*, Sep. 2010, pp. 1–6.

[32] H. Badihi, Y. Zhang, and H. Hong, "A review on application of moni-toring, diagnosis, and fault-tolerant control to wind turbines," in *Proc. Conf. Control Fault-Tolerant Syst. (SysTol)*, Oct. 2013, pp. 365–370.

[33] M. J. Kabir, A. M. T. Oo, and M. Rabbani, "A brief review on offshore wind turbine fault detection and recent development in condition monitoring based maintenance system," in *Proc. Australas. Universities Power Eng. Conf. (AUPEC)*, Sep. 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 Trans. Ind. Electron.*, vol. 62, no. 10, pp. 6536–6545, Oct. 2015.

[35] W. Qiao and D. Lu, "A survey on wind turbine condition monitoring and fault diagnosis—Part II: Signals and signal processing methods," *IEEE Trans. Ind. Electron.*, vol. 62, no. 10, pp. 6546–6557, Oct. 2015.

[36] J. Tautz-Weinert and S. J. Watson, "Using SCADA data for wind turbine condition monitoring—A review," *IET Renew. Power Gener.*, 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, Sep. 2017.

[38] C. Li, J. V. de Oliveira, M. Cerrada, D. Cabrera, R. V. Sanchez, and G. Zurita, "A systematic review of fuzzy formalisms for bearing fault diagnosis," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 7, pp. 1362–1382, Jul. 2019.

[39] H. Seyr and M. Muskulus, "Decision support models for operations and maintenance for offshore wind farms: A review," *Appl. Sci.*, vol. 9, no. 2, p. 278, Jan. 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. 29857–29881, 2020.

[41] N. Jammu and P. Kankar, "A review on prognosis of rolling element bearings," *Int. J. Eng. Sci. Technol.*, 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," *Mech. Syst. Signal Process.*, vol. 42, nos. 1–2, pp. 314–334, Jan. 2014.

[43] S. Baltazar, C. Li, H. Daniel, and J. Valente de Oliveira, "A review on neurocomputing based wind turbines fault diagnosis and prognosis," in *Proc. Prognostics Syst. Health Manage. Conf. (PHM-Chongqing)*, Oct. 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 *Proc. IEEE Power Electron. Mach. Wind Appl.*, Jun. 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," *Appl. Mech. Mater.*, vol. 197, pp. 206–210, Sep. 2012, doi: 10.4028/www.scientific.net/amm.197.206.

[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, Jul. 2013.

[47] I. Zuccolin, "KEC T78 gearbox planetary failure report," Kruger Energy Inc., Montréal, QC, Canada, 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," *Model., Identificat. Control*, vol. 33, no. 2, pp. 69–86, 2012, doi: 10.4173/mic.2012.2.4.

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

[50] A. Parr, "MSD quality system documentation, MSD initial repair report 042214," Kruger Energy Inc., Montréal, QC, Canada, Tech. Rep. 5, 2015.

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

[52] L. M. Popa, B. B. Jensen, E. Ritchie, and I. Boldea, "Condition monitoring of wind generators," in *Proc. 38th IAS Annu. Meeting Conf. Rec. Ind. Appl. Conf.*, vol. 3, 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 Trans. Sustain. Energy*, vol. 3, no. 3, pp. 432–444, Jul. 2012.

[54] M. Sundaresan, M. Schulz, and A. Ghoshal, "Structural health monitoring static test of a wind turbine blade," Nat. Renew. Energy Lab., Golden, CO, USA, Tech. Rep. NREL/SR-500-28719, 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," *Meas. Sci. Technol.*, vol. 19, no. 12, Dec. 2008, Art. no. 122001.

[56] M. Rezamand, M. Kordestani, R. Carriveau, D. S.-K. Ting, and M. Saif, "A new hybrid fault detection method for wind turbine blades using recursive PCA and wavelet-based PDF," *IEEE Sensors J.*, vol. 20, no. 4, pp. 2023–2033, Feb. 2020.

[57] S. Sankar, M. Nataraj, and P. V. Raja, "Failure analysis of bearing in wind turbine generator gearbox," *J. Inf. Syst. Commun.*, 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 Trans. Ind. Electron.*, vol. 55, no. 4, pp. 1813–1822, Apr. 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 Trans. Ind. Appl.*, vol. 40, no. 3, pp. 735–739, May 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 Trans. Mechatronics*, vol. 25, no. 3, pp. 1468–1478, Jun. 2020.

