



# A generalized cauchy method for remaining useful life prediction of wind turbine gearboxes

He Liu<sup>a</sup>, Wanqing Song<sup>a,\*</sup>, Yuhui Niu<sup>a</sup>, Enrico Zio<sup>b,c,d</sup>

<sup>a</sup> School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China

<sup>b</sup> Energy Department, Politecnico di Milano, Via La Masa 34/3, 20156, Italy

<sup>c</sup> Department of Medical Physics and Computer Science, Bashkir State Medical University, Lenina st. 3, 450008 Ufa, Russia

<sup>d</sup> MINES ParisTech, PSL Research University, CRC, Sophia Antipolis, France



## ARTICLE INFO

### Article history:

Received 18 July 2020

Received in revised form 13 October 2020

Accepted 16 November 2020

### Keywords:

Generalized Cauchy process

Long-range dependent

Fractal

Gearbox degradation

Remaining useful life

## ABSTRACT

The accurate estimate of the Remaining Useful Life (RUL) of mechanical tools is a fundamental problem in Engineering. This prediction often implies the knowledge and application of sophisticated mathematical methods based on fractal and Long-Range Dependence (LRD) stochastic processes. However, the existing RUL prediction methods based on stochastic model cannot simultaneously consider the fractal and LRD characteristics of the equipment degradation process. This paper describes a new RUL prediction model based on the Generalized Cauchy (GC) process, which is a stochastic process with independent parameters. That is, the GC process uses the fractal dimension  $D$  and Hurst index  $H$  to describe the fractal and LRD characteristics of the degradation sequence, respectively. Then, the GC process is taken as the diffusion term, describing the uncertainty of the degradation sequence, to establish the GC degradation model, and the power law and exponential forms are used to describe the nonlinear drift of the degradation sequence. The stochastic volatility of the degradation sequence causes the equipment RUL unable to be predicted for a long time. This article uses the largest Lyapunov index to reveal the maximum prediction range of RUL. The analysis of actual equipment degradation verifies the effectiveness of the degradation model based on power law drift and GC process. The prediction results of the comparative case show that the prediction performance of the GC degradation model is better than Brownian motion, fractional Brownian motion, and long short-term memory neural network.

© 2020 Elsevier Ltd. All rights reserved.

## 1. Introduction

In recent years, how to solve the problems of environmental pollution and the reduction of non-renewable energy have become increasingly crucial. As a renewable green energy source, the wind power has received much more attention [1,2]. The gearbox, one of the most important components of wind turbine, whose failure will lead to the loss of wind power generation. Therefore, predictive maintenance of the gearbox is key to improving the reliable operation and power generation of wind turbine. The accurate RUL prediction provides the basis for effective and optimized forecast maintenance.

\* Corresponding author.

E-mail addresses: [nh324310@163.com](mailto:nh324310@163.com) (H. Liu), [swqls@126.com](mailto:swqls@126.com) (W. Song), [n18321023297@163.com](mailto:n18321023297@163.com) (Y. Niu), [enrico.zio@polimi.it](mailto:enrico.zio@polimi.it) (E. Zio).

The RUL prediction is mainly based on physics models and data analysis methods [3–5]. The physics-based method relies on the knowledge and control of the equipment physical structure, so that it is restricted to the specific properties of the single equipment. This makes it impossible to obtain a general RUL prediction model [6,7]. The data-based RUL prediction method has instead gradually becoming the core technology of equipment predictive maintenance. As a class of data-driven methods, for RUL prediction several methods based on neural networks [8–10], support vector machines [11,12], Kalman filter [13], etc. have been proposed. However, these methods require a large amount of training data, resulting in a computationally highly expensive RUL prediction [14,15].

Alternative data-driven methods, based on stochastic process, are recently proposed (see e.g. [16–18]). However, these methods based on Brownian motion (Bm) [19], inverse Gaussian process [20], gamma process [21] and other stochastic degradation models [22–24] rely on the Markov assumption. Although these stochastic models do not require a complicated training process to fit the degradation information of equipment data. A challenging problem, for these models, is that the equipment degradation is a Long-Range Dependence (LRD) process [25]. So that, the incremental independence of the Markov model is difficult to fit the LRD characteristics of the equipment degradation process. The fitting performance between the degradation model and the equipment data is important for prediction accuracy of RUL. Therefore, how to describe a degradation model with LRD characteristics is a key issue to improve the prediction accuracy of RUL.

A Long Short-Term Memory (LSTM) neural network can provide a feasible solution for the RUL prediction. Based on a special recursive structure, LSTM uses past and current output to describe the degradation process with LRD characteristics [26,27]. However, it is not easy to give accurate RUL prediction by the complex structure and training process of the LSTM neural network. In addition, the complex and variable load of wind turbines in the degradation process of the gearbox shows nonlinear and non-stationary characteristics. Therefore, a more general stochastic degradation model with LRD characteristics has been applied to RUL prediction [28,29]. For instance, the fractional Brownian motion (fBm) degradation model based on the stochastic model does not require consideration of various physical factors affecting the degradation sequence, nor does it require a complex training process [30,31,5]. It only requires the sequence generated by the stochastic model to complete the characteristic analysis and fitting of the degradation data, and then realize the equipment RUL prediction.

In this model, the equipment degradation data is considered as a nonlinear fractal time series. Local irregularity is described by the fractal index  $\alpha$ , and global correlation characteristic is described by the long correlation index  $\beta$ . However, as a widely used LRD degradation model, the fBm is unable to describe the fractal time series with strong local irregularity and strong global correlation [32–34] due to the linear relationship  $\beta = 2 - \alpha$ . In order to separately describe the local and global characteristics of fractal time series, M. Li [35,36] defines a GC process that uses the fractal dimension  $D$  and the Hurst index  $H$  to describe local irregularities and global correlations, respectively. Therefore, the RUL prediction model based on the GC process, the GC degradation model, not only solves the problem of LSTM neural network requiring a large number of training samples, but also provides a scheme for describing complex degradation data. That is, compared with the LSTM neural network, the accuracy and length of the RUL prediction obtained by the GC degradation method have better advantages in the case of the same sample. In addition, the sequence generated by the GC process depends on the model parameters of the historical degradation data. The GC degradation model only needs to calculate the model parameters once to complete the RUL prediction, which is also called multi-step prediction. However, the LSTM neural network needs to be trained and predicted at each point until the predicted value exceeds the fault threshold. Therefore, the RUL prediction process of GC degradation model has lower computational complexity than LSTM neural network.

The LRD characteristics of the GC process are generally revealed by the Autocorrelation Function (ACF) that is defined by the fractal dimension  $D$  and the Hurst index  $H$ . That is, the GC process is an LRD process if the ACF is not integrable or summable at infinity. At this time, the ranges of the  $D$  and the  $H$  are defined in  $H \in (0.5, 1)$  and  $D \in [1, 2)$ .

In order to take into account both the randomness and the LRD characteristic of the degradation process, the GC process is taken as a diffusion term to describe a degradation model. The diffusion term reveals the uncertainty of the equipment degradation process. The trend of the degradation process is mainly determined by the drift term. Therefore, the drift term of the GC degradation model is taken in the form of exponential drift and power rate drift, to overcome the nonlinearity of the degradation process. Then, a difference equation [37] and the incremental distribution are used to obtain an iterative form of the GC degradation model. Besides, the incremental distribution of the GC process is obtained from the GC sequence generated by the fractional linear system theory [38], which relies on Gaussian assumptions and probability statistics.

The non-linearity and volatility of equipment degradation data make the results of RUL prediction uncertain. Therefore, the few results predicted by the GC degradation model for equipment degradation data cannot describe the actual RUL. In order to improve the efficiency of this model, we propose to use the Monte Carlo method and the weak convergence theory [39,40], so that we can easily obtain the approximate analytical formula of the probability density function of the RUL predicted value. The uncertainty of the gearbox degradation leads to a decrease in the RUL prediction accuracy with increasing prediction time. The Lyapunov exponent will also be used to calculate the maximum prediction range of the degradation data, ensuring accuracy the RUL prediction result of the gearbox [41].

On the other hand, the degradation data are significantly affected by the external environment, so some data preprocessing is necessary. In this paper, data are pre-processed by using time domain feature factors [42,27] and Kernel Principal Component Analysis (KPCA) [43,44].

In the case study, several gearbox degradation data sets [26] are used to verify the validity of the proposed GC degradation model. The gearbox degradation sequence, processed by KPCA, is used as the input sequence for parameter estimation. The variogram estimator [45], the detrended fluctuation analysis [46] and the maximum likelihood method [47] will be con-

sidered to calculate the parameters of the GC degradation model. The Bm, the fBm degradation model and the LSTM neural network is also applied to make a comparison and to show the better efficiency of proposed method. The RUL prediction results are analyzed by the Maximum Relative Error (MRE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Health Degree (HD) and Score of Accuracy (SOA) [48,26].

This paper is organized as follows: Section 2 analyzes the characteristics of gearbox degradation data. The GC incremental modeling process is given in Section 3. Section 4 describes the GC degradation model and introduces the parameter estimation method. Section 5 uses the GC degradation model to predict the RUL of the gearbox. The conclusion is given in Section 6.

## 2. Analysis of gearbox degradation data

### 2.1. KPCA

Principal component analysis is the most commonly used feature extraction method [49]. It performs dimensionality reduction on samples to obtain fusion sequences with obvious features. However, the essence of principal component analysis is to process the data through linear mapping, and it is not possible to obtain good results when processing nonlinear degradation data. KPCA combined with kernel function and component analysis, which can not only effectively process nonlinear data, but also provide more feature information. In this paper, KPCA is used to analyze various time domain feature factor sequences  $F = [F_1, F_2, \dots, F_i]$  (as shown in Table 1 [42,27]) of gearbox degradation data. The data processing process of KPCA is given in detail in [44].

Dimensional factor is greatly affected by energy and is insensitive to the weak trends of the initial degradation stage. Traditional 1 dimensionless factor can accurately extract the weak degradation of the initial stage, but the extraction accuracy reduces gradually as the degradation continues. A new class of dimensionless factors has been proposed, which is insensitive to energy and can accurately extract the degradation of the initial stage. According to the cumulative contribution rate, KPCA selects the main components that contain degradation information above 95% as the input sequence for the next step.

### 2.2. Fractal Characteristics

The fractal characteristics of time series usually use fractal index  $\alpha$  and long correlation index  $\beta$  to describe [36]:

$$R_{xx}(0) - R_{xx}(\tau) \sim c_1 |\tau|^\alpha, \quad (\tau \rightarrow 0) \tag{1}$$

$$R_{xx}(\tau) \sim c_1 \tau^{-\beta}, \quad (\tau \rightarrow \infty) \tag{2}$$

where  $R_{xx}(\cdot)$  is the ACF of the time series,  $\tau$  is the time lag,  $c_1$  is constant. Eqs. (1) and (2) are established under the conditions of  $\tau \rightarrow 0$  and  $\tau \rightarrow \infty$ , respectively, which means that  $\alpha$  and  $\beta$  describe local volatility and overall correlation. The ranges of  $\alpha$  and  $\beta$  are defined in  $0 < \alpha \leq 2$  and  $0 < \beta < 1$ . The smaller the value, the more obvious the local fluctuation and the higher the overall correlation of the fractal time series.

