

Research Article

The Chaotic Prediction for Aero-Engine Performance Parameters Based on Nonlinear PLS Regression

Chunxiao Zhang¹ and Junjie Yue²

¹ College of Science, Civil Aviation University of China, Tianjin 300300, China

² Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China

Correspondence should be addressed to Chunxiao Zhang, cxzhang@cauc.edu.cn

Received 3 May 2012; Revised 18 July 2012; Accepted 19 July 2012

Academic Editor: Zhiwei Gao

Copyright © 2012 C. Zhang and J. Yue. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The prediction of the aero-engine performance parameters is very important for aero-engine condition monitoring and fault diagnosis. In this paper, the chaotic phase space of engine exhaust temperature (EGT) time series which come from actual air-borne ACARS data is reconstructed through selecting some suitable nearby points. The partial least square (PLS) based on the cubic spline function or the kernel function transformation is adopted to obtain chaotic predictive function of EGT series. The experiment results indicate that the proposed PLS chaotic prediction algorithm based on biweight kernel function transformation has significant advantage in overcoming multicollinearity of the independent variables and solve the stability of regression model. Our predictive NMSE is 16.5 percent less than that of the traditional linear least squares (OLS) method and 10.38 percent less than that of the linear PLS approach. At the same time, the forecast error is less than that of nonlinear PLS algorithm through bootstrap test screening.

1. Introduction

Aeroengine condition monitoring is applied as a better way to estimate aero-engine condition and reliability amongst most of the commercial airlines [1, 2]. Predicting and detecting abnormal behavior of main aeroengine performance parameters is of great importance in the aeroengine condition monitoring. With the development of modern industry, one typical aeroengine becomes a more and more integrative and complex system, of which working condition and fault are very difficult to predict and judge, and the influence and harm caused by system faults are also more serious than ever. Therefore, there will be a growing demand for effective prediction methods of long-term trend of aeroengine system to predict

abnormal behavior of significant performance parameters and potential faults. Then we can take preventive maintenance measure to avoid faults occurrence, shorten aircraft down time, thereby reducing maintenance cost and improving aeroengine life.

The aeroengine system is a complex system with nonlinear and dynamic chaos action, so it is difficult to establish an exact analytic mathematical model of performance parameters. As a result, analysis usually depends on time series obtained from observation. But in practical monitoring the availability of obtaining characteristic data is very limited, and one fault usually shows various changes in performance parameters, and a same change may result from different faults [3, 4]. However, extracting multiple fault characteristic of parameters usually brings complexity and redundancy, which may bring difficulty for the following prediction. The theory of chaotic prediction is a mathematics method to deal with inexactness, uncertainty, and incompleteness, which is remarkable efficiency in aspects of data reduction and feature extraction.

Studies on chaos prediction can be traced back to Parker and Chua (1989) [5]. More sophisticated approaches of nonlinear chaotic time series prediction mainly rely on neural network [6], support vector theory [7], fuzzy rules [8], a simple least square algorithm [9]. In recent years, considerable interest has been aroused in the study of nonlinear chaos prediction with the actual observation of aeroengine dynamic system [10, 11]. These papers summarize the prediction model of nonlinear chaos system based on ordinary least squares (OLS); however, there are some practical problems in those models such as serious multicollinearity and poor robustness of regression function.

Partial least squares (PLS) can overcome information overlaps caused by multicollinearity since it adopts canonical correlation analysis, which improves accuracy of nonlinear system modeling [12]. What is more, PLS adopts cross validity method, which rejects redundancy principal component and enhances regression equations robustness. Hence, in our investigation the designed chaotic algorithm is based on the nonlinear partial least squares (NPLS) regression and the prediction research on the time series.

2. Data Pretreatment

Typical aeroengine performance parameters are exhaust gas temperature (EGT), low-spool rotor speed (N1), high-spool rotor speeds (N2), and fuel flow (WF) [13]. These four measurements are often called the four basic parameters, which can be recorded in the airborne equipment on most new and old aeroengines. Note that we take the most important performance parameter, EGT, as example to analyze in this paper; the other parameters can be fitted in the same method.

2.1. Similarly Correcting EGT Sequence

According to operating principle of the aeroengine, the data of main performance parameters of the same aeroengine in different external flight conditions such as the external atmosphere total temperature, total pressure, MACH numbers are quite different, so the original engine exhaust temperature EGT data usually cannot be directly used for comparison analysis. To

solve this problem, similar theory is adopted to eliminate the influence of external conditions on the aeroengine EGT sequence [14].

$$\text{EGT}_{\text{cor}} = \frac{\text{EGT}_{\text{obs}}}{\theta T_2}, \quad (2.1)$$

where EGT_{obs} is the original EGT data, EGT_{cor} is the corrected value of EGT_{obs} , $\theta T_2 = (\text{TAT} + 273.15)/288.15$, TAT is the total air temperature.

