Reliability estimation of reciprocating seals based on multivariate dependence analysis and its experimental validation

CHAO ZHANG¹, (Member, IEEE), RENTONG CHEN¹, SHAOPING WANG¹, YUJIE QIAN¹, AND MILETA M. TOMOVIC²

¹School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191 China
²Engineering Technology Department, Old Dominion University, Norfolk, VA 23529 USA

Corresponding author: Chao Zhang (e-mail: czhangstar@gmail.com).

This study was supported by the National Natural Science Foundation of China (51875015, 51620105010, 51675019), and Natural Science Foundation of Beijing Municipality (L171003).

ABSTRACT Accurate reliability estimation for reciprocating seals is of great significance due to their wide use in numerous engineering applications. This work proposes a reliability estimation method for reciprocating seals based on multivariate dependence analysis of different performance indicators. Degradation behavior corresponding to each performance indicator is first described by the Wiener process. Dependence among different performance indicators is then captured using D-vine copula, and a weight-based copula selection method is utilized to determine the optimal bivariate copula for each dependence relationship. A two-stage Bayesian method is used to estimate the parameters in the proposed model. Finally, a reciprocating seal degradation test is conducted, and the proposed reliability estimation approach is validated by test data. Results show that the proposed model is accurate and effective in estimating the reliability of reciprocating seals.

INDEX TERMS D-vine copula, dependence analysis, reciprocating seal, reliability estimation

I. INTRODUCTION

HYDRAULIC reciprocating seals can prevent leakage, contain pressure, and exclude contamination during reciprocating motion. With advantages including low cost, good sealing performance, high oil resistance, and high thermal resistance, reciprocating seals are widely used in numerous engineering applications [1] including in aerospace, medical, marine, and automotive industries. Reciprocating seal failure can lead to hydraulic oil leakage and other types of malfunctions which can cause huge economic loss and potentially catastrophic consequences. For this reason, reliability estimation of reciprocating seals is of great importance and has drawn increasing attention from scholars.

Obtaining adequate failure data for mechanical components with high reliability and a long lifespan is difficult. In this situation, degradation testing or accelerated degradation testing is commonly conducted to evaluate the reliability of various products. Ma et al. studied a step-stress accelerated degradation test of highly reliable products and proposed the M-optimality criterion to improve mechanism equivalence [2]. Wang et al. presented an optimal design plan for accelerated degradation testing with multiple stresses and conducted sensitivity analysis [3]. Si et al. [4] classified the reliability estimation methodology for degradation testing or accelerated degradation testing into the three categories of knowledge-based approaches, physical model-based approaches, and data-driven approaches. Knowledge-based approaches perform reliability estimation based on an expert system or a fuzzy system, and require special knowledge about the particular components or systems combined with failure data [5], [6]. Physical model-based approaches study the physical behavior of the degradation process, and carry out reliability estimation for critical physical components using mathematical or physical models of the degradation phenomenon (including crack by fatigue, wear, and corrosion) [7], [8]. Data-driven approaches aim at transforming the data provided by status monitoring systems into relevant models which can describe the degradation behavior, and these approaches include the machine learning method [9]–[11] and stochastic model-based method [12]–[14]. For reciprocating seals, as the failure mechanism is highly complicated and the operation condition is usually harsh, a comprehensive...
physical-based degradation model is often difficult to establish [15]–[17]. In addition, as adequate failure data for a specific type of reciprocating seal are often hard to obtain, a data-driven approach is preferable for reliability estimation.

Leakage rate, friction force, and contact temperature are three important performance indicators for reciprocating seals which can reflect their operation and degradation status. These three performance indicators are also easy to collect [18], [19], and are preferred in many engineering applications. This study uses the Wiener process, a commonly used data-driven approach, to describe the degradation process of each performance indicator. For many types of mechanical components with complex failure mechanisms, different performance indicators will be statistically dependent due to various commonly shared factors including environmental stress, material properties, and various operational stresses. For reciprocating seals, as illustrated in the following section, leakage rate and friction force are both affected by the two failure modes of reciprocating seals: wear and aging. In addition, contact temperature can reflect the lubrication status of the contact area. Leakage rate and friction force are both affected by contact temperature, and vice versa. Therefore, it is necessary to take the dependence among the three performance indicators into consideration to formulate the degradation model for reciprocating seals.

Compared with traditional joint bivariate distribution, copulas offer more flexibility in capturing the dependence between different marginal distributions. They model marginal behavior and dependence structure separately, in which the marginal distributions do not have to belong to identical distribution families [20]. For this reason, copulas are widely used in dependence analysis for various research applications including financial phenomena [21], [22], energy management [23], [24], accelerated life testing, and reliability analysis [25], [26]. Copulas have also been used to describe the dependence between different performance indicators of mechanical components. Pan et al. used two performance indicators governed by the Wiener process, utilizing Frank copula to describe their dependence and build a degradation model [27]. Chen et al. characterized the failure dependence between subsystems of mechanical systems and constructed a reliability improvement model based on relative failure rate [28]. In both studies, the type of copula function was presumed to be known in advance. In many situations, however, it is difficult to determine which type of copula is the most appropriate initially. Various studies have focused on methods to determine an appropriate copula. Pan used Akaike information criterion (AIC) to select the best copula when characterizing the dependence between different degradation paths of products [29]. Deviance information criterion (DIC) was used by Zhang et al. to find the optimal copula and build an accelerated life testing model for solid lubricated bearings [30].

