



Plastic gear remaining useful life prediction using artificial neural network

B. H. Kien¹ · D. Iba¹ · Y. Tsutsui¹ · D. Taga¹ · N. Miura¹ · T. Iizuka¹ · A. Masuda¹ · I. Moriwaki¹

Received: 26 March 2021 / Accepted: 25 September 2021 / Published online: 21 October 2021
© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature 2021

Abstract

Prognostic and health management (PHM) of plastic gears has attracted attention due to an increasing performance of plastic gears, uncovering potential applications in the industry, especially in vehicle transmissions. Meanwhile, health indicator (HI) construction and remaining useful life (RUL) estimation are two key elements to efficiently perform PHM. In this paper, a health indicator generator (HIG) based on an artificial neural network (ANN) is constructed. The HIG is learned from training data extracted from plastic gears' raw vibration data using the Fourier decomposition method (FDM) in a sensitive frequency band (SFB) and labeled using a change-point detection algorithm (CDA). Three prediction strategies, including linear regression (LR), estimation of parameters for Weibull distribution (EWD), HI combined average RUL (HI-ARUL), are deployed using HI generated from HIG to predict the RUL of the plastic gear. The results show that the generated HI is sensitive to the early failure of plastic gears and is capable of applying an efficient and precise diagnosis method, which can be performed during the whole working time of plastic gear with prediction errors smaller than 7%.

Vorhersage der Restnutzungsdauer von Kunststoffzahnradern mit künstlichem neuronalem Netz

Zusammenfassung

Das Prognose- und Gesundheitsmanagement (PHM) von Kunststoffzahnradern hat durch die zunehmende Leistungsfähigkeit von Kunststoffzahnradern und deren Einsatzmöglichkeiten in der Industrie, insbesondere in Fahrzeuggetrieben, Aufmerksamkeit erregt. In der Zwischenzeit sind die Konstruktion des Zustandsindikators (HI) und die Schätzung der Restnutzungsdauer (RUL) zwei Schlüsselemente für eine effiziente Durchführung von PHM. In diesem Artikel wird ein Gesundheitsindikatorgenerator (HIG) basierend auf einem künstlichen neuronalen Netz (ANN) konstruiert. Das HIG wird aus Trainingsdaten gelernt, die aus den Rohschwingungsdaten von Kunststoffzahnradern mit der Fourier-Zerlegungsmethode (FDM) in einem empfindlichen Frequenzband (SFB) extrahiert und mit einem Änderungspunkt-Erkennungsalgorithmus (CDA) gekennzeichnet werden. Drei Vorhersagestrategien, einschließlich linearer Regression (LR), Schätzung von Parametern der Weibull-Verteilung (EWD), HI kombinierter Durchschnitt RUL (HI-ARUL), werden unter Verwendung von HI, das aus HIG generiert wird, eingesetzt, um die RUL des Kunststoffzahnrad vorherzusagen. Die Ergebnisse zeigen, dass der erzeugte HI empfindlich auf den frühen Ausfall von Kunststoffzahnradern reagiert. Daher ist es möglich, ein effizientes und präzises Diagnoseverfahren anzuwenden, das während der gesamten Arbeitszeit von Kunststoffzahnradern mit Vorhersagefehlern von weniger als 7% durchgeführt werden kann.

1 Introduction

Gears are widely used as essential machine elements to transmit power and motion [1, 2]. Many studies on metallic gears have been done, including design, manufacturing operation, and monitoring [3–5]. However, for mechanical engineers, plastic gears are a fascinating mean of reducing drive-cost, weight, noise, and ability to operate without lu-

✉ B. H. Kien
kien@pml.mech.kit.ac.jp

¹ Kyoto Institute of Technology,
Matsugasaki, Kyoto 606-8585, Japan

brication. Plastic gears also allow for smaller, more efficient transmissions in many products. Today, the drive capacity and working accuracy of plastic gears have been boosted due to better molding techniques combined with flexible geometric designs [6, 7]. Despite the technical advantages of plastic gears over metallic gears, the differences between the two are motivations for further investigations. Bravo et al. [8] indicated that the major part of the differences is due to plastics having an elastic modulus of approximately 100 times lower than most steels and 30 times lower than aluminum. Therefore, plastic gears mesh with increased areas of contact between teeth during engagement [9]. In addition, the potential of using plastic gears in engineering applications is limited due to the poor heat conduction, i.e. varying temperature during operation, low dimensional stability due to shrinkage, moisture absorption, low-load capacity, and thermal expansion [10]. The above reasons result in the failure behavior of plastic gears entirely different from the metallic gears. For example, different modes of failure are observed, e.g. melting of gear teeth, deflection of teeth, deformation of tooth profile, debris formation, cracks, tooth fracture and surface pitting [11]. In the next few years, if metallic gears are widely replaced with plastic gears in transmissions, e.g., electromechanical actuators in cars, the prognostic and health management (PHM) of plastic gears should be further investigated to improve reliability in finished products.

PHM of machine elements has long received researchers' attention for its crucial role in improving reliability and reducing cost in the industrial systems [12, 13]. PHM generally combines condition monitoring, fault diagnostics, fault prognostics, and decision support [14]. The remaining useful life (RUL) prediction is regarded as one of the most crucial issues in performing PHM. It commonly takes two key steps to efficiently implement RUL prediction: extraction of degradation or health status features from the original sensor data and an approximation of the prediction model to estimate RUL. During the working time from the running-in period to failure, the degradation information of rotating machinery spreads out to the surrounding environment. It can be recorded by vibration, acoustic or temperature sensors, etc. [15]. Among those types of sensor data, vibration data is most popular because it is easy to measure and provides rich dynamic information reflecting health conditions. However, vibration data usually contains noise, e.g. vibration data obtained from a gear operation test rig not only includes fundamental meshing frequencies but also nonlinearity frequencies of the test rig, rolling-elements noise of bearings, motor vibration, and its harmonics. Thus, it is necessary to apply feature extraction methods to extract robust features reflecting the health status of monitored objects from their original vibration data.

The extraction of degradation or health features from vibration data is performed using three typical methods: time-domain, frequency-domain, and time-frequency-domain [16]. Time-domain methods calculate statistical features from time-series data using methods such as root mean square (RMS), Kurtosis factor (KF), absolute mean amplitude (AMA), and peaks. Frequency-domain methods employ fast Fourier transform (FFT) to express time-domain data in a frequency-amplitude relationship. The high-active frequencies, which are strongly related to failure, can be effortlessly recognized. The time-frequency-domain methods investigate original data using the time-frequency representation transformations such as short-time Fourier transform (STFT), Wavelet transform (WT), Wigner-Ville transform (WVT). The extracted degradation or health features are used to generate health indicators that reflect the rotating machinery's state during working time.

The health indicator (HI) has been generated to display degradation profiles in previous approaches. Mahamad et al. [17] fitted the Weibull hazard rate function to measurement values RMS and KF extracted from vibration data of bearings. Ali et al. [18] released a new measurement based on RMS and KF values of vibration data, namely Root mean square entropy estimator (RMSEE), which exhibits fewer fluctuations. In their work, Weibull Distribution (WD) was used to fit RMS, KF, and RMSEE for tracking the degradation of bearings. Zhang et al. [19] employed Wavelet packet decomposition to decompose vibration signals into low-frequency and high-frequency parts, converted to frequency-domain using FFT. The peaks of frequencies of the decomposed signals were selected to track the degradation of a blower. Wu et al. [20] calculated power values on a sensitive frequency band (SFB) of bearings as a degradation indicator. Ji et al. [21] used a principal component analysis (PCA) for feature extraction to eliminate useless information and noise. Zhang et al. [22] translated the multi-dimension sensor data of aircraft engines to the dimensionless health index using a single-layer perceptron called a data generator. The construction of HI is thoroughly reviewed by Wang et al. [23].

After the HI is extracted, a prediction strategy is implemented to estimate RUL. There are mainly two categories of prediction strategy: model-based methods and data-driven methods. Using a model-based method, Bechhoefer et al. [24] approximated the coefficient of a Paris's Law equation for gear RUL estimation. Zheng [25] estimated three-parameters of Weibull Distribution (WD) using the graphical method and genetic algorithm for failure analysis of a series of CNC lathes. Li et al. [26] used the standard sequential probability ratio test method for Weibull life distribution combined with the hidden Markov model (HMM) to describe gearboxes' deterioration process. If precise models are established, the model-based methods can satisfy the

prediction results. However, due to the complexity of structures in the real world, it is hard to build physical models. Data-driven methods are efficient and easy to deploy since data-driven machine health monitoring is popular due to the widespread deployment of low-cost sensors and their connection to the Internet [27]. In fact, data-driven methods are designed to mine data without any understanding of the physical processes of phenomena. Hence, in case of scarce data, the data-driven methods are not efficient, i.e. although the models can perform well on the validation data set, they would not be reliable in case of scarce data. Using a data-driven model, Tian [28] developed a method based on an Artificial Neural Network (ANN) to achieve a more accurate RUL of pump bearings. Guo et al. [29] employed a recurrent neural network based-health indicator (RNN-HI) for RUL prediction of bearings. Nistane [30] compare prognostic models based on four approaches: Gated recurrent neural network (GRNN), classification and regression tree (CART), and autoregressive-moving average models (ARMA) in prognostics of rolling element bearings. Dong et al. [31] optimized a modified version of SVM, namely LS-SVM, for bearing degradation process prediction.

