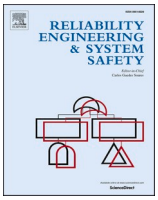
Reliability Engineering and System Safety 217 (2022) 108114

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress



Knowledge-based data augmentation of small samples for oil condition prediction

Yan Pan [a], Yunteng Jing [a], Tonghai Wu [a,b,*], Xiangxing Kong [c]

[a] Key Laboratory of Education Ministry for Modern Design and Rotor-Bearing System, Xi'an Jiaotong University, Xi'an, 710049, China
[b] Xi'an Jinghui Information Technology Co., Ltd, Xi'an, 710049, China
[c] Aero Engine Academy of China, Beijing, 101304, China

ARTICLE INFO

Keywords:
Data augmentation
Uncertainty
Wiener process
Particle filtering
Oil condition monitoring

ABSTRACT

Due to insufficient monitoring data, the reliability and accuracy of oil condition predictions are not guaranteed. Data-driven models provide data augmentation with small samples to solve this problem. However, the absence of degradation mechanisms would introduce unpredictable uncertainties in a long-term prediction. To address this, a data augmentation method is proposed for improved prediction by integrating degradation mechanisms and monitoring data. Primarily, a degradation model is established considering the degradation mechanisms. The model parameters are estimated with the time-vary probability distribution of the monitoring data. Therefore, the evidential variables are used to describe parameters with small samples. Then, the detailed parameters are estimated by integrating small-sample and prior parameters. With this well-trained model, the augmented data can be obtained with a particle filtering method for prediction. For validation, both the sparse and truncated samples from real-world monitoring are used to demonstrate the superiority of the proposed method. The high predicted accuracy demonstrates that the reliability of oil condition prediction can be guaranteed even with small samples.

1. Introduction

Oil condition monitoring (OCM) can provide traceable information on lubrication degradation in rotating machines. Modeling oil degradation with consistent monitoring serves as the fundamental way to oil state characterization [1] and remaining useful life (RUL) prediction [2]. Generally, the reliability and accuracy of prediction highly depend on the quantity of the historical data. However, due to the off-line operation and the gradual degradation of oil properties, the oil monitoring data contains dispersed and sparse small samples. The sparsity of the data is problematic for the extraction of the signal [3]. Therefore, the prediction in OCM has been limited by uncertainty due to such small samples. Data augmentation allows for solving such problems considering that it can reduce the epistemic uncertainty arising from the lack of data or knowledge [4].

For the oil degradation process, two types of samples are frequently encountered: sparse and truncated samples [5]. Sparse samples possess significant gaps between adjacent data points due to low sampling frequency or under-sampling [6]. Truncated samples hamper data reliability due to the lack of full-life data [7]. Additional uncertainties in prediction from such samples also include: 1) over-fitting; 2) data imbalance; 3) outliers; 4) poor parameter optimization [8]. To overcome these limitations, many strategies have been proposed at data, model, and parameter levels. At the data level, Monte Carlo (MC) sampling [9] and Virtual Sample Generation (VSG) [10] are adopted for data generation or augmentation. For the model level, generation models or regularization constraints [11] are used to enhance the extraction from small samples. Furthermore, optimization algorithms specified for small samples are developed at the parameter level, including the expectation-maximization (EM) algorithm [7] and restricted maximum likelihood estimation (MLE) algorithm [12]. These data augmentation methods can be applied to two types of problems arising from 1) the gaps in adjacent samples; 2) the distribution of small samples [13].

The gaps originate from the lack of sufficient data or the occurrence of missing data. The data-driven method is an efficient solution for filling gaps in adjacent samples, which can be classified into two categories, namely, interpolation [14] and extrapolation [5]. 1) Interpolation targets to fill individual data using the inherent or random data attributes. For example, He [15] applied nonlinear interpolation VSG to enhance the energy prediction on small samples. To prevent data

* Corresponding author.
E-mail address: wt-h@163.com (T. Wu).

https://doi.org/10.1016/j.ress.2021.108114
Received 3 May 2021; Received in revised form 19 September 2021; Accepted 27 September 2021
Available online 3 October 2021

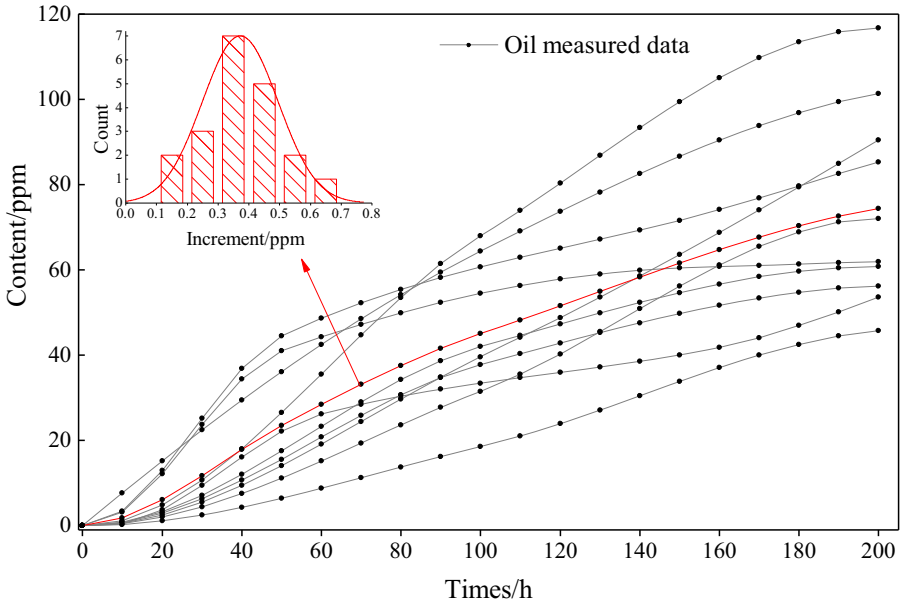


Fig. 1. The oil monitoring data subjected to the Wiener process.

imbalance in small samples, Li [16] proposed a trend-diffusion and tree-structure-based prediction approach for small sample sets, which employed random megatrend-diffusion to estimate bounds and used a heuristic mechanism to fill gaps between the samples. 2) Extrapolation realizes data augmentation based on prior training data over a certain period. This strategy is similar to the data-driven method for RUL, which predicts the time-series data under the Bayesian framework and the Markov property. The Bayesian framework connects the prior probability and the posterior probability, and Imai [17] demonstrated its effectiveness for marginal data augmentation with the prior information. The Markov property, which indicates the conditional probability distribution of a future state solely depends on the current state, is fundamental for Markov Chain Monte Carlo (MCMC) estimation with small samples [18]. In machine learning, data augmentation has been proven as an efficient method for fault identification with limited data [19]. Yoo [20] applied Generative Adversarial Networks (GAN) to combine actual test data and CAE data for small-sample prediction. These data-driven methods can obtain consistent and accurate augmented data to fill gaps. However, the uncertain distribution depends on the availability of the training set which is insufficient in small samples.

Empirical or mechanism knowledge can provide an indication of the small-sample distribution. Physics-based modeling describes the degradation process then constructs the sample distribution, so these methods can remedy the drawback in determining which distribution to adopt in data-driven models. Among them, empirical and physical models have emerged. For example, Liu [21] assumed that the degradation process was subjected to gamma or Gaussian processes in small-sample RUL. Furthermore, they explored the Wiener process with a Gaussian-distributed diffusion factor to describe the small-sample degradation [22]. Zhao [23] utilized the gray model (GM) to predict the reliability life data of wire rope under the small sample size condition. Addressing unknown parameters in distribution, two methods have been proposed. The method based on data augmentation can provide plausible solutions for parameter estimation, such as the MC sampling with MLE [9] and the modified MLE [12]. However, the inconsistent distribution of augmented data would deviate from the original distribution. Consequently, another strategy is to directly estimate the approximate parameters from small samples using interval estimation instead of point estimation. Liu [24] proposed the evidential variable to describe small samples with interval parameters, which has a higher ability to express parameter uncertainty. Nevertheless, it fails to obtain precise and detailed parameter estimation using overly wide intervals with small samples. Since physical-based models are susceptible to data variation, data-driven models are demanded to compensate and obtain more accurate results. On the other hand, the degradation mechanism can provide the basis for modeling. Hence, it introduces both the prior knowledge of state evolution while the time-varying data updates the parameters.

To improve predictive performance with small samples in OCM, a new data augmentation model integrating mechanism knowledge and monitoring data is developed in this work. A degradation model is constructed based on oil properties, then the initialization and update of the model parameters are realized with the monitoring data. Afterwards, a particle filtering (PF) method is applied to achieve the generation of augmented data for prediction. The major contributions are summarized as follows.

1) A new data augmentation model is proposed to integrate the advantages of physics-based and data-driven models to reduce the epistemic uncertainty in oil condition prediction with small samples. The data augmentation method provides complements for predictions.
2) To obtain accurate prediction data, a two-step data augmentation method is proposed. First, the initial augmentation is taken with small samples using evidence variables; second, the reliable augmentation is realized with the training data and the initial augmented data.
3) The random errors of the stochastic process are eliminated with the PF algorithm. For accurate prediction, the sequential importance sampling (SIS) strategy with the time-varying probability density function (PDF) is applied.