2.2. Eliminating Outliers According to Layida Criterion

If outliers exist in the data, they will seriously affect the prediction accuracy of an algorithm. Layida criterion is the most commonly used for discrimination outliers, of which the precondition is that there is sufficient data [15]. For time series $\{X_i\} (i = 1, \dots, n)$, if n is large enough (normally $n > 80$), and only random disturbance exists in all data. Assume that the random time series is following a normal distribution whose mean is μ and deviation is σ^2 , so the probability that the deviation $(X_i - \mu)$ falls in $\pm 3\sigma$ ranges is about 0.3%. According to Layida criterion, if the absolute deviation of a time series data $X_i (i = 1, \dots, n)$ is larger than 3σ , then X_i can be judged as an outlier and should be removed.

Let the number of sample be N , the observed value of time series is $\{x_i\} (i = 1, \dots, N)$, the sample mean and standard deviation is \bar{x} and S , respectively. If the sample point x_i meets $|x_i - \bar{x}| > 3S$, then x_i is an outlier which will be eliminated. \bar{x} and S should be computed again after all the outliers are removed.

3. Chaos Prediction Model Based on Nonlinear Partial Least Squares (PLS) Regression

Suppose observed time series is $\{x_n\} (n = 1, 2, \dots, N)$, the chaotic phase space s_n is reconstructed by [16].

$$s_n = (x_n, x_{n-\tau}, \dots, x_{n-(m-1)\tau}) \in R^m, \quad n = (m-1)\tau + 1, (m-1)\tau + 2, \dots, N, \quad (3.1)$$

where τ is the delay time, m is the embedding dimension and N is the length of the time series.

3.1. Characterizing Chaos through Lyapunov Exponents

Through the years, the existence of chaos has been characterized in time series by means of several methods, among others: analysis of Fourier spectra, fractal power spectra, entropy, fractal dimension, and calculation of Lyapunov exponents. However, several of these methods have proven not to be very efficient [16]. The Lyapunov exponents provide all the information about the local and global complexity of chaos, therefore the measurement of Lyapunov exponents has been an effective method to judge whether a time series is caused by a chaotic dynamical system.

The Lyapunov exponents $\lambda_1, \lambda_2, \dots, \lambda_d$ are the average rates of expansion ($\lambda_i > 0$) or contraction ($\lambda_i < 0$) near a limit set of a dynamical system [16]. In fact, the LE is a quantitative

measure of the divergence of several trajectories in the system. There is one LE for each direction in the d -dimensional phase space where the system is embedded. The variable d represents the Lyapunov dimension or phase space dimension. The LE does not change when the initial conditions of a trajectory are modified or when some perturbation occurs in the system. If at least one LE is positive, the system presents chaotic motion [17]; hence if chaos exists in a dynamical system, the largest Lyapunov exponent must be positive.

We will calculate the largest Lyapunov exponents according to the algorithm offered by Wolf et al. (1985) [17]; the numerical results are shown in Table 1.

3.2. Chaos Prediction Model

Let state vector at time T be s_T , we obtain the nearest neighbor points $s_{\alpha_1}, s_{\alpha_2}, \dots, s_{\alpha_K}$ by calculating the Euclidean distance between the target points s_T and any one of reconstruction vectors [18], where

$$s_{\alpha_i} = (x_{\alpha_i}, x_{\alpha_i-\tau}, \dots, x_{\alpha_i-(m-1)\tau}), \quad i = 1, 2, \dots, K. \quad (3.2)$$

Let X be the sample matrix of reconstructed variable,

$$X = \begin{bmatrix} x_{\alpha_1} & x_{\alpha_1-\tau} & \cdots & x_{\alpha_1-(m-1)\tau} \\ x_{\alpha_2} & x_{\alpha_2-\tau} & \cdots & x_{\alpha_2-(m-1)\tau} \\ \vdots & \vdots & \vdots & \vdots \\ x_{\alpha_K} & x_{\alpha_K-\tau} & \cdots & x_{\alpha_K-(m-1)\tau} \end{bmatrix} \in R^{K \times m}. \quad (3.3)$$

Let y be the prediction variable, then the nonlinear chaos prediction model is described as

$$y = f(x_1, x_2, \dots, x_m), \quad (3.4)$$

where x_i is the i th column of X , $i = 1, 2, \dots, m$, $y = [x_{\alpha_1+p}, x_{\alpha_2+p}, \dots, x_{\alpha_K+p}] \in R^K$, p is the length of prediction step.

