IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 24, 2020, accepted November 4, 2020, date of publication November 16, 2020,
date of current version December 29, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3037975

Bearing Data Model of Correlation Probability Box Based on New G-Copula Function

LIANGCAI DONG[1], YING LIU[1], HONG TANG[2], AND YI DU[3]
[1]Department of Industrial Engineering, Shanghai Maritime University, Shanghai 201306, China
[2]State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body, Hunan University, Changsha 410083, China
[3]City College, Kunming University of Science and Technology, Kunming 650051, China

Corresponding author: Hong Tang (1151894071@qq.com)

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. Fault diagnosis method based on vibration signals has been studied for many years. Considering complementary information of the vibration signals from different directions, this article proposed an applied model of a correlation probability box based on G-Copula function for diagnosing bearing faults. First, to avoid constructing binary Copula function directly from the definition of binary Copula function, a new function is defined, and a construction method of binary G-Copula function is proposed based on the new function. Then, the correlation probability box model is established based on a joint cumulative distribution of the G-Copula function to increase the independent of the input data in the support vector machine (SVM) model, and the aggregated widths of the correlation probability box model can be used to monitor a development of the bearing failure. Finally, the experimental results showed that the proposed method obtain the better classification accuracy than other data processing study.

INDEX TERMS Bearing failure, copula function, correlation probability box, support vector machine, classification.

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 modern industry [1]. However, ''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 additional empirical effort [2]''.

The idea of probability box (p-box) was ''originally put forward to express pure 'epistemic uncertainty' with 'interval', and has experienced cross research with fuzzy theory [3], DS evidence theory [4], Boolean logic reasoning based on traditional probability theory [5], Kolmogorov method with sparse samples [6], 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, expert system and fuzzy set theory'' [7]. The applications of p-box involve failure probability assessment of fault system [8], uncertainty assessment of dynamic response of vibration system [9], reliability design of automobile gearbox in the absence of experimental data [10], finite element modeling and parameter optimization of rocket edge shell structure [11], multi parameter uncertainty correlation in damping oscillator [12], mechanical reliability system architecture and assessment [13], error accumulation expression and evaluation of measurement system [14], etc. For the p-box in an application of the bearing diagnosis, the p-box model can well achieve the fusion of the irreducible and epistemic uncertainty [15], [16]. However, data cross between the different p-boxes reduces correct recognition rates of a classifier.

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

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

"In 2009, we proposed firstly a new way for the bearing fault diagnosis based on the p-box theory by collecting uncertainty of the time-domain signals of the bearing, because the p-box theory owns significant advantages in dealing with the uncertainty of bearing signals. However, in previous work on bearing fault diagnosis, for the sake of mathematical simplicity, the random variable of the p-box was considered to be independent [15]. In practical engineering, the correlations among uncertain parameters are objective; although there is duplication between correlation information, there is also complementary information associated with the correlation information. However, research on how to establish the correlation model between the p-boxes is rare, and no complete method is available to measure the aggregated uncertainty information of correlation models." Copula function, as a tool to describe the dependence rules between variables, overcomes the shortcomings of traditional linear correlation coefficient in studying the nonlinear relationship of variables [17]. To reduce the computational complexity caused by complex correlation, Zhang *et al.* [18] proposed a method of association decision fusion based on regular Fuji copula. Nyaga *et al.* [19] introduced a high flexibility copula modeling method based on the sensitivity and specificity of binary beta distribution. On the basis of the random correlation of wind power ramp characteristics, Cui *et al.* [20] established a conditional probability prediction model based on Copula theory. Based on copula theory for correlation investigations of wind farms, Wang and Luo [21] proposed a two-stage filtration method to evaluate different types of copulas. Li *et al.* [22] derived a quantitative reproducibility score with a mixed copula model to fit a curve. To counteract low resolution, poor illumination and noise issues, copula probability models based on Gabor wavelets for face recognition were proposed by Li *et al.* [23]. The copula parameter estimator and bootstrap confidence intervals were derived by Calabrese and Osmetti [24]. A wind power forecasting error model based on dynamic copula theory was proposed by Li *et al.* [25] to obtain accurate error intervals for forecasts. Liu and Liu [26] applied copula theory to build the joint probability distribution between a power and capacity of an energy storage system. Based on the stochastic interdependence between random variables such as the wind speed and load demand, the probability distribution of correlated random variables was established by Yu and Ghadimi [27]. A random deviation from a Dempster Shafer structure (DSS) or the p-box can be considered, and take the interval formed by the cumulative likelihood and the quantile of the cumulative confidence function as the sample [28]. Hence, these theories of constructing a correlation model can also be applied to the p-box because of the corresponding transformation relationship between the p-box structure and the correlation structure.

Based on the inversion of Sklar's theorem, the binary copula function is directly obtained from the binary joint distribution function. This article based the function method proposes a new class of functions i.e., functions G; based on the definition of function G, a new method of constructing binary Copula function is proposed i.e., G-Copula function, which enriches the methods of constructing binary Copula functions. To reduce data cross between the different p-boxes, a model of probability box based on a joint cumulative distribution of the G-Copula function is described in the current study.

The remainder of this study is organized as follows. Basic notions of the p-box and the Copula function are first given in Section II. In Section III, the definition of function G is given, and some notation and concepts related to function G are introduced to study the construction of G-Copula function, followed by a correlation p-box (cp-box) model based on the joint cumulative distribution of G-Copula function. Fitting effects of the Frank-Copula, Clayton-Copula and G-Copula functions for bearing data are demonstrated based on deviation squares; a comparison is made with other diagnostic algorithms and in detail the novelty of the proposed cp-box method is explained in Section IV. Finally, conclusions are shown in Section V.

II. RELATED WORK

A. BASIC NOTIONS OF P-BOX

A cumulative probability distribution function (CDF) of a random variable X is not be expressed by a single curve, as the estimated value $\widehat{x}$ of X is not a single scalar value [2], [15]. Considering estimated value $\widehat{x} \in \left[\underline{\widehat{x}}_i, \overline{\widehat{x}}_i\right]$, upper and lower bounds of the CDF is given by:

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

$$\underline{F}(x) = \underline{P}(X \leq x) \tag{1b}$$

where $\overline{F}(x)$ and $\underline{F}(x)$ are upper and lower bounds of the CDF, respectively; $\underline{P}$ denotes a lower probability measure. $\left[\underline{F}(x), \overline{F}(x)\right]$ is called as the p-box; then the random variable X with 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.

In Fig. 1, the area of ① corresponds to the lower probability measure $\underline{P}(X > x)$ of Eq. (1a); the area of ④ corresponds to $\underline{F}(x) = \underline{P}(X \leq x)$; the area of $1-$① corresponds to $\bar{F}(x) = 1 - \underline{P}(X > x)$, 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. BASIC NOTIONS OF COPULA FUNCTION

A vector of random variables $X = (X_1, X_2, \ldots, X_n)^{\mathrm{T}}$ with marginal distribution functions $(F_{X_1}(x_1), F_{X_2}(x_2), \ldots, F_{X_n}(x_n))$ is considered, and the marginal distribution functions can be satisfied by following expression [29]:

$$F(x_1, x_2, \ldots, x_n) = C(F_{X_1}(x_1), F_{X_2}(x_2), \ldots, F_{X_n}(x_n)) \tag{2}$$

where $F(x_1, x_2, \ldots, x_n)$ is the joint cumulative distribution function, and C is the Copula function. If the marginal distribution function $F_{X_i}(x_i)(i = 1, 2, \ldots, n)$ of the random

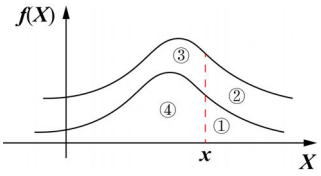

FIGURE 1. Schematic diagram of the p-box definition.

variable is a continuous function, then C can be uniquely determined. The probability density function $f(x_1, x_2, \ldots, x_n)$ of $F(x_1, x_2, \ldots, x_n)$ is given by

$$F(x_1, x_2, \ldots, x_n) = c(F_{X_1}(x_1), F_{X_2}(x_2), \ldots, F_{X_n}(x_n)) \times \prod_{i=1}^{n} f_{X_i}(x_i) \quad (3)$$

where $c(F_{X_1}(x_1), F_{X_2}(x_2), \ldots, F_{X_n}(x_n))$ and $f_{X_i}(x_i)(i = 1, 2, \ldots, n)$ are the probability density functions of the Copula and $X_i(i = 1, 2, \ldots, n)$, respectively.

III. CONSTRUCTIONS OF G-COPULA FUNCTION AND CP-BOX MODEL

A. CONSTRUCTION OF FUNCTION G

Two accelerometers are considered in this study, each of which can obtain bearing vibration signals. "In most cases, bearing signals collected from a sensor follow a random distribution because of the measurement errors of the sensor, the different positions of the measurements and the variability in working conditions [15]." Therebefore, the signals based one accelerometer is considered as a random variable, then a vector of random variables (X, Y) can be obtained from two accelerometers.

