

Received July 22, 2020, accepted July 28, 2020, date of publication August 3, 2020, date of current version August 27, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3013875

Tapered Roller Bearing Failure Diagnosis Based on Improved Probability Box Model

HONG TANG¹, HONGLIANG DAI¹, AND YI DU^{1,2}

¹State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body, Hunan University, Changsha 410083, China

²City College, Kunming University of Science and Technology, Kunming 650051, China

Corresponding author: Yi Du (duyi12844059@sina.cn)

This work was supported in part by the Creative Research Groups Foundation of China under Grant 51621004, in part by the National Natural Science Foundation of China under Grant 51365020, in part by the Hunan Provincial Civil-Military Integration Industry Development under Grant [2018]23, in part by the Open Foundation of China Ship Scientific Research Center (CSSRC) under Grant 702SKL201705, and in part by the State Key Laboratory of Advanced Design Manufacturing for Vehicle Body under Grant 71865009.

ABSTRACT Bearing failure often occurs in rotating machinery. The fault diagnosis method based on the vibration signals has been studied for many years. This paper proposed an improved probability box (ip-box) modeling method for diagnosing bearing faults. The major theoretical principles involved with the probability box (p-box) modeling methods and a projection method. Since a larger aggregated width results in the p-box not being conducive to a fault identification and diagnosis, the mean of the focal element interval and the amount of data fluctuation between the adjacent focal elements were used as additional information. Then, the additional information was added to the ip-box model by the cooperative optimization method. Finally, the experimental results showed that the classification performance of a support vector machine (SVM) trained with eight measured values from the ip-box was significantly improved.

INDEX TERMS Probability box, cooperative optimization, projection method, fault diagnosis, SVM.


I. INTRODUCTION

The fault diagnosis technology of bearings has become an important means and key technology to ensure the safety and stability of production systems in the development of a modern industry [1]. Data obtained in practical engineering contains uncertainty because of restrictions on the working environment or test cost [2]. The existence of the uncertainty limits an application of the traditional probability model for the fault diagnosis technology of bearings [3], [4]. An interval model is one of the common methods to describe the uncertainty of bearing signals, but it can obtain only the range of the bearing data, not its probability distribution, and the statistical information of the bearing data cannot be perfectly used.

Tang *et al.* [5] introduced the theory of the p-box into mechanical fault diagnosis to solve the above problems. The idea of p-boxes was originally put forward to express pure ‘epistemic uncertainty’ with ‘interval’ and has experienced cross research with fuzzy theory [6], evidence theory of Dempster structure [7], Boolean logic reasoning based on traditional probability theory [8], Kolmogorov method with sparse samples [9], etc. The p-box theory not only integrates

random algorithms such as Bayesian reasoning and evidence theory, but also artificial intelligence algorithms such as neural network [10], expert system and fuzzy set theory [11]. For the p-box in an application of the bearing diagnosis, vibration signal is considered as a random variable, and its random distribution type is analyzed to establish the p-box model; then features of the p-box are obtained by measuring the aggregated uncertainty of the p-box to input into SVM model. However, the focal element interval of the p-box may not be the optimal width for each p-box model, which may increase the data crossover between p-boxes, thereby reducing the independence of input data in the SVM model.

Because the interval model is the basis of the p-box theory, the width of the p-box can be significantly reduced by reducing the confidence, but the estimation of parameters will be seriously distorted. In the bearing fault diagnosis, the p-box must be discrete into focal elements. Therefore, there may be data redundancy in the focal elements from the same p-box. However, through nonlinear programming, the phenomenon of partial repetition between focal elements can be reduced. Based on a convolution quadrature method, Fahmy developed a novel numerical modeling optimization technique [12]. The interval evaluation of parametrized functions was defined by Marendet *et al.* [13] adapting

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

numerical constraint programming techniques to quantified inequalities. A piecewise linear approximation method, a derivative algorithm, and a duality-based algorithm are proposed by Zhang and Li [14] for solving interval quadratic programming. Based on nonlinear interval programming and the nested optimization solving strategy, Wang *et al.* [15] proposed an interval uncertain optimization method for artillery structural dynamic responses considering robustness and interval economy. An improved spatial branch-and-bound algorithm was proposed by Zhao *et al.* [16] to efficiently solve the reformulated bilinear programming problem. Besides, Zeng *et al.* [17] proposed a new multiple attribute decision-making method based on the nonlinear programming methodology. Based on the normal fluctuation range of each cyber-physical energy system state, a bilevel nonlinear programming model was proposed by Wang *et al.* [18]. More the nonlinear programming models in [19]–[22]. However, the treatment of the focus element interval is rarely found. With the deepening of research, the nonlinear programming process of the projection method is relatively simple in the interval values and satisfies the accuracy of programming for the p-box focus element.

For the focal element interval correction of the p-box, this paper discretizes the p-box into several focal element intervals and constructs the whole envelope interval of focal element interval sets. Then, the maxima and minima are searched in the focal element interval by an optimization method, and the complete search path is saved. By evaluating the relationship between the subinterval of the focal element and the search path, the crossover point along the negative gradient direction and the bound of the focal element interval is obtained as the initial point of the subinterval search path. Then, the extremum of the focal element interval is solved by the projection method. The method optimizes all the focal element interval sets in a collaborative way, avoiding the repeated calculation of an internal search, reducing the aggregated width of the p-box, and improving the overall calculation efficiency. The mean value of the interval and data fluctuation of the adjacent focal element interval is added as additional information to the p-box, which can further reduce the aggregated width to avoid the overlap between p-box models. Then, the area difference and bound difference between the p-box and ip-box are defined, and the improvement degree of the p-box is analyzed quantitatively. Finally, the validity of the proposed method is demonstrated based on the experiment data. It should be noted that the current method is used to optimize the width of the focal element interval for the p-box model, i.e., ip-box model, using the nonlinear programming process of the projection method.

II. RELATED WORK

A. BASIC NOTIONS OF P-BOX

A signal collected by an accelerometer varies with the time because of the measurement error of the sensor, the different position of the measurement, and the different of the working condition, then the signal can be considered as a random

variable X . A cumulative probability distribution function (CDF) of the random variable X does not be expressed by a single curve, as the estimated value \hat{x} of X is not a single scalar value [23]. Considering i th estimated value $\hat{x}_i \in [\hat{x}_i^-, \hat{x}_i^+]$, the upper and lower bounds of the CDF are given by:

$$\bar{F}(x) = 1 - P(X > x) \quad (1a)$$

$$F(x) = P(X \leq x) \quad (1b)$$

where $\bar{F}(x)$ and $F(x)$ are the upper and lower bounds of the CDF, respectively; P denotes a lower probability measure. $[F(x), \bar{F}(x)]$ is called as the p-box, then the random variable X with the uncertainty is limited in it. For a single scalar value, Eqs. (1a) and (1b) are equivalent, and the CDF can be expressed by a single curve; for an interval, the schematic diagram of CDF can be plotted in Fig. 1.

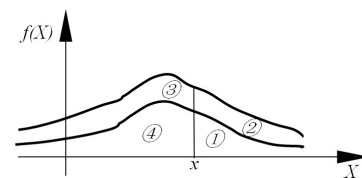


FIGURE 1. Schematic diagram of the p-box definition.

In Fig. 1, the area of ① corresponds to the lower probability measure $P(X > x)$ of Eq. (1a); the area of ④ corresponds to $F(x) = P(X \leq x)$; the area of $1 - \text{①}$ corresponds to $\bar{F}(x) = 1 - P(X > x)$, i.e., the sum of the areas of ②, ③ and ④. The difference between Eqs. (1a) and (1b) is the area of ② and ③, which is not equivalent.

B. P-BOX MODELING METHOD OF BEARING TIME-DOMAIN SIGNAL

The bearing signal has strong nonstationary randomness, which includes not only the irreducible uncertainty which cannot be reduced by further empirical study brought by the accuracy of the equipment (although the equipment may be better machined) but also the epistemic uncertainty brought by the operator during the collection of bearing data, it can generally be reduced by an additional empirical effort [23]. The p-box model can well achieve the fusion of the irreducible and epistemic uncertainty. According to the different forms of bearing signals, three p-box modeling methods were proposed in the previous work [5].

The probability and statistics toolbox of MATLAB is used to analyze whether the bearing data samples satisfy commonly used random distributions. If random distribution type is met, a distribution-type p-box modeling method (DTPMM) can be used. Normal distribution for the bearing data is considered. Then, the p-box can be determined by limiting the random distribution parameter in an interval mean $\mu \in [\mu^{\text{left}}, \mu^{\text{right}}]$ and standard deviation $\sigma \in [\sigma^{\text{lower}}, \sigma^{\text{upper}}]$ with a mass m , which is given by

$$\left\{ \left[\mu_i^{\text{left}}, \sigma_i^{\text{lower}} \right], m_i \right\}, \left(1 \leq i \leq \frac{n}{2} \right) \quad (2a)$$

$$\left\{ \left[\mu_i^{\text{left}}, \sigma_i^{\text{upper}} \right], m_i \right\}, \left(\frac{n}{2} \leq i \leq n \right) \quad (2b)$$

$$\left\{ \left[\mu_i^{\text{right}}, \sigma_i^{\text{lower}} \right], m_i \right\}, \left(1 \leq i \leq \frac{n}{2} \right) \quad (2c)$$

$$\left\{ \left[\mu_i^{\text{right}}, \sigma_i^{\text{upper}} \right], m_i \right\}, \left(\frac{n}{2} \leq i \leq n \right) \quad (2d)$$

And “the averaging discretization method and the outer discretization method can be used to discretize a parameter of the p-box model in any case” [5].

