Research Article

Multistep Degradation Tendency Prediction for Aircraft Engines Based on CEEMDAN Permutation Entropy and Improved Grey–Markov Model

Wei Jiang(1), Jianzhong Zhou(2), Yanhe Xu(2), Jie Liu(2,3), and Yahui Shan(2)

1Faculty of Mechanical and Material Engineering, Huaiyin Institute of Technology, Huai’an 223003, China
2School of Hydropower and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China
3Nondestructive Detection and Monitoring Technology for High Speed Transportation Facilities, Key Laboratory of Ministry of Industry and Information Technology, Nanjing 210016, China

Correspondence should be addressed to Wei Jiang; jw_hust@163.com

Received 16 May 2019; Revised 25 September 2019; Accepted 11 October 2019; Published 31 October 2019

Academic Editor: Yong Xu

Copyright © 2019 Wei Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As an essential component and core power source of aircraft, the operational stability of aeroengine has important impact on system safety and reliability. Accurate degradation tendency prediction on an engine can not only improve its operational stability but also significantly reduce the maintenance costs. In this paper, a novel forecasting method that combines CEEMDAN permutation entropy and improved Grey–Markov model is proposed to perform multistep degradation tendency prediction of aircraft engines. In order to accurately quantify the degradation level of engines, a new integrated degradation index (IDI) is innovatively designed by multidimensional sensory data. And then, because of high speed and excellent performance, CEEMDAN algorithm is specifically employed to decompose the generated IDI series to eliminate the potential influence of stochastic fluctuations. Aiming at the complexity of intrinsic mode functions (IMFs) generated by CEEMDAN, an IMFs reconstruction strategy based on permutation entropy is developed to better characterize the degradation states. Finally, on the basis of above achievements and for higher forecasting efficiency and accuracy, an improved Grey–Markov model combined with the moving window algorithm, which is unique, is constructed to realize multistep degradation trend prediction of engines. The proposed method is applied to the degradation tendency prediction of aircraft engines. The experimental results validate the effectiveness and superiority of the proposed method, and it is more suitable for engineering applications in comparison with other methods.

1. Introduction

With the gradual improvement of mechanical system’s integration and complexity, reasonable and complete health state monitoring is of great significance for ensuring stable and reliable operation of equipment [1–3]. As an essential component and core power source of aircraft, the operational stability of aeroengine has important impact on system safety and maintenance costs [4–6]. In recent years, with the proposal of condition-based maintenance (CBM), the fault response mode has converted from passive treatment to active prevention [7, 8]. CBM contributes to identify the operational status of equipment and avoid unnecessary downtime maintenance. Because of these advantages, it gradually becomes one of the most commonly used maintenance pattern and attracts more and more focuses of researchers [9, 10]. More specifically, degradation tendency prediction plays an important role in the implementation of CBM, which is helpful to discover abnormal operation states before fault occurs and effectively decrease the failure rate and maintenance costs [11, 12].

Generally, the implementation of degradation tendency prediction can be mainly divided into two stages, i.e., degradation indicator construction and development trend forecasting. Throughout the entire process of prediction, an appropriate indicator needs to be constructed to quantify the degradation levels of an aircraft engine, and it can be regarded as the foundation of the subsequent degradation
trend forecasting. With the accumulation of running time, large amounts of sensory data from different positions are collected for analysis [13, 14]. On this basis, how to build a degradation indicator by adequately utilizing these data is the primary problem to be solved in the task of tendency prediction. For this reason, some researches focusing on constructing a suitable index that can effectively indicate the extent of deterioration have been carried out. For instance, Kral first pointed out that the status of used oil can be used to analyze the operation conditions of the vehicle cooling system [15]. Volponi utilized the fuel flow rate to evaluate the health state of gas turbine [16]. Gebrael analyzed the states of rolling element bearing by the collected vibration signals [17]. However, these indicators are mainly based on single sensory signal, which means that some important signals containing abundant degradation information would be ignored. Besides, because of the growing diversity of collected signals, it is difficult to select a representative signal from large amounts of sensor signals that can accurately reflect the degradation state of equipment. Thus, the construction of an excellent degradation index fully utilizing different sensory data is still a difficult point in the research of degradation tendency prediction.

Degradation tendency prediction aims at obtaining the evolution of degradation in the future and supplying adequate data basis for decision-making. In general, the current models for trend prediction can be divided into three types, including knowledge-based models, physics-based models, and data-driven models. In particular, the practical application of former two models would face more restrictions due to the difficulties in obtaining relevant knowledge and establishing a suitable physical model. Data-driven models, which fully utilize the acquired monitoring data, can effectively achieve the purpose of degradation trend prediction without the help of domain knowledge and physical rules [18, 19]. Ma used a stacked sparse autoencoder with multilayer self-learning to forecast the remaining useful life of engine unit [4]. Fu adopted an improved least squares support vector machine model to predict the state development trend of hydroelectric generating unit [20]. Grey–Markov model, one of data-driven models, is widely applied in the prediction problem of systems with uncertain structure or characteristics due to its simple principle and excellent performance. Zhou presented the application of a Grey–Markov model with incidence analysis in the degradation trend forecasting for energy conversion equipment [21]. The results confirmed that the proposed method realized the satisfied performance. In Reference [22], an improved Grey–Markov model based on wavelet transform was developed to achieve accurate prediction of China’s energy supply and demand. Until now, the advantages of Grey–Markov over other prediction models, such as convenient parameters training, low computing time, and high forecasting accuracy, have been validated by massive experimental researches [23, 24]. Based on this, the Grey–Markov model is conducive to obtain optimal solutions. However, there are still two inherent drawbacks in this prediction model. Firstly, the Grey–Markov model is established based on the complete training samples, which would lead to lower accuracy with the increasing of prediction time and is unreasonable when there are exponential and chaotic data in training sets [25]. In addition, the single-step prediction pattern of Grey–Markov may cause the decrease of computation efficiency. Secondly, due to stochastic volatility and inherent complexity of original sensory data, it is difficult that relying on just the single Grey–Markov model to accurately forecast the degradation trend. In order to obtain better predicted results, it is necessary to first analyze the characteristics of raw data. Thus, the multiscale decomposition algorithm is introduced into Grey–Markov to improve the prediction performance of the single model. Various decomposition methods, such as wavelet transform (WT) [26], empirical mode decomposition (EMD) [27], and ensemble empirical mode decomposition (EEMD) [28], are adopted to decompose the original data series to reduce the influence of irregular volatility on forecasting results. Compared with the approaches above, a new type of decomposition method, named “complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN),” has attracted a huge amount of attention due to its excellent performance and high efficiency [29–31]. Qu developed a wind speed forecasting method based on CEEMDAN and an improved backpropagation neural network (BPNN), and the experiment results indicated that CEEMDAN could efficiently solve the problem of data fluctuations [30]. Therefore, it is a valuable subject that a novel prediction method combining the merits of CEEMDAN and the Grey–Markov model should be constructed to further enhance the forecasting performance.

