

ESN-based Combination Method for Turbine Condition Prediction

Yanyan Liu¹, Yangming Guo² and Aihua Wang¹

¹Harbin University of Science and Technology, Harbin 150080, China

²Northwestern Polytechnical University, Xian 710072, China

Abstract—As a nonlinear time series prediction method, echo state network (ESN) attracts more attention because of its good approximation capability for the nonlinear system. Aiming at the characters of nonlinear time series in the Turbine’s condition prediction analysis, such as including noise and presenting chaos, etc. a combination method based on ESN was proposed. Firstly, the noise contained in nonlinear time series was reduced by the wavelet analysis. Then the training sample data were yielded via phase space reconstruction of the time series. After reducing the dimension of the training sample data by principal component analysis, all principal data were sent into the ESN prediction model. An actually dynamic pressure time series of aircraft power were conducted. The experiments compare the proposed method with traditional ESN prediction model on prediction accuracy and time cost. The results show that the proposed method can both effectively improve the prediction accuracy and efficiency, and it is an effective nonlinear time series prediction method in practice.

Keywords—Turbine; health prediction; echo state network (ESN); wavelet analysis; principal components analysis (PCA)

I. INTRODUCTION

Turbine system is the key system to ensure the safe and reliable operation of the aircraft, as well as the completion of the overall flight mission. Therefore, the theoretical research and engineering practice related with the Turbine monitoring and fault diagnosis technology has always been highly focused. In recent years, the technology of fault prediction based on state monitoring has provided a new method for the health status analysis and the maintenance of the engine [1]. However, due to the highly complexity of the Turbine system, it is usually highly difficult to construct a precise mathematical model for the whole system. However, the nonlinear time series is capable of indicating the characteristics of the running engine and affecting the control of the power system. Therefore, it is one of the most wildly adopted methods to predict and analyze the operational and health state of the engine [2].

Prediction based on time series is a statistical method. The developing trend of the data with time is predicted according to certain rules based on the past and current data. This method is widely utilized in the areas of fault prediction and diagnosis, control, and etc[3]. However, the operating environment of the Turbine is rather complicated and adverse; thus good prediction result is not able to be obtained with the adoption of the traditional fault diagnosis and prediction method. In this paper, a technical scheme is proposed which is based on the

combination of several diagnostic approaches. Furthermore, the various characteristics of objects to be diagnosed are also incorporated by the proposed method. Overall, better prediction results could be obtained with the application of the method proposed in this work. Firstly, the obtained time series are pre-processed with the adoption of the wavelet analysis aiming to reduce the noise, then the principal components analysis (PCA) is applied in order to reduce the dimension of the space. After being processed, the data is fed into the Echo State Networks, ESN [4]; through this prediction model, the prediction and analysis process is completed.

In a paper, simulation and analysis are performed for the combined prediction method based on ESN. The analysis here is based on the real observed nonlinear time series of the dynamic pressure for the engine. As indicated by the results, the proposed prediction method is capable of obtaining good predicting results.

II. THE COMBINATION METHOD BASED ON ESN

It is indicated by previous researchers that the prediction model based on ESN is capable of providing good predictions when dealing with nonlinear time series. Nevertheless, a larger amount of simulation time is usually required if this model is adopted; furthermore, the predicting accuracy of the ESN model is likely to be reduced by the noise pollution in the measured data [5]. Thus, in this paper, the wavelet analysis denoising method and principal component analysis (PCA) are combined here in order to improve the predicting performance. The presented combined prediction method in this work is based on the traditional ESN prediction model, and the characteristics of predicting the nonlinear time series are also incorporated. The principle of the combination method based on ESN is presented in Figure 1.

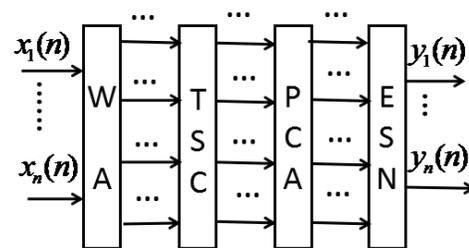


FIGURE I. PRINCIPLE OF ESN-BASED COMBINATION METHOD.

Here, WA: Wavelet Analysis; TSC: Training Sample Construction; PCA: Principal Component Analysis.

