IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON ADVANCES IN PROGNOSTICS AND SYSTEM HEALTH MANAGEMENT

Received March 31, 2019, accepted April 23, 2019, date of publication May 1, 2019, date of current version May 15, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2914236

Data-Driven Remaining Useful Life Prediction Considering Sensor Anomaly Detection and Data Recovery

LIANSHENG LIU, (Member, IEEE), QING GUO, (Member, IEEE), DATONG LIU, (Senior Member, IEEE), AND YU PENG, (Member, IEEE)
School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150080, China

Corresponding author: Datong Liu (liudatong@hit.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61803121 and in part by the China Postdoctoral Science Foundation under Grant 2019M651277.

ABSTRACT Prognostics and health management (PHM) is being adopted more and more in the modern engineering systems. As one of the most important technologies in the PHM domain, remaining useful life (RUL) prediction has attracted much attention from the researchers in the scholar and industrial field. Although many methods have been proposed to improve the prediction result, the problem of sensor anomaly detection and data recovery has not been considered together. To achieve this object, the data-driven RUL prediction framework considering sensor anomaly detection and data recovery is proposed, which is expected to improve the performance of RUL prediction caused by sensor anomaly. The selected sensor data are first detected to decide whether they are anomalous. If the data of this selected sensor are normal, they are continuously adopted as the input of the RUL prediction algorithm. But, if the data are anomalous, they will be recovered by the related algorithm. The recovered data will be utilized as the input of the RUL prediction algorithm. In the proposed framework, mutual information, Kernel principal component analysis (KPCA), least square-support vector machine (LS-SVM), and Gaussian process regression (GPR) are utilized. Both simulation data and practical data are used to evaluate the performance of the proposed method. Compared with abandoning the anomalous sensor data, the recovered data can indeed help to enhance the RUL prediction result.

INDEX TERMS Remaining useful life, mutual information, sensor anomaly detection, data recovery for prognostics.

I. INTRODUCTION

To assess the system condition in advance, one promising technology is the Prognostics and Health Management (PHM) [1]. Not only can the health of the system and critical components be assessed, but also the Remaining Useful Life (RUL) can be achieved by appropriate methods [2]. Based on the information provided by PHM, the condition-based maintenance can be realized. In this way, the system life-cycle cost can be reduced and its availability can also be improved. Therefore, PHM has become one of the most promising researches in recent years [3]–[5].

In the research area of PHM, there are mainly two kinds of methods, which are the model-based method and the data-driven method [6], [7]. To achieve precise and stable prediction result, the model-based method is appropriate. The reason is that the physical model can reflect the system behavior accurately. However, the physical model of complex system is not easily formulated [8]. Therefore, the model-based method cannot be realized in many modern industry systems. In contrast, the data-driven method which depends on the condition monitoring data of the system is easier to be implemented [9]. With the progress of sensor technology, industry internet, internet of things, etc., more and more sensing data of the monitored system are available. Hence, the data-driven method has become the promising technology for the engineering system.

Many data-driven PHM methods have been proposed to improve the prediction result for different applications.

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

To achieve better uncertainty, the model noises are adjusted by the short-term prediction and correction loop, which also has the advantage of less computation requirement [10]. Wang *et al.* [11] propose one kind of three-stage intelligent method, in which the variation mode decomposition-based trend detection and self-weight algorithm are both adopted. Experimental results show that this method can be used to implement monitoring adaptively. The deep belief network has been adopted to enhance the monitoring result of analog circuit and bearing, respectively [12], [13]. In addition, the Accelerated Degradation Testing (ADT) can be utilized to obtain sufficient data for system condition analysis [14], [15]. With the help of ADT, the RUL prediction result can be enhanced at some degree.

As the traditional research subject, the related works about the rotating machine are abundant. Li *et al.* [16] present a comprehensive review on the related fuzzy formalisms to diagnose the bearing fault. The novel convolutional neural network and feature enhancement method are introduced to improve bearing fault [17], [18]. Guo *et al.* [19] propose one kind of hybrid model to realize the deep fault in the rotating machine. The different fault size and position of the bearing can be achieved by the vertical-horizontal synchronized root mean square index [20]. The feature extraction of bearing fault can be achieved accurately using the method of time-frequency manifold sparse reconstruction [21]. The long short-term memory (LSTM) recurrent neural network is adopted successfully to evaluate bearing performance degradation by utilizing the fault propagation information [22]. These works provide innovative methods mainly based on signal processing. However, the original sensing data or sensor anomaly are not considered.

On the theme of sensing data analysis, some works have been implemented to select the most appropriate sensor data or detect sensor anomaly for prognostics. The improved permutation entropy is proposed to select the appropriate sensors for RUL prediction in [23]. Take the aircraft engine as an example, its condition monitoring data are detected to reach better prediction result [24]. For the aerospace application, the sensor anomaly detection is also important and should pay attention to enhance its reliability [25]. For the unmanned aerial vehicle, Guo *et al.* [26] propose one kind of fault detection method based on the local regulated optimization method. If the sensing data are affected by its anomalous or fault condition, its output can be recovered to maintain a good result [27]. Hence, it is important to consider the influence of anomalous sensor data on RUL prediction result.

However, to the best of our knowledge, there is no data-driven framework which has considered the sensor anomaly detection and data recovery for RUL. Especially for the online application scenario, the RUL prediction algorithm is run continuously to assess the system condition. If one sensor data become anomalous, it may bring serious influence on the prediction result. To avoid this problem, we propose the data-driven framework which detect and recover the target sensor together for RUL prediction. If the target sensor data are normal, they will be continuously adopted as the input of RUL prediction algorithm. If the target sensor data are detected as anomaly, they will be recovered by other available sensors data. Then, the recovered sensor data are adopted as the input of RUL prediction algorithm. This proposed framework is based on mutual information, Kernel Principal Component Analysis (KPCA), Least Square - Support Vector Machine (LS-SVM), and Gaussian Process Regression (GPR). Compared with abandoning the anomalous sensor data, the result using recovered sensor data is expected to be improved.

The remaining of this paper is arranged as follows. Section II presents the proposed data-driven framework and the related algorithms briefly. Section III introduces the utilized data sets for the evaluation experiments. Section IV gives the experimental results in detail. Section V draws the conclusion and provides the future work.

II. THE PROPOSED METHOD

In this section, the proposed data-driven framework considering sensor anomaly detection and data recovery is firstly introduced. Then, the related theories are presented. Finally, the metrics for measuring anomaly detection, data recovery and RUL prediction are provided.