The strong nonstationary randomness of the bearing signal may cause the bearing data samples to not satisfy a random distribution type, so the DTPMM is inapplicable. A dimensionless p-box modeling method (DPMM) can obtain the dimensionless values from the raw bearing data, which avoids the data clutter and improves the regularity. The dimensionless values may satisfy the random distribution type. The skewness features and the kurtosis features are the third and fourth moments of the bearing signals, respectively, which can be expressed as:

$$SK_i = \frac{\frac{1}{n} \sum_{j=1}^n [\alpha_{ij} - \bar{\alpha}_i]^k}{\frac{1}{n} \sum_{j=1}^n \alpha_{ij}^k}, \quad (i = 1, \dots, m), (k = 3, 4) \quad (3)$$

where $SK_i(i = 1, \dots, m)(k = 3)$ and $SK_i(i = 1, \dots, m)(k = 4)$ are the skewness feature and kurtosis features of the i th row vector of bearing signals, respectively, and $\bar{\alpha}_i$ is the mean of the i th row vector. Then, a vector of skewness features and a vector of kurtosis features can be obtained. If these vectors satisfy a random distribution type, the DTPMM can be applied.

The DPMM needs to extract the features from the raw bearing data, which results in information loss in the probability and statistics of the raw bearing data. To preserve the characteristic (i.e., the information of the probability and statistics), a raw-data p-box modeling method (RDPMM) is proposed. The RDPMM can directly establish the p-box model for the raw bearing data based on the definition of the p-box. First, the maxima and minima from each column vector of bearing signals can be obtained by the RDPMM. Then, a Dempster-Shafer structure is obtained based on the maxima and minima vectors and discretized according to the same basic probability assignment. Finally, the upper and lower bounds of the p-box can be approximated by discrete sampling.

III. IMPROVEMENT AND COMPARISON OF P-BOX MODELING METHODS

In this section, we mainly improve the influence of the overlapped interior region of the intervals on the p-box, and then the area difference and bound difference between the p-box and ip-box models are analyzed quantitatively.

A. IP-BOX MODELING METHOD

The aggregated width of the p-box model has a great influence on the result of pattern recognition. In some cases,

a calculated p-box will also be the best possible result in the sense that the bounds can be no tighter without excluding some of the overlapped interior regions of the intervals.

When the wider p-box model is discretized into a large number of focal elements, there are overlap regions between the focal elements, which not only is detrimental to the correct classification rates but also reduces the calculation efficiency [24]. However, although these focal elements overlap each other, they are all included in a Dempster-Shafer structure. Using the gradient project method based on the characteristics, a collaborative optimization method can be proposed to improve the width of the p-box. According to the sampling frequency n , the p-box model can be discretized into a Dempster-Shafer structure with the following expression:

$$\{[\bar{x}_i, \underline{x}_i], m_i\}, (i = 1, 2, \dots, n)$$

where $[\bar{x}_i, \underline{x}_i]$ is the i th focal element interval and $m_i(i = 1, \dots, n)$ is the i th mass value. Here, $[\bar{x}_i, \underline{x}_i]$ can be expressed as A_i . Although there are overlapping regions between A_{i-1} , A_i , and A_{i+1} , they are within the definition domain of a random variable X . Consider a union A_U is related to the focal element interval A by the following expression:

$$A_U = A_1 \cup A_2 \cup \dots \cup A_i \dots \cup A_n \quad (4)$$

where $A_i(1, 2, \dots, n)$ denotes the i th focal element interval. Define the set of enveloping all A_i as an interval envelope A_{envelop} . Then, the A_U must be included in the interval envelope A_{envelop} :

$$A_U \subset A_{\text{envelop}} \quad (5)$$

where the minimum interval envelope A_{envelop} of A_U can be expressed as:

$$A_{\text{envelop}} = ([\min(\bar{x}_1), \max(\underline{x}_1)], \dots, [\min(\bar{x}_i), \max(\underline{x}_i)], \dots, [\min(\bar{x}_n), \max(\underline{x}_n)]) \quad (6)$$

The interval envelope A_{envelop} is obtained by the discretized p-box model, where a large number of overlap regions exist in the intervals [25]. By evaluating the overlap regions between A_i and A_{envelop} , the initial gradient value and gradient direction of the subregions can be determined to reduce the overlap regions and repeated calculation of the interval, as shown in Fig. 2.

Consider Fig. 2(a). Search the feasible direction in the bound $\bar{x} \leq x \leq \underline{x}$ along the negative gradient direction of the distribution functions of p-box $-\nabla F((x_1, x_2)_k)$ based on $\min F_X(x; \Psi)$, where Ψ is a random distribution parameter and x satisfies $\bar{x} \leq x \leq \underline{x}$ [26]. When the bound $x_2 = \underline{x}_2$, $-\nabla F((x_1, x_2)_k)$ is projected on the bound to generate a new search direction d_k . The modeling steps of the method can be expressed as follows:

Step One: Consider x_1 as the initial feasible point, and set $k = 1$.

Step Two: Transform the bound $\bar{x} \leq x \leq \underline{x}$ into the form of inequality $Ax \leq b$, then A and b are respectively decomposed

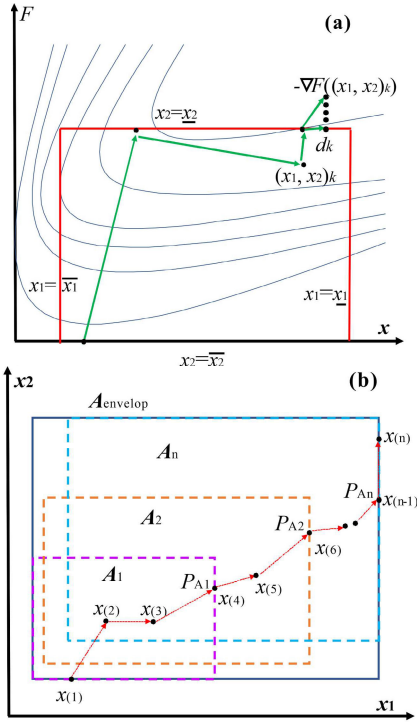


FIGURE 2. The cooperative optimization method. (a) The negative gradient direction of $F_X(x, \Psi)$; (b) the cooperative optimization path.

into $[A_1 \ A_2]^T$ and $[b_1 \ b_2]^T$ at the point x_k , where

$$A_1 x_k = b_1, A_2 x_k > b_2 \quad (7)$$

Step Three: Define $M = A_1$. If M is an empty set, let $P = I$, where I is the unit matrix; otherwise, P must satisfy the following expression:

$$P = I - M^T(MM^T)^{-1}M \quad (8)$$

Step Four: Let $d_k = -P\nabla F(x_k)$. If $d_k \neq 0$, go to step six; if $d_k = 0$, go to step five.

Step Five: If M is an empty set, then stop the calculation, and obtain x_k ; otherwise, obtain u by the following expression:

$$u = (MM^T)^{-1}M\nabla F(x_k) \quad (9)$$

where if $u \geq 0$, then stop the calculation, and obtain a global best point x_k ; if u contains a negative component, then selects a negative component. For example, A_1 is corrected by u_j . Remove the row corresponding to u_j in A_1 , and return to step three.

Step Six: Solve $\min F(x_k + \lambda d_k)$, s.t: $0 \leq \lambda \leq \lambda_{\max}$, where λ_{\max} can be obtained by the following expression:

$$\lambda_{\max} = \begin{cases} \infty, & d' \geq 0 \\ \min \left\{ \frac{b'_i}{d'_i} \mid d'_i < 0 \right\}, & d' < 0 \end{cases} \quad (10)$$

where b' and d' can be expressed as follows:

$$b' = b_2 - A_2 x_k \quad (11a)$$

$$d' = A_2 d_k \quad (11b)$$

and $F_{X_1}(x_1; \Psi_1)$ can be obtained by a search. Then, calculate x_{k+1} by the following expression:

$$x_{k+1} = x_k + \lambda_k d_k \quad (12)$$

Set $k = k + 1$, and return step two.

To understand the above modeling method, consider Fig. 2(b). A_1, A_2 and A_n are obtained from the discretized p-box model. In addition, $x_{(1)}$ is the initial feasible point. Here, $x_{(1)} \sim x_{(n)}$ is the search path, where the direction is the negative gradient direction; $x_{(1)} \sim x_{(n-1)}$ is the search path inside bound, and $x_{(n-1)} \sim x_{(n)}$ is the search path after gradient projection. For A_1, A_2 and A_n , the partial search path of $x_{(1)} \sim x_{(n)}$ is also their internal search path because of the overlapped interior region of the intervals. For example, $x_{(2)} \sim x_{(4)}$ is the internal search path of A_1 , and $x_{(2)} \sim x_{(6)}$ is the internal search path of A_2 . When A_1, A_2 and A_n are searched, filtering the repeated internal search paths and searching directly from different search points will avoid many repeated search calculations. The extreme point of A_1 can be calculated by the gradient project method. Point P_{A1} is the initial point to search directly, which avoids searching interior regions; for the same reason, search A_2 with the point P_{A2} .

The crossover point can be obtained between A_i and the search path $(x_1^{\min}, x_2^{\min}, \dots, x_k^{\min})$ by calculating the minimum value of $F(x)$ in the domain of definition of A_{envelop} and saving $(x_1^{\min}, x_2^{\min}, \dots, x_k^{\min})$ and gradient information. The last crossover point $P_{A_i}^{\min}$ along the search direction is selected as the initial point. Calculate the optimal solution of $\min F_X(x; \Psi)$, s.t: $\bar{x} \leq x \leq \underline{x}$, and assign it to the upper bound of the p-box \bar{F}_i . Similarly, the optimal solution from the negative function $-F(x)$ of $F(x)$ can be obtained and assigned to the lower bound \underline{F}_i .

