Intelligent Deep Learning Method for Forecasting the Health Evolution Trend of Aero-Engine With Dispersion Entropy-Based Multi-Scale Series Aggregation and LSTM Neural Network

WEI JIANG¹,², NAN ZHANG¹,², XIAOMING XUE¹,², YANHE XU³, (Member, IEEE), JIANZHONG ZHOU³, AND XINZI WANG²

¹Jiangsu Key Laboratory of Advanced Manufacturing Technology, Huaiyin Institute of Technology, Huai’an 223003, China
²Faculty of Mechanical and Material Engineering, Huaiyin Institute of Technology, Huai’an 223003, China
³School of Hydropower and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

Corresponding authors: Wei Jiang (jw_hust@163.com) and Xiaoming Xue (taxueli@hyit.edu.cn)

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 51709122 and Grant 51809099, and in part by the Open fund for Jiangsu Key Laboratory of Advanced Manufacturing Technology under Grant HGAMTL-1713.

ABSTRACT Accurate health evolution trend forecasting of aero-engine is essential for operation reliability and maintenance costs of aeronautical equipment. In this study, an intelligent deep learning method, systematically blending the dispersion entropy-based multi-scale series aggregation scheme and long short term memory (LSTM) neural network, is proposed for forecasting the health evolution trend of aero-engine. Firstly, a comprehensive measurement of health levels, namely, integrated health state index (IHSI), is developed with high-dimensional dataset. Secondly, the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) is exploited to decompose the IHSI sequence into several multi-scale series to further capture the internal characteristics of original sequence. Subsequently, multi-scale series aggregation assisted with dispersion entropy analysis theory is conducted for obtaining the aggregated sub-series (ASS). Finally, the ASS are served as the inputs of LSTM network to complete the health evolution trend forecasting of aero-engine. To demonstrate the effectiveness of the proposed method, six approaches are present for the comparisons of forecasting performance. The experimental results indicate that the proposed method can effectively measure the health evolution process of aero-engine and further obtain more accurate trend forecasting results.

INDEX TERMS Aero-engine, health evolution trend forecasting, health state index, CEEMDAN, multi-scale series aggregation, dispersion entropy, LSTM neural network.

I. INTRODUCTION

With the improvement of systemic integration degree and automation level, complete and reasonable health state monitoring has an important significance in ensuring stable operation of equipment and meanwhile contributes to achieve the condition-based maintenance (CBM) [1]–[3]. Aero-engine, as a kernel part of aeronautical equipment, plays an important role in system safety and operation reliability [4]–[6]. Due to the significance of engine for aeronautical machine, necessary monitoring and analysis of health state for engines should be effectively implemented during the whole cycle life of equipment, which can extract real-time state information of engine and avoid serious security incidents. More specifically, health evolution trend forecasting is an essential step for the implementation of monitoring and status analysis, which is helpful to evaluate the operation states in the future and for the purpose of making maintenance plan in advance [7]–[9].

In order to accurately forecast the health evolution trend of aero-engine, an appropriate measurement for effectively quantifying the health levels of engine should be built in advance. Note that plenty of data from various parts are acquired during runtime for subsequent analysis [10]. Accordingly, how to make full use of these data to build
a health state measurement is the primary problem to be solved. Aiming at this problem, some researches on building a suitable indicator for the effective description of health levels have been conducted. For example, adopted the power condition to characterize the performance levels of battery packs [11]. Kral obtained the operation status of vehicle cooling systems based on the analysis of oil conditions [12]. Gebrael utilized the collected vibration signals to analyze the status of rolling bearing [13]. However, the above studies usually adopt the single physical signal for analysis, in which the reflection of health status is only from an one-sided aspect and some signals containing health state information may be neglected. Furthermore, a wide variety of signals are collected by acquisition system, which causes great difficulties for the selection of representative signal. Therefore, the construction of an excellent health state index is still a difficult point to be overcome urgently for health evolution trend prediction of aero-engine.

Health evolution trend forecasting is fundamentally a problem of time series prediction, i.e., obtaining the future health conditions. On the whole, the prediction models mainly include three categories, i.e., knowledge-based, physics-based and data-driven models. Actually, due to the difficulties of obtaining prior knowledge and physical principles, the former two prediction models usually suffer many dilemmas in practical applications. Data-driven models, taking full advantage of monitoring data, can realize the accurate prediction without prior knowledge and physical rules [14], [15]. In recent years, as a popular data-driven model, artificial neural network (ANN) have been widely used for time series forecasting and achieved many successful cases [16], [17]. Adhikari established an ANN-based model to predict the financial time series [18]. Khashei adopted back propagation neural network (BPNN) to complete the task of time series prediction [19]. Tian utilized an ANN model to forecast the remaining useful life of bearing [20]. However, the change of time series in practical engineering is mostly non-linear and non-stationary, and usually shows evident characteristics of fluctuations. For this reason, it is quite difficult to forecast series fluctuation accurately and reliably. Because the change of health status is a gradual accumulation process, health trend forecasting not only relies on the current condition but on the condition of the past. If only consider the condition at present, the information contained in the condition at an earlier time would be ignored. Aiming at this problem, a new deep learning model, i.e., recurrent neural network (RNN), is developed and constructs the associations between hidden layer nodes, where the memory of recent states can be stored [21]. Because of memory characteristics, RNN can handle the related data in the time domain. Many researches have demonstrated the effectiveness of RNN for time series prediction [22], [23]. An as improved version of RNN, long short term memory (LSTM) neural network has been applied in time series prediction, language progressing and other fields, and attracted more and more attentions [24], [25]. Due to the “gate” structure of LSTM, the useless information can be strictly filtered and more valuable information can be extracted from historical dataset.