A. DATA-DRIVEN FRAMEWORK

The data-driven RUL prediction framework dealing with sensor anomaly detection and data recovery is proposed, which includes the sensor anomaly detection and data recovery. The flowchart of the proposed framework is shown as follows.

In the above figure, the raw sensors data are acquired by the employed sensors in the system. If the sensors data (i.e., $S_1, \ldots, S_n$) which contain the system degradation information can be selected and analyzed appropriately (i.e., $S_1^{'}, \ldots, S_m^{'}$), the relatively optimal RUL prediction result can be realized.

However, if one of those utilized sensors data for RUL prediction becomes anomalous, the prediction result may be influenced at some degree. In the proposed method, the target sensor data will be detected to decide whether its data are anomalous. As shown in Fig. 1, if the data are normal, they are continuously utilized as the input of RUL prediction. If the data become anomalous, they will be recovered. Then, those recovered data with the other available sensors data (i.e., $S_1^{'}, \ldots, S_{q-1}^{'}, S_{q-1}^{'}, \ldots, S_m^{'}$) will be utilized for the following RUL prediction.

To implement the proposed method, some typical algorithms for sensor anomaly detection, data recovery and RUL prediction are essential. As illustrated in the framework, these three kinds of algorithms are KPCA, LS-SVM, and GPR, respectively. Some previous studies using these three algorithms have proved their effectiveness for the corresponding target. The details will be introduced in the following subsections. To be specific, these three methods can be replaced by other available algorithms for realizing relatively optimal result. The merit of this framework is that RUL can be run with acceptable prediction result in case of sensor anomaly.

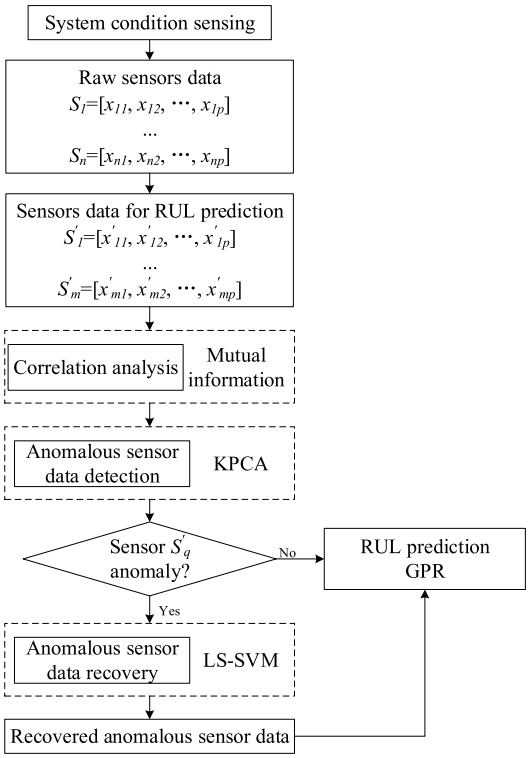


FIGURE 1. Data-driven framework for system RUL prediction considering sensor anomaly detection and data recovery.

B. MUTUAL INFORMATION

To analyze the correlation between two sensors data of system condition, the quantitative indicator of mutual information is utilized. Its definition is based on Shannon entropy theory. Thus, how to get the Shannon entropy is firstly introduced. If the variable is discrete, its Shannon entropy can be determined by [28]

$$H = -\sum_{i=1}^{N} p_i(x) \log p_i(x), \tag{1}$$

where $p_i(x)$ is the probability each variable element, and N is the quantity of the discrete variable. The logarithm base in (1) determines the unit of the Shannon entropy. When the logarithm base is 2, the corresponding unit is *bit*. When the logarithm base is e, the corresponding unit is *nat*.

To calculate the mutual information between two sensors data, one essential item is to get the conditional entropy $H(X|Y)$. Its definition is given by

$$\begin{aligned} H(X|Y) &= \sum_{j=1}^{m} p(y_j)\, H(X|Y = y_j) \\ &= -\sum_{j=1}^{m}\sum_{i=1}^{n} p(y_j, x_i) \log p(x_i|y_j). \end{aligned} \tag{2}$$

Similarly, the $H(Y|X)$ can be reached by the following equation.

$$\begin{aligned} H(Y|X) &= \sum_{i=1}^{n} p(x_i)\, H(Y|X = x_i) \\ &= -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j) \log p(y_j|x_i) \end{aligned} \tag{3}$$

Based on the above definitions, mutual information can be achieved by

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \\ &= \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}. \end{aligned} \tag{4}$$

The dependency among the aforementioned definitions is demonstrated by Fig. 2.

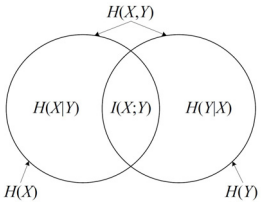


FIGURE 2. Relationship among mutual information, entropy and conditional entropy [27].

In this study, mutual information is used for measuring the relationship between two sensors data. The reference data for the target anomaly detection and recovery are determined according to the numerical value of mutual information.

C. ANOMALY DETECTION ALGORITHM

As verified in our previous study [29], KPCA has the advantage of realizing anomaly detection of the target sensor data. For training data, they are firstly mapped into the high-dimensional feature space. Then, Principal Component Analysis (PCA) is adopted to construct the feature space.

For training data $X_n = [x_1 x_2 \ldots x_n] \in \mathbb{R}^{n\times m}$, its sample of data vector is $x_i \in \mathbb{R}^{m\times 1} (1 \leq i \leq n)$. The training data are mapped by ϕ: $x \in \mathbb{R}^{m\times 1} \rightarrow \phi(x) \in F$. In this way, the target high-dimensional feature space is achieved. Then, PCA can be used to determine the vector f for the following transformation

$$t = (\phi(x))^T f. \tag{5}$$

If $\phi(x)$ has the features of mean centred and variance scale. The corresponding optimization can be formulated by

$$\begin{aligned} \max J_{KPCA}(f) &= \max \frac{1}{n-1}\sum_{i=1}^{n} t_i^2 \\ &= \max \frac{1}{n-1}\sum_{i=1}^{n} ((\phi(x_i))^T f)^2 \\ s.t.\, f^T f &= 1. \end{aligned} \tag{6}$$

The vector f and $t_n = [t_1, t_2, \ldots, t_n]^T$ are loading vector and score vector, respectively. There are coefficients α_j, $1 \le j \le n$ of the loading vectors, which are defined by

$$s.t. f^T f = 1. \tag{7}$$

The optimization problem can be transformed by substituting (7) into (6), as given by

$$\max J_{KPCA}(\alpha) = \max \frac{1}{n-1} \sum_{i=1}^{n} ((\phi(x_i))^T \sum_{j=1}^{n} \alpha_j \phi(x_j))^2$$
$$s.t. (\sum_{i=1}^{n} \alpha_i \phi(x_i))^T (\sum_{i=1}^{n} \alpha_i \phi(x_i)) = 1, \tag{8}$$

where $\alpha = [\alpha_1 \alpha_2 \ldots \alpha_n] \in \mathbb{R}^{n\times 1}$. Some available kernel functions (e.g., sigmoid kernel, polynomial kernel, and radial basis kernel) can be adopted.

