Journal of Quality Technology
A Quarterly Journal of Methods, Applications and Related Topics

ISSN: 0022-4065 (Print) 2575-6230 (Online) Journal homepage: https://www.tandfonline.com/loi/ujqt20

A data-level fusion approach for degradation modeling and prognostic analysis under multiple failure modes

Abdallah Chehade, Changyue Song, Kaibo Liu, Abhinav Saxena & Xi Zhang

To cite this article: Abdallah Chehade, Changyue Song, Kaibo Liu, Abhinav Saxena & Xi Zhang (2018) A data-level fusion approach for degradation modeling and prognostic analysis under multiple failure modes, Journal of Quality Technology, 50:2, 150-165, DOI: 10.1080/00224065.2018.1436829

To link to this article: https://doi.org/10.1080/00224065.2018.1436829

Published online: 09 Apr 2018.

Submit your article to this journal

Article views: 724

View related articles

View Crossmark data

Citing articles: 5 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=ujqt20

JOURNAL OF QUALITY TECHNOLOGY
2018, VOL. 50, NO. 2, 150–165
https://doi.org/10.1080/00224065.2018.1436829

Check for updates

A data-level fusion approach for degradation modeling and prognostic analysis under multiple failure modes

Abdallah Chehade[a], Changyue Song[b], Kaibo Liu[b], Abhinav Saxena[c], and Xi Zhang[d]

[a]Department of Industrial and Manufacturing Systems Engineering, University of Michigan–Dearborn, Dearborn, Michigan; [b]Department of Industrial and Systems Engineering, University of Wisconsin–Madison, Madison, Wisconsin; [c]Machine Learning Lab, Software Sciences and Analytics, GE Global Research Center, San Ramon, California; [d]Department of Industrial Engineering and Management, Peking University, Beijing, China

ABSTRACT

Operating units, in practice, often suffer from multiple modes of failure, and each failure mode has a distinct influence on the service life cycle path of a unit. The rapid development of sensor and communication technologies has enabled multiple sensors to simultaneously monitor and track the health status of a unit in real time. However, one challenging question that remains to be resolved is how to leverage data from multiple sensors for better degradation modeling and prognostic analysis, especially when there are multiple failure modes. Currently, many of the existing approaches in prognostics either (a) fail to capture the dependency between sensors and instead focus on analyzing each sensor independently or (b) fail to incorporate the failure-mode diagnosis for better degradation modeling and prognostics during condition monitoring. To address the limitations in the existing literature, we propose a data-level fusion methodology to construct a composite failure-mode index, named FM-INDEX, via the fusion of multiple sensor data. Our goal is to utilize the FM-INDEX to better characterize the failure mode of an operating unit in real time, thus leading to better degradation modeling and prognostic analysis. A case study that involves the degradation data set of an aircraft gas turbine engine with two potential failure modes is conducted to numerically evaluate the performance of our proposed method compared to other techniques in the related literature.

KEYWORDS

big data; condition monitoring; data fusion; health index; multiple sensors

1. Introduction

Unexpected failures often result in production downtime and delayed schedules, which may lead to severe economic losses, customer dissatisfaction, and safety issues. As a result, it is critically important to accurately measure the health status of a unit (e.g., machine, tool, equipment) in real time to avoid unexpected failures. To achieve this goal, condition monitoring techniques that aim to fully understand and track the degradation status of a unit to better estimate its remaining lifetime have been widely used (Saha, Goebel, and Christophersen 2009; Ye, Tang, and Xie 2011). With the availability of such prognostic information, a condition-based maintenance strategy (Mobley 2002; Pulcini 2008) can then be implemented to significantly improve production efficiency, reduce inventory and operational costs, and enhance customer loyalty.[1]

The rapid development of sensor and communication technologies has enabled a big-data environment with multiple sensors simultaneously deployed to monitor complex systems, which provides a great opportunity for condition monitoring. However, most of the existing approaches focus on single sensor signals for degradation modeling and prognostic analysis. These approaches rely on the assumption that the single sensor fully characterizes the underlying physical transition of the degradation process. Unfortunately, such a simplified assumption may not be valid in many real-world applications and often leads to inaccurate or unreliable prognostic results (Brotherton et al. 2002; Saxena et al. 2008a).

While the big data collected from multiple sensors provide an unprecedented opportunity to better understand and monitor the health status of an operating unit in real time, they also pose unique challenges. These challenges range from the variety of possible degradation profiles, the complexity of potential failure mechanisms, sensor correlations/interdependencies, computational costs, and more. For the purpose of degradation modeling and prognostic analysis, we focus on the following two questions: (1) how to accurately diagnose the failure mode causing the unit to degrade during condition monitoring and (2) how to combine relevant sensor data. To answer

CONTACT Xi Zhang xi.zhang@coe.pku.edu.cn 5 Yiheyuan Road, Beijing, China, 100871.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/ujqt.
[1]For consistency, bold face is used for vectors and matrices. All vectors are assumed to be column vectors by default.
© 2018 American Society for Quality

these questions, data-fusion approaches have been widely used. An overview of data-fusion approaches to condition monitoring, fault diagnosis, and prognostics can be found in Jardine, Lin, and Banjevic 2006. Data fusion can be classified into two main categories (Hall and Llinas 1997) based on the implementation level: data-level fusion and decision-level fusion. Specifically, data-level fusion combines multiple sensor data or extracted features into a one-dimensional health index that can be further used for decision making (e.g., degradation modeling and prognostics), whereas decision-level fusion involves the integration of prediction results from separate analyses of each individual sensor data (Hastie et al. 2005). While the decision-level fusion is straightforward to implement, it ignores the dependency of multiple sensor data and requires repeated computations based on each individual sensor data. Thus, such an approach often leads to biased results, making its application to big data quite limited.

Unlike decision-level fusion, data-level fusion has shown promise as an effective solution for prognostics and has recently attracted more attention within the literature (Liu and Shi 2015). However, most of the existing data-level fusion methods for degradation modeling and prognostics assume there is only one potential failure mode or ignore the differences between the failure modes. Unfortunately, as shown in many real-world applications, multiple failure modes that may have distinct influences on the degradation process usually exist. Thus, an essential challenge here is to accurately identify the failure mode and predict the remaining lifetime along the life cycle of a unit in real time. To achieve accurate failure mode diagnosis, several techniques have been proposed in the literature, such as Hidden Markov Models (Bunks, McCarthy, and Al-Ani 2000), Artificial Intelligence (Ali, Rahman, and Hamzah 2014), and Physical-Based Modeling Approaches (Simani, Fantuzzi, and Patton 2013). These techniques, however, mainly focus on analyzing data from a single sensor. While some studies have considered the failure mode diagnosis based on multiple sensor data, they either (a) employ simple voting schemes that ignore the dependencies between the sensors (Heger and Pandit 2004; Kobayashi and Simon 2007; Salahshoor, Mosallaei, and Bayat 2008) or (b) fail to incorporate the failure mode diagnosis for better degradation modeling and prognostic analysis during condition monitoring (Isermann 1997; Wang and Wang 2000; Volponi et al. 2004; Basir and Yuan 2007; Sarkar, Jin, and Ray 2011; Jiang, Wang, and Tsung 2012; De Ketelaere, Hubert, and Schmitt 2015). In particular, there are two fundamental requirements that need to be satisfied: (1) the failure mode diagnosis needs to be conducted and updated continuously throughout the life cycle of a unit and (2) the fault diagnostic result should become more accurate as the unit approaches the end of its life. This will help ensure proper maintenance planning and avoid unexpected failure.

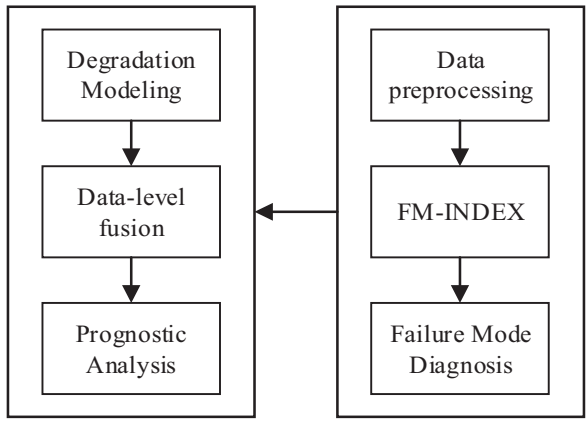

Figure 1. A flow chart of the proposed data-level fusion approach for degradation modeling and prognostics when there are multiple failure modes.

To the best of our knowledge, the existing literature still lacks a generic data-fusion approach that can be effectively used for degradation modeling and prognostics for systems with multiple sensors and multiple failure modes. Thus, this article seeks to fill this literature gap by developing a data-level fusion methodology via the combination of multiple sensor data under multiple failure modes. Specifically, our proposed method first constructs a failure mode index, named FM-INDEX, which allows for the online estimation of the probability that a unit belongs to a certain failure mode. Once developed, the FM-INDEX is integrated with existing data-level fusion approaches (e.g., Liu, Chehade, and Song 2015; Liu and Huang 2016) to address the challenges of degradation modeling and prognostics when there are multiple failure modes. Figure 1 shows a schematic diagram of the proposed data-fusion approach.

The rest of the article is structured as follows: Section 2 first introduces the Commercial Modular Aero-Propulsion System Simulation (C-MAPSS) that was developed by NASA (Frederick, DeCastro, and Litt 2007). Then the article reviews a degradation data set (Saxena et al. 2008b) that was generated by C-MAPSS to simulate the degradation process of aircraft gas turbine engines. Section 3 proposes a data-level fusion method for failure mode diagnosis during condition monitoring, which is tailored to the needs of degradation modeling and prognostics. Section 4 demonstrates the effectiveness of the proposed method in terms of its failure mode diagnosis, degradation modeling, and remaining life prediction using the data set introduced in Section 2. Section 5 draws a conclusion and discusses future research directions.

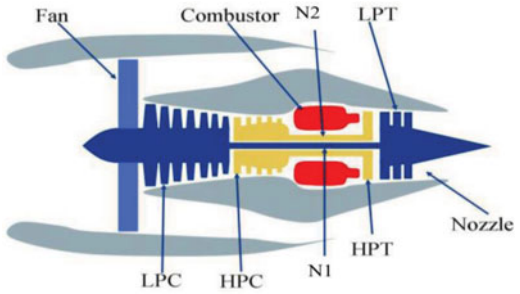

Figure 2. Simplified engine diagram simulated by C-MAPSS (Saxena et al. 2008b).

2. Data set overview

C-MAPSS is a tool for simulating realistic, large commercial turbofan engines (Frederick, DeCastro, and Litt 2007), which has been widely used in the literature of condition monitoring and degradation analysis (Ramasso and Saxena 2014). Figure 2 shows a schematic diagram of the aircraft gas turbine engine that is simulated by C-MAPSS.