All of the studies discussed have utilized bivariate copulas which capture the dependence between two marginal distributions. For reciprocating seals, as mentioned above, there are three commonly used performance indicators, namely, leakage rate, contact temperature, and friction force. As more performance indicators contain more information on the operation and degradation status of the seal [31], a degradation model which considers three performance indicators and the dependence among them will likely be more accurate in estimating reliability. Multivariate copulas including multivariate Gaussian copula and t-copula have been utilized to describe the dependence of multiple variables. However, multivariate copulas require that each two variables follow the same dependence structure. For reciprocating seals, the dependence structure of the marginal distributions corresponding to each two performance indicators may not be the same. In this case, a degradation model for reciprocating seals based on multivariate copula function may not be accurate.

On the basis of bivariate copula function and hierarchical structure, vine-copula, which is a multivariate modeling method, offers more flexibility and accuracy for capturing the dependence of three or more variables [32]. Based on conditional probability, vine-copula decomposes joint distribution into a multiplication model consisting of multiple bivariate copulas along with corresponding conditional probability [33], [34]. Vine-copula utilizes a topological graph to illustrate the dependence among variables, and is often referred to as regular vine (R-vine). D-vine copula is a type of R-vine copula in which dependence of different pairs of variables can be described by different types of bivariate copula. This makes D-vine more flexible and accurate in capturing the dependence among multiple variables.

Focused on the reciprocating seal, which has high reliability and a long lifetime with multiple performance indicators, this work studies reciprocating seal reliability estimation and formulates a degradation model. The failure behavior and degradation process corresponding to each performance indicator is first described by Wiener process, and relevant marginal distributions are obtained. Secondly, based on D-vine copula and corresponding hierarchical structure, bivariate copula functions are used to describe the dependence among the three performance indicators and build a degradation model. A Bayesian method is also used to select optimal copula for each bivariate dependence relationship based on weight calculation. Finally, parameters in the proposed model are estimated based on a two-stage Bayesian method, and the reliability of reciprocating seals can be estimated.

The remainder of this paper is organized as follows. In Section II, the failure mechanism of reciprocating seals is briefly introduced, and the degradation model based on D-vine copula is built. Two-stage Bayesian framework is also constructed to estimate the unknown parameters. In Section III, degradation testing of reciprocating seals is carried out, and the proposed model is validated based on experimental data. Finally, conclusions are provided in Section IV.

II. MODEL DEVELOPMENT
A. FAILURE MECHANISM OF RECIPROCATING SEALS

A schematic diagram of the application of a reciprocating seal in a hydraulic actuator is illustrated in Fig. 1. The reciprocating seal is installed in a specific groove with a preload to the shaft, and a dust seal is generally used as an auxiliary component. In one actuation of the actuator, during the outstroke, some hydraulic fluid may be adhered to the rod and dragged out of the cylinder. During the instroke, some fluid is dragged into the cylinder. If the amount of fluid dragged out of the cylinder during the outstroke is bigger than that taken into the cylinder during the instroke, leakage is likely to occur. During either instroke or outstroke, a sufficient lubricating film between the seal and the surface of the rod is required, which can support the seal and prevent exceedingly high temperature, as well as reduce friction and prevent excessive wear [19].

For reciprocating seals, friction force, leakage rate, and contact temperature can be used as performance indicators [31]. When a reciprocating seal is in operation, the material of the seal will gradually wear out, and the topography of the contact surface will also change. Therefore, the sealing effect will inevitably fluctuate and this will cause a corresponding change in the leakage rate. In addition, topography changes of the contact surface will also lead to variations in friction force when the rod is moving [35].

The reciprocating seal is usually made of rubber, and thus aging is also one of its most significant failure modes. The elongation and hardness of the reciprocating seal is significantly affected by aging. Change to the material properties of the seal will also lead to the change of topography in the contact area, and thus affect the leakage rate and friction force when the reciprocating seal is in operation [31].

Contact temperature is another important performance indicator [31], [36], [37]. Heat is generated by friction between the rod and the seal, and the temperature can affect the material properties of the seal, affecting the topography of the contact area. Heat can also affect the viscosity of hydraulic oil [36]. When the viscosity decreases, the support ability of the film will decrease, which can result in dry friction and increase the leakage rate. When the viscosity rises, the friction force will increase, and the leakage rate will also be affected [35].

From the discussion above, it is reasonable to conclude that these three performance indicators are dependent. In addition, as the dependence between either two performance indicators could be different, the D-Vine copula is used to describe the dependence among the three performance indicators and build the degradation model.

B. RELIABILITY FUNCTION FOR EACH PERFORMANCE INDICATOR OF RECIPROCATING SEAL

As discussed above, friction force, contact temperature, and leakage rate are three dependent performance indicators of reciprocating seals which can be used to describe the degradation process. Wiener process is based on Brownian motion and is applicable for reflecting the cumulative effect of degradation failure. The process is commonly used for describing the degradation process of various types of mechanical products in practice [4], [12]. In this work, Wiener process is utilized to describe the degradation process for each performance indicator. Assuming that the number of reciprocating seals being tested is N, and the number of measurements for each seal in a pre-defined time interval is M. The kth performance indicator of the ith seal at the jth measurement is denoted as \( Y_k(i,j) \). Here and in the following context, \( k = 1, 2, 3, i = 1, 2, ..., N, \) and \( j = 1, 2, ..., M \). Based on Wiener process, the expression of each performance indicator can be given by

\[
Y_k(i,j) = \mu_k \Lambda (t_{ij}) + \sigma_k B (\Lambda (t_{ij}))
\]

(1)

where \( \mu_k \) denotes the drift coefficient reflecting the degradation rate, \( \sigma_k \) is the volatility parameter, \( B (\cdot) \) is the standard Brownian Motion, and \( \Lambda (t_{ij}) \) is the time scale function reflecting the nonlinearity of degradation paths which is usually monotonic. Specifically, \( \Lambda (t_{ij}) \) is assumed to be \( \Lambda (t_{ij}) = t_{ij}^{\eta_k} \) in this work, where \( \eta_k \) is the power exponent in the time scale function.