Based on the analysis above, an intelligent deep learning method, systematically blending the dispersion entropy (DE)-based multi-scale series aggregation scheme and LSTM neural network, is proposed in this paper to forecast the health evolution trend of aero-engine. More specifically, a comprehensive measurement of health levels, namely, integrated health state index (IHSI), is constructed with high-dimensional dataset. Different from the indicator adopted in [11]–[13], the constructed IHSI fully retains useful health status information and reflect the condition of engines synthetically. Subsequently, CEEMDAN algorithm is exploited to decompose the IHSI sequence into several multi-scale series to further capture the internal characteristics of original sequence. Because of excellent decomposition performance, CEEMDAN contributes to obtain intrinsic mode function (IMF) components with single frequency and provides data foundation for next research. Then, multi-scale series aggregation assisted with DE analysis theory is conducted for obtaining the aggregated sub-series (ASS). With the obtained ASS, there will be fewer sequences comparing with original multi-scale series, and prediction efficiency and accuracy will also be improved. Finally, the ASS are served as the inputs of LSTM network to complete the health evolution trend forecasting of aero-engine. Because of excellent ability of information filtering, LSTM is helpful for the improvement of final forecasting accuracy. To demonstrate the effectiveness of the proposed method, six approaches are present for the comparisons of forecasting performance. The experimental results indicate that the proposed method can effectively measure the health evolution process of aero-engine and further obtain more accurate trend forecasting results.

The rest of this paper is arranged as follows: Section 2 briefly introduces the CEEMDAN and LSTM neural network. Section 3 provides the detailed descriptions of the proposed method. In Section 4, experiment and results analysis are present. Finally, we conclude this study in Section 5.

II. BACKGROUND KNOWLEDGE
A. CEEMDAN ALGORITHM
To adaptively analyze the time-frequency characteristics of signal, an empirical mode decomposition (EMD) method was developed by Huang et al. [26]. With this method, the original signal can be decomposed into several multi-scale series, i.e., IMFs, which reveal the features of signal from different time scales [27]. Based on the detailed implementation procedures of EMD described in [27], the raw signal can be reconstructed by n IMFs and one residue \( r_n(t) \):

\[
x(t) = \sum_{i=1}^{n} \text{IMF}_i(t) + r_n(t)
\]

where \( x(t) \) is the original signal and \( \text{IMF}_i(t) \) denotes the obtained \( i \)-th IMF after signal decomposition with EMD.
Although it has been demonstrated that EMD shows excellent performance in analyzing the non-stationary signal, there are some inherent limitations that make some restriction on the application of EMD, such as mode mixing problem and end point effect. Aiming at these problems, a noise-assisted signal analysis approach, named ensemble empirical mode decomposition (EEMD), was proposed by Wu and Huang [28]. However, it lacks the ability to completely eliminate Gaussian white noise after signal reconstruction, and the high time consumption caused by the added noise make great influence on the application of the decomposition approach. Based on these, a complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) method was developed to make up for the deficiencies of EEMD [29]. It resolves the problem of mode mixing, and the reconstructed error is always zero. In the meantime, the computational costs of CEEMDAN are significantly lower than that of EEMD. The implementation steps of CEEMDAN can be summarized below:

1. With the EMD method, each \( x_i(t) = x(t) + \varepsilon_0 w_i(t) \), \( i = 1, 2, \ldots, I \) can be decomposed, where \( \varepsilon_0 \) denotes a noise factor, and \( w_i(t) \) denotes Gaussian white noise with zero mean and homogeneity of variance. Then, the first IMF of original signal using CEEMDAN can be obtained:

   \[
   \overline{IMF}_1(t) = \frac{1}{I} \sum_{i=1}^{I} IMF_{i1}(t)
   \]  

2. Calculate the first residue,

   \[
   r_1(t) = x(t) - \overline{IMF}_1(t)
   \]

3. Decompose residue \( r_1(t) + \varepsilon_1 E_1(w_1(t)) \) to calculate the second IMF,

   \[
   \overline{IMF}_2(t) = \frac{1}{I} \sum_{i=1}^{I} E_1(r_1(t) + \varepsilon_1 E_1(w_1(t)))
   \]

where \( E_j(\cdot) \) denotes the operator of \( j \)-th IMF obtained by EMD.

4. Step (2) to (3) are repeated until all IMFs have been addressed. The final residue is calculated below,

   \[
   r_m(t) = x(t) - \sum_{j=1}^{m} \overline{IMF}_j(t)
   \]

where \( m \) is the number of IMFs. Thus, through the procedures mentioned above, the raw signal \( x(t) \) can be reconstructed as follows:

   \[
   x(t) = \sum_{j=1}^{m} \overline{IMF}_j(t) + r_m(t)
   \]

Equation (6) presents a reconstruction form of raw signal. Note that the IMFs indicate the features of signal at different time scales, and the residue changes more smoothly and is helpful for the improvement of forecasting accuracy.

