

A Prediction Method for Aero-engine Health Management Based on Nonlinear Time Series Analysis

Qiang Huang, Haixia Su, Jian Wang*, Weixing Huang, Guigang Zhang
. Hi-tech Innovation Center
Institute of Automation, Chinese Academy of Science
Beijing, China
jian.wang@ia.ac.cn

Jiayang Huang
Shanghai Engineering Research Center of Civil Aircraft
Health Monitoring
Shanghai, China
huangjiayang@comac.cc

Abstract—Aero-engine is the heart of the aircraft. If a failure of aero-engine occurs during the flight, it will be a direct threat to flight safety of the aircraft, so the aero-engine health management came into being, and the prediction is a very important part of it. This paper was focused on the prediction methods of health management. Firstly, we introduced the research status of the aero-engine prediction methods, and then proposed a prediction method of nonlinear time series analysis using C-C method and BP-Adaboost algorithm, at last, a simulation example was given to illustrate the validity of the method. Experimental results indicated that the method has the advantage of high prediction precision, and to some extent, it can provide a reference for the maintenance plan.

Keywords- Aero-engine; Prediction; C-C method; BP-Adaboost algorithm

I. INTRODUCTION

At present, the aviation field is facing severe cost pressures. The United States government report shows that the costs of aircraft operation and support may be more than ten times the initial price. In this context, we need a new aircraft platform to reduce the aircraft operating costs and improve the accuracy of the aircraft maintenance. Therefore, Prognosis and Health Management (PHM) came into being.

PHM is to ensure continuous monitoring and assessing the functional health of the aircraft, to predict failure or the components close to the service life, and to use this information to improve operational decision-making, increase the safety and reliability of the system, reduce the maintenance costs of the system, enhance the accuracy of maintenance and make the condition based maintenance come true.

Aero-engine is the core power system of the aircraft, so its security and reliability are related to the flight safety. Especially for the civil aviation aircraft carrying numerous passengers, it is significant to discover potential safety hazards and predict the impending failure for ensuring the life safety of the passengers and avoiding great economic losses.

There are a lot of prediction methods of aero-engine, but they can be divided into three categories: data-driven modeling, physics-based modeling and knowledge-based reasoning. In data-driven modeling, the prediction system can directly obtain measurement data from the sensors embedded

in each system, process the data, and predict the future data with the underlying principles mined by data analysis methods from the extracted feature data. BP (Back Propagation) neural network [1], RBF (Radial Basis Function) neural network [2], relevance vector machine [3], support vector machine [4], time series analysis belong to data-based methods. In physics-based modeling, the predictive function of the system depends on model analysis and reasoning. The physics-based modeling can describe the nature of the object, but it is difficult to establish accurate mathematical/physical model, which limits the application of these methods to some extent. The knowledge-based reasoning does not need to establish accurate mathematical/physical model. They combine the experience and knowledge of experts in multiple disciplines to build the prediction system. At present, the mainly application forms of knowledge-based reasoning are expert system and fuzzy logic. As the expert system can't break through the bottlenecks of knowledge acquisition and knowledge representation, the method is limited in practical applications.

Since the applications of physics-based modeling and knowledge-based reasoning have some limitations, data-driven modeling is used more and more widely in the prediction field. In view of this, this paper adopts the method of nonlinear time series analysis (data-driven modeling) to predict the wear condition of aero-engine and analyzes the results to verify the validity of the method.

II. MAIN CONTENTS AND RESEARCH STATUS OF AERO-ENGINE PREDICTION

The research contents of aero-engine prediction can be summarized as performance prediction, failure prediction and residual life prediction [5,6]. Performance prediction is the process of "no objective" to predict the performance of the engine performance deterioration, which runs through the whole life of the engine, and it is the basis for realization of failure prediction and remaining life prediction. Failure prediction works when performance degradation of engine occurs, which can predict the time point and mode of the failure occurrence by judging engine's property tendency. Residual life prediction, which is based on the performance prediction, assesses how much time the engine will appear functional failure at current state by analyzing the results of failure prediction. Fault prediction and residual life prediction

are two specific and objective extensions of the performance prediction, and they analyze the possible failure or the remaining useful life based on the performance prediction. Accordingly, performance prediction, fault prediction and residual life prediction are both independent and mutually influenced in some aspects.

A. Research Status of Aero-engine Performance Prediction

For aero-engine overall performance prediction, performance parameters directly reflect its performance, so it is important to build a good model for evaluating the performance degradation degree of aero-engine. Rolls-Royce used Kalman filter and least squares method to achieve the analysis and prediction of gas path performance parameter. Wang Z X [7] studied turbofan engine rotating components and overall performance degradation based on physical failure model. Song Y X [8] established the prediction model of engine performance parameters by using multiple linear regression.