Definition 1: A binary function $C(u, v)$: $I^2 \to I = [0, 1]$ is called Copula, which satisfies: (1) boundary condition: $C(u, 0) = C(0, v) = 0$, $C(u, 1) = u$ and $C(1, v) = v$; (2) 2-increasing: $V_C([u_1, u_2] \times [v_1, v_2]) = C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$ for $\forall 0 \leq u_1 \leq u_2 \leq 1$, where $V_C([u_1, u_2] \times [v_1, v_2])$ is called the volume of C on rectangle $[u_1, u_2] \times [v_1, v_2]$ [30]. This volume is the second-order difference of C on rectangle $[u_1, u_2] \times [v_1, v_2]$, i.e., $V_C([u_1, u_2] \times [v_1, v_2]) = \Delta_{v_1}^{v_2} \Delta_{u_1}^{u_2} C(u, v)$.

Definition 2: Let φ be a continuous and absolutely decreasing convex function in $[0, 1] \to [0, \infty]$, where $\varphi(1) = 0$. The generalized inverse function of φ is expressed as $\varphi^{[-1]}$ within $[0, \infty] \to [0, 1]$, which can be defined as:

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & 0 \leq t \leq \varphi(0) \\ 0, & \varphi(0) \leq t \leq \infty \end{cases} \quad (4)$$

Then C is called Archimedes Copula, i.e., $C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v))$, and function φ is called the generator of Archimedes Copula function C.

Definition 3: Let $g(x)$ be function G, which satisfies the following conditions: (1) $g(x)$ is the increasing concave function within $[0, 1]$, i.e., $g^{\cdot}(x) \geq 0$, $g^{\cdot\cdot}(x) \leq 0$, $g(0) = 0$, and $g(1) = 1$, where superscript '$\cdot$' indicates derivation, similar in other places; (2) Let $2g^{\cdot}(x) + xg^{\cdot\cdot}(x) \leq 0$ for all $x \in [0, 1]$.

Definition 4: Let $F(x)$ and $G(y)$ be marginal distribution of the random variables X and Y, respectively, and its joint distribution function is Copula $C(u, v)$. For Copula $C(u, v)$, an upper dependence coefficient λ_U and a lower tail dependence coefficient λ_L are respectively defined as:

$$\lambda_U = \lim_{u \to 1^-} \frac{1 - 2u + C(u, u)}{1 - u} \quad (5a)$$

$$\lambda_L = \lim_{u \to 0^+} \frac{C(u, u)}{u} \quad (5b)$$

For λ_U and λ_L within $(0, 1]$, between X and Y satisfy upper tail dependence and the lower tail dependence; as λ_U and λ_L are zero, between X and Y are upper tail independence and the lower tail independence.

This article is based on function G to study the construction of binary Copula function, so the research of function G is important. Before processing, some notation and concepts related to function G must be introduced in here (see Appendix A for the proof):

(1) the weighted linear combination of function G is still function G, i.e., $g_i(x)$, $(i = 1, 2, \ldots, n)$ is a function G, then $\sum_{i=1}^{n} \lambda_i g_i(x)$ is still a function G, where $\sum_{i=1}^{n} \lambda_i = 1$ and $\lambda_i > 0$;

(2) the expansion transformation of function G is still function G, i.e., $g_1(x)$ is a function G, then $g_2(x) = g_1(\alpha x)/g_1(\alpha)$ is still a function G for $\forall 0 < \alpha \leq 1$;

(3) any composition of function G is still function G, i.e., let $g_1(x)$ and $g_2(x)$ be functions G, then $g(x) = g_2(g_1(x))$ is still function G.

Then, the procedure of a construction method of function G is described by the following steps. Firstly, based on the definition 3, the following expression is considered:

$$2y^{\cdot} + xy^{\cdot\cdot} = a(x) \quad (6)$$

where $a(x) \leq 0$ for all $x \in [0, 1]$. Then, its homogeneous linear differential equation is expressed as $2y^{\cdot} + xy^{\cdot\cdot} = 0$; two special solutions of the homogeneous equation are $y_1 = 1/x$ and $y_2 = 1$, respectively. The general solution of Eq. (6) satisfies $y = C_1(x)\frac{1}{x} + C_2(x)$; substituting the general solution into Eq. (6), we have $2C_2^{\cdot}(x) + C_1^{\cdot\cdot}(x) + xC_2^{\cdot\cdot}(x) = a(x)$; the supplementary condition is considered in here, i.e., $C_1^{\cdot}(x)y_1 + C_2^{\cdot}(x)y_2 = 0$, followed by equations:

$$\begin{cases} C_1^{\cdot}(x)\dfrac{1}{x} + C_2^{\cdot}(x) = 0 \\ 2C_2^{\cdot}(x) + C_1^{\cdot\cdot}(x) + xC_2^{\cdot\cdot}(x) = a(x) \end{cases} \quad (7)$$

Solving Eq. (7), we have $C_1(x) = -\int xa(x)\mathrm{d}x + \gamma_1$ and $C_2(x) = \int a(x)\mathrm{d}x + \gamma_2$, where γ_1 and γ_2 are constants. Finally, the general solution of $g(x)$ can be expressed as $g(x) = C_1(x)\frac{1}{x} + C_2(x)$.

Employing above procedure, functions G, i.e., $g(x)$, can be given based on the general solution of Eq. (6), as $a(x) \leq 0$ can be obtained within $[0, 1]$.

B. CONSTRUCTION OF G-COPULA FUNCTION

Based on function G, a new method for constructing binary Copula function can be given by following procedure. Let $g(x)$ is a function G, i.e., $g(x)$ satisfies the boundary conditions:

$$C(u,0) = u \cdot 0 \cdot g(u) + ug(0) - ug(u)g(0) = 0 \tag{8a}$$
$$C(0,v) = 0 \cdot v \cdot g(0) + 0 \cdot g(v) - 0 \cdot g(0)g(u) = 0 \tag{8b}$$
$$C(u,1) = ug(u) + ug(1) - ug(u)g(1) = ug(u) + u - ug(u) = u \tag{8c}$$
$$C(1,v) = 1 \cdot vg(1) + 1 \cdot g(v) - g(1)g(v) = v + g(v) - g(v) = v \tag{8d}$$

and satisfies 2-increasing: for $\forall u_1 \le u_2$ and $u_1, u_2, v_1, v_2 \in I$, we have

$$\begin{aligned} &V_C([u_1,u_2]\times[v_1,v_2]) \\ &= C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \\ &= (v_2-v_1)(u_2g(u_2) - u_1g(u_1)) + (g(v_2) - g(v_1)) \\ &\quad + [u_1(g(u_1)-u_1) - u_2(g(u_2)-u_2)] \end{aligned} \tag{9}$$

For $\forall u_1 \le u_2, v_1 \le v_2$, we have $(v_2 - v_1)(u_2g(u_2) - u_1g(u_1)) \ge 0$ and $g(v_2) - g(v_1) \ge 0$, because $g(x)$ is increasing within $[0, 1]$. Let $\omega = (u_1g(u_1) - u_1) - (u_2g(u_2) - u_2)$, and $h(x) = xg(x) - x$ for all $x \in [0, 1]$, where $h(x)$ is monotonically decreasing. Then $\omega = h(u_1) - h(u_2)$, $h'(x) = g(x) + xg'(x) - 1$ and $h''(x) = 2g'(x) + xg''(x)$. Based on the condition $2g'(x) + xg''(x) \le 0$ for all $x \in [0, 1]$, $h'' \le 0$ can be inferred; then $h'(x)$ is monotonically decreasing within $[0, 1]$, and $h'(x)$ is less than or equal $h'(0)$ in which $h'(0) = -1 < 0$. For $\omega = h(u_1) - h(u_2)$ and $u_1 \le u_2$, ω is greater than or equal zero; for $\forall u_1 \le u_2$, $\forall v_1 \le v_2$ and $u_1, u_2, v_1, v_2 \in I$, we have $V_C([u_1, u_2] \times [v_1, v_2]) \ge 0$; then $C(u, v)$ is 2-increasing. Hence, the $C(u, v)$ is a binary Copula function, i.e., G-Copula function, and can be expressed by $C(u, v) = uvg(u) + ug(v) - ug(u)g(v)$, where the function g is the generator of G-Copula function.