Despite numerous studies about RUL prediction of rotating machinery, the construction of HI and RUL prediction for plastic gears has been largely unknown. This paper constructs an automatic HI generator (HIG) from the real-time vibration data of plastic gears based on ANN architecture. The generated HI is suitable for monitoring plastic gear conditions during operation. However, to estimate RUL of plastic gears, it is necessary to deploy prediction strategies. There are numerous RUL prediction strategies based on physical models or data-driven models as mentioned above in order to highlight the efficiency of generated HI we selected simple and common models including linear regression (LR), estimation of parameters for Weibull distribution (EWD), and a proposed data-driven strategy using HI and average RUL (HI-ARUL) from historical run-to-failure tests of plastic gears. The errors between prediction RUL and actual RUL are compared to evaluate the efficiency of the selected methods. The signal processing method for vibration data to extract the degradation status of plastic gear and the labeling method for training data

to achieve desirable HI are two highlight contributions in this paper. The main concepts are presented in four main sections:

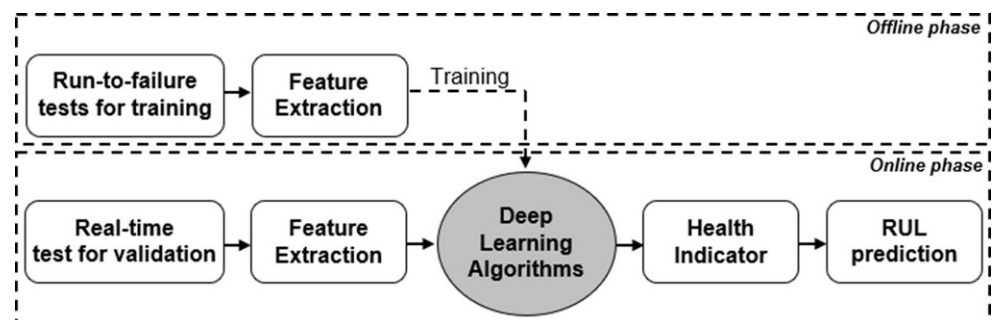
- In Sect. 2, a traditional prognostic framework using deep learning for RUL estimation in real-time is depicted.
- In Sect. 3, background methods are introduced, including the Fourier decomposition method, change-point detection algorithm, ANN architecture, linear regression, and Weibull distribution estimation.
- In Sect. 4, experimental works are described, including data acquisition, the feature extraction method, the labeling and pre-processing method for input data, and training of the ANN.
- In Sect. 5, results of the generated HI and RUL prediction using three strategies are discussed.

2 Prognosis framework

Numerous prognostic methods and systems have been published. However, it depends on specific cases of collected data and monitored objects which suitable methods are considered and selected. A large number of run-to-failure tests and efficient vibration data pre-processing techniques for plastic gears are the reasons for our selected methods. In this article, we employ a deep learning-based model for HI generating. Then, three prediction strategies are performed to estimate RUL. The procedure of the proposed method includes two phases as depicted in Fig. 1.

In the offline procedure: vibration signals are collected from run-to-failure tests of plastic gears for training. The robust features are extracted from vibration data using proposed techniques. In this scenario, the raw vibration signals are first decomposed and reconstructed in narrow frequency bands using the Fourier decomposition method (FDM). The absolute mean amplitude (AMA) of reconstructed signals are calculated and then assessed to select the most robust features, which are low-noise and can reflect the health status of plastic gears. The extracted robust features are pre-processed and labeled using data normalization and change-point detection algorithm (CDA) methods before inputting to an ANN. The trained ANN is considered a HIG, which

Fig. 1 Main procedures of a prognostic method using deep learning-based health indicator construction and physical-based RUL prediction



automatically generates HI of plastic gears from real-time vibration data.

In the online procedure: the real-time run-to-failure tests of plastic gears are carried out to evaluate the proposed prognostic method's performance. Correspondingly, robust features are extracted from the two tests. The trained HIG predicts the label of new input features every minute. The predicted labels are considered HI of the plastic gear, which approximates 0 when the plastic gear is healthy, reaches 1 once the plastic gear has broken, and has an increasing trend when the plastic gear runs into degradation. The generated HI is not only able to detect early deviations from the healthy state but also capable of forecasting time to failure of the plastic gear by using three prediction strategies: linear regression (LR), estimation of parameters for Weibull distribution (EWD), health indicator—average remaining useful life (HI-ARUL).

3 Background methods

3.1 Fourier decomposition method

Vibration signals obtained from a gear operation test rig include DC, shaft frequency, fundamental meshing frequency, rolling-elements noise of bearings, motor vibration, and its harmonics [32]. The extraction of features from raw vibration signal data has met challenges due to an enormous amount of noise and unnecessary information. The selection of robust components in collected data plays a decisive role in achieving precise diagnostic and prognostic tasks. This obstacle leads the authors to use the Fourier decomposition method (FDM) to decompose vibration signals into multi-components, which are then assessed to select robust components reflecting plastic gears' situation. The implementation process of the FDM is described in Fig. 2.

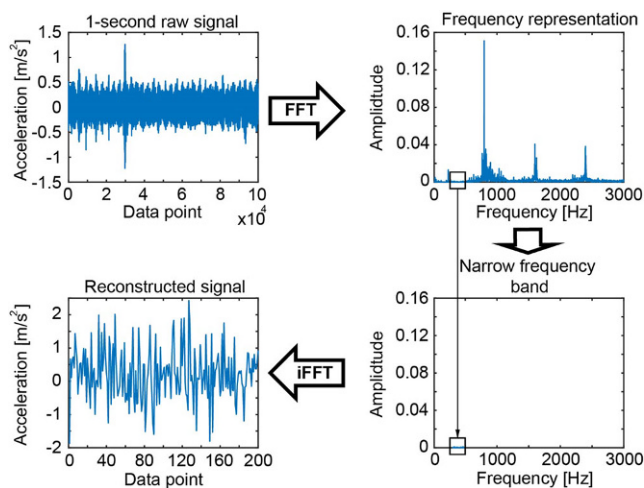


Fig. 2 The Fourier decomposition method implementation process

The figure shows a one-second vibration signal collected from the gear operation test rig at a 100,000 Hz sampling rate representing 100,000 data points. The input signal was represented in frequency-domain using fast Fourier transform (FFT). A narrow frequency band was selected to reconstruct the signal using the inverse FFT (iFFT). A reconstructed signal consists of 200 data points reflecting time-domain characteristics of frequencies in the corresponding frequency band. Reconstructed signals were then assessed to determine the desirable narrow frequency band. In this paper, the FDM was employed to explore a sensitive frequency band (SFB) of plastic gears in Sect. 4.2. Then, robust features are calculated from the SFB, which reflect early failure situations of plastic gear and are low-noise.

3.2 Change-point detection algorithm

The creation of change-point detection algorithm (CDA) is a long-established issue, and numerous papers proposed various approaches [33]. The objective of a CDA is to discover the abrupt changes in time-series data. A CDA divides a series data into multiple parts that are searched for the parts containing the abrupt change. For a detailed explanation, we assume time series data with length m represented by circles, as shown in Fig. 3.

The vector of data m contains a change point if it can be split into two segments with length m_1 and m_2 satisfying Eq. 1:

$$C(m_1) + C(m_2) + \tau < C(m) \quad (1)$$

where τ is a threshold value, and C represents a cost function, which could be probability density or “means,” “variance,” or “linear” value of data. By selection of a threshold value and cost function, the abrupt change-point can be detected. In this paper, we define threshold values and compare mean values of data segments for CDA setting to discover a rising tendency in health data of plastic gears when degradation happened during run-to-failure tests in Sect. 4.3.

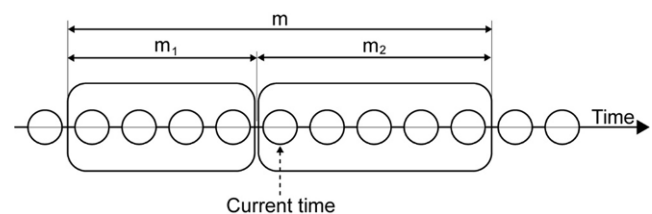


Fig. 3 The time series data represented by circles

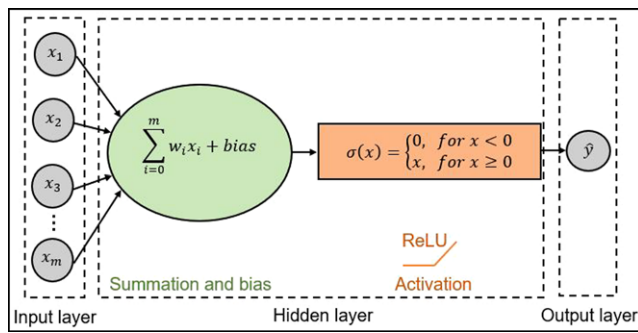


Fig. 4 A multilayer perceptron including three layers using Rectified Linear Unit (ReLU) activation

3.3 Artificial neural network

Artificial neural networks (ANNs) have emerged since the 1980s from developments in cognitive and computer science research [34]. However, applications of ANNs have become popular as a result of increasing computing power, increasing data size, and advanced deep learning research. A popular ANN model—Rumelhart’s multilayer perceptron (MLP), consists of at least three layers of neurons: an input layer, a hidden layer, and an output layer, as can be seen in Fig. 4.

The computation of the network is performed by calculating the weighted sum of the input vector with weight vector and bias. The weighted sum is then fed into the activation function to achieve an output. The computation of an ANN can be given as:

$$\hat{y} = \sigma (w^T \cdot x + b) \tag{2}$$

where: $x = [x_1, x_2, \dots, x_m]^T$ is input variable matrix with T samples and m features, w corresponding weight matrix, b is bias vector, and \hat{y} is output response variable vector. If the input variables are vibration data obtained from the accelerometer, the dimension of input matrix m will equal 1. The activation function σ defines the output of the neural network in a range of values avoiding the infinity value. The activation also adds nonlinearity to the output, which makes ANN capable to reflect complex nonlinear relationship in real-world applications. There are three common activation functions: ReLU, sigmoid, and tanh. Dependent on the designable target and characteristic of data, the activation function has to be suitably selected.

We employed an ANN to construct a data generator to generate HI from new input data automatically in this paper. The extracted features from vibration data can be considered as x , the label of extracted features is the response vector \hat{y} . The training dataset consists of extracted features and labels, which are used to train the network. The trained network’s weights and biases are used to predict the labels

of new input variables. The predicted label is considered as the HI of the plastic gear. To achieve desirable HI, we propose a labeling method, which is described in Sect. 4.5, and appointed ReLU as the activation function. The most commonly used activation function returns 0 if it receives any negative input, but for the positive value, it returns the value back. Hence, with the proposed labeling method, when the testing input data is lower than almost training input data, the trained ANN predicts a negative label. However, the activation returns the negative predicted label as 0, which is suitable for healthy state of plastic gears. If the predicted label is positive, the activation returns it back.