### 2.3. Maximum prediction range

The non-stationary and non-linear characteristics reveal that the trend of fractal time series has obvious uncertainty, which leads to the decline of long-term prediction accuracy. In this paper, the maximum prediction range of the fractal time series is analyzed with the Lyapunov exponent in chaos theory [42,41]. The RUL of the gearbox is accurately predicted within this range. The maximum prediction range  $\eta$  is as follows:

$$\eta = \frac{1}{\kappa} \tag{3}$$

where  $\kappa$  is maximum Lyapunov exponent, which is obtained by the small data amount method [50].

## 3. Incremental modeling of GC process

In the field of fractal time series, the stationary Gaussian process in the form of ACF described by Eq. (4) is called the GC process [33].

**Table 1**  
Time domain feature factor sequence.

Symbol	$F_1$	$F_2$	$F_3$	$F_4$	$F_5$
Feature factor	Skip-over	Similarity	Margin	Waveform	Kurtosis
Symbol	$F_6$	$F_7$	$F_8$	$F_9$	
Feature factor	Skewness	Peak	Pulse	Vibration intensity	

$$R_{xx}(\tau) = \left(1 + |\tau|^{4-2D}\right)^{-\frac{1-H}{2-D}} \tag{4}$$

where  $D$  is the fractal dimension and  $H$  is the Hurst parameter,

3.1. The properties of the GC process

In the GC process,  $D$  and  $H$  can be described by  $\alpha$  and  $\beta$  in the following way:

$$D = 2 - \frac{\alpha}{2}, H = 1 - \frac{\beta}{2} \tag{5}$$

Therefore, the fractal characteristics indicate that  $D$  and  $H$  reveal the local irregularities and global LRD characteristics of the GC process when  $1 \leq D < 2, 0.5 < H < 1$ , respectively. The correlation of the GC process can be observed by the ACF curve, as shown in Fig. 1(a). The GC process can flexibly describe fractal time series due to the independent nature of parameters, such as low fractal and weak LRD, high fractal and strong LRD. The fBm with a linear relationship can only describe low fractal and strong LRD or high fractal and weak LRD sequences, as shown in Fig. 1(b).

When  $H$  takes a higher value, the ACF curve declines slowly with a significant tailing phenomenon. According to Taqqu law, time series with heavy tailing can also exhibit LRD characteristics. It is verified that the LRD characteristics of GC process become stronger with the increase of  $H$ .

In general, the LRD process is a self-similar process, but the GC process only meets a weak self-similar property, also known as local self-similar. The global self-similarity of the GC process is obtained by Lamperti transformation. The ACF of the GC process after the Lamperti transformation has the following form (see Fig. 2):

$$R_{yy}(\tau) = \left\{ 1 + \left[ \ln \left( 1 + \frac{\tau}{t} \right) \right]^{4-2D} \right\}^{2H-2} \tag{6}$$

3.2. Modeling of GC incremental distribution

The GC process relies on the Gaussian assumption. Its increment follows a Gaussian distribution and can be obtained using probability statistics.

Ortigueira fractal linear system theory [38] shows that convolution through white noise and filters can produce a stationary time series:

$$x(t) = w(t) * h(t) = \int_0^t h(t - \tau)w(\tau)d\tau \tag{7}$$

where  $w(t)$  is white noise,  $h(t)$  is the impulse function, and  $x(t)$  is the generated stationary time series. In this article, a non-stationary Gaussian white noise and the impulse function generated by the ACF of the GC process are convolved to obtain the GC sequence. The specific form is as follows:

$$X(t) = w(t) * F^{-1} \left[ F \left( \left( 1 + |\tau|^{4-2D} \right)^{-\frac{1-H}{2-D}} \right)^{0.5} \right] \tag{8}$$

where  $F(\cdot)$  and  $F^{-1}(\cdot)$  are Fourier transform and inverse transform, respectively. The GC sequence generated by different parameters is shown in Fig. 3.

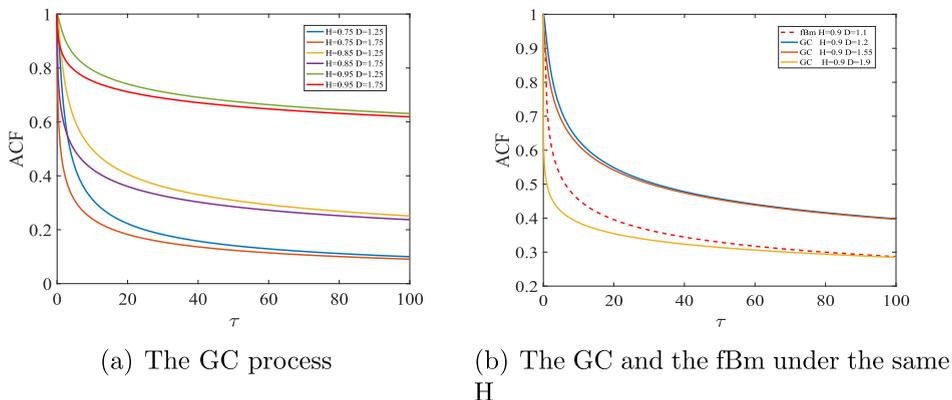


Fig. 1. ACF curve.

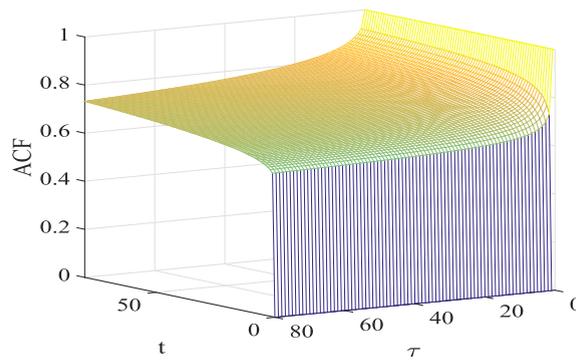


Fig. 2. Three-dimensional ACF curve of Lamperti transformation.

Make multiple differences for the generated GC sequence, and obtain the incremental distribution of the GC process by probability statistics:

$$\Delta GC(t) = [GC(t + \tau) - GC(t)] \sim N(0, \delta_\tau) \tag{9}$$

Fig. 4 shows the specific modeling process of GC incremental distribution.

It should be noted that the larger the simulation numbers, the closer the incremental fitting of the GC sequence is to the true distribution. However, if the simulation numbers are too large, the real-time applicability may become worse and the fitting value may be too high. If the real-time applicability of high effect is too pursued, the GC incremental distribution may have a low fitting value. This article discusses the optimal number of simulation through the analysis of the incremental distribution generation process. It ensures that the real-time applicability and the fitting effect of the GC incremental distribution are considered simultaneously. Table 2 reveals the time required for different simulation numbers. Fig. 5 shows the distribution of GC increments for different simulation numbers.

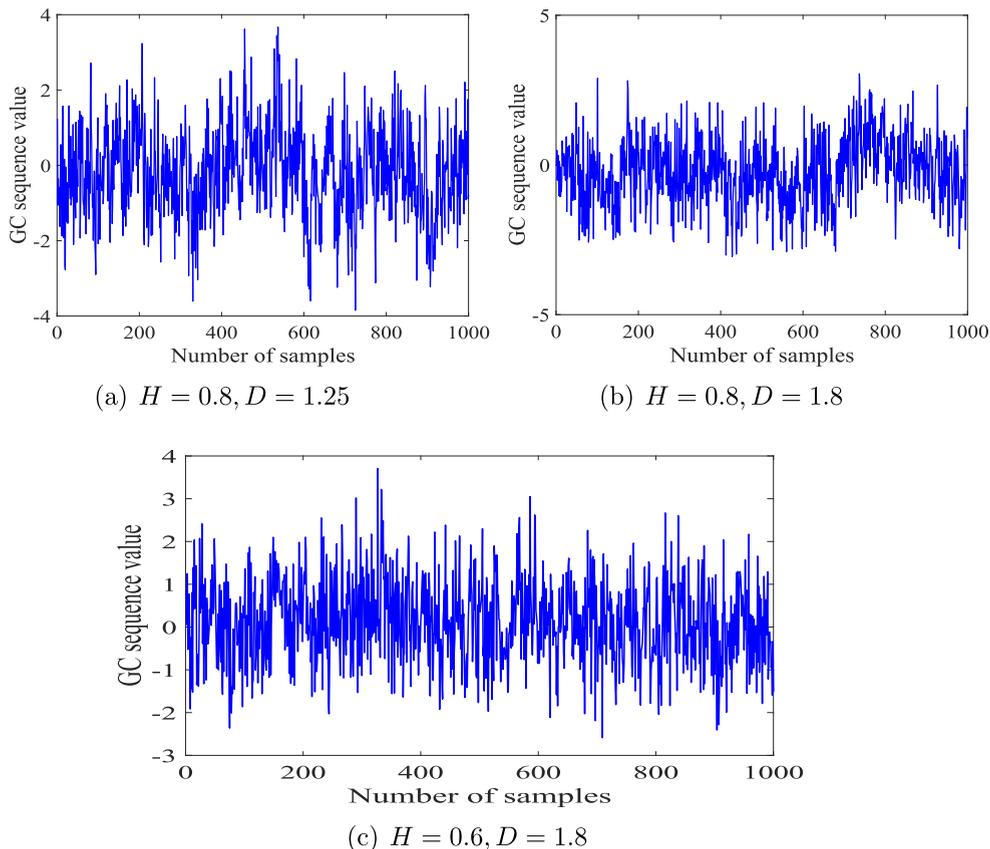


Fig. 3. GC numerical sequence.

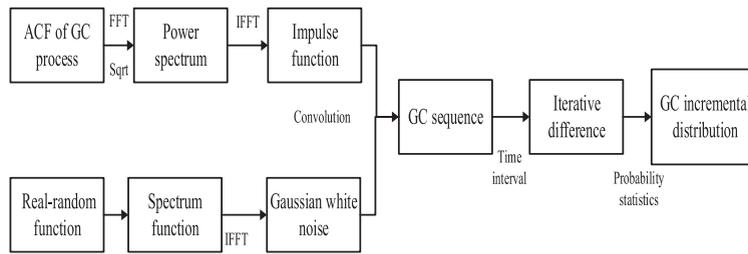


Fig. 4. GC incremental distribution modeling process.