Similarly, G-Copula function can be also given by following procedure: (1) boundary conditions are the same as Eq. (8); (2) 2-increasing: for $\forall u_1 \le u_2$ and $u_1, u_2, v_1, v_2 \in I$, we have

$$\begin{aligned} &V_C([u_1,u_2]\times[v_1,v_2]) \\ &= C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \\ &= (v_2-v_1)(u_2g(u_2) - u_1g(u_1)) \\ &\quad + (u_2-u_1)(g(v_2)-g(v_1)) + u_1(g(u_1v_2) - g(u_1v_1)) \\ &\quad - u_2(g(u_2v_2) - g(u_2v_1)) \end{aligned} \tag{10}$$

For $\forall u_1 \le u_2, v_1 \le v_2$, we have $(v_2 - v_1)(u_2g(u_2) - u_1g(u_1)) \ge 0$ and $(u_2 - u_1)(g(v_2) - g(v_1)) \ge 0$, because $g(x)$ is increasing within $[0, 1]$. Let $\omega = u_1(g(u_1v_2) - g(u_1v_1)) - u_2(g(u_2v_2) - g(u_2v_1))$ and $h(x) = x(g(xv_2) - g(xv_1))$ for all $x \in [0, 1]$. Then $h(u_1) - h(u_2)$ is ω; $[g(xv_2) + xv_2g'(xv_2)] - [g(xv_1) + xv_1g'(xv_1)]$ is $h'(x)$. Let $u(x, y)$ be $g(xy) + xyg'(xy)$ for $\forall x, y \in [0, 1]$, then $h'(x)$ satisfies $u(x, v_2) - u(x, v_1)$, and $u_y(x, y)$ satisfies $x(2g'(xy)+xyg''(xy))$. Based on the condition $2g'(x) + xg''(x) \le 0$ for all $x \in [0, 1]$, $2g'(xy) + xyg''(xy) \le 0$ is obtained, and $u_y(x, y)$ is less than or equal zero for $\forall x, y \in [0, 1]$, i.e., $u(x, y)$ is monotonically decreasing about y for all x. Based on $v_1 \le v_2$, $h'(x)$ is less than or equal zero, and satisfies $u(x, v_2) - u(x, v_1)$ for all $x \in [0, 1]$. Therebefore, $h(x)$ is monotonically decreasing within $[0, 1]$, and ω is greater than or equal zero in which $\omega = h(u_1) - h(u_2)$. For $\forall u_1 \le u_2$, $\forall v_1 \le v_2$ and $u_1, u_2, v_1, v_2 \in I$, we have $V_C([u_1, u_2] \times [v_1, v_2]) \ge 0$, then $C(u, v)$ satisfies 2-increasing. Hence, the $C(u, v)$ is also G-Copula function, and can be expressed by $C(u, v) = uvg(u) + ug(v) - ug(uv)$.

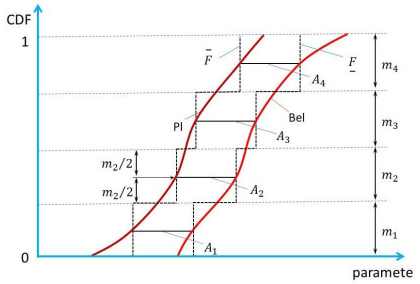

FIGURE 2. Average discretization method of the cp-box.

C. ESTABLISHMENT OF CP-BOX MODEL BASED ON G-COPULA FUNCTION

An unknown parameter of G-Copula function can be estimated by moment method or plausibility function. Then, confidence interval for the joint cumulative distribution of G-Copula function can be constructed based on the Kolmogorov-Smirnov method [31], [32]. Considering the confidence level α, the confidence bounds can be expressed as:

$$\underline{F}_\beta(\beta) = \max\{\mathrm{Pl}(\beta) - \varepsilon, 0\} \tag{11}$$
$$\overline{F}_\beta(\beta) = \max\left\{\underline{F}_\beta(\beta) + \varepsilon, 1\right\} \tag{12}$$

where $\mathrm{Pl}(\beta)$ is the plausibility bound of the joint cumulative distribution, ε is confidence capacity and satisfies $\varepsilon = \sqrt{\ln(2/\beta)/2}$. The cp-box model can be composed by the plausibility upper bound $\overline{F}_\beta(\beta)$ and belief lower bound $\underline{F}_\beta(\beta)$.

According to the average discretization method, subintervals of equal length $1/n$ are used in this study [15]. Based on Fig. 2, a focal element interval A with a mass function $m(A)$ for the cp-box can be expressed as:

$$A_{i,j} = \left[\mathrm{Pl}_i^{-1}\left(\frac{1}{n}\left(j-\frac{1}{2}\right)\right), \mathrm{Bel}_i^{-1}\left(\frac{1}{n}\left(j-\frac{1}{2}\right)\right)\right], \quad j = 1, \ldots, n \quad m_i(A_{i,j}) = \frac{1}{n} \tag{13}$$

where Pl_i^{-1} and Bel_i^{-1} are inverse functions of the plausibility (Pl) function and the belief (Bel) function, respectively.

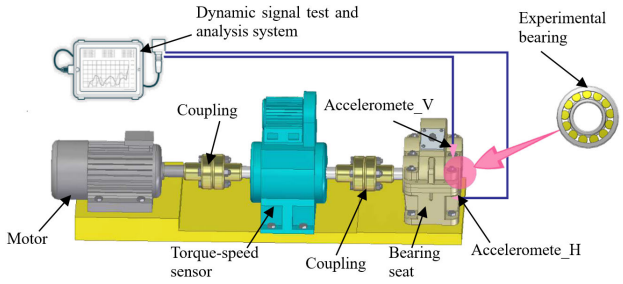

FIGURE 3. Schematic block diagram of the rotating machinery test rig.

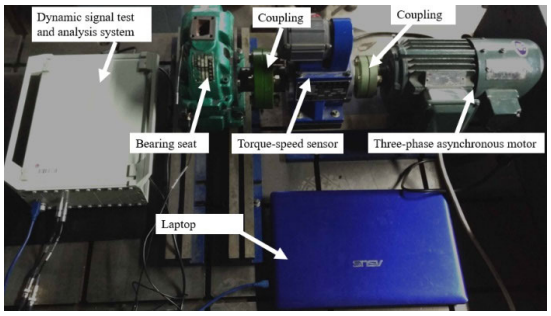

FIGURE 4. Layout of the rotating machinery test rig.

A discrete DSS is related to the interval with mass value by the following expression:

$$\left([\overline{x_i}, \underline{x_i}], \frac{1}{n}\right), \quad (i = 1, 2, \ldots, n)$$

where "$\overline{x_i} \leq \underline{x_i}$; additionally, $\underline{x_i} \neq \underline{x_{i-1}}$ when $\overline{x_i} = \overline{x_{i-1}}$. The width of different focal element interval 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 the cp-box can be obtained by accumulating all uncertainty probability of the DSS. The basic feature called an aggregated width can be expressed as":

$$\omega_1 = \sum_{i=1}^{n} (m_i \times |\overline{x_i} - \underline{x_i}|) \tag{14}$$

"where $\overline{x_i}$ and $\underline{x_i}$ are the upper and lower bounds of the focal element interval respectively; m_i is the mass value of the corresponding focal element interval".

"Similarly, in order to measure more cp-box information, more features can be expressed as":

$$\omega_2 = \sum_{i=1}^{n} (m_i \times \log_2 |\overline{x_i} - \underline{x_i}|) \tag{15}$$

$$\omega_3 = \sum_{i=1}^{n} (m_i \times \log_2(1 + |\overline{x_i} - \underline{x_i}|)) \tag{16}$$

$\omega_i (i = 2, 3)$ is also the scalar values.

IV. RESULTS AND ANALYSIS

To demonstrate the validity of the method proposed in this article, the method is applied to a pattern recognition system of the experimental signal of rotating machinery including the bearing fault. Schematic block diagram of the rotating machinery test rig as shown in Fig. 3.

The physical layout of the test rig is shown in Fig. 4. The three-phase asynchronous motor (Siemens, 3∼, 3.0hP) is connected to the main shaft (Φ30 mm) through the coupling. The torque-speed sensor is installed between the shaft to provide load. Two accelerometers (PCB ICP M603C01) are arranged in the horizontal (i.e., _H) and vertical (i.e., _V) directions to pick up the vibration signals. The experimental

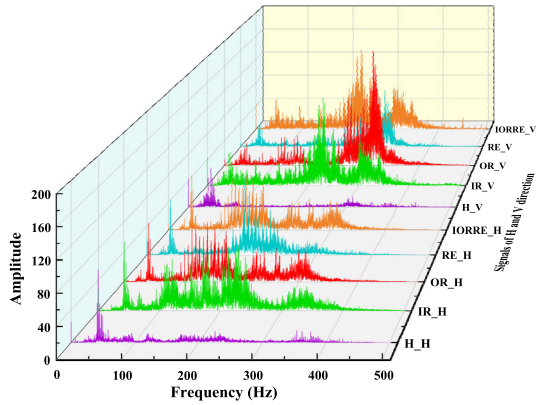

FIGURE 5. The amplitude of the bearing signals under eight conditions.

bearing is a 30305 SKF tapered roller bearing; the number of balls is 13, rolling element diameter 9.06 mm, contact angle 28°, bearing pitch diameter 44.6 mm, respectively. Single faults, which are 0.5 mm deep and 0.5 mm wide, are processed in the inner race, the outer race and the rolling elements to simulate bearing failure by using wire cut electrical discharge machining technology [33], [34]. The sampling frequency and sampling time are 1000 Hz and 5.5s, respectively. The average value of motor speed is 1085.7 r/min. Before proceeding, a handheld sensor (YE5501 sensitivity calibrator) was used to calibrate the ICP accelerometers; a laptop with LabVIEW software was used to control the test rig.

Tang *et al.* [15] point out "the uncertainty of bearing signal (i.e. the unknown signal transmission path) is the main feature of bearing signal, because it always exists and changes from the generation of health signal to the formation of fault signal; the collection of bearing signal uncertainty provides a new method for rolling bearing diagnosis". Therebefore, for the bearing signals without preprocessing techniques (i.e., wavelet transform of the vibration signals, high-pass filtration, bandpass filtration, etc.), the amplitude-frequency signals under five conditions along the horizontal and vertical directions are shown in Fig. 5, where "H stands for healthy bearing, IR stands for inner race fault, OR stands for outer race fault, RE stands for rolling element fault, and IORRE stands for inner race, outer race with rolling element faults".