The main contributions of this work is the development of a multistep degradation tendency prediction method for aircraft engines based on CEEMDAN permutation entropy (PE) and the improved Grey–Markov model with moving window (IGMMW). And on the subject of detail, a new integrated degradation index (IDI) is constructed by multidimensional sensory data for the accurate quantification of engine degradation levels. Meanwhile, CEEMDAN algorithm is first used to decompose the IDI series to eliminate the effect of data fluctuations. Then, an intrinsic mode functions (IMFs) reconstruction strategy based on PE theory is innovatively designed to reduce the complexity of decomposed components. Finally, for the sake of higher forecasting efficiency and accuracy, a novel prediction model, namely, IGMMW, is developed to forecast the degradation trend of engine units. The general implementation of the proposed method can be divided into four steps, i.e., IDI series construction, series decomposition, IMFs reconstruction, and tendency prediction. Firstly, an IDI for the measurement of engine deterioration levels is built by using different sensor data and an appropriate data fusion method. Compared with the single signal adopted in [15–17], the generated IDI fully retains valuable degradation information contained in various signals and achieves the mapping from high-dimensional signal space to one-dimensional index space. Secondly, the CEEMDAN is utilized to adaptively decompose the generated IDI series to further eliminate the potential influence of stochastic fluctuations,
which is a remarkable improvement in comparison with other algorithms adopted in [26–28] due to its excellent decomposition performance. Subsequently, a PE-based reconstruction strategy is designed to achieve the reduction of IMFs’ complexity, i.e., several IMFs to fewer reconstructed IMFs (RIMFs). With the idea of aggregation based on PE values, there will be fewer decomposition components and the forecasting accuracy and efficiency will also be improved. Finally, based on the obtained RIMFs, an IGMMW prediction model is developed to efficiently forecast the future degradation trend of engines. Because of the combination of moving window method, the problem of circular update for sequences being modeled, which occurs in [21–24], can be solved well. Besides, the adaptive parameter in moving window, i.e., the step size, is helpful to the implementation of multistep prediction to further improve the computational efficiency. The proposed method is used for the degradation trend prediction of aircraft engines, in which the sensory signals are measured from different parts of engine units. The experimental results confirm the effectiveness and superiority of the proposed method, and it is more suitable for engineering applications in comparison with other methods.

The rest of this paper is organized as follows. In Section 2, the essential background knowledge about CEEMDAN, Grey theory, and Markov chain modeling mechanism is reviewed. The proposed method is introduced in Section 3. In Section 4, the proposed method is used to predict the degradation trend of aircraft engines and the experimental results are analyzed and discussed in detail. Finally, general conclusions are given in Section 5.

2. Preliminaries

2.1. Complete Ensemble Empirical Mode Decomposition with Adaptive Noise. In order to deal with the analysis of nonstationary signal, an empirical mode decomposition (EMD) algorithm was proposed by Huang in 1998 [32]. The method is an adaptive signal time-frequency-domain analysis technique and decomposes the raw signals into a series of IMFs. Specifically, each IMF component reflects the different characteristics of raw signal at different time scales, which usually satisfies two conditions below: (a) in the entire time series, the number of extreme points is always equal to that of zero-crossings or the difference of number being not more than one; (b) the average value of envelope at any time point, defined by the local maxima and the local minimum, should be zero [30]. With the definition of IMFs mentioned above and relevant hypothesis, the original signal can be decomposed into several IMFs and one residue component by EMD:

$$s(t) = \sum_{i=1}^{n} IMF_i(t) + r_n(t),$$  (1)

where $s(t)$ denotes the raw signal, $IMF_i(t)$ is the $i$-th IMF component obtained by EMD method, $n$ is the number of IMFs, and $r_n(t)$ is the final residue function, which represents the mean tendency of data sequence.

Although EMD method has significant advantages in analyzing the nonstationary signal, there are some inherent limitations that make great influence on the performance of EMD, such as the mode mixing problem and the end-point effect [31]. In order to eliminate these problems in EMD, a noise-assisted signal analysis approach, named “ensemble empirical mode decomposition (EEMD),” was developed by Wu and Huang in 2009 [33]. However, the EEMD algorithm cannot eliminate the effect of Gaussian white noise on the reconstructed signal, and the high computational costs due to the added noise make a great restriction to the application of the decomposition method. Because of the above defects in EEMD, the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) algorithm was designed to improve the performance of EEMD [34]. The method effectively eliminates the mode mixing phenomenon in the IMFs, and the reconstruction error is always zero. Meanwhile, compared with EEMD, the computation time of CEEMDAN can be decreased remarkably. Suppose $E_j(t)$ is the operator of $j$-th IMF component that is decomposed by EMD, and let $w(t)$ be Gaussian white noise with zero mean and homogeneity of variance. The general procedures of CEEMDAN are given as follows:

1. Decompose each $s_i(t) = s(t) + \varepsilon_i w_i(t), i = 1, 2, \ldots, I$ using EMD to obtain the first IMF, in which $\varepsilon_i$ is a noise coefficient. Meanwhile, define the first mode component decomposed by CEEMDAN as

$$\overline{IMF}_1(t) = \frac{1}{I} \sum_{j=1}^{I} IMF_{1j}(t).$$  (2)

2. Calculate the first residue $r_1(t)$:

$$r_1(t) = s(t) - \overline{IMF}_1(t).$$  (3)

3. Decompose residue $r_1(t) + \varepsilon_1 E_1(w_1(t))$ to calculate the second mode denoted by $\overline{IMF}_2(t)$:

$$\overline{IMF}_2(t) = \frac{1}{I} \sum_{j=1}^{I} E_1(r_1(t) + \varepsilon_1 E_1(w_1(t))).$$  (4)

4. Repeat steps (2) and (3) until all of the IMFs are analyzed. The final residue can be defined as:

$$r_m(t) = s(t) - \sum_{j=1}^{m} \overline{IMF}_j(t),$$  (5)

where $m$ is the total number of IMFs decomposed by CEEMDAN. Therefore, based on the steps mentioned above, the original signal $s(t)$ can be expressed as
\[ s(t) = \sum_{j=1}^{m} I M F_j(t) + r_m(t). \]  

(6)

Based on equation (6), the signal can be decomposed into \( m \) IMFs and a residue, which provides an exact way of reconstruction for raw data. It is worth noting that the obtained IMFs reflect the characteristics of original signal at different timescales, and the residue is smoother and contributes to the reduction of prediction error.

2.2. Grey System Modeling and Prediction. Grey theory is usually used to deal with the prediction problems which have insufficient and uncertain information [35]. On the basis of Grey theory, the Grey prediction model has been developed and applied in different fields successfully [36, 37]. In general, the Grey model can be presented as GM(\( u, v \)), where \( u \) is the order of the differential equation and \( v \) is the number of variables. Because of superior computation efficiency, GM(1, 1) is the most widely used in practical applications [37]. The general steps of GM(1, 1) are illustrated as follows:

1. Suppose the raw series can be written as
   \[ x^{(0)} = (x^{(0)}(t_1), x^{(0)}(t_2), \ldots, x^{(0)}(t_n)), \]  
   where \( x^{(0)}(t_i) \) is the system output at time \( t_i \).
2. Based on the raw series \( x^{(0)} \) and the accumulated generating operator (AGO), a new series \( x^{(1)} \) can be generated to show the tendency of exponential growth [35]:
   \[ x^{(1)} = (x^{(1)}(t_1), x^{(1)}(t_2), \ldots, x^{(1)}(t_n)), \]  
   where
   \[ x^{(1)}(t_k) = \sum_{m=1}^{k} x^{(0)}(t_m), \quad k = 1, 2, \ldots, n. \]  
3. The whitening differential equation of GM(1, 1) can be built as follows [35]:
   \[ \frac{dx^{(1)}}{dt} + ax^{(1)} = b, \]  
   where \( a \) is the so-called development coefficient and \( b \) is known as the endogenous control coefficient, and they are two constants determined by the original series.

With the generated series \( x^{(1)} \) and the least squares method, the coefficients of equation (10) can be estimated as
   \[ [a, b]^T = (B^T B)^{-1} B^T Y, \]  
   where
   \[ B = \begin{bmatrix}
   -\frac{1}{2} \left[ x^{(1)}(t_1) + x^{(1)}(t_2) \right] & 1 \\
   -\frac{1}{2} \left[ x^{(1)}(t_2) + x^{(1)}(t_3) \right] & 1 \\
   \vdots & \vdots \\
   -\frac{1}{2} \left[ x^{(1)}(t_{n-1}) + x^{(1)}(t_n) \right] & 1
   \end{bmatrix}, \]
   \[ Y = \begin{bmatrix}
   x^{(0)}(t_2), x^{(0)}(t_3), \ldots, x^{(0)}(t_n)
   \end{bmatrix}^T. \]  

(12)

(4) Based on equation (10), the solution of \( x^{(1)}(t_k) \) can be obtained as
   \[ \tilde{x}^{(1)}(t_k) = \left( x^{(0)}(t_1) - \frac{b}{a} \right) e^{-at_{k-1}} + \frac{b}{a}, \]  
   \[ e^{-at_k} - e^{-at_{k-1}}. \]  
   \[ (14) \]

(5) Finally, with the subtraction operation, the predicted value of raw data point \( x^{(0)}(t_k) \) can be acquired as follows:
   \[ \tilde{x}^{(0)}(t_k) = \tilde{x}^{(1)}(t_k) - \tilde{x}^{(1)}(t_{k-1}) = \left( x^{(0)}(t_1) - \frac{b}{a} \right) e^{-at_k} - e^{-at_k} - e^{-at_{k-1}}. \]  

(15)

2.3. Markov Chain Modeling Mechanism. Without the consideration of random fluctuations in the original series, the forecasting accuracy of GM(1, 1) is low when the sequence being modeled fluctuates sharply [25]. In order to reveal the inherent laws of data fluctuation and improve the prediction precision, Markov chain, a particular stochastic process theory, is used to modify the prognostic residue series of GM(1, 1) and effectively improve the prediction accuracy [38, 39]. The detailed modeling mechanism of Markov chain is described below.

