



# Technology development and commercial applications of industrial fault diagnosis system: a review

Chengze Liu<sup>1</sup> · Andrzej Cichon<sup>2</sup> · Grzegorz Królczyk<sup>3</sup> · Zhixiong Li<sup>3</sup>

Received: 24 May 2021 / Accepted: 12 September 2021 / Published online: 18 October 2021  
© The Author(s) 2021

## Abstract

Machinery will fail due to complex and tough working conditions. It is necessary to apply reliable monitoring technology to ensure their safe operation. Condition-based maintenance (CBM) has attracted significant interest from the research community in recent years. This paper provides a review on CBM of industrial machineries. Firstly, the development of fault diagnosis systems is introduced systematically. Then, the main types of data in the field of the fault diagnosis are summarized. After that, the commonly used techniques for the signal processing, fault diagnosis, and remaining useful life (RUL) prediction are discussed, and the advantages and disadvantages of these existing techniques are explored for some specific applications. Typical fault diagnosis products developed by corporations and universities are surveyed. Lastly, discussions on current developing situation and possible future trends are in the CBM performed.

**Keywords** Condition-based monitoring · Fault diagnosis system · Signal processing · Diagnostics and prognostics · Commercialization

## 1 Introduction

### 1.1 Background

In modern industry, machines develop towards being more complicated, and intelligent, and are subject to growingly demanding operation conditions. Slight performance deterioration or security risks may bring serious consequences, which can lead to sudden breakdown or even devastating accident with enormous financial losses and casualties if not detected early [1–7]. For example, an air crash caused by mechanical failures (e.g., an engine fault), a wind turbine that collapsed due to mechanical failures, and Volvo repeatedly recalling cars due to mechanical failures. In order to keep the machinery running with high reliability and

maintain a low downtime, it is very important to identify the existence and severity of faults in the machinery accurately.

In the literature, maintenance techniques are separated into three types [8], including the breakdown maintenance, preventive maintenance, and condition-based maintenance. The breakdown maintenance is a strategy that is applied to repair the components/machinery only after a fault has occurred. Therefore, it would not be able to avoid any faults. Preventive maintenance performs maintenance activities at a fixed-length interval regardless of the practical machine condition [8]. The time interval is generally determined by experience and original equipment manufacturer recommendations, which will cause over-maintenance when the periodical interval is too short or lead to an unexpected failure when the chosen interval is too long. Breakdown maintenance and preventive maintenance are disappearing from the real industrial application because of these problems. CBM (also called predictive maintenance) is a maintenance procedure that takes maintenance actions based on the information indicated in the condition of machinery instead of regular time interval, which achieves objectives of cost reduction and reliability improvement. The main advantage of CBM is to recommend maintenance activity only when there is evidence of abnormal behaviors of a machine for avoiding unnecessary maintenance tasks and to not interrupt normal operations [9–11]. As a result, CBM has

✉ Zhixiong Li  
zhixiong.li@ieee.org

<sup>1</sup> School of Engineering, Ocean University of China, Qingdao 266100, China

<sup>2</sup> Faculty of Electrical Engineering, Opole University of Technology, 45-758 Opole, Poland

<sup>3</sup> Faculty of Mechanical Engineering, Opole University of Technology, 45-758 Opole, Poland

attracted more and more attentions from academic researchers and industrial operators over the past few years.

Figure 1 presents three important processes of CBM [12]:

- (1) Data acquisition, to collect and store useful data from targeted physical assets.
- (2) Data processing, to process and analyze the data collected in step 1.
- (3) Maintenance decision-making, to take useful maintenance actions.

The key of a CBM program is the maintenance decision-making where maintenance actions are recommended through diagnosis and prognosis. Fault diagnosis aims at identifying the fault mode of the machinery after detection and prognostics commonly oriented towards identifying and quantifying the fault. The latter is also capable of predicting the process of degradation. Thus, a maintenance decision is determined with reliable prediction. It may be noted that prognostics is much more efficient than diagnostics to achieve maintenance of machinery. While correct diagnosis significantly reduces the downtime by detecting a fault in incipient stage and identifying the faulty location, prognosis directly estimates how soon and how likely a fault will occur. However, prognostics is usually difficult to acquire a 100% prediction result. Diagnostics can be a complementary tool to provide maintenance decision support when prediction approach fails and a fault occurs.

## 1.2 Development of fault diagnosis system

To ensure the safe operation of the machine, fault diagnosis system is exploited. With progresses of condition monitoring theory and detection technology, especially network technology, fault diagnosis system can be roughly grouped into three categories: single fault diagnosis system, distribution fault diagnosis system, and remote fault diagnosis system. These systems will be discussed in the following three subsections, respectively.

### 1.2.1 Single fault diagnosis system

Single fault diagnosis system aims to provide real-time help for users, which allows the health state of a specific machinery to be evaluated independently of any connected assets. It is a common situation to find multiple faults in a single component; there is no one-to-one relationship between the

fault symptom and the fault itself. Moreover, the degree and response rate of each fault are different. As a result, we should integrate various techniques to ensure the accuracy and effectiveness of fault diagnosis in a single fault diagnosis system. Hsueh et al. [13] proposed a novel methodology to monitor the condition of a three-phase induction motor. They applied the empirical wavelet transform as a preprocessor to transform the raw signal into two-dimensional grayscale images and used a deep convolutional neural network to automatically extract robust features from the grayscale images to diagnose the faults. Zhang et al. [14] developed a method based on permutation entropy, ensemble empirical mode decomposition, and support vector machines to detect motor bearing faults. Liu et al. [15] combined the least squares support vector machines (LSSVM) and empirical mode decomposition to improve the accuracy of bearing fault diagnosis. However, machines are composed of a series of connected components that continuously interact with one another. In order to simplify diagnosis analysis, single fault diagnosis systems ignore the complexity and uncertainty caused by the interactions among components and detect each component in isolation, which may reduce confidence in the outputs of the system, or even lead to misdiagnosis.

### 1.2.2 Distributed fault diagnosis system

The characteristics of production equipment in modern industry are upsizing, complex, continuous, and automation, which generate severe challenges to effective maintenance. Distributed fault diagnosis mode has received much attention, because it combines diagnostic information from different components to improve the reliability of the condition monitoring of each individual component and the whole system. A distributed fault diagnosis system is comprised of several different diagnostic sub-systems, which decomposes the fault diagnosis task of the entire system into the fault diagnosis task of part sub-systems according to the idea “disassemble-synthesizing.” Each sub-system independently diagnoses the fault in their local areas. If the sub-system cannot achieve the task, the information from the entire system is used to resolve the problem. Jiang et al. [16] proposed a distributed monitoring scheme based on multivariate statistical analysis and Bayesian method for large-scale plant-wide processes. Shahnazari et al. [17] designed a distributed detection and isolation architecture to detect the faults of heating, ventilation, and air conditioning systems. Chen et al. [18] proposed a distributed fast fault diagnosis method based on deterministic learning theory for multi-machine power system fault detection. They established a knowledge bank and gradually updated it. However, distributed fault diagnosis systems apply computer local area network to transfer information. Different large-scale systems are required to repeatedly construct their own local area network.

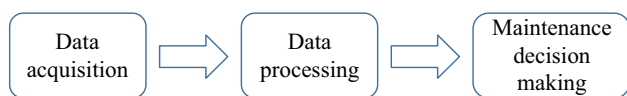


Fig. 1 General CBM process

### 1.2.3 Remote fault diagnosis system

The progresses of network and communication technology provide the opportunity to develop remote fault diagnosis system. Compared with traditional diagnosis system, remote fault diagnosis system has many good characteristics such as open architecture, resource sharing, and high efficiency. Remote fault diagnosis systems apply Internet to transfer the fault information to the central maintenance station. Then, the diagnosis result and maintenance suggestions are sent to users via the Internet. Consequently, users can access the performance of machinery from anywhere in the world by remote fault diagnosis system. For the applications of remote fault diagnosis system, readers can refer to Refs. [19–21]. However, the current fault diagnosis systems are difficult to share and exchange information because of different architectures. To eliminate information island and improve the cross-platform interoperability, some international standards and advanced technologies have been proposed. Wang et al. [22] proposed a remote fault diagnosis system that took Extensible Markup Language (XML) as a core and exploited it to encode diagnostic data. Zhao et al. [23] introduced the progress of remote fault diagnosis system based on the service-oriented architecture (SOA). The architecture is the integration of multiple technologies (e.g., Web Services, Smart Client, and XML).

## 2 Data acquisition

For any maintenance practice, data acquisition is one of the most important steps. Data collection can be divided into two categories: condition monitoring data and event data. The condition monitoring data are the measurements related to the health condition of the targeted machinery. The event data include the information about the maintenance adjustments and operational changes (e.g., installation, repair, oil change, etc.) [12]. Event data and condition monitoring data are equally important in CBM. However, the latter has gotten more attention which the collection of event data in practical application is often ignored. One possible reason is that the event data are not regarded as the same value as the condition monitoring data during the monitoring process. This is incorrect since the event data are critical for researchers/engineers in consideration of system redesign and improvement of condition indicators. Hence, it is a must to combine the event data and condition monitoring data to build a better CBM model that accurately identify the health condition of machineries. More details about the event data can be found in [12, 24, 25].

So far, according to the different mechanisms of monitoring condition and sensors, various condition monitoring data can be applied to indicate the condition of mechanical

equipment like vibration, current, acoustic emission, temperature, and oil debris analysis.

### 2.1 Vibration

Since vibration signals contain the dynamic characteristics of machinery condition, it has become a most widely used and effective method to evaluate operation and machinery condition in recent years. Failures can produce changes in the vibration signal. For example, a crack in a bearing generates a shock impulse every time the crack contacts another part of the machine. Then, the location and severity of fault can be clearly identified by the vibration signal [26]. Vibration signals are usually acquired by the vibration test equipment, such as displacement sensors, speed sensors, and acceleration sensors. The main advantage of vibration monitoring is the ability to diagnosis different types of faults, either mechanical or electrical faults. Moreover, inexpensive sensors, immediate measurement, and the ability to pinpoint the damage component and its location are other benefits of the vibration analysis [27, 28]. However, vibration measurement requires access to the machine, which is hard to realize for complex and severe conditions such as corrosion and elevated temperature. Another issues of vibration monitoring are that special training is required to accurately install vibration sensors and the output is easily be interfered by external noise [29].

### 2.2 Stator current

When a rotor system has faults (e.g., the misalignment of the shaft, bearing failures, etc.), it generates additional torque ripple. The motor will create a corresponding electrical torque to balance the torque ripple [30]. Thus, faults of targeted physical asset could be reflected by current signals. Current signal does not require additional sensors to be mounted on or next to the measured machinery, which can be directly obtained by trapping into the existing voltage and current transformers that are already installed as part of the protection system [31]. This appealing merit therefore facilitates the fault diagnosis in long-distance cases. One major disadvantage of the current analysis is that fault components are subtle in the current signal where the dominant components are supply frequency components. Then, for downstream mechanisms (e.g., gearboxes) in an electromechanical machine, the torque change caused by faults may have little impact on motor current. Furthermore, many problems will be associated in the case of current spectrum analysis, for example, the characteristic harmonics caused by air-gap variation, harmonics of eccentricity caused by the construction of the motor, and harmonics due to variations of the load and the supply frequency [31, 32].

### 2.3 Acoustic emission

Acoustic emission (AE) analysis has become an effective condition monitoring tool. AE was originally used for non-destructive testing of static structures, and it has been extended to condition monitoring of machinery (e.g., shaft cracks and composite material spalling, fracture, and delamination of in rolling bearings) in recent years [33, 34]. In machinery monitoring application, AE is defined as transient elastic waves caused by the interface of components in relative motion [10]. Since the frequency response of acoustic emission is higher than vibration's, which is within 100 kHz and 1 MHz, significant merit of AE monitoring over vibration methods is the ability to capture surface and subsurface slight damage and detect early fault [35]. However, the applications of AE analysis in the fault diagnosis are partly limited due to the difficulty in processing, interpreting, and classifying the obtained data. Moreover, AE signal suffers from severe attenuation and reflections before reaching the sensors, so AE sensors are required to be close to its source [10, 36].

### 2.4 Temperature

It is well known that temperature is one of the most powerful parameters for indicating the healthy condition of machinery. Every component has a temperature range during normal operation. Users can compare the actual temperature with the range to judge whether a fault has occurred. Nevertheless, an increase in temperature can be caused by various factors, such as change in working load and speed and degradation of lubricant oil; even if the temperature rise can be recognized, users should determine the cause of the temperature rise [35]. In addition, traditional temperature detection method is not sensitive for early fault. Infrared thermography (IRT) is a novel method that remotely measures the temperature of an object and provides the thermal image. Infrared detectors are the key of IRT system, which detects the infrared radiation emitted by an object in a non-intrusive way and exploits Stefan-Boltzmann's law to obtain the temperature [37]. However, IRT analysis should take time to heat up motor and process thermal images. Another disadvantage of IRT is that it experiences high system costs [37, 38].

### 2.5 Oil debris monitoring

Oil debris monitoring is a well-established way to evaluate the quality of oil and monitor the wear condition of internal oil-wetted components, which can be roughly grounded into two sub-categories: oil condition monitoring and wear

debris detection. The former interests in the degradation of oil properties caused by oxidation, thermal, and shear effects to determine whether the oil is suit for further use, and the latter focus on the components, contents, and morphology of debris generated and its distribution for confirming wear state of components [9, 39]. Many debris detection techniques have been developed for investigating the health conditions of machines in the past few decades, which are separated into offline detection and online detection. Some offline methods are inefficient and cannot provide the wear state in real time, so online analysis becomes a hotspot in state analysis. There are some different types of online debris detection according to the measurement principles, for instance, optical method, inductive method, resistive-capacitive method, and acoustic methods [40]. Compared to other parameters such as vibration, oil debris analysis is earlier to identify loss of mechanical integrity and can monitor the evolution of the wear process [27]. In addition, it has several distinguished advantages, e.g., close relationship with wear surface profile, long persistence of information, and powerful anti-interference capability [41]. Unfortunately, oil debris monitoring is only applicable to systems that have a recirculating lubricating fluid loop and cannot identify the fault from specific components which have a common metal element, such as bearings and gears.

### 2.6 Epilog

Due to different detection principle, each monitoring technology has its own advantages and disadvantages as shown in Table 1. For example, vibration analysis is the most widely used technology in condition monitoring field, but it requires special training for accurately installing sensors and is not easy to detect early fault of machinery; AE monitoring can overcome the latter problem but suffers from severe attenuation and reflections of signals before reaching the sensors; current detection does not need additional sensors but limited by low signal-to-noise ratio; while temperature technology and oil debris technology are only used in specified conditions or auxiliary use, multi-sensor information is the direction of development of condition monitoring in the future. This paper will not cover the details of information fusion. One point the authors would like to make is that multi-sensor system combines all observation information based on combined optimization criterion to obtain the consistency interpretation and description to observation environment and create a new result at the same time [27]. It aims at using multiple sensors to improve the estimation precision of condition. More detailed discussion and application for data fusion have been reported in the publications [27, 42, 43].

**Table 1** Advantages and disadvantages of different monitoring parameters

Monitoring parameters	Advantages	Disadvantages
Vibration	<ul style="list-style-type: none"> <li>• At relatively low cost</li> <li>• Provides immediate measurement</li> <li>• Can detect different types of faults</li> <li>• Pinpoints the damage component and its location</li> </ul>	<ul style="list-style-type: none"> <li>• An intrusive method</li> <li>• Easily disturbed by of external noise</li> <li>• Requires special training for accurately installing sensors</li> <li>• Difficult to detect early fault</li> </ul>
Current stator	<ul style="list-style-type: none"> <li>• A non-intrusive method</li> <li>• At relatively low cost</li> <li>• Suitable for long-distance cases</li> <li>• Does not need additional sensors</li> </ul>	<ul style="list-style-type: none"> <li>• Sometimes provides low signal-to-noise ration</li> <li>• Requires effective signal processing technology</li> </ul>
Acoustic emission	<ul style="list-style-type: none"> <li>• Can detect early fault</li> <li>• Can capture surface and subsurface slight damage</li> <li>• Provides high signal-to-noise ration</li> </ul>	<ul style="list-style-type: none"> <li>• Difficult to process, interpret, and classify the obtained data</li> <li>• Suffers from severe attenuation and reflections of signals before reaching the sensors</li> </ul>
Temperature	<ul style="list-style-type: none"> <li>• A non-intrusive method</li> <li>• Provides thermal image</li> </ul>	<ul style="list-style-type: none"> <li>• At relatively high cost</li> <li>• Requires a large volume of time for heating up motor and processing thermal images</li> </ul>
Oil/Debris	<ul style="list-style-type: none"> <li>• Can detect loss of mechanical integrity at early stage</li> <li>• Closely related to wear surface profile</li> <li>• Has powerful anti-interference capability</li> </ul>	<ul style="list-style-type: none"> <li>• Limited to machinery with closed-loop lubricating system</li> <li>• Difficult to identify the fault from specific components which have a common metal element</li> </ul>

### 3 Signal processing

It is almost impossible to directly recognize the type of faults owing to the variability and richness of the original signals. Hence, signal processing is used to cleaning, transforming, and modeling data with the goal of minimizing or eliminating noise and extracting important information related to faults. Signal analysis techniques can be classified into three categories according to the different domains: time domain analysis, frequency domain analysis, and time–frequency analysis. A summary of these techniques is given in Table 2, which provides their advantages and disadvantages to help researchers who work in the field of signal processing select appropriate signal processing tools. This paper focuses on the performance of vibration signal in signal processing, fault diagnosis, and life prediction in the sense that vibration analysis is the most mature technique in condition monitoring. In the following sections, signal processing techniques are discussed in detail.

#### 3.1 Time domain analysis

When the running conditions of machinery deviate from the normal condition, the time domain statistical features of the signal will be different from the normal condition. Moreover, the features will be different under different fault mode. Consequently, the time domain features contain abundant fault information, and they can be served as sensitive character used to analyze the condition of machinery.

Table 3 shows the commonly used features in time domain, including dimensional statistical parameters and non-dimensional statistical parameters. Statistics analysis has been broadly used in condition detection and fault

diagnosis. Williams et al. [44] combined high-frequency resonance technology and traditional vibration metrics (e.g., root mean square, peak value, kurtosis, and crest factor) to detect damage in rolling element bearings. Lei et al. [45] developed two diagnostic parameters for diagnosing faults of planetary gearboxes. Although statistics analysis is easy to implement and can evaluate the performance of machinery, it is difficult to directly expose the occurrence and location of faults. In addition, feature selection is usually a challenge when applying statistics method to analyze signals.

Another popular time domain technique is time synchronous average (TSA). It exploits the ensemble average of the signal separated by the exact period to obtain the signal components of interest, and any others will be reduced asymptotically towards zero. Zhang et al. [46] applied TSA and wavelet transform to identify the gear failure. However, TSA requires a long signal and corresponding rotation mark signal, and signal characteristics cannot be revealed by TSA under varying speeds due to the phase accumulation error. To combat the weakness of the TSA, Xiao et al. [47] proposed an improved dynamic time synchronous averaging and extracted the periodic feature signal from the fluctuated vibration signal to diagnose the gear fault. They used the dynamic time warping to estimate the phase accumulation error among the envelop signal segments and further applied it to compensate the phase accumulation error between the intrinsic mode function segments of the reconstructed signal.

