



Remaining useful life prediction based on a multi-sensor data fusion model



Naipeng Li ^{a,b}, Nagi Gebraeel ^c, Yaguo Lei ^{a,*}, Xiaolei Fang ^d, Xiao Cai ^a, Tao Yan ^a

^a Key Laboratory of Education Ministry for Modern Design and Rotor-Bearing System, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China

^b Shanxi Key Laboratory of Advanced Manufacturing Technology, North University of China, Taiyuan, Shanxi 030051, China

^c H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, 765 Ferst Drive, Atlanta, GA 30332, USA

^d Edward P. Fitts Department of Industrial and Systems Engineering, North Carolina State University, Raleigh, NC 27695, USA

ARTICLE INFO

Keywords:

Prognostic degradation modeling
Remaining useful life prediction
Big data
Multi-sensor fusion
State-space model

ABSTRACT

With the rapid development of Industrial Internet of Things, more and more sensors have been used for condition monitoring and prognostics of industrial systems. Big data collected from sensor networks bring abundant information resources as well as technical challenges for remaining useful life (RUL) prediction. The major technical challenges include how to select informative sensors and fuse multi-sensor data to improve the prediction performance. To deal with the challenges, this paper proposes a RUL prediction method based on a multi-sensor data fusion model. In this method, the inherent degradation process of the system state is expressed using a state transition function following a Wiener process. Multi-sensor signals are explicated as various proxies of the inherent system degradation process using a multivariate measurement function. The system state is estimated by fusing multi-sensor signals using particle filtering. Informative sensors are selected by a prioritized sensor group selection algorithm. This algorithm first prioritizes sensors according to their individual performances in RUL prediction, and then selects an optimal sensor group based on their combined performances. The effectiveness of the proposed method is demonstrated using a simulation study and aircraft engine degradation data from NASA repository.

1. Introduction

Prognostic degradation modeling utilizes degradation signals to predict the remaining useful life (RUL) of industrial systems, which is later used to guide maintenance-related decisions [1]. A degradation signal can be derived from raw sensor signals and sometimes involves some signal transformations [2]. Majority of the prognostic degradation modeling literature has focused on modeling a single degradation signal [3,4]. Different types of modeling approaches have been used in this area, ranging from stochastic process models [5,6] to artificial intelligence techniques [7]. With the rapid development of Industrial Internet of Things, more and more sensors have been used for condition monitoring of industrial systems. Big data collected from multi-sensor networks bring abundant information resources for RUL prediction. However, they also bring technical challenges for RUL prediction, because they require to conduct an information fusion process among different sensor signals. Most existing studies regarding degradation modeling with multi-sensor fusion aimed to derive a single composite/aggregate health index (HI) [8–12], and modeled the aggregated HI as a single degradation signal using the approaches mentioned above.

Although many models have shown promise when tested on specific applications, they often suffer from the lack of generality and scalability. Different from many conventional models, this paper developed a modeling framework for multi-sensor applications that extends the current literature along two dimensions. First, the proposed model provides a generalizable approach that explicitly distinguishes the underlying health state and the observed multi-sensor degradation signals. Second, our model leverages the underlying correlation structures driving how individual signals evolve relative to each other.

Literature focusing on predicting RUL with multi-sensor fusion can be classified into three main categories according to their fusion levels: data-level fusion, feature-level fusion, and decision-level fusion [13]. Data-level fusion usually targets raw data. It involves the fusion of raw sensor signals into an aggregate HI or their combination through a multi-input prognostic model. For example, Liu et al. [9] proposed a signal-to-noise ratio metric to guide HI construction with multi-sensor fusion. Fang et al. [14] identified informative sensors via penalized location-scale regression and predicted RUL by fusing the informative sensor signals using multivariate functional principle component analysis (MFPCA). The second fusion strategy, feature-level fusion, revolves

* Corresponding author.

E-mail address: yaguo@xjtu.edu.cn (Y. Lei).

around fusing the features extracted from raw sensor signals. This usually involves choosing appropriate feature extraction tools such as statistical signal processing and dimensionality reduction. For example, Tse et al. [10] fused nine statistical features extracted from vibration signals into a HI using principal component analysis (PCA). Lei et al. [11] assigned different weights to multiple features and fused them into a HI using a self-organizing map neural network. Chen et al. [15] used kernel PCA to extract nonlinear features from signals, and applied a recurrent neural network to fuse the features and predict RUL. The third strategy is called decision-level fusion. As the name implies, this strategy inputs sensor signals into different models to generate decision results, and then fuses different decisions together. Sun et al. [16] linked the univariate reliability functions of different Wiener process models to form a joint distribution using a multivariate copula function. Liu et al. [17] integrated the prediction results of multiple support vector regression models by introducing dynamic weights for each sub-model.

A large component of the above works assume that system failure is completely characterized by a feature or an aggregate HI [8-13,16,17] derived from raw signals. In some remaining literature [14,15,18], RUL prediction is conducted by mapping the relationship between raw signals and lifetimes via regression models or artificial intelligent techniques. A common assumption in most of the literature mentioned herein is that the health state of the system is exactly equivalent to the degradation signals or their derived HI. Arguably, this is an oversimplification because the health state of an industrial system is generally unobservable, and the signals are just sensor responses of the system which may involve partial or even no information of the system state.

This paper focuses on a general industrial problem, i.e., how to predict RUL of a complex industrial system based on its multi-sensor degradation signals without any information about its physical health state. To deal with this problem, we develop a generalizable modeling framework based on the basic idea in Fig. 1. Multi-sensor signals are treated as different proxies of the ground truth health state. Multi-sensor fusion is conducted based on a multi-sensor data fusion model (MSDFM), wherein system states and multi-sensor signals are analyzed separately. A state transition function is used to describe how the hidden system state evolves over time. Failure is defined as the system state crossing a failure threshold, instead of the amplitude of the observed signal. A multivariate measurement function is used to model the degradation behavior of sensor signals as a function of the system state. The measurement function involves the correlations between different sensor signals as well as their different trends.

In real applications, only a subset of sensors are sufficient to describe the degradation process. Other sensors may be uninformative or present redundant information for the degradation process. Therefore, informative sensor selection is also a very beneficial procedure for RUL prediction with multi-sensor fusion. One common strategy in sensor selection is to inspect degradation trends of sensor signals visually [9]. Another technique is group penalization [14], whereby a penalty term is added into a criterion to shrink the number of sensors for RUL prediction. The relationship between the number of sensors and the

uncertainty of RUL was studied in [19], and a suitable number was determined via tradeoff between accuracy and cost. A more straightforward strategy is to try all possible sensor combinations and select the optimal one that provides the most accurate RUL prediction result [20]. Different from existing techniques, we propose a new sensor selection approach referred to as prioritized sensor group selection (PSGS) that leverages a historical training dataset to select the most informative sensors. This approach first prioritizes the sensors according to their individual prediction accuracy. Then, the most informative sensors are selected by analyzing the prediction accuracy of different sensor groups. The major contributions of this paper can be summarized as follows:

- 1) We develop a generalizable state-space model referred to as MSDFM for multi-sensor fusion prognostics. Different from commonly used degradation models, this model distinguishes the degradation behavior of the system state and the measurement behavior of multi-sensor signals. RUL is defined based on the degradation process of virtual states instead of signals. This model gives a more reasonable interpretation about the failure of systems than degradation models completely dependent on signals.
- 2) We develop a sensor selection algorithm referred to as PSGS. The major novelty of this algorithm is that it prioritizes sensors according to their individual prediction performance. Sensors with higher prediction accuracy are input into the RUL prediction process preferentially. This strategy reduces the amount of combinations by specifying a priority order and ensures the efficiency of each selected sensor by evaluating their individual performance.

The remainder of this paper is organized as follows. Section 2 introduces the framework of conventional RUL prediction methods using state-space models. Section 3 presents our proposed MSDFM-based RUL prediction method. In Section 4, a simulation study is conducted to demonstrate the effectiveness of the proposed method. Section 5 evaluates the proposed method using a degradation dataset of aircraft engines available from NASA repository. Some conclusions are drawn in Section 6.

2. RUL prediction using state-space models

State-space models have been used for RUL prediction in cases of a single sensor applications [21,22]. They generally include three major steps: model development, system state estimation, and RUL prediction.