For the obtained series \( \tilde{x}^{(0)}(t_1), \tilde{x}^{(0)}(t_2), \ldots, \tilde{x}^{(0)}(t_n) \) by GM(1, 1), the relative error between the predicted value and the original value can be calculated:
   \[ \omega(t_k) = \frac{|x^{(0)}(t_k) - \tilde{x}^{(0)}(t_k)|}{x^{(0)}(t_k)} \times 100\%, \quad k = 1, 2, \ldots, n. \]  

(16)

Then, the range of the relative error \([\min \omega(t_k), \max \omega(t_k)]\) can be split into \( S \) intervals with equal length, which are called as \( S \) states. Each state is a section of the range \([\min \omega(t_k), \max \omega(t_k)]\), i.e.,
   \[ \omega(t_k) \in [R_{i,j}, U_{i,j}], \quad j = 1, 2, \ldots, S, \]  

(17)

where
\[
R_{t_k} = \min_{t_k} \omega(t_k) + \frac{j - 1}{S} \left( \max_{t_k} \omega(t_k) - \min_{t_k} \omega(t_k) \right).
\]

(18)

\[
U_{t_k} = \min_{t_k} \omega(t_k) + \frac{j}{S} \left( \max_{t_k} \omega(t_k) - \min_{t_k} \omega(t_k) \right).
\]

(19)

In the theory of Markov chain modeling, the transition probability from state \(i\) to \(j\) by \(m\) steps can be deduced as

\[
p_{ij}^{(m)} = \frac{M_{ij}(m)}{M_i}, \quad i, j = 1, 2, \ldots, S,
\]

(20)

where \(M_{ij}(m)\) is the transition times that occurred from state \(i\) to state \(j\) by \(m\) steps and \(M_i\) is the number of data whose relative errors belong to state \(i\).

With these \(m\)-step transition probabilities, the state transition probability matrix \(P(m)\) can be constructed to effectively restrain the effects of random fluctuations:

\[
P(m) = \begin{bmatrix}
p_{11}^{(m)} & p_{12}^{(m)} & \cdots & p_{1S}^{(m)} \\
p_{21}^{(m)} & p_{22}^{(m)} & \cdots & p_{2S}^{(m)} \\
\vdots & \vdots & \ddots & \vdots \\
p_{S1}^{(m)} & p_{S2}^{(m)} & \cdots & p_{SS}^{(m)}
\end{bmatrix}.
\]

(21)

Subject to

\[
p_{ij}^{(m)} \geq 0, \quad \sum_{j=1}^{S} p_{ij}^{(m)} = 1.
\]

(22)

The probability matrix \(P(m)\) reveals the transition laws between different states, which is the modeling foundation of Markov chain. Through the more detailed procedures mentioned in [39], the future state transition on the basis of the current state, i.e., state \(l\) denoted by \(S_l\), can be estimated. The lower boundary and upper boundary of this state are denoted as \(R_{l1}\) and \(U_{l1}\), respectively. Finally, the corresponding predicted value by GM(1, 1), denoted by \(\hat{x}^{(0)}(t_k)\), can be modified by Markov chain according to the following formula:

\[
\hat{h}^{(0)}(t_k) = \hat{x}^{(0)}(t_k) \left[ 1 + 0.5(R_{l1} + U_{l1}) \right].
\]

(23)

### 3. The Proposed Degradation Tendency Prediction Method

In this paper, a novel multistep method based on CEEMDAN permutation entropy and IGMMW model, systematically blending the signal decomposition technique and intelligent prediction technology, is proposed for the degradation tendency prediction of aircraft engines. This section includes four parts: the construction of integrated degradation index, the reconstruction of IMFs using PE theory, improved Grey–Markov model with moving window, and the general procedure of the proposed method.

#### 3.1. Construction of Integrated Degradation Index

For the sake of quantifying the degradation degree of engines effectively, a proper degradation index should be constructed. With the help of a large number of available sensory data and the idea of linear transformation, a new integrated degradation index (IDI) is innovatively proposed in this paper, which achieves the mapping from multidimensional data domain to one-dimensional index domain.

Suppose \(W_1\) of \(U_1 \times V\) matrix and \(W_2\) of \(U_2 \times V\) matrix are two groups of multidimensional sensor dataset, which represent the faulty and healthy states of engines, respectively. \(U_1\) and \(U_2\) are the sizes of datasets under faulty and healthy conditions, and \(V\) is the dimension of sensory dataset. With \(W_1\) and \(W_2\), a \(V \times 1\) matrix \(T\) can be designed to build the relationship of mapping between multidimensional sensory data and one-dimensional IDI as

\[
T = (W^TW)^{-1}W^TG,
\]

(24)

where \(W = [W_1^TW_2]^T\), \(G = [G_1^T, G_2]^T\), \(G_1\) represents a \(1 \times U_1\) zero vector, and \(G_2\) represents a \(1 \times U_2\) unit vector. With the constructed matrix \(T\) and the historical dataset \(Q\) collected from the sensor, the IDI denoted by \(d\) can be obtained:

\[
d = 1 - T \cdot Q.
\]

(25)

Note that the value of IDI changes between 0 and 1, and “0” represents healthy state and “1” represents faulty state. The calculation of IDI can be essentially regarded as a process of multidimensional data fusion, and it provides an effective way to accurately describe the degradation levels of aircraft engines.

#### 3.2. Reconstruction of IMFs Using PE Theory

The sensory signals collected from the online monitoring system are susceptible to operation environment and background noise. For this reason, the corresponding decomposition results with CEEMDAN may consist of many IMFs, which will enhance the complexity of model training and reduce the forecasting accuracy. Thus, on the premise of retaining all effective components, an IMFs reconstruction strategy using PE theory is first developed for the decrease of the number of IMFs. The reconstruction process is described in detail below.

For the IMF components after decomposition \((c_1(t), c_2(t), \ldots, c_i(t), \ldots, c_m(t))\) \((c_i(t)\) represents the \(i\)-th IMF component and \(m\) is the number of IMFs), the corresponding phase space reconstruction vector of \(c_i(t)\) can be expressed based on the the Takens–Maine theorem as

\[
C_i(t) = (c_i(t), c_i(t + \tau), \ldots, c_i(t + (n - 1)\tau)),
\]

(26)

where \(\tau\) is the time delay and \(n\) is the embedded dimension. On this basis, the elements of \(C_i(t)\) are rearranged by \(n\) number of real values in ascending order, which meets

\[
c_i(t + (j_1 - 1)\tau) \leq c_i(t + (j_2 - 1)\tau) \leq \cdots \leq c_i(t + (j_n - 1)\tau).
\]