**Table 2**  
Time required for each part of the incremental distribution generation process.

Simulation number	Hurst index	Fractal dimension	GC sequence	Differential process	Statistical process
30,000			4.813s	0.442s	0.101s
20,000			1.383s	0.434s	0.096s
10,000	0.045s	0.019s	0.382s	0.412s	0.09s
5000			0.107s	0.387s	0.086s

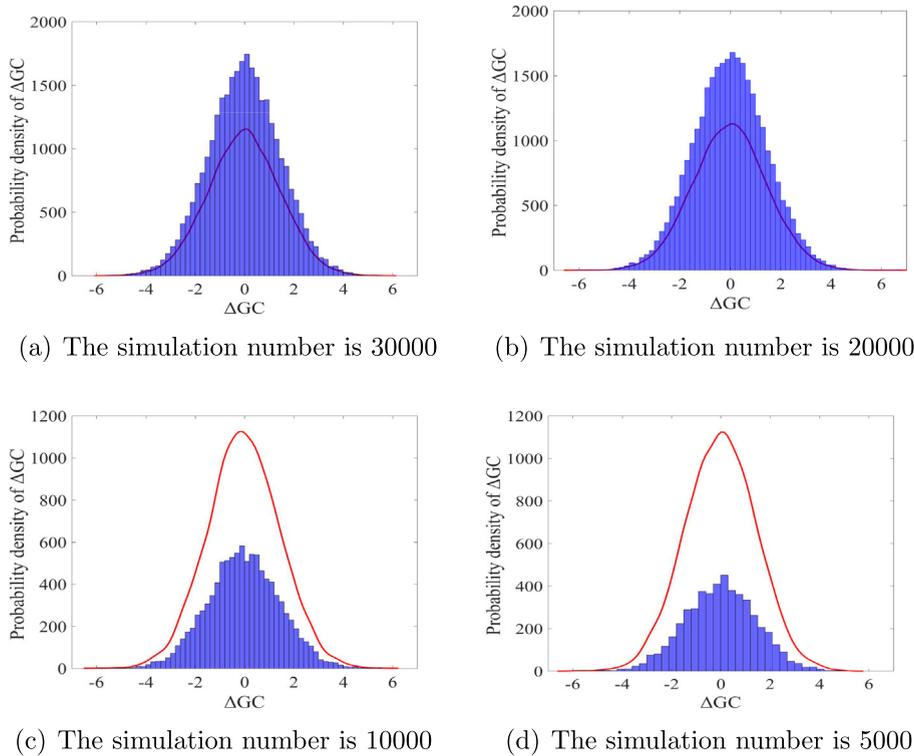


Fig. 5. The incremental distribution of the GC sequence. The red curve represents the assumed Gaussian distribution. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The results in Table 2 and Fig. 5 reveal a problem: when the simulation numbers are too small, although the time is shorter, the GC increment distribution obtained by statistics is lower than the actual distribution; when the simulation numbers are too large, the GC incremental distribution takes a longer time and is greater than the actual distribution. Therefore, an appropriate simulation numbers can not only obtain a GC incremental distribution that is close to the true distribution, but also ensure better real-time applicability. Through experimental analysis, this paper generates a GC sequence of 20000 simulations to get the incremental distribution.

#### 4. Degradation modeling and RUL prediction based on GC process

##### 4.1. GC degradation model

The degradation state of the gearbox at time  $t$  can be characterized by [31]:

$$X(t) = \mu \int_0^t \varphi(t; \theta) dt + \delta B_H(t) \tag{10}$$

where  $\mu \int_0^t \varphi(t; \theta) dt$  represents the nonlinear drift term of the overall degradation trend of the degradation process,  $\mu$  is the drift coefficient,  $\theta$  is a vector with unknown parameters,  $\delta$  is a constant diffusion coefficient and  $B_H(t)$  is the fBm. The generalized form of the degradation model is obtained using Ito's theorem [51,52], that is,  $\mu$  and  $\delta$  are replaced by the time functions  $\phi(x)$  and  $\omega(x)$ , and the interference term  $B_H(t)$  is represented by  $\psi(x)$ . Eq. (11) can be rewritten as:

$$X(t) = \phi(t) \int_0^t \varphi(t; \theta) dt + \omega(t)\psi(t) \tag{11}$$

The fBm degradation model uses fBm as an interference term [31]. Based on the principle, this article considers the GC process as a random interference term with LRD properties. The general expression of the GC degradation model is as follows:

$$X(t) = \phi(t) \int_0^t \varphi(t; \theta) dt + \omega(t)GC(t) \tag{12}$$

Simplify the GC degradation model and obtain the following form:

$$X(t) = \mu\varphi(t) + \delta GC(t) \tag{13}$$

where  $\mu$  and  $\delta$  are the drift and diffusion parameters of the GC degradation model.  $\varphi(t)$  is a time-dependent function used to describe the overall trend of the degradation process equivalent to  $\int_0^t \varphi(t; \theta)$ . It can be described by linear function, power rate function and exponential drift function [53]. The GC degradation model can be rewritten into the following three forms:

$$M0 : X(t) = X(0) + \mu t + \delta GC(t), \tag{14}$$

$$M1 : X(t) = X(0) + \mu t^b + \delta GC(t), \tag{15}$$

$$M2 : X(t) = X(0) + \mu e^{bt} + \delta GC(t), \tag{16}$$

where  $M0, M1, M2$  represents linear, power law and exponential drifts, respectively, and  $b$  represents the variability of the nonlinear degradation process. Since the degradation process of gearbox is a nonlinear time series, the linear function cannot be described well. Fig. 6 shows the simulation paths for three forms of drift. In this paper, the power rate and exponential function are used to describe the nonlinear drift of the GC degradation model. The nonlinearity of the exponential term and power law drift makes it difficult to obtain the analytical form of  $b$ . This paper uses the Nelder-Mead simplex algorithm and fminsearch function to solve the estimated value of  $b$ .

##### 4.2. Parameter estimation of GC degradation model

###### 4.2.1. Parameter estimation of Hurst index and fractal dimension

There are many mature techniques for estimating  $D$  and  $H$ . For example, the box dimension method, Root mean square method and Power spectral density method are used to calculate the estimated value of  $D$ ; The rescaled range method, wavelet analysis method and Hilbert-Huan method are used to calculate the estimated value of  $H$ . However, there are always some deviations in the above methods. This article will use variogram estimator and detrended fluctuation analysis methods to estimate  $D$  and  $H$ , respectively. The specific estimation process is described in [45,46].

###### 4.2.2. Parameter estimation of drift parameter and diffusion parameter

For the parameter estimation of  $\mu$  and  $\delta$ , this paper uses the maximum likelihood estimation method. In the estimation process, the ACF of the GC process is used to calculate the value of  $\mu$  and  $\delta$ , whose specific process is as follows.

Step 1: The observation data from  $t_0$  to  $t_k$  is used as a sample input sequence and written as a vector form  $\mathbf{X} = [X_0, X_1, \dots, X_k]^T$ , and its increment  $\mathbf{x} = [X_1 - X_0, X_2 - X_1, \dots, X_k - X_{k-1}]^T$  follows a multi-dimensional normal distribution  $\mathbf{x} \sim N(\mu\varphi, \delta^2\mathbf{Q})$  according to the Gaussian assumption of the GC process, where  $\varphi = [\varphi(t_1) - \varphi(t_0), \varphi(t_2) - \varphi(t_1), \dots, \varphi(t_k) - \varphi(t_{k-1})]^T$  is the mean vector, and  $\sigma^2\mathbf{Q}$  is a covariance matrix,  $Q_{ij} = \left(1 + (|i - j|\tau)^{4-2D}\right)^{-\frac{1-H}{2-D}}$  represents the position of each element in  $\mathbf{Q}$ .

Step 2: Calculate the logarithmic form of the likelihood function  $f(\mathbf{x}; \mu, \delta)$  according to the principle of maximum likelihood method  $(\hat{\mu}, \hat{\delta}) = \arg \max_{\mu, \delta} f(\mathbf{x}; \mu, \delta)$ .

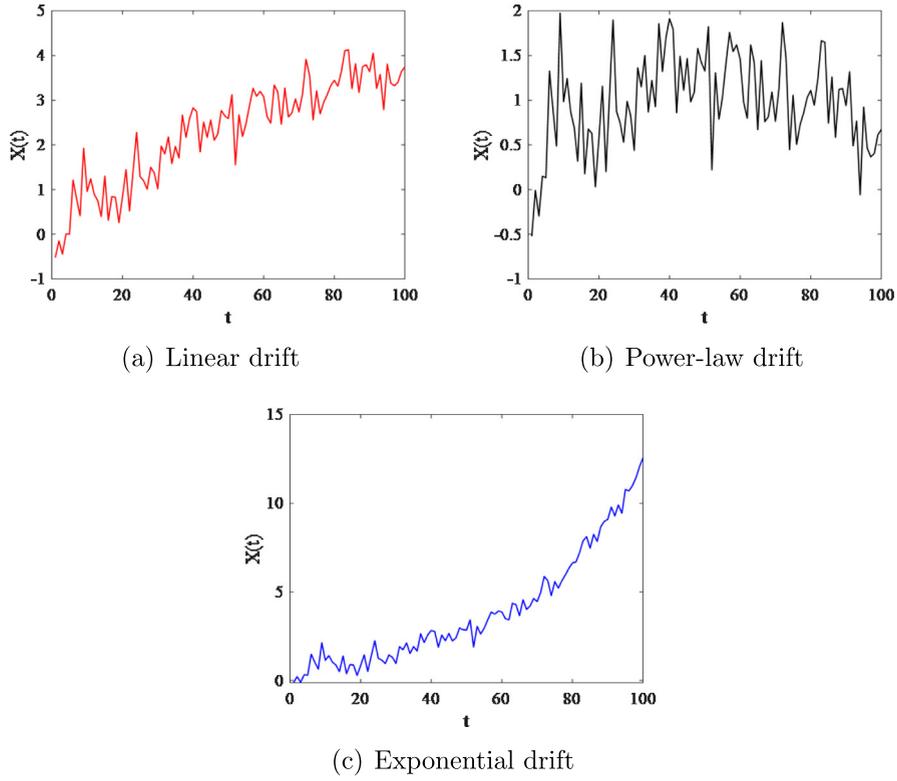


Fig. 6. Path simulation with different drift.

$$l(\mathbf{x}; \mu, \delta) = -\frac{k}{2} \ln(2\pi) - k \ln \delta - \frac{1}{2} \ln |\mathbf{Q}| - \frac{1}{2\delta^2} (\mathbf{x} - \mu\varphi)^T \mathbf{Q}^{-1} (\mathbf{x} - \mu\varphi) \quad (17)$$

Step 3: The logarithmic form of the likelihood function is:

$$l(\mathbf{x}; \mu, \delta) = -\frac{k}{2} \ln(2\pi) - k \ln \delta - \frac{1}{2} \ln |\mathbf{Q}| - \frac{1}{2\delta^2} (\mathbf{x} - \mu\varphi)^T \mathbf{Q}^{-1} (\mathbf{x} - \mu\varphi) \quad (18)$$