In practice, the observed data sets of independent variable and dependent variable can usually be obtained. However, the model form of their relations is unknown especially in case of higher dimensions, and the problem becomes more complicated when the relationship between the dependent variable and independent variables is nonlinear. Usually (3.4) is fitted by additive model or multiplicative model. For the sake of computation convenience, we consider the addition model of each independent variable; that is,

$$y = f_1(x_1) + f_2(x_2) + \cdots + f_m(x_m) + \varepsilon. \quad (3.5)$$

At present, the nonlinear relationship between the original variable usually is into a quasi-linear relationship, after then, the prediction model is established by an effective linear theory method. Spline function and kernel function are commonly used in the transformation of the basic functions.

Table 1: Comparative analysis of the optimal results based on various forecasting methods.

Forecasting method	Embedding dimension (m)	Delay time (τ)	Subset numbers	Largest Lyapunov exponent	NMSE
Cubic spline interpolation	9	3	5	0.25	0.6400
Uniform kernel	8	19	2	0.16	0.5529
Epanechnikov kernel	8	19	2	0.17	0.5529
Bi-weight kernel	8	19	2	0.38	0.5431
Tri-weight kernel	9	3	5	0.29	0.5918
Gaussian kernel	3	10	4	0.32	0.6325
PLS	9	3	0	0.15	0.6469
OLS	30	16	0	0.098	0.7081

(1) Cubic Spline Function Transformation

Let $\hat{f}_j(x_j)(j = 1, 2, \dots, m)$ be the cubic spline fitting function, then,

$$y = \beta_0 + \hat{f}_1(x_1) + \hat{f}_2(x_2) + \dots + \hat{f}_m(x_m) + \varepsilon, \quad (3.6)$$

where

$$\hat{f}_j(x_j) = \sum_{l=0}^{M_j+2} \beta_{j,l} \Omega_3 \left(\frac{x_j - \xi_{j,l-1}}{h_j} \right), \quad (3.7)$$

$$\Omega_3 \left(\frac{x_j - \xi_{j,l-1}}{h_j} \right) = z_{j,l} = \frac{1}{3!h_j^3} \sum_{k=0}^4 (-1)^k \binom{4}{k} (x_j - \xi_{j,l-3+k})^3, \quad (3.8)$$

$$\xi_{j,l-1} = \min(x_j) + (l-1)h_j,$$

$$h_j = \frac{\max(x_j) - \min(x_j)}{M_j} \quad (l = 0, 1, \dots, M_j + 2), \quad (3.9)$$

where $\beta_0, \beta_{j,l}$ are the model parameters to be determined; $\xi_{j,l-1}, h_j, M_j$ are range points, segment length and section number are divided on the variable x_j , respectively.

The minimum observed value of variables is recorded as $\min(x_j)$ and the maximum as $\max(x_j)$, then the nonlinear prediction function relationship between independent variables and dependent variable can be transformed as the following equation:

$$y = \beta_0 + \sum_{j=1}^m \hat{f}_j(x_j) + \varepsilon = \beta_0 + \sum_{j=1}^p \sum_{l=0}^{M_j+2} \beta_{j,l} \Omega_3 \left(\frac{x_j - \varepsilon_{j,l-1}}{h_j} \right) + \varepsilon. \quad (3.10)$$

Equations (3.6)–(3.8) show that the relationship between y and $z_{j,l}$ is linear, and the relationship between x_j and $z_{j,l}$ is nonlinear, so the chaos prediction model with cubic spline function transformation is a quasi-linear regression model (see in (3.10)).

(2) Kernel Function Transformation

The principle of Kernel function transformation is same as spline function transformation; that is, a kernel function transformation based on one dimensional nonlinear function $f_j(x_j)$, and the cubic spline function $\{\Omega_3((x_j - \xi_{j,l-1})/h_j)(l = 0, 1, \dots, M_j + 2)\}$ is replaced by basic function $\{K((x_j - \xi_{j,l-1})/h_j)(l = 0, 1, \dots, M_j + 2)\}$. Thus, the chaos prediction model with kernel function transformation is as follows [19]:

$$\hat{y} = \beta_0 + \sum_{j=1}^p \hat{f}_j(x_j) = \beta_0 + \sum_{j=1}^p \sum_{l=0}^{M_j+2} \beta_{j,l} K\left(\frac{x_j - \xi_{j,l-1}}{h_j}\right). \quad (3.11)$$

The commonly used kernel functions include uniform kernel function, Epanechnikov kernel function, biweight kernel function, tri-weight kernel function and Gaussian kernel function.

3.3. Nonlinear Partial Least Squares Regression (NPLS)

However, there exist some problems when establishing the above quasi-linear regression model based on ordinary least squares method, because the dimension of regression model will increase significantly after a function change, which may lead to the result that the number of sample points is less than the number of transformed variables; on the other hand, the data of reconstructed variable x_1, x_2, \dots, x_m are come from the same series, and the nonlinear correlation exists between each other. So it is not appropriate to employ the ordinary least squares method to estimate the model parameters when the multicollinear relationship among the new transformed independent variables is found. One of the effective methods to solve this problem is to use partial least squares method (PLS) to estimate parameter of the predictive model.

