Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

An SW-ELM Based Remaining Useful Life Prognostic Approach for Aircraft Engines

Dingzhou Peng * **Shen Yin** * **Kuan Li** * **Hao Luo** *

* *School of Astronautics, Harbin Institute of Technology, Harbin, 150001, China (e-mail: dingzhoup@163.com).*

Abstract: With the rapid development of prognostics and health management (PHM), the prognostic of the remaining useful life (RUL) is gradually being used for performance management and optimization. The aerospace industry is particularly in need of this, for instance, the remaining life expectancy of aircraft engines is of great significance to guarantee the safety and reliability. However, it is hard to establish the physical model of aircraft engines with the complex degradation process, which motivates the data-driven solution to RUL prediction. In this paper, a data-driven RUL prognostic approach is proposed for aircraft engines. Key performance indicators are extracted from sensor variables through principal component analysis. The summation wavelet-extreme learning machine is used to predict the KPIs' degradation process by iterative method, and then KPIs' degradation states are determined by subtractive-maximum entropy fuzzy clustering to calculate the RUL of engines. To validate the prediction model, aircraft engine degradation data are used for model simulation. Compared with other algorithms, the proposed method delivers superior prediction performance.

Keywords: Remaining useful life, Data-driven, Summation wavelet-extreme learning machine, Iterative prediction, Clustering

1. INTRODUCTION

In recent years, PHM for aircraft engines have become more and more popular with researchers, aiming to provide early warning of faults, extend system life and diagnose intermittent faults. In general, RUL prognostic methods can be divided into the following categories: model based, data-driven and hybrid model based prognostic method (Li et al. (2018)). Currently, model-based prognostics methods mainly include particle filtering (Jouin et al. (2016)), Weibull distribution (Ali et al. (2015)), reduced order sliding mode observer (Yang and Yin (2019)), etc. Data-driven methods include hidden markov models (Zhang et al. (2005)), recurrent neural networks (Yam et al. (2001)), support vector machines (Gebraeel et al. (2004)), etc. For RUL prediction, the following issues have to be considered: (1) the quantity, type and quality of the required data, (2) the impact of noise on the data, (3) the number of fault modes that can be handled, (4) whether new fault types can be handled (Sikorska et al. (2011)). As for aircraft engine systems, it is hard to implement the model based prognostic method due to the inherited complex degraded processes such as nonlinearity, randomness and non-stationarity (Xu et al. (2013)). Fortunately, there are many reliable sensors installed on the engines for data collection, based on which the fault information can be provided by real-time monitoring (Yin et al. (2019)). Therefore the data-driven methods can model the system's failure mechanism well, and provide better and RUL prediction performance (Yin et al. (2016)).

Currently, data-driven RUL estimation methods can be divided into two groups(Si et al. (2011)): (1) Direct RUL prognostic. The method can directly predicts the value of RUL based on the input data. In Yuan et al. (2016), the LSTM is used to directly predict the RUL of aircraft engines. However, It doesn't reveal the degradation process well, and predicted results relie largely on smooth and monotonic features. (2) Variable degradation based prognostic. In the method, the key performance indicators (KPIs) obtained by feature extraction can well characterize the degradation. RUL can be calculated by determining when the predicted KPIs reach the thresholds.

Since the prediction process of the second method is more intuitive and persuasive, this paper makes RUL prediction based on this method. A large number of sensor data of engines could generate redundant information, which could increase the computation load or affect the prognostics results. Therefore, there are two key questions:

(1) How to extract features and build the degradation model?

(2) How to exploit multiple features to judge system degraded states?

In order to solve these problems, a data-driven RUL prognostic approach, based on and summation wavelet-extreme learning machine (SW-ELM), is proposed in this paper. The main contribution of this paper can be summarized in the following 3 aspects:

- To reduce the data dimension and obtain useful features, principal component analysis (PCA) is adopted for feature extraction;
- To accelerate the computation and improve the prediction accuracy, SW-ELM is used to model the

degradation process of the engine, and the iterative prediction method is adopted to achieve the long-term prediction of degradation process;

- The degradation states of the engines can be determined by subtractive-maximum entropy fuzzy clustering (S-MEFC). Failure thresholds obtained by S-MEFC can be used to determine when the KPIs' prediction process ends.

The rest of this paper is organized as follows. Section II proposes the RUL prediction model for aircraft engines which contains feature extraction, degradation prediction model, multi-step prediction and threshold-based degradation state estimation. Section III validates the proposed model on aircraft engines data. Section IV concludes the work.

