

Article

Aircraft Engine Prognostics Based on Informative Sensor Selection and Adaptive Degradation Modeling with Functional Principal Component Analysis [†]

Bin Zhang ^{1,2}, Kai Zheng ¹, Qingqing Huang ^{3,*}, Song Feng ¹ , Shangqi Zhou ¹ and Yi Zhang ¹

¹ School of Advanced Manufacturing Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; zhangbin@cqupt.edu.cn (B.Z.); zhengkai@cqupt.edu.cn (K.Z.); fengsong@cqupt.edu.cn (S.F.); zsq11915@163.com (S.Z.); zhangyi@cqupt.edu.cn (Y.Z.)

² The State Key Laboratory of Mechanical Transmissions, Chongqing University, Chongqing 400044, China

³ School of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

* Correspondence: qingq.huang@gmail.com; Tel.: +86-(023)62461709

[†] This paper is an extended version of the paper published in Zhang, B.; Zheng, K.; Luo, J.F.; Zhang, Y. Informative Sensor Selection and Health Indicator Construction for Aircraft Engines Prognosis. In Proceedings of the 2019 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC), Beijing, China, 15–17 August 2019. (in press).

Received: 15 December 2019; Accepted: 6 February 2020; Published: 9 February 2020



Abstract: Engine prognostics are critical to improve safety, reliability, and operational efficiency of an aircraft. With the development in sensor technology, multiple sensors are embedded or deployed to monitor the health condition of the aircraft engine. Thus, the challenge of engine prognostics lies in how to model and predict future health by appropriate utilization of these sensor information. In this paper, a prognostic approach is developed based on informative sensor selection and adaptive degradation modeling with functional data analysis. The presented approach selects sensors based on metrics and constructs health index to characterize engine degradation by fusing the selected informative sensors. Next, the engine degradation is adaptively modeled with the functional principal component analysis (FPCA) method and future health is prognosticated using the Bayesian inference. The prognostic approach is applied to run-to-failure data sets of C-MAPSS test-bed developed by NASA. Results show that the proposed method can effectively select the informative sensors and accurately predict the complex degradation of the aircraft engine.

Keywords: prognostics; aircraft engine; sensor selection; degradation modeling; functional principal component analysis; Bayesian inference

1. Introduction

As the heart of an aircraft, the engine consists of several subsystems with millions of parts, and its health condition directly impacts operation and safety of the whole aircraft system. While the reliability of aero-engines has been improved over the years, factors such as fatigue and wear will inevitably cause the health condition of an engine to degrade with usage, reducing the overall performance of the powered aircraft [1]. Therefore, a number of sensors of varying types have been widely used to monitor the health degradation of aircraft engines, which creates a multi-sensor environment for operational health analysis and maintenance decision-making. Over the past decades, prognostics that utilizes monitoring sensors information to predict the future health and estimates the remaining useful life (RUL) before failure/end-of-life has attracted increasing attention from academic researchers and industrial operators [2]. For its potential in enhancing reliability and maintenance efficiency while

reducing unnecessary maintenance and minimizing operational costs, prognostics have been an active research field for aircraft engine applications [3].

Generally, the existing methods pertaining to aircraft engine prognostics can be classified into two categories: model-based and data-driven methods. For model-based methods, a mathematical model that can describe the engine health degradation process that is normally required to be constructed according to physical failure characteristics before prediction. Some typical model-based methods have been developed for engine health estimation and RUL prediction, such as the Markov model-based [4,5] and particle filtering-based method [6,7]. While model-based methods enable improved accuracy of engine prognostics, certain assumptions and simplifications of the adopted models may pose limitations on their practical deployment. With the rapid development of data mining techniques and the growing availability of health monitoring data, data-driven methods attract increasing attention. Data-driven methods utilize the information extracted or learned from observed data to identify the degradation behavior and predict the future health condition without using any particular physical model [8,9]. In this view, data-driven methods may be the more applicable prognostic solution for complicated systems, as aero-engines that have limited knowledge of physics-of-failure but have massive multi-sensor monitoring data. Among the data-driven prognostic methods, machine learning and statistical approaches are two popular branches. The common machine learning techniques used for prognostics include support vector regression [10], artificial neural network [11,12], fuzzy logic [13], etc. Machine learning methods are capable of dealing with prognostic issues of complex systems, but the predicted results are hard to be explained because of the lack of transparency. Besides, it is difficult for machine learning methods to handle the uncertainty in health degradation and long-term prognostics. In contrast, statistical approaches fit probabilistic models to the health monitoring data and are effective in managing the uncertainty of the degradation process as well as its influence on future predictions. Thus, statistical prognostics are recently becoming dominant [2].

In statistical data driven prognostics, the available health monitoring data that contains useful health degradation information is modeled by statistical models to perform prognostics [14]. Initially, regression-based statistical models are applied to model the available data and predict future health. In [15], a hybrid multi-variate adaptive regression splines method was proposed for the aircraft engine prognosis. To account unit-to-unit variability, Zaidan et al. [16] formulated a Bayesian hierarchical model with linear random coefficient model for engine prognostics. A parametric p th order polynomial model was proposed by Liu et al. [17,18] and was validated on a degradation dataset of an aircraft gas turbine engine. Another group of frequently used statistical models are stochastic process models that can well capture the degradation dynamics of a system. For its nice mathematical properties, the Wiener process has been intensively studied in degradation modeling and prognostics for engines and other engineered systems [19]. Considering monotone non-decreasing degradation in practice, Le Son et al. [20] modeled deterioration of gas turbine engines as a non-homogeneous Gamma process. Peng et al. [21] further researched modeling of time-varying degradation rates with an inverse Gaussian process. In practice, statistical approaches are mainly concentrated on univariate degradation modeling and prognostics, so a composite health index (HI) that captures health degradation of an engine is normally needed to be constructed by fusion of the multi-sensor monitoring data. Wang et al. [22] proposed a HI construction method by a linear data transformation method of multi-sensor data. In [23], the failure time and the multi-sensor data are combined via a novel latent linear model to construct a generic HI. Theoretically, integration of more sensors may be beneficial for prognostics. However, different sensors usually have different levels of relevance to the degradation process. Also, some sensors may be even unrelated to the underlying degradation process due to high noise, low sensitivity to the health degradation, inconsistency among the population, etc. In addition, fusion of too many sensors may incur a heavy data processing burden and unnecessary costs. As a result, it is better to select the degradation-relevant sensors for HI construction, and hereafter, this kind of sensors are denoted as informative sensors in this paper. Coble et al. [24] proposed three metrics to characterize the suitability of sensors for prognostics. Yu [6] proposed an informative sensor selection method

based on logistic regression with penalization regularization for aircraft engine health prognostics. In [25], Liu et al. proposed an improved permutation entropy method for selection of informative sensors. To identify multiple stages before failure, Chehade et al. studied informative sensor selection and fusion using statistical hypothesis testing [26]. Besides, the task of informative sensor selection was integrated with the statistical modeling in [23] and [27] for engine prognostics.

Through fusion of the selected informative sensors into a composite HI, many statistical methods have been proposed for prognostics of complex systems, including aircraft engines. However, most of the statistical approaches are based on parametric models, such as a linear model [16,23,28], exponential model [17,27], quadratic model [24], etc. These methods are effective when the engine degradation pattern can be described with a specific model, but there are numerous cases where the degradation trajectory may not be appropriately fitted by a parametric model [1,3,16]. Therefore, a more flexible degradation modeling and prognosis technique is needed for practical applications. Furthermore, the existing literature relevant to the selection of informative sensors is basically focused on the relevance of a sensor to the degradation process. While metrics of monotonicity [10,25], correlation and robustness [2,29] have been commonly utilized to evaluate the relevance of sensors for prognostics. There is a lack of efforts to consider the degree to which the sensors of a population of systems have the same underlying shape [24]. Specifically, for a fleet of engines operating under one operational condition and with one failure mode, it is highly desired that the sensor demonstrating a consistently increasing or decreasing trend for all the engines to be selected as the informative sensor. To address these limitations, a statistical method with informative sensor selection is proposed for prognostics of the aircraft engine in this paper. In the presented method, a new metric that evaluates the degree of consistent trend of a sensor for a population of systems is proposed and used to select informative sensors to construct the engine HI. Then, the functional principal component analysis (FPCA) method is applied to adaptively model the engine HI and combined with Bayesian inference for future prognostics.

The remainder of this paper is organized as follows. Section 2 briefly gives the related theories. Section 3 presents a detailed description of the proposed prognostic approach. Section 4 demonstrates the developed method with a case study on the degradation of a turbofan engine. Section 5 discusses some key issues of the proposed approach. Finally, Section 6 summarizes the research and future work.

2. Related Theories

In this section, the theories utilized in this study are briefly introduced.

2.1. Basics of FPCA

Functional data analysis concerns the analysis of data that are in the form of a function with statistical methods. Functional data are of intrinsically infinite dimension, and thus, dimension reduction based on some kind of basis function is crucial for analysis of such data. Unlike the parametric basis function such as B-spline, Fourier, and wavelet functions, FPCA directly derives the basis function from the data and has been a useful tool for functional data analysis [30–32].