3.4 Linear regression

Many existing algorithms are based on deep neural models, which can model the non-linear complex relationship between the vibration data and RUL. However, in some cases, where knowledge of a suitable degradation model is unavailable, the linear model is the most natural choice to use [35]. In this study, linear regression is applied to estimate Weibull distribution parameters using the graphical method in Sect. 3.5. Moreover, in Sect. 5.2, the RUL of plastic gear is estimated by fitting a linear model using the generated HI. The linear model parameters are optimized by minimizing the residual sum of squares between the observed values in the dataset and the values predicted by the linear approximation.

3.5 Weibull distribution estimation

The Weibull Distribution (WD) has been widely studied since its introduction in 1951 by Professor Wallodi Weibull. The WD can reflect the fatigue strength and fatigue life of mechanical products and their parts under random loads [36]. WD functions are defined in Table 1 by three parameters: $\beta > 0$ is the shape parameter, $\eta > 0$ is the scale parameter, and $\gamma > 0$ is the location parameter. The data vector at time t is let as x_t . There are plenty of works that

Table 1 The Weibull distribution functions

Probability density function (pdf)	$f(x_t, \beta, \eta, \gamma) = \frac{\beta}{\eta} \left(\frac{x_t - \gamma}{\eta}\right)^{\beta-1} \exp\left(-\left(\frac{x_t - \gamma}{\eta}\right)^\beta\right)$
Cumulative distribution function (cdf)	$f(x_t, \beta, \eta, \gamma) = 1 - \exp\left(-\left(\frac{x_t - \gamma}{\eta}\right)^\beta\right)$
Reliability function	$R(x_t, \beta, \eta, \gamma) = \exp\left(-\left(\frac{x_t - \gamma}{\eta}\right)^\beta\right)$
Failure rate function	$\lambda(x_t, \beta, \eta, \gamma) = \frac{\beta}{\eta} \left(\frac{x_t - \gamma}{\eta}\right)^{\beta-1}$

Fig. 5 The influence of three parameters on the cdf

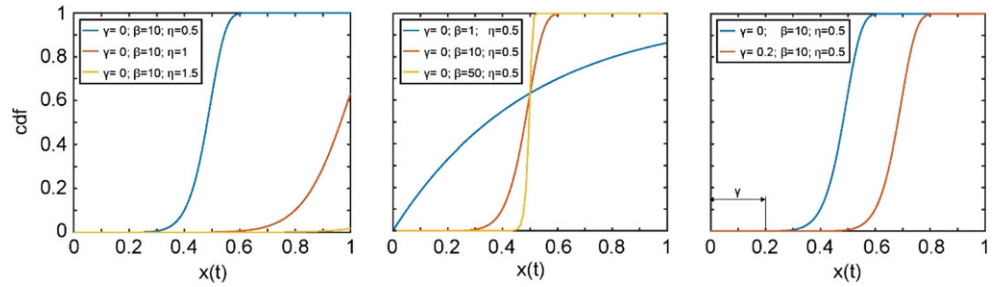
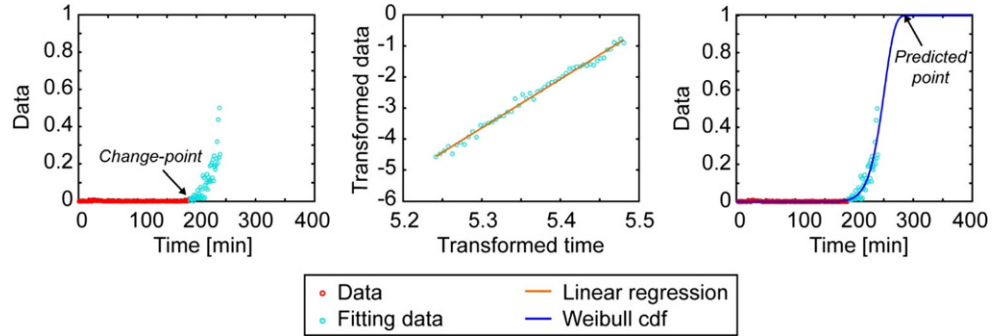


Fig. 6 An example of using graphical method for estimation of parameters for Weibull cdf



applied WD to fit the degradation profiles of machinery parts for RUL estimation, e.g. [24, 25]. In this paper, we employ a Weibull cumulative distribution function (cdf) to perform a physical-based prediction method. Fig. 5 shows the shapes of the cdf with changing parameters. The location parameter γ locates the distribution along the time axis. Changing the value of γ has the effect of “sliding” the distribution to the left ($\gamma < 0$) or to the right ($\gamma > 0$). We commonly use the distribution at the start time $t = 0$. Thus, the location parameter γ commonly is set as 0.

The common Weibull cdf model is reduced including two parameters by setting $\gamma = 0$, which is given as:

$$f(x_t, \beta, \eta, \gamma) = 1 - \exp\left(-\left(\frac{x_t}{\eta}\right)^\beta\right) \tag{3}$$

If both size of the Eq. 3 are transformed by $\ln(1/(1 - x_t))$, we get:

$$\ln\left(\frac{1}{1 - f(x_t, \beta, \eta, \gamma)}\right) = \left(\frac{x_t}{\eta}\right)^\beta \tag{4}$$

so that:

$$\ln\left[\ln\left(\frac{1}{1 - f(x_t, \beta, \eta, \gamma)}\right)\right] = \beta \ln x_t - \beta \ln \eta \tag{5}$$

If we let $Y = \ln\left[\ln\left(\frac{1}{1 - f(x_t, \beta, \eta, \gamma)}\right)\right]$, $X = \ln x_t$, and $c = -\beta \ln \eta$, then Eq. 5 represents a simple linear regression function corresponding to:

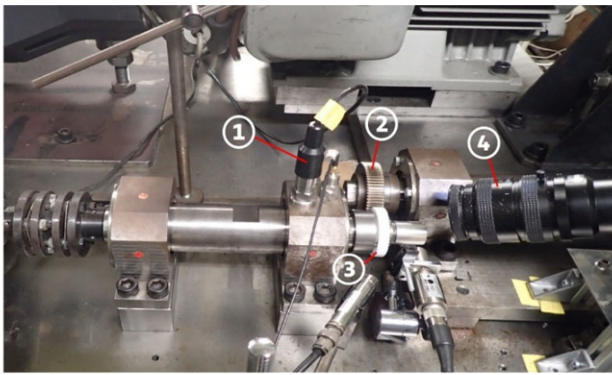
$$Y = \beta X + c \tag{6}$$

Hence, the two parameters of Weibull cdf can be estimated by using LR. This method is a common estimation method for WD called the graphical method [37]. Fig. 6 describes a mechanism of using the graphical method to shape Weibull cdf for fitting assumed health data. In the first step, the change point is identified to collect data for fitting, which is transformed to a logarithmic scale. Then, LR is deployed on transformed data to determine two parameters of the Weibull cdf. The predicted failure time point is when the Weibull cdf reaches 1. The WD is implemented after developing an HI of plastic gears in Sect. 5.3.

4 Experimental works

4.1 Data acquisition

The vibration data used in this paper are collected from a gear operation test rig shown in Fig. 7. In this figure, ① is an accelerometer, ② is the driving steel gear, ③ is the plastic test gear, and ④ is a high-speed camera monitoring crack at the tooth root of the test gear. POM (Polyoxymethylene) spur gears are the research objects, module is 1.0 mm, and



① Accelerometer ② Master gear ③ Test gear ④ High-speed camera

Fig. 7 The plastic gear operation test rig [38]

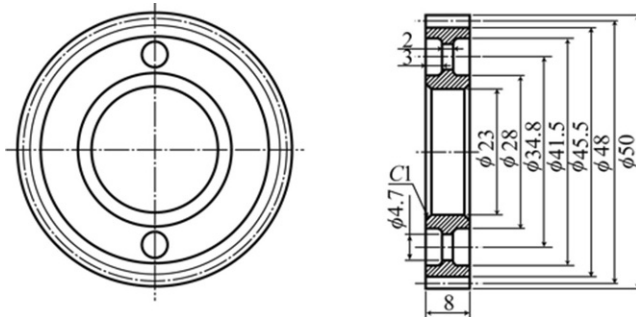


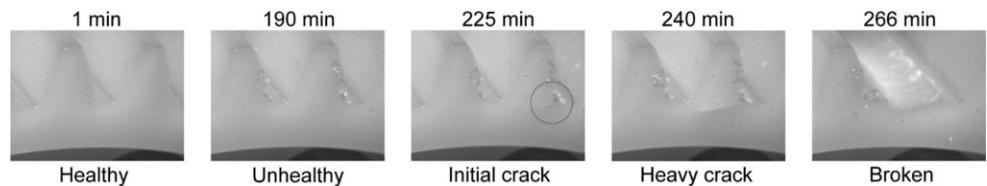
Fig. 8 The testing plastic gear specification

the number of teeth is 48, as can be seen in Fig. 8. The number of teeth of the driving steel spur gear is 67.

The testing condition was kept constant at 1000rpm rotational speed and 7Nm of torque loading applied to the plastic gear. Each run-to-failure test of the plastic gears was carried out from the initial stage until one tooth has broken. The time to failure can be precisely captured by the high-speed camera, e.g. at 266min in Fig. 9. The images captured before the break can only approximately reflect the current health status of the plastic gear.

In this research, seven run-to-failure tests were conducted to train and evaluate the proposed ANN. The initial crack time and time to failure (when at least one tooth

Fig. 9 Captured images from the high-speed camera during a run-to-failure test



has broken) were determined by monitoring the high-speed camera images. Results are listed in Table 2.

In each run-to-failure test, the data acquisition system collected one-second vibration data every one minute at 100,000Hz sampling frequency. Hence, each one-second vibration data consists of 100,000 data points. As the rotational speed is 1000rpm, the number of teeth of plastic gear is 48, the rotational frequency results as 16.67Hz, and the fundamental gear meshing frequency (GMF) would be 800.16Hz.