B. LSTM NEURAL NETWORK

The LSTM neural network is a deep learning method with wide range of application, where inputs and outputs can change with time. Because of its excellent performance on dealing with long-term dependency problem, LSTM neural network has been used in the fields of time series forecasting and attracted lots of attention [24], [25]. Specifically, the memory cell in network, which realizes the function of temporal state storage, is the foundation of the whole architecture. Based on this memory cell, the information can be updated through three gates, including input gate, forget gate and output gate. The implementation process of LSTM can be summarized as follows [30]: (1) The input activations are controlled by the input gate, and the new information can be stored by memory cell when the input gate is activated. (2) The extraneous information can be removed by the forget gate, and the past cell conditions can be forgotten when the forget gate is activated. (3) The output activations are controlled by the output gate, and the final status can be updated by the latest cell output when the output gate is activated.

For the purpose of health state tendency prediction, the LSTM neural network adopts a three-layers structure in this paper. In addition, suppose that the \( x = (x_1, x_2, \ldots, x_T) \) and \( y = (y_1, y_2, \ldots, y_T) \) represent the historical data and corresponding predicted data, respectively. So, the forecasting results can be obtained by the following formulas [31]:

\[
\begin{align*}
    i_t & = \sigma (W_{ix} x_t + W_{ik} k_{t-1} + W_{ia} a_{t-1} + b_i) \\
    f_t & = \sigma (W_{fx} x_t + W_{fk} k_{t-1} + W_{fa} a_{t-1} + b_f) \\
    a_t & = \sigma (W_{ax} x_t + W_{ak} k_{t-1} + b_a) \\
    k_t & = \sigma (W_{sk} k_{t-1} + b_k) \\
    y_t & = W_{yk} k_t + b_y
\end{align*}
\]

where \( i_t \) is the input gate, \( W \) is the weight matrix, \( k_t \) is the vectors for activating the block of memory, \( a_t \) is the vectors for activating the memory cell. \( f_t \) denotes the forget gate, \( a_t \) denotes the output gate, \( b \) denotes the bias and "\( \cdot \)" denotes dot product. Besides, \( \sigma (\cdot) \) represents sigmoid function:

\[
\sigma (x) = \frac{1}{1 + e^{-x}}
\]

\( g (\cdot) \) and \( h (\cdot) \) represent centered logistic functions:

\[
\begin{align*}
    g (x) & = \frac{4}{1 + e^{-x}} - 2, \quad x \in [-2, 2] \\
    h (x) & = \frac{2}{1 + e^{-x}} - 1, \quad x \in [-1, 1]
\end{align*}
\]

The basic structure of LSTM is depicted in Fig. 1.

III. THE PROPOSED HEALTH EVOLUTION TREND FORECASTING METHOD

In this paper, an intelligent deep learning method, systematically blending the DE-based multi-scale series aggregation scheme and LSTM network, is proposed in this paper to predict the health evolution trend of aero-engine. This section
contains the following: the construction of integrated health state index, the aggregation of multi-scale series based on dispersion entropy analysis and the implementation steps of the proposed method.

**A. CONSTRUCTION OF INTEGRATED HEALTH STATE INDEX**

To accurately and comprehensively evaluate the health levels of aero-engine, it is necessary to construct a proper health state index. With a large number of available sensory signals, a comprehensive index, named integrated health state index (IHSI), is innovatively developed based on an idea of liner transformation in this paper. In fact, the construction of IHSI can be regarded as a transformation from high-dimensional dataset to one-dimensional series.

Suppose $Q_1$ of $U_1 \times V$ matrix and $Q_2$ of $U_2 \times V$ matrix denote two high-dimensional datasets, which represent the faulty and healthy states of engine, respectively. Based on $Q_1$ and $Q_2$, a transformation matrix $T$ can be generated to build a mapping relationship between the original high-dimensional sensory dataset and the constructed one-dimensional IHSI sequence as follows:

$$T = \left( Q^T Q \right)^{-1} Q^T G$$

where $Q = [Q_1; Q_2]^T$, $G = [G_1, G_2]^T$, $G_1$ denotes a $1 \times U_1$ zero vector and $G_2$ denotes a $1 \times U_2$ unit vector. On the basis of the obtained matrix $T$ and the collected dataset $R$, the IHSI, written $h$, can be calculated as follows:

$$h = T \cdot R$$

Specifically, the value of IHSI is limited in $[0, 1]$, and “1” represents health condition and “0” represents fault condition. Thus, the constructed IHSI can provide a feasible way to comprehensively measure the health levels of aero-engine.

**B. AGGREGATION OF MULTI-SCALE SERIES BASED ON DISPERSION ENTROPY ANALYSIS**

Because of the impact of operational conditions and environment noise, the sensory signals usually show the significant non-stationary and non-linear features and the corresponding decomposition results with CEEMDAN always contain several IMFs. This phenomenon can make difficulties to effectively train the prediction model and improve the forecasting accuracy. Aiming at this problem, a multi-scale series aggregation scheme assisted with DE analysis theory is developed in this paper to retain all effective information of IMFs and meanwhile obviously decrease the number of IMFs. The process of DE-based multi-scale series aggregation can be summarized below.

