IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 11, 2019, accepted May 24, 2019, date of publication May 28, 2019, date of current version June 13, 2019.

*Digital Object Identifier 10.1109/ACCESS.2019.2919663*

# A Symbolic Regression Based Residual Useful Life Model for Slewing Bearings

**PENG DING[1], QINRONG QIAN[1], HUA WANG[1], AND JIANYONG YAO[2]**
[1]School of Mechanical and Power Engineering, Nanjing Tech University, Nanjing 211816, China
[2]School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

Corresponding author: Hua Wang (wanghua@njtech.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51875273, and in part by the Project of Jiangsu Provincial Six Talent Peaks under Grant GDZB-033.

**ABSTRACT** Slewing bearings are vital functional components of large machinery. It is of far reaching significance to study their life prediction and health management. Many studies are based on data-driven approaches. However, part of them in the form of ''black-box'' lack actual physical meanings due to opacity model structures and have difficulty in choosing optimal parameters. Few kinds of literature focus on explicit model relationships for slewing bearings' life models. In this paper, a novel approach based on symbolic regression is proposed with the aim of exploring slewing bearings' explicit life models in depth and to predict residual useful life (RUL). The proposed method integrates the strengths of multiple signals describing a comprehensive response to slewing bearings' health and various genetic programming (GP) algorithms modeling life expressions. In addition, independent, hybrid, and piecewise strategies are introduced and explicit model relationships with respect to degradation indicators (DIs) are established via GPs. To verify the proposed method, three run-to-failure experiments under discrepant operating conditions of slewing bearings are carried out. Prediction results demonstrate that models generated by epigenetic linear genetic programming (ELGP) under hybrid and piecewise modeling strategy with similarity-based combination strategy perform best. More importantly, their life expressions are more succinct and intelligible than in other situations.

**INDEX TERMS** Predictive models, prognostics and health management, remaining life assessment, slewing bearings, symbolic regression.

## I. INTRODUCTION

slewing bearings are widely used in industrial machineries (e.g., wind turbine generators, tunnel-boring machines, tower cranes and military technology). They serve as connections in slewing systems and bear axial force, radial force and overturning moment under harsh outdoor environments or improper installation. If failure occurs, the entire machine will be out of service and even a great accident might happen. Therefore, a detailed and effective technique of residual useful life (RUL) prediction is urgently needed to reduce downtime of industrial machineries and maintenance cost. This paper is devoted to introduce a novel prognostics strategy for slewing bearings' RUL prediction under discrepant operating conditions.

Prognostics methods can be mainly divided into four categories [1]: physics model-based, statistical model-based, artificial intelligence (AI) and hybrid approaches. Kunc and Prebil [2], Kunc *et al.* [3] analyzed the contact between raceways and rolling bodies by finite element methods and it can be applied to predict the raceway life and analysis the contact strength of a single-row, four-point contact slewing bearing. This approach was based on first principle of damage and failure mechanisms and required a sufficient research and complete understanding for mechanical components. It is not easy for complex systems or components to completely grasp damage mechanism [4]. Another way in the form of special models relies on statistics, mainly talking about reliability problem of RUL. Feng presented a novel approach [5] of parameters estimation based on Weibull distribution for life prediction of slewing bearings but failed to provide online prediction. Statistical-based methods usually rely on mass data. They are also considered to be data-driven based approaches and another part of these ones is based on AI algorithms, AI can establish diagnosis or prognosis

The associate editor coordinating the review of this manuscript and approving it for publication was Chuan Li.

models from signals directly without specific prior knowledge. Kim *et al.* [6] used support vector machine (SVM) classifiers to estimate health status of bearing degradation process and provided long-term bearings' degradation assessment and prediction. However, this RUL prediction lacked connections with actual physical degradation process. Lu *et al.* [7] discussed least square support vector machines (LSSVM) improved by particle swarm algorithm (PSO) in slewing bearing's degradation prognosis but it was limited to the same working condition. Guo *et al.* [8] forecasted RUL of bearings by means of particle filter (PF) algorithm after indicator construction by recurrent neural network algorithm which highly depended on massive data. In the field of system health prognostics, data-driven approaches have been widely used and show effectiveness for complex systems or mechanical parts in the absence of complete degradation mechanism [4]. However, some of data-driven models including the above ones, especially AI ones, have some defects: Firstly, prediction or diagnosis results were influenced greatly by part parameters inside them and it was not easy to find the most suitable ones that may cause difficulties for operators who were not proficient in this algorithm and influenced the ability of generalization [9]; Second was the opacity of model structure which was short of practical physical meanings [10], [11], it was just like a black-box and you may be hindered from truly understanding the life model or degradation process. For scenarios requiring high reliability, this uncertainty and opacity may be fatal. The motivation of this paper is to explore this problem.

Another factor that influences prediction accuracy of data-driven approaches is quality and quantity of input data. In traditional bearings, vibration data was considered as effective signal [6], [8], [12], [13] for its ability to reflect the degradation trend. Owing to the low speed and heavy load characteristics, large-size slewing bearings are of great diversity from small bearings. If only vibration signal is applied in slewing bearings' condition monitoring, it may not perform as well as bearings and a comparison is given in Section IV. Caesarendra and vokelj [14], [15] both combined vibration signals with acoustic emission techniques to diagnose slewing bearings' fault. Ding *et al.* [16], [17] applied vibration, temperature and torque signals to assess health status of slewing bearings. Above multiple signal based solutions performed better than single signal ones. That is to say single signal may lose fault information compared with multi signal.

In view of two factors above, vibration, temperature and torque signal are discussed in this paper and symbolic regression [18] based prediction methods which are the first attempt in the field of rotating machinery are applied later. They mainly focus on the exploration of explicit mapping relationships between RUL and DIs and provide a reliable and effective life prediction way.

Threefold contributions are summarized: (1) Life models describing explicit relationships between slewing bearings' health and DIs can be obtained which is helpful to further study internal structures of life models and semi-empirical life models with determined functional structures could be summarized or identified through massive experiments in the future; (2) Researchers can obtain a preliminary identification of the most representative damage signal from this symbolic life expressions (it is noted that sensitive indicators need to be verified next and they are just preliminary judgments) which could be helpful to select suitable monitoring signals; (3) RUL estimation just relies on the formulas like Eq. (12), as shown at the bottom of page 11. which has low computational pressure. Hence, the calculation cost can be reduced greatly. Researchers might establish lightweight expert systems in view of this merit in the near future. Remaining sections of this paper are organized as follows: Section II presents signal processing strategies. Section III describes symbolic regression methods and their principle. Verifications related to run-to-failure experiments are discussed in Section IV. Section V summarizes the paper.

## II. SIGNAL PROCESSING STRATEGIES

A good degradation indicator is one of the main factors influencing the prognostics result. In this section, three parts of signal processing strategies, determining DIs, are presented in the following subsections. Also, three-dimension of signal (vibration, temperature and torque) is utilized to continuously and integrally depict the slewing bearings' degradation process which is a combination of our past study [7], [19].

### A. SIGNAL DENOISING USING EEMD-SVD

In this section, we introduce ensemble empirical mode decomposition (EEMD) [20] and singular value decomposition (SVD) [21] algorithms for signal denoising. SVD is a powerful tool applied in many domains (e.g., features extraction and dimensionality reduction), also serving as a signal denoising method in rolling bearings widely. EEMD is the development of EMD which avoids inherent shortcomings via the assistance of white noise. It can decompose the signal into components from high to low frequency and usually, noise is stored in high frequency. We take advantage of this feature and combine SVD algorithm for signal denoising. The detailed comparison with wavelet denoising algorithms is discussed in Section 4.2.2. Detailed steps based on EEMD-SVD algorithms are as follows.

*Step I:* Decompose the raw signal into IMFs and residue using EEMD method as Eq. (1).

$$s(t) = \sum_{n=1}^{n} c_n(t) + r(t) \tag{1}$$

where $s(t)$ is raw signal, $c_n(t)$ and $r(t)$ are IMFs and residue, respectively.

*Step II:* Determine the number $m$ of low frequency IMFs and form the trend component with the residue as Eq. (2).

$$tr(t) = \sum_{m=1}^{m} c_m(t) + r(t) \tag{2}$$

where $tr(t)$ is trend component and $c_m(t)$ are low frequency IMFs.

*Step III:* Then the raw signal is divided into trend components and remaining high frequency components as shown in Eq. (3) where $k + m{=}n$.

$$s(t) = \sum_{k=1}^{k} c_k(t) + tr(t) \tag{3}$$

where $c_k(t)$ are high frequency IMFs.

*Step IV:* Each of the remaining high frequency component is reconstructed using SVD with difference spectrum [22] that determines the number of singular values.

*Step V:* Generate final denoising-signal $s^*(t)$ as Eq. (4) using the reconstructed and the trend components.

$$s^*(t) = \sum_{k=1}^{k} c_k^*(t) + tr(t) \tag{4}$$

where $s^*(t)$ and $c_k^*(t)$ are denoising-signal and reconstructed components, respectively.

### B. FEATURES EXTRACTION & SELECTION

Three main types of features: time domain, frequency domain and time-frequency domain features are utilized in this paper. Detailed formula expressions are the same with [16].

