IEEE Access

Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Health Monitoring of Landing Gear Retraction/extension System Based on Optimized Fuzzy C-means Algorithm

JIE CHEN[1], SENYAO CHEN[1], ZHENBAO LIU[1], CAIKUN LUO[2], ZHENGDONG JING[1], QINGSHAN XU[1]

[1] The School of Civil Aviation, Northwestern Polytechnical University, Xi'an 710072, P. R. China
[2] The School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, P. R. China

Corresponding author: S.Y. Chen (e-mail: 782850331@qq.com).

ABSTRACT The landing gear retraction/extension(R/E) system has a critical impact on the safety of aircraft's take-off and landing, and its health status is important to decision-making of the system's prognostics and health management (PHM). With flight parameter data, a method to monitor the health of aircraft landing gear R/E system is proposed based on the improved fuzzy c-means algorithm (FCM).The landing gear health status are classified by the FCM cluster, and the granularity principle and density function are utilized to optimize the number of FCM algorithm clusters and the initial clustering center to better classification. At the same time, an optimized multi-dimensional scaling algorithm (MDS) is used to obtain low-dimensional features which facilitate cluster analysis. And the algorithm comparison and case analysis results show that the proposed method has a great health monitoring effect.

INDEX TERMS Landing gear retraction/extension system, fuzzy c-means, MDS, Health monitoring

I. INTRODUCTION

The health status of the landing gear R/E system directly affects the take-off and landing performance of the aircraft and flight safety [1], which plays an important role in preventing aircraft structure damage, slowing aircraft flutter, improving occupant comfort, and ensuring aircraft flight safety [2]. On May 21, 2018, an A330-200 aircraft of Saudi Airlines was on the way from Madinah to Dhaka, Bangladesh. Due to the failure of the nose landing gear, it was urgently diverted. In recent years, condition based maintenance (CBM) has been vigorously developed, compared with the traditional breakdown maintenance and preventive maintenance, the maintenance cost of the aircraft life cycle is reduced, and in which, the health status acquisition is very important. PHM [3-6] is an upgrade and development of CBM to meet the requirements of autonomous support and independent diagnosis. Usually there are three methods to realize PHM: model-based, data-driven and probability statistics, in which the data-driven method requires less prior knowledge of the target system. Based on the collected data, the hidden information can be mined by data processing and analysis, and then the shortcomings of the model-based and knowledge-based way is avoided [7].

In recent years, health monitoring technology has been rapidly developed in the aeronautics field, such as aircraft structures [8], inertial navigation systems [9], engine [10], and landing gear systems. Based on expert system, Yang [11] proposed a fault prognosis method of system component based on characteristic analysis and designed a landing gear R/E control diagnosis system. Chen [12] used the H ∞ filtering to realize the landing gear R/E system failure prognosis by utilizing the operational principle, failure characteristic analysis and the major failure modes. Although several articles [11-12] have been devoted to landing gear R/E systems, the model-based expert system method requires detailed analysis of the system internal mechanism, which is inefficient.

With the development of flight data extraction technology, the PHM based on data-driven has been advanced. Byington [13] proposed a method of health statement assessment based on neural network after learning the knowledge of landing gear actuator components. And He [14] studied the effect of changes in component parameter to extract the fault characteristics, and proposed a method to monitor the system health based on neural network. But with the massive and complex flight data, traditional neural network methods may lead to local optimal solutions.

Clustering analysis [15-16] is an important classification of unsupervised learning in data-driven methods. The general clustering algorithm is hard clustering, such as the traditional hard-core clustering algorithm K-means algorithm. Hard clustering separates the connections and is difficult to express the transition between samples, which is easy to fall into local optimum [17]. However, the landing gear R/E system is a process of gradual degradation, the traditional hard core clustering obviously can't express the transition status. The concept of fuzzy membership degree in the fuzzy clustering theory is proposed to reflect the relationship more objectively between sample data.

Due to these problems, a new data-driven health monitoring method is proposed in this paper. Firstly, the health status representation of the landing gear R/E system is obtained by the time domain feature extraction. Secondly, the manifold learning is used to improve the global aspects of the MDS to reduce the initial status features to obtain low-dimensional features that facilitate cluster analysis. And then the FCM clustering algorithm is served as the status classifier to realize system's health monitoring. What's more, during clustering, the granularity principle is utilized to find the best number of clusters and the density function which can improve the initial cluster centers is used to reduce the possibility of falling into local optimum. Finally the algorithm comparison and case analysis results show that the improved health monitoring method effectively complete the health monitoring of landing gear R/E system.

II. FAILURE AND PERFORMANCE EFFECT ANALYSIS OF LANDING GEAR R/E SYSTEM

A. LANDING GEAR RETRACTION/EXTENSION SYSTEM

The aircraft landing gear R/E system [18] is used to retract the landing gear after the aircraft has taken off, and to extract it before the aircraft landed. This system mainly consists of three parts:

(1) The selector valve is a three-position four-way electromagnetic reversing valve for switching the hydraulic oil circuit to achieve the system's retracting and releasing action.

(2) The front unlocking actuator cylinder realizes the unlocking and locking of the system. While the piston rod is extended, the landing gear is unlocked, vice versa. The front landing gear R/E actuator is used to retract or lower the front landing gear. When the piston rod of the actuator is extended, the front landing gear is lowered, vice versa.

(3) The main landing gear R/E actuator cylinder is used to retract or lower the main landing gear: when the piston rod of the main landing gear actuator is extended, the main landing gear is retracted, vice versa. When the main gear lower position lock actuator piston rod is extended, the landing gear is locked, vice versa. And the main gear upper position lock actuator piston rod is the opposite.

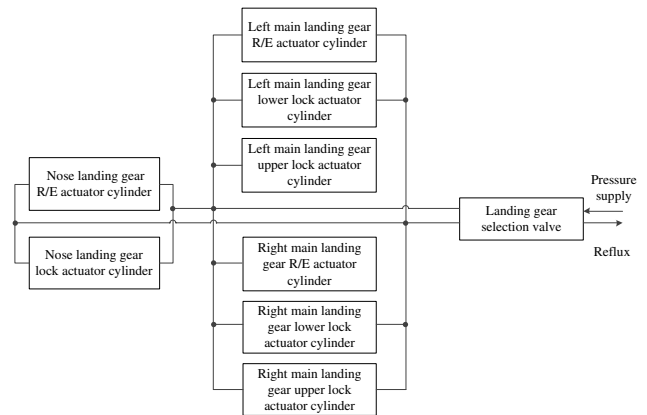The basic system configuration of the landing gear R/E system is shown as Fig.1.1.