(27)
According to the above equation, map $C_i(t)$ into a group of symbols:

$$S_p(ξ) = (j_1, j_2, \ldots, j_n), \quad (28)$$

where $ξ = 1, 2, \ldots, k$ and $k \leq n!$, $S_p(ξ)$ is one of the $n!$ arrangements. Calculate the probabilities of these symbols denoted as $p_1, p_2, \ldots, p_k$ ($\sum_{ξ=1}^{k} p_ξ = 1$), and then, the PE value of IMF component $c_i(t)$ can be acquired according to the following formula:

$$H_p(n) = \frac{1}{\ln(n!)} \sum_{ξ=1}^{k} p_ξ \ln p_ξ, \quad (29)$$

where $1/\ln(n!)$ is the regularization coefficient and $H_p(n)$ is bounded in $[0, 1]$. In essence, PE has significant advantages in measuring the randomness of series. Thus, with the idea of similarity-based combination, the IMFs after reconstruction, expressed as RIMFs, can be generated according to the designed criterion as follows:

$$\text{RIMF} = \sum_{i=1}^{i+j} c_i(t), \quad \text{s.t.} \quad |H_{pi} - H_{p(i+j)}| \leq \frac{2(H_{p\text{max}} - H_{p\text{min}})}{m}. \quad (30)$$

With the proposed IMFs reconstruction method, the original IMFs decomposed by CEEMDAN algorithm can be classified into several groups adaptively, and the PE values of IMFs in each group are limited to one specific interval. For this reason, the constructed RIMFs not only preserve all of the components’ information but also obviously reduce the number of IMFs. On this foundation, the RIMFs can be used as the inputs of the prediction model to effectively improve the forecasting efficiency and accuracy.

3.3. Improved Grey–Markov Model with Moving Window.

In order to eliminate the influence of exponential and chaotic data on modeling and enhance the prediction performance of Grey–Markov model, an improved Grey–Markov model with moving window (IGGMW) is innovatively proposed to perform the multistep degradation tendency prediction of aircraft engines. With the moving window algorithm, the prediction model can be circularly reconstructed based on adjacent datasets, and thus, the forecasting accuracy can be further improved. The grey modeling mechanism based on moving window is shown in Figure 1, where $n$ is the length of original modeling series, $N$ is the number of data points to be predicted, $L$ is the step size of moving window, and $a$ and $b$ represent the coefficients depicted in equation (10).

Different from the traditional Grey–Markov model, the IGMWM adopts a novel structure that the modification of prediction results by Markov chain is embedded into each grey modeling process based on moving window. The flowchart of the trend forecasting based on the IGMWM is depicted in Figure 2, and the detailed steps are given as follows:

Step 1: initialize the modeling times $g = 1$.
Step 2: with the Grey modeling mechanism based on moving window, GM(1, 1) can be established to predict the IDI series segments. According to equations (11) to (13), the coefficients, i.e., $a_g$ and $b_g$, can be obtained.
Step 3: based on the acquired coefficients, the predicted values of IDI series segments, denoted by $(\tilde{d}^g(n + (g - 1)L + 1), \tilde{d}^g(n + (g - 1)L + 2), \ldots, \tilde{d}^g(n + gL))$, are calculated by equation (15). It is worth noting that the step size $L$ would make significant influence on the forecasting results. For this reason, the parameter $L$ should be estimated effectively by comparing the forecasting performance of model as we increase the capacity in step size.
Step 4: the predicted values in the previous step need to be modified by the Markov chain modeling mechanism, and the future values of IDI series segments after modification, expressed as $(\tilde{d}^g(n + (g - 1)L + 1), \tilde{d}^g(n + (g - 1)L + 2), \ldots, \tilde{d}^g(n + gL))$, can be acquired.
Step 5: replacing original predicted values with the modified values, the new series can be generated for the next modeling.
Step 6: if $g < N/L$, then repeat steps (2) to (5) with $g = g + 1$. Otherwise, the forecasting process has been completed. Finally, the values of complete IDI series to be predicted with IGMMW, i.e., $(\tilde{d}(n + 1), \tilde{d}(n + 2), \ldots, \tilde{d}(n + N))$, can be obtained as the final prediction results.

It can also be analyzed from Figure 2 that the forecasting precision is further improved because of the dynamic update of modeling series with modified values. In addition, the adaptive model parameter, i.e., the step size $L$, contributes to the implementation of multistep prediction to improve the computation efficiency. Therefore, the proposed IGMMW has great performances on forecasting accuracy and efficiency and can be used for the degradation tendency prediction of aircraft engines.

3.4. Procedure of the Proposed Method. In this paper, a multistep degradation tendency prediction method for aircraft engines is developed based on CEEMDAN-PE theory and the IGMMW model. The flowchart of the proposed method is shown in Figure 3, and the general procedures are summarized as follows:

Step 1: the different types of sensory data are collected by the online monitoring system
Step 2: the IDI series is constructed with the acquired sensory data to evaluate the degradation levels of engines
Step 3: the constructed IDI series is divided into two parts, including training set and testing set
Step 4: for the training set, the IDI series can be decomposed into several IMFs using CEEMDAN algorithm
Step 5: reconstruct the IMFs with the proposed PE-based reconstruction strategy and obtain the RIMFs
Step 6: for each RIMF, an optimal IGMMW can be built to predict the future IDI series.

Step 7: accumulate the results of all IGMMW models and the final predicted IDI series can be obtained.

Step 8: the testing set is utilized to evaluate the accuracy of final prediction results.

4. Experiments and Results Analysis

4.1. Experimental Setup and Model Performance Evaluation.

In this section, the “Prognostic Data Challenge Problem 2008” datasets are selected to demonstrate the effectiveness of the proposed degradation tendency prediction method, which consist of multivariate sensor signals collected from different parts of aircraft engines [40]. The schematic illustration of an engine model, for clarity, is presented in Figure 4, which mainly contains two turbines, two compressors, a fan, a combustor, and a nozzle [40]. Specifically, the sensory dataset for each cycle includes unit identifier, cycle index, 3 operation setting parameters, and 21 kinds of monitoring data [40]. According to the setting parameters, the operation states of engines can be roughly divided into 6 categories [41], as depicted in Table 1. Due to the diversity of raw data, it is unreasonable that the monitoring data are
directly used to measure the degradation status. The research task of this paper is to design a novel indicator to accurately describe the degradation levels of engine and effectively predict the development trend of engine degradation.