Subsequently, three selection metrics are conducted for suitable features. One is monotonicity, applying to all features, which is a little different from quantitative description of monotonicity [23]. Considering the imprecision and incompleteness of that description and supposing the irreversibility of a degenerate process for slewing bearings, we apply a qualitative trend description for monotonicity. That is, the trend is consistent over a wide range, allowing volatility within the short interval. For IMF energy features, we adopt another two metrics: correlation [24] and kurtosis [7] defined as Eq. (5-6). Correlation calculating similarity between relative important features and candidate features, kurtosis describing waveform peak and indicating the damage inside the machine are applied to select the suitable scale space of time-frequency features that can well reflect degradation process and reduce redundancy to some extent.

$$corr = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n} (X_i - \overline{X})^2 \sum_{i=1}^{n} (Y_i - \overline{Y})^2}} \tag{5}$$

where $X_i$ is a series of root mean square (RMS) value extracted from the original signal, $Y_i$ is a series of feature. $\bar{X}$, $\bar{Y}$ are their mean values.

$$K = \frac{E(X - \overline{X})^4}{\sigma^4} = \frac{1}{n} \sum_{i=1}^{n} \left[ \frac{X(t) - \overline{X}}{\sigma} \right]^4 \tag{6}$$

where $\bar{X}$, $\sigma$ are mean and standard deviation of the series of $X(t)$. The screening principles are as follows: First apply the monotonicity metric to filter all the indicators and obtain the required indicators. then calculate the correlation between RMS and time-frequency features, only value of correlation greater than 0.85 will be retained. The remaining time-frequency features will be further evaluated under kurtosis and the relatively bigger ones are winners at last.

### C. FEATURES FUSION VIA LPP ALGORITHM

A feature fusion method, which is a new attempt for slewing bearings, named locality preserving projection (LPP) [25] is presented in this part. LPP, one of the manifold learning algorithms [26], optimally preserves the neighborhood structure of the data and extensively apply in the domain of exploratory data analysis, feature fusion, machine vision, and pattern identification. It can greatly preserve the local information from high dimensional features which is more useful than global information based fusion algorithms explained amply in [27] and has low computational cost in the comparison with other manifold learning algorithms. However, nearest neighbor nodes and $\beta$ in heat kernel are two parameters need to be predefined in LPP which may influence the trajectory trend of DIs. Detailed steps of LPP can be found in [25].

## III. THREE SYMBOLIC REGRESSION METHODS AND THEIR PRINCIPLES

Symbolic regression methods for modeling slewing bearings' life model are introduced and presented in this section. Genetic programming (GP) algorithm [28] and its two variants: multi genes genetic programming (MGGP) algorithm [29], [30] and epigenetic linear genetic programming (ELGP) algorithm [10] are employed to study slewing bearings' life models.

### A. INTRODUCTION OF THREE TYPES OF GP

GP, a branch of evolution algorithms, can obtain solutions under the fitness function pressure without preset model structure. Inspired by the Darwinian principle of natural selection and survival of the fittest, GP is a solution based on global optimization and extensively used in the domain such as nonlinear function approximation, system identification, pattern recognition, machine learning, etc. Hence, we introduce this thinking into the study of slewing bearings' life model. It can directly obtain life model expressions and figure out the relationships between RUL and DIs compared with "black box" algorithms. Based on this feature, forms of expression of different DIs to the degradation process can be clearly and quantitatively analyzed and researchers could further study the degradation process of slewing bearings.

However, in [31] and [32], researchers point out the traditional GP has some weaknesses: (1) bloat which means there exists no improvement of fitness as trees grow; (2) when dealing with complex problems, GP often performs bad; (3) large search space may cause the increase in computational cost and complicated models. As a variant of GP, the only difference between MGGP and GP is the number of gens or trees. It can improve the accuracy trading off the complexity of models. It seems that modeling by MGGP is

more precise than traditional GP, but the greater the gens number is, the longer the expression of model is. The users will have difficulty understanding this tediously long model expressions. In fact, part of GP's shortcomings mentioned above still exist in MGGP. ELGP as a novel GP, which will be introduced in detail in the next subsection, has innovations as follows:1) it adopts stack-based [33] approach to express models; 2) ELGP owns the ability of local optimization by introducing the layer of epigenetic [34]; 3) multi-objective optimization is utilized in ELGP to make the model intelligible and succinct. Figure 1 of detail steps from GP and ELGP is presented below:

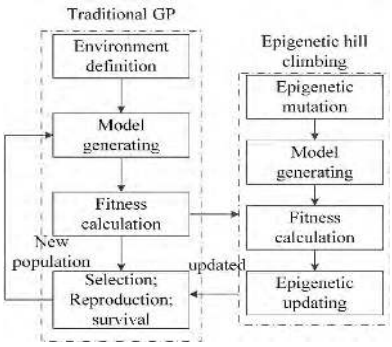


**FIGURE 1. Procedure of traditional GP is on the left, right is epigenetic hill climbing procedure.**

### *B. DETAILED DESCRIPTION ON ELGP*

The detailed description of ELGP is centered around its three salient innovations against traditional GP.

#### 1) STACK-BASED MODEL REPRESENTATION

ELGP utilizes stack-based representation to epigenetic layers. Compared with tree-based representation which may be invalid owing to the change of instructions, it can guarantee the validity of the syntax all the time under any instruction. Robustness of this approach depends mainly on ignoring the executions that have a larger size than the current stack. For example, if $x+$ operator wants to implement and the $x$ is the only element on the stack, the result is meaningless. Another rule is that it only remains the expression on the top of the stack. Here are some examples in Eq. (7).

$$\begin{aligned} i_1 &= [x_1x_2-] \Rightarrow M_1 : (x_1 - x_2) \\ i_2 &= [x_1x_2 - + \times /] \Rightarrow M_2 : (x_1 - x_2) \\ i_3 &= [x_3 + x_1 \times x_1x_2-] \Rightarrow M_3 : (x_1 - x_2) \end{aligned} \tag{7}$$

Three genotypes express the same model: $x_1$-$x_2$. From $i_2$ we can find the execution $+ \times /$ are ignored due to the insufficient stack size; $i_3$ only remains the last expression due to the second rule.

#### 2) EPIGENETIC LAYER

Epigenetic layer which is ignored by traditional GP is used to play a role of an on/off marker controlling the genotypes. This corresponding sequence of on/off markers is referred to as an epigenome. For a clearer understanding of layer of epigenetic, a few examples can be found in Eq. (8):

$$\begin{aligned} i_3 \Rightarrow i_3' &= \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ & & x_3 + x_1 \times x_1x_2- & & & & \end{bmatrix} \Rightarrow M_3' : (x_3 - x_2) \\ i_3 \Rightarrow i_3'' &= \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ & & x_3 + x_1 \times x_1x_2- & & & & \end{bmatrix} \Rightarrow M_3'' : (x_3x_1) \\ i_3 \Rightarrow i_3''' &= \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ & & x_3 + x_1 \times x_1x_2- & & & & \end{bmatrix} \Rightarrow M_3''' : (x_1 - x_2) \end{aligned} \tag{8}$$

From the above representations, 1/0 means on/off and the binary numbers above the genotype is an epigenome (epigenetic layer). The epigenetic markers are initialized randomly. In the evolution of the epigenetic layer, epigenetic hill climbing (EHC), which the whole processing is in the right half of Figure 1 and stems from local optimization algorithm: random hill climbing algorithm, are applied. It can make each element of epigenome has a probability to mutate (i.e. from 1 to 0, or 0 to 1) presenting in Figure 2 and turns to the smaller fitness (fitness $f_i$ is being minimized) value with succinct models. The evaluation mechanism of individual in EHC is based on the model's accuracy and complexity [10]. Only the lower complexity with smaller or equal fitness can pass this evolution, otherwise, the new individual will be discarded.

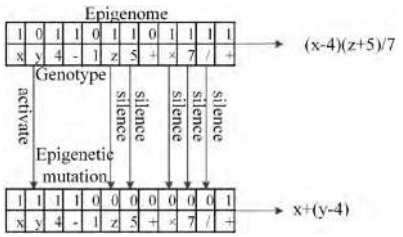


**FIGURE 2. Epigenetic mutation: Activation means the element of epigenome changes from 0 to 1 and vice versa.**

#### 3) MULTI-OBJECTIVE OPTIMIZATION

To obtain succinct and accurate models, ELGP apply multi-objective optimization: accuracy, complexity, and age of individuals in the evolution. The first two objectives ensure the final model interpretability and accuracy, the third objective, age of individual which is the number of generations since the oldest ancestor was born, avoids premature convergence. Combined with three objectives, models via ELGP can be more precise and succinct.

## IV. EXPERIMENT VERIFICATION

In this section, run-to-failure experiments of slewing bearings under three working conditions are carried out to certificate the proposed methods above and we will discuss this symbolic regression based approaches in depth.