This simulator generates and records 21 sensor signals of aircraft engines in their service life cycles. The description of the 21 sensors is given in Table 1. Furthermore, each operating unit is subject to one of two potential failure modes: failure due to the high-pressure compressor (HPC) or failure due to the engine's fan. Specifically, each engine starts with an unknown initial wear condition and proceeds to degrade until failure. Due to the complexity of the engine system and the dynamics of the degradation process, the damage profile of an engine cannot be directly inferred based on flight duration or physical models. Also, the simulation model for generating the data set is not available to the users. Thus, the underlying assumption here is that users have to rely solely on the multiple sensor data, which is collected from each engine to estimate the current health status and make predictions about future behavior.

Table 1. Detailed sensor descriptions (Saxena et al. 2008b).

Symbol	Description	Units
T2	Total temperature at fan inlet	°R
T24	Total temperature at LPC outlet	°R
T30	Total temperature at HPC outlet	°R
T50	Total temperature at LPT outlet	°R
P2	Pressure at fan inlet	psia
P15	Total pressure in bypass-duct	psia
P30	Total pressure at HPC outlet	psia
Nf	Physical fan speed	rpm
Nc	Physical core speed	rpm
epr	Engine pressure ratio (P50/P2)	—
Ps30	Static pressure at HPC outlet	psia
phi	Ratio of fuel flow to Ps30	pps/psi
NRf	Corrected fan speed	rpm
NRc	Corrected core speed	rpm
BPR	Bypass ratio	—
farB	Burner fuel-air ratio	—
htBleed	Bleed enthalpy	—
Nf_dmd	Demanded fan speed	rpm
PCNfR_dmd	Demanded corrected fan speed	rpm
W31	HPT coolant bleed	lbm/s
W32	LPT coolant bleed	lbm/s

The collected degradation data set contains 100 training units. To obtain the failure mode for the training units, we leveraged another data set (called "control group") from the same source (Saxena et al. 2008b), which contains units that are known to fail due only to the HPC failure. Specifically, we compared the trends (considered to be an increasing/decreasing trend if the initial observation is smaller/larger, respectively, than the last observation before failure) of the multiple sensor data (e.g., P30, phi, BPR, W31, W32) from these two data sets and found that there are 44 training units from this data set consistently showing different trend information compared to the control group. As a result, our preliminary analysis concluded that 56 training units fail due to the fault at the HPC (we denote this failure mode type with the label "−1"), and the remaining 44 training units fail due to the fault at the fan (we denote this failure mode type with the label "1"). The data set also contains another 100 testing units with unknown failure modes. Each training unit runs until it fails, while each testing unit stops operation at some point before failure. There is also a file recording the actual remaining lifetime of the 100 testing units. Thus, the key challenge here is to accurately estimate the remaining lifetime of these testing units based on the available observations. A more detailed description of the data set can be found in (Saxena et al. 2008b).

3. Methodology development

This section proposes a data-level fusion methodology for degradation modeling and prognostics when there are multiple failure modes. As we mentioned in Section 1, different failure modes may have distinct influences on the degradation path of a unit. To illustrate this point, Figure 3 shows an example of the data set introduced in Section 2. This figure displays the degradation profiles of two different sensors for two units that fail due to two different failure modes. It clearly shows that the sensor measurements tend to be highly correlated and that some sensors are sensitive to certain types of failure modes (e.g., sensor 2) while others (e.g., sensor 1) are not. This example motivates the use of data-level fusion approaches that take advantage of multiple and dependent sensor information to better infer the failure mode of a unit along its service life cycle. If such failure mode information is known, then it can be readily integrated with the existing data-level fusion models for better degradation modeling and prognostics.

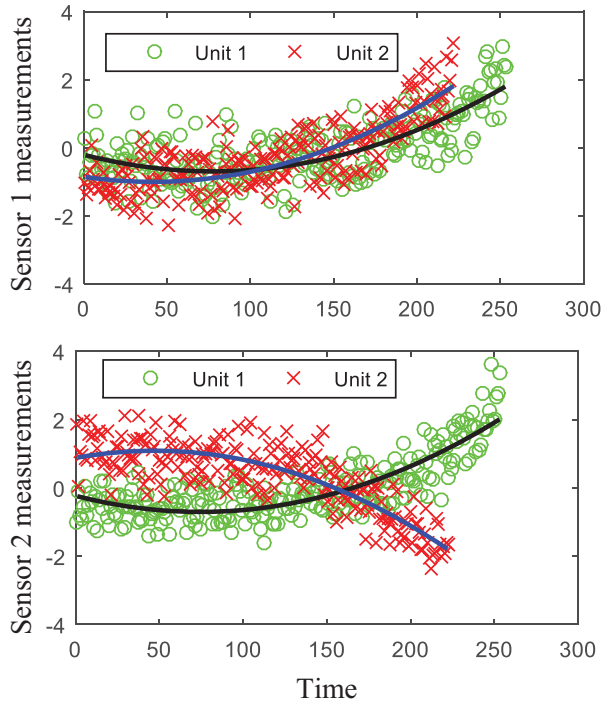

Figure 3. An example of the degradation profiles of two sensors for two units. The circles correspond to the sensor measurements for unit 1 that fail due to failure mode 1, whereas the crosses correspond to the sensor measurements for unit 2 that fail due to failure mode –1. The lines represent a second-order polynomial fit for the sensor measurements.

3.1. Data-level fusion for failure mode diagnosis

In this subsection we describe the proposed data-level fusion model for failure mode diagnosis by constructing a composite index, named FM-INDEX, as a combination of multiple sensor data.

3.1.1. Model formulation

Let $\boldsymbol{L}_{i,.,t} = [L_{i,1,t}, \ldots, L_{i,s,t}] \in \mathbb{R}^{1\times s}$ be the vector of sensor measurements of s sensors for unit i at time t for $t = 1, \ldots, n_i$. Here, $\mathbb{R}$ is the set of real numbers, $L_{i,j,t}$ is the measurement of sensor j for unit i at time t, and n_i is the number of available observations for unit i. In addition, we denote z_i to be the true failure mode that training unit i belongs to, m to be the total number of training units, and $V(f(\cdot), z)$ to be the loss function that quantifies the deviation between the predicted label given the fusion model $f(\cdot)$ and the ground truth label z. To achieve a failure mode diagnosis, one possible idea is to construct a fusion model $f_t(\boldsymbol{L}_{i,.,t})$ that predicts the failure mode for unit i such that the loss function $V(f_t(\boldsymbol{L}_{i,.,t}), z_i)$ can be minimized over the entire training data set:

$$\text{obj} = \min_{f_t} \frac{1}{m} \sum_{i=1}^{m} \sum_{t=1}^{n_i} V(f_t(\boldsymbol{L}_{i,.,t}), z_i). \quad (1)$$

While this approach is intuitively sound, it produces a fusion model $f_t(\boldsymbol{L}_{i,.,t})$ at each observation time t, which is difficult to implement for new testing units during condition monitoring. As a result, we propose the following optimization problem with one fusion model $f(\boldsymbol{L}_{i,.,t})$ that is independent of the observation time t:

$$\text{obj} = \min_{f} \frac{1}{m} \sum_{i=1}^{m} \sum_{t=1}^{n_i} a_{i,t} V(f(\boldsymbol{L}_{i,.,t}), z_i) + \lambda R(f). \quad (2)$$

Here, we define $x_{i,t} = f(\boldsymbol{L}_{i,.,t})$ to be the FM-INDEX for unit i at time t. Thus, an ideal case is that the constructed FM-INDEX, $x_{i,t}$, can always retrieve the true failure mode label z_i (i.e., achieving the minimum value in the loss function $V(x_{i,t}, z_i)$). $R(f)$ is a regularization function to prevent overfitting, λ is the regularization parameter, and $a_{i,t}$ is called the loss weight for unit i at time t, which is a key parameter of our proposed data-fusion model in Eq. [2]. In the next subsection, we show how to construct the loss weights $\{a_{i,t}\}$ such that the failure mode diagnosis can be tailored to the needs of degradation modeling and prognostics. One of the main advantages of the proposed approach is the flexibility in choosing the loss function. Furthermore, under certain parameters, the proposed method can be effectively transformed into several known and powerful classifiers such as the logistic regression (Freedman 2009) and the Support Vector Machine (SVM) (Chang and Lin 2011). More details about the loss function and $R(f)$ are discussed in Section 3.1.3.

3.1.2. Setting the loss weights $\{a_{i,t}\}$

Denote M_k to be the set of training units that belong to failure mode k. In particular, we consider the following three unique requirements when setting the loss weights $\{a_{i,t}\}$:

1. Equal importance of the units from the same failure mode;
2. equal importance for different failure modes; and
3. a more accurate failure mode diagnostic result as the unit approaches failure.

Requirement (1) means that for each unit that belongs to the same failure mode, its importance should be equally considered in the proposed data-fusion model. Thus, the following constraint must be satisfied:

$$\sum_{t=1}^{n_i} a_{i,t} = c_k, \quad \forall i \in M_k, \forall k, \quad (3)$$

where c_k denotes a constant number whose value depends on failure mode k.

In practice, as the number of training units from different failure modes may not be the same, requirement (2)

must also be satisfied to deal with this imbalanced classification problem:

$$\sum_{i\in M_k}\sum_{t=1}^{n_i} a_{i,t} = |M_k| * c_k = c, \quad \forall k, \tag{4}$$

where c denotes a constant value and $|M_k|$ is the cardinality of the set M_k (i.e., the number of training units that fail due to failure mode k). This constraint ensures equal contributions from each failure mode in constructing the FM-INDEX.

Considering that the remaining life prediction becomes more critical as a unit approaches the end of its life, requirement (3) must be satisfied as well. We impose this constraint by setting $\{a_{i,t}\}$ as a nondecreasing series:

$$a_{i,t+1} \geq a_{i,t}, \quad i = 1, \ldots, m, t = 1, 2, \ldots, n_i - 1. \tag{5}$$

In this way, a greater penalty will be given to the loss function as time increases to ensure a more accurate failure mode diagnostic result as the unit approaches the end of its life. Specifically, if we assume $a_{i,t}$ follows an arithmetic series ($2a_{i,t} = a_{i,t+1} + a_{i,t-1}$), then we can show that the values of the loss weights for training units that fail due to failure mode k satisfy the following equation:

$$a_{i,t} = a_{i,1} + (t-1)\frac{\frac{2m}{|M_k|} - 2a_{i,1}n_i}{(n_i-1)n_i}, t = 1, \ldots, n_i, \forall i \in M_k. \tag{6}$$

In the above equation, c_k in Eq. [3] is set to be $\frac{m}{|M_k|}$ and c in Eq. [4] is set to be m.