In this paper, a modified chaos prediction approach based on nonlinear PLS regression with basic function transformation is developed. The cubic spline function and various kernel function transformation are used in our model, and the comparative analysis between them is presented in Section 5.

3.4. Prediction Evaluation

Normally, prediction accuracy is evaluated by the mean squared error MSE,

$$\text{MSE} = \frac{1}{P} \sum_{p=1}^P (x_{T+p} - \hat{x}_{T+p})^2, \quad (3.12)$$

where \hat{x}_{T+p} is predict value, x_{T+p} is observed value. However MSE is concerned with the order of the magnitude of the observed data. In this paper, the normalized mean square error NMSE is used as the evaluation criterion of prediction.

$$\text{NMSE} = \frac{\text{MSE}}{\sigma^2}, \quad (3.13)$$

where σ^2 is sample variance of the prediction sequence. Our investigation adopts the mean of 20 predicted values NMSE in order to overcome contingency.

4. Algorithm and Algorithm Flowchart

Step 1. The observational data of aeroengine performance parameter EGT are collected and similarly corrected, and then the outliers are removed according to Layida criterion.

Step 2. For the selected embedding dimension m ($1 \leq m \leq m_{\max}$) and delay time τ ($1 \leq \tau \leq \tau_{\max}$), the chaos phase space on pretreated EGT time series is reconstructed, and the appropriate nearby points are selected to construct independent variable matrix \mathbf{X} and dependent variable vector \mathbf{y} .

Step 3. Choose section number M ($1 \leq M \leq M_{\max}$); each row vector x_j ($j = 1, 2, \dots, m$) in independent variable matrix \mathbf{X} is transformed to \mathbf{Z}_j by various basis function transformation such as (3.8), then

$$\mathbf{Z}_j = (z_{j,0}, z_{j,1}, \dots, z_{j,M+2}). \quad (4.1)$$

Step 4. The new independent variable \mathbf{Z}_j and dependent variable \mathbf{y} are standardized and marked as $(\tilde{\mathbf{Z}}, \tilde{\mathbf{y}}) = (\tilde{\mathbf{Z}}_1, \tilde{\mathbf{Z}}_2, \dots, \tilde{\mathbf{Z}}_m, \tilde{\mathbf{y}})$, so the prediction model is described as

$$\tilde{\mathbf{y}} = \sum_{j=1}^p \sum_{l=0}^{M+2} \alpha_{j,l} \tilde{z}_{j,l} + \varepsilon, \quad (4.2)$$

where $\alpha_{j,l}$ ($j = 1, 2, \dots, m; l = 0, 1, \dots, M + 2$) are the parameter of regression model (4.2), ε is the stochastic error and follows the normal distribution with zero mean and the same variance.

Step 5. The partial least squares (PLS) is applied to build regression model (4.2). The number of appropriate PLS components is extracted by effective cross-method [20]; estimate the regression coefficients $\alpha_{j,l}$ ($j = 1, 2, \dots, m; l = 0, 1, \dots, M + 2$), and obtain the regression coefficients $\beta_0, \beta_{j,l}$ ($j = 1, 2, \dots, m; l = 0, 1, \dots, M + 2$) of original model (3.5) by returning $\alpha_{j,l}$.

Step 6. If estimation of prediction model for all the combination of M are completed, then go to step 7, otherwise go to step 3.

Step 7. Calculate the predicted value according to prediction function (3.6) and calculate the corresponding predicted NMSE.

Step 8. If the calculation of the combination of all the embedding dimension m and delay time τ is completed, then go to step 9, otherwise go to step 2.

Step 9. Find out the optimal prediction model with the smallest NMSE, and obtain the corresponding dimension m , delay time τ and section number M .

Algorithm flow chart is shown in Figure 1.

