Chinese Journal of Aeronautics, (2014),27(5): 1086–1096

Chinese Society of Aeronautics and Astronautics
& Beihang University
Chinese Journal of Aeronautics

cja@buaa.edu.cn
www.sciencedirect.com

Remaining useful life prognostics for aeroengine based on superstatistics and information fusion

Liu Junqiang *, Zhang Malan, Zuo Hongfu, Xie Jiwei

College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Received 1 November 2013; revised 30 December 2013; accepted 22 April 2014
Available online 6 September 2014

KEYWORDS

Degradation;
Information fusion;
Kalman filtering;
Performance;
Prognostics;
Remaining useful life;
Superstatistics

Abstract Remaining useful life (RUL) prognostics is a fundamental premise to perform condition-based maintenance (CBM) for a system subject to performance degradation. Over the past decades, research has been conducted in RUL prognostics for aeroengine. However, most of the prognostics technologies and methods simply base on single parameter, making it hard to demonstrate the specific characteristics of its degradation. To solve such problems, this paper proposes a novel approach to predict RUL by means of superstatistics and information fusion. The performance degradation evolution of the engine is modeled by fusing multiple monitoring parameters, which manifest non-stationary characteristics while degrading. With the obtained degradation curve, prognostics model can be established by state-space method, and then RUL can be estimated when the time-varying parameters of the model are predicted and updated through Kalman filtering algorithm. By this method, the non-stationary degradation of each parameter is represented, and multiple monitoring parameters are incorporated, both contributing to the final prognostics. A case study shows that this approach enables satisfactory prediction evolution and achieves a markedly better prognosis of RUL.

1. Introduction

The statistics analysis[1] on China civil aviation incidents from 2006 to the first half of 2012 shows that, the number of flight incidents in the first half of 2012 appears a year-on-year growth of 26%. Importantly, the number of these incidents simply caused by engine cut-off rises 83%. Apparently, as the heart of an aircraft, aeroengine directly determines flight safety, thus needs to be repaired or replaced promptly. However, premature repair or replacement will inevitably lead to an increase in airlines' operation cost. Therefore, the condition-based maintenance (CBM) should be performed at a perfect moment, which can be identified according to the balance between safety and efficiency. To achieve this goal, more and more attention has been paid to remaining useful life (RUL) prognostics for aeroengine.

Considering the fact that taking time-consuming and repeated tests for an aeroengine is not affordable, research relying on monitoring parameters has been conducted in RUL prognostics using a variety of methodologies. The basis

* Corresponding author. Tel.: +86 25 84896260.
E-mail address: liujunqiang@nuaa.edu.cn (J. Liu).

Peer review under responsibility of Editorial Committe of CJA.

ELSEVIER Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.cja.2014.08.013

of these methodologies is going through a transition from single parameter to multiple parameters, which is shown in the following discussion.

In order to infer remaining useful life, fatigue crack growth function is created based on Paris model in Ref.[2]. A certain component is adopted for RUL prognostics by applying Bayesian framework in updating the parameters of its exponential degradation model.[3] Furthermore, exhaust gas temperature margin (EGTM) is taken as a measure of the engine's hidden health state, and the estimation is carried out through combining Bayesian approach with immune particle swarm optimization algorithm.[4] According to the evolution of EGTM, a performance degradation model is developed in Ref.[5]. However, all the methodologies mentioned above only take into account single parameter thus suffering from inaccuracy, which makes the prognosis of RUL unacceptable for CBM. Hence, to take the increasing complexity and integration of aeroengine into consideration, multiple parameters are utilized to achieve higher prediction accuracy in other studies. Multivariate time series analysis method is developed to estimate the performance reliability of the system by means of multiple failure modes.[6] Principal components method is adopted in the performance degradation analysis in Ref.[7]. With several accelerating variables, a stochastic process describing degradation is combined into a generalized cumulative damage approach, and new accelerated life test models are presented in Ref.[8]. Multivariate Wiener process is made use of in Ref.[9], assuming the variables are independent from each other. While, the correlation between multiple performance parameters is taken into account in the establishment of reliability assessment model in Ref.[10], and the accuracy of this approach is testified via simulation. Furthermore, multiple degradation characteristics are applied in reliability analysis in Ref.[11].

Besides this transition, the specific techniques and methods of the prognostics are being developed from traditional ones to artificial intelligence. Traditionally, prognostics have been achieved through data fitting and regression analysis.[12,13] While, artificial intelligence is shown in the application of Kalman filtering algorithm,[14] neural networks combined with chaotic particle swarm optimization,[15] Bayesian state estimation combined with state-space method et al.[16] In comparison, artificial intelligence methods are superior in terms of automatic calculation.

Nevertheless, the methodologies described above have not taken into account the non-stationary degradation process of the system. There is a pressing need to demonstrate the realistic degradation characteristics in order to achieve higher prediction accuracy.

Non-stationary process is defined as a stochastic process whose joint probability distribution changes when shifted in time or space. Generally, the distribution function of the observed degradation data of aeroengine shows no change, but the parameters of the function differ within different time windows. To be specific, the observed data changes slowly in early degradation and rapidly in severe degradation which indicates failure. Obviously, the degradation process of the engine is absolutely non-stationary.[17–19] Research on the variation of the parameters of the distribution function for non-stationary time series is exactly the application of superstatistics theory, also known as ''statistics of statistics''. This theory is originally proposed by Christian Beck et al in 2003, and then utilized to describe complex non-equilibrium systems. Superstatistics theory is initially and mainly applied to dynamical systems, and then adopted in the analysis of cosmic rays, regional climate, anomaly detection of network traffic and other fields. The most significant advantage of resorting to this theory is that it enables the presentation of the non-stationary characteristics of the degrading system.

In order to overcome the limitations of the existing methodologies, this paper presents a novel approach to RUL prognostics by means of superstatistics and information fusion. In this method, the non-stationary degradation processes of multiple parameters are incorporated in the establishment of the prognostics model, and the calculation is achieved by Kalman filtering algorithm, thereby enabling high prediction accuracy.

2. Prognostics framework

The novel approach to RUL prognostics is proposed on the basis of the considerations described as follows:

(1) For the purpose of tracking the realistic degradation evolution process of the engine, the non-stationary characteristics of each monitoring parameter should be taken into account.
(2) To avoid the disadvantages, such as inaccuracy, caused by methodologies which simply depend on single parameter, multiple monitoring parameters should be fused based on their correlations.
(3) As the observed data is contaminated with noise, it is better to add noise factor into calculation in order to guarantee prediction accuracy.
(4) Considering the time-varying characteristics of the entire non-stationary degradation process, RUL determination should be made in accordance with the consecutively introduced observations.

According to the understandings presented above, RUL prognostics framework is constructed in Fig. 1.

3. Degradation description based on superstatistics and information fusion

Generally speaking, the inherent health state of aeroengine is invisible, but with the increase of service time, it will be reflected on the variation of the performance monitoring parameters, including gas path performance monitoring parameters, lubrication oil parameters and vibration monitoring parameters.[20] Experientially, gas path performance monitoring parameters play a decisive role in measuring the underlying health state, because with the degradation of the system, the temperature and pressure on gas path components will gradually approach their threshold values, respectively. Accordingly, this paper adopts seven characterization parameters[21] from gas path components for prognostics. These characterization parameters are total temperature at low pressure compressor (LPC) outlet (T_{24}, °R), total temperature at high pressure compressor (HPC) outlet (T_{30}, °R), total temperature at low pressure turbine (LPT) outlet (T_{50}, °R), total pressure in bypass duct (P_{15}, psia), total pressure at HPC outlet (P_{30}, psia), physical fan speed (N_{f}, r/min) and physical core speed (N_{c}, r/min).