The bearing signals contain a large amount of statistical information, such as the mean, mode, and variance, which can be used [27]. Adding this information to the p-box model is conducive to the improvement of the area of the p-box model. The mean x_{average} of the bearing data can express the amount of data fluctuation between the adjacent focal elements [28]. The additional information is related to the mean x_{average} by the following expressions:

$$\bar{u} = \left\{ \frac{\bar{x}_2 - \bar{x}_1}{\bar{x}_{\text{average}}}, \frac{\bar{x}_3 - \bar{x}_2}{\bar{x}_{\text{average}}}, \dots, \frac{\bar{x}_i - \bar{x}_{i-1}}{\bar{x}_{\text{average}}}, \dots, \frac{\bar{x}_n - \bar{x}_{n-1}}{\bar{x}_{\text{average}}} \right\} \quad (13a)$$

$$\underline{u} = \left\{ \frac{\underline{x}_2 - \underline{x}_1}{\underline{x}_{\text{average}}}, \frac{\underline{x}_3 - \underline{x}_2}{\underline{x}_{\text{average}}}, \dots, \frac{\underline{x}_i - \underline{x}_{i-1}}{\underline{x}_{\text{average}}}, \dots, \frac{\underline{x}_n - \underline{x}_{n-1}}{\underline{x}_{\text{average}}} \right\} \quad (13b)$$

where \bar{u} and \underline{u} are the supremum and infimum of the focal element interval. Then, a Dempster-Shafer structure is related to the additional information by the following expression:

$$\{([\bar{u}_1, \underline{u}_1], m_1), ([\bar{u}_2, \underline{u}_2], m_2), \dots, ([\bar{u}_i, \underline{u}_i], m_i), \dots, ([\bar{u}_n, \underline{u}_n], m_n)\}$$

where $\bar{u}_i \leq u_i$ and where $u_i \neq u_{i-1}$ whenever $\bar{u}_i = \bar{u}_{i-1}$. Therefore, the upper bound and lower bound can be approximated by \bar{F} and \underline{F} , respectively.

B. ANALYZING P-BOX AND IP-BOX

The area difference between the p-box and ip-box can be calculated by defining the distance between the area and bound of the p-box and ip-box. Consider two random distribution parameters ψ_1 and ψ_2 , one from the p-box model $[\bar{F}_{\psi_1}(x), \underline{F}_{\psi_1}(x)]$ and the other from ip-box model $[\bar{F}_{\psi_2}(x), \underline{F}_{\psi_2}(x)]$ [29]. The area between the p-box upper bound $\bar{F}_{\psi_1}(x)$ and the ip-box upper bound $\bar{F}_{\psi_2}(x)$ can be calculated by the integral:

$$\bar{S} = \int_{-\infty}^{\infty} |\bar{F}_{\psi_1}(x) - \bar{F}_{\psi_2}(x)| \Delta x \tag{14}$$

where \bar{S} is the area between the p-box upper bound $\bar{F}_{\psi_1}(x)$ and the ip-box upper bound $\bar{F}_{\psi_2}(x)$. Similarly, the area between the p-box lower bound $\underline{F}_{\psi_1}(x)$ and the ip-box lower bound $\underline{F}_{\psi_2}(x)$ can be calculated by the integral:

$$\underline{S} = \int_{-\infty}^{\infty} |\underline{F}_{\psi_1}(x) - \underline{F}_{\psi_2}(x)| \Delta x \tag{15}$$

where \underline{S} is the area between the p-box lower bound $\underline{F}_{\psi_1}(x)$ and the ip-box upper bound $\bar{F}_{\psi_2}(x)$. The absolute area difference S is related to \bar{S} and \underline{S} by the following expression:

$$S = \bar{S} + \underline{S} \tag{16}$$

The absolute area difference S can express the area difference between the p-box model and ip-box model as a whole, but the distance of the bounds cannot be expressed. Consider one point β in the cumulative probability range of 0 to 1. Draw a horizontal line parallel to the horizontal axis through β [30]. The horizontal line and p-box (or ip-box) will produce multiple crossover points, where the crossover points of $[\bar{F}_{\psi_1}(x), \underline{F}_{\psi_1}(x)]$ are $\bar{F}_{\psi_1}^{-1}(\beta)$ and $\underline{F}_{\psi_1}^{-1}(\beta)$, respectively, and the crossover points of $[\bar{F}_{\psi_2}(x), \underline{F}_{\psi_2}(x)]$ are $\bar{F}_{\psi_2}^{-1}(\beta)$ and $\underline{F}_{\psi_2}^{-1}(\beta)$, respectively. Then, the absolute bound difference $\Delta L(\beta)$ can be defined as

$$\Delta L(\beta) = \Delta L_1(\beta) + \Delta L_2(\beta) \tag{17}$$

where $\Delta L_1(\beta)$ and $\Delta L_2(\beta)$ can be calculated by the following expressions:

$$\Delta L_1(\beta) = |\bar{F}_{\psi_1}^{-1}(\beta) - \bar{F}_{\psi_2}^{-1}(\beta)| \tag{18a}$$

$$\Delta L_2(\beta) = |\underline{F}_{\psi_1}^{-1}(\beta) - \underline{F}_{\psi_2}^{-1}(\beta)| \tag{18b}$$

S and $\Delta L(\beta)$ can quantify the difference between the p-box and the ip-box, and they are used to compare and analyze the results of the two modeling methods. For the same reason, a relative area difference D between $[\bar{F}_{\psi_1}(x), \underline{F}_{\psi_1}(x)]$ and

$[\bar{F}_{\psi_2}(x), \underline{F}_{\psi_2}(x)]$ can be defined as:

$$D = 100 \times \frac{S}{\sum_{-\infty}^{\infty} |\bar{F}_{\psi_1}(x) - \underline{F}_{\psi_1}(x)| \Delta x} \% \tag{19}$$

A relative bound difference D_L between $[\bar{F}_{\psi_1}(x), \underline{F}_{\psi_1}(x)]$ and $[\bar{F}_{\psi_2}(x), \underline{F}_{\psi_2}(x)]$ can be defined as:

$$D_L = 100 \times \frac{\Delta L(\beta)}{\sum_{-\infty}^{\infty} |\bar{F}_{\psi_2}(x) - \underline{F}_{\psi_2}(x)| \Delta x} \% \tag{20}$$

IV. MODELING AND DISCUSSION

A. DATA COLLECTION

The vibration signals in Ref. [5] were used to verify the effectiveness of the improved method. The drive end bearing type is a 30305 SKF tapered roller bearing. The signal acquisition system is NI PXI-1042Q high-performance acoustic vibration testing system. The sensor is a PCB M603C01 ICP acceleration transducer; the sampling frequency is 10240, and the motor speed is 800r/min; the sample numbers is then set to 60 [5]. The experiment platform, accelerometers arrangement and photograph of fault bearings are shown in Fig. 3, where the grooves with 0.5mm deep and 0.5mm wide were processed in the bearing inner raceway, outer raceway, and rolling element by using wire-cut electrical discharge machining technology.

The vibration data and bearing fault cases can be expressed as follows: “The power spectra of the bearing signals under eight conditions are shown in Fig. 4, where H stands for healthy bearing, IR stands for inner race fault, OR stands for outer race fault, RE stands for rolling element fault, IOR stands for inner and outer race faults, IRRE stands for the inner race and rolling element faults, ORRE stands for the outer race and rolling element faults, and IORRE stands for the inner race, outer race, and rolling element faults”; the power spectra were calculated by the MATLAB 2017b in which the Hanning window was used to record the data, then data of the power spectra were extracted for plotting power spectra using the OriginPro 2018.

H, IR, OR and RE correspond to single fault signals of the rolling element bearing, and IOR, IRRE, ORRE, and IORRE correspond to compound fault signals in Fig. 4. Based on the vibration spectrum, we found that “the basic power spectra are not generally informative in Fig. 4, partly because they are in the presence of strong masking signals from other machine components but also because the angle of the load from the radial plane varies with the position of each rolling element in the bearing as the ratio of the local radial to the axial load changes” [1], [31]. It should be noted that the uncertainty of the bearing signals is the main characterization to bearing signals, because it always exists and changes from the beginning of health signals to the formation of fault signals. Hence, collecting uncertainty of the bearing time-domain signals provides a new way for rolling element bearing diagnostics.

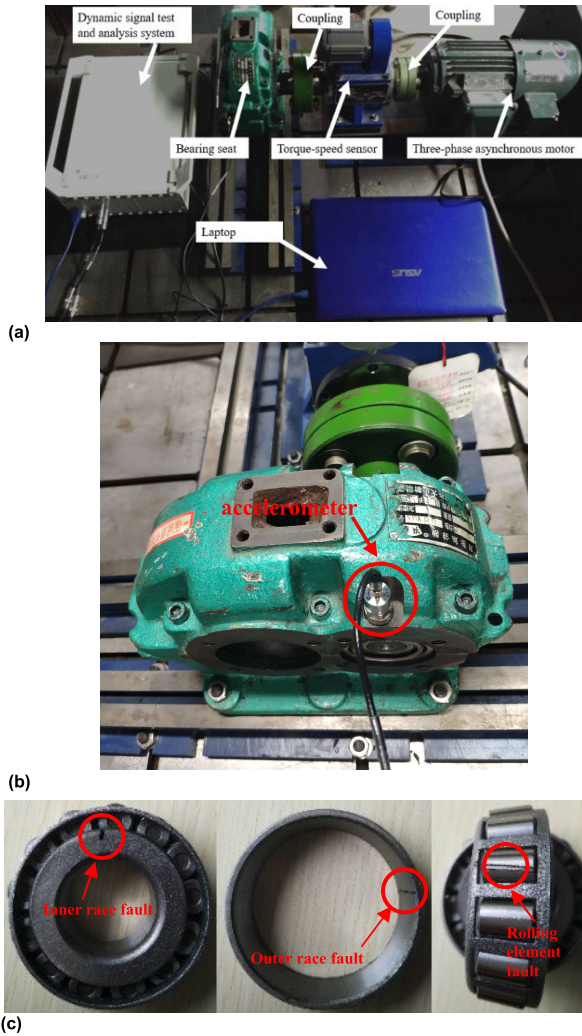


FIGURE 3. The experiment platform [5]; (a) The overall layout of the experiment platform; (b) accelerometer arrangement; (c) photograph of fault bearings.