For the multi-scale series IMFs obtained by CEEMDAN (c_1(t), c_2(t), ..., c_l(t), ..., c_m(N)) represents the i-th IMF component, $N$ is sequence length and $m$ is the number of IMFs), each component $c_i(t)$ can be mapped into $d_i(t) = (d_i(1), d_i(2), ..., d_i(N))$ by normal cumulative distribution function,

$$d_i(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c_i(t)} e^{-\frac{\tau^2}{2\sigma^2}} d\tau$$

where $\mu$ is the expectation, $\sigma$ is the standard deviation and the value of $d_i(t)$ changes between 0 and 1. And then, each element in $d_i(t)$, denoted as $d_i(j)$ ($j = 1, 2, ..., N$), can be transformed to an one-dimensional integer sequence from 1 to $S$ by a linear mapping model:

$$z_i^j = \text{round} (s \cdot d_i(j) + 0.5)$$

where $S$ is the number of category and $z_i^j$ is the $j$-th sequence. On this basis, define an embedding vector $\vec{z}_i^{\xi, s}$ as follows:

$$\vec{z}_i^{\xi, s} = (z_i^1, z_i^2 + \xi, ..., z_i^N + \xi (S - 1))$$

where $\xi$ is embedding dimension and $\omega$ is time delay. Based on Eq. (20), each sequence $\vec{z}_i^{\xi, s}$ will be mapped into a dispersion mode $\Pi_{1 \leq j \leq S}$. Specifically, $z_i^j = v_0$, $z_i^j + \omega = v_1$, ..., $z_i^j + (\xi - 1)\omega = v_{\xi - 1}$.

Subsequently, the frequency of each mode $\Pi_{1 \leq j \leq S}$ is calculated as:

$$p \left( \Pi_{1 \leq j \leq S} \right) = \frac{\text{Number} \left( \Pi_{1 \leq j \leq S} \right)}{N - (\xi - 1) \omega}$$

According to the above equation, the DE value of each multi-scale series c_i(t) can be obtained by the equation below:

$$H (c, \xi, s, \omega) = - \sum_{i=1}^{\xi} p \left( \Pi_{1 \leq j \leq S} \right) \ln \left( p \left( \Pi_{1 \leq j \leq S} \right) \right)$$

Note that the value of DE is bounded in $[0, 1]$. In essence, DE can effectively measure the randomness of sequences. Based on this property, the original multi-scale series can be
recombined to generate the aggregated sub-series, denoted as ASS, according to the following criterion:

\[
ASS = \sum_{i}^{i+j} c_i(t), \text{ s.t. } |H_i - H_{i+j}| \leq \frac{2(H_{\max} - H_{\min})}{m} \tag{23}
\]

With the proposed multi-scale series aggregation scheme, a group of ASS can be acquired by analyzing DE distributions of raw multi-scale series. In essence, the generated ASS not only completely retain the information of original series, but significantly decrease the number of series. Based on these, the constructed ASS can be served as the inputs of forecasting model for further improvement of prediction accuracy and efficiency.

C. IMPLEMENTATION STEPS OF THE PROPOSED METHOD

In order to accurately forecast the health evolution trend of aero-engine, an intelligent deep learning method combining DE-based multi-scale series aggregation and LSTM neural network is proposed. The flowchart of the proposed method is present in Fig. 2, and the implementation steps general procedures can be described in detail below.

Step 1: With the signal acquisition system, monitoring data during the whole lifecycle of aero-engine should be collected.

Step 2: Based on the collected datasets, the IHSI sequence can be built to measure the health levels of aero-engine.

Step 3: The generated IHSI sequence should be randomly split into training set and testing set. Specifically, training set
is used to search optimal parameters of LSTM neural network and further obtain the forecasting results with high accuracy.

Step 4: Based on CEEMDAN decomposition algorithm, the IHSI sequence of training set can be decomposed into several multi-scale series, i.e., IMFs.

Step 5: All the IMFs are aggregated with the developed DE-based multi-scale series aggregation scheme and obtain a series of ASS.

Step 6: An optimal LSTM model is trained for each ASS to forecast the sequence of the corresponding ASS.

Step 7: Sum over predicted sequences and then the final forecasting results are available.

Step 8: Compare testing set and forecasting results and the performance of prediction method can be evaluated.

IV. EXPERIMENT AND RESULTS ANALYSIS
A. DESCRIPTIONS OF EXPERIMENTAL DATA AND ERROR CRITERIA

To demonstrate the effectiveness of the proposed method, a classical dataset derived from the problem of prognostics and health management in 2008 is used and the corresponding results can be further analyzed and evaluated in this section. Specifically, the dataset is collected from an aero-engine model shown in Fig. 3, and consists of unit number, cycle index, operational setting parameters and 21 sensory signals for each cycle [32]. As described in Reference [33], there are 6 kinds of operational conditions of aero-engine according to different groups of setting parameters. With a view of the difference between different sensory data, it is unreasonable to directly adopt these data for the measurement of health levels of engine. Therefore, a comprehensive measurement that can accurately quantify the health levels of engine is urgently needed, and on this basis the health evolution trend of aero-engine can be effectively predicted.

In addition, three commonly used error criteria are utilized to analyze the performance of prediction method, including mean absolute error (MAE), root mean square error (RMSE) and mean relative percentage error (MRPE) [34], [35]. MAE and RMSE are two scale-dependent measurements, which can reflect the proximity and the departure between the forecasting results and actual data, respectively. Besides, MRPE is a percentage measurement which can describe the average prediction ability of forecasting method. Definitions of these three criteria are listed as follows:

\[
\text{MAE} = \frac{1}{k} \sum_{i=1}^{k} |Y_i - \hat{Y}_i| \quad (24)
\]

\[
\text{RMSE} = \sqrt{\frac{1}{k} \sum_{i=1}^{k} (Y_i - \hat{Y}_i)^2} \quad (25)
\]

\[
\text{MRPE} = \frac{1}{k} \sum_{i=1}^{k} \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \times 100\% \quad (26)
\]

where \(k\) denotes the number of testing samples, \(Y_i\) and \(\hat{Y}_i\) are the \(i\)-th actual value and predicted value, respectively. Furthermore, note that all experiments are executed in MATLAB 2014.