In order to obtain useful information from historical data and realize the meaningful performance prediction, a lot of scholars have carried out the research of time series analysis [9]. In this paper, the prediction method is the method of time series analysis. Firstly, the C-C method is used to reconstruct the phase space, and then the BP-Adaboost algorithm is used to predict the wear condition of aero engine. Detailed algorithm is introduced later.

B. Research Status of Aero-engine Failure Prediction

According to the history and current status, failure prediction predicts the future state of aero-engine, sounds the alarm when the engine is in abnormal state. Brotherton [10] combined the neural network and genetic algorithm to establish the failure prediction system of the aero-engine gas path. Zhao H L [11] evaluated the fault possibility of each part of the engine in the future by Monte Carlo simulation technology.

C. Research Status of Aero-engine Residual Life Prediction

As the increase of service time, the aero-engine will go through the process of performance degradation until disabler. The process from the current state of the engine to the functional failure state is called the remaining life. At present, the main prediction methods are based on the failure physical model/performance degradation model and the method based on machine learning. P.J. García Nieto [12] realized the residual life prediction of aero-engine by hybrid PSO-SVM-based method. Jones [13] proposed that the prediction of residual life of engine under the Bayesian framework.

III. PREDICTION MODEL OF NONLINEAR TIME SERIES BASED ON C-C METHOD AND BP-ADABOOST ALGORITHM

A. C-C Method

In order to obtain more useful information from the time series, the phase space reconstruction of the time series data is needed. The time delay τ and embedding dimension m are two important parameters for phase space reconstruction [14, 15].

For small data sets, it is very effective to use the C-C method with the correlation integral to estimate the time delay τ and the time window length τ_ω ($\tau_\omega = (m-1)\tau_d$) [15, 16, 17].

According to the theorem of Takens [17], the phase space is reconstructed as:

$$X = \{X_i = [x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau}], i=1, 2, 3, \dots, M\} \quad (1)$$

where τ is the time delay, m is the embedding dimension, N is the length of time series, and $M = N - (m-1)\tau$ is the number of points in the phase space.

The C-C method uses the correlation integral to obtain the time delay τ and embedding dimension m , and the correlation integral is described as follows. [16, 17]

$$C(m, n, r, t) = \frac{2}{M(M-1)} \sum_{1 \leq i \leq j < M} \theta(r - d_{ij}), r > 0 \quad (2)$$

$$d_{ij} = |X_i - X_j| \quad (3)$$

$$\theta(r - d_{ij}) = \begin{cases} 0 & r - d_{ij} < 0 \\ 1 & r - d_{ij} > 0 \end{cases} \quad (4)$$

The correlation integral is a cumulative distribution function, and its specific meaning is the probability that the distance between any two points in the phase space is less than r . Define statistical detection:

$$S(m, N, r, t) = C(m, N, r, t) - C^m(1, N, r, t) \quad (5)$$

The time sequence is divided into t equal dimension sub sequences, and the t is the reconstruction time delay, as:

$$\begin{cases} x(1) = \{x_1, x_{i+1}, \dots, x_{\lfloor N/t \rfloor - t + 1}\} \\ x(2) = \{x_2, x_{i+2}, \dots, x_{\lfloor N/t \rfloor - t + 2}\} \\ \dots \\ x(t) = \{x_t, x_{i+t}, \dots, x_{\lfloor N/t \rfloor - t + t}\} \end{cases} \quad (6)$$

To calculate the statistics of each time sub sequence with $N \rightarrow \infty$:

$$S_1(m, r, t) = \frac{1}{t} \sum_{s=1}^t [C_s(m, r, t) - C_s^m(1, r, t)] \quad (7)$$

According to the corresponding maximum and minimum values of the radius r , the difference is obtained as:

$$\Delta S_1(m, t) = \max\{S_1(m, N, r_j, t)\} - \min\{S_1(m, N, r_j, t)\} \quad (8)$$

The first zero crossing of $\Delta S_1(m, t)$ is the optimal time delay τ_d . According to the principle of asymptotic distribution, let $m = 2, 3, 4, 5$, $r_j = \frac{i\sigma}{2}$ (σ is the standard deviation of the sequence), $i = 1, 2, 3, 4$, then we can obtain:

$$\bar{S}_1(t) = \frac{1}{16} \sum_{m=2}^5 \sum_{j=1}^4 S_1(m, r_j, t) \quad (9)$$

$$\Delta \bar{S}_1(t) = \frac{1}{4} \sum_{m=2}^5 \Delta S_1(m, t) \quad (10)$$

$$S_{1cor}(t) = \Delta \bar{S}_1(t) + |\bar{S}_1(t)| \quad (11)$$

The first local minimum of $S_{1cor}(t)$ is the time delay window τ_ω . We can obtain the embedding dimension m by $\tau_\omega = (m-1)\tau_d$.