3.1.3. Hinge-loss functions

Many of the existing loss functions (e.g., square loss, logistic loss, hinge loss) can be used in Eq. [2] to identify the failure mode during condition monitoring (Hastie et al. 2005). Here, we choose to focus on the hinge-loss function as a demonstration. Recall that in Section 2, our problem of interest has two potential failure modes; thus we concentrate on the two failure mode scenarios. A discussion on the extensions to the multiple failure modes is given at the end of this subsection. In addition, to highlight our main idea, we focus on the linear fusion model as an illustration (i.e., $x_{i,t} = f(\mathbf{L}_{i,.,t}) = \mathbf{L}_{i,.,t}\mathbf{w} + b$). Without loss of generality, such a linear model can be easily extended to nonlinear models by either the creation of nonlinear features or the use of kernel methods (Kim and Scott 2010).

Specifically, the hinge loss function can be expressed as :

$$V(x_{i,t}, z_i) = \xi_{i,t} = \max(0, 1 - z_i x_{i,t}), \tag{7}$$

where $\xi_{i,t}$ is called the misclassification error. Based on the hinge-loss function, the failure mode of training unit i is classified correctly at time t (i.e., with zero loss, $\xi_{i,t} = 0$) if $z_i x_{i,t} \geq 1$; otherwise, it is misclassified with an error $\xi_{i,t} = 1 - z_i x_{i,t}$. It is known that the models solely relying on the hinge-loss function are subject to overfitting. To avoid this problem, a regularization function $R(f) = f^2 = \mathbf{w}^T\mathbf{w}$ is usually integrated with the hinge-loss function. It can be shown that with the regularized hinge-loss function, the proposed data-level fusion model in Eq. [2] simplifies to the weighted SVM approach (Chang and Lin 2011):

$$\begin{aligned} \text{obj} = \min_{\boldsymbol{w},b,\xi_{i,t}} & C\sum_{i=1}^{m}\sum_{t=1}^{n_i}(a_{i,t} * \xi_{i,t}) + 0.5\mathbf{w}^T\mathbf{w} \\ \text{s.t. } & z_i x_{i,t} \geq 1 - \xi_{i,t}, \xi_{i,t} \geq 0, \\ & i = 1, 2, \ldots, m, t = 1, 2, \ldots, n_i \end{aligned} \tag{8}$$

(See Appendix A for details.)

Here, C is a tuning parameter that controls the relative importance of minimizing the weighted misclassification errors $\sum_{i=1}^{m}\sum_{t=1}^{n_i}(a_{i,t} * \xi_{i,t})$ and the regularization term $\mathbf{w}^T\mathbf{w}$. In practice, the value of C can be chosen by cross validation and the weighted SVM problem can be solved efficiently by many algorithms, such as the gradient descent methods (Nocedal and Wright 1999).

The two failure mode scenarios discussed above can also be extended to the multiple failure mode scenarios (see Hsu and Lin 2002 for more details). For example, one possibility is to design multiple FM-INDICES and apply techniques like the one versus rest or one versus one strategies (Bishop 2006).

3.1.4. Failure mode probability model

Here, we discuss how to leverage the proposed FM-INDEX, $x_{i,t}$, to estimate the probability of the failure mode. In the existing literature, it is a common practice to use the sigmoid function to transform the outputs of a classification model to a probability function. The technique is called Platt scaling, which was proposed by John Platt (2000) in the context of support vector machines. Currently this technique is available in a lot of computer software packages (e.g., MATLAB). Here, we adopt the sigmoid function $g(x_{i,t}) = \frac{1}{1+e^{Ax_{i,t}+B}}$ in Platt (2000) to develop the failure mode (FM) probability model:

$$P(Z_{i,t} = k|x_{i,t}) = \begin{cases} \frac{1}{1+e^{Ax_{i,t}+B}}, \text{for } k = 1 \\ \frac{e^{Ax_{i,t}+B}}{1+e^{Ax_{i,t}+B}}, \text{for } k = -1 \end{cases}, \tag{9}$$

where $Z_{i,t}$ denotes the predicted failure mode for unit i at time t; k is the failure mode label; and the parameters A and B can be estimated by the maximum likelihood estimation approach. Without loss of generality, this proposed FM probability model in Eq. [9] can be also extended to multiple failure modes (e.g., see Hahn (2006)).

3.1.5. Real-time failure mode probability estimation

Previously, we developed an FM-INDEX and a probability estimation of the failure mode for a specific observation. Here, we show how to calculate the failure mode probability for a testing unit based on its available observations.

In the reliability-based approach, survival models are often used to find the failure time; competing risk analysis is then performed to infer the failure mode probability Ma and Krings (2008). However, this approach often focuses on estimating the remaining life distribution (RLD) for the entire population instead of each individual unit still operating in the field. Furthermore, the derived FM probability estimation depends on the selected failure time window. In fact, it is possible that the estimated failure times of two units from different failure modes exhibit similar results, whereas the in situ sensor data of these two units already present different characteristics. Therefore, in this article, we instead choose a different path that predicts the FM probability of a unit in real time directly based on its in situ sensor measurements. Specifically, we first calculate the conditional probability $P(Z_{i,t} = k|x_{i,t})$ in Eq. [9] for unit i at time t. Then we integrate the probability estimations up to the current time n_i as shown below:

$$P(Z_i = k|\boldsymbol{L}_{i,..,}) = P(Z_i = k|\boldsymbol{x}_{i,.}) = \sum_{t=1}^{n_i} c_{i,t} P(Z_{i,t} = k|x_{i,t}). \tag{10}$$

Here, $c_{i,t}$ is the weight coefficient of the probability estimation $P(Z_{i,t} = k|x_{i,t})$ for unit i at time t. Similar to the loss weights $\{a_{i,t}\}$ in Section 3.1.2, we propose to assign higher weights to the most recent observations:

$$c_{i,t} \geq c_{i,t-1} \geq 0, \quad t = 1, 2, \ldots, n_i - 1. \tag{11}$$

In addition, according to the law of probability the following constraint should be satisfied:

$$\sum_{t=1}^{n_i} c_{i,t} = 1, \quad t = 1, 2, \ldots, n_i, \tag{12}$$

Similar to Eq. [6], if we assume $c_{i,t}$ follows an arithmetic series ($2c_{i,t} = c_{i,t+1} + c_{i,t-1}$), then we can determine the value of $c_{i,t}$ using the following arithmetic series:

$$c_{i,t} = c_{i,1} + (t-1)\frac{2 - 2c_{i,1}n_i}{(n_i - 1)n_i}, \quad t = 1, \ldots, n_i. \tag{13}$$

Considering that the fault diagnostic result may not be reliable at the early stage of degradation, we can further apply the idea of a moving window to the arithmetic series in Eq. [13]. Then, the value of $c_{i,t}$ is calculated by:

$$c_{i,t} = \begin{cases} c_{i,(n_i-r+1)} + (t - n_i + r - 1)\frac{2-2rc_{i,(n_i-r+1)}}{(r-1)r}, & t = (n_i - r + 1), \ldots, n_i \\ 0, & t = 1, \ldots, (n_i - r) \end{cases}, \tag{14}$$

where r is the moving window size.

3.1.6. Flow chart of the proposed data-level fusion model for failure mode diagnosis

Figure 4 illustrates the flow chart of the proposed failure mode diagnostic approach. To summarize, we first construct a composite index, named FM-INDEX, that characterizes the failure mode of a unit based on its multiple sensor data. Then, with the constructed FM-INDEX, the FM probability model in Eq. [9] produces a probability estimation of the failure mode for each observation. Finally, we combine the probability estimations from the available observations of a unit to produce a single failure mode probability estimation for the unit as shown in Eq. [10].

3.2. Degradation modeling and health index construction under multiple failure modes

From the motivating example in Figure 3, we can see that different sensors may contain partial and dependent information about the same unit. In such cases, data fusion methods that take advantage of the sensor data dependencies to better characterize the condition of a unit are often used. For example, Liu and Huang (2016) proposed to construct a health index via the combination of multiple sensor data when there is only one failure mode. They identified two important properties that the health index should possess to ensure successful degradation modeling and prognostics:

> Property 1. Given the same failure mode, the variance in the failure threshold of the developed health index should be minimal.
>
> Property 2. Given the same failure mode, the model fitting errors for the developed health index should be minimal.

In particular, Liu and Huang (2016) proposed a linear fusion model to construct a health index, $h_{i,t} = \mathbf{L}_{i,.,t}\mathbf{w}_h$. Here, $h_{i,t}$ is the developed health index for unit i at time t, and $\mathbf{w}_h$ is the fusion coefficients for combining multiple sensor data $\mathbf{L}_{i,.,t}$. By optimizing the identified properties, Liu and Huang (2016) further showed that the optimal weight $\mathbf{w}_h$ can be efficiently solved via

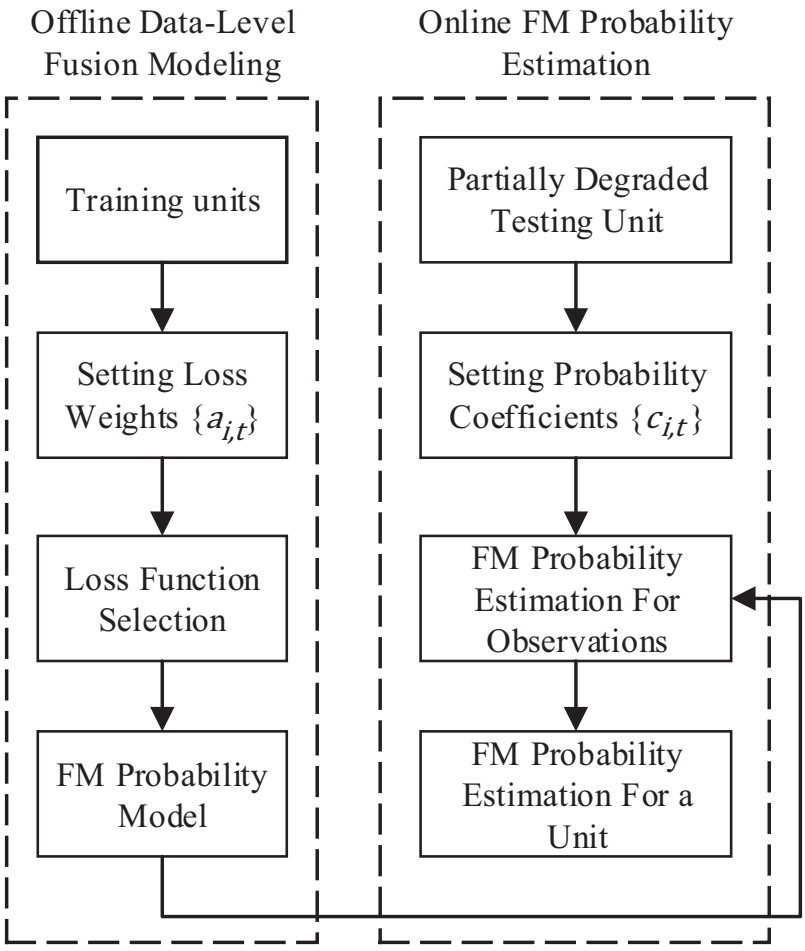

Figure 4. Flow chart of the proposed data-level fusion model for failure mode diagnosis.

quadratic programming. While such a data-level fusion model has demonstrated a significant improvement for degradation modeling and prognostics, it is limited by the single failure mode assumption. To address this issue, we propose to construct a conditional health index for each failure mode k (i.e., $h_{i,t}^{(k)} = \mathbf{L}_{i,.,t}\mathbf{w}_h^{(k)}$), which can be further integrated with the developed failure mode diagnostic procedure. Specifically, $\mathbf{w}_h^{(k)}$ can be derived by using the proposed method in Liu and Huang (2016), based only on the training units that fail due to failure mode k. Our goal is for the constructed health index to better characterize the degradation status of a unit when compared to relying on any original sensor data, thus leading to an improved remaining life prediction.