B. IHSI SEQUENCE CONSTRUCTION

As mentioned in Reference [32], some sensory signals contain valuable health status information of aero-engine while others do not. We tend to select these signals that can effectively present the development trend of health status to establish the IHSI sequence. Following the strategy of signal filtration depicted in [33], seven sensory signals, as listed in Table 1, can be chosen for subsequent study.

Based on these selected signals and Eq. (16), matrices \(T_i\) \((i = 1, 2, \cdots, 6)\) can be built for six different operational conditions. Specially, \(Q_1\) and \(Q_2\) should be established in advance for the construction of transformation matrices. In this paper, \(Q_1\) is constructed based on sensory data under faulty states, where the remaining cycle life \(C_r\) \((C_r = C_o - C_w, C_o\) is the current operation cycle and \(C_w\) is the whole cycle) is bounded in \([-3, 0]\). Analogously, \(Q_2\) is built under healthy states, where \(C_r\) is less than -200. Based on \(T_1 \sim T_6\) and historical dataset, an IHSI sequence can be obtained by Eq. (17) as depicted in Fig. 4. It can be found from this figure that the constructed IHSI sequence shows a clear tendency of health development with the increase of cycle index. Consequently, the developed IHSI can be viewed as an effective and comprehensive measurement for quantifying health levels of aero-engine. Moreover, the first 150 health indexes in the IHSI sequence \((h(1), h(2), \cdots, h(150))\) can be chosen to

![FIGURE 3. The system diagram of an aero-engine model.](image-url)

| Serial Number | Signal Description                  | Unit  |
|---------------|------------------------------------|-------|
| 1             | Total temperature at low-pressure | °R    |
|               | compressor outlet                  |       |
| 2             | Total temperature at high-pressure | °R    |
|               | compressor outlet                  |       |
| 3             | Total temperature at low-pressure  | °R    |
|               | turbine outlet                     |       |
| 4             | Total pressure at high-pressure    | psia  |
|               | compressor outlet                  |       |
| 5             | Static pressure at high-pressure   | Psa   |
|               | compressor outlet                  |       |
| 6             | Ratio of fuel flow to Ps30         | pps/psia |
| 7             | Bypass ratio                       | -     |
train the model and the others \((h(151), h(152), \ldots, h(210))\) to evaluate the final forecasting results.

### C. IHSI SEQUENCE DECOMPOSITION AND DISPERSION ENTROPY-BASED MULTI-SCALE SERIES AGGREGATION

From Fig. 4, it can be found that the constructed IHSI sequence tend to fluctuate significantly so that the important information contained in the sequence would be submerged. In order to further capture the internal characteristics of original IHSI sequence and obtain the prediction results with high accuracy, CEEMDAN algorithm is used to decompose the IHSI sequence into several multi-scale series, i.e., IMFs. Fig. 5 presents decomposition results of IHSI sequence in detail, which consists of seven IMF components and one residue component.

Due to the potential impact of multi-scale series’ complexity on model training and forecasting efficiency, a multi-scale series aggregation scheme assisted with DE analysis theory is innovatively developed to generate new ASS. The DE distribution of all IMFs and residue component are shown in Fig. 6, and Table 2 lists the corresponding values of these components in detail. From Fig. 6, it can be found that the values of IMFs decrease monotonously until that the value of residue component is zero, which indicates that the randomness of decomposed multi-scale series show a declined tendency. With the aggregation formula (23), the interval length of DE for the recombination of multi-scale series, i.e., \(2(H_{\text{max}} - H_{\text{min}})/m\), can be calculated as 0.223. Based on Eq. (23) and actual distribution of DE, all multi-scale series can be divided into 4 groups, and then 4 ASS are aggregated accordingly as present in Table 3. It can be found that the DE value of IMF1 is within in \([0.669, 0.892]\), and thus it is served to construct the first ASS component. The similar procedures are implemented to reconstruct the ASS2, ASS3 and ASS4. Fig. 7 depicts the generated ASS components, which reflects the different characteristics of four ASS. Specifically, ASS1 is a non-stationary series with highest fluctuation frequency whereas RIMF4 shows a smooth trend during the whole cycle life. Besides, there are fewer sub-series comparing with original multi-scale series, which contributes to the improvement of prediction efficiency and accuracy. Therefore, the aggregated ASS can be directly employed as the inputs of forecasting model to forecast the health evolution trend of aero-engine.

### D. HEALTH EVOLUTION TREND FORECASTING OF AERO-ENGINE

As described in Fig. 2, on the basis of the obtained ASS, the forecasting values for all ASS can be calculated by LSTM neural network. Fig. 8 presents the detailed comparisons between actual values and forecasting values of LSTM.
network for 4 ASS. From the figure, it can be seen that the significant non-stationarity of ASS1 and ASS2 cause much deviation between actual values and the corresponding prediction values, while the forecasting sequences of ASS3 and ASS4 can accurately fit the real sequences due to their gradual variation and stationary property. With the obtained forecasting values of ASS1 to ASS4, the final results of the proposed intelligent deep learning method for forecasting the health evolution trend of aero-engine can be further calculated by the cumulative operation of all predicted sequences. Fig. 9 shows the comparisons between original IHSI sequence and final forecasting series in detail, which indicates that the predicted results achieve the great prediction performances on accuracy and fitting degree.