In functional data analysis, the given data $Y_i(t)$ ($i = 1, 2, \dots, n$) is assumed as an independent and identically distributed (i.i.d) realization of a stochastic process $Y(t)$ that is in L^2 and defined on the interval $[0, \Gamma]$. The expectation and covariance function of $Y(t)$ are respectively $\mu(t) = E(Y(t))$ and $G(s, t) = E((Y(s) - \mu(s))(Y(t) - \mu(t)))$, where $E(\cdot)$ is the expectation operator. Under mild assumptions, Mercer's theorem implies that the spectral decomposition of $G(s, t)$ leads to:

$$G(s, t) = \sum_{k=1}^{\infty} \lambda_k \phi_k(s) \phi_k(t), t \in [0, \Gamma] \quad (1)$$

where $\lambda_1 \geq \lambda_2 \geq \dots$ are the ordered eigenvalues and $\phi_k(t)$ are the corresponding eigenfunctions or functional principal components (FPCs). These terms are obtained by solving the following eigen-decomposition problem:

$$\int_{\Gamma} G(s, t) \phi_k(t) ds = \lambda_k \phi_k(t), t \in [0, \Gamma]$$

$$\text{s.t.} \begin{cases} \int_{\Gamma} \phi_k^2(t) dt = 1 \\ \int_{\Gamma} \phi_l(t) \phi_k(t) dt = 0, l \neq k \end{cases} \quad (2)$$

Then the Karhunen–Loève theorem states that the functional data $Y_i(t)$ ($i = 1, 2, \dots, n$) are represented by FPCs of the underlying stochastic process covariance function as Equation (3). In the statistical literature, this method has been coined FPCA:

$$Y_i(t) = \mu(t) + \sum_{k=1}^{\infty} \varepsilon_{ik} \phi_k(t), t \in [0, \Gamma] \quad (3)$$

where ε_{ik} are the FPC scores calculated as Equation (4), and ε_{ik} are random variables that are independent across i and uncorrelated across k with $E(\varepsilon_{ik}) = 0$ and $E(\varepsilon_{ik}^2) = \lambda_k$:

$$\varepsilon_{ik} = \int_{\Gamma} (Y_i(t) - \mu(t)) \phi_k(t) dt, t \in [0, \Gamma] \quad (4)$$

Based on the above analysis, the functional data is expressed as the sum of infinite terms of FPCs by FPCA. Each FPC independently depicts one mode of variations in the functional data, thus the variance (Var) of the data $Y_i(t)$ ($i = 1, 2, \dots, n$) is the sum of all the eigenvalues of the covariance function as follows:

$$Var(Y_i(t)) = Var(Y_i(t) - \mu(t)) = \sum_{k=1}^{\infty} Var(\varepsilon_{ik}) = \sum_{k=1}^{\infty} \lambda_{ik} \quad (5)$$

However, only a small number of eigenvalues are commonly significantly nonzero in practice. For eigenvalues which are approximately zero, the corresponding FPC scores will also be approximately zero. Therefore, $Y_i(t)$ ($i = 1, 2, \dots, n$) can be approximated by the first K -truncated terms as:

$$Y_i(t) = \mu(t) + \sum_{k=1}^K \varepsilon_{ik} \phi_k(t), t \in [0, \Gamma] \quad (6)$$

To ensure that sufficient variance of the functional data is explained by the truncated expansion, K is usually chosen using the fraction of variance explained (FVE) method as Equation (7) and the threshold η is often a value higher than 85% [30]:

$$K = \min_k \left(\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^{\infty} \lambda_i} \geq \eta \right) \quad (7)$$

2.2. Bayesian Inference

Bayesian inference is based on the Bayes' theorem, which allows one to formally incorporate prior knowledge or experience into computing statistical probabilities. In this way, the inference takes the accumulated information as a priori, and then uses the observed data to update the priori to a posterior. Currently, Bayesian inference has been widely applied in reliability engineering [8,16,21].

For a continuous random variable y and related random parameters θ , Bayes' theorem states that their joint probability can be written in two ways as [33]:

$$p(\theta, y) = p(y|\theta)p(\theta) = p(\theta|y)p(y) \quad (8)$$

Eliminating the joint probability $p(y, \theta)$ and rearranging a bit, we obtain Bayesian inference for the parameters θ :

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \propto p(y|\theta)p(\theta) \quad (9)$$

In the literature, $p(\theta|y)$ and $p(\theta)$ are respectively defined as the posterior and prior distribution of the random parameters θ , and $p(y|\theta)$ is denoted as the likelihood function. The $p(y|\theta)$ can be understood as the probability of observing samples of the random variable y conditioned on the parameters θ . Based on Equation (9), Bayesian inference updates the parameters distribution from a priori to a posterior with the likelihood function of the observed data. Thus, the uncertainty in the parameters can be reduced and accuracy of further inference using the parameters can be improved.

3. The Proposed Method

This section is devoted to introducing the proposed method. The prognostic approach consists of informative sensor selection, HI construction, degradation modeling, and health prognostics. Flowchart of the method is shown in Figure 1.

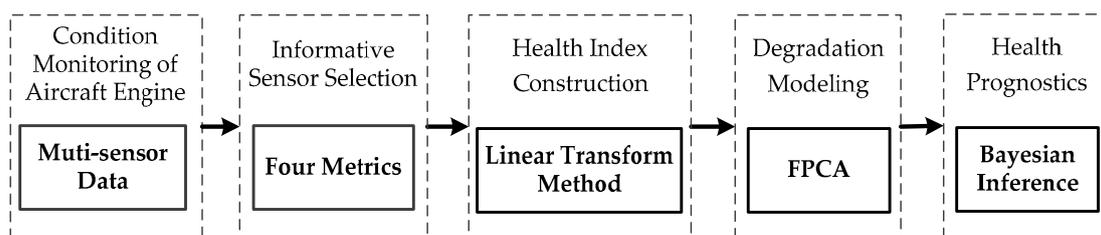


Figure 1. Flowchart of the proposed method. Abbreviations: FPCA, functional principal component.

3.1. Informative Sensor Selection

In this subsection, four metrics are presented to select informative sensors for complex systems of aircraft engines with a number of monitoring sensors.

In practice, system health deterioration is inherently a stochastic process. Metrics for informative sensor selection should be defined upon separating a sensor measurement into its trend and the residual. This decomposition can be carried out locally with smoothing methods or globally with parametric or nonparametric methods. To make full use of the sensor measurements for the whole degrading process modeling and to preserve local characteristics of the degradation, local smoothing methods may be a better option. In this paper, the trend of a sensor measurement is obtained using the moving average method. With the sensor trend and residual, three metrics for degradation feature selection were defined in [29] to improve prognostics effectiveness and efficiency. For the sake of completeness and clarity, the three metrics are reformulated for sensor selection as:

$$s(t_j) = s_T(t_j) + s_R(t_j), \text{Corr}(S) = \frac{\left| N \sum_j s_T(t_j) t_j - \sum_j s_T(t_j) \sum_j t_j \right|}{\sqrt{\left[N \sum_j s_T(t_j)^2 - \left(\sum_j s_T(t_j) \right)^2 \right] \left[N \sum_j t_j^2 - \left(\sum_j t_j \right)^2 \right]}} \quad (10)$$

$$\text{Mon}(S) = \frac{\left| \sum_j \delta(s_T(t_j) - s_T(t_{j-1})) - \sum_j \delta(s_T(t_{j-1}) - s_T(t_j)) \right|}{N-1}, \text{Rob}(S) = \frac{1}{N} \sum_j \exp\left(-\left| \frac{s_R(t_j)}{s_T(t_j)} \right| \right)$$

where $s(t_j)$ is the measurement of the sensor S at the time t_j ($j = 1, 2, \dots, N_j$) with the trend of $s_T(t_j)$ and residual $s_R(t_j)$, $\delta(\cdot)$ is the simple unit step function.

Among the three metrics, correlation (Corr) measures the linearity between the interested sensor and the usage time, monotonicity (Mon) assesses consistently increasing or decreasing trend of the sensor, and robustness (Rob) reflects the tolerance of the sensor to noises and outliers. The sensor with higher scores should be selected as the informative sensor, for the three metrics are all positively correlated with the relevance of a sensor to the system degradation.

From Equation (10), it can be concluded that these metrics are defined on the sensor from one individual system and the same sensor from other systems of the population is not accounted. For a population of systems (such as a group of aircraft engines) that operate under the same condition and with one failure mode, the informative sensor is highly desired to have a consistent increasing or decreasing trend for all the systems. Additionally, the sensor with a small variation of measurements upon system failures and a large range till failure is also desired for prognostics. Thus, a new metric of predictability (Pre) is formulated for sensor selection as:

$$Pre(S) = Exp\left(-\frac{std(s_f)}{|mean(s_f) - mean(s_s)|}\right) \quad (11)$$

where s_f and s_s are respectively the measurements at the failure and start instant of the sensor S for a population of systems.

The new metric Pre takes the failure and start values of a sensor from the entire population into account, and thus the trend consistency of a sensor among a population is given consideration. Also, Pre is positively related to the performance of the sensor with the range of [0,1], and favorites the sensor with well-clustered failure values and large range. That is the sensor with a high score of Pre should be selected as the informative sensor.

For prognostic parameter selection, Coble et al. presented three metrics in [24] and the prognosability (Pro) metric is quite similar to the proposed Pre. However, the two metrics are different in measuring the range of one sensor. In the metric Pro, the range is the mean range from start to failure for a population of systems. Sensors with well-clustered failure values and large range are encouraged by this measure. While in the Pre, as in Equation (11), the sensor measurements at the start and failure instant of a population are respectively averaged to the given range of the sensor. The metric Pre selects sensors with both well-clustered failure and start values as well as large range.

To select the informative sensor for aircraft engines, two steps are included using the presented goodness metrics. In the first step, binary-value and constant-value sensors are eliminated by simple visual inspection, because these sensors do not make sense for prognosis. As can be seen from these definitions, one goodness metric only partially measures suitability of a sensor for prediction and sensor selection based on only one metric will be biased. In the second step, sensors are evaluated by four metrics simultaneously, and then the informative sensors selected with drop out strategy are fused to construct a HI for health prognosis of aircraft engines.

3.2. Health Index Construction

For a data-driven prognostics, extraction of the health signatures and background knowledge from massive training/testing sensors is required. Based on the selected informative sensors, the HI construction is considered in this subsection.

The transformation of the selected informative sensors into one-dimensional HI is a process of information fusion, which enables a general measure to characterize the health condition of a system and also the degradation of different systems from the same population to be similar. The HI can be constructed using many information fusion techniques, such as linear data transformation method [13,17,22], PCA and Euclidean distance measure [20], and logistic regression [6,11]. For its

effectiveness and ease of use, the linear data transformation method is also employed to construct the HI for aero-engines in this paper.

Suppose the selected informative sensors are d -dimensional and the two sensor data sets that represent the aircraft engine failed, and healthy states are \mathbf{Q}_0 of $M_0 \times d$ and \mathbf{Q}_1 of $M_1 \times d$ matrix, where M_0 and M_1 are, respectively, the data sizes for engine failed and healthy states. Generally, data sets for healthy and failed states can be collected from the training data sets at the beginning and at the end of the run-to-failure tests. With these two data sets, a transformation matrix \mathbf{T} of $d \times 1$ is obtained and then the one-dimensional HI \mathbf{y} for any d -dimensional matrix \mathbf{Q} is:

$$\begin{aligned} \mathbf{y} &= \mathbf{Q}\mathbf{T} \\ \mathbf{T} &= (\mathbf{Q}_{off}^T \mathbf{Q}_{off})^{-1} \mathbf{Q}_{off}^T \mathbf{S}_{off} \end{aligned} \quad (12)$$

where $\mathbf{Q}_{off} = [\mathbf{Q}_0; \mathbf{Q}_1]^T$, $\mathbf{S}_{off} = [\mathbf{S}_0; \mathbf{S}_1]^T$, \mathbf{S}_0 is a $1 \times M_0$ zero vector and \mathbf{S}_1 is a $1 \times M_1$ unity vector.