B. BP-Adaboost

Adaboost algorithm takes a weighted method, increases the weights of weak predictors with smaller training error and reduces the weights of weak predictors with larger training error, so several weak predictors are combined effectually. BP-Adaboost algorithm uses some BP networks as weak predictors and groups the weak predictors to form a strong predictor to improve the predicting precision [18]. The algorithm flow chart is shown in Fig. 1.

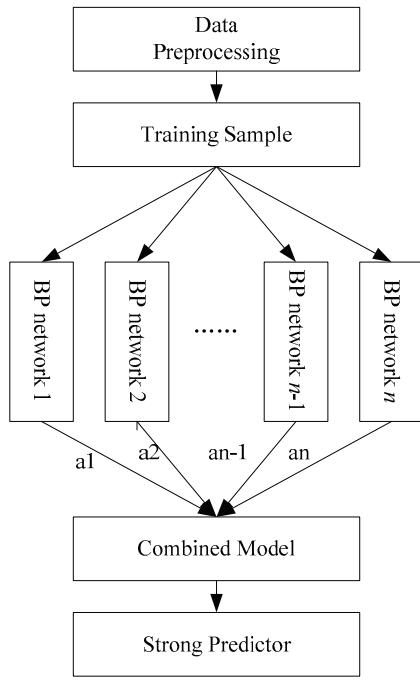


Figure 1. Flow chart of BP-Adaboost algorithm

C. Prediction Model of Nonlinear Time Series Based on C-C Method and BP-Adaboost Algorithm

As a result, the model combined with C-C method and BP-Adaboost algorithm is proposed in this paper for the nonlinear time series prediction of small data. Firstly, the C-C method is used to reconstruct phase space. Secondly, BP-Adaboost algorithm is applied to train the reconstructed samples. After the two steps, the strong predictor is formed. Prediction model is shown in Fig. 2.

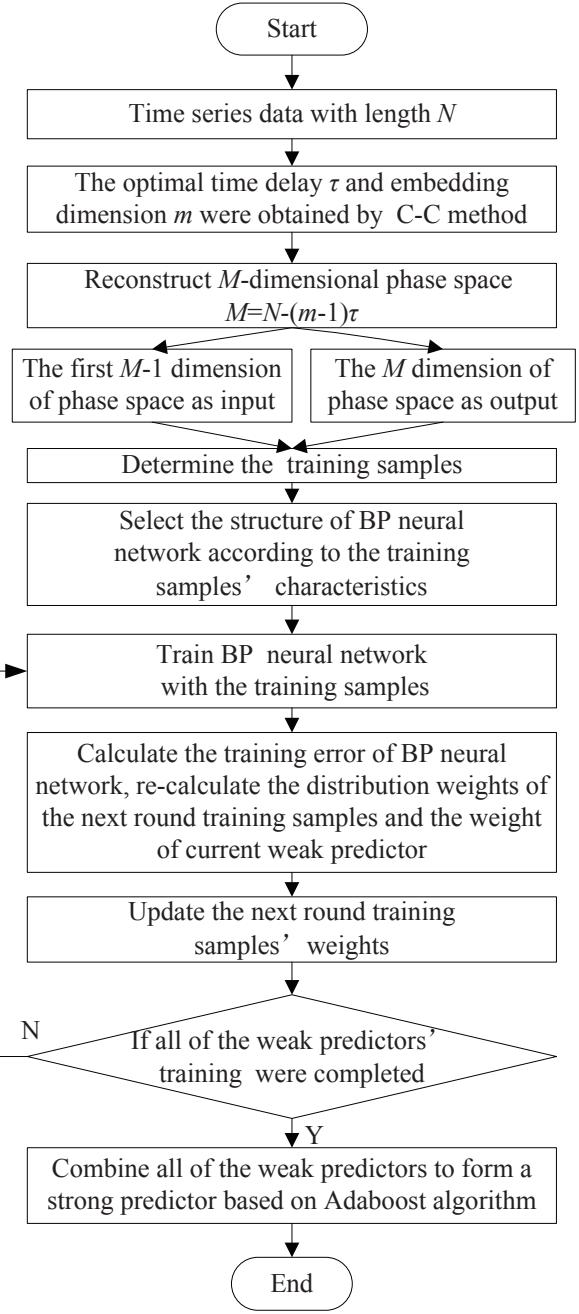


Figure 2. Prediction model of nonlinear time series based on C-C method and BP-Adaboost algorithm

IV. PREDICTION OF AERO-ENGINE WEAR CONDITION BASED ON C-C METHOD AND BP-ADABOOST ALGORITHM

Data comes from oil spectral analysis of a certain type of aero-engine from one to another oil change, and sorted by time (68). At present, the aero-engine wear condition's analysis is mainly by monitoring the Fe content, so only the Fe content time series is analyzed in this paper. The experiment process was shown in Fig. 3.