[61] J. Z. Sikorska, M. Hodkiewicz, and L. Ma, "Prognostic modelling options for remaining useful life estimation by industry," *Mech. Syst. Signal Process.*, vol. 25, no. 5, pp. 1803–1836, Jul. 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 J.*, vol. 18, no. 12, pp. 4990–5001, Jun. 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 Trans. Rel.*, vol. 68, no. 1, pp. 330–341, Mar. 2019.

[64] M. Kordestani, M. Saif, M. E. Orchard, R. Razavi-Far, and K. Khorasani, "Failure prognosis and applications—A survey of recent literature," *IEEE Trans. Rel.*, early access, Sep. 17, 2019, doi: 10.1109/TR.2019.2930195.

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

[66] C. A. A. Magniez, "Combining information flow and physics-of-failure in mechatronic products," Technische Univ. Eindhoven, Eindhoven, The Netherlands, Tech. Rep. 978-90-386-0925-6, 2007.

[67] Y. Li, T. R. Kurfess, and S. Y. Liang, "Stochastic prognostics for rolling element bearings," *Mech. Syst. Signal Process.*, vol. 14, no. 5, pp. 747–762, Sep. 2000.

[68] C. J. Li and H. Lee, "Gear fatigue crack prognosis using embedded model, gear dynamic model and fracture mechanics," *Mech. Syst. Signal Process.*, vol. 19, no. 4, pp. 836–846, Jul. 2005.

[69] C. H. Oppenheimer and K. A. Loparo, "Physically based diagnosis and prognosis of cracked rotor shafts," *Proc. SPIE*, vol. 4733, pp. 122–132, Jun. 2002.

[70] R. F. Orsagh, J. Sheldon, and C. J. Klenke, "Prognostics/diagnostics for gas turbine engine bearings," in *Proc. ASME Turbo Expo, Collocated Int. Joint Power Gener. Conf.*, 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 *Proc. IEEE Aerosp. Conf.*, Mar. 2006, p. 8.

[72] J. Qiu, B. B. Seth, S. Y. Liang, and C. Zhang, "Damage mechanics approach for bearing lifetime prognostics," *Mech. Syst. Signal Process.*, vol. 16, no. 5, pp. 817–829, Sep. 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, Aug. 2009.

[74] D. Breteler, C. Kaidis, T. Tinga, and R. Loendersloot, "Physics based methodology for wind turbine failure detection, diagnostics & prognostics," in *Proc. EWEA Annu. Event*, 2015, pp. 1–9.

[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 *Proc. IEEE Conf. Prognostics Health Manage. (PHM)*, Jun. 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, Jun. 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," *Int. J. Mater. Mech. Eng.*, vol. 3, no. 1, pp. 6–15, 2014.

[78] M. Florian and J. Sørensen, "Wind turbine blade life-time assessment model for preventive planning of operation and maintenance," *J. Mar. Sci. Eng.*, vol. 3, no. 3, pp. 1027–1040, Sep. 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 *Proc. IEEE Int. Conf. Control Appl.*, Oct. 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. 47133–47143, 2020.

[82] Y. Lei, *Intelligent Fault Diagnosis and Remaining Useful Life Prediction of Rotating Machinery*. Oxford, U.K.: Butterworth-Heinemann, 2016.

[83] M. Morshedizadeh, M. Kordestani, R. Carriveau, D. S.-K. Ting, and M. Saif, "Improved power curve monitoring of wind turbines," *Wind Eng.*, vol. 41, no. 4, pp. 260–271, Aug. 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 Renew. Power Gener.*, vol. 12, no. 9, pp. 1025–1033, Jul. 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 Convers. Manage.*, vol. 134, pp. 361–372, Feb. 2017.

[86] M. Carolin Mabel and E. Fernandez, "Analysis of wind power generation and prediction using ANN: A case study," *Renew. Energy*, vol. 33, no. 5, pp. 986–992, May 2008.

[87] J. S. Lal Senanayaka, H. Van Khang, and K. G. Robbersmyr, "Autoencoders and recurrent neural networks based algorithm for prognosis of bearing life," in *Proc. 21st Int. Conf. Electr. Mach. Syst. (ICEMS)*, Oct. 2018, pp. 537–542.

[88] Q. Cui, Z. Li, J. Yang, and B. Liang, "Rolling bearing fault prognosis using recurrent neural network," in *Proc. 29th Chin. Control Decis. Conf. (CCDC)*, May 2017, pp. 1196–1201.

[89] J. Deutsch and D. He, "Using deep learning-based approach to predict remaining useful life of rotating components," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 48, no. 1, pp. 11–20, Jan. 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, Nov. 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, Nov. 2016.

[92] D. Zurita, J. A. Carino, M. Delgado, and J. A. Ortega, "Distributed neuro-fuzzy feature forecasting approach for condition monitoring," in *Proc. IEEE Emerg. Technol. Factory Autom. (ETFA)*, Sep. 2014, pp. 1–8.