Step 4: Calculate Eq. (20) for partial derivatives of  $\mu$  and  $\delta$ , and set the partial derivative value to 0,

$$\begin{aligned} \frac{\partial l}{\partial \mu} &= \frac{\partial \left( -\frac{k}{2} \ln(2\pi\delta^2) - \frac{1}{2} \ln |\mathbf{Q}| - \frac{1}{2\delta^2} (\mathbf{x} - \mu\varphi)^T \mathbf{Q}^{-1} (\mathbf{x} - \mu\varphi) \right)}{\partial \mu} \\ &= \frac{\partial \left( -\frac{1}{2\delta^2} (\mathbf{x} - \mu\varphi)^T \mathbf{Q}^{-1} (\mathbf{x} - \mu\varphi) \right)}{\partial \mu} \\ &= \frac{\partial \left( -\frac{1}{2\delta^2} (\mathbf{x}^T \mathbf{Q}^{-1} \mathbf{x} - \mu\varphi^T \mathbf{Q}^{-1} \mathbf{x} - \mathbf{x}^T \mathbf{Q}^{-1} \mu\varphi + \mu^2 \varphi^T \mathbf{Q}^{-1} \varphi) \right)}{\partial \mu} \\ &= -\frac{(-\varphi^T \mathbf{Q}^{-1} \mathbf{x} - \mathbf{x}^T \mathbf{Q}^{-1} \varphi + 2\mu\varphi^T \mathbf{Q}^{-1} \varphi)}{2\delta^2} = 0 \end{aligned} \quad (19)$$

$$\begin{aligned} \frac{\partial l}{\partial \delta^2} &= \frac{\partial \left( -\frac{k}{2} \ln(2\pi\delta^2) - \frac{1}{2} \ln |\mathbf{Q}| - \frac{1}{2\delta^2} (\mathbf{x} - \mu\varphi)^T \mathbf{Q}^{-1} (\mathbf{x} - \mu\varphi) \right)}{\partial \delta^2} \\ &= -\frac{k}{2\delta^2} + \frac{(\mathbf{x} - \mu\varphi)^T \mathbf{Q}^{-1} (\mathbf{x} - \mu\varphi)}{2(\delta^2)^2} = 0 \end{aligned} \quad (20)$$

Step 5: Solve Eqs. (21) and (22) obtain the maximum likelihood estimates of  $\mu$  and  $\delta$ :

$$\hat{\mu} = \frac{\varphi^T \mathbf{Q}^{-1} \mathbf{x}}{\varphi^T \mathbf{Q}^{-1} \varphi} \quad (21)$$

$$\hat{\delta} = \sqrt{\hat{\sigma}^2} = \left( \frac{1}{k} \frac{(\mathbf{x}^T \mathbf{Q}^{-1} \mathbf{x}) (\varphi^T \mathbf{Q}^{-1} \varphi) - (\varphi^T \mathbf{Q}^{-1} \mathbf{x})^2}{\varphi^T \mathbf{Q}^{-1} \varphi} \right)^{\frac{1}{2}} \tag{22}$$

The randomness and non-linearity of fractal time series make small changes in parameter estimates have a significant impact on the prediction convergence, as shown in Fig. 7. Therefore, it is necessary to discuss the convergence of parameter estimators. It can better guarantee the accuracy of parameter estimates and the prediction convergence of the GC degradation model.

Substitute  $x = \mu\varphi + \delta\Delta GC(t)$  into  $\hat{\mu} = \frac{\varphi^T \mathbf{Q}^{-1} x}{\varphi^T \mathbf{Q}^{-1} \varphi}$

$$\hat{\mu} = \mu + \delta \frac{\varphi^T \mathbf{Q}^{-1} \Delta GC(t)}{\varphi^T \mathbf{Q}^{-1} \varphi} \tag{23}$$

$E[\hat{\mu}] = \mu$  means that the estimator  $\hat{\mu}$  of the drift parameter is an unbiased estimator. Calculate the variance of Eq. (25):

$$\text{Var}[\hat{\mu}] = \delta^2 \frac{\varphi^T \mathbf{Q}^{-1} \mathbf{Q} \mathbf{Q}^{-1} x}{(\varphi^T \mathbf{Q}^{-1} \varphi)^2} = \frac{\delta^2}{\varphi^T \mathbf{Q}^{-1} \varphi} \tag{24}$$

$\text{Var}[\hat{\mu}] \leq Ck^{2H-2}$  can be obtained by Gerschgorin Circle theorem, where C is a constant. Because the range of H is limited to (0.5, 1),  $\text{Var}[\hat{\mu}] \rightarrow 0$  when  $k \rightarrow \infty$ . Therefore, the estimator  $\hat{\mu}$  of the drift parameter is convergent. Performing the same calculation process on the estimator  $\hat{\delta}$  of the diffusion parameter can verify that it is also convergent and unbiased.

### 4.3. RUL prediction

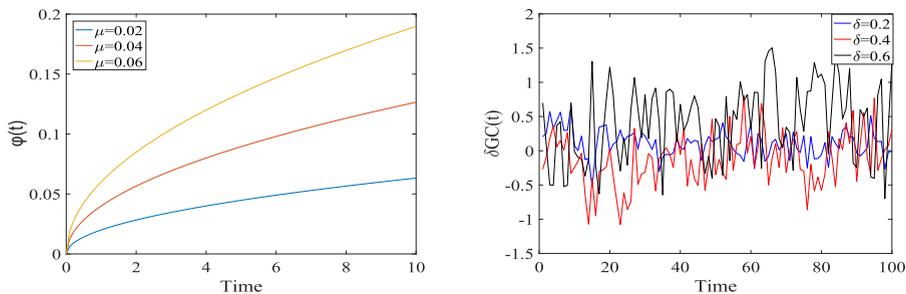
In order to achieve predictive maintenance of the gearbox, this paper sets an initial failure threshold. When the degradation state exceeds this threshold, it starts to predict the RUL of the gearbox. Fig. 8(a) and (b) define the RUL prediction process in which the degradation process shows an upward and downward trend, respectively.

Therefore, the next degradation state at time  $t_k$  can be expressed in the iterative difference form of the GC degradation model, which is defined as follows:

$$\begin{aligned} X(t_k + l_k) &= X(t_k) + \mu[\varphi(t_k + l_k) - \varphi(t_k)] \\ &\quad + \delta[GC(t_k + l_k) - GC(t_k)] \end{aligned} \tag{25}$$

where  $l_k$  is the time increment, and  $GC(t_k + l_k) - GC(t_k) = \Delta GC(t) \sim N(0, \delta_\tau)$ .

The time series generated by the GC process is non-stationary and nonlinear. The generation process is random. Thus, the few results predicted by the GC degradation model for gearbox degradation data cannot describe the actual RUL. The non-linear drift term of the GC degradation model makes it difficult to derive the probability density distribution of the predicted value of RUL. Based on the Gaussian assumption and weak convergence assumption of the GC process [39,40], the approximate analytical formula of the probability density function can be derived. Then, the Monte Carlo method is used to generate a large number of RUL prediction results. The maximum value of the probability density distribution of RUL prediction results represents the predicted value of RUL (see Fig. 9). Fig. 10 shows the whole process of GC degradation model predicting RUL.



(a) The influence of drift parameter on the degradation trend (b) The influence of diffusion parameter on the volatility of degradation

Fig. 7. The influence of drift parameters and diffusion parameters on the degradation process.

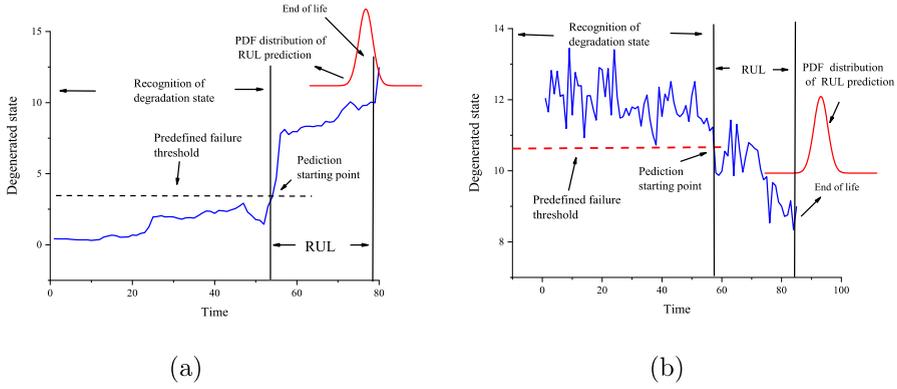


Fig. 8. Definition of RUL prediction.

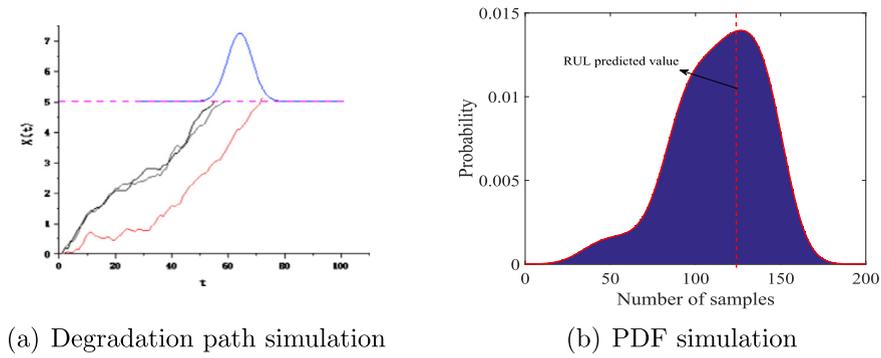


Fig. 9. RUL prediction principle of GC degradation model.