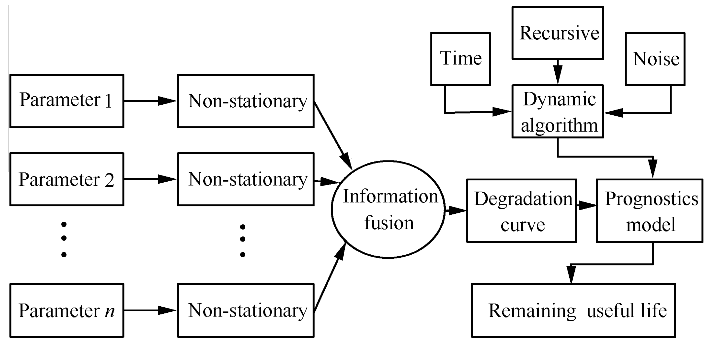

Fig. 1 RUL prognostics framework for aeroengine.

Due to the complexity of the engine, it is necessary to take stationarity test for performance parameters. Stationarity refers to the characteristics of a stochastic process whose statistical properties, such as the mean and the variance, do not change over time or position.[22,23] Otherwise, it is called non-stationary process which usually shows a linear or cyclical trend. Stationarity test is to identify whether the observed time series of these monitoring parameters are stationary or not. The analysis of the observed data from T_{24} indicates that the temperature shows an increase of 0.5 °R during the first 200 cycles (blue dotted line), while 1 °R within the following 100 cycles (red dotted line). Obviously, the mean and the variance of this degradation process change significantly over time, meaning that the observed time series of T_{24} are non-stationary. The non-stationary time series from T_{24} is given in Fig. 2.

If the prognostics approach starts with unprocessed non-stationary time series regardless of the existing anomaly, there will be two problems. First, the realistic performance degradation level cannot be accurately revealed, because the anomalous behavior of the system is not taken into account while constructing training sample set. Second, due to the uncertainty of the anomaly, it is hard to infer the degradation trend thus leading to inaccurate prediction.

To address these problems, superstatistics theory is combined into the novel approach. Superstatistics theory refers to the statistics of statistics, and it is performed to study the original time series by investigating the statistical properties of the stationary time series generated through division of the non-stationary time series. Slow variable is a relatively core concept of superstatistics theory, and also an intuitive expression of it. The so called slow variable refers to the powerful fluctuations of a system in a large time scale, which determines the inherent change of the system. In the anomaly detection of network traffic, superstatistics theory has been taken into full application. Firstly, the time window is divided according to the distribution of the time series. Secondly, distribution model is constructed based on the statistical properties of the observed data within each time window. Ultimately, the parameters of the statistical model are identified as slow variables, which are used to detect anomaly. This case makes it clear that traditional applications of superstatistics theory have to assume a distribution model for the observed data, and then take the parameters of the model as slow variables to realize anomaly detection. Hence, there are two challenges for traditional approaches. First, an appropriate distribution model does not always exist for all kinds of time series; second, the tedious inspection process of the distribution model may result in computational complexity. To avoid such problems, superstatistics can be utilized differently in this paper according to the following procedure. In the first place, a segmentation algorithm is utilized to divide the time window of the non-stationary time series into several time windows where the time series are stationary. Then, the change rates of the time series within different time windows are directly taken as slow variables to pinpoint anomaly. The break points used to divide the whole time window are obtained by means of Bernaola Galvan's heuristic segmentation algorithm (BG algorithm).[24,25] The equations of BG algorithm are given by

$$\begin{cases} S_{\mathrm{D}}(i) = \left[\dfrac{(N_1(i)-1)S_1(i)^2 + (N_2(i)-1)S_2(i)^2}{N_1(i)+N_2(i)-2}\right]^{\frac{1}{2}} \\ \qquad \times \left[\dfrac{1}{N_1(i)} + \dfrac{1}{N_2(i)}\right]^{\frac{1}{2}} \\ T(i) = \left|\dfrac{u_1(i)-u_2(i)}{S_{\mathrm{D}}(i)}\right| \end{cases} \tag{1}$$

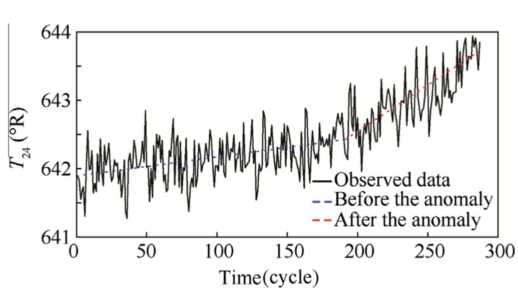

Fig. 2 Non-stationary time series from T_{24}.

where $N_1(i)$ and $N_2(i)$ denote the numbers of time points of the left and right part of point i(i is a positive integer, $i = 1, 2, \ldots, N-1$, and N is the number of time points.), respectively; $u_1(i)$ and $u_2(i)$ are the mean values of each part; S_1 (i) and $S_2(i)$ are the standard deviations of each part; $S_{\mathrm{D}}(i)$ is called combined standard deviation; $T(i)$ denotes test statistics which indicates the difference of the two parts, and the bigger $T(i)$ is, the more different the distribution characteristics of the two parts are.

This algorithm is targeted at performing anomaly detection within the whole time window by identifying the maximum $T(i)$ which corresponds to the break point i. Then i is used to divide the whole time window into two, meaning the non-stationary time series become two parts of relatively stationary ones. If the time series within each time window are still non-stationary, the segmentation should be performed again. The number of time windows depends on the practical application and whether the time series of each sub-window are stationary or not.

The break points generated from BG algorithm can be utilized for segmentation, but not qualified to serve as anomalies which indicate the underlying performance change of the system. To determine the actual anomaly of each parameter, slow variables can be calculated according to

$$S_{\mathrm{V}} = \frac{\Delta S}{\Delta t} = \frac{S_j - S_i}{t_j - t_i} \tag{2}$$

where S_{V} is slow variable, S_j the observed date at time t_j, S_i the observed date at time t_i and ΔS the variation of the observed data of each monitoring parameter within the corresponding time interval Δt; $j > i(i, j = 1, 2, \ldots, N)$, N is the number of time points. Theoretically, the smaller the time interval is, the more details of slow variables can be revealed. However, due to the noise impact, the fluctuations of slow variables will increase with the narrowing of the time interval. For the purpose of obtaining a relative variation (from one steady state to another steady state transition) of slow variables as well as avoiding the impact caused by noise, the time window should be adjusted appropriately. And then, the slow variables within the entire time window can be calculated by sliding time window Δt.

Taking monitoring parameter T_{24} for example, the break points (calculated by Eq. (1)) are the 132th cycle (break point 1), the 215th cycle (break point 2) and the 271th cycle (break point 3), shown in Fig. 3. The abscissa is time window, and the ordinate denotes slow variable. For a simplified illustration, the time interval Δt equals 10 cycles, and 30 sub-windows are evenly chosen from all the time windows. It is shown that the slow variables prior to break point 2 obey a normal distribution, and the slow variables posterior to break point 2 obey another normal distribution with a mean value differing by 0.08 from that of the former one. Therefore, the anomaly of T_{24} is break point 2 which corresponds to the 215th cycle. Break point 1 and break point 3 are identified as fake anomalies, because they are located in the steady fluctuations of the slow variables.