A. Predicting Principle of the Nonlinear Time Series for ESN

ESN is a special type of recurrent neural network. A huge dynamic reservoir (DR) is adopted by the ESN which is composed by hundreds of, even thousands of, hidden layer neurons. This DR is utilized to replace the intermediate layer for the traditional recurrent neural network. For the traditional neural network, a number of problems are incurred by the adoption of gradient descent learning algorithm, such as slow convergence, falling into local minimum and complicate training process. Nevertheless, due to the introduction of DR computing model, these problems can be readily avoided [5]. An illustration of the structure of ESN is presented in Figure 2.

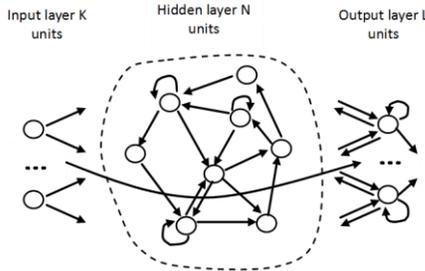


FIGURE II. STRUCTURE OF ESN

For the ESN architecture illustrated in Figure 2, the numbers of input units, hidden layer units and output units are assumed that k , n and l respectively. At time t , the sequences for the states of the input units, the hidden layer and the output units are presented as follow respectively:

$$\begin{cases} \mathbf{u}(t) = [u_1(t), u_2(t), \dots, u_k(t)] \\ \mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_n(t)] \\ \mathbf{y}(t) = [y_1(t), y_2(t), \dots, y_l(t)] \end{cases} \quad (1)$$

For the ESN, the basic relationships can be expressed as:

$$\begin{cases} \mathbf{x}(t+1) = \tan \text{sig}(\mathbf{W}_{n \times k}^{\text{in}} \mathbf{u}(t+1) + \mathbf{W}_{n \times n} \mathbf{x}(t) + \mathbf{W}_{n \times l}^{\text{back}} \mathbf{y}(t)) \\ \mathbf{y}(t+1) = \mathbf{W}_{l \times (k+n+1)}^{\text{out}}(\mathbf{u}(t+1), \mathbf{x}(t+1), \mathbf{y}(t)) \end{cases} \quad (2)$$

where $\mathbf{W}_{n \times k}^{\text{in}}$ denotes the connection weight matrix for the input layer, $\mathbf{W}_{n \times l}^{\text{back}}$ represents the feedback weighting matrix from the output layer to the hidden layer, $\mathbf{W}_{l \times (k+n+1)}^{\text{out}}$ indicates the connection weight matrix for the output layer, $\mathbf{W}_{n \times n}$ means the connection weight matrix for DR. For $\mathbf{W}_{n \times n}$, it is usually maintained the sparse connection between 1% and 5%, while the spectral radius is generally less than 1. Thus, a dynamical memorizing ability and a certain degree of stability are possessed by the DR method.

The training process for ESN can be divided into the following steps:

- 1) The initialization of the matrices such as $\mathbf{W}_{n \times k}^{\text{in}}$, $\mathbf{W}_{n \times n}$, $\mathbf{W}_{n \times l}^{\text{back}}$;
- 2) Determination of the sequence $\mathbf{x}(t)$ for the hidden layer and the output weight matrix $\mathbf{W}_{l \times (k+n+1)}^{\text{out}}$ with the utilization of the sampled data. Here, the matrices $\mathbf{W}_{n \times k}^{\text{in}}$, $\mathbf{W}_{n \times n}$ and $\mathbf{W}_{n \times l}^{\text{back}}$ are randomly generated, before the construction of the ESN. Furthermore, these matrices are fixed once they are generated. As to the weight connection matrix $\mathbf{W}_{l \times (k+n+1)}^{\text{out}}$ which is connected to the output, it is supposed to be generated with the utilization of the Linear regression method through network training [6]. In this way, the trained ESN can be further for the modeling and prediction of the nonlinear time series.

B. Noise Reduction of the Time Series Based on Wavelet Analysis

For the noise reduction based on wavelet analysis, the signal is supposed to be decomposed according to multi-resolution. Then, the signal in time-domain is represented as the sum of a number of fractions in time domain. These fractions are supposed to signals with various sub frequencies. Furthermore, the signal sequence is refined through multi scale analysis from multiple levels. Then the useful signal can be distinguished from the noise [7]. The principle of noise reduction using wavelet analysis is presented as follows.