Comparing the bearing signals obtained along the horizontal and vertical directions, we infer that the bearing data collected under the same bearing conditions have duplication and complementary information features in Fig. 5. This result may be because the accelerometer arranged in the vertical direction is located in the nonload-bearing region of the bearing pedestal and the radial clearance between the rolling elements and the inner race and outer race passing through this region is very small; thus, the impact of the rolling elements and outer race fault and the impact of the rolling elements and inner race fault yield similar signals. However, the accelerometer arranged in the horizontal direction is close to the load-bearing region of the bearing pedestal, and the radial clearance between the rolling elements and the inner and outer races changes considerably due to the bearing load. Hence, the bearing signals in the horizontal and vertical directions are similar, but there are also obvious differences.

A. ESTABLISHMENT OF MARGINAL DISTRIBUTION

The health bearing data along the horizontal and vertical directions are considered as the random variable X and Y, respectively. To reduce fluctuations of the signal over time, the health bearing data are divided into 30 groups and the average value of each group is calculated. Then, acceleration sequences of the health bearing data for the horizontal and vertical directions are $\{x_i, i = 1, 2, \ldots, 30\}$ and $\{y_i, i = 1, 2, \ldots, 30\}$, respectively; eight data are used as segmentation points, and the whole number axis is divided into nine segments. empirical distributions can be given by:

$$F_n(x) = \begin{cases} 0, & x < 0.0031; \\ 1/15, & 0.0031 \le x < 0.0146; \\ 1/6, & 0.0146 \le x < 0.0186; \\ 1/3, & 0.0186 \le x < 0.0529; \\ 13/30, & 0.0529 \le x < 0.0647; \\ 2/3, & 0.0647 \le x < 0.1182; \\ 13/15, & 0.1182 \le x < 0.2163; \\ 29/30, & 0.2163 \le x < 0.3821; \\ 1, & x \ge 0.3821; \end{cases},$$

$$G_n(y) = \begin{cases} 0, & y < -0.0851; \\ 1/10, & -0.0851 \le y < 0.0184; \\ 1/6, & 0.0184 \le y < 0.0285; \\ 11/30, & 0.0285 \le y < 0.0623; \\ 1/2, & 0.0623 \le y < 0.0902; \\ 2/3, & 0.0902 \le y < 0.1472; \\ 13/15, & 0.1472 \le y < 0.2337; \\ 29/30, & 0.2337 \le y < 0.3596; \\ 1, & y \ge 0.3596; \end{cases}$$

The joint distribution of two random variables is generated by their respective marginal distribution and corresponding copula function based on the Sklar's theorem [30]. The health bearing data are analyzed, and frequency histograms are simulated along the horizontal and vertical directions in the health bearing data. Unknown parameters of the distributions can be estimated by moment method; and error sum of squares for theoretical values and empirical values of the distributions can be calculated in which the distribution corresponding to the minimum error sum of squares is the distribution function of the health bearing data. Then, the frequency histograms are simulated along the horizontal and vertical directions as shown in Fig. 6.

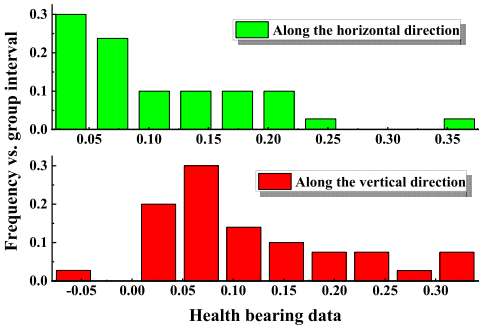

FIGURE 6. The frequency histograms of the health bearing data; top: along the horizontal direction; bottom: along the vertical direction.

TABLE 1. Parameter Estimation and Error Sum of Squares for the Distribution Functions of the Bearing Data Along Different Directions.

Health bearing data	Theoretical distribution	Estimation of distribution parameters	Error sum of squares
Horizontal direction	Exp(λ)	9.5974	12.2061
	$Ga(\theta, \mu)$	(1.3213, 12.7073)	12.6311
Vertical direction	$Ga(\theta, \mu)$	(1.2028, 10.7352)	12.1226
	$\chi^2(n)$	0.1120	3.6994

Based on Fig. 6, the exponential distribution Exp(λ) and the gamma distribution $Ga(\theta, \mu)$ are respectively used to simulate the overall density function of the health bearing data along the horizontal direction; the gamma distribution $Ga(\theta, \mu)$ and the chi-square distribution $\chi^2(n)$ are respectively used to simulate the overall density function of the data along the vertical direction. Using sample mean instead of population mean and sample variance instead of population variance, the parameters of each distribution are estimated by moment method; then the distribution function of X and Y corresponding to the minimum sum of error squares of each distribution and the corresponding empirical distribution of the bearing data is selected. Parameter estimation and error sum of squares for the distribution functions of the bearing data along different directions are listed in Table 1.

In Table 1, for the bearing data along the horizontal direction, the error sum of squares for the exponential distribution and its empirical distribution is the smallest; the error sum of squares for the chi-square distribution and its empirical distribution is the smallest along the vertical direction. Therefore, the exponential distribution and the chi-square distribution are selected to establish the marginal distributions of

TABLE 2. Three Types of Copula Functions.

Function type	Function form	Generator	Parameter range
Frank-Copula	$-\frac{1}{\eta}\ln\left(1+\frac{\left(e^{-\eta u}-1\right)\left(e^{-\eta v}-1\right)}{e^{-\eta}-1}\right)$	$\varphi(t)=-\ln\left(\frac{e^{-\eta t}-1}{e^{-\eta}-1}\right)$	$\eta\in(-\infty,0)\cup(0,\infty)$
Clayton-Copula	$\max\left(\left(u^{-\alpha}+v^{-\alpha}-1\right)^{-\frac{1}{\alpha}},0\right)$	$\varphi(t)=\frac{1}{\alpha}\left(t^{-\alpha}-1\right)$	$\alpha\in(-1,0)\cup(0,\infty)$
G-Copula	$\frac{\beta}{6}u^3v-\frac{\beta}{6}u^3v^2+\frac{\beta}{6}uv^2-\frac{\beta}{6}uv+uv$	$g(x)=\frac{\beta}{6}\left(x^2-1\right)+1$	$\beta\in\left(-\infty,0\right]$

the health bearing data. Based on the establishment of the marginal distribution for the health bearing data, marginal distributions of the other bearing conditions (i.e., IR, OR, RE and IORRE) were established; the calculation results showed that chi-square distribution, exponential distribution, chi-square distribution and gamma distribution were selected to establish the marginal distribution of IR, OR, RE and IORRE respectively.

B. PARAMETER ESTIMATION OF COPULA FUNCTION MODEL

Archimedes Copula was widely used in many fields because of its simple structure, convenient construction. In this study, Frank-Copula, Clayton-Copula function and G-Copula function are used to measure the correlation of the health bearing data along the horizontal and vertical directions. The generator and parameter range of three types of Copula functions are listed in Table 2.

Estimating the unknown parameters in Frank-Copula, Clayton-Copula and G-Copula functions, concrete expressions of the Copula functions can be constructed. Based on the acceleration sequences $\{x_i, i = 1, 2, \ldots, 30\}$ and $\{y_i, i = 1, 2, \ldots, 30\}$, Kendall rank correlation coefficient $\widehat{\tau}$ of samples for (X, Y) can be calculated by:

$$\widehat{\tau}=\frac{2}{n(n-1)}\sum_{1\le i<j\le n}\operatorname{sign}(x_i-x_j)(y_i-y_j) \tag{17}$$

where $\operatorname{sign}(x)$ is a sign function, and calculated by:

$$\operatorname{sign}(x)=\begin{cases}1, & x>0\\ 0, & x=0\\ -1, & x<0\end{cases} \tag{18}$$

Kendall rank correlation coefficient of (X, Y) for the health bearing data is $\widehat{\tau} = 0.5310$. For general binary Copula function, the overall Kendall rank correlation coefficient τ and Copula function $C(u, v)$ can be expressed as:

$$\tau=1-4\iint_{I^2}\frac{\partial C(u,v)}{\partial u}\frac{\partial C(u,v)}{\partial v}\mathrm{d}u\mathrm{d}v \tag{19}$$

For the health bearing data, the overall Kendall rank correlation coefficient τ is $-\frac{\beta}{18}$ based on the G-Copula function. For special Copula functions (Archimedes Copula), the overall Kendall rank correlation coefficient τ and the generator $\varphi(t)$ of Copula function $C(u, v)$ can be expressed as:

$$\tau=1+4\int_0^1\frac{\varphi(t)}{\varphi^{\cdot}(t)}\mathrm{d}t \tag{20}$$

Using Frank-Copula and Clayton-Copula functions, we can calculate the overall Kendall rank correlation coefficients of the (X, Y) i.e., $\tau=\frac{\alpha}{\alpha+2}$ and $\tau=1+\frac{4[D_1(\eta)-1]}{\eta}$, where $D_1(\eta)$ is given by:

$$D_1(\eta)=\frac{k}{\eta}\int_0^{\eta}\frac{t}{e^t-1}\mathrm{d}t \tag{21}$$

Based on the nonparametric estimation, Kendall rank correlation coefficient of the sample is used as the estimated value of the overall Kendall rank correlation coefficient, i.e., $\widehat{\tau}=\tau$. Then, the estimated values of parameters are obtained in different Copula functions. The estimated values of Frank-Copula, Clayton-Copula and G-Copula functions are calculated as: $\eta = 6.3301$, $\alpha = 2.2644$ and $\beta = -8.8465$, respectively.