4.2 Feature extraction and selection

Extraction of robust features from raw and noisy vibration data plays a crucial role in tracking the degradation or health status of a gear. Typically, measurement values are efficiently extracted by using methods such as peaks, absolute mean amplitude (AMA), root mean square (RMS), Kurtosis factor (KF) [16], as can be seen in Table 3.

However, the characteristics of plastic gears are different from metal gears, e.g. thermal expansion, high deformation during operation, often dry-running operation [11]. As a result, the feature extraction from vibration data of plastic gear has met challenges. For example, Fig. 10 shows noisy features extracted from a run-to-failure test of plastic gear using the four mention statistical methods.

Thus, we propose a feature extraction method using FDM based on the following reasons. In gear meshing operation, the main vibration causes are fundamental gear meshing frequency and its harmonics. Any errors in gear manufacture or assembly or degradation will result in frequency sidebands surrounded GMF reflecting one per revolution modulation [39]. Fig. 11 shows frequency-domain representation of one-second vibration data for a healthy and a cracked plastic gear.

As can be seen, three orders of GMF harmonic, their sideband, and the nonlinearity frequencies of the test rig, e.g. 24Hz, are the main vibration frequencies. As shown in the enlarging plot, there is an occurrence of fundamental frequencies as shaft frequency in low-frequency band and

Table 2 The failure time monitored by the high-speed camera

Failure	Training test					Evaluation test	
	Test #1	Test #2	Test #3	Test #4	Test #5	Test A	Test B
Initial crack [min]	144	139	144	125	142	225	133
Broken [min]	182	172	174	167	172	266	163

Table 3 The four common statistic methods for feature extraction

Peak	$MAX x(i) $
Absolute mean amplitude (AMA)	$\frac{1}{N} \sum_i^N x(i) $
Root mean square (RMS)	$\sqrt{\frac{1}{N} \sum_i^N x(i)^2}$
Kurtosis factor (KF)	$\frac{1}{N} \sum_i^N x(i) ^4 / \left(\sqrt{\frac{1}{N} \sum_i^N x(i)^2} \right)^4$

an increasing amplitude of frequency sidebands surrounded GMFs when the crack happened. This phenomenon was also reported in simulation modeling of gear fault based on stiffness [40]. The emergence of increasing amplitude in a low-frequency band when the crack happened lead the authors to utilize specific narrow frequency bands for tracking the deterioration of plastic gears. A specific narrow frequency band, which reflects early failure and shows low noise, can be considered as SFB. The scheme of the proposed feature extraction method is shown in Fig. 12. From frequency-domain of vibration data, narrow frequency bands are selected to reconstruct data using inverse FFT. The AMA of reconstructed signals are assessed and compared to determine the SFB.

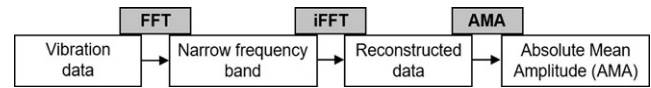


Fig. 12 A feature extraction scheme using FDM and AMA

Fig. 13 shows AMA values of reconstructed data from specific frequency bands. We investigated a frequency band from [0Hz, 1700Hz], which covers two fundamental GMF at 800.16Hz, 1600.32Hz, and their sidebands.

As can be seen, the result of AMA of reconstructed data from three frequency bands [1 Hz, 100Hz], [700Hz, 900Hz], [1500Hz, 1700Hz] are noisy. In the other cases, the degradation of plastic gear results in an increase of AMA of reconstructed signals from remain frequency bands during the test. Among the remaining frequency bands, the AMA of reconstructed data from frequency band [300Hz, 500Hz] shows a low-noise, early changing tendency, and low amplitude. Especially, the small value of AMA in the selected frequency band [300Hz, 500Hz] is evidence of a small effect from the main vibration frequencies. Therefore, we selected the AMA of reconstructed data from frequency band [300Hz, 500Hz] as robust extracted features of plastic gear for the next procedures.

Fig. 10 A noisy features extracted from raw vibration data using statistical methods

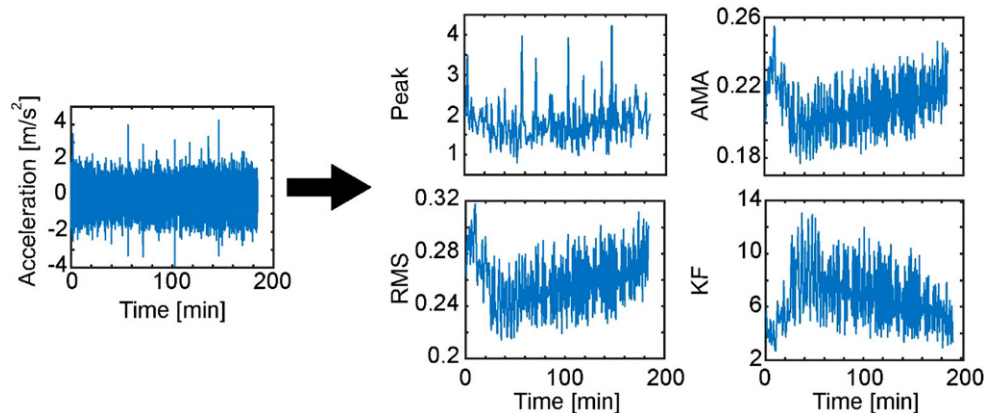


Fig. 11 Frequency-domain representations of one-second vibration data for a healthy and a cracked plastic gear

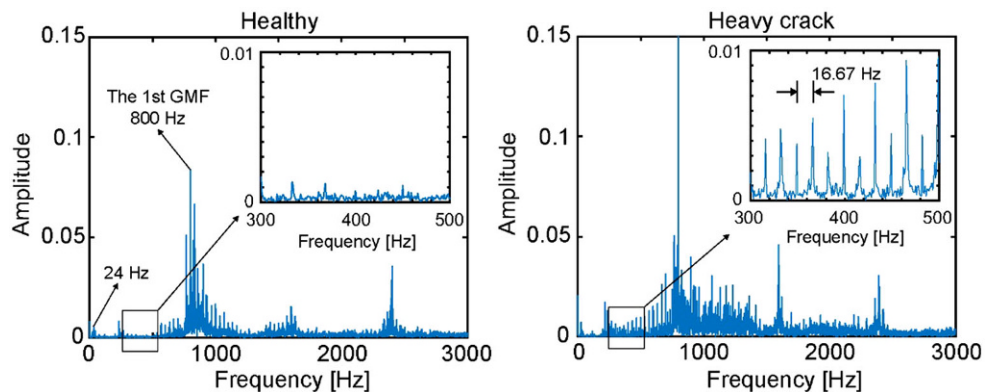
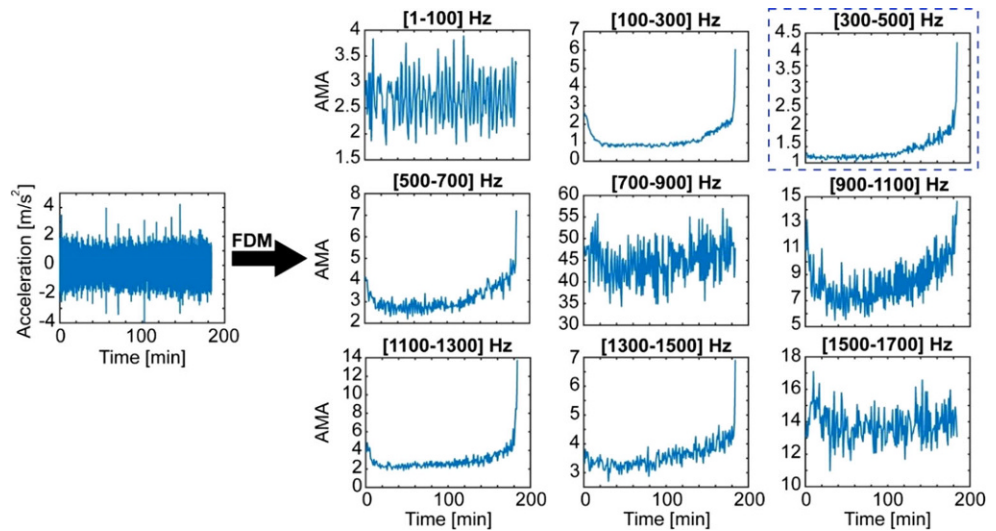


Fig. 13 AMA of reconstructed data in various specific narrow frequency bands



4.3 Change-point detection of health condition

The extracted AMA from frequency band [300 Hz, 500 Hz] includes three main significant areas, as can be seen in Fig. 14. There is an increasing tendency of AMA before the initial crack time, in the following designated as the changing area. The initial crack time is approximately captured by the high-speed camera in the initial crack area [41]. Meanwhile, the break is precisely detected by the high-speed camera when at least one tooth of plastic gear has broken.

Until now, there is a lack of explanation about the appearance of a change in health status extracted from vibration data before the initial crack happened in plastic gears. Several previous studies noted that wear failure, thermal failure, and fatigue failure are typical failure types that happened during a run-to-failure test of plastic gear [11, 42]. Further-

more, plastic gears work without lubrication, which may cause other types of failure before fatigue crack happened. Despite challenges in the explanation about the relationship between vibration monitoring and specific failure types of plastic gears, CDA shows an efficient performance in tracking changes in health data. To improve the quality of CDA application, we calculate average absolute mean amplitude (AAMA), which is given:

$$AAMA(i) = \frac{1}{n} \sum_{i-n}^i AMA(i) \tag{7}$$

where ($n=5$) is the sliding window length. The current value of AAMA is the average of five preceding AMA. This calculation makes data smooth based on historical data without using unknown future data. Using the principle of CDA explained in Sect. 2.1, the change point of data can be detected by comparison of the mean values of data segments using a specific threshold. Table 4 summarizes the detected change-point time of seven run-to-failure tests when the threshold decreases from 10–0.001. In this table, CDA discovers a constant change point if the threshold gets a high value, i.e., threshold 10 or 5. Besides, if the threshold takes a small value, i.e., threshold ≤ 0.01 , CDA reveals early change points.