### *A. EXPERIMENTAL FACILITIES AND LOADING PROCESSING*

Two test rigs will be introduced in this part. One is laboratory slewing bearings test rig, which is suitable for relatively small

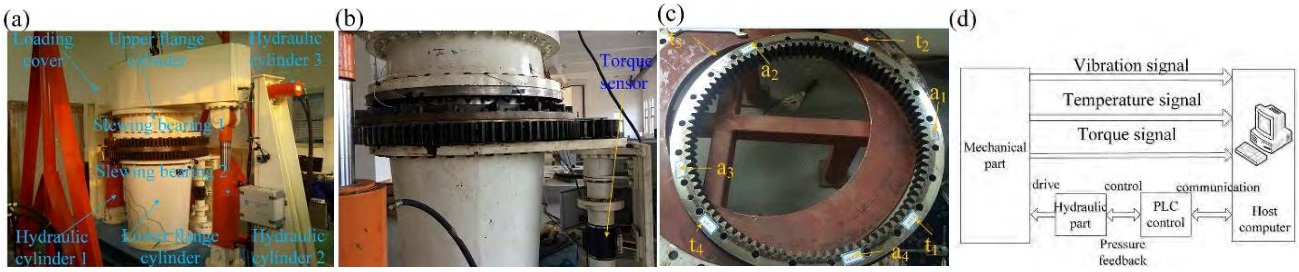


**FIGURE 3.** Lab test system and corresponding sensor installations: (a) lab test stan; (b) installation of torque sensor; (c) installation of acceleration $a_{1-4}$ and temperature $t_{1-4}$ sensor; (d) flow of the whole test stand.

**TABLE 1.** Loading processing of experiment A&B.

| NO. | Axial force | | Overturning moment | | Rotating speed | | Duration | |
|---|---|---|---|---|---|---|---|---|
| Experiment | A | B | A | B | A | B | A | B |
| 1 | 24(KN) | 82.5 | 60(KN·m) | 34.5 | 4(rpm) | 2 | 30(minute) | |
| 2 | 48 | 165 | 120 | 69 | 4 | 2 | 30 | |
| 3 | 72 | 247.5 | 180 | 103.5 | 4 | 2 | 30 | |
| 4 | 96 | 330 | 240 | 138 | 4 | 4 | run to failure | |

size slewing bearings' experiments. Two sets of run-to-failure data are generated from this rig. Structures of test rig and layouts of sensors are represented in Figure 3. It contains mechanical part, hydraulic loading part and measuring and controlling part. The hydraulic load part shown in Figure 3(a) is mainly composed of hydraulic cylinders and a hydraulic motor. The torque sensor is mounted on the gear of the drive mechanism presented in Figure 3(b). Accelerometers with the sample frequency is 2048Hz and temperature sensors (PT100) are installed every 90 degrees at the outer ring and oil fill holes from Figure 3(c). Figure 3(d) gives a brief flowchart of the entire experiment.

Two experiments under this lab test rig using the same type slewing bearing which the raceway diameter is 730 millimeters were carried out in June (Experiment A) and in August (Experiment B). Both of them were run for 12 days and 152 hours in Nanjing, China respectively. Main difference between the two experiments are ambient temperature which ambient temperature of Experiment A is much lower than that of Experiment B and loading process recorded in detail in Table 1 below. Also, slewing bearings' assembly and installation foundation will influence the life of slewing bearings to a certain extent.

Another industrial test rig is suitable for slewing bearings with large size. Its structure is shown in Figure 4. The principle is similar to the lab one before. Experiment C from this rig lasted 72 days from March to June in Shanghai, China.

Slewing bearings under this industrial test rig is installed in 3MW wind turbines and their raceway diameter is 2410 millimeters. Loading processing is as Table 2 below.

### B. STUDY ON RUL PREDICTION OF SLEWING BEARINGS

This section is devoted to detailly discuss the RUL prediction of slewing bearings. Figure 5 below present the entire

**FIGURE 4.** Industrial test rig.

**TABLE 2.** Loading processing of experiment C.

| NO. | Axial force | Overturning moment | Rotating speed | Duration |
|---|---|---|---|---|
| 1 | 0(KN) | 0(KN·m) | 0.5(rpm) | 30(minute) |
| 2 | 115.9 | 1074.3 | 1.5 | 60 |
| 3 | 231.7 | 2148.6 | 1.5 | 240 |
| 4 | 347.6 | 3222.9 | 1.5 | 240 |
| 5 | 463.8 | 4297.2 | 1.5 | run to failure |

process intuitively. Utilizing signal processing strategies and symbolic regression methods as discussed before, we combine two slewing bearings' life models established by GPs algorithms to forecast RUL of the third slewing bearing. It is noted that signal processing strategies are the same in the train sample A and B and the validation set.

#### 1) DATA PROCESSING

Three sets of experiments data including vibration temperature and torque signal can be obtained from the lab and industrial test rig. In Figure 6, we found that noise is still widespread. Hence de-noising processing is of great significant for signal processing. In this paper, we utilize

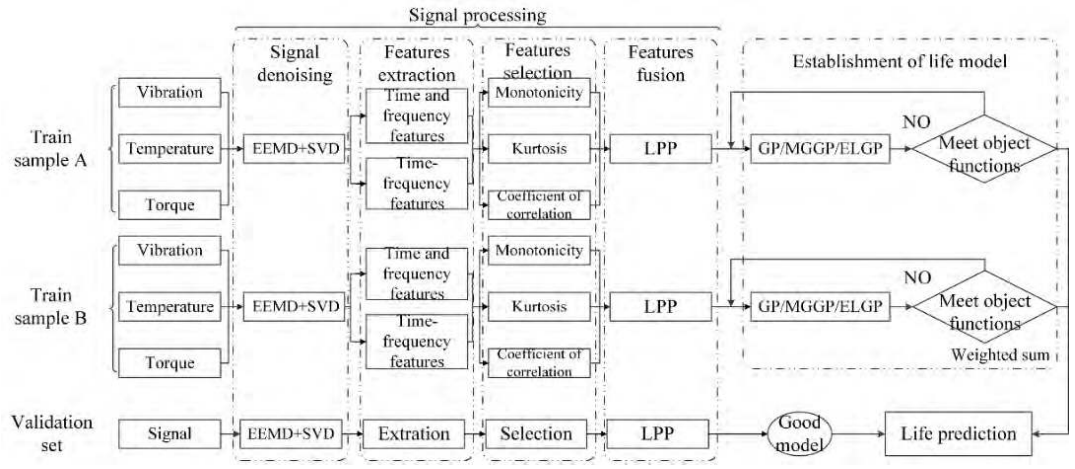


**FIGURE 5.** The whole process for the verification of proposed methods.

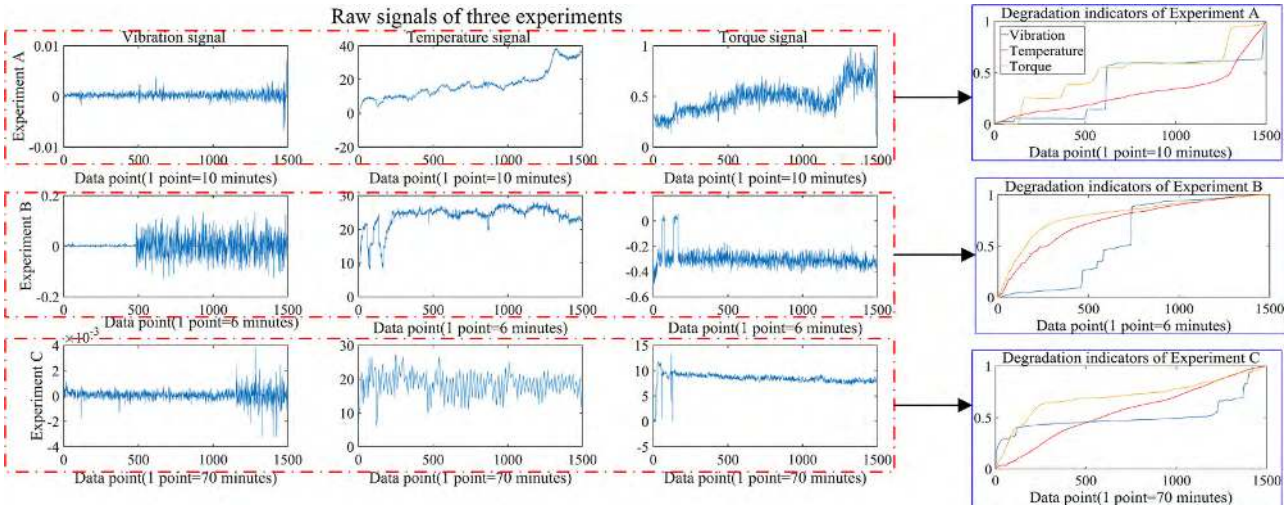


**FIGURE 6.** Raw signal and DIs of three experiments.

EEMD-SVD based de-noising algorithms. Compared with other traditional de-noising algorithms (wavelet denoising, EMD denoising, SVD denoising, etc.), EEMD-SVD can denoise signal at different frequencies with adaptive number of singular values determining by difference spectrum.

Table 3 is a simple comparison with wavelet de-noising algorithm based on Experiment A and two judging criteria: signal-to-noise ratio (SNR) and mean square error (MSE) are given. It is clearly that the performance of EEMD-SVD is better than wavelet algorithm in signal denoising and few parameters in the algorithm need to be predefined.