Autoregressive (AR) model has also been proved to be a powerful method in signal processing. The idea of AR model is to fit the data to a parametric time series model and extract features based on this parametric model [12]. The commonly used models include autoregressive model, moving average model, and autoregressive moving average model. Since the



**Table 2** Comparison of various signal processing techniques

Domain	Techniques	References	Advantages	Disadvantages
Time	Statistics analysis	[44, 45]	<ul style="list-style-type: none"> <li>• Has simple compute process</li> <li>• Can evaluate the performance of machinery</li> </ul>	<ul style="list-style-type: none"> <li>• Difficult to obtain detail about the fault</li> <li>• No clear guidelines on the selection of features</li> </ul>
	TSA	[46, 47]	<ul style="list-style-type: none"> <li>• Can improve signal-to-noise ratio by suppressing the asynchronous components</li> </ul>	<ul style="list-style-type: none"> <li>• Requires a long signal and corresponding rotation mark signal</li> <li>• Difficult to obtain signal characteristics under varying speed</li> </ul>
	AR model	[48, 49]	<ul style="list-style-type: none"> <li>• Does not need knowledge of mathematical model and fault mechanisms</li> <li>• Has high sensibility for condition variation</li> </ul>	<ul style="list-style-type: none"> <li>• Requires a large volume of time for determine appropriate order of model</li> <li>• Inapplicability to non-stationary signals</li> </ul>
Frequency	Power spectrum	[50–52]	<ul style="list-style-type: none"> <li>• Provides frequency information</li> </ul>	<ul style="list-style-type: none"> <li>• Unable to provide phase information</li> </ul>
	Higher order spectrum	[29, 53, 54]	<ul style="list-style-type: none"> <li>• Suppresses Gaussian noise</li> <li>• Retains phase information</li> </ul>	<ul style="list-style-type: none"> <li>• Suffers from non-Gaussian noise interference</li> <li>• Heavy computation burden</li> </ul>
	Cepstrum	[55, 56]	<ul style="list-style-type: none"> <li>• High capability to simplify and extract the periodic components</li> <li>• Can separate transmission path from the real signal</li> </ul>	<ul style="list-style-type: none"> <li>• Suffers from intense background noise at early stage</li> <li>• Unable to provide obvious features due to the average effect of FT</li> </ul>
	Envelope spectrum	[57–60]	<ul style="list-style-type: none"> <li>• Provides more fault information than original signal</li> <li>• Remarkable ability in extracting the characteristic frequencies</li> <li>• Provides spectral and temporal representation of modulating signal</li> </ul>	<ul style="list-style-type: none"> <li>• No standard method to choose the central frequency and bandwidth of the bandpass</li> </ul>
Time–frequency	STFT	[12, 61, 62]	<ul style="list-style-type: none"> <li>• Easy to understand</li> <li>• Can offer local features of signal</li> </ul>	<ul style="list-style-type: none"> <li>• Lacks adaptability</li> <li>• Not appropriate to analyze highly transient phenomena in signals</li> </ul>
	Wavelet transform	[24, 63–66]	<ul style="list-style-type: none"> <li>• Has a zooming and adaptive windowing capability suited for non-stationary signal</li> <li>• Effective in self-similar or fractal signals</li> </ul>	<ul style="list-style-type: none"> <li>• No standard method to choose wavelet basis</li> <li>• Suffers from energy leakage</li> </ul>
	WVD	[67–69]	<ul style="list-style-type: none"> <li>• Has the highest time–frequency resolution among all the time–frequency distributions</li> </ul>	<ul style="list-style-type: none"> <li>• Inevitably hindered by the cross-term interferences</li> <li>• Suppresses cross-term at the expense of deteriorated time–frequency resolution</li> </ul>
	EMD	[70–75]	<ul style="list-style-type: none"> <li>• Adaptive capability to represent arbitrary signals</li> <li>• Provides fine time–frequency resolution</li> <li>• Free of cross-term interferences</li> </ul>	<ul style="list-style-type: none"> <li>• Difficult to resolve signal components due to end effects and mode mixing</li> </ul>
	EEMD	[14, 76, 77]	<ul style="list-style-type: none"> <li>• Adaptive capability to represent arbitrary signals</li> <li>• Provides fine time–frequency resolution</li> <li>• Can suppress mode mixing</li> </ul>	<ul style="list-style-type: none"> <li>• Unable to eliminate mode mixing completely</li> <li>• No standard method to choose the number of trails and the amplitude of added noise</li> </ul>
	LMD	[78–82]	<ul style="list-style-type: none"> <li>• Adaptive capability to represent arbitrary signals</li> <li>• Provides fine time–frequency resolution</li> <li>• Retains more variations of the frequency and amplitude in signals</li> </ul>	<ul style="list-style-type: none"> <li>• Difficult to resolve signal components due to end effects and mode mixing</li> </ul>

parameters of AR model are sensitive to the condition variation, it can be used to reflect the condition variation of a dynamic system. Furthermore, AR models can be applied to diagnose the fault without constructing mathematical models or studying the fault mechanisms. Wang et al. [48] extracted

the detective features from the voltage data sequence using the ARMA model. Several researches have applied AR models combining with others methods to fight the problem of AR model in processing non-stationary signals. For example, Han et al. [49] proposed a joint approach based on

**Table 3** Expression and properties of time domain features

Categories	Features	Expressions	Properties
Dimensional statistical parameters	Mean	$\bar{x} = \frac{\sum_{i=1}^N x_i}{N}$	The average value of a signal
	Root mean square	$x_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$	The measure of power contained in a signal
	Standard deviation	$x_{std} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$	The indicator of the amount of variation or dispersion from the average
Non-dimensional statistical parameters	Peak value	$x_p = \max(x_i)$	The indicator of change in a signal due to occurrence of impacts
	Kurtosis	$x_{kur} = \frac{1}{N} \sum_{i=1}^N \frac{(x_i - \bar{x})^4}{x_{std}^4}$	The measure of whether the distribution is peaked or flat related to a normal distribution
	Skewness	$x_{ske} = \frac{1}{N} \sum_{i=1}^N \frac{(x_i - \bar{x})^3}{x_{std}^3}$	The measure of lack of symmetry about its mean
	Crest factor	$CF = \frac{x_p}{x_{rms}}$	The measure of the spikiness of a signal
	Shape factor	$SF = \frac{x_{rms}}{ \bar{x} }$	The function of the redressed signal average

variational mode decomposition and AR model parameters for fault detection of rolling bearing. They decomposed the vibration signal into a series of stationary component signals, and constructed an AR model for each component mode.

### 3.2 Frequency domain analysis

Frequency domain analysis gives information about the signal in the frequency domain. The significant merit of frequency domain analysis over time domain is the ability to distinguish and isolate frequency components of interest. As shown in Fig. 2, various frequency domain techniques have been applied in signal processing.

#### 3.2.1 Power spectrum

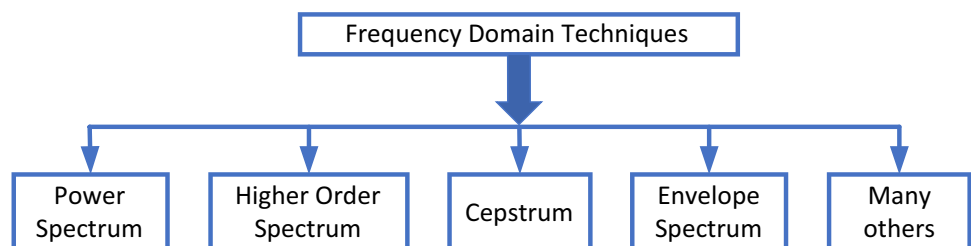
The most commonly applied approach in frequency domain analysis is power spectrum. To verify the performance of power spectrum in fault diagnosis, Liang et al. [28] attached five accelerometers at the ends of the induction motor to measure vibration of the induction motor. The power spectrum of an induction without and with broken rotor bar fault under 0% and 100% motor load conditions is shown in Fig. 3. There are no visible sidebands for broken rotor fault

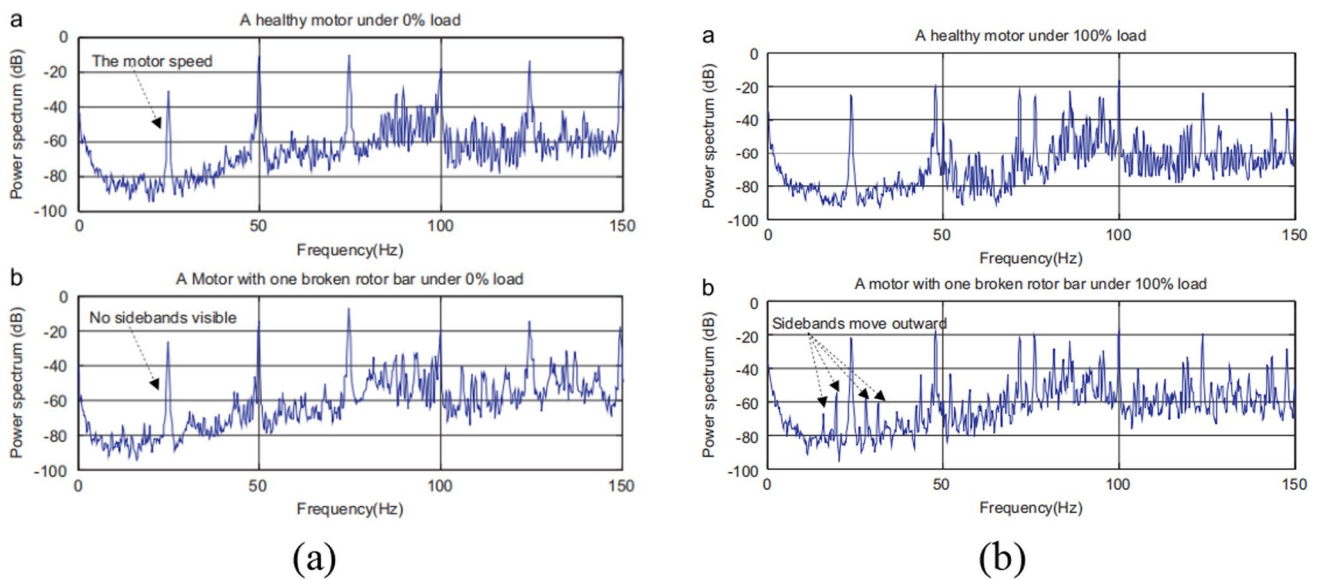
under 0% motor load, in the sense that the slip is too small to be identified. Figure 3b shows clear sidebands for the same fault when the load is grown to 100%. It can be seen that the broken rotor bar fault can be clearly diagnosed by power spectrum, provided a certain amount load is exerted on the motor. More application examples of power spectrum in mechanical fault diagnosis and condition monitoring can be found in [50–52]. One major drawback of power spectrum is losing phase information.

#### 3.2.2 Higher order spectrum

The merit of HOS is its ability to suppress Gaussian noise in signal detection, parameter estimation, classification, etc. In addition, HOS preserves the phase information. Hence, HOS provides more information than the power spectrum, in the case of non-Gaussian signals, can detect nonlinear couplings, and explains the origin of certain peaks in the power spectrum [24]. To achieve more accurate results, higher order spectrum has been applied in fault diagnosis. Bi et al. [53] applied the improved variational mode decomposition and bi-spectrum algorithm to distinguish the states of the valve clearance. Guo et al. [29] analyzed the vibration of the planetary gearbox based on wavelet packet energy

**Fig. 2** Frequency domain techniques



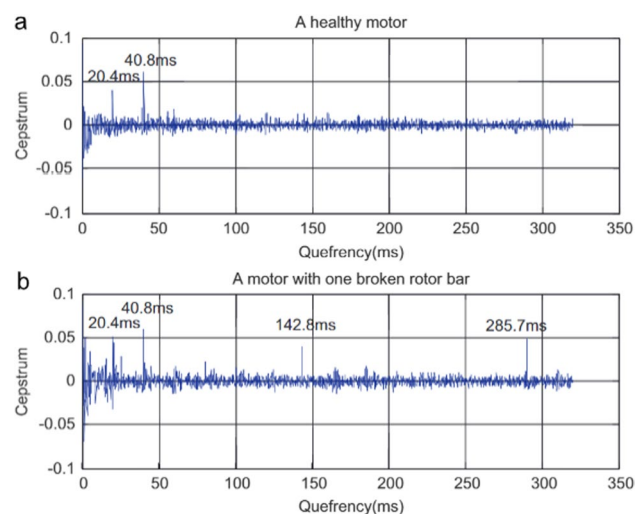


**Fig. 3** The power spectrum of induction motor without and with broken rotor bar fault (**a** 0% load; **b** 100% load) [28]

and modulation signal bi-spectrum and pointed out that the method was effective and feasible to identify the early fault diagnosis. Huang et al. [54] compared the performance of the conventional bi-spectrum and the modulation signal bi-spectrum in detecting rotor faults and showed that the variant is more effective.

### 3.2.3 Cepstrum

From Fig. 3, it can be observed that a lot of harmonics and sidebands appear in the power spectrum of the broken bar fault. Although these provides us some extra information about the faulty source, understanding what those harmonics and sidebands are and how they are related to each other are a problem. Cepstrum analysis is capable of simplifying and extracting the periodic components in spectrum of the vibration signal. Then, cepstrum analysis can transform the relation of the components from convolution form into addition form, which separates the transmission path from the real signal [55]. Therefore, Liang et al. [28] applied cepstrum to analysis the fault of induction motor as shown in Fig. 4. For the induction motor without broken bar fault, the cepstrum only presents the fundamental rotating quefrequency and its harmonics but the cepstrum of the motor with broken bar faults shows several extra sidebands information related to the fault. For most cases of early machinery failure, the spectrum feature is often buried in intense background noises. Moreover, fault feature would be further weakened by the average effect of Fourier transform after cepstrum processing. In order to overcome the drawback, Li et al. [56] and Zhang et al. [55] proposed a local cepstrum technique for



**Fig. 4** The cepstrum of induction motor without (**a**) and with (**b**) broken bar fault [28]

diagnosing gearbox faults, which enhances the capability of extracting periodical features.

### 3.2.4 Envelope analysis

A well-known approach applied to extract defect frequency components from the signal is envelope analysis. Before carrying out envelope demodulation, the signal is usually bandpass filtered at a frequency region where there is a high signal-to-noise ratio. Then, Hilbert transform is performed to obtain the envelop spectrum [50]. It has been found that envelope signals contain many more fault-related



information than the original signals and can be used both for spectral and temporal representation of modulating signal. Chacon et al. [57] combined wavelet packet analysis and Hilbert transform to achieve the bearing incipient fault detection. Abd-el-Malek et al. [58] proposed a method based on envelope analysis and reliably diagnosed broken rotor bars. However, central frequency and bandwidth of the bandpass filter are usually specified based on the experience of researchers, and their values will significantly affect the accuracy of analysis result. In order to overcome the problem, some researchers apply several methods to decompose signal into a set of single component amplitude-modulated and frequency-modulated signals as a preprocessor, thus avoiding demodulation error. For example, Du et al. [59] employed empirical mode decomposition to preprocess bearing vibration signals and applied the Wigner-Ville distribution and envelope analysis to recognize bearing fault types. Wang et al. [60] combined variational mode decomposition and envelope analysis to detect incipient bearing faults.

### 3.3 Time–frequency analysis

On the one hand, time domain methods cannot obtain the frequency information of the signal. On the other hand, frequency domain methods cannot reveal the local features in both time and frequency domains simultaneously. Moreover, most of the traditional techniques are based on the stationary assumption, which is incapable of analyzing non-stationary signals. Joint time–frequency analysis is an effective method to resolve these problems, which represents the signal in a time–frequency-amplitude/energy density three-dimensional space to reveal the frequency components and their time variation characteristics for more accurate diagnostics [83]. Up to now, various time–frequency approaches have been proposed, including linear time–frequency distributions, e.g., short-time Fourier transform (STFT) and wavelet transform; bilinear time–frequency distributions like Wigner-Ville distribution (WVD) and its variants; and adaptive non-parametric approaches such as empirical mode decomposition (EMD), ensemble empirical mode decomposition (EEMD), and local mean decomposition (LMD). We will give a discussion on these techniques in the following sections.

#### 3.3.1 Short-time Fourier transform

In essence, linear time–frequency methods are a process of decomposing signals into a weighted sum of a series of components in both time and frequency domains, which is different from bilinear time–frequency distribution in the fact that they are free from cross-term interferences. However, the time and frequency resolutions cannot reach the best simultaneously owing to the effect of Heisenberg

uncertainty principle [84]. The idea of STFT is to segregate a signal into parts with short-time windows and apply FFT to each part. The window can move along time; thus, STFT adds a time variable to the Fourier spectrum and then the time-varying nature of a signal can be revealed by the local spectrum.

STFT is the earliest time–frequency approach; many researchers have applied it to fault signature process [12, 61, 62]. It overcomes the defect that traditional Fourier transform cannot reveal the local features of a signal. Once the window size is chosen, however, STFT only provides constant time–frequency resolution. Consequently, STFT is not suitable to be applied to process non-stationary signals with high change at the scale of the window (e.g., impulses).

#### 3.3.2 Wavelet transform

Wavelet transform uses wavelets instead of sine function as the basis and applies the scale parameter and the time parameter to express the signal in a series of signal components with different frequencies at different time [12, 83]. Wavelet analysis can adjust window size through the selection of the mother wavelet and approximation scales, thus overcoming the disadvantage of STFT.

Wavelet analysis has been successfully applied to machinery fault diagnosis. Gangsar et al. [63] studied the impact of different wavelet function to the fault diagnosis of induction motor. Jaber et al. [64] proposed a signal analysis approach based on the discrete wavelet transform and artificial neural network for industrial robot joints fault detection. Talhaoui et al. [65] utilized discrete wavelet transform to process stator current signals of induction machine and obtained the envelope spectrum via Hilbert transform for broken rotor bars. For other references on the theoretical background of wavelet transform, application of the wavelet to machinery fault diagnosis, and new research trend, readers can refer to [24, 66]. The variable time and frequency resolution of wavelet analysis make it have a great performance for non-stationary signal processing. Although wavelet analysis is capable to iteratively decompose the approximation signals, it cannot further process the detail signals. In order to solve this problem, wavelet packet transform is proposed to increase the frequency resolution for high-frequency components so that the detail signals can also be iteratively decomposed [83]. As a result, the signal is transformed into multiple equal frequency bands, which can get a better time and frequency resolution. Up to now, a lot of wavelet basis have been proposed. However, the standard method of choosing wavelet basis has not been established. Then, the energy leakage is inevitable because of the fact that wavelet analysis is an adjustable windowed FT essentially.

### 3.3.3 Wigner-Ville distribution

The Wigner-Ville distribution is the basis of almost all bilinear time–frequency distribution, which is not based on signal segmentation so that has the best resolution. However, for multi-component signals, it is inevitably disturbed by cross-terms, and auto-terms and cross-terms may overlap on the time–frequency plane, which will make it more difficult to identify the time–frequency features [83]. Therefore, the WVD is not suitable to be applied to analyze the non-stationary signal directly. Figure 5 shows the WVD of a synthetic signal. It can analyze the time–frequency structures of two components with the highest resolution. It can be observed that the signal is seriously interfered by cross-terms. The cross-terms may make it difficult to understand the signal structure without a priori knowledge of the signal.