2. RUL PREDICTION MODEL

2.1 Feature Extraction

Sensor degradation data of aircraft engines are mainly divided into two trends: 1) the growing trend 2) the attenuation trend. To obtain representative features, PCA can be used to perform dimensionality reduction for two kind of sensor data respectively. The principal component with an information retention rate of over 90% can be used as the key performance indicator (KPI) for engine degradation. Finally, two KPIs obtained can be used to characterize the degradation process. KPI1 is increasing over time, while KPI2 is decreasing.

2.2 Degradation Prediction Model

To describe the degradation process of the engines well, it is important to find a suitable prediction model. Many studies have shown that multi-layer feedforward neural networks have better general approximation ability. However, the networks have the disadvantages of easy convergence to local minimum and slow learning process. Motivated by this, some researchers have proposed an improved model called SW-ELM (Javed et al. (2014)) (see Fig.1) that is based on Extreme Learning Machine (ELM) (Huang et al. (2004)). It only needs a batch learning for the training process, and has fast learning speed with good generalization performance.

The characteristics of SW-ELM are summarized as follows:

(1) Each node of the hidden layer contains two different activation functions: f_1 and f_2. The node output is the average of dual functions $\bar{f} = (f_1 + f_2)/2$.

$$f_1 = \theta(X) = log[x + \sqrt{1 + x^2}] \tag{1}$$

$$f_2 = \phi(X) = cos(5x) \cdot e^{-0.5x^2} \tag{2}$$

(2) To provide a suitable initial state for the SW-ELM, two parameter initialization methods are used here. The first one is the wavelet heuristic adjustment (Oussar and Dreyfus (2000)). The scaling and translation values are adjusted according to the interval of input data. This initialization method ensures that the activation function f_2 covers the whole input data. The second one is the Nguyen-Widrow initialization method (Nguyen and Widrow (1990)). It can determine the parameters (weights and biases) of the input-hidden layer by the range of the input data.

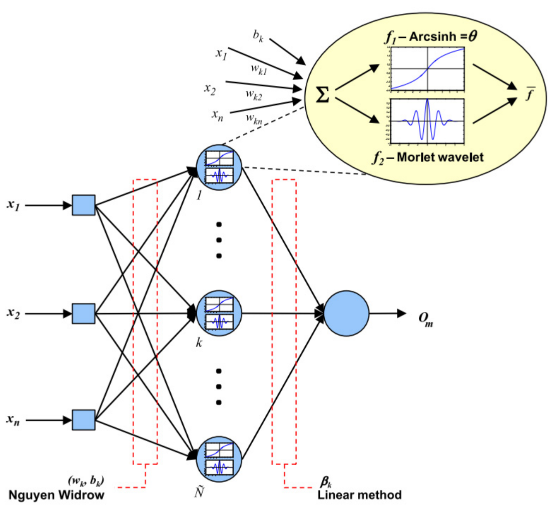


Fig. 1. Structure of SW-ELM (Javed et al. (2014))

The detailed implementation of SW-ELM can be referred to Javed et al. (2014). In this paper, SW-ELM is used to predict the value of the next moment of a KPI. Therefore, the input of SW-ELM is a piece of KPI data before time t, and the output is the value of a KPI at time t. The parameters of the input-hidden layer can be initialized by the method of Nguyen and Widrow (1990). The relationship between the hidden-output layer can be formulated in the following matrix

$$H_{avg}\beta = T \tag{3}$$

where H_{avg}, obtained by input data of training set, is the output matrix of the hidden layers. T is the target matrix of predicted data. Every row vector of H_{avg} and T correspond to an instance of the training data. β is the weight parameter of the hidden-output layer. Therefore, based on train set of a KPI, the least squares solution for β can be given as follows

$$\widehat{\beta} = H^{\dagger}_{avg}T = (H^T_{avg}H_{avg})^{-1}H^T_{avg}T \tag{4}$$

Algorithm 1 The Procedures of SW-ELM

Input: Train data, the number of nodes
Output: $w, b, \widehat{\beta}$

1: Slice time series of a KPI for model training
2: Initialize the network parameters by Oussar and Dreyfus (2000); Nguyen and Widrow (1990)
3: Calculate $H^{\dagger}_{avg}$
4: Obtain $\widehat{\beta}$ by (4)

It should be noted that KPIs' data need to be first normalized to the interval [0,1] by min-max standardization. What's more, before training a KPI's SW-ELM model, a time sliding window needs to be used to slice the time series of the KPI to obtain the input data of the model, and the next data of the time window is used as the predicted data. After trained, SW-ELM can be used to predict their value of next moment through inputing a time series of two KPIs respectively.