As can be easily derived, the HI obtained by the linear transformation method as Equation (12) is varying approximately between 1 and 0. With the usage of the aircraft engine, the HI has a decreasing trend. This HI contains health condition information extracted from multi-dimensional informative sensors of the monitored aircraft engine. It can be used to construct background health knowledge in the offline process and to further conduct the online prediction process. Although the linear transformation is discussed here and will be employed for the case studies, other information fusion methods can also be used to construct the HI with the selected informative sensors.

3.3. Degradation Modeling by FPCA

In view that the health degradation is a stochastic process with uncertainties from physical degradation dynamics, usage variations and other effects, the degradation pattern of an aircraft engine is complex and unknown. To adaptively study the aircraft engine degradation, the HI that describes the health deterioration of an engine is assumed as discrete sampling values of the functional data and is modeled with the FPCA method in this subsection.

Suppose there is a population of aircraft engines, then the degradation pattern of one aircraft engine is an independent realization of the degradation process of the whole population. Based on the FPCA analysis method briefly introduced in Section 2, the degradation of one engine can be represented by the mean function and FPCs derived adaptively from the population stochastic degradation process $X(t)$ ($t \in I$), where I is the service time interval of the population of engines (the observation interval for the engine/system with the longest possible lifetime). Further, considering the fact that the HI indirectly demonstrates the health degradation of the engine, so the measurement error is one important aspect that should be accounted for. To be specific, for a population of n aircraft engines with N_i measurements for the i -th engine, the degradation of the i -th engine is modeled with its HI $y_i(t_{ij})$ ($i = 1, 2, \dots, n; j = 1, 2, \dots, N_i$) by FPCA as:

$$\begin{aligned} y_i(t_{ij}) &= x_i(t_{ij}) + e_{ij} \\ x_i(t) &= \mu(t) + \sum_{k=1}^K \varepsilon_{ik} \phi_k(t) \end{aligned} \quad (13)$$

where $x_i(t_{ij})$ and e_{ij} are the unobservable degradation and measurement error of the i -th engine at time t_{ij} , $x_i(t)$ is the underlying functional data with the mean function $\mu(t)$ and the FPCs $\phi_k(t)$ s for the i -th engine, and ε_{ik} are the related FPC scores.

With the FPCA-based degradation modeling as Equation (13), the degradation characteristics of the whole engine population is represented by the mean function and the first K FPCs, while the degradation peculiarity of an engine is captured by its specific FPC scores. The mean function $\mu(t)$ describes the common degrading trend of all the engines from an engine population, and the first few FPCs reflect the main varying modes of the population degradation process. As in [27,34], the measurement errors e_{ij} are practically assumed to be i.i.d with the normal distribution $N(0, \sigma^2)$ and are also independent of FPC scores in this study. Also, note that no pre-specified parametric form is

needed to be assumed for $\mu(t)$ and the $\phi_k(t)s$, but they are adaptively derived from the engines HI in the following.

For the mean function $\mu(t) = E(X(t))$, it is estimated by local linear smoothers using degradation observations $y_i(t_{ij})$ ($i = 1, 2, \dots, n; j = 1, 2, \dots, N_i$) from all n engines as:

$$\hat{\mu}(t) = \alpha_0(t) + \min_{\alpha_0, \alpha_1} \sum_i \sum_j \kappa_1\left(\frac{t_{ij}-t}{h_\mu}\right) (y_i(t_{ij}) - \alpha_0 - \alpha_1(t - t_{ij}))^2 \quad (14)$$

where $\kappa_1(\cdot)$ is a univariate kernel function and h_μ is the smoothing bandwidth.

For estimations of the $\phi_k(t)s$, the covariance function $G(s, t) = E((X(s) - \mu(s))(X(t) - \mu(t)))$ of the engine degradation process $X(t)$ should be firstly estimated. For the n engines, the following holds:

$$\begin{aligned} V_i(t_{ij}, t_{il}) &= (y_i(t_{ij}) - \hat{\mu}(t_{ij}))(y_i(t_{il}) - \hat{\mu}(t_{il})) \\ E(V_i(t_{ij}, t_{il})) &= V(t_{ij}, t_{il}) = G(t_{ij}, t_{il}) + \sigma^2 \delta_{jl} \end{aligned} \quad (15)$$

where δ_{ij} is the Kronecker delta function.

It is easy to see that the diagonal (i.e., $j = l$) of the raw covariance function $V(t_{ij}, t_{il})$ calculated with the degradation observations are contaminated with measurement errors. Estimation of the covariance function $G(s, t)$ using the local linear surface smoothing should be without the diagonal as:

$$\hat{G}(s, t) = \beta_0(s, t) + \min_{\beta_0, \beta_{11}, \beta_{12}} \sum_i \sum_{1 \leq j \neq l \leq N_i} \kappa_2\left(\frac{t_{ij}-s}{h_G}, \frac{t_{il}-t}{h_G}\right) (V_i(t_{ij}, t_{il}) - \beta_0 - \beta_{11}(s - t_{ij}) - \beta_{12}(t - t_{il}))^2 \quad (16)$$

where $\kappa_2(\cdot)$ is a bivariate kernel function and h_G is the smoothing bandwidth.

Then eigen-decomposition following Equation (2) by discretizing the $\hat{G}(s, t)$ is performed to obtain the eigenvalues $\hat{\lambda}_k$ ($k = 1, 2, \dots$) and the corresponding FPCs. Spline interpolation is further utilized to obtain the continuous $\hat{\phi}_k(t)s$.

For the variance σ^2 of the measurement errors, it is related to the diagonal values of $V(t_{ij}, t_{il})$ and $G(s, t)$. To mitigate boundary effects, its estimation is:

$$\hat{\sigma}^2 = \max\left(\frac{2}{I} \int_{I_1} (\hat{V}(t) - \tilde{G}(t)) dt, 0\right) \quad (17)$$

where $\hat{V}(t)$ is the local linear smoother of diagonal values of $V(t_{ij}, t_{il})$, $\tilde{G}(t)$ is the diagonal values of $\hat{G}(s, t)$, and the interval $I_1 = [I/4, 3I/4]$.

Considering the measurement errors in the degradation model, to estimate FPC scores with Equation (4) will lead to bias. To remedy that, the conditional expectation method proposed by Yao et al. [35] is utilized here to estimate the ε_{ik} :

$$\hat{\varepsilon}_{ik} = \hat{\lambda}_k \hat{\phi}_{ik}^T \hat{\Sigma}_i^{-1} (\mathbf{y}_i - \hat{\boldsymbol{\mu}}_i) \quad (18)$$

where $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{iN_i})^T$ is the engine HI observing vector, $\hat{\boldsymbol{\mu}}_i = (\hat{\mu}_{i1}, \hat{\mu}_{i2}, \dots, \hat{\mu}_{iN_i})^T$ and $\hat{\phi}_{ik} = (\hat{\phi}_{i1}, \hat{\phi}_{i2}, \dots, \hat{\phi}_{iN_i})^T$ are the mean and FPC vectors interpolated from the mean function and FPCs, and $\hat{\Sigma}_i$ is with the entry (j, l) as $\hat{\Sigma}_i(j, l) = \hat{G}(t_{ij}, t_{il}) + \hat{\sigma}^2 \delta_{jl}$.

With the estimated parameters, the uncontaminated degradation of the i -th engine is modeled by FPCA as:

$$\hat{x}_i(t_{ij}) = \hat{\mu}(t_{ij}) + \sum_{k=1}^K \hat{\varepsilon}_{ik} \hat{\phi}_k(t_{ij}) \quad (19)$$

Based on the FPCA modeling analysis, degradation of an engine is statistically represented with the common trend and a few varying terms that reflect the main dynamics of the degradation process. For estimation of the model parameters, all degradation observations of the engine population are fused and local linear smoothing is needed. In this study, the Gaussian kernel function is applied for both curve and surface smoother, and the related smoothing bandwidth is determined by the one-leave-out cross-validation strategy.

3.4. Prognostics with Bayesian Inference

For one aero-engine that is still operating, its future health is of significance for condition-based maintenance and health management. Therefore, Bayesian inference is combined with the degradation model for engine prognostics in this Subsection.

With the degradation modeling by FPCA as detailed in the previous Subsection, the degradation trajectory of one in-service engine can be described as:

$$x(t) = \mu(t) + \sum_{k=1}^K \varepsilon_k \phi_k(t) \quad (20)$$

Assume that we have monitored the in-service engine at a vector of time $\mathbf{t} = (t_1, t_2, \dots, t_H)$ with the observed HI $\mathbf{y}_{1:H} = (y_1, y_2, \dots, y_H)^T$, where t_H denotes the latest monitoring time. Once the mean function $\mu(t)$ and FPCs $\phi_k(t)$ s are estimated with HI of historical run-to-failure engines of the population as elaborated in Section 3.3, the trend and the main variations of the in-service engine degradation can be obtained by interpolations with its monitoring times. However, the FPC scores $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_K)^T$ are to be inferred.

In practice, system health deterioration is usually a gradual process under the effects of various internal and external environmental factors, and the degradation process can be assumed as a Gaussian process. Then the uncorrelated FPC scores ε_k ($k = 1, 2, \dots, K$) will respectively follow the normal distributions $N(0, \lambda_k)$ with λ_k being the ordered eigenvalues of the covariance function of the degradation process. Taking FPC scores estimated from the historical run-to-failure engines as the prior distribution $p_0(\boldsymbol{\varepsilon})$ and considering the likelihood function $p(\mathbf{y}_{1:H}|\boldsymbol{\varepsilon})$ of observing $\mathbf{y}_{1:H}$ [36], we have:

$$\begin{aligned} p_0(\boldsymbol{\varepsilon}) &\sim N(0, \boldsymbol{\Lambda}), \boldsymbol{\Lambda} = \text{diag}(\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_K) \\ p(\mathbf{y}_{1:H}|\boldsymbol{\varepsilon}) &= \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp\left(-\frac{\mathbf{e}^T \mathbf{e}}{2\hat{\sigma}^2}\right), \mathbf{e} = \mathbf{y}_{1:H} - \boldsymbol{\mu}_{1:H} - \boldsymbol{\Phi} \boldsymbol{\varepsilon} \end{aligned} \quad (21)$$

where $\boldsymbol{\mu}_{1:H}$ and $\boldsymbol{\Phi}$ with entry (h, k) as $\hat{\phi}_k(t_h)$ ($h = 1, 2, \dots, H; k = 1, 2, \dots, K$) are respectively interpolated with the monitoring time vector $\mathbf{t} = (t_1, t_2, \dots, t_H)$ from the $\hat{\mu}(t)$ and $\hat{\phi}_k(t)$ s estimated from the run-to-failure engines, and $\hat{\sigma}^2$ is the estimated measurement error.