For the purpose of obtaining a comprehensive indicator to measure the hidden health state, the relationship between the statistical properties of the observed data and the inherent health state of the engine needs to be established by means of information fusion[26] for these multiple parameters. In that

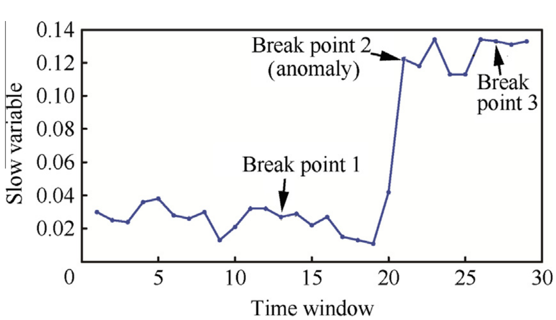

Fig. 3 Non-stationary time series from T_{24}.

case, the performance degradation can be ultimately tracked. Information fusion is the process of integrating multiple data into a consistent and useful representation. Once the anomaly detection for each parameter is carried out, the average anomalies of multiple parameters can be taken as the final break points to divide the time window. After the division is completed, the fusion can be executed according to the stationary time series within each time window. The comprehensive indicator is defined as health indicator (HI), assuming that HI equals 1 when there is no degradation, and HI equals 0 when the engine is in severe degradation. The theories of multiple linear regression has been studied and developed maturely, so given the statistical characteristics of the observed data, the parameters are fused through multiple linear regression model[27,28] (a detailed explanation of applying linear method to information fusion will be given in the case study) given by

$$\begin{cases} y_t^1 = a_1^0 + \boldsymbol{X}(t)\boldsymbol{A}_1^{\mathrm{T}} \\ y_t^2 = a_2^0 + \boldsymbol{X}(t)\boldsymbol{A}_2^{\mathrm{T}} \\ \qquad \vdots \\ y_t^n = a_n^0 + \boldsymbol{X}(t)\boldsymbol{A}_n^{\mathrm{T}} \end{cases} \tag{3}$$

where y_t^1 denote HI time series prior to the 1st anomaly; y_t^2 denote HI time series between the 1st anomaly and the 2nd anomaly; y_t^n denote HI time series between the $(n-1)$th anomaly and the nth anomaly, and n is the number of anomalies; $\boldsymbol{X}(t)$ is the vector of multiple parameters $[x_1(t), x_2\ (t), \ldots, x_l(t)]$, and l is the number of parameters; $a_1^0, \boldsymbol{A}_1$ are the coefficients of the model corresponding to y_t^1, and $\boldsymbol{A}_1 = [a_1^1, a_1^2, \ldots, a_1^l]$; $a_2^0, \boldsymbol{A}_2$ are the coefficients of the model corresponding to y_t^2, and $\boldsymbol{A}_2 = [a_2^1, a_2^2, \ldots, a_2^l]$; $a_n^0, \boldsymbol{A}_n$ are the coefficients of the model corresponding to y_t^n, and $\boldsymbol{A}_n = [a_n^1, a_n^2, \ldots, a_n^l]$. All the coefficients can be obtained via training sample set and sample data. For training[21], we assume HI equals 1 for the first five hours, 0 for the last five hours, and $\boldsymbol{k}_{A_n}$ for the five hours near the nth anomaly ($\boldsymbol{k}_{A_n} = (1 - m_{A_n})\boldsymbol{I}$), and m_{A_n} denotes the average change rate at the nth average anomaly, $\boldsymbol{I}$ is unit matrix. Training sample set $\boldsymbol{M}$ is expressed by

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{X}_{\boldsymbol{k}^1} & \boldsymbol{k}^1 \\ \boldsymbol{X}_{\boldsymbol{k}_{A_1}} & \boldsymbol{k}_{A_1} \\ \boldsymbol{X}_{\boldsymbol{k}_{A_2}} & \boldsymbol{k}_{A_2} \\ \vdots & \vdots \\ \boldsymbol{X}_{\boldsymbol{k}_{A_n}} & \boldsymbol{k}_{A_n} \\ \boldsymbol{X}_{\boldsymbol{k}^0} & \boldsymbol{k}^0 \end{bmatrix} \tag{4}$$

where $\boldsymbol{X}_{\boldsymbol{k}^1}$ is the first five sets of observed data of the parameters whose HI at the first five time points equals 1; $\boldsymbol{X}_{\boldsymbol{k}_{A_n}}$ is the five sets of observed data of the parameters whose HI at the five time points near the nth anomaly equals $1 - m_{A_n}$; and $\boldsymbol{X}_{\boldsymbol{k}^0}$ is the last five sets of observed data of the parameters whose HI at the last five time points equals 0; $\boldsymbol{k}^1 = [1, 1, 1, 1, 1]^{\mathrm{T}}$, $\boldsymbol{k}_{A_n} = (1 - m_{A_n})\boldsymbol{I}$ and $\boldsymbol{k}^0 = [0, 0, 0, 0, 0]^{\mathrm{T}}$. With the implementation of sample training, the coefficients $a_n^0, \boldsymbol{A}_n$ can be acquired, thereby enabling the calculation of HI.

As anomaly detection is carried out before the fusion, HI time series are capable of characterizing the non-stationary process of multiple parameters. This can be testified by analyzing the slow variables of HI. The slow variables obey one normal distribution prior to the anomaly and another normal distribution posterior to the anomaly (seen in the case study). Such distribution of the slow variables is sufficient to manifest the significant variation in the change rate of the performance degradation, meaning that the degradation process is non-stationary with an anomaly. If the fusion is performed without any anomaly information, the fused HI would be unable to describe the non-stationary characteristics of the degradation process, because the direct fusion does not take into account the observed data near anomaly when constructing training sample set.

In conclusion, the fused HI manages to not only demonstrate the non-stationary characteristics of monitoring parameters, but also incorporate such characteristics of each parameter by means of information fusion. In this way, the realistic degradation curve can be drawn more accurately.

4. Prognostics

4.1. Prognostics model

The novel approach to RUL prognostics consists of two major steps: obtaining the non-stationary performance degradation curve of the engine and establishing prognostics model and implement prediction. As the first step can be completed by methods introduced above, this section focuses on the second step. To that end, state-space method combined with Kalman filtering algorithm[29] is utilized for modeling and calculation.

State-space method is composed of state equation (Eq. (5)) and observation equation (Eq. (6)). In this paper, state equation presents the transition of its parameters from one moment to the next, and observation equation connects the fused HI with time sequence. The equations are shown as follows:

$$\boldsymbol{x}_t = \boldsymbol{F}\boldsymbol{x}_{t-1} + w_t\boldsymbol{I} \tag{5}$$

$$y_t = \boldsymbol{H}\boldsymbol{x}_t + m_t \tag{6}$$

where $\boldsymbol{x}_t$ is the underlying state vector of the system at time $t, \boldsymbol{x}_t = \left[b_t^0, b_t^1, \ldots, b_t^m\right]^{\mathrm{T}}$, $b_t^0, b_t^1, \ldots, b_t^m$ denote time varying parameters of the prognostics model; y_t is the HI at time t; w_t is measurement noise; m_t is process noise; w_t and m_t are independent from each other and assumed to be Gaussian noise, $w_t \sim N(0,Q)$, $m_t \sim N(0,U)$, $E[w_t m_t] = 0$; $\boldsymbol{H}$ is observation matrix and $\boldsymbol{F}$ is state transition matrix, where $\boldsymbol{H} = [t^0, t^1, \ldots, t^m]$ and $\boldsymbol{F} = \boldsymbol{I}_{(n+1)\times(n+1)}$.