In order to suppress the cross-terms interferences, researchers have proposed various solution like the pseudo Wigner-Ville distribution (PWVD), Cohen class distributions, and affine class distributions. Guan et al. [67] studied planetary gearbox fault diagnosis based on the Cohen class distribution. Fan et al. [68] applied EMD-PWVD to transform vibration signals into contour time–frequency images and combined with FCM clustering to detect bearing faults. Li et al. [69] used the smoothed pseudo Wigner-Ville distribution to detect induction motor failures. However, the Wigner-Ville distribution’s variants suppress cross-terms

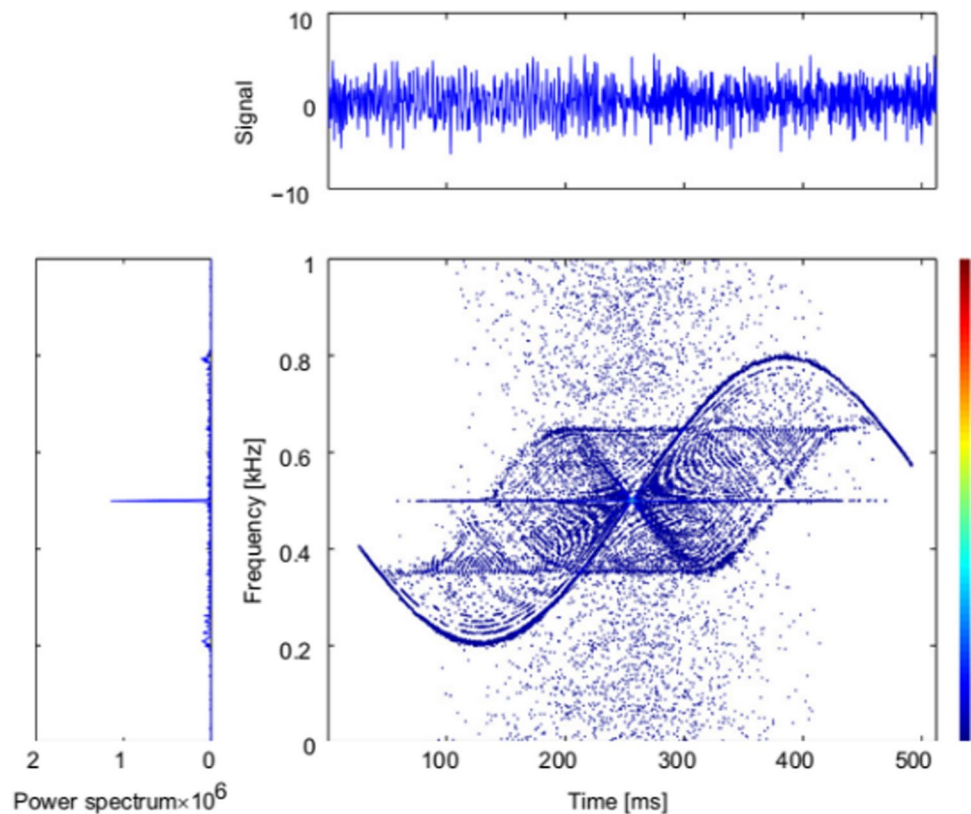
will lead to reduced time–frequency resolution and may create extra interference.

### 3.3.4 Empirical mode decomposition

The calculation of instantaneous frequency is one of the major problems in time–frequency distribution construction, and EMD is extensively used to accurately estimate the instantaneous frequency. For mono-component signals, the instantaneous frequency is computed by the derivative of phase relative to time, but most signals are composed of multiple components in real applications. EMD can decompose the signal into a series of complete and almost orthogonal components, called intrinsic mode function (IMF)[85]. An IMF is a function that satisfies the following two conditions: (1) in the whole data set, the number extrema and the number of zero-crossings must either equal or differ at most by one, and (2) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero [70].

EMD has several merits. Firstly, EMD is a self-adaptive signal analysis approach; there is no need to construct any basis function to match the signal characteristic structure. Then, EMD is free from cross-term interferences, which is because the original signal is represented as a linear superposition of a series of IMFs. Many researchers have applied EMD to fault diagnosis of bearing and gearbox. Ben Ali

**Fig. 5** Wigner-Ville distribution of a synthetic signal [83]



et al. [71] extracted fault features by EMD energy entropy and used it to train the artificial neural network to detect bearing faults. They found that the method can identify the severity of the fault successfully. He et al. [72] applied a hybrid method based on EMD, Fast ICA, and a sample entropy measure to gear defect detection. Xu et al. [74] used time-varying filtering for EMD and a high-order energy operator to identify bearing defect. Zhao et al. [73] applied an improved approach of orthogonal empirical mode decomposition to extract the fault feature of gearbox. Although EMD have a good performance in analyzing non-stationary signals, it also has some weakness such as end effects, sifting stop criterion, and mode mixing. Feng et al. [75] and Bokde et al. [70] reviewed the EMD algorithm and its problems and the corresponding solutions. Then, it described the possible research directions in the future.

### 3.3.5 Ensemble empirical mode decomposition

In the mode mixing problem, different frequency components are decomposed into the same IMF or the same frequency components are decomposed into different IMFs, which allow the EMD fails to represent the fault characteristics of a signal accurately. EEMD was developed from EMD by Wu and Huang to overcome the problem of mode mixing [76]. The principle of the EEMD can be given as follows. The added white noise of finite amplitude populates the whole time–frequency space uniformly, and a signal is provided to this background. The result of each trial is composed of a signal and an added white noise. In the ensemble mean of sufficient trials, the noise can be completely reduced due to the fact that it is different in separate trials. With more and more trials being added in the ensemble, the last remaining part is the signal component of interest [70, 83, 85].

To verify the effectiveness of EEMD in overcoming the mode mixing drawback, Lei et al. [85] applied EEMD and EMD to analyze a simulation signal, which is a sine signal attached by small impulses. Figure 6 shows that the decomposition result of EMD suffers from mode mixing. The sine signal and the impulses are decomposed into the same IMF, and the sine signal is decomposed into two IMFs. However, EEMD decomposed the simulation signal into two IMFs accurately, which succeeded in representing the real characteristics of the signal. Studies on EEMD applied to fault diagnosis have been increasing steadily in the past few years. Tabrizi et al. [76] developed a method based on EEMD and wavelet packet decomposition for fault detection of rolling element bearings. Zhang et al. [14] presented a novel hybrid model which combined EEMD and permutation entropy in motor bearing fault diagnosis. Luo et al. [77] proposed a hybrid system based on EEMD for fault diagnosis of rolling element bearings and noticed that the method had better classification accuracy than its original version. In order to

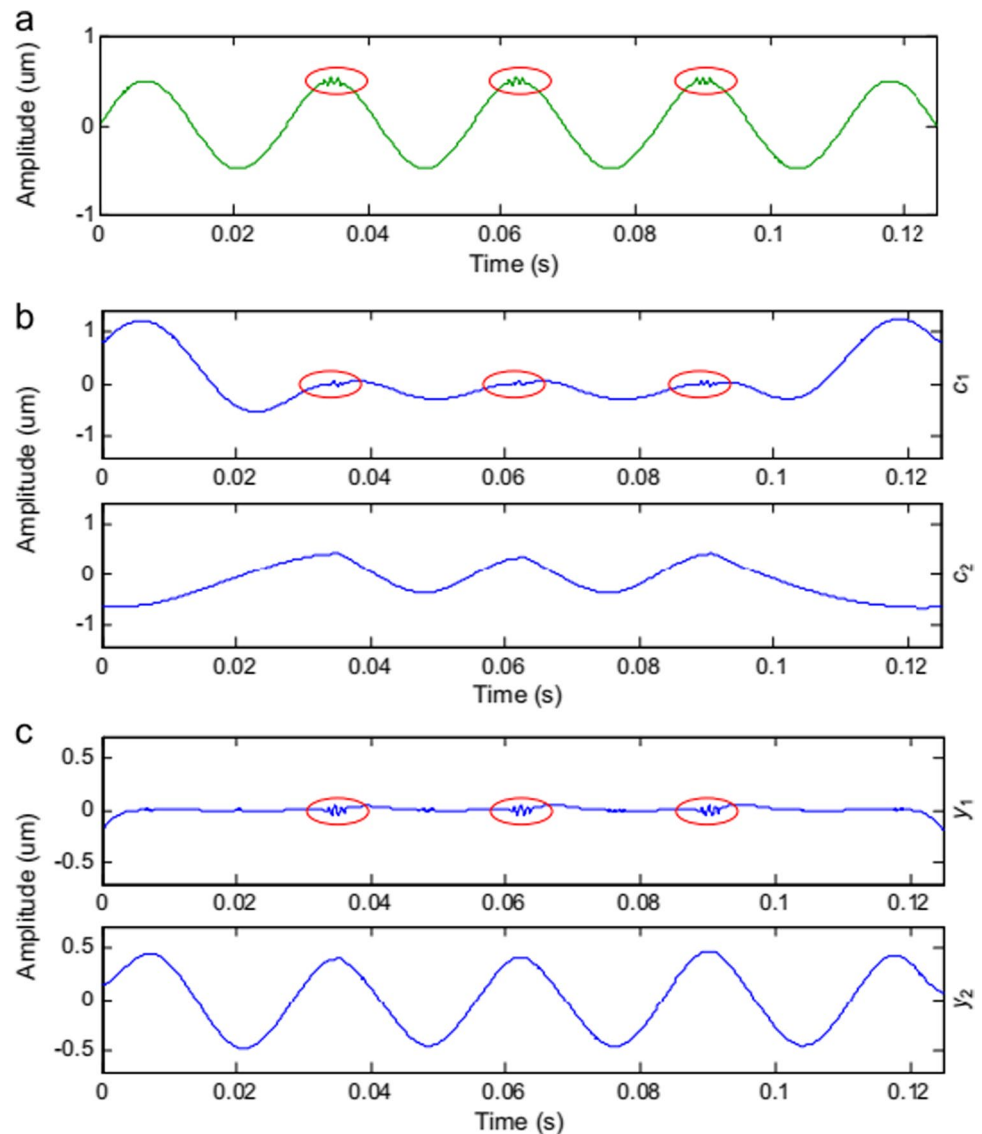
improve the accuracy of EEMD, two factors should be concerned: the number of ensemble trials and the amplitude of added noise. If the amplitude is too small, then it may not improve mode mixing. On the other hand, if the amplitude is too large, it will produce several redundant IMFs. Moreover, it should be noticed that too many trials would add burden to the computational procedure [14, 85]. However, there is still no available standard that can determine these parameters.

### 3.3.6 Local mean decomposition

LMD was proposed by Smith to analyze non-stationary signals in 2005, which has the similar principles as EMD [80]. LMD is applied to decompose a multicomponent signal into a sequence of product functions (PFs) with the physical significance. Each PF is the product of amplitude envelope signal and a pure frequency-modulated signal [81]. One main advantage of LMD is that it can directly obtain the instantaneous amplitude and the instantaneous frequency of the signal without Hilbert transform. Then, LMD does not use the cubic spline but uses smoothed local means and local magnitudes to fit lower and upper envelope in the iteration process, so as to avoid the envelope errors. Moreover, LMD obtained more concentrated information compared with EMD [79, 81].

Until now, LMD-based approaches have been widely applied in the field of condition monitoring of various machineries. Cheng et al. [78] used LMD to identify gear and bearing work condition. Then, LMD is compared with EMD and the results show the superiority of the LMD approach. Liu et al. [79] proposed a hybrid fault diagnosis approach using LMD and the second-generation wavelet de-noising to extract fault features of gearboxes and rolling bearings. However, like EMD, LMD still suffers from the end effects and mode mixing. The former can alleviate through extending the waveform based on spectral coherence [50]. To solve the problem of mode mixing of LMD, Wang et al. [80] applied the ensemble local mean decomposition (ELMD) for fault diagnosis of the gearbox and noted that the method achieved a good result. Essentially, ELMD is a noise-assisted LMD, using the statistical characteristic of Gaussian white noise to improve the distribution of extreme points in original signal. When the ensemble mean of PF components in each trial is calculated, the added noise in each PF can be eliminated automatically, thus solving the mode mixing problem. However, ELMD is limited by the number of added white leading to the fact that the noise cannot be canceled out completely, which further cause the growth in reconstruction errors. Based on the above problems, Wang et al. [81] added white noise in pairs to optimize ELMD for composite fault diagnosis of gearboxes.

**Fig. 6** **a** A simulation signal, **b** IMFs decomposed by EMD, and **c** IMFs decomposed by EEMD [85]



## 4 Diagnostics

Machinery unavoidably generates various faults due to long-term operating under complex and severe condition such as heavy load, high speed, and corrosion. Accurate fault diagnosis of the machinery is important to avoid or minimize the unplanned breakdown and catastrophic accidents. Fault diagnosis is also called pattern recognition, which is a procedure of mapping the information obtained in the feature space to machine faults in the fault space [12]. In recent decades, using fault diagnosis approaches to monitor the health conditions of machines has attracted much attention. A lot of technologies have been developed. Unfortunately, it is difficult for users to assess each specific method and its variants, respectively. Motivated by this, several publications [12, 86–90] reviewed a large volume of effective fault diagnosis methods and separated them

into various categories for discussing collectively similar techniques as shown in.

Table 4. Through studying and comparing the meaning and coverage of different categories from these papers, we classify the fault diagnosis techniques into three categories: physical models, knowledge-based models, and artificial intelligence models. Table 5 illustrates the advantages and disadvantages of some commonly used fault diagnosis approaches in order.

### 4.1 Physical models

Physical models quantitatively characterize the behavior of a failure mode using physical laws, which implies a thorough understanding of the system behavior in response to stress, at both macroscopic and microscopic levels [90]. Through quantifying the differences between



**Table 4** Classification approaches related to fault diagnosis

References	Years	Categories of fault diagnosis approaches
Jardine et al. [12]	2006	<ul style="list-style-type: none"> <li>• Statistical approaches</li> <li>• AI approaches</li> <li>• Other approaches</li> </ul>
Cibulka et al. [86]	2012	<ul style="list-style-type: none"> <li>• Model-based approaches</li> <li>• Signal processing-based approaches</li> </ul>
Gao et al. [87]	2015	<ul style="list-style-type: none"> <li>• Model-based methods</li> <li>• Signal-based methods</li> <li>• Knowledge-based methods</li> <li>• Hybrid methods</li> <li>• Active fault diagnosis methods</li> </ul>
Xu et al. [88]	2017	<ul style="list-style-type: none"> <li>• Knowledge-driven methods</li> <li>• Data-driven methods</li> <li>• Value-driven methods</li> </ul>
Wei et al. [89]	2019	<ul style="list-style-type: none"> <li>• Fault frequency-based methods</li> <li>• Artificial intelligence-based methods</li> </ul>
Dai et al. [90]	2019	<ul style="list-style-type: none"> <li>• Model-based methods</li> <li>• Knowledge-based methods</li> <li>• Signal-based methods</li> </ul>

measurements from the real process and the outputs of the model, the condition of a fault can be identified accurately. The main advantage of physical model is the ability to provide most accurate condition estimation when the model is developed with complete knowledge of system behavior and appropriate parameters. In addition, the output of physical model can be easily understood. Each physical model is usually specific to an application, so this paper does not further subdivide physical models. Jung et al. [91] combined model-based residuals and incremental anomaly classifiers to identify unknown faults of internal combustion engine. Gao et al. [92] developed a physical model based on coil sub-element for the fault detection of winding short-circuit in a direct-drive permanent magnet synchronous motor. They can evaluate motor performances under various winding short-circuit faults without changing the internal structure of the model. However, for a complex system, it is difficult to obtain an accurate model due to detailed and complete knowledge of system behavior required. Another weakness of physical model is that different models need to be established for different applications [87, 90]. Therefore, the application in fault diagnosis of physical model is limited.

## 4.2 Knowledge-based models

Knowledge-based approaches evaluate the similarity between the observed condition and a database of previously defined failures to deduce the health of machinery [93]. Sub-categories are separated into expert systems and fuzzy systems.

### 4.2.1 Expert systems

Expert system takes advantage of the computer to solve the complex problems normally solved by experts, which generally consists of a knowledge base and an inference engine. The knowledge base contains all facts, procedures, and rules (e.g., precise IF–THEN statements), which are accumulated through experience from one or more experts over a number of years [94, 95]. The inference engine used the knowledge base to analyze each case. Owing to the merits such as simple to development, easy to understand and transparent reasoning, expert systems have been successfully used in fault diagnosis. For example, Hussain et al. [94] developed an expert system to diagnose power circuit breakers and on-load tap changers. They found that the expert system can not only identify the health of the testing device but also locate the cause of each anomaly. Xu et al. [96] proposed a new belief rule-based expert system to identify fault modes that may co-exist in marine diesel engines. Xu et al. [95] combined expert system and Bayesian network to analyze fault types of generation system. In order to be powerful, expert systems must offer only one output for each set of any possible combination of inputs. However, as the growth of the number of inputs and expected outputs, the number of rules required also increases, which can lead to “combinatorial explosion.” Moreover, the knowledge base needs to be updated with increasing knowledge are obtained [93, 97].

### 4.2.2 Fuzzy systems

Fuzzy systems also apply IF–THEN rules obtained from knowledge of experts to solve problems, but unlike expert systems that use true or false as a logic to precisely define sets and related membership, they partition a feature space into fuzzy sets and use imprecise rules for reasoning [87, 88]. Since one fuzzy rule can replace a large number of conventional rules, fuzzy systems need fewer rules to achieve inference than expert systems. Fuzzy logic is rarely used as the main approach for fault diagnosis, which is usually combined with other methods for improving the performance of diagnostics. Adaptive neuro-fuzzy inference system (ANFIS) can take full advantage of the learning ability of neural network and inference ability of fuzzy logic and its fuzzy membership functions and rules are obtained by self-learning rather than reliance on experience [90, 98, 99]. Chen et al. [98] applied fuzzy entropy of local mean decomposition and adaptive neuro-fuzzy inference system to classify fault patterns of planetary gears. They used fuzzy entropy to reflect the complexity and irregularity of each PF components. Parey et al. [99] introduced a method based on adaptive neuro-fuzzy inference system for a single stage spur gearbox damage diagnosis.



**Table 5** Advantages and disadvantages of diagnostic techniques

Categories	Techniques	Advantages	Disadvantages
Physical models	Physical models	<ul style="list-style-type: none"> <li>• Provide most accurate and precise estimates among all diagnosis techniques</li> <li>• Outputs can be easily understood</li> </ul>	<ul style="list-style-type: none"> <li>• Require detailed and complete knowledge of system behavior</li> <li>• Provide specific models for different applications</li> <li>• Difficult to establish an accurate model for a complex system</li> </ul>
Knowledge-based models	Expert systems	<ul style="list-style-type: none"> <li>• Develop the knowledge base by human experts</li> <li>• Easy to implement</li> <li>• Provide a transparent reasoning process</li> </ul>	<ul style="list-style-type: none"> <li>• Rely entirely on knowledge of human experts</li> <li>• Require precise input and a lot of rules</li> </ul>
	Fuzzy systems	<ul style="list-style-type: none"> <li>• Develop the knowledge base by human experts</li> <li>• Easy to implement</li> <li>• Provide a transparent reasoning process</li> </ul>	<ul style="list-style-type: none"> <li>• Rely entirely on knowledge of human experts</li> <li>• Require precise input and a lot of rules</li> </ul>
Artificial intelligence models	KNNs	<ul style="list-style-type: none"> <li>• Have a simple architecture</li> <li>• Easy to implement</li> </ul>	<ul style="list-style-type: none"> <li>• Have high computational complexity for high dimension</li> <li>• No clear guidelines on the selection of <math>k</math> value and distance metric</li> </ul>
	ANNs	<ul style="list-style-type: none"> <li>• Can model complex, multi-dimensional, and non-linear systems</li> <li>• Have high performance in approximation, classification, and noise immunity by non-linear information processing</li> <li>• Provide condition information of machinery without a priori knowledge</li> <li>• Can use any type of input data</li> <li>• Remarkable ability in multivariate analysis</li> </ul>	<ul style="list-style-type: none"> <li>• Suffer from vanishing or exploding gradient problems</li> <li>• Require a lot of training data</li> <li>• Not clear about how decisions are made in a network</li> <li>• Require largely trial and error to determine the most appropriate model</li> <li>• No standard method to determine optimal structure and parameters</li> </ul>
	SVMs	<ul style="list-style-type: none"> <li>• Can obtain robust and accurate results with high dimension or non-linear input data</li> <li>• Suitable for small samples and real-time analysis</li> <li>• Can guarantee good generalization performance</li> </ul>	<ul style="list-style-type: none"> <li>• No standard method to choose the kernel function</li> <li>• Need to adjust parameters for specific problems</li> <li>• Classification and regression results by SVM as point estimates only</li> </ul>
	Deep learning models	<ul style="list-style-type: none"> <li>• Can learn features automatically</li> <li>• More capable of learning complex structure from data due to the deep architecture</li> <li>• Do not need the feature extractor</li> </ul>	<ul style="list-style-type: none"> <li>• Require a large volume of data</li> <li>• Take a long time for training</li> <li>• No physical meaning</li> </ul>

### 4.3 Artificial intelligence models

The Fourth Industrial Revolution and the industrial Internet have enhanced the importance and complexity of machinery, which make it difficult to maintain the diagnosis accuracy and ensure the sensitivity to faults by physical models and knowledge-based models [100]. Artificial intelligence approaches apply the information from previously collected data to monitor the current damage state and estimating the future trend instead of building models based on the expert experience or failure mechanisms. The most popular AI techniques in the field of fault diagnosis include k-nearest neighbor (KNN), artificial neural network (ANN), support vector machine (SVM), and deep learning.