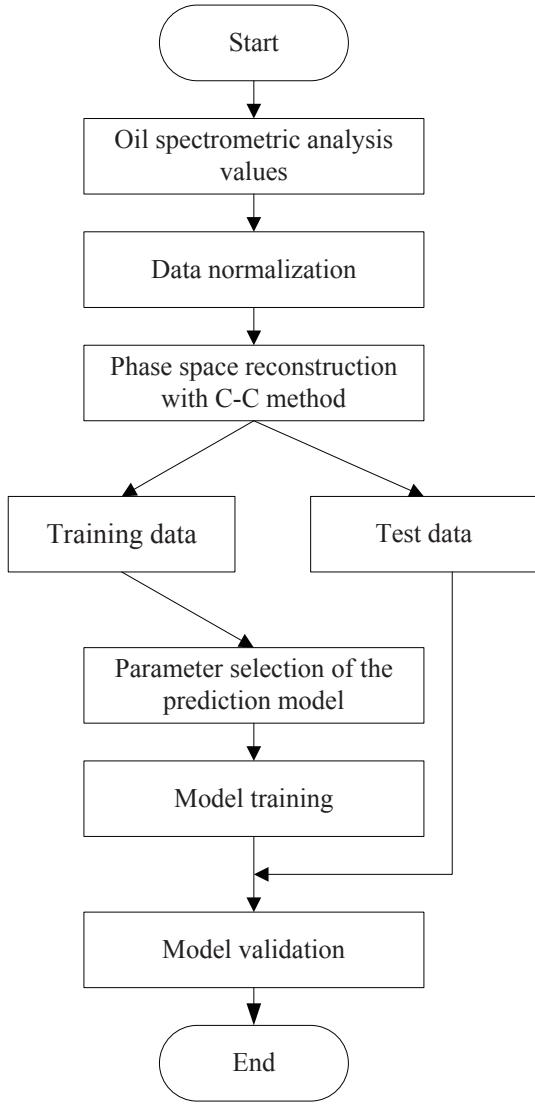


Figure 3. diagram of experiment process

A. Data Preprocessing

The time series data of Fe content is shown in Fig. 4. The data is normalized to improve the modeling accuracy and the method of normalization is shown as follows:

$$x(i) = s + (l - s)(x(i) - x_{\min}) / (x_{\max} - x_{\min}) \quad (12)$$

The $x(i)$ on the left side of the equal sign is the normalized data. The x_{\max} and x_{\min} are the maximum and minimum values of the original data; the $x(i)$ on the right side of the equal sign is the original data; the s and l are the maximum and minimum values of the normalized data. The Fe content time series $\{x_i\}$ is obtained after normalization, and the data can be restored to the original range by anti-normalization.

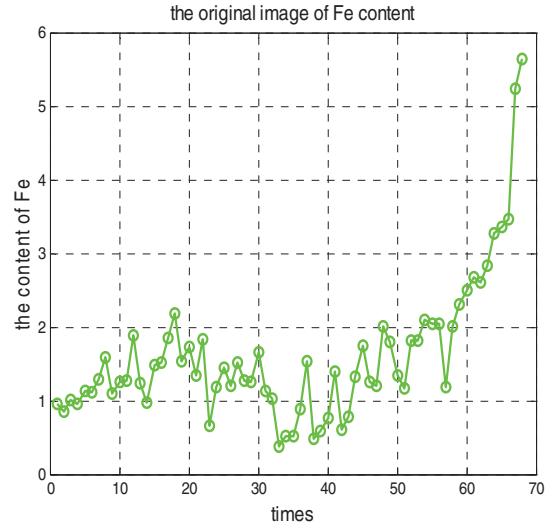


Figure 4. original data curve

B. Phase Space Reconstruction

The time delay τ and embedding dimension m are two important parameters for phase space reconstruction. The results shown in Fig.5 are obtained by C-C method.