The data with noise included is denoted as $f(t) = r(t) + e(t)$, where $r(t)$ indicates the original signal, $e(t)$ denotes the Noise signal. Then the discrete wavelet analysis is performed on $f(t)$ based on the basic wavelet $\phi(t)$. The process is presented as follows

$$\begin{aligned} \Psi(j, m) &= 2^{-\frac{j}{2}} \int f(t) \bar{\phi}_{j,m}(t) dt = 2^{-\frac{j}{2}} \int [r(t) + e(t)] \bar{\phi}_{j,m}(t) dt \\ &= 2^{-\frac{j}{2}} \int r(t) \bar{\phi}_{j,m}(t) dt + 2^{-\frac{j}{2}} \int e(t) \bar{\phi}_{j,m}(t) dt \\ &= R(j, m) + E(j, m) \end{aligned} \quad (3)$$

where, j indicates the periodic scale of the wavelet, m denotes the time factor that reflects the translation of wavelet in time domain, $\bar{\phi}(t)$ represents the Complex conjugate function for $\phi(t)$, while $R(j, m)$ means the coefficient of the recovered signal after wavelet translation, $E(j, m)$ is the noise signal after wavelet transform coefficient. Given the translation of binary wavelet, the relationship of the singularity for signal $f(t)$ and the index α of Lipchitz is denoted as follows:

$$|\Psi(j, m)| \leq \omega 2^{j\alpha} \quad (4)$$

where, ω is a constant value.

As indicated by formula (4), $\alpha > 0$ for useful signal, the maximum modulus value for the wavelet transform increases with the scale. Nevertheless, for the noise signal, $\alpha \leq 0$; then the maximum modulus of the wavelet transform will decrease with the increase of the scale. Therefore, different thresholds need to be set at different scales; then, we are supposed to compare the maximum modulus with the threshold. If the maximum modulus is smaller than the threshold, then it is set to be 0. Otherwise, it is maintained as this maximum modulus is supposed to be incurred by the wavelet transform of the signal. Then, a wavelet transform reconstruction process can be performed, in order to reduce the noise.

C. Dimension Reduction Based on PCA

The PCA is an effective method for dimension reduction and feature extraction. This method is based on the Karhunen-Loeve (K-L) decomposition. A set of new latent variables are constructed to reduce the dimension of the space for the original data. Then the main information is extracted from the new mapping space. Thus, it is easier for us to understand the characteristics of the original data space. Since statistical feature vectors are orthogonal to each other, then the relevance between the variables can be eliminated. Furthermore, variables for the mapping space are constructed by the linear combinations of the variables in the original data. Thus, the main information of the original data is maintained while reducing the space dimension [9]. The main steps for the PCA are presented as follows:

- 1) Standardization of the sample matrix X of the input, then a new matrix X^* is obtained;
- 2) Establishing the correlation matrix, i.e., $S = X^{*T} X^*$;
- 3) Calculating the λ and unit feature vector β for the correlation matrix S , and the λ_i are to be ordered as $\lambda_1 \geq \lambda_2 \geq \dots \lambda_n$;
- 4) The variance contribution of the h are to be calculated, the variance contribution rate of λ_i is calculated as:

$$\eta_i = \left| \lambda_i / \sum_{i=1}^h \lambda_i \right| \times 100\% \quad (5)$$

5) Set a threshold, then the cumulative contribution of the η is to be calculated. If η is larger than the pre-specified threshold, then the first P values are taken as a new feature, which is used to represent the main information contained in the original variables;

6) Extracting the unit eigenvector $\hat{\beta} = [\beta_1, \beta_2, \dots, \beta_p]$ which corresponds to the p ; then Principal component matrix $Z = X\hat{\beta}$ can be determined.

Hence, the dimension reduction of the target data is completed, and the majority information of the original data is preserved.

III. ESN BASED COMBINATION METHOD IN THE HEALTH ANALYSIS OF TURBINE

For the traditional ESN model, there exist some drawbacks when dealing with the prediction of nonlinear time series. Thus, a combination method based on ESN is presented in this paper which combines the wavelet denoising method and principal component analysis. Furthermore, the characteristics of the nonlinear time series for an Turbine are also incorporated. With the application of wavelet analysis in this method, the noise in the time series can easily be smoothed through the application of the wavelet analysis; hence, the precision of the prediction can be greatly improved. Furthermore, the dimension can be reduced with the utilization of the principal component analysis while the speed of the prediction can also be improved.