#### 4.3.1 k-Nearest neighbor

The KNN is one of the most simply implemented AI algorithms, which classifies the objects according to the principle that the instances within a dataset will usually exist in close proximity to other instances with similar properties [26]. Key elements to KNN algorithm include the following: the value of  $k$  and the distance metric, which may greatly affect the algorithm performance. The most commonly used distance metric is Euclidean distance, and other distance functions such as Manhattan, Mahalanobis, and Minkowski can also obtain similar results [52, 101, 102]. However, KNN has high computational complexity for high dimensional data because of the fact that the performance of KNN relies on the number of dimensions.

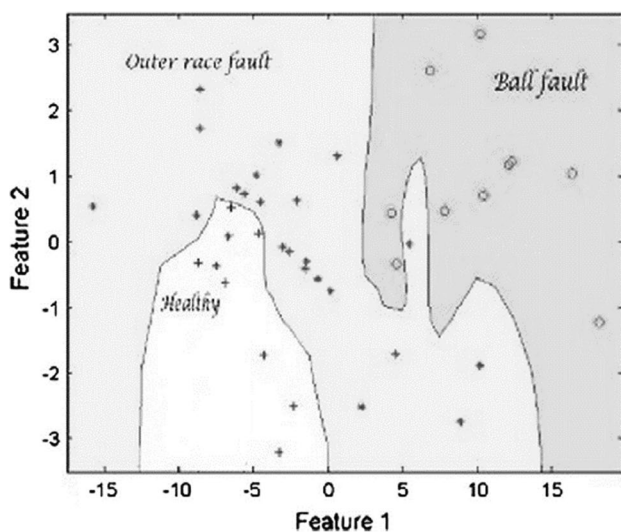
The KNN has been applied in the fault diagnosis. Safzadeh et al. [26] used KNN to identify the condition of bearing based on vibration signal and load signal. As shown in Fig. 7, three real states of bearing were analyzed: bearing healthy condition, ball fault, and outer race way fault. It can be seen that KNN is useful to detect the position of the faults. Toma et al. [101] used genetic algorithm to select the most optimal features to reduce complexity of KNN. They compared properties of KNN, decision tree, and random forest by the motor current signals, and observed that all classification algorithms successfully evaluate the bearing faults. Glowacz et al. [102] applied KNN, K-means clustering, and the linear perceptron to detect early stator faults in a single-phase induction motor. Islam et al. [103] combined KNN and genetic algorithm to identify the bearing fault condition. They found that the proposed method outperforms the existing average distance-based methods with regard to classification accuracy.

### 4.3.2 Artificial neural network

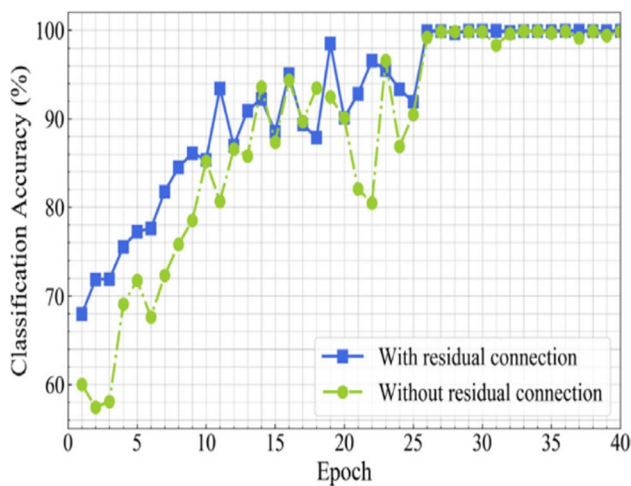
ANN is a computational model that mimics the working process of human brains which has powerful pattern classification and faults recognition capabilities. Network architectures applied for fault diagnosis can be separated as follows: (a) static (i.e., feed-forward) network in which the inputs for each layer only rely on the outputs of the previous layer and (b) dynamic network in which the inputs to a specific layer depend on the outputs of the previous layer and the previous iterations of the network itself [97]. Most ANN approaches proposed to date have been based on static networks, including the multi-layer perceptron (MPL) [52, 104–106], radial basis function network (RBF)

[97, 107, 108], and general regression neural network (GRNN) [25]. Several dynamic networks (e.g., recurrent neural network) have been developed for fault diagnosis. Since recurrent neural network (RNN) store temporal information by the additional feedback in the form of time-delayed inputs, it is suitable to apply both the historical conditions and sensing data to diagnose with low model complexity [97, 109].

Most papers applied a MPL to monitor the condition of machinery. Moosavian et al. [110] extracted features from the power spectral density values of signals, and compared the performance of ANN and KNN in fault diagnosis of bearing. Han et al. [104] used improved EMD and a MPL neural network to improve the performance of fault diagnosis. Glowacz [105] used the nearest neighbor classifier, backpropagation neural network (BPNN) and modified classifier based on words coding to identify the real state of a three-phase induction motor by acoustic signals. Taimoor et al. [106] exploited the extended Kalman filter to update the weight parameters of MLP neural network for improving the fault diagnosis capabilities, which is applied to detect an aircraft actuators and sensors fault. Compared with MLP, RBF trains quicker than MPL. Zhou et al. [107] combined unscented Kalman filter and RBF to detect fault in the pumping unit. Jin et al. [108] applied radial basis function neural network with power spectrum of Welch method to bearing fault diagnosis and further discussed the limit performance of the neural network. RNN outperforms MPL and RBF due to the ability to consider temporal dependencies via local or global feedback connections in the network. An et al. [111] proposed a novel model based on RNN and transfer learning to classify the health condition of bearing, which has the ability of processing variable size sequences under different working conditions. However, one main limitation of RNN is that it is difficult to store information for a long time. To solve this problem, researches have proposed some variants such as long short-term memory networks [112, 113] and echo state networks [114]. ANNs have a good performance in approximation, classification, and noise-immunity of complex systems. However, they still have their own problems. As the depth of the network increases, neural networks easily suffer from vanishing or exploding gradient problems, which causes the training process difficult to converge. To alleviate this issue, Zhang et al. [115] employed RNN with residual connection to learn representative features. They compared the performance of the RNN with and without residual connection. As shown in Fig. 8, the method with residual connection provides higher classification accuracy in almost every epoch, and residual connection make the training process more stable in the last few epochs, which contributes to improve the accuracy of fault diagnosis. ANNs need a large amount of training data, but it is not always available in reality [97]. Moreover, standard



**Fig. 7** Illustration of class borders for KNN classification on bearing condition [26]



**Fig. 8** Effect of residual connection on classification accuracy [115]

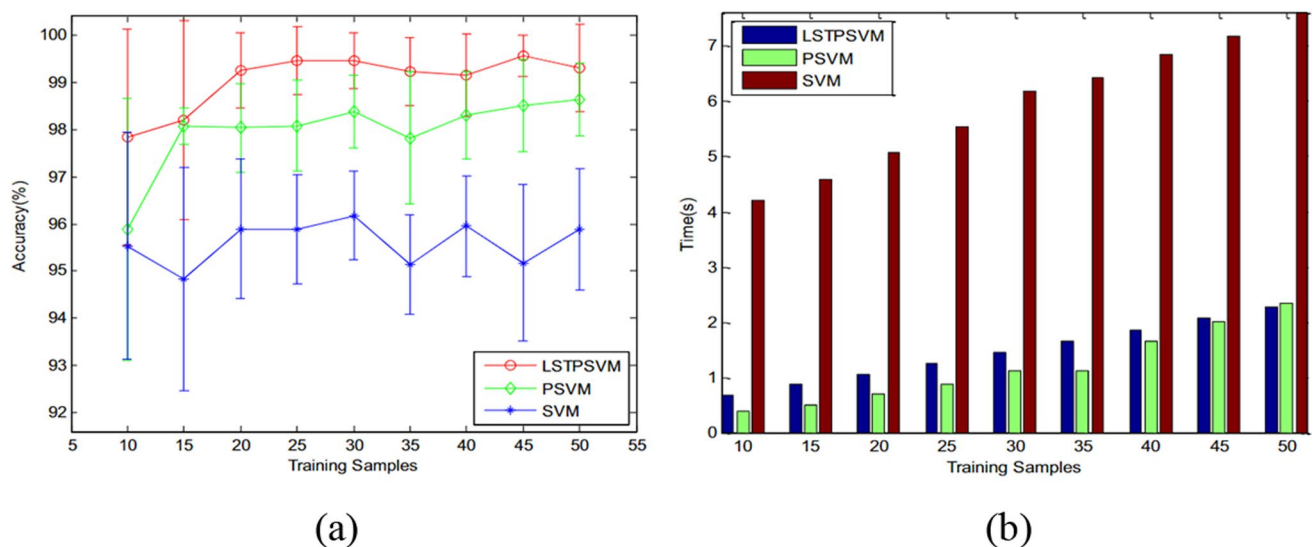
method of choosing ANN's structure and parameters is still a challenge.

### 4.3.3 Support vector machine

SVM is an AI method based on statistical learning theory proposed by Vapnik in the early 1990s, which attempts to find an optimal separating hyperplane with the maximum distance between the plane and the nearest data to classify the data [52, 109, 116]. In the standard SVM, it is assumed that the data is divided into two classes: positive and negative. Users can refer to [25, 117] for more fundamentals and the basic about the standard SVM.

SVM models have been successfully applied to fault diagnosis in existing publications. Singh et al. [118] applied

Stockwell transform and SVM to detect bearing fault in a three-phase induction motor. Han et al. [119] proposed a fault diagnosis method based on the improved Fast-ICA algorithm, the wavelet packet energy spectrum and SVM, which is used to recognize the slight damage and fracture of a bearing. Several improved SVMs have been developed to satisfy the demands in real application. Ma et al. [120] proposed a novel algorithm based on the scattering transform and the least squares recursive projection twin support vector machine (LSPTSVM). The algorithm can overcome problem of traditional approaches which are noise sensitive in feature extraction. They compared the performance of this algorithm with the proximal support vector machine (PSVM) and SVM in bearing fault diagnosis. The results are shown in Fig. 9; for various numbers of training samples, LSPTSVM provides the highest accuracy and the smallest variance. Moreover, the calculation time of LSPTSVM is close to that of PSVM and approximately 1/4 that of traditional SVM. It means that the classification performance of LSPTSVM is better than that of PSVM and SVM. Liu et al. [121] applied particle swarm optimization to optimize unknown parameters of wavelet support vector machine (WSVM) for monitoring the condition of rolling element bearings. They found that the WSVM achieved a greater accuracy than the traditional SVM. Xu and Chen [116] proposed an intelligent fault identification approach of bearings using improved least squares support vector machine (LS-SVM). Moosavian et al. [122] applied ANN and LS-SVM for fault diagnosis of spark plug in an internal combustion engine and used D-S evidence theory to increase the fault detection accuracy. SVM has better performance in terms of dealing with small simple size than ANN. However, when the number of training sample is small but the number of features is huge,



**Fig. 9** Classification accuracy (a) and computation times (b) of LSTPSVM, PSVM, and SVM [120]

then it is not necessary that all available features are of equal importance in the classification context. Ghosh et al. [123] proposed a method based on import vector classifier to select an optimal set of features, so as to get good classification performance. Huang et al. [124] imports feature clustering to enhance support vector machine recursive feature elimination for feature selection. Jalalian et al. [125] used potential SVM and Gaussian dynamic time warping to eliminate the fixed-length limitation of feature vectors in training data for enhancing classification performance. The performance of SVM is highly dependent on the selected kernel function, but standard methods of selecting kernel function have not been established. Then, more advanced search techniques should be developed to improve the simplicity and accuracy of parameter estimations [25, 109].

#### 4.3.4 Deep learning

Machine learning methods like ANN and SVM requires the feature is extracted and selected manually by users, which largely depends on the knowledge of signal processing and user experience. Nevertheless, it is difficult to know what features should be provided to model for a complex machinery. Deep learning algorithm is capable to overcome the problem mentioned above, which automatically learn features by the deep architectures composed of multiple levels of non-linear operations, so as to allow a system to learn complex functions for mapping the input data to the output data directly with a small error [52, 126]. The difference of traditional machine learning and deep learning is shown in Fig. 10.

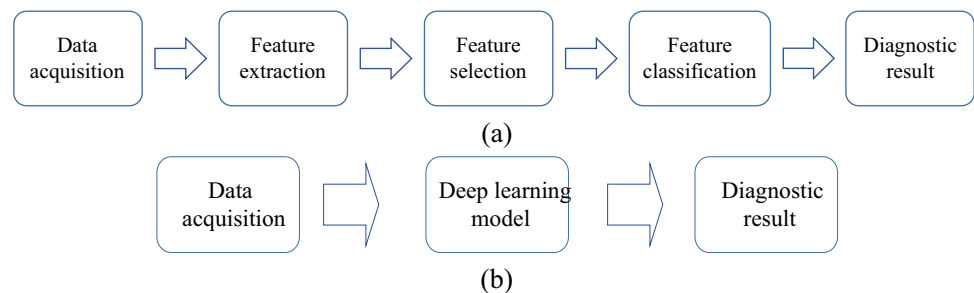
Recent models based on deep learning (e.g., autoencoder, deep belief network) have been proven successful in fault diagnosis. Autoencoder consists of two parts: encoder network and decoder network. The former transforms the input data to a low-dimensional space and the latter reconstructs the inputs from the corresponding codes [52]. Shao et al. [127] proposed a novel deep autoencoder feature learning approach for monitoring faults of gearbox and roller bearing. Deep belief network (DBN) is constructed by multi-layer restricted Boltzmann machine, which can be used to approximate complex nonlinear function with small error.

Tao et al. [128] applied deep belief network to adaptively fuse multi-feature data and diagnose various bearing faults. The result showed that the method based on DBN obtained higher identification accuracy than SVM, KNN, and BPNN. Tran et al. [129] investigated DBN for diagnosing valves faults in reciprocating compressors and noticed that their approach was more powerful than relevant vector machine and backpropagation neuron networks. However, several powerful deep learning models require fixed-size inputs like images. Qin et al. [130] proposed a novel deep learning framework, namely, attention-based discrete sequence anomaly detection, to extract features from a series of relatively long and variable length sequence, in which the attention mechanism is used to improve the interpretability. Data plays a very important role in deep learning algorithms. The need of data is a challenge to implement deep learning in practical applications. Moreover, it is noticed that the deep architecture increases the number of parameters, thus generating the risk of over-fitting [131, 132].

## 5 Prognostics

Prognostics calculates the remaining useful life (RUL) of an asset based on condition monitoring information to provide sufficient lead time for maintenance planning. The RUL is defined as the time left before the health condition of asset crosses a failure threshold [133]. At present, the widely used methods for failure threshold determination are based on some ISO standards, such as the ISO7919 and ISO8688 series, or some standards specially designed for certain industries, such as VDI/3834 for wind turbines [25]. For example, according to the recommendation of the ISO 8688–2, Zhang et al. [134] predetermined the cutting tool life by the flank wears value. RUL prediction plays a significant role in a CBM program. A suitable prediction technology is expected to simplify the prognostic modeling and produce accurate prediction results. There are two major issues related to the RUL prediction: (1) how to forecast the RUL based on the condition monitoring data and (2) how to evaluate the predict accuracy of different technologies. Many technologies applied to RUL estimation have been published

**Fig. 10** **a** Traditional machine learning and **b** deep learning.





in recent years. As shown in There is no unified standard for RUL prediction metrics. According to different requirements of researchers and operators, various RUL prediction metrics have been developed to evaluate the prediction results from different aspects. Therefore, users are suggested to choose proper prediction metrics based on their own requirement. There are some useful metrics developed by [25, 139, 140]. This section will not discuss the prediction metrics.

Table 6, references [8, 12, 25, 97, 109, 135–138] have described commonly used RUL prediction methods and classified them into different categories based on different criteria. In order to avoid confusing readers in the classification, the existing prognostics models are divided into four main categories in this paper based on their basic techniques and methodologies: physical models, knowledge-based models, stochastic models, and AI models. These techniques have their respective advantages and disadvantages, as summarized in Table 7. The following sub-sections emphasize the basis and achievement of these techniques in recent years. There is no unified standard for RUL prediction metrics. According to different requirements of researchers and operators, various RUL prediction metrics have been developed to evaluate the prediction results from different aspects.

**Table 6** Classification approaches related to RUL prediction

References	Years	Categories of RUL prediction approaches
Jardine et al. [12]	2006	<ul style="list-style-type: none"> <li>• Statistical approaches</li> <li>• Model-based approaches</li> <li>• AI approaches</li> </ul>
Heng et al. [135]	2009	<ul style="list-style-type: none"> <li>• Traditional reliability approaches</li> <li>• Physics-based models</li> <li>• Data-driven models</li> <li>• Integrated approaches</li> </ul>
Sikorska et al. [97]	2011	<ul style="list-style-type: none"> <li>• Knowledge-based models</li> <li>• Life expectancy models</li> <li>• ANN</li> <li>• Physical models</li> </ul>
Xu et al. [136]	2013	<ul style="list-style-type: none"> <li>• Model-based approaches</li> <li>• Experience-based approaches</li> <li>• Data-driven approaches</li> </ul>
Lee et al. [137]	2014	<ul style="list-style-type: none"> <li>• Model-based approaches</li> <li>• Data-driven approaches</li> <li>• Hybrid prognostic approaches</li> </ul>
Kan et al. [109]	2015	<ul style="list-style-type: none"> <li>• Model-based methods</li> <li>• Data-driven models</li> <li>• Combination models</li> </ul>
Ben Ali et al. [8]	2015	<ul style="list-style-type: none"> <li>• Physical models</li> <li>• Expert systems</li> <li>• Data-driven models</li> </ul>
Bai et al. [138]	2016	<ul style="list-style-type: none"> <li>• Model-based approaches</li> <li>• Data-driven approaches</li> </ul>
Lei et al. [25]	2018	<ul style="list-style-type: none"> <li>• Physics model-based approaches</li> <li>• Statistical model-based approaches</li> <li>• AI approaches</li> <li>• Hybrid approaches</li> </ul>

Therefore, users are suggested to choose proper prediction metrics based on their own requirement. There are some useful metrics developed by [25, 139, 140]. This section will not discuss the prediction metrics.