Three domain features are extracted from original signal. From our signal processing strategy, not all the features are suitable for the next processing. Features selection based on the metrics discussed in Section 2 will be implemented. For all features, we first choose the ones with obvious trend, excluding those who continue to float. Then, we select the time-frequency features based on the correlation and kurtosis

**TABLE 3.** Comparison between EEMD-SVD and wavelet(db4).

| Denoising algorithms | Criteria | Vibration | Temperature | Torque |
|---|---|---|---|---|
| EEMD-SVD | SNR | 0.5197 | 23.2934 | 7.4237 |
| | MSE | 6.1356e-04 | 0.5679 | 0.0594 |
| Wavelet(db4) | SNR | 0.6255 | 21.1485 | 7.2976 |
| | MSE | 6.2446e-04 | 0.7271 | 0.0610 |

criteria and choose features with high similarity and kurtosis. There is no uniform standard to indicate which indicators are most responsive to the degradation trend of slewing bearings. The intrinsic dimension of features here is determined by packing numbers [35], it is calculated that all intrinsic dimensions of high dimension feature sets are close to one, so intrinsic dimension is set to one. Then we adapt features fusion via LPP algorithm to represent degradation trends of

the three slewing bearings. In the algorithm, we define the range of the nearest neighbor is bigger than the intrinsic dimension and smaller than the number of features [36]. Therefore, six nodes are taken in the adjacency graph and parameter $\beta$ in the heat kernel above is set to 0.05. After fusion, DIs shown in Figure 6 are normalized to remove statistical error in repeated data and make subsequent symbol regression algorithms easier to converge.

From Figure 6, we can find health conditions of the same slewing bearing varies greatly under different working conditions. Therefore, it is not easy to predict different types of slewing bearings. Usually, only one set of data including training sets and validation sets is used in many papers which study bearings' or slewing bearings' life prediction. The results of those literatures are usually ideal because the training set and the validation set are derived from a set of data and even have an overlapping. It is unreasonable to verify life models with data from training sets. In our study, we propose a novel method based on symbolic regression and can obtain life model expressions directly. Through analysis of the structure of life models under different working conditions, mapping relationships between DIs and RUL can be found.

### 2) MODELING STRATEGIES AND HYPOTHESIS

Two modeling strategies are applied: One is to establish a life model for each degradation indicator, we assume that the contribution of each degradation indicator to RUL is same because of the uncertainty of quantitative effect of each indicator to RUL, hence final model can be obtained by combining three models corresponding with each degradation indicator using weighted average. We call this method independent modeling ignoring the interaction between signals. Another is hybrid modeling which establishes life models with three DIs simultaneously. This hybrid modeling strategy considers the interaction between diverse signal under the pressure of natural selection of GPs. The two strategies are all based on three GP algorithms talked above.

We believe that although the overall trend of degradation is different but major factors mapping to RUL and the local structure of factors are fixed. This local mapping relationship is magnified or reduced due to differences in operating conditions. Another hypothesis is that if we can get enough life models under diverse situations. A reasonable combination of them, which will greatly improve the accuracy of forecast. It is worth noting that we both apply weighted average and similarity based methods for a quantitative comparison of precision. When the sample size is large, we do recommend the weights based on the similarity between samples which consists of working condition, loading process, sample size, degenerate trajectories after signal processing and so on. Based on the two above hypotheses and GPs algorithms, life model expressions of full life cycle from Experiments A and B are first obtained to test Experiment C. In order to study the life models conveniently, simple function sets showing in Table 4 are applied in this paper and the rest of GPs' parameter settings can be found above.

**TABLE 4. Parameter settings for GPs.**

| Parameters | GP/MGGP | ELGP |
|---|---|---|
| Population size | 2000 | 2000 |
| Max generations | 10000 | 10000 |
| Function sets | {+, -, ×, /, ^, sin, cos, exp, log} | {+, -, ×, /, ^, sin, cos, exp, log} |
| Selection strategy | Tournament | Age-fitness Pareto [37] |
| Inputs | DIs | |
| Outputs | RUL | |

**TABLE 5. Performance of GPs under different modeling strategies.**

| Metric | $R^2$ | | RMSE | |
|---|---|---|---|---|
| Strategies | Independent | Hybrid | Independent | Hybrid |
| GP | 0.7210 | 0.8657 | 227.9669 | 145.9360 |
| M4GGP | 0.8193 | 0.8557 | 183.4605 | 163.9290 |
| M8GGP | 0.8284 | 0.7671 | 178.7517 | 175.1026 |
| M12GGP | 0.8181 | 0.7980 | 184.0529 | 193.9763 |
| ELGP | 0.7925 | 0.9009 | 196.5773 | 135.8817 |

### 3) VALIDATION AND DISCUSSION

In this subsection, we do predictions for three existing experiments. First, raw signal and three DIs from Experiment C presented in Figure 6 are used for verification. The degradation of large-size slewing bearings is quite different from small ones, especially in temperature and torque signal. Based on the second hypothesis we proposed before, we combine the two models of the full life cycle from Experiment A and B with the same weights and two different modeling strategies (hybrid and independent) are conducted for verification.

It is intuitively to discovery that the hybrid modeling method is better than the independent modeling method depicted in Figure 7 clearly. Residual useful life prediction of hybrid modeling strategy considering the interaction under the pressure of fitness function are more precise than independent ones. Also, when the prediction time goes on, the hybrid models are closer to real RUL and ELGP model performs best. To quantify the performances between modeling strategies and GPs, Table 5 are listed below.

In general, the performances of the independent models are worse than hybrid models. The only difference between the two modeling methods is the interaction between different signal components. In the independent models, life model is established by three DIs independently and the interaction between the signals is not considered. The hybrid models considering the coupling between the different signals are more precise. Inspired by this detection, we infer different signals are interactive. That means when the temperature increases, the vibration of the slewing bearing may be more severe. If the coupling between signals can be understood and considered in the prognosis, RUL prediction may be more

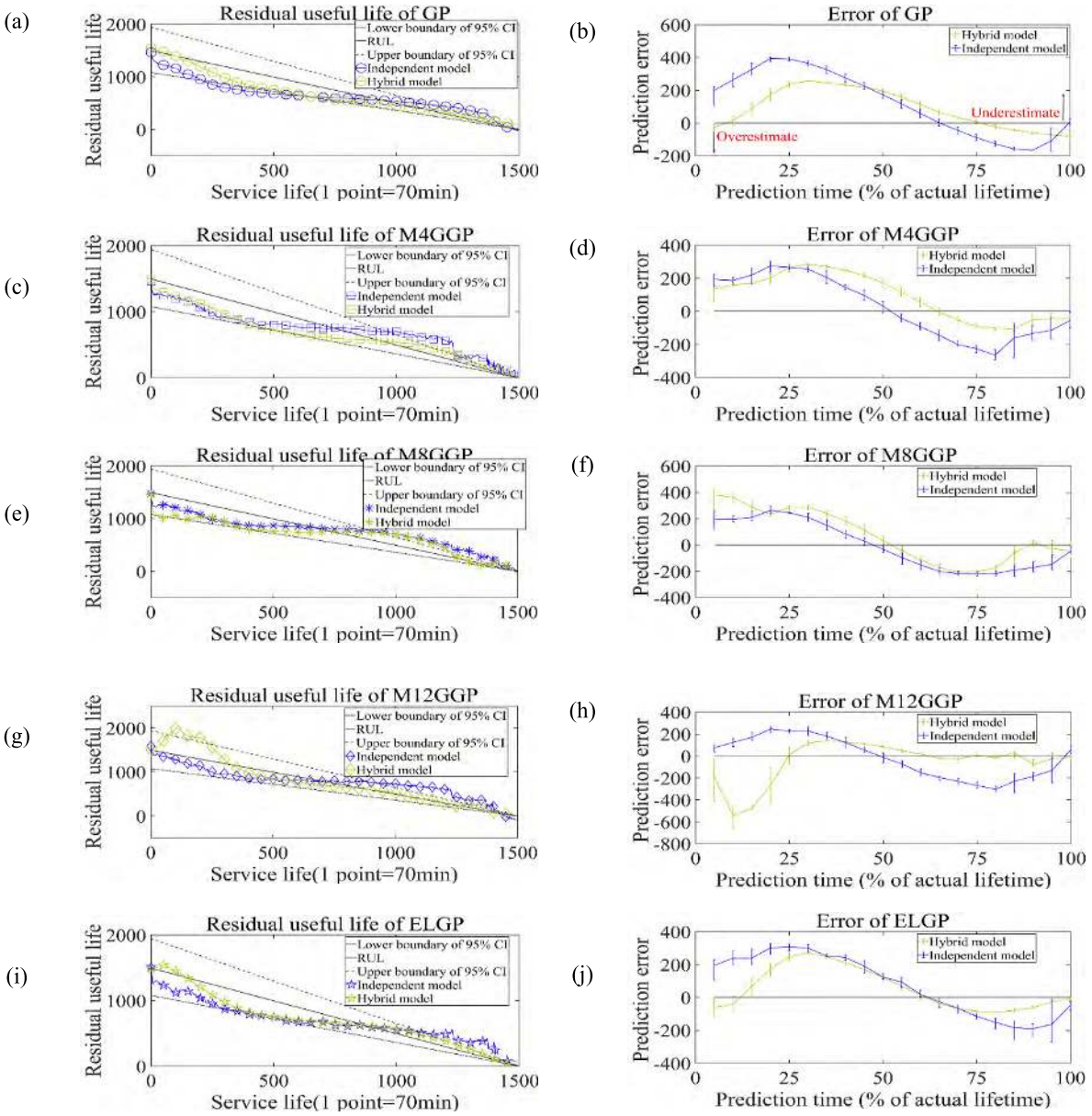


**FIGURE 7.** Predictions and errors under two modeling strategies: (a), (c), (e), (g) and (i) are RUL prediction of GP, M4GGP, M8GGP, M12GGP and ELGP respectively; (b), (d), (f), (h) and (j) are RUL prediction error of GP, M4GGP, M8GGP, M12GGP and ELGP respectively.