In order to effectively assess the performance of the prediction model, three generally adopted error criteria are selected to measure the prediction accuracy, including mean absolute error (MAE), mean absolute percentage error (MAPE), and coefficient of determination $R^2$ [18, 20]. MAE is a measurement of proximity between the real value and predicted value, and MAPE can reflect the average forecasting ability of the model. Within the interval $[0, 1]$, $R^2$ can

$$MAE = \frac{1}{k} \sum_{j=1}^{k} |y_j - \hat{y}_j|$$

$$MAPE = \frac{1}{k} \sum_{j=1}^{k} \left( \frac{|y_j - \hat{y}_j|}{y_j} \right) \times 100\%$$

$$R^2 = 1 - \frac{\sum_{i=1}^{k} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{k} (y_i - \bar{y})^2}$$
depict the evolution trend of performance degradation to construct the IDI series. Through the process of signals selection described in Reference [41], 7 monitoring signals listed in Table 2 can be finally determined in this study.

With these selected signals, the IDI series can be generated to describe the degradation levels of engine. Based on equation (21), transformation matrices $T_i$ ($i = 1, 2, \ldots, 6$) are established for six different operation states. For this, $W_1$ and $W_2$ need to be constructed in advance under these states. In this study, $W_1$ is built with the collected data under fault conditions, where the residual cycle life (the difference between the operational cycle and the whole cycle of an engine unit) is within $[-3, 0]$. Similarly, $W_2$ is created under healthy conditions, where the residual cycle life is smaller than $-200$. According to these transformation matrices and historical datasets, a one-dimensional IDI series can be constructed and presented in Figure 5. More specifically, from this figure, we can find that the IDI series shows a gradual health deterioration process with the accumulation of operation cycle. In conclusion, the constructed IDI can be served as an effective measurement for degradation status of aircraft engines. Besides, for the following study, the first 150 points of IDI series ($d(1), d(2), \ldots, d(150)$) are used as the training set and the remaining ($d(151), d(152), \ldots, d(210)$) as the testing set.

### 4.3. IDI Series Decomposition and IMFs Reconstruction

Figure 5 shows that the generated IDI series fluctuates violently. In order to eliminate the influence of irregular fluctuations of original series, CEEMDAN is employed to decompose the IDI sequence for the decrease of the non-stationary characteristic. The results are presented in Figure 6, in which IDI series is decomposed into 8 independent IMFs and one residue.

To highlight the superiority of CEEMDAN, EMD is exploited to decompose the IDI series for comparison. The decomposed results based on EMD are shown in Figure 7, in which IDI series is decomposed into 10 IMFs and one residue. Compared with the results of CEEMDAN, there are obvious phenomena of mode mixing in the obtained IMFs by EMD and the layers of EMD are more than that of CEEMDAN, which tend to cause the reduction of forecasting accuracy and computation speed to some extent.

Due to the significant influence of IMFs complexity to model training and prediction accuracy, an IMFs reconstruction strategy based on PE theory is investigated to obtain the simplified RIMF components. The PE distributions of the CEEMDAN decomposed results are shown in Figure 8, and the corresponding PE values of these 8 IMFs and one residue are listed in Table 3. We can observe from the figure that the PE values of decomposed results show a decreasing trend as the decomposition conducted. More specifically, the PE value is progressively decreased from 0.895 of IMF1 to 0 of $r_0$, which indicates the complexities of decomposed results are gradually reduced. With the IMFs reconstruction method described in Section 3.2, the interval length for the classification of IMFs, namely, $2(H_{p_{\text{max}}} - H_{p_{\text{min}}})/m$, can be set to 0.199. The
Table 2: Description of the selected monitoring signals for the construction of IDI series.

<table>
<thead>
<tr>
<th>Index</th>
<th>Monitoring signal</th>
<th>Abbreviation</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Total temperature at low-pressure compressor outlet</td>
<td>T24</td>
<td>°R</td>
</tr>
<tr>
<td>2</td>
<td>Total temperature at high-pressure compressor outlet</td>
<td>T30</td>
<td>°R</td>
</tr>
<tr>
<td>3</td>
<td>Total temperature at low-pressure turbine outlet</td>
<td>T50</td>
<td>°R</td>
</tr>
<tr>
<td>4</td>
<td>Total pressure at high-pressure compressor outlet</td>
<td>P30</td>
<td>psia</td>
</tr>
<tr>
<td>5</td>
<td>Static pressure at high-pressure compressor outlet</td>
<td>Ps30</td>
<td>psia</td>
</tr>
<tr>
<td>6</td>
<td>Ratio of fuel flow to Ps30</td>
<td>Phi</td>
<td>pps/psia</td>
</tr>
<tr>
<td>7</td>
<td>Bypass ratio</td>
<td>BPR</td>
<td>—</td>
</tr>
</tbody>
</table>

Figure 5: Generated IDI series.

Figure 6: CEEMDAN decomposed results of the generated IDI series.
reconstruction of IMFs is analyzed based on equation (28), and the corresponding results are illustrated in Table 4. As listed in the table, IMF2 and IMF3 are selected to generate the RIMF2 based on the specific interval of PE, i.e., [0.488, 0.687]. The similar processes are conducted for the construction of other RIMFs, including RIMF1, RIMF3, and RIMF4. The obtained RIMFs based on the CEEMDAN decomposed IMFs are presented in Figure 9. It can be found from the figure that four RIMFs have very different time characteristics. In particular, RIMF1 has the highest frequency and nonstationarity while RIMF4 is stable in the whole of lifecycle. Besides, it is only four RIMFs that make the number of IMFs reduced significantly. Therefore, the generated RIMFs can be served as the input of the developed IGMMW model to forecast the degradation tendency of aircraft engines to further improve the prediction efficiency and accuracy.