D. DATA RECOVERY ALGORITHM

For data recovery of the anomalous sensor data, LS-SVM regression is utilized. As evaluated in [27], LS-SVM can achieve relatively acceptable result of data recovery. By selecting the appropriate input of training data, the recovered sensor data with better precision and stability can be realized.

LS-SVM evolves from Support Vector Machine (SVM). In the traditional research area, SVM is mainly adopted for solving the regression problem. The inequal constraints of SVM are transformed into the equality constraints and the squared error becomes a loss function during the training process. The details of this algorithm are described as follows.

The training data for the LS-SVM algorithm are denoted by $S = \{(x_i, y_i)|i = 1, 2, 3, \cdots, N\}, x_i \in R^n, y_i \in R$. x_i are input data and y_i are the corresponding output data. The optimal problem can be described by

$$\min J(w, \xi) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{k=1}^{N} \xi_i^2, \tag{9}$$
$$y_i = w^T \varphi(x_i) + b + \xi_i, \quad i = 1, 2, 3, \cdots, N, \tag{10}$$

where $\varphi(\cdot)$ refers to nonlinear mapping function. b denotes the bias and ξ_i indicates the error. $J(\cdot)$ represents the loss function. γ refers to the adjustable constant. The mapping function is adopted to determine the features from original space. Then, training data are mapped into the high-dimension feature space. Finally, the problem of nonlinear regression can be solved.

Based on (9), the corresponding Lagrangian function can be illustrated by

$$\begin{bmatrix} 0 & \overrightarrow{1}^T \\ \overrightarrow{1} & K + \gamma^{-1} I \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{11}$$

where $\overrightarrow{1} = [1, 1, \cdots, 1]^T$, $\alpha = [\alpha_1, \alpha_2, \cdots, \alpha_N]^T$, and K refers to the kernel matrix. The elements in K can be expressed by

$$K(i, j) = k(x_i, x_j) = \varphi(x_i)^T \varphi(x_j). \tag{12}$$

The widely utilized kernel function is the radial basis function, which is defined by

$$k(x_i, x_j) = \exp(-\|x_i - x_j\|/2\delta^2). \tag{13}$$

By solving (9), in which $A = K + \gamma^{-1} I$, the values of b and α can be realized by

$$b = \frac{\overrightarrow{1}^T A^{-1} y}{\overrightarrow{1}^T A^{-1} \overrightarrow{1}}, \tag{14}$$
$$\alpha = A^1 (y - b\overrightarrow{1}). \tag{15}$$

In this way, the regression function of LS-SVM can be achieved by

$$y(x) = \sum_{i=1}^{N} \alpha_i k(x, x_i) + b. \tag{16}$$

In this study, LS-SVM is adopted to reconstruct the data of anomalous sensor using other available sensors data. Those reconstructed sensor data will be utilized as the input of RUL prediction algorithm. Compared with abandoning the anomalous sensor data as the input of prediction algorithm, the prediction result is expected to be improved at some degree.

E. RUL PREDICTION ALGORITHM

Many algorithms have been proposed to realize RUL prediction of complex system. To summarize the related works, some algorithms are to optimize the parameters of the algorithm. Another kind of typical algorithm is to fuse different methods to achieve better prediction result. In this study, we choose the well-known GPR algorithm to realize RUL prediction, which has been widely adopted in different scenarios.

GPR is based on the Gaussian Process (GP), which is one kind of stochastic process. Many practical industrial applications can be formulated by GP. Therefore, GPR can also be adopted to formulate the related problems. For the input data $D = \{x_n\}_{n=1}^{N}, x \in R^d$, the corresponding functions $f(x_1), \ldots, f(x_N)$ are consisted of some random variables. These random variables also comply with the joint Gaussian distribution. The $f(x_1), \ldots, f(x_N)$ are utilized to formulate GP, which are illustrated by

$$f(x) \sim GP(m(x), k(x_i, x_j)), \tag{17}$$
$$m(x) = E[f(x)], \tag{18}$$
$$k(x_i, x_j) = E[(f(x_i) - m(x_j))(f(x_i) - m(x_j))], \tag{19}$$

where $m(x)$ and $k(x_i, x_j)$ represent mean function and covariance function, respectively.

For practical application, $f(x)$ is influenced by noise, as given by

$$y = f(x) + \varepsilon, \tag{20}$$

where $\varepsilon \in N(0, \sigma_n^2)$ is white noise. The ε does not depend on $f(x)$. If $f(x)$ is adopted for the model, the corresponding y is also GP, as expressed by

$$y \sim GP(m(x), k(x_i, x_j + \sigma_n^2 \delta_{ij})), \tag{21}$$

where δ_{ij} refers to the Dirac function.

The following assumptions are utilized to realize the final calculation process. Let $D_1 = \{x_i, y_i\}_{i=1}^{N}$ and $D_2 = \{x_i^*, y_i\}_{i=1}^{N^*}$ be training data and test data, respectively, in which $x_i, x_i^* \in R^d$. In the above illustration, the dimension of data is represented by d. The mean vectors of training data and test data are m and m^*, respectively. The output of test data is denoted by $f(x_*)$. y refers to the training vector. Therefore, f_* and y comply with the joint Gauss distribution, as given by

$$\begin{pmatrix} y \\ f_* \end{pmatrix} \sim \left(\begin{bmatrix} m \\ m_* \end{bmatrix}, \begin{pmatrix} C(X,X)K(X,X_*) \\ K(X_*,X)K(X_*,X) \end{pmatrix} \right), \tag{22}$$

where the covariance matrix $C(X, X)$ is derived from the training data. $C(X, X) = K(X, X) + \delta_{ij} I$ is the corresponding formulation and δ_{ij} is contaminated with noise. $I \in R^{N \times N}$ refers to the unit matrix, $K(X_*, X_*)$ indicates the covariance of test data and $K(X, X_*) \in R^{N \times N^*}$ is the matrix which denotes the covariance.