FIGURE 1.1 **The basic system configuration of the landing gear R/E system.**

B. SIMULATION ANALYSIS OF SYSTEM FAILURE

From Fig. 1.1, the actuator cylinder is used to extend and retract the landing gear, and the system power is provided by the hydraulic source, so that the system failure is closely related to hydraulic pressure. In this paper, aiming at the landing gear extension stage, AMESim software [19-20] is used to simulate two common failure modes in system, which is the oil leakage and oil filter plugging.

The leakage degree of the actuator cylinder can be represented by the leakage coefficient set in the AMESim. As the leakage occurs, the dynamic pressure and flow rate in chamber will change, and then the leakage can be detected by the pressure and flow signal of the actuator cylinder.

Set the initial leakage coefficient as 0.01L/min/bar to represent the normal condition as well as the variation from 0.01L/min/bar to 0.02L/min/bar to represent the system performance degradation, and the cavity pressure, actuator cylinder flow and displacement variation are shown in Fig. 1.2-4.
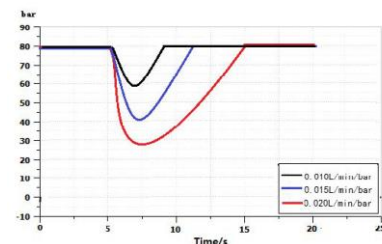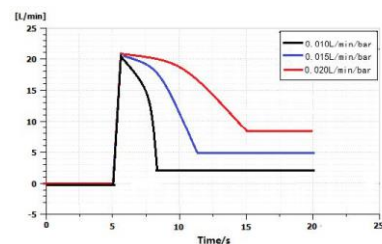

FIGURE 1.2 **The pressure curves**
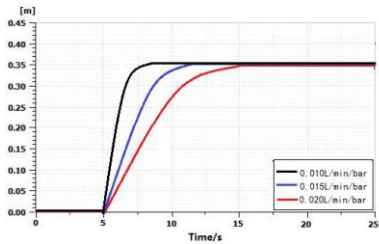

FIGURE 1.3 **The flow curves**

FIGURE 1.4 The displacement curves

As the Fig.1.2 shows, when the leakage coefficient is 0.01L/min/bar, the cavity pressure is the highest with landing gear's extension, and as the leakage increases, the pressure gradually decreases. And as the Fig.1.3-4 shows, the flow is the lowest when the leakage coefficient is 0.01L/min/bar, and with the leakage increasing, the flow is also increasing to meet the drive, as well as the increasing of the actuation time.

Oil filter plugging is also one of the failure modes in the R/E system. If the oil filter cannot be replaced in time, the flow will be insufficient and the system function will decline. Moreover, through the blockage of the oil filter, it is also possible to indirectly cause the wear of the components or contamination of the oil.

Oil filter plugging can be simulated by adjusting the equivalent pore size of the oil filter element. On the basis of one hydraulic system report which is from our research project, the equivalent pore size is set from 0.5 to 1 mm, and the simulation results are shown in Fig.1.5-6.
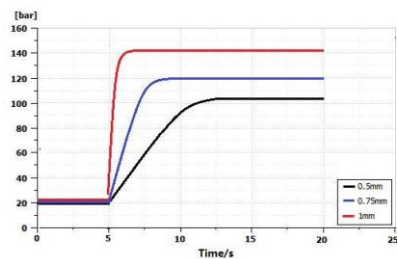

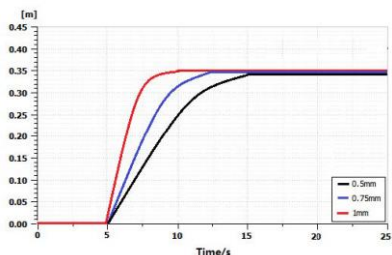**FIGURE 1.5 Pump output pressure**

FIGURE 1.6 Actuator displacements in different pore size

As the Fig.1.5-6 shows, when the oil filter pore size is varied from 0.5 to 1mm, the displacement of the actuator cylinder becomes less, and the pump outlet pressure becomes larger.

In all, the leakage and oil filter blockage have a great impact on system operation. And the simulation results show that the main hydraulic pressure value is highly sensitive to

the system performance degradation or failure such as system pressure loss, pressure fluctuation and system instability, so that the hydraulic pressure is closely related to the health status of the system.

III. STATUS FEATURE EXTRACTION AND DIMENSION REDUCTION

For the case airplane in this paper, the "main hydraulic pressure" sampling frequency is 1 Hz, which obviously limits to extract feature from the frequency domain, but it is feasible in time domain.

A. HEALTH STATUS FEATURE EXTRACTION

When the system is degrade or failed, the amplitude and waveform of the signal will change. Time domain feature of sampling data can be used to identify the system status better, and the statistical indexes of the time domain include mean, variance, standard deviation, root mean square, etc. And their detailed descriptions are shown in the Table I.

TABLE I
TIME DOMAIN FEATURE PARAMETERS

Feature Name	Statistical Feature Formula
Mean	$X_{mean} = \dfrac{1}{L}\sum_i x_i$
Variance	$X_{var} = \dfrac{1}{L}\sum_i (x_i - \bar{x})^2$
Standard deviation	$X_{std} = \sqrt{\dfrac{1}{L}\sum_i (x_i - \bar{x})^2}$
Peak-to-peak value	$X_{pk} = \max(x_1, x_2, \ldots x_L) - \min(x_1, x_2, \ldots x_L)$
Kurtosis parameter	$X_{ku} = \dfrac{1}{X_{rms}^4}\sum_i (x_i - \bar{x})^4$
Margin factor	$X_{clf} = X_{max} / X_{smr}$
Crest factor	$X_{cf} = X_{max} / X_{rms}$
Feature Name	Statistical Feature Formula
Kurtosis	$X_k = \dfrac{1}{L}\sum_i \dfrac{(x_i - \bar{x})^4}{X_{std}}$
Form factor	$X_{sf} = X_{rms} / \dfrac{1}{L}\sum_i |x_i|$
Coefficient of Variation	$X_{cv} = X_{std} / X_{mean}$
Skewness	$X_{sk} = \dfrac{1}{X_{rms}^3}\sum_i (x_i - \bar{x})^3$
Maximum	$X_{max} = \max(x_1, x_2, \ldots x_L)$
Minimum	$X_{min} = \min(x_1, x_2, \ldots x_L)$
Root-mean-square	$X_{rms} = \sqrt{\dfrac{1}{L}\sum_i x_i^2}$