Based on Bayesian inference outlined in Section 2.2, the posterior distribution $p(\boldsymbol{\varepsilon}|\mathbf{y}_{1:H})$ of the FPC scores $\boldsymbol{\varepsilon}$ of the in-service engine can be updated as:

$$p(\boldsymbol{\varepsilon}|\mathbf{y}_{1:H}) \propto p_0(\boldsymbol{\varepsilon})p(\mathbf{y}_{1:H}|\boldsymbol{\varepsilon}) \quad (22)$$

From Equation (21), the $p_0(\boldsymbol{\varepsilon})$ and $p(\mathbf{y}_{1:H}|\boldsymbol{\varepsilon})$ are conjugate multivariate normal distributions. The $p(\boldsymbol{\varepsilon}|\mathbf{y}_{1:H})$ can be analytically derived to also follow multivariate normal distribution as follows:

$$\begin{aligned} p(\boldsymbol{\varepsilon}|\mathbf{y}_{1:H}) &\sim N(\boldsymbol{\mu}_H^\boldsymbol{\varepsilon}, \boldsymbol{\Sigma}_H^\boldsymbol{\varepsilon}) \\ \boldsymbol{\Sigma}_H^\boldsymbol{\varepsilon} &= \left(\frac{\boldsymbol{\Phi}^T \boldsymbol{\Phi}}{\hat{\sigma}^2} + \boldsymbol{\Lambda}^{-1}\right)^{-1}, \boldsymbol{\mu}_H^\boldsymbol{\varepsilon} = \boldsymbol{\Sigma}_H^\boldsymbol{\varepsilon} \frac{\boldsymbol{\Phi}^T}{\hat{\sigma}^2} (\mathbf{y}_{1:H} - \boldsymbol{\mu}_{1:H}) \end{aligned} \quad (23)$$

Further inserting the posteriori distribution of the FPC scores ε into Equation (20), a real-time prognosis for the in-service engine based on the latest observations $\mathbf{y}_{1:H}$ is:

$$\begin{aligned} x(t) &\sim N(\mu_H^x(t), \sigma_H^{2x}(t)) \\ \mu_H^x(t) &= \mu(t) + \boldsymbol{\phi}^T(t) \boldsymbol{\mu}_H^\varepsilon, \sigma_H^{2x}(t) = \boldsymbol{\phi}^T(t) \boldsymbol{\Sigma}_H^\varepsilon \boldsymbol{\phi}(t) \end{aligned} \quad (24)$$

where $\mu(t) = \hat{\mu}(t)$ and $\boldsymbol{\phi}(t) = (\hat{\phi}_1(t), \hat{\phi}_2(t), \dots, \hat{\phi}_K(t))^T$ are respectively the mean value and FPC vector interpolated from the mean function and K FPCs at a future time t .

4. Case Studies and Results Analysis

The main purpose of this section is to demonstrate the validity and performance of the proposed prognostics approach with the case study on an aircraft gas turbine engine.

4.1. Sensor Data of Aircraft Engine

The sensor data of the aircraft gas turbine engine degradation is generated from the commercial modular aero-propulsion system simulation (C-MAPSS) developed at NASA [37] and published online for research investigations. Each time series signal represents a different degradation instance of the dynamic simulation of the same engine population and consists of multi-sensor measurements. For each cycle of a degradation instance, 21 sensor measurements, as listed in Table 1, were recorded. The multi-sensor data was contaminated with noises and each engine started with different initial health conditions and manufacturing variations, which was unknown. With limited knowledge about the true physical model, we relied solely on the multi-sensor data from the training and testing engines to understand the engine degradation process. The data can be downloaded from NASA data repository [38].

In particular, the data set FD001 that contains 100 training and testing engines is used in this study. The multi-sensor measurements for each training engine were collected until failure, whereas the multi-sensor measurements for each in-service unit were truncated at some random point before its failure. Although these engines were simulated under one operational condition and one failure mode, the engine with multiple usage conditions and failure modes may also be analyzed using the proposed method with necessary processing as in [20,39].

Table 1. Sensor measurements of the aircraft engine.

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

4.2. Results and Analysis on Informative Sensor Selection

Before evaluation and selection, a rough screening of the 21 sensors tells that the sensors T2, P2, P15, epr, farB, htBleed, Nf_dmd, and PCNfR_dmd are of binary or constant value and are excluded. With the remaining 13 sensors of the 100 training engines, the discussed four metrics and the two metrics of Pro and trendability (Tre) in [24] are calculated as in Table 2. For metrics of Mon, Corr, and Rob that are defined on sensors from one engine, statistics of the mean and standard deviation (std) are obtained based on results of the 100 engines. From the results, four sensors Nf, Nc, NRf, and NRc were found to be of significant difference from the other nine sensors.

Table 2. Evaluation results of the sensors for the training data set of FD001. Abbreviations: Mon, monotonicity; Corr, correlation; Rob, robustness; Pre, predictability; Pro, prognosability; Tre, trendability.

Symbol	Mon		Corr		Rob		Pre	Pro	Tre
	Mean	std	Mean	std	Mean	std			
T24	0.0715	0.0397	0.8365	0.0391	0.9997	2.1×10^{-5}	0.8048	0.8048	0.9679
T30	0.0515	0.0395	0.8241	0.0397	0.9982	1.1×10^{-4}	0.7619	0.7624	0.9709
T50	0.0929	0.0456	0.8751	0.0294	0.9980	1.5×10^{-4}	0.8709	0.8709	0.9642
P30	0.0756	0.0533	0.8660	0.0382	0.9995	3.0×10^{-5}	0.8381	0.8381	0.9660
Nf	0.1517	0.0554	0.7986	0.1639	1.0000	6.3×10^{-5}	0.6543	0.6543	0.9667
Nc	0.1482	0.1137	0.7535	0.1797	0.9997	2.2×10^{-5}	0.2632	0.3250	0.9198
Ps30	0.1317	0.0531	0.8818	0.0262	0.9985	8.9×10^{-5}	0.8671	0.8671	0.9668
phi	0.0997	0.0494	0.8744	0.0312	0.9996	2.6×10^{-5}	0.8466	0.8466	0.9693
NRf	0.1565	0.0541	0.8027	0.1592	1.0000	5.2×10^{-7}	0.6630	0.6630	0.9650
NRc	0.1574	0.1211	0.7623	0.2004	0.9997	1.7×10^{-5}	0.1820	0.3106	0.9074
BPR	0.0743	0.0451	0.8546	0.0305	0.9983	1.0×10^{-4}	0.8419	0.8419	0.9719
W31	0.0559	0.0348	0.8544	0.0334	0.9982	1.2×10^{-4}	0.8289	0.8289	0.9704
W32	0.0661	0.0405	0.8558	0.0366	0.9982	1.1×10^{-4}	0.8091	0.8091	0.9682

For the two sensors Nf and NRf, they have high variation in metric Corr and relative low value in metric Pre. Thus, these two sensors should not be selected as the informative sensors based on the

analysis in Section 3.1. In addition, the two sensors score low values of metric Pro proposed in [24], verifying the effectiveness of the proposed metric Pre. Nevertheless, the metric Tre proposed also in [24] to characterize the trendability of a population of sensors scores both high for the two sensors. To make sense of interpretation, signals of the two sensors for the 100 training engines are shown in Figure 2. It shows that the two sensors are not correlated well with the usage of the engine and their ranges are both very small (actually less than 0.5) from start to failure, which validates these sensors have high variation in metric Corr and score low value in metric Pre and Pro. In previous studies using the C-MAPSS data, including [9,20], the above two sensors were also wiped out as non-informative sensors.

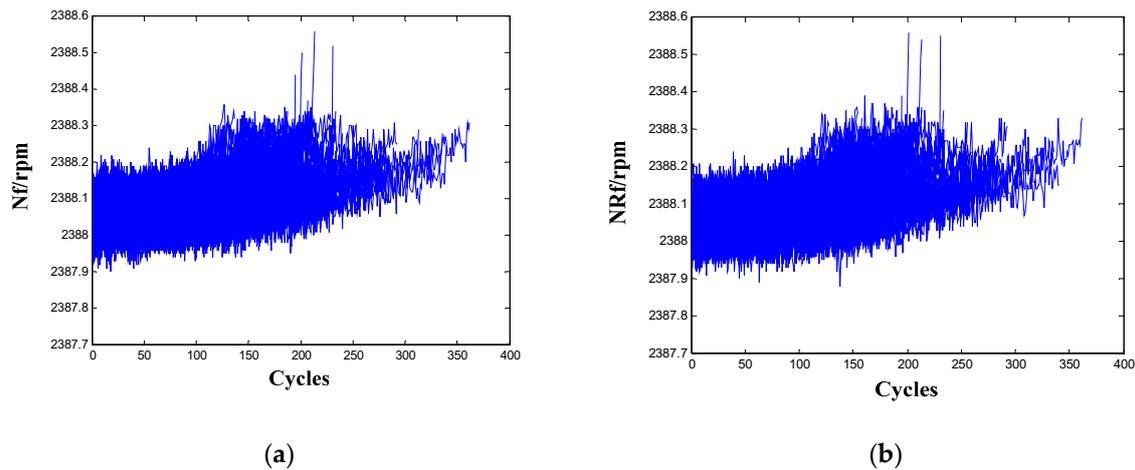


Figure 2. Sensor signals for the 100 training engines: (a) Nf; (b) NRf.

From the results in Table 2, the sensors Nc and NRC have high variations in both metric Mon and Corr, and their Pre scores are even lower. Also, the Pro and Tre scores of the two sensors are relatively lower as compared to other sensors. Thus, the two sensors also should not be selected as informative sensors. To interpret the evaluation results, signals of the two sensors for the 100 engines are given in Figure 3. It is observed that the two sensor signals demonstrate a very inconsistent degrading trend for the population of training engines that operate under the same condition and with one failure mode. There are increasing, decreasing, as well as oscillating patterns in the two sensors, which explains the high variations in metric Mon and Corr, as well as the low scores in metric Pre, Pro, and Tre.