4.4. Degradation Tendency Prediction of Aircraft Engines. Based on the generated RIMFs and the proposed IGMMW prediction model shown in Figure 2, the degradation trend of aircraft engines can be predicted by the steps given in Figure 3. It is worth noting that there are two important parameters in the process of the IGMMW model construction, including the step size $L$ of moving window and
the number of states \( S \) in Markov chain, which would make significant influence on the final prediction results. Thus, these two parameters should be estimated effectively by the training set to build the optimal IGMMW model. With the idea mentioned in Section 3.3, we investigate how the models behave as we increase the capacity both in step size and number of states. Figure 10 shows the evolution of MAE, MAPE, and \( R^2 \) as the increase of \( L \) (from 1 to 6) and \( S \) (from 3 to 9). From the figure, it can be found that suitable values of \( L \) and \( S \) are helpful to obtain more accurate prediction results, whereas too big or too small values would cause the low precision and weak fitting. Therefore, the parameters are determined to construct the optimal IGMMW model as \( L = 3 \) and \( S = 5 \).

According to the constructed optimal IGMMW model, the forecasting values of all RIMFs reconstructed from the CEEMDAN decomposition results can be obtained. Figure 11 presents the comparison results of actual IDI and forecasting values of IGMMW model for these four RIMFs. From the figure, it can be found that there are obvious deviations between the IGMMW forecasting values of RIMF1 and the corresponding real values due to the strong nonstationarity, whereas the IGMMW model achieves the better prediction performances for the RIMF3 and RIMF4 with gentle change. Based on the obtained forecasting results of RIMFs, the final prediction results of future degradation tendency for aircraft engines can be calculated by the accumulation of all forecasting values. The comparison results of actual IDI series and final forecasting values of the proposed CEEMDAN-PE-IGMMW method are shown in Figure 12. It can be seen that the predicted results have excellent performance for fitting the original IDI sequence.

Besides, seven other models, including EMD-PE-IGMMW, CEEMDAN-IGMMW, extreme learning machine (ELM), support vector regression (SVR), autoregressive integrated moving average model (ARIMA), and backpropagation neural network (BPNN), are used for degradation trend prediction to validate the superiority of the proposed method. The forecasting results of these eight models are evaluated based on three error criteria illustrated in Section 4.1. As Figure 13 shows, for different purposes, the

### Table 3: PE values of the CEEMDAN decomposed IMFs.

<table>
<thead>
<tr>
<th>IMF1</th>
<th>IMF2</th>
<th>IMF3</th>
<th>IMF4</th>
<th>IMF5</th>
<th>IMF6</th>
<th>IMF7</th>
<th>IMF8</th>
<th>( r_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.895</td>
<td>0.687</td>
<td>0.489</td>
<td>0.388</td>
<td>0.262</td>
<td>0.188</td>
<td>0.167</td>
<td>0.166</td>
<td>0.000</td>
</tr>
</tbody>
</table>

### Table 4: Analysis of the reconstruction of IMFs.

<table>
<thead>
<tr>
<th>RIMFs</th>
<th>IMFs contained</th>
<th>Interval of PE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIMF1</td>
<td>IMF1</td>
<td>[0.696, 0.895]</td>
</tr>
<tr>
<td>RIMF2</td>
<td>IMF2, IMF3</td>
<td>[0.488, 0.687]</td>
</tr>
<tr>
<td>RIMF3</td>
<td>IMF4, IMF5</td>
<td>[0.189, 0.388]</td>
</tr>
<tr>
<td>RIMF4</td>
<td>IMF6, IMF7, IMF8, ( r_0 )</td>
<td>[0, 0.188]</td>
</tr>
</tbody>
</table>

![Figure 9: RIMFs reconstructed from the CEEMDAN decomposed IMFs.](image)
model comparisons can be classified into three parts. Specifically, the first is designed to demonstrate the superiority of CEEMDAN, the second to illustrate the effectiveness of the IMFs reconstruction method based on PE theory, and the last to verify the superior performance of IGMMW model in degradation tendency prediction. Moreover, the
comparisons between these three parts can further prove the feasibilities of series decomposition and IMFs reconstruction for improving the forecasting accuracy.

Figure 14 depicts the detailed degradation tendency prediction results of eight models. Furthermore, the forecasting accuracies of the eight models are presented in Figure 15, and the detailed comparisons of the error criteria (MAE, MAPE, and $R^2$) are listed in Table 5. Compared with the predicted results shown in Figure 14, the forecasting values of the proposed method are closer to the actual IDI values than that of other methods. More specifically, with the analysis of error criteria given in Figure 15 and Table 5, it can be found that the MAE and MAPE of the proposed method are 0.047 and 6.691%, which are significantly smaller than that of other approaches. Meanwhile, the $R^2$ of the proposed method is 0.982, compared with other seven methods, which are 0.978, 0.972, 0.831, 0.723, 0.499, 0.503, and 0.474, respectively.

It can be concluded from the figure that (1) Compared with other models, the proposed CEEMDAN-PE-IGMMW method achieves the best fitting performance between the forecasting results and the original IDI series, while ARIMA shows the worst. (2) The series decomposition using CEEMDAN contributes to effectively eliminate the phenomenon of mode mixing existing in EMD to further improve the forecasting accuracy. (3) With the IMFs reconstruction strategy based on PE theory, the prediction precision can be significantly improved in the whole forecasting period. (4) The hybrid prediction models, including CEEMDAN-PE-IGMMW, EMD-PE-IGMMW, and CEEMDAN-IGMMW, have better forecasting results than other five single models, which illustrates that series decomposition is helpful to enhance the prediction performance of single models. (5) The forecasting ability of IGMMW for most of the jumping IDI points is better than that of other single models, which indicates the sensitivity of
Figure 14: Comparisons of the forecasting results of eight models in the degradation tendency prediction.
the constructed model for data fluctuations. However, it is worth noting that the proposed method falls into a dilemma the same as other approaches, i.e., the lower accuracy at some cycle indexes with drastic volatilities of IDI data, such as 163rd, 180th, and 207th cycle index. The main reason is that the results modification using Markov chain is conducted based on historical data to be modeled so that the ability to correct is limited. This paper concentrates on developing a novel method that can achieve the more accurate prediction for degradation tendency in the whole time scales compared with other existing methods. As a result, there is less attention on the local results of the method. Despite this, it can be observed from Figure 14 that the proposed method still presents lower error at these cycle indexes compared with other prediction models.