In this paper, we considered three specific thresholds at 10, 1, and 0.01 reflecting change-point detection from late to early. Fig. 15 presents an example of the implementation of CDA in Test #1 corresponding to three thresholds. In the figure, the red solid line is AAMA, and the blue solid lines are the mean values of divided segments of data by CDA.

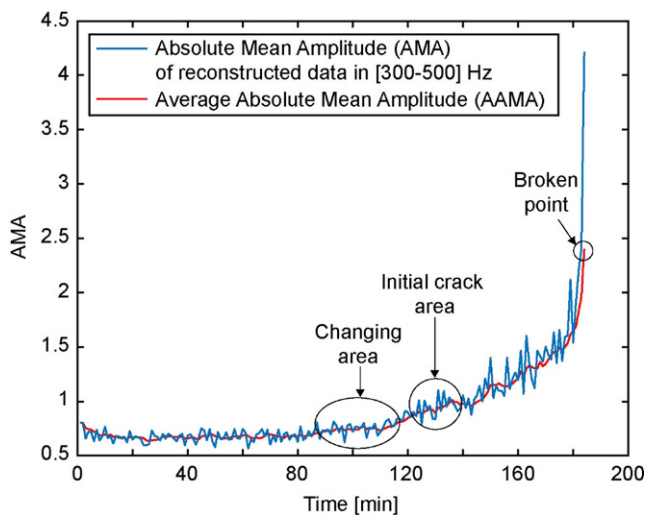
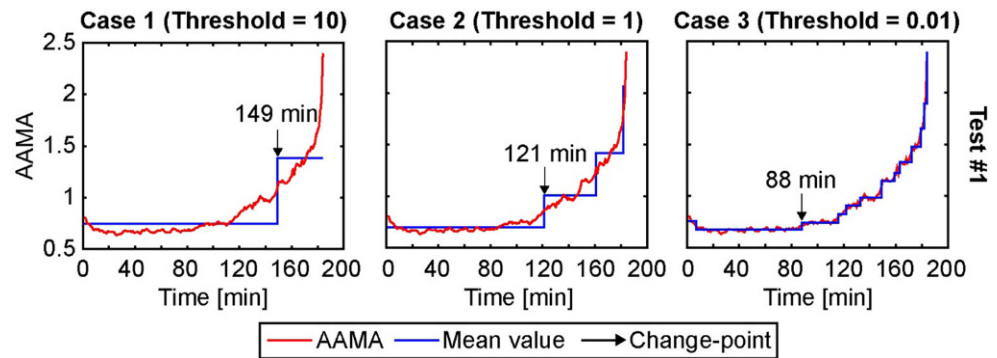


Fig. 14 The situations of plastic gear related to AMA and AAMA

Table 4 The detected change-points of training dataset

Threshold	Change-point time [min]						
	Test #1	Test #2	Test #3	Test #4	Test #5	Test A	Test B
10	149	142	129	150	131	218	135
5	149	142	129	142	131	218	135
1	121	131	123	131	104	204	106
0.5	121	129	109	131	104	192	102
0.1	89	126	94	107	97	191	90
0.05	89	106	88	76	97	180	89
0.01	88	106	86	78	95	180	81
0.005	87	106	86	78	95	168	81
0.001	86	105	86	79	95	168	81

Fig. 15 An example of the change-point detection by using CDA



4.4 Data normalization

Data normalization is an important step that guarantees the learning of the deep learning model reaches stable convergence of network weights and biases. Unscaled input data can result in a slow learning process or exploding gradients causing the failure of the learning process. In this work, the AAMA was normalized to be within a range of [0, 1], namely NAAMA, using the min/max normalization method. Besides, a scaler of the training dataset was established to normalize new real-time testing data with unknown extreme values. In other words, the new real-time testing data will be normalized to be within a range of [0, 1] using minimum and maximum values of the historical training data set. In the unusual case, if there exist extreme values in a new data set that is out of the training data set, it will be necessary to update and re-train the HIG. The scheme of real-time testing data normalization is depicted

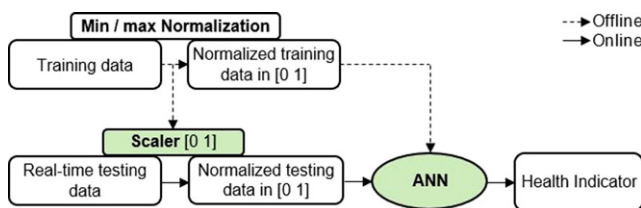


Fig. 16 The real-time data normalization scheme

in Fig. 16. The result of NAAMA from five run-to-failure tests from the data normalization process is shown in Fig. 17.

4.5 Learning of ANN

In this paper, we utilize supervised learning of ANN as a regressor to map the relationship between NAAMA and labels, which required data composed of labels. This means that NAAMA and design labels of training data are used to train the ANN. Then, the trained ANN is used to predict the labels of new input NAAMA.

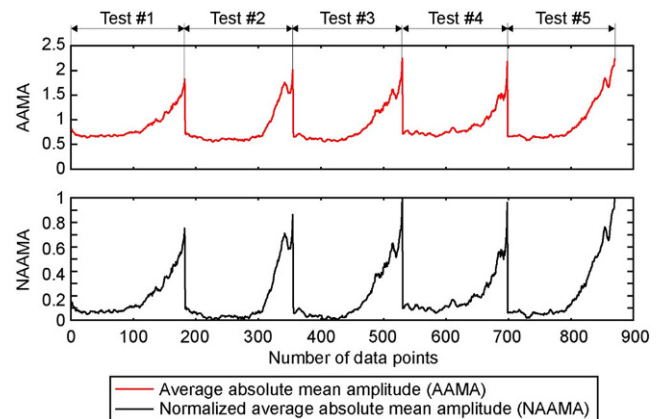


Fig. 17 Data normalization results of the training dataset

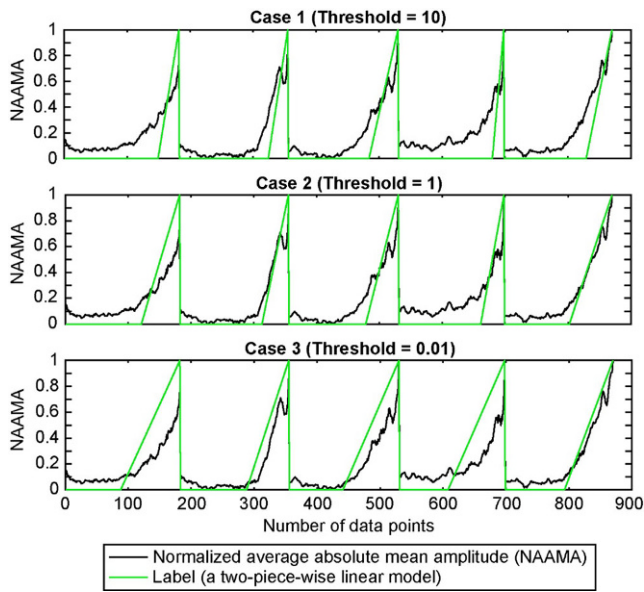


Fig. 18 Labeling method for training data in three cases

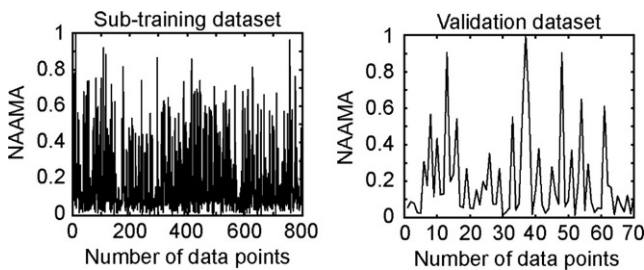
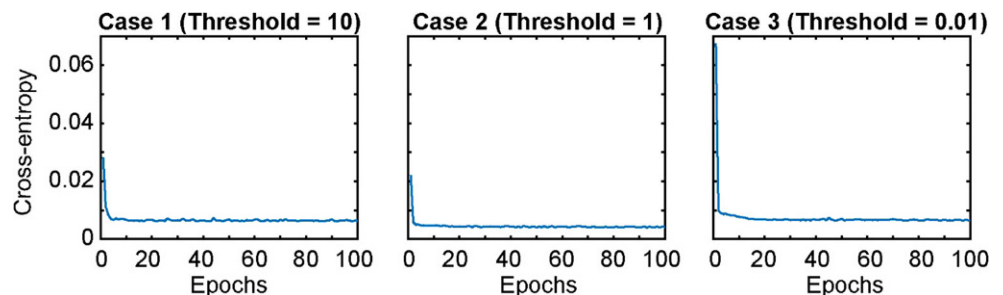


Fig. 19 Random sub-training dataset and validation dataset

In the labeling method, to label NAAMA, we proposed a two-piece-wise linear model, which consists of two sections: healthy states of plastic gear from the first stage until change-point time detected using CDA, and performance degradation from the change-point time until plastic gear has broken. In this method, the broken time point is precisely identified by the high-speed camera. While the change points have been caught by CDA using three specific thresholds. Fig. 18 shows three cases of the labeling method for five run-to-failure tests corresponding to three change-point determinations.

Fig. 20 Cross-entropy loss results of learning



Randomization of training data ensures that the learning of ANNs is generalized. Hence, after labeling, the training dataset was randomized and divided into two groups with a ratio of 92% for the sub-training dataset and 8% for the validation dataset, as can be seen in Fig. 19.

The batch size was taken as 16. The backpropagation algorithm and gradient descent method were used to optimize the learning procedure [43]. The training time consumed around three seconds on an i7 6700 processor at 3.4GHz with 16GB memory. Cross-entropy loss function and root mean square error (RMSE) are two metrics used to evaluate the learning performance of ANN. Loss results express the fast convergence of learning after less than 10 iterations, as can be seen in Fig. 20.