To achieve f_*, the following three equations are essential.

$$f_*|X, y, X_* \sim N(\overline{f_*}, cov(f_*)), \tag{23}$$

$$\begin{aligned} \overline{f_*} &= E[f_*|X, y, X_*] \\ &= m + K(X_*, X)C(X, X)^{-1}(y - m), \end{aligned} \tag{24}$$

$$cov(f_*) = K(X_*, X_*) - C(X, X)^{-1}K(X, X_*). \tag{25}$$

By using $m(x)$ and $k(x_i, x_j)$, GP model can be fully formulated. During RUL prediction in this study, the regression capability of GPR algorithm is adopted, in which the mean is set to be 0. To realize the covariance function, the neural network is selected. In this way, the related hyperparameters in the aforementioned functions can reach the relatively optimal results.

F. RELATED METRICS

False Positive Ratio (*FPR*), False Negative Ratio (*FNR*) and Accuracy (*ACC*) are widely adopted to measure the performance of anomaly detection. Their definitions are illustrated by the following three equations.

$$FPR = \frac{FN}{TP + FN} \times 100\% \tag{26}$$

$$FNR = \frac{FP}{FP + TN} \times 100\% \tag{27}$$

$$ACC = \frac{TP + TN}{FP + FN + TN + TP} \tag{28}$$

In (26), FN refers to the amount of falsely detected normal data and $TP + FN$ denotes the quantity of all normal data for implementing anomaly detection. Hence, FPR denotes the ratio of normal data detected falsely. In (27), FP indicates the amount of falsely detected normal data and $FP+TN$ refers to the total number of anomalous data for carrying out anomaly detection. Therefore, FNR denotes how many anomalous data are detected falsely. Using the aforementioned explanation, the meaning of ACC can be understood that it denotes the ratio of both normal data and anomalous data detected correctly.

Relative Error (*RE*) and Root Mean Squared Error (*RMSE*) are usually adopted to measure the performance of the data recovery algorithm. These two definitions are given by

$$RE = \frac{\left| R_i - \hat{R}_i \right|}{R_i} \times 100\%, \tag{29}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(R_i - \hat{R}_i \right)^2}. \tag{30}$$

In the above two equations, R_i is the real data and $\hat{R}_i$ denotes the corresponding recovered data. If the numerical value of *RE* is smaller, the accuracy of recovered data at different points is more excellent. If the numerical value of *RMSE* is smaller, it implies that the stability of the whole recovered data is more excellent.

Mean Absolute Error (*MAE*) and Root Mean Square Error (*RMSE*) are two well-known metrics for evaluating the performance of RUL prediction, as defined by

$$MAE = \frac{1}{N} \sum_{k=1}^{N} |P_k - R_k|, \tag{31}$$

$$RMSE = \sqrt{\frac{\sum_{k=1}^{N} (P_k - R_k)^2}{N}}. \tag{32}$$

In (31) and (32), P_k denotes the predicted result, R_k refers to the actual values, and N refers to the total cycle number. If *MAE* and *RMSE* values are small, it means that the predicted result has better performance of accuracy and stability, respectively.

III. DATA DESCRIPTION

In this section, the utilized simulation data and practical dat for evaluating the proposed data-driven framework of RUL prediction are introduced.

A. SIMULATION DATA SET

The PHM 2008 Conference challenge data which contain the RUL information of aircraft engine are utilized [30]. Fig. 3 shows the structure of the aircraft engine, which is mainly consisted of High-Pressure Compressor (HPC), Low-Pressure Compressor (LPC), High-Pressure Turbine (HPT), Low-Pressure Turbine (LPT), etc.

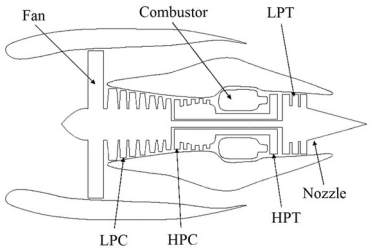


FIGURE 3. The simulated aircraft engine structure [31].

To implement the simulation of the aircraft engine, the tool of Commercial Modular Aero-Propulsion System Simulation (C-MAPSS) is adopted. In C-MAPSS, both the open loop and the closed loop can be realized. The connection of different subsystems is given in Fig. 4.

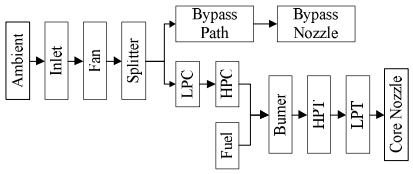


FIGURE 4. The simulated aircraft engine structure [31].

The engine has one kind of build-in control system. It is composed of a fan-speed controller, a few regulators and a few limiters, which are utilized to keep the aircraft engine work under some certain limits. The typical sets in this simulation environment has high similarity with the real scenario. For sensing its condition, a number of 21 sensors are utilized to sense its condition, as illustrated in Table 1.

TABLE 1. Sensors employed in the simulation [31].

Index	Symbol	Description	Units
1	T2	Total temperature at fan inlet	°*R*
2	T24	Total temperature at LPC outlet	°*R*
3	T30	Total temperature at HPC outlet	°*R*
4	T50	Total temperature at LPT outlet	°*R*
5	P2	Pressure at fan inlet	psia
6	P15	Total pressure in bypass-duct	psia
7	P30	Total pressure HPC outlet	psia
8	Nf	Physical fan speed	rpm
9	Nc	Physical core speed	rpm
10	Epr	Engine Pressure ratio	-
11	Ps30	Static pressure at HPC outlet	psia
12	Phi	Ratio of fuel flow to Ps30	pps/psi
13	NRf	Corrected fan speed	rpm
14	NRc	Corrected core speed	rpm
15	BPR	Bypass ratio	-
16	farB	Burner fuel-air ratio	-
17	htBleed	Bleed enthalpy	-
18	Nf_dmd	Demanded fan speed	rpm
19	PCNfR_dmd	Demanded corrected fan speed	rpm
20	W31	HPT coolant bleed	lbm/s
21	W32	LPT coolant bleed	lbm/s

Among these 21 sensors, only some sensors data contain its degradation information. However, if one of sensors data used for RUL prediction becomes anomalous, it may lead to seriously wrong RUL prediction result. If these anomalous sensor data are detected and dropped, the RUL prediction result may also deviate from the actual value seriously.

B. PRACTICAL DATA SET

The condition monitoring data of Flyable Electromechanical Actuator (FLEA) is adopted in this study. This data set has been utilized for some studies [32], [33]. Its structure is illustrated in Fig. 5 in details.