## 5.1 Physical models

Physical models estimate the RUL of machinery by solving a deterministic equation or set of equations derived from extensive empirical data [97]. The parameters of physical models are generally identified by scientific knowledge and specific laboratory or field experimentation. The advantage of this method in prognostics is that it can provide confidence limits. Paris-Erdogan (PE) model is one of the most popular physical models in the RUL prediction of machinery, which was first used to identify the crack magnitudes in [25]. Then, various versions are proposed to estimate the RUL of machinery [97, 141, 142]. There are still some other physics model-based approaches in the field of machinery prognostics. For example, Hu et al. [143] proposed a RUL forecasting method based on the Norton law to study the degradation of turbine blade and aluminum electrolytic capacitor. Chao et al. [144] developed a novel method based on time-dependent crack growth models and used it for turbo propulsion systems condition prognostics. However, an industrial system generally has a complex structure that includes a lot of components. It is difficult to establish an accurate model to describe the behaviors of all potential components for a complex system.

## 5.2 Knowledge-based models

Knowledge-based models evaluate the RUL of machinery by using the specific expertise and experience in long-term accumulation. It is further grouped into expert systems and fuzzy systems.

### 5.2.1 Expert systems

Expert system is an experience-based system aiming to aid non-specialist users in evaluating the RUL of machinery by designing reasoning and decision-making mechanism. One main advantage of expert system is its ability to establish reasoning for a specific result. However, it is noticed that expert systems are not feasible to provide exact RUL output and confidence limits, which limits their application in the field of predicting RUL. Expert system is usually combined with other prognostic models to assess the RUL of industrial equipment. For example, Xiahou et al. [145] combines expert knowledge and condition monitoring data for RUL prediction of bearing under the belief function theory framework. Alamaniotis et al. [146] applied expert's experience



**Table 7** Advantages and disadvantages of prognostic techniques

Categories	Techniques	Advantages	Disadvantages
Physical models	Physical models	<ul style="list-style-type: none"> <li>• Provide most accurate and precise estimates among all prediction techniques</li> <li>• Outputs can be easily understood</li> <li>• Provide confidence limits</li> </ul>	<ul style="list-style-type: none"> <li>• Require detailed and complete knowledge of system behavior</li> <li>• Provide specific models for different applications</li> <li>• Difficult to establish an accurate model for a complex system</li> </ul>
Knowledge-based models	Expert systems	<ul style="list-style-type: none"> <li>• Develop the knowledge base by human experts</li> <li>• Establish reasoning for a specific result</li> <li>• Easy to implement</li> </ul>	<ul style="list-style-type: none"> <li>• Unable to provide exact RUL output and confidence limits</li> <li>• Rely entirely on knowledge of human experts</li> <li>• Require precise input and a lot of rules</li> </ul>
	Fuzzy systems	<ul style="list-style-type: none"> <li>• Require fewer rules than expert systems</li> <li>• Deal with imprecise or incomplete data</li> <li>• Provide confidence limits on the output</li> </ul>	<ul style="list-style-type: none"> <li>• Require domain experts to develop the fuzzy rules</li> </ul>
Stochastic models	PHMs	<ul style="list-style-type: none"> <li>• Easy to develop</li> <li>• Provide confidence limits</li> </ul>	<ul style="list-style-type: none"> <li>• Require historical data which is pertaining to failure modes</li> <li>• No standard method to choose parameters</li> <li>• Select parameters manually</li> </ul>
	HMMs and HSMMs	<ul style="list-style-type: none"> <li>• Suitable for modeling degradation process of non-linear and non-stationary system</li> <li>• Can model spatial and temporal data</li> <li>• Do not need specific knowledge of failure mechanism progression</li> <li>• Manage incomplete data sets</li> <li>• Provide confidence limits</li> </ul>	<ul style="list-style-type: none"> <li>• Require a lot of training data</li> <li>• Heavy computation workload</li> </ul>
	KFs	<ul style="list-style-type: none"> <li>• Basic KF is computationally efficient</li> <li>• Suitable for incomplete and noisy measurements</li> <li>• Variants can accommodate non-linear system</li> </ul>	<ul style="list-style-type: none"> <li>• Some variants diverge easily</li> <li>• Require Gaussian noise assumption</li> <li>• Variants is expensive in computational time</li> </ul>
	PFs	<ul style="list-style-type: none"> <li>• Can model multivariate and dynamic processes</li> <li>• Suitable for non-linear system, non-linear and/or non-Gaussian noise</li> <li>• Offer higher accuracy than other filtering algorithms</li> </ul>	<ul style="list-style-type: none"> <li>• Suffer from sample degeneracy</li> <li>• Require a lot of data and computation time</li> </ul>
Artificial intelligence models	ANNs	<ul style="list-style-type: none"> <li>• Can model complex, multi-dimensional, and non-linear systems</li> <li>• Have high performance in approximation, classification, and noise-immunity by non-linear information processing</li> <li>• Provide condition information of machinery without a priori knowledge</li> <li>• Can use any type of input data</li> <li>• Remarkable ability in multivariate analysis</li> </ul>	<ul style="list-style-type: none"> <li>• Require a lot of training data</li> <li>• Not clear about how decisions are made in a network</li> <li>• Require largely trial and error to determine the most appropriate model</li> <li>• No standard method to determine optimal structure and parameters</li> </ul>
	SVMs	<ul style="list-style-type: none"> <li>• Can obtain robust and accurate results with high dimension or non-linear input data</li> <li>• Suitable for small samples and real-time analysis</li> <li>• Can guarantee good generalization performance</li> </ul>	<ul style="list-style-type: none"> <li>• No standard method to choose the kernel function</li> <li>• Need to adjust parameters for specific problems</li> <li>• Classification and regression results by SVM as point estimates only</li> <li>• Difficult to provide the RUL directly</li> </ul>
	GPR models	<ul style="list-style-type: none"> <li>• Suitable for handling high-dimensional data and small size sample</li> <li>• Can model the degradation process of nonlinear dynamic systems</li> <li>• Allow non-parametric learning of a regression function from noisy data</li> </ul>	<ul style="list-style-type: none"> <li>• Heavy computational demand</li> <li>• Difficult to find optimal value of parameters</li> </ul>

and expertise to compensate for a potential lack of historical data for RUL prediction of power plant components.

### 5.2.2 Fuzzy systems

Fuzzy systems can deal with incomplete/imprecise data and complex systems. In the RUL estimation, fuzzy systems can provide confidence limits on the output. Majidian and Saidi [147] applied fuzzy logic and a neural network to forecast the RUL of boiler tubes. Although they found that the neural network was easier to develop, the prediction results of fuzzy system have more advantages. Researchers usually combine fuzzy logic with other techniques for getting better prognosis performance. Kang et al. [148] applied fuzzy evaluation-Gaussian process regression model to estimate the RUL in the case of limited data. Cheng et al. [149] developed a novel RUL prediction method using adaptive neuro-fuzzy inference system (AFNIS) and particle filtering to automatically detect gearboxes faults in wind turbines.

## 5.3 Stochastic models

Mechanical systems usually degrade stochastically. Therefore, their degradation processes can be modeled as stochastic processes. The uncertainty of machinery degradation processes is mainly caused by four variability sources: the temporal variability, the unit-to-unit variability, the nonlinear variability, and the measurement variability [150]. An appropriate stochastic process model is supposed to include the four variability sources simultaneously. At present, various stochastic models have been developed to estimate the RUL of machinery. This section focuses on several commonly used stochastic models in the field of RUL prediction.

### 5.3.1 Proportional hazards model

Proportional hazards model (PHM) was first proposed by Cox, which is one of the most popular models of prognostics [97]. It assumes that the hazard rate of a system is composed of two multiplicative factors, i.e., a baseline hazard function and a covariate function [25]. Qiu et al. [151] proposed a health indicator construction algorithm to characterize bearing degradation and applied support vector regression (SVR) and Weibull proportional hazards model to predict bearing RUL. Du et al. [152] built a PHM to calculate the failure risk of the lubricating oil. Man et al. [153] proposed a novel method based on a joint modeling framework to forecast the RUL of systems subject to hard failures. A Wigner process is applied to model stochastic degradation signals and the PH model is applied to model time-to-event data. Wang et al. [154] applied kernel principal component analysis and Weibull proportional hazards model to assess the reliability of bearings. The result verified the effectiveness of the

method in predicting the machinery RUL. Other application can be found in the following Refs. [155–157]. For a PHM, however, it is difficult to obtain sufficient data that contained rich information about machinery condition. Additionally, PHM needs other methods (e.g., Markov model) to describe the covariate functions, which further increases the computation workload [195].

### 5.3.2 Hidden Markov models and semi-hidden Markov models

Hidden Markov model (HMM) is a stochastic method based on the principle of Markov chains for modeling signals that evolve through a finite number of states [109]. Compared with Markov model, not all states in HMM can be observed directly, so the corresponding transition probability cannot be assigned directly. An HMM is characterized by the following: (1) the number of states in the model; (2) the number of distinct observation symbols per state; (3) the state transition probability distribution; (4) the observation symbol probability distribution in state; and (5) the initial state distribution [97]. HMM was first applied to RUL prediction by Bunks et al. [158]. The main advantages of HMM is that it enables modeling of both spatial and temporal phenomena. If the number of states is enough, HMM can classify time series data without expertise. In addition, HMM is applicable to nonlinear and non-stationary systems [97]. Therefore, HMM is obtaining more attention in recent years [97, 109, 159–161]. Du et al. [159] proposed a technique based on HMM for estimating the RUL of lubricant oil. They assumed that the process of lubricant oil degradation can be modeled by a HMM with three states. Soualhi et al. [160] developed a probabilistic approach based on HMM to model the degradation states of a system for the prediction of impending faults. They suggested that a prognostic method is not only limited to the prediction of RUL but is also improved to estimate the risk of future failure. Tao et al. [161] combined long short-term memory network and hidden Markov model to calculate the RUL of tool and the associated confidence interval. However, the Markov chain assumptions limit the practicability of the technology. Hidden semi-Markov model (HSMM) is an improvement to the HMM, which is not bound by the assumption of the Markov chain. Moreover, it allows for the modeling of state duration with an explicit distribution that need not be exponential and thus is more powerful in estimation RUL [97]. Zhu et al. [162] developed an improved HSMM by learning the duration parameters and RUL distribution database and used it for RUL prediction of tool. Liu et al. [163] applied a HSMM to obtain estimation of degradation state and the distribution of RUL and proved its high accuracy in tool wearing diagnosis and prognosis. One main weakness of all forms of Markov model is heavy computation workload, even for the simplest models with

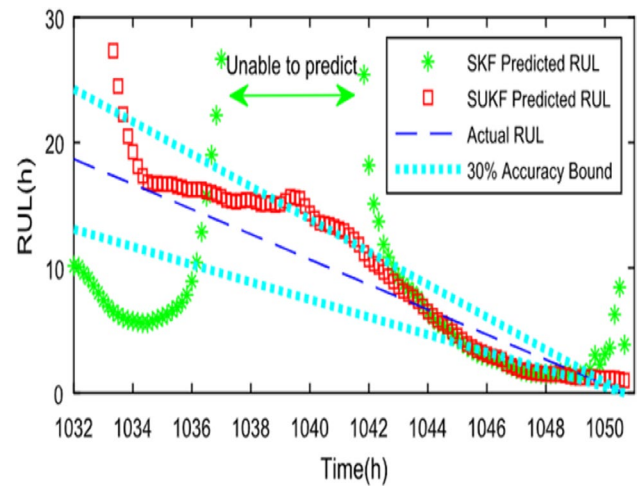
few states. Another limitation on the application of Markov model is the selection of state sequence and parameters [97].

### 5.3.3 Kalman filter

Kalman filter is an effective recursive digital processing technique, which is applied to estimate the state of a dynamic system by minimizing mean squared error from a series of incomplete data with noise [97]. Kalman filter has been extensively used with satisfactory performance in the field of RUL prediction [150, 164, 165]. Traditionally, Kalman filter is used to describe a linear degradation process. Extended Kalman filter (EKF) is the most popular Kalman variant in the state estimation of non-linear system. It performs state estimation through the linearization procedure of local approximation to the current mean and covariance [109]. However, EKF transforms the noise into non-Gaussian and thus invalidating one of the original assumptions, so it performs poorly when trying to approximate non-Gaussian processes. In addition, EKF is expensive in computational time because all covariance and model parameters are required to be recalculated in each iteration and then the filter can also diverge easily [97]. To solve problems mentioned, several modified versions of EKF has been proposed, such as unscented Kalman filter (UKF) and Monte-Carlo Kalman filter (MCKF) [166, 167]. Cui et al. [168] proposed a novel approach based on switching unscented Kalman filter (SUKF) for bearing RUL prediction. They selected measurement error as the standard deviation of RMS in the degradation stage in order to make the filtering results of condition monitoring data smoother. Figure 11 presents the results of RUL prediction. It can be seen that traditional switching Kalman filter (SKF) cannot predict RUL during the period when the condition monitoring data shows a downward trend. However, the results provided by SUKF are very close to the actual RUL and most of the prediction results fall within the 30% accuracy bound. Unfortunately, these variants are still limited by the Gaussian noise assumption.

### 5.3.4 Particle filter

Particle filter (PF) is an alternative to Kalman filter for estimation the posterior distribution in Bayesian network models, which is applicable to non-linear system with non-Gaussian noise, and it does not degrade the filter performance [109]. With enough samples, the forecasting accuracy of PF is higher than either the EKF or UKF. Because of these advantages, there has been great interest in using the PF method to predict the RUL of machinery. Li et al. [169] proposed a dual filter prediction method based on LSSVM and unscented particle filtering (UPF). To identify the effectiveness of the proposed dual filters fusion method



**Fig. 11** SUKF and SKF predicted RULs compared with actual RUL [168]

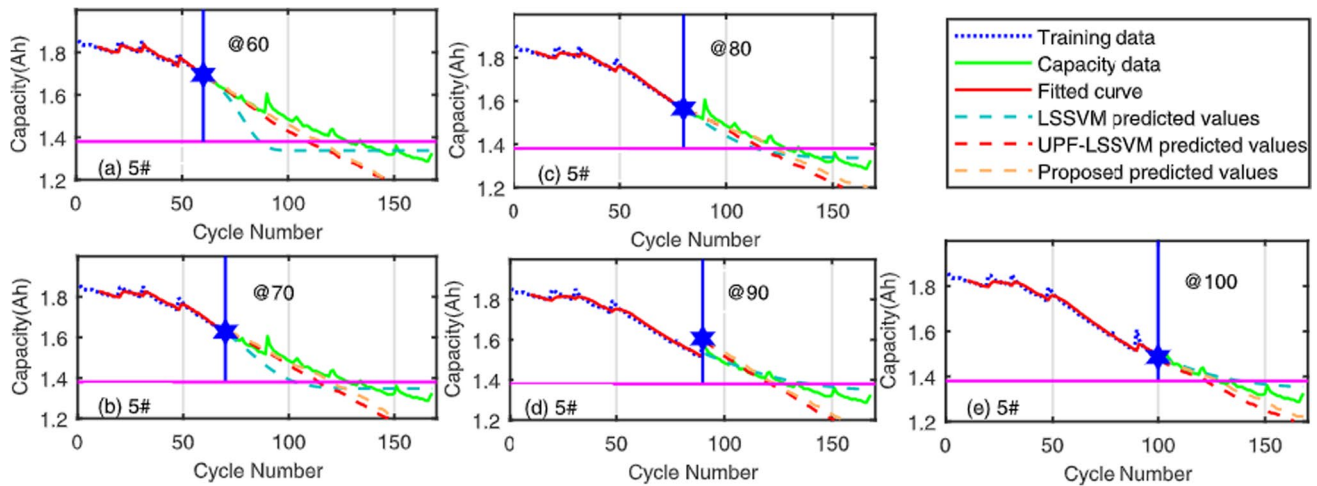
for lithium-ion battery RUL prediction, they compare it with UPF-LSSVM and LSSVM. Figure 12 shows the RUL prediction results at different prediction starting points. It can be seen that the dual UPF-LSSVM provides smaller predicted RUL fluctuation and prediction error with the increase of training data. Qian et al. [170] applied an enhanced particle filter method to predict RUL of rolling bearings and compared the performance of this method with traditional particle filter and SVR. A detailed discussion on the application of RUL prognosis using particle filters is given in [97, 171]. Unfortunately, PF suffers from one long-standing limitation: sample degeneracy. After a few iterations in the particle propagation process, the weight will concentrate on a few particles only and most particles will have the negligible weight [172]. Consequently, the particles degrade into poor distribution and lots of computational resources are wasted as to update particles that are ineffective for condition estimation. Researchers have proposed two commonly used approaches to address the problem: resampling and selection of importance density function [97, 172, 173].

## 5.4 Artificial intelligence

AI attempts to characterize the machinery degradation patterns using implicit information obtained by signal processing or data mining, so as to make maintenance decisions automatically. Many AI techniques have been used to predict RUL of machinery over the few years.

### 5.4.1 Artificial neural network

ANNs can directly or indirectly calculate the RUL of machinery by observing information without specific knowledge of the problem. Among various types of networks,



**Fig. 12** The RUL prediction results with different prediction starting point and prediction methods [169]

feed-forward neural network (FFNN) is the most popular of all kinds of neural networks. Most papers used FFNN to learn the relationship between the health indicators and the RUL [174–176]. Zhang et al. [177] used correlation analysis methods to extract the indicators of health status from the partial incremental capacity curves and established two ANN to estimate the state of health and RUL of battery synchronously. The RUL prediction results are shown in Fig. 13. The real RUL is represented by the dotted line, and the predicted RUL is represented by the solid line. As we can see, they show good consistence with each other. It means that the ANN model can provide great generalization ability and high accuracy for RUL estimation. Bastami et al. [178] used wavelet packet transform to extract features and MPL to predict RUL of bearings. General regression neural network is incredibly fast to train, and it can be used to estimate a continuous distribution. Huang et al. [179] applied the genetic algorithm/GRNN for proactive assessments of lifetime of a wafer-handling robot arm. Compared with RBFN and MPL, RNN has stronger ability to process nonlinear dynamic information by local/global feedback connection in the network. Zhang et al. [180] used a long short-term memory recurrent neural network to learn the long-term dependencies among the degraded capacities of lithium-ion batteries for estimating the RUL.

#### 5.4.2 Support vector machine

SVM was originally applied in the research of pattern recognition and was not used to the nonlinear regression estimation and time series prediction until the introduction of Vapnik's  $\epsilon$  insensitive loss function [109]. Various modified versions of original SVM algorithm have been developed for RUL prediction. Maior et al. [181] applied EMD and wavelet transform to improve input quality and used particle swarm

optimized support vector machines to prediction the RUL of bearing. Dong et al. [182] proposed a novel approach based on principal component analysis and least square SVM to achieve bearing degradation prediction. Ordonez et al. [183] combined an auto-regressive integrated moving average (ARIMA) model and SVM to predict the RUL of aircraft engines. SVR is the common application form of SVM in the field of prognostics. Benkedjough et al. [184] proposed a bearing life prediction method with the combination of SVR and the isometric feature mapping reduction technique. From Fig. 14, it is noticeable that the RUL in the middle of predictions is under the real RUL value, so this method is suitable for making maintenance interventions before the real time of a failure. Other applications of SVR can be found in [185–187].