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

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

[95] B. Chen, P. C. Matthews, and P. J. Tavner, "Automated on-line fault prognosis for wind turbine pitch systems using supervisory control and data acquisition," *IET Renew. Power Gener.*, vol. 9, no. 5, pp. 503–513, Jul. 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 Mach. Theory*, vol. 137, pp. 509–526, Jul. 2019.

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

[98] A. Malhi, R. Yan, and R. X. Gao, "Prognosis of defect propagation based on recurrent neural networks," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 3, pp. 703–711, Mar. 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, Dec. 2016.

[100] Y. Xie and T. Zhang, "The application of echo state network and recurrent multilayer perceptron in rotating machinery fault prognosis," in *Proc. IEEE Chin. Guid., Navigat. Control Conf. (CGNCC)*, Aug. 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, May 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," in *Proc. ICEE-B*, Oct. 2017, pp. 1–6.

[103] S. E. Kramti, J. Ben Ali, L. Saidi, M. Sayadi, and E. Bechhoefer, "Direct wind turbine drivetrain prognosis approach using elman neural network," in *Proc. 5th Int. Conf. Control, Decis. Inf. Technol. (CoDIT)*, Apr. 2018, pp. 859–864.

[104] M. Xia, T. Li, T. Shu, J. Wan, C. W. de Silva, and Z. Wang, "A two-stage approach for the remaining useful life prediction of bearings using deep neural networks," *IEEE Trans. Ind. Informat.*, vol. 15, no. 6, pp. 3703–3711, Jun. 2019.

[105] X. Li, W. Zhang, and Q. Ding, "Deep learning-based remaining useful life estimation of bearings using multi-scale feature extraction," *Rel. Eng. Syst. Saf.*, vol. 182, pp. 208–218, Feb. 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 Mach. Theory*, vol. 133, pp. 229–249, Mar. 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 Trans. Ind. Electron.*, vol. 58, no. 9, pp. 4353–4364, Sep. 2011.

[108] F. Cheng, L. Qu, and W. Qiao, "A case-based data-driven prediction framework for machine fault prognostics," in *Proc. IEEE Energy Convers. Congr. Expo. (ECCE)*, Sep. 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 *Proc. 9th IEEE Int. Symp. Diag. Electr. Mach., Power Electron. Drives (SDEMPED)*, Aug. 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 Trans. Ind. Informat.*, early access, Mar. 12, 2018.

[111] S. Dey and J. A. Stori, "A Bayesian network approach to root cause diagnosis of process variations," *Int. J. Mach. Tools Manuf.*, vol. 45, no. 1, pp. 75–91, Jan. 2005.

[112] M. Rausand and A. Høyland, *System Reliability Theory: Models, Statistical Methods, and Applications*, vol. 396. Hoboken, NJ, USA: Wiley, 2004.

[113] P. O'Connor and A. Kleyner, *Practical Reliability Engineering*. Hoboken, NJ, USA: Wiley, 2012.

[114] Q. Miao and V. Makis, "Condition monitoring and classification of rotating machinery using wavelets and hidden Markov models," *Mech. Syst. Signal Process.*, vol. 21, no. 2, pp. 840–855, Feb. 2007.

[115] M. Rezamand, R. Carriveau, D. S.-K. Ting, M. Davison, and J. J. Davis, "Aggregate reliability analysis of wind turbine generators," *IET Renew. Power Gener.*, vol. 13, no. 11, pp. 1902–1910, Aug. 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 *Proc. Int. Conf. Chem., Mater. Food Eng.* Paris, France: Atlantis Press, 2015, pp. 198–200.

[117] S. Hong and Z. Zhou, "Application of Gaussian process regression for bearing degradation assessment," in *Proc. 6th Int. Conf. New Trends Inf. Sci., Service Sci. Data Mining (ISSDM)*, 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," *Mech. Syst. Signal Process.*, vol. 134, Dec. 2019, Art. no. 106302.

[119] X.-H. Zhang and J.-S. Kang, "Hidden Markov models in bearing fault diagnosis and prognosis," in *Proc. 2nd Int. Conf. Comput. Intell. Natural Comput.*, Sep. 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," *Proc. Inst. Mech. Eng., C, J. Mech. Eng. Sci.*, vol. 227, no. 8, pp. 1853–1863, Aug. 2013.

[121] R. K. Singleton, E. G. Strangas, and S. Aviyente, "Extended Kalman filtering for remaining-useful-life estimation of bearings," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1781–1790, Mar. 2015.