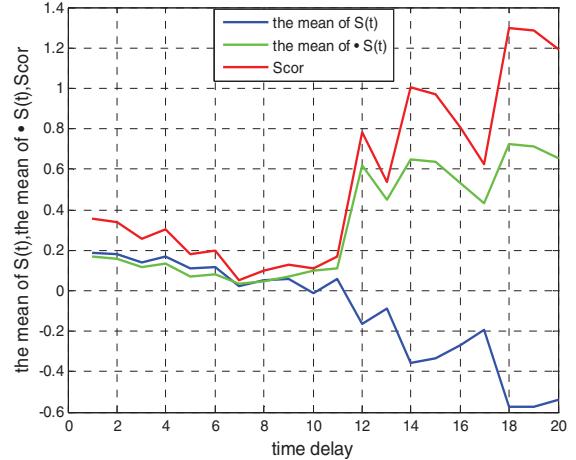


Figure 5. the results from C-C method

In Fig. 5, the first tiny point of $\Delta S(t)$ appears at $t = 3$, and the global minimum of S_{cor} appears at $t = 7$, so the optimal time delay $t_d = 3$ and the time window length $t_w = 7$ are gained by C-C method. It is easy to get that $\tau = 3$ and $m = 3$ by formula $\tau_w = (m-1)\tau_d$.

After getting the optimal time delay τ and embedding dimension m , the original data can be reconstructed as:

$$X = \left\{ X_i = [x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau}] \right\}, i = 1, 2, 3, \dots, M \quad (13)$$

From embedding theorem, there is a smooth mapping $\hat{f} : R^m \rightarrow R^m$ in the reconstructed phase space, so

$$X_{i+h} = \hat{f}(X_i) \quad (14)$$

where the X_{i+h} is the h -step evolution state of the X_i .

In this paper, $h = 1$, so the inputs of the weak predictors are $X_i = [x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau}]$ and the outputs are the m dimension of the next state point, i.e. $Y_i = x_{i+(m-1)\tau+1}$, $i = 1, 2, 3, \dots, N - (m-1)\tau - 1$. According to the above method, the 61 input and output samples are obtained, and the training samples are from 1 to 55, and the test samples are the left 56 to 61.

C. Parameter Selection of The Prediction Model

In this paper, BP neural networks are used as weak predictors. It is necessary to determine the number of input layer nodes, the number of hidden layers and nodes, and the number of output layer nodes, etc. According to the input/output samples' dimension, the number of input layer nodes and output layer nodes of every BP neural network are selected respectively as 3 and 1. The number of hidden layers is selected as 1 and the number of hidden layer nodes are chosen as 10 by formula $m = \sqrt{n+l} + \alpha$ (n is the number of input layer nodes, l is the number of output layer nodes, α is a constant between 1 and 10). The transfer function between the input layer and the hidden layer is logsig function, and the transfer function of the hidden layer to the output layer is chosen as the purelin transfer function, and the training function is trainlm function.

The Adaboost algorithm is adopted in this paper to combine BP weak predictors into a strong predictor, and the number of weak predictors is 10.

D. Results

The above algorithm is achieved by MATLAB, and the results are presented in Fig. 6.

As the results presented in Fig. 6, both for training samples and test samples, BP-Adaboost algorithm shows strong fitting and generalization ability, and high prediction precision. It can not only predict the change trend of Fe content in the oil, but also can predict the value of Fe content in the oil with high accuracy.

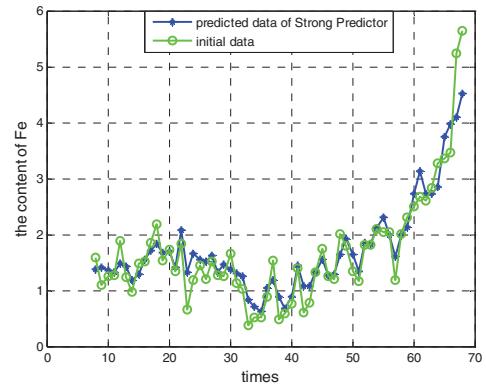


Figure 6. the results from the prediction model

In order to illustrate the superiority of the BP-Adaboost algorithm, this article compares the output of stronger predictor with the output of weak predictor, and the compare results are shown in Fig. 7 to Fig. 17.

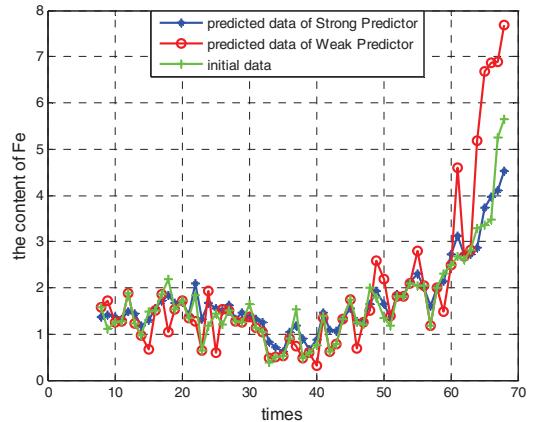


Figure 7. the comparison chart of strong predictor's output and weak predictor 1's output

