Energy Engineering

Tech Science Press

Doi:10.32604/ee.2025.073843

ARTICLE

A Fusion Optimization Method for Remaining Useful Life Prediction of Wind Turbine Gearboxes

Wei Chen, Zhi Wei*, Tingting Pei, Jianghao Zhu and Yang Wu

School of Automation and Electrical Engineering, Lanzhou University of Technology, Lanzhou, 730050, China

*Corresponding Author: Zhi Wei. Email: 232080802001@lut.edu.cn

Received: 26 September 2025; Accepted: 25 November 2025; Published: 27 May 2026

ABSTRACT: Wind turbine gearboxes are critical components in large-scale power generation systems, and their unexpected failures often result in significant economic losses, long downtime, and decreased energy efficiency. Accurate prediction of their Remaining Useful Life (RUL) is therefore vital for enhancing operational reliability, implementing condition-based maintenance, and optimizing lifecycle management. However, existing approaches often neglect the memory effect in degradation processes and fail to establish an effective interaction between stochastic degradation modeling and *RUL* prediction. To address these challenges, this study proposes a novel fusion method that integrates a stochastic degradation model with an intelligent prediction framework. The degradation model employs Fractional Brownian Motion (FBM) to capture long-range dependence and memory effects in gearbox performance, while the prediction framework leverages an enhanced recurrent neural network optimized through evolutionary mechanisms. By linking degradation modeling with *RUL* prediction through parameter optimization, the proposed method strengthens the interaction between physical degradation and data-driven prediction. Simulation results based on gearbox datasets demonstrate that the proposed approach significantly improves *RUL* prediction performance, achieving a 23.2% reduction in $\overline{RMSE}$, a 26.7% improvement in $\overline{SF}$, and a 3.3% increase in R^2 compared with traditional RNN and LSTM models, highlighting its potential for practical deployment in wind farm operations to support proactive maintenance scheduling and enhance system reliability.

KEYWORDS: Wind turbine gearbox; remaining useful life prediction; stochastic degradation modeling; fractional Brownian motion; neural networks; condition-based maintenance

1 Introduction

Gearboxes play a crucial role as mechanical transmission components within wind turbines, enduring prolonged exposure to volatile loads and intricate environmental factors. Consequently, they are particularly susceptible to issues like fatigue wear and crack propagation, leading to degradation and potential failure [1]. As such, the precise estimation of the Remaining Useful Life (RUL) of these gearboxes and the dynamic evaluation of their health status represent pivotal technologies for enhancing the intelligent operation and maintenance standards of wind turbines [2].

Current research on wind turbine gearbox RUL prediction primarily focuses on two approaches: stochastic process degradation modeling and data-driven methods. In stochastic process degradation modeling, Peng et al. [3] introduced a performance degradation model based on the transformed inverse Gaussian process, incorporating explanatory variables and random effects to forecast the RUL of deteriorating equipment. This method demonstrated accurate and practical degradation process modeling and

RUL prediction when compared against real gearbox data. Limon and Yadav [4] predicted gearbox RUL under multi-stress conditions using a degradation model based on the monotonic gamma process and Bayesian inference. A degradation model [5] utilizing a two-stage gamma distribution in conjunction with a bidirectional long short-term memory neural network was introduced for monitoring gearbox degradation, as well as predicting and estimating RUL. Djeziri et al. [6] suggested a hybrid modeling strategy that combines stochastic process modeling with cluster analysis to address the limitations of purely data-driven methods in capturing early degradation patterns. Mu et al. [7] utilized the Wiener process to develop an RUL prediction model for wind turbine gearbox temperature data, validating the model's feasibility with operational data. Liu et al. [8] described an RUL prediction method based on the generalized Cauchy process to characterize equipment degradation behavior, outperforming traditional inverse Gaussian processes in prediction accuracy through comparative analysis. While these methods offer interpretability and can model life based on failure mechanisms, they often rely on simplified assumptions like independent increments, constant parameters, and linear degradation, limiting their adaptability to the complex, non-linear degradation patterns of wind turbine gearboxes during long-term operation. Additionally, classical degradation models may overlook the strong memory and temporal dependence on equipment operation, leading to limited recognition ability, especially in early degradation stages.

In terms of data-driven methods, deep learning techniques such as convolutional neural networks [9,10], gated recurrent units [11,12], and temporal convolutional networks [13–15] have been employed for RUL prediction with promising results. Xiang et al. [16] proposed a compact adaptive deep learning network with a regression layer for RUL prediction, validated using wind turbine operational data. Reference [17] utilized principal component analysis to fuse and reduce high-dimensional feature data, optimizing support vector machine parameters with a genetic algorithm to construct a gearbox RUL prediction framework. Zhao et al. [18] developed a kernel density-based RUL prediction method considering environmental influences on component life to enhance prediction accuracy. Chen et al. [19] enhanced RUL prediction accuracy by improving the similarity method, combining hierarchical clustering and genetic algorithms for component classification training. Zhang et al. [20] forecasted motor bearing RUL through multi-sensor data fusion and the multi-head attention mechanism algorithm, showcasing deep learning's potential in RUL prediction. Pan et al. [21] modeled wind turbine gearbox performance degradation via multi-sensor data fusion, highlighting data-driven methods' advantages in assessing equipment health. An RUL prediction model was proposed in [22] based on performance degradation indicators for rolling bearing health management, illustrating multi-level data analysis potential. While these methods automatically extract features from multi-source time-series data and address non-linear and high-dimensional data modeling requirements, they may struggle with challenges such as data imbalance, feature discrepancies, and varying degradation stages, leading to limited generalization ability and a high risk of overfitting. Moreover, their "black-box" nature hinders their application in scenarios with stringent reliability requirements.

In this study, a novel approach is proposed for predicting the RUL of wind turbine gearboxes by integrating Fractional Brownian Motion (FBM) stochastic degradation process modeling with a Quantum Genetic Chain-Coded Bidirectional Recurrent Neural Network (QGCC-BiRNN). The FBM process captures long memory and adjustable roughness, effectively characterizing the nonlinear evolution of gearbox degradation and enhancing prediction stability and physical consistency. The Quantum Genetic Chain-Coded (QGCC) strategy incorporates quantum coding to augment the bidirectional modeling capacity of the Bidirectional Recurrent Neural Network (BiRNN), facilitating differential modeling of diverse sensor data sources and feature-rich fusion. By combining stochastic degradation processes with data-driven methodologies, this approach offers a robust solution for assessing reliability and managing the health of wind turbine gearboxes.

2 Stochastic Degradation Model of Wind Turbine Gearbox Based on FBM Process

The FBM is a continuous non-Markovian process that exhibits long-term dependence and autocorrelation [23]. Its increments are fixed and correlated, introducing a long-range correlated structure. The Hurst exponent, H, is a statistic widely used in equipment degradation modeling to quantify the long-term dependence of time series. When $0 < H < 0.5$, the degradation sequence exhibits mean-reverting behavior, indicating a strong rebound tendency in the system. Conversely, when $H = 0.5$, the degradation sequence shows typical memorylessness, conforming to the characteristics of a random walk. In the case of $0.5 < H < 1$, the system degradation process has obvious persistence, and the future evolution trend is more likely to continue in the current degradation direction.