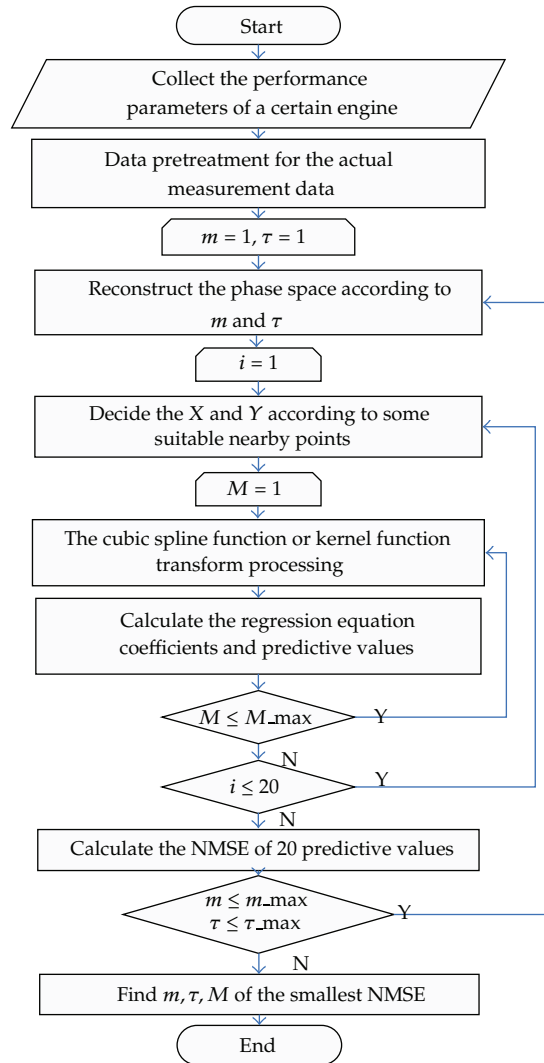


Figure 1: Algorithm flowchart.

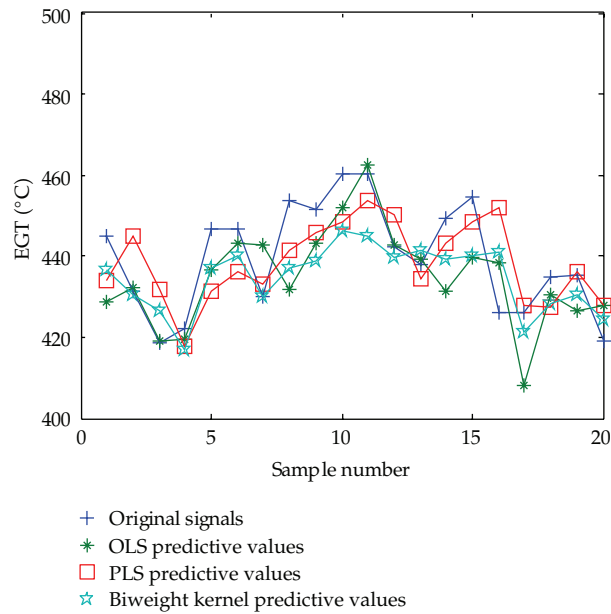
5. Results and Analysis

The in-the-wing ACARS data of four PW4077D aeroengines have been collected for one year, and about 1400 EGT data of each engine were selected to form a time series used in numerical example analysis.

The maxim value of embedding dimension m_{\max} , the delay time τ_{\max} , and the segment number M_{\max} in the above algorithm are estimated after a large amount of experiment. In our research, the range of m is among 1–30, τ is 1–20, M is 1–8, respectively. The cubic spline function and five kernel function introduced in Section 2 are employed for transformation, and we minimize chaotic prediction NMSE based on PLS algorithm.

Table 2: The optimal results after selecting independent variable by Bootstrap method based on NPLS.

Forecasting method	Embedding dimension (m)	Delay time (τ)	Subset numbers	NMSE
Cubic spline interpolation	9	3	5	0.6480
Uniform kernel	28	6	4	0.5753
Epanechnikov kernel	9	3	7	0.5917
Bi-weight kernel	8	19	2	0.5606
Triweight kernel	9	3	5	0.5786
Gaussian kernel	3	10	4	0.6510

**Figure 2:** Comparison of original values and predictive values.

The numerical results of the first engine are shown in Table 1, in which line 1 to line 6 show the prediction results based on various NPLS, line 7 to line 8 show the prediction results based on the common PLS and OLS.

As can be seen in Table 1, all the largest Lyapunov exponents are positive and it is clarified that the nonlinear and chaotic natures exist in EGT sequence. The first seven NMSE calculated based on PLS are all less than that based on OLS, which indicates that there is an obvious advantage of PLS over OLS in fitting chaotic prediction models. The NMSE of PLS based on various nonlinear transform methods is even less than that of free transformed PLS, and it indicates that there is a greater advantage of chaotic prediction based on PLS in forecasting aeroengine EGT. PLS prediction algorithm based on biweight Kernel transformation is the optimal chaotic prediction algorithm, estimated optimal embedding dimension is 8, the delay time is 19, and the number of nonlinear transformation section is 2.