According to Eqs. (5) and (6), the prognostics model is given by

$$\hat{y}_t = b_t^0 + b_t^1 t \ldots + b_t^m t^m + \delta \tag{7}$$

where $\hat{y}_t$ is the predicted HI at time t; δ is the error of the prognostics model, calculated as the average error of the convergence phase in the best fitting process by Kalman filtering. To estimate the time varying parameters (the state vector of state-space method) $b_t^0, b_t^1, \ldots, b_t^m$ of the prognostics model, Kalman filtering algorithm is employed for calculation.

Kalman filtering[30,31] is a recursive algorithm consisting of two phases: predicting (Eq. (8)) and updating (Eq. (9)). The first phase is to estimate the state and the covariance of the current time step based on the state and the covariance from the previous time step. The second phase is to combine the current observations with the state estimate for a refined state, and update the current state and covariance with improved ones. The state estimate in the first phase is termed as a priori state estimate, whereas a posteriori state estimate in the second phase. Kalman filtering algorithm is capable of predicting the time varying parameters of the prognostics model, as well as achieving best fitting (to minimize the variance of the predicted value and the actual value) to reduce the error δ. The phases can be described as follows:

$$\begin{cases} \hat{\boldsymbol{x}}_{t|t-1} = \boldsymbol{F}\hat{\boldsymbol{x}}_{t-1|t-1} \\ \boldsymbol{p}_{t|t-1} = \boldsymbol{F}\boldsymbol{p}_{t-1|t-1}\boldsymbol{F}^{\mathrm{T}} + Q\boldsymbol{I} \end{cases} \tag{8}$$

$$\begin{cases} \hat{\boldsymbol{x}}_{t|t} = \hat{\boldsymbol{x}}_{t|t-1} + K_t\delta_t\boldsymbol{I} \\ \boldsymbol{p}_{t|t} = \boldsymbol{p}_{t|t-1} - K_t\boldsymbol{H}\boldsymbol{p}_{t|t-1} \end{cases} \tag{9}$$

where $\hat{\boldsymbol{x}}_{t|t-1}$ is the a posteriori state estimate at time t given observations up to and including at time $t-1$, $E[\boldsymbol{x}_t - \hat{\boldsymbol{x}}_{t|t-1}] = \boldsymbol{0}$; $\boldsymbol{p}_{t|t-1}$ is the a posteriori error covariance, $\boldsymbol{p}_{t|t-1} = \mathrm{cov}(\boldsymbol{x}_t - \hat{\boldsymbol{x}}_{t|t-1})$; K_t is KalmanGain, $K_t = \boldsymbol{p}_{t|t-1}\boldsymbol{H}^{\mathrm{T}}/S_t$, $S_t = \boldsymbol{H}\boldsymbol{p}_{t|t-1}\boldsymbol{H}^{\mathrm{T}} + U$; δ_t is the error of prediction, $\delta_t = |y_t - \hat{y}_t|, \hat{y}_t$ denotes prediction of y_t.

In Kalman filtering algorithm, the initial state $\hat{\boldsymbol{x}}_{0|0}$ and error covariance $\boldsymbol{p}_{0|0}$ are derived by prior knowledge at the very beginning. After the recursive calculation is triggered, the state and the error covariance at time t can be inferred according to the state along with the error covariance at time $t-1$ and the observations up to time $t-1$ (including at time $t-1$). With the introduction of the observations at time t, the a posteriori estimate of the state and the covariance at time t can be obtained and then updated. The whole process of Kalman filtering algorithm is illustrated in Fig. 4. By this means, the recursive calculation is carried out with appropriate convergence. Consequently, the time varying parameters can be predicted and the error δ_t can be identified via best fitting (presented in details in case study).

To sum up, the advantages of state-space method combined with Kalman filtering algorithm can be presented as follows: firstly, it is capable of creating an appropriate prognostics model with time varying parameters; secondly, an accurate prediction of the time varying parameters can be carried out; lastly, the error can be reduced based on best fitting.

4.2. Steps

The implementation of RUL prognostics in this paper involves a number of steps. To begin with, the time window of the non-stationary degradation process of each parameter is divided into several time windows, where the time series are stationary. Then, according to the divided time windows, the degradation processes are fused to model the performance degradation of the system. After that, prognostics model is constructed by applying state-space method along with Kalman filtering algorithm. At the same time, the time varying parameters of

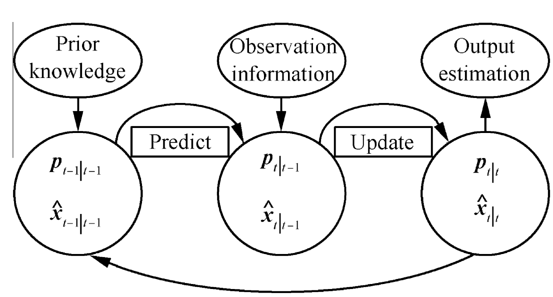

Fig. 4 Schematic of Kalman filtering algorithm.

the model can be predicted and updated by the algorithm. By this means, the prognostics can be finally achieved.

To illustrate a step-by-step implementation of the approach, a flowchart is given in Fig. 5.

More specifically, the steps are elaborated as follows:

Step 1. Detect the anomalies of all seven parameters using Eqs. (1) and (2); then, take the average anomaly as the break point to divide the time window of non-stationary time series. In this step, the value of $\boldsymbol{k}_{A_n}$ (in training sample set) is obtained along with the break point according to

$$P_{\mathrm{avg}_n} = \frac{1}{l}\sum_{i=1}^{l} P_{i_n} \tag{10}$$

$$m_{\mathrm{avg}_n} = \frac{1}{l}\sum_{i=1}^{l} m_{i_n} \tag{11}$$

where P_{avg_n} is the nth average anomaly for all parameters; m_{avg_n} is the average change rate at the nth average anomaly for all parameters; P_{i_n} is the nth anomaly of parameter i; m_{i_n} is the proportion of the decrease or increase of observed value for parameter i from no degradation to the nth anomaly to the decrease or increase of observed value for parameter i over the entire degrading process, i = 1, 2,…, l, l is the number of multiple parameters, and n is the number of anomalies and m_{i_n} is given by

$$m_{i_n} = \frac{\frac{1}{5}\sum_{t=1}^{5} S_{i_t} - \frac{1}{5}\sum_{t=n-2}^{n+2} S_{i_t}}{\frac{1}{5}\sum_{t=1}^{5} S_{i_t} - \frac{1}{5}\sum_{t=N-4}^{N} S_{i_t}} = \frac{\sum_{t=1}^{5} S_{i_t} - \sum_{t=n-2}^{n+2} S_{i_t}}{\sum_{t=1}^{5} S_{i_t} - \sum_{t=N-4}^{N} S_{i_t}} \tag{12}$$

where $\frac{1}{5}\sum_{t=1}^{5} S_{i_t}$ is the mean of the observed data for parameter i near the initial degradation; $\frac{1}{5}\sum_{t=n-2}^{n+2} S_{i_t}$ is the mean of the observed data for parameter i near the nth anomaly, n is the number of anomalies; $\frac{1}{5}\sum_{t=N-4}^{N} S_{i_t}$ is the mean of the observed data for parameter i near the ultimate degradation, N is the number of the time points; notice that 5 time points are adopted in calculating the mean value of different degradation phase (same number of time points are applied in similar application[21]).