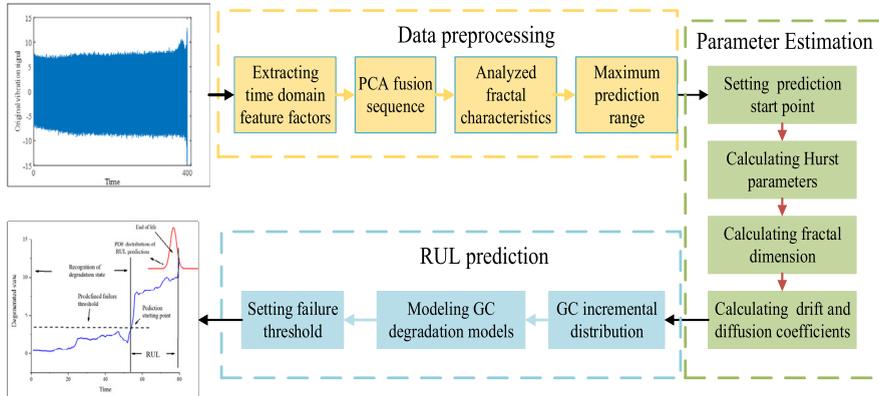


Fig. 10. Incremental distribution of GC sequences.

$$f_{L_k|M}(l_k) \cong \frac{[\delta(l_k)]^{\frac{1}{2}}}{\sqrt{2\pi}[\delta(0)]^{\frac{1}{2}} \int_0^{l_k} [\delta(\tau)]^{\frac{1}{2}} d\tau} \times \left\{ \frac{w - X(t_k) - \mu\varphi(l_k + t_k) + \mu\varphi(t_k)}{\int_0^{l_k} [\delta(\tau)]^{\frac{1}{2}} d\tau} + \frac{\mu d\varphi(l_k + t_k)}{[\delta(l_k)]^{\frac{1}{2}}} \right\} \times \exp \left[ -\frac{[w - X(t_k) - \mu\varphi(l_k + t_k) + \mu\varphi(t_k)]^2}{2[\delta(0)]^{\frac{1}{2}} \int_0^{l_k} [\delta(\tau)]^{\frac{1}{2}} d\tau} \right] \quad (26)$$

## 5. Case study

### 5.1. Case 1

To verify the validity of the GC degradation model, the gearbox degradation data sets collected from the gear contact fatigue test rig are used for simulation. The test bench is the FZG test bench of the company STRAMA, which consists of a torque controller, a cooling lubrication controller, a gear test box and an operating system. The structure of the test rig and gearbox is shown in Fig. 11. In the gearbox degradation experiment, the oil flow is 4L/h; the cooling temperature is 70°C; the material of the gear is 40 Cr and the vibration signal of the gearbox degradation is collected by the accelerometer fixed on the gearbox. The test stops when the maximum amplitude of the collected vibration signal exceeds the threshold. In order to verify the generality of the GC degradation model, this case study uses the GC degradation model for two sets of gearbox degradation data in different environments. The specific parameters are shown in Table 3.

#### 5.1.1. Preprocessing of gearbox degradation data

This case uses the time domain feature factors  $F_i$  and KPCA to analyze the original vibration signal shown in Fig. 12. The gearbox degradation trend after the extraction of  $F_i$  is shown in Fig. 13. Using KPCA for further analysis, the main components with a cumulative contribution rate of more than 95% are selected. The process is shown in Fig. 14.

Fig. 14 means that 95% of the information on the degradation of the two sets of gearboxes is included in the first 5 or 4 main components. Therefore, the KPCA sequence that fuse the first 5 or 4 main components is obtained, as shown in Fig. 15. It can be seen that the degradation sequence has obvious inflection at points 554 and 375. At this instance, the working phase of the gearbox is divided into normal operation and slow degradation phase.

Eq. (5) indicates that the parameters  $\alpha$  and  $\beta$  describing fractal characteristics can be obtained from the estimated values of  $D$  and  $H$ . The fractal characteristic parameters of the KPCA sequence are shown in Table 4, and the maximum prediction range is also given in Table 4.

#### 5.2. Validity analysis of GC degradation model

In this case, we use the GC degradation model (M1 and M2) to predict the RUL of Gearbox1. The degradation state of the gearbox corresponding to point 554 in the KPCA sequence is preset as the first failure threshold. The actual RUL of the gearbox has a length from the predicted start point 544 and predicted end point 600, indicating that the predicted length is within the maximum prediction range. Therefore, the GC degradation model with LRD characteristics can make the RUL prediction results more accurate.

In the previous RUL prediction process, the RUL of all points is predicted by using a set of parameters of prediction starting points. The nonlinear disturbance and drift term of the degradation model lead to small changes in parameters that have a significant impact on the prediction accuracy. In this case, the metabolic method is used to iterate the input sequence to calculate the model parameters of each set of prediction starting points. After obtaining the model parameters of the initial prediction starting point, the data before the prediction sample is deleted and the latest data is added until the next prediction starting point is included. The prediction samples are kept in equal dimensions to calculate the model parameters of the next prediction starting point, as shown in Fig. 16. From the perspective of the length of the paper, this case only gives some parameters of a degradation model, as shown in Table 5. Fig. 17 shows the RUL prediction results of Gearbox1.

The SOA, HD, RMSE and MAPE are introduced to analyze the predictive performance of the M1 and M2 degradation model in Table 6. The box plot is also used in this case to give the relative error between the actual and predicted RUL, as shown in Fig. 18, which gives the prediction error of the two models in more detail.

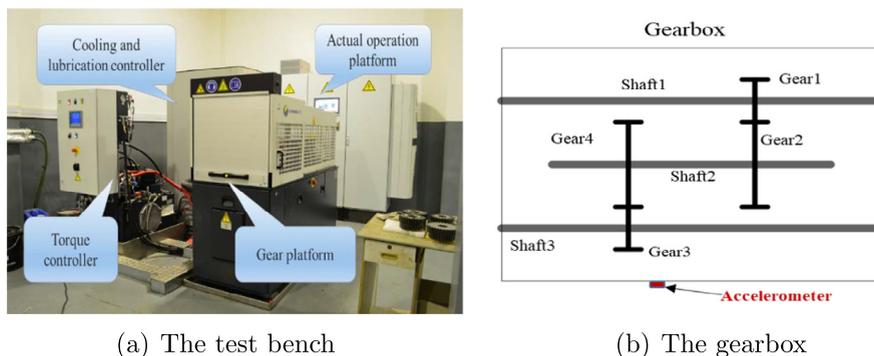
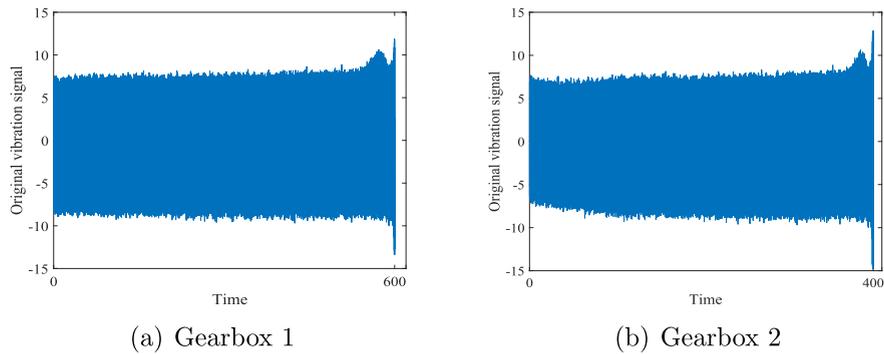


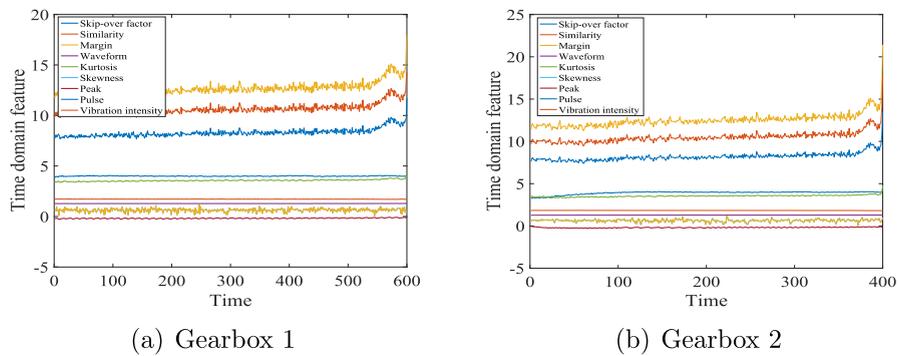
Fig. 11. Structure of test bench and gearbox.

**Table 3**  
Experimental parameters of gearbox degradation.

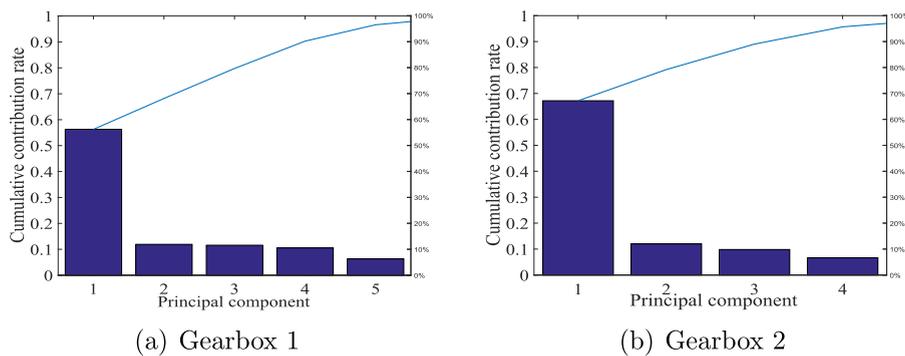
Data sets	Lifetime	Sampling	Operating
Gearbox 1	600 min	Time:10s Period:20s	Speed: 500r/min Torque:1400 N*m
Gearbox 2	400 min	Time:10s Period:50s	Frequency:50 kHz



**Fig. 12.** Original vibration signal of gearbox degradation.



**Fig. 13.** Degeneration trend of time domain feature factors.



**Fig. 14.** Cumulative contribution rate of principal components.

The analysis of the error results in Table 6 shows that the GC degradation model can effectively predict the RUL of gearbox. The values of RMSE and MAPE are inversely proportional to the prediction performance of the GC degradation model; HD and SOA values are directly proportional to the fitting performance of the model. As a result, we can deduce that the prediction performance of the M1 model is better than M2, which verifies that the GC process obey the heavy-tailed character-

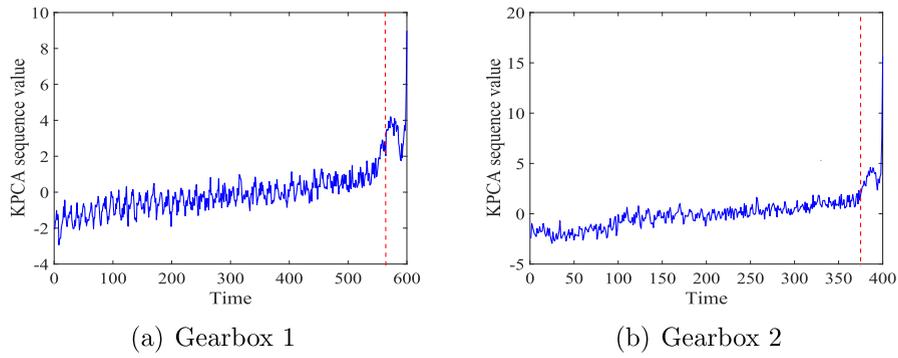


Fig. 15. KPCA sequence.

Table 4  
Fractal characteristics of KPCA sequence.