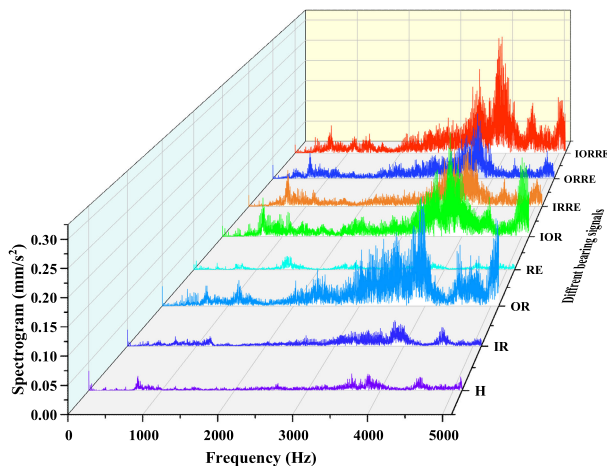


FIGURE 4. The power spectra of the experimental bearing acceleration signals.

In a previous study, we showed that the p-box modeling method can suitably avoid the decoupling process for bearing signals [5].

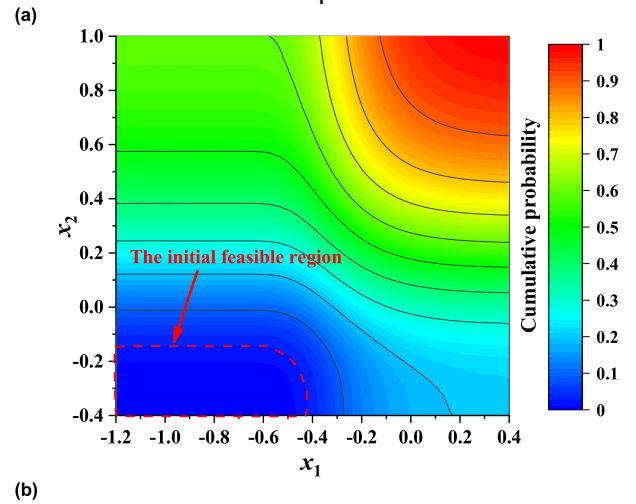
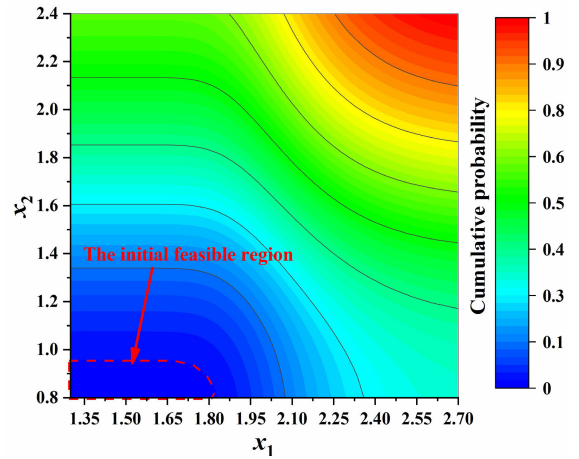


FIGURE 5. The p-boxes based on the DPMM; (a) kurtosis p-box of IORRE; (b) skewness p-box of IORRE.

B. COMPARISON BETWEEN P-BOX AND IP-BOX BASED ON DPMM

The random distribution of these features from the IORRE data can be verified to establish a p-box based on the DPMM, as shown in Fig. 5, where the results for the random distributions of the kurtosis and skewness features satisfy normal distributions, and their number of focal elements is 1000. The remaining bearing signals (i.e., H, IR, OR, RE, IOR, IRRE and ORRE) can also be verified to establish corresponding p-boxes based on the DPMM.

Kurtosis and skewness p-boxes of IORRE are the set of all nondecreasing functions from the reals within 0 to 1, as shown in Fig. 5, where the initial feasible points for the kurtosis and skewness of p-boxes is on the boundary. The initial feasible point $(x_1=1.3, x_2=0.8)$ and $(x_1=-1.2, x_2=-0.4)$ used in this paper are selected for the kurtosis and skewness of p-boxes, respectively. Consider the optimization problem $\min F_X(x; \Psi)$, s.t: $\bar{x} \leq x \leq \underline{x}$; and set $k = 1$. To analyze the influence of different focal element numbers on area differences, some values of focal elements are set, and they are represented in the array $N=[20, 30, \dots, 1000]$. The additional information is calculated with Eq. (13) and added into the

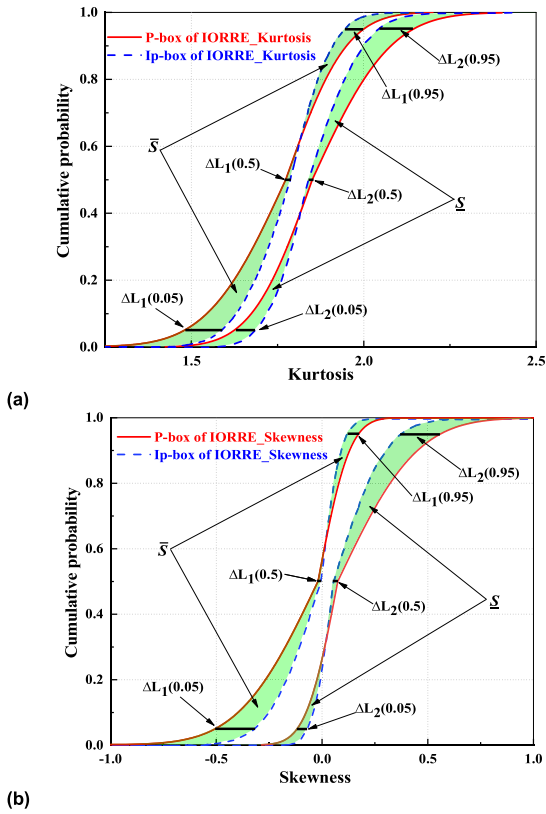


FIGURE 6. The p-boxes and ip-boxes of IORRE based on the kurtosis and skewness; (a) the p-box vs. the ip-box for the kurtosis; (b) the p-box vs. the ip-box for the skewness.

kurtosis and skewness p-boxes of IORRE. When $N=1000$, the p-boxes and ip-boxes are compared as shown in Fig. 6.

The distance of bounds for the ip-box is smaller than that of the p-box, as shown in Fig. 6. This may be because the overlapped region of the focal element intervals is reduced by filtering the repeated internal search paths to obtain the optimal modeling path based on the ip-box modeling method. The iterations of the ip-boxes and p-boxes are listed in Table 1.

TABLE 1. The ip-boxes vs. the p-box in terms of iterations.

Type	P-box	Ip-box	Relative efficiency (ip-box vs. p-box)
Kurtosis	32670	15730	48.14%
Skewness	25580	15390	60.16%

From Table 1, we infer that the iterations of the skewness p-box are lower than the iterations of the kurtosis p-box because more time is required to calculate a greater-order moment. The iterations of the kurtosis ip-box account for 48.14% of the iterations of the kurtosis p-box; the iterations of the skewness ip-box account for 60.16% of the iterations of the skewness p-box. The difference in the relative efficiency is affected by the search regions of the optimized path, and the overlapping of search regions increases the relative efficiency of the iterations. From Fig. 5, we infer that there are many overlapping parts in search regions when solving the

maximum and minimum kurtosis p-box and skewness p-box, but the overlapping regions of the kurtosis p-box are smaller than the overlapping regions of the skewness p-box, which reduces the relative efficiency of the kurtosis (ip-box vs. p-box). Similarly, the p-boxes of H, IR, OR, RE, IOR, IRRE, and ORRE based on the DPMM are analyzed, and they are similar to the p-box of IORRE.

C. EFFECT OF FOCUS ELEMENTS

The focal element intervals of the ip-box model are smaller than those of the p-box model; that is, the data of the ip-box model are closer to the mean value, as shown in Fig. 6. When the cumulative probability density of \bar{S} is 0.6, the area \bar{S} of the upper part is smaller than that of the lower part (Fig. 5(a)). When the cumulative probability density of \underline{S} is 0.4, the area \underline{S} of the upper part is larger than that of the lower part. When the cumulative probability density is 0.5, \bar{S} is approximately antisymmetric to \underline{S} (Fig. 5(a)). When the cumulative probability density of \bar{S} is 0.6, the area \bar{S} of the upper part is smaller than that of the lower part (Fig. 5(b)). When the cumulative probability density of \underline{S} is 0.3, the area \underline{S} of the upper part is larger than that of the lower part. Similarly, \bar{S} is approximately antisymmetric to \underline{S} (Fig. 5(b)). Using $\bar{S} + \underline{S}$, the curve of area difference between the ip-box and p-box based on the kurtosis and skewness of IORRE is obtained under the array $N=[20, 30, \dots, 1000]$, as shown in Fig. 7.