The rest of the paper is organized as follows. Section 2 provides an overview of the methods applied in this work. The integrated modeling strategy that combines the physics-based model and the data-driven model is introduced in Section 3. Section 4 describes the two-stage data augmentation process. In Section 5, the verification of the proposed model is presented. The analysis and discussion of the results are given in Section 6. Section 7 contains the conclusion.

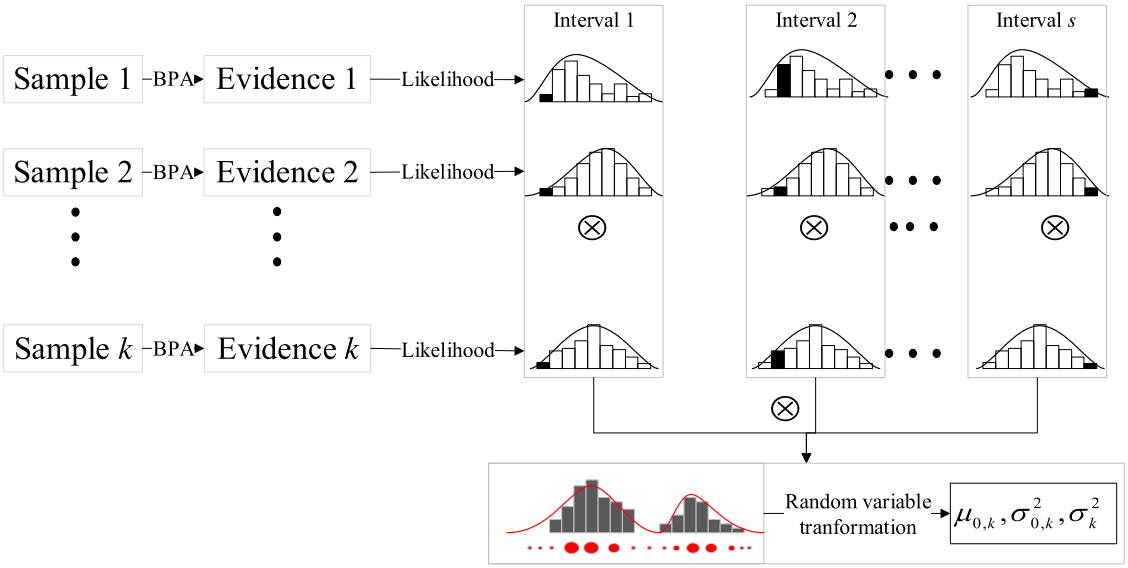


Fig. 2. An illustration of the evidential variables.

2. Related methods

In this section, related methods about the Wiener process, parameter estimation, and particle filtering are reviewed and discussed.

2.1. Wiener process

The general formulation of the Wiener process [25] that describes the system uncertainty consists of two parameters: the drift coefficient and the diffusion coefficient, as shown in Eq. (1),

$$x(t) = x(0) + bt + \sigma_B B(t), \tag{1}$$

where $x(t)$ denotes a measurement at moment t; $x(0)$ denotes the initial state; b is the drift coefficient; and σ_B is the diffusion coefficient. $B(t)$ is a standard Brownian motion that satisfies three basic properties: $B(0) = 0$; for any moment $0 \le \tau \le t$, $B(t) - B(\tau)$ satisfies the Gaussian distribution $N(0, t-\tau)$; $B(t)$ has independent increments.

The oil degradation over time can be described as a stochastic process. As shown in Fig. 1, ten groups of oil monitoring data are proven to satisfy the Wiener process in Eq. (1).

2.2. Parameter estimation

Due to the inevitable uncertainties in stochastic processes, parameter estimation with small samples should be able to: 1) address uncertainties; 2) adopt prior information.

2.2.1. Evidential variables

When the interval of the parameter is divided into multiple groups, the estimations fall into the corresponding intervals instead of definite points. Let each estimate be the evidential variable [24], and then the complete distribution can be obtained based on the Dempster-Shafer (D-S) theory. The illustration of the evidential variables is shown in Fig. 2.

Suppose that independent samples x_1 and x_2 constitute the identification framework and all possible propositions construct a power set $2^X = \{\varnothing, \{x_1\}, \{x_2\}, \{x_1, x_2\}\}$. The basic probability assignment (BPA) describes the belief degree of sub-propositions in the identification framework. The belief mass satisfies the following conditions.

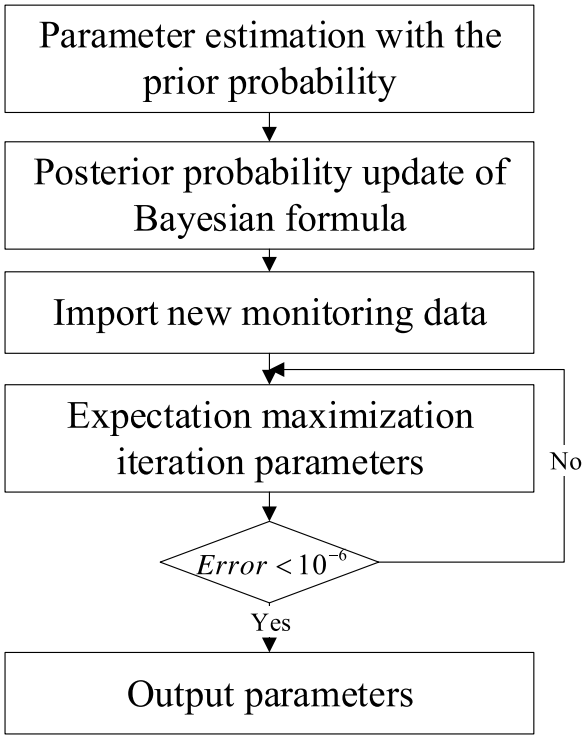


Fig. 3. The iteration process of EM algorithm.

$$\begin{aligned} &m(A) \ge 0; \forall A \in 2^X, \\ &m(\varnothing) = 0, \\ &\sum_{A \in 2^X} m(A) = 1. \end{aligned} \tag{2}$$

Two belief mass, or pieces of evidence, m_1 and m_2 can be combined as shown in Eq. (3),

$$m_r(\mathbf{\Theta}_i) = \frac{\sum_{B \cap C = \mathbf{\Theta}_i} m_1(B) m_2(C)}{1 - \sum_{B \cap C = \varnothing} m_1(B) m_2(C)}, \tag{3}$$

where B and C are corresponding sub-intervals, Θ_i is the combined interval and $B \ne \varnothing, C \ne \varnothing$.

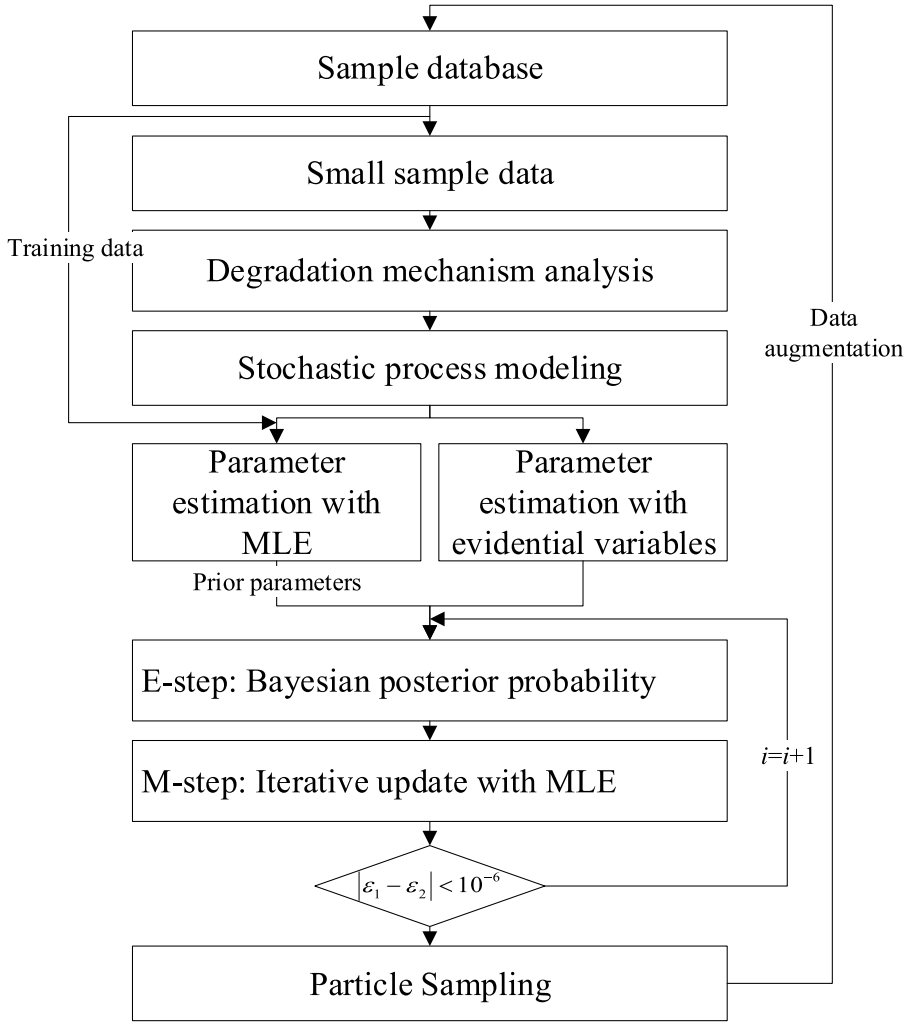


Fig. 4. Flowchart of augmentation-based prediction procedure.