A. Description of the Investigated Example

Pressure measurement is an important means for the purpose of controlling and health monitoring of the engine [10]. In practice, the dynamic pressure is an important feature for the characterization of flight conditions for the aircraft, and it is also one of the most important parameters for controlling the power output. In this paper, the prediction of the measured time series data of dynamic pressure is presented in Figure 3 through the application of the combination method based on ESN. Furthermore, the results are also compared with that obtained with the adoption of the traditional ESN based prediction method.

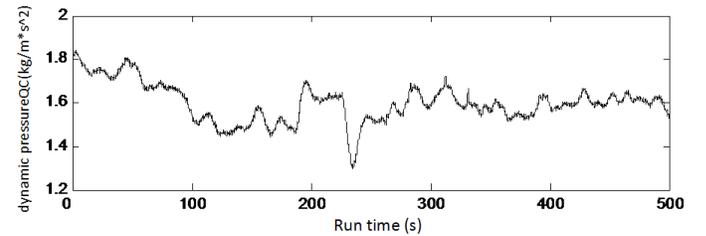


FIGURE III. MEASURED DATA FOR THE DYNAMIC PRESSURE

As indicated by the results, the dynamic pressure data is a nonlinear chaotic time series [11]. If the analysis is to be performed with the application of the ESN based prediction model, the phase space reconstruction of nonlinear time series is required to be reconstructed. Then, the sampling data can be obtained. From the data, nonlinear model which is close to the dynamic system can be determined through training. In this paper, the C-C method [12] is adopted to determine the parameters which can be used for the reconstruction of phase space.

B. Precision Verification and Procedures

In this paper, three standards are introduced to reflect the precision of prediction from different perspectives, i.e., the mean square error (EMSE), normalized root mean square error (ENRMSE) and the normalized absolute error (ENAE). The definitions of the above variance evaluation functions are provided as follows:

Given s samples, let $y(t)$ be the expected output at t (i.e., the observed value), $\hat{y}(t)$ be the predicted output at t (i.e., the predicted value), \bar{y} be the mean of the expected output. Then, the standard deviation is calculated as $\sigma_y = \left(\frac{1}{s} \sum_{t=1}^s (y(t) - \bar{y})^2 \right)^{\frac{1}{2}}$.

1) Mean squared error EMSE

$$E_{MSE} = \frac{1}{s} \sum_{t=1}^s [\hat{y}(t) - y(t)]^2 \quad (6)$$

2) Standard root mean square error ENRMSE

$$E_{NRMSE} = \frac{\left(\frac{1}{s} \sum_{t=1}^s (\hat{y}(t) - y(t))^2 \right)^{\frac{1}{2}}}{\sigma_y} \quad (7)$$

3) Normalized absolute error ENAE

$$E_{NAE} = \frac{\sum_{t=1}^s |\hat{y}(t) - y(t)|}{\sum_{t=1}^s y(t)} \quad (8)$$

where, EMSE indicates the mean square of the error of the prediction, ENRMSE reflects the average deviation of the predicted values compared with the observed values, while the EMSE means the normalization of the average absolute error for the predicted value with the observed mean normalized [13].

In this paper, the steps of the combined model are illustrated as follows:

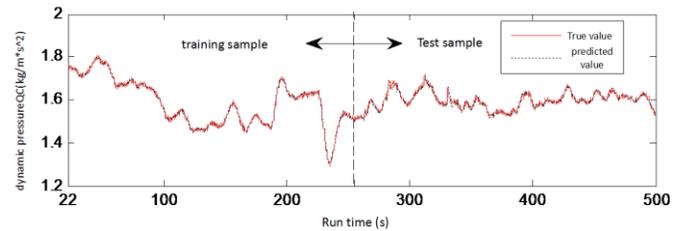
- 1) Performing wavelet analysis on the sample data from original time series for the purpose of reducing noise;
- 2) Reconstructing the phase space in order to generate the sample data;
- 3) Reducing the dimension through the application of the principal component analysis on the sample data;
- 4) Feeding the sample data with reduced dimension into the model, and completing the training process of the ESN;
- 5) Performing predicting process by adoption of the trained ESN prediction model