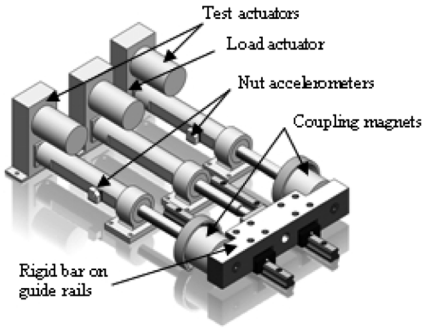


FIGURE 5. The actual structure of FLEA [32].

In the FLEA, there are three types of actuators. The first is the nominal actuator. The second is injected with faults. The third is with dynamic load. This load can be switched in two terms of the healthy condition and the faulty test. There are 16 sensors deployed on the FLEA to sense its condition. The specific descriptions of these sensors are given in Table 2.

TABLE 2. The deployed sensors on FLEA [32].

Index	Description	Type	Quantity
1	Load cell	Omega LC703-150	1
2	Accelerometer	Endevco 7253C	2
3	Thermocouple	T type	4
4	Rotary encoder	Ultramotion E5DIFF	2
5	Linear potentiometer	UltraMotion precision	1
6	Voltage Sensor	custom	3
7	Current sensor	custom	3

Similar to the simulation data set, only some sensors data among these 16 sensors contain the degradation information of FLEA. To avoid the influence of anomalous sensor data on RUL prediction of FLEA, the effectiveness of the proposed method is evaluated.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

Both simulation data and practical data are utilized to evaluate the proposed method in this section. Experimental results are illustrated and discussed in detail, especially the comparison experiments between RUL prediction with the recovered sensor data and RUL prediction without the recovered sensor data.

A. EXPERIMENTAL RESULTS AND DISCUSSION OF SIMULATION DATA SET

Based on our study in [23], the optimal RUL prediction of the aircraft can be achieved by using seven-dimensional sensing data. Those utilized sensors among the 21 sensors are 3#, 4#, 8#, 9#, 14#, 15# and 17#. Therefore, the following experiment will be carried out based on these seven sensors data. By using (4), the mutual information of those 7 sensors are summarized in TABLE 3.

TABLE 3. Mutual information of selected simulation data.

Sensor	3#	4#	8#	9#	14#	15#	17#
3#	4.6740	4.0270	2.5950	4.0353	3.9717	4.1129	1.3478
4#	4.0270	4.6104	2.5964	3.9717	3.9009	4.0421	1.3529
8#	2.5950	2.5964	3.1496	2.5614	2.4906	2.6246	0.7188
9#	4.0353	3.9717	2.5614	4.6187	3.9092	4.0576	1.2321
14#	3.9717	3.9009	2.4906	3.9092	4.5479	3.9896	1.2506
15#	4.1129	4.0421	2.6246	4.0576	3.9896	4.6820	1.3368
17#	1.3478	1.3529	0.7188	1.2321	1.2506	1.3368	1.7409

To carry out the sensor anomaly detection, sensor 15# is supposed to work in fault condition and its data are anomalous. To achieve better anomaly detection result, the cross-validation experiments of different mutual information between other sensors data and sensor 15# data are carried out to determine the reference sensors data for training KPCA. when the reference sensors are 3#, 4# and 9#, the relatively optimal anomaly detection can be achieved, as given in Fig. 6. If Squared Prediction Error (SPE) of test data is larger than that of the training data, these test data are detected as anomaly data. Otherwise, they are detected as normal data.

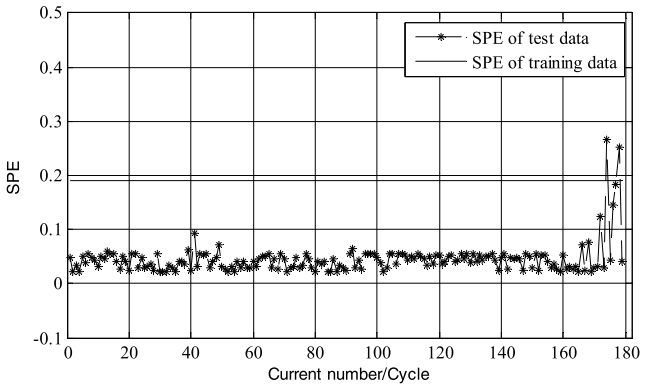


FIGURE 6. Sensor 15# data anomaly detection by sensors 3#, 4# and 9# data using KPCA.

In this experiment, all the detected sensor data are normal. Thus, we adopt the *FPR* metric to analyze the experimental result. *FPR* can reflect the ratio of normal data detected falsely. Therefore, it can be adopted to measure the experimental results of this anomaly detection result. In this experiment, only 2 normal data are detected falsely. The total number of detected data is 179. Hence, the corresponding *FPR* of this experiment is 1.12%.

In general, if the value of *FPR* is smaller than 10%, the performance of the anomaly detection algorithm is excellent. If its data become anomalous, they will be detected accordingly. To verify the proposed data-driven framework, sensor 15# data will be recovered by sensors 3#, 4# and 9# data.

Experimental result of recovered sensor 15# data is shown in Fig. 7.

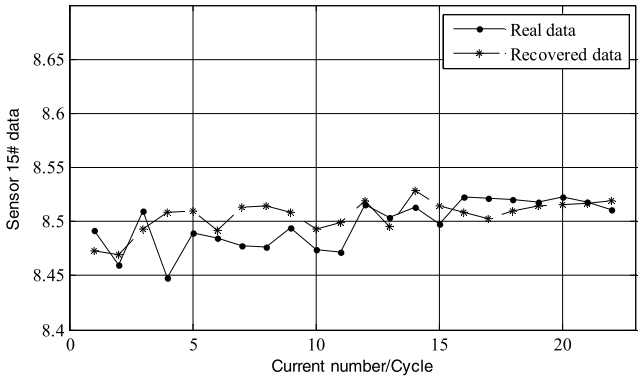


FIGURE 7. Recovered sensor 15# data by sensors 3#, 4# and 9# data using LS-SVM.

The recovered data curve is close to the real data curve. Therefore, the performance of the data recovery algorithm is relatively excellent. By numerical analysis, the maximal *RE*, minimal *RE* and *RMSE* of the recovered data are 0.72%, 0.04%, and 0.02, respectively. Then, those recovered sensor data are adopted as the input of RUL prediction. The prediction result is given in Fig. 8.