RMSE computes the difference between actual labels and predicted labels to validate performance of trained ANN is given by:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{8}$$

where y_i is the actual value, \hat{y}_i is the predicted value, and n is the number of observations. In Fig. 21, the green line is the actual labels of the input NAAMA in the validation dataset, while the pink line is predicted labels computed by trained ANN. The labeling method in Case 3 is able to reliably determine the early change point, however, the small RMSE of 0.390 proved that the learning process of the ANN with the labeling method in Case 1 is more accurate. Additionally, the evaluation of RMSE in Table 5 considering three cases of labeling method and various sizes of ANN, reveals the optimal configuration of this ANN. As mention in Sect. 3.3, a three layer ANN is sufficient for the large majority of problems. For example, Model #1 in the table consists of one input layer with 64 units, one hidden layer with 64 units, and one output layer with one unit. With an increase in the number of hidden layers, such as Model #2–#5, the change of RMSE is negligible. This result proved that a simple ANN with three basic layers is sufficient for data in this study.

Fig. 21 RMSE between predicted labels and actual labels of NAAMA

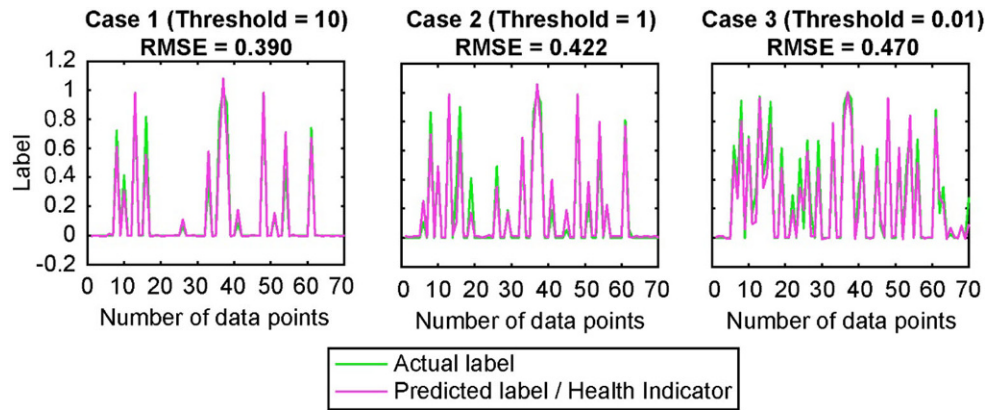


Table 5 Effect of the size of ANN architecture on RMSE

Model	Number of neural in layers	RMSE		
		Case 1	Case 2	Case 3
#1	64–64–1	0.390	0.422	0.470
#2	64–64–32–1	0.388	0.431	0.464
#3	64–64–32–16–1	0.389	0.420	0.469
#4	64–64–32–16–8–1	0.376	0.426	0.473
#5	64–64–32–16–8–4–1	0.386	0.418	0.470

5 Results and discussion

In this section, the HI of the plastic gear is first introduced. The generated HI can be used to detect early deviations from the healthy state and is capable to estimate the RUL of plastic gears using three prediction strategies.

5.1 Health indicator of plastic gear

With the purpose of generating HI for plastic gears, two real-time run-to-failure tests, namely Test A and Test B, were pre-processed as an input for the trained ANN, called HIG. Firstly, NAAMA was extracted from the raw vibration data of the two tests. Then the extracted NAAMA was fed into the HIG to predict the label. If the predicted label efficiently reflects the health status of plastic gear, it will be suitable to be considered as HI of the plastic gear. Fig. 22 shows the result of HI generated from the two real-time tests using the HIG. The black solid line represents the input NAAMA. The green solid line is the actual label and the pink solid line is the output predicted label. In which, the actual label is the piece-wise model generated by the proposed labeling method as the explanation in Sect. 4.5. The actual label reflects perfectly health status of the plastic gear including the healthy situation with value 0, the break

Fig. 22 Generated HI of Test A and Test B in three cases of labeling method

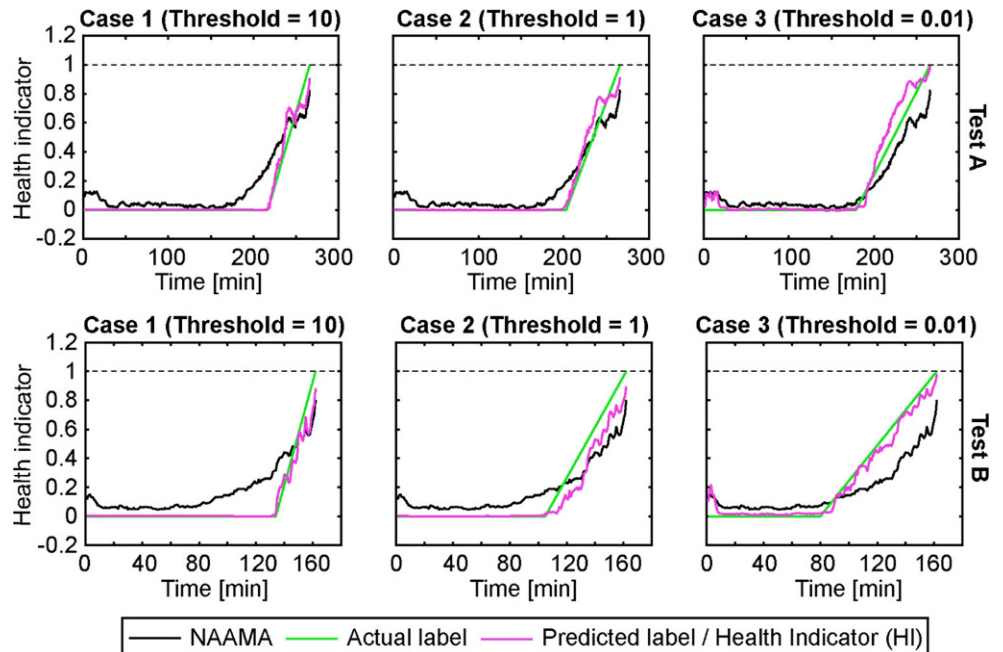
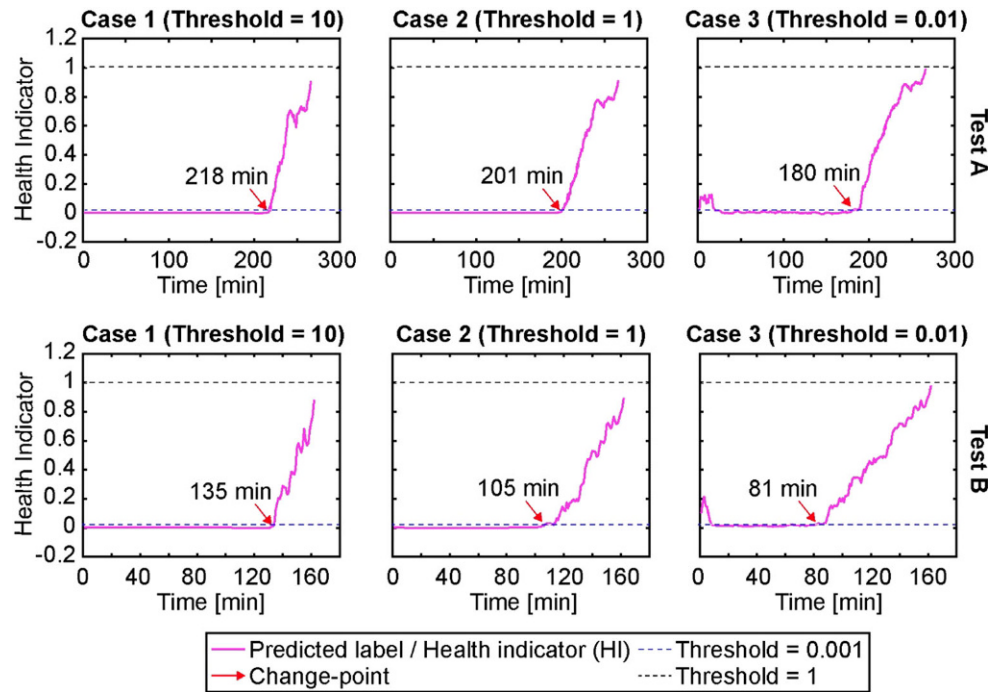


Fig. 23 The change-points of predicted labels



with value 1, and the linear degradation status with values from 0 up to 1. However, the generating of actual labels only could be determined using CDA when the run-to-failure tests finished. Thus, the actual label is used to evaluate the predicted label considering the possibility of using the predicted label as the HI of the plastic gear.

As can be seen that the NAAMA can reflect the condition of plastic gear with low values for healthy status and increasing values during degradation period. However, the predicted label of NAAMA expresses outstanding characteristics to be considered as HI as following.

- The first characteristic: The predicted label can be used to efficiently represent the healthy situation of plastic gears with the value approximated the absolute 0 value. In addition, with the labeling method Case 3, the predicted la-

bel may reflect the early stage of plastic gear with a high wear rate [42].

- The second characteristic: the change-point of the predicted label reveals deviations from the healthy state when it takes a value greater than a threshold, as shown in Fig. 23. The threshold should be greater than 0 and small enough to early detect the change of the HI from 0, e.g. the minimum threshold 0.001 of using CDA as results in Sect. 4.3, The change point of the predicted label is equivalent to the change point of the actual label detected by using CDA, as shown in Table 6.
- The third characteristic: it is hard to design an absolute precise health indicator, which equals 1 when a plastic gear is broken due to the imperfections in gear manufacturing and in experimental settings, e.g. gear geometric errors or the difference in the orientation and position

Table 6 Comparison between change-points of the predicted labels and the actual labels

Indicator	Test A			Test B		
	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
Predicted label	218	201	180	135	105	81
Actual label	218	204	180	135	106	81

Table 7 Comparison between values of NAAMA and the predicted labels

Indicator	Test A			Test B		
	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
NAAMA	0.826	0.826	0.826	0.801	0.801	0.801
Predicted label	0.908	0.913	0.994	0.881	0.895	0.980
Actual label	1	1	1	1	1	1

of gears between an experimental situation and a perfect situation. However, at the time to failure, the predicted labels are greater than the corresponding NAAMA and approach absolute 1 value, as seen in Table 7. This specification reveals the potential of accuracy improvement of prognostic tasks based on value thresholds of HI.