To model the evolution of the degradation process, one of the commonly used methods is the general mixed-effect degradation model proposed by Lu and Meeker (1993):

$$h_t = \eta(\boldsymbol{\varphi}, \boldsymbol{\theta}, t) + \varepsilon_t, \tag{15}$$

where $\eta(\cdot)$ is the parametric form of the degradation model; h_t is the health index that measures the underlying degradation status at time t; $\boldsymbol{\varphi}$ is a vector of fixed-effect parameters that represent common characteristics of the population; $\boldsymbol{\theta}$ is a vector of random-effect parameters that characterize the unit-to-unit variability; and ε_t is an error term that represents the measurement noises. Depending on the parametric form of $\eta(\cdot)$, this degradation model can be used to describe different functional forms according to the evolution of the conditional health index. Without loss of generality, here we focus on the pth order polynomial degradation model as a demonstration:

$$h_{i,t}^{(k)} = \sum_{\alpha=0}^{p} \theta_{i,\alpha}^{(k)} t^{\alpha} + \varepsilon_{i,t}^{(k)} = \boldsymbol{\Gamma}_t \boldsymbol{\theta}_i^{(k)} + \varepsilon_{i,t}^{(k)} \tag{16}$$

where p is the order of the polynomial model; $h_{i,t}^{(k)}$ is the constructed conditional health index for unit i at time t given failure mode k; $\boldsymbol{\theta}_i^{(k)} = [\theta_{i,0}^{(k)}, \ldots, \theta_{i,p}^{(k)}]^T$ is the random-effect parameters for unit i given that it degrades due to failure mode k and is often assumed to follow a multivariate normal distribution, $\boldsymbol{\theta}_i^{(k)} \sim N_{p+1}(\mathbf{u}_k^0, \boldsymbol{\Sigma}_k^0)$; $\varepsilon_{i,t}^{(k)}$ is the random noise and follows $N(0, \sigma_k^2)$; and $\boldsymbol{\Gamma}_t = [1, t, \ldots, t^p]$. Here, we focus on the polynomial degradation model for two reasons. First, many degradation models discussed in the existing literature can be transformed into Eq. [16, such as the exponential form models (Gebraeel et al. 2005; Liu, Gebraeel, and Shi 2013;

Yan et al. 2016) and the random coefficient-based models (Lu, Park, and Yang 1997; Wang 2000). Second, we implement the Bayesian approach that leverages the in situ data from multiple sensors of an operating unit to online update the parameters of the polynomial model. In this way, the updated model becomes better suited to the unique degradation characteristics of the operating unit to ensure good extrapolation performance for prognostics.

3.3. Estimation of the remaining life distribution (RLD)

In this section we propose a prognostic method to predict the remaining lifetime of a partially degraded unit in real time when there are multiple failure modes. Specifically, this method integrates the Bayesian updating approach and the failure mode diagnostic procedure proposed in Section 3.1.

For prognostic analysis, a common assumption is that a unit fails once its degradation signal crosses a predefined failure threshold. Consequently, a fundamental requirement here is to accurately predict the future evolution of the degradation signal for a unit up to the failure threshold. However, the central challenge is that the failure mode may affect the degradation path and thus the failure time of a unit. To address this issue, we propose to first offline estimate the values of $\mathbf{u}_k^0$, $\mathbf{\Sigma}_k^0$, and σ_k^2 in Eq. [16] based on the training units that fail due to failure mode k. Then, based on the in situ data $\boldsymbol{L}_{i,..}$ collected from unit i, we calculate the conditional health index $\mathbf{h}_{i,.}^{(k)}$ and further calibrate the random-effect parameters $\boldsymbol{\theta}_i^{(k)}$ by using the Bayesian updating approach in Gebraeel (2006), conditional on unit i failing due to failure mode $Z_i = k$. Specifically, for $\boldsymbol{\theta}_i^{(k)} \sim N_{p+1}(\mathbf{u}_k^0, \mathbf{\Sigma}_k^0)$, we can show that the conditional posterior distribution of $\boldsymbol{\theta}_i^{(k)}$ still follows a multivariate normal distribution given the constructed conditional health index up to the current observation time n_i, $\mathbf{h}_{i,.}^{(k)} = [h_{i,1}^{(k)}, \ldots, h_{i,n_i}^{(k)}]' \in \mathbb{R}^{n_i \times 1}$:

$$\boldsymbol{\theta}_i^{(k)} | \mathbf{h}_{i,.}^{(k)} \sim N_{p+1}\left(\mathbf{u}_k^1, \mathbf{\Sigma}_k^1\right) \tag{17}$$

where

$$\mathbf{u}_k^1 = \left(\frac{\mathbf{\Psi}'_i \mathbf{\Psi}_i}{\sigma_k^2} + (\mathbf{\Sigma}_k^0)^{-1}\right)^{-1} \left(\frac{\mathbf{\Psi}'_i \boldsymbol{h}_{i,.}^{(k)}}{\sigma_k^2} + \left(\mathbf{\Sigma}_k^0\right)^{-1} \mathbf{u}_k^0\right),$$

$$\mathbf{\Sigma}_k^1 = \left(\frac{\mathbf{\Psi}'_i \mathbf{\Psi}_i}{\sigma_k^2} + (\mathbf{\Sigma}_k^0)^{-1}\right)^{-1},$$

and

$$\mathbf{\Psi}_i \in R^{n_i \times (p+1)} = \begin{bmatrix} 1 & \ldots & 1 \\ \ldots & \ldots & \ldots \\ 1 & \ldots & t^p \\ \ldots & \ldots & \ldots \\ 1 & \ldots & n_i^p \end{bmatrix}$$

(see Appendix B for details).

Here, we consider the normal distribution for the random-effect parameters since it leads to a convenient and analytical solution when updating the random-effect parameters in real time. In fact, this assumption has been widely adopted in the existing literature (Gebraeel et al. 2005; Liu, Chehade and Song 2015; Liu and Huang 2016; Yan et al. 2016). Yet, it is also possible to use other distributions or even some simulation-based approaches to compute the posterior distribution and remaining lifetime. In such cases, our proposed ideas for estimating the failure mode can still be applied. Thus, to highlight our main ideas, we focus on only the normal distribution for the random-effect parameters in the remainder of the article.

Based on the updated random-effect parameters $\boldsymbol{\theta}_i^{(k)} | \mathbf{h}_{i,.}^{(k)}$, we derive the RLD. In particular, since $\tilde{h}_{i,n_i+t}^{(k)} = \theta_{i,0}^{(k)} + \cdots + \theta_{i,p}^{(k)} * (n_i + t)^p + \varepsilon_{i,n_i+t}^{(k)}$ follows a normal distribution with mean $\tilde{u}_{i,n_i+t,k} = [1, \ldots, (n_i + t)^p]\mathbf{u}_k^1$ and variance $\tilde{\sigma}_{i,n_i+t,k}^2 = [1, \ldots, (n_i + t)^p]\mathbf{\Sigma}_k^1[1, \ldots, (n_i + t)^p]' + \sigma_k^2$, we can calculate the conditional cumulative distribution function (cdf) for the remaining lifetime $\tilde{T}_i$ based on the conditional health index $\mathbf{h}_{i,.}^{(k)}$ given that unit i fails due to failure mode $Z_i = k$:

$$P(\tilde{T}_i \le t | \mathbf{L}_{i,..}, Z_i = k) = P\left(\tilde{T}_i \le t | \mathbf{h}_{i,.}^{(k)}\right)$$

$$= P\left(\tilde{h}_{i,n_i+t}^{(k)} \ge D_h^{(k)} | \mathbf{h}_{i,.}^{(k)}\right) = \Phi\left(\frac{\tilde{u}_{i,n_i+t,k} - u_{h,k}^d}{\sqrt{\tilde{\sigma}_{i,n_i+t,k}^2 + v_{h,k}^d}}\right), \tag{18}$$

where $\Phi(\cdot)$ is the standard normal cdf and $D_h^{(k)}$ is the failure threshold for the conditional health index due to failure mode k. It is assumed to follow a normal distribution with mean $u_{h,k}^d$ and variance $v_{h,k}^d$, which can be estimated by the sample mean and the sample variance of the last observations before failure for training units that fail due to failure mode k.