C. CONTRAST OF COPULA FUNCTION MODEL

After the unknown parameters are estimated in different Copula functions, it is necessary to establish a proper Copula function model. The optimum Copula function model can be obtained by comparing different Copula function. Transform the empirical distribution of the health bearing data along the horizontal and vertical directions into a sequence (u_i, v_i), where $u_i = F_n(x_i)$, $v_i = G_n(y_i)$, $i = 1, 2, \ldots, n$, and n denotes the sample size. The correlation of the health bearing data along the horizontal and vertical directions is analyzed by comparing the square deviation between Copula function and empirical Copula function. Calculate the value of the i-th empirical distribution function $C_n(u_i, v_i)$, which can be calculated by:

$$C_n(u_i,v_i)=\frac{1}{n}\sum_{i=1}^{n}I\left\{F_n(x_i)\le u, G_n(y_i)\le v\right\},\quad u,v\in[0,1] \tag{22}$$

The sum of the deviation squares between Copula function value and empirical Copula function value can be expressed as:

$$d^2=\sum_{i=1}^{n}\left(C_n(u_i,v_i)-C(u_i,v_i)\right)^2 \tag{23}$$

TABLE 3. Comparisons of Copula Model.

Function type	Function form	d^2
Frank-Copula	$-\frac{1}{6.3301}\ln\left(1+\frac{\left(e^{-6.3301u}-1\right)\left(e^{-6.3301v}-1\right)}{e^{-6.3301}-1}\right)$	0.0121
Clayton-Copula	$\max\left(\left(u^{-2.2644}+v^{-2.2644}-1\right)^{\frac{1}{2.2644}},0\right)$	0.0240
G-Copula	$-\frac{8.8465}{6}u^3v+\frac{8.8465}{6}u^3v^2-\frac{8.8465}{6}uv^2+\frac{8.8465}{6}uv+uv$	0.0113

TABLE 4. The Upper Tail Correlation Coefficient λ_U and the Lower Tail Correlation Coefficient λ_L Between the Horizontal Direction and the Vertical Direction.

Correlation coefficient	H	IR	OR	RE	IORRE
λ_U	0.8324	0.7938	0.7031	0.5431	0.7921
λ_L	0.5901	0.6642	0.7736	0.8011	0.7841

Then, for three Copula functions selected in this study, the sum of the deviation squares d^2 between the simulated and real values can be calculated. The fitting degree of Copula function increases with the decrease of d^2. Compare the three Copula functions, and the comparison results are listed in Table 3.

For X and Y, the deviation squares d^2 between G-Copula function and empirical Copula function is 0.0113, which is the smallest in the three Copula functions; the result indicated that G-Copula function own the optimum fitting effect for the health bearing data along different directions.

Similarly, we analyzed the other bearing conditions (i.e., IR, OR, RE and IORRE), and the experimental results showed that the fitting results were similar to those of the healthy bearing data. Hence, G-Copula function are used to measurement the bearing data correlation between the horizontal direction and the vertical direction. According to the expression of G-Copula function, the upper tail correlation coefficient λ_U and the lower tail correlation coefficient λ_L between the horizontal direction and the vertical direction are calculated respectively in the bearing data, as shown in Table 4.

The upper and lower tails for the bearing data along the horizontal and vertical directions, with obvious tail correlation, are gradually dependent in Table 4. It shows that when the output value of one direction increases or decreases to a certain extent, it will cause the output value of another direction to fluctuate.

D. MEASUREMENT CP-BOX MODEL

The concrete expression of G-Copula function is applied in this case based on the estimated parameter, because G-Copula function is better to measure the correlation of the bearing data. Considering the confidence intervals of the joint cumulative distribution in which the confidence level is 95%, the focal element interval $A_{i,j}$ can be obtained (calculated with Eq. (12)). Accumulating values of the mass function with $A_{i,j}$, the cp-boxes of H, IR, OR, RE and IORRE can be obtained as shown in Fig. 7.

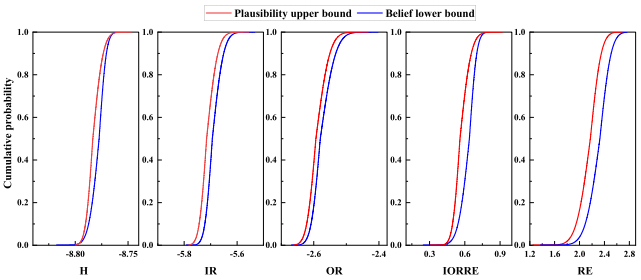

FIGURE 7. The cp-boxes of the different bearing conditions.

TABLE 5. The Information of the Aggregated Uncertainty Measurement for the cp-Boxes.

Bearing status	H	IR	OR	RE	IORRE
ω_1	0.0108	0.0147	0.0378	0.0894	0.0462
ω_2	-7.6589	-4.8384	-2.9885	-3.8976	-3.4465
ω_3	0.1081	0.1172	0.1234	0.1786	0.1393

The cp-boxes are composed of the computed on the basis of the plausibility upper and belief lower bounds as shown in Fig. 7. It should be noted that there is no data cross between the different cp-boxes, which means that the cp-boxes are better to distinguish in the next pattern recognition system. Before proceeding, it is necessary to measure those cp-boxes. the information of the aggregated uncertainty measurement are obtained using Eqs. (13) to (15), as listed in Table 5.

The values of the aggregated width ω_1 are less than one in Table 5, then the values of ω_2 are less than zero because of an effect of $\log_2$ (see Eq. (15)). The information of the aggregated uncertainty measurement can be considered as features of those cp-boxes to input a pattern recognition system.

E. PATTERN RECOGNITION FOR CP-BOX FEATURES

This case addresses the SVM validation in which the cp-box feature is present in Table 5. "The correct recognition rate is used as the target. A total of 100 cp-boxes were obtained from the bearing signals from each condition. Thus, 500 cp-boxes were obtained for all five conditions. The feature vector sets from the 500 cp-boxes were calculated. For a Python environment-based SVM classifier, a sigmoid function was selected as the kernel function, and the parameters C and gamma were set to [1e3, 5e3, 1e4, 5e4, 1e5] and [1e-4, 5e-4, 1e-3, 5e-3, 1e-2, 1e-1], respectively. The parameters C and gamma were automatically optimized [15]. The input data were divided into training, validation and testing sets; the number of feature vectors for each set can be represented as

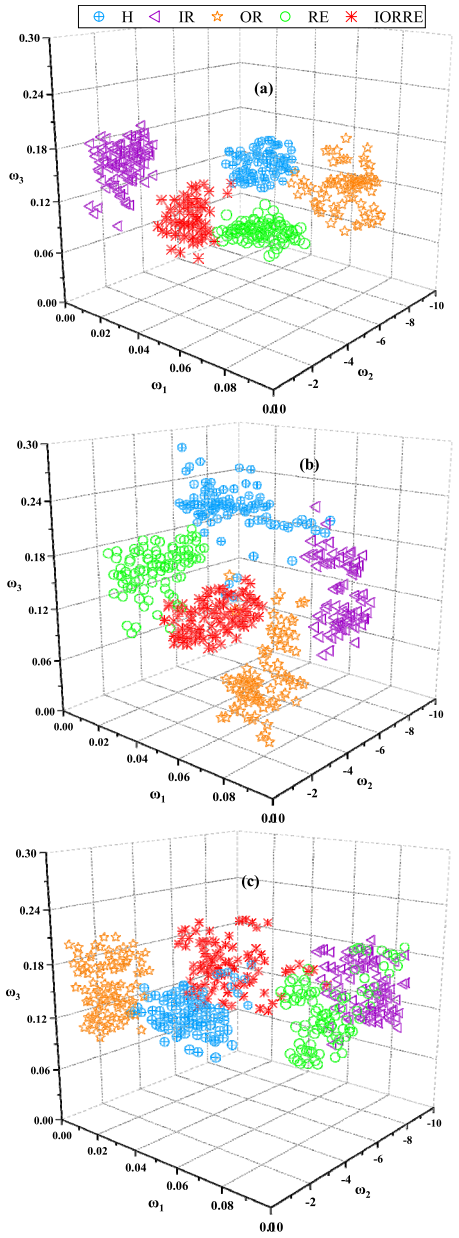

FIGURE 8. The classification performances of the different cp-boxes; (a) cp-box based G-Copula function; (b) cp-box based Frank-Copula function; (c) cp-box based Clayton-Copula function.

{300, 100, 100}''. Then, 10-fold cross-validation was used before the SVM validation. Fig. 8 presents case distribution of five data sets on three-dimensional space, which is made up of ω_1, ω_2 and ω_3. Additionally, the classification performances of the cp-boxes based on Frank-Copula and Clayton-Copula functions are also displayed in Fig. 8.