[122] C. K. R. Lim and D. Mba, "Switching Kalman filter for failure prognostic," *Mech. Syst. Signal Process.*, vols. 52–53, pp. 426–435, Feb. 2015.

[123] J. Herp, M. H. Ramezani, M. Bach-Andersen, N. L. Pedersen, and E. S. Nadimi, "Bayesian state prediction of wind turbine bearing failure," *Renew. Energy*, vol. 116, pp. 164–172, Feb. 2018.

[124] J. Wang and R. X. Gao, "Multiple model particle filtering for bearing life prognosis," in *Proc. IEEE Conf. Prognostics Health Manage. (PHM)*, Jun. 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 *Proc. IEEE Autotestcon*, Sep. 2010, pp. 1–6.

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

[127] M. A. Djeziri, S. Benmoussa, and R. Sanchez, "Hybrid method for remaining useful life prediction in wind turbine systems," *Renew. Energy*, vol. 116, pp. 173–187, Feb. 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 *Proc. North Amer. Power Symp. (NAPS)*, Sep. 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 Trans. Sustain. Energy*, vol. 9, no. 1, pp. 157–167, Jan. 2018.

[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," *Renew. Energy*, vol. 129, pp. 260–270, Dec. 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 Trans. Rel.*, vol. 60, no. 1, pp. 14–20, Mar. 2011.

[132] C. Sun, Z. Zhang, and Z. He, "Research on bearing life prediction based on support vector machine and its application," *J. Phys., Conf. Ser.*, vol. 305, Jul. 2011, Art. no. 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, Nov. 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, Apr. 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," *Digit. Signal Process.*, vol. 27, pp. 159–166, Apr. 2014.

[136] S. Hong, Z. Zhou, E. Zio, and W. Wang, "An adaptive method for health trend prediction of rotating bearings," *Digit. Signal Process.*, vol. 35, pp. 117–123, Dec. 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 Trans. Ind. Electron.*, vol. 61, no. 6, pp. 2864–2874, Jun. 2014.

[138] J. Ben 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," *Mech. Syst. Signal Process.*, vols. 56–57, pp. 150–172, May 2015.

[139] A. Soualhi, K. Medjaher, and N. Zerhouni, "Bearing health monitoring based on Hilbert–Huang transform, support vector machine, and regression," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 1, pp. 52–62, Jan. 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 Trans. Ind. Informat.*, vol. 12, no. 3, pp. 924–932, Jun. 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 Trans.*, vol. 37, no. 6, pp. 543–557, Jun. 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," *Mech. Syst. Signal Process.*, vol. 35, nos. 1–2, pp. 219–237, Feb. 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, May 2016.

[144] X. Jin, Y. Sun, Z. Que, Y. Wang, and T. W. S. Chow, "Anomaly detection and fault prognosis for bearings," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 9, pp. 2046–2054, Sep. 2016.

[145] H. Jiang, J. Chen, and G. Dong, "Hidden Markov model and nuisance attribute projection based bearing performance degradation assessment," *Mech. Syst. Signal Process.*, vols. 72–73, pp. 184–205, May 2016.

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

[147] S. A. Aye and P. S. Heyns, "An integrated Gaussian process regression for prediction of remaining useful life of slow speed bearings based on acoustic emission," *Mech. Syst. Signal Process.*, vol. 84, pp. 485–498, Feb. 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. 10986–10995, 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, Jan. 2018.

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

[151] L. Zhao and X. Wang, "A deep feature optimization fusion method for extracting bearing degradation features," *IEEE Access*, vol. 6, pp. 19640–19653, 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," *Renew. Energy*, vol. 145, pp. 642–650, Jan. 2020.

[153] M. Elforjani and S. Shanbr, "Prognosis of bearing acoustic emission signals using supervised machine learning," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5864–5871, Jul. 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. 24521–24532, 2018.

[155] M. Rezamand, M. Kordestani, M. E. Orchard, R. Carriveau, D. S.-K. Ting, and M. Saif, "Improved remaining useful life estimation of wind turbine drivetrain bearings under varying operating conditions (VOC)," *IEEE Trans. Ind. Informat.*, early access, May 7, 2020, doi: 10.1109/TII.2020.2993074.

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

[157] S. A. Billings, *Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*. Hoboken, NJ, USA: Wiley, 2013.

[158] R. De Maesschalck, D. Jouan-Rimbaud, and D. L. Massart, "The Mahalanobis distance," *Chemometrics Intell. Lab. Syst.*, 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," *Mech. Syst. Signal Process.*, vol. 38, no. 2, pp. 515–533, Jul. 2013.