Assuming $X(t)$ adheres to an FBM process, the mathematical representation for its stochastic degradation is as follows:

$$X(t) = X(0) + \lambda t + \alpha \int_0^t \eta(\gamma;\beta)\, d\gamma + \sigma B_H(t) \tag{1}$$

In the formula, $X(0)$ represents the initial degradation state of the device, and it is assumed that $X(0) = 0$; λt characterizes the linear degradation trend; $\int_0^t \eta(\gamma;\beta)\, d\gamma$ reflects the nonlinear degradation trend; λ and α are random variable used to capture the individual differences caused by factors such as structural differences and environmental fluctuations; β and σ are common parameters; $B_H(t)$ is a fractional Brownian motion used to simulate the random disturbances with memory characteristics during the degradation process. To reflect the possible statistical dependence between the linear degradation term and the nonlinear degradation term, it is assumed that the corresponding random variables jointly follow a two-dimensional normal distribution. This assumption not only helps to characterize the collaborative variation characteristics between the two types of degradation components but also provides a theoretical basis for the subsequent derivation of the kernel density function of the remaining useful life and the estimation of relevant parameters.

The system's lifetime T is defined as the time at which the degradation amount $X(t)$ first reaches the failure threshold ω and its specific definition is as follows:

$$T = \inf\{X(t) \geq \omega | X(0) = 0\} \tag{2}$$

According to Eq. (1), the distribution of the degradation amount at any time t is given by $X(t) \sim N\left(m(t), \sigma^2 t^{2H}\right)$, where $m(t) = \lambda t + \alpha \int_0^t \eta(\gamma;\beta)\, d\gamma$. Employing the approximation form and combining it with Eq. (2), the approximate expression of the density function under the first-passage time condition is derived as:

$$f_T(t) \approx C \cdot t^{-(1+H)} \exp\left(-\frac{(\omega - m(t))^2}{2\sigma^2 t^{2H}}\right) \tag{3}$$

in the formula: C is the normalization constant; H is the Hurst exponent; ω is the failure threshold.

Furthermore, the expression of the final Kernel Density Estimation (KDE) function of the remaining useful life can be derived:

$$f_{RUL-KDE}(t) = \frac{1}{nh} \sum_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(t - T_i)^2}{2h^2}\right) \tag{4}$$

where n represents the sample size; h represents the kernel bandwidth; T_i satisfies $X(T_i) = \omega$.

3 Quantum Genetic Chain-Coded Bidirectional Recurrent Neural Network

QGCC leverages the superposition and entanglement properties of quantum states to consolidate numerous candidate gene configurations into single or multiple quantum states [24]. Through evolutionary quantum operations akin to genetic algorithms, superior genes progressively prevail, expediting the search and optimization process. Conversely, BiRNN employs two-directional RNNs to comprehend the present instant comprehensively. By amalgamating historical data and forthcoming cues, BiRNN can enhance the precision in modeling the dynamic attributes of the complete sequence.

Let the normalized time series input to the QGCC-BiRNN at time t be represented as $x^t = [x_1^t, x_2^t, \cdots, x_n^t]^T (x_i^t \in [a_i, b_i] = [0,1])$, with the quantum state of the input layer at time t denoted by $|x^t\rangle = [|x_1{}^t\rangle, |x_2{}^t\rangle, \cdots, |x_n{}^t\rangle]^T$. The calculation of $|x_i{}^t\rangle$ is given by:

$$\begin{aligned} |x_i^t\rangle &= [\cos\alpha_i^t, \sin\alpha_i^t]^T \\ &= \cos\left(\frac{2\pi(x_i^t - a_i)}{b_i - a_i}\right)|0\rangle + \sin\left(\frac{2\pi(x_i^t - a_i)}{b_i - a_i}\right)|1\rangle \\ &= \left[\cos\left(\frac{2\pi(x_i^t - a_i)}{b_i - a_i}\right), \sin\left(\frac{2\pi(x_i^t - a_i)}{b_i - a_i}\right)\right]^T \\ &= [\cos(2\pi x_i^t), \sin(2\pi x_i^t)]^T \end{aligned} \tag{5}$$

in the formula, α_i^t represents the phase of the element $|x_i^t\rangle$; a_i and b_i represents the upper and lower limits of the values of the time series x_i^t.

The formula for calculating the hidden layer state in the QGCC-BiRNN network at time t is given by:

$$\begin{aligned} |h_j^t\rangle &= \left[\cos\beta_j^t, \sin\beta_j^t\right]^T \\ &= |h_j^t\rangle_x + |h_j^t\rangle_h + |h_j^t\rangle_y \\ &= \prod_{i=1}^n \left(R\left(\theta_{ij}^t\right)|x_i^t\rangle\right) + \prod_{j=1}^p \left(R\left(\delta_{jj}^t\right)|h_j^{t-1}\rangle\right) + \prod_{k=1}^m \left(R\left(\gamma_{kj}^t\right)|y_k^t\rangle\right) \end{aligned} \tag{6}$$

where β_j^t represents the phase of $|h_j^t\rangle$; $|h_j^t\rangle_x$, $|h_j^t\rangle_h$ and $|h_j^t\rangle_y$ represent the three linear components at time t, which are respectively controlled by the input layer at time t, the hidden layer at time $t-1$ and the output layer at time t; $R(\cdot)$ represents the quantum phase-shift gate; θ_{ij}^t, δ_{jj}^t and γ_{kj}^t are the corresponding phases at time t.

Upon further analysis, $|h_j^t\rangle_x$, $|h_j^t\rangle_h$, $|h_j^t\rangle_y$ can be expressed as follows:

$$|h_j^t\rangle_x = \prod_{i=1}^n \left(R\left(\theta_{ij}^t\right)|x_i^t\rangle\right) = \prod_{i=1}^n \cos\left(\alpha_i^t + \theta_{ij}^t\right)|0\rangle + \prod_{i=1}^n \sin\left(\alpha_i^t + \theta_{ij}^t\right)|1\rangle \tag{7}$$

$$|h_j^t\rangle_h = \prod_{j=1}^p \left(R\left(\delta_{jj}^t\right)|h_j^{t-1}\rangle\right) = \prod_{j=1}^p \cos\left(\beta_j^{t-1} + \delta_{jj}^t\right)|0\rangle + \prod_{j=1}^p \sin\left(\beta_j^{t-1} + \delta_{jj}^t\right)|1\rangle \tag{8}$$

$$|h_j^t\rangle_y = \prod_{k=1}^m \left(R\left(\gamma_{kj}^t\right)|y_k^t\rangle\right) = \prod_{k=1}^m \cos\left(\varepsilon_k^t + \gamma_{kj}^t\right)|0\rangle + \prod_{k=1}^m \sin\left(\varepsilon_k^t + \gamma_{kj}^t\right)|1\rangle \tag{9}$$