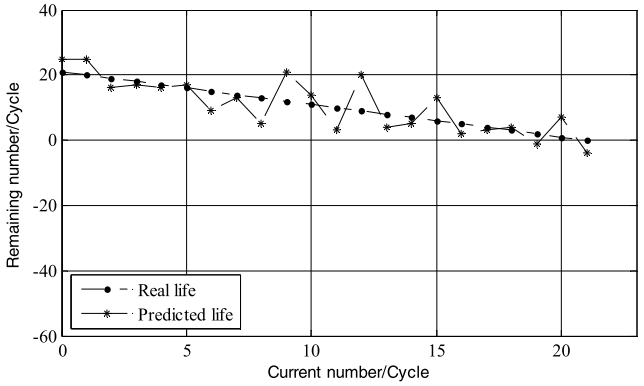


FIGURE 8. RUL prediction by sensors 3#, 4# and 9# data and the recovered sensor 15# data using GPR.

To evaluate the experimental result in Fig. 8, the metrics of *MAE* and *RMSE* are calculated. The corresponding values of these two metrics are 4.14 cycles and 5.02 cycles, respectively. If the anomalous sensor 15# data are not utilized as the input of GPR, the RUL prediction result is illustrated in Fig. 9.

The values of *MAE* and *RMSE* in this experiment are 21.68 cycles and 25.74 cycles, respectively. Compared with

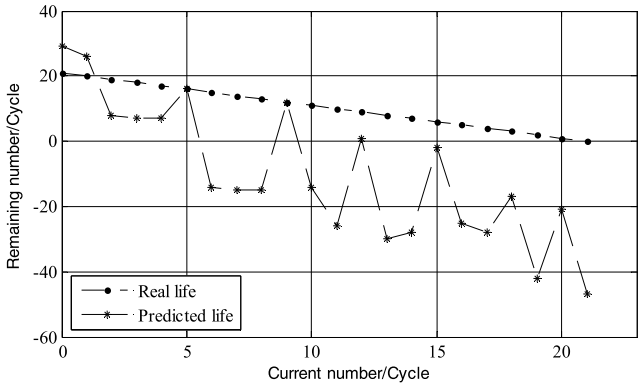


FIGURE 9. RUL prediction by sensors 3#, 4#, and 9# data using GPR.

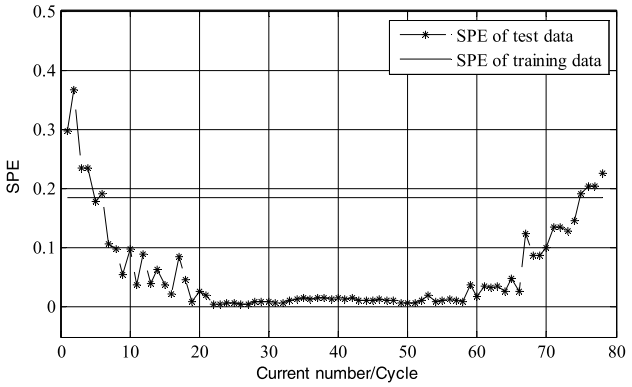


FIGURE 10. Sensor 8# data anomaly detection by sensors 3# and 11# data using KPCA.

the two experimental results in Fig. 8 and Fig. 9, the recovered sensor 15# data have directly positive influence on the RUL prediction result. Therefore, the proposed data-driven framework of RUL prediction with sensor anomaly detection and data recovery has been verified by this simulation data set.

B. EXPERIMENTAL RESULTS AND DISCUSSION OF PRACTICAL DATA SET

Similar with experiments implemented using the simulation data, the optimal RUL prediction of the FLEA can be achieved by using four-dimensional data. The corresponding sensor numbers are 3#, 8#, 11# and 13#. Therefore, the following evaluation experiment will be carried out based on these four sensors data. By using (4), the mutual information of those 7 sensors are summarized in Mutual information among these four sensors are calculated, as given in TABLE 4.

TABLE 4. Mutual information of selected practical sensors data.

Sensor	3#	8#	11#	13#
3#	4.4603	3.9563	3.7208	2.1044
8#	3.9563	4.8122	4.0795	2.5104
11#	3.7208	4.0795	4.5835	2.2842
13#	2.1044	2.5104	2.2842	2.9037

In the following evaluation, sensor 8# data are supposed to be anomalous. As the evaluation process in the above subsection, the cross-validation experiments from big mutual information to small mutual information between other sensors data and sensor 8# data are also implemented to select the reference sensors data for training KPCA. The relatively optimal detection result is illustrated in the following Fig. 10. The corresponding reference sensors are 3# and 11#.

The total number of sensor 8# data detected is 78 and the number of falsely detected data is 9. Thus, the value of *FPR* is 11.54%. Although this *FPR* is relatively large, the confidence of anomaly detection can be regulated to a small value. Then,

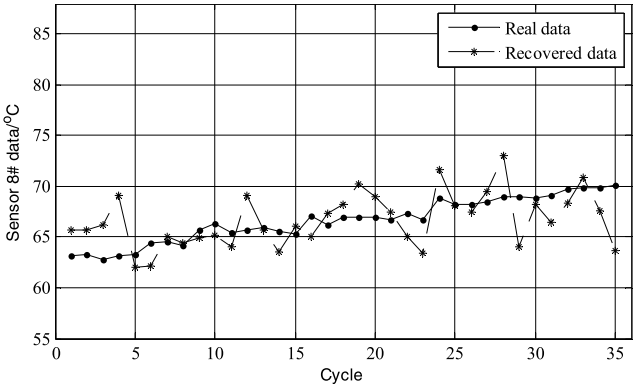


FIGURE 11. Recovered sensor 8# data by sensors 3# and 11# data using LS-SVM.

the sensors 3# and 11# data are adopted to recover sensor 8# data. Experimental result is illustrated in Fig. 11.

By numerical analysis, the maximal *RE,* minimal *RE* and *RMSE* of the recovered data are 9.05%, 0.24%, and 2.56, respectively. The recovered data do not deviate from the real data seriously. Those recovered sensor data are adopted as the input of RUL prediction. The prediction result is given in Fig. 12.

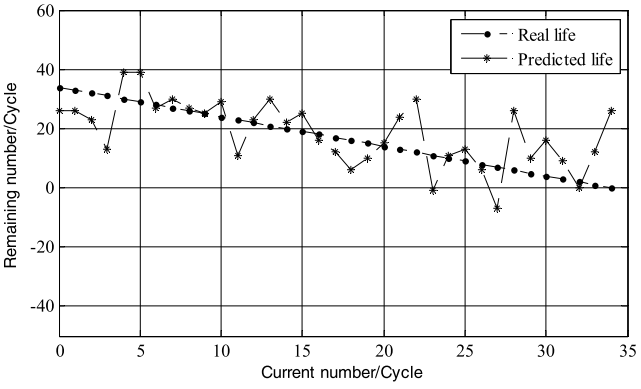


FIGURE 12. RUL prediction by sensors 3# and 11# data and the recovered sensor 8# data using GPR.