The increment of the \( k \)th performance indicator of the \( i \)th sample in time interval \([t_{ij-1}, t_{ij}]\) can be calculated by \( \Delta Y_k(i,j) = Y_k(i,j) - Y_k(i,j-1) \). In the following context, we use \( X_1 \) to denote \( \Delta Y_1(i,j) \), \( X_2 \) to denote \( \Delta Y_2(i,j) \), \( X_3 \) to denote \( \Delta Y_3(i,j) \), and \( X_k \) to denote \( \Delta Y_k(i,j) \). Based on the definition of Wiener process, the \( \Lambda (t_{ij}) \) follows normal distribution, as given by

\[
X_k \sim N (\mu_k (t_{ij}^{\eta_k} - t_{ij-1}^{\eta_k}), \sigma_k^2 (t_{ij}^{\eta_k} - t_{ij-1}^{\eta_k}))
\]

(2)

Using \( \Delta \Lambda (t_{ij}) \) to denote \( t_{ij}^{\eta_k} - t_{ij-1}^{\eta_k} \), then the probability density function (pdf) of \( X_k \) can be expressed as

\[
f_k(X_k) = \frac{1}{\sqrt{2\pi \sigma_k^2 \Delta \Lambda (t_{ij})}} \cdot \exp \left\{ \frac{-(X_k - \mu_k \Delta \Lambda (t_{ij}))^2}{2\sigma_k^2 \Delta \Lambda (t_{ij})} \right\}
\]

(3)

The cumulative distribution function of \( X_k \) can be given by

\[
F_k(X_k) = \Phi \left( \frac{X_k - \mu_k \Delta \Lambda (t_{ij})}{\sigma_k \sqrt{\Delta \Lambda (t_{ij})}} \right)
\]

(4)

A pre-specified threshold level for each performance indicator is \( d_k \), and the failure time \( T_k \), which represents the first passage time (FPT) over the threshold, can be given by \( T_k = \inf \{ t : Y_k(t) \geq d_k \} \). The FPT follows inverse Gaussian distribution, i.e., \( T_k \sim IG (\delta_k, \lambda_k) \). With \( \delta_k = d_k/\mu_k \) and \( \lambda_k = (d_k/\sigma_k)^2 \), the reliability function corresponding to the \( k \)th performance indicator can be given by

\[
R_k(t) = \Phi \left( \frac{d_k - \mu_k \Lambda (t)}{\sigma_k \sqrt{\Lambda (t)}} \right) - \exp \left( \frac{2\mu_k d_k}{\sigma_k^2} \right) \Phi \left( \frac{-d_k + \mu_k \Lambda (t)}{\sigma_k \sqrt{\Lambda (t)}} \right)
\]

(5)
C. RELIABILITY ESTIMATION FOR RECIPROCATING SEALS BASED ON D-VINE COPULA

1) Multivariate degradation model establishment

When there is correlation among different performance indicators, copula function is an effective method to combine marginal distributions of each performance indicator and build the coupling model of degradation behavior. When the number of marginal distributions is more than two, based on the Sklar’s theorem [38], multivariate copula function, such as multivariate Gaussian copula or t-copula, can be used to construct the joint distribution, as given by

$$F(x_1, x_2, x_3) = C(F_1(x_1), F_2(x_2), F_3(x_3))$$  \hspace{1cm} (6)

However, the number of available multivariate copulas is limited. A multivariate copula also assumes that the dependence structure of each pair of its marginal distributions is the same, which might not be the case for the three performance indicators of reciprocating seals. Combining bivariate copula function with hierarchical structure, D-vine copula utilizes a topological graph model to illustrate the dependence between different marginal distributions, and decomposes joint distribution into a multiplication model which consists of multiple bivariate copulas and corresponding conditional probability [32]. For three-dimensional variables, the joint probability density function can be decomposed as

$$f(x_1, x_2, x_3) = f_1(x_1) \cdot f_2|1 (x_2|x_1) \cdot f_3|1,2 (x_3|x_1, x_2)$$ \hspace{1cm} (7)

In addition, based on the condition probability formula and Sklar’s theory, \(f_2|1 (x_2|x_1)\) and \(f_3|1,2 (x_3|x_1, x_2)\) can be composed in a further form as follows

$$f_2|1 (x_2|x_1) = \frac{f(x_1, x_2)}{f_1(x_1)} = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1), f_2(x_2)$$

$$f_3|1,2 (x_3|x_1, x_2) = \frac{f_1,3|2 (x_1, x_3|x_2)}{f_1|1,2 (x_1|x_2)} = c_{1,3|2}(F_1|1,2 (x_1|x_2), F_3|2 (x_3|x_2)) \cdot f_3|2 (x_3|x_2)$$ \hspace{1cm} (9)

In (9), \(F_1|1,2 (x_1|x_2)\) and \(F_3|2 (x_3|x_2)\) are conditional distribution functions, and can be expressed by

$$F_1|1,2 (x_1|x_2) = \frac{\partial C_{12}(F_1(x_1), F_2(x_2))}{\partial F_2(x_2)}$$ \hspace{1cm} (10)

$$F_3|2 (x_3|x_2) = \frac{\partial C_{23}(F_2(x_2), F_3(x_3))}{\partial F_2(x_2)}$$ \hspace{1cm} (11)