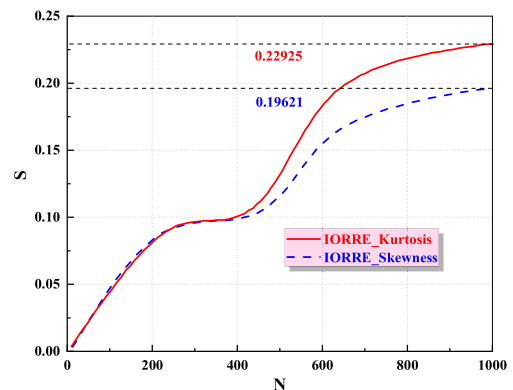


FIGURE 7. The area difference varies with the number of focal elements.

As the number of focal elements increases, the area difference increases (Fig. 7). This is because Pl_i and Bel_i can be approximated by \bar{F} and \underline{F} in a stepwise, respectively. When the number of focal elements is less than 400, the growth rates of area differences for the kurtosis and the skewness are similar; when the number of focal elements is more than 400, the growth rates of area differences of the kurtosis are larger than that of the skewness. This is because the skewness is a smaller-order moment of a probability function than the kurtosis. From Fig. 5, we infer that $\nabla F_{Kurtosis} > \nabla F_{Skewness}$. Then, for $\min F_X(x; \Psi)$, s.t: $\bar{x} \leq x \leq \underline{x}$, as the number of focal elements increases, the kurtosis is more advantageous than the skewness. The number of focal elements is 1000, and the area difference between the p-box and the ip-box based on

TABLE 2. The relative area differences and the relative bound differences for IORRE.

Type	Relative area difference	Relative bound difference		
		$\beta=0.05$	$\beta=0.5$	$\beta=0.95$
Kurtosis	33.13%	73.82%	73.82%	73.82%
Skewness	31.72%	72.47%	72.47%	72.47%

the kurtosis of IORRE is 0.22925; the area difference of the skewness is 0.19621 (Fig. 7). However, the distance of the bounds cannot be expressed by the area difference. To express the bound difference $\Delta L(\beta)$, the p-box and the ip-box are truncated with $\beta = [0.05, 0.5, 0.95]$, as shown in Fig. 6. The bound differences between the ip-box and p-box are larger on both sides and smaller in the middle. A relative area difference and a relative bound difference are calculated by Eqs. (19) and (20), as listed in Table 2, respectively, because the relative difference can better reflect the reliability of measurement than the absolute difference.

From the relative area differences in Table 2, we infer that the improvement degree of the kurtosis is better than that of the skewness. The relative boundary differences of kurtosis and skewness are close to 0 under $\beta = 0.5$, which is less than the values for $\beta = 0.05$ and $\beta = 0.95$. For $\beta = 0.05$ and $\beta = 0.95$, the relative boundary difference of kurtosis and skewness is approximately equal, which proves that \bar{S} is approximately antisymmetric to \underline{S} (Fig. 6(b)). This may be because the modeling data satisfy a normal distribution, where $(\mu_{left}, \sigma_{lower})$ must be antisymmetric to $(\mu_{right}, \sigma_{upper})$ under $\beta = 0.5$ and $(\mu_{left}, \sigma_{upper})$ must be antisymmetric to $(\mu_{right}, \sigma_{lower})$ under $\beta = 0.5$. However, the numerical calculation will produce errors, so there is only approximate antisymmetry for \bar{S} and \underline{S} . From Table 2, we infer that this small error can be ignored. For H, IR, OR, RE, IOR, IRRE, and ORRE, the relative area difference and relative bound difference based on the DPMM are listed in Table 3.

Comparing Tables 2 and 3, the improvement degree of the p-boxes for single fault signals of the rolling element bearing (i.e., the p-boxes of H, IR, OR and RE) is smaller than that of the p-boxes for compound fault signals (i.e., the p-boxes of IOR, IRRE, ORRE, and IORRE). This may be because the

overlapping interior region of the Dempster-Shafer structure for the p-box of compound fault signals is larger than that of single fault signals. From Table 3, we infer that the improvement degrees of the kurtosis and skewness p-boxes for H are the smallest because the healthy bearing signals envelop a minimal amount of uncertainty, and the p-box of H could be tighter at the beginning.

D. COMPARISON BETWEEN P-BOX AND IP-BOX BASED ON RDPMM

It is worth emphasizing that the cooperative optimization method is inapplicable for the RDPMM because the p-boxes can be established by the RDPMM based on raw bearing data, which avoids the need to verify the suitability of a random distribution. Then, the additional information can be added directly into the p-box of the RDPMM (calculated with Eq. (13)), which reduces the amount of IORRE data fluctuation between the adjacent focal elements. The p-box of the RDPMM and the ip-box of the RDPMM based on the IORRE were established as shown in Fig. 8.

The interval range of each focal element interval is not equal $(|\bar{x}_i - \underline{x}_i| \neq |\bar{x}_{i+k} - \underline{x}_{i+k}|)$, which results in a different improvement degree for each focal element (Fig. 8). The larger the raw focal element interval $[\bar{x}_i, \underline{x}_i]$ is, the larger the improvement degree is. The aggregated width of the ip-box is smaller than that of the p-box, as shown in Fig. 8. The ip-box and the p-box need 714301 iterations, respectively, indicating that the number of iterations for the ip-box has not decreased. This is because the modeling data of the p-box and the ip-box for the RDPMM are from the raw bearing data. In a comparison with the number of iterations 32670 for the kurtosis p-box of IORRE, the relative efficiency (the kurtosis p-box vs. the RDPMM p-box) is 4.57%. In a comparison with the number of iterations 25580 for the skewness p-box of IORRE, the relative efficiency (the skewness p-box vs. the RDPMM p-box) is 3.58%. Similarly, in a comparison with the number of iterations 15730 for the kurtosis ip-box of IORRE, the relative efficiency (the kurtosis ip-box vs. the RDPMM ip-box) is 2.20%; in a comparison with the number of iterations 15390 for the skewness ip-box of IORRE, the relative efficiency (the skewness ip-box vs. the RDPMM ip-box)

TABLE 3. The relative area differences and the relative bound differences for H, IR, OR, RE, IOR, IRRE and ORRE.

Kurtosis		H	IR	OR	RE	IOR	IRRE	ORRE
Relative area difference (%)		8.12	23.11	20.19	20.31	29.12	30.18	31.15
	$\beta=0.05$	19.11	73.44	73.36	73.19	71.97	72.11	74.13
	$\beta=0.5$	0.88	0.9	0.89	0.91	0.89	0.92	0.91
Relative bound difference (%)	$\beta=0.95$	19.10	73.45	73.35	73.21	71.97	72.14	74.15
Skewness		H	IR	OR	RE	IOR	IRRE	ORRE
Relative area difference (%)		7.33	19.12	19.29	18.13	29.11	28.28	30.19
	$\beta=0.05$	17.98	70.61	71.11	69.99	70.77	69.39	71.72
	$\beta=0.5$	0.85	0.87	0.88	0.88	0.86	0.87	0.88
Relative bound difference (%)	$\beta=0.95$	17.97	70.62	71.11	69.97	70.79	69.38	71.69

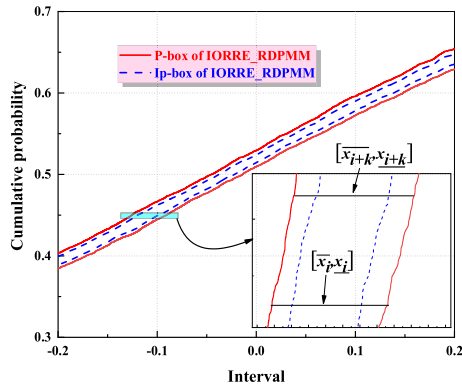


FIGURE 8. The p-box vs. the ip-box for the RDPMM.

is 2.15%. The results for other bearing signals were similar to IORRE. Thus, although the aggregated width of the ip-box for the RDPMM is small, the number of iterations is not reduced.

V. INTELLIGENT DETECTION FOR ROLLING BEARING

A. OVERLAPPED REGIONS BETWEEN DIFFERENT BEARING CONDITIONS

For the different bearing conditions, the p-boxes and the ip-boxes are established based on the skewness as shown in Fig. 9.

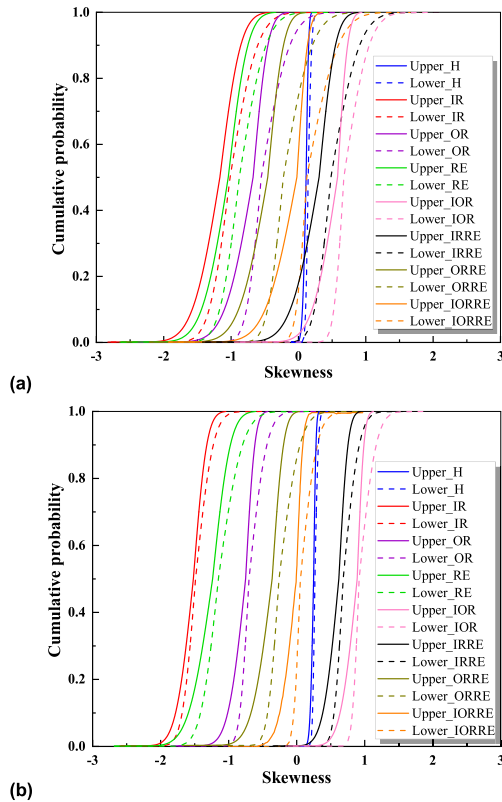


FIGURE 9. The p-boxes and the ip-boxes based on the skewness; (a) the p-boxes of the different bearing conditions; (b) the ip-boxes of the different bearing conditions.