In practice, the failure mode for a unit is often unknown a priori. Thus, we integrate Eq. [18] with the FM probability estimation $P(Z_i = k | \mathbf{L}_{i,..})$ in Section 3.1.5 to predict the RLD $F_{\tilde{T}_i | \boldsymbol{L}_{i,..}}(t)$ given the collected *in situ* sensor data $\mathbf{L}_{i,..}$:

$$F_{\tilde{T}_i|\mathbf{L}_{i,..}}(t) = P(\tilde{T}_i \le t|\mathbf{L}_{i,..}) = \sum_{k\in K}\{P(\tilde{T}_i \le t|\mathbf{L}_{i,..}, Z_i = k) * P(Z_i = k|\mathbf{L}_{i,..})\} = \sum_{k\in K}\left\{P\left(\tilde{h}^{(k)}_{i,n_i+t} \ge D^{(k)}_h|\mathbf{h}^{(k)}_{i,.}\right) * P(Z_i = k|\mathbf{x}_{i,.})\right\}, \quad (19)$$

where the set K includes all the potential failure modes. Since the remaining lifetime for a testing unit should be greater than zero, we further consider the truncated cdf conditioning on $\tilde{T}_i \ge 0$:

$$P(\tilde{T}_i \le t|\mathbf{L}_{i,..}, \tilde{T}_i \ge 0) = \frac{P(0 \le \tilde{T}_i \le t|\mathbf{L}_{i,..})}{P(\tilde{T}_i \ge 0|\mathbf{L}_{i,..})} = \frac{\sum_{k\in K}\{P(Z_i = k|\mathbf{L}_{i,..}) * (P(\tilde{T}_i \le t|\mathbf{L}_{i,..}, Z_i = k) - P(\tilde{T}_i \le 0|\mathbf{L}_{i,..}, Z_i = k))\}}{1 - \sum_{k\in K}\{P(Z_i = k|\mathbf{L}_{i,..}) * P(\tilde{T}_i \le 0|\mathbf{L}_{i,..}, Z_i = k)\}}. \quad (20)$$

Here, because the truncated cdf in Eq. [20] is skewed, we can use the median as the point estimator for the remaining lifetime. Numerically, this is equivalent to finding the observation time t such that $P(\tilde{T}_i \le t|\mathbf{L}_{i,..}, \tilde{T}_i \ge 0) = 0.5$.

4. Case study

In this section, we evaluate our proposed data-level fusion method for failure mode diagnosis and prognostic analysis based on the data set introduced in Section 2. We also consider the following two benchmark methods: (1) the data-level fusion method in Liu and Huang (2016) introduced in Section 3.2, which ignores the effect of multiple failure modes and (2) each of the original sensor data with the consideration of the effect of multiple failure modes.

4.1. Data preprocessing and sensor selection

Before implementing the proposed data fusion model, we must first determine which sensors to use. To ensure a fair comparison with the existing literature, we adopt a similar sensor-selection criterion as in Liu and Huang (2016). The criterion selects a sensor if it shows a consistently increasing or decreasing degradation trend among all the training units that fail due to the same failure mode. Accordingly, the following 14 sensors (i.e., $s = 14$) from Table 1 are selected: T24, T30, T50, P30, Nf, Nc, Ps30, phi, NRf, NRc, BPR, htBleed, W31, and W32. Given that these sensors measure different characteristics of the engine with different measurement units (e.g., psia, rpm, lbm/s), we further take a log transformation and then standardize the data set before the data fusion as in Liu and Huang (2016).

4.2. Failure mode diagnosis

In this subsection, we demonstrate how to construct the proposed FM-INDEX. We also conduct a sensitivity analysis on the weights of the constructed FM-INDEX. We then evaluate and compare the failure mode diagnostic accuracy by using the FM-INDEX and each selected sensor data. Finally, we illustrate how to estimate the probability of the failure mode for a testing unit in real time.

4.2.1. Parameter setup

In this case study, we choose the value of the loss weight $\{a_{i,t}\}$ by assuming it follows the arithmetic series in Eq. [6]. In addition, we select the hinge-loss function in Eq. [7] with the regularization function $R(f) = f^2 = \mathbf{w}^T\mathbf{w}$. By solving the optimization problem in Eq. [8], the optimal weights $\mathbf{w}^*$ and the intercept term b^* for constructing the FM-INDEX are obtained and shown in Table 2.

Recall that the two failure modes occur either at the HPC or at the fan. From Table 2, we can see that the sensors that directly monitor the HPC or the fan tend to exhibit higher weights, which is consistent with our intuition. For example, the sensors with the largest four weights are phi, P30, BPR, and NRc. According to the description in Table 1, these sensors directly measure the characteristics of the HPC or the fan.

4.2.2. Sensitivity study

In this subsection, we conduct a sensitivity analysis on the optimal weights $\mathbf{w}^*$ of the FM-INDEX. Specifically, we consider the missing data scenario by randomly hiding certain percentage of the observations. We then solve the optimization problem in Eq. [8] to study the sensitivity

Table 2. The optimal weights and the intercept term for combining the selected sensor data when constructing the FM-INDEX.

Name	T24	T30	T50	P30	Nf	Nc	Ps30	phi
Value	−0.2757	−0.2347	−0.4254	2.3447	−1.7481	−1.5456	−0.7011	3.1123
Name	NRf	NRc	BPR	htBleed	W31	W32	intercept	
Value	−1.6992	−1.8486	−2.213	−0.2213	0.4162	0.4461	−2.1222	

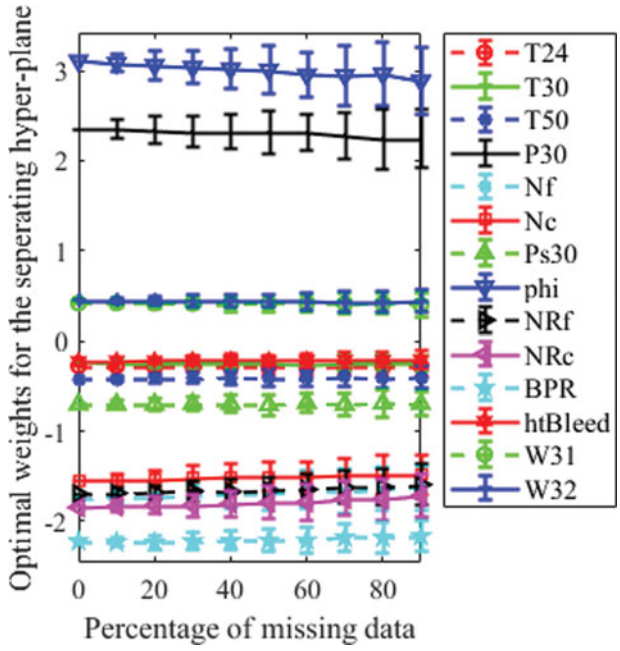

Figure 5. The optimal weights learned at different degrees of missing data with their 95 percent confidence intervals. Note that the optimal weights for 0% missing data are given in Table 2.

of the derived optimal weights $\mathbf{w}^*$ at different degrees of missing data.

Figure 5 shows the changes in the optimal weights $\mathbf{w}^*$ based on different degrees of missing data with corresponding 95 percent confidence intervals. For example, the label "30%" refers to the case where 30 percent of the training data is randomly removed prior to learning the optimal weights. Here the confidence intervals are obtained by simulating 100 replications with the same percentage of missing data.

From Figure 5, we can see that the optimal weights for constructing the FM-INDEX are relatively stable at different degrees of missing data, especially when the missing amount of missing data is less than or equal to 20 percent. This indicates that our proposed data-level fusion model allows for certain amounts of missing data without great expense to the accuracy of the learned optimal weights.

4.2.3. FM-INDEX demonstration

Figure 6 plots the degradation signals of each selected sensor data and the constructed FM-INDEX for all 100 training units. The light gray and dark gray show the units that belong to failure mode 1 and −1, respectively. The overlapping areas, which are highlighted in black, represent the undistinguished regions of the two failure modes.

From Figure 6, we can see that the FM-INDEX better separates the units from these two failure modes than each selected sensor data did. This is mainly because the FM-INDEX leverages the dependent information from multiple sensor data; thus, it is more informative than the individual sensor data for failure mode diagnosis. In addition, the overlapping areas become much smaller (i.e., failure mode diagnostic results become more accurate) as the units approach failure. This observation

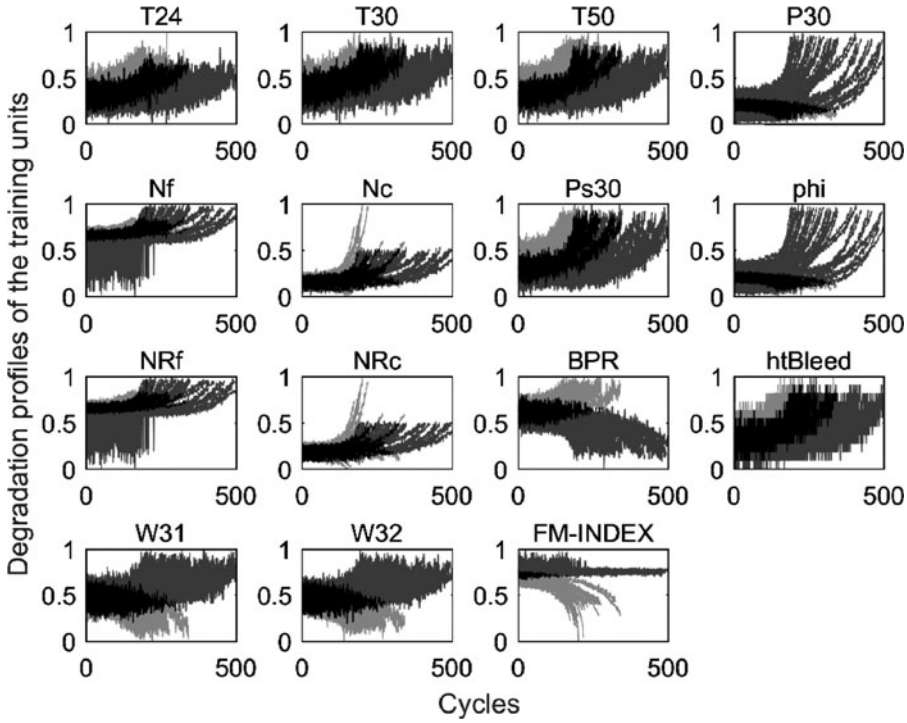

Figure 6. An illustration of the degradation signals of each selected sensor data and the constructed FM-INDEX for the 100 training units. The light gray refers to failure mode 1, the dark gray refers to failure mode −1, and the black refers to the overlap.

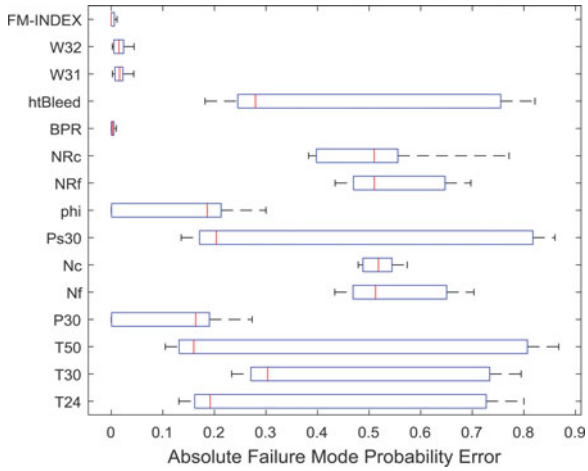

Figure 7. The boxplots of the absolute errors in the FM probability estimation over the training units by using a leave-one-out cross-validation.

coincides with our model constraints, in which we consider a nondecreasing series in the key parameter $\{a_{i,t}\}$ when constructing the FM-INDEX.