PLS regression model includes all of the original selected variables, and there is multi-collinearity in the independent variables, which can affect the prediction accuracy. The model parameters estimated based on PLS is of a pretty complex nonlinear character,

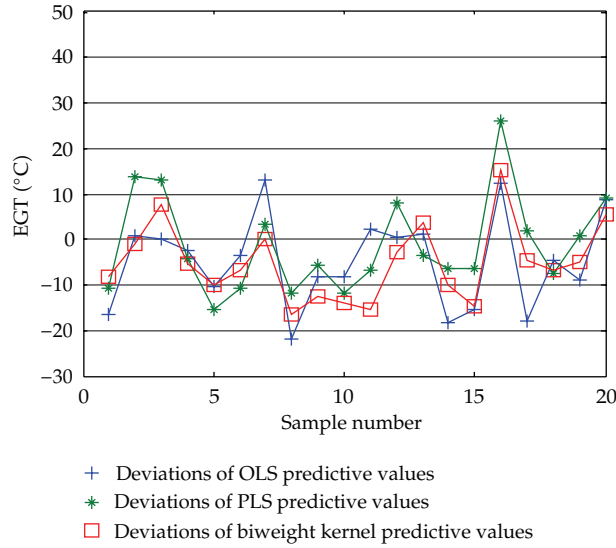


Figure 3: Deviations comparison.

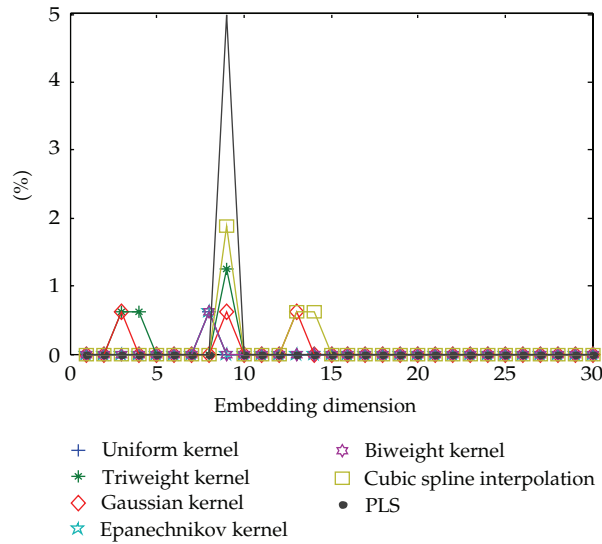


Figure 4: Embedding dimension analysis of some forecasting methods.

so the independent variables cannot be selected by the parameter test of linear methods. In this paper, the nonparametric statistical method (Bootstrap [20]) is adopted to select the independent variables transformed with kernel function and Cubic spline function, and to find the minimum NMSE; the optimal results of the first engine are shown in Table 2.

Table 2 shows the predictive results based on various NPLS algorithm after selecting independent variable by Bootstrap method. Only the optimal predictive NMSE based on triweight Kernel function transformation decreases, and the minimum NMSE in Table 2 is larger than that in Table 1; it indicates that better NMSE cannot be achieved by Bootstrap test for selecting independent variables.

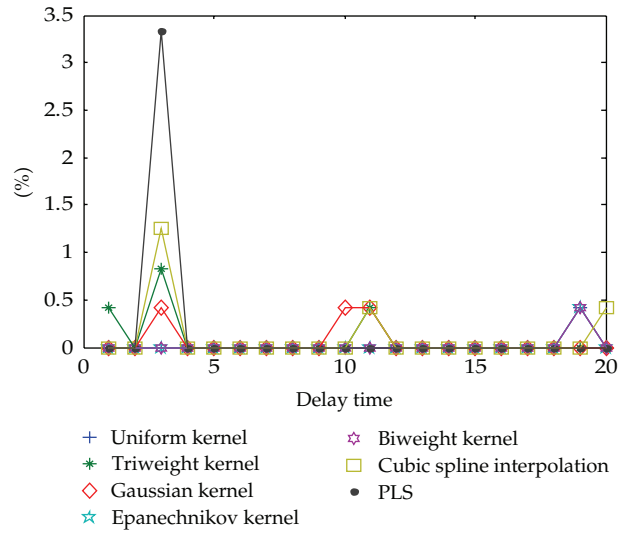


Figure 5: Delay time analysis of some forecasting method.

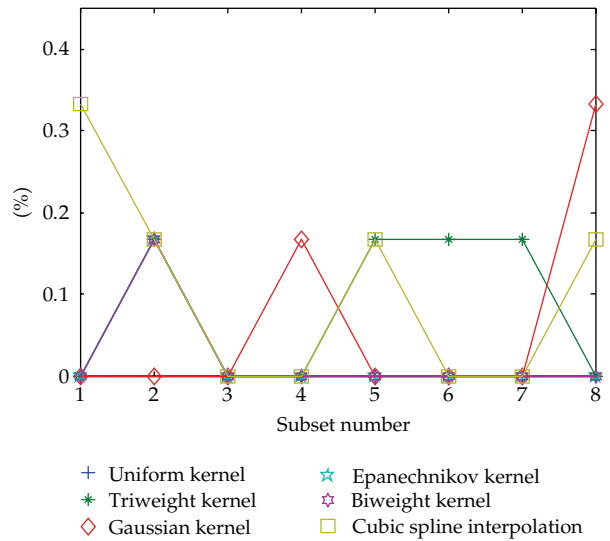


Figure 6: Subset numbers analysis of some forecasting methods.