2.1. Model development

In the context of state-space modeling, the inherent degradation process of the system state is described using a state transition function, and the observed degradation signal is expressed using a measurement function. The general expression of a state-space model is as follows:

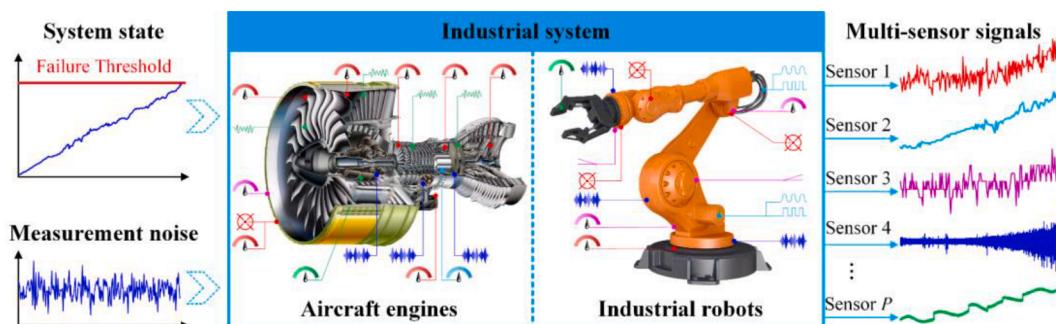


Fig. 1. Basic idea of the modeling framework.

$$\begin{cases} x_k &= f(x_{k-1}, \omega_{k-1}) \\ y_k &= h(x_k, v_k) \end{cases} \quad (1)$$

where $f(\cdot)$ and $h(\cdot)$ are the state transition function and measurement function, respectively; x_k and y_k represent the system state and the degradation signal at time t_k , respectively. ω_{k-1} and v_k denote the state transition noise and the measurement noise, respectively.

2.2. System state estimation

The state estimation process is performed using a two-step Bayesian framework [23].

- 1) **Prediction:** Predict the prior probability of the current state via the Chapman-Kolmogorov equation:

$$p(x_k|y_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|y_{1:k-1})dx_{k-1} \quad (2)$$

where $p(x_k|x_{k-1})$ can be acquired from the state transition function in (1), and $y_{1:k-1}$ denotes the historical monitoring signals up to t_{k-1} .

- 1) **Update:** Obtain the posterior probability of the current state by updating the prior probability.

$$p(x_k|y_{1:k}) = \frac{p(y_k|x_k)p(x_k|y_{1:k-1})}{p(y_k|y_{1:k-1})} \quad (3)$$

where $p(y_k|x_k)$ can be acquired from the measurement function in (1), and

$$p(y_k|y_{1:k-1}) = \int p(y_k|x_k)p(x_k|y_{1:k-1})dx_k \quad (4)$$

The steps above result in the optimal Bayesian solution that is difficult to determine analytically. To estimate the system state approximately, a family of filtering algorithms including Kalman filtering (KF), extended Kalman filtering (EKF), and PF [24] have been developed. Particle filtering is considered the de facto approach for

nonlinear/non-Gaussian problems, which is the case considered in this paper.

2.3. RUL prediction

RUL amounts to the time it takes the system state to cross a pre-specified failure threshold from its current state.

$$L = \inf\{l : x(l+t_k) \geq D|x_k\} \quad (5)$$

where L denotes the RUL at t_k , $\inf\{\cdot\}$ is a function denoting the inferior limit of a variable, $x(l+t_k)$ is the system state at $l+t_k$, and D is the failure threshold.

Analytical expression of the RUL distribution can be obtained in cases where the state transition function follows some specific functional forms or processes [25,26]. In this study, we choose a widely used function in degradation modeling, i.e., the linear Wiener process, whose RUL follows an inverse Gaussian distribution.

3. RUL prediction using the MSDFM

General problem statement: how to predict RUL of a complex industrial system based on its multi-sensor degradation signals without any information about its physical health state.

To deal with this problem, a MSDFM-based RUL prediction method is developed in this section. Fig. 2 outlines the flowchart of the proposed method. At first, a MSDFM is developed to describe the degradation process of the inherent system state and its relationship with multi-sensor signals. After that, model parameters are estimated according to offline training data. Then, a sensor selection algorithm named PSGS is proposed to select informative sensors. Finally, the system state and degradation rate are estimated jointly using a PF algorithm [27], and RUL is predicted based on the estimated results. Each step is explained as follows.

3.1. Model development

For industrial systems whose state of health is unobservable, it is impossible to get the actual degradation process of states. What seems certain is that a system generally operates starting from a healthy stage and degrades over time until it reaches a completely failed stage. RUL

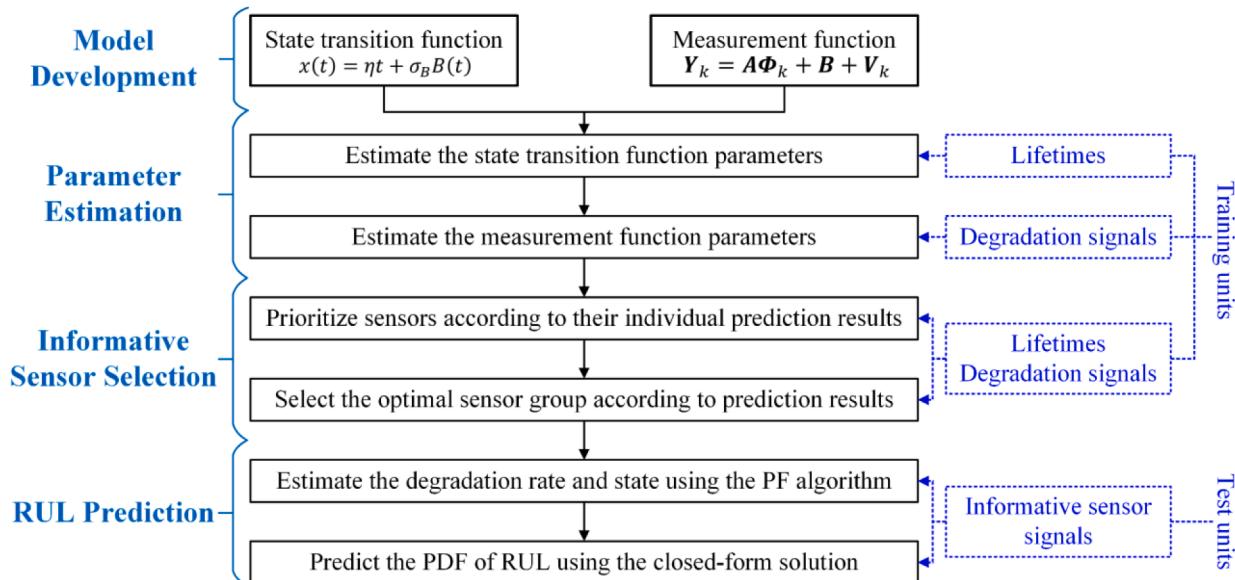


Fig. 2. Flowchart of the proposed MSDFM-based RUL prediction method. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

decreases linearly along with the operation time. Inspired by this phenomenon, we define a virtual state as a linear Wiener process increasing from 0 to 1, where 0 corresponds to a baseline “as-good-as-new” healthy stage, and 1 corresponds to a completely failed stage. The failure threshold is correspondingly defined as $D = 1$. The degradation process is formulated as follows:

$$x(t) = \eta t + \sigma_B B(t) \quad (6)$$

where $x(t)$ is the system state whose initial value is 0, η is the degradation rate that is denoted as a random variable following normal distribution $N(\mu_\eta, \sigma_\eta^2)$. $B(t)$ is a standard Browning motion, and σ_B is the standard deviation of the diffusion term.

The sensor number is represented as P . The measurement values of the P sensors at t_k , i.e., $Y_k = (y_k^1, y_k^2, \dots, y_k^P)$, can be described using a multivariate measurement function:

$$\begin{bmatrix} y_k^1 \\ y_k^2 \\ \vdots \\ y_k^P \end{bmatrix} = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & a_P \end{bmatrix} \begin{bmatrix} \varphi(x_k, \theta_1) \\ \varphi(x_k, \theta_2) \\ \vdots \\ \varphi(x_k, \theta_P) \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_P \end{bmatrix} + \begin{bmatrix} v_k^1 \\ v_k^2 \\ \vdots \\ v_k^P \end{bmatrix} \quad (7)$$

where x_k is the state value at t_k . A general expression $\varphi(\cdot)$ is used to describe the mapping relationship between states and sensor signals. Parameter vector θ_p denotes a set of parameter values in function $\varphi(\cdot)$ corresponding to sensor p . a_p and b_p are the scale and location parameters of sensor p , which control the scale range and the intercept of the sensor signal, respectively. v_k^p is the measurement noise following a normal distribution $N(0, \sigma_p^2)$. To make sure the validity of our methodology, $\varphi(\cdot)$ must satisfy the following three conditions:

- 1) $\varphi(x)$ is a strictly monotonic (increasing or decreasing) function.
- 2) For each sensor, model parameters in vector θ_p are deterministic values.
- 3) There always exists a positive number M , which satisfies $|\varphi(x)| \leq M$, $x \in [0, 1]$.