Furthermore, since the true failure modes of the training units are known, we can evaluate the accuracy of the FM probability estimation. Specifically, we define the absolute error in the FM probability estimation for training unit i as:

$$p_i^{(err)} = |1 - P(Z_i = z_i \,|\mathbf{L}_{i,..,})|. \tag{21}$$

The comparison results by using our proposed FM-INDEX and each selected sensor data as shown in Figure 7, which provides boxplots of the absolute errors in the FM probability estimation over the training units, using a leave-one-out cross-validation technique. The result shows that the FM-INDEX has a lower median FM probability error than each individual sensor data.

4.2.4. Real-time failure mode probability estimation

In this section, we illustrate the failure mode probability estimation for a unit as more data are collected in real time by using the proposed method in Section 3.1.5. Here, the coefficients $\{c_{i,t}\}$ are chosen by following the arithmetic series in Eq. [14] with a moving window size $r = 30$. Figure 8 shows the changes of the probability estimation $P(Z_i = -1|\mathbf{L}_{i,..,})$ for unit i that is known to degrade due to failure mode -1.

From Figure 8, we can see that as more observations are collected, the probability estimation of the failure mode improves and $P(Z_i = -1|\mathbf{L}_{i,..,})$ approaches one. Furthermore, since it is difficult to identify the failure mode at an early stage of degradation, the figure shows fluctuations in the probability estimation at the beginning.

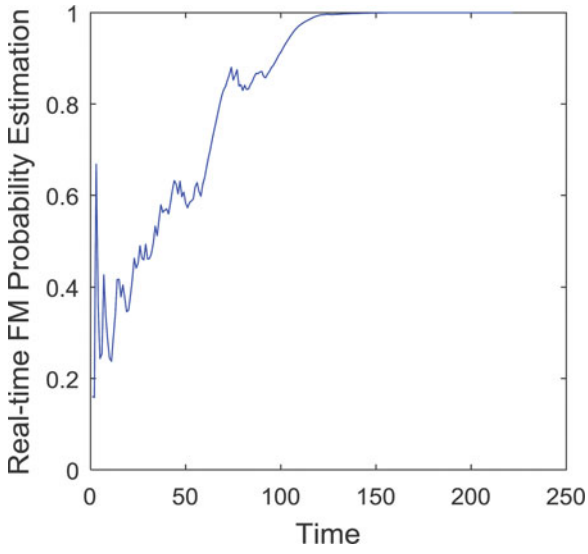

Figure 8. An illustration of the probability estimation $P(Z_i = -1|\mathbf{L}_{i,..,})$ for a validation unit that fails due to failure mode -1 as more data are collected in real time.

4.3. Estimation of the RLD

We estimate the remaining lifetime of the testing units by implementing the proposed method in Section 3.3. Recall that we assume the random-effect parameters follow a multivariate normal distribution. In this case study, we conducted the Henze-Zirkler test to check the normality assumption based on the training data set for each failure mode. For the constructed conditional health index due to failure mode 1, we achieved a p-value of 0.49, and for the constructed conditional health index due to failure mode 2, we achieved a p-value of 0.15. Thus, these results further validate that it is satisfactory to adopt the normal distribution for the random-effect parameters.

To evaluate the prognostic performance of the proposed and benchmark methods, we consider the error metric in Eq. [22], err_i, which is defined as the relative error between the estimated and the actual remaining lifetime for unit i.

$$err_i = \frac{(n_i + \tilde{T}_i) - (n_i + T_i)}{n_i + T_i} = \frac{\tilde{T}_i - T_i}{n_i + T_i}, \tag{22}$$

where n_i is the time index of the last observation for testing unit i when it stops further usage prior to failure; T_i is the actual remaining lifetime for unit i; and $\tilde{T}_i$ is the estimated remaining lifetime for unit i.

Figure 9 shows the comparison results among the absolute value of the mean percentage error (%) for the proposed method, the single HI, and the best performing single sensor. In this example, the "100" label refers to those scenarios where only the testing units with an actual remaining lifetime equal to or less than 100 are taken into consideration. For the single sensors, we estimate the failure mode and remaining lifetime given only the data

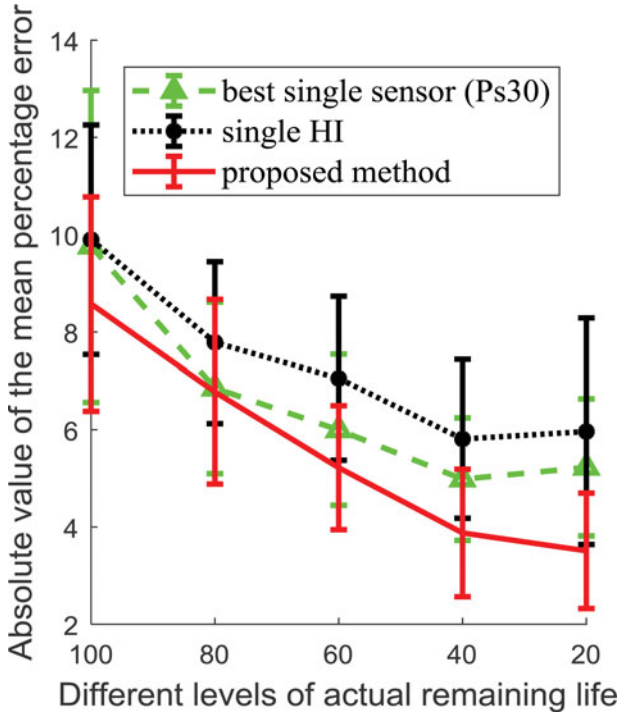

Figure 9. Comparison of results of the absolute value of the mean percentage error (%) for the proposed method versus the other two benchmark methods. The bars correspond to the 95 percent confidence intervals.

from the sensor of interest (i.e., ignoring the data from the other sensors). Based on the plots of the sensor data in Figure 6, we can see that the degradation signals exhibit an exponential functional form. Thus, in the benchmark methods, we consider the following degradation model for the log-transformed sensor measurement $L_{i,j,t}$ given the failure mode k:

$$L_{i,j,t} = \theta_{i,j,0}^{(k)} + \theta_{i,j,1}^{(k)} t + \theta_{i,j,2}^{(k)} t^2 + \varepsilon_{i,j,t}^{(k)}, \tag{23}$$

where $\varepsilon_{i,j,t}^{(k)}$ is the error term and is often assumed to follow the normal distribution, which can be tested based on the fitted-degradation model. For the single HI, we adopt the method by Liu and Huang (2016) to estimate the remaining lifetime, which ignores the presence of multiple failure modes. In addition, to achieve a fair comparison the random failure threshold approach in Eq. [18] is considered for all three methods.

Based on Figure 9, we can see that: (a) our proposed approach outperformed the other two benchmark methods; (b) in general, the predictions become more accurate as the unit approaches failure in our proposed approach; and (c) the best single sensor (Ps30) outperforms the single HI constructed by the model in Liu and Huang (2016). Observation (a) mainly results from two reasons. First, our proposed data fusion method takes advantage of data from multiple dependent sensors, thus leading to a better remaining lifetime prediction. Second, our proposed approach considers the distinct influence of different failure modes on the life cycle path of a unit. One possible reason for Observation (b) is that less actual remaining lifetime typically indicates that more observational data have been collected, and thus we are more confident about the predicted failure mode diagnosis and the constructed degradation model. In addition, we choose both the loss weights $\{a_{i,t}\}$ and the coefficients $\{c_{i,t}\}$ to linearly increase with time; thus more weights are given to the observations closer to failure. However, because only partial degradation signals are provided for the testing units, different sets of units are used here to evaluate the prediction error at different levels of the actual remaining lifetime. As a result, there is no guarantee of having a monotonically decreasing trend in Figure 9. Finally, the main reason behind Observation (c) is that the single HI does not consider the presence of two failure modes by assuming all units fail due to the same failure mode—thus it leads to poor prediction performance. This result further demonstrates the importance of our proposed method for online prediction of the failure mode when estimating the RLD.

5. Discussion and conclusion

With the rapid development of condition monitoring and sensor techniques, multiple sensors have been widely used to simultaneously monitor the degradation status of a unit. In such a big data environment, each sensor may contain only partial and dependent information about the health status of a unit, and sensors may exhibit distinct degradation characteristics due to different failure modes.

The main contribution of this article is to address the challenges of degradation modeling and prognostics in big data environments for systems with multiple sensors as well as multiple failure modes. To do so, we first constructed an FM-INDEX that effectively and accurately diagnosed the failure mode for a unit during condition monitoring, which was tailored to the needs of degradation modeling and prognostic analysis. Next, we extended the existing data-level fusion techniques for prognostic analysis to multiple failure mode scenarios by integrating the failure mode diagnosis in the remaining life estimation. The developed methodology was tested and validated using multiple sensor signals from aircraft gas turbine engines that contain two potential failure modes (Saxena et al. 2008b). Our case study showed that the developed FM-INDEX better distinguishes the units from the two failure modes than each original sensor data did, that failure mode diagnostic results become more accurate as a unit approaches failure, and that

the remaining life prediction of the proposed method outperforms the related benchmarks.

There are several important topics for future research. First, it is worth exploring the performance of the developed FM-INDEX by using nonlinear fusion functions (e.g., via kernel methods). Second, we derive the FM-INDEX in this article based on historical records offline via a frequentist approach. For future research, it is worth investigating how to online update the FM-INDEX via a Bayesian approach.