2.2.2. Parameter update

The introduction of prior knowledge can facilitate small-sample parameter estimation, and the Bayesian formula can bridge the prior and posterior probabilities. The prior probabilities are derived from training samples, and the posterior probabilities can be calculated as in Eq. (4).

$$P(\Theta_k|Y_{1:k}) = \frac{P(Y_{1:k}|\Theta_k)P(\Theta_k)}{\sum_{i=1}^{n} P(\Theta_i)P(Y_{1:i}|\Theta_i)}, \tag{4}$$

where $P(\Theta_k|Y_{1:k})$ is the posterior probability, $P(Y_{1:k}|\boldsymbol{\Theta}_k)$ is the likelihood probability, and $P(\Theta_k)$ is the prior probability.

The EM algorithm maximizes the likelihood expectation of the hidden variables. This algorithm has been employed for parameter estimation of small samples with truncated or censored data [7]. In the parameter update process, the expectation is calculated based on the prior probability in the E step; and the maximum likelihood expectation is searched for in the M-step. The specific process is shown in Fig. 3.

2.3. Particle filtering

The Bayesian theory provides a framework for using prior knowledge, and then the parameter update describes the evolution of oil degradation. The evolved model distribution should be represented by sampling to realize data augment. PF with SIS [26] allows extracting the posterior distribution by Bayesian filtering. The non-linear and non-Gaussian sampling satisfies the requirement for parameters update with small samples.

The basic idea of PF is state detection by sampling particles in the state equation and weighting them with observation. Considering the evolving distribution of the degradation model, PF is suitable to trace the oil state because it can incorporate real-time measurements and instantly update the state [27].

3. Integrated degradation modeling strategy

For predicting the oil degradation in OCM, the physics-based modeling method is studied by involving the chemical reaction and material balance principles. For the PDF, the mechanism model is converted into a Wiener process on the state-space model of the oil degradation [28].

3.1. Degradation mechanism analysis

Oil degradation essentially originates from the chemical reaction that involves a critical equilibrium point between the generation of products and the corresponding loss of reactants. The equations based on mechanism analysis can be used to describe such processes including the Arrhenius equation [29], kinetic reaction equation [30], material balance equation [31], etc. Moreover, oil degradation is often accompanied by the production of wear debris. These particles flowing in the oil present a dynamic concentration equilibrium due to filtration, settlement, and lubricant consumption [32]. Therefore, a dynamically

time-varying concentration equilibrium can be established in OCM.

The general mathematical expression of the physical and chemical degradation mechanisms is shown in Eq. (5).

$$\frac{d(c)}{dt} = -k(r(c) - p(c)), \tag{5}$$

where c denotes the measurement indicator, t denotes the sampling time, $r(c)$ corresponds to the indicator increment variable, and $p(c)$ denotes the indicator decrement variable.

The corresponding expression for $c(t)$ can be obtained by substituting the mechanism equation. For example, when the mechanism analysis is provided by the first-order kinetic reaction or the abrasive particle concentration equilibrium in the steady-state [32], the solution of Eq. (5) can be expressed as:

$$c(t) = k(1 - \exp(-\lambda t)) + c(0), \tag{6}$$

where $c(t)$ is the monitoring data at time t and $c(0)$ is the initial value; k, λ are random variables.

3.2. Degradation model

Variables k and λ can be estimated by interpolated least squares [1] with enough samples. Considering extensive random fluctuations exhibited by small samples, the stochastic process is used to describe these variables. The computation can be simplified by logarithmic transformation, suppose that $x(t) = \ln(k - c(t))$ and $b = -\lambda$, then Eq. (6) can be converted to the state-space model (SSM) with time-varying drift coefficients as shown in Eq. (7), where b denotes the degradation rate of the system and subject to the Gaussian distribution $N(\mu_0, \sigma_0^2)$. Then the SSM is established as follows.

$$\begin{cases} x_k = x_{k-1} + b_{k-1}(t_k - t_{k-1}) + \omega_{k-1} \\ y_k = x_k + \nu_k \end{cases}, \tag{7}$$

where x_k denotes the state value at moment t_k; y_k is the observation at moment t_k; $b_{k-1} \sim N(\mu_0, \sigma_0^2)$; $\omega_{k-1} \sim N(0, \sigma_B^2 \Delta t_{k-1})$, $\nu_k \sim N(0, \sigma_\nu^2)$ are respectively the state transition error and the observation error.

When the observation error is neglected, the smoothed observation series $\{\widetilde{y}_k - \widetilde{y}_{k-1} - b_{k-1}\Delta t_{k-1}, k = 1, 2, \cdots, K\}$ obeys the following distribution.

$$P(Y_{1:k}|\mu_0, \sigma_0^2, \sigma_B^2) = \frac{1}{\prod_{j=1}^{k}\sqrt{2\pi\sigma^2(t_j - t_{j-1})}} \exp\left[-\sum_{j=1}^{k}\frac{\left(\Delta\widetilde{y}_{j-1} - b(t_j - t_{j-1})\right)^2}{2\sigma^2(t_j - t_{j-1})}\right], \tag{8}$$

where $\sigma^2 = \sigma_0^2(t_j - t_{j-1}) + \sigma_B^2$ and $\Delta\widetilde{y}$ is the difference between adjacent smoothed observations.

Furthermore, the observation series $\{(\widetilde{y}_k - \widetilde{y}_{k-1})/\Delta t_{k-1}, k = 1, 2, \cdots, K\}$ obeys the following distribution:

$$P(Y_{1:k}|\mu_0, \sigma_0^2) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(b - \mu_0)^2}{2\sigma_0^2}\right]. \tag{9}$$

4. Two-stage small-sample prediction

The flowchart of the augmentation-based prediction procedure is shown in Fig. 4, which mainly consists of a two-stage procedure. 1) In The initial data augment stage, the degradation model is constructed and the estimation of the initial parameters with evidential variables. 2) In the updated data augment stage, combining with the initial prior parameters, the parameters are updated and then data generation is taken by PF.

4.1. The initial data augmentation

The degradation model involves two types of parameters, shift coefficient $b_{k-1} \sim N(\mu_0, \sigma_0^2)$ and diffusion coefficient $\omega_{k-1} \sim N(0, \sigma_B^2 \Delta t_{k-1})$. The evidential variables, using intervals instead of points and apply the likelihood function to construct BPA, are combined to determine the probability. To obtain the continuous output, the evidential variables are transformed into random variables. Finally, data augmentation is accomplished based on the initialization parameters.

4.1.1. Priors parameter initialization

The distribution of the degradation model is illustrated in Eq. (8), in which the likelihood function is constructed as $L(\Theta y_k) = P(Y_{1:k}\mu_0, \sigma_0^2, \sigma_B^2)$. An evidence chain is constructed for the data at different moments and the parameter is assigned to the corresponding interval. For example, the j-th parameter is divided into s intervals, $\{[\Theta_1^j, \Theta_2^j], (\Theta_2^j, \Theta_3^j], \cdots, (\Theta_s^j, \Theta_{s+1}^j]\}$, where $\Theta = \{\mu_0, \sigma_0^2, \sigma_B^2\}$. At the k-th moment, the likelihoods that respectively fall into the i th interval and the overall intervals can be approximately expressed as:

$$\begin{aligned} L(\Theta_i|y_k) &= \int_{\theta_i}^{\theta_{i+1}} L(\Theta|y_k)d\Theta \approx \frac{L(\Theta_{i+1}|y_k) + L(\Theta_i|y_k)}{2}(\Theta_{i+1} - \Theta_i), \\ L(\Theta|y_k) &= \int_{\Theta_{min}}^{\Theta_{max}} L(\Theta|y_k)d\Theta \approx \sum_{i=1}^{s} L(\Theta_i|y_k). \end{aligned} \tag{10}$$

The number of intervals and the interval boundary values $\{\Theta_i, i = 1, \cdots, s\}$ are set according to the estimation with training data or expertise. The parameter set Θ in Eq. (10) involves three parameters $\{\mu_0, \sigma_0^2, \sigma_B^2\}$ with mutually independent intervals. To collect the complete parameter information, all possible interval combinations are traversed.

The BPA for the k-th moment at the i th interval is

$$m_k(\Theta_\mathbf{i}) = \frac{L(\Theta_\mathbf{i}|y_k)}{L(\Theta|\mathbf{y_k})}, \tag{11}$$

where $m_k(\Theta_i)$ denotes the probability that the parameters fall into the i th interval at the k-th moment.

After the BPA, the probability $m_r(\Theta_i)$ that the collected small samples at moment k belong to the i th interval can be performed by Eq. (3). The above process combines the previous joint probability $m_r(\Theta_i)$ with the new BPA, then the belief degree of the overall intervals can be obtained recursively.