precise than before. This is also confirmed from the side of this article using a variety of signals to establish the life model is reasonable. In hybrid models, prediction precision of ELGP is obviously higher than that of the other two and the expressions of models are more succinct and compact than MGGPs owing to the epigenetic layer and the multi-objective optimization mechanism. Models from GP, M4GGP are better than M8GGP and M12GGP and the performance of the training set is getting better as genes increase, indicating that there exists overfitting. Under hybrid modeling strategy, GP models discards the vibration components in the evolution of the population, it's unreasonable to establish models only with temperature and torque components because vibration index is one of important components [7], [14], [15] reflecting the degradation processing for slewing bearings. Therefore, the performance of GP models in Table 5 is worse than ELGP models.

To further verify the effectiveness and improve the precision of the proposed methods, data from lab test rig will also be verified under ELGP algorithm and hybrid modeling strategy. In the attempts we find that it doesn't perform well when we directly weight average life model expressions from full life cycle. Three factors may lead to the poor generalization we consider: (1) Individual health of slewing bearings varies greatly which can be found from the raw signal of three experiments in this research and historical run-to-failed experiment data is scarce; (2) The ability for representation of single symbolic regression based model is limited, it means

**TABLE 6.** Life model expressions of experiment A from lab test rig.

| Stage | Life models of Experiment A | |
|---|---|---|
| I | $RUL(v,t,n)= 4.38(\cos(\exp(((\sin(n)-t)+4.48)))-0.26)-n+1470.3\cos(\sin((4.77t)))$ | (9-1) |
| II | $RUL(v,t,n) = v+((10.42\cos(58.78\sin(\sin(t))))+(v+(\exp(((\sin(\exp(n))\times(-3.07t))+7.46149))+v)))$ | (9-2) |
| III | $RUL(v,t,n) = 4.25\sin(6.80(27.45t))+\exp((\sin(8.77((n-t)-0.19))+5.67))-v$ | (9-3) |
| IV | $RUL(v,t,n) = v-\cos(v)+(\exp(((3.86218-t)\times\sin((-7.78n))))+287.75)\times\cos((-1.59t))$ | (9-4) |

**TABLE 7.** Life model expressions of experiment B from lab test rig.

| Stage | Life models of Experiment B | |
|---|---|---|
| I | $RUL(v,t,n)= \cos(32.13n)+\exp(7.68-\exp(v+\cos(n+3.30)))-0.07$ | (10-1) |
| II | $RUL(v,t,n) = 20.82\cos(-0.22+\exp(\exp(t)))+t+(-0.077-(\exp(\exp(n+1.1)-\sin(\cos(n)))-\exp(7.27)))$ | (10-2) |
| III | $RUL(v,t,n) = 1.79326+(2.33(-4.63\sin((-35.18(v+t)))))+1248.88\sin(3.03t)$ | (10-3) |
| IV | $RUL(v,t,n) = 2.13+8.52(\cos(\exp((4.51n)))\times n)+\exp(6.79)-\exp(n\times(5.76+n))$ | (10-4) |

**TABLE 8.** Life model expressions of experiment C from industrial test rig.

| Stage | Life models of Experiment C | |
|---|---|---|
| I | $RUL(v,t,n) = (212.09(v\times\sin(\exp(n))))+\exp(\cos(n)+6.29)+72.48\sin(-11.44t)$ | (11-1) |
| II | $RUL(v,t,n)= v+(v+\exp((\cos(\sin(((0.91-n)+((t-1)\times(10.81v)))))+6.08)))$ | (11-2) |
| III | $RUL(v,t,n) = t-(v-(\exp(v-(\sin(n)\times(-12.15\cos(t))))-(654.294-174.85t)))$ | (11-3) |
| IV | $RUL(v,t,n) = 57.6\cos((n+143.67)\times\cos(v))-\exp(5n)+281.75+v-1.45-\exp(3.74t)$ | (11-4) |

that descriptions and explanations of a complex degradation from slewing bearings' full life cycle are hard to obtain; (3) Combination strategy among training samples need to be adjusted according to the characteristics of experiment data. Take a careful observation of DIs shown in Figure 6, vibration, temperature and torque indicators vary evidently near the 250$^{th}$, 750$^{th}$ and 1250$^{th}$ points. Great changes of degradation processes before and after these points are likely to occur. Mapping relationships of potential functions corresponding to different degradation stages are obviously different. Therefore, multi-stage modeling is helpful to improve the representational ability of a single model and reduce over-fitting. Meanwhile, it is clearly that big data techniques achieve tremendous improvements in data-driven based fault diagnosis and RUL prediction which results in part from the complete sample learning space that can cover verification sets sufficiently. In predictions of rare learning samples, to select learning samples more related to validation sets can improve generalization ability of data-driven model. Hence, we attempt to piecewise modeling at four different stages according to these sudden change points and combine life models by taking the similarity between training samples and testing samples into account. Table 6-8 lists life model expressions from lab and industrial test rigs below. It is noted that $v$, $t$, and $n$ represents vibration, temperature and torque DIs generated by LPP.

**TABLE 9.** Performances of the proposed methods and linear regression models on three experiments.

| Methods | Experiment | $R^2$ | RMSE |
|---|---|---|---|
| Proposed method | A | 0.8030(0.7107) | 191.5651(232.1260) |
| | B | 0.8412(0.6796) | 171.9573(244.2849) |
| | C | 0.9216(0.9178) | 120.8763(123.7438) |

Due to the large individual difference between slewing bearings, this combination strategy takes into account the similarity among three experiments and it is determined based on the Euclidean distance of degradation trajectories. A set of cross validations based on the proposed approaches under the same modeling strategies are shown below:

Figure 8 and Table 9 clearly depict RUL prediction and detailed performances under cross validations. It is noted that values in brackets of Table 9 represents weight-average combination methods. This piecewise modeling and combination methods considering the differences between samples can effectively obtain more accurate predictions of three experiments within huge differences compared with weighted

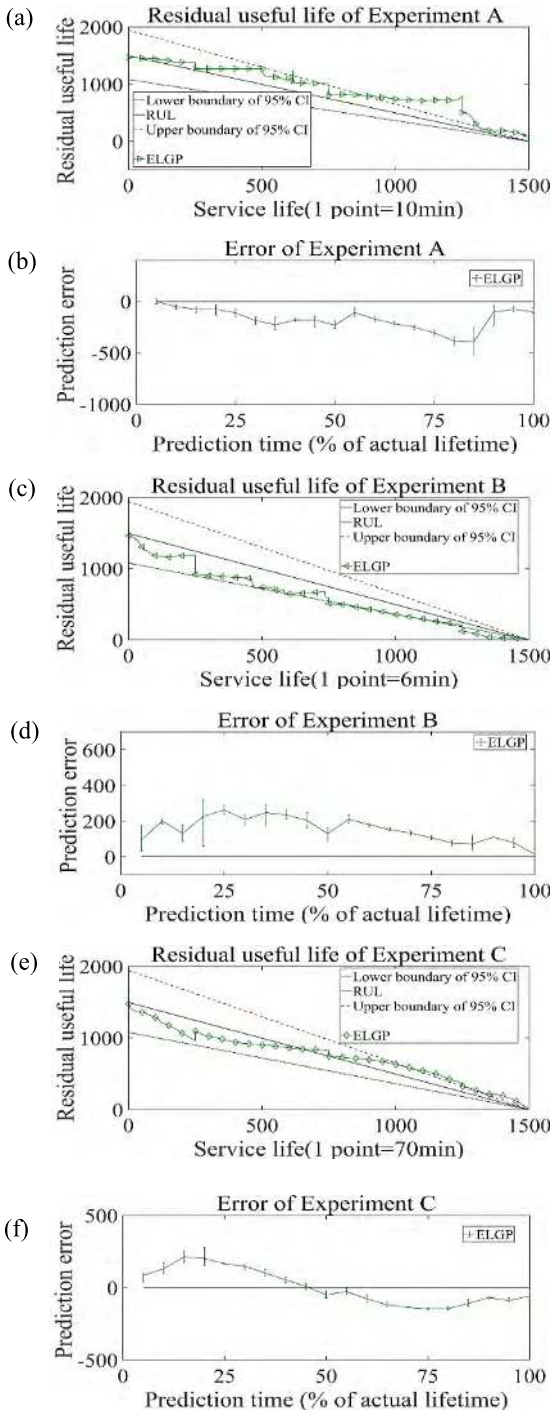


**FIGURE 8.** Cross validations of the lab and industrial test systems: (a), (c) and (e) are RUL prediction of Experiment A, B and C; (b), (d) and (f) are RUL prediction error of Experiment A, B and C.

average modeling method of the full life cycle talked before. Cross validations can greatly prove the effectiveness of our prediction strategies for slewing bearings' RUL prediction.