Nomenclature

$L_{i,j,t}$	the sensor measurements of sensor j for unit i at time t
$\mathbf{L}_{i,.,t}$	the sensor measurements of all the sensors for unit i at time t
$\mathbb{R}$	the set of real numbers
$V(Z,z)$	the loss function that quantifies the deviation of the predicted label Z from the true label z
z_i	the true failure-mode label for unit i
n_i	the number of available observations for unit i; also referred to as the time index of the last observation
$f(.)$	the proposed fusion model for constructing the FM-INDEX
$x_{i,t}$	the FM-INDEX for unit i at time t
$R(f)$	the regularization function that prevents overfitting of the fusion model $f(.)$
λ	the regularization parameter
$a_{i,t}$	the loss weight for unit i at time t
M_k	the set of training units that fail due to failure mode k
c_k	the sum of the loss weights for a training unit that belongs to M_k
c	the sum of the loss weights for the set of training units M_k
$\xi_{i,t}$	the misclassification error for unit i at time t
$\mathbf{w}$	the weights that combine multiple sensor data to construct the FM-INDEX
b	the intercept term of the constructed fusion model $f(.)$
C	the tuning parameter that controls the importance of minimizing the misclassification errors
$Z_{i,t}$	the predicted failure mode for unit i at time t
$P(Z_{i,t}=k\|\mathbf{L}_{i,.,t})$	the probability that unit i at time t degrades due to failure mode k given the sensor data $\mathbf{L}_{i,.,t}$ for unit i at time t
A, B	the slope and intercept terms of the logistic regression, respectively
Z_i	the predicted failure mode for unit i
$P(Z_i=k\|\mathbf{L}_{i,.,.})$	the probability that unit i degrades due to failure mode k given the available in situ data $\mathbf{L}_{i,.,.}$ for unit i
$c_{i,t}$	the weight coefficient of the probability estimation $P(Z_{i,t}=k\|\mathbf{L}_{i,.,t})$ for unit i at time t
r	the moving window size
$\eta(.)$	the functional form of the degradation model
$\boldsymbol{\varphi}$	a vector of fixed-effect parameters that represents the common characteristics for the population
$\boldsymbol{\theta}$	a vector of random-effect parameters that characterize the unit-to-unit variability
ε_t	an error term that represents the measurement noises
$\boldsymbol{\Gamma}_t$	a vector of time-dependent covariates
p	the order of the polynomial model
$\varepsilon^{(k)}_{i,j,t}$	the random noise in the measurements of sensor j at time t given that unit i is degrading due to failure mode k
$h^{(k)}_{i,t}$	the conditional health index for unit i at time t given that unit i degrades due to failure mode k
$\mathbf{w}^{(k)}_h$	the fusion coefficients for combining multiple sensor data for the conditional health index due to failure mode k
$\mathbf{u}^0_k, \mathbf{u}^1_k$	the prior and the posterior means of $\boldsymbol{\theta}_i$ given that unit i degrades due to failure mode k
$\boldsymbol{\Sigma}^0_k, \boldsymbol{\Sigma}^1_k$	the prior and the posterior variances of $\boldsymbol{\theta}_i$ given that unit i degrades due to failure mode k
$\boldsymbol{\Psi}_i$:	the design matrix for unit i
$\tilde{u}_{i,n_i+t,k}$	the mean of the projected conditional health index $\tilde{h}^{(k)}_{i,n_i+t}$
$\tilde{\sigma}^2_{i,n_i+t,k}$	the variance of the projected conditional health index $\tilde{h}^{(k)}_{i,n_i+t}$
$D^{(k)}_h$	the failure threshold for the conditional health index given that the unit will fail due to failure mode k
$u^d_{h,k}$	the mean of the failure threshold $D^{(k)}_h$

$v_{h,k}^{d}$ the variance of the failure threshold $D_h^{(k)}$

$p_i^{(err)}$ the absolute error in the failure-mode probability estimation for training unit i

err_i the relative error between the estimated and the actual remaining lifetime for unit i

About the authors

Dr. Chehade is an assistant professor in the Department of Industrial and Manufacturing Systems, University of Michigan–Dearborn. His email address is achehade@umich.edu.

Mr. Song is a graduate student in the Department of Industrial and Systems Engineering, University of Wisconsin–Madison. His email address is csong39@wisc.edu.

Dr. Liu is an assistant professor in the Department of Industrial and Manufacturing Systems, University of Wisconsin–Madison. His email address is kliu8@wisc.edu. He is a member of ASQ.

Dr. Saxena is a researcher at GE Global Research. His email address is a asaxena@ge.com.

Dr. Zhang is an associate professor in the Department of Industrial Engineering and Management, Peking University. His email address is xi.zhang@coe.pku.edu.cn. He is a member of ASQ.

Funding

The authors gratefully acknowledge the support provided by the National Science Foundation Grant CMMI-1435809, the Office of Naval Research under Grant N00014-17-1-2261, and the National Science Foundation of China Grants 71571003 and 71471005.

References

Ali, Y., R. Rahman, and R. Hamzah. 2014. Acoustic emission signal analysis and artificial intelligence techniques in machine condition monitoring and fault diagnosis: A review. *Journal Teknologi* 69 (2):121–26. doi: 10.11113/jt.v69.3121

Basir, O., and X. Yuan. 2007. Engine fault diagnosis based on multi-sensor information fusion using Dempster–Shafer evidence theory. *Information Fusion* 8 (4):379–86. doi:10.1016/j.inffus.2005.07.003.

Bishop, C. 2006. *Pattern recognition and machine learning*. New York City, NY: Springer.

Brotherton, T., P. Grabill, D. Wroblewski, R. Friend, B. Sotomayer, and J. Berry. 2002. A testbed for data fusion for engine diagnostics and prognostics. *IEEE Aerospace Conference Proceedings* 6:3029–42. doi: 10.1109/AERO.2002.1036145

Bunks, C., D. McCarthy, and T. Al-Ani. 2000. Condition-based maintenance of machines using hidden Markov models. *Mechanical Systems and Signal Processing* 14 (4):597–612. doi:10.1006/mssp.2000.1309.

Chang, C., and C. Lin. 2011. LIBSVM: A library for support vector machines. *ACM Transactions of Intelligent Systems and Technology* 2 (3):1–27. doi:10.1145/1961189.1961199.

De Ketelaere, B., M. Hubert, and E. Schmitt. 2015. Overview of PCA-based statistical process monitoring methods for time-dependent, high-dimensional data. *Journal of Quality Technology* 47 (4):318–35. doi:10.1080/00224065.2015.11918137.

Frederick, D., J. DeCastro, and J. Litt. 2007. *User's guide for the commercial modular aero-propulsion system simulation (C-MAPSS)*. Technical Memorandum No. TM-2007-215026, NASA, Washington, DC.

Freedman, D. 2009. *Statistical models: Theory and practice*. Cambridge, UK: Cambridge University Press.

Gebraeel, N., M. Lawley, R. Li, and J. Ryan. 2005. Residual-life distributions from component degradation signals: A Bayesian approach. *IIE Transactions* 37 (6):543–57. doi:10.1080/07408170590929018.

Gebraeel, N. 2006. Sensory-updated residual life distributions for components with exponential degradation patterns. *IEEE Transactions on Automation Science and Engineering* 3 (4):382–93. doi:10.1109/TASE.2006.876609.

Hahn, E. 2006. Link function selection in stochastic multicriteria decision making models. *European Journal of Operational Research* 172 (1):86–100. doi:10.1016/j.ejor.2004.09.041.

Hall, D., and J. Llinas. 1997. An introduction to multisensor data fusion. *Proceedings of the IEEE* 85 (1):6–23.

Hastie, T., R. Tibshirani, J. Friedman, and J. Franklin. 2005. The elements of statistical learning: Data mining, inference and prediction. *Mathematical Intelligencer* 27 (2): 83–85. doi:10.1007/BF02985802.

Heger, T., and M. Pandit. 2004. Optical wear assessment system for grinding tools. *Journal of Electronic Imaging* 13 (3): 450–61. doi:10.1117/1.1760757.

Hsu, C., and C. Lin. 2002. A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks* 13 (2):415–25. doi:10.1109/72.991427.

Isermann, R. 1997. Supervision, fault-detection and fault-diagnosis methods—An introduction. *Control Engineering Practice* 5 (5):639–52. doi:10.1016/S0967-0661(97)00046-4.

Jardine, A., D. Lin, and D. Banjevic. 2006. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing* 20 (7):1483–1510. doi:10.1016/j.ymssp.2005.09.012.

Jiang, W., K. Wang, and F. Tsung. 2012. A variable-selection-based multivariate EWMA chart for process monitoring and diagnosis. *Journal of Quality Technology* 44 (3):209–30. doi:10.1080/00224065.2012.11917896.

Kim, J., and C. Scott. 2010. L2 kernel classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32 (10):1822–31. doi:10.1109/TPAMI.2009.188.

Kobayashi, T., and D. Simon. 2007. Hybrid Kalman filter approach for aircraft engine in-flight diagnostics: Sensor fault detection case. *Journal of Engineering for Gas Turbines and Power* 129 (3):746–54. doi:10.1115/1.2718572.

Liu, K., A. Chehade, and C. Song. 2015. Optimize the signal quality of the composite health index via data fusion for degradation modeling and prognostic analysis. *IEEE Transactions on Automation Science and Engineering* 14 (3): 1504–14. doi:10.1109/TASE.2015.2446752.

Liu, K., N. Gebraeel, and J. Shi. 2013. A data-level fusion model for developing composite health indices for

degradation modeling and prognostic analysis. *IEEE Transactions on Automation Science and Engineering* 10 (3):652–64. doi:10.1109/TASE.2013.2250282.

Liu, K., and S. Huang. 2016. Integration of data fusion methodology and degradation modeling process to improve prognostics. *IEEE Transactions on Automation Science and Engineering* 13 (1):344–54. doi:10.1109/TASE.2014.2349733.

Liu, K., and J. Shi. 2015. Internet of Things (IoT)-enabled system informatics for service decision making: Achievements, trends, challenges, and opportunities. *IEEE Intelligent Systems* 30 (6):18–21.

Lu, C., and W. Meeker. 1993. Using degradation measures to estimate a time-to-failure distribution. *Technometrics* 35 (2):161–74. doi:10.1080/00401706.1993.10485038.

Lu, J., J. Park, and Q. Yang. 1997. Statistical inference of a time-to-failure distribution derived from linear degradation data. *Technometrics* 39 (4):391–400. doi:10.1080/00401706.1997.10485158.

Ma, Z., and A. W. Krings. 2008. Competing risks analysis of reliability, survivability, and prognostics and health management (PHM). *Aerospace Conference, 2008 IEEE*, 1–21. doi: 10.1109/AERO.2008.4526636

Mobley, R. 2002. *An introduction to predictive maintenance*. Amsterdam, Netherlands: Butterworth-Heinemann.

Nocedal, J., and S. Wright. 1999. *Numerical optimization*. New York City, NY: Springer.

Platt, J. 2000. Probabilistic outputs for support vector machines and comparison to regularized likelihood methods. In *Advances in Large Margin Classifiers*, eds. A. Smola, P. Bartlett, B. Scholkopf, and D. Schuurmans, 61–74. Cambridge, MA: MIT Press.

Pulcini, G. 2008. Repairable system analysis for bounded intensity functions and various operating conditions. *Journal of Quality Technology* 40 (1):78–96. doi:10.1080/00224065.2008.11917714.

Ramasso, E., and A. Saxena. 2014. Performance benchmarking and analysis of prognostic methods for CMAPSS datasets. *International Journal of Prognostics and Health Management* 5 (2):1–15.

Saha, B., K. Goebel, and J. Christophersen. 2009. Comparison of prognostic algorithms for estimating remaining useful life of batteries. *Transactions of the Institute of Measurement and Control* 31 (3–4):293–308. doi:10.1177/0142331208092030.