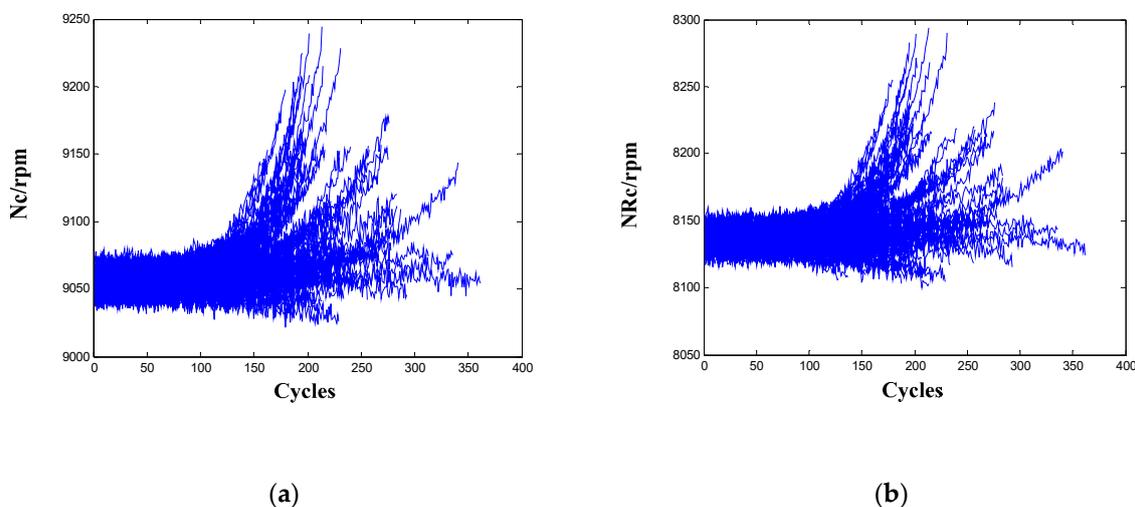


Figure 3. Sensor signals for the 100 training engines: (a) Nc; (b) NRC.

For the three metrics Pre, Pro, and Tre that consider the trend consistency of a sensor for a population operate under the same condition and with one failure mode, some comparisons can be made from Table 2. The metric Tre in [24] is effective to sift out sensors without a consistent trend for the population of engines by a lower score, it cannot sift out the small range sensor, whose correlation with the degradation is poor. The proposed metric Pre and the metric Pro in [24] are both effective for wiping out these two kinds of sensors by evident lower scores. Furthermore, the Pre punishes the first kind of sensors with a lower score than that of the Pro, thus it is much easier to identify non-informative sensors with the proposed Pre. Sensors for the training engines from FD003 with the same failure mode as FD001 (i.e., HPC failure) are also evaluated and the results are in Table 3. Similar results can be drawn as those of FD001, which further shows the validity of the proposed metric Pre.

Table 3. Evaluation results of the sensors for the training data set of FD003 with the same failure mode of FD001.

Symbol	Mon		Corr		Rob		Pre	Pro	Tre
	Mean	std	Mean	std	Mean	std			
T24	0.0711	0.0427	0.8415	0.0350	0.9997	2.7×10^{-5}	0.7723	0.7723	0.9708
T30	0.0670	0.0420	0.8261	0.0367	0.9982	1.0×10^{-4}	0.7991	0.7991	0.9707
T50	0.1020	0.0499	0.8781	0.0253	0.9980	1.0×10^{-4}	0.8542	0.8542	0.9691
P30	0.0868	0.0404	0.8741	0.0283	0.9995	3.0×10^{-5}	0.8348	0.8348	0.9714
Nf	0.1545	0.0514	0.8334	0.1098	1.0000	6.2×10^{-7}	0.7024	0.7040	0.9664
Nc	0.1339	0.1143	0.7656	0.1827	0.9997	1.8×10^{-5}	0.2071	0.2946	0.9198
Ps30	0.1371	0.0506	0.8847	0.0217	0.9985	8.9×10^{-5}	0.8878	0.8878	0.9709
phi	0.1020	0.0501	0.8838	0.0213	0.9996	2.8×10^{-5}	0.8659	0.8659	0.9649
NRf	0.1530	0.0551	0.8256	0.1225	1.0000	6.6×10^{-7}	0.6990	0.6998	0.9665
NRc	0.1388	0.1098	0.7479	0.2379	0.9997	1.9×10^{-5}	0.1386	0.2760	0.9183
BPR	0.0740	0.0473	0.8583	0.0305	0.9983	1.1×10^{-4}	0.8487	0.8489	0.9660
W31	0.0549	0.0368	0.8532	0.0376	0.9982	1.2×10^{-4}	0.8388	0.8388	0.9691
W32	0.0607	0.0351	0.8544	0.0330	0.9982	8.3×10^{-5}	0.8280	0.8280	0.9719

From Table 2, there is no significant difference for the evaluation results of the other nine sensors, so they are selected as the informative sensors for the HI construction of the engines. To contrast with the non-informative sensors, signals of two selected sensors are illustrated in Figure 4.

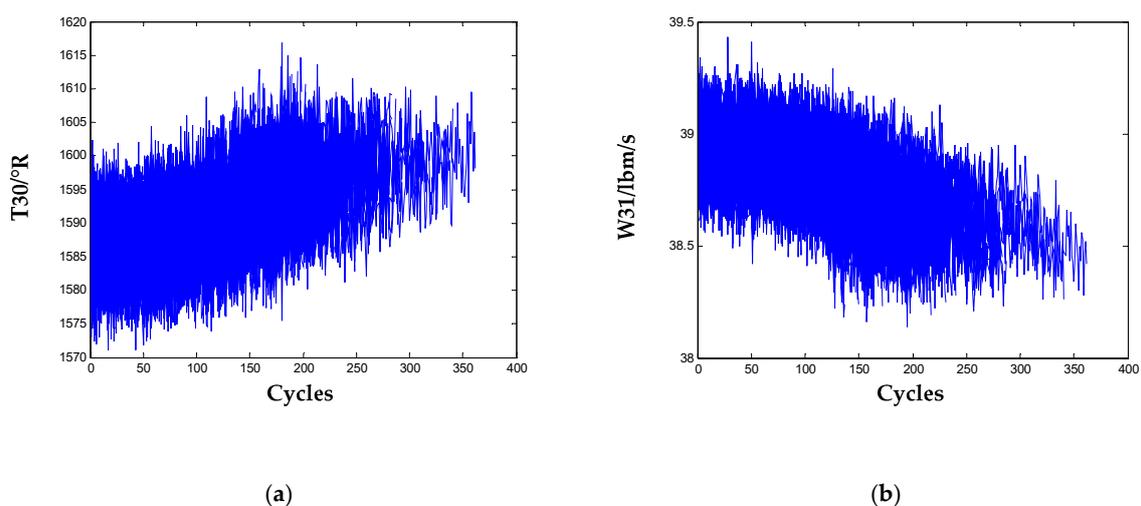


Figure 4. Sensor signals for the 100 training engines: (a) T30; (b) W31.

4.3. Results and Analysis on Degradation Modeling

With the selected nine-dimensional informative sensors, HI that describes the degradation of the engine is constructed as Equation (12). Data matrices for healthy and failed states of the engine are

respectively created using the nine-sensor measurements of the initial five cycles and the failure cycle of all the engines. The HI of one engine and all the training engines are displayed in Figure 5. It is observed that the HI of all the engines are with the same decreasing trend and roughly vary between 1 and 0, which means a smaller HI relates to a less healthy condition of the engine.

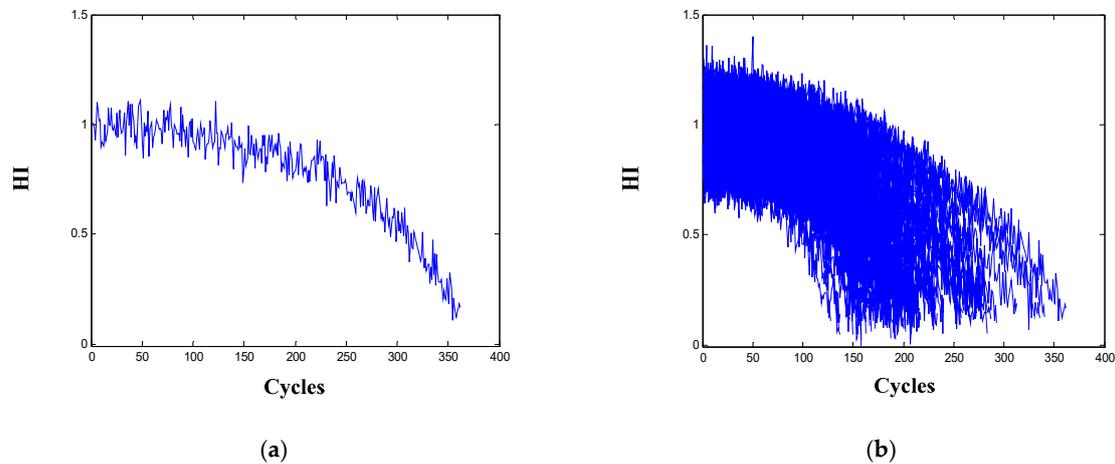


Figure 5. Health index (HI) for: (a) one engine; (b) all the 100 training engines.

Based on the HI of the 100 training engine systems, parameters of the FPCA-based degradation model can be estimated as stated in Section 3.3 and the main ones are shown in Figure 6. From Figure 6a, it is observed that the estimated mean function adaptively captures the general degrading trend of the 100 engines population. To ensure that sufficiently variance is retained in the truncated expansion as in Equation (19), the threshold η is set as 0.95 and the truncated terms K is 7 using Equation (7). That is, the first seven FPCs are needed to explain more than 95% of the variations of the engine degradation process, and each of the FPCs reflects one distinct varying mode of the engine deterioration process. After modeling by the FPCA method, the background health information of the 100 engines is abstracted into few parameters, i.e., the estimated mean function and the first seven FPCs, which are the common characteristics shared by engines from the training population, and the corresponding FPC scores specify the degradation peculiarity of one specific engine.

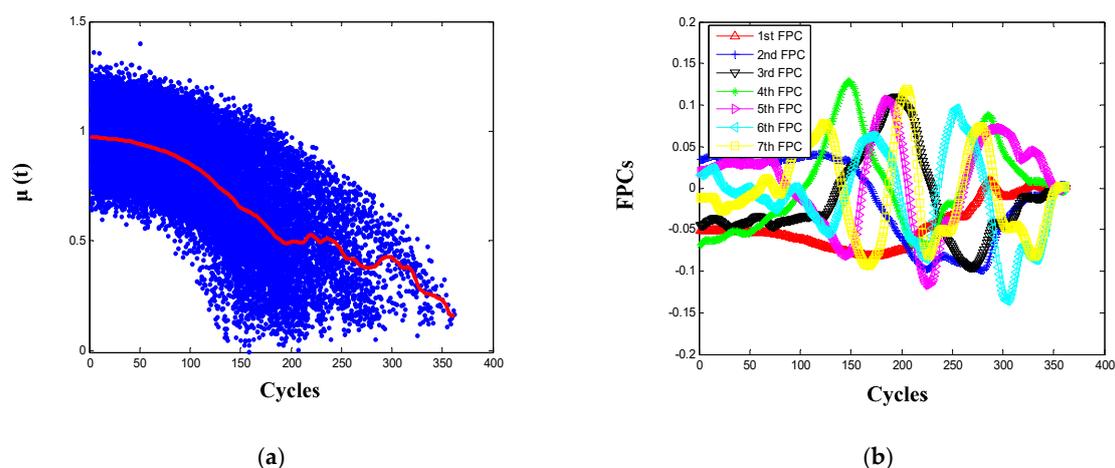


Figure 6. Estimation results: (a) the mean function; (b) the first seven functional principal components (FPCs).