	$\alpha$	$\beta$	$\kappa$	$\eta$
Gearbox 1	1.2176	0.6692	0.0125	80
Gearbox 2	1.1226	0.505	0.0145	68

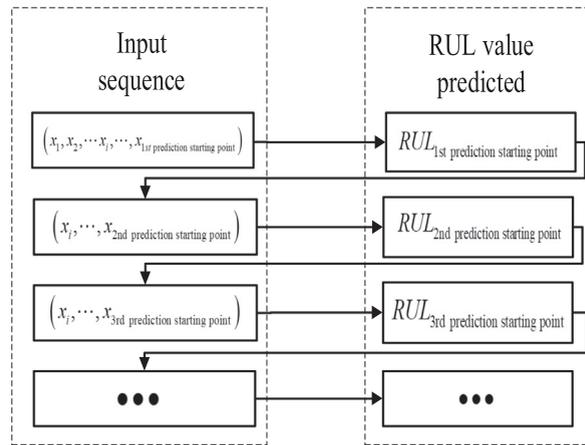
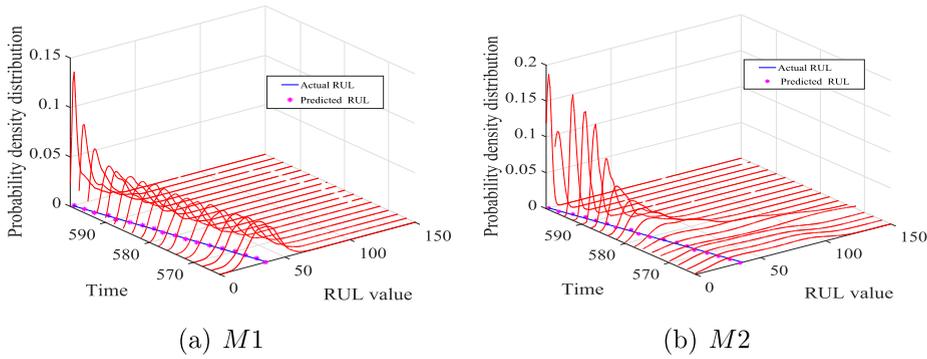


Fig. 16. Principles of metabolism method.

Table 5  
Parameter estimation of GC degradation models (M1 and M2).

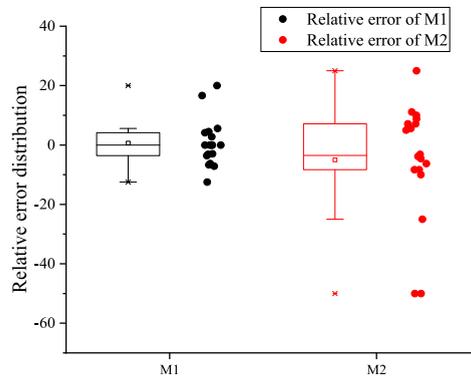
Model	Predict starting point	$H$	$D$	$\mu$	$\delta$	$b$
M1	564	0.638	1.221	0.0356	0.477	0.919
	568	0.636	1.222	5.95E-03	0.503	1.369
	572	0.6623	1.192	6.25E-03	0.5	1.397
	576	0.659	1.188	0.003	0.491	1.067
	580	0.719	1.17	0.006	0.555	0.924
	584	0.708	1.157	0.016	0.521	0.702
	588	0.708	1.172	0.579	0.5755	0.396
	592	0.726	1.181	1.172	0.641	0.216
M2	596	0.716	1.174	1.351	0.620	0.198
	564	0.638	1.221	0.514	0.467	0.016



**Fig. 17.** Comparison of RUL prediction results of GC degradation models (M1 and M2). The red curve shows the probability density distribution of RUL prediction value. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Table 6**  
Prediction results analysis of M1 and M2.

	SOA	HD	RMSE	MAPE
M1	0.9513	0.9902	1.0274	0.0532
M2	0.9129	0.9206	1.7795	0.1385



**Fig. 18.** Relative error distribution between predicted and actual RUL.

istic of decay with power rate. Fig. 18 shows that the relative error distribution of MI is relatively concentrated. Besides, the GC process that obeys the power law decay has stable prediction performance.

5.3. General analysis of GC degradation model

The case 2 uses the Gearbox2 under different sampling conditions to verify the generality of the proposed model. The fBm and the Bm model are classic LRD and incremental independent stochastic degradation models respectively. The LSTM is a classic neural network algorithm. They are introduced as comparison model to reveal the advantages of the GC degradation model. The modeling process is shown in [31,54,55]. In this case, the degradation state corresponding to point 375 is set as the first failure threshold. The sample sequence length is  $n = 100$ . The stochastic model parameters of Gearbox2 are shown in Table 7. Because of the uncertainty of prediction, Fig. 19 shows the probability density distribution of the GC, the fBm, the LSTM and Bm prediction RUL.

**Table 7**  
Parameter estimation of GC degradation models (M1 and M2).

Model	Predict starting point	$H$	$D$	$\mu$	$\delta$	$b$
GC		0.5313	1.3048	0.5286	0.5446	0.1991
fBm	375	0.5313	-	0.5715	0.1326	0.1762
Bm		-	-	0.4228	0.2839	0.3069

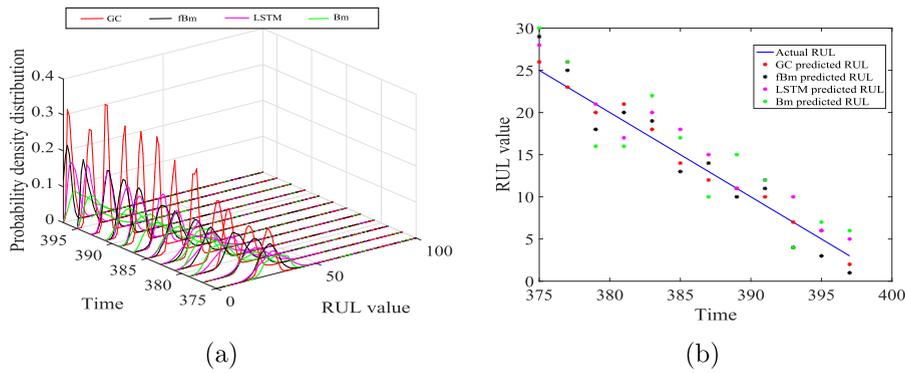


Fig. 19. RUL prediction comparison.

**Table 8**  
Prediction results analysis of M1 and M2.

	SOA	HD	RMSE	MAPE
GC	0.9640	0.9790	1.0000	0.0866
fBm	0.8351	0.8934	2.2546	0.2149
LSTM	0.8311	0.8829	2.3629	0.2095
Bm	0.6761	0.7325	3.5707	0.3258

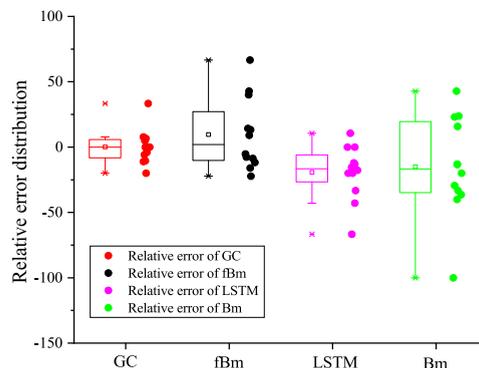


Fig. 20. Relative error between predicted and actual RUL.

The prediction results of Gearbox2 data are still analyzed using several evaluation indicators mentioned in case 1, as shown in Table 8 and Fig. 20. The analysis results prove the generality of the GC degradation model, which can be widely applied to the RUL prediction of the gearbox. Also, it shows that the prediction performance of the GC degradation model is better than the fBm, the LSTM and the Bm.

5.4. Case 2

In order to objectively prove the generality of the GC degradation model, the FEMTO-ST organization publishes the full life data of rolling bearing in the PHM 2012 Data Challenge. The data are obtained by experiments on the PRONOSTIA test bench, which are applied to this case. The PRONOSTIA test bench mainly has a rotating part (gearbox and two motors), a degradation generating part (mainly the generation of radial force) and a measuring part (vibration sensor, etc.). Fig. 21 shows the specific structure of the platform. During the operation of the PRONOSTIA platform, the vibration sensor composed of two accelerometers is sampled at a frequency of 25.6 kHz, and the vibration signal is recorded every 0.1 min with a sampling time of 0.1 s. The FEMTO-ST organization provides three sets of run-to-failure data for different load environments. In this case, 2463 samples collected under the working conditions of 1800 rpm and 4000 N are selected as the verification sequence of the GC degradation model.

Similar to the operating environment of gear, the bearing degradation process is also disturbed by factors such as noise and has a complex trend, as shown in Fig. 22. Following the data preprocessing process of Case 1, this case uses the time domain feature factors  $F_i$  to extract bearing degradation trends. Some characteristic factors shown in Fig. 23 have similar

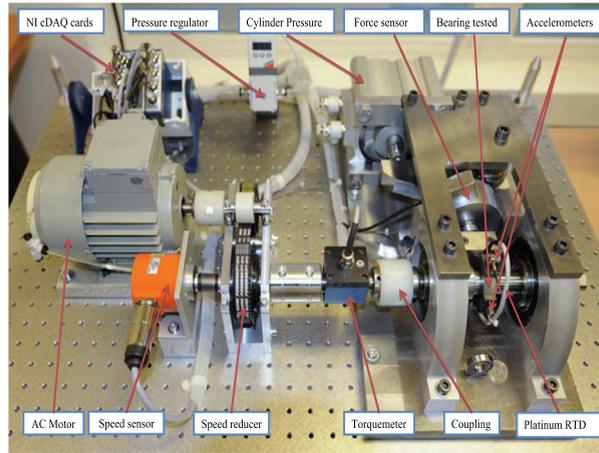


Fig. 21. The structure of the PRONOSTIA platform.

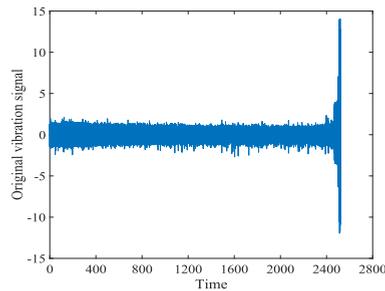


Fig. 22. Original vibration signal of bearing degradation.

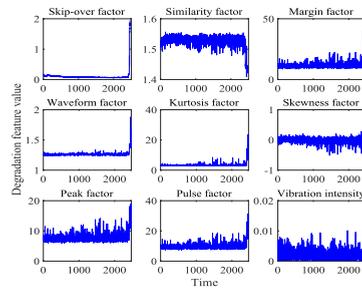


Fig. 23. Bearing degradation trend after feature factor extraction.

degradation trend. Therefore, this case uses the KPCA method to analyze the 9 degradation trends and eliminate redundant indicators to obtain a degradation sequence that integrates different feature trends. In this case, four feature factors with a cumulative contribution of more than 95% are selected as the main components of the KPCA method, as shown in Fig. 24.

The variogram estimator and the detrended fluctuation analysis methods are used to calculate the parameters D and H of the degradation sequence in Fig. 24(b). The fractal dimension and long correlation index obtained by D and H are shown in Table 9.

In this case, the end of bearing life threshold is set to 16. From the bearing degradation sequence given in Fig. 24, it can be found that the bearing has no obvious degradation in the first half of the operation process. The fluctuation of the characteristic value belongs to normal bearing wear. Therefore, this case chooses 2400 points as the starting point to predict the RUL. The degradation characteristics at this time have changed significantly, and the actual RUL is also within the maximum prediction range. At this instance, the RUL value predicted by the GC degradation model can better reflect the health of the bearing.