Besides, for comparison, five other prediction methods, including EMD-DE-LSTM, CEEMDAN-LSTM, LSTM, extreme learning machine (ELM) and BPNN, are also used for health evolution trend forecasting of aero-engine. More specifically, the comparison between the proposed method and EMD-DE-LSTM is conducted to validate the excellent decomposition performance of CEEMDAN approach, the comparison between the proposed method and CEEMDAN-LSTM tends to show the superiority of multi-scale series aggregation scheme assisted with DE analysis theory, and the comparison between LSTM and other two single models tries to demonstrate the better prediction performance of LSTM. Finally, the comparison results between the proposed method and other five methods can further indicate the feasibility and superiority of combining CEEMDAN decomposition algorithm, DE-based multi-scale series aggregation scheme and LSTM for the improvement of prediction accuracy.
FIGURE 8. The forecasting results of LSTM network for 4 ASS: (a) ASS1 (b) ASS2 (c) ASS3 (d) ASS4.

FIGURE 9. The results of the proposed health evolution trend forecasting method.

TABLE 4. Forecasting accuracy comparison of six methods.

| Methods           | MAE     | RMSE   | MRPE(%) |
|-------------------|---------|--------|---------|
| The proposed method | 0.030   | 0.042  | 3.621   |
| EMD-DE-LSTM       | 0.043   | 0.049  | 5.523   |
| CEEMDAN-LSTM      | 0.067   | 0.092  | 7.714   |
| LSTM              | 0.082   | 0.126  | 10.697  |
| ELM               | 0.125   | 0.191  | 14.272  |
| BPNN              | 0.147   | 0.253  | 20.616  |

Fig. 10 presents the detailed comparison results of six methods in health evolution trend forecasting for aero-engine.

In addition, the forecasting accuracies of six methods are depicted in Fig. 11, and the corresponding values of three
error criteria, including MAE, RMSE and MRPE, are listed in Table 4. From Figure 10, it can be observed that the predicted results of the proposed method have better performance for fitting the actual IHSI series in the entire time-scale than that of other methods. More specifically, based on the analysis of error criteria depicted in Fig. 11 and
Table 4, we observe that MAE, RMSE and MRPE of the proposed method are 0.030, 0.042 and 3.621%, which are obviously less than that of other prediction methods.

The following conclusions can be summarized on the basis of the above results as: (1) Compared with other approaches, the proposed forecasting method based on multi-scale series aggregation and LSTM network has the best prediction performance, whereas BPNN has the worst. This is because the optimal model parameters of BPNN are difficult to determine. (2) The combination prediction methods, including the proposed method, EMD-DE-LSTM and CEEMDAN-LSTM, can achieve the higher forecasting accuracy than other three single methods, which illustrate the effectiveness and necessity of series decomposition for improving the final prediction accuracy. (3) With CEEMDAN algorithm, the mode mixing phenomenon occurring in the decomposed series of EMD can be effectively eliminated so that the predicted error are further decreased. (4) Because of the developed DE-based multi-scale series aggregation scheme, the forecasting performance of the proposed method can be significantly enhanced. (5) Different from other single methods, LSTM realizes the higher accuracy and meanwhile has the better ability for tracking the jumping IHSI points, which indicates the feasibility and superiority of adopting the LSTM network in the proposed forecasting method. Consequently, the proposed intelligent deep learning method not only contributes to accurate measurement of health levels for aero-engine, but can achieve the superior forecasting performance comparing with other traditional method.

V. CONCLUSION

In this paper, an intelligent deep learning method, systematically blending the DE-based multi-scale series aggregation scheme and LSTM neural network, is proposed for forecasting the health evolution trend of aero-engine. Compared with existing approaches, the proposed method has three significant advantages. First, a comprehensive measurement of health levels, i.e., IHSI, is constructed with high-dimensional dataset for accurately quantifying the health status of aero-engine. Second, aiming at the fluctuations of the IHSI series, CEEMDAN is applied to decompose the IHSI sequence to further capture the internal characteristics of original sequence, and additionally an DE-based multi-scale series aggregation scheme is innovatively developed to retain all effective information of multi-scale series and meanwhile obviously decrease the number of series. Lastly, the aggregated sequences, i.e., ASS, are served as the inputs of LSTM neural network to forecast the health evolution trend of aero-engine.

To demonstrate the effectiveness of the proposed method, six different approaches are applied in the analysis of a well-known case study. The experimental results indicate that the proposed method can effectively measure the health evolution process of aero-engine and further obtain more accurate trend forecasting results. Furthermore, because of the fluctuations and non-stationarity of actual operation data, the forecasting strategy on the basis of combination pattern is much better served for actual engineering than single forecasting models.