#### 4) INTERPRETABILITY OF SYMBOLIC REGRESSION BASED LIFE MODEL

Next, life model structures will be discussed in depth. Thanks to the merits of ELGP in the compactness and accuracy of life modeling, we delve into hybrid life models of ELGP under the full life cycle and different life stages. To observe the model structure conveniently, we ignore the parameter size only considering the impact of the model structure. Replace all parameters from the three hybrid models of Experiment A, B and C's full life cycle into c as shown in Eq. (12-14), as shown at the bottom of the next page.

Temperature and torque components are almost in the exponential form and appear at higher frequency which is consistent with expressions at diverse life stages shown in Table 6-8 (since GP is a random search algorithm, we do not consider inconsistencies among the small structures here and the exponential form above is a common result of several attempts) proving the first hypothesis of this paper. It also means that the two indicators could be more responsive to RUL. Occurrences of vibration components from expressions are less than temperature and torque and vibration components are discarded in GP hybrid modeling processing although this abandonment will lead to a decrease in accuracy. We suppose that the sensitivity of the vibration signal reflecting degradation performance of slewing bearing is not high. This means even when the slewing bearings' vibration is very violent, it is not proper to judge severe damage of slewing bearing. Compare the two experiments from the same test system, it is obviously that Experiment B in August when the ambient temperature is much higher than June lasted about 152 hours which is far less than duration time in June. We infer this discrepancy is caused by loading process and external temperature. The greater the load is and the more energy it produces inside the slewing bearing, Therefore, temperature of the slewing bearing will rise. In addition, from the physical point of view, when temperature of slewing bearings in service rises to the degree of the melting of the grease, the lubrication become very poor. Hence, friction between raceways is aggravated and it makes the driving torque increase violently. If this situation continues, the slewing bearing is very likely to fail. Combining the duration time of the same slewing bearing's experiment talked before, we speculate that it is more accurate to judge the damage degree of slewing bearing by observing the change of temperature and torque indicators than vibration indicator only. Overall, temperature and torque indicators may be more sensitive than vibration indictor when RUL of the slewing bearing is estimated and the life model using three indicators is more precise which will be verified in the next subsection. In the view of other institutions' researches [38], [39], experts in the domain of slewing bearings' condition monitoring agree with the viewpoint that vibration signal is too weak and not sensitive at low speed operations accompanied by strong noise which proves the interpretability of the life expressions from our proposed methods. Expressions obtained by symbolic regression methods allow researchers to clearly see the mapping relationship between DIs and RUL of slewing bearings, which are incomparable to those of the traditional ''black box'' algorithms. However, more precise predictions rely on more experiment data to some extent. If enough

**TABLE 10. Performance of diverse signal based on ELGP.**

| Diverse signal | $R^2$ | RMSE |
|---|---|---|
| vibration | 0.4078 | 332.1131 |
| temperature | 0.8224 | 181.8700 |
| torque | 0.8249 | 180.5737 |
| vib and tem | 0.8348 | 175.3901 |
| vib and tor | out of range | 663.8295 |
| tem and tor | 0.8579 | 162.7566 |
| vib, tem and tor | 0.9009 | 135.8817 |

**TABLE 11. Performances of diverse DIs under hybrid modeling strategies.**

| DIs | $R^2$ | RMSE |
|---|---|---|
| PCA | 0.7676 | 208.0478 |
| LPPMD | 0.7951 | 195.3380 |
| LPP | 0.9009 | 135.8817 |

samples are available, accuracy of RUL prediction will be greatly improved. In general, we consider this study has some enlightening significance, practical application value and prediction results and methods are relatively good ones compared with our previous works [7], [19].

### 5) COMPARISONS UNDER DIFFERENT SITUATION

This subsection presents three comparisons between different signals, DIs and other life models to deeply and comprehensively explain the superiority of the proposed methods.

Vibration is the traditional signal for condition monitoring and life prediction of mechanical parts. Experts in the field of slewing bearings utilize more than vibration signal [14]–[16] for their life prediction or fault diagnosis. We suppose that failure information will be lost if vibration signal is used alone for condition monitoring of slewing bearings and temperature or torque signals are more representative of degeneracy of slewing bearings. Comparative results for RUL prediction of Experiment C among three indicators are presented in Figure 9 and Table 10.

From Figure 9 and Table 10: (1) multiple signals based life prediction can greatly improve the accuracy of slewing bearings; (2) In the predictions of single signal, temperature and torque signals performs much better than vibration; (3) Life models considering both temperature and torque signals are superior to the others.

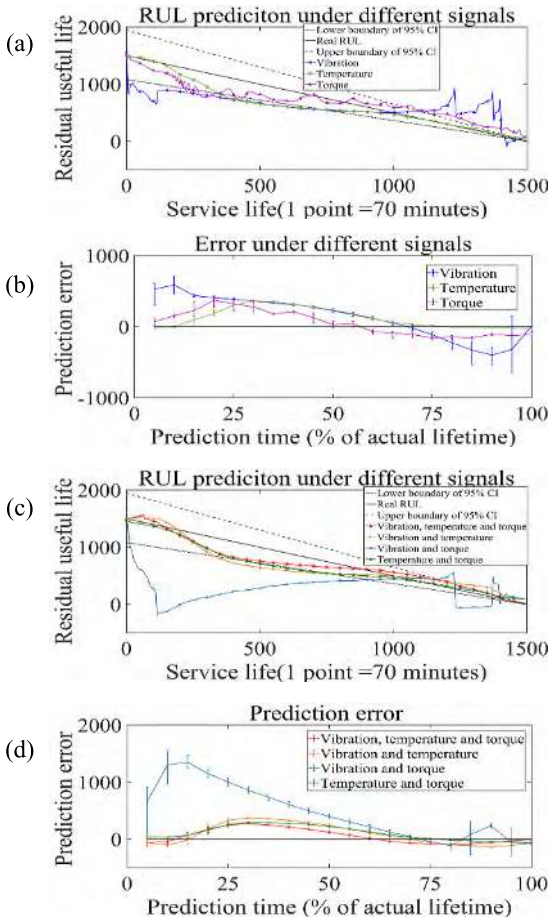


**FIGURE 9. Prediction and error under diverse DIs: (a) and (b) are RUL prediction and error for single signal driven methods; (c) and (d) are RUL prediction and error for multi signal driven methods.**

Next, to inspect the goodness of LPP algorithm for features fusion, a comparison between principal component analysis (PCA) and an improved LPP algorithm based on mahalanobis distance (LPPMD) [40] with the same predefined parameters of LPP will be presented as following and corresponding DIs are depicted in Figure 10 and Figure 11. It is note that the first principal components with contribution value greater than 85% are designed as DIs.

RUL prediction and error between the three kinds of DIs are described in Figure 12 and a detailed comparison of performances are listed in Table 11 below. It is clearly that LPP based DIs outperform the other two ones. In reference [40], manifold learning algorithms improved by mahalanobis

$$RUL(v,t,n) = sin(c/t) + ((cos(((cos(exp(exp(cos((t+c)))))v)/sin(c)))(((v/t)\exp(c))n)) + \boxed{\exp((\cos((ct))cos(sin(\sin(n))))(c\cos(t))} \quad (12)$$

$$RUL(v,t,n) = cos(c(ccos(n))) - (ccos(((cos(t)(t2v)) - (n + exp(exp(n)))))v + c + (cos(cos(c) + \boxed{\exp((\log(\cos(\cos(\cos(\sin(c)+t))) + \log(\cos((t^2)))n)n) - c)} \quad (13)$$

$$RUL(v,t,n) = ((cos(cv) \times exp(c)) + -n) + \boxed{\exp(ccos(tn)) - n)} + \boxed{\exp(c(n-c)))))} \quad (14)$$

**TABLE 12.** Prediction performances of cross validations between SVM, GPR.

| Metric | $R^2$ | | | RMSE | | |
|---|---|---|---|---|---|---|
| Experiment | A | B | C | A | B | C |
| SVM (Quadratic) | Out of range | 0.38 | 0.82 | 465.49 | 339.86 | 179.79 |
| SVM (Cubic) | Out of range | Out of range | Out of range | 876.52 | 815.19 | 486.53 |
| SVM (Fine Gaussian) | 0.64 | 0.78 | 0.45 | 260.34 | 200.69 | 318.71 |
| SVM (Medium Gaussian) | 0.63 | 0.53 | 0.82 | 263.38 | 296.62 | 180.00 |
| SVM (Coarse Gaussian) | Out of range | Out of range | 0.95 | 532.01 | 431.75 | 97.73 |
| GPR (Rational Quadratic) | 0.66 | 0.69 | 0.84 | 253.33 | 241.59 | 171.52 |
| GPR (Squared Exponential) | 0.66 | 0.85 | 0.43 | 252.92 | 168.41 | 327.01 |
| GPR (Matern 5/2) | 0.70 | 0.79 | 0.71 | 238.16 | 198.01 | 230.79 |
| GPR (Exponential) | 0.62 | 0.60 | 0.88 | 267.24 | 271.83 | 152.24 |