The main hydraulic pressure curves in the landing gear R/E stage of the 261 vehicles are shown in Fig.2.1-2, and each vehicle is distinguished by different color.
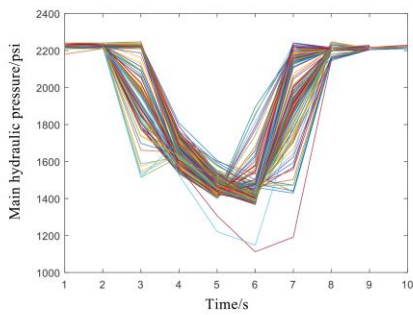
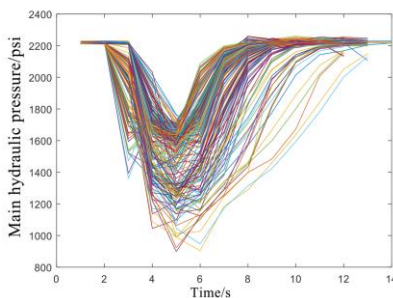**FIGURE 2.1 Pressure curves in retracting stage**

FIGURE 2.2 Pressure curves in extending stage

According to the feature parameter in Table I, the parameters of the main hydraulic pressure variation are simulated. And to save the article space, the partial parameters are shown in Fig.2.3-9.
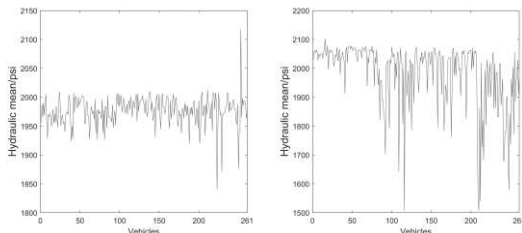

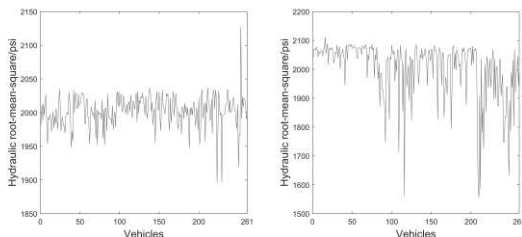**FIGURE 2.3 Mean of pressure in the R/E stage**

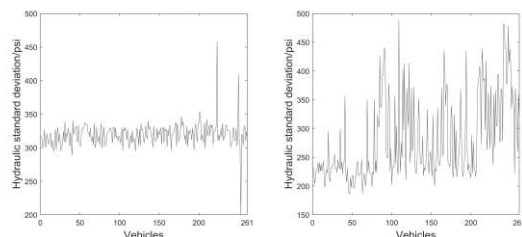**FIGURE 2.4 Root-mean- square of pressure in the R/E Stage**

FIGURE 2.5 Standard deviation of pressure in the R/E stage
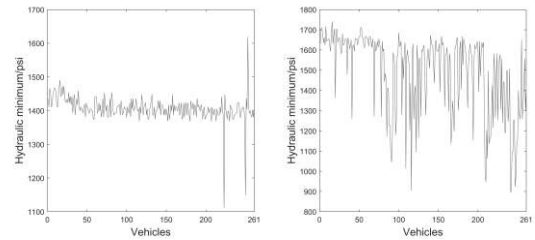

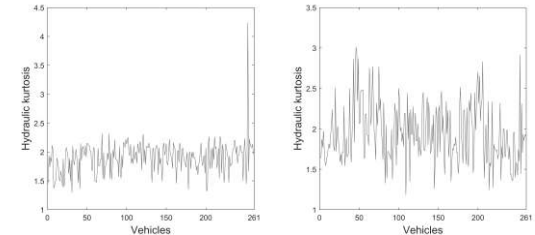**FIGURE 2.6 Minimum of pressure in the R/E stage**

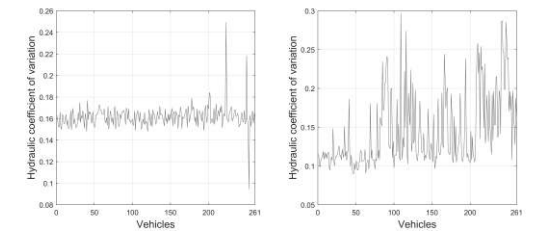**FIGURE 2.7 Kurtosis parameter of pressure in the R/E stage**

FIGURE 2.8 Coefficient of pressure in the R/E stage
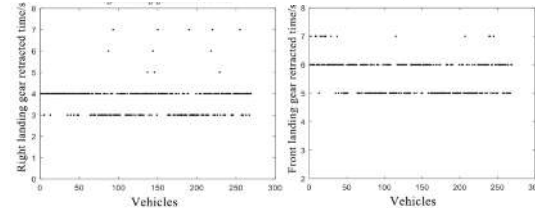
FIGURE 2.9 Nose (left) and right (right) landing gear retracted time

B. HEALTH STATUS FEATURE DIMENSION REDUCTION

In order to extract information effectively, it is necessary to use the main hydraulic pressure feature value to establish a high-dimensional space to fully express system health status. But massive data contain a lot of redundant information inevitability, which affects the efficiency of health status recognition. Therefore, the improved MDS algorithm is introduced for dimension reduction. Compared with the traditional linear dimension reduction algorithms, such as PCA (Principal Component Analysis) and LDA (Linear Discriminant Analysis), MDS has better fidelity effect on the nonlinear relationship in the original data, and the dimension reduction precision is higher.

The principle of multidimensional scaling [21] is to create a suitable low-dimensional space according to the similarity between paired data, and in this low-dimensional space, the distance and similarity between sample data are as consistent as possible.

According to the principle of MDS, Euclidean distance is used to describe the difference between sample data, and due

to this, data in high-dimensional space can be mapped into low-dimensional space. However, the Euclidean distance has some shortcomings in reflecting the overall data. When raw data have complex geometric features, as the Fig.2.10 shows, point a and c are in the same manifold spaces, and point a, b are not. However, the Euclidean distance between point a and b is smaller than a and c, which shows that the Euclidean distance has limitations in dealing with the data relationship. So the distance measure between samples must be considered in two aspects:

1. Local consistency, that is, the distance between adjacent data is small.
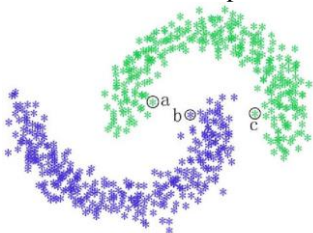2. Global consistency, that is, the distance between points on the same manifold space is small.