REFERENCES

[1] D. Wang and K.-L. Tsui, “Brownian motion with adaptive drift for remaining useful life prediction: Revisited,” Mech. Syst. Signal Process., vol. 99, pp. 691–701, Jan. 2018.
[2] C.-C. Lin, D.-J. Deng, C.-H. Kuo, and L. Chen, “Concept drift detection and adaptation in big imbalance industrial IoT data using an ensemble learning method of offline classifiers,” IEEE Access, vol. 7, pp. 56198–56207, 2019.
[3] W. Fu, K. Shao, J. Tan, and K. Wang, “Fault diagnosis for rolling bearings based on composite multiscale fine-sorted dispersion entropy and SVM with hybrid mutation SCA-HHO algorithm optimization,” IEEE Access, vol. 8, pp. 13086–13104, 2020.
[4] S.-P. Zhu, H.-Z. Huang, W. Peng, H.-K. Wang, and S. Mahadevan, “Probabilistic physics of failure-based framework for fatigue life prediction of aircraft gas turbine discs under uncertainty,” Rel. Eng. Syst. Safety, vol. 146, pp. 1–12, Feb. 2016.
[5] Y. Wang, Y. Shi, M. Cai, W. Xu, T. Pan, and Q. Yu, “Study of fuel-controlled aircraft engine for fuel-powered unmanned aerial vehicle: Energy conversion analysis and optimization,” IEEE Access, vol. 7, pp. 109246–109258, 2019.
[6] M.-A. Beaudoin and K. Bezdinian, “Analytical lump model for the non-linear dynamic response of bolted flanges in engine-casing cases,” Mech. Syst. Signal Process., vol. 115, pp. 14–28, Jan. 2019.
[7] A. Heng, S. Zhang, A. C. C. Tan, and J. Mathew, “Rotating machinery prognostics: State of the art, challenges and opportunities,” Mech. Syst. Signal Process., vol. 23, no. 3, pp. 724–739, Apr. 2009.
[8] B. Qi, Y. Wang, P. Zhang, C. Li, and H. Wang, “A novel deep recurrent belief network model for trend prediction of transformer DGA data,” IEEE Access, vol. 7, pp. 80069–80078, 2019.
[9] Y. Li, H. Liu, and Z. Ma, “Trend extraction of rail corrugation measured dynamically based on the relevant low-frequency principal components reconstruction,” Meas. Sci. Technol., vol. 27, no. 10, Oct. 2016, Art. no. 105005.
[10] Y. Zhang, G. O. H. Peng, J. K. Banda, S. Dasgupta, M. Husband, R. Su, and C. Wen, “An energy efficient power management solution for a fault-tolerant more electric engine/aircraft,” IEEE Trans. Ind. Electron., vol. 66, no. 7, pp. 5663–5675, Jul. 2019.
[11] S.-L. Wang, D.-I. Stroe, C. Fernandez, L.-Y. Xiong, Y.-C. Fan, and W. Cao, “A novel power state evaluation method for the lithium battery packs based on the improved external measurable parameter coupling model,” J. Cleaner Prod., vol. 242, Jan. 2020, Art. no. 118506.
[12] J. Král, B. Konecný, J. Král, K. Madac, G. Ferdoko, and V. Molnar, “Degradation and chemical change of longlife oils following intensive use in automobile engines,” Measurement, vol. 50, pp. 34–42, Apr. 2014.
[13] N. Z. Gebrael, M. A. Lawley, R. Li, and J. K. Ryan, “Residual-life distributions from component degradation signals: A Bayesian approach,” IIE Trans., vol. 37, no. 6, pp. 543–557, Jun. 2005.
[14] L. Liu, Q. Guo, D. Liu, and Y. Peng, “Data-driven remaining useful life prediction considering sensor anomaly detection and data recovery,” IEEE Access, vol. 7, pp. 58336–58345, 2019.
[15] S. Guth and T. P. Sapsis, “Machine learning predictors of extreme events occurring in complex dynamical systems,” Entropy, vol. 21, no. 10, p. 925, Sep. 2019.
[16] D. Pradeepkumar and V. Ravi, “Forecasting financial time series volatility using particle swarm optimization trained quantile regression neural network,” Appl. Soft Comput., vol. 58, pp. 35–52, Sep. 2017.
[17] J. Wang and J. Wang, “Forecasting stock market indexes using principle component analysis and stochastic time effective neural networks,” Neurorobocomputing, vol. 156, pp. 68–78, May 2015.
[18] R. Adhikari and R. K. Agrawal, “A combination of artificial neural network and random walk models for financial time series forecasting,” Neural Comput. Appl., vol. 24, no. 6, pp. 1441–1449, May 2014.
[19] M. Khashei and M. Bijari, “Fuzzy artificial neural network (p, d, q) model for incomplete financial time series forecasting,” J. Intell. Fuzzy Syst., vol. 26, no. 2, pp. 831–845, 2014.
[20] Z. Tian, L. Wong, and N. Safaei, “A neural network approach for remaining useful life prediction utilizing both failure and suspension histories,” Mech. Syst. Signal Process., vol. 24, no. 5, pp. 1542–1555, Jul. 2010.
[21] T. Lin, B. G. Horne, and C. L. Giles, “How embedded memory in recurrent neural network architectures helps learning long-term temporal dependencies,” *Neural Netw.*, vol. 11, no. 5, pp. 861–868, Jul. 1998.

[22] M. Lopez-Martin, B. Carro, A. Sanchez-Esguevillas, and J. Lloret, “Network traffic classifier with convolutional and recurrent neural networks for Internet of Things,” *IEEE Access*, vol. 5, pp. 18042–18050, 2017.

[23] P.-A. Chen, L.-C. Chang, and F.-J. Chang, “Reinforced recurrent neural networks for multi-step-ahead flood forecasts,” *J. Hydrol.*, vol. 497, pp. 71–79, Aug. 2013.

[24] M. Sundermeyer, H. Ney, and R. Schluter, “From feedforward to recurrent LSTM neural networks for language modeling,” *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 23, no. 3, pp. 517–529, Mar. 2015.