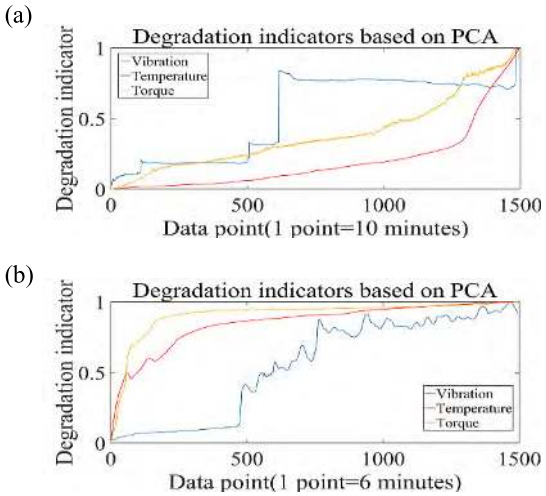


**FIGURE 10.** DIs of PCA: (a) Experiment A; (b) Experiment B.

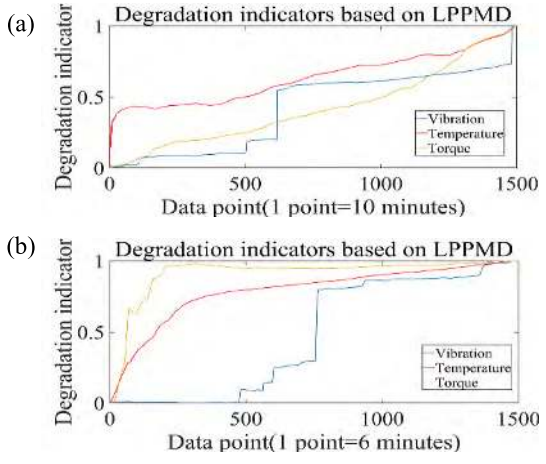


**FIGURE 11.** DIs of LPPMD: (a) Experiment A; (b) Experiment B.

distance obtain good results of fault diagnosis. However, the RUL prediction of slewing bearings is undesirable.

The third comparison with SVM and gauss process regression (GPR) are presented in Table 12. In order to ensure the reliability of the comparison, signal processing methods and

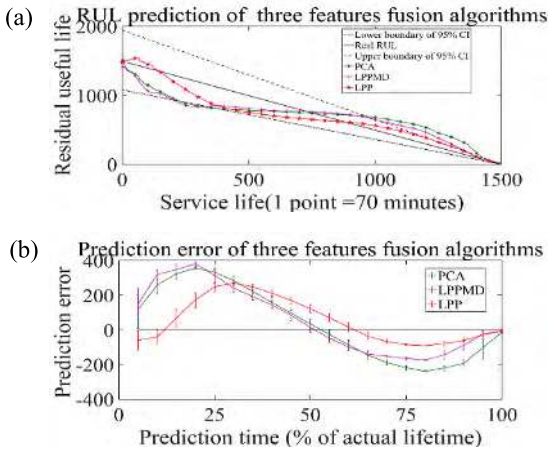


**FIGURE 12.** Prediction and error based on different DIs: (a) RUL prediction based on three kinds of DIs; (b) RUL predictionerror based on three kinds of DIs.

modeling strategies are consistent with our proposed methods. Also, we utilize different kernel functions to mitigate the influence of model preset parameters.

SVM and GPR both have difficulty in the selection of optimal kernel functions and other hyperparameters. The performances of life models vary enormously along with the change of validation sets. It's hard to determine the correct kernel function or model structure for specific situations. These models are unintelligible and difficult in determination of kernel function or model structure. This uncertainty often makes the prediction results instable and unreliable. Although there also exists predefined parameters in GPs, they have little influence on the generalization of prediction models after several attempts with different parameters.

From Table 12, Only two situations (SVM with coarse gaussian kernel function and GPR with squared exponential kernel function) perform better than the proposed life model expressions. The "black box" algorithms do perform well in many cases including rolling bearings' and slewing bearings' life prediction. However, we consider it is uneasy to figure out the true reason to choose the correct predefined model structure or parameters and to understand the real degradation process. That's the weakness of the so-called "black-box"

algorithms. Moreover, we do not completely deny all of such methods but one thing is worth affirming: ''black box'' based life models are so unintelligible and lack physical meanings that it is difficult to understand the quantitative relationship between degradation process and indicators. It hinders people from truly obtaining the most sensitive degradation indicator and relatively less important indicator to RUL. We utilize symbolic regression based method which is never considered before and life model established by ELGP in hybrid modeling strategy can obtain a relatively good prediction and explain the degradation process to some extent which is the aspect that ''black-box'' algorithms can't reach.

Via the three groups of comparisons above, Appropriate DIs are of great importance to life prediction of slewing bearings, our study applies three-dimension based indicators which can extract fault information to some extent and LPP based feature fusion method greatly preserves the local structure of high dimensional features and outperforms the others in the comparison. We think these are half of the contribution to the ideal prediction results. Symbolic regression based modeling method directly obtains life model expressions and relationship between indicators and RUL can be analyzed. Although some ''black box'' algorithms are better than ELGP in precision of prediction, we think it is unsuitable to focus on precision only, physical meaning and interpretability for life models or functional relationships among them are of equal importance. This is the biggest contribution of our research.

## V. CONCLUSIONS

This paper presented a complete set of signal processing methods based on multiple signals. Then reasonable DIs were achieved to describe health condition. We utilized symbolic regression methods to establish life model expressions of slewing bearings based on DIs and analyzed the life models by comparing models under different working conditions. It was a big innovation in the domain of life prediction and significantly outperformed other conventional ''black box'' algorithms. At last good predictive results were presented mainly owing to comprehensive signal processing strategies and ELGP models under reasonable modeling strategies including hybrid and piecewise modeling and similarity based combination strategies.

We find (1) there was an interaction between different signals for life model of slewing bearings after compare independent models and hybrid models. Life models considering interaction between different signals were more precise than opposite ones. But quantitative description of this interaction was not discussed here; (2) In the performance comparisons, ELGP under hybrid and piecewise modeling and similarity based combination strategies can obtain more compact model expressions and better predictive results than others in establishing complex nonlinear models as shown in real-measured experiments; (3) Temperature and torque components were usually in exponential form and we speculated that the two indicators were more sensitive and effective than vibration components by analyzing life model structures from three experiments and comparative experiments.

In the future, an on-line dynamic detection and prediction system which the training sample library is updated online and a semi-empirical life model expression through massive symbolic regression based life model expressions will be established. In addition, the following key points are advised to study in detail: (1) Relatively fixed functional relations between RUL and DIs under different working conditions; (2) Coupling interactions among signals for slewing bearing life prediction; (3) High-precision health evaluation under small samples and variable working conditions.