FIGURE 2.10 Manifold spaces

Therefore, considering the shortcomings of MDS and the advantage of manifold distance, a new MDS dimension reduction algorithm based on manifold distance measure is presented in this Section. Calculate the line segment length on the manifold space between the data sample points x_i and x_j as

$$L(x_i, x_j) = \rho^{d_{ij}} - 1 \qquad (2\text{-}1)$$

where d_{ij} is the Euclidean distance between x_i and x_j, and $\rho > 1$ is the scaling factor.

And the manifold distance can be expressed as:

$$D(x_i, x_j) = \min \sum_{k=1}^{l-1} L(p_k, p_{k+1}) \qquad (2\text{-}2)$$

where $1 \le k \le l - 1$.

Suppose characteristic data $X = [x_1, x_2, \ldots, x_N]$ is represented as $Y = [y_1, y_2, \ldots, y_N]$ and the manifold distance is $D(x_i, x_j)$ in the low dimensional space. So the dimension reduction loss function is $E = \sum (D(x_i, x_j) - D(y_i, y_j))^2$, and the optimized MDS algorithm is as follows:

Step 1: Calculate the distance matrix P according the formula 2-2, and find the difference matrix $D = P^2$. P is the distance matrix of the input feature set:

$$P = \begin{bmatrix} d_{11} & \cdots & d_{1n} \\ \vdots & \ddots & \vdots \\ d_{n1} & \cdots & d_{nn} \end{bmatrix} \qquad (2\text{-}3)$$

where d_{ij} represents the Euclidean distance between the i-th sample and the j-th sample.

Step 2: Construct the inner product matrix $B = -\frac{1}{2} JDJ$.

Where $J = I - \frac{1}{N} ee'$, N is the amount of the objects, I is the N-order unit matrix, and e is an all 1 column vector of length n.

Step 3: Extract m eigenvalues $\lambda_1 \ldots \lambda_m$ and corresponding m eigenvectors $v_1 \ldots v_m$ in matrix B.

Step 4: Obtain the dimension reduction result: the eigenvectors corresponding to the first k largest eigenvalues.

III. OPTIMIZED FUZZY C-MEANS CLUSTERING ALGORITHM

A. FCM CLUSTERING

The data-driven method is used to monitor the landing gear R/E system health status in this paper. The algorithm based on the objective function exactly fits this feature to process the flight data, and FCM [22] clustering is the typical one.

Define the data sample as x_k, the data element as $A = \{x_1, x_2, \ldots, x_n\}$, the number of x_k as s, and the feature vector value set of the data sample as $p(x_k) = (x_{k1}, x_{k2}, \ldots, x_{ks})$. The goal of fuzzy c-means clustering is to cluster n raw data samples into c subsets $A_1, \ldots, A_c$ by the membership function, which is shown as

$$\left\{ \mu_{ik} \mid \mu_{ik} = [0,1]; \sum_{i=1}^{c} \mu_{ik} = 1; 0 < \sum_{i=1}^{n} \mu_{ik} < 1, \forall i \right\} \qquad (3\text{-}1)$$

The principle of the sum of intra-class weighted squared errors is introduced as objective function of the FCM clustering algorithm, which is shown as follows:

$$J(U, V) = \sum_{k=1}^{n} \sum_{i=1}^{c} (\mu_{ik})^m (d_{ik})^2 \qquad (3\text{-}2)$$

where, $U = [\mu_{ik}]^{c \times n}$ is the fuzzy partition matrix. $V = \{v_1, v_2, \ldots, v_c\}$ is the fuzzy clustering center of various classes. μ_{ik} represents the membership of the k-th sample in the i-th class, d_{ik} is the Euclidean distance between sample x_k and clustering center v_i, and $m \in [1, +\infty)$ is the fuzzy weighting coefficient. The smaller the value of $J(U, V)$, the better the effect of clustering. The distance between the sample and the cluster center can be shown as

$$(d_{ik})^2 = \|x_k - v_i\|_A = (x_i - p_i)^T A(x_k - p_i) \qquad (3\text{-}3)$$

where, A is an $S \times S$-order symmetric positive definite matrix.

As the objective function is established, the fuzzy clustering analysis is transformed to an optimal programming problem, which is shown as

$$\min(J(U, V)) = \sum_{k=1}^{n} \min \left\{ \sum_{i=1}^{c} (\mu_{ik})^m (d_{ik})^2 \right\} \qquad (3\text{-}4)$$

And μ_{ik} should satisfy the formula 3-1.

Calculate the Eq.3-4 by Lagrange multiplier method to get

$$F = \sum_{i=1}^{c} (\mu_{ik})^m (d_{ik})^2 + \lambda (\sum_{i=1}^{c} \mu_{ik} - 1) \qquad (3\text{-}5)$$

And then use the derivative method to solve the Eq.3-5 and get the following relationship lastly:

$$\begin{cases} \mu_{ik} = \left[\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}} \right)^{\frac{2}{m-1}} \right]^{-1}, I_k = \varnothing \\ \mu_{ik} = 0, \forall i \in I_k^*, \sum_{i \in I_k} \mu_{ik} = 1, I_k \neq \varnothing \end{cases} \quad (3\text{-}6)$$

Therefore, the cluster center $V^{(l)}$ is:

$$v_i = \frac{\sum_{k=1}^{n} (\mu_{ik})^m \times x_k}{\sum_{k=1}^{n} (\mu_{ik})^m} \quad (3\text{-}7)$$

With the above principle, the modeling steps of the FCM algorithm can be obtained as follows:

Step 1: Establish the objective function Eq.3-4.

Step 2: Set m as the fuzzy weighting coefficient (which is set as 2 generally), c as number of cluster classes, ε as the basis for iterative decision (which is set as 0.01 generally), b as the maximum number of iterations. Initialize number of iterations $l = 0$, and the cluster center $V^{(l)}$.

Step 3: Calculate the membership matrix according to the Eq.3-6.

Step 4: Update the cluster center by the Eq.3-7.

Step 5: Judge whether $\left\| V^{(l+1)} - V^{(l)} \right\| < \varepsilon$ is true, if it's true, terminate the calculation and output cluster results, and if not, let $l = l + 1$, turn to step 3.

B. THE OPTIMIZATION OF FCM

Since the determination principle of the cluster center is to give the initial clustering center and then derive the result according to the iterative operation, in which the noise points may be selected and may cause the algorithm to fall into local optimum. What's more, the number of clusters is required at the beginning due to the FCM algorithm, which puts forward a priori requirements for clustering data, and may undermine the unsupervised nature of the algorithm. So if the number of clustering categories is not properly selected, it will greatly affect the accuracy of fuzzy clustering results.