To demonstrate the performance of the proposed method, the true parametric model of exponential (Exp) $y(t) = a \times \exp(bt) + c$ [37], power law (Pow) as in [17] and 3rd polynomial (Poly3) as in [20]

are also used to fit the HI of the simulated engines. Degradation modeling results for one engine are shown in Figure 7. It can be observed that the FPCA method and the three parametric models fit the degradation pattern well, but the proposed degradation model in Equation (19) reveals more local degradation dynamics of the engine. For quantitative comparison, sum of square error (SSE), the coefficient of determination (R-square) [40], and root mean square error (RMSE) are statistically calculated for the 100 engines as in Table 4.

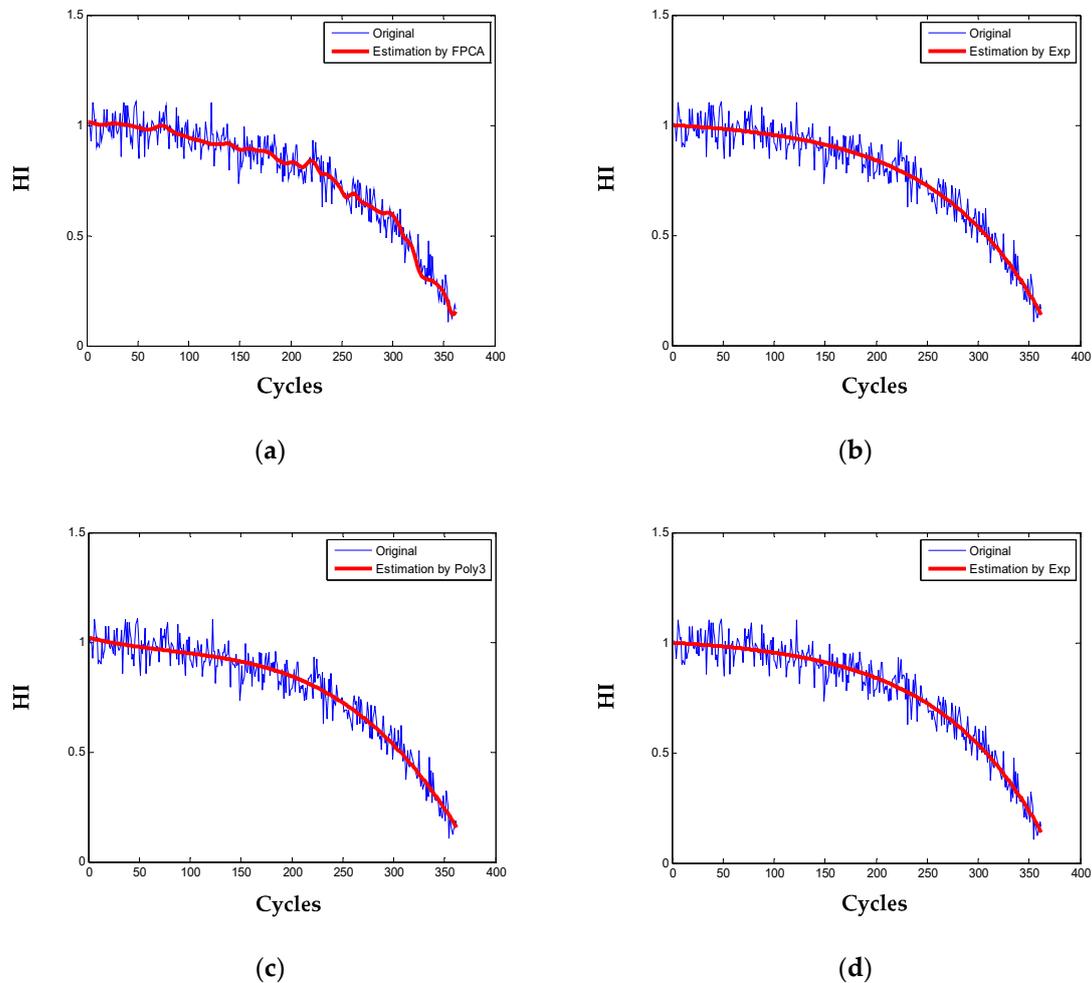


Figure 7. Degradation modeling for one engine: (a) the proposed functional principal component analysis (FPCA); (b) the fitting by Exponential; (c) the fitting by 3rd order polynomial; (d) the fitting by Power law.

Table 4. Comparisons of the degradation modeling results. Abbreviations: SSE, sum of square error; R-square, the coefficient of determination; RMSE, root mean square error.

Method	SSE		R-Square		RMSE	
	Mean	std	Mean	std	Mean	std
FPCA	0.6597	0.1571	0.9305	0.0224	0.0565	0.0026
Exp	0.6423	0.1567	0.9324	0.0219	0.0562	0.0025
Poly3	0.6673	0.1592	0.9210	0.0219	0.0570	0.0026
Pow	0.6736	0.1658	0.9194	0.0222	0.0575	0.0027

Results in Table 4 show that the proposed method performs better than the Pow and Poly3 models, and demonstrates comparable performance to the true Exp model. However, the proposed method

models the degradation adaptively without any assumption on the parametric form of the degradation pattern. This will be more significant when little knowledge is known about the latent degradation trend. Also, all the 100 engine degradation observations from the training set are pooled to estimate the parameters of the proposed degradation model. While in the parametric models, fittings are carried out with one individual engine independently, common information about the degradation process of the engine population is prone to be lost. In the following, the FPCA method is combined with Bayesian inference to predict the aero-engine health and comparison are made with the true Exp degradation model.

4.4. Results and Analysis on Health Prognostics

For validation of the presented method for long-term health prognostics, the multiple sensor data of the 100 testing engines from FD001 is processed following the same procedures of the training engines as detailed above.

Through linear transformation of the nine selected informative sensors with the coefficients learned from the training engines, the HI of all the testing engines and of two engines are shown in Figure 8. With the operation of one engine, its health condition degrades and the HI declines as the same pattern of the training engines.

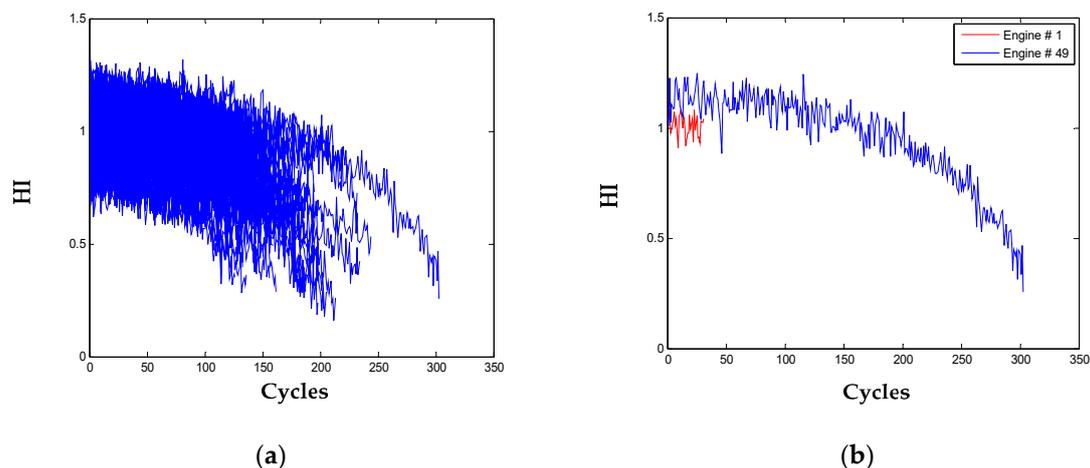


Figure 8. HI for: (a) all the 100 testing engines; (b) two engines.

With the estimated results from the training engines as the prior distribution $p_0(\epsilon)$ for FPC scores $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_7)^T$, the ϵ of a testing engine is updated to a posteriori distribution $p(\epsilon|\mathbf{y}_{1:H})$ with its latest HI observations $\mathbf{y}_{1:H}$ using Equation (21) as discussed in Section 3.4. Statistically, the 100 testing engines are truncated from about 20% to 96% of their true useful lives. In the following, the testing engines with enough observations are chosen to demonstrate the prognostic inference.

The testing engine # 49 consumed about 96% of its true useful life (i.e., 324 cycles) before being truncated and there are in total 303 observations. Prognostics of the engine is performed at 35%, 50%, 80%, and 100% of its health monitoring history with, respectively, $H = 106, 152, 242,$ and 303 historical observations. The health prediction results are as Figure 9. When there is not enough historical observations (in cases of 35% and 50%), the parametric model of Exp is poorly fitted and the engine long-term health is over or under estimated. While the proposed method predicts long-term health of the engine more accurately by utilization of the mean function and the first seven FPCs learned from the training engines. With the approaching of its failure (in cases of 80% and 100%), both methods track and predict the engine health degradation with high accuracy since there accumulates sufficient historical observations to fit or update the model parameters. Also note that the proposed method reveals more local dynamics of the engine degradation, which is beneficial for explaining of the degradation variation.