Then \(f(x_1, x_2, x_3)\) in (7) can be decomposed as

$$f(x_1, x_2, x_3) = f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3)$$

$$\cdot c_{12}(F_1(x_1), F_2(x_2))$$

$$\cdot c_{23}(F_2(x_2), F_3(x_3))$$

$$\cdot c_{1,3|2}(F_1|1,2 (x_1|x_2), F_3|2 (x_3|x_2))$$ \hspace{1cm} (12)

Based on (12), the joint probability density function can be calculated by three marginal distributions, i.e. \(f_1(x_1)\), \(f_2(x_2)\) and \(f_3(x_3)\), and three bivariate copula functions, i.e. \(C_{12}, C_{23}\) and \(C_{1,3|2}\). Subscripts \(c_{12}, c_{23}\) and \(c_{1,3|2}\) are density functions corresponding to \(C_{12}, C_{23}\) and \(C_{1,3|2}\), respectively.

The decomposition procedure of joint probability density function can be illustrated using a hierarchical structure, as shown in Fig. 2. In this figure, Layer 1 shows the three marginal probability density functions of the three performance indicators. As discussed above, during the operation process, temperature variation in the contact zone could be caused by the friction force, and temperature variation could result in a change of material properties, thus affecting the degradation of friction. In addition, the change of material properties could also affect the status of lubricating film and the leakage rate. Therefore, contact temperature is
placed in the second position in Layer 1 of Fig. 2, and is connected to the other two performance indicators, friction force and leakage rate. Layer 2 represents the dependent model between adjacent performance indicators by using bivariate copulas \( C_{12} \) and \( C_{23} \). Subscript \( C_{12} \) describes the dependence relationship between friction force and contact temperature, and \( C_{23} \) describes the dependence relationship between contact temperature and leakage rate. In Layer 3, conditional bivariate copula \( C_{13/2} \) describes the conditional correlation between performance indicators 1 and 3 given performance indicator 2. Without loss of generality, friction force and leakage rate are also placed into the second position in Layer 1 of Fig. 2 respectively, and the corresponding results are calculated and used as comparison.

The degradation model for reciprocating seals considering the dependence among three performance indicators can then be formulated as

\[
X_k \sim N (\mu_k \Delta (t_{ij}), \sigma_k^2 \Delta (t_{ij}))
\]

\[
Y_k (t_{ij}) = \sum_{s=0}^3 \Delta Y_k (t_{is})
\]

\[
f (X_1, X_2, X_3) = \prod_{k=1}^3 f (X_k; \alpha_k)
\]

\[
\cdot c_{12} (F_1 (X_1), F_2 (X_2); \theta_{12})
\]

\[
\cdot c_{23} (F_2 (X_2), F_3 (X_3); \theta_{23})
\]

\[
\cdot c_{13/2} (F_{1|2} (X_1|X_2), F_{3|2} (X_3|X_2); \theta_{13/2})
\]

where \( \alpha_k = (\mu_k, \sigma_k, q_k) \) represents the unknown parameters corresponding to the marginal distribution of each performance indicator, and \( \theta_{12}, \theta_{23}, \text{and} \theta_{13/2} \) are the dependence parameters of the corresponding copula function, respectively.

With the pre-specified thresholds for the three performance indicators, the reliability function can be given by

\[
R (t) = P (Y_1 (t) < d_1, Y_2 (t) < d_2, Y_3 (t) < d_3)
\]

(14)

The reliability can then be obtained by Monte Carlo method, and the meantime to failure (MTTF) can be calculated by

\[
E (T) = \int_0^{\infty} R (t) \; dt
\]

(15)

Combining degradation data and corresponding parameter estimation method, this model could be utilized to describe the degradation process of reciprocating seals and obtain related reliability information.

2) Bayesian copula selection

Once the degradation model of reciprocating seal is built based on D-vine copula, suitable bivariate copula functions must be determined. This work uses a Bayesian method to locate the optimal copulas for \( C_{12}, C_{23}, \text{and} C_{13|2} \), respectively, then the copula candidate which has the largest posterior probability is selected [39]. In this approach, values of parameters in the proposed degradation model are not required for copula selection.

Let \( C = (C_1, C_2, \ldots, C_L) \) denote the set of copula candidates, and \( H_l \) represents the event of the given sample data \( D \) coming from copula \( C_l (C_l \subset C, 1 \leq l \leq L) \). The posterior probability \( P (H_l|D) \) is the probability that it satisfies \( H_l \) with the given data \( D \). Using Kendall’s \( \tau \), the posterior probability can be expressed as [39]

\[
P (H_l|D, I) = \frac{\int_{-1}^1 P (H_l, \tau|D, I) \; d\tau}{\int_{-1}^1 \frac{P (D|H_l, \tau, I) \cdot P (H_l|\tau, I) \cdot P (\tau|I)}{P (D|I)} \; d\tau}
\]

(16)

where \( P (D|H_l, \tau, I) \) is the likelihood function, \( P (H_l|\tau, I) \) is the prior probability that \( H_l \) is valid, \( P (\tau|I) \) is the prior probability for Kendall’s \( \tau, I \) represents other relevant information, and \( P (D|I) \) is the normalization constant. In (16), \( P (D|H_l, \tau, I) \) is dependent on Kendall’s \( \tau \), and can be calculated by [39]

\[
P (D|H_l, \tau, I) = \prod_{j=1}^M c_l (u_j, v_j|\tau)
\]

(17)

where \( c_l (u_j, v_j|\tau) \) is the density of copula function \( C_l \).