In view of this, the following improvements are proposed:

1) IMPROVE THE OPTIMAL NUMBER OF CLUSTERS WITH GRANULARITY PRINCIPLE

The granularity represents a measure of the particle size of the object, which is introduced into the fuzzy cluster to represent the measure of the sample data information. Physical granularity is the basis for the division of specific physical quantities, while information granularity represents the metrics for refinement of sample data.

The degree of the information particle coupling $Cd(c)$ indicates the degree of tightness between the particles in the class. The smaller the coupling degree of the information particles is, and the closer the relationship between the particles in the class is, which is expressed as

$$Cd(c) = \frac{1}{n} \sum_{i=1}^{c} \sum_{j=1}^{n} \mu_{ij}^m d_{ij}^2, i = 1, 2 \ldots n \quad (3\text{-}8)$$

The information particle separation degree $Sd(c)$ indicates the separation property of the inter-class particles. The greater the separation degree is, the weaker the correlation between the inter-class particles and the better the separation degree are, which is calculated as

$$Sd(c) = \frac{\sum_{i,k=1, i \neq k}^{c} d_{ik}^2}{c(c-1)/2}, i, k = 1, 2, \ldots c \quad (3\text{-}9)$$

According to the degree of the information particle coupling and information particle separation, define $Gd(c)$ as the effectiveness function:

$$Gd(c) = \alpha Cd(c) + (1 - \alpha) \frac{1}{Sd(c)} \quad (3\text{-}10)$$

where α and $1 - \alpha$ are weighting factors to balance the degree of the information particle coupling and information particle separation. Between the degree of the information particle coupling and the information particle separation, the smaller value takes a larger weight, vice versa, thus the effects of the two degrees on the effectiveness function can be balanced. Therefore, if we want to obtain a more appropriate cluster number, the value of the effectiveness function should be as small as possible. And the enumeration method can be used to calculate the value of the effectiveness function under different cluster numbers to determine the optimal cluster number.

2) OPTIMIZE THE INITIAL CLUSTER CENTER POSITION

For the same data set, if the initial cluster center position is selected unreasonably, the algorithm may be partially optimal. The location of the cluster center is closely related to the sample space particle density, and this idea is used to optimize the initial clustering center of the data sample. Define $\rho_i^{(0)}$ as the particle density at point x_i in the data set:

$$\rho_i^{(0)} = \sum_{k=1}^{n} \frac{1}{1 + f_d \left\| x_i - x_k \right\|^2} \quad (3\text{-}11)$$

where

$$f_d = \frac{1}{r_d^2} \quad (3\text{-}12)$$

From the Eq.3-11, the more data near the point x_i will make the value of the density function greater. r_d is the effective radius of the data, its size is related to the distribution of the sample data, which means the average square root distance between the data points:

$$r_d^2 = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{k=1}^{n} \left\| x_i - x_k \right\|^2 \quad (3\text{-}13)$$

Finally, the initial cluster center optimization procedure is summarized as follows:

Step 1: Set c as the number of cluster categories, n as the total number of samples, and initialize the number of iterations $l = 1$.

Step 2: Calculate the density function value of each sample data.

Step 3: Select the density function maximum

$$D_1 = \max\left\{\rho_i^{(0)}, i = 1, 2, \ldots, n\right\}$$ as the first fuzzy clustering initial center.

Step 4: If $l > c$, the algorithm end, and if not, turn to step 5 to find the next cluster center.

Step 5: Since the density function value near the existing cluster center is very large. In order to avoid the repeated cluster centers, the relationship of the density function is adjusted as

$$\rho_i^{(k)} = \rho_i^{(k-1)} - D_{k-1} \frac{1}{1 + f_d \left\| x_i - x_{k-1} \right\|^2}, k = 2, 3, \ldots c \quad (3\text{-}14)$$

Step 6: $l = l + 1$, and turn to step 4.

C. CLUSTERING PERFORMANCE INDEX

Clustering performance index [23] is important to measure clustering effects, which can help optimize the goals more clearly to obtain the more qualified clustering result. And the clustering goal is to make intra-class similarity high and inter-class similarity low.

There are two types of clustering performance index. One is to compare the clustering result with a reference model, e.g. external index. And the other is to directly examine the clustering result without referring to any model, e.g. internal index.

1) EXTERNAL INDEX

Aiming at data set $D = \{x_1, x_2, \ldots, x_m\}$, we can assume that the clusters given by the reference model are divided into $C = \{C_1, C_2, \ldots, C_m\}$, and the clusters given by clustering are divided into $C^* = \{C_1^*, C_2^*, \ldots, C_m^*\}$. And let λ and λ^* as the data categories corresponding to C and C^* respectively. Define p^0 as total correct recognition rate

$$p^0 = \frac{\sum_{i=1}^{N} a_i}{N}, \text{ (if } \lambda = \lambda^*, a_i = 1, \text{ and if not, } a_i = 0) \quad (3\text{-}15)$$

where, N is the total number of sample data.

Define p^l as the recognition rate to category l:

$$p^l = \frac{\sum_{i=1}^{n_m^*} b_i}{n_m}, \text{ (if } \lambda = \lambda^*, b_i = 1, \text{and if not, } b_i = 0) \quad (3\text{-}16)$$

2) INTERNAL INDEX

Divide the clustering result clusters into $C = \{C_1, C_2, \ldots, C_m\}$, and define:

$$avg(C) = \frac{2}{|C|(|C|-1)} \sum_{1 \le i \le j \le |C|} dist(x_i, x_j) \quad (3\text{-}18)$$

$$diam(C) = \max_{1 \le i \le j \le |C|} dist(x_i, x_j) \quad (3\text{-}19)$$

$$d_{\min}(C_i, C_j) = \min_{x_i \in C_i, x_j \in C_j} dist(x_i, x_j) \quad (3\text{-}20)$$

$$d_{cen}(C_i, C_j) = dist(\mu_i, \mu_j) \quad (3\text{-}21)$$

where $dist(*)$ is used to calculate the distance between two samples. μ represents the center of clusters, and its calculation method is

$$\mu = \frac{1}{|C|} \sum_{1 \le i \le |C|} x_i \quad (3\text{-}22)$$

And $avg(C)$ represents the average distance between samples in cluster C, $diam(C)$ represents the farthest distance between samples in cluster C, $d_{\min}(C_i, C_j)$ represents the closest distance between samples, and $d_{cen}(C_i, C_j)$ represents the distance of the center of the cluster C_i and C_j.