The fBm, the LSTM and the Bm are also introduced as comparison models in this case to verify the effectiveness of the GC model. The parameter estimates of the stochastic degradation model are shown in Table 10. The probability distribution of RUL obtained by the four different prediction methods is shown in Fig. 25(a), and the predicted value of RUL is shown in Fig. 25(b). In order to objectively compare the prediction performance of the four methods, SOA, HD, RMSE, MAPE and relative error are used to analyze the error between the predicted value and the actual RUL value, as shown in Table 11 and Fig. 26.

From the parameter value given in Table 10, the Hurst index  $H$  and the fractal dimension  $D$  of the bearing degradation sequence are approximately linear relationships  $H = 2 - D$ . Therefore, the error analysis of the GC degradation model is not much better than the fBm model. In Table 6 of Case 1, the relationship between the Hurst index  $H$  and the fractal dimension  $D$  of the gearbox degradation sequence tends to be independent. As a result, the prediction of the GC degradation model is significantly better than the fBm model. These two cases verify that the GC degradation model has better predictive performance in a variety of degradation sequences. The principle of LSTM neural network requires a lot of training data to get

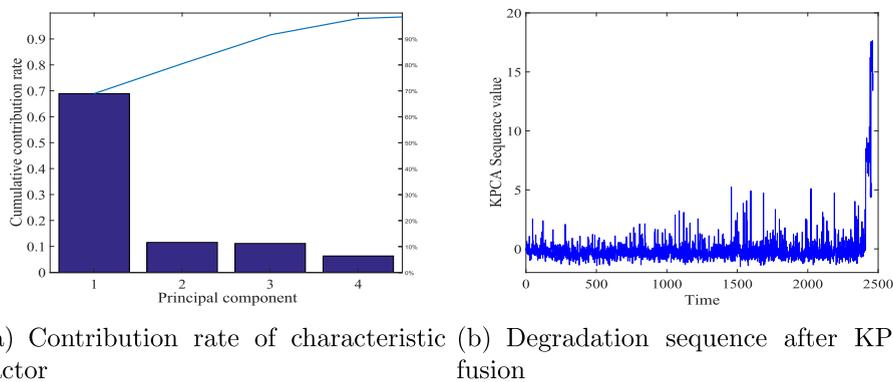


Fig. 24. KPCA analysis process.

Table 9  
Fractal characteristics of bearing degradation sequence.

$\alpha$	$\beta$	$\kappa$	$\eta$
1.2762	0.733	0.0044	227

Table 10  
Prediction performance of different methods.

Model	$H$	$D$	$\mu$	$\delta$	$b$
GC	0.6302	1.3120	$6.66e-10$	1.5504	4.4215
fBm	0.6302	-	$3.10e-10$	0.3645	4.5699
Bm	-	-	$7.16e-06$	0.7725	0.9307

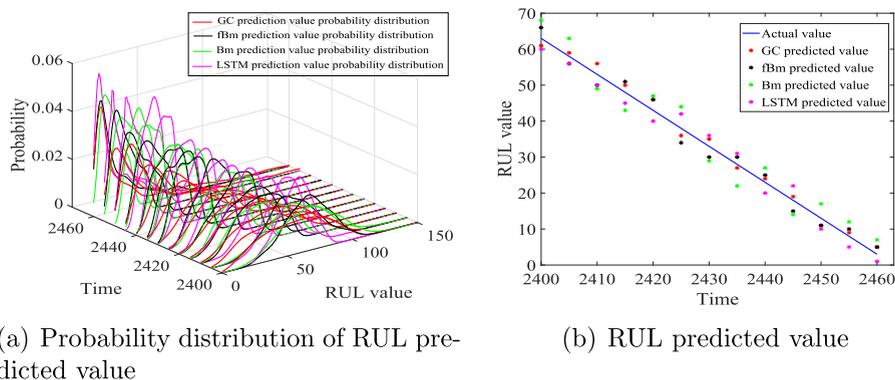
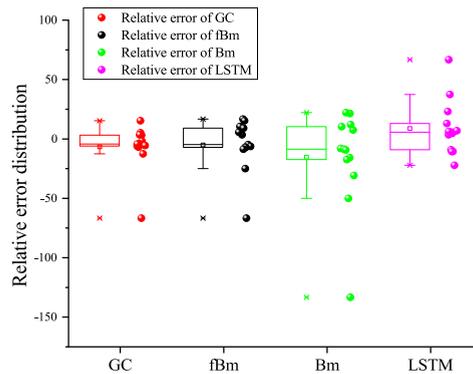


Fig. 25. Predicted results of bearing RUL.

**Table 11**  
Prediction results analysis of M1 and M2.

	SOA	HD	RMSE	MAPE
GC	0.5387	0.9897	1.5014	0.1085
fBm	0.4847	0.9793	2.1389	0.1433
Bm	0.3223	0.9396	4.5993	0.2668
LSTM	0.4918	0.9734	3.0509	0.1692



**Fig. 26.** Relative error between predicted value and actual RUL.

accurate prediction results. However, the small amount of training data given in this case is far from reaching the training accuracy required by the neural network. Thus, the GC degradation model can obtain higher prediction accuracy in the case of a small amount of training data. The Bm model cannot describe the time dependence of the degradation sequence, its prediction performance is far lower than that of the GC degradation model.

## 6. Conclusion

In this paper, the GC degradation model with long-range dependence and fractal has been described and applied for the prediction of the remaining useful life of gearbox. The main conclusions are as follows:

1. The autocorrelation function curves under different parameters reveal that the generalized Cauchy process is more flexible than the single-parameter model in describing fractal time series.
2. The preprocessing results of the equipment degradation data verify the effectiveness of kernel principal component analysis. The characteristics of the degradation sequence show that the fractal dimension and long-range dependence of the degradation data in different environments are different. The Hurst index and the fractal dimension are independent of each other, which can better allow the generalized Cauchy process to describe the degradation sequence. This can ensure the prediction accuracy of the remaining service life within the maximum prediction range.
3. The remaining useful life prediction of the equipment degradation data shows that the prediction performance of the GC degradation model in the form of nonlinear power rate drift is better than the exponential drift form, which verifies the heavy tail property of the generalized Cauchy process with power rate decay. The generality of the generalized Cauchy degradation model is verified by predicting the remaining useful life of the equipment degradation data sets with different conditions. At the same time, the comparison case of various remaining useful life prediction methods verifies that the generalized Cauchy degradation model has superior prediction performance.

## CRediT authorship contribution statement

**He Liu:** Conceptualization, Formal analysis, Software, Writing - original draft. **Wanqing Song:** Data curation, Methodology, Validation, Supervision. **Yuhui Niu:** Software, Validation, Writing - review & editing. **Enrico Zio:** Validation, Writing - review & editing.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.ymsp.2020.107471>.