#### 5.4.3 Gaussian process regression

Gaussian process (GP) is defined as a cumulative damage process of random variables with joint multivariate Gaussian distribution, which can achieve non-parametric learning of regression functions from noisy data [25, 109]. Gaussian process regression (GPR) is one of the applications of GP. It weights targets with respect to distance between training and test input to predict the output. In contrast to approaches mentioned, GPR is suitable for dealing with the RUL estimation issue of small data sets and multi-dimensional operating space [188]. Aye et al. [189] used affine mean Gaussian process regression (AMGPR) to predict the RUL of slow speed bearing. This method provided an excellent fit to data by an integration of simple mean and covariance functions. Figure 15 presents that the AMGPR traced the actual whole life of the bearing quite closely and most prediction results all within 95% confidence interval. Jia et al. [190] combined GPR and probability predictions to predict

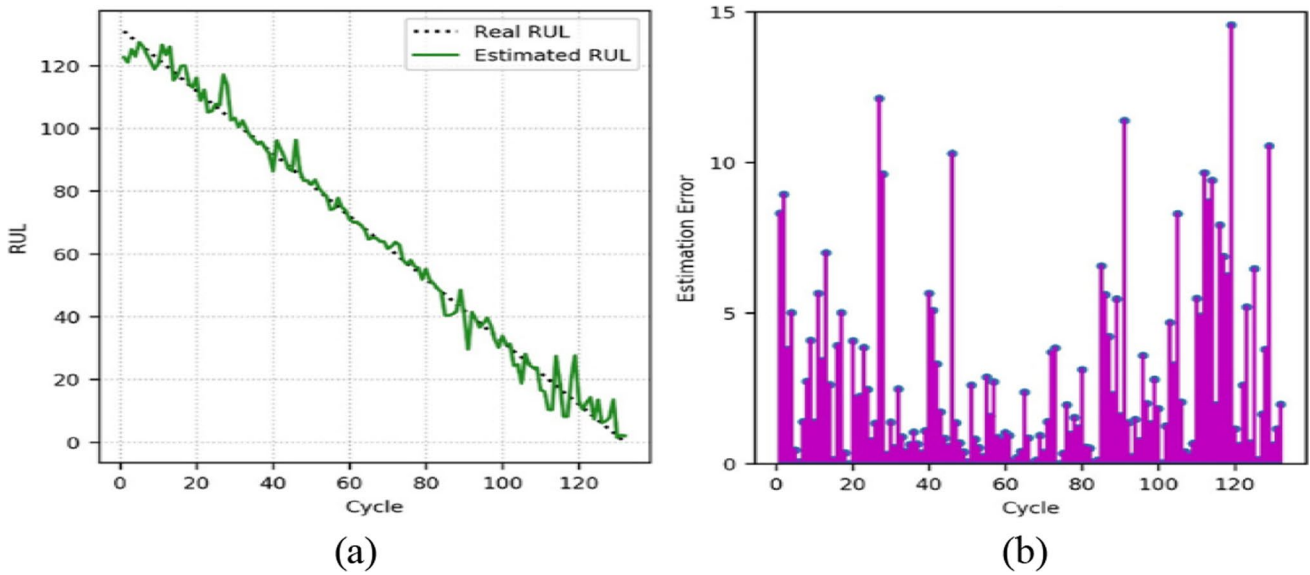


Fig. 13 The RUL prediction results of ANN model [177]

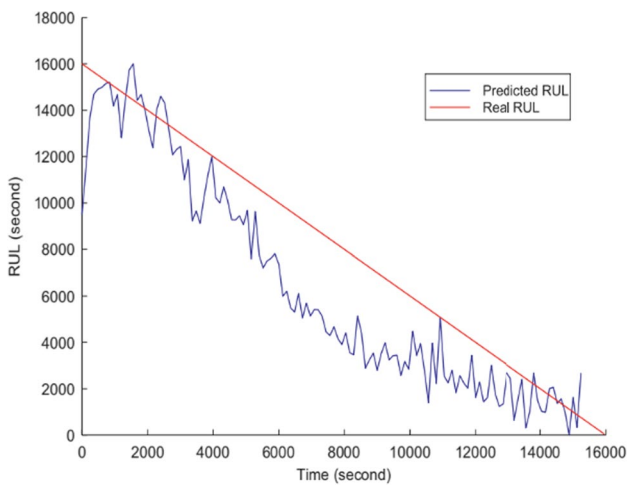


Fig. 14 RUL prediction result of a bearing based on SVR [184]

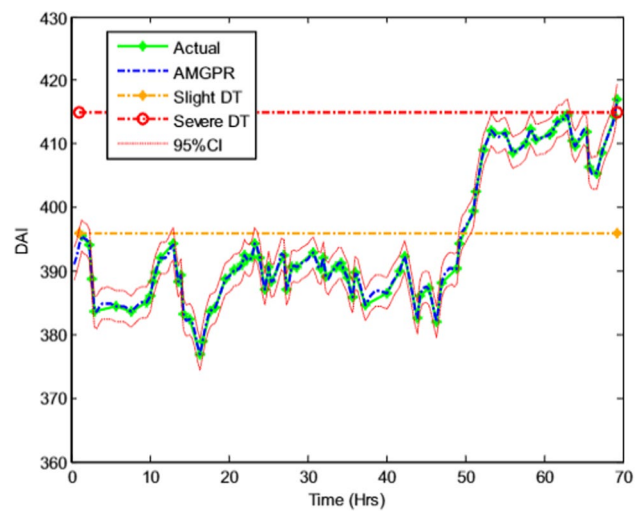


Fig. 15 RUL prediction result for bearing with 95% CI [189]

the short-term state of health of lithium-ion batteries and applied GPR model to assess the RUL by the mapping relationship between state of health and RUL. Kang et al. [148] used Gaussian process regression based on fuzzy evaluation to achieve RUL estimation of the lithium battery. They found that the proposed approach can avoid over-fitting in the case of finite data. Ismail et al. [191] proposed a GPR-based method to extract the degradation behavior of insulated gate bipolar transistor for faults prognosis. Readers can find more theoretical details and applications of this approach in Refs. [109, 192, 193]. One major problem of GPR is the huge computational demand due to its non-parametric nature.

All in all, the aforementioned techniques for RUL prediction will be subject to available training data, which is the

key for the success of a CBM system. Generally, useful data sets are very limited in many applications, although many numerical simulation and experimental methods have been introduced to generate as much as possible data sets. Existing technical limitations of the most solutions in practice still remains in the following aspects: (a) RUL prediction where limited data are available; (b) RUL prediction under the big-data situation; (c) RUL prediction of a single component involving multiple faults; (d) how to manage the uncertainties in RUL prediction; and (e) in the case when you have little information environment, a little time, you have few numbers of data to live in Environment. These challenging



tasks must be addressed in the future to develop practical RUL prediction techniques.

## 6 Commercialization of fault diagnosis system

Fault diagnosis technology originated in the end of the 1960s, so as to reduce the incidence of mechanical fault through improving design methodology and to meet improved reliability, greater safety, and financial savings. The USA established mechanical fault prevention group (MFPG) in 1967 to solve many accidents caused by machinery failures since the Apollo program. The purpose of MFPG is to effectively interchange technical information among segments of scientific and engineering communities for getting a better understanding of the processes of mechanical failures [194]. Then, the UK established machine health and condition monitoring association (MHMG & CMA) to study the fault diagnosis technology in the 1960s–1970s, mainly focused on friction and wear, automotive and aircraft generator monitoring, and diagnosis. Japan developed productive maintenance since 1971 and reached a leading position in the steel, chemical, and railway fields [12, 50, 195]. Other countries also gradually pay attention to the research of fault diagnosis, e.g., study on the failure detection of the marine diesel engines in Switzerland, the Swedish AGEMA infrared thermography, and the vibration monitoring system of the Danish B & K [195, 196]. At the middle of the 1980s, fault diagnosis technology had entered a new stage (i.e., intelligence fault diagnosis) with the development of artificial intelligence such as neural network and the application of computer technology. However, China began late in studying fault diagnosis. Some universities and institutes did not absorb advanced diagnostic theory and technology until the late 1970s and developed its own fault diagnosis device by researching new detection methods and summarizing the experience [50, 197].

Up to now, the Internet of Things (IOT), cloud storage, dynamic data analysis, and other advanced technologies have played an increasingly important role in CBM. Recent study has shown that IOT and big data analysis based on cloud platform can improve the efficiency of predictive maintenance by 25–30%. Moreover, the compound annual growth rate of predictive maintenance will be 39% from 2016 to 2022 according to the report on global predictive maintenance issued by IOT Analytics. Under this increase rate, the market size will reach 73.45 billion yuan in 2022 [198]. Consequently, it is critical for manufacturers provided maintenance service to develop advanced fault diagnosis system and expand the market as soon as possible. Table 8 provides a summary of some available and popular fault diagnosis systems.

## 7 Conclusions and future challenges

Diagnosis and prognosis are necessary actions in industries to estimate the condition of machinery and optimize the usage of machinery. By predicting the failure possibility of components or the entire system, downtime and economic loss can be reduced as much as possible. CBM is an effective and robust maintenance strategy used to avoid over-maintenance or under-maintenance. This paper has reviewed recent research and development in machinery diagnosis and prognostics following the three processes of the CBM program, namely, data acquisition, data processing, and maintenance decision-making. In the data acquisition section, five detection technologies are discussed. Each technology has its own advantages and disadvantages due to the different monitoring principle. The signal processing section reviews signal processing methods in existing publications from the theoretical background and real application and gives a list of the advantages and disadvantages for these technologies. The diagnostics section summarizes the related publications by separating them into three categories, i.e., physical models, knowledge-based models, and AI models. In the prognostics section, RUL prediction techniques are roughly classified following four categories and their achievements are discussed. It is also noticed that some commercial corporations and universities take enormous efforts to develop fault diagnosis systems, and acquire dramatic achievement in recent years.

Although much advancement has been achieved in the CBM, there are still some aspects which require to be further developed. The last of this paper aims to give the challenges and opportunities in this field, which is hoped to point out the future research directions and provide some suggestions for researchers.

### (1) Lack of high-quality data

Data acquisition is a challenging task in condition monitoring, especially for deep learning. As a rule of thumb, the number of samples should be at least ten times bigger than the number of parameters in a deep learning model [213]. With the increase of the installed sensors, the volume of the collected data is rapidly grown than ever before, but it also comes to bring negative effects. In many applications, lots of factors may pollute the collected data, such as sensor placement, the interruption of data transmission, and machine vibration [131]. The quality of data is more important than its quantity. When the incorrect data are directly used to train the diagnosis model, it will produce unreliable diagnosis result. Therefore, it is necessary to develop effective approach to clear anomaly data and further improve its quality, such as clustering algorithms and

**Table 8** Commercially available fault diagnosis systems

Supplier	Product	Properties	Monitoring technology
Bently [199]	3500 series	<ul style="list-style-type: none"> <li>• A flexible, scalable platform for continuous, online monitoring of most critical machinery as well as all other machine assets used in plant-wide processes</li> <li>• Provides automatic protection value by tripping monitored machines</li> <li>• Provides continuous condition monitoring information for use in a proactive maintenance program</li> <li>• Incorporates the industry's most extensive selection of machinery measurement parameters combined with software configuration for virtually all machinery monitoring scenarios</li> </ul>	<ul style="list-style-type: none"> <li>• Vibration</li> <li>• Temperature</li> </ul>
	Trendmaster series	<ul style="list-style-type: none"> <li>• Links hundreds of permanently mounted transducers and measurement points to a single-cable sensor network</li> <li>• Optimized for cost-effective monitoring of machines of intermediate criticality, i.e., machines that do not demand immediate shutdown protection, but for which monthly portables-based monitoring would be insufficient</li> </ul> <p>Suitable for use in the most hazardous environments (i.e. explosive gases)</p>	<ul style="list-style-type: none"> <li>• Vibration</li> <li>• Temperature</li> </ul>
	AnomAlert series	<ul style="list-style-type: none"> <li>• Testing for and recording existing and any potential faults on or within motors, generators, or the equipment they belong to</li> <li>• Applies model-based method</li> </ul>	<ul style="list-style-type: none"> <li>• Current and voltage</li> </ul>
	Ranger Pro series	<ul style="list-style-type: none"> <li>• Leverages IoT to improve uptime, productivity, and safety with wireless vibration sensor</li> <li>• Can work in hazardous environments</li> <li>• Has long battery life</li> </ul>	<ul style="list-style-type: none"> <li>• Vibration</li> <li>• Temperature</li> </ul>
Collins Aerospace [200]	HUMS series	<ul style="list-style-type: none"> <li>• Automatically monitors hundreds of aircraft signals through a network of advanced sensors and onboard computers</li> <li>• Integrates current and historical flight data into a comprehensive database with full reporting capabilities</li> </ul>	<ul style="list-style-type: none"> <li>• Vibration</li> </ul>
STI [201]	CMCP series	<ul style="list-style-type: none"> <li>• Suitable for cooling towers, fans, motors, turbines, compressors, and pump vibration monitoring</li> <li>• Allows remote viewing (or editing) by any internet connected PC or Smart Phone</li> </ul>	<ul style="list-style-type: none"> <li>• Vibration</li> <li>• Temperature</li> </ul>
Rockwell [202]	Dynamix series	<ul style="list-style-type: none"> <li>• Provides machinery protection and condition monitoring of rotating and reciprocating assets</li> <li>• Approved for electrical safety and hazardous areas</li> <li>• Offers expansion modules for additional capabilities</li> </ul>	<ul style="list-style-type: none"> <li>• Vibration</li> </ul>
PCH Engineering [203]	PCH series	<ul style="list-style-type: none"> <li>• Simple integration to existing monitoring systems (e.g., SCADA)</li> <li>• Detects misalignment and bearing faults in rotating machinery</li> <li>• Stores vibration data locally or online</li> </ul>	<ul style="list-style-type: none"> <li>• Vibration</li> </ul>
SKF [204]	Multilog series	<ul style="list-style-type: none"> <li>• Provides a complete system for initiation of machinery shutdown, early fault detection, and diagnosis</li> <li>• Provides automated advice for correcting existing or impending conditions</li> </ul>	<ul style="list-style-type: none"> <li>• Vibration</li> </ul>

**Table 8** (continued)

Supplier	Product	Properties	Monitoring technology
Alta solution [205]	AS series	<ul style="list-style-type: none"> <li>• An open architecture web-based machinery condition monitoring system</li> <li>• Provides live machine status, dynamic waveforms (time and spectrum), and historical data trending</li> <li>• Provides modular system for easy upgrading and replacement</li> </ul>	<ul style="list-style-type: none"> <li>• Vibration</li> </ul>
B&K Vibro [206]	VC-8000 series	<ul style="list-style-type: none"> <li>• Provides one full year of condition monitoring data even without a CM server</li> <li>• Allows you to access the VC-8000 machinery protection system at any time, from anywhere in the world</li> <li>• Suitable for critical rotating and reciprocating machinery</li> </ul>	<ul style="list-style-type: none"> <li>• Vibration</li> <li>• Acoustic</li> <li>• Temperature</li> </ul>
Tangzhi Science and Technology [207]	JK460 series	<ul style="list-style-type: none"> <li>• Developed for critical components of wind turbines</li> <li>• Based on generalized resonance/resonance demodulation fault diagnosis technology</li> <li>• More sensitive to early fault due to compound monitoring of vibration and shock</li> </ul>	<ul style="list-style-type: none"> <li>• Vibration</li> </ul>
Inventec Appliances [208]	EN-8000 series	<ul style="list-style-type: none"> <li>• Suitable for online dynamic monitoring and fault diagnosis of large rotating equipment</li> <li>• Provides trend analysis and back-up database</li> </ul>	<ul style="list-style-type: none"> <li>• Vibration</li> <li>• Temperature</li> </ul>
Beijing Ryongson International [209]	LC-8000 series	<ul style="list-style-type: none"> <li>• Can display various technical parameters such as acceleration, speed, displacement, and kurtosis</li> <li>• Suitable for automatic diagnosis of rotating machinery and AC asynchronous motors</li> </ul>	<ul style="list-style-type: none"> <li>• Vibration</li> </ul>
China Orient Institute of Noise and Vibration [210]	DASP series	<ul style="list-style-type: none"> <li>• Provides offline and online analysis</li> <li>• Uses neural network to offer preliminary diagnosis results</li> </ul>	<ul style="list-style-type: none"> <li>• Vibration</li> <li>• Temperature</li> </ul>
Southeast University [211]	MFD series	<ul style="list-style-type: none"> <li>• Applies global logic reasoning and local neural network to offer fault information and suggestions</li> <li>• Save on-site information in the event of unforeseen accidents</li> <li>• Realize the whole cycle parallel collection of multiple vibration signals according to the speed signal</li> </ul>	<ul style="list-style-type: none"> <li>• Vibration</li> </ul>
Wuhan University of Technology [212]	MSD-3 series	<ul style="list-style-type: none"> <li>• Uses instantaneous speed method to diagnose faults of diesel engine</li> </ul>	<ul style="list-style-type: none"> <li>• Vibration</li> <li>• Oil debris monitoring</li> </ul>

Bayesian model. Moreover, few suppliers would like to publish their run-to-failure data due to military secret or commercial competition [25]. We can find that most experiments were carried out through the bearing data published by Case Western Reserve University and University of Cincinnati; some others were conducted by self-established test-bed [126]. In practical applications, various faults occur, and the proposed techniques may not necessarily perform well on field operating machine. To facilitate the development of monitoring technologies, the public database should be established that collects a lot of datasets specially produced for demonstration of condition monitoring.

- (2) Lack of standard method to choose signal processing technology

Different technologies of signal processing have their respective advantages and disadvantages and perform differently in different cases. There is no clear way to select, design, or implement a signal processing technology in real applications. Some authors, e.g., Moosavian et al. [110], Bastami et al. [178], and Maior et al. [181], have not discussed the reason why they have preferred to implement their solution using their selected technology. This could be a result of a drive within experience to ensure that these technologies as logically and accurately as possible. However, researchers, who may be beginners to condition monitoring, are not

always clear as to which technology will work better, so they need to take a lot of time to review related literature and do corresponding trials. Future research should focus on developing a standard scoring system that ranks the performance of different techniques. After identifying the type of collected data, system applies various evaluation parameters (e.g., computing time, the ability to process non-linear signal) to assess and rank different techniques, which can give researchers an insight into the performance of each technology to choose a proper one.

### (3) Improvement of interpretability of the deep learn algorithm

Although deep learning algorithms have achieved good results in the field of fault diagnosis and life prediction, an open issue of black box for deep hierarchical networks still confuses researchers. It is difficult or even impossible to have physical explanations of the model's outputs. Moreover, as model grows in size, the structure and parameters of model can become a complicated issue. They are constructed by experimental trials once and once again rather than the strictly theoretical background [131, 214]. To improve the interpretability of the deep learning algorithms, two research directions are recommended to be concerned [131]. (a) Different from the ANN, the statistical learning theories, such as SVM and HMM, are beneficial to construct models with easily-understand outputs due to the rigorous theory grounds. (b) The process of learning features by deep learning is similar to the filtering process. As a result, adaptive filter theory might be used to explain the physical meaning of deep learning models, and visualization technologies are expected to intuitively express what the models have learned from the input data. However, these still require a lot of researcher effort.

### (4) Development of a universal platform

The upsurge and progressive maturity of new information and communication technologies used to industrial processes and products has propelled “smartization” of manufacturing industries. In this context, some fault diagnosis systems have been developed, which monitors abnormal behaviors of specific assets [43]. They always follow the three main steps of a CBM program, i.e., data acquisition, signal processing, and maintenance decision-making. In order to make full use of resources, we believe that it is necessary to develop a universal platform to deal with the data collected from different sources in future CBM research [215, 216]. For instance, data may come in various forms, such as vibration signals, current signals, and AE signals. For different types of data, the platform allows extracting relevant knowledge from targeted assets by automati-

cally selecting proper intelligent monitoring and data fusion strategies, as well as by applying proper fault diagnosis and life prediction technologies.