Based on the above four Equations, the internal index of the cluster can be derived: Davies-Bouldin in Index (DBI) and Dunn Index(DI).

$$DBI = \frac{1}{k} \sum_{i=1}^{k} \max_{j \ne i} \left(\frac{avg(C_i) + avg(C_j)}{d_{cen}(C_i, C_j)} \right) \quad (3\text{-}23)$$

$$DI = \min_{1 \le i \le k} \left\{ \min_{j \ne i} \left(\frac{d_{\min}(C_i, C_j)}{\max_{1 \le i \le k} diam(C_l)} \right) \right\} \quad (3\text{-}24)$$

The DBI represents the intra-class precision and the smaller the better. And then the DI represents the degree of separation between classes, and the larger the better.

IV. HEALTH MONITORING OF LANDING GEAR R/E SYSTEM

A. HEALTH MONITORING PROCESS OF LANDING GEAR R/E SYSTEM

The health monitoring process of the landing gear R/E system mainly includes three aspects: health status extraction, status feature dimension reduction, and status recognition. That means, extract the feature of the main hydraulic pressure value, use the improved MDS algorithm to reduce the dimension, and use the improved FCM algorithm to classify and identify the status. This whole monitoring process is shown in Fig.4.1.
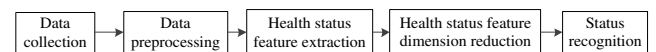

FIGURE 4.1 The Health monitoring process of the landing gear R/E system

B. CASE SIMULATION AND RESULT ANALYSIS

The original data is derived from the QAR data of 261 vehicles of a certain type of aircraft for three years, and part of them is shown in the Fig.4.2. The performance of the landing gear system is seriously degraded. And the original data of the aircraft landing gear retracting system is divided into three categories: normal data, degraded data and fault data. To count the original data of 261 vehicles, there are a total of 207 vehicles in the normal state, 49 vehicles in the degraded state, and 5 vehicles in the fault state.

1) DETERMINATION OF THE NUMBER OF FCM CLUSTERING CATEGORIES

Set the fuzzy weighting coefficient m to 2, the threshold ε to 0.01, the maximum number of iterations max_b to 100, the maximum number of classifications max_c to 10.Accroding to the principle of granularity, the coupling weight α is slightly larger than the separation weight $1-\alpha$ since the degradation of the system is less, and set α to 0.6. When the number of clusters c takes different values, the data obtained by dimensionality reduction is substituted into formula 3-10 to get the corresponding effectiveness function $Gd(c)$ value shown in the Table II. And the relationship image between the number of clusters c and $Gd(c)$ is shown in Fig.4.2.

TABLE II
GD(C) VALUE CORRESPONDING TO THE DIFFERENT CLUSTER NUMBERS C

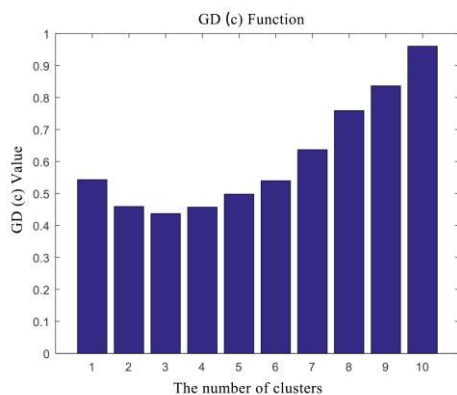| The number of clusters c	Gd(c) value
2	0.4792
3	0.4371
4	0.4572
5	0.4978
6	0.5398
7	0.6369
8	0.7590
9	0.8365
10	0.9605

FIGURE 4.2 The relationship between the number of clusters c and Gd(c)

From the results, the effectiveness function value reaches the minimum one when the number of clusters is 3, so that the effectiveness function is reasonable. The improved FCM algorithm can better solve the problem that the FCM algorithm cannot automatically get the optimal number of clusters and is sensitive to initialization. Compared with the traditional FCM algorithm, it has better clustering effect and the convergence is faster.

2) CLUSTERING IN DIFFERENT DIMENSIONALITY REDUCTION METHODS

Combined four different dimensionality reduction methods with FCM, the membership function and clustering results can be obtained and shown in the Fig.4.3-10.Where, Fig. 4.3 -6 are the membership of each category by four different dimensionality reduction methods, and Fig.4.7-10 are clustering results in each dimensionality reduction method. And for the convenience of writing and reading, the two cooperative algorithms (dimension reduction algorithm and FCM algorithm) are denoted as xxx_FCM, such as PCA_FCM.

From Fig.4.3-10, we can see that: based on the PCA_FCM algorithm, the five failures of the original data are separated, but a small part of the health and sub-health data has poor separation. For the LDA_FCM algorithm, the membership function is not sufficiently separated, and only 4 of the 5 vehicles in the fault categories are recognized. Based on the MDS_FCM algorithm, fault data can be fully identified, and the separation of health and sub-health data is also good. And based on the improved MDS_FCM algorithm, the fight parameter data can be clearly separated and the membership function value is so close to 1. The fault data can be fully identified, and the effect of separating health and sub-health data is totally obvious. As a result, the improved MDS algorithm can better identify the information in the eigenvalues of the original data, and has a high separation degree for different health states.
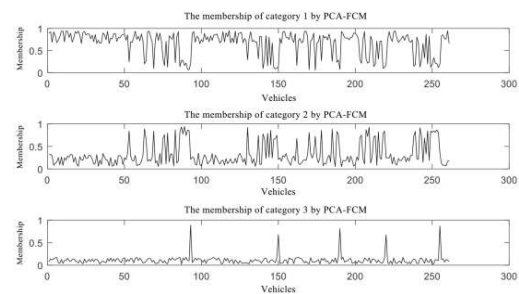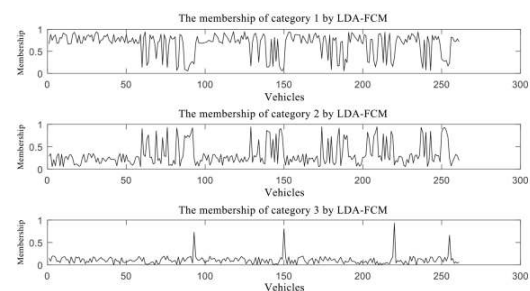

FIGURE 4.3 The membership of each category by PCA-FCM

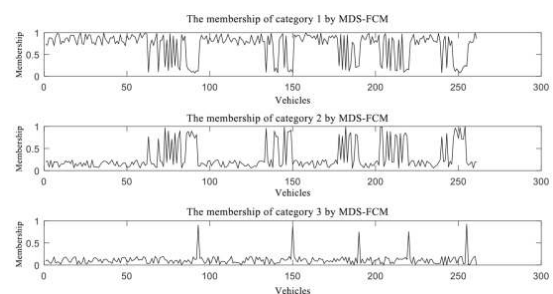FIGURE 4.4 The membership of each category by LDA-FCM