To evaluate the experimental result in Fig. 12, the metrics of *MAE* and *RMSE* are calculated. The value of *MAE* is 7.66 cycles and the value of *RMSE* is 9.81 cycles. Although these two values are relatively large, the comparison experiment which discards sensor 8# data for the RUL prediction algorithm directly may have worse prediction result.

For the RUL prediction experiments without sensor 8# data, the result is shown in Fig. 13.

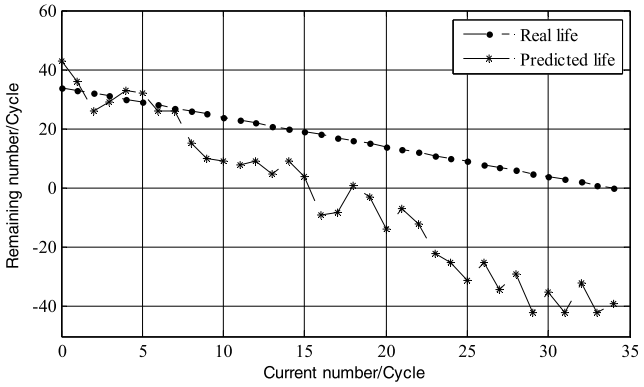


FIGURE 13. RUL prediction by sensors 3# and 11# data using GPR.

The values of *MAE* and *RMSE* in this experiment are 21.74 cycles and 25.95 cycles, respectively. The prediction values of RUL deviate from the real values seriously. Compared with the two experimental results in Fig. 12 and Fig. 13, the recovered sensor 8# data also have directly positive influence on the RUL prediction result of FLEA. These experimental results are similar with those of simulation data set. Hence, the proposed data-driven framework of RUL prediction is also verified by this practical data set.

C. COMPARISON ANALYSIS AND DISCUSSION

By observing the aforementioned experiments, it can be seen that the whole framework can indeed bring positive effectiveness for RUL prediction by recovering the anomalous sensor data. Compared with discarding the anomalous sensor 15# data as the input of RUL, *MAE* and *RMSE* of RUL prediction result are improved 80.90% and 80.49% using the simulation data set. For the practical data set, the recovered sensor 8# data can improve the *MAE* and *RMSE* of RUL prediction result 64.77% and 62.20%. Therefore, the proposed framework can realize better accurate and stable prediction results when the utilized sensor data are anomalous.

V. CONCLUSION AND FUTURE WORK

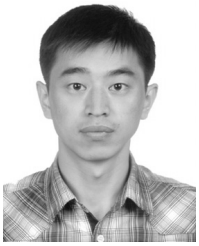
One kind of data-driven framework for RUL prediction considering sensor anomaly detection and data recovery is proposed in this study. The selected sensors data for implementing the target sensor data detection and recovery are based on the value of mutual information. Both the anomaly detection and recovered data can reach relatively positive result. For the RUL prediction, the recovered data can indeed help to enhance the prediction result deeply. The effectiveness of this proposed framework is evaluated by both simulation data and practical data.

The current study just considers the *FPR* metric in the term of anomaly detection. In future, different kinds of sensor anomaly detection will be implements, including bias, drift, etc. On the other hand, different algorithms for anomaly detection, data recovery and RUL prediction will also be considered in the proposed framework to extend its application scenario.

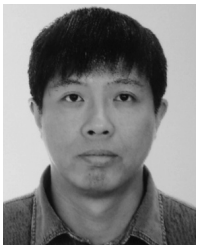
REFERENCES

[1] D. Wang, K.-L. Tsui, and Q. Miao, "Prognostics and health management: A review of vibration based bearing and gear health indicators," *IEEE Access*, vol. 6, pp. 665–676, 2018.
[2] G. Niu, *Data-Driven Technology for Engineering Systems Health Management: Design Approach, Feature Construction, Fault Diagnosis, Prognosis, Fusion and Decisions*. Beijing, China: Springer, 2017.
[3] F. Sun, N. Wang, X. Li, and W. Zhang, "Remaining useful life prediction for a machine with multiple dependent features based on Bayesian dynamic linear model and copulas," *IEEE Access*, vol. 5, pp. 16277–16287, 2017.
[4] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," *Mech. Syst. Signal Process.*, vol. 115, pp. 213–237, Jan. 2019.
[5] R. Zhao, D. Wang, R. Yan, K. Mao, F. Shen, and J. Wang, "Machine health monitoring using local feature-based gated recurrent unit networks," *IEEE Trans. Ind. Electron.*, vol. 65, no. 2, pp. 1539–1548, Feb. 2018.
[6] L. Liu, S. Wang, D. Liu, Y. Zhang, and Y. Peng, "Entropy-based sensor selection for condition monitoring and prognostics of aircraft engine," *Microelectron. Rel.*, vol. 55, nos. 9–10, pp. 2092–2096, 2015.
[7] H. Zhang, Q. Miao, X. Zhang, and Z. Liu, "An improved unscented particle filter approach for lithium-ion battery remaining useful life prediction," *Microelectron. Rel.*, vol. 81, pp. 288–298, Feb. 2018.
[8] L. Liu, S. Wang, D. Liu, and Y. Peng, "Quantitative description of sensor data monotonic trend for system degradation condition monitoring," in *Proc. Prognostics Syst. Health Manage. Conf. (PHM-Chengdu)*, Chongqing, China, Oct. 2016, pp. 1–5.
[9] Y. Zhang, L. Liu, Y. Peng, and D. Liu, "An electro-mechanical actuator motor voltage estimation method with a feature-aided Kalman filter," *Sensors*, vol. 18, no. 12, p. 4190, 2018.
[10] W. Yan, B. Zhang, G. Zhao, J. Weddington, and G. Niu, "Uncertainty management in Lebesgue-sampling-based diagnosis and prognosis for lithium-ion battery," *IEEE Trans. Ind. Electron.*, vol. 64, no. 10, pp. 8158–8166, Oct. 2017.
[11] Y. Wang, Z. Wei, and J. Yang, "Feature trend extraction and adaptive density peaks search for intelligent fault diagnosis of machines," *IEEE Trans. Ind. Informat.*, vol. 15, no. 1, pp. 105–115, Jan. 2019.
[12] G. Zhao, X. Liu, B. Zhang, Y. Liu, G. Niu, and C. Hu, "A novel approach for analog circuit fault diagnosis based on deep belief network," *Measurement*, vol. 121, pp. 170–178, Jun. 2018.
[13] Z. Chen and W. Li, "Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1693–1702, Jul. 2017.
[14] X.-Y. Li, J.-P. Wu, H.-G. Ma, X. Li, and R. Kang, "A random fuzzy accelerated degradation model and statistical analysis," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 3, pp. 1638–1650, Jun. 2018.
[15] H. Liao and Z. Tian, "A framework for predicting the remaining useful life of a single unit under time-varying operating conditions," *IIE Trans.*, vol. 45, pp. 964–980, Sep. 2013.
[16] C. Li, J. L. V. de Oliveira, M. C. Lozada, D. Cabrera, V. Sanchez, and G. Zurita, "A systematic review of fuzzy formalisms for bearing fault diagnosis," *IEEE Trans. Fuzzy Syst.*, to be published.
[17] H. Wang, S. Li, L. Song, and L. Cui, "A novel convolutional neural network based fault recognition method via image fusion of multi-vibration-signals," *Comput. Ind.*, vol. 105, pp. 182–190, Feb. 2019.
[18] H. Wang, P. Wang, L. Song, B. Ren, and L. Cui, "A novel feature enhancement method based on improved constraint model of online dictionary learning," *IEEE Access*, vol. 7, pp. 17599–17607, Jan. 2019.
[19] X. Guo, C. Shen, and L. Chen, "Deep fault recognizer: An integrated model to denoise and extract features for fault diagnosis in rotating machinery," *Appl. Sci.*, vol. 7, no. 1, p. 41, 2017.