With the above mention characteristics, we use the predicted label as the HI for plastic gear. The generated HI not only reflects situations of plastic gear during the working time but also is capable to release efficient and precise diagnosis results even when using common and simple RUL prediction strategies.

5.2 RUL estimation

Three strategies were applied to predict the RUL of plastic gears with the HI generated from two evaluation tests. The HI in the healthy situation of plastic gear approximately equals 0 and is not useful for prediction tasks. The prediction strategies could be performed using the HI collected from the change-point until the current time when the prediction is made, namely “fitting HI”. Fig. 24 illustrates the utilization of three prediction strategies in the same scenarios: Test A, labeling method Case 3 with change-point of HI is at 180min and the current time is at 200min for the prediction of the time to failure.

Three prediction strategies are described by the following mechanisms.

The LR strategy matches a linear model close to the fitting HI. The predicted time to failure is decided when the linear model reaches threshold 1, as can be seen in Fig. 24a.

The EWD strategy utilizes the graphical method to estimate two parameters of Weibull cdf from fitting HI data. The predicted time to failure is estimated when the Weibull cdf model equals 1, as can be seen in Fig. 24b.

The HI-ARUL strategy employs HI and average RUL of conducted run-to-failure tests to estimate the time to failure, as can be seen in Fig. 24c. Naturally, the RUL of plastic gear by percentage can be calculated from the generated HI using Eq. 9:

$$RUL_{Pred_Percentage} = (1 - HI) \times 100\% \tag{9}$$

Where $RUL_{Pred_Percentage}$ is the predicted RUL in percentage and HI is the current health indicator. However, the predicted RUL can also be specifically calculated by time unit according to:

$$RUL_{Pred_Time} = RUL_{Pred_Percentage} \times RUL_{Average} \tag{10}$$

Where RUL_{Pred_Time} is the predicted RUL in time unit and $RUL_{Average}$ is the average RUL measuring by time calculated from historical run-to-failure tests. Table 8 shows the average RUL calculated from five run-to-failure tests considering three cases of change-point detection. In this table, the RUL of plastic gear in each case is computed as the subtraction of time to failure and change-point time. Accordingly, in three cases of change-point detection, the average RUL of plastic gears is 33.4, 51.4, or 82.8 min, respectively.

The time to failure predicted by the HI-ARUL strategy can be calculated as:

$$Time_{Pred_Broken} = Time_{Current} + RUL_{Pred_Time} \tag{11}$$

Where, $Time_{Pred_Broken}$ is the predicted time to failure, $Time_{Current}$ is the current time when the prediction is made. For the example depicted in Fig. 24c, as the change-point time of HI is 180 min, the prediction is made at 200 min, and the HI at 200 min equals 0.2711, the results are:

$$RUL_{Pred_Percentage} = (1 - 0.2711) \times 100\% = 72.89\% \tag{12}$$

$$RUL_{Pred_Time} = 72.89\% \times 82.8 = 60.35(\text{minutes}) \tag{13}$$

$$Time_{Pred_Broken} = 200 + 60.35 = 260.35(\text{minutes}) \tag{14}$$

Fig. 24 An example for the prediction of time to failure using three strategies

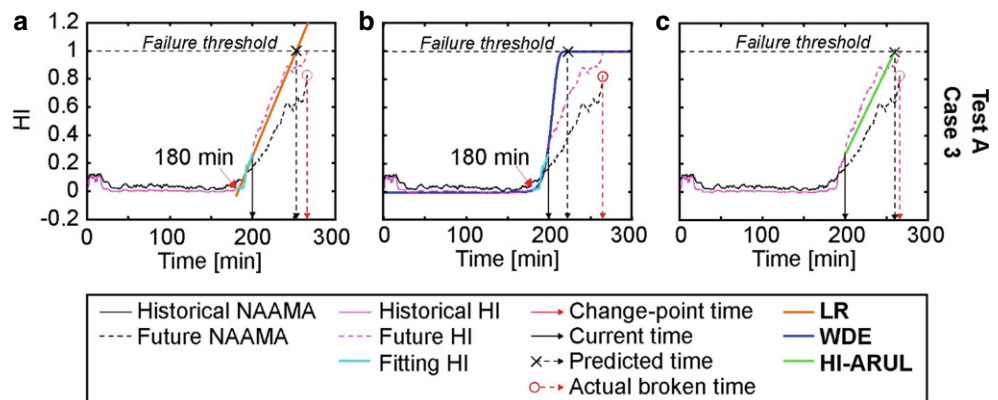


Table 8 Average RUL of five run-to-failure tests corresponding three cases

Test	Broken time [min]	Change-point time [min]			RUL [min]		
		Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
Test #1	182	149	121	88	33	61	94
Test #2	172	142	131	106	30	41	66
Test #3	174	129	123	86	45	51	88
Test #4	167	150	131	78	17	36	89
Test #5	172	130	104	95	42	68	77
Average RUL					33.4	51.4	82.8

Table 9 Prediction errors using LR for Test A

Case	Change-point time [min]	Error [%]								
		190 min	200 min	210 min	220 min	230 min	240 min	250 min	260 min	266 min
1	218	–	–	–	<i>0.75</i>	<i>4.13</i>	<i>4.13</i>	<i>1.88</i>	<i>1.50</i>	<i>2.63</i>
2	201	–	–	<i>3.76</i>	<i>2.63</i>	<i>4.88</i>	<i>4.51</i>	<i>2.63</i>	<i>0</i>	<i>1.13</i>
3	180	–	<i>4.13</i>	<i>8.64</i>	<i>8.27</i>	<i>7.89</i>	<i>6.76</i>	<i>5.26</i>	<i>3.38</i>	<i>2.25</i>

Table 10 Prediction errors using LR for Test B

Case	Change-point time [min]	Error [%]								
		90 min	100 min	110 min	120 min	130 min	140 min	150 min	160 min	162 min
1	135	–	–	–	–	–	<i>2.47</i>	<i>9.87</i>	<i>8.02</i>	<i>6.79</i>
2	105	–	–	–	<i>46.9</i>	<i>37.6</i>	<i>12.9</i>	<i>9.26</i>	<i>8.02</i>	<i>7.41</i>
3	81	–	<i>6.79</i>	<i>3.08</i>	<i>3.70</i>	<i>6.79</i>	<i>4.93</i>	<i>4.93</i>	<i>4.93</i>	<i>4.93</i>

Table 11 Prediction errors using EWD for Test A

Case	Change-point time [min]	Error [%]								
		190 min	200 min	210 min	220 min	230 min	240 min	250 min	260 min	266 min
1	218	–	–	–	<i>10.9</i>	<i>8.27</i>	<i>4.13</i>	<i>0.37</i>	<i>6.01</i>	<i>8.27</i>
2	201	–	–	<i>13.1</i>	<i>9.02</i>	<i>5.64</i>	<i>2.25</i>	<i>1.50</i>	<i>6.01</i>	<i>8.27</i>
3	180	<i>14.6</i>	<i>16.9</i>	<i>14.3</i>	<i>10.5</i>	<i>7.14</i>	<i>3.76</i>	<i>0.75</i>	<i>2.63</i>	<i>4.13</i>

Table 12 Prediction errors using EWD for Test B

Case	Change-point time [min]	Error [%]								
		90 min	100 min	110 min	120 min	130 min	140 min	150 min	160 min	162 min
1	135	–	–	–	–	–	<i>0.00</i>	<i>15.4</i>	<i>16.6</i>	<i>15.4</i>
2	105	–	–	<i>0.00</i>	<i>2.46</i>	<i>2.46</i>	<i>4.32</i>	<i>8.02</i>	<i>11.7</i>	<i>12.3</i>
3	81	–	–	<i>10.5</i>	<i>1.23</i>	<i>0.61</i>	<i>1.23</i>	<i>3.70</i>	<i>6.79</i>	<i>6.79</i>

The actual time to failure captured by the high-speed camera in test A is at 266 min. Finally, the predicted RUL by time unit of plastic gear can be calculated as:

$$RUL_{Pred_Time} = Time_{Pred_Broken} - Time_{Current} \tag{15}$$

To compare the effectiveness of the three prediction strategies mentioned above, Table 9, 10, 11, 12, 13 and 14 record errors between the predicted time to failure and the actual time to failure during the working time of the two real-time tests. The prediction is making every ten minutes from change-point time until the end of the test using and

comparing three cases of labeling method. The error can be simply calculated as:

$$Error = \left(1 - \frac{|Time_{Pred_Broken} - Time_{Actual_Broken}|}{Time_{Actual_Broken}} \right) \times 100\% \tag{16}$$

Where $Time_{Actual_Broken}$ is the actual time to failure captured by the high-speed camera. Values of errors are smaller than 5%, which express a good prediction performance, are shown in italics. However, it is impossible to predict the broken time of plastic gears before the change point hap-

Table 13 Prediction errors using HI-ARUL for Test A

Case	Change-point time [min]	Error [%]								
		190 min	200 min	210 min	220 min	230 min	240 min	250 min	260 min	266 min
1	218	–	–	–	5.80	5.39	5.71	2.04	1.45	1.15
2	201	–	–	4.24	5.22	5.28	5.17	1.51	1.99	1.68
3	180	1.02	2.12	3.84	5.68	5.61	5.86	2.24	1.19	0.19

Table 14 Prediction errors using HI-ARUL for Test B

Case	Change-point time [min]	Error [%]								
		90 min	100 min	110 min	90 min	130 min	140 min	90 min	160 min	162 min
1	135	–	–	–	–	–	1.10	1.71	4.33	2.45
2	105	–	–	1.34	1.92	6.15	1.30	1.84	5.21	3.33
3	81	1.92	2.44	1.17	2.60	6.63	0.72	1.43	3.74	1.03

pened when the HI approximately equals 0. Hence, in these scenarios, the results are shown as dashes.