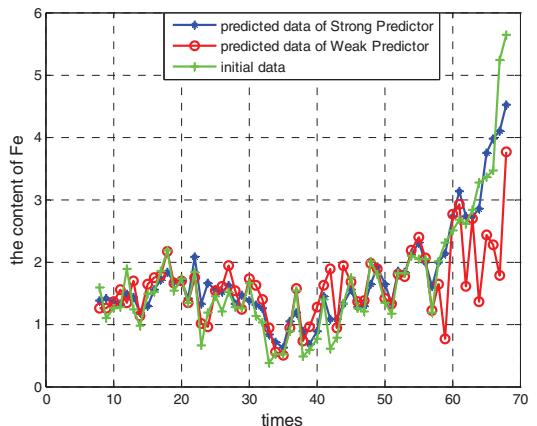


Figure 8. the comparison chart of strong predictor's output and weak predictor 2's output

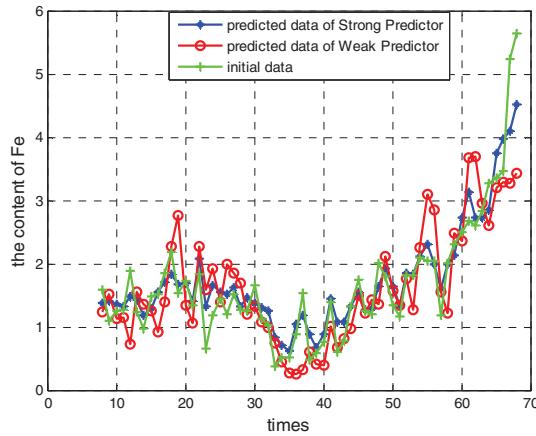


Figure 9. the comparison chart of strong predictor's output and weak predictor 3's output

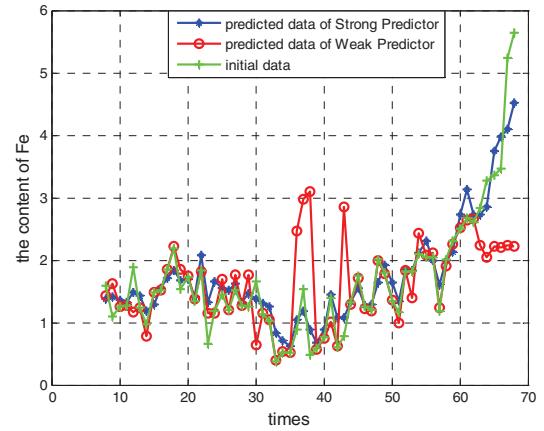


Figure 12. the comparison chart of strong predictor's output and weak predictor 6's output

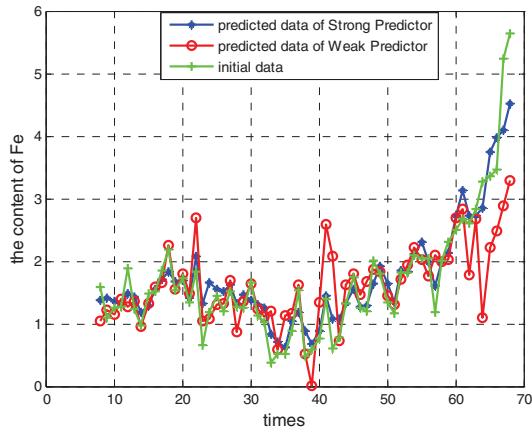


Figure 10. the comparison chart of strong predictor's output and weak predictor 4's output

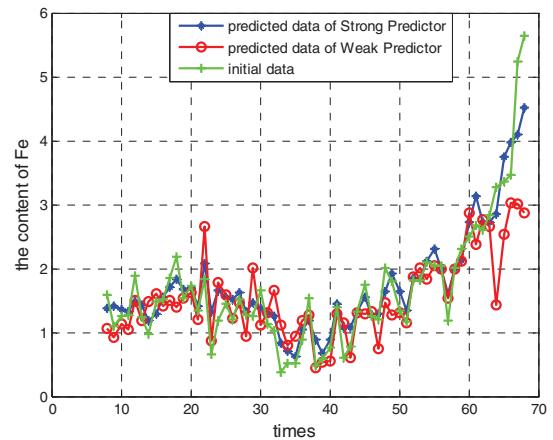


Figure 13. the comparison chart of strong predictor's output and weak predictor 7's output