Condition 1 and 2 ensure the one-to-one relationship between the state value x_k and its response $\varphi(x_k, \theta_p)$. Condition 3 ensures that $\varphi(x_k)$ is a limited function within interval $[0, 1]$. All the three conditions make sure that $y_k^p = \varphi(x_k, \theta_p)$ has an inverse function $x_k = \psi(y_k^p, \theta_p)$. It is easy to construct a function satisfying the above three conditions, such as the polynomial function x^c , ($c > 0$), the exponential function $\exp(cx)$, ($c \neq 0$), the logarithmic function $\ln(c(x+d))$, ($c \neq 0, d > 0$), the trigonometric function $\sin(cx)$, ($0 < c \leq \pi/2$), and their combinations.

Since degradation signals of different sensors are various in generation mechanism, transmission paths, fault sensitivities, etc., their signal curves generally present different degradation behaviors. The variation of the degradation behaviors can be formulated by choosing different functions $\varphi(\cdot)$ or assigning different parameters a_p , b_p , and θ_p . In addition, since different sensors monitor the same system degradation, they may suffer similar interference from the outside environment. Therefore, they also present high correlation with each other, especially in their measurement noises. To describe the correlation between different sensors, the noise vector $V_k = (v_k^1, v_k^2, \dots, v_k^P)$ is formulated as an independent identical P -dimensional multivariate normal distribution $N(0, \Sigma)$ with

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{2,1} & \cdots & \sigma_{P,1} \\ \sigma_{1,2} & \sigma_2^2 & \cdots & \sigma_{P,2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1,P} & \sigma_{2,P} & \cdots & \sigma_P^2 \end{bmatrix} \quad (8)$$

where σ_p^2 ($p = 1, 2, \dots, P$) is the variance of the measurement noise of sensor p , σ_{ij} ($i \neq j$) is the covariance of the measurement noises between sensor i and j .

Let $\Phi_k = (\varphi(x_k, \theta_1), \varphi(x_k, \theta_2), \dots, \varphi(x_k, \theta_P))'$, $A = \text{diag}(a_1, a_2, \dots, a_P)$ which is a positive diagonal matrix with elements a_1, a_2, \dots, a_P , $B = (b_1, b_2, \dots, b_P)'$. The following MSDFM is developed.

$$\begin{cases} x_k = x_{k-1} + \eta \Delta t_{k-1} + \omega_{k-1} \\ Y_k = A\Phi_k + B + V_k \end{cases} \quad (9)$$

where the state transition function is a discrete difference expression of (6), $\Delta t_k = t_k - t_{k-1}$ is the time interval, ω_{k-1} is a random sampling from $N(0, \sigma_B^2 \Delta t_k)$, and the measurement function is simplified from (7).

3.2. Parameter estimation

Assume that we have obtained the degradation signals of N training units. For unit n , the measurements of each sensor are recorded at the same time sequence $T_n = (t_{n,1}, t_{n,2}, \dots, t_{n,K_n})'$, where K_n denotes the total number of samples and t_{n,K_n} is the last sampling time, i.e., the failure time of unit n . The sampling sequence of sensor p is indicated as $(y_{n,1}^p, y_{n,2}^p, \dots, y_{n,K_n}^p)'$, where $y_{n,k}^p$ denotes the observation of sensor p unit n at time $t_{n,k}$. To avoid the over-parameterization problem, we develop an estimation methodology that reduce the number of independent variables of each criterion as far as possible.

In the state transition function, the unknown parameters include μ_η , σ_η^2 , and σ_B^2 . These parameters can be estimated using maximum likelihood estimation (MLE). The lifetime probability density function (PDF) follows the inverse Gaussian distribution [25]:

$$f(t) = \frac{D}{\sqrt{2\pi t^2(\sigma_\eta^2 t^2 + \sigma_B^2 t)}} \exp\left(-\frac{(t\mu_\eta - D)^2}{2(\sigma_\eta^2 t^2 + \sigma_B^2 t)}\right) \quad (10)$$

We have known the lifetimes of N training units $T = (t_{1,K_1}, t_{2,K_2}, \dots, t_{N,K_n})$. Let $\tilde{\sigma}_B^2 = \sigma_B^2/\sigma_\eta^2$. Parameter $\tilde{\sigma}_B^2$ is estimated by maximizing the following criterion.

$$\begin{aligned} \mathcal{L}(\tilde{\sigma}_B^2 | T) &= -\sum_{n=1}^N \ln(t_{n,K_n}) - \frac{1}{2} \sum_{n=1}^N \ln(t_{n,K_n}^2 + \tilde{\sigma}_B^2 t_{n,K_n}) \\ &\quad - \frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma_\eta^2(\tilde{\sigma}_B^2)) - \frac{N}{2} \end{aligned} \quad (11)$$

The MLEs of μ_η and σ_η^2 are calculated by inputting the estimation result $\tilde{\sigma}_B^2$ into

$$\hat{\mu}_\eta(\tilde{\sigma}_B^2) = \sum_{n=1}^N \frac{D}{t_{n,K_n} + \tilde{\sigma}_B^2} \Bigg/ \sum_{n=1}^N \frac{t_{n,K_n}}{t_{n,K_n} + \tilde{\sigma}_B^2} \quad (12)$$

$$\hat{\sigma}_\eta^2(\tilde{\sigma}_B^2) = \frac{1}{N} \sum_{n=1}^N \frac{(\hat{\mu}_\eta(\tilde{\sigma}_B^2) t_{n,K_n} - D)^2}{t_{n,K_n}^2 + \tilde{\sigma}_B^2 t_{n,K_n}} \quad (13)$$

The estimation result of σ_B^2 is calculated by $\hat{\sigma}_B^2 = \hat{\sigma}_\eta^2 \tilde{\sigma}_B^2$. The details of the derivation are given in Appendix A.

As is shown in (7), there are P different sensors in the measurement function. Each sensor performs different degradation behavior. Mean-

while, outside interference from the same environment causes correlated measurement noises among different sensors. The unique degradation behavior of each sensor is corresponding to its own parameters a_p , b_p , and θ_p . The correlation between different sensors are determined by the covariance matrix Σ . Parameters of the measurement function are estimated following two steps. In the first step, each sensor signal is analyzed separately to estimate the model parameters a_p , b_p , and θ_p . In the second step, all sensor signals are analyzed jointly to estimate the covariance matrix Σ .

Let us first focus on an individual sensor signal, the noise v_k^p can be reduced by smoothing signals using a local regression method [28]. The smoothed signal is approximated as

$$\tilde{y}_{n,k}^p = a_p \varphi(x_{n,k}, \theta_p) + b_p \quad (14)$$

where $x_{n,k}$ is the state value of unit n at t_k .

Based on our definition to the system state and measurement function, the parameter vector θ_p can be estimated using the following criterion.

$$\min_{\theta_p} \sum_{n=1}^N \left(\frac{1}{K_n} \sum_{k=1}^{K_n} \left(\psi \left(\frac{\tilde{y}_{n,k}^p - b_p(\theta_p)}{a_p(\theta_p)}, \theta_p \right) - \mu_{\eta} t_{n,k} \right)^2 \right) \quad (15)$$

Parameter a_p and b_p are calculated by inputting the estimation result of θ_p into the following two equations. The details of derivation are given in Appendix B.

$$\hat{a}_p(\theta_p) = \frac{1}{N} \sum_{n=1}^N \frac{\tilde{y}_{n,K_n}^p - \tilde{y}_{n,1}^p}{\varphi(1, \theta_p) - \varphi(0, \theta_p)} \quad (16)$$

$$\hat{b}_p(\theta_p) = \frac{1}{N} \sum_{n=1}^N \frac{\tilde{y}_{n,1}^p \varphi(1, \theta_p) - \tilde{y}_{n,K_n}^p \varphi(0, \theta_p)}{\varphi(1, \theta_p) - \varphi(0, \theta_p)} \quad (17)$$

Up to now, parameters corresponding to a single sensor have been estimated. The following process is to estimate the covariance matrix of measurement noises. According to (14), the measurement noises can be approximated using $v_{n,k}^p = y_{n,k}^p - \tilde{y}_{n,k}^p$. The measurement noise matrix of unit n is

$$V_n = \begin{bmatrix} v_{n,1}^1 & v_{n,1}^2 & \cdots & v_{n,1}^P \\ v_{n,2}^1 & v_{n,2}^2 & \cdots & v_{n,2}^P \\ \vdots & \vdots & \ddots & \vdots \\ v_{n,K_n}^1 & v_{n,K_n}^2 & \cdots & v_{n,K_n}^P \end{bmatrix} = [V_n^1 \quad V_n^2 \quad \cdots \quad V_n^P] \quad (18)$$

Its covariance matrix is denoted as Σ_n . The estimation result of Σ is calculated as the mean of $\{\Sigma_n, n = 1, 2, \dots, N\}$. Until now, all parameters have been estimated.