According to the prediction errors listed in Table 13 and 14, HI-ARUL expresses outstanding prediction results. The HI-ARUL method can perform a prediction at the very early stage, when the change-point time of HI has been detected, with errors are 1.02% and 1.92%. The errors are slightly changed during the whole working time of plastic gears. However, the fluctuation area in degradation data [18] is unavoidable causing the higher error. For example, in Table 14, the error of prediction at 130 min is 6.63% for Test B using the labeling method Case 3. Additionally, the accuracy of this method depends on the average RUL of the dataset. Test A has a long time of test (266 min), which is different from almost all tests, causing higher errors than Test B. Table 13 proves the robustness of the proposed method in comparison with the other two prediction strategies even when the real-time input test is different from the tests for training.

Although LR and EWD are established methods used for RUL prediction of rotating machinery but have met challenges. For example, LR can be adaptive with Test A but not Test B when the HI is largely different from the linear model. The EWD is based on linear regression of graphical method to estimate parameters. A slight error of linear approximation in the graphical method can cause a significant error in shaping WD. Additionally, in the initial stage of prediction, e.g., in Table 12, errors are high when the EWD method is performed at the current time of 90 or 100 min even when the change-point time is at 81 min resulting in blank results.

6 Conclusion

Health indicator construction is the key to achieve a precise RUL prediction, which is crucial to perform prognostics and health management of rotating machinery for failure

prevention and maintenance cost reduction. In this paper, HI of plastic gears were automatically generated from two real-time run-to-failure tests by HIG based on ANN. The HIG was trained on input data using data pre-processing techniques (FDM and AMA for robust feature extraction, CDA for labeling method). Besides, the sensitivity to failure of the HIG was considered corresponding to three cases of the labeling method. With the most sensitive level (Case 3), the HI is used not only for early failure detection but also to predict the time to failure of plastic gears from the initial stage of performance degradation. Among three considered prediction strategies, the HI-ARUL expresses an outstanding performance with acceptable errors (<5%) and the RUL predictability from the early change point of HI during the whole working time of plastic gear. Although the construction and evaluation of the data generator are based on a limited amount of data from seven run-to-failure tests, the proposed method reveals applicability in monitoring and maintenance for plastic gears. Especially, the transferring of proposed techniques to an end-to-end structure based on ANN, which automatically outputs HI and RUL estimation, can be convenient for users.

References

1. Mobley RK (2001) Gears and gears drives. In: Plant engineer's Handbook. Butterworth-Heinemann, Boston
2. Litvin FL (2016) Theory and practice of gearing and transmission. Mechanisms and machine science, vol 34. Marco Ceccarelli, Cassino
3. Mao K (2007) Gear tooth contact analysis and its application in the reduction of fatigue wear. *Wear* 262:1281–1288
4. Basan R, Franulovic M, Previl I, Crnjacic-Zic N (2011) Analysis of strain-life fatigue parameters and behavior of different group of metallic materials. *Int J Fatigue* 33:484–491
5. Patil SS, Karuppanan S, Atanasovska I, Wahab AA (2014) Contact stress analysis of helical gear pairs, including frictional coefficients. *Int J Mech Sci* 85:205–211
6. Davis JR (2005) Gear materials, properties, and manufacture. ASM International, Materials Park

7. Elmquist JL (2014) Deciding when to go plastic. *Gear Technology* July 2014:46–47
8. Bravo A, Koffi D, Toubal L, Erchiqui F (2015) Life and damage mode modeling applied to plastic gears. *Eng Fail Anal* 58:113–133
9. Yelle H, Koffi D, Guilbault R (1999) Comportement à l'usure et à l'échauffement des engrenages en plastique en fonction des matériaux de la paire. In: *World congress on gearing and power transmission*, pp 713–722
10. Walton D, Shi YW (1989) A comparison of ratings for plastic gears. *Proc Inst Mech Eng Part C Mech Eng Sci* 203(1):31–38
11. Jain M, Patil S, Ghosh SS (2019) A review on failure characteristics of polymeric gears. In: *AIP Conference Proceeding*, vol 2148
12. Heng A, Zhang S, Tan ACC, Mathew J (2009) Rotating machinery prognostics: state of the art, challenges and opportunities. *Mech Syst Signal Process* 23(3):724–739
13. Lee J, Wu F, Zhao W, Ghaffari M, Liao L, Siegel D (2014) Prognostics and health management design for rotary machinery systems—reviews, methodology and applications. *Mech Syst Signal Process* 42(1/2):314–334
14. Benkedjough T, Medjaher K, Zerhouni N, Rechak S (2013) Remaining useful life estimation based on nonlinear feature reduction and support vector regression. *Eng Appl Artif Intell* 26:1751–1760
15. Bogue R (2013) Sensors for condition monitoring: a review of technologies and applications. *Sens Rev* 33(4):295–299
16. Xiang S, Qin Y, Zhu C, Wang Y, Chen H (2020) Long short-term memory neural network with weight amplification and its application into gear remaining useful life prediction. *Eng Appl Artif Intell* 91:103587
17. Mahamad AK, Saon S, Hiyama T (2010) Predicting remaining useful life of rotating machinery based artificial neural network. *Comput Math Appl* 60:1078–1087
18. Ali JB, Morello BC, Saidi L, Malinowski S, Fnaiech F (2015) Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network. *Mech Syst Signal Process* 56/57:150–172
19. Zhang Z, Wang Y, Wang K (2013) Fault diagnosis and prognosis using wavelet packet decomposition, Fourier transform and artificial neural network. *J Intell Manuf* 24:1213–1227
20. Wu B, Li W, Qiu M (2017) Remaining useful life prediction of bearing with vibration signals based on a novel indicator. *Shock Vib*. <https://doi.org/10.1155/2017/8927937>
21. Ji S, Han X, Hou Y, Song Y, Du Q (2020) Remaining useful life prediction of airplane engine based on PCA-BLSTM. *Sensor* 20(16):4537
22. Zhang J, Wang P, Yan R, Gao RX (2018) Long short-term memory for machine remaining life prediction. *J Manuf Syst* 48(C):78–86
23. Wang D, Tsui KL, Miao Q (2018) Prognostics and health management: a review of vibration based bearing and gear health indicators. *IEEE Access* 6:665–676
24. Bechhoefer E, Morton B (2012) Condition monitoring architecture: to reduce total cost of ownership. In: *IEEE Conf. on Prognostics and Health Management*, pp 1–9
25. Zheng R (2015) Parameter estimation of three-parameter Weibull distribution and its application in reliability analysis. *J Vib Shock* 34(5):78–81
26. Li X, Cai J, Zuo H, Chen X, Mao H, Xu Y (2017) Reliability sequential compliance method for a partially observable gear system subject to vibration monitoring. *J Vibroeng* 19(5):3313–3334
27. Zhao R, Yan R, Chen Z, Mao K, Wang P, Gao RX (2019) Deep learning and its applications to machine health monitoring. *Mech Syst Signal Process* 115:213–237
28. Tian Z (2009) An artificial neural network method for remaining useful life prediction of equipment subject to condition monitoring. *J Intell Manuf* 23(2):143–148
29. Guo L, Li N, Jia F, Lei Y, Lin J (2017) A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing* 240:98–109
30. Nistane VM (2020) Comparative performance of prognostics for remaining useful life of bearing. *Noise Vib Worldw* 51(11):208–222
31. Dong S, Luo T (2013) Bearing degradation process prediction based on the PCA and optimized LS-SVM model. *Measurement* 46(9):3143–3152
32. Iba D, Ohmori S, Hongu J, Nakamura M, Moriwaki I (2013) Failure detection of plastic gears based on comparison of Fourier coefficients of a gear mesh vibration model by rectangle pulse train and frequency analysis of acceleration response. *Trans Japan Soc Mech Eng* 79:5138–5148 (in Japanese)
33. Desobry F, Davy M, Doncarli C (2005) An online kernel change detection algorithm. *IEEE Trans Signal Process* 53(8):2961–2974
34. Yang ZR, Yang Z (2014) 6.01—Artificial neural networks. *Compr Biomed Phys* 6:1–17
35. Babu GS, Zhao P, Li XL (2016) Deep convolutional neural network based regression approach for estimation of remaining useful life. *Lecture notes in computer science*, vol 9642
36. Tao P, Haiyan J, Yong X (2011) Weibull distribution parameters for fault feature extraction of rolling bearing. In: *Chinese Control and Decision Conference (CCDC)*, pp 69–74
37. Nwobi F, Ugomma CA (2014) A comparison of methods for the estimation of Weibull distribution parameters. *Metodoloski Zv* 11(1):65
38. Ishii Y, Iba D, Miyamoto S, Miura N, Iizuka T, Masuda A, Sone A, Moriwaki I (2018) Automatic acquisition system of meshing vibration data and image pictures of crack propagation. In: *Proceedings of the machine design and tribology division meeting in JSME* (in Japanese)
39. Sawali N (2016) Vibration sideband modulations and harmonics separation of a planetary helicopter gearbox with two different configurations. *Adv Acoust Vib*. <https://doi.org/10.1155/2016/9820768>
40. Chaari F, Baccar W, Abbes MS, Haddar M (2008) Effect of spalling or tooth breakage on gearmesh stiffness and dynamic response of a on-stage spur gear transmission. *Eur J Mech A Solids* 27:691–705
41. Kien BH, Iba D, Ishii Y, Tsutsui Y, Miura N, Iizuka T, Masuda A, Sone A, Moriwaki I (2019) Crack detection of plastic gears using a convolutional neural network pre-learn from images of meshing vibration data with transfer learning. *Forsch Geb Ingenieurwes* 83(3):645–653
42. Ghazali WM, Idris DMND, Sofian AH, Siregar JP, Aziz IAA (2017) A review on failure characteristics of polymer gear. *MATEC Web Conf* 90(01029):1–10
43. Vogl TP, Mangis JK, Rigler AK, Zink WT, Alkon DL (1988) Accelerating the convergence of the back-propagation method. *Biol Cybern* 59:257–263