The performance of the SVM model can be affected greatly by the input data. With regards to bearing fault diagnosis, the different fault signal (targeting fault) can be identified if the input data are independent. The classification performance of the cp-boxes based G-Copula function is better than the cp-boxes based Frank-Copula and Clayton-Copula functions, as shown in Fig. 8. For G-Copula, Frank-Copula and Clayton-Copula functions, correctly classified cases are 96.1%, 88.1% and 79.4%, respectively. This is because that there is no data cross between the different cp-boxes based G-Copula function, which increases the independent of the input data.

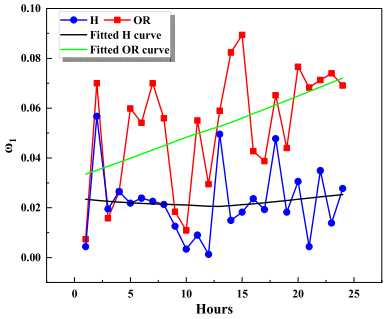

FIGURE 9. Condition monitoring based on the aggregated width.

Additionally, the aggregated width ω_1 can be applied to prevent the bearing failure. The vibration signals of OR are used as an example, because the outer race with the centripetal force exerted by the rolling element is easily failure [35]–[37]. Having recorded vibration data of H and OR for 24 hours, the corresponding aggregated width ω_1 has been obtained from the cp-boxes based G-Copula function, and plotted in Fig. 9.

An exponential function is used to fit the curves of H and OR in Fig. 9. The fitted OR curve has the rising trend which can be considered as a representment of the development for the bearing outer race failure. As complementary, the fitted H curve is also plotted in Fig. 9, which has a relative stable trend.

F. COMPARISON OF DIFFERENT DATA PROCESSING METHODS

''A comparison of different data processing study should be undertaken to demonstrate the advantage of the method proposed by this article''. For vibration signals, wavelet analysis method, a p-box modeling method of raw bearing data and a composite multiscale fuzzy entropy method have been demonstrated to be effective for bearing fault diagnosis [15], [38], [39]. There should be some rules in the comparison as follows: (1) the same data should be applied to different algorithms; (2) the experimental data used should be converted to unit values; (3) the same SVM model is chosen as a pattern recognition tool; (4) the results of classification performance are expressed in percentile (%).

The wavelet analysis has been applied in signal processing for more than fifty years [38]. For the raw vibration acceleration signals along the horizontal direction, the wavelet DB4 is used to filter the first two level detail signals with 50% period expansion. Fig. 10 presents the time domain signals before and after noise reduction.

Morlet wavelet is used for wavelet multiscale decomposition i.e.,

$$\psi(t) = \sqrt[4]{\frac{1}{\pi}} \cos(kt)\, \mathrm{e}^{-\frac{t^2}{2}} \tag{24}$$

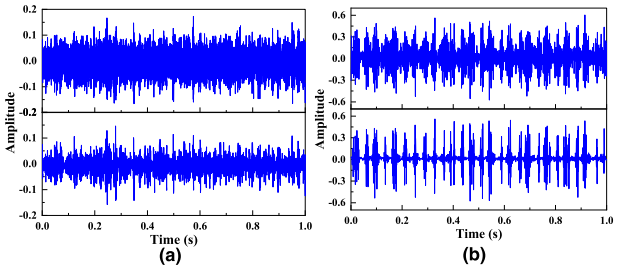

FIGURE 10. The time domain signals before and after noise reduction; (a) H; (b) OR.

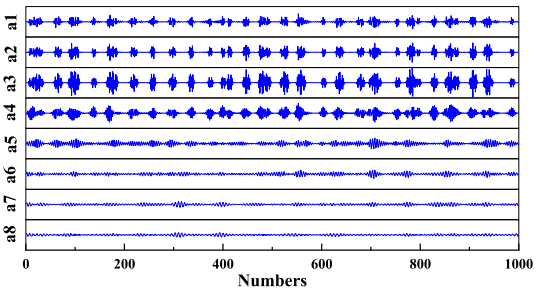

FIGURE 11. Wavelet multiscale decomposition for the OR signals after noise reduction.

where the value of wave number k is five in this study. The result of wavelet multiscale decomposition for the OR signals after noise reduction is shown in Fig. 11, where the scale factor $ai = \{1 : 1.25 : 9.75\}$.

Before proceeding, the raw vibration acceleration signals should be demonstrated to be stationary signals, because amplitude spectrum analysis cannot be applied directly in nonstationary signals. For the discrete Fourier transform, the number of sampling points per segment 4096 (i.e., 2^{12}) is used based the current sampling frequency and sampling time, and the minimum frequency resolution is $1/4.096 = 0.244$Hz. Hence, according to the minimum frequency resolution and characteristic frequency of the outer race (see Appendix B), the motor speed variation should be limited in 4 r/min. Fig. 12 presents the motor speed variation with the time.

The smoothed motor speed fluctuation is limited in 4 r/min as shown in Fig. 12. To extract the fault characteristic frequency at a certain motor speed, the low frequency signals in Fig. 11 is reconstructed, and its amplitude spectrum is shown in Fig. 13.

The characteristic frequency of OR at a certain motor speed is 97.1 Hz as shown in Fig. 13, which means that the effective OR signals can be reflected greatly by the low frequency signals. Based on the above wavelet analysis, the effective bearing signals along the horizontal and vertical directions for H, OR, IR, RE and IORRE were obtained. Then, time-domain statistical features (i.e., range, mean value, standard deviation, skewness, kurtosis and crest factor) of the effective bearing signals were extracted to input into the SVM model. The correct classification of faults and the detailed accuracy of each class are given in Table 6.

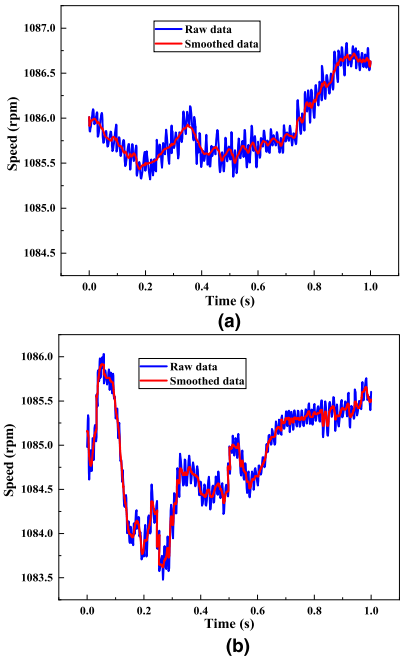

FIGURE 12. The motor speed variation with the time; (a) H; (b) OR.

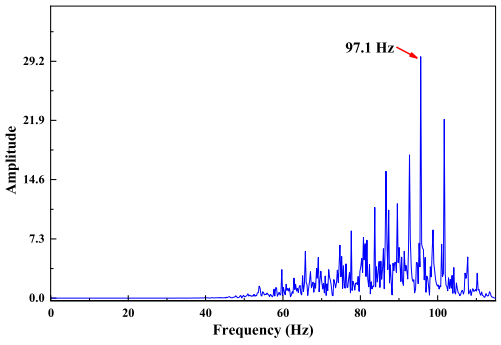

FIGURE 13. The amplitude spectrum of OR.

TABLE 6. Evaluation of the Success of the Numeric Prediction and the Detailed Accuracy by Class of the SVM Model.

Classified as	TP rate	FP rate	Precision	Recall	F-measure
H	0.85	0.0036	0.9714	0.85	0.9067
IR	0.75	0.0143	0.8824	0.75	0.8108
OR	0.75	0.0486	0.6881	0.75	0.7177
RE	0.75	0.0450	0.7042	0.75	0.7264
IORRE	0.75	0.0379	0.7389	0.75	0.7444

Parameters	Values (SVM)
Correctly classified cases	77.5%
Incorrectly classified cases	22.5%
Kappa statistic	0.7429
Mean absolute error	0.225
Root mean squared error	0.2278
Relative absolute error	77.89%
Root relative squared error	77.97%

"The correct recognition rate is 77.5% in Table 6. However, compared to the correct recognition rate 96.1% from the proposed method in this article, there is still room for

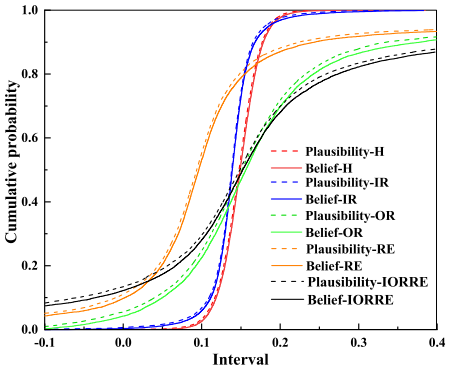

FIGURE 14. The results of the p-box modeling method of raw bearing data.