3.3. Informative sensor selection

In real applications, some sensors may be uninformative or provide redundant information for the degradation process. These sensors may reduce the accuracy of RUL prediction or increase computational cost. Therefore, it is necessary to evaluate the contribution of each sensor and select informative ones that are expected to provide more accurate prediction results. The task of sensor selection is to find a sensor group with which the prognostic methodology can get the most accuracy result. A straightforward way is to try all possible groups and select the optimal one according to their performance. In theory, P sensors totally generate $2^P - 1$ random groups (including the cases with a single sensor). The number of groups increases exponentially with the sensor number. It will be time consuming to try all possible groups in high-dimensional cases. To reduce the computational cost and guarantee the validity of the selection result, we propose a PSGS algorithm, whose

flowchart is as shown in Fig. 3. This algorithm first prioritizes P sensors according to their individual RUL predictions, and then combines them into P sensor groups following the prioritization. The sensor group that provides the most accurate RUL prediction is finally selected as the optimal one. As shown in Fig. 3, it is the basis to predict RUL using a random possible group in the PSGS flowchart. Thus, we first develop the algorithm of RUL prediction using a random sensor group.

3.3.1. RUL prediction using a random sensor group

Table 1 gives the algorithm of RUL prediction using a random sensor group. Without loss of generality, we assume that Ω is a random subset of the P sensor indexes, which may include a single sensor or several sensors. The PF with a fuzzy resampling algorithm [27] is employed to estimate the degradation rate and system state jointly.

At the initial time, a series of particles $\{\eta_0^i, x_0^i\}_{i=1:N_s}$ are generated involving the degradation rate particle η_0^i and the state particle x_0^i , where N_s is the particle number. The particle weights are initialized as $\{w_0^i = 1/N_s\}_{i=1:N_s}$. The degradation rate particle η_0^i is sampled from the normal distribution $N(\hat{\mu}_\eta, \hat{\sigma}_\eta^2)$, and the state particle is initialized as $x_0^i = 0$. At each time, the state particle is one-step transmitted following the state transition function.

$$x_k^i = x_{k-1}^i + \eta_{k-1}^i \Delta t_k + \omega_{k-1}^i \quad (19)$$

The particle weights are updated [23] according to the observations of the sensor group Ω using

$$w_k^i = \frac{w_{k-1}^i}{\sqrt{(2\pi)^{|\Omega|} |\Sigma_\Omega|}} \times \exp \left(-\frac{\left((Y_{\Omega,n,k} - A_\Omega \Phi_{\Omega,k}^i - B_\Omega) \right)^T \Sigma_\Omega^{-1} (Y_{\Omega,n,k} - A_\Omega \Phi_{\Omega,k}^i - B_\Omega)}{2} \right) \quad (20)$$

where $|\Omega|$ denotes the number of sensors in the group, $Y_{\Omega,n,k}$ is the observations of unit n at $t_{n,k}$ from the group Ω . A_Ω , B_Ω , $\Phi_{\Omega,k}^i$, and Σ_Ω are the submatrices of A , B , Φ_k^i , and Σ respectively, formed by the elements corresponding to the sensors in Ω . For example, if $\Omega = (1, 2, 3)$, then

$$Y_{\Omega,n,k} = (y_{n,k}^1, y_{n,k}^2, y_{n,k}^3)', A_\Omega = \begin{bmatrix} \hat{a}_1 & 0 & 0 \\ 0 & \hat{a}_2 & 0 \\ 0 & 0 & \hat{a}_3 \end{bmatrix}, B_\Omega = (\hat{b}_1, \hat{b}_2, \hat{b}_3)', \Phi_{\Omega,k}^i = (\varphi(x_k^i, \hat{\theta}_1), \varphi(x_k^i, \hat{\theta}_2), \varphi(x_k^i, \hat{\theta}_3))', \text{ and } \Sigma_\Omega = \begin{bmatrix} \hat{\sigma}_1^2 & \hat{\sigma}_{1,2} & \hat{\sigma}_{1,3} \\ \hat{\sigma}_{2,1} & \hat{\sigma}_2^2 & \hat{\sigma}_{2,3} \\ \hat{\sigma}_{3,1} & \hat{\sigma}_{3,2} & \hat{\sigma}_3^2 \end{bmatrix}.$$

The particles are resampled using the fuzzy resampling algorithm [27] and the particle weights are reset to be $w_k^i = 1/N_s$. The resampled particles are denoted as $\{\eta_k^i, x_k^i\}_{i=1:N_s}$. The median of the particles $\hat{\eta}_k$ and \hat{x}_k are employed as the estimation results of the degradation rate and the system state, respectively. The RUL at t_k is the time left before the system state increases from \hat{x}_k to D . It can be treated as the lifetime of a stochastic process increasing with a rate $\hat{\eta}_k$ from 0 towards a threshold $D - \hat{x}_k$. Therefore, based on the lifetime PDF in (10), the analytical solution of the RUL PDF can be expressed as follows:

$$f(l|\hat{x}_k, \hat{\eta}_k) = \frac{\left(D - \hat{x}_k \right)}{\sqrt{2\pi \hat{\sigma}_B^2 l^3}} \exp \left(-\frac{\left(l\hat{\eta}_k - D + \hat{x}_k \right)^2}{2\hat{\sigma}_B^2 l} \right) \quad (21)$$

3.3.2. PSGS algorithm for informative sensor selection

The detailed process of the proposed PSGS algorithm is as follows. At first, each of the P sensor signals is input into the RUL prediction algo-

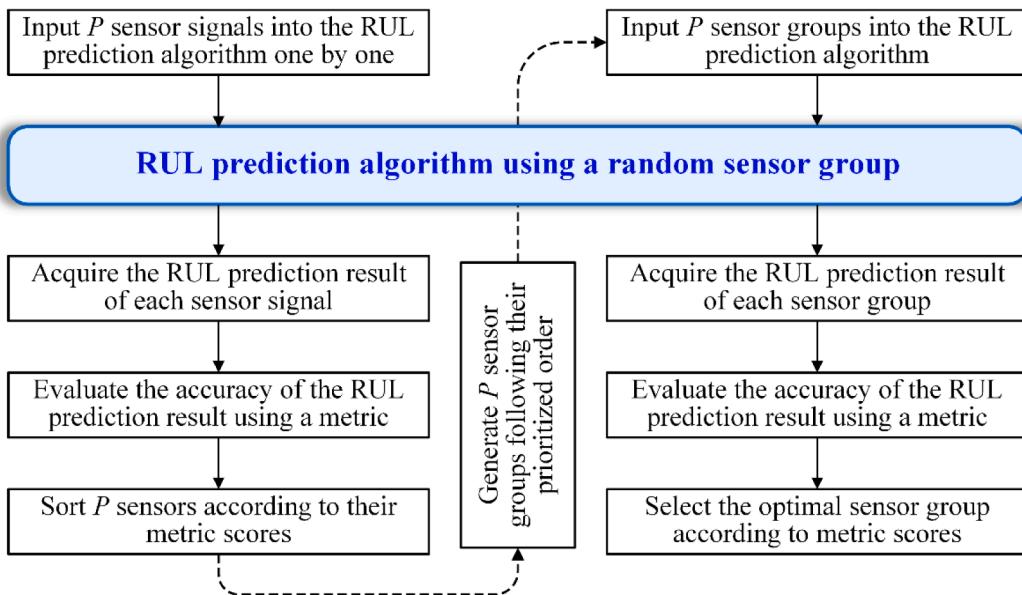


Fig. 3. Flowchart of the PSGS algorithm.

Table 1
RUL prediction using a random sensor group.

Input: Sensor signals Y_n of training unit n ; subset of sensor indexes Ω ; estimated model parameters.