Substituting Eqs. (7)–(9) into Eq. (6), the final calculation formula for the hidden layer state $|h_j^t\rangle$ can be obtained, as shown below:

$$|h_j^t\rangle = \left[\cos\beta_j^t, \sin\beta_j^t\right]^T$$

$$
\begin{aligned}
&= \prod_{i=1}^{n}\left(R\left(\theta_{ij}^{t}\right)\left|x_{i}^{t}\right\rangle\right)+\prod_{j=1}^{p}\left(R\left(\delta_{jj}^{t}\right)\left|h_{j}^{t-1}\right\rangle\right)+\prod_{k=1}^{m}\left(R\left(\gamma_{kj}^{t}\right)\left|y_{k}^{t}\right\rangle\right)\\
&= \left\{\prod_{i=1}^{n}\cos\left(\alpha_{i}^{t}+\theta_{ij}^{t}\right)+\prod_{j=1}^{p}\cos\left(\beta_{j}^{t-1}+\delta_{jj}^{t}\right)+\prod_{k=1}^{m}\cos\left(\varepsilon_{k}^{t}+\gamma_{kj}^{t}\right)\right\}|0\rangle\\
&+\left\{\prod_{i=1}^{n}\sin\left(\alpha_{i}^{t}+\theta_{ij}^{t}\right)+\prod_{j=1}^{p}\sin\left(\beta_{j}^{t-1}+\delta_{jj}^{t}\right)+\prod_{k=1}^{m}\sin\left(\varepsilon_{k}^{t}+\gamma_{kj}^{t}\right)\right\}|1\rangle\\
&= \begin{bmatrix}\prod_{i=1}^{n}\cos\left(\alpha_{i}^{t}+\theta_{ij}^{t}\right)+\prod_{j=1}^{p}\cos\left(\beta_{j}^{t-1}+\delta_{jj}^{t}\right)+\prod_{k=1}^{m}\cos\left(\varepsilon_{k}^{t}+\gamma_{kj}^{t}\right)\\ \prod_{i=1}^{n}\sin\left(\alpha_{i}^{t}+\theta_{ij}^{t}\right)+\prod_{j=1}^{p}\sin\left(\beta_{j}^{t-1}+\delta_{jj}^{t}\right)+\prod_{k=1}^{m}\sin\left(\varepsilon_{k}^{t}+\gamma_{kj}^{t}\right)\end{bmatrix}
\end{aligned}
\tag{10}
$$

Finally, the calculation formula for the output layer state $\left|y_{k}^{t}\right\rangle$ at time t is as follows:

$$
\begin{aligned}
\left|y_{k}^{t}\right\rangle &= \left[\cos\varepsilon_{k}^{t},\sin\varepsilon_{k}^{t}\right]^{T}\\
&= \prod_{j=1}^{p}\left(R\left(\varphi_{jk}^{t}\right)\left|\beta_{j}^{k}\right\rangle\right)\\
&= \prod_{j=1}^{p}\cos\left(\beta_{j}^{t}+\varphi_{jk}^{t}\right)|0\rangle+\prod_{j=1}^{p}\sin\left(\beta_{j}^{t}+\varphi_{jk}^{t}\right)|1\rangle
\end{aligned}
\tag{11}
$$

The QGCC-BiRNN network, formulated using Eqs. (5)–(11), possesses the capability to dynamically modify weight parameters. This network autonomously optimizes its structure based on the output layer's state, improving the alignment between input features and the global memory mechanism. Consequently, it enhances the accuracy of modeling intricate nonlinear relationships and demonstrates better nonlinear approximation capabilities.

4 RUL Prediction Framework Based on FBM Process and QGCC-BiRNN

The wind turbine gearbox's RUL prediction model, as presented in this study, combines FBM degradation modeling and a bidirectional recursive neural network with quantum genetic chain coding. The prediction framework, illustrated in Fig. 1, encompasses data preprocessing, multi-source sensor data fusion, FBM stochastic process parameter estimation, and the RUL prediction interacts with the FBM process.

4.1 Data Preprocessing of Wind Turbine Gearboxes

Enhancing the precision and reliability of the predictive model for estimating the RUL of gearboxes necessitates the preprocessing of the initially gathered data. Typically, data from wind turbine gearboxes encompass a range of sensor readings, including rotational speed, torque, and vibration, which frequently exhibit noise and outliers. The main stages of data preprocessing are as follows:

(1) Detection and elimination of outliers

In this study, the Local Outlier Factor (LOF) algorithm is utilized for outlier detection to enhance data reliability [25]. The LOF algorithm is adept at detecting anomalous samples with notably lower local density compared to their neighboring environment. By appropriately defining the k-value and threshold, disruptive data points can be detected and eliminated, thereby improving the precision of model training.

(2) Normalization processing

The Min-Max normalization technique is utilized to linearly scale the gathered gearbox degradation data to the range $[0,1]$, ensuring a standardized numerical scale for different physical quantities. Let the original data sequence be represented as $\{x_1(t), x_2(t), \cdots, x_n(t)\}$, with its maximum value denoted as $x_{\max}$ and its minimum value denoted as $x_{\min}$. The normalized data expression $\overline{x}_i(t)$ is calculated as:

$$\overline{x_i}(t) = \frac{x_i(t) - x_{\min}}{x_{\max} - x_{\min}}, i = 1, 2, \cdots, n \tag{12}$$