Fig. 9 gives the upper and lower CDFs of the different bearing conditions. The p-boxes based on the skewness

values of the different bearing conditions are not easy to be distinguished in Fig. 9(a) because of the overlapped regions between the different p-boxes; the overlapped regions are not conducive to the intelligent detection of machine learning. Based on the ip-box modeling method, the overlapped regions are reduced by filtering the repeated internal search paths to obtain the optimal modeling path as shown in Fig. 9(b). Based on the kurtosis and the RDPMM, the p-boxes and the ip-boxes were established to contrast the overlapped regions; the experimental results showed that the overlapped regions of the ip-boxes were smaller than that of the p-boxes. Reducing the overlapped regions can increase the discrimination of each ip-box, then improve the recognition rate of the ip-box. The effectiveness of the ip-box is demonstrated as follows.

B. FEATURE-BASED DEMPSTER-SHAFER STRUCTURE

In the pattern recognition system, the ip-box must be discrete into a Dempster-Shafer structure which is composed of the improved focal element intervals and mass function. The width of different focal element intervals can be obtained through the uncertainty measurement method. The weight of a single focal element interval can be obtained by multiplying the corresponding mass value because different focal element interval is independent of each other. A basic feature of ip-box can be obtained by accumulating every uncertainty probability of the Dempster-Shafer structure. The basic feature called the aggregated width can be expressed as:

$$\omega_1 = \sum_{i=1}^n (m_i \times |\alpha_{\bar{x}_i} - \alpha_{x_i}|) \tag{21}$$

where $\alpha_{\bar{x}_i}$ and α_{x_i} are the upper and lower bounds of the focal element interval respectively; m_i is the mass value of the corresponding focal element interval. The aggregated widths can be obtained by measuring the aggregated uncertainty of different ip-boxes, as shown in Fig. 10.

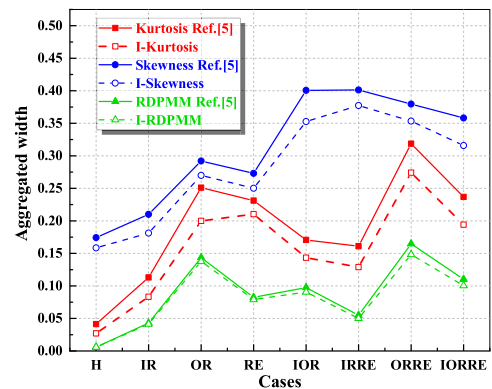


FIGURE 10. The aggregated width for different ip-boxes.

From Fig. 10, the improved aggregated width is smaller than that in Ref. [5], which proves that this paper method is effective. For the improved method, the aggregated width of the RDPMM ip-boxes is the smallest, and that of the skewness ip-boxes is the largest. It may mean that the RDPMM ip-boxes are more conducive to pattern recognition because

TABLE 4. The interval value of the aggregated uncertainty measurement.

Cases	Kurtosis		Skewness		RDPMM	
	ω_5	ω_6	ω_5	ω_6	ω_5	ω_6
H	[0.0088,0.0484]	[0.0041,0.0539]	[0.0097,0.0492]	[0.0048,0.0599]	[0.2117,0.2175]	[0.2000,0.2077]
IR	[0.0835,0.2835]	[0.5129,0.7757]	[0.0835,0.3535]	[0.5729,0.8457]	[0.3345,0.3754]	[0.2789,0.3304]
OR	[0.0136,0.2208]	[0.3333,0.6879]	[0.0196,0.2908]	[0.4133,0.7579]	[1.5219,1.6568]	[1.3068,1.4901]
RE	[0.1451,0.3532]	[0.5522,0.8246]	[0.2851,0.4932]	[0.6122,0.8946]	[1.0576,1.1351]	[0.8326,0.9514]
IOR	[0.0063,0.1514]	[0.2533,0.4631]	[0.0076,0.2814]	[0.3133,0.5331]	[1.3556,1.4451]	[1.0776,1.1892]
IRRE	[0.2396,0.3386]	[0.2073,0.8985]	[0.2996,0.3986]	[0.2773,0.9785]	[0.9076,0.9571]	[0.7246,0.7981]
ORRE	[0.0063,0.2551]	[0.3979,0.8320]	[0.0079,0.4251]	[0.4679,0.9220]	[1.0763,1.2296]	[0.8252,1.0257]
IORRE	[0.0913,0.1371]	[0.1177,0.4344]	[0.1083,0.2771]	[0.1977,0.5344]	[1.7574,1.8579]	[1.4819,1.6116]

TABLE 5. Confusion matrix of the kurtosis ip-boxes.

Classified as	H	IR	OR	RE	IOR	IRRE	ORRE	IORRE
H	195	5	0	0	0	0	0	0
IR	0	188	9	3	0	0	0	0
OR	0	0	187	0	10	3	0	0
RE	0	7	12	179	2	0	0	0
IOR	0	0	2	0	180	10	8	0
IRRE	0	6	0	4	0	189	0	1
ORRE	0	2	7	0	4	0	187	0
IORRE	5	1	0	3	0	2	1	188

the improvement of the aggregated width can reduce the degree of overlap between ip-boxes. Similarly, in order to measure more ip-box information, more features can be expressed as:

$$\omega_2 = \sum_{i=1}^n (m_i \times \log_2 |\alpha_{\bar{x}_i} - \alpha_{\underline{x}_i}|) \tag{22a}$$

$$\omega_3 = \sum_{i=1}^n (m_i \times \log_2 (1 + |\alpha_{\bar{x}_i} - \alpha_{\underline{x}_i}|)) \tag{22b}$$

$$\omega_4 = - \sum_{i=1}^n m_i \times \log_2 (\varepsilon_2 - \varepsilon_1) \tag{22c}$$

$$\omega_5 = \left[\sum_{i=1}^n m_i \alpha_{\bar{x}_i}, \sum_{i=1}^n m_i \alpha_{\underline{x}_i} \right] \tag{22d}$$

$$\omega_6 = \left[\sum_{i=1}^n \Theta \alpha_{\bar{x}_i}, \sum_{i=1}^n \Theta \alpha_{\underline{x}_i} \right] \tag{22e}$$

where ε_1 and ε_2 are the means of the probability statistics associated with the lower and upper bounds of the cp-box, respectively; $\omega_i (i = 2, 3, 4)$ and ω_1 are similar, and are scalar values. Θ is a conditional value similar to the degree of confidence, which satisfies $\sum_{i=1}^{n-1} m_i \leq \Theta \leq \sum_{i=1}^n m_i$. The value 0.95 was used for Θ in the study [30]. $\omega_j (j = 5, 6)$ is the interval value of uncertainty measurement information of the ip-boxes, as listed in Table 4.

C. PATTERN RECOGNITION WITH DIFFERENT METHODS

“In the present approach, the influence of various loads and speeds on the ip-box leads to shifting lower and upper bounds of the ip-box to the negative or positive values on the abscissa axis, but the aggregated width of the ip-box is unaffected because the signal patterns remain unchanged. This is also one of the advantages that the ip-box can include the uncertainties of the bearing signals. Hence, the ip-box still has the advantages of improvement under different loads and speeds and has the robustness of anti-interference. Additionally, the diagnostic results are unaffected as long as the aggregated width remains unchanged. For simplicity reasons, we can ignore the influence of different loads and speeds on the ip-box to study the classification performance of different ip-box modeling methods.”

A total of 1600 p-boxes were used in this study, where 60% were training set, 20% were verification set and 20% were test set. The Python environment-based SVM classifier was configured as described in Ref. 5.

Based on the SVM, the results of the kurtosis ip-boxes were displayed in a two-dimensional confusion matrix [32], as listed in Table 5.

From Table 5, we infer that the SVM classified 195, 188, 187, 179, 180, 189, 187, and 188 cases correctly for H, IR, OR, RE, IOR, IRRE, ORRE, and IORRE, respectively. Similarly, the confusion matrix of the skewness and the RDPMM ip-boxes can be obtained

The SVM for the skewness classified 193, 190, 186, 176, 177, 180, 186, and 181 cases correctly for H, IR, OR,

TABLE 6. Detailed accuracy by class of the ip-boxes.

Kurtosis	TP rate	FP rate	Precision	Recall	F-measure
H	0.975	0.0036	0.975	0.975	0.975
IR	0.94	0.015	0.8995	0.94	0.9193
OR	0.935	0.0214	0.8618	0.935	0.8969
RE	0.895	0.0071	0.9471	0.895	0.9203
IOR	0.9	0.0114	0.9184	0.9	0.9091
IRRE	0.945	0.0107	0.9265	0.945	0.9356
ORRE	0.935	0.0064	0.9541	0.935	0.9444
IORRE	0.94	0.0007	0.9947	0.94	0.9666
Skewness	TP rate	FP rate	Precision	Recall	F-measure
H	0.965	0.0043	0.9698	0.965	0.9674
IR	0.95	0.015	0.9005	0.95	0.9246
OR	0.93	0.0236	0.8493	0.93	0.8878
RE	0.88	0.0114	0.9167	0.88	0.898
IOR	0.885	0.0129	0.9077	0.885	0.8962
IRRE	0.9	0.0143	0.9	0.9	0.9
ORRE	0.93	0.0107	0.9254	0.93	0.9277
IORRE	0.905	0.0014	0.9891	0.905	0.9452
RDPMM	TP rate	FP rate	Precision	Recall	F-measure
H	1	0.0036	0.9756	1	0.9877
IR	1	0.0021	0.9852	1	0.9926
OR	1	0.0036	0.9756	1	0.9877
RE	0.985	0.0021	0.985	0.985	0.985
IOR	0.98	0.0021	0.9849	0.98	0.9825
IRRE	0.955	0.0029	0.9795	0.955	0.9671
ORRE	0.975	0.0021	0.9848	0.975	0.9799
IORRE	0.965	0.0014	0.9897	0.965	0.9772

RE, IOR, IRRE, ORRE, and IORRE, respectively. For the RDPMM ip-boxes, the SVM correctly predicted 200, 200, 200, 197, 196, 191, 195 and 193 cases. The ip-boxes obtained by the RDPMM have greater accuracy in correctly predicting the bearing condition with a combined bearing component fault than the DPMM. It may be because that the impact of the aggregated width. For the RDPMM and the DPMM, the detailed accuracy of each class is shown in Table 6.