The belief degree $m_r(\Theta)$ means the probabilities that the parameter value belongs to the possible parameter intervals $\{[\Theta_1, \Theta_2], (\Theta_2, \Theta_3], \cdots, (\Theta_s, \Theta_{s+1}]\}$. To obtain the continuous parameter value, it is necessary to transform the evidential variables into random variables, as shown in Eq. (12).

$$\widehat{\Theta}_k = \int_0^{+\infty} \Theta f(\Theta)d\Theta \approx \sum_{i=1}^{s} \frac{(\Theta_{i+1} + \Theta_i)m_r(\Theta_i)}{2}, \tag{12}$$

where $\widehat{\Theta}_k$ denotes the parameter estimation for the k-th samples at the corresponding moment and $f(\Theta)$ denotes the probability density of the corresponding interval.

4.1.2. Sample generation

Suppose that the neighboring samples in a small-sample set have consistent parameters, the initial data augmentation is based on the initialized parameters $\widehat{\Theta}_k$. Define sparse data $\boldsymbol{Y}_s = \{y_n, n = 1, 2, \cdots, N_t\}$ and augmented data $\boldsymbol{Y}_e = \{y_k, k = n + i\}$, where i is the number of data augmentation, N_t denotes the number of small samples, y_n are the small samples. The model parameters corresponding to a certain moment can be estimated with the evidential variables, namely $\Phi_s = \{\Theta_\mathbf{n}, n = 1, 2, \cdots, N_t\}$. The gaps of adjacent samples are filled by Eq. (6) with $\Theta_\mathbf{k} = \Theta_\mathbf{n}$, where $n = 1, 2, \cdots, N_t, k = n + i$.

For truncated data $\boldsymbol{Y}_c = \{y_n, n = 1, 2, \cdots, N_t\}$ and augmented data $\boldsymbol{Y}_e = \{y_k, k = N_t + i\}$, where i is the number of data augmentation. The model parameters are estimated by evidential variables with small samples y_n, then $\Phi_s = \{\Theta_n, n = 1, 2, \cdots, N_t\}$ is obtained. The truncated data are predicted with the assumption that $\Theta_k = \Theta_{N_t}$, which is substituted into Eq. (6), thus truncated data augmentation is achieved.

4.2. The updated data augmentation

The initial augmented data are obtained by applying SSM and initialized parameters, defined as $\boldsymbol{Y}_e = \{y_k, k = 1, 2, \cdots, N_e\}$, where N_e is the length of samples. The initialized parameters are the approximate estimated values, but the distribution of the model evolves with the introduction of the augmented data. Thus, the model parameters need to be updated. The updated process involves two steps:

1) The parameters are updated by the EM algorithm.
2) Based on the updated parameters, PF is used to generate the data.

4.2.1. Parameter update strategy

The sequence $\{(\widetilde{y}_k - \widetilde{y}_{k-1})/\Delta t_{k-1}\}$, $k = 1, 2, \cdots, N_e$, is constructed from the smoothed training data to compute the drift coefficient parameters $\{\mu_0, \sigma_0^2\}$. Furthermore, the diffusion coefficient parameters are calculated by $\sigma_B^2 = \sigma^2 - \sigma_0^2 \Delta t_{k-1}$, where σ^2 denotes the variance of the smoothed series $\Delta\widetilde{y}_{1:k-1} = y_{2:k} - y_{1:k-1}$. The prior parameter set $\{\mu_0, \sigma_0^2, \sigma_B^2\}$ is obtained by applying MLE with the training data.

The EM algorithm is applied to update the parameters based on the prior parameters. The EM algorithm includes the E-step and M-step.

In the E-step, the expectation is generated from the Bayesian formula.

$$P(\mu_0, \sigma_0^2, \sigma_B^2 | Y_{1:k}) = \frac{P(Y_{1:k} | \mu_0, \sigma_0^2, \sigma_B^2) P(\mu_0, \sigma_0^2)}{P(Y_{1:k})} \propto P(Y_{1:k} | \mu_0, \sigma_0^2, \sigma_B^2) P(\mu_0, \sigma_0^2). \tag{13}$$

In the M-step, the parameters iteration is performed by MLE.

$$\widehat{\Theta}_k^{i+1} = \operatorname{argmax} \ln\left(\Theta | \widehat{\Theta}_k^i\right). \tag{14}$$

where $\boldsymbol{\Theta}_k = [\mu_{0,k}, \sigma_{0,k}^2, \sigma_{B,k}^2]$, denotes the prior parameters computed by the k-th monitoring data; $\widehat{\boldsymbol{\Theta}}_k = [\widehat{\mu}_{0,k}, \widehat{\sigma}_{0,k}^2, \widehat{\sigma}_{B,k}^2]$, indicates the updated parameters of the k-th monitoring data; i denotes the current number of iteration steps.

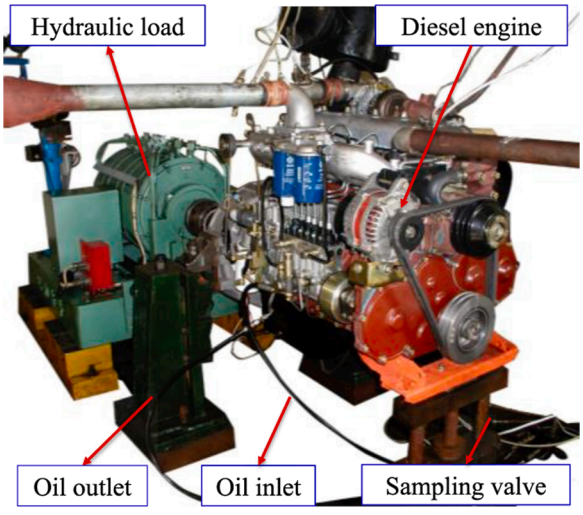


Fig. 5. The sampling and the test equipment.

Eq. (13) represents the joint probability as Gaussian distributions, whose expectation and variance can be derived from the analytical solution as shown in Eq. (15).

$$\begin{aligned} \widehat{\mu}_{0,k}^{i+1} &= \frac{\mu_0 \sigma_B^2 + y_k \sigma_0^2}{t_k \sigma_0^2 + \sigma_B^2}, \\ \widehat{\sigma}_{0,k}^{2\ i+1} &= \frac{\sigma_0^2 \sigma_B^2}{t_k \sigma_0^2 + \sigma_B^2}. \end{aligned} \tag{15}$$

The log-likelihood function of the posterior probability in the M-step is

$$\begin{aligned} lnP(Y_{1:k}, \mu_0, \sigma_0^2, \sigma_B^2 | \Theta_k) &= lnP(Y_{1:k} | \mu_0, \sigma_0^2, \sigma_B^2, \Theta_k) + lnP(\mu_0, \sigma_0^2 | \Theta_k) \\ &= -\frac{k+1}{2} ln 2\pi - k ln\sigma - \frac{1}{2} ln \sum_{j=1}^{k} (t_j - t_{j-1}) \\ &- \sum_{j=1}^{k} \frac{(y_j - y_{j-1} - b(t_j - t_{j-1}))^2}{2\sigma^2 (t_j - t_{j-1})} - k ln\sigma_{0,k} - \frac{(b - \mu_{0,k})^2}{2\sigma_{0,k}^2}. \end{aligned} \tag{16}$$

By solving the partial differential equation for σ_B^2 in Eq. (16) and set it to zero, then σ_B^2 can be expressed as the function with $\mu_{0,k}$ and $\sigma_{0,k}^2$.

$$\widehat{\sigma}_{B,k}^{2\ i+1} = \frac{1}{k} \sum_{j=1}^{k} \frac{(y_j - y_{j-1})^2 - 2\mu_{0,k}(y_j - y_{j-1})(t_j - t_{j-1}) + (t_j - t_{j-1})^2 \left(\mu_{0,k}^2 + \sigma_{0,k}^2\right)}{(t_j - t_{j-1})}. \tag{17}$$

The initial augmented data $\boldsymbol{Y}_e = \{y_k, k = 1, 2, \cdots, N_e\}$ are substituted into Eqs. (15) and (17), and the parameters $\widehat{\mu}_0, \widehat{\sigma}_0^2, \widehat{\sigma}^2$ are updated iteratively.

4.2.2. Sample generation

With the updated parameters $\widehat{\mu}_0, \widehat{\sigma}_0^2, \widehat{\sigma}^2$, the expectation of particles in the PF is used as the predicted data. After parameter update, the distribution of the model becomes:

$$P(Y_{1:k} | \mu_0, \sigma_0^2, \sigma_B^2) = \frac{1}{\prod_{i=1}^{k} \sqrt{2\pi \widehat{\sigma}_{0,i}^2 \Delta t_i^2 + \widehat{\sigma}_{B,i}^2 \Delta t_i}} \exp\left[-\sum_{i=1}^{k} \frac{\left(\Delta\widetilde{y}_i - \widehat{\mu}_{0,i} \Delta t_i\right)^2}{2\left(\widehat{\sigma}_{0,i}^2 \Delta t_i^2 + \widehat{\sigma}_{B,i}^2 \Delta t_i\right)} \right]. \tag{18}$$

SIS is performed on the updated distribution, and the particle set $\{x_k^i, i = 1, 2, \cdots, N_s\}$ is generated randomly, where N_s is the number of particles. The state particles are recursively expressed as follows:

$$x_k^i = x_{k-1}^i + b_{k-1}^i (t_k - t_{k-1}) + w_{k-1}^i. \tag{19}$$

The weights of the particles is then updated by substituting the observations into Eq. (20):

$$\widetilde{\omega}_k^i = P\left(y_k \middle| x_k^i, \widehat{\sigma}_\nu^2\right) = \frac{1}{\sqrt{2\pi \widehat{\sigma}_\nu^2}} \exp\left[-\frac{(y_k - x_k^i)^2}{2\widehat{\sigma}_\nu^2} \right], \tag{20}$$

where the observed variance $\widehat{\sigma}_\nu^2$ is obtained from the smoothed data with MLE.