Step 2. Divide the non-stationary time series of each parameter into stationary ones based on the average anomaly (Eq. (10)) generated in Step 1; assign $\boldsymbol{k}_{A_1}, \boldsymbol{k}_{A_2}, \ldots, \boldsymbol{k}_{A_n}$ into training sample set $\boldsymbol{M}$ (Eq. (4)) for the training of the observed data to calculate the coefficients $a_n^0, \boldsymbol{A}_n$ of the linear regression model (Eq. (3)) within different time windows.

Step 3. Assign the coefficients $a_n^0, \boldsymbol{A}_n$ and observed data from training set into Eq. (3) to obtain run-to-die (from no degradation to severe degradation) HI time series.By now, the fused HI time series are non-stationary. The break points (potential anomalies) can be obtained by Eq. (1), and the final anomaly can be identified by analyzing the slow variables. According to the distribution of the slow variables, it can be testified that the anomaly of HI time series is consistent with monitoring parameters.

Step 4. Take run-to-die HI as an input of state-space method (Eqs. (5) and (6)) and Kalman filtering algorithm (Eqs. (8) and (9)) to fulfill best fitting, thus minimizing the error δ_t. Most importantly, the prognostics model is constructed in this step.

Step 5. Assign the coefficients $a_n^0, \boldsymbol{A}_n$ and observed data from test set into Eq. (3) to obtain running (the degradation state is unknown) HI time series.

Step 6. Take running HI as an input of state-space method (Eqs. (5) and (6)) and Kalman filtering algorithm (Eqs. (8) and (9)) to estimate the time varying parameters of the prognostics model.According to the results of Step 4 and Step 6 the prognostics model Eq. (7) can be ultimately constructed with identified parameters. With the observations introduced, this approach is capable of achieving an accurate prognosis of RUL.

5. Case study

(1) Data sources

The required data of this case study comes from a NASA turbofan engine degradation simulation for 100 engines in Ref.[32], produced by commercial modular aero-propulsion system simulation (C-MAPSS) damage propagation model. The data set consists of three subsets including training set, test set, and RUL set. Both training set and test set include 24 monitoring parameters of all the simulated engines. The difference of these two sets is that training set refers to run-to-die data for each engine which is recorded from no degradation to failure, whereas test set refers to running data recorded from unknown state to another unknown state. RUL set provides

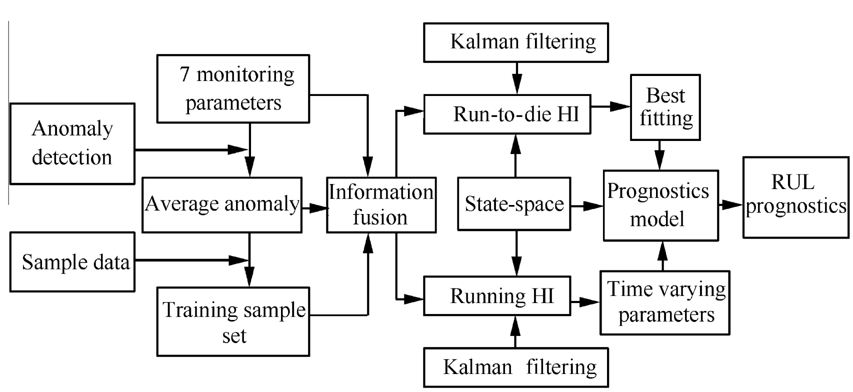

Fig. 5 Flowchart of the prognostics.

the corresponding realistic remaining useful life of each engine under test set, and of course a standard to measure the accuracy of our approach.

In our prognostics, engine No. 2 is chosen as the prediction object. The initial values of the inputs are obtained by means of statistical and fitting methods, $\boldsymbol{p}_{0|0} = [0.001{,}0.001{,}0.001]$ and $\hat{\boldsymbol{x}}_{0|0} = [0.9039, 0.0020, -0.0000]^{\mathrm{T}}$.

(2) Implementation

Step 1. According to Eqs. (1) and (2), anomaly detection can be carried out with results listed in Table 1. To meet the requirements of practical application and the restriction that the time series of each time window must be stationary, Eq. (1) is utilized to generate four time windows, and Eq. (2) is used to identify the fake anomalies among the obtained break points.

Table 1 gives the P_{i_1} and m_{i_1} of each parameter. Parameter P_{15} is not included in this Table, because the monitoring data remain almost the same during the observation. According to Table 1, the average anomaly P_{avg_1}, which will be taken as the break point to divide the time window, is located at the 217th cycle (Eq. (10)), and the corresponding m_{avg_1} equals 0.507, meaning $\boldsymbol{k}_{A_1} = [0.493{,}0.493{,}0.493{,}0.493{,}0.493]^{\mathrm{T}}$.

Step 2. Once the break point (anomaly) and k are obtained, training sample set can be constructed (Eq. (4)) as below

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{X}_{\boldsymbol{k}^1} & \boldsymbol{k}^1 \\ \boldsymbol{X}_{\boldsymbol{k}_{A_1}} & \boldsymbol{k}_{A_1} \\ \boldsymbol{X}_{\boldsymbol{k}^0} & \boldsymbol{k}^0 \end{bmatrix}$$

where $\boldsymbol{k}_{A_1} = [0.493{,}0.493{,}0.493{,}0.493{,}0.493]^{\mathrm{T}}$, $\boldsymbol{k}^1 = [1,1,1,1,1]^{\mathrm{T}}$, $\boldsymbol{k}^0 = [0,0,0,0,0]^{\mathrm{T}}$,

$$\boldsymbol{X}_{\boldsymbol{k}^1} = \begin{bmatrix} 641.89 & 1583.84 & 1391.28 & 21.60 & 554.53 & 2388.01 & 9054.72 \\ 641.82 & 1587.05 & 1393.13 & 21.61 & 554.77 & 2387.98 & 9051.31 \\ 641.55 & 1588.32 & 1398.96 & 21.60 & 555.14 & 2388.04 & 9054.24 \\ 641.68 & 1584.15 & 1396.08 & 21.61 & 554.25 & 2387.98 & 9058.01 \\ 641.73 & 1579.03 & 1402.52 & 21.60 & 555.12 & 2388.03 & 9058.15 \end{bmatrix}$$

$$\boldsymbol{X}_{\boldsymbol{k}_{A_1}} = \begin{bmatrix} 642.97 & 1591.44 & 1411.63 & 21.61 & 553.75 & 2388.09 & 9071.14 \\ 642.46 & 1583.46 & 1408.68 & 21.61 & 553.01 & 2388.08 & 9063.95 \\ 643.30 & 1585.11 & 1407.97 & 21.61 & 553.22 & 2388.10 & 9061.10 \\ 642.91 & 1592.69 & 1412.70 & 21.61 & 553.09 & 2388.12 & 9070.12 \\ 642.88 & 1592.52 & 1412.91 & 21.61 & 553.28 & 2388.02 & 9074.41 \end{bmatrix}$$

$$\boldsymbol{X}_{\boldsymbol{k}^0} = \begin{bmatrix} 643.78 & 1602.03 & 1429.67 & 21.61 & 551.46 & 2388.16 & 9084.13 \\ 643.91 & 1601.35 & 1430.04 & 21.61 & 551.96 & 2388.22 & 9089.87 \\ 643.67 & 1596.84 & 1431.17 & 21.61 & 550.85 & 2388.20 & 9098.67 \\ 643.44 & 1603.63 & 1429.57 & 21.61 & 551.61 & 2388.18 & 9102.01 \\ 643.85 & 1608.50 & 1430.84 & 21.61 & 551.66 & 2388.20 & 9109.36 \end{bmatrix}$$