Table 6 provides information about the true positive (TP) rate, false positive (FP) rate, precision, recall, and F-measure values for the eight classes using the SVM. From Table 6, we refer that the mean of the RDPMM F-measure values is larger than the means of the kurtosis and skewness F-measure values, and the mean of the kurtosis F-measure values is larger than the means of the skewness F-measure values.

Teymourlouieet al. [33] and Yaddaden et al. [34] point out “the basic quality measure offered by the error rate is no longer appropriate: errors are not simply present or absent; they come in different sizes”. Then, for the different ip-boxes modeling methods, the measurement errors in classification results are given in Table 7.

From Table 7, we infer that the RDPMM ip-box is the best according to all five metrics: it has the smallest value for each error measure and the largest Kappa statistic. The kurtosis ip-box is the second-best by all five metrics. The skewness ip-box is the third-best by all five metrics.

Employing the p-box modeling methods, the kurtosis, skewness, and RDPMM p-boxes were established, and its features were extracted for pattern recognition; the experimental results showed that the classification performances of

TABLE 7. Evaluation of the success of the numeric prediction.

Parameters	Values (SVM)		
	Kurtosis	Skewness	RDPMM
Correctly classified instances	93.31%	91.81%	98.25%
Incorrectly Classified Instances	6.69%	8.19%	1.75%
Kappa statistic	0.92	0.91	0.98
Mean absolute error	0.0669	0.0819	0.0175
Root mean squared error	0.071	0.0867	0.0237
Relative absolute error	0.9378	0.9227	0.9874
Root relative squared error	0.9381	0.9232	0.9876

the kurtosis, skewness, and RDPMM p-boxes were 83.4%, 81.5%, and 95.2%, respectively. Compared with the p-box, the improvement degree of the classification performance for the kurtosis ip-boxes is 9.91%, that of the skewness ip-boxes is 10.31%, and that of the RDPMM ip-boxes is 3.05%. It is because the overlapped regions between the ip-boxes are reduced.

D. COMPARISON OF DIFFERENT DATA PROCESSING METHODS

A comparison of different data processing studies should be undertaken to demonstrate the advantage of the method proposed by this paper. Composite multiscale fuzzy entropy is an effective method to analyze the complexity of time series in bearing fault diagnosis [35]. It can not only reflect the complexity characteristics of time series from multiple scales but also has the advantages of short data and good robustness. Fig. 11 presents the results of the composite multiscale

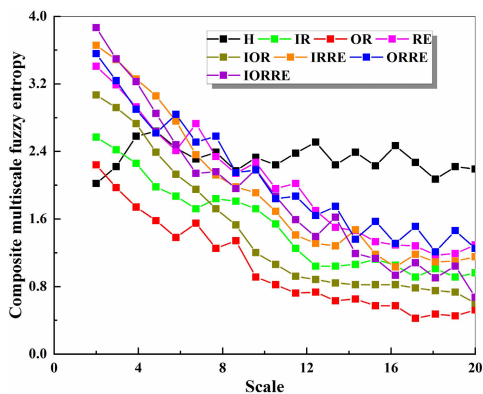


FIGURE 11. Composite multiscale fuzzy entropy of the different bearing data.

fuzzy entropy for each bearing condition based on the current data, where some values were applied as follows: Largest scale is 20, embedding dimension 2, a gradient of exponent function 2, and similarity tolerance 0.15SD (SD denotes the standard deviation of raw bearing data).

In Fig. 11, the fuzzy entropy of H is larger on the relatively large scale, and changes gently with the increase of the scale values; the curve of composite multiscale fuzzy entropy for other bearing conditions shows the obvious decreasing trend. In this contrastive study, the steps used in this method can be described as the follows: Firstly, Total 1600 samples were used in this study, i.e., there were 200 samples for each bearing condition; the feature set was obtained by calculating the values of composite multiscale fuzzy entropy for each bearing condition. Then, 60% of features were training set, 20% were verification set and 20% were test set. Finally, the correct classification of faults can be given by the SVM model, as listed in Table 8.

TABLE 8. A comparison of current and composite multiscale fuzzy entropy.

Method	Kurtosis	Skewness	RDPMM	Composite multiscale fuzzy entropy
Total correct recognition rate	93.31%	91.81%	98.25%	83.5%

For the composite multiscale fuzzy entropy, the experimental results showed that the total correct recognition rate is 83.5% in Table 8. However, compared to the correct recognition rates 93.31%, 91.81%, and 98.25% from the proposed methods in this paper, there is still room for improvement. It is because the method of composite multiscale fuzzy entropy requires additional empirical effort in the bearing fault diagnosis [35]. Additionally, for the calculation time of finding fault, the proposed method in this paper consumes 55 seconds, which is 15% faster than the traditional method.

VI. CONCLUSION

This study presents a procedure for obtaining tighter p-boxes by reducing coincidence intervals using the cooperative optimization method and adding additional information. The machine learning method (i.e., SVM) gives the classification accuracy based on the ip-boxes.

For the iterations of the p-box and ip-box, the relative efficiencies of the DPMM and the RDPMM were calculated. The results showed that the relative efficiency, which denotes the RDPMM ip-box vs. the kurtosis ip-box, was 2.20%. For the skewness ip-box, the relative efficiency was 2.15%. Therefore, the computational efficiency of the skewness ip-box is best.

To quantitatively assess the difference between the p-box and ip-box, relative area difference and relative bound difference of the different cumulative probability density values were obtained based on the kurtosis and skewness. The experimental results showed that the aggregated width of the DPMM p-box was improved, and the improvement degree of the kurtosis p-box was larger than that of the skewness p-box. Consequently, the kurtosis ip-box is more conducive for bearing fault diagnosis than the skewness p-box without considering the efficiency of calculation.

To compare the modeling method of the p-box and the ip-box for the kurtosis, skewness, and the RDPMM, an SVM was used as the classification algorithm. The experimental results showed that the improvement degree of the classification performance for the kurtosis ip-boxes was 9.91%, that of the skewness ip-boxes was 10.31%, and that of the RDPMM ip-boxes was 3.05%. However, the total correct classification rates of the RDPMM ip-boxes were higher than those of the kurtosis and the skewness ip-boxes.

REFERENCES

- [1] K. Worden, W. J. Staszewski, and J. J. Hensman, "Natural computing for mechanical systems research: A tutorial overview," *Mech. Syst. Signal Process.*, vol. 25, no. 1, pp. 4–111, Jan. 2011, doi: 10.1016/j.ymssp.2010.07.013.
- [2] Y. Lei, Z. Liu, D. Wang, X. Yang, H. Liu, and J. Lin, "A probability distribution model of tooth pits for evaluating time-varying mesh stiffness of pitting gears," *Mech. Syst. Signal Process.*, vol. 106, pp. 355–366, Jun. 2018, doi: 10.1016/j.ymssp.2018.01.005.
- [3] I. S. dos Santos, C. R. Hall Barbosa, M. T. F. da Silva, and O. B. D. Oliveira Filho, "Multi-resolution wavelet analysis for noise reduction in impulse puncture voltage measurements," *Measurement*, vol. 153, Mar. 2020, Art. no. 107416, doi: 10.1016/j.measurement.2019.107416.
- [4] S. Biswal and A. Ramaswamy, "Finite element model updating of concrete structures based on imprecise probability," *Mech. Syst. Signal Process.*, vol. 94, pp. 165–179, Sep. 2017, doi: 10.1016/j.ymssp.2017.02.042.
- [5] H. Tang, Y. Du, and H.-L. Dai, "Rolling element bearing diagnosis based on probability box theory," *Appl. Math. Model.*, vol. 80, pp. 944–960, Apr. 2020, doi: 10.1016/j.apm.2019.10.068.
- [6] V. Nithya, R. Sakthivel, R. Sakthivel, and F. Alzahrani, "Dissipative-based non-fragile filtering for fuzzy networked control systems with switching communication channels," *Appl. Math. Comput.*, vol. 373, May 2020, Art. no. 125011, doi: 10.1016/j.amc.2019.125011.
- [7] F. Yang, H. Wei, and P. Feng, "A hierarchical dempster-shafer evidence combination framework for urban area land cover classification," *Measurement*, vol. 151, Feb. 2020, Art. no. 105916, doi: 10.1016/j.measurement.2018.09.058.
- [8] S. Sozzo, "Effectiveness of the quantum-mechanical formalism in cognitive modeling," *Soft Comput.*, vol. 21, no. 6, pp. 1455–1465, 2017, doi: 10.1007/s00500-015-1834-y.