C. Predicting Method Based on Traditional ESN

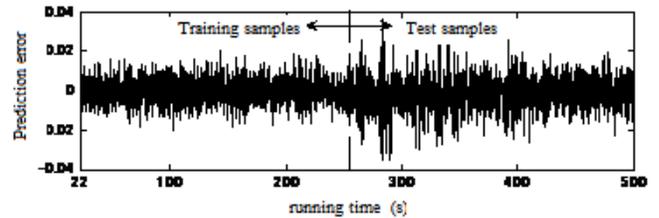
The C-C method is utilized to reconstruct the phase space for the measured nonlinear time series of dynamic pressure data, while the delay time and the embedding dimension are chosen to be 74 and 6 respectively. The delay vector for the reconstructed dynamic pressure is denoted as $X(t) = (x(t), x(t-14), L, x(t-(6-1) \times 14))$

Thus, the number of nodes for the input layer of the ESN network is assumed to be 6, and the number of neurons in the output layer equals to 1. Then, the number of neurons in the

hidden layer is assumed to be 300, the connection matrix of the hidden layer is 3%, while the spectral radius of the sparse matrix is set to be 0.5. The training of prediction model based on ESN can be performed while using the constructed training samples. Here, a single step and five steps predictions are performed for the hydrodynamic parameters. Here, the first half of the data (usually 50%, i.e., data on the left of the dashed line) is used for the training of the ESN network, while the remaining data is adopted as the test data. The prediction and corresponding prediction error are presented in Figure 4 and Figure 5 respectively.

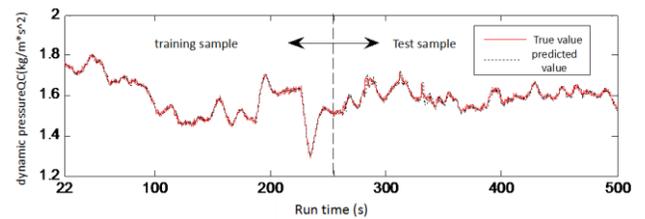


(a) Prediction result of a single step for the dynamic pressure using the ESN based prediction model

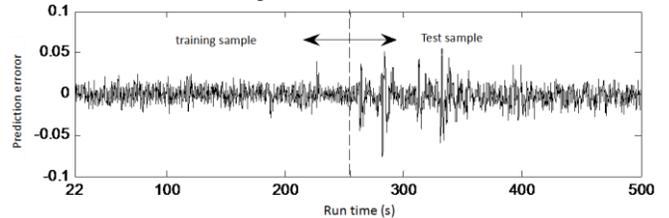


(b) Corresponding predicting error for the one step prediction using ESN based prediction model

FIGURE IV. ONE STEP PREDICTION OF DYNAMIC PRESSURE USING THE ESN BASED MODEL



(a) Prediction of the dynamic pressure for 5 steps using the ESN based prediction model



(b) Corresponding prediction error of the dynamic pressure for 5 steps using the ESN based prediction model

FIGURE V. STEPS PREDICTION CURVE OF DYNAMIC PRESSURE WITH THE ESN MODEL

IV. APPLICATIONS

In order to apply the combination method based on ESN proposed in this paper, the noise of corresponding dynamic pressure data is anticipated to be reduced with the application

of the wavelet analysis. Here, the time series data with the noise reduced are presented in Figure 6.

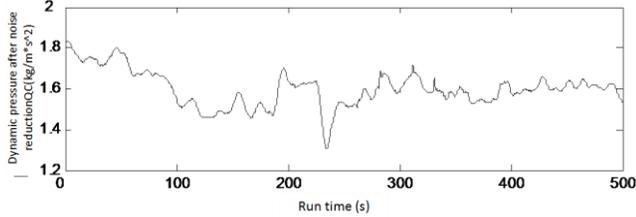


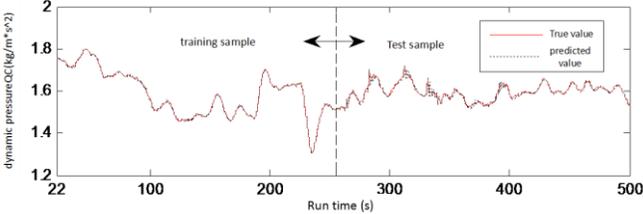
FIGURE VI. DYNAMIC PRESSURE DATA AFTER NOISE REDUCING

Then, the C-C method is used to reconstruct the phase space for the dynamic pressure data, then, the delay time is