- Generate a series of initial state and degradation rate particles associated with particle weights.
- **For** $k = 1 : K_n$
 - **For** $i = 1 : N_s$
 - Transmit the state particle one-step ahead following (19).
 - Update the particle weight using (20).
 - **End for**
 - Resample particles using the fuzzy resampling algorithm.
 - Normalize particles weights.
 - Calculate the RUL PDF using (21).
- **End for**

Output: RUL prediction result.

rithm in Table 1 to predict the RUL of N training units. Then, the P sensors are prioritized according to their individual performance. Different kinds of metrics have been developed to evaluate the performance of RUL prediction [1]. The weighted absolute relative error (WARE) is employed in this study, which is calculated using the following expression:

$$\text{WARE} = \frac{1}{N} \sum_{n=1}^N \frac{1}{K_n} \left(\sum_{k=1}^{K_n} \frac{t_{n,k} \text{Er}_{n,k}}{t_{n,K_n}} \right) \quad (22)$$

where $\text{Er}_{n,k}$ is the absolute relative error (ARE) of unit n at $t_{n,k}$.

$$\text{Er}_{n,k} = \frac{\left| t_{n,K_n} - \hat{t}_{n,k} - t_{n,k} \right|}{t_{n,K_n}} \times 100\% \quad (23)$$

where t_{n,K_n} is the actual lifetime of unit n , and $\hat{t}_{n,k}$ is the expectation of the predicted RUL at $t_{n,k}$.

A lower WARE score corresponds to a more accurate RUL prediction result, which means that the sensor is more helpful for RUL prediction. Thus, the P sensors are sorted in ascending order following their WARE scores. Sensors in front of the order will be chosen preferentially. Following this rule, we generate P different sensor groups by adding one at each time. In particular, the P sensor groups are $(1), (1, 2), (1, 2, 3), \dots, (1, 2, 3, \dots, P)$, respectively, where the number represents the sensor

index in the prioritized order. After that, each of the P sensor groups is input into the algorithm in Table 1 to predict the RUL of training units again. The WARE scores of the P sensor groups are also calculated using (22). Finally, the sensor group that occupies the lowest WARE score is selected as the optimal one. The selected sensor subset will be employed to predict the RUL of test units.

3.4. RUL prediction

The RUL of a test unit is also predicted using the algorithm in Table 1. We only need to input the informative sensor signals of the test unit into this algorithm. The selected sensor subset is denoted as S . In (20), the observed sensor signals are changed as $Y_{S,n,k}$, and the parameters in the multivariate measurement function become A_S , B_S , $\Phi_{S,k}^i$, and Σ_S , respectively.

4. Simulation case study

4.1. Simulation data generation

In this section, a simulation case study is conducted to demonstrate our proposed method. To illustrate the generality of our method, two degradation datasets are simulated. System states of dataset 1 degrade following a linear degradation process as shown in (9), with prespecified parameters $\eta \sim N(0.05, 1e-6)$, $\sigma_B^2 = 6e-4$. States in dataset 2 follows a nonlinear degradation process:

$$x_k = x_{k-1} + \alpha \beta t_{k-1}^{\beta-1} \Delta t_{k-1} + \omega_{k-1} \quad (24)$$

The model parameters are prespecified as $\alpha \sim N(0.003, 4e-7)$, $\beta = 2$, $\sigma_B^2 = 6e-4$. Each dataset includes 100 degradation trajectories. Each trajectory increases from the initial value 0 and stops after crossing 1. Degradation processes are monitored by 10 sensors. Three among them are informative ones and the remaining seven are non-informative. Degradation signals of 10 sensors are generated by the following steps.

- 1) Generate 10 correlated degradation trajectories for each unit, whose parameters η , α and ω_{k-1} follow the same normal distribution but have different correlations with the actual degradation trajectory. The correlation coefficients of three informative sensors are randomly sampled from the uniform distribution $U(0.8, 0.9)$, while

- those for the remaining seven non-informative sensors are sampled from $U(0, 0.5)$.
- 2) Generate a set of model parameters for the measurement function (7), where a_p and b_p are randomly sampled from $N(3, 1)$, and c_p are randomly sampled from $U(1, 3)$. Measurement noises of the 10 sensor signals follows the same normal distribution $N(0, 0.02)$. Their correlation matrix R is randomly generated using the method in [29]. Correspondingly, the covariance matrix of the measurement noises is equal to $\Sigma = 0.02R$.
 - 3) Input the 10 degradation trajectories from step 1 into (7) and generate the degradation signals of 10 sensors.

Two sets of state degradation trajectories and their sensor signals are shown in Fig. 4. Among the 100 units of each dataset, 90 are randomly selected as training units and the remaining 10 are used as test units in the following analysis.

4.2. Sensor selection of the simulation data

Training samples of the two datasets are input into the MSDFM in (9) separately, and model parameters are estimated using our proposed method. To consistent with the simulation process, the polynomial function $\varphi(x, \theta_p) = x^{c_p}$ is employed in the measurement function. To illustrate the working process of our method in dealing with nonlinear state degradation processes, we display the estimation results of parameter c_p in Fig. 5. It is seen that the estimation results of linear degradation processes are close to prespecified values. The parameters of nonlinear degradation processes are overestimated. The reason can be explained as follows. Our method defines a virtual health state following linear degradation processes. For nonlinear state degradation cases, the nonlinearity of state degradation is transformed into the measurement function. Therefore, parameter c_p in dataset 2 acquires higher estimation results than actual values. Due to the transformation of the nonlinearity, our method is able to deal with various state degradation behaviors.

To visualize the sensor selection process, we present the prioritized order of 10 sensors and the WARE scores of sensor groups in Fig. 6. It is seen that three informative sensors are ranked at the top of the line in both datasets. In addition, the third group acquires the lowest WARE score in both cases, which means that the group composed of three informative sensors performs best in RUL prediction. Therefore, three informative sensors are selected successfully by our method.

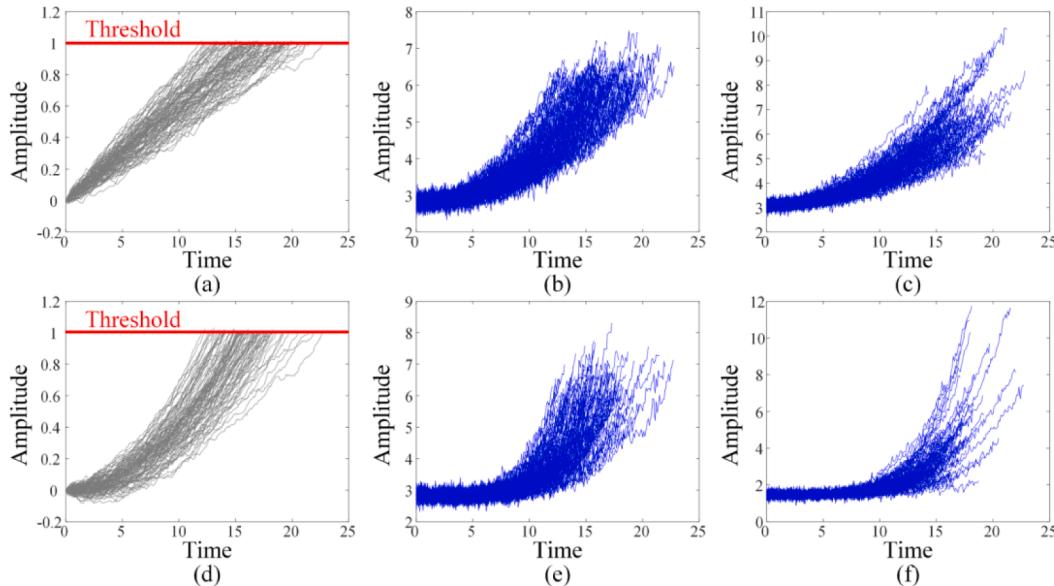


Fig. 4. Simulated degradation processes: (a) state degradation trajectories of dataset 1, (b) informative sensor signals of dataset 1, (c) noninformative sensor signals of dataset 1, (d) state degradation trajectories of dataset 2, (e) informative sensor signals of dataset 2, (f) noninformative sensor signals of dataset 2.

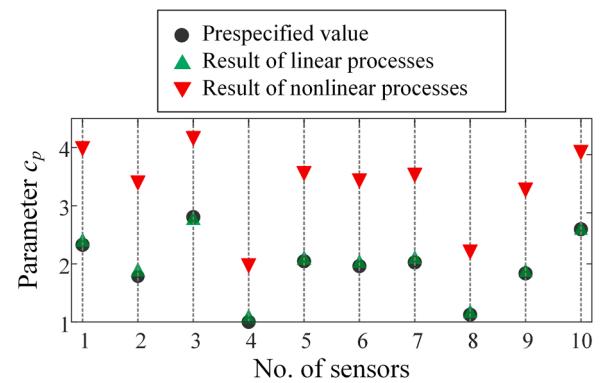


Fig. 5. Estimation results of c_p for both linear and nonlinear state degradation processes.

4.3. RUL prediction of the simulation data

To analyze the influence of sensor selection and fusion on RUL prediction, we compare the results of three variants by using different sensor groups. The first method designated “no fusion” only uses the first sensor that performs best in Fig. 6. The second method designated “no selection” uses all 10 sensors. Our method designated “selection & fusion” for distinguish uses three informative sensors.