## REFERENCES

[1] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, ''Machinery health prognostics: A systematic review from data acquisition to RUL prediction,'' *Mech. Syst. Signal Process.*, vol. 104, pp. 799–834, May 2018.
[2] R. Kunc and I. Prebil, ''Numerical determination of carrying capacity of large rolling bearings,'' *J. Mater. Process. Technol.*, vols. 155–156, pp. 1696–1703, Nov. 2004.
[3] R. Kunc, A. erovnik, and I. Prebil, ''Verification of numerical determination of carrying capacity of large rolling bearings with hardened raceway,'' *Int. J. Fatigue*, vol. 29, nos. 9–11, pp. 1913–1919, 2007.
[4] J. Zhu, N. Chen, and W. Peng, ''Estimation of bearing remaining useful life based on multiscale convolutional neural network,'' *IEEE Trans. Ind. Electron.*, vol. 66, no. 4, pp. 3208–3216, Apr. 2019.
[5] Y. Feng, X. Huang, J. Chen, H. Wang, and R. Hong, ''Reliability-based residual life prediction of large-size low-speed slewing bearings,'' *Mech. Mach. Theory*, vol. 81, pp. 94–106, Nov. 2014.
[6] H.-E. Kim, A. C. C. Tan, J. Mathew, and B. K. Choi, ''Bearing fault prognosis based on health state probability estimation,'' *Expert Syst. Appl.*, vol. 39, no. 5, pp. 5200–5213, 2012.
[7] C. Lu, J. Chen, R. Hong, Y. Feng, and Y. Li, ''Degradation trend estimation of slewing bearing based on LSSVM model,'' *Mech. Syst. Signal Process.*, vols. 76–77, pp. 353–366, Aug. 2016.
[8] L. Guo, N. Li, F. Jia, Y. Lei, and J. Lin, ''A recurrent neural network based health indicator for remaining useful life prediction of bearings,'' *Neurocomputing*, vol. 240, pp. 98–109, May 2017.
[9] G. Wang, Z. He, X. Chen, and Y. Lai, ''Basic research on machinery fault diagnosis—What is the prescription,'' *Chin. J. Mech. Eng.*, vol. 49, no. 1, pp. 63–72, 2013.
[10] W. La Cava, K. Danai, L. Spector, P. Fleming, A. Wright, and M. Lackner, ''Automatic identification of wind turbine models using evolutionary multiobjective optimization,'' *Renew. Energy*, vol. 87, pp. 892–902, Mar. 2016.
[11] R. Liu, B. Yang, E. Zio, and X. Chen, ''Artificial intelligence for fault diagnosis of rotating machinery: A review,'' *Mech. Syst. Signal Process.*, vol. 108, pp. 33–47, Aug. 2018.
[12] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni, and G. Tripot, ''A data-driven failure prognostics method based on mixture of Gaussians hidden Markov models,'' *IEEE Trans. Rel.*, vol. 61, no. 2, pp. 491–503, Jun. 2012.
[13] L. Liao, ''Discovering prognostic features using genetic programming in remaining useful life prediction,'' *IEEE Trans. Ind. Electron.*, vol. 61, no. 5, pp. 2464–2472, May 2014.
[14] W. Caesarendra, P. B. Kosasih, A. K. Tieu, A. S. Moodie, and B.-K. Choi, ''Condition monitoring of naturally damaged slow speed slewing bearing based on ensemble empirical mode decomposition,'' *J. Mech. Sci. Technol.*, vol. 27, no. 8, pp. 2253–2262, 2013.
[15] M. Z. S. Â¡vokelj and I. Prebil, ''EEMD-based multiscale ICA method for slewing bearing fault detection and diagnosis,'' *J. Sound Vib.*, vol. 370, pp. 394–423, May 2016.
[16] P. Ding, H. Wang, Y. F. Dai, J. Chen, H. Zhang, and F. Z. Sun, ''MDCCS based multistage life prediction of slewing bearing with a novel performance description: An improved variational mode decomposition approach,'' *Exp. Tech.*, vol. 43, no. 3, pp. 341–358, 2019.
[17] P. Ding, H. Wang, and Y. Dai, ''A clustering-based framework for performance degradation prediction of slewing bearing using multiple physical signals,'' *ASCE J. Risk Uncertain. Eng. Syst. B, Mech. Eng.*, vol. 5, no. 2, 2019, Art. no. 020908.

[18] D. A. Augusto and H. J. C. Barbosa, "Symbolic regression via genetic programming," in *Proc. 6th Brazilian Symp. Neural Netw. (SBRN)*, Rio de Janeiro, Brazil, 2000, pp. 173–178.

[19] H. Wang, M. Tang, and X. Huang, "Smart health evaluation of slewing bearing based on multiple-characteristic parameters," *J. Mech. Sci. Technol.*, vol. 28, no. 6, pp. 2089–2097, 2014.

[20] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mech. Syst. Signal Process.*, vol. 35, nos. 1–2, pp. 108–126, Feb. 2013.

[21] E.-L. Chen, X. Zhang, Y.-J. Shen, and X.-M. Cao, "Fault diagnosis of rolling bearings based on SVD denoising and blind signals separation," *J. Vib. Shock*, vol. 31, no. 23, pp. 185–190, 2012.

[22] X. Zhao and B. Ye, "Selection of effective singular values using difference spectrum and its application to fault diagnosis of headstock," *Mech. Syst. Signal Process.*, vol. 25, no. 5, pp. 1617–1631, 2011.

[23] L. Guo, Y. Lei, N. Li, T. Yan, and N. Li, "Machinery health indicator construction based on convolutional neural networks considering trend burr," *Neurocomputing*, vol. 292, pp. 142–150, May 2018.

[24] X. Chen, Z. Shen, Z. J. He, C. Sun, and Z. W. Liu, "Remaining life prognostics of rolling bearing based on relative features and multivariable support vector machine," *Proc. Inst. Mech. Eng. C, J. Eng. Mech. Eng. Sci.*, vol. 227, no. 12, pp. 2849–2860, 2013.

[25] X. He and P. Niyogi, "Locality preserving projections," in *Proc. Adv. Neural Inf. Process. Syst.*, 2002, vol. 16, no. 1, pp. 186–197.

[26] Y. Wang, G. Xu, L. Liang, and K. Jiang, "Detection of weak transient signals based on wavelet packet transform and manifold learning for rolling element bearing fault diagnosis," *Mech. Syst. Signal Process.*, vols. 54–55, pp. 259–276, Mar. 2015.

[27] J. Yu, "Bearing performance degradation assessment using locality preserving projections and Gaussian mixture models," *Mech. Syst. Signal Process.*, vol. 25, no. 7, pp. 2573–2588, 2011.

[28] J. R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. Cambridge, MA, USA: MIT Press, 1992.

[29] Q. Ouyang and W. Lu, "Monthly rainfall forecasting using echo state networks coupled with data preprocessing methods," *Water Resour. Manag.*, vol. 32, no. 2, pp. 659–674, 2018.

[30] S. J. Hadi and M. Tombul, "Monthly streamflow forecasting using continuous wavelet and multi-gene genetic programming combination," *J. Hydrol.*, vol. 561, pp. 674–687, Jun. 2018.

[31] M. A. Haeri, M. M. Ebadzadeh, and G. Folino, "Statistical genetic programming for symbolic regression," *Appl. Soft Comput.*, vol. 60, pp. 447–469, Nov. 2017.

[32] M. O'Neill, L. Vanneschi, S. Gustafson, and W. Banzhaf, "Open issues in genetic programming," *Genet. Program. Evolvable Mach.*, vol. 11, nos. 3–4, pp. 339–363, 2010.

[33] T. Perkis, "Stack-based genetic programming," in *Proc. 1st IEEE Conf. Evol. Comput.*, Orlando, FL, USA, vol. 1, Jun. 1994, pp. 148–153.

[34] W. La Cava, T. Helmuth, K. Danai, and L. Spector, "Genetic programming with epigenetic local search," in *Proc. 16th Genetic Evol. Comput. Conf. (GECCO)*, Madrid, Spain, 2015, pp. 1055–1062.

[35] B. Kégl, "Intrinsic dimension estimation using packing numbers," in *Proc. 16th Annu. Neural Inf. Process. Syst. Conf.*, Vancouver, BC, Canada, 2002, pp. 697–704.

[36] T. Song, B. P. Tang, and L. I. Feng, "Fault diagnosis method for rotating machinery based on manifold learning and K-nearest neighbor classifier," *J. Vib. Shock*, vol. 32, no. 5, pp. 149–153, 2013.

[37] M. Schmidt and H. Lipson, "Age-fitness Pareto optimization," in *Proc. 12th Annu. Genetic Evol. Comput. Conf.*, New York, NY, USA, 2010, pp. 543–544.

[38] W. Caesarendra, B. Kosasih, A. K. Tieu, and C. A. S. Moodie, "Circular domain features based condition monitoring for low speed slewing bearing," *Mech. Syst. Signal Process.*, vol. 45, no. 1, pp. 114–138, 2014.

[39] Y. H. Kim, A. C. C. Tan, and V. Kosse, "Condition monitoring of low-speed bearings—A review," *Austral. J. Mech. Eng.*, vol. 6, no. 1, pp. 61–68, 2008.

[40] B. Yao, P. Zhen, L. Wu, and Y. Guan, "Rolling element bearing fault diagnosis using improved manifold learning," *IEEE Access*, vol. 5, pp. 6027–6035, 2017.


**PENG DING** received the M.S. degree in mechatronic engineering from Nanjing Tech University, in 2019. His current research interests include signal processing, condition monitoring, life prediction, and statistical modeling.

**QINRONG QIAN** is currently pursuing the M.S. degree in mechanical engineering with Nanjing Tech University. Her current research interest includes the useful life prediction of slewing bearings.

**HUA WANG** received the B.S. and Ph.D. degrees in mechanical engineering from Harbin Engineering University, Harbin, China, in 2001 and 2006, respectively. He has been holding a short-term visiting position to the Dublin Institute of Technology, Ireland, since 2010, and is also as a Visiting Scholar with the Department of Mechanical and Industrial Engineering, University of Massachusetts, USA, from 2014 to 2015. He is currently a Full Professor of mechanical engineering with Nanjing Tech University. His research interests include machinery condition monitoring and fault diagnosis, mechanical signal processing, intelligent fault diagnostics, and remaining useful life prediction.

He is the Senior Member of the China Mechanical Engineering Society. He is also a Reviewer for several SCI-indexed journals in this field.

**JIANYONG YAO** received the B.Tech. degree from the Tianjin University, Tianjin, China, in 2006, and the Ph.D. degree in mechatronics from Beihang University, Bejing, China, in 2012. He was a Visiting Exchange Student with the School of Mechanical Engineering, Purdue University, from 2010 to 2011.

In 2012, he joined the School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing, China, where he is currently as a Full Professor. His current research interests include servo control of mechatronic systems, adaptive and robust control, fault detection, and the accommodation of dynamic systems.