The result of the prediction is

$$\widehat{y}_k = \widehat{x}_k = \sum_{i=1}^{N_s} (\omega_k^i x_k^i), \tag{21}$$

where $\omega_k^i = \widetilde{\omega}_k^i / \sum_{i=1}^{N_s} \widetilde{\omega}_k^i$.

To prevent particle impoverishment during SIS, the number of valid particles is evaluated after each update step,

$$N_{eff} = \frac{1}{\sum_{i=1}^{Ns} (w_k^i)^2}. \tag{22}$$

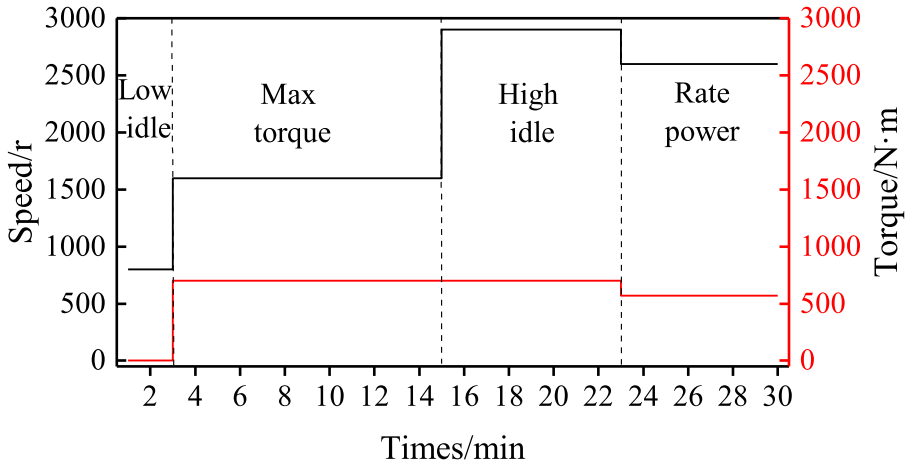


Fig. 6. Condition of engine reliability bench.

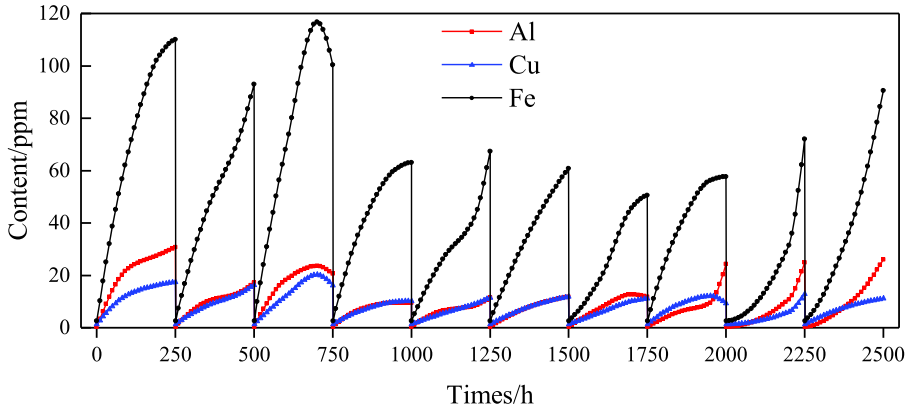


Fig. 7. The original data of wear element content in engine oil monitoring.

Resampling is required when the number of efficient particles N_{eff} is less than the assigned value, and the resampling weights are reset to $w_k^i = 1/N$.

5. Case study

The data for model verification was derived from 2500 h full life-cycle engine oil monitoring, in which the oil change interval was set to 250hr. We obtained a database of 250 sets of engine oil with a total of ten operating cycles. The experimental data were obtained from the 2500 h reliability bench test, on the EQD210–10 diesel engine, conducted by the Dongfeng Motor Company. The specification and grade of engine oil were 10 w/40 and CD respectively. Oil samples were collected through a valve installed at the oil drain port as shown in Fig. 5.

5.1. Sample database

To match the dynamics of the diesel engine, the bench test was set up according to the loads specified in the standard [33]. The mixed loading condition of the bench experiment is shown in Fig. 6. There are four conditions: 1) 3 min at low idling speed, 2) 12.5 min at the maximum torque, 3) 7 min at high idling speed, and 4) 7.5 min at the rated condition. The speed and load are changed evenly within 1 min. Each cycle lasted 30 min, and the total number of cycles is 5000.

Different from instantaneous engine dynamics characteristics, the debris in oil carried the accumulated wear from the initial state under circulating oil circuit. To obtain enough observation data, the oil was collected every 10 hr from the same part and tested instantaneously. Spectral analysis was performed to detect the concentration of metal debris. Considering the composition of typical engine wear parts, the contents of Fe, Cu and Al in oil were selected for monitoring. The three wear elements have a consistent trend in concentration as shown in Fig. 7. The robustness of data from the Fe element is more suitable for validation.

All 250 sets of oil data are used as the training set to calculate the prior parameters. To verify the accuracy and precision from data augmentation, two types of test sets were used, 1) sparse test set, which selects identical intervals or 1/3 of the data randomly from one operation cycle; 2) truncated test set, which selects the first 40%, 60% and 80% data from one operation cycle.

5.2. Augmentation-based prediction with small samples

The wear concentration maintains a dynamic equilibrium during the lubrication cycle. The model only involves the wear mechanism in the steady-state period as it neglects the break-in period and the regular oil change intervals. As shown in Eq. (1), the Wiener process is obtained after the exponential transformation of the mechanism equation. The procedures for prediction with the combined mechanism-data-driven model are shown in Algorithm 1.

1) *Sparse sample case*: Apply the evidential variables method for parameter estimation with ten groups of sparse test sets, then obtain small-sample parameter sets $\{\tilde{\mu}_0, \tilde{\sigma}_0^2, \tilde{\sigma}_B^2\}$. The initial augmented data is generated with SSM. The parameters of the training samples are taken by MLE, forming the prior parameter set $\{\mu_0, \sigma_0^2, \sigma_B^2\}$. Then

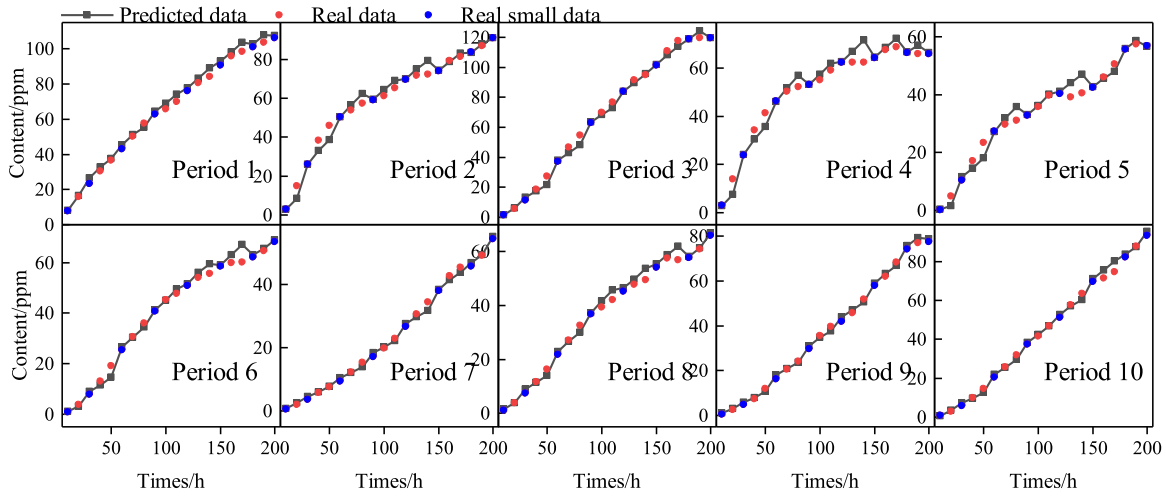


Fig. 8. Comparison of data augmentation with sparse sample.