Finally, the computing time of eight prediction models in degradation tendency forecasting is given in Table 6. Compared with the computing time listed in this table, the time of the IGMMW model is 2.01 s, which is less than that of other models. The reason is that the parameters of IGMMW model can be set in advance and never need to be adjusted during the prediction process. Besides, the computing time of the developed CEEMDAN-PE-IGMMW model is 2.33 s, which is slightly more than that of IGMMW and obviously less than that of other six models, including two combination models and four single models. The phenomena indicate that the forecasting method combining with CEEMDAN algorithm, PE theory, and IGMMW model has higher calculation efficiency. Meanwhile, it is worth noting that MAE, MAPE, and $R^2$ of the CEEMDAN-PE-IGMMW model are 0.047, 6.691%, and 0.982, respectively, which are significantly better than that of IGMMW model. This paper aims to develop a novel prediction method that can effectively increase the forecasting accuracy. To summarize, the proposed CEEMDAN-PE-IGMMW model not only has the merit of high computation speed comparing with traditional models but also contribute to the improvement of forecasting accuracy shown in Table 5.

![Figure 15: Analysis of degradation tendency prediction results in terms of MAE, MAPE, and $R^2$.](image)

### Table 5: Analysis of forecasting accuracy of eight models.

<table>
<thead>
<tr>
<th>Models</th>
<th>MAE</th>
<th>MAPE (%)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEEMDAN-PE-IGMMW</td>
<td>0.047</td>
<td>6.691</td>
<td>0.982</td>
</tr>
<tr>
<td>EMD-PE-IGMMW</td>
<td>0.062</td>
<td>8.721</td>
<td>0.978</td>
</tr>
<tr>
<td>CEEMDAN-IGMMW</td>
<td>0.076</td>
<td>9.938</td>
<td>0.972</td>
</tr>
<tr>
<td>IGMMW</td>
<td>0.089</td>
<td>10.911</td>
<td>0.831</td>
</tr>
<tr>
<td>ELM</td>
<td>0.097</td>
<td>13.527</td>
<td>0.723</td>
</tr>
<tr>
<td>SVR</td>
<td>0.123</td>
<td>17.277</td>
<td>0.499</td>
</tr>
<tr>
<td>ARIMA</td>
<td>0.169</td>
<td>20.965</td>
<td>0.503</td>
</tr>
<tr>
<td>BPNN</td>
<td>0.167</td>
<td>21.319</td>
<td>0.474</td>
</tr>
</tbody>
</table>

### Table 6: Comparison of computing time of eight models in degradation tendency prediction.

<table>
<thead>
<tr>
<th>Models</th>
<th>Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEEMDAN-PE-IGMMW</td>
<td>2.33</td>
</tr>
<tr>
<td>EMD-PE-IGMMW</td>
<td>2.65</td>
</tr>
<tr>
<td>CEEMDAN-IGMMW</td>
<td>2.89</td>
</tr>
<tr>
<td>IGMMW</td>
<td>2.01</td>
</tr>
<tr>
<td>ELM</td>
<td>3.23</td>
</tr>
<tr>
<td>SVR</td>
<td>14.57</td>
</tr>
<tr>
<td>ARIMA</td>
<td>10.20</td>
</tr>
<tr>
<td>BPNN</td>
<td>9.32</td>
</tr>
</tbody>
</table>
5. Conclusions

In this paper, a multistep prediction method based on CEEMDAN permutation entropy and the IGMMW model, systematically integrating the ideas of series decomposition and model update, is proposed to forecast the degradation tendency of aircraft engines. The proposed method has three unique merits. Firstly, a new one-dimensional IDI is constructed by using the multidimensional sensory data to accurately quantify the degradation levels of engines. Secondly, because of the significant advantages of CEEMDAN in signal decomposition, CEEMDAN is used to decompose the generated IDI series to eliminate the influence of data fluctuations. Furthermore, an IMFs reconstruction strategy based on PE theory is designed to reduce the complexity of decomposed components. Finally, an IGMMW model is innovatively developed to predict the degradation tendency to further improve the forecasting accuracy and computation efficiency.

A well-known sensory dataset of aircraft engines is used to demonstrate the effectiveness and superiority of the proposed prediction method. The experimental results confirm that the proposed multistep method contributes to achieve the accurate and rapid prediction of degradation trend for aircraft engines. Compared with the traditional single forecasting models, the predicted idea based on multistage mode is more suitable for practical engineering applications.

Acronyms

- CBM: Condition-based maintenance
- WT: Wavelet transform
- BPNN: Backpropagation neural network
- EMD: Empirical mode decomposition
- EEMD: Ensemble empirical mode decomposition
- CEEMDAN: Complete ensemble empirical mode decomposition with adaptive noise
- PE: Permutation entropy
- IGMMW: Improved Grey–Markov model with moving window
- IDI: Integrated degradation index
- IMFs: Intrinsic mode functions
- RIMFs: Reconstructed IMFs
- GM(1,1): Grey model
- AGO: Accumulated generating operator
- MAE: Mean absolute error
- MAPE: Mean absolute percentage error
- ELM: Extreme learning machine
- SVR: Support vector regression
- ARIMA: Autoregressive integrated moving average model.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the National Key R&D Program of China (Nos. 2016YFC0402205 and 2016YFC040910), the National Natural Science Foundation of China (NSFC) (Nos. 51579107, 51079057, and 51809099), and the Natural Science Foundation of Huazhong University of Science and Technology (No. 2017KFXJJ209).

References