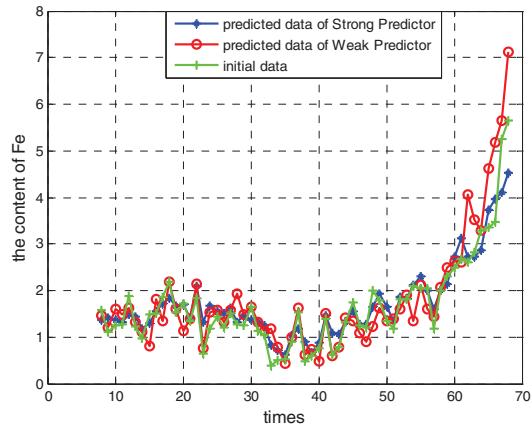


Figure 11. the comparison chart of strong predictor's output and weak predictor 5's output

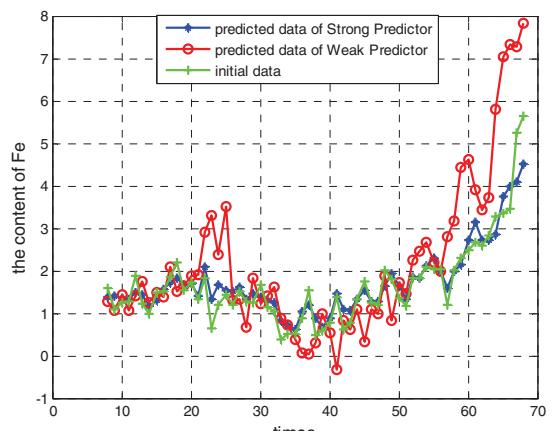


Figure 14. the comparison chart of strong predictor's output and weak predictor 8's output

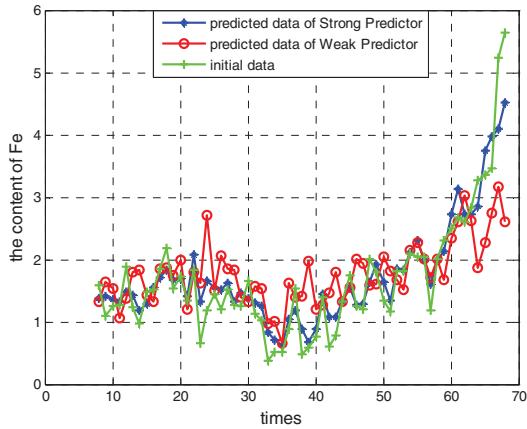


Figure 15. the comparison chart of strong predictor's output and weak predictor 9's output

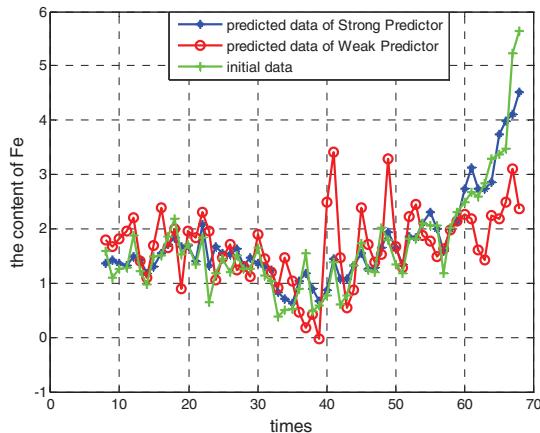


Figure 16. the comparison chart of strong predictor's output and weak predictor 10's output

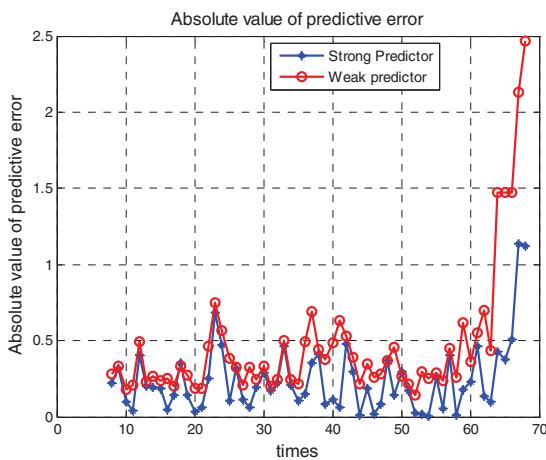


Figure 17. the comparison chart of the prediction error of strong predictor and the average prediction error of weak predictors

From Fig. 7-Fig. 16, it is easy to see that the prediction error of every BP weak predictor is relatively larger, which is related to the characteristic of BP neural network's limited fitting ability. But BP-Adaboost algorithm effectively combines each weak predictor, and greatly improves the prediction accuracy, which is obviously reflected in the Fig. 17. Therefore, BP-Adaboost is efficacious for aero-engine wear condition prediction. From Fig. 17, it can be seen that the prediction error of BP-Adaboost algorithm is increased with the increase of the prediction time. This is related to that small scale time series can only be medium or short-term prediction. From what has been analyzed above, it is concluded that BP-Adaboost algorithm can effectively improve the prediction accuracy of BP neural network, so it is suitable for the medium and short-term prediction of aero-engine wear condition.