The prior information in (16) includes \( P (H_l|\tau, I) \) and \( P (\tau|I) \). Here \( I \) represents other relevant information, and is defined as follows:

1) \( I_1 \): The prior on \( \tau \) is additional knowledge regarding the dependence between two performance indicators. In this case, \( \tau \) belongs to set \( \Theta \), and each outcome of \( \tau \in \Theta \) is equally likely. In this work, no prior information is available, so \( \Theta \) can be simply assumed to be \([-1, 1]\), providing

\[
P (\tau|I_1) = \begin{cases} \frac{1}{\lambda (\Theta)}, & \tau \in \Theta \\ 0, & \tau \notin \Theta \end{cases}
\]

(18)

where \( \lambda (\Theta) \) denotes the Lebesgue measure of \( \Theta \).

2) \( I_2 \): For a given \( \tau \), all copula families satisfying \( \tau \in \Omega_l \) are equally probable. Here, \( \Omega_l \) is the domain of Kendall’s \( \tau \) for the \( l \)th copula \( C_l \).

Eq. (18) can then be further expressed as

\[
P (H_l|D, I) = \frac{1}{P (D|I)} \int_{-1}^1 \prod_{j=1}^M c_l (u_j, v_j|\tau) \cdot \frac{1}{\lambda (\Theta)} \; d\tau
\]

(19)

\[
\left( \frac{1}{P (D|I)} \cdot P (H_l|\tau, I_2) \cdot \frac{1}{\lambda (\Theta)} \right) \int_{\Omega_l \cap \Theta} \prod_{j=1}^M c_l (u_j, v_j|\tau) \; d\tau
\]

where \( P (D|I) \) and \( P (H_l|\tau, I_2) \) are both constant. Let

\[
W_l = \frac{1}{\lambda (\Theta)} \cdot \int_{\Omega_l \cap \Theta} \prod_{j=1}^M c_l (u_j, v_j|\tau) \; d\tau
\]

(20)
and the weight $Q_l$ for each copula function $C_l$ in the set of copula candidates $C$ can be expressed as

$$Q_l = \frac{W_l}{\sum_{m=1}^{L} W_m} \quad (m = 1, 2, ..., L) \quad (21)$$

The copula function $C_l$ in the set of copula candidates $C$ with the largest weight can then be selected for $C_{12}$, $C_{23}$ and $C_{13|2}$, respectively.

3) Bayesian inference

Unknown parameters $\alpha_k = (\mu_k, \sigma_k, q_k)$, $\theta_{12}$, $\theta_{23}$, and $\theta_{13|2}$ in (13) are estimated by degradation data, then the reliability function can be obtained, and mean time to failure (MTTF) as well as other indices regarding reliability can be calculated. Based on (13), the log-likelihood function can be derived, as given by

$$\ln L = \sum_{k=1}^{3} \sum_{i=1}^{N} \sum_{j=1}^{M} \ln [f_k (X_k) ; \alpha_k]$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{M} \ln [c_{12} (F_1 (X_1), F_2 (X_2)) ; \theta_{12}]$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{M} \ln [c_{23} (F_2 (X_2), F_3 (X_3)) ; \theta_{23}]$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{M} \ln [c_{13|2} (F_1|2 (X_1|X_2), F_3|2 (X_3|X_2)) ; \theta_{13|2}]$$

(22)

Generally, the maximum likelihood estimates (MLEs) of unknown parameters $(\alpha, \theta_{12}, \theta_{23}, \theta_{13|2})$ can be obtained by maximizing the log-likelihood function. However, as the dimension of the parameter space in the proposed model is high, obtaining the estimated value by traditional MLE method may be difficult. This work utilizes inference functions for margins (IFM) method as proposed by Joe [40], which is a computationally attractive alternative to MLE for estimating parameters in multivariate copula models. This approach mainly divides the log-likelihood function into two parts: the contribution denoted by $L_m$ from the log-likelihood of each marginal distribution $L_k, k = 1, 2, 3,$ and the contribution denoted by $L_c$ from the dependence structure in data represented by copula. This provides

$$\ln L_m = \sum_{k=1}^{3} \sum_{i=1}^{N} \sum_{j=1}^{M} \ln [f_k (X_k) ; \alpha_k] \quad (23)$$

$$\ln L_c = \sum_{i=1}^{N} \sum_{j=1}^{M} \ln [c_{12} (F_1 (X_1), F_2 (X_2)) ; \theta_{12}]$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{M} \ln [c_{23} (F_2 (X_2), F_3 (X_3)) ; \theta_{23}]$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{M} \ln [c_{13|2} (F_1|2 (X_1|X_2), F_3|2 (X_3|X_2)) ; \theta_{13|2}]$$

(24)

The IFM approach estimates the parameters of each marginal distribution and each copula function in two stages. In the first stage, the parameters regarding each marginal distribution, i.e., $\alpha_k = (\mu_k, \sigma_k, q_k)$, are estimated from the corresponding $L_m$. In the second stage, the parameters of copula function are obtained through $L_c$. The estimated parameters of each marginal distribution and the copula function in the first stage are used to calculate input in the second stage. For comparison, a full Bayesian joint estimation is also calculated, which has only one stage and does not divide the parameters into two groups.