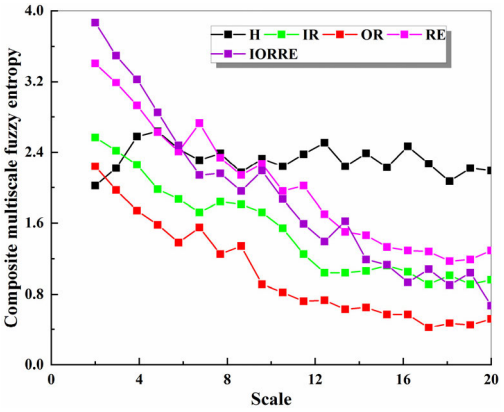

FIGURE 15. Composite multiscale fuzzy entropy of the different bearing data.

improvement. It is may be because that the feature extraction of bearing signal leads to the loss of statistical information beyond the feature''.

''The p-box modeling method of raw bearing data can directly establish the p-box model for the raw bearing data based on the definition of the p-box [15]. First, the maxima and minima from each column vector of bearing signals from the horizontal and vertical directions can be obtained by the p-box modeling method of raw bearing data. Then, the 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. Fig. 14 presents the p-boxes of H, OR, IR, RE and IORRE within $[-0.1, 0.4]$ based on the modeling method of raw bearing data''.

There is data cross between the different p-boxes in Fig. 14. The information of the aggregated uncertainty measurement for the p-boxes are obtained using Eqs. (13) to (15), and inputted into the SVM model. the experimental results showed that correct recognition rate is 86.1%, which is smaller than the correct recognition rate 96.1% from the proposed method in this article.

''Composite multiscale fuzzy entropy is an effective method to analyze the complexity of time series in bearing fault diagnosis. 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. 15 presents the results of the composite multiscale fuzzy entropy for each bearing condition based on the current data, where the value of largest scale is 20, embedding dimension 2, gradient of exponent function 2, and similarity tolerance 0.15SD (SD denotes standard deviation of raw bearing data), respectively [39]''.

In Fig. 15, ''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 following: Firstly, Total 1000 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% 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; the experimental results showed that the total correct recognition rate is 88.6%, which is smaller than the correct recognition rates 96.1% from the proposed methods in this article. It is because that the method of composite multiscale fuzzy entropy requires additional empirical effort in the bearing fault diagnosis [39]''.

V. CONCLUSION

This study presents a procedure for the detection of bearing faults by classifying them using a machine learning model, namely, the SVM model. The different Copula functions were used to analyze the correlation structures between random variables. Then, the cp-box models were established, and inputted into the SVM model. Meanwhile, some conclusions can be obtained as following:

A new G-Copula function has been established based on a new function G proposed in this article; in the research of the correlation structures of the bearing data, the experimental results showed that the fitting performance of G-Copula function was better than that of Frank-Copula and Clayton-Copula functions.

The cp-box models have been established based on the joint cumulative distribution of G-Copula function; In the classification performance analysis of the SVM model, the experimental results showed that there is no data cross between different the cp-boxes, which increases the independent of the input data for the SVM model. Additionally, the aggregated widths of the cp-boxes can be used to prevent effectively the bearing failure, which provided great benefits for applications of the condition monitoring program.

In a contrastive study of the correct recognition rate, the experimental results showed that the aggregated uncertainty measurement information of the cp-boxes yield the better classification accuracy than other diagnostic algorithms.

APPENDIX A

Proof concept (1): It is known that $g_i(x)$ is a function G, which satisfies all properties of function G; let $h(x) = \sum_{i=1}^{n} \lambda_i g_i(x)$, then

$$h'(x) = \sum_{i=1}^{n} \lambda_i g_i'(x) \geq 0 \tag{A1}$$

$$h''(x) = \sum_{i=1}^{n} \lambda_i g_i''(x) \leq 0 \tag{A2}$$

i.e., $h(x)$ is an increasing concave function on [0, 1], and

$$h(0) = \sum_{i=1}^{n} \lambda_i g_i(0) = 0 \tag{A3}$$

$$g(1) = \sum_{i=1}^{n} \lambda_i g_i(1) = \sum_{i=1}^{n} \lambda_i = 1 \tag{A4}$$

$$\begin{aligned} 2h'(x) + xh''(x) &= 2\sum_{i=1}^{n} \lambda_i g_i'(x) + x\sum_{i=1}^{n} \lambda_i g_i''(x) \\ &= \sum_{i=1}^{n} \lambda_i \left(2g_i'(x) + x\, g_i''(x)\right) \leq 0, \\ & x \in [0, 1] \end{aligned} \tag{A5}$$

Therefore, $h(x) = \sum_{i=1}^{n} \lambda_i g_i(x)$ is still function G.

Proof concept (2): It is known that $g_1(x)$ is a function G, which satisfies all properties of function G. Then,

$$g_2'(x) = \alpha g_1'(\alpha x)/g_1(\alpha) \geq 0 \tag{A6}$$

$$g_2''(x) = \alpha^2 g_1''(\alpha x)/g_1(\alpha) \leq 0 \tag{A7}$$

i.e., $g_2(x)$ is an increasing concave function on [0, 1], and

$$g_2(0) = g_1(0)/g_1(\alpha) = 0 \tag{A8}$$

$$g_2(1) = g_1(\alpha \cdot 1)/g_1(\alpha) = 1 \tag{A9}$$

$$\begin{aligned} 2g_2'(x) + xg_2''(x) &= 2\frac{\alpha g_1'(\alpha x)}{g_1(\alpha)} + x\frac{\alpha^2 g_1''(\alpha x)}{g_1(\alpha)} \\ &= \frac{\alpha}{g_1(\alpha)}\left[2g_1'(\alpha x) + \alpha x g_1''(\alpha x)\right] \end{aligned} \tag{A10}$$

Due to $0 < \alpha \leq 1$ and $0 \leq \alpha x \leq 1$, therefore

$$\frac{\alpha}{g_1(\alpha)} > 0, \quad 2g_1'(\alpha x) + \alpha x g_1''(\alpha x) \leq 0 \tag{A11}$$

$$2g_2'(x) + xg_2''(x) \leq 0, \quad \forall x \in [0, 1] \tag{A12}$$

Then, $g_2(x) = g_1(\alpha x)/g_1(\alpha)$ is still function G.

Proof concept (3): It is known that $g_1(x)$ and $g_2(x)$ are the function G, which satisfies all properties of function G. Then,

$$g'(x) = g_2'(g_1(x))\, g_1'(x) \geq 0 \tag{A13}$$

$$g''(x) = g_2''(g_1(x))\left(g_1'(x)\right)^2 + g_2'(g_1(x))\, g_1''(x) \leq 0 \tag{A14}$$

i.e., $g(x)$ is an increasing concave function on [0, 1], and

$$g(0) = g_2(g_1(0)) = 0 \tag{A15}$$

$$g(1) = g_2(g_1(1)) = 1 \tag{A16}$$

$$\begin{aligned} &2g'(x) + xg''(x) \\ &\quad = 2g_2'(g_1(x))\, g_1'(x) \\ &\qquad + x\left[g_2''(g_1(x))\left(g_1'(x)\right)^2 + g_2'(g_1(x))\, g_1''(x)\right] \\ &\quad = g_2'(g_1(x))\left[2g_1'(x) + xg_1''(x)\right] \\ &\qquad + xg_2''(g_1(x))\left(g_1'(x)\right)^2 \end{aligned} \tag{A17}$$

Due to $g_1(x)$ and $g_2(x)$ are the function G, then

$$2g_1'(x) + xg_1''(x) \leq 0 \tag{A18}$$

$$g_2'(g_1(x)) \geq 0, \; g_2''(g_1(x)) \leq 0 \tag{A19}$$

$$2g'(x) + xg''(x) \leq 0 \tag{A20}$$

Then, $g(x) = g_2(g_1(x))$ is still function G.

APPENDIX B

The characteristic frequencies of the different bearing faults can be expressed as:

$$f_i = \frac{1}{2} N_B f_r \left(1 + \frac{D_b \cos\theta}{D_c}\right) \tag{B1}$$

$$f_o = \frac{1}{2} N_B f_r \left(1 - \frac{D_b \cos\theta}{D_c}\right) \tag{B2}$$

$$f_{re} = \frac{1}{2} f_r \left(\frac{D_c}{D_b}\right)\left[1 - \left(\frac{D_b \cos\theta}{D_c}\right)^2\right] \tag{B3}$$

where f_i, f_o and f_{re} are the characteristic frequencies of the inner race, outer race and rolling element, respectively; N_B is the number of balls in the bearing, f_r the rotating frequency of the bearing, D_b the ball diameter, D_c the pitch diameter, and θ the contact angle, respectively.

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] S. Ferson, V. Kreinovich, L. Grinzburg, D. Myers, and K. Sentz, "Constructing probability boxes and Dempster-Shafer structures," Sandia Nat. Laboratories, Livermore, CA, USA, Tech. Rep., 2003, doi: 10.2172/809606.

[3] C. M. Eastman, "Introduction to fuzzy arithmetic: Theory and applications," *Int. J. Approx. Reasoning*, vol. 1, no. 1, pp. 141–143, Jan. 1987.

[4] G. Shafer, "The combination of evidence," *Int. J. Intell. Syst.* , vol. 1, no. 3, pp. 155–179, 1986.

[5] T. Hailperin, *Boole's Logic and Probability*. Amsterdam, The Netherlands: North-Holland, 1986.