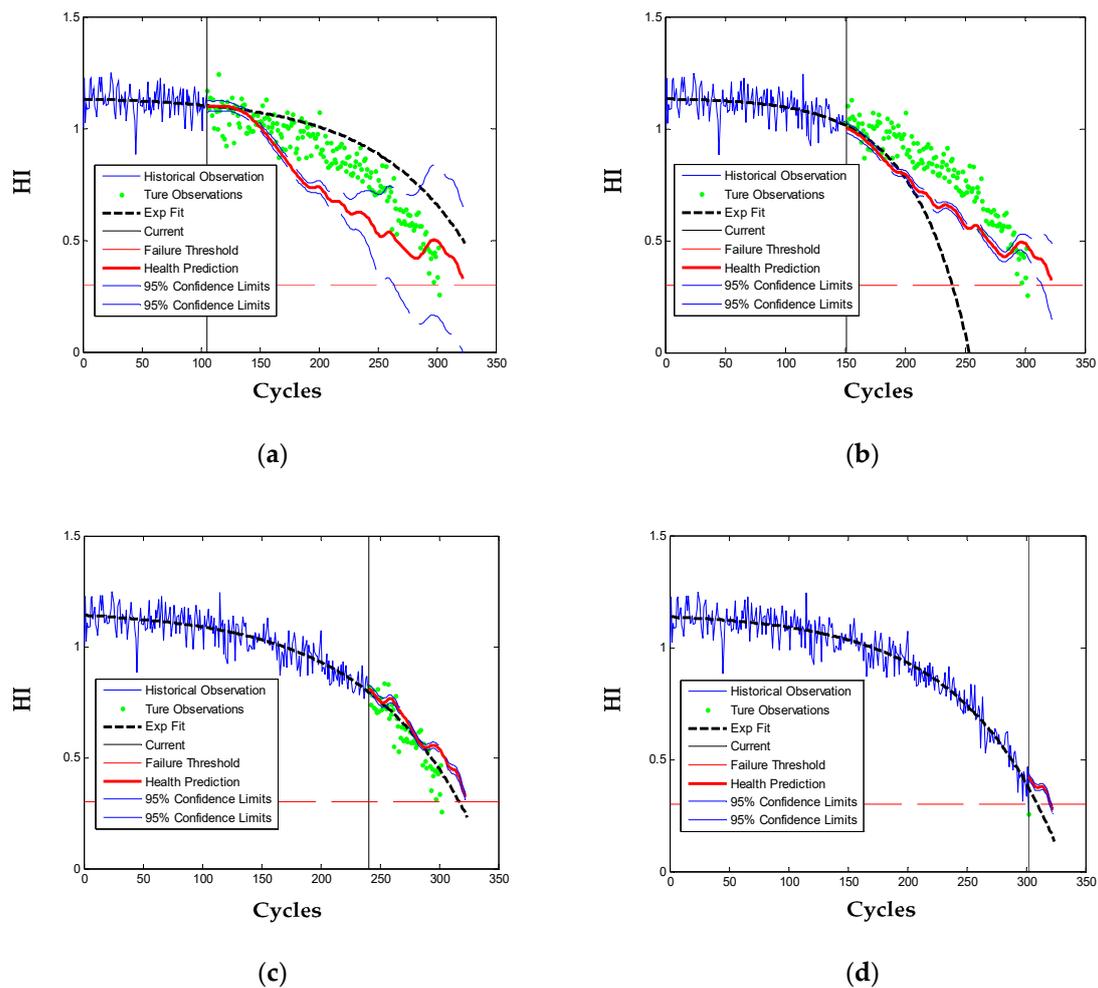


Figure 9. HI prediction for the testing engine # 49 at four percentages of the observations: (a) 35%; (b) 50%; (c) 80%; (d) 100%.

Besides the testing engine # 49, health of the other nine testing engines which are observed to consume more than 90% of their true useful lives are also predicted at four monitoring phases (i.e., 35%, 50%, 80%, and 90%) of their health monitoring histories. The HI prediction RMSEs for all the 10 engines are statistically compared in Table 5. It can be observed that the proposed method outperforms the true Exp model in the very long-term prediction, when the engines are in the early phase of service and only 35% of observations are used. With such limited historical HI, some engines are even wrongly fitted, such as # 20 and # 34. For the case of 50% observations, the proposed method also predicts better, except for engine # 20 and # 81. When the engines run into their later stages (in cases of 80% and 90%), performance of the proposed method is comparable to the true Exp model as there are sufficient HI observations to fit the parametric model.

To further illustrate the RUL prediction ability of the proposed approach, HI of the 100 testing engines are predicted with FPC scores updated by the last recorded observations and are extrapolated to the pre-defined failure threshold (i.e., $HI_f = 0.3$) to estimate the final failure cycles. Results of the 100 testing engines are plotted in Figure 10. For the proposed statistical method, both the mean and median of 100 simulations are given. For the engine whose HI cannot be properly fitted to the Exp model, its life is estimated as 206 cycles, which is the mean life of the 100 training engines. It can be observed that the predicted RULs by the proposed method and Exp model are close to the actual values in the region where the RUL value is small. That is because when the engine is working closer to the failure, the degradation is enhanced and can be captured for better prognostics. However, when

the engines are far from failure, the mean and median prediction results of the proposed method are more accurate than the parametric Exp model.

Table 5. Comparison of health prognostic results for the 10 testing engines.

Engine #	RMSE							
	35%		50%		80%		90%	
	Exp	FPCA	Exp	FPCA	Exp	FPCA	Exp	FPCA
20	3.4×10^4	0.0800	0.0591	0.0880	0.0570	0.0737	0.0664	0.0910
31	0.1568	0.1345	0.9053	0.1395	0.0638	0.0954	0.0532	0.0829
34	3.4×10^3	0.1688	1.4623	0.1506	0.0649	0.0981	0.0677	0.0921
35	0.1599	0.0818	0.1701	0.0730	0.1091	0.0813	0.0535	0.0727
42	1.5788	0.0960	0.1034	0.0971	0.0710	0.0863	0.0534	0.0659
49	0.1345	0.1149	0.9053	0.1119	0.0638	0.0752	0.0532	0.0795
68	2.4974	0.1100	0.5035	0.1246	0.0621	0.0959	0.0844	0.0962
76	20.285	0.1038	0.1593	0.1264	0.0701	0.0969	0.0594	0.0666
81	0.3026	0.1242	0.1005	0.1292	0.1672	0.1216	0.0624	0.0936
82	3.1×10^4	0.1052	0.5006	0.1391	0.0735	0.0903	0.0671	0.0661
mean	6.8×10^3	0.1119	0.4869	0.1179	0.0796	0.0915	0.0611	0.0807

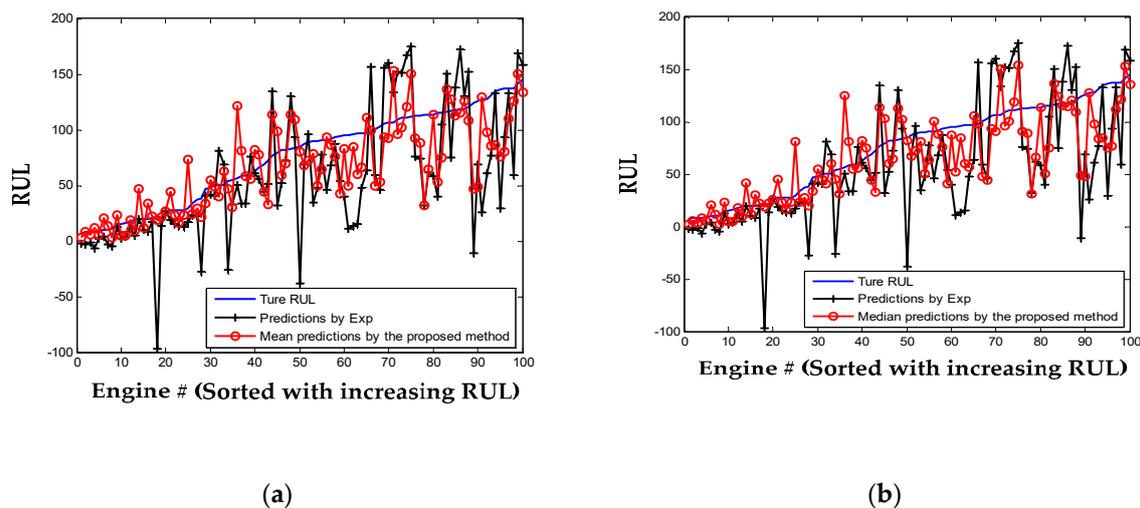


Figure 10. Remaining useful life (RUL) predictions for the 100 testing engines: (a) the mean by the proposed method; (b) the median by the proposed method.

The common score used to evaluate the performance of prognostic method for the C-MAPSS data set as Equation (25) and prediction RMSE are also calculated and listed in Table 6. The far lower score of the proposed method (both mean and median) as well as the smaller RMSE indicate that the performance of the proposed prognostic method is better than the Exp model. Also, narrower prediction errors range and a higher number of correct-predictions are achieved using the proposed method. From the above analysis, the proposed adaptive method is more robust than the parametric Exp method, especially when one engine is in an incipient service stage of its life.

$$Score = \sum_{i=1}^n S_i, S_i = \begin{cases} \exp\left(-\frac{d_i}{13}\right) - 1, & \text{when } d_i < 0 \\ \exp\left(\frac{d_i}{10}\right) - 1, & \text{when } d_i \geq 0 \end{cases} \quad (25)$$

where d_i = predicted RUL of the i -th engine—true RUL of the i -th engine ($i = 1, 2, \dots, n$).

Table 6. Comparison of RUL prediction results for the 100 testing engines.

Method	Score	RMSE	Rang of Prediction Errors	Under-Predictions	Correct-Predictions	Over-Predictions
Exp	67,352	45.40	[−135, 63]	76	1	23
FPCA	mean	3092	[−82, 65]	65	3	32
	median	3567	[−83, 68]	67	5	28

5. Discussion

In this paper, a data-driven statistical prognostic approach for informative sensor selection and adaptive degradation modeling based on FPCA is introduced. Effectiveness of the proposed method is validated by the case study on an aircraft engine simulation experiment. Nevertheless, some key issues should be tackled for application of the presented method to real engines and other systems.

For the proposed approach to be effective for real engines, enough run-to-failure sensor data should be provided, and this is a common requirement for a data-driven approach. In real aircraft engine operation, the degradation of one engine usually begins after one long normal working stage, the so-called delay-time phenomenon. Modeling of the potential-to-failure stage is of more significance. Thus, it is better to apply the proposed method upon the detection of incipient faults. For few instances of real engine failures, and that is often the case, the jointly Gaussian assumption may no longer hold for the parameter estimation. To partially tackle this problem, some kind of bootstrap methods can be useful. As for the extension of the proposed method, especially the adaptive degradation modeling by FPCA, to monitor other systems, efforts should be put especially to the construction of a HI that relates to the degradation of these systems, as the linear data transformation method used in this paper may not be applicable.