2.3 Multi-step Predictions

Due to the need for the long-term prediction of the engine degradation process, it is necessary to predict KPIs in multiple steps. Five long-term prediction methods proposed in Gouriveau and Zerhouni (2012) include iterative, parallel, direct, multiple-input several multiple-outputs and DirRec prediction. Among them, the iterative method is the most common and convenient one. It takes the predicted value of the previous moment as the input of the next-time prediction process. Due to its easy implementation with fewer constraints, the iterative multi-step prediction method is adopted in this paper. The schematic diagram is shown in Fig.2. $\{f^1, [\theta^1]\}$ represents a SW-ELM prediction model, p is the length of the input data.

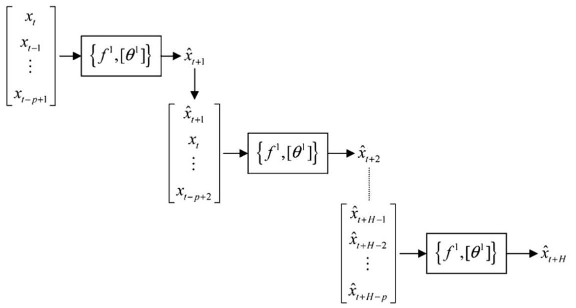


Fig. 2. Iterative model (Gouriveau and Zerhouni (2012))

Based on SW-ELM prediction model, the iterative multi-step prediction method is used to obtain predicted values of two KPIs over a long period of time.

2.4 Threshold-based Degradation State Estimation

The collected sensor data are time series data without labels during the engine life cycle. To cope with this problem, unsupervised classification methods can be used to classify unlabeled data, and the clustering results are then used to judge the degradation state of the system. In this paper, KPIs are clustered according to an unsupervised classification algorithm called subtraction-maximum entropy fuzzy clustering (S-MEFC) proposed in Javed et al. (2013). The algorithm firstly obtains the initial clustering center by subtractive clustering, and then optimizes the initial clustering center by Li and Mukaidono (1995). S-MEFC is summarized in Algorithm 2

Since the two KPIs are monotonic during degradation of the engine, the last class can be considered as the fault state (see Fig.3(a)). After obtaining the cluster centers, this paper proposes a novel method to judge whether predicted data of KPIs could reach the fault state. According to the conservative estimation strategy, when one of the KPIs reaches the corresponding failure threshold in Fig.3(b), the prediction process is terminated and the RUL can be calculated.

2.5 RUL Prediction Model

By summarizing the above methods, the RUL prediction model of aircraft engines can be built. The algorithmic process of the predictive model is shown in Algorithm 3. Its schematic diagram is shown in Fig.4. It shuold be noted that the two SW-ELMs are respectively used to predict the multi-step values of two KPIs by iterative prediction method. In this process, prediction and judgment are carried out simultaneously. In other words, after the values

---

Algorithm 2 The Procedures of S-MEFC

Input: Training data of KPIs, fuzzy parameters σ and end threshold ϵ

Output: Cluster centers V^{new}

1: Acquire c cluster centers by subtractive clustering $V^{old} = \{V_j^{old}\}_{j=1}^{c}$(Chiu (1994))

2: Compute membership matrix U (Li and Mukaidono (1995))

$$\mu_{ij} = \frac{e^{-D^2_{SE_{ij}}/2\sigma^2}}{\sum_{k=1}^{c} e^{-D^2_{SE_{ik}}/2\sigma^2}} \quad \forall i, j \tag{5}$$

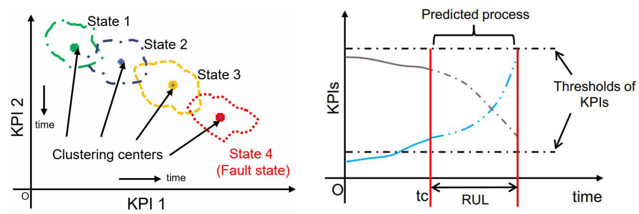
where $D^2_{SE_{ij}}$ and $D^2_{SE_{ik}}$ are the squares of the Euclidean distance from the i-th point to the j-th point and the k-th cluster center, respectively.