[20] L. Cui, J. Huang, and F. Zhang, "Quantitative and localization diagnosis of a defective ball bearing based on vertical–horizontal synchronization signal analysis," *IEEE Trans. Ind. Electron.*, vol. 64, no. 1, pp. 8695–8705, Nov. 2017.
[21] X. Ding and Q. He, "Time–frequency manifold sparse reconstruction: A novel method for bearing fault feature extraction," *Mech. Syst. Signal Process.*, vol. 80, pp. 392–413, Dec. 2016.
[22] B. Zhang, S. Zhang, and W. Li, "Bearing performance degradation assessment using long short-term memory recurrent network," *Comput. Ind.*, vol. 106, pp. 14–29, Apr. 2019.
[23] L. Liu, S. Wang, D. Liu, and Y. Peng, "Quantitative selection of sensor data based on improved permutation entropy for system remaining useful life prediction," *Microelectron. Rel.*, vol. 75, pp. 264–270, Aug. 2017.
[24] L. Liu, D. Liu, Y. Zhang, and Y. Peng, "Effective sensor selection and data anomaly detection for condition monitoring of aircraft engines," *Sensors*, vol. 16, no. 5, p. 623, 2016.
[25] L. Liu, D. Liu, and Y. Peng, "Detection and identification of sensor anomaly for aerospace applications," in *Proc. Annu. Rel. Maintainability Symp. (RAMS)*, Tucson, AZ, USA, Jan. 2016, pp. 1–6.
[26] K. Guo, L. Liu, S. Shi, D. Liu, and X. Peng, "UAV sensor fault detection using a classifier without negative samples: A local density regulated optimization algorithm," *Sensors*, vol. 19, no. 4, p. 771, 2019.
[27] L. Liu, D. Liu, Q. Guo, Y. Peng, and J. Liang, "SDR: Sensor data recovery for system condition monitoring," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf. (I2MTC)*, Houston, TX, USA, May 2018, pp. 1–6.
[28] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul./Oct. 1948.
[29] L. Liu, M. Liu, Q. Guo, D. Liu, and Y. Peng, "MEMS Sensor data anomaly detection for the UAV flight control subsystem," in *Proc. IEEE SENSORS*, New Delhi, India, Oct. 2018, pp. 1–4.
[30] A. Saxena, K. Goebel, D. Simon, and N. Eklund, "Damage propagation modeling for aircraft engine run-to-failure simulation," in *Proc. Int. Conf. Prognostics Health Manage. (PHM)*, Denver, CO, USA, Oct. 2008, pp. 1–9.
[31] D. K. Frederick, J. A. DeCastro, and J. S. Litt, "User's guide for the commercial modular aero-propulsion system simulation (C-MAPSS)," NASA/ARL, New York, NY, USA, Tech. Manual TM2007-215026, 2007.
[32] E. Balaban, A. Saxena, S. Narasimhan, I. Roychoudhury, and K. Goebel, "Experimental validation of a prognostic health management system for electro-mechanical actuators," in *Proc. AIAA Infotech@ Aerosp.*, 2011, pp. 1–13.
[33] E. Balaban *et al.*, "Prognostic health-management system development for electromechanical actuators," *J. Aerosp. Inf. Syst.*, vol. 12, no. 3, pp. 329–344, 2015.



LIANSHENG LIU (M'12) received the B.Sc. degree in measurement technology and instrumentation and the M.Sc. and Ph.D. degrees in instrumentation science and technology from the Harbin Institute of Technology (HIT), in 2006, 2008, and 2017, respectively, where he is currently an Assistant Professor. From 2012 to 2014, he was with McGill University as a Visiting Ph.D. Student supported by the China Scholarship Council. His research interests include prognostics and health management, anomaly sensor data detection, and cyber physical systems. His research aims to develop advanced approaches for complex industrial system monitoring.



QING GUO (M'16) received the B.Sc. degree from the Department of Radio Engineering, Beijing University of Posts and Telecommunications, in 1985, and the M.Sc. and Ph.D. degrees in information and communication engineering from the Harbin Institute of Technology (HIT), in 1990 and 1998, respectively, where he is currently a Professor and the Dean of the School of Electronics and Information Engineering. His research interests include satellite communications and deep space communications.

DATONG LIU (M'11–SM'16) received the B.Sc. degree in automatic test and control and the M.Sc. and Ph.D. degrees in instrumentation science and technology from the Harbin Institute of Technology (HIT), in 2003, 2005, and 2010, respectively, where he was with the Department of Computer Science and Technology, from 2001 to 2003, and is currently a Professor with the School of Electronics and Information Engineering. From 2013 to 2014, he visited the University of Arizona as a Scholar. His research interests include automatic test, machine learning for test data processing, data-driven prognostics and health management, lithium-ion battery management, and complex system health management.

YU PENG (M'10) received the B.Sc. degree in measurement technology and instrumentation and the M.Sc. and Ph.D. degrees in instrumentation science and technology from the Harbin Institute of Technology (HIT), in 1996, 1998, and 2004, respectively, where he is currently a Professor and also the Deputy Dean of the School of Electronics and Information Engineering. His research interests include automatic test technologies, virtual instruments, system health management, and reconfigurable computing.