IEEE *Access*
Multidisciplinary : Rapid Review : Open Access Journal

FIGURE 4.5 The Membership of each category by MDS-FCM
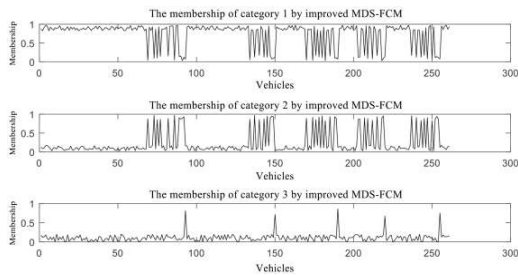

FIGURE 4.6 The membership of each category by optimized MDS-FCM
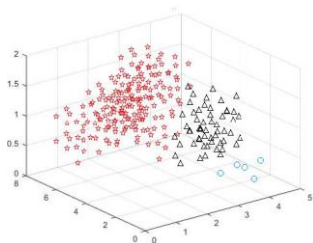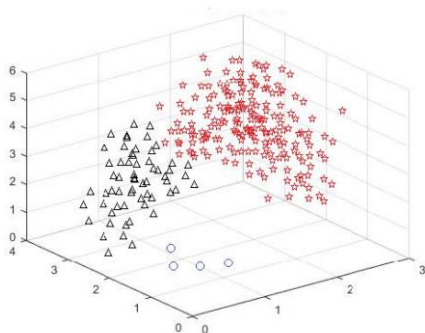

FIGURE 4.7 Clustering results in PLA-FCM

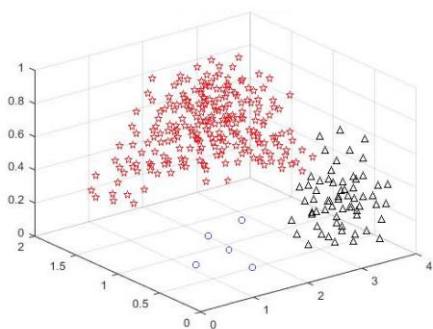**FIGURE 4.8 Clustering results in LDA-FCM**

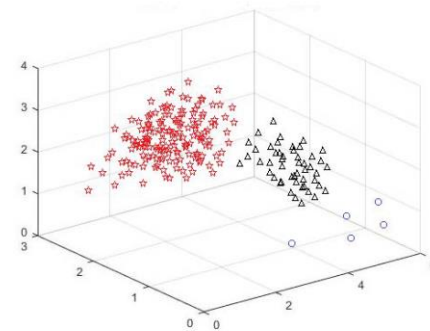**FIGURE 4.9 Clustering results in MDS-FCM**

FIGURE 4.10 Clustering results in optimized MDS-FCM

For the clustering results of the four algorithms, the statistical results of the total correct recognition rate, failure recognition rate, sub-health recognition rate, the false alarm rate that misidentified the health state to sub-health state, and the internal evaluation indexes DBI and DI are shown in Table III-V:

TABLE III
STATISTICS OF CLUSTERING RESULTS

The number of each algorithm correct recognition	Health data	Sub-health data	Fault data
The number of PCA_FCM correct recognition	184	38	5
The number of LDA_FCM correct recognition	168	33	4
The number of MDS_FCM correct recognition	190	40	5
The number of improved MDS_FCM correct recognition	203	47	5

TABLE IV
CORRECT RATE STATISTICS OF CLUSTERING RESULTS

Clustering algorithm	Total correct recognition rate	Failure recognition rate	Sub-health recognition rate	The false alarm rate
PCA_FCM	89.2%	100%	83.6%	32.7%
LDA_FCM	82.3%	80%	77.5%	44.1%
MDS_FCM	92.3%	100%	87.5%	24.5%
Improved MDS_FCM	97.7%	100%	95.9%	7.8%

The results of the recognition rates of several clustering methods are shown in the Table III and IV. Comparing the recognition rates of the four algorithms, it can be seen that the improved MDS_FCM algorithm has the highest correct rate for identifying the health status of the landing gear R/E process, reaching 97.7%, and failures can be fully identified. The recognition rate of sub-health vehicles can reach 95.9%, which is particularly prominent compared to the other three algorithms, and the false alarm rate is only 7.8%, which is also significantly lower than the other three algorithms. It shows that the improved MDS_FCM algorithm can effectively extract the key information in the eigenvalues. While achieving dimensionality reduction and discarding

redundant information, the non-linear information in the high-dimensional data is retained, and the purpose of utilizing the manifold distance to optimize the MDS algorithm is achieved.

TABLE V
CLUSTERING INTERNAL EVALUATION INDEXES

Clustering algorithm	DBI	DI
PCA_FCM	0.31	4.31
LDA_FCM	0.35	3.76
MDS_FCM	0.25	4.98
Improved MDS_FCM	0.16	6.34

The internal evaluation indexes DBI and DI are summarized in the Table V, and the results show that the improved MDS_FCM method can better concentrate the same kind of data, and make different types of data more separated.

V. CONCLUSION

Based on the characteristics of flight parameter data, this paper extracts the time domain characteristic value of the main hydraulic pressure in the landing gear R/E period, uses different algorithms to reduce the characteristic features, and then uses the FCM algorithm to identify the sample data category. An improved MDS dimensionality reduction algorithm is given to optimize the shortcomings of the PCA, LDA, and MDS dimensionality reduction methods. Aiming at the limitations of the FCM algorithm, the granularity principle and density function are respectively utilized to improve the number of FCM algorithm clusters and the initial cluster center. And the results show that the improved algorithm greatly improves the accuracy of the clustering algorithm, and the algorithm model established in this paper can effectively assist the health evaluation and decision making of the landing gear R/E systems.

REFERENCES

[1] L. Chen, Analysis on the Kinematic and Dynamic Performance of Landing Gear, Nanjing University of Aeronautics and Astronautics, Nanjing, 2007.

[2] R. Fan, Analysis and Test for Retraction-Extension Dynamic Performance of a Civil Aircraft Main Landing Gear, Nanjing University of Aeronautics and Astronautics, Nanjing, 2012.

[3] B. Sun, R. Kang, and J.S. Xie, "Research and application of the prognostic and health management system," Systems Engineering and Electronics, vol.29, no.10, 2007.

[4] David Lira Nuñez, Milton Borsato, "OntoProg: An ontology-based model for implementing Prognostics Health Management in mechanical machines," Advanced Engineering Informatics, vol.38, pp.746-759, 2018.