The coefficients within different time windows are presented as follows:

$$(a_1^0, a_1^1, \ldots, a_1^7) = (335.848, -0.117, -0.003, -0.013, 0.784, 0.128, -0.148, 0.003)$$

$$(a_2^0, a_2^1, \ldots, a_2^7) = (0, -0.035, 0.001, -0.020, 38.560, 0.037, -0.339, 0.001)$$

To testify the applicability of multiple linear regression model for the fusion of observed information, the statistics[27] produced during the whole process are carried out and listed in Table 2.

In Table 2, it is shown that two phases are utilized for multiple linear regression in order to implement information fusion. The statistics which serve as criteria to judge whether the regression model fit the circumstance include four parts: the R^2 statistic, the F statistic, an estimate of the error variance and $\bar{r}$ (the average residual error). R^2 is one minus the ratio of the error sum of squares to the total sum of squares, and the closer this value is to 1, the more appropriately the model fits the data. F is the test statistic generated in the regression, for a significant linear regression relationship between the response variable and the predictor variables. The bigger F value is, the more significant the model is. The critical value of F statistic in this case is between 9.33 and 9.37 according to F distribution critical value table. Additionally, the estimated v_{error}(error variance) and $\bar{r}$ are small enough to be accepted in the fusion process. It is obvious that all the statistics of both phases are well qualified to testify that the multiple linear regression model fits for the observed data and the information fusion can be implemented reasonably.

Table 1 Anomaly P_{i_1} and change ratio m_{i_1} of the observed parameters.

Item	P_{i_1} (cycle)	m_{i_1}	Item	P_{i_1} (cycle)	m_{i_1}
T_{24}	215	0.637	P_{30}	228	0.457
T_{30}	218	0.499	N_{f}	209	0.500
T_{50}	215	0.556	N_{c}	218	0.394

Table 2 Statistics produced in the multiple linear regression.

Phase	R^2	F	v_{error}	$\bar{r}$
1st phase	0.9820	15.5594	0.0058	0.0284
2nd phase	0.9852	33.2688	0.0030	0.0239

Step 3. Assign the coefficients within different time windows into Eq. (3), and then the information fusion for run-to-die HI time series can be implemented with the introduction of the observations from training set. The fused run-to-die HI time series are shown in Fig. 6.

The slow variables of the HI are presented in Fig. 7. It's shown that the distributions of HI's slow variables prior to and posterior to the 21st time window differ with each other by a difference of 0.05 in mean values. Obviously, the change rate of the degradation has changed abruptly in this time window. What should be noted is that the 21st time window happens to be near to the 217th cycle (average anomaly), which means that the fused HI time series based on superstatistics manage to reflect the engine's non-stationary performance degradation precisely.

Step 4. The result of best fitting via state-space method combined with Kalman filtering algorithm is presented in Fig. 8 (green for predicted HI, and red for fused HI).

Based on the best fitting for observations, the prognostics model can be constructed as

$$\hat{y}_t = b_t^0 + b_t^1 t + b_t^2 t^2 + \delta$$

where $\delta = \pm 0.138$. Note that δ is average error calculated according to δ_t generated in the fitting process.

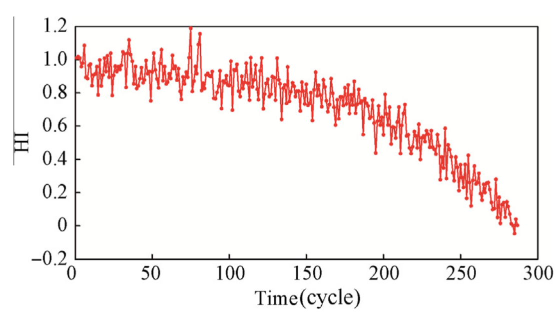

Fig. 6 HI time series under training set.

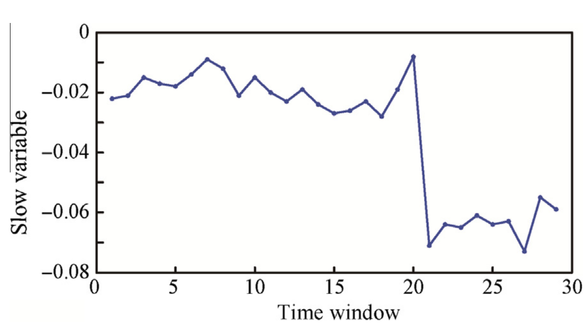

Fig. 7 Slow variables of HI.

Step 5. Replace the observations from training set with observations from test set in Step 3, and then the information fusion for running HI time series can be carried out. The fused running HI time series are presented in Fig. 9.

Step 6. By assigning the obtained HI in Step 5 into Kalman filtering algorithm and state-space method, the time varying parameters of the prognostics model can be estimated. Ultimately, the prognosis of RUL can be accomplished with the introduction of time. The prediction results are shown in Fig. 10. It's presented that the failure of engine No.2 under test set is estimated to occur after 90 cycles. Compared to the realistic RUL of 98 cycles, the proposed approach is followed by a prediction error of 8.163%.

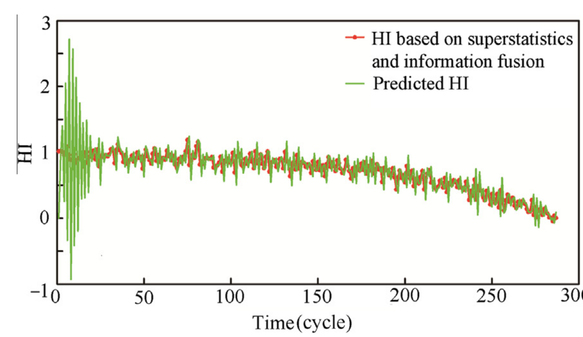

Fig. 8 Best fitting for HI time series under training set.

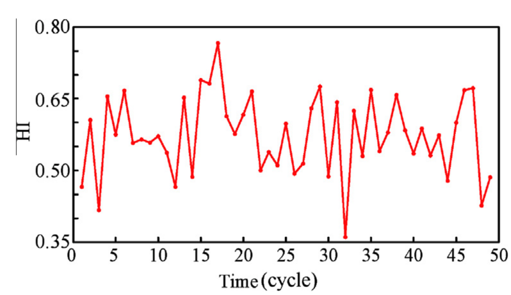

Fig. 9 HI time series under test set.