Stage 1. Estimation of parameters regarding the marginal distribution

In each marginal distribution for the performance indicators of reciprocating seals, the unknown parameters are the corresponding parameters in Wiener process. The likelihood function of the $k$th performance indicator is given by

$$L_k = \prod_{i=1}^{N} \prod_{j=1}^{M} [f_k (X_k)]$$

$$= \prod_{i=1}^{N} \prod_{j=1}^{M} \frac{1}{\sqrt{2\pi \sigma_k^2 \Delta t_{ij}}} \cdot \exp \left\{ -\frac{(X_k - \mu_k \Delta t (t_{ij}))^2}{2\sigma_k^2 \Delta t (t_{ij})} \right\} \quad (25)$$

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/.
Then, the log-likelihood $L_m$ can be given by

$$
\ln L_m = \sum_{k=1}^{3} \sum_{i=1}^{N} \sum_{j=1}^{M} \ln [f_k (X_k) ; \alpha_k]
$$

$$
= - \frac{1}{2} \sum_{k=1}^{3} \sum_{i=1}^{N} \sum_{j=1}^{M} \ln (2\pi \sigma^2_k \Delta \Lambda (t_{ij}))
$$

$$
- \frac{1}{2} \sum_{k=1}^{3} \sum_{i=1}^{N} \sum_{j=1}^{M} (X_k - \mu_k \Delta \Lambda (t_{ij}))^2
$$

(26)

The parameter set $\alpha_k = (\mu_k, \sigma_k, \theta_k)$ of marginal distribution can be estimated by Bayesian method. In the Bayesian approach the unknown parameters are treated as random variables and their probabilistic models are illustrated by posterior distributions. The posterior distribution of $\alpha$ can be given by

$$
\pi (\alpha | \Delta Y (t_{ij})) \propto \ln L_m (\Delta Y (t_{ij}) | \alpha) \cdot \pi (\alpha) \quad (27)
$$

where $\pi (\alpha)$ denotes the prior distribution of unknown parameters. In this study, non-informative prior distribution is utilized for $\alpha_k$ in which $\ln L_m (\Delta Y (t_{ij}) | \alpha)$ is the log-likelihood of the degradation model and $\pi (\alpha | \Delta Y (t_{ij}))$ denotes the posterior distribution. Markov chain Monte Carlo (MCMC) method is used here to construct a Markov chain, which constructs the stationary distribution as the prior distribution and produces samples of the posterior distribution through this chain. The parameter estimation process based on MCMC is carried out in the software OpenBUGS.

Stage 2. Estimation of parameters regarding copula function

In the first stage, the parameters of Wiener process for each performance indicator have already been obtained, and will be used in the second stage to calculate sample data of $F_k (t_{ij})$ for each performance indicator. The likelihood function $L_c$ denotes the dependence structure in data represented by D-vine copula, as shown in (24). Subscripts $\theta_{12}, \theta_{23}$ and $\theta_{13|2}$ are the parameters reflecting the dependent relationship in copula functions $C_{12}, C_{23},$ and $C_{13|2}$, respectively. Here, $\theta_{12}$ is taken as an example to illustrate the parameter estimation process in Stage 2. The unknown parameters are also estimated by Bayesian approach. The log-likelihood function corresponding to the dependence between performance indicators 1 and 2 can be given by

$$
\ln L_{c12} (\theta_{12}) = \sum_{i=1}^{N} \sum_{j=1}^{M} \ln [c_{12} (F_1 (X_1), F_2 (X_2)) ; \theta_{12}]
$$

(28)

The posterior distribution of $\theta_{12}$ can be expressed as

$$
\pi (\theta_{12} | F_1 (X_1), F_2 (X_2)) \propto \ln L_{c12} (F_1 (X_1), F_2 (X_2) | \theta_{12}) \cdot \pi (\theta_{12})
$$

(29)

where $\pi (\theta_{12})$ denotes the prior distribution of $\theta_{12}$. In this work, non-informative prior distribution is utilized for $\theta_{12}$. Here, $\ln L_{c12} (F_1 (X_1), F_2 (X_2) | \theta_{12})$ is the log-likelihood of the degradation model and $\pi (\theta_{12} | F_1 (X_1), F_2 (X_2))$ denotes the posterior distribution. The parameter estimation process is also based on MCMC and is carried out in the software OpenBUGS.

The calculation process for estimating $\theta_{23}$ is similar. When estimating $\theta_{13|2}, F_{12} (X_1 | X_2),$ and $F_{32} (X_3 | X_2)$ can be calculated by $F_1 (X_1), F_2 (X_2),$ and $F_3 (X_3)$ based on (10) and (11), after the optimal copulas of $C_{12}$ and $C_{23}$ are determined. Non-informative prior distribution is also utilized for $\theta_{23}$ and $\theta_{13|2}$. The remaining calculation process is also similar to the estimation process of $\theta_{12}$ and $\theta_{23}$, and the entire parameter estimation process is shown in Fig. 3.

Once the parameters have been estimated, reliability function can be calculated by (14) with Monte Carlo simulation, and MTTF can be obtained by (15). Other indices regarding reliability of reciprocating seals can also be determined.

III. MODEL VALIDATION

A. INTRODUCTION TO THE EXPERIMENT

Degradation testing of a certain model of reciprocating seal was carried out to verify the proposed model. The test rig is shown in Fig. 4(a), and a schematic of the test rig is provided in Fig. 4(b). For this experiment, the servo valve controls the driving cylinder which drives the test cylinder. The driving cylinder and the test cylinder were vertically installed and individually powered by the hydraulic pressure generated by different hydraulic circuits. The test cylinder reciprocated on the sliding guide rail, while the rod of the test cylinder remained stationary. A high precision displacement sensor was installed in the driving cylinder to control the dis-
placment. The temperature and pressure of the hydraulic oil in the test cylinder chamber were measured by temperature and pressure sensors. An electrical resistance strain gauge (a force sensor) was used to measure the friction force of the reciprocating seal, and a micro temperature sensor was utilized to monitor the temperature on the contact surface. The leaked hydraulic oil was collected in a measuring cup under the test cylinder.