[5] R. Li, Wim J.C. Verhagen, Richard Curran, "Toward a methodology of requirements definition for prognostics and health management system to support aircraft predictive maintenance," Aerospace Science and Technology, vol.102, 2020.

[6] S. K. Zeng, Michael G. Pecht, J. Wu, "Status and Perspectives of Prognostics and Health Management Technologies," Acta Aeronautica ET Astronautica Sinica, vol.26, no.5, pp.626-632,2005.

[7] M. L. Suo, B. L. Zhu and D. Zhou, "Neighborhood grid clustering and its application in fault diagnosis of satellite power system," Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, vol.233, pp.1270-1283, 2018.

[8] D. Y. Gao, Z. J. Wu ,L. Yang. "Structural health monitoring for long-term aircraft storage tanks under cryogenic temperature," Aerospace Science and Technology, vol.92, pp.881-891, 2019.

[9] C. Kamali, Amitabh Saraf, and A. A. Pashilkar, "Online Health Monitoring of Simplex Inertial Navigation System of a High Performance Aircraft," IFAC Proceedings Volumes, vol.45, no.1, pp.1-6, 2012.

[10] Seref Demirci,Chingiz Hajiyev,and Andreas Schwenke, "Fuzzy logic-based automated engine health monitoring for commercial aircraft," Aircraft Engineering & Aerospace Technology, vol.80, no.5, pp.516-525, 2008.

[11] Y. Yang, Aircraft landing gear extension and retraction control system diagnostics prognostics and health management, PhD Thesis, Cranfield University, UK, 2012.

[12] J. Chen, C. B. Ma, and D. Song, "Multiple failure prognosis of landing gear retraction/extension system based on H filtering," Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, vol.229, no.8, pp.1543-1555, 2015.

[13] C. S. Byington, M. Wastson, and D. Edwards, "Data-driven neural network methodology to remaining life predictions for aircraft actuator components," IEEE aerospace conference proceedings, vol.6, pp.3581–3589, 2004.

[14] L. He, L. Y. Liang, C. B. Ma, "Multiple Failure Simulation and Health Evaluation of Aircraft Landing Gear Hydraulic Retraction/Extension System," Journal of Northwestern Polytechnical University, vol.6, pp.990-995, 2016.

[15] Slawomir Wierzchoń, Mieczyslaw Kłopotek, "Modern Algorithms of Cluster Analysis," Springer, Cham, Switzerland, 2018.

[16] Paul Phillips, Dominic Diston and Andrew Starr, "Perspectives on the commercial development of landing gear health monitoring systems," Transportation Research Part C: Emerging Technologies, vol.19, pp.1339–1352, 2011.

[17] J. J. Liu, M. Zhang, H. Wang, "Sensor Fault Detection and Diagnosis Method for AHU Using 1-D CNN and Clustering Analysis," Computational Intelligence and Neuroscience, 2019.

[18] K. Y. Yeung, C. Fraley, A. Murua, et al, "Model-based clustering and data transformations for gene expression data," Bioinformatics, vol.17, no.10, pp.11, 2001.

[19] Manimaran and Hiremath, "Dynamic modeling, simulation and experimental investigation on cryogenic tank pressurization system," International Journal of Modelling and Simulation, vol.40, no.4, pp.291-307, 2020.

[20] Nicolae Vasiliu, Daniela Vasiliu, Constantin CĂLINOIU, et al, "Simulation of Fluid Power Systems with Simcenter Amesim," CRC Press, Boca Raton, 2018.

[21] D. Shang, P. J. Shang, and L. Liu, "Multidimensional scaling method for complex time series feature classification based on generalized complexity-invariant distance," Nonlinear Dynamics, vol.95, no.4 pp.2875-2892, 2019.

[22] Y. Zhou and Q. C. Ren, "Fuzzy c-means clustering algorithm for performance improvement of ENN," Cluster Computing, vol.22, no.5, pp.11163-11174, 2019.

[23] Biao. Li, B. P. Tang, L. Deng, and X. X. Yu, "Multiscale dynamic fusion prototypical cluster network for fault diagnosis of planetary gearbox under few labeled samples," Computers in Industry, vol. 123, 2020.

AUTHORS

JIE CHEN received the B.S. degree in electronic information engineering, M.S. degree in power electronics and power transmission, and Ph.D. degree in control science and engineering from Northwestern Polytechnical University, Xi'an, China, in 2004, 2007, and 2011. Since 2015, he has been an Associate Professor with the school of aeronautics, Northwestern Polytechnical University, Xi'an, China. His research interests include airworthiness technology and management, prognostics and

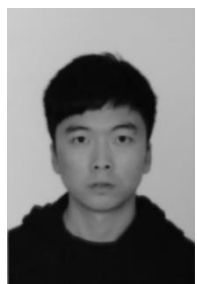health management, and aircraft fault diagnosis. He has authored many papers on high-quality journals.

SENYAO CHEN received the B.S. degree in electronic information engineering from Northwestern Polytechnical University, Xi'an, China in 2018, where he is currently pursuing the M. S. degree in aeronautical engineering.

He participates in some national projects. His research interests include management, prognostics and health management, and aircraft fault diagnosis

ZHENBAO LIU (M'11) received the B.S. and M.S. degree in electrical engineering and automation from Northwestern Polytechnical University, Xi'an, China, in 2001 and 2004, respectively, and the Ph.D. degree in electrical engineering and automation from the University of Tsukuba, Tsukuba, Japan, in 2009. He is currently a Professor with Northwestern Polytechnical University. He was a Visiting Scholar with Simon Fraser University, Canada, in 2012. His research interests include UAV, prognostics and health management, and aircraft fault diagnosis.

CAIKUN LUO received the B.S. degree in electronic information engineering and M.S. degree in aeronautical engineering from Northwestern Polytechnical University, Xi'an, China in 2016 and 2019,

He participates in some national projects. His research interests include management, prognostics and health management, and aircraft fault diagnosis.

ZHENGDONG JING received the B.S. degree in electronic information engineering from Northwestern Polytechnical University, Xi'an, China in 2018, where he is currently pursuing the M. S. degree in aeronautical engineering.

He participates in some national projects. His research interests include management, prognostics and health management, and aircraft fault diagnosis.

QINGSHAN XU received the B.S. degree in electronic information engineering from Nanchang Hangkong University, Nanchang, China in 2018, where he is currently pursuing the M. S. degree in aeronautical engineering.

He participates in some national projects. His research interests include management, prognostics and health management, and aircraft fault diagnosis