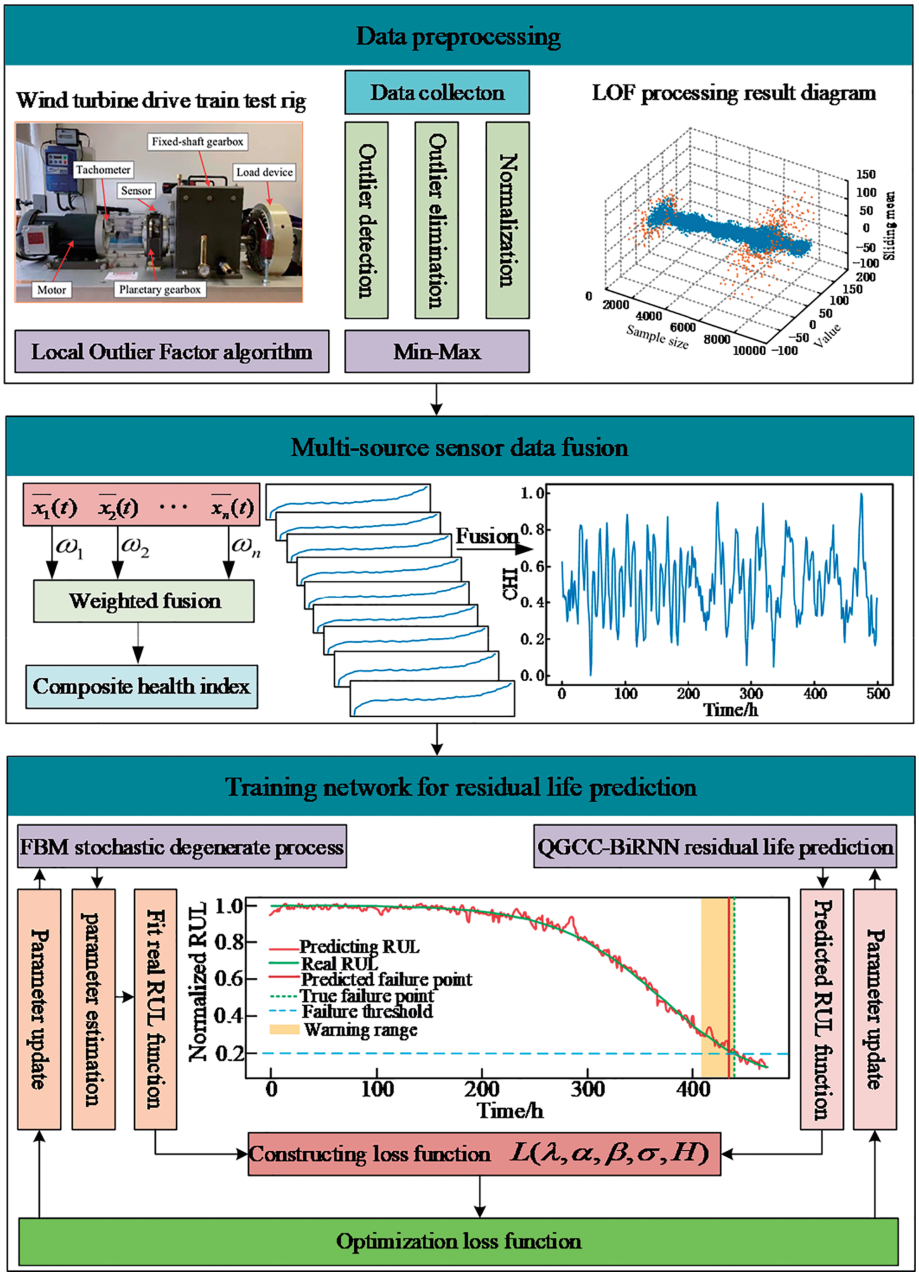


Figure 1: Block diagram of remaining useful life prediction

4.2 Multi-Source Sensor Data Fusion

Assuming there are n sensors gathering state degradation data from various parts of the wind turbine gearbox, let $\overline{x}_i(t)$ represents the degradation state information collected by the i-th sensor at time t. After normalization, a weighted fusion method is then applied to integrate this information and generate a Composite Health Indicator (CHI) for monitoring the data from multiple sensors. The specific formulation is as follows:

$$CHI(t) = \sum_{j=1}^{n} \overline{x}_j(t)\,\omega_j \tag{13}$$

where ω_j represents the weighting coefficient of the j-th sensor, which is used to measure the weight of this sensor in the process of forming health indicators.

4.3 Parameter Estimation of FBM Stochastic Process

Maximum likelihood estimation is used to estimate the correlation parameters of the FBM process. First define $m(t;\theta) = \lambda t + \alpha \int_0^t \eta(\gamma;\beta)\,d\gamma$, and rewrite the degenerate model as $X(t) = m(t;\theta) + \sigma B_H(t)$; a vector form of observations $\boldsymbol{X} = [X(t_1), X(t_2), \cdots, X(t_n)]^T$, $\boldsymbol{m}(\theta) = [m(t_1;\theta), m(t_2;\theta), \cdots, m(t_n;\theta)]^T$, then $\boldsymbol{X} \sim N\left(\boldsymbol{m}(\theta), \sigma^2 \sum_H\right)$, where $\sum_H = [\sum_H]_{ij} = 1/2\left(t_i^{2H} + t_j^{2H} - \left|t_i - t_j\right|^{2H}\right)$ is the covariance matrix of the FBM. Follow these steps to complete the parameter estimation process:

(1) Construction of Logarithmic Likelihood Function

To assess the likelihood of gearbox data given the parameters of the FBM model, a log-likelihood function is formulated in the following manner:

$$\log L(\theta, \sigma, H) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log\left(\left|\sigma^2 \sum_H\right|\right) - \frac{1}{2}(\boldsymbol{X} - \boldsymbol{m}(\theta))^T \left(\sigma^2 \sum_H\right)^{-1} (\boldsymbol{X} - \boldsymbol{m}(\theta)) \tag{14}$$

where $\boldsymbol{X}$ represents the observation vector; the function $\boldsymbol{m}(\theta)$ denotes the model mean function; $\sum_H$ represents the parameterized covariance matrix, whose structure is determined by the Hurst exponent H. The log-likelihood function assesses the likelihood of gearbox data generation given a specific model configuration and serves as the foundation for subsequent parameter estimation.

(2) Estimation of Trend Parameters $\theta = (\lambda, \alpha, \beta)$

The optimal trend parameters under H are derived by minimizing the quadratic component within the likelihood function. This approach is analogous to conducting weighted least squares estimation on sample residuals, incorporating a known covariance matrix, to enhance the precision of parameter estimation.

$$\min_{\lambda,\alpha,\beta} (\boldsymbol{X} - \boldsymbol{m}(\theta))^T \left(\sum_H^{-1}\right) (\boldsymbol{X} - \boldsymbol{m}(\theta)) \tag{15}$$

(3) Estimated Fluctuation Intensity σ^2

According to the maximum-likelihood criterion, the fluctuation intensity is uniformly estimated by Eq. (16). This estimation quantifies the average variability in the data while accounting for trend parameters and is a crucial component of the likelihood function.

$$\widehat{\sigma^2} = \frac{1}{n}\left(\boldsymbol{X} - \boldsymbol{m}\left(\hat{\theta}\right)\right)^T \sum_H^{-1} \left(\boldsymbol{X} - \boldsymbol{m}\left(\hat{\theta}\right)\right) \tag{16}$$

(4) Estimation of Hurst Exponent H

To enhance the model's characterization of long-range dependent structures, the Hurst exponent H is optimally estimated as follows:

$$\hat{H} = \arg\min_{H} \left\{ \log\left(\left|\sum\nolimits_{H}\right|\right) + \frac{1}{\widehat{\sigma^2}} \left(\boldsymbol{X} - \boldsymbol{m}\left(\hat{\theta}\right)\right)^{T} \sum\nolimits_{H}^{-1} \left(\boldsymbol{X} - \boldsymbol{m}\left(\hat{\theta}\right)\right) \right\} \tag{17}$$

The optimization task involves determining the Hurst exponent H that most effectively characterizes the covariance matrix structure of data under the premise of fixed trend parameter $\theta = (\lambda, \alpha, \beta)$ and the fluctuation intensity parameter σ^2. This aims to capture the long-range dependence and internal correlation within time series data.

4.4 Interaction between Residual Life Prediction and FBM Process

A nonlinear mapping of the degradation state $X(t)$ is defined to form a Composite Health Index $CHI(t) = g\left(X(t); \vartheta_{feature}\right)$. This health index $CHI(t)$ serves as the input for the QGCC-BiRNN network for time series prediction. Subsequently, a residual life estimation value, namely: $\widehat{RUL(t)} = f_{QGCC-BiRNN}\left(\left\{CHI(s)_{s=0}^{t}\right\}; \vartheta_{QGCC}\right)$, is obtained. The predicted results $\widehat{RUL(t)}$ are used to optimize the FBM parameters to make the degradation model more accurate. The detailed procedure is outlined below:

(1) Construction of loss function

$$L(\lambda, \alpha, \beta, \sigma, H) = \left[\widehat{RUL}(t; \lambda, \alpha, \beta, \sigma, H) - RUL_{real}(t)\right]^2 \tag{18}$$

The loss function quantifies the disparity between the predicted *RUL* and the actual *RUL*. By incorporating the life prediction outcomes of the QGCC-BiRNN network into the optimization objective, adjustments to the degradation model parameters are made based on prediction errors. This process aims to enhance the alignment between FBM parameters and the observed degradation pattern.