Figure 2 shows the comparison of the prediction values based on OLS, PLS, and biweight Kernel transformation chaos algorithm with the original value. The PLS algorithm based on biweight Kernel function transformation has the optimal NMSE in Figure 2. However, not all of the predictive value has the minimum forecast NMSE.

The deviation between the predictive value and original value based on the above three algorithms are shown in Figure 3. The fluctuations of the deviations curve based on biweight Kernel method with the optimal NMSE is the smallest, and larger predicted error is not appeared in its deviation sequence.

Table 3: Optimal results of three engines based on various forecasting methods.

Forecasting method	Second aeroengine				Third aeroengine				Fourth aeroengine			
	m	τ	Subset numbers	NMSE	m	τ	Subset numbers	NMSE	m	τ	Subset numbers	NMSE
Cubic spline interpolation	8	5	4	0.6041	9	2	5	0.3354	17	3	5	0.3800
Uniform kernel	12	2	4	0.5411	7	2	2	0.3741	10	3	4	0.3529
Epanechnikov kernel	12	2	4	0.5706	9	2	2	0.3417	17	3	4	0.3329
Bi-weight kernel	12	2	4	0.5143	9	2	2	0.2574	17	3	4	0.2431
Triweight kernel	11	3	3	0.5717	9	3	4	0.2805	19	5	5	0.3918
Gaussian kernel	5	9	3	0.5675	14	2	4	0.4521	17	2	3	0.4325
PLS	26	3	0	0.6876	17	3	0	0.4688	28	3	0	0.4192
OLS	12	7	0	0.7515	8	11	0	0.5742	16	7	0	0.4937

Through the numerical results of the first engine we also find that the search range of designed algorithm is enormous and it takes a long time, so we hope to further reduce search range and the algorithm complexity. The percentages of the number of NMSE in Table 1 which falls in $[\min(\text{NMSE}), \min(\text{NMSE}) + 0.03]$ are statistically analyzed, and the results are shown in Figures 4, 5 and 6.

As can be seen from Figure 4 to Figure 6, the percentage of NMSE falling in $[\min(\text{NMSE}), \min(\text{NMSE}) + 0.03]$ is lower for different value of m , τ , M , which indicates that the probability of obtaining suboptimal solution by our algorithm is very low, and also illustrates that the selection of variables is of significant importance to improve the prediction accuracy. From Figure 4 we can see that when the dimension of predicted model is over 15, there is no optimal NMSE which falls within the range above mentioned, So the range of m in the algorithm can be selected among ranging from 1 to 15. Figures 5 and 6 also show that the predicted NMSE distribution of various delay time and section number are more dispersed, so the search range of τ and M cannot be minimized.

The brief computational results of the other three engines are as in Table 3.

From the above calculation results we can see that the algorithm based on various nonlinear PLS (Cubic Spline Interpolation, Uniform Kernel, Epanechnikov Kernel, biweight Kernel, Tri-weight Kernel, and Gaussian Kernel) is superior to that based on PLS, and the algorithm based on PLS is superior to that based on OLS. The results in Tables 1 and 3 show that the predictive results derived from the algorithm based on biweight Kernel are the optimal among the four aeroengines. And there are no significant difference between the first aeroengines and the second aeroengines in the same aircraft in predictive NMSE, so we can conclude that the proposed algorithm is stable for predicting the aeroengine parameter EGT.

Tables 1 and 3 also show that the optimal values of the embedding dimension (m), delay time (τ), subset numbers (M) searched by various NPLS algorithm are approximately the same, but they are quite different among the four aeroengines.

6. Conclusion

The chaos prediction model based on NPLS through Kernel function or Cubic Spline function transformation is developed and is applied in predicting the aeroengine performance

parameters EGT. Numerical results show that the proposed model is able to effectively predict the performance parameter EGT with a higher degree of accuracy. The optimal embedding dimension is 8, the delay time is 19, and the number of segments is 2. NPLS chaotic prediction algorithm based on biweight Kernel function transformation is the optimal EGT prediction algorithm because of the smallest predictive NMSE. Our prediction NMSE is 16.5 percent less than that of the traditional linear least squares (OLS) method and 10.38 percent less than that of the linear PLS approach. The NPLS algorithm of chaotic prediction with selection variables through Bootstrap test fails to make NMSE further reduced. And we find that reducing calculation time complexity by lessening search range of embedding dimension will not affect the prediction accuracy of EGT. It is proved that the proposed algorithm is robust by the predictive results from four aeroengines.