3: Adjust cluster centers

$$V_j^{new} = \frac{\sum_{i=1}^{N} \mu_{ij} \cdot x_i}{\sum_{i=1}^{N} \mu_{ij}} \quad \forall j \tag{6}$$

4: Repeat Step 2 and Step 3 until the end condition is met

$$\| v^{new} - v^{old} \| < \epsilon \tag{7}$$



(a) Clustering diagram (b) The judgment process

Fig. 3. Threshold-based degradation state estimation

of KPIs are predicted at one moment, the prediction process should be stopped if one of predicted data of KPIs reaches the corresponding failure threshold.

Therefore, for a given piece of sensor data before time t, the model can obtain the KPIs by PCA, then use SW-ELMs to predict the system's degradation process, and finally give the RUL of the system at time t.

Algorithm 3 RUL Prediction Model

Input: Train set and test set

Output: RUL of engines

1: Preprocess data and obtain train and test data of two KPIs.

2: Train a SW-ELM model for two KPIs respectively after slicing the training data.

3: Cluster the training data and obtain failure thresholds for KPIs.

4: Provide one engine's initial time series of KPIs during testing.

5: Use trained SW-ELMs to predict KPIs' degradation data by iterative method and determine when the predicted data reache the failure thresholds.

6: Calculate the RUL of the engine.

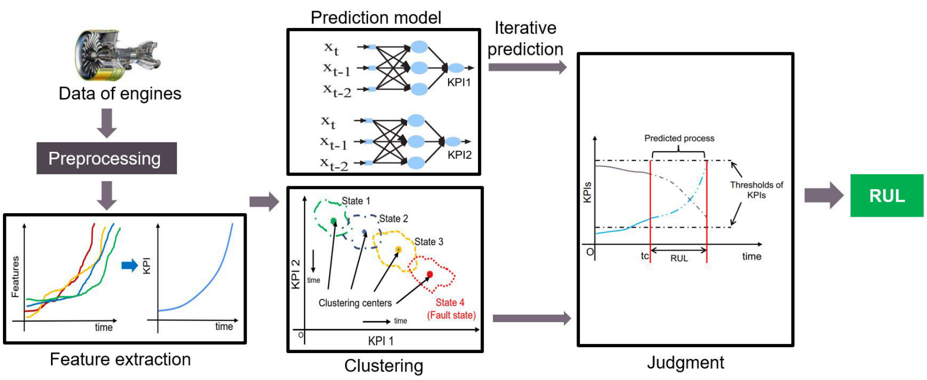


Fig. 4. RUL prediction model

3. EXPERIMENTS

3.1 Data Introduction

The aircraft engine degradation data (Saxena et al. (2008)), provided by NASA, is used to verify the proposed method. These data are generated by simulating turbofan engines degradation process. In this paper, the dataset from the first failure mode (FD001) includes 21 kinds of sensor data for the life cycle of the engines with different initial conditions. The training data contain complete life cycle data of 100 engines fault evolutions, while the test data only give initial evolution data for 100 engines over a period of time. In addition, the RULs of engines corresponding to the test data are provided for comparative analysis of the final predicted results.

3.2 Data Preprocessing

Firstly, all sensors are numbered from 1 to 21, and then the data are filtered with moving average filter of 50 lengths to reduce the noise impact. In order to select variables with clear trends of change, these sensor variables are selected with the standard deviation greater than 0.01. In general, the sensor variables of engines with a certain trend can indicate the degradation process well. However, some variables are sometimes increasing and sometimes decreasing, which could have a bad impact on prediction process. Therefore, based on their trend statistics in train set, these variables should be removed with two trends. Finally, the sensor variables, numbered 2, 3, 4, 7, 8, 11, 12, 13, 15, 17, 20, 21, are screened out. They are used to calculate the values of two KPIs.

3.3 Prediction Results

After feature extraction, the information retention rates can reach over 95% for the principle components of two types of sensor data, which can verify that the two KPIs can be used to characterize the degradation process. All parameters are shown in Table 1. Note that n_s is (input,hidden,output nodes) for SW-ELM. It shows that SW-ELM can predict a KPI of the next moment through a KPI time series of length 26, and n_s of the two KPIs' SW-ELM are the same. The others are parameters of S-MEFC. The clustering results can be obtained by clustering two KPIs of 100 engines of train set (see Fig.5). KPI1 is monotonically decreasing over time, while KPI2 is opposite. In the figure, these clusters can be distinguished well. Therefore, the last cluster center in the lower right corner can be set as the failure thresholds that are 5.6557 and -4.0773.