(2) Establishment of gradient formula

$$\frac{\partial L}{\partial \vartheta_{FBM}} = \frac{\partial L}{\partial \widehat{RUL}} \cdot \frac{\partial \widehat{RUL}}{\partial CHI} \cdot \frac{\partial CHI}{\partial X(t)} \cdot \frac{\partial X(t)}{\partial \vartheta_{FBM}} \tag{19}$$

The formula above uses the chain rule to obtain the partial derivative of the final loss function L to the *FBM* process parameter $\vartheta_{FBM} = (\lambda, \alpha, \beta, H)$, which ensures that the *FBM* modeling has learnability in the life prediction model and can adjust the trend generated by itself reversely through the feedback information at the *RUL* end.

(3) Establish update rules and complete parameter reverse update

$$\vartheta^{(k+1)}{}_{QGCC} = \vartheta^{(k)}{}_{QGCC} - \eta_1 \frac{\partial L}{\partial \vartheta_{QGCC}} \tag{20}$$

$$\vartheta^{(k+1)}{}_{FBM} = \vartheta^{(k)}{}_{FBM} - \eta_2 \frac{\partial L}{\partial \vartheta_{FBM}} \tag{21}$$

The parameters of the QGCC-BiRNN network and FBM degenerate model are updated in reverse order in the respective formulas, where η_1 and η_2 represent the weight coefficients for backward updates. Eq. (20) optimizes the weights of the QGCC-BiRNN network to enhance time series modeling capabilities, and Eq. (21) adjusts the FBM degradation model based on feedback from the prediction error loss function to align it more closely with the degradation evolution pattern observed in gearbox data. This dual-channel

parameter optimization approach improves the model's adaptability to degradation processes, facilitating the integrated optimization of data-driven and stochastic process modeling.

5 Example Analysis

5.1 Examples Illustrate

To assess the efficacy of the proposed approach, we simulated and validated vibration signal data from a gearbox obtained during an acceleration experiment conducted on an experimental platform as described in reference [26]. The dimensions and details of gear elements are shown in Table 1, and the test conditions of the experiment are shown in Table 2.

Table 1: Gear element dimensions and details

Gear element	Number of teeth	Mate teeth	Root diameter (mm)	Helix angle	Face width (mm)	Ratio
Ring gear	99	39	1047	7.5L	230	/
Planet gear	39	99	372	7.5L	227.5	/
Sun pinion	21	39	186	7.5R	220	5.71
Intermediate gear	82	23	678	14R	170	/
Intermediate pinion	23	82	174	14L	186	3.57
High-speed gear	88	22	400	14L	110	/
High-speed pinion	22	88	100	14R	120	4.0

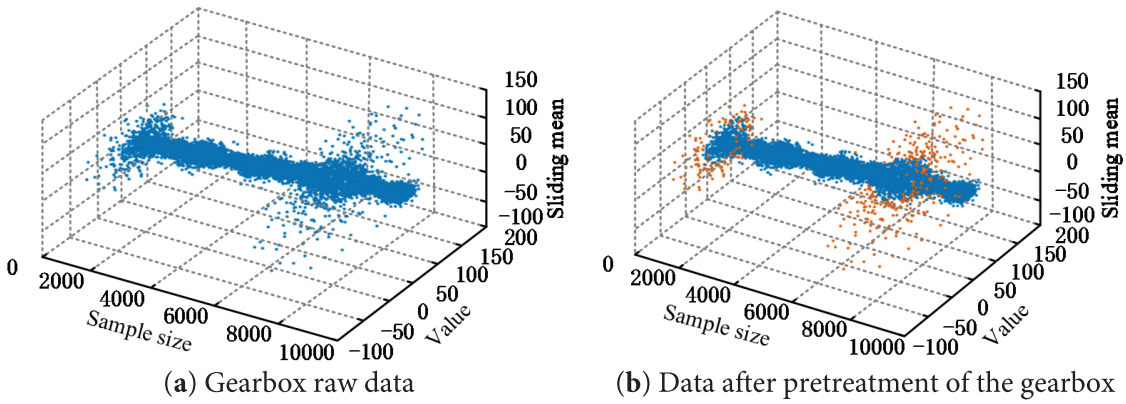
Table 2: Test condition

Main shaft speed (rpm)	Nominal HSS speed (rpm)	Electric power (% of rated)	Duration (min)
22.09	1800	50%	10

5.2 Residual Life Prediction

The gearbox data was preprocessed using the LOF algorithm to remove abnormal data points. The preprocessing outcomes are depicted in Fig. 2.

The preprocessed data on gearbox degradation undergoes normalization and fusion to generate a CHI, as depicted in Fig. 3. Subsequently, degradation process modeling is conducted based on the evolving pattern of the composite health index in conjunction with the FBM process. Following smoothing, the residual prediction label for the wind turbine gearbox is established, serving as the authentic life label for training QGCC-BiRNN, illustrated in Fig. 4.



(**a**) Gearbox raw data (**b**) Data after pretreatment of the gearbox

Figure 2: Pretreatment result of gearbox degradation data

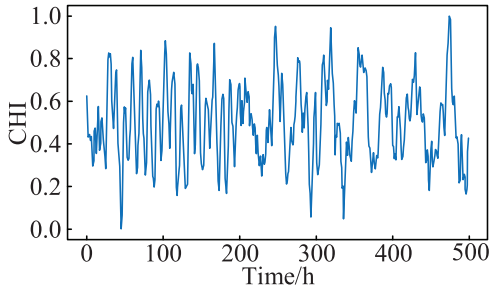


Figure 3: Trend chart of composite health indicators

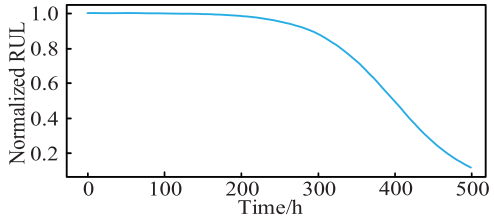


Figure 4: Performance degradation curve of gearbox based on FBM process

The failure threshold is defined as 0.2, denoting gearbox failure when both the actual and predicted gearbox lifespans are below this threshold. Normalized gearbox data is fed into the QGCC-BiRNN network for predicting gearbox lifespan, with the corresponding results depicted in Fig. 5.

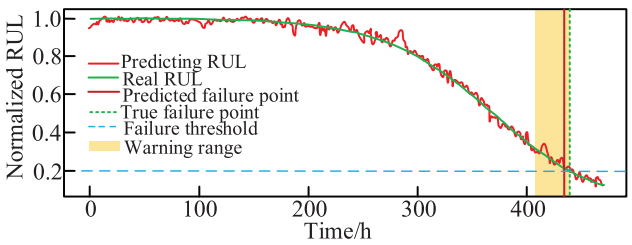


Figure 5: Schematic diagram of QGCC-BiRNN prediction results

The figure above illustrates the predictive performance of the QGCC-BiRNN model in estimating the RUL of wind power gearboxes in comparison to degradation curves derived from the FBM process. Overall, the model's prediction closely aligns with the degradation pattern of wind turbine gearboxes based on the FBM process, demonstrating consistent accuracy across various time intervals. Particularly noteworthy is the model's adeptness at capturing pivotal points of rapid RUL decline during the intermediate and advanced stages of degradation, showcasing its proficiency in time series modeling. Notably, the model accurately identifies the onset of system degradation, closely matching the actual position on the prediction curve, underscoring its capability to recognize initial degradation trends. Furthermore, the prediction curve exhibits minimal oscillation or significant deviations from the actual trend, indicating that the QGCC-BiRNN model bolsters the stability of time series modeling by integrating quantum gene chain coding and a bidirectional neural network architecture.