To compare the prediction performance of three methods quantitatively, the ARE values of the 10 test units are calculated using (23). Their mean and variance at 10%, 20%, 30%, ..., 90% percentiles of lifetime are calculated and displayed in Fig. 7. It is seen that “no selection” presents higher ARE values during the lifetime. The mean and variance scores of “no fusion” and “fusion & selection” decreases gradually in dataset 1. In dataset 2, their prediction errors begin to decrease after half of the lifetime. This is corresponding to the degradation trend presented in Fig. 4(e), where obvious degradation performs after half of the lifetime. In both cases, “fusion & selection” acquires the lowest ARE values from 60% to 90% of lifetime. It implies that our method improves the accuracy of RUL prediction by selecting informative sensors and integrating their useful information. In addition, its ARE values are in the same level for both two datasets. This phenomenon demonstrates that our method can provide accurate and stable prediction results for both linear and nonlinear state degradation processes.

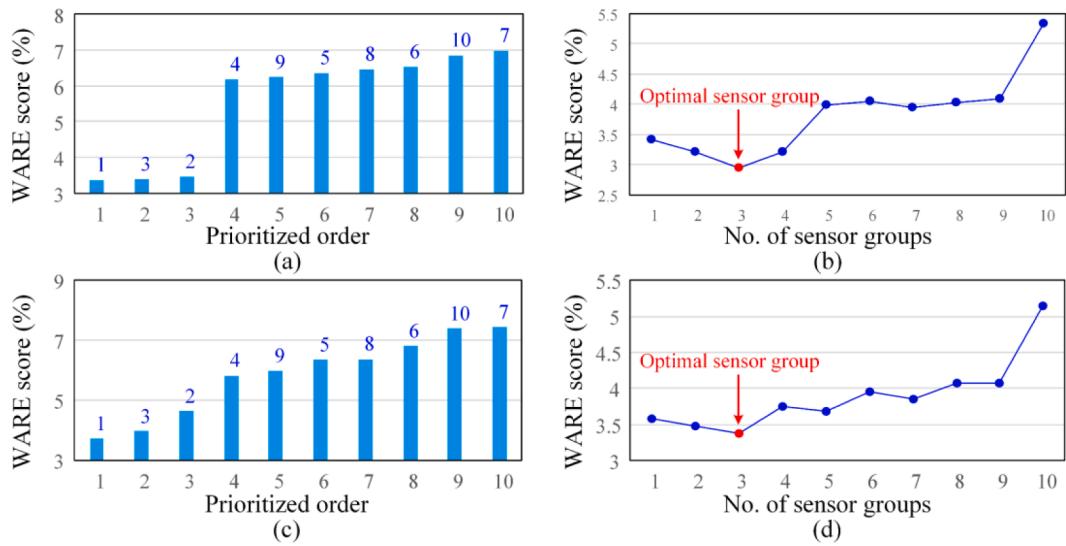


Fig. 6. WARE scores: (a) prioritized order of sensors in dataset 1, (b) scores of sensor groups in dataset 1, (c) prioritized order of sensors in dataset 2, (d) scores of sensor groups in dataset 2.

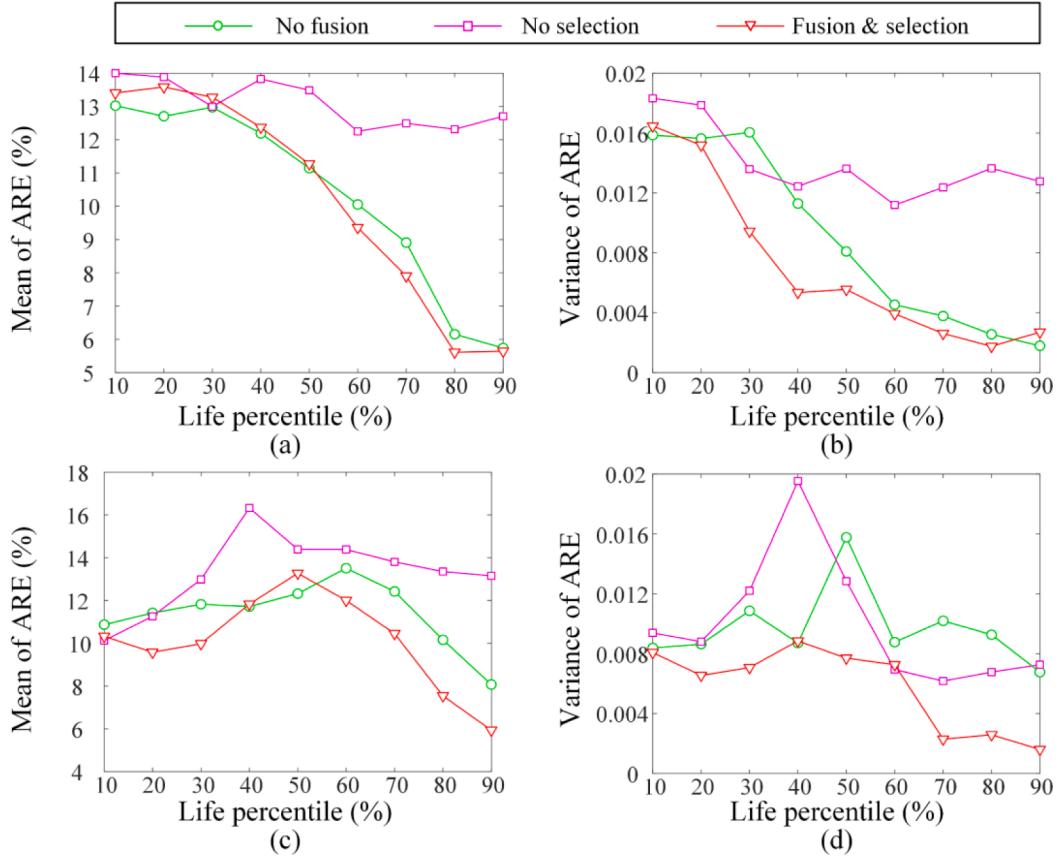


Fig. 7. Mean and variance of ARE scores: (a) mean of dataset 1, (b) variance of dataset 1, (c) mean of dataset 2, (d) variance of dataset 2.

5. Experimental demonstration

5.1. Experimental data introduction

To demonstrate the effectiveness of the proposed method, an aircraft engine degradation dataset provided in NASA prognostics data repository [30] is employed in this section. This dataset is composed of the following resources: degradation signals of 100 training engines during

the whole lifetime; degradation signals of 100 test engines that are terminated at random time points prior to failure; and the actual lifetimes of the 100 test engines. Totally 21 sensor signals were used to monitor the degradation of engines. The 21 sensor signals of the 100 training engines are shown in Fig. 8. It is obvious that some sensors present no information for engine degradation, such as 1, 5, 6, 10, 16, 18, and 19. It is difficult to estimate the health state of aircraft engines according the signals of these uninformative sensors. The following

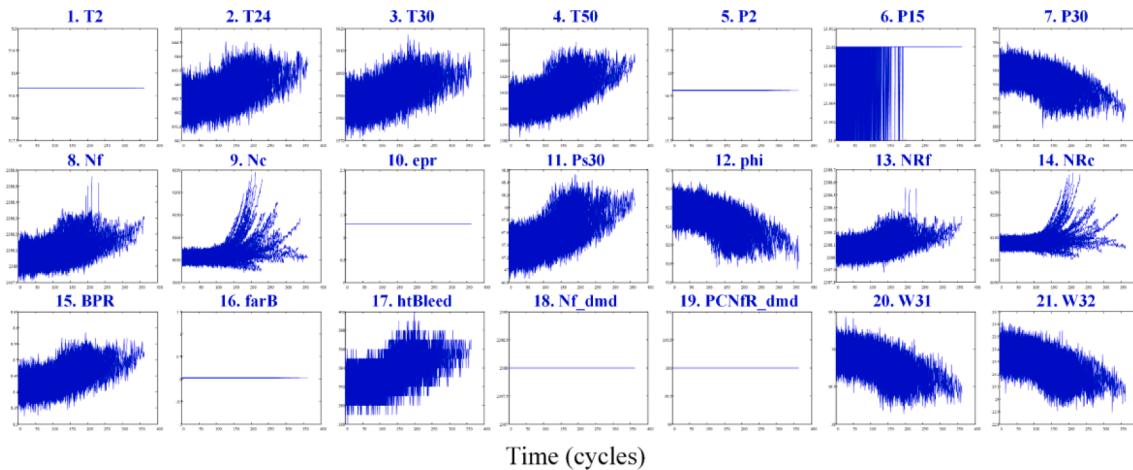


Fig. 8. Degradation signals of the 21 sensors of 100 training engines.

subsections will first select the optimal sensor group and prediction the RUL of 100 test engines using our proposed method as well as three other competitors.