Table 1. Parameter list

Parameters	n_s	r_a	σ	ϵ
Values	(26,43 1)	0.78	0.68	0.1

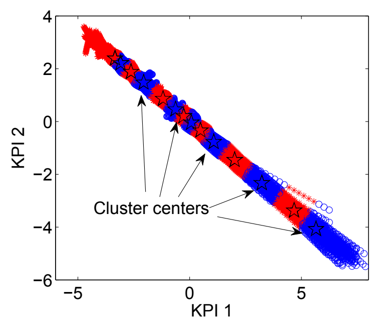


Fig. 5. Cluster results

The RUL prediction process for the 9th test engine is shown in Fig.6. When KPI2 reaches its threshold, degradation prediction process is terminated and then the RUL is equal to the length of the predicted data.

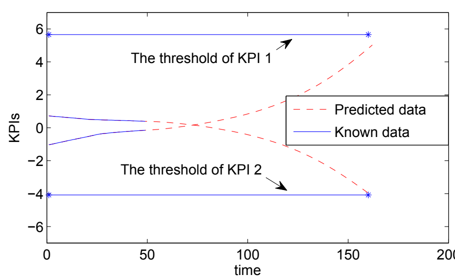


Fig. 6. The prediction process of 9th test engine

3.4 Result Analysis

The trained model is used to predict the RULs of 100 test engines. The predicted results are shown in Fig.7. It is obvious that the model can deliver accurate RUL prediction for most engines. In addition, when the real RUL is small, the prediction result is more accurate. This

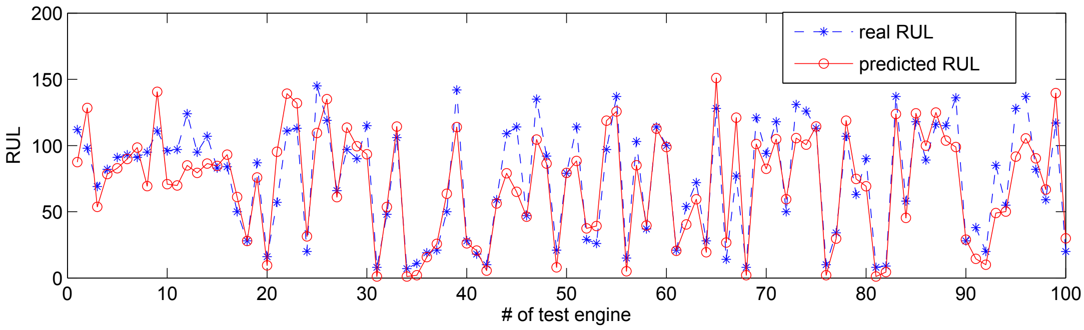


Fig. 7. The predicted RUL of 100 test engines

is because the error caused by the iterative process will be smaller for RUL prediction of a short-life engine.

To evaluate the prediction results, an evaluation indicator was proposed in Saxena et al. (2008). In general, the score increases exponentially with the increasing error, the trend can be approximately described by the following scoring functions.

$$s = \begin{cases} \sum_{i=1}^{n} e^{-\left(\frac{d}{a_1}\right)} - 1 & when \quad d < 0 \\ \sum_{i=1}^{n} e^{\left(\frac{d}{a_2}\right)} - 1 & when \quad d \geq 0 \end{cases} \tag{8}$$

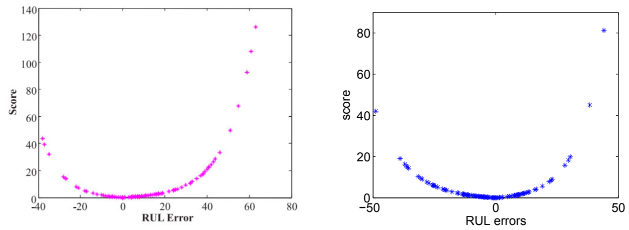
where s is the score and n is the number of test engines, $d = \widehat{t_{\text{RUL}}} - t_{RUL}$ (predicted RUL − real RUL), $a_1 = 13$, $a_2 = 10$. In addition, the common evaluation indicator for predictive models is as follows.

$$R^2 = 1 - \frac{\sum_i \left(RUL_i - \widehat{RUL_i}\right)^2}{\sum_i \left(RUL_i - \overline{RUL_s}\right)^2} \tag{9}$$

where RUL_i is the ith predicted RUL, $\widehat{RUL_i}$ is the real RUL. $\overline{RUL_s}$ is the average of the real RUL. The closer R^2 is to 1, the better the model predictive performance is.