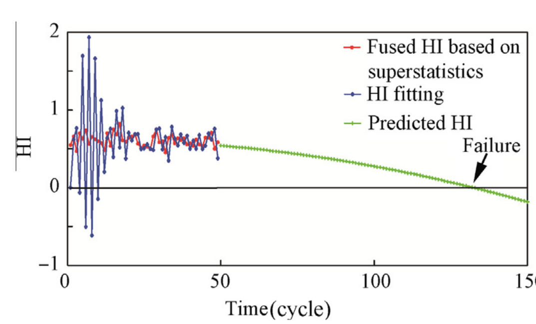

Fig. 10 RUL prognostics.

In order to verify the accuracy of the approach proposed in this paper which is based on superstatistics and information fusion (noted by SIF_Kalman in the following description), a comparison of it to the other two traditional methodologies is made in the following discussion. The traditional methodologies are least squares[33] based on information fusion (noted by least squares) and Kalman filtering algorithm based on information fusion (noted by IF_Kalman). Least Squares is a method to minimize the sum of squared residuals, a residual being the difference between the fitted value to a corresponding observed value. The following discussion is mainly presented in terms of fitting, prediction error and prediction evolution process.

(1) Fitting and prediction error

The prediction results by least squares and IF_Kalman are illustrated in Fig. 11(a) and (b), separately. Comparing Fig. 10 with Fig. 11(a) and (b) and according to Table 3, Least squares generates the best fitting but worst prediction error , meaning an over fitting phenomenon occurs in the procedure. Under the circumstances of over fitting, Least squares method generally leads to poor predictive performance because of its exaggeration in minor fluctuations of the data, thereby not available for prediction in engineering. For the rest two methods, SIF_Kalman shows a prediction error of 8.16% and a fitting error of 9.32%, both smaller than that of IF_Kalman. Therefore, SIF_Kalman is more appropriate for the prognosis of RUL.

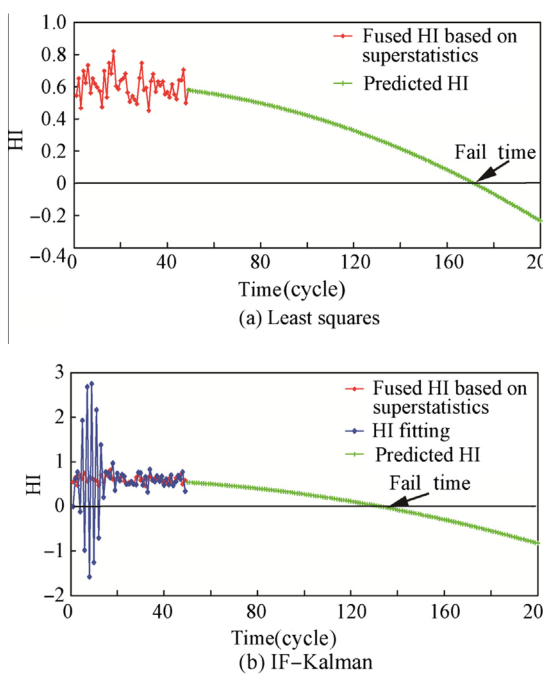

Fig. 11 RUL prognostics achieved by least squares and IF_Kalman.

(2) Prediction evolution process

Practically, with the introduction of monitoring parameters, the prediction results of all the three methods would gradually approach the actual RUL, which explains the concept of prediction evolution. For example, assuming that the entire service life of an engine is 1000 cycles, the RUL predicted at the 750th cycle (0.75 standard life) is far more accurate than the RUL predicted at the 500th (0.5 standard life)cycle. However, the prediction evolution process of each method differs, which is shown in Fig. 12.

Standard life in Fig. 12 is calculated according to

$$L_S = L_O/L_E \tag{13}$$

where L_S denotes standard life, L_O the operation life when RUL prognostics is implemented and L_E the entire useful life.

The standard life of 0.75 corresponds to the 21st time window of HI's slow variables (see in Fig. 7), indicating that anomalous behavior occurs in the degradation. Before the anomaly, the prediction evolution processes of all the methods can be analyzed as follows. For least squares, the prediction error is unacceptable and the approaching speed is far too slow; what's worse, the centerline of the fluctuations (blue dotted line between 0.5 and 0.75) deviates far away from the actual useful life line (red solid line). For IF_Kalman and SIF_Kalman, their prediction curves fluctuate around the actual useful life line, but the fluctuations of the former are more volatile than the latter. Hence, the prediction performance of SIF_Kalman is the best prior to anomaly. For the situations after anomaly, the conclusion is almost the same. For least squares, the prediction curve approaches the red line more rapidly; however, the prediction error has not been improved due to the poor prediction performance at the earlier stage. For IF_Kalman, the fluctuations become less volatile; while, the approaching process is not so ideal because the degradation trend is hard to track when the non-stationary characteristics of the degradation are not taken into account, so the prediction is not so accurate. For SIF_Kalman, the approaching is the most rapid, the fluctuations are the most acceptable, and the prediction is the most accurate, all because of the two

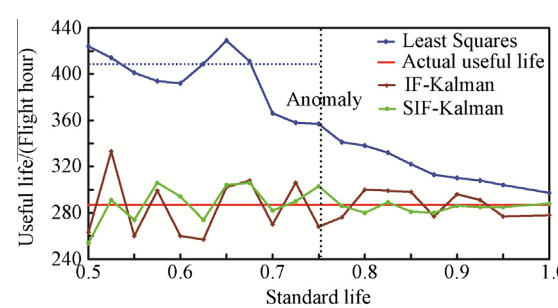

Fig. 12 Comparison of prediction evolution processes of three methods.

Table 3 Predicted RUL of each method.

Method	Actual RUL (cycle)	Predicted RUL (cycle)	Fitting error (%)	Prediction error (%)
Least squares	98	122	6.33	24.49
IF_Kalman	98	83	9.85	15.31
SIF_Kalman	98	90	9.32	8.16

reasons. Firstly and most importantly, the realistic non-stationary degradation can be revealed by HI which is obtained by means of superstatistics and information fusion; secondly, the time varying parameters of the prognostics model can be estimated precisely. Therefore, SIF_Kalman is also superior to traditional methods posterior to anomaly.

In conclusion, SIF_Kalman fits for the prognosis of RUL because of small fitting and prediction error and satisfactory prediction evolution process. Importantly, all such good prediction performance should owe to the application of superstatistics, information fusion and appropriate algorithm.

6. Conclusions

In this paper, a novel approach to RUL prognostics of aeroengine subject to non-stationary degradation is developed, and the calculation is realized by using state-space method and Kalman filtering algorithm. The main accomplishments of the approach are as follows:

(1) The non-stationary characteristics of monitoring parameters are taken into account, which contributes to the accurate demonstration of the realistic degradation of the engine.
(2) Limitations caused by single parameter-based methodologies are avoided, because multivariate observed information is made full use of.
(3) As the observed information is contaminated when collected, monitoring parameters are incapable of characterizing the actual health state of the system directly. However, Kalman filtering algorithm takes noise factor into consideration, thereby enabling higher accuracy and more satisfactory prediction evolution process. Additionally, the algorithm simply requires observed data instead of enormous unavailable failure data, making the implementation available.

The case study shows that an accurate prognosis of RUL can be achieved, which provides a guide for operators to realize CBM, thereby assisting in guaranteeing civil aviation safety and minimizing operation cost. For further study, the application of superstatistics theory will be developed, e.g. the degradation amount distribution of performance parameters of aeroengine can be applied to RUL prognostics by combining stochastic process with super statistics.