[25] R. He, Y. Liu, K. Wang, N. Zhao, Y. Yuan, Q. Li, and H. Zhang, “Automatic cardiac arrhythmia classification using combination of deep residual network and bidirectional LSTM,” *IEEE Access*, vol. 7, pp. 102119–102135, 2019.

[26] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proc., Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

[27] Z. Qu, W. Mao, K. Zhang, W. Zhang, and Z. Li, “Multi-step wind speed forecasting based on a hybrid decomposition technique and an improved back-propagation neural network,” *Renew. Energy*, vol. 133, pp. 919–929, Apr. 2019.

[28] Z. Wu and N. E. Huang, “Ensemble empirical mode decomposition: A noise-assisted data analysis method,” *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, Jan. 2009.

[29] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, “A complete ensemble empirical mode decomposition with adaptive noise,” in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 2011, pp. 4144–4147.

[30] X. Shi, Z. Chen, H. Wang, D. Y. Yeung, W. Wong, and W. Woo, “Convolutional LSTM network: A machine learning approach for precipitation now casting,” in *Proc. Adv. Neural Inform. Process. Syst.*, vol. 28, 2015, pp. 802–810.

[31] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, “Long short-term memory neural network for traffic speed prediction using remote microwave sensor data,” *Transp. Res. C: Emerg. Technol.*, vol. 54, pp. 187–197, May 2015.

[32] A. Saxena, K. Goebel, D. Simon, and N. Eklund, “Damage propagation modeling for aircraft engine run-to-failure simulation,” in *Proc. Int. Conf. Prognostics Health Manage.*, Denver, CO, USA, Oct. 2008, pp. 6–9.

[33] T. Wang, J. Yu, D. Siegel, and J. Lee, “A similarity-based prognostics approach for remaining useful life estimation of engineered systems,” in *Proc. Int. Conf. Prognostics Health Manage.*, Denver, CO, USA, Oct. 2008, pp. 6–9.

[34] W. Fu, K. Wang, C. Zhang, and J. Tan, “A hybrid approach for measuring the vibrational trend of hydroelectric unit with enhanced multi-scale chaotic series analysis and optimized least squares support vector machine,” *Trans. Inst. Meas. Control*, vol. 41, no. 15, pp. 4436–4449, Nov. 2019.

[35] Y. Jiang, S. Liu, L. Peng, and N. Zhao, “A novel wind speed prediction method based on robust local mean decomposition, group method of data handling and conditional kernel density estimation,” *Energy Convers. Manage.*, vol. 200, Nov. 2019, Art. no. 112099.

**WEI JIANG** received the B.S. degree in mechanical engineering and the Ph.D. degree in water resources and hydropower engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2014 and 2019, respectively. He is currently a Lecturer with the Faculty of Mechanical and Material Engineering, Huaiyin Institute of Technology. His research interests include fault diagnosis, trend prognosis, and intelligent management for mechanical systems.

**YANHE XU** (Member, IEEE) received the B.S. degree from the Xi’an University of Technology, Xi’an, China, in 2012, and the Ph.D. degree in water resources and hydropower engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2019. He is currently a Lecturer with the School of Hydropower and Information Engineering, Huazhong University of Science and Technology. His research interests include control theory, system identification, and modeling of power generation systems.

**WEI JIANG** received the B.S. degree in mechanical engineering and the Ph.D. degree in water resources and hydropower engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2014, and the Ph.D. degree in water resources and hydropower engineering from the Huazhong University of Science and Technology, Wuhan, in 2019. He is currently a Lecturer with the Faculty of Mechanical and Material Engineering, Huaiyin Institute of Technology. His research interests include control theory, system identification, and modeling of power generation systems.

**XIANGMING XUE** received the B.S. degree in mechanical and electronic engineering from the Changchun University of Technology, Changchun, China, in 2010, and the M.S. and Ph.D. degrees in water resources and hydropower engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2012 and 2016, respectively. He is currently a Lecturer with the Faculty of Mechanical and Material Engineering, Huaiyin Institute of Technology. His research interests include vibration signal processing, intelligent fault diagnosis, and modeling of power generation systems.

**YANHE XU** (Member, IEEE) received the B.S. degree from the Xi’an University of Technology, Xi’an, China, in 2012, and the Ph.D. degree in automatic control from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1982. He is currently a Professor with the School of Hydropower and Information Engineering, Huazhong University of Science and Technology. His research interests include control theory, system identification, and modeling of power generation systems.

**JIANZHONG ZHOU** was born in Wuhan, China, in December 1959. He received the B.S. degree in automatic control from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1982. He is currently a Professor with the School of Hydropower and Information Engineering, HUST. His research interests include modeling, control, and operation theory in hydraulic power plants.

**XINZI WANG** received the B.S. degree in electronic engineering from Wuhan Textile University, Wuhan, China, in 2014. She is currently an Assistant with the Faculty of Mechanical and Material Engineering, Huaiyin Institute of Technology. Her research interests include fault diagnosis, signal processing, and intelligent management for electronic systems.

**NAN ZHANG** received the B.S. degree in electronic information engineering from Wuhan Textile University, Wuhan, China, in 2014, and the Ph.D. degree in water resources and hydropower engineering from the Huazhong University of Science and Technology, Wuhan, in 2019. He is currently a Lecturer with the Faculty of Mechanical and Material Engineering, Huaiyin Institute of Technology. His research interests include control theory, system identification, and modeling of power generation systems.