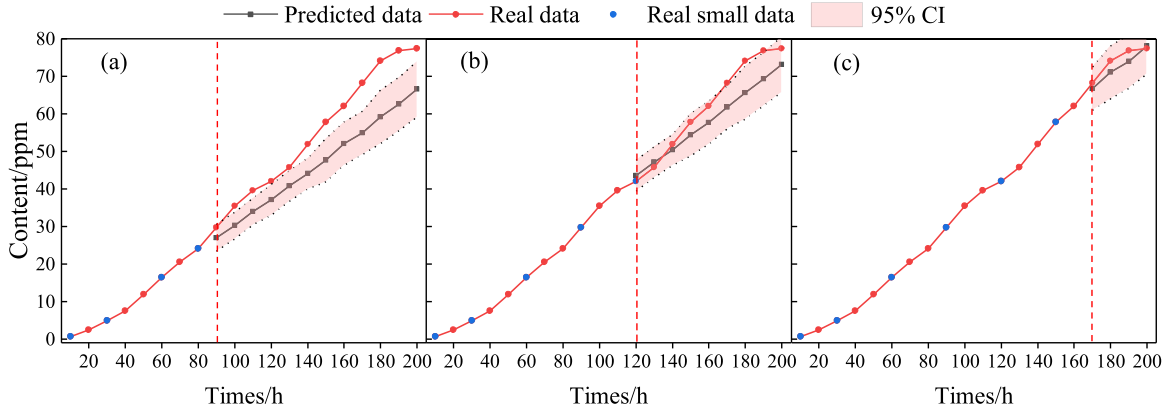


Fig. 9. Comparison of data augmentation with truncated sample, (a) 40% data augment, (b) 60% data augment, (c) 80% data augment.

apply the EM algorithm to update the parameters $[\widehat{\mu}_0, \widehat{\sigma}_0^2, \widehat{\sigma}_B^2]$. PF is adopted based on the updated parameter set, and the expectations of sampled particles are used as the augmented data. As shown in Fig. 8, the predicted data in black shows consistency with the real data in red based on the small sample data in blue. It can be seen that the real data presents obviously fluctuant and the predicted data can be accurately augmented to fill the gaps between the adjacent samples. The results show a degradation trend consistent with the real values, which accurately describes the wear evolution.

2) *Truncated sample case*: The augmented steps for truncated data are the same as those for sparse data. When the truncated data is predicted based on sparse samples, it involves two steps: 1) the sparse data in the truncated samples is augmented; 2) the augmented data in the truncated samples is predicted. It is shown that the method can achieve data augmentation based on small samples. Then 40%, 60%, and 80% of the full life-cycle monitoring data are selected for the test. To evaluate the prediction performance, convergence is introduced to describe the results. The results are shown in Fig. 9. The prediction accuracy improves with the increase of monitoring data, and the confidence interval (CI) tends to converge, which indicates the more reliable prediction results. It can be shown that the better converge with the more training data, which verifies the high prediction accuracy with sufficient data.

5.3. Performance evaluation

In order to assess predictive performance, two characters of temporal and spatial scales needed to be quantified, including 1) quantify how accurately an algorithm performs at a given time relative to prediction, 2) if the performance converges (i.e., satisfies accuracy) quantify how fast does it converge [26]. Therefore, two evaluation indexes were adopted, 1) cumulative relative accuracy (CRA), 2) convergence speed (CS). CRA is used to evaluate the accuracy of the overall data augmentation process, whose values closer to 1 indicate better performance. CS is the distance between origin and centroid of area under the curve, reflecting the prediction convergence rate. The smaller index denotes a faster convergence rate.

The cumulative relative accuracy is calculated by Eq. (23).

$$\mathrm{CRA} = \frac{1}{K}\sum_{k=1}^{K}\left(1 - \frac{\left|\widehat{l}_k - l_k\right|}{l_k}\right), \tag{23}$$

where K is the length of the sequence; $\widehat{l}_k$ and l_k represent the augmented data and the real data at moment t_k, respectively.

The convergence speed is computed from Eq. (24).

$$\begin{aligned} &\mathrm{CS} = \sqrt{x_{cen}^2 + y_{cen}^2}, \\ &x_{cen} = \frac{\sum_{k=1}^{K-1}\left(t_{k+1}^2 - t_k^2\right)RA_k}{2\sum_{k=1}^{K-1}(t_{k+1} - t_k)RA_k}, y_{cen} = \frac{\sum_{k=1}^{K-1}(t_{k+1} - t_k)RA_k^2}{2\sum_{k=1}^{K-1}(t_{k+1} - t_k)RA_k}, \end{aligned} \tag{24}$$

where $RA_k = 1 - \frac{|\widehat{l}_k - l_k|}{l_k}$ is the relative accuracy of the augmented data at

Table 1
CRA with sparse prediction.

No.	1	2	3	4	5	6	7	8	9	10
CRA	0.9574	0.9353	0.9491	0.9339	**0.8636**	0.9387	0.9272	0.9412	0.9435	0.9351

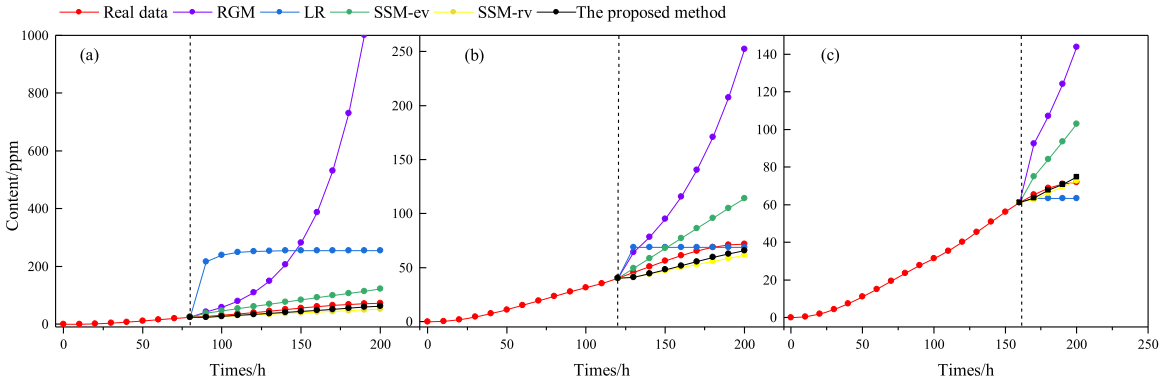


Fig. 10. Comparison of four methods for data augmentation, (a) 40% data augment, (b) 60% data augment, (c) 80% data augment.

Table 2
Comparison of CRA and CS for prediction with small samples.

Indicator	CRA			CS		
	40%	60%	80%	40%	60%	80%
RGM	–	–	0.3223	–	–	125.0
LR	–	0.8312	0.9141	–	122.4	139.5
SSM-ev	0.4696	0.6889	0.7218	128.2	122.6	124.5
SSM-rv	0.7552	0.8419	0.9712	127.2	124.5	122.0
The proposed method	**0.8272**	**0.8742**	**0.9781**	**109.7**	**107.2**	**105.4**

"-" represents the negative value due to the predicted value being greater than twice the real value.

t_k, and (x_{cen}, y_{cen}) is the center of mass of the area under the relative accuracy between t_{k+1} and t_k.

The CRA of data augment with sparse samples are all greater than 0.85 for the 10 groups of samples as shown in Table 1.

Considering the exponential degradation of the oil condition, exponential-based prediction models were selected for comparison in the verification process. Physics-based and data-driven models, including rolling gray model (RGM) [34], logistic regression (LR) [35], SSM with evidence variables (SSM-ev) [22], and SSM with random variables (SSM-rv) [36], were used for data augmentation with small samples. The SSM-rv adopted the augmented data in the first stage, whose major difference with the proposed method is the methods of obtaining likelihood probabilities. The data augmentation was performed with the same sample set and predicted values of 40%, 60%, and 80% of samples were compared, as shown in Fig. 10. It can be seen that the predicted values of the four methods gradually converge to the real values with the increasing samples. The RGM, whose prediction is dependent on the average of the adjacent samples, has the poorest performance caused by epistemic uncertainty due to the absence of samples. The LR has a high prediction accuracy with enough samples (80%), in other words, the small samples (40%) are not suitable for the LR without parameters update. The SSM-ev presents a poorer performance than the SSM-rv that applies the augmentation data, indicating that data augmentation can effectively improve the prediction accuracy. In addition, interval parameter estimates of the evidence variables present worse precision than point estimates of random variables. Comparing SSM-rv and the proposed method reveals that the prior knowledge induces a better prediction with the data augmentation. The proposed method presents the best prediction accuracy throughout the evaluation.

Algorithm 1
Procedures of augmentation-based prediction.

Step 1:Parameter estimation with training samples using MLE, then form the prior parameter set $\{\mu_0, \sigma_0^2, \sigma_B^2\}$.
Step 2:Parameter estimation with test samples using Eqs. (10)–(12) to obtain the initial parameter set $\{\widetilde{\mu}_0, \widetilde{\sigma}_0^2, \widetilde{\sigma}_B^2\}$.
Step 3:Import the initial parameters into the model and apply Eq. (7) for initial augmentation, then generated $Y_e = \{y_k, k = 1, 2, \cdots, N_e\}$.
Step 4:Iteratively update parameter using Eqs. (15) and (17).
Step 5:The iteration ends until $|L_t - L_{t-1}| \leq \varepsilon$, where $\varepsilon = 10^{-6}$, then obtain the updated parameter set $[\widehat{\mu}_0, \widehat{\sigma}_0^2, \widehat{\sigma}_B^2]$.
Step 6:Sampling is performed using Eq. (21) to obtain the predicted data.

For quantitatively characterizing the performance, the CRA and CS are respectively calculated by Eqs. (23) and (24), and the results are shown in Table 2. The proposed method obtains the highest CRA and the best CS for all the verification samples. This further demonstrates that the proposed method presents accurate predictions and simultaneously decreases the random fluctuations.

6. Discussion

The high prediction accuracy of the method is attributed to the reduction of epistemic uncertainty through data augmentation [4]. The novelty of the method involves three aspects: 1) interpretable modeling of the evolving process of the embedding mechanism, 2) parameter updates using evidence variables with small samples, 3) epistemic uncertainty elimination with PF algorithm and parameter iteration, respectively.