Acknowledgments

The authors were grateful to the anonymous reviewers for their critical and constructive review of the manuscript. This study was co-supported by the State Key Program of National Natural Science of China (No. 61232002), the Joint Funds of the National Natural Science Foundation of China (No. 60939003), China Postdoctoral Science Foundation (Nos. 2012M521081, 2013T60537), the Fundamental Research Funds for the Central Universities of China (No. NS2014066), Postdoctoral Science Foundation of Jiangsu Province of China (No. 1301107C), and Philosophy and Social Science Research Projects in Colleges and Universities in Jiangsu of China (No. 2014SJD041).

References

1. China Academy of Civil Aviation Science and Technology. *Statistical analysis report on incidents in civil aviation of China [Internet]*; 2012[2013-09-12]. Available from: http://safety.caac.gov.cn/admin/news/outview.do?pk = 100000093653107 [Chinese].
2. Lu JC, Meeker W. Using degradation measures to estimate a time-to-failure distribution. *Technometrics* 1993;**35**(2):161–74.
3. Gebraeel N. Residual-life distributions from component degradation signals: a bayesian approach. *IIE Trans* 2005;**37**(6):543–57.
4. Ren SH, Zuo HF. A combined prediction method for the residual-life of civil aviation engines based on performance degradation. *Mech Sci Technol Aerosp Eng* 2011;**30**(1):23–9 [Chinese].
5. Fu XY, Zhong SS, Zhang MN. An improved method for aeroengine residual life prediction. *J Harbin Inst Technol* 2013;**45**(5):51–5 [Chinese].
6. Lu S, Lu H, Kolarik WJ. Multivariate performance reliability prediction in real-time. *Reliability Eng Syst Saf* 2000;**72**(1):39–45.
7. Yang K, Yang GB. Performance degradation analysis using principal component method. *Proceedings of the annual reliability and maintainability symposium*; 1997 Jan 13–16; Philadelphia, PA, USA; 1997. p. 136–41.
8. Park C, Padgett WJ. Stochastic degradation models with several accelerating variables. *IEEE Trans Reliability* 2006;**55**(2):379–90.
9. Barker CT, Newby MJ. Optimal non-periodic inspection for a multivariate degradation model. *Reliability Eng Syst Saf* 2009;**94**(1):33–43.
10. Wang P, Coit DW. Reliability prediction based on degradation modeling for systems with multiple degradation measures. *Reliability and maintainability annual symposium (RAMS)*; 2004 Jan 26–29; South Windsor, CT, USA; 2004. p. 302–7.
11. Zhong QH, Zhang ZH, Liang SJ. Reliability analysis approach based on multivariate degradation data. *Syst Eng Theory Pract* 2011;**31**(3):544–51 [Chinese].
12. Crk V. Reliability assessment from degradation data. *Proceedings of reliability and maintainability annual symposium*; 2000 Jan 24–27; Los Angeles, CA, USA; 2000. p. 155–61.
13. Wang P. System reliability prediction based on degradation modeling considering field operating stress scenarios [dissertation]. New Jersey: The State University of New Jersey; 2003.
14. Ren SH. Research on methods of performance reliability assessments and life on wing prediction for civil aeroengine [dissertation]. Nanjing: Nanjing University of Aeronautics and Astronautics; 2010 [Chinese].
15. Zhang Y, Zuo HF, Ren SH, Zhu L. Research on an algorithm for prediction of aeroengine performance deterioration. *China Mech Eng* 2011;**22**(9):1009–12 [Chinese].
16. Sun JZ, Zuo HF. Gas path analysis based on multi-sources diagnostics information fusion. *J Aerosp Power* 2013;**28**(8):1885–96 [Chinese].
17. Beck C, Cohen EGD. Superstatistics. *Phys A: Stat Mech Appl* 2003;**332**:267–75.
18. Pickands J. Statistical inference using extreme order statistics. *Ann Stat* 1975;**3**(1):119–31.
19. Yang Y, Hu HP, Xiong W, Ding F. Novel non-stationary time series anomaly detection model based on superstatistics theory. *Comput Sci* 2011;**38**(6):93–117 [Chinese].
20. Wang HW, Wu HQ. Residual useful life prediction for aircraft engine based on information fusion. *J Aerosp Power* 2012;**27**(12):2749–55, Chinese.
21. Sun JZ, Zuo HF, Wang WB, Michael GP. Application of a state space modeling technique to system prognostics based on a health index for condition-based maintenance. *Mech Syst Signal Process* 2012;**28**:585–96.
22. Christian B, Ezechiel GDC, Harry LS. From time series to superstatistics. *Phys Rev: E- Stat Nonlinear Soft Matter Phys* 2005;**72**(5):1–7.

23. Yang Y. A novel network traffic anomaly detection model based on superstatistics theory [dissertation]. Wuhan: Huazhong University of Science and Technology; 2010 [Chinese].
24. Wang Y. The model of partial squares regression after outlier detection. *Sci Technol Eng* 2011;**11**(19):4556–8, Chinese.
25. Feng GL, Gong ZQ, Dong WJ, Li JP. Abrupt climate change detection based on heuristic segmentation algorithm. *Acta Physica Sinica* 2005;**54**(11):5494–9, Chinese.
26. Michael P. A prognostics and health management roadmap for information and electronics-rich system. *Microelectron Reliability* 2010;**50**(3):317–23.
27. Audone B, Giunta G. Multiple linear regression to detect shielding effectiveness degradations. *Electromagnetic Compatibility (EMC) international symposium*; 2008 Sep 8–12; Hamburg, Germany; 2008. p. 1–6.
28. Lin ZC, Wu WJ. Multiple linear regression analysis of the overlay accuracy model. *IEEE Trans Semiconductor Manuf* 1999;**12**(2):229–37.
29. Kalman RE. A new approach to linear filtering and prediction theory. *Trans ASME J Basic Eng* 1960;**82**(Series D):35–46.
30. Li LM, Gong WB, Liu HJ, Yu JP. A carrier tracking algorithm based on adaptive extended Kalman filter. *Acta Aeronautica et Astronautica Sinica* 2012;**33**(7):1319–28 [Chinese].
31. Quan W, Xu L, Zhang HJ, Fang JC. Interlaced optimal-request and unscented Kalman filtering for attitude determination. *Chin J Aeronaut* 2013;**26**(2):449–55.
32. Saxena A, Goebel K. *Turbofan engine degradation simulation data set [Internet]*; 2008[2013-09-12]. Available from: http://ti.arc.nasa.gov/project/prognostic-data-repository.
33. Lu F, Huang JQ, Kong TX. Gas-path fault diagnosis for aero-engine based on variable weighted least squares. *J Aerosp Power* 2011;**26**(10):2376–81 [Chinese].

Liu Junqiang is currently a lecturer in College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, China. His major research focuses on prognostic and health management (PHM) for aero-engine and transportation information engineering & control.

Zhang Malan is currently a master student in College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, China. Her major research focuses on PHM for aero-engine and transportation information engineering & control.

Zuo Hongfu is currently a full Professor in College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, China. His major research focuses on reliability & maintainability engineering and transportation information engineering & control.

Xie Jiwei is currently a master student in College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, China. His major research focuses on PHM for aero-engine and transportation information engineering & control.