To enhance the model's practicality, this study introduces a dual-threshold warning mechanism based on predictive outcomes. The yellow interval in the figure denotes the warning range. When the predicted RUL falls below 0.3 but exceeds 0.2, it enters the warning range; if it drops below 0.2, a failure warning is issued [27]. The predictive plot illustrates a gradual decline in the model's RUL leading up to the True Failure Point, with timely activation of the warning zone, indicating the model's capability to anticipate failure trends. In practical applications, early warnings may entail resource wastage, whereas delayed responses can escalate faults. The model demonstrates relatively precise identification of gearbox performance failure criticality, triggering warnings sufficiently in advance to strike a balance between early warning and accuracy.

Fig. 6 illustrates the distribution of estimated RUL values for the gearbox following KDE. The probability density curve of the predictions exhibits a noticeable right-skewed distribution, with prominent peaks primarily clustered within intermediate to high RUL intervals. This trend suggests a conservative bias in the model's predictions during the equipment's non-complete degradation phase, effectively addressing the issue of significant errors in predictions observed in traditional models during this stage of gearbox degradation. The overall distribution curve appears smooth, devoid of undesirable features such as discrete or double peaks, indicating the model's output stability and consistent statistical significance in the prediction outcomes.

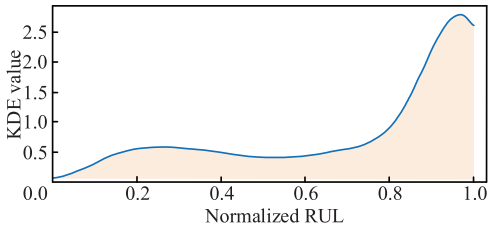


Figure 6: Kernel density function for residual life prediction

The KDE plot illustrates that the predicted values span the entire range from 0 to 1, suggesting high prediction coverage across various stages of degradation without any concentration bias towards a specific stage. Notably, there remains a discernible prediction density in the late degradation phase near failure, indicating the model's capability to effectively discern equipment states nearing failure.

Fig. 7 presents the variation of the FBM+QGCC-BiRNN model's error with the number of training epochs. From the convergence trend, it is evident that the training error decreases sharply during the initial phase, indicating that the model rapidly learns the dominant degradation feature patterns within the first few iterations. As training progresses, the convergence curve gradually flattens and stabilizes around zero after approximately 30 epochs, suggesting that the model has effectively converged. By integrating the physical

constraints of the degradation process captured by the FBM with the adaptive feature extraction mechanism of the QGCC-BiRNN, the model achieves rapid convergence and maintains a low error level within a limited number of iterations. This convergence curve clearly demonstrates the effectiveness and stability of the proposed hybrid model.

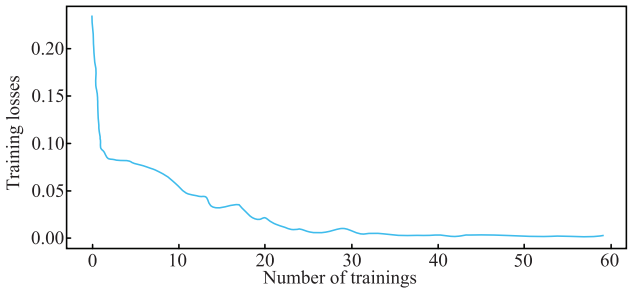


Figure 7: The convergence curve of the FBM+QGCC-BiRNN model

5.3 Error Analysis and Model Comparison

To assess the predictive accuracy of the QGCC-BiRNN model, Root Mean Squared Error (*RMSE*), Scoring Function (*SF*) and coefficient of determination (R^2) [28] are introduced to evaluate prediction errors. The calculation formula is as follows:

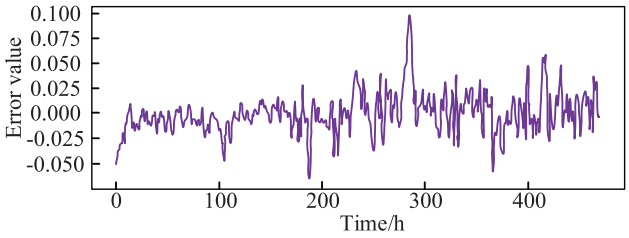
$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(RUL_{i,predict} - RUL_{i,real}\right)^2} \tag{22}$$

$$SF = \begin{cases} \sum_{i=1}^{n}\left(e^{\frac{RUL_{i,pred}-RUL_{i,real}}{13}} - 1\right), & \left(RUL_{i,predict} - RUL_{i,real}\right) < 0 \\ \sum_{i=1}^{n}\left(e^{\frac{RUL_{i,pred}-RUL_{i,real}}{10}} - 1\right), & \left(RUL_{i,predict} - RUL_{i,real}\right) \geq 0 \end{cases} \tag{23}$$

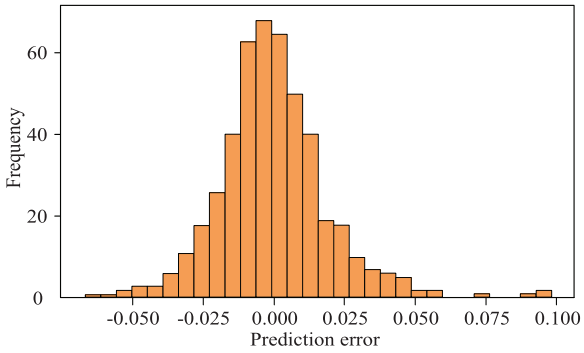
$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(RUL_{i,real} - RUL_{i,predict}\right)^2}{\sum_{i=1}^{n}\left(RUL_{i,real} - \frac{1}{n}\sum_{i=1}^{n} RUL_{i,real}\right)^2} \tag{24}$$

The smaller the *RMSE* and *SF* values, the better the prediction effect of the model. The closer the coefficient of determination (R^2) is to 1, the more accurately the model captures the relationship between the predicted and actual values.

In Fig. 8a, the prediction error curve remains in a low range in the initial stage, then rises slightly, and then tends to decline after reaching a certain peak in the middle and late stages of degradation. This error variation law accords with the evolution characteristics of actual equipment from health to failure: in the early stage, due to the stable operation of the wind turbine gearbox system, the prediction error is relatively small; while near the failure point, due to the increased fluctuation of the sensing signal, the prediction error increases slightly. Fig. 8b shows that the prediction error of the gearbox life of the whole wind turbine is concentrated around. Fig. 8a,b jointly reflects that the prediction model proposed in this paper can maintain overall stability in the life prediction process of wind turbine gearbox, reflecting its strong anti-interference ability to data noise.