First, data-driven models, whether distribution-based [8], sampling-based [37] or Bayesian model-based [38], generate data essentially by exploring the external data laws without interpretation. In the paper, the degradation mechanism is characterized by differential equations with SSM modeling. The model parameters are updated with actual monitoring data, realizing data augmentation integrated mechanisms and data. The dynamic update of the model describes the degradation process during the monitoring cycle. The identical degradation process is instead of various time-varying intervals for the consistent description of the real state degradation.

Second, considering small-sample conditions, likelihood probabilities are obtained using evidential variables in parameter updates, unlike

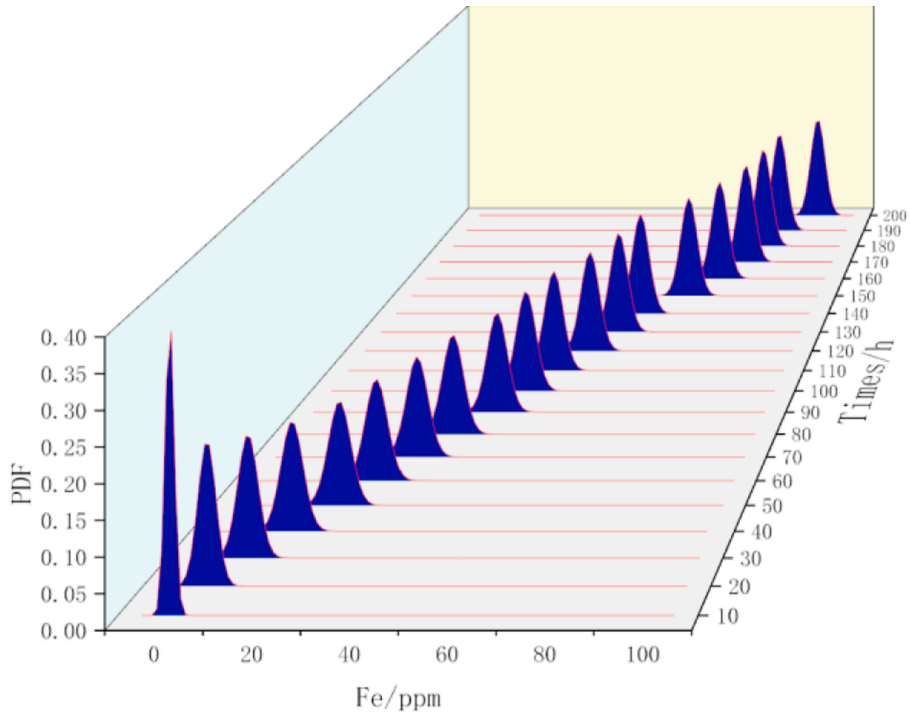


Fig. 11. Illustration of the probability distribution evolution.

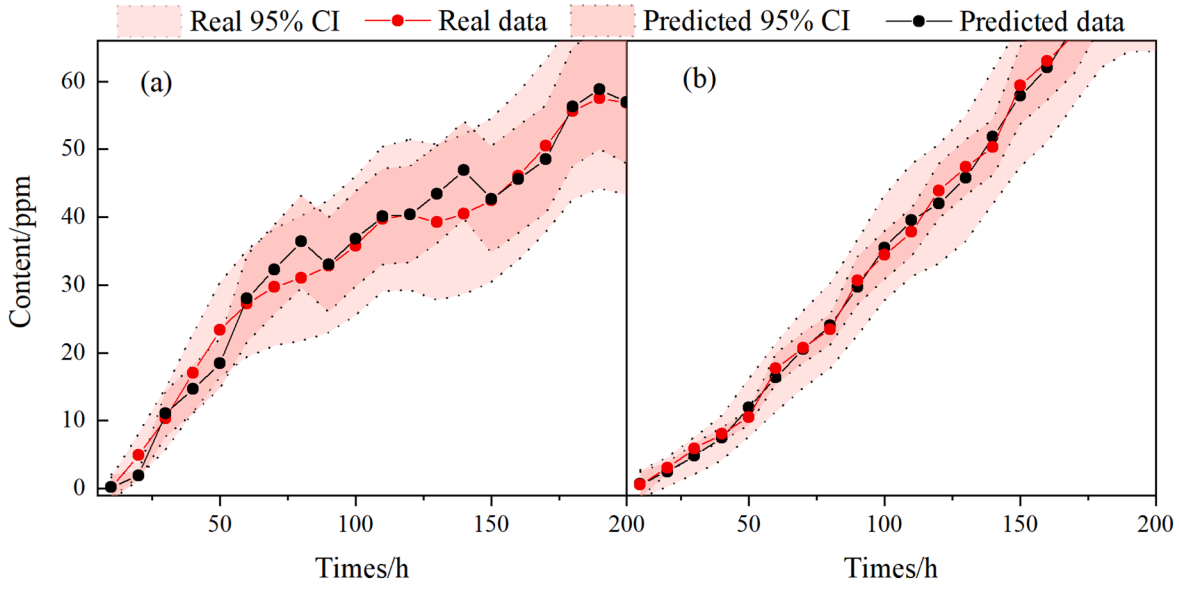


Fig. 12. Comparison of 95% confidence intervals, (a) CAR=0.8636, (b) CAR=0.9574.

the unreliable estimation with random variables in traditional methods [36,39]. Using random variables to describe the model parameter, a complete probability density is needed. Thus the reliable results are only effective for enough samples. With the D-S evidence variable approach, only the basic probability assignments (BPAs) for parameter intervals need to be estimated. Furthermore, a two-step data augmentation strategy is proposed to improve the prediction performance.

Third, two methods of eliminating epistemic uncertainty are described. One is the application of PF [27]. To describe the prediction performance of the sampling with PF, the prediction PDF is visually illustrated in Fig. 11, where the expectation of distribution represents the monitoring data. The results demonstrate the effectiveness of the improved model in reducing random errors of the stochastic process.

The other method is the update of variances with prior knowledge. It can be observed that the variances of the diffusion and drift factors in the Wiener process gradually decrease, accompanying the EM parameter iterations with the prior parameters [36]. We can transform Eq. (15) into

$$\widehat{\sigma}_0^2 = \frac{\sigma_0^2\sigma_A^2}{\sigma_0^2+\sigma_A^2} = \frac{\sigma_0^2}{\sigma_0^2/\sigma_A^2+1} = \frac{\sigma_A^2}{\sigma_A^2/\sigma_0^2+1}, \tag{25}$$

where either the variance σ_0^2 of the prior parameter or the variance σ_A^2 of the measured parameter is greater than the variance $\widehat{\sigma}_0^2$ of the updated parameter. The variance decreases with increasing iterations, which is the primary contributor to uncertainty reduction. The 95% CI for parameter estimation with real and augmented data is shown in Fig. 12.

It can be seen that the 95% CI of the predicted data gradually converges, and the predictions present increasing reliability.

Overall, there are still some limitations to the proposed method:

1) The updating PDF is obtained through the iteration of the analytic solution, which cannot be effectively solved due to complex integration. To match the real non-linear degradation, sampling and filtering methods, which obtain the solution by approximation, are more generalizable for the application of the proposed method.
2) It is difficult to guarantee the effectiveness of a signal indicator in OCM. As a symbolic indicator, the Fe element was chosen to characterize the wear state in this paper. However, it was not guaranteed to be a comprehensive representation of the oil state. Therefore we aim to eliminate cognitive uncertainty by fusing multiple indicators for reliability prediction.

7. Conclusion

To improve the reliability and accuracy of the prediction in OCM, we have proposed a data augmentation model in which both the physical degradation and data probability are considered with the monitoring data. Primarily, the model is constructed by incorporating the Winner process into the mechanism equations. A two-stage strategy is established to update the model parameters for the model training and the data augmentation, respectively. The performance of the model is validated with the real small samples from engine bench tests. The main conclusions are as follows.

1) A knowledge-based data augmentation of small samples for oil condition prediction is proposed, in which both the certain oil degradation and the uncertain data distribution are jointly concerned.
2) A two-stage strategy is established for the model parameter update by integrating interval estimation with small samples and point estimation with augmented data. The updated parameters provide an accurate prediction with the PF method.
3) The proposed two-stage data augmentation shows reliable and high accuracy in prediction with both sparse and truncated data.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors appreciate the financial support from the National Natural Science Foundation of China (No. 51975455) and the support of Xi'an JingHui Information Technology Co., Ltd.