Friction force, contact temperature, and leakage rate signal were collected and used to assess the degradation process of the reciprocating seals. After being conditioned by a signal conditioning module, the signal was collected by a data collection system. Data collection and processing software was programmed in National Instruments LabWindows CVI®, and data collection was automatically performed every four hours. For friction force and contact temperature signal, the sample rate was 200 Hz and each collection lasted for 20 seconds. The failure of each reciprocating seal being tested was manually judged by test operators based on friction force, contact temperature, and leakage rate. Once a reciprocating seal failed, it was dismantled from the test rig. A new seal was then installed and the test was resumed. The operating condition of the test sample is provided in Table 1.

### B. ESTIMATION OF PARAMETERS REGARDING EACH PERFORMANCE INDICATOR

Eight samples were tested in the experiment, and the test lasted for 500 hours. Plots of the three performance indicators of four samples are provided in Fig. 5.

As discussed in Section II.B, Wiener process is utilized to describe the degradation process of the three performance indicators of reciprocating seals, as shown in (1). Parameters in $\alpha_k = (\mu_k, \sigma_k, q_k)$ in the Wiener process must be estimated in the first stage of IFM method. The prior distributions of $\alpha_k$ are set as non-informative prior distribution, and the posterior distributions of these parameters are computed by OpenBUGS. When the MCMC simulation starts, the Gelman-Rubin ratio is utilized to check if the distribution has come to convergence from the sampled values of posterior distribution. Three Markov chains start with different initial values and the number of iterations is 100,000. Plots of Gelman-Rubin ratio are provided in Fig. 6. It can be observed in the figure that the Gelman-Rubin statistics of all parameters quickly become stable and converge after 500 iterations. A total of 80,000 iterations, i.e., from 20001 to 10000, are used to calculate the mean value of the unknown parameters in $\alpha_k$.

Posterior distributions of the parameters in $\alpha_k$ for the three performance indicators are shown in Fig. 7, and the mean value and the standard deviation for each parameter are shown in Table 2 and Table 3, respectively. Using this information, the marginal distributions of the three performance indicators and conditional cumulative distribution functions can then be calculated based on (3), (10) and (11), which are used as the input in the second stage of Bayesian inference.

### C. ESTIMATION OF PARAMETERS REGARDING DEPENDENCE AND COPULA DETERMINATION

Copula functions are used to describe the dependence among the three performance indicators and Bayesian copula selection method is utilized to choose the most suitable copula function. In this work the copula candidates include Clayton, Gumbel, Gaussian, and Farlie-Gumbel-Morgenstern (FGM) copulas. Clayton copula and Gumbel copula belong to Archimedean copula family, and Gaussian copula belong to elliptical copula family. As FGM copula does not belong to Archimedean or elliptical copula family, these four copula candidates are typical. The calculation results of weights and corresponding ranking for $C_{12}$, $C_{23}$, and $C_{13|2}$ are shown in Table 4. It can be seen that for $C_{12}$, Clayton copula has the highest weight, so Clayton copula is adopted to describe the dependence between friction force and contact temperature. For $C_{23}$, Gaussian copula has the highest weight, so Gaussian copula is used to describe the dependence between contact temperature and leakage rate. For $C_{13|2}$, Clayton copula also has the highest weight, and it is used to describe the dependence between the two conditional cumulative distribution functions, i.e., $F_{1|2}(x_1|x_2)$ and $F_{3|2}(x_3|x_2)$.

The scatterplots among the indicators and $F_{1|2}(x_1|x_2)$ and $F_{3|2}(x_3|x_2)$ are shown in Fig. 8. It can be seen from Fig.

<table>
<thead>
<tr>
<th>TABLE 1. Operating condition of test sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating conditions</td>
</tr>
<tr>
<td>Stroking length</td>
</tr>
<tr>
<td>Frequency</td>
</tr>
<tr>
<td>Temperature</td>
</tr>
<tr>
<td>Velocity</td>
</tr>
<tr>
<td>Pressure in the cylinder chamber</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 2. Mean value of unknown parameters of the three performance indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance indicator</td>
</tr>
<tr>
<td>Friction force</td>
</tr>
<tr>
<td>Leakage rate</td>
</tr>
<tr>
<td>Contact temperature</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 3. Standard deviation of unknown parameters of the three performance indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance indicator</td>
</tr>
<tr>
<td>Friction force</td>
</tr>
<tr>
<td>Leakage rate</td>
</tr>
<tr>
<td>Contact temperature</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 4. Weight calculation results for copula selection and corresponding ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight</td>
</tr>
<tr>
<td>Clayton</td>
</tr>
<tr>
<td>Gumbel</td>
</tr>
<tr>
<td>Gaussian</td>
</tr>
<tr>
<td>FGM</td>
</tr>
</tbody>
</table>
8 that the dependence between friction force and contact
temperature has lower tail dependence, and the dependence
between $F_{1/2} (x_1|x_2)$ and $F_{3/2} (x_3|x_2)$ also has lower tail
dependence. While the dependence between contact temp-
perature and leakage does not have any tail dependence. These
scatterplots are consistent with the copula selection results.