6. Conclusions

To model and track the complex degradation pattern of aircraft engines for accurate and efficient prognosis, a novel method for informative sensor selection and adaptive degradation track is proposed in this paper. The deterioration sensitive sensors can be selected by the presented metrics and be fused to construct a health index that describes the degradation of an aircraft engine. Taking the degradation index of one engine as functional data, the degradation process of the engine population is then adaptively modeled by the FPCA, and future health is predicted with Bayesian inference. Experimental studies are performed on the sensor dataset of aircraft gas turbine engines, and the results verify that the proposed method can effectively select the informative sensors to model and predict the complex degradation process of the aircraft engine. The failure threshold is set to a fixed value in this study; however, large variation of the failure value requires a random failure threshold to be pursued in the future.

Author Contributions: Conceptualization, B.Z. and Q.H.; methodology, B.Z.; software, B.Z.; validation, B.Z., Q.H., and K.Z.; formal analysis, S.F.; investigation, B.Z. and S.Z.; resources, B.Z.; data curation, B.Z.; writing—original draft preparation, B.Z.; writing—review and editing, B.Z.; visualization, Y.Z.; supervision, Y.Z.; project administration, B.Z.; funding acquisition, Q.H. and B.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (NSFC) (Grant No. 51605065 & No. 51905065), the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJ1704109), and the Open Found of The State Key Laboratory of Mechanical Transmissions (Grant No. SKLMT-KFKT-201809).

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Volponi, A.J. Gas turbine engine health management: Past, present, and future trends. *J. Eng. Gas Turbines Power* **2014**, *136*, 051201. [[CrossRef](#)]

2. Lei, Y.G.; Li, N.P.; Guo, L.; Li, N.B.; Yao, T.; Lin, J. Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mech. Syst. Signal Process.* **2018**, *104*, 799–834. [[CrossRef](#)]
3. Hanachi, H.; Mechefske, C.; Liu, J.; Banerjee, A.; Chen, Y. Performance-based gas turbine health monitoring, diagnostics, and prognostics: A survey. *IEEE Trans. Reliab.* **2018**, *67*, 1340–1363. [[CrossRef](#)]
4. Wang, P.; Gao, R.X. Markov nonlinear system estimation for engine performance tracking. *J. Eng. Gas Turbines Power* **2016**, *138*, 091201. [[CrossRef](#)]
5. Al-Dahidi, S.; Di Maio, F.; Baraldi, P.; Zio, E. Remaining useful life estimation in heterogeneous fleets working under variable operating conditions. *Reliab. Eng. Syst. Saf.* **2016**, *156*, 109–124. [[CrossRef](#)]
6. Yu, J.B. Aircraft engine health prognostics based on logistic regression with penalization regularization and state-space-based degradation framework. *Aerosp. Sci. Technol.* **2017**, *68*, 345–361. [[CrossRef](#)]
7. Skordilis, E.; Moghaddass, R. A double hybrid state-space model for real-time sensor-driven monitoring of deteriorating systems. *IEEE Trans. Autom. Sci. Eng.* **2019**. [[CrossRef](#)]
8. Mosallam, A.; Medjaher, K.; Zerhouni, N. Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction. *J. Intell. Manuf.* **2016**, *27*, 1037–1048. [[CrossRef](#)]
9. Wang, C.S.; Lu, N.Y.; Cheng, Y.H.; Jiang, B. A data-driven aero-engine degradation prognostic strategy. *IEEE Trans. Cybern.* **2019**. [[CrossRef](#)]
10. Khelif, R.; Chebel-Morello, B.; Malinowski, S.; Laajili, E.; Fnaiech, F.; Zerhouni, N. Direct remaining useful life estimation based on support vector regression. *IEEE Trans. Ind. Electron.* **2016**, *64*, 2276–2285. [[CrossRef](#)]
11. Lu, F.; Wu, J.D.; Huang, J.Q.; Qiu, X.J. Aircraft engine degradation prognostics based on logistic regression and novel OS-ELM algorithm. *Aerosp. Sci. Technol.* **2019**, *84*, 661–671. [[CrossRef](#)]
12. Huang, C.G.; Huang, H.Z.; Li, Y.F. A Bi-Directional LSTM prognostics method under multiple operational conditions. *IEEE Trans. Ind. Electron.* **2019**, *66*, 8792–8802. [[CrossRef](#)]
13. Li, Z.X.; Goebel, K.; Wu, D.Z. Degradation modeling and remaining useful life prediction of aircraft engines using ensemble learning. *J. Eng. Gas Turbines Power* **2019**, *141*, 041008. [[CrossRef](#)]
14. Si, X.S.; Wang, W.; Hu, C.H.; Zhou, D.H. Remaining useful life estimation—a review on the statistical data driven approaches. *Eur. J. Oper. Res.* **2011**, *213*, 1–14. [[CrossRef](#)]
15. Lasheras, S.F.; Nieto, P.J.G.; De Cos Juez, F.J.; Bayón, R.M.; Suárez, V.M.G. A hybrid PCA-CART-MARS-based prognostic approach of the remaining useful life for aircraft engines. *Sensors* **2015**, *15*, 7062–7083. [[CrossRef](#)] [[PubMed](#)]
16. Zaidan, M.A.; Mills, A.R.; Harrison, R.F.; Fleming, P.J. Gas turbine engine prognostics using Bayesian hierarchical models: A variational approach. *Mech. Syst. Signal Process.* **2016**, *70–71*, 120–140. [[CrossRef](#)]
17. Liu, K.B.; Huang, S. Integration of data fusion methodology and degradation modeling process to improve prognostics. *IEEE Trans. Autom. Sci. Eng.* **2016**, *13*, 344–354. [[CrossRef](#)]
18. Song, C.Y.; Liu, K.B. Statistical degradation modeling and prognostics of multiple sensor signals via data fusion: A composite health index approach. *IJSE Trans.* **2018**, *50*, 853–867. [[CrossRef](#)]
19. Zhang, Z.X.; Si, X.S.; Hu, C.H.; Lei, Y.G. Degradation data analysis and remaining useful life estimation: A review on Wiener-process-based methods. *Eur. J. Oper. Res.* **2018**, *271*, 775–796. [[CrossRef](#)]
20. Le Son, K.; Fouladirad, M.; Barros, A. Remaining useful lifetime estimation and noisy gamma deterioration process. *Reliab. Eng. Syst. Saf.* **2016**, *149*, 76–87. [[CrossRef](#)]
21. Peng, W.W.; Li, Y.F.; Yang, Y.J.; Mi, J.; Huang, H.Z. Bayesian degradation analysis with inverse Gaussian process models under time-varying degradation rates. *IEEE Trans. Reliab.* **2017**, *66*, 84–96. [[CrossRef](#)]
22. Wang, P.F.; Youn, B.D.; Hu, C. A generic probabilistic framework for structural health prognostics and uncertainty management. *Mech. Syst. Signal Process.* **2012**, *28*, 622–637. [[CrossRef](#)]
23. Kim, M.; Song, C.; Liu, K. A generic health index approach for multisensor degradation modeling and sensor selection. *IEEE Trans. Autom. Sci. Eng.* **2019**, *16*, 1426–1437. [[CrossRef](#)]
24. Coble, J.; Hines, J.W. Applying the general path model to estimation of remaining useful life. *Int. J. Progn. Health Manag.* **2011**, *2*, 71–82.
25. Liu, L.S.; Wang, S.J.; Liu, D.T.; Peng, Y. Quantitative selection of sensor data based on improved permutation entropy for system remaining useful life prediction. *Microelectron. Reliab.* **2017**, *75*, 264–270. [[CrossRef](#)]
26. Chehade, A.; Shi, Z.Y. Sensor fusion via statistical hypothesis testing for prognosis and degradation analysis. *IEEE Trans. Autom. Sci. Eng.* **2019**, *16*, 1774–1787. [[CrossRef](#)]
27. Fang, X.L.; Paynabar, K.; Gebraeel, N. Multistream sensor fusion-based prognostics model for systems with single failure modes. *Reliab. Eng. Syst. Saf.* **2017**, *159*, 322–331. [[CrossRef](#)]

28. Ye, Z.S.; Xie, M. Stochastic modelling and analysis of degradation for highly reliable products. *Appl. Stoch. Models Bus. Ind.* **2015**, *31*, 16–32. [[CrossRef](#)]
29. Zhang, B.; Zhang, L.J.; Xu, J.W. Degradation feature selection for remaining useful life prediction of rolling element bearings. *Qual. Reliab. Eng. Int.* **2016**, *32*, 547–554. [[CrossRef](#)]
30. Shang, H.L. A survey of functional principal component analysis. *AStA Adv. Stat. Anal.* **2014**, *98*, 121–142. [[CrossRef](#)]
31. Wang, J.L.; Chiou, J.; Müller, H.G. Functional data analysis. *Annu. Rev. Stat. Its Appl.* **2016**, *3*, 257–295. [[CrossRef](#)]
32. Chung, S.; Kontar, R. Functional principal component analysis for extrapolating multi-stream longitudinal data. *arXiv Preprint* **2019**, arXiv:1903.03871.
33. Ghahramani, Z. Bayesian non-parametrics and the probabilistic approach to modelling. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2013**, *371*, 20110553. [[CrossRef](#)] [[PubMed](#)]
34. Fang, X.L.; Zhou, R.R.; Gebraeel, N. An adaptive functional regression-based prognostic model for applications with missing data. *Reliab. Eng. Syst. Saf.* **2015**, *133*, 266–274. [[CrossRef](#)]
35. Yao, F.; Müller, H.G.; Wang, J.L. Functional data analysis for sparse longitudinal data. *J. Am. Stat. Assoc.* **2005**, *100*, 577–590. [[CrossRef](#)]
36. Zhou, R.R.; Serban, N.; Gebraeel, N. Degradation applied to residual lifetime prediction using functional data analysis. *Ann. Appl. Stat.* **2011**, *5*, 1586–1610. [[CrossRef](#)]
37. Saxena, A.; Goebel, K.; Simon, D.; Eklund, N. Damage Propagation for Aircraft Engine Run-To-Failure Simulation. In Proceedings of the International Conference on Prognostics and Health Management, Denver, CO, USA, 6–9 October 2008.
38. Saxena, A.; Goebel, K. Turbo Fan Engine Degradation Simulation Dataset. In *NASA Ames Prognostics Data Repository*; NASA Ames: Moffett Field, CA, USA, 2012. Available online: <http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository> (accessed on 15 August 2018).
39. Singh, S.K.; Kumar, S.; Dwivedi, J.P. A novel soft computing method for engine RUL prediction. *Multimed. Tools Appl.* **2019**, *78*, 4065–4087. [[CrossRef](#)]
40. Nieto, P.J.G.; García-Gonzalo, E.; Lasheras, S.F.; De Cos Juez, F.J. Hybrid PSO–SVM-based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability. *Reliab. Eng. Syst. Saf.* **2015**, *138*, 219–231. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).