[6] M. J. Frank, R. B. Nelsen, and B. Schweizer, "Best-possible bounds for the distribution of a sum—A problem of kolmogorov," *Probab. Theory Rel. Fields*, vol. 74, no. 2, pp. 199–211, 1987, doi: 10.1007/BF00569989.

[7] H. T. Nguyen and E. A. Walker, "A first course in fuzzy logic," *Math. Gazette*, vol. 81, no. 491, 2005.

[8] D. Berleant and J. Zhang, "Bounding the times to failure of 2-component systems," *IEEE Trans. Rel.*, vol. 53, no. 4, pp. 542–550, Dec. 2004, doi: 10.1109/TR.2004.837315.

[9] F. Tonon, "Using random set theory to propagate epistemic uncertainty through a mechanical system," *Rel. Eng. Syst. Saf.*, vol. 85, nos. 1–3, pp. 169–181, Jul. 2004, doi: 10.1016/j.ress.2004.03.010.

[10] P. Limbourg and R. Savic, "Fault tree analysis in an early design stage using the Dempster-Shafer theory of evidence," *Risk, Rel. Soc. Saf.*, vol. 1, no. 3, pp. 99–105, 2007.

[11] M. Oberguggenberger, J. King, and B. Schmelzer, "Imprecise probability methods for sensitivity analysis in engineering. In 5th International Symposium on Imprecise Probability: Theories and Applications," *Prague, Czech Republic*, vol. 6, pp. 1130–1138, 2007.

[12] S. Destercke, D. Dubois, and E. Chojnacki, "Unifying practical uncertainty representations —I: Generalized p-boxes," *Int. J. Approx. Reasoning*, vol. 49, no. 3, pp. 649–663, Nov. 2008, doi: 10.1016/j.ijar.2008.07.003.

[13] L. G. Crespo, S. P. Kenny, and D. P. Giesy, "Reliability analysis of polynomial systems subject to p-box uncertainties," *Mech. Syst. Signal Process.*, vol. 37, nos. 1–2, pp. 121–136, May 2013, doi: 10.1016/j.ymssp.2012.08.012.

[14] M. C. M. Troffaes, E. Miranda, and S. Destercke, "On the connection between probability boxes and possibility measures," *Inf. Sci.*, vol. 224, pp. 88–108, Mar. 2013, doi: 10.1016/j.ins.2012.09.033.

[15] 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.

[16] T. Hong, D. Yi, D. Jiaman, and L. Liqiang, "Compound fault diagnosis based on probability box theory," in *Proc. 9th Int. Conf. Modeling, Identificat. Control (ICMIC)*, Jul. 2017, pp. 336–341.

[17] G. Papaefthymiou and D. Kurowicka, "Using copulas for modeling stochastic dependence in power system uncertainty analysis," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 40–49, Feb. 2009.

[18] S. Zhang, L. N. Theagarajan, S. Choi, and P. K. Varshney, "Fusion of correlated decisions using regular vine copulas," *IEEE Trans. Signal Process.*, vol. 67, no. 8, pp. 2066–2079, Apr. 2019, doi: 10.1109/TSP.2019.2901379.

[19] V. N. Nyaga, M. Arbyn, and M. Aerts, "CopulaDTA: An r package for copula-based bivariate beta-binomial models for diagnostic test accuracy studies in a Bayesian framework," *J. Stat. Softw.*, vol. 82, no. 1, pp. 1–27, 2017, doi: 10.18637/jss.v082.c01.

[20] M. Cui, V. Krishnan, B.-M. Hodge, and J. Zhang, "A copula-based conditional probabilistic forecast model for wind power ramps," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3870–3882, Jul. 2019, doi: 10.1109/TSG.2018.2841932.

[21] Y. Wang and Y. Luo, "Research of wind power correlation with three different data types based on mixed copula," *IEEE Access*, vol. 6, pp. 77986–77995, 2018, doi: 10.1109/ACCESS.2018.2884539.

[22] Q. Li, J. B. Brown, H. Huang, and P. J. Bickel, "Measuring reproducibility of high-throughput experiments," *Ann. Appl. Statist.*, vol. 5, no. 3, pp. 1752–1779, Sep. 2011.

[23] C. Li, Y. Huang, and Y. Xue, "Dependence structure of Gabor wavelets based on copula for face recognition," *Expert Syst. Appl.*, vol. 137, pp. 453–470, Dec. 2019, doi: 10.1016/j.eswa.2019.05.034.

[24] R. Calabrese and S. A. Osmetti, "A new approach to measure systemic risk: A bivariate copula model for dependent censored data," *Eur. J. Oper. Res.*, vol. 279, no. 3, pp. 1053–1064, Dec. 2019, doi: 10.1016/j.ejor.2019.06.027.

[25] L. Li, S. Miao, Q. Tu, S. Duan, Y. Li, and J. Han, "Dynamic dependence modelling of wind power uncertainty considering heteroscedastic effect," *Int. J. Electr. Power Energy Syst.*, vol. 116, Mar. 2020, Art. no. 105556, doi: 10.1016/j.ijepes.2019.105556.

[26] W. Liu and Y. Liu, "Energy storage sizing by copula modelling joint distribution for wind farm to be black-start source," *IET Renew. Power Gener.*, vol. 13, no. 11, pp. 1882–1890, Aug. 2019, doi: 10.1049/iet-rpg.2018.6154.

[27] D. Yu and N. Ghadimi, "Reliability constraint stochastic UC by considering the correlation of random variables with copula theory," *IET Renew. Power Gener.*, vol. 13, no. 14, pp. 2587–2593, Oct. 2019, doi: 10.1049/iet-rpg.2019.0485.

[28] S. Ferson, "Dependence in probabilistic modeling, Dempster-Shafer theory and probability bounds analysis," Sandia National Lab., Livermore, CA, USA, Tech. Rep., 2004.

[29] R. B. Nelsen, *An Introduction to Copulas*. Berlin, Germany: Springer, 2006.

[30] R. Schefzik, T. L. Thorarinsdottir, and T. Gneiting, "Uncertainty quantification in complex simulation models using ensemble copula coupling," *Stat. Sci.*, vol. 28, no. 4, pp. 616–640, Nov. 2013, doi: 10.1214/13-STS443.

[31] P. B. Stark, "Risk-limiting postelection audits: Conservative *P*-values from common probability inequalities," *IEEE Trans. Inf. Forensics Security*, vol. 4, no. 4, pp. 1005–1014, Dec. 2009, doi: 10.1109/TIFS.2009.2034190.

[32] S. K. Mahjour, A. A. S. Santos, M. G. Correia, and D. J. Schiozer, "Developing a workflow to select representative reservoir models combining distance-based clustering and data assimilation for decision making process," *J. Petroleum Sci. Eng.*, vol. 190, Jul. 2020, Art. no. 107078, doi: 10.1016/j.petrol.2020.107078.

[33] X. Yan and M. Jia, "Application of CSA-VMD and optimal scale morphological slice bispectrum in enhancing outer race fault detection of rolling element bearings," *Mech. Syst. Signal Process.*, vol. 122, pp. 56–86, May 2019, doi: 10.1016/j.ymssp.2018.12.022.

[34] T. Lin, G. Chen, W. Ouyang, Q. Zhang, H. Wang, and L. Chen, "Hyperspherical distance discrimination: A novel data description method for aero-engine rolling bearing fault detection," *Mech. Syst. Signal Process.*, vol. 109, pp. 330–351, Sep. 2018, doi: 10.1016/j.ymssp.2018.01.009.

[35] X. Chen, W. Xu, Y. Liu, and M. R. Islam, "Bearing corrosion failure diagnosis of doubly fed induction generator in wind turbines based on stator current analysis," *IEEE Trans. Ind. Electron.*, vol. 67, no. 5, pp. 3419–3430, May 2020, doi: 10.1109/TIE.2019.2917418.

[36] H. Tang, H. L. Dai, and Y. Du, "Tapered roller bearing failure diagnosis based on improved probability box model," *IEEE Access*, vol. 8, pp. 151452–151464, 2020, doi: 10.1109/ACCESS.2020.3013875.

[37] H. Tang, Z. Yuan, H. Dai, and Y. Du, "Fault diagnosis of rolling bearing based on probability box theory and GA-SVM," *IEEE Access*, vol. 8, pp. 170872–170882, 2020, doi: 10.1109/ACCESS.2020.3024792.

[38] G. G. Yen and K.-C. Lin, "Wavelet packet feature extraction for vibration monitoring," *IEEE Trans. Ind. Electron.*, vol. 47, no. 3, pp. 650–667, Jun. 2000.

[39] 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.ymssp.2016.09.010.

LIANGCAI DONG received the Ph.D. degree from Shanghai Maritime University, Shanghai, China, in 2009.

He is currently an Associate Professor and the Director of the Department of Industrial Engineering, Shanghai Maritime University. His research interests include intelligent optimization and scheduling, intelligent simulation systems, and machine learning.

YING LIU received the B.A. degree in physical distribution management from the Suzhou University of Science and Technology, Jiangsu, China, in 2019. She is currently pursuing the M.A.Eng. degree with Shanghai Maritime University. Her current research interests include digital twin, intelligent optimization and scheduling, and machine learning.

HONG TANG is currently pursuing the Ph.D. degree with the College of Mechanical and Vehicle Engineering, Hunan University, China.

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

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.