## References

- [1] Y. Qin, X. Wang, Z. Jingqiang, The optimized deep belief networks with improved logistic sigmoid units and their application in fault diagnosis for planetary gearboxes of wind turbines, *IEEE Trans. Industr. Electron. PP* (2018) 1, <https://doi.org/10.1109/TIE.2018.2856205>.
- [2] S.T. Kandukuri, A. Klausen, H.R. Karimi, K.G. Robbersmyr, A review of diagnostics and prognostics of low-speed machinery towards wind turbine farm-level health management, *Renew. Sustain. Energy Rev.* 53 (2016) 697–708, <https://doi.org/10.1016/j.rser.2015.08.061>.
- [3] F. Cheng, L. Qu, W. Qiao, L. Hao, Enhanced particle filtering for bearing remaining useful life prediction of wind turbine drivetrain gearboxes, *IEEE Trans. Industr. Electron.* 66 (6) (2019) 4738–4748, <https://doi.org/10.1109/TIE.2018.2866057>.
- [4] J. Wu, K. Hu, Y. Cheng, H. Zhu, X. Shao, Y. Wang, Data-driven remaining useful life prediction via multiple sensor signals and deep long short-term memory neural network, *ISA Trans.* 97 (2020) 241–250, <https://doi.org/10.1016/j.isatra.2019.07.004>.
- [5] H. Zhang, Z. Mo, J. Wang, Q. Miao, Nonlinear-drifted fractional brownian motion with multiple hidden state variables for remaining useful life prediction of lithium-ion batteries, *IEEE Trans. Reliab.* 69 (2) (2020) 768–780, <https://doi.org/10.1109/TR.2019.2896230>.
- [6] X.-S. Si, W. Wang, C.-H. Hu, D.-H. Zhou, Remaining useful life estimation – a review on the statistical data driven approaches, *Eur. J. Oper. Res.* 213 (1) (2011) 1–14, <https://doi.org/10.1016/j.ejor.2010.11.018>.
- [7] B. Sun, Y. Li, Z. Wang, Y. Ren, Q. Feng, D. Yang, M. Lu, X. Chen, Remaining useful life prediction of aviation circular electrical connectors using vibration-induced physical model and particle filtering method, *Microelectron. Reliab.* 92 (2019) 114–122, <https://doi.org/10.1016/j.microrel.2018.11.015>.
- [8] X. Li, W. Zhang, H. Ma, Z. Luo, X. Li, Data alignments in machinery remaining useful life prediction using deep adversarial neural networks, *Knowl.-Based Syst.* 197 (2020) 105843, <https://doi.org/10.1016/j.knsys.2020.105843>.
- [9] B. Wang, Y. Lei, N. Li, N. Li, A hybrid prognostics approach for estimating remaining useful life of rolling element bearings, *IEEE Trans. Reliab.* 69 (1) (2020) 401–412, <https://doi.org/10.1109/TR.2018.2882682>.
- [10] B. Yang, R. Liu, E. Zio, Remaining useful life prediction based on a double-convolutional neural network architecture, *IEEE Trans. Industr. Electron.* 66 (12) (2019) 9521–9530, <https://doi.org/10.1109/TIE.2019.2924605>.
- [11] C. Ordóñez, F. Sánchez Lasheras, J. Roca-Pardiñas, F.J.d.C. Juez, A hybrid arima svm model for the study of the remaining useful life of aircraft engines, *J. Comput. Appl. Math.* 346 (2019) 184–191, <https://doi.org/10.1016/j.cam.2018.07.008>.
- [12] M. Yan, X. Wang, B. Wang, M. Chang, I. Muhammad, Bearing remaining useful life prediction using support vector machine and hybrid degradation tracking model, *ISA Trans.* 98 (2020) 471–482, <https://doi.org/10.1016/j.isatra.2019.08.058>.
- [13] M. Baptista, E.P. Henriques, I. de Medeiros, J. Malere, C. Nascimento, H. Prendinger, Remaining useful life estimation in aeronautics: combining data-driven and kalman filtering, *Reliab. Eng. Syst. Saf.* 184 (2019) 228–239, <https://doi.org/10.1016/j.res.2018.01.017>.
- [14] B. Wang, Y. Lei, N. Li, T. Yan, Deep separable convolutional network for remaining useful life prediction of machinery, *Mech. Syst. Signal Process.* 134 (2019) 106330, <https://doi.org/10.1016/j.ymsp.2019.106330>.
- [15] E. Zio, G. Peloni, Particle filtering prognostic estimation of the remaining useful life of nonlinear components, *Reliab. Eng. Syst. Saf.* 96 (3) (2011) 403–409, <https://doi.org/10.1016/j.res.2010.08.009>.
- [16] Y. Lei, N. Li, J. Lin, A new method based on stochastic process models for machine remaining useful life prediction, *IEEE Trans. Instrum. Meas.* 65 (12) (2016) 2671–2684, <https://doi.org/10.1109/TIM.2016.2601004>.
- [17] N. Li, Y. Lei, L. Guo, T. Yan, J. Lin, Remaining useful life prediction based on a general expression of stochastic process models, *IEEE Trans. Industr. Electron.* 64 (7) (2017) 5709–5718, <https://doi.org/10.1109/TIE.2017.2677334>.
- [18] H. Sun, D. Cao, Z. Zhao, X. Kang, A hybrid approach to cutting tool remaining useful life prediction based on the wiener process, *IEEE Trans. Reliab.* 67 (3) (2018) 1294–1303, <https://doi.org/10.1109/TR.2018.2831256>.
- [19] H. Wang, X. Ma, Y. Zhao, An improved wiener process model with adaptive drift and diffusion for online remaining useful life prediction, *Mech. Syst. Signal Process.* 127 (2019) 370–387, <https://doi.org/10.1016/j.ymsp.2019.03.019>.
- [20] W. Peng, S.-P. Zhu, L. Shen, The transformed inverse gaussian process as an age and state-dependent degradation model, *Appl. Math. Model.* 75 (2019) 837–852, <https://doi.org/10.1016/j.apm.2019.07.004>.
- [21] S. Limon, O.P. Yadav, Predicting remaining lifetime using the monotonic gamma process and bayesian inference for multi-stress conditions, *Procedia Manuf.* 38 (2019) 1260–1267, <https://doi.org/10.1016/j.promfg.2020.01.218>.
- [22] Z. Chen, Y. Li, T. Xia, E. Pan, Hidden markov model with autocorrelated observations for remaining useful life prediction and optimal maintenance policy, *Reliab. Eng. Syst. Saf.* 184 (2019) 123–136, <https://doi.org/10.1016/j.res.2017.09.002>.
- [23] C.-P. Lin, J. Cabrera, F. Yang, M.-H. Ling, K.-L. Tsui, S.-J. Bae, Battery state of health modeling and remaining useful life prediction through time series model, *Appl. Energy* 275 (2020) 115338, <https://doi.org/10.1016/j.apenergy.2020.115338>.
- [24] M.H. Ling, H.K.T. Ng, K.L. Tsui, Bayesian and likelihood inferences on remaining useful life in two-phase degradation models under gamma process, *Reliab. Eng. Syst. Saf.* 184 (2019) 77–85, <https://doi.org/10.1016/j.res.2017.11.017>.
- [25] H. Zhang, D. Zhou, M. Chen, X. Xi, Predicting remaining useful life based on a generalized degradation with fractional brownian motion, *Mech. Syst. Signal Process.* 115 (2019) 736–752, <https://doi.org/10.1016/j.ymsp.2018.06.029>.
- [26] Y. Qin, S. Xiang, Y. Chai, H. Chen, Macroscopic-microscopic attention in lstm networks based on fusion features for gear remaining life prediction, *IEEE Trans. Industr. Electron. PP* (2019) 1, <https://doi.org/10.1109/TIE.2019.2959492>.
- [27] S. Xiang, Y. Qin, C. Zhu, Y. Wang, H. Chen, Long short-term memory neural network with weight amplification and its application into gear remaining useful life prediction, *Eng. Appl. Artif. Intell.* 91 (2020) 103587, <https://doi.org/10.1016/j.engappai.2020.103587>.
- [28] H. Zhang, M. Chen, X. Xi, D. Zhou, Remaining useful life prediction for degradation processes with long-range dependence, *IEEE Trans. Reliab.* 66 (4) (2017) 1368–1379, <https://doi.org/10.1109/TR.2017.2720752>.
- [29] Z. Zhang, X.-S. Si, C. Hu, An age and state dependent nonlinear prognostic model for degrading systems, *IEEE Trans. Reliab.* 64 (4) (2015) 1214–1228.
- [30] M. Li, L. JiaYue, On the predictability of long-range dependent series, *Math. Probl. Eng.* 2010 (2010), <https://doi.org/10.1155/2010/397454>.
- [31] X. Xi, M. Chen, H. Zhang, D. Zhou, An improved non-markovian degradation model with long-term dependency and item-to-item uncertainty, *Mech. Syst. Signal Process.* 105 (2018) 467–480, <https://doi.org/10.1016/j.ymsp.2017.12.017>.
- [32] M. Li, Fractal time series—a tutorial review, *Math. Probl. Eng.* 2010 (2010), <https://doi.org/10.1155/2010/157264>.
- [33] M. Li, Multi-fractional generalized cauchy process and its application to teletraffic, *Physica A* 550 (2020) 123982, <https://doi.org/10.1016/j.physa.2019.123982>.
- [34] Z. Meng, J. Li, N. Yin, Z. Pan, Remaining useful life prediction of rolling bearing using fractal theory, *Measurement* 156 (2020) 107572, <https://doi.org/10.1016/j.measurement.2020.107572>.
- [35] M. Li, J.-Y. Li, Generalized cauchy model of sea level fluctuations with long-range dependence, *Physica A* 484 (2017) 309–335, <https://doi.org/10.1016/j.physa.2017.04.130>.
- [36] M. Li, S. Lim, W. Zhao, Long-range dependent network traffic: a view from generalized cauchy process, 2020.
- [37] H. Liu, W. Song, M. Li, A. Kudreyko, E. Zio, Fractional levy stable motion: finite difference iterative forecasting model, *Chaos, Solit. Fractals* 133 (2020) 109632, <https://doi.org/10.1016/j.chaos.2020.109632>.

- [38] M.D. Ortigueira, Introduction to fractional linear systems. Part 2. discrete-time case, *IEE Proceedings – Vision, Image Signal Process.* 147 (1) (2000) 71–78, <https://doi.org/10.1049/ip-vis:20000273>.
- [39] Y. Cheng, H. Zhu, K. Hu, J. Wu, X. Shao, Y. Wang, Reliability prediction of machinery with multiple degradation characteristics using double-wiener process and monte carlo algorithm, *Mech. Syst. Signal Process.* 134 (2019) 106333, <https://doi.org/10.1016/j.ymssp.2019.106333>.
- [40] X.-S. Si, W. Wang, M.-Y. Chen, C.-H. Hu, D.-H. Zhou, A degradation path-dependent approach for remaining useful life estimation with an exact and closed-form solution, *Eur. J. Oper. Res.* 226 (1) (2013) 53–66, <https://doi.org/10.1016/j.ejor.2012.10.030>.
- [41] Z. Tian, Chaotic characteristic analysis of network traffic time series at different time scales, *Chaos, Solit. Fractals* 130 (2020) 109412, <https://doi.org/10.1016/j.chaos.2019.109412>.
- [42] Y. Li, W. Song, J. Chen, Long correlation fault trend prediction for rolling bearings based on a dimensionless parameter, *Noise Vib. Control* 38 (2018) 141–145.
- [43] Y. Si, H.R. Karimi, H. Gao, Modelling and optimization of a passive structural control design for a spar-type floating wind turbine, *Eng. Struct.* 69 (2014) 168–182, <https://doi.org/10.1016/j.engstruct.2014.03.011>.
- [44] Q. Zhang, P. Li, X. Lang, A. Miao, Improved dynamic kernel principal component analysis for fault detection, *Measurement* 158 (2020) 107738, <https://doi.org/10.1016/j.measurement.2020.107738>.
- [45] M. Li, S.C. Lim, Modeling network traffic using generalized cauchy process, *Physica A* 387 (11) (2008) 2584–2594, <https://doi.org/10.1016/j.physa.2008.01.026>.
- [46] Y. Liu, G. Yang, L. Ming, H. Yin, Variational mode decomposition denoising combined the detrended fluctuation analysis, *Signal Process.* 125 (aug.) (2016) 349–364.
- [47] W. Song, C. Cattani, C.-H. Chi, Multifractional brownian motion and quantum-behaved particle swarm optimization for short term power load forecasting: an integrated approach, *Energy* 194 (2020) 116847, <https://doi.org/10.1016/j.energy.2019.116847>.
- [48] Y. Pan, R. Hong, J. Chen, W. Wu, A hybrid dbn-som-pf-based prognostic approach of remaining useful life for wind turbine gearbox, *Renewable Energy* 152 (2020) 138–154, <https://doi.org/10.1016/j.renene.2020.01.042>.
- [49] A. Kouadri, M. Hajji, M.-F. Harkat, K. Abodayeh, M. Mansouri, H. Nounou, M. Nounou, Hidden markov model based principal component analysis for intelligent fault diagnosis of wind energy converter systems, *Renewable Energy* 150 (2020) 598–606, <https://doi.org/10.1016/j.renene.2020.01.010>.
- [50] M.T. Rosenstein, J.J. Collins, C.J. De Luca, A practical method for calculating largest lyapunov exponents from small data sets, *Physica D* 65 (1) (1993) 117–134, [https://doi.org/10.1016/0167-2789\(93\)90009-P](https://doi.org/10.1016/0167-2789(93)90009-P).
- [51] M. Gradinaru, I. Nourdin, F. Russo, P. Vallois, m-order integrals and generalized itô's formula; the case of a fractional brownian motion with any hurst index, *Annales de l'Institut Henri Poincaré (B), Probab. Stat.* 41 (4) (2005) 781–806, <https://doi.org/10.1016/j.anihpb.2004.06.002>.
- [52] B. Pei, Y. Xu, J.-L. Wu, Stochastic averaging for stochastic differential equations driven by fractional brownian motion and standard brownian motion, *Appl. Math. Lett.* 100 (2020) 106006, <https://doi.org/10.1016/j.aml.2019.106006>.
- [53] Z.S. Ye, L.-P. Chen, L.C. Tang, M. Xie, Accelerated degradation test planning using the inverse gaussian process, *IEEE Trans. Reliab.* 63 (3) (2014) 750–763.
- [54] Y. Zhang, R. Xiong, H. He, M.G. Pecht, Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries, *IEEE Trans. Veh. Technol.* 67 (7) (2018) 5695–5705, <https://doi.org/10.1109/TVT.2018.2805189>.
- [55] Z. Zhang, X. Si, C. Hu, Y. Lei, Degradation data analysis and remaining useful life estimation: a review on wiener-process-based methods, *Eur. J. Oper. Res.* 271 (3) (2018) 775–796, <https://doi.org/10.1016/j.ejor.2018.02.033>.