[160] D. M. Hawkins and D. H. Olwell, *Cumulative Sum Charts and Charting for Quality Improvement*. Cham, Switzerland: Springer, 2012.

[161] Y. Hu *et al.*, "A prediction method for the real-time remaining useful life of wind turbine bearings based on the Wiener process," *Renew. Energy*, vol. 127, pp. 452–460, Nov. 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, Dec. 2018.

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

[164] D. A. Pola *et al.*, "Particle-filtering-based discharge time prognosis for lithium-ion batteries with a statistical characterization of use profiles," *IEEE Trans. Rel.*, vol. 64, no. 2, pp. 710–720, Jun. 2015.

[165] F. Tamssaouet, K. T. P. Nguyen, K. Medjaher, and M. E. Orchard, "Degradation modeling and uncertainty quantification for system-level prognostics," *IEEE Syst. J.*, early access, Apr. 20, 2020, doi: 10.1109/JSYST.2020.2983376.

Milad Rezamand received the bachelor's degree in mechanical engineering from the Shahid Bahonar University of Kerman, Kerman, Iran, in 2008, the master's degree in aerospace engineering from the K. N. Toosi University of Technology, Tehran, Iran, in 2012, and the Ph.D. degree in mechanical engineering from the University of Windsor, Windsor, ON, Canada, in 2019.

He is currently a Post-Doctoral Fellow with the University of Windsor. His research interests include fault diagnostics, prognosis, and predicting the remaining useful life. His ongoing research includes estimating the remaining useful life of wind turbine components.

Mojtaba Kordestani (Senior Member, IEEE) received the bachelor's degree in electronic engineering from the Malek Ashtar University of Technology, Shahin Shahr, Iran, in 2002, the master's degree in control engineering from Islamic Azad University, Tehran, Iran, in 2008, and the Ph.D. degree in electrical engineering from the University of Windsor, Windsor, ON, Canada, in 2018.

He is currently a Post-Doctoral Fellow with the University of Windsor. His research interests are in control, estimation, fault diagnostics, and failure prognosis. He has published about 40 refereed journal articles and conference papers in these areas.

Rupp Carriveau (Member, IEEE) received the bachelor's degree in civil structural engineering from the University of Windsor, Windsor, ON, Canada, in 1994, and the M.A.Sc. and Ph.D. degrees in fluids engineering from Western University, London, ON, in 2000 and 2004, respectively.

He is also the Director of the Turbulence and Energy Laboratory and the Centre for Energy and Water Advancement, University of Windsor.

Dr. Carriveau is also a member of The American Society of Mechanical Engineers (ASME). He also serves on the Editorial Boards of *Wind Engineering*, *Advances in Energy Research*, and the *International Journal of Sustainable Energy*. He is also the President of the Underwater Energy Storage Society.

David S.-K. Ting received the B.A.Sc. degree from the University of Manitoba, Winnipeg, MB, Canada, in 1989, and the M.A.Sc. and Ph.D. degrees in turbulent combustion from the University of Alberta, Edmonton, AB, Canada, in 1992 and 1995, respectively.

He explored fluid–structure interactions at McGill University, Montreal, QC, Canada, prior to joining the University of Windsor, Windsor, ON, Canada, as a Faculty Member with the Department of Mechanical, Automotive and Materials Engineering. He is also the Founder of the Turbulence & Energy Laboratory, University of Windsor. He has cosupervised over 75 graduate students, coauthored more than 140 journal articles, authored four textbooks, and coedited eight books.

Dr. Ting is also a member of ASME.

Marcos E. Orchard (Member, IEEE) received the M.S. and Ph.D. degrees in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2005 and 2007, respectively.

He is currently a Professor with the Department of Electrical Engineering, Universidad de Chile, Santiago, Chile, where he is also an Associate Researcher with the Advanced Center for Electrical and Electronic Engineering (UTFSM). He has authored or coauthored more than 100 articles. His current research interests include failure prognosis, electromobility, mining industry, and finance.

Dr. Orchard is also a fellow of the Prognostic and Health Management Society (PHMS).

Mehrdad Saif (Senior Member, IEEE) received B.S., M.S., and D.Eng. degrees in electrical engineering from Cleveland State University, Cleveland, OH, USA, in 1982, 1984, and 1987, respectively.

He has been the Dean of the Faculty of Engineering, University of Windsor, Windsor, ON, Canada, since July 2011. His research interests are in systems and control; estimation, and observer theory; model-based fault diagnostics; condition monitoring, diagnostics, and prognostic; and application of these areas to automotive, power, autonomous systems, and other complex engineering systems. He has published over 350 refereed journal articles and conference papers and an edited book in these areas.