(**a**) Trend chart of prediction error



(**b**) Prediction error distribution histogram

Figure 8: Residual life prediction error curve and error distribution histogram

Fig. 9 demonstrates that the RMSE values stabilize at low levels, indicating the model's efficacy in error control for overall predictions. Particularly during the end-of-life stage with RUL approaching zero, the prediction errors do not exhibit significant amplification, underscoring the model's robustness. The *SF* index, a metric sensitive to prediction timeliness and accuracy, offers a comprehensive assessment of the wind turbine gearbox's life prediction. The stable *SF* values depicted in the figure, devoid of abrupt fluctuations, suggest that the prediction outcomes accurately capture the true degradation pattern. These findings further affirm the overall superiority of the FBM-QGCC-BiRNN model in the RUL prediction task for wind turbine gearboxes.

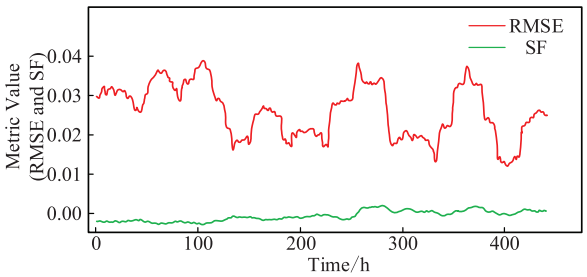
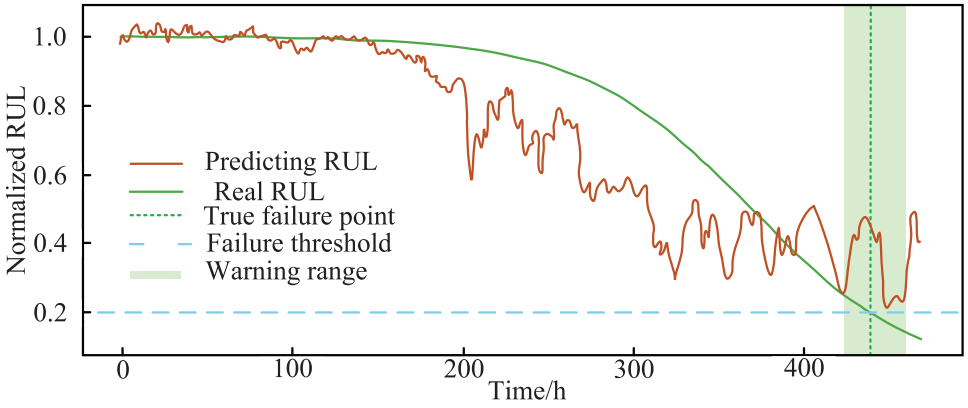


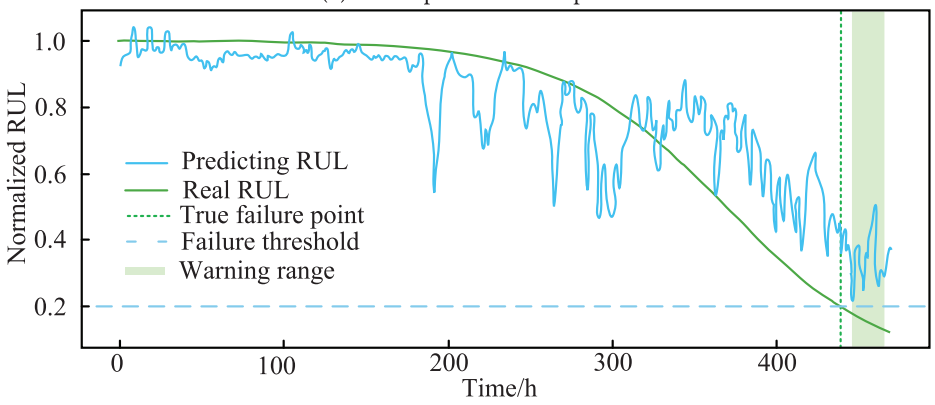
Figure 9: RMSE and *SF* curve

To further verify the effectiveness of the proposed method, comparative experiments were conducted between the FBM+QGCC-BiRNN model and conventional RNN and LSTM models. The experimental results are presented in Fig. 10a,b. Combined with the results shown in Figs. 5, 8 and 9, it can be observed that the proposed model maintains a small and stable prediction error throughout the entire degradation cycle. The predicted curve closely aligns with the real degradation trajectory across the full lifecycle, exhibiting a smooth overall trend without significant oscillations. This indicates that the model can accurately

capture the degradation evolution of the gearbox's performance. In particular, during the failure stage, the predicted failure point almost coincides with the true failure point, demonstrating the model's high temporal consistency and accuracy in end-of-life prediction.



(**a**) RNN operation result plot



(**b**) LSTM operation result plot

Figure 10: Running curves of traditional models RNN and LSTM

As summarized in Table 3, the quantitative results further confirm the superiority of the proposed approach. The RMSE of the FBM+QGCC-BiRNN model is 0.0354, which is lower than that of RNN (0.0413) and LSTM (0.0461). The *SF* is 1.42×10^{-3}, substantially outperforming RNN (1.73×10^{-3}) and LSTM (1.86×10^{-3}). The R^2 reaches 0.9847, higher than those of RNN (0.9635) and LSTM (0.9514). These results demonstrate that the proposed model not only achieves superior prediction accuracy compared with traditional deep learning methods but also exhibits stronger stability and trend-tracking capability, thereby validating the overall effectiveness and robustness of the FBM+QGCC-BiRNN model.

Table 3: Comparative analysis of performance parameters

	$\overline{RMSE}$	$\overline{SF}$	R^2
FBM+QGCC-BiRNN	0.0354	1.42×10^{-3}	0.9847
RNN	0.0413	1.73×10^{-3}	0.9635
LSTM	0.0461	1.86×10^{-3}	0.9514
Improvement (%)	23.2%	26.7%	3.3%

6 Conclusions

This study proposed a fusion optimization method that combines a FBM-based stochastic degradation model with a QGCC-BiRNN to predict the RUL of wind turbine gearboxes. The proposed approach effectively integrates the physical interpretability of FBM with the strong temporal feature extraction and learning capability of the QGCC-BiRNN, enabling accurate and reliable characterization of gearbox degradation behavior.

The FBM-based stochastic model effectively captured the nonlinear and long-range dependency in the degradation process, providing a smooth and physically consistent representation of gearbox deterioration. The proposed FBM+QGCC-BiRNN model achieved a 23.2% reduction in $\overline{RMSE}$, a 26.7% improvement in $\overline{SF}$, and a 3.3% increase in R^2 compared with conventional RNN and LSTM models, confirming its superior accuracy and stability in RUL prediction. Furthermore, the QGCC-BiRNN demonstrated strong responsiveness during the initial degradation phase, effectively mitigating the "flat-line convergence" problem encountered in traditional recurrent architectures while maintaining high prediction precision throughout the mid- and late-life stages.

Overall, the FBM+QGCC-BiRNN framework provides a reliable and interpretable technical approach for the health management of wind turbine gearboxes. Future research will focus on integrating multi-source condition monitoring data and developing lightweight architectures to enable real-time deployment in industrial wind farm environments.

Acknowledgement: None.

Funding Statement: This work was supported by the National Natural Science Foundation of China (No. 51767017), the Major Project of Gansu Provincial Joint Research Fund (No. 25JRRA1143), and the Key Research and Development Program of Gansu Province (No. 25YFGA032).