- [9] F. Giovanneschi, "Dictionary learning for adaptive GPR landmine classification," *IEEE T. Geosci. Remote*, vol. 57, no. 12, pp. 10036–10055, 2019, doi: [10.1109/TGRS.2019.2931134](https://doi.org/10.1109/TGRS.2019.2931134).
- [10] R. Xin, J. Zhang, and Y. Shao, "Complex network classification with convolutional neural network," *Tsinghua Sci. Technol.*, vol. 25, no. 4, pp. 447–457, Aug. 2020, doi: [10.26599/TST.2019.9010055](https://doi.org/10.26599/TST.2019.9010055).
- [11] S. Kabir, T. K. Geok, M. Kumar, M. Yazdi, and F. Hossain, "A method for temporal fault tree analysis using intuitionistic fuzzy set and expert elicitation," *IEEE Access*, vol. 8, pp. 980–996, 2020, doi: [10.1109/ACCESS.2019.2961953](https://doi.org/10.1109/ACCESS.2019.2961953).
- [12] M. A. Fahmy, "Modeling and optimization of anisotropic viscoelastic porous structures using CQBEM and moving asymptotes algorithm," *Arabian J. Sci. Eng.*, vol. 44, no. 2, pp. 1671–1684, Feb. 2019, doi: [10.1007/s13369-018-3652-x](https://doi.org/10.1007/s13369-018-3652-x).
- [13] A. Marendet, "A standard branch-and-bound approach for nonlinear semi-infinite problems," *Eur. J. Oper. Res.*, vol. 282, no. 2, pp. 438–452, 2020, doi: [10.1016/j.ejor.2019.10.025](https://doi.org/10.1016/j.ejor.2019.10.025).
- [14] Q. Zhang and Z. Li, "Development of an interval quadratic programming water quality management model and its solution algorithms," *J. Cleaner Prod.*, vol. 249, Mar. 2020, Art. no. 119319, doi: [10.1016/j.jclepro.2019.119319](https://doi.org/10.1016/j.jclepro.2019.119319).
- [15] L. Wang, "Interval optimization for structural dynamic responses of an artillery system under uncertainty," *Eng. Optimiz.*, vol. 52, no. 2, pp. 343–366, 2020, doi: [10.1080/0305215X.2019.1590563](https://doi.org/10.1080/0305215X.2019.1590563).
- [16] C. Zhao, C. Wan, and Y. Song, "An adaptive bilevel programming model for nonparametric prediction intervals of wind power generation," *IEEE Trans. Power Syst.*, vol. 35, no. 1, pp. 424–439, May 2020, doi: [10.1109/TPWRS.2019.2924355](https://doi.org/10.1109/TPWRS.2019.2924355).
- [17] S. Zeng, S. M. Chen, and K. Y. Fan, "Interval-valued intuitionistic fuzzy multiple attribute decision making based on nonlinear programming methodology and TOPSIS method," *Inform. Sci.*, vol. 506, pp. 424–442, 2020, doi: [10.1016/j.ins.2019.08.027](https://doi.org/10.1016/j.ins.2019.08.027).
- [18] H. Wang, A. Meng, Y. Liu, X. Fu, and G. Cao, "Unscented Kalman filter based interval state estimation of cyber physical energy system for detection of dynamic attack," *Energy*, vol. 188, Dec. 2019, Art. no. 116036, doi: [10.1016/j.energy.2019.116036](https://doi.org/10.1016/j.energy.2019.116036).
- [19] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Math. Program.*, vol. 106, no. 1, pp. 25–57, Mar. 2006, doi: [10.1007/s10107-004-0559-y](https://doi.org/10.1007/s10107-004-0559-y).
- [20] H. C. Xie and D. H. Liao, "An enhanced nonlinear interval number programming method considering correlation of interval variables," *Struct. Multidiscip.*, vol. 60, no. 5, pp. 2021–2033, 2019, doi: [10.1007/s00158-019-02307-6](https://doi.org/10.1007/s00158-019-02307-6).
- [21] M. Hohmann, J. Warrington, and J. Lygeros, "A moment and sum-of-squares extension of dual dynamic programming with application to nonlinear energy storage problems," *Eur. J. Oper. Res.*, vol. 283, no. 1, pp. 16–32, May 2020, doi: [10.1016/j.ejor.2019.10.041](https://doi.org/10.1016/j.ejor.2019.10.041).
- [22] D. Pylorof, E. Bakolas, and K. S. Chan, "Design of robust Lyapunov-based observers for nonlinear systems with sum-of-squares programming," *IEEE Control Syst. Lett.*, vol. 4, no. 2, pp. 283–288, 2020, doi: [10.1109/LCSYS.2019.2925511](https://doi.org/10.1109/LCSYS.2019.2925511).
- [23] S. Ferson, *Constructing Probability Boxes and Dempster-Shafer Structures*. Albuquerque, NM, USA: Sandia National Laboratories, 2003.
- [24] H. Tang, "Compound fault diagnosis based on probability box theory," in *Proc. ICMIC*, 2017, pp. 336–341.
- [25] A. Pizzoloto, A. Sharma, K. Maute, A. Sciacovelli, and V. Verda, "Multi-scale topology optimization of multi-material structures with controllable geometric complexity—Applications to heat transfer problems," *Comput. Methods Appl. Mech. Eng.*, vol. 357, Dec. 2019, Art. no. 112552, doi: [10.1016/j.cma.2019.07.021](https://doi.org/10.1016/j.cma.2019.07.021).
- [26] S. Zhou and Q. Yue, "Appointment scheduling for multi-stage sequential service systems with stochastic service durations," *Comput. Oper. Res.*, vol. 112, Dec. 2019, Art. no. 104757, doi: [10.1016/j.cor.2019.07.015](https://doi.org/10.1016/j.cor.2019.07.015).
- [27] G. Quintana Carapia and I. Markovsky, "Input parameters estimation from time-varying measurements," *Measurement*, vol. 153, Mar. 2020, Art. no. 107418, doi: [10.1016/j.measurement.2019.107418](https://doi.org/10.1016/j.measurement.2019.107418).
- [28] T. Hong, D. Yi, D. Jiaman, and L. Liqiang, "An improved probability box discretization method for mechanical fault diagnosis," in *Proc. 36th Chin. Control Conf. (CCC)*, Jul. 2017, pp. 11011–11016.
- [29] J. Xu, Y. Tang, W. Yang, F. Li, and L. Shi, "Event-triggered minimax state estimation with a relative entropy constraint," *Automatica*, vol. 110, Dec. 2019, Art. no. 108592, doi: [10.1016/j.automatica.2019.108592](https://doi.org/10.1016/j.automatica.2019.108592).
- [30] S. H. Razavi Hajiagha, H. Amoozad Mahdiraji, and M. Tavana, "A new bi-level data envelopment analysis model for efficiency measurement and target setting," *Measurement*, vol. 147, Dec. 2019, Art. no. 106877, doi: [10.1016/j.measurement.2019.106877](https://doi.org/10.1016/j.measurement.2019.106877).
- [31] T. Gao, S. Cao, L. Hou, and Y. Hou, "An experimental study on the nonlinear vibration phenomenon of a rotor system subjected to barrel roll flight and coupled rub-impact faults," *Measurement*, vol. 153, Mar. 2020, Art. no. 107406, doi: [10.1016/j.measurement.2019.107406](https://doi.org/10.1016/j.measurement.2019.107406).
- [32] H. Witten Ian and F. Eibe, *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed. Berlin, Germany: Morgan Kaufmann, 2005, pp. 26–27.
- [33] M. Teymourlouie, A. Zaeri, M. Nematbakhsh, M. Thimm, and S. Staab, "Detecting hidden errors in an ontology using contextual knowledge," *Expert Syst. Appl.*, vol. 95, pp. 312–323, Apr. 2018, doi: [10.1016/j.eswa.2017.11.034](https://doi.org/10.1016/j.eswa.2017.11.034).
- [34] Y. Yaddaden, M. Adda, A. Bouzouane, S. Gaboury, and B. Bouchard, "User action and facial expression recognition for error detection system in an ambient assisted environment," *Expert Syst. Appl.*, vol. 112, pp. 173–189, Dec. 2018, doi: [10.1016/j.eswa.2018.06.033](https://doi.org/10.1016/j.eswa.2018.06.033).
- [35] J. Zheng, H. Pan, and J. Cheng, "Rolling bearing fault detection and diagnosis based on composite multiscale fuzzy entropy and ensemble support vector machines," *Mech. Syst. Signal Process.*, vol. 85, pp. 746–759, Feb. 2017, doi: [10.1016/j.ymsp.2016.09.010](https://doi.org/10.1016/j.ymsp.2016.09.010).



HONG TANG received the B.E. degree in mechanical engineering from the Kunming University of Science and Technology, China. He is currently pursuing the Ph.D. degree with the College of Mechanical and Vehicle Engineering, Hunan University, China.

His research interests include engineering structure optimization design and reliability, intelligent materials and structural mechanics, and NVH. He is also a member of the Machinery Association.



HONGLIANG DAI received the Ph.D. degree from Shanghai Jiao Tong University, Shanghai, China, in 2005.

He is currently a Professor, a Ph.D. Supervisor, and the Vice President of the School of Mechanical and Vehicle Engineering, Hunan University. He is the Technical Expert of the China New Energy Vehicle Battery Echelon Utilization Alliance. He has published more than 110 SCI papers. His research interests include intelligent materials and structural mechanics, NVH, and engineering structure optimization design and reliability. He is the Reviewer of more than 20 international journals, such as the *International Journal of Solids and Structures*, *Applied Mechanics*, *European Mechanics*, *Composite Structures*, *Thermal Fabric*, and the *International Journal of Pressure Vessels and Piping*.



YI DU received the Ph.D. degree from the Kunming University of Science and Technology, Kunming, China, in 2012.

He is currently a Professor and the Deputy Director of the Mechanical Department, Kunming University of Science and Technology. His research interests include mechanical fault diagnosis, information fusion, and uncertainty analysis.

• • •