The parameters in copula functions $C_{12}$, $C_{23}$, and $C_{13|2}$
are estimated in the second stage of IFM method. The prior distributions of $\theta_{12}$, $\theta_{23}$, and $\theta_{13|2}$ are all set as non-
informative prior distribution, and the posterior distributions
of these parameters are computed by OpenBUGS. In this process, the Gelman-Rubin ratio is also utilized to check whether the distribution has come to convergence, and the results are shown in Fig. 9, along with the posterior distribution of the three parameters. The mean value and standard deviation of \( \theta_{12} \), \( \theta_{23} \), and \( \theta_{13|2} \) are provided in Table 5.

Akaike information criterion (AIC) is then used to verify the accuracy of Bayesian copula selection method. The results are shown in Table 6, where it can be seen that Clayton copula, Gaussian copula, and Clayton copula are also selected as the optimal copula for \( C_{12} \), \( C_{23} \), and \( C_{13|2} \), respectively. The selection results are the same as the results based on Bayesian copula selection method. However, the parameters \( \theta_{12} \), \( \theta_{23} \), and \( \theta_{13|2} \) in all copula candidates must be estimated in advance in order to obtain corresponding AIC value, which could be time consuming.

### TABLE 5. Mean value and standard deviation of parameters in Copula function

| Copula | \( C_{12} \) (Clayton) | \( C_{23} \) (Gaussian) | \( C_{13|2} \) (Clayton) |
|--------|------------------------|------------------------|------------------------|
| Mean value | 1.6241 | 0.9834 | 8.8080 |
| Standard deviation | 0.0423 | 0.0732 | 1.1133 |
| Kendall’s tau | 0.4481 | 0.8838 | 0.8150 |

### TABLE 6. AIC calculation results for copula selection and corresponding ranking

| Copula | \( C_{12} \) | \( C_{23} \) | \( C_{13|2} \) |
|--------|-------------|-------------|-------------|
| AIC | Ranking | AIC | Ranking | AIC | Ranking |
| Clayton | 654.22 | 1 | 215.30 | 3 | 230.51 | 1 |
| Gumbel | 716.51 | 3 | 140.27 | 2 | 407.55 | 3 |
| Gaussian | 693.99 | 2 | 114.06 | 1 | 343.24 | 2 |
| FGM | 779.53 | 4 | 774.20 | 4 | 783.30 | 4 |

### D. RELIABILITY ESTIMATION FOR RECIPROCATING SEALS BASED ON D-VINE COPULA

When all unknown parameters in Wiener process and Copula functions are estimated, the reliability function can be calculated and the mean time to failure (MTTF) can be obtained. The reliability function when the three performance indicators are independent and the reliability function when the dependence among the three performance indicators are described by three-variable Gaussian copula function are also calculated. Comparison of the three reliability functions is provided in Fig. 10. It can be seen from the figure that the reliability when dependence among the three performance indicators are described by three-variable Gaussian copula function is a little higher than that when the three performance indicators are independent, while the reliability when the dependence among the three performance indicators are
IV. CONCLUSION

A reliability estimation model for reciprocating seals based on multivariate dependence analysis was presented in this paper. Degradation behavior corresponding to each performance indicator of reciprocating seals was described by a Wiener process. The dependence among the three performance indicators was then described by D-vine copula function, and a Bayesian copula selection approach was proposed to determine the optimal copula and build a degradation model. Parameter estimation was conducted based on a two-stage Bayesian method. Degradation testing for reciprocating seals was then carried out and validation based on test data was conducted. Validation results show that the proposed model is effective in estimating the reliability of reciprocating seals and can achieve a superior goodness of fit. Suggestions for future work include using time-varying copulas to describe the dependence and locating appropriate functions as evolution parameters in the copula function.

REFERENCES


SHAOPING WANG the B.S., M.S., and Ph.D. degrees in mechanical engineering from Beihang University, Beijing, China, in 1988, 1991, and 1994 respectively. She is currently serving as a professor in School of Automation Science and Electrical Engineering, Beihang University, Beijing, China.

Her research interests include reliability engineering, network reliability, reliability testing, and flight control. She has authored or co-authored over 80 papers in journals and conference proceedings. She currently holds more than 20 patents.

Dr. Wang was a recipient of the Yangtze River Scholar in 2013, and the IEEE International Conference on Industrial Engineering and Applications Best Paper Award in 2017.

YUJIE QIAN received the B.S. degree in mechanical engineering from Beihang University, Beijing, China, in 2019. He is currently pursuing the M.S. degree in mechanical engineering at Beihang University, Beijing, China.

His research interests include fault diagnosis and prognostics of hydraulic components.

MILETA M. TOMOVIC received the BS in mechanical engineering from University of Belgrade, in 1979, MS in mechanical engineering from MIT, in 1981, and PhD in mechanical engineering from University of Michigan in 1991. He is currently serving as Chair of Engineering Technology Department, F. Batten College of Engineering and Technology, Old Dominion University, Norfolk, VA.

Prior to joining ODU he had seventeen years of teaching and research experience at Purdue University, with emphasis on development and delivery of manufacturing curriculum, conducting applied research, and engagement with Indiana industry. His research interests include design, manufacturing systems and processes, product lifecycle management, system dynamics and controls. He has co-authored one textbook on materials and manufacturing processes that has been adopted by over 50 national and international institutions of higher education. In addition, he has authored or co-authored over 60 papers in journals and conference proceedings, focused on applied research related to design and manufacturability issues, as well as issues related to mechanical engineering technology education.