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: Wei Chen and Zhi Wei; data collection: Zhi Wei; analysis and interpretation of results: Zhi Wei, Tingting Pei and Jianghao Zhu; draft manuscript preparation: Zhi Wei and Yang Wu. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

1. Pichika SVVSN, Yadav R, Geetha Rajasekharan S, Praveen HM, Inturi V. Optimal sensor placement for identifying multi-component failures in a wind turbine gearbox using integrated condition monitoring scheme. Appl Acoust. 2022;187(3):108505. doi:10.1016/j.apacoust.2021.108505.
2. Salameh JP, Cauet S, Etien E, Sakout A, Rambault L. Gearbox condition monitoring in wind turbines: a review. Mech Syst Signal Process. 2018;111(4):251–64. doi:10.1016/j.ymssp.2018.03.052.
3. Peng W, Zhu SP, Shen L. The transformed inverse Gaussian process as an age- and state-dependent degradation model. Appl Math Model. 2019;75(5):837–52. doi:10.1016/j.apm.2019.07.004.
4. Limon S, Yadav OP. Predicting remaining lifetime using the monotonic gamma process and Bayesian inference for multi-stress conditions. Procedia Manuf. 2019;38:1260–7. doi:10.1016/j.promfg.2020.01.218.
5. Li Z, Zhou J, Nassif H, Coit D, Bae J. Fusing physics-inferred information from stochastic model with machine learning approaches for degradation prediction. Reliab Eng Syst Saf. 2023;232(1):109078. doi:10.1016/j.ress.2022.109078.
6. Djeziri MA, Benmoussa S, Sanchez R. Hybrid method for remaining useful life prediction in wind turbine systems. Renew Energy. 2018;116(4):173–87. doi:10.1016/j.renene.2017.05.020.
7. Mu S, Su Y, Jing K, Li C, Wang T. Remaining life prediction of wind turbine bearing based on Wiener process. IOP Conf Ser Mater Sci Eng. 2020;788(1):012089. doi:10.1088/1757-899x/788/1/012089.
8. Liu H, Song W, Niu Y, Zio E. A generalized cauchy method for remaining useful life prediction of wind turbine gearboxes. Mech Syst Signal Process. 2021;153(6):107471. doi:10.1016/j.ymssp.2020.107471.
9. Takahashi R, Matsubara T, Uehara K. Data augmentation using random image cropping and patching for deep CNNs. IEEE Trans Circuits Syst Video Technol. 2020;30(9):2917–31. doi:10.1109/tcsvt.2019.2935128.
10. de Oliveira da Costa PR, Akçay A, Zhang Y, Kaymak U. Remaining useful lifetime prediction via deep domain adaptation. Reliab Eng Syst Saf. 2020;195(060):106682. doi:10.1016/j.ress.2019.106682.
11. Encalada-Davila A, Moyon L, Tutiven C, Puruncajas B, Vidal Y. Early fault detection in the main bearing of wind turbines based on gated recurrent unit (GRU) neural networks and SCADA data. IEEE/ASME Trans Mechatron. 2022;27(6):5583–93. doi:10.1109/tmech.2022.3185675.
12. Renström N, Bangalore P, Highcock E. System-wide anomaly detection in wind turbines using deep autoencoders. Renew Energy. 2020;157(1):647–59. doi:10.1016/j.renene.2020.04.148.
13. Rezamand M, Kordestani M, Orchard ME, Carriveau R, Ting DS, Saif M. Improved remaining useful life estimation of wind turbine drivetrain bearings under varying operating conditions. IEEE Trans Ind Inf. 2021;17(3):1742–52. doi:10.1109/tii.2020.2993074.
14. Merainani B, Laddada S, Bechhoefer E, Chikh MAA, Benazzouz D. An integrated methodology for estimating the remaining useful life of high-speed wind turbine shaft bearings with limited samples. Renew Energy. 2022;182(17–18):1141–51. doi:10.1016/j.renene.2021.10.062.
15. Xiang S, Qin Y, Liu F, Gryllias K. Automatic multi-differential deep learning and its application to machine remaining useful life prediction. Reliab Eng Syst Saf. 2022;223(15):108531. doi:10.1016/j.ress.2022.108531.
16. Xiang S, Qin Y, Luo J, Wu F, Gryllias K. A concise self-adapting deep learning network for machine remaining useful life prediction. Mech Syst Signal Process. 2023;191:110187. doi:10.1016/j.ymssp.2023.110187.
17. Yang Y. Research on predicting the RUL of wind turbine gearboxes using SVM optimized by GA [dissertation]. Jilin, China: Jilin University; 2017. (In Chinese).
18. Zhao B, Li JJ, Shi H, Ren QL, Kang H. Remaining useful life prediction of component kernel density estimation considering environmental changes. J Vib Shock. 2024;43(8):145–54. (In Chinese). doi:10.13465/j.cnki.jvs.2024.08.016.
19. Chen Y, Rao Y, Cai Z, Wang Z. Remaining useful lifetime prediction and economic reserve strategy of equipment components based on improved similarity. J Syst Eng Electron. 2021;43(9):2688–96. (In Chinese).
20. Zhang W, Zhang T, Jia M, Cai J. Prediction of remaining life of motor bearings using multi-sensor fusion and MHA-LSTM. Instrumentation. 2024;45(3):84–93. (In Chinese).
21. Pan Y, Hong R, Chen J, Wu W. A hybrid DBN-SOM-PF-based prognostic approach of remaining useful life for wind turbine gearbox. Renew Energy. 2020;152(1–2):138–54. doi:10.1016/j.renene.2020.01.042.

22. Yang C, Ma J, Wang X, Li X, Li Z, Luo T. A novel based-performance degradation indicator RUL prediction model and its application in rolling bearing. ISA Trans. 2022;121(5):349–64. doi:10.1016/j.isatra.2021.03.045.
23. Kasinathan R, Kasinathan R, Chalishajar D, Sandrasekaran V, Baleanu D. Trajectory controllability of impulsive neutral stochastic functional integrodifferential equations driven by fBm with noncompact semigroup via Mönch fixed point. Qual Theory Dyn Syst. 2024;23(2):72. doi:10.1007/s12346-023-00917-6.
24. Chen Y, Li F, Wang J, Tang B, Zhou X. Quantum recurrent encoder-decoder neural network for performance trend prediction of rotating machinery. Knowl Based Syst. 2020;197(24):105863. doi:10.1016/j.knosys.2020.105863.
25. Horyń C, Nowak-Brzezińska A. Automatic block size optimization in the LOF algorithm for efficient anomaly detection. Appl Soft Comput. 2025;170(1):112675. doi:10.1016/j.asoc.2024.112675.
26. Liu D, Cui L, Cheng W. A review on deep learning in planetary gearbox health state recognition: methods, applications, and dataset publication. Meas Sci Technol. 2024;35(1):012002. doi:10.1088/1361-6501/acf390.
27. Chen D, Qin Y, Qian Q, Wang Y, Liu F. Transfer life prediction of gears by cross-domain health indicator construction and multi-hierarchical long-term memory augmented network. Reliab Eng Syst Saf. 2023;230:108916. doi:10.1016/j.ress.2022.108916.
28. Mu H, Zheng J, Hu C, Zhao R, Dong Q. Remaining useful life prediction of multivariate degradation equipment based on CDBN and BiLSTM. Chin J Aeronaut. 2022;43(7):301–12. (In Chinese).