5.2. Informative sensor selection

According to the nonlinear degradation trend of sensor signals in Fig. 8, we choose three different functions for $\varphi(x, \theta_p)$, including the polynomial function x^{c_p} , the exponential function $\exp(c_p x)$, and the logarithmic function $\ln(c_p(x + 1))$. The metric in (15) is used to evaluate the goodness of model fitting quantitatively. The polynomial function acquires the lowest fitting error among the three competitors. Thus, it is more suitable for this experimental case study. For other specific problems, a suitable function can be selected using the same strategy. To present the details of sensor selection, Fig. 9(a) provides the WARE scores calculated using each sensor, and Fig. 9(b) shows the WARE scores calculated using 21 sensor groups. According to the metric scores in Fig. 9(a), all of the seven uninformative sensors 1, 5, 6, 10, 16, 18, and 19 are ranked at the end of the order. Sensors that present obvious degradation trend are ranked in front of the order, such as 11, 12, 2, and

20. In Fig. 9(b), the WARE scores jumps from sensor group 15, which indicates that the prediction error increases rapidly when each of the last seven uninformative sensors is involved. From the enlarged view of the first 14 groups we can see that, the group composed of four sensors 11, 12, 20, and 2 acquires the smallest WARE score, which means that it provides the most accurate RUL prediction results for 100 training units. Thus, this group is selected as the optimal sensor group for RUL prediction.

5.3. RUL prediction

The three methods used in the simulation case study are employed to predict the RUL of 100 test units. We also choose a benchmark method [14] for comparison, which selects informative sensors via penalized location scale regression, and fuses multi-sensor signals using MFPCA. AREs of their prediction results are calculated using (23). The mean and variance of their AREs at 10%, 20%, 30%, ..., 90% percentiles of lifetime are calculated and displayed in Fig. 10. It is seen that, as time goes on, the mean and variance curves of all methods decrease gradually, which implies that they provide more accurate and stable prediction results when more degradation signals are available. In addition, our proposed method presents better performance than other three methods from 40% of lifetime. In general, our method acquires the most accurate and stable prediction result among most of the lifetime.

The superiority of our method compared with other three competitors can be explained with the following reasons. The “no fusion” method predicts RUL using a single sensor, which may provide incomplete degradation information resulting in inaccurate prediction results. Conversely, the “no selection” method fuses all sensor signals for RUL prediction. However, non-informative sensors may reduce the performance of the prediction method. Fang’s method is more concerned about the overall performance of a sensor group. It is difficult to guarantee the effectiveness of each sensor in the selected sensor group. Our method selects informative sensors following their priority order and fuses multi-sensor data using a generalizable multivariate state-space modeling framework. It promises the validity of sensor selection and fusion. As a result, it acquires the most accurate and stable RUL prediction result among the four competitors.

6. Conclusions

RUL prediction with multi-sensor signals is a more challenging issue than the cases of a single degradation signal. Most studies try to deal with this issue by combining multiple sources of sensor signals into a composite HI or mapping the correlation between signals and RUL values. They define failure based on the amplitude of sensor signals.

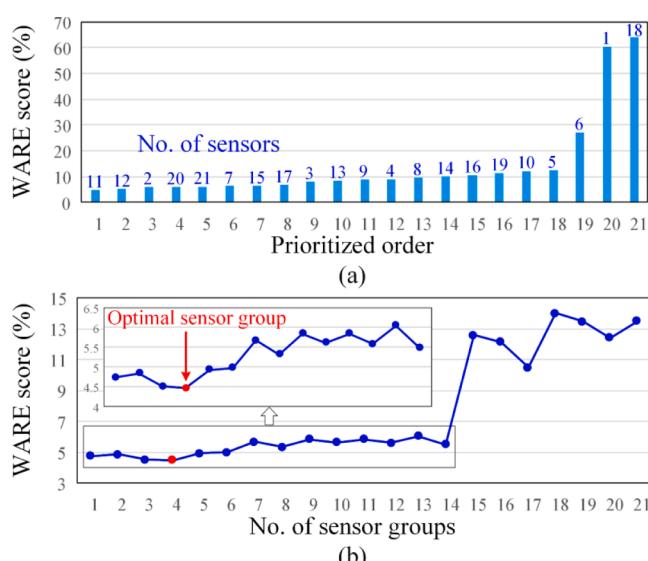


Fig. 9. WARE scores of prediction results: (a) results predicted using each of the 21 sensors, and (b) results predicted using 21 sensor groups. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

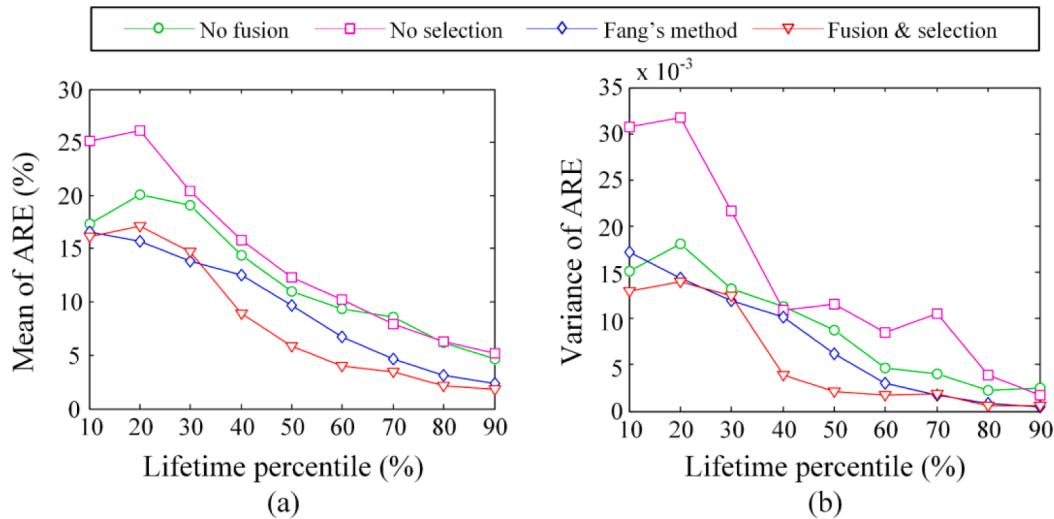


Fig. 10. Mean and variance of ARE scores for the aircraft engine data: (a) mean, and (b) variance.

Different from existing studies, we try to explicate the behavior of multi-sensor signals from the perspective of state-space modeling, i.e., multi-sensor signals are different proxies of the hidden system state. The failure is determined by the underlying system state instead of the sensor signals. This modeling framework provides a new and more straightforward explanation for the relationship between the degradation process of the system state and the behaviors of the sensor signals. Based on this novel idea, we propose a MSDFM-based method for RUL prediction. A PSGS algorithm is also proposed to select informative sensors from multi-stream sensor networks. The effectiveness of the proposed method is validated using aircraft engine degradation data provided by NASA prognostics repository.

In this work, we only focus on industrial systems whose health state is completely unobservable. The underlying system state is assumed as a virtual health state following a linear Wiener process. This assumption is actually a concession to unobservability of system states. In real applications, health states of some systems are semi-observable via shutdown inspection, such as cutting tool wear of a milling machine and fatigue crack growth of a gearbox. The state observations will provide useful information for degradation modeling and RUL prediction. It is a more

realistic and challenging issue when offline state observations are coupled with online multi-sensor monitoring signals. In the future, we will conduct research regarding multi-sensor data driven RUL prediction for this kind of semi-observable systems.

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

This research was supported by National Key R&D Program of China (2018YFB1306100), NSFC-Zhejiang Joint Fund for the Integration of Industrialization and Informatization (U1709208), National Natural Science Foundation of China (52005387, 61673311), the Opening Project of Shanxi Key Laboratory of Advanced Manufacturing Technology (XJZZ201902), and the Fundamental Research Funds for the Central Universities.