V. CONCLUSION

In this paper we have successively reviewed the research status of aero-engine health management prediction methods, introduced the C-C method and BP-Adaboost algorithm, and used the data comes from oil spectral analysis of a certain type of aero-engine from one to another oil change to verify whether the algorithm is effect or not. From the results, we could draw the conclusion that the time series prediction method combined with the C-C method and BP-Adaboost algorithm is very good for aero-engine wear condition prediction. In the future, we need to use the method to predict other performance parameters, and to find and improve the shortcomings of the method. Finally, we hope to see that the method will be applied to the aircraft health management system to achieve fault prediction of the whole machine and condition based maintenance.

ACKNOWLEDGMENT

This research was supported by: (1) Special Project for Civil-Aircraft, MIIT; (2) Natural Science Foundation of Sichuan Province Education Department under Grant 16ZB0035.

REFERENCES

- [1] Goely A K, Saxenax S, Bhanotz S. A fast learning algorith for training feed forword neural networks[J]. International Journal of Systems Science, 2006, 37(10): 709-722.
- [2] Narendra K S. Identification and control for dynamic system using neural network[J]. IEEE Trans on Neural Networks, 1990, 1(1): 4-27.
- [3] Tipping M E. Sparse bayesian learning and relevance vector machine[J]. Journal of Machine Learning Research, 2001, 24(1): 211-244.
- [4] Vapnir. The nature of statistical learning[M]. New York: Springer, 1995.
- [5] Kalgren P W, Byington C S, Roemer M J. Defining PHM, a lexical evolution of maintenance and logistics[C]// Autotestcon, 2006 IEEE. IEEE, 2006: 353-358.
- [6] Sheppard J W, Kaufman M A, Wilmer T J. IEEE standards for prognostics and health management[J]. Aerospace and Electronic Systems Magazine, IEEE, 2009, 24(9): 34-41.
- [7] Wang Zhanxue, Liu Zengwen, Ye Xinnong. Effect of component deterioration on turbofan engine performance[J]. Journal of Aerospace Power, 2007, 22(5): 792-796.
- [8] Song Yunxue, Zhang Kexin, Shi Yongsheng. Research on aeroengine performance parameters forecast based on multiple linear regression

- forecasting method[J]. Journal of Aerospace Power, 2009, 24(2): 427-431.
- [9] Jaw L C, Friend R. ICEMS: a Platform for Advanced Condition-Based Health Management[C]// Aerospace Conference, 2001, IEEE Proceedings, Big Sky, MT, 2001. IEEE, 2001: 2909-2914.
- [10] Brotherton T, Jahns G, Jacobs J, et al. Prognosis of faults in gas turbine engines[C]// Aerospace Conference Proceedings, 2000 IEEE. IEEE, 2000, 6: 163-171.
- [11] Zhao H L, Liu Y W. Forecasting for aero-engine failure risk based on Monte Carlo simulation[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(3): 545-550.
- [12] P.J. García Nieto, E. García-Gonzalo, F. Sánchez Lasheras, et al. Hybrid PSO-SVM-based method for forecasting of the remaining useful life for aircraft engines and evaluation of its reliability [J]. Reliability Engineering and System Safety, 2015, 138: 219-231.
- [13] Jones S M, Reveley M. An overview of the NASA aviation safety program assessment process[J]. 3rd AIAA Aviation Technology, Integration, and Operations (ATIO), AIAA Paper, 2003, 6706.
- [14] Rosenstein M T, Collins J J, De L C, et al. Reconstruction expansion as a geometry-based framework for choosing proper delay times[J]. Nonlinear Phenomena, 1994, 73(1-2):82-89.
- [15] Kim H S, Eykholt R, Sales J D, et al. Nonlinear dynamics, delay times, and embedding windows[J]. Physica D: Nonlinear Phenomena, 1999, 127(1-2):48-60.
- [16] Sprott J C. How common is chaos phys[J]. Lett A, 1993, 173(1): 21-24.
- [17] Sprott J C. Automatic generation of strange attractors[J]. Computers & Graphics, 1993, 17(3): 325-332.
- [18] Jun Wang, Jiayuan Yu. Scientific creativity research based on generalizability theory and BP_Adaboost RT[J]. Procedia Engineering, 2011, 15: 4178-4182.