Salahshoor, K., M. Mosallaei, and M. Bayat. 2008. Centralized and decentralized process and sensor fault monitoring using data fusion based on adaptive extended Kalman filter algorithm. *Measurement* 41 (10):1059–76. doi:10.1016/j.measurement.2008.02.009.

Sarkar, S., X. Jin, and A. Ray. 2011. Data-driven fault detection in aircraft engines with noisy sensor measurements. *Journal of Engineering for Gas Turbines and Power* 133 (8):783–89. doi:10.1115/1.4002877.

Saxena, A., J. Celaya, E. Balaban, K. Goebel, B. Saha, S. Saha, and W. Schwabacher. 2008a. Metrics for evaluating performance of prognostic techniques. Paper presented at the 2008 International Conference on Prognostics and Health Management, Denver, CO, October 6. doi: 10.1109/PHM.2008.4711436

Saxena, A., K. Goebel, D. Simon, and N. Eklund. 2008b. Damage propagation modeling for aircraft engine run-to-failure simulation. Paper presented at the 2008 International Conference on Prognostics and Health Management, Denver, CO, October 6. doi: 10.1109/PHM.2008.4711414

Simani, S., C. Fantuzzi, and R. Patton. 2013. *Model-based fault diagnosis in dynamic systems using identification techniques*. London: Springer.

Volponi, A., T. Brotherton, R. Luppold, and D. Simon. 2004. *Development of an information fusion system for engine diagnostics and health management*. Cleveland, OH: NASA.

Wang, H., and J. Wang. 2000. Fault diagnosis theory: Method and application based on multisensor data fusion. *Journal of Testing and Evaluation* 28 (6):513–18. doi:10.1520/JTE12143J.

Wang, W. 2000. A model to determine the optimal critical level and the monitoring intervals in condition-based maintenance. *International Journal of Production Research* 38 (6):1425–36. doi:10.1080/002075400188933.

Yan, H., K. Liu, X. Zhang, and J. Shi. 2016. Multiple sensor data fusion for degradation modeling and prognostics under multiple operational conditions. *IEEE Transactions on Reliability* 65 (3):1416–26. doi:10.1109/TR.2016.2575449.

Ye, Z., L. Tang, and M. Xie. 2011. A burn-in scheme based on percentiles of the residual life. *Journal of Quality Technology* 43 (4):334–45. doi:10.1080/00224065.2011.11917868.

Appendix A

This appendix shows that our proposed data fusion model $\min_{\boldsymbol{w},b}\frac{1}{m}\sum_{i=1}^{m}\sum_{t=1}^{n_i}a_{i,t}V(f(\mathbf{L}_{i,.,t}),z_{i,t})+\lambda R(f)$ with the hinge-loss function $V(x_{i,t},z_{i,t})=\max(0,1-z_{i,t}x_{i,t})$, where $x_{i,t}=f(\mathbf{L}_{i,.,t})=\mathbf{L}_{i,.,t}\mathbf{w}+b$ and $R(f)=||f||^2=\mathbf{w}^T\mathbf{w}$, is equivalent to the weighted SVM approach: $\min_{\boldsymbol{w},b,\xi_{i,t}}C\sum_{i=1}^{m}\sum_{t=1}^{n_i}(a_{i,t}*\xi_{i,t})+0.5\mathbf{w}^T\mathbf{w}$, s.t. $z_{i,t}x_{i,t}\geq 1-\xi_{i,t}$ and $\xi_{i,t}\geq 0, i=1,2,\ldots,m, t=1,2,\ldots,n_i$.

Let $\xi_{i,t}=V(x_{i,t},z_{i,t})=\max(0,1-z_{i,t}x_{i,t})$. Then, $\xi_{i,t}\geq 0$ and $\xi_{i,t}\geq 1-z_{i,t}x_{i,t}$. Consequently, $\min_{\boldsymbol{w},b}\max(0,1-z_{i,t}x_{i,t})+\lambda\mathbf{w}^T\mathbf{w}$ is equivalent to $\min_{\mathbf{w},b,\xi_{i,t}}\xi_{i,t}+\lambda\mathbf{w}^T\mathbf{w}$, subject to that $\xi_{i,t}\geq 0$ and $\xi_{i,t}\geq 1-z_{i,t}x_{i,t}$.

In this way, $\min_{\mathbf{w},b}\frac{1}{m}\sum_{i=1}^{m}\sum_{t=1}^{n_i}(a_{i,t}*\max(0,1-z_{i,t}x_{i,t}))+\lambda\mathbf{w}^T\mathbf{w}$ is equivalent to $\min_{\mathbf{w},b,\xi_{i,t}}\frac{1}{m}\sum_{i=1}^{m}\sum_{t=1}^{n_i}(a_{i,t}*\xi_{i,t})+\lambda\mathbf{w}^T\mathbf{w}$, s.t. $z_{i,t}x_{i,t}\geq 1-\xi_{i,t}$, and $\xi_{i,t}\geq 0,\ i=1,2,\ldots,m, t=1,2,\ldots,n_i$. Finally, by scaling the objective function with a factor of $1/(2\lambda)$ and setting $C=1/(2m\lambda)$, we can rewrite the objective function as $\min_{\mathbf{w},b,\xi_{i,t}}C\sum_{i=1}^{m}\sum_{t=1}^{n_i}(a_{i,t}*\xi_{i,t})+0.5\mathbf{w}^T\mathbf{w}$.

Appendix B

This appendix shows that given the failure mode $Z_i=k$, the posterior distribution of $\boldsymbol{\theta}_i^{(k)}$ still follows a multivariate normal distribution with mean $\mathbf{u}_k^1=(\frac{\boldsymbol{\Psi}_i'\boldsymbol{\Psi}_i}{\sigma_k^2}+(\boldsymbol{\Sigma}_k^0)^{-1})^{-1}(\frac{\boldsymbol{\Psi}_i'\mathbf{h}_{i,.}^{(k)}}{\sigma_k^2}+(\boldsymbol{\Sigma}_k^0)^{-1}\mathbf{u}_k^0)$ and variance $\boldsymbol{\Sigma}_k^1=(\frac{\boldsymbol{\Psi}_i'\boldsymbol{\Psi}_i}{\sigma_k^2}+(\boldsymbol{\Sigma}_k^0)^{-1})^{-1}$.

Given the failure mode $Z_i=k$, the probability density function of $\boldsymbol{\theta}_i^{(k)}$ can be expressed as $P(\boldsymbol{\theta}_i^{(k)})\propto e^{-\frac{1}{2}(\boldsymbol{\theta}_i^{(k)}-\mathbf{u}_k^0)'(\boldsymbol{\Sigma}_k^0)^{-1}(\boldsymbol{\theta}_i^{(k)}-\mathbf{u}_k^0)}$. Thus,

$$P\left(\boldsymbol{\theta}_i^{(k)}|\mathbf{h}_{i,.}^{(k)}\right)\propto P\left(\mathbf{h}_{i,.}^{(k)}|\boldsymbol{\theta}_i^{(k)}\right)P\left(\boldsymbol{\theta}_i^{(k)}\right)\propto$$

$$e^{-\frac{1}{2\sigma_k^2}\left(\mathbf{h}_{i,.}^{(k)}-\boldsymbol{\Psi}_i\boldsymbol{\theta}_i^{(k)}\right)'\left(\mathbf{h}_{i,.}^{(k)}-\boldsymbol{\Psi}_i\boldsymbol{\theta}_i^{(k)}\right)-\frac{1}{2}\left(\boldsymbol{\theta}_i^{(k)}-\mathbf{u}_k^0\right)'(\boldsymbol{\Sigma}_k^0)^{-1}\left(\boldsymbol{\theta}_i^{(k)}-\mathbf{u}_k^0\right)} \propto$$
$$e^{\boldsymbol{\theta}_i^{(k)\prime}\left(\frac{\boldsymbol{\Psi}'_i\boldsymbol{\Psi}_i}{\sigma_k^2}+(\boldsymbol{\Sigma}_k^0)^{-1}\right)\boldsymbol{\theta}_i^{(k)}-2\left(\frac{\boldsymbol{h}_{i,.}^{(k)}\boldsymbol{\Psi}_i}{\sigma_k^2}+(\mathbf{u}_k^0)'(\boldsymbol{\Sigma}_k^0)^{-1}\right)\boldsymbol{\theta}_i^{(k)}}.$$

It is known that the Gaussian family is the conjugate prior for Gaussian distribution; thus the posterior also follows a multivariate normal distribution. As a result, $\boldsymbol{\theta}_i^{(k)}|\mathbf{h}_{i,.}^{(k)} \sim N_{p+1}(\mathbf{u}_k^1, \boldsymbol{\Sigma}_k^1)$ and $P(\boldsymbol{\theta}_i^{(k)}|\mathbf{h}_{i,.}^{(k)}) \propto e^{-\frac{1}{2}(\boldsymbol{\theta}_i^{(k)}-\mathbf{u}_k^1)'(\boldsymbol{\Sigma}_k^1)^{-1}(\boldsymbol{\theta}_i^{(k)}-\mathbf{u}_k^1)}$.

Comparing the above two equations for $P(\boldsymbol{\theta}_i^{(k)}|\mathbf{h}_{i,.}^{(k)})$, we obtain the results $(\boldsymbol{\Sigma}_k^1)^{-1} = (\frac{\boldsymbol{\Psi}'_i\boldsymbol{\Psi}_i}{\sigma_k^2} + (\boldsymbol{\Sigma}_k^0)^{-1})$ and $(\mathbf{u}_k^1)'(\boldsymbol{\Sigma}_k^1)^{-1} = (\frac{\mathbf{h}_{i,.}^{(k)}\boldsymbol{\Psi}_i}{\sigma_k^2} + (\mathbf{u}_k^0)'(\boldsymbol{\Sigma}_k^0)^{-1})$. This finishes the proof that $\boldsymbol{\theta}_i^{(k)}|\boldsymbol{h}_{i,.}^{(k)} \sim N_{p+1}(\boldsymbol{u}_k^1, \boldsymbol{\Sigma}_k^1)$, where $\boldsymbol{u}_k^1 = (\frac{\boldsymbol{\Psi}'_i\boldsymbol{\Psi}_i}{\sigma_k^2} + (\boldsymbol{\Sigma}_k^0)^{-1})^{-1}(\frac{\boldsymbol{\Psi}_i\boldsymbol{h}_{i,.}^{(k)}}{\sigma_k^2} + (\boldsymbol{\Sigma}_k^0)^{-1}\boldsymbol{u}_k^0)$ and $\boldsymbol{\Sigma}_k^1 = (\frac{\boldsymbol{\Psi}'_i\boldsymbol{\Psi}_i}{\sigma_k^2} + (\boldsymbol{\Sigma}_k^0)^{-1})^{-1}$.