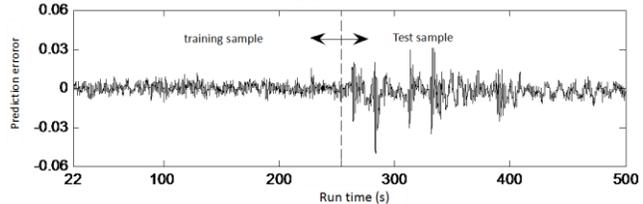
TABLE I. PRINCIPAL COMPONENTS AND VARIANCE CONTRIBUTION OF THE DYNAMIC PRESSURE PHASE SPACE

Serial number	Characteristic value	variance Contribution rate	Serial number	Characteristic value	variance Contribution rate
1	5.4388	0.90646	4	0.011723	0.0019538
2	0.46416	0.077359	5	0.0054133	0.00090222
3	0.075893	0.012649	6	0.0040479	0.00067465

As indicated by the results in table 1, the variance contribution of the first 2 eigenvalues is 0.9838 which is already greater than the pre-specified value of 0.95. Hence, the first 2 principal components are selected to construct the delay vector.



(a) One step prediction with the application of the method proposed in this paper



(b) One step prediction error of dynamic pressure

FIGURE VII. ONE STEP PREDICTION CURVE OF DYNAMIC PRESSURE WITH THE PROPOSED METHOD

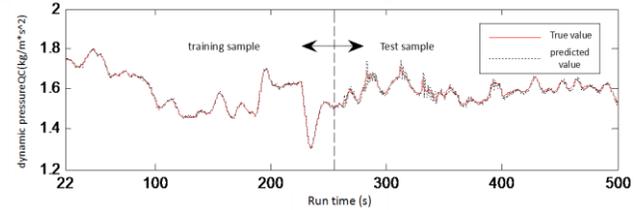
Through the application of the principal component analysis, the dimension for the phase space is reduced from 6 to 2. Then, when constructing the ESN network, the number of nodes for the input layer equals to 2, while the numbers of nodes for the hidden and output layers are the same as the selection process for the ESN in Section 2.3. This is also applicable for the determination of the other parameters. Similarly, the first half of the data is utilized to build the model, while the remaining is utilized for the prediction process with the adoption of the existing model. The model is used to for the analysis of the nonlinear time series data (after noise reducing) for further prediction. The prediction results for the one step and 5 steps are presented in Figures 7 and 8 respectively. In the figure, the data on the left side of the

obtained as 14 while the embedding dimension equals to 6. Hence, the delayed vector for the dynamic pressure is represented as:

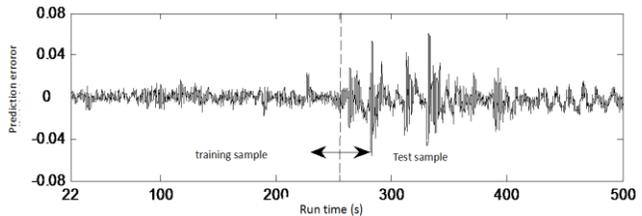
$$X(t) = (x(t), x(t-14), \dots, x(t-(6-1) \times 14)) .$$

After that, the principal component analysis can be applied for the purpose of reducing the dimension of the reconstructed delay vector. In the experiment in this paper, the principal elements are selected if the cumulative variance contribution rate of these principal elements meet the relationship $\eta \geq 0.95$. The corresponding eigenvalues and their variance contribution rates are presented in Table 1.

dashed line is the training sample data, while the data on the right side is the sample data for testing.



(a) 5 steps prediction of dynamic pressure



(b) 5 step prediction error of dynamic pressure

FIGURE VIII. STEPS PREDICTION CURVE OF DYNAMIC PRESSURE WITH THE PROPOSED METHOD

In this section, the prediction results with the application of the traditional ESN based prediction method and the combination method are listed in Table 2.