**Acknowledgements** This research is supported by the National Science Foundation of China (NSFC) under the grant no. 51979261, the Taishan Scholar of Shandong of China under the grant no. tsqn201812025), and Narodowego Centrum Nauki of Poland under the grant no. 2020/37/K/ST8/02748.

**Availability of data and materials** All data and materials used to produce the results in this article can be obtained upon request from the corresponding authors.

## Declarations

**Ethical approval** The authors declare that there is no ethical issue applied to this article.

**Consent to participate** The authors declare that all authors have read and approved to submit this manuscript to IJAMT.

**Consent to publish** The authors declare that all authors agree to sign the transfer of copyright for the publisher to publish this article upon on acceptance.

**Competing interests** The authors declare that there is no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Kumar A, Vashishtha G, Gandhi CP et al (2021) Novel convolutional neural network (NCNN) for the diagnosis of bearing defects in rotary machinery. *IEEE T Instrum Meas* 70:1–10
2. Glowacz A (2021) Fault diagnosis of electric impact drills using thermal imaging. *Measurement* 171:108815
3. Xu Y, Li ZX, Wang SQ et al (2021) A hybrid deep-learning model for fault diagnosis of rolling bearings. *Measurement* 169:108502
4. Jiang WX, Wang HH, Liu GJ et al (2021) A Novel method for mechanical fault diagnosis of underwater pump motors based on power flow theory. *IEEE T Instrum Meas* 70:1–17
5. Chen Z, Han Y, Zhao Y et al (2019) Mission reliability evaluation based on operational quality data for multistate manufacturing systems. *Int J Prod Res* 57:1840–1856

6. Glowacz A, Tadeusiewicz R, Legutko S et al (2021) Fault diagnosis of angle grinders and electric impact drills using acoustic signals. *Appl Acoust* 179:108070
7. He Y, Zhao Y, Han X et al (2020) Functional risk-oriented health prognosis approach for intelligent manufacturing systems. *Reliab Eng Syst Safe* 203:107090
8. Ben Ali J, Chebel-Morello B, Saidi L et al (2015) Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network. *Mech Syst Signal Pr* 56–57:150–172
9. Dong CZ, Bas S, Catbas FN (2019) A completely non-contact recognition system for bridge unit influence line using portable cameras and computer vision. *Smart Struct Syst* 24(5):617–630
10. Kandukuri ST, Klausen A, Karimi HR et al (2016) A review of diagnostics and prognostics of low-speed machinery towards wind turbine farm-level health management. *Renew Sust Energ Rev* 53:697–708
11. Tian ZG, Jin TD, Wu BR et al (2011) Condition based maintenance optimization for wind power generation systems under continuous monitoring. *Renew Energ* 36(5):1502–1509
12. Jardine AK, Lin DM, Banjevic D (2006) A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech Syst Signal Pr* 20(7):1483–1510
13. Hsueh YM, Ittangihal VR, Wu WB et al (2019) Fault diagnosis system for induction motors by CNN using empirical wavelet transform. *Symmetry-Basel* 11(10):15
14. Zhang XY, Liang YT, Zhou JZ et al (2015) A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM. *Measurement* 69:164–179
15. Liu XF, Bo L, Luo HL (2015) Bearing faults diagnostics based on hybrid LS-SVM and EMD method. *Measurement* 59:145–166
16. Jiang QC, Huang B (2016) Distributed monitoring for large-scale processes based on multivariate statistical analysis and Bayesian method. *J Process Contr* 46:75–83
17. Shahnazari H, Mhaskar P, House JM et al (2019) Distributed fault diagnosis of heating, ventilation, and air conditioning systems. *Aiche J* 65(2):640–651
18. Chen TR, Hill DJ, Wang C (2020) Distributed fast fault diagnosis for multimachine power systems via deterministic learning. *IEEE T Ind Electron* 67(5):4152–4162
19. Dong CZ, Catbas FN (2021) A review of computer vision-based structural health monitoring at local and global level. *Struct Health Monit* 20(2):692–743
20. Dong CZ, Bas S, Catbas FN (2020) Investigation of vibration serviceability of a footbridge using computer vision-based methods. *Eng Struct* 224:111224
21. Dong CZ, Celik O, Catbas FN (2019) Marker free monitoring of the grandstand structures and modal identification using computer vision methods. *Struct Health Monit* 18(5–6):1491–1509
22. Wang WB, Tse PW, Lee J (2007) Remote machine maintenance system through Internet and mobile communication. *Int J Adv Manuf Tech* 31(7–8):783–789
23. Zhao FG, Chen J, Dong GM et al (2010) SOA-based remote condition monitoring and fault diagnosis system. *Int J Adv Manuf Tech* 46(9):1191–1200
24. Henriquez P, Alonso JB, Ferrer MA et al (2014) Review of automatic fault diagnosis systems using audio and vibration signals. *IEEE T Syst Man Cy-S* 44(5):642–652
25. Lei YG, Li NP, Guo L et al (2018) Machinery health prognostics: a systematic review from data acquisition to RUL prediction. *Mech Syst Signal Pr* 104:799–834
26. Safizadeh MS, Latifi SK (2014) Using multi-sensor data fusion for vibration fault diagnosis of rolling element bearings by accelerometer and load cell. *Inform Fusion* 18:1–8
27. Duan ZH, Wu TH, Guo SW et al (2018) Development and trend of condition monitoring and fault diagnosis of multi-sensors information fusion for rolling bearings: a review. *Int J Adv Manuf Tech* 96(1–4):803–819
28. Liang B, Iwnicki SD, Zhao Y (2013) Application of power spectrum, cepstrum, higher order spectrum and neural network analyses for induction motor fault diagnosis. *Mech Syst Signal Pr* 39(1):342–360
29. Guo JC, Shi ZQ, Li HY et al (2018) Early fault diagnosis for planetary gearbox based wavelet packet energy and modulation signal bispectrum analysis. *Sensors-Basel* 18(9):12
30. Li F, Pang X, Yang Z (2019) Motor current signal analysis using deep neural networks for planetary gear fault diagnosis. *Measurement* 145:45–54
31. Chen XW, Feng ZP (2020) Induction motor stator current analysis for planetary gearbox fault diagnosis under time-varying speed conditions. *Mech Syst Signal Pr* 140:24
32. Gangsar P, Tiwari R (2020) Signal based condition monitoring techniques for fault detection and diagnosis of induction motors: a state-of-the-art review. *Mech Syst Signal Pr* 144
33. Eftekharnjad B, Carrasco MR, Charnley B et al (2011) The application of spectral kurtosis on acoustic emission and vibrations from a defective bearing. *Mech Syst Signal Pr* 25(1):266–284
34. Kilundu B, Chiementin X, Duez J et al (2011) Cyclostationarity of acoustic emissions (AE) for monitoring bearing defects. *Mech Syst Signal Pr* 25(6):2061–2072
35. de Azevedo HDM, Araujo AM, Bouchonneau N (2016) A review of wind turbine bearing condition monitoring: state of the art and challenges. *Renew Sust Energ Rev* 56:368–379
36. Couturier J, Mba D (2008) Operational bearing parameters and acoustic emission generation. *J Vib Acoust* 130(2):5
37. Bagavathiappan S, Lahiri B, Saravanan T et al (2013) Infrared thermography for condition monitoring—a review. *Infrared Phys Techn* 60:35–55
38. Glowacz A (2019) Fault diagnosis of single-phase induction motor based on acoustic signals. *Mech Syst Signal Pr* 117:65–80
39. Sheng S (2016) Monitoring of wind turbine gearbox condition through oil and wear debris analysis: a full-scale testing perspective. *Tribol T* 59(1):149–162
40. Ding Y, Wang Y, Xiang J (2016) An online debris sensor system with vibration resistance for lubrication analysis. *Rev Sci Instrum* 87(2):025109
41. Hong W, Cai W, Wang S et al (2018) Mechanical wear debris feature, detection, and diagnosis: a review. *Chin J Aeronaut* 31(5):867–882
42. Khaleghi B, Khamis A, Karray FO et al (2013) Multisensor data fusion: a review of the state-of-the-art. *Inform Fusion* 14(1):28–44
43. Diez-Olivan A, Del Ser J, Galar D et al (2019) Data fusion and machine learning for industrial prognosis: trends and perspectives towards Industry 4.0. *Inform Fusion* 50:92–111
44. Williams T, Ribadeneira X, Billington S et al (2001) Rolling element bearing diagnostics in run-to-failure lifetime testing. *Mech Syst Signal Pr* 15(5):979–993
45. Lei YG, Kong DT, Lin J et al (2012) Fault detection of planetary gearboxes using new diagnostic parameters. *Meas Sci Technol* 23(5)
46. Zhang L, Hu NQ (2019) Time domain synchronous moving average and its application to gear fault detection. *IEEE Access* 7:93035–93048
47. Xiao HF, Zhou XJ, Shao YM (2016) Application of an improved dynamic time synchronous averaging method for fault diagnosis in conditions of speed fluctuation and no tachometer. *P I Mech Eng C-J Mec* 230(14):2517–2531



48. Wang L, Pan J, Gao YF et al (2019) Incipient fault diagnosis of limit switch based on a ARMA model. *Measurement* 135:473–480
49. Han T, Jiang D (2016) Rolling bearing fault diagnostic method based on VMD-AR model and random forest classifier. *Shock Vib* 2016:5132046
50. Liu ZP, Zhang L (2020) A review of failure modes, condition monitoring and fault diagnosis methods for large-scale wind turbine bearings. *Measurement* 149
51. Cai JH (2014) Fault diagnosis of rolling bearing based on empirical mode decomposition and higher order statistics. *P I Mech Eng C-J Mec* 229(9):1630–1638
52. Liu R, Yang B, Zio E et al (2018) Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mech Syst Signal Pr* 108:33–47
53. Bi XY, Cao SQ, Zhang DM (2019) Diesel engine valve clearance fault diagnosis based on improved variational mode decomposition and bispectrum. *Energies* 12(4):661
54. Huang BS, Feng GJ, Tang XL et al (2019) A performance evaluation of two bispectrum analysis methods applied to electrical current signals for monitoring induction motor-driven systems. *Energies* 12(8):23
55. Zhang X, Zhou R, Zhang W (2019) Improved local cepstrum and its applications for gearbox and rolling bearing fault detection. *Meas Sci Technol* 30(7)
56. Li B, Zhang XN, Wu JL (2017) New procedure for gear fault detection and diagnosis using instantaneous angular speed. *Mech Syst Signal Pr* 85:415–428
57. Chacon JLF, Kappatos V, Balachandran W et al (2015) A novel approach for incipient defect detection in rolling bearings using acoustic emission technique. *Appl Acoust* 89:88–100
58. Abd-el-Malek M, Abdelsalam AK, Hassan OE (2017) Induction motor broken rotor bar fault location detection through envelope analysis of start-up current using Hilbert transform. *Mech Syst Signal Pr* 93:332–350
59. Du WL, Wang ZY, Gong XY et al (2016) Optimum IMFs selection based envelope analysis of bearing fault diagnosis in plunger pump. *Shock Vib* 2016:8
60. Wang HD, Deng SE, Yang JX et al (2020) Parameter-adaptive VMD method based on BAS optimization algorithm for incipient bearing fault diagnosis. *Math Probl Eng* 2020:15
61. Wang LH, Zhao XP, Wu JX et al (2017) Motor fault diagnosis based on short-time fourier transform and convolutional neural network. *Chin J Mech Eng-En* 30(6):1357–1368
62. Burriel-Valencia J, Puche-Panadero R, Martinez-Roman J et al (2018) Fault diagnosis of induction machines in a transient regime using current sensors with an optimized Slepian window. *Sensors-Basel* 18(1)
63. Gangsar P, Tiwari R (2019) Diagnostics of mechanical and electrical faults in induction motors using wavelet-based features of vibration and current through support vector machine algorithms for various operating conditions. *J Braz Soc Mech Sci* 41(2)
64. Jaber AA, Bicker R (2016) Fault diagnosis of industrial robot gears based on discrete wavelet transform and artificial neural network. *Insight* 58(4):179–186
65. Talhaoui H, Menacer A, Kessal A et al (2017) Experimental diagnosis of broken rotor bars fault in induction machine based on Hilbert and discrete wavelet transforms. *Int J Adv Manuf Tech* 95(1–4):1399–1408
66. Chen J, Li Z, Pan J et al (2016) Wavelet transform based on inner product in fault diagnosis of rotating machinery: a review. *Mech Syst Signal Pr* 70–71:1–35
67. Guan YP, Liang M, Neculescu DS (2019) Velocity synchronous bilinear distribution for planetary gearbox fault diagnosis under non-stationary condition. *J Sound Vib* 443:212–229
68. Fan HW, Shao SJ, Zhang XH et al (2020) Intelligent fault diagnosis of rolling bearing using FCM clustering of EMD-PWVD vibration images. *IEEE Access* 8:145194–145206
69. Li W, Mechefske CK (2016) Detection of induction motor faults: a comparison of stator current, vibration and acoustic methods. *J Vib Control* 12(2):165–188
70. Bokde N, Feijoo A, Villanueva D et al (2019) A review on hybrid empirical mode decomposition models for wind speed and wind power prediction. *Energies* 12(2)
71. Ben Ali J, Fnaiech N, Saidi L et al (2015) Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. *Appl Acoust* 89:16–27
72. He Q, Ren X, Jiang G et al (2014) A hybrid feature extraction methodology for gear pitting fault detection using motor stator current signal. *Insight* 56(6):326–333
73. Zhao L, Huang DR, Qin Y (2015) Fault Diagnosis for gearbox based on improved empirical mode decomposition. *Shock Vib* 2015:9
74. Xu YB, Cai ZY, Ding K (2018) An enhanced bearing fault diagnosis method based on TVF-EMD and a high-order energy operator. *Meas Sci Technol* 29(9):19
75. Feng Z, Zhang D, Zuo MJ (2017) Adaptive mode decomposition methods and their applications in signal analysis for machinery fault diagnosis: a review with examples. *IEEE Access* 5:24301–24331
76. Tabrizi A, Garibaldi L, Fasana A et al (2015) Early damage detection of roller bearings using wavelet packet decomposition, ensemble empirical mode decomposition and support vector machine. *Meccanica* 50(3):865–874
77. Luo M, Li C, Zhang X et al (2016) Compound feature selection and parameter optimization of ELM for fault diagnosis of rolling element bearings. *ISA T* 65:556–566
78. Cheng JS, Yang Y, Yang Y (2012) A rotating machinery fault diagnosis method based on local mean decomposition. *Digit Signal Process* 22(2):356–366
79. Liu Z, He Z, Guo W et al (2016) A hybrid fault diagnosis method based on second generation wavelet de-noising and local mean decomposition for rotating machinery. *ISA T* 61:211–220
80. Wang L, Liu Z, Miao Q et al (2018) Time-frequency analysis based on ensemble local mean decomposition and fast kurtogram for rotating machinery fault diagnosis. *Mech Syst Signal Pr* 103:60–75
81. Wang Z, Wang J, Cai W et al (2019) Application of an improved ensemble local mean decomposition method for gearbox composite fault diagnosis. *Complexity* 2019:1564243
82. Li YB, Si SB, Liu ZL et al (2019) Review of local mean decomposition and its application in fault diagnosis of rotating machinery. *J Syst Eng Electron* 30(4):799–814
83. Feng ZP, Liang M, Chu FL (2013) Recent advances in time-frequency analysis methods for machinery fault diagnosis: a review with application examples. *Mech Syst Signal Pr* 38(1):165–205
84. Parchami M, Zhu WP, Champagne B et al (2016) Recent developments in speech enhancement in the short-time Fourier transform domain. *IEEE Circ Syst Mag* 16(3):45–77
85. Lei YG, Lin J, He ZJ et al (2013) A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mech Syst Signal Pr* 35(1–2):108–126
86. Cibulka J, Ebbesen MK, Hovland G et al (2012) A review on approaches for condition based maintenance in applications with induction machines located offshore. *Model Ident Control* 33(2):69–86
87. Gao Z, Cecati C, Ding SX (2015) A survey of fault diagnosis and fault-tolerant techniques—Part I: fault diagnosis with model-based and signal-based approaches. *IEEE T Ind Electron* 62(6):3757–3767