The results of two papers can be used for comparative analysis here. In Ramasso et al. (2012), an EVIPRO-KNN algorithm was used to predict RUL but without a score curve. In Javed et al. (2015), predictability based features selection was used to improve predictive performance with a score curve (see Fig.8(a)). The score curve of the proposed method is compared with the one in Javed et al. (2015) (see Fig.8(b)). As can be seen from Fig.8, the prediction errors of proposed model is more concentrated around zero. More importantly, it has fewer positive errors and smaller maximum positive error. The positive error indicates that the predicted value is greater than the true value. If this condition always happens, it could affect management's early maintenance for the engine and cause more serious consequences. Therefore, this proposed model has better security.

Moreover, other evaluation indicators are shown in Table 2. Note that *paper*1 and *paper*2 represent Ramasso et al. (2012) and Javed et al. (2015), respectively. It can be found that the error interval of the proposed model in this paper is significantly smaller. Compared with *paper*2, the total score of the proposed model can be reduced by half, and its R^2 is higher. These show that the prediction



(a) Method by Javed et al. (2015) (b) Method by this paper

Fig. 8. Scores for 100 test engines

results of the proposed model are more accurate. In addition, compared with *paper*2's time consumption of several minutes, the model is able to complete the operation in seconds, which shows that the model has better real-time performance. Therefore, these analysis results can prove that the proposed novel model can deliver better predictive performance with faster computation speed.

Table 2. Analysis of prediction results

Criteria	*paper*1	*paper*2	this paper
Interval of RUL errors	[-85,120]	[-39,60]	[-49,48]
Total score	None	1046	515
R^2	None	0.641	0.807
Time	None	3m 54sec	6sec

In order to further determine the stability of the prediction model, this paper conducted 10 tests on the prediction model and evaluated the prediction results according to the scoring function. The scores are shown in Fig.9. It can be seen that the total scores are within the interval [500,800] and the fluctuation range is about 300 points. The scores would change greatly with small error fluctuations because the scoring function is exponentially related with the prediction error. Therefore, in fact, the fluctuation of the prediction error is not obvious. The proposed model can deliver stable prediction results.

4. CONCLUSION

In this paper, an SW-ELM based remaining useful life prognosis approach is proposed for aircraft engines. Firstly, two KPIs are extracted from sensor data by PCA. Then a series of methods are used to predict the degradation process of system KPIs and the value of RUL, such as SW-ELM, iterative prediction and S-MEFC. Finally, the prediction model was tested using turbofan engine data

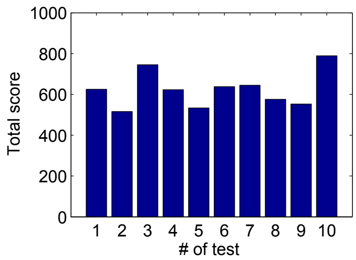


Fig. 9. Total score of 10 tests

from PHM challenge 2008. Compared with the simulation results from other literatures, it can be proved that the proposed model can deliver superior prediction performance with faster computation and higher stability.

In future work, other feature selection or extraction methods can be considered, such as recursive feature elimination can be used to screen out more important variables. Furthermore, deep learning methods can be considered for variable degradation based RUL prognostic in the future.

ACKNOWLEDGEMENTS

This work is supported by the National Defense Basic Scientific Research Program of China under Grant JCKY2017212C005.

REFERENCES

Ali, J.B., Chebel-Morello, B., Saidi, L., Malinowski, S., and Fnaiech, F. (2015). Accurate bearing remaining useful life prediction based on weibull distribution and artificial neural network. *Mechanical Systems and Signal Processing*, 56, 150–172.

Chiu, S.L. (1994). Fuzzy model identification based on cluster estimation. *Journal of Intelligent & fuzzy systems*, 2(3), 267–278.

Gebraeel, N., Lawley, M., Liu, R., and Parmeshwaran, V. (2004). Residual life predictions from vibration-based degradation signals: a neural network approach. *IEEE Transactions on industrial electronics*, 51(3), 694–700.