TABLE II. PREDICTION ERROR AND TIME COST OF DYNAMIC PRESSURE WITH THE PROPOSED METHOD AND ESN MODEL

Prediction step	Prediction method	Prediction time /s	Prediction error		
			E_{MSE}	E_{NRMSE}	E_{NAE}
Paper method	ESN	5.844	7.9609e-5	0.224500	0.004162
	Paper method	3.500	6.4349e-5	0.204810	0.003473
5 step	ESN	5.031	0.000175	0.333570	0.005967
	Paper method	3.594	0.000139	0.301890	0.005319

As shown by Figures 4, 5, 7, and 8, also indicated by the results in table 2, for the same time series data for the dynamic pressure, the prediction time is reduced from 5.031 s (time consumed by the traditional ESN based method) to 3.594 s (time required by the method in this paper) for the 5 steps prediction. While for the one-step prediction, the prediction time is reduced from 5.844s to 3.500s. Thus, the cumulated prediction rate is increased by 66.97%. From the perspective of prediction accuracy, the obtained mean square error, mean square error or the normalized absolute error of nonlinear time series prediction for dynamic pressure is significantly improved with the adoption of the combination method based on ESN.

Overall, due to the incorporation of the wavelet denoising and principal component analysis approaches, higher prediction accuracy is capable of being obtained compared with the adoption of the traditional ESN based method. Furthermore, the required time for the training of the ESN is also significantly reduced, this indicates faster prediction speed is obtained. Thus, the proposed combination method is able to be applied to a lot of scenarios.

V. CONCLUSION

In this paper, for the Turbine condition and analysis of the health characteristics, the measured nonlinear time series are of their own properties. Considering the prediction error and real-time property, a combination method based on ESN is presented. First of all, this method is based on the ESN model, while the wavelet analysis and principal component analysis approaches are incorporated for the analysis of the time series prediction. The noise denoising is implemented by using the wavelet analysis, while the dimension reduction is achieved by using the principal component analysis approach. As analyzed in the paper, the prediction accuracy and speed can be significantly improved.

In this paper, the measured time series data of the dynamic pressure is processed by the proposed model in this work and the traditional method. The steps of the prediction process are provided while the implementation process is discussed thoroughly. Furthermore, the prediction accuracy and the prediction speed of different methods are compared. As indicated by the simulation results, the prediction effect of this proposed method in this paper is much better, while the nonlinear time series is much closer. and the prediction accuracy of the proposed method in this paper is improved and meanwhile, the corresponding training and prediction time is reduced to some extent. Hence, the proposed method in this paper is a more practical and valuable prediction method.

REFERENCES

- [1] GuoYangming, CaiXiaobin, ZhangBaozhen, et al. Review of prognostics and health management technology[J]. Computer Measurement & Control.2008,16(9):1213-1216,1219.(in chinese)
- [2] JiangCaihong, SunZhiyan, Wang Xi. Critical technologies for Turbine prognostics and health management systems development[J].Journal of Turbospace Power,2009, 24(11) :2589-2594. (inchinese)
- [3] Sorjamaa A, Hao J, Reyhani N, Ji Y, etc. Methodology for long-term prediction of time series[J]. Neurocomputing. 2007, 70: 2861-2869.

- [4] Jaeger H. The echo state approach to analyzing and training recurrent neural networks[R], German: German National Research Center for Information Technology, Tech. Rep. 148, 2001.
- [5] FuLinjuan, ZhaiZhengjun, GuoYangming. Research on multivariate prediction model based on echo state networks[J].Computer Measurement & Control, 2009.17(7):1356-1358,1361
- [6] LeiZhengwei, XuZhangsui, Mi Dong,etal. Optimization analysis of neural network forecast performance[J]. Computer Measurement & Control, 2004, 12(1):15-16,20.
- [7] ChenGuo. A new de-noising method for rotor faults signal[J]. Journal of Vibration Engineering, 2007, 20(3) : 285-290.
- [8] Mallat S, Huang W. Singularity detection and processing with wavelets[J]. IEEE Trans. On Information Theory, 1992, 38(2): 617-643.
- [9] XiJianhui, Han Min. Prediction of multivariate time series based on principal component analysis and neural networks [J]. Control Theory & Application, 2007, 24(5) : 719-724.
- [10] ZhaoWentao, LiQiong. Research on Turbine health monitoring and diagnosis methods[J].Aircraft Design, 2010,30(6):30-34.
- [11] SongYunxue, PengPeng, ShiYongsheng. Chaos arithmetic for civil Turbine forecasting[J].Journal of Turbospace Power, 2011,26(3):675-679.
- [12] QinYiqing, CaiWeidong, YangBingru. Research on phase space reconstruction of nonlinear time series[J].Journal of System Simulation, 2008, 20(11) : 2969-2973.
- [13] FuLinjuan. Research on prediction of time series in complex nonlinear system[D]. Xi'an: Northwestern Polytechnical University,2009.