88. Xu Y, Sun Y, Wan J et al (2017) Industrial big data for fault diagnosis: taxonomy, review, and applications. *IEEE Access* 5:17368–17380
89. Wei Y, Li Y, Xu M et al (2019) A review of early fault diagnosis approaches and their applications in rotating machinery. *Entropy-Switz* 21(4):409
90. Dai JY, Tang J, Huang SZ et al (2019) Signal-based intelligent hydraulic fault diagnosis methods: review and prospects. *Chin J Mech Eng-En* 32(1):22
91. Jung D, Ng KY, Frisk E et al (2018) Combining model-based diagnosis and data-driven anomaly classifiers for fault isolation. *Control Eng Pract* 80:146–156
92. Gao C, Lv K, Si J et al (2019) A physical faulty model based on coil sub-element for direct-drive permanent magnet synchronous motor with stator winding short-circuit faults. *IEEE Access* 7:151307–151319
93. Yang BS, Lim DS, Tan ACC (2005) VIBEX: an expert system for vibration fault diagnosis of rotating machinery using decision tree and decision table. *Expert Syst Appl* 28(4):735–742
94. Hussain A, Lee SJ, Choi MS et al (2015) An expert system for acoustic diagnosis of power circuit breakers and on-load tap changers. *Expert Syst Appl* 42(24):9426–9433
95. Xu BB, Li HH, Pang WT et al (2019) Bayesian network approach to fault diagnosis of a hydroelectric generation system. *Energy Sci Eng* 7(5):1669–1677
96. Xu XJ, Yan XP, Sheng CX et al (2020) A belief rule-based expert system for fault diagnosis of marine diesel engines. *IEEE T Syst Man Cy-S* 50(2):656–672
97. Sikorska JZ, Hodkiewicz M, Ma L (2011) Prognostic modeling options for remaining useful life estimation by industry. *Mech Syst Signal Pr* 25(5):1803–1836
98. Chen X, Cheng G, Li H et al (2016) Diagnosing planetary gear faults using the fuzzy entropy of LMD and ANFIS. *J Mech Sci Technol* 30(6):2453–2462
99. Parey A, Singh A (2019) Gearbox fault diagnosis using acoustic signals, continuous wavelet transform and adaptive neuro-fuzzy inference system. *Appl Acoust* 147:133–140
100. He Y, Chen Z, Zhao Y et al (2020) Mission reliability evaluation for fuzzy multistate manufacturing system based on an extended stochastic flow network. *IEEE T Reliab* 69(4):1239–1253
101. Toma RN, Prosvirin AE, Kim JM (2020) Bearing fault diagnosis of induction motors using a genetic algorithm and machine learning classifiers. *Sensors-Basel* 20(7)
102. Glowacz A, Glowacz Z (2017) Diagnosis of stator faults of the single-phase induction motor using acoustic signals. *Appl Acoust* 117:20–27
103. Islam R, Khan SA, Kim J-m (2016) Discriminant feature distribution analysis-based hybrid feature selection for online bearing fault diagnosis in induction motors. *J Sensors* 2016:1–16
104. Han H, Cho S, Kwon S et al (2018) Fault diagnosis using improved complete ensemble empirical mode decomposition with adaptive noise and power-based intrinsic mode function selection algorithm. *Electronics-Switz* 7(2):11
105. Glowacz A (2018) Acoustic based fault diagnosis of three-phase induction motor. *Appl Acoust* 137:82–89
106. Taimoor M, Aijun L (2020) Adaptive strategy for fault detection, isolation and reconstruction of aircraft actuators and sensors. *J Intell Fuzzy Syst* (Preprint) 1–20
107. Zhou W, Li XL, Yi J et al (2019) A novel UKF-RBF method based on adaptive noise factor for fault diagnosis in pumping unit. *IEEE T Ind Inform* 15(3):1415–1424
108. Jin Z, Han Q, Zhang K et al (2020) An intelligent fault diagnosis method of rolling bearings based on Welch power spectrum transformation with radial basis function neural network. *J Vib Control* 26(9–10):629–642
109. Kan MS, Tan ACC, Mathew J (2015) A review on prognostic techniques for non-stationary and non-linear rotating systems. *Mech Syst Signal Pr* 62–63:1–20
110. Moosavian A, Ahmadi H, Tabatabaefar A et al (2013) Comparison of two classifiers; K-nearest neighbor and artificial neural network, for fault diagnosis on a main engine journal-bearing. *Shock Vib* 20(2):263–272
111. An Z, Li S, Xin Y et al (2019) An intelligent fault diagnosis framework dealing with arbitrary length inputs under different working conditions. *Meas Sci Technol* 30(12):125107
112. Park P, Di Marco P, Shin H et al (2019) Fault detection and diagnosis using combined autoencoder and long short-term memory network. *Sensors-Basel* 19(21):17
113. Habbouche H, Benkedjouh T, Amirat Y et al (2021) Gearbox failure diagnosis using a multisensor data-fusion machine-learning-based approach. *Entropy-Switz* 23(6):20
114. Long JY, Zhang SH, Li C (2020) Evolving deep echo state networks for intelligent fault diagnosis. *IEEE T Ind Inform* 16(7):4928–4937
115. Zhang Y, Zhou T, Huang X et al (2021) Fault diagnosis of rotating machinery based on recurrent neural networks. *Measurement* 171:108774
116. Xu H, Chen G (2013) An intelligent fault identification method of rolling bearings based on LSSVM optimized by improved PSO. *Mech Syst Signal Pr* 35(1–2):167–175
117. Tan XH, Bi WH, Hou XL et al (2011) Reliability analysis using radial basis function networks and support vector machines. *Comput Geotech* 38(2):178–186
118. Singh M, Shaik AG (2019) Faulty bearing detection, classification and location in a three-phase induction motor based on Stockwell transform and support vector machine. *Measurement* 131:524–533
119. Han L, Li CW, Guo SL et al (2015) Feature extraction method of bearing AE signal based on improved FAST-ICA and wavelet packet energy. *Mech Syst Signal Pr* 62–63:91–99
120. Ma S, Cheng B, Shang Z et al (2018) Scattering transform and LSPTSVM based fault diagnosis of rotating machinery. *Mech Syst Signal Pr* 104:155–170
121. Liu Z, Cao H, Chen X et al (2013) Multi-fault classification based on wavelet SVM with PSO algorithm to analyze vibration signals from rolling element bearings. *Neurocomputing* 99:399–410
122. Moosavian A, Khazaei M, Najafi G et al (2015) Spark plug fault recognition based on sensor fusion and classifier combination using Dempster-Shafer evidence theory. *Appl Acoust* 93:120–129
123. Ghosh S, Wang YZ (2015) Feature import vector machine: a general classifier with flexible feature selection. *Stat Anal Data Min* 8(1):49–63
124. Huang XJ, Zhang L, Wang BJ et al (2018) Feature clustering based support vector machine recursive feature elimination for gene selection. *Appl Intell* 48(3):594–607
125. Jalalian A, Chalup SK (2013) GDTW-P-SVMs: variable-length time series analysis using support vector machines. *Neurocomputing* 99:270–282
126. Hoang DT, Kang HJ (2019) A survey on deep learning based bearing fault diagnosis. *Neurocomputing* 335:327–335
127. Shao HD, Jiang HK, Zhao HW et al (2017) A novel deep autoencoder feature learning method for rotating machinery fault diagnosis. *Mech Syst Signal Pr* 95:187–204
128. Tao J, Liu Y, Yang D (2016) Bearing fault diagnosis based on deep belief network and multisensor information fusion. *Shock Vib* 2016:1–9
129. Tran VT, AlThobiani F, Ball A (2014) An approach to fault diagnosis of reciprocating compressor valves using Teager-Kaiser energy operator and deep belief networks. *Expert Syst Appl* 41(9):4113–4122

130. Qin Z-Q, Ma X-K, Wang Y-J (2020) ADSAD: an unsupervised attention-based discrete sequence anomaly detection framework for network security analysis. *Comput Secur* 99:102070
131. Lei YG, Yang B, Jiang XW et al (2020) Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mech Syst Signal Pr* 138:39
132. Khan S, Yairi T (2018) A review on the application of deep learning in system health management. *Mech Syst Signal Pr* 107:241–265
133. Li N, Lei Y, Lin J et al (2015) An improved exponential model for predicting remaining useful life of rolling element bearings. *IEEE T Ind Electron* 62(12):7762–7773
134. Zhang X, Lu X, Li W et al (2021) Prediction of the remaining useful life of cutting tool using the Hurst exponent and CNN-LSTM. *Int J Adv Manuf Tech* 112(7):2277–2299
135. Heng A, Zhang S, Tan ACC et al (2009) Rotating machinery prognostics: state of the art, challenges and opportunities. *Mech Syst Signal Pr* 23(3):724–739
136. Xu JP, Wang YS, Xu L (2014) PHM-oriented integrated fusion prognostics for aircraft engines based on sensor data. *IEEE Sens J* 14(4):1124–1132
137. Lee J, Wu F, Zhao W et al (2014) Prognostics and health management design for rotary machinery systems—reviews, methodology and applications. *Mech Syst Signal Pr* 42(1–2):314–334
138. Bai GX, Wang PF (2016) Prognostics using an adaptive self-cognizant dynamic system approach. *IEEE T Reliab* 65(3):1427–1437
139. Saxena A, Celaya J, Saha B et al (2010) Metrics for offline evaluation of prognostic performance. *Int J Prog Health Manage* 1(1):4–23
140. Yang F, Habibullah MS, Zhang T et al (2016) Health index-based prognostics for remaining useful life predictions in electrical machines. *IEEE T Ind Electron* 63(4):2633–2644
141. Qian Y, Yan R, Gao RX (2017) A multi-time scale approach to remaining useful life prediction in rolling bearing. *Mech Syst Signal Pr* 83:549–567
142. Haile MA, Riddick JC, Assefa AH (2016) Robust particle filters for fatigue crack growth estimation in rotorcraft structures. *IEEE T Reliab* 65(3):1438–1448
143. Hu Y, Baraldi P, Di Maio F et al (2016) Online performance assessment method for a model-based prognostic approach. *IEEE T Reliab* 65(2):718–735
144. Chan KS, Enright MP, Moody JP et al (2012) Life prediction for turbopropulsion systems under dwell fatigue conditions. *J Eng Gas Turbine Power* 134(12)
145. Xiahou TF, Zeng ZG, Liu Y (2021) Remaining useful life prediction by fusing expert knowledge and condition monitoring information. *IEEE T Ind Inform* 17(4):2653–2663
146. Alamaniotis M, Grelle A, Tsoukalas LH (2014) Regression to fuzziness method for estimation of remaining useful life in power plant components. *Mech Syst Signal Pr* 48(1–2):188–198
147. Majidian A, Saidi MH (2007) Comparison of fuzzy logic and neural network in life prediction of boiler tubes. *Int J Fatigue* 29(3):489–498
148. Kang WJ, Xiao JY, Xiao MQ et al (2020) Research on remaining useful life prognostics based on fuzzy evaluation-Gaussian process regression method. *IEEE Access* 8:71965–71973
149. Cheng F, Qu L, Qiao W (2018) Fault prognosis and remaining useful life prediction of wind turbine gearboxes using current signal analysis. *IEEE T Sustain Energ* 9(1):157–167
150. Lei Y, Li N, Lin J (2016) A new method based on stochastic process models for machine remaining useful life prediction. *IEEE T Instrum Meas* 65(12):2671–2684
151. Qiu G, Gu Y, Chen J (2020) Selective health indicator for bearings ensemble remaining useful life prediction with genetic algorithm and Weibull proportional hazards model. *Measurement* 150
152. Du Y, Wu TH, Zhou SX et al (2020) Remaining useful life prediction of lubricating oil with dynamic principal component analysis and proportional hazards model. *P I Mech Eng J-J Eng* 234(6):964–971
153. Man J, Zhou Q (2018) Prediction of hard failures with stochastic degradation signals using Wiener process and proportional hazards model. *Comput Ind Eng* 125:480–489
154. Wang F, Chen X, Dun B et al (2017) Rolling bearing reliability assessment via kernel principal component analysis and Weibull proportional hazard model. *Shock Vib* 2017
155. Hu J, Chen P (2020) Predictive maintenance of systems subject to hard failure based on proportional hazards model. *Reliab Eng Syst Safe* 196
156. Verhagen WJC, De Boer LWM (2018) Predictive maintenance for aircraft components using proportional hazard models. *J Ind Inf Integr* 12:23–30
157. Duan C, Makis V, Deng C (2018) An integrated framework for health measures prediction and optimal maintenance policy for mechanical systems using a proportional hazards model. *Mech Syst Signal Pr* 111:285–302
158. Bunks C, McCarthy D, Al-Ani T (2000) Condition-based maintenance of machines using Hidden Markov Models. *Mech Syst Signal Pr* 14(4):597–612
159. Du Y, Wu TH, Makis V (2017) Parameter estimation and remaining useful life prediction of lubricating oil with HMM. *Wear* 376:1227–1233
160. Soualhi A, Clerc G, Razik H et al (2016) Hidden Markov models for the prediction of impending faults. *IEEE T Ind Electron* 63(5):3271–3281
161. Tao ZR, An QL, Liu GY et al (2019) A novel method for tool condition monitoring based on long short-term memory and hidden Markov model hybrid framework in high-speed milling Ti-6Al-4V. *Int J Adv Manuf Tech* 105(7–8):3165–3182
162. Zhu KP, Liu TS (2018) Online tool wear monitoring via hidden semi-markov model with dependent durations. *IEEE T Ind Inform* 14(1):69–78
163. Liu TS, Zhu KP, Zeng LC (2018) Diagnosis and prognosis of degradation process via hidden semi-Markov model. *IEEE-Asme T Mech* 23(3):1456–1466
164. Duan B, Zhang Q, Geng F et al (2020) Remaining useful life prediction of lithium-ion battery based on extended Kalman particle filter. *Int J Energ Res* 44(3):1724–1734
165. Baraldi P, Mangili F, Zio E (2012) A Kalman filter-based ensemble approach with application to turbine creep prognostics. *IEEE T Reliab* 61(4):966–977
166. Yildiz R, Barut M, Zerdali E (2020) A comprehensive comparison of extended and unscented Kalman filters for speed-sensorless control applications of induction motors. *IEEE T Ind Inform* 16(10):6423–6432
167. Chen N, Majda AJ (2019) A new efficient parameter estimation algorithm for high-dimensional complex nonlinear turbulent dynamical systems with partial observations. *J Comput Phys* 397:22
168. Cui L, Wang X, Xu Y et al (2019) A novel switching unscented Kalman filter method for remaining useful life prediction of rolling bearing. *Measurement* 135:678–684
169. Li X, Ma Y, Zhu J (2021) An online dual filters RUL prediction method of lithium-ion battery based on unscented particle filter and least squares support vector machine. *Measurement* 184:109935

170. Qian YN, Yan RQ (2015) Remaining useful life prediction of rolling bearings using an enhanced particle filter. *IEEE T Instrum Meas* 64(10):2696–2707
171. Jouin M, Gouriveau R, Hissel D et al (2016) Particle filter-based prognostics: review, discussion and perspectives. *Mech Syst Signal Pr* 72:2–31
172. Li T, Sun S, Sattar TP et al (2014) Fight sample degeneracy and impoverishment in particle filters: a review of intelligent approaches. *Expert Syst Appl* 41(8):3944–3954
173. Yuan S, Chen J, Yang W et al (2017) On-line crack prognosis in attachment lug using Lamb wave-deterministic resampling particle filter-based method. *Smart Mater Struct* 26(8):085016
174. Laredo D, Chen Z, Schütze O et al (2019) A neural network-evolutionary computational framework for remaining useful life estimation of mechanical systems. *Neural Netw* 116:178–187
175. Sbarufatti C, Corbetta M, Giglio M et al (2017) Adaptive prognosis of lithium-ion batteries based on the combination of particle filters and radial basis function neural networks. *J Power Sources* 344:128–140
176. Dong CZ, Li L, Yan J, Zhang Z, Pan H, Catbas FN (2021) Pixel-level fatigue crack segmentation in large-scale images of steel structures using an encoder–decoder network. *Sensors* 21(12):4135
177. Zhang S, Zhai B, Guo X et al (2019) Synchronous estimation of state of health and remaining useful lifetime for lithium-ion battery using the incremental capacity and artificial neural networks. *J Energy Storage* 26:100951
178. Bastami AR, Aasi A, Arghand HA (2019) Estimation of remaining useful life of rolling element bearings using wavelet packet decomposition and artificial neural network. *IJST-T Electr Eng* 43(1):233–245
179. Huang YC, Yang ZS, Liao HS (2019) Labeling confidence values for wafer-handling robot arm performance using a feature-based general regression neural network and genetic algorithm. *Appl Sci-Basel* 9(20):13
180. Zhang YZ, Xiong R, He HW et al (2018) Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries. *IEEE T Veh Technol* 67(7):5695–5705
181. Maior CBS, das Chagas Moura M, Lins ID, (2019) Particle swarm-optimized support vector machines and pre-processing techniques for remaining useful life estimation of bearings. *Eksploat Niezawodn* 21(4):610–618
182. Dong S, Luo T (2013) Bearing degradation process prediction based on the PCA and optimized LS-SVM model. *Measurement* 46(9):3143–3152
183. Ordóñez C, Lasheras FS, Roca-Pardinas J et al (2019) A hybrid ARIMA-SVM model for the study of the remaining useful life of aircraft engines. *J Comput Appl Math* 346:184–191
184. Benkedjouh T, Medjaher K, Zerhouni N et al (2013) Remaining useful life estimation based on nonlinear feature reduction and support vector regression. *Eng Appl Artif Intel* 26(7):1751–1760
185. Liu J, Vitelli V, Zio E et al (2015) A novel dynamic-weighted probabilistic support vector regression-based ensemble for prognostics of time series data. *IEEE T Reliab* 64(4):1203–1213
186. Khelif R, Chebel-Morello B, Malinowski S et al (2016) Direct remaining useful life estimation based on support vector regression. *IEEE T Ind Electron* 64(3):2276–2285
187. Wei J, Dong G, Chen Z (2017) Remaining useful life prediction and state of health diagnosis for lithium-ion batteries using particle filter and support vector regression. *IEEE T Ind Electron* 65(7):5634–5643
188. Avendano-Valencia LD, Chatzi EN, Tcherniak D (2020) Gaussian process models for mitigation of operational variability in the structural health monitoring of wind turbines. *Mech Syst Signal Pr* 142:23
189. Aye SA, Heyns PS (2017) An integrated Gaussian process regression for prediction of remaining useful life of slow speed bearings based on acoustic emission. *Mech Syst Signal Pr* 84:485–498
190. Jia JF, Liang JY, Shi YH et al (2020) SOH and RUL prediction of lithium-ion batteries based on gaussian process regression with indirect health indicators. *Energies* 13(2):20
191. Ismail A, Saidi L, Sayadi M et al (2020) Remaining useful lifetime prediction of thermally aged power insulated gate bipolar transistor based on Gaussian process regression. *T I Meas Control* 42(13):2507–2518
192. Huber MF (2014) Recursive Gaussian process: on-line regression and learning. *Pattern Recogn Lett* 45:85–91
193. Liu J, Chen ZQ (2019) Remaining useful life prediction of lithium-ion batteries based on health indicator and Gaussian process regression model. *IEEE Access* 7:39474–39484
194. Society for Machinery Failure Prevention Technology. Available: <https://www.mfpt.org/about/>. Accessed 11 Aug 2021
195. Sun FC, Hu DW, Liu HP (2014) Foundations and practical applications of cognitive systems and information processing. Springer, Berlin, Heidelberg
196. Sheng C, Li Z, Qin L et al (2011) Recent progress on mechanical condition monitoring and fault diagnosis. *Procedia Eng* 15:142–146
197. Niemann H, Kjølstad Poulsen N, Mirzaei M et al (2018) Fault diagnosis and condition monitoring of wind turbines. *Int J Adapt Control* 32(4):586–613
198. RootCloud. Available: <https://www.rootcloud.com/cmscontent/949.html>. Accessed 11 Aug 2021
199. Bently Nevada. Available: <https://www.bakerhughesds.com/bently-nevada/monitoring-systems>. Accessed 11 Aug 2021
200. Collins Aerospace. Available: <https://www.collinsaerospace.com/what-we-do/Helicopters/Commercial-Helicopters/Health-Usage-Management-Systems>. Accessed 11 Aug 2021
201. STI Vibration Monitoring Inc. Available: [https://www.stiweb.com/Monitoring\\_Systems\\_s/18.htm](https://www.stiweb.com/Monitoring_Systems_s/18.htm). Accessed 11 Aug 2021
202. Rockwell Automation. Available: <https://www.rockwellautomation.com/en-us/products/hardware/allen-bradley/condition-monitoring/1444-dynamix.html>. Accessed 11 Aug 2021
203. PCH Engineering. Available: <https://www.pch-engineering.dk/389/sil-2-vibration-monitors>. Accessed 11 Aug 2021
204. Svenska Kullager-Fabriken. Available: <https://www.skf.com/cn/products/condition-monitoring-systems/machine-protection-systems/imx-m#cid-469892>. Accessed 11 Aug 2021
205. Alta Solutions. Available: <https://www.altasol.com/as-7000-machinery-protection-system/>. Accessed 11 Aug 2021
206. Brüel & Kjær Vibro. Available: <https://www.bkvvibro.com/en/products/vc-8000machineprotectionsystem.html>. Accessed 11 Aug 2021
207. Tangzhi Science&Technology. Available: <http://www.tangzhi.com/fengsolution>. Accessed 11 Aug 2021
208. Inventec Appliances Corporation. Available: <http://a.gongkong.com/customer/envada/8000.asp>. Accessed 11 Aug 2021
209. Beijing Ryongson International. Available: <http://www.tmlc.cn/Mobile/Goods/show/id/301.html>. Accessed 11 Aug 2021
210. China Orient Institute of Noise & Vibration. Available: <http://www.coinv.com/page/pid/147.html>. Accessed 11 Aug 2021
211. Research Center of Condition Monitoring and Fault Diagnosis, Southeast University. Available: <http://www.rccmf.com/>. Accessed 11 Aug 2021
212. Shanghai Sanjin Technology Developmenty Co.Ltd. Available: <http://www.shsanjin.com/zidonghua/show/id/25.html>. Accessed 11 Aug 2021
213. Nasir V, Sassani F (2021) A review on deep learning in machining and tool monitoring: methods, opportunities, and challenges. *Int J Adv Manuf Tech* 115(9):2683–2709

214. Peng Y, Dong M, Zuo MJ (2010) Current status of machine prognostics in condition-based maintenance: a review. *Int J Adv Manuf Tech* 50(1–4):297–313
215. Li ZX, Yan XP, Guo ZW et al (2012) A new intelligent fusion method of multi-dimensional sensors and its application to tribosystem fault diagnosis of marine diesel engines. *Tribol Lett* 47(1):1–15
216. Yan XP, Xu XJ, Sheng CX et al (2018) Intelligent wear mode identification system for marine diesel engines based on multi-level belief rule base methodology. *Meas Sci Technol* 29(1):13

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.