References

[1] Pan Y, Jing Y, Wu T, Kong X. An integrated data and knowledge model addressing aleatory and epistemic uncertainty for oil condition monitoring. Reliab Eng Syst Saf 2021;210:107546. https://doi.org/10.1016/j.ress.2021.107546.
[2] Du Y, Wu T, Zhou S, Makis V. Remaining useful life prediction of lubricating oil with dynamic principal component analysis and proportional hazards model. Proc Inst Mech Eng, Part J 2020;234:964–71. https://doi.org/10.1177/1350650119874560.
[3] Morgan I, Liu H. Computational analysis of sparse datasets for fault diagnosis in large tribological mechanisms. IEEE Trans Syst, Man, Cybern C 2011;41:617–29. https://doi.org/10.1109/TSMCC.2010.2073703.
[4] Jiang C, Zhang Z, Han X, Liu J. A novel evidence-theory-based reliability analysis method for structures with epistemic uncertainty. Comput Struct 2013;129:1–12. https://doi.org/10.1016/j.compstruc.2013.08.007.
[5] Colwell RK, Chao A, Gotelli NJ, Lin S-Y, Mao CX, Chazdon RL, et al. Models and estimators linking individual-based and sample-based rarefaction, extrapolation and comparison of assemblages. J Plant Ecol 2012;5:3–21. https://doi.org/10.1093/jpe/rtr044.
[6] Greenland S, Schwartzbaum JA, Finkle WD. Problems due to small samples and sparse data in conditional logistic regression analysis. Am. J. Epidemiol. 2000;151:531–9. https://doi.org/10.1093/oxfordjournals.aje.a010240.
[7] Lee G, Scott C. EM algorithms for multivariate Gaussian mixture models with truncated and censored data. Comput Stat Data Anal 2012;56:2816–29. https://doi.org/10.1016/j.csda.2012.03.003.
[8] Li D-C, Lin W-K, Chen C-C, Chen H-Y, Lin L-S. Rebuilding sample distributions for small dataset learning. Decis Support Syst 2018;105:66–76. https://doi.org/10.1016/j.dss.2017.10.013.
[9] Doucet A, Godsill S, Andrieu C. On sequential Monte Carlo sampling methods for Bayesian filtering. Stat Comput 2000;10:197–208. https://doi.org/10.1023/A:1008935410038.
[10] Kang G, Wu L, Guan Y, Peng Z. A virtual sample generation method based on differential evolution algorithm for overall trend of small sample data: used for lithium-ion battery capacity degradation data. IEEE Access 2019;7:123255–67. https://doi.org/10.1109/ACCESS.2019.2937550.
[11] Demir-Kavuk O, Kamada M, Akutsu T, Knapp E-W. Prediction using step-wise L1, L2 regularization and feature selection for small data sets with large number of features. BMC Bioinformatics 2011;12:412. https://doi.org/10.1186/1471-2105-12-412.
[12] Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 1997;53:983–97. https://doi.org/10.2307/2533558.
[13] Li D-C, Lin L-S, Chen C-C, Yu W-H. Using virtual samples to improve learning performance for small datasets with multimodal distributions. Soft Comput 2019;23:11883–900. https://doi.org/10.1007/s00500-018-03744-z.
[14] Colwell RK, Mao CX, Chang J. Interpolating, extrapolating, and comparing incidence-based species accumulation curves. Ecology 2004;85:2717–27. https://doi.org/10.1890/03-0557.
[15] He Y-L, Wang P-J, Zhang M-Q, Zhu Q-X, Xu Y. A novel and effective nonlinear interpolation virtual sample generation method for enhancing energy prediction and analysis on small data problem: a case study of Ethylene industry. Energy 2018;147:418–27. https://doi.org/10.1016/j.energy.2018.01.059.
[16] Li D-C, Chen C-C, Chang C-J, Lin W-K. A tree-based-trend-diffusion prediction procedure for small sample sets in the early stages of manufacturing systems. Expert Syst Appl 2012;39:1575–81. https://doi.org/10.1016/j.eswa.2011.08.071.
[17] Imai K, van Dyk DA. A Bayesian analysis of the multinomial probit model using marginal data augmentation. J Econom 2005;124:311–34. https://doi.org/10.1016/j.jeconom.2004.02.002.
[18] Lüdtke O, Ulitzsch E, Robitzsch A. A Comparison of Penalized Maximum Likelihood Estimation and Markov Chain Monte Carlo Techniques for Estimating Confirmatory Factor Analysis Models With Small Sample Sizes. Front Psychol 2021;12:615162. https://doi.org/10.3389/fpsyg.2021.615162.
[19] Yu K, Lin TR, Ma H, Li X, Li X. A multi-stage semi-supervised learning approach for intelligent fault diagnosis of rolling bearing using data augmentation and metric learning. Mech Syst Signal Process 2021;146:107043. https://doi.org/10.1016/j.ymssp.2020.107043.
[20] Yoo Y, Jung U-J, Han YH, Lee J. Data Augmentation-Based Prediction of System Level Performance under Model and Parameter Uncertainties: role of Designable Generative Adversarial Networks (DGAN). Reliab Eng Syst Saf 2021;206:107316. https://doi.org/10.1016/j.ress.2020.107316.
[21] Liu D, Wang S, Zhang C, Tomovic M. Bayesian model averaging based reliability analysis method for monotonic degradation dataset based on inverse Gaussian process and Gamma process. Reliab Eng Syst Saf 2018;180:25–38. https://doi.org/10.1016/j.ress.2018.06.019.
[22] Liu D, Wang S. A degradation modeling and reliability estimation method based on Wiener process and evidential variable. Reliab Eng Syst Saf 2020;202:106957. https://doi.org/10.1016/j.ress.2020.106957.
[23] Zhao D, Gao C, Zhou Z, Liu S, Chen B, Gao J. Fatigue life prediction of the wire rope based on grey theory under small sample condition. Eng Fail Anal 2020;107:104237. https://doi.org/10.1016/j.engfailanal.2019.104237.
[24] Liu D, Wang S, Tomovic MM, Zhang C. An evidence theory based model fusion method for degradation modeling and statistical analysis. Inf Sci (Ny) 2020;532:33–60. https://doi.org/10.1016/j.ins.2020.04.042.
[25] Belhaj Salem M, Fouladirad M, Deloux E. Prognostic and Classification of Dynamic Degradation in a Mechanical System Using Variance Gamma Process. Mathematics 2021;9:254. https://doi.org/10.3390/math9030254.
[26] Blitzstein J, Diaconis P. A Sequential Importance Sampling Algorithm for Generating Random Graphs with Prescribed Degrees. Internet Math 2011;6:1494. https://doi.org/10.1080/15427951.2010.557277.
[27] Lei Y, Li N, Lin J. A New Method Based on Stochastic Process Models for Machine Remaining Useful Life Prediction. IEEE Trans Instrum Meas 2016;65:2671–84. https://doi.org/10.1109/TIM.2016.2601004.
[28] Vališ D, Žák L, Pokora O, Lánský P. Perspective analysis outcomes of selected tribodiagnostic data used as input for condition based maintenance. Reliab Eng Syst Saf 2016;145:231–42. https://doi.org/10.1016/j.ress.2015.07.026.
[29] Pu W, Zhang Q, Zhang W, Ren S, Chen Z, Tian T. Flash temperature and anti-wear tribofilm growth mechanisms by asperity contact in top-ring/liner conjunction of IC engines. Tribol Int 2020;146:106186. https://doi.org/10.1016/j.triboint.2020.106186.
[30] Sharma GK, Chawla OP. Modelling of lubricant oil alkalinity in diesel engines. Tribol Int 1988;21:269–74. https://doi.org/10.1016/0301-679X(88)90004-7.
[31] (Lei)(Johan) Yang L, Padding JT, Kuipers JAM. Modification of kinetic theory of granular flow for frictional spheres, Part I: two-fluid model derivation and numerical implementation. Chem Eng Sci 2016;152:767–82. https://doi.org/10.1016/j.ces.2016.05.031.

[32] Fan B, Li B, Feng S, Mao J, Xie Y-B. Modeling and experimental investigations on the relationship between wear debris concentration and wear rate in lubrication systems. Tribol Int 2017;109:114–23. https://doi.org/10.1016/j.triboint.2016.12.015.
[33] GB/T 19055-2003 Reliability test methods for motor vehicle engines n.d. http://www.nssi.org.cn/nssi/front/6230403.html (accessed June 30, 2021).
[34] Chang S-C, Lai H-C, Yu H-C. A variable P value rolling Grey forecasting model for Taiwan semiconductor industry production. Technol Forecast Soc Change 2005;72:623–40. https://doi.org/10.1016/j.techfore.2003.09.002.
[35] Bergtold JS, Yeager EA, Featherstone AM. Inferences from logistic regression models in the presence of small samples, rare events, nonlinearity, and multicollinearity with observational data. J Appl Stat 2018;45:528–46. https://doi.org/10.1080/02664763.2017.1282441.
[36] Si X-S, Wang W, Chen M-Y, Hu C-H, Zhou D-H. A degradation path-dependent approach for remaining useful life estimation with an exact and closed-form solution. Eur J Oper Res 2013;226:53–66. https://doi.org/10.1016/j.ejor.2012.10.030.
[37] Starling JK, Mastrangelo C, Choe Y. Improving Weibull distribution estimation for generalized Type I censored data using modified SMOTE. Reliab Eng Syst Saf 2021;211:107505. https://doi.org/10.1016/j.ress.2021.107505.
[38] Beldjoudi C, Kernane T, Maroufy HE. Bayesian inference for a susceptible-exposed-infected-recovered epidemic model with data augmentation. Math Popul Stud 2020;27:232–58. https://doi.org/10.1080/08898480.2019.1656491.
[39] Jin G, Matthews DE, Zhou Z. A Bayesian framework for on-line degradation assessment and residual life prediction of secondary batteries inspacecraft. Reliab Eng Syst Saf 2013;113:7–20. https://doi.org/10.1016/j.ress.2012.12.011.