Gouriveau, R. and Zerhouni, N. (2012). Connexionist-systems-based long term prediction approaches for prognostics. *IEEE Transactions on Reliability*, 61(4), 909–920.

Huang, G.B., Zhu, Q.Y., Siew, C.K., et al. (2004). Extreme learning machine: a new learning scheme of feedforward neural networks. *Neural networks*, 2, 985–990.

Javed, K., Gouriveau, R., and Zerhouni, N. (2013). Novel failure prognostics approach with dynamic thresholds for machine degradation. In *IECON 2013-39th Annual Conference of the IEEE Industrial Electronics Society*, 4404–4409. IEEE.

Javed, K., Gouriveau, R., and Zerhouni, N. (2014). Sw-elm: A summation wavelet extreme learning machine algorithm with a priori parameter initialization. *Neurocomputing*, 123, 299–307.

Javed, K., Gouriveau, R., and Zerhouni, N. (2015). A new multivariate approach for prognostics based on extreme learning machine and fuzzy clustering. *IEEE transactions on cybernetics*, 45(12), 2626–2639.

Jouin, M., Gouriveau, R., Hissel, D., Péra, M.C., and Zerhouni, N. (2016). Particle filter-based prognostics: Review, discussion and perspectives. *Mechanical Systems and Signal Processing*, 72, 2–31.

Li, R.P. and Mukaidono, M. (1995). A maximum-entropy approach to fuzzy clustering. In *Proceedings of 1995 IEEE International Conference on Fuzzy Systems.*, volume 4, 2227–2232. IEEE.

Li, X., Ding, Q., and Sun, J.Q. (2018). Remaining useful life estimation in prognostics using deep convolution neural networks. *Reliability Engineering & System Safety*, 172, 1–11.

Nguyen, D. and Widrow, B. (1990). Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. In *1990 IJCNN International Joint Conference on Neural Networks*, 21–26. IEEE.

Oussar, Y. and Dreyfus, G. (2000). Initialization by selection for wavelet network training. *Neurocomputing*, 34(1-4), 131–143.

Ramasso, E., Rombaut, M., and Zerhouni, N. (2012). Joint prediction of continuous and discrete states in time-series based on belief functions. *IEEE transactions on cybernetics*, 43(1), 37–50.

Saxena, A., Goebel, K., Simon, D., and Eklund, N. (2008). Damage propagation modeling for aircraft engine run-to-failure simulation. In *2008 international conference on prognostics and health management*, 1–9. IEEE.

Si, X.S., Wang, W., Hu, C.H., and Zhou, D.H. (2011). Remaining useful life estimation–a review on the statistical data driven approaches. *European journal of operational research*, 213(1), 1–14.

Sikorska, J., Hodkiewicz, M., and Ma, L. (2011). Prognostic modelling options for remaining useful life estimation by industry. *Mechanical systems and signal processing*, 25(5), 1803–1836.

Xu, J., Wang, Y., and Xu, L. (2013). Phm-oriented integrated fusion prognostics for aircraft engines based on sensor data. *IEEE Sensors Journal*, 14(4), 1124–1132.

Yam, R., Tse, P., Li, L., and Tu, P. (2001). Intelligent predictive decision support system for condition-based maintenance. *The International Journal of Advanced Manufacturing Technology*, 17(5), 383–391.

Yang, H. and Yin, S. (2019). Reduced order sliding mode observer-based fault estimation for markov jump systems. *IEEE Transactions on Automatic Control*.

Yin, S., Ding, S.X., and Zhou, D. (2016). Diagnosis and prognosis for complicated industrial systemspart i. *IEEE Transactions on Industrial Electronics*, 63(4), 2501–2505.

Yin, S., Rodriguez-Andina, J.J., and Jiang, Y. (2019). Real-time monitoring and control of industrial cyber-physical systems: With integrated plant-wide monitoring and control framework. *IEEE Industrial Electronics Magazine*, 13(4), 38–47.

Yuan, M., Wu, Y., and Lin, L. (2016). Fault diagnosis and remaining useful life estimation of aero engine using lstm neural network. In *2016 IEEE International Conference on Aircraft Utility Systems (AUS)*, 135–140. IEEE.

Zhang, X., Xu, R., Kwan, C., Liang, S.Y., Xie, Q., and Haynes, L. (2005). An integrated approach to bearing fault diagnostics and prognostics. In *Proceedings of the 2005, American Control Conference, 2005.*, 2750–2755. IEEE.