Appendix A

Since the lifetime PDF follows the inverse Gaussian distribution in (10), the model parameters depending on the lifetimes of training units are

$$\mathcal{L}(\mu_\eta, \sigma_\eta^2, \sigma_B^2 | \mathbf{T}) = -\sum_{n=1}^N \ln(t_{n,K_n}) - \frac{1}{2} \sum_{n=1}^N \ln(\sigma_\eta^2 t_{n,K_n}^2 + \sigma_B^2 t_{n,K_n}) - \frac{N}{2} \ln(2\pi) - \sum_{n=1}^N \frac{(\mu_\eta t_{n,K_n} - D)^2}{2(\sigma_\eta^2 t_{n,K_n}^2 + \sigma_B^2 t_{n,K_n})} \quad (25)$$

Let $\tilde{\sigma}_B^2 = \sigma_B^2/\sigma_\eta^2$. The log-likelihood function is transformed to

$$\mathcal{L}(\mu_\eta, \sigma_\eta^2, \tilde{\sigma}_B^2 | \mathbf{T}) = -\sum_{n=1}^N \ln(t_{n,K_n}) - \frac{1}{2} \sum_{n=1}^N \ln(t_{n,K_n}^2 + \tilde{\sigma}_B^2 t_{n,K_n}) - \frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma_\eta^2) - \sum_{n=1}^N \frac{(\mu_\eta t_{n,K_n} - D)^2}{2\sigma_\eta^2 (t_{n,K_n}^2 + \tilde{\sigma}_B^2 t_{n,K_n})} \quad (26)$$

Calculate the first derivatives of the log-likelihood function with respect to μ_η and σ_η^2 , and let them equal zero. The MLEs of μ_η and σ_η^2 as calculated as (12) and (13), respectively. Submit (12) and (13) into (26). It is reduced into (11) that is only dependent on $\tilde{\sigma}_B^2$. Parameter $\tilde{\sigma}_B^2$ is estimated by maximizing the criterion. The MLEs of μ_η , σ_η^2 and σ_B^2 are calculated by inputting $\tilde{\sigma}_B^2$ into corresponding equations.

Appendix B

Three conditions for function $\varphi(\cdot)$ in 3.1 ensure that x_k and $\varphi(x_k, \theta_p)$ are one-to-one mapping. Thus, the state value can be approximated using the following inverse function:

$$\hat{x}_{n,k}(\theta_p) = \psi \left(\frac{\tilde{y}_{n,k}^p - b_p(\theta_p)}{a_p(\theta_p)}, \theta_p \right) \quad (27)$$

Since the actual state values are unobservable, it is impossible to estimate the model parameters by fitting the approximated state value with the actual value. Here we choose an easy-to-implement methodology to estimate the model parameters. From (6) it is observed that, the state sequence $(x_{n,1}, x_{n,2}, \dots, x_{n,K_n})'$ follows a multivariate normal distribution with mean trend $(\mu_{\eta} t_{n,1}, \mu_{\eta} t_{n,2}, \dots, \mu_{\eta} t_{n,K_n})'$. Therefore, the parameters θ_p can be estimated by minimizing the square error between the approximated state values and their mean trend as shown in (15).

According to our definition to the system state, it satisfies $x_{n,1} = 0$ and $x_{n,K_n} = 1$. By submitting them into (14), we can express a_p and b_p as (16) and (17), respectively.

References

- [1] Lei Y, Li N, Guo L, Li N, Yan T, Lin J. Machinery health prognostics: a systematic review from data acquisition to RUL prediction. *Mech Syst Signal Process* 2018; 104:799–834.
- [2] Wang D, Tsui K-L, Miao Q. Prognostics and health management: a review of vibration based bearing and gear health indicators. *IEEE Access* 2018;6:665–76.
- [3] Li N, Gebraeel N, Lei Y, Bian L, Si X. Remaining useful life prediction of machinery under time-varying operating conditions based on a two-factor state-space model. *Reliab Eng Syst Saf* 2019;186:88–100.
- [4] Peng W, Li Y-F, Yang Y-J, Huang H-Z, Zuo MJ. Inverse Gaussian process models for degradation analysis: a Bayesian perspective. *Reliab Eng Syst Saf* 2014;130: 175–89.
- [5] Si XS. An adaptive prognostic approach via nonlinear degradation modeling: application to battery data. *IEEE Trans Ind Electron*. 2015;62:5082–95.
- [6] Le Son K, Fouladirad M, Barros A. Remaining useful lifetime estimation and noisy gamma deterioration process. *Reliab Eng Syst Saf* 2016;149:76–87.
- [7] Tao T, Zio E, Zhao W. A novel support vector regression method for online reliability prediction under multi-state varying operating conditions. *Reliab Eng Syst Saf* 2018;177:35–49.
- [8] Yang F, Habibullah MS, Zhang T, Xu Z, Lim P, Nadarajan S. Health index-based prognostics for remaining useful life predictions in electrical machines. *IEEE Trans Ind Electron* 2016;63:2633–44.
- [9] Liu K, Chehade A, Song C. Optimize the signal quality of the composite health index via data fusion for degradation modeling and prognostic analysis. *IEEE Trans Autom Sci Eng* 2017;14:1504–14.
- [10] Tse PW, Wang D. Enhancing the abilities in assessing slurry pumps' performance degradation and estimating their remaining useful lives by using captured vibration signals. *J Vib Control* 2017;23:1925–37.
- [11] Lei Y, Li N, Gontarz S, Lin J, Radkowski S, Dybala J. A model-based method for remaining useful life prediction of machinery. *IEEE Trans Reliab*. 2016;65: 1314–26.
- [12] Liu Y, Hu X, Zhang W. Remaining useful life prediction based on health index similarity. *Reliab Eng Syst Saf*. 2019;185:502–10.
- [13] Liu K, Huang S. Integration of data fusion methodology and degradation modeling process to improve prognostics. *IEEE Trans Autom Sci Eng* 2016;13:344–54.
- [14] Fang X, Paynabar K, Gebraael N. Multistream sensor fusion-based prognostics model for systems with single failure modes. *Reliab Eng Syst Saf* 2017;159:322–31.
- [15] Chen J, Jing H, Chang Y, Liu Q. Gated recurrent unit based recurrent neural network for remaining useful life prediction of nonlinear deterioration process. *Reliab Eng Syst Saf* 2019;185:372–82.
- [16] Sun F, Liu L, Li X, Liao H. Stochastic modeling and analysis of multiple nonlinear accelerated degradation processes through information fusion. *Sens* 2016;16:1242.
- [17] Liu J, Vitelli V, Zio E, Seraoui R. A novel dynamic-weighted probabilistic support vector regression-based ensemble for prognostics of time series data. *IEEE Trans Reliab* 2015;64:1203–13.
- [18] Listou Ellefsen A, Bjørlykhaug E, Aasøy V, Ushakov S, Zhang H. Remaining useful life predictions for turbofan engine degradation using semi-supervised deep architecture. *Reliab Eng Syst Saf* 2019;183:240–51.
- [19] Wei M, Chen M, Zhou D. Multi-sensor information based remaining useful life prediction with anticipated performance. *IEEE Trans Reliab* 2013;62:183–98.
- [20] Ramasse E. Investigating computational geometry for failure prognostics. *Int J Progn Health Manag* 2014;5:1–18.
- [21] Orchard ME, Hevia-Koch P, Zhang B, Tang L. Risk measures for particle-filtering-based state-of-charge prognosis in lithium-ion batteries. *IEEE Trans Ind Electron* 2013;60:5260–9.
- [22] Hao S, Yang J, Bérenguer C. Nonlinear step-stress accelerated degradation modelling considering three sources of variability. *Reliability Eng Syst Saf* 2018; 172:207–15.
- [23] Arulampalam MS, Maskell S, Gordon N, Clapp T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans Signal Process* 2002; 50:174–88.
- [24] Jouin M, Gouriveau R, Hissel D, Péra M-C, Zerhouni N. Particle filter-based prognostics: review, discussion and perspectives. *Mech Syst Signal Process* 2016; 72:73:2–31.
- [25] Chikara RS, Folks L. The inverse gaussian distribution. theory, methodology, and applications. New York: Marcel Dekker; 1989.
- [26] Li N, Lei Y, Guo L, Yan T, Lin J. Remaining useful life prediction based on a general expression of stochastic process models. *IEEE Trans Ind Electron* 2017;64:5709–18.
- [27] Li N, Lei Y, Yan T, Li N, Han T. A Wiener-process-model-based method for remaining useful life prediction considering unit-to-unit variability. *IEEE Trans Ind Electron* 2019;66:2092–101.
- [28] Cleveland WS, Devlin SJ. Locally weighted regression: an approach to regression analysis by local fitting. *J Am Stat Assoc* 1988;83:596–610.
- [29] Numpacharoen K, Atsawarungruskit A. Generating correlation matrices based on the boundaries of their coefficients. *PLoS ONE* 2012;7:e48902.
- [30] Saxena A, Goebel K. Turbofan engine degradation simulation data set, NASA Ames prognostics data repository. Moffett Field, CA: NASA Ames Research Center; 2008.