In particular, the conservative baselines of main aeroengine performance parameters are given in ECM program provided by jet engine manufacturers [14]. Set the baseline value of EGT be EGT_{base} , the deviation between EGT_{cor} and EGT_{base} can be defined as

$$\Delta EGT = EGT_{cor} - EGT_{base}. \quad (6.1)$$

The trend plot of the deviation (ΔEGT) sequence in continuous flights is routinely used by airlines for the performance monitoring and diagnostics [13]. On the basis of our prediction algorithm, the EGT_{cor} of upcoming flights can be predicted, and then a layout for trending of the deviation of main performance parameters for upcoming flights can be provided. Experienced power maintainer or engineer look at a given trend shift in the measurement deviations can identify significant behavior disparities, evaluate the aeroengine condition, and isolate the faulty module or component using the ECM trend plot reports and related PW4000 fingerprint.

In summary, accurately forecasting the aeroengine performance parameters is of great importance in the aeroengine condition monitoring. The anomalies of the aeroengine performance parameters can be detected in time by the predictive value calculated from our algorithm, then maintenance measures are taken early to prevent the aeroengine from failure effectively, which assists to reduce aircraft down time, and improve the reliability of the aeroengine.

Acknowledgment

This work is supported by the Fundamental Research Funds for the Central Universities under Grant no. ZXH2012 K005.

References

- [1] E. Ortiz, A. Babbar, and V. L. Syrmos, "Extreme value theory for engine health monitoring and diagnosis," in *Proceedings of the IEEE International Conference on Control Applications (CCA '09)*, pp. 1069–1074, Saint Petersburg, Russia, July 2009.
- [2] A. Babbar, E. M. Ortiz, V. L. Syrmos, and M. M. Arita, "Advanced diagnostics and prognostics for engine health monitoring," in *Proceedings of the IEEE Aerospace Conference*, March 2009.
- [3] Z. Gao, X. Dai, T. Breikin, and H. Wang, "Novel parameter identification by using a high-gain observer with application to a gas turbine engine," *IEEE Transactions on Industrial Informatics*, vol. 4, no. 4, pp. 271–279, 2008.

- [4] Z. Gao, T. Breikin, and H. Wang, "High-gain estimator and fault-tolerant design with application to a gas turbine dynamic system," *IEEE Transactions on Control Systems Technology*, vol. 15, no. 4, pp. 740–753, 2007.
- [5] T. S. Parker and L. O. Chua, *Practical Numerical Algorithms for Chaotic Systems*, Springer, New York, NY, USA, 1989.
- [6] R. Gençay and T. Liu, "Nonlinear modelling and prediction with feedforward and recurrent networks," *Physica D*, vol. 108, no. 1-2, pp. 119–134, 1997.
- [7] W. Niu, G. Wang, Z. Zhai, and J. Cheng, "Prediction of chaotic time series based on rough sets and support vector machine," *International Journal of Digital Content Technology and Its Applications*, vol. 5, no. 9, pp. 145–152, 2011.
- [8] L. X. Wang and J. M. Mendel, "Generating fuzzy rules by learning from examples," *Institute of Electrical and Electronics Engineers. Transactions on Systems, Man, and Cybernetics*, vol. 22, no. 6, pp. 1414–1427, 1992.
- [9] D. Kugiumtzis, O. C. Lingjærde, and N. Christophersen, "Regularized local linear prediction of chaotic time series," *Physica D*, vol. 112, no. 3-4, pp. 344–360, 1998.
- [10] G. Ding and S. Zhong, "Approximation capability analysis of parallel process neural network with application to aircraft engine health condition monitoring," in *Proceedings of the 4th International Symposium on Neural Networks (ISNN '07)*, vol. 4493 of *Lecture Notes in Computer Science*, pp. 66–72, 2007.
- [11] C. Zhang and J. Yue, "Application of an improved adaptive chaos prediction model in aero-engine performance parameters," *WSEAS Transactions on Mathematics*, vol. 11, no. 2, pp. 114–124, 2012.
- [12] L. Zunxiong and L. Jianhui, "Chaotic time series multi-step direct prediction with partial least squares regression," *Journal of Systems Engineering and Electronics*, vol. 18, no. 3, pp. 611–615, 2007.
- [13] E. J. Mo, M. S. Jie, C. S. Kim, and K. W. Lee, "Gas turbine engine fault isolation based fuzzy inference logic with ECM trend plot report," in *Proceedings of the International Conference on Control, Automation and Systems (ICCAS '07)*, pp. 454–458, Seoul, Korea, October 2007.
- [14] Pratt&Whitney Company, "ECM II Training manual," Section 2, 1-34, 1994.
- [15] D. M. Hawkins, *Identification of Outliers*, Chapman & Hall, London, UK, 1980.
- [16] L. D. Iasemidis, K. Tsakalis, J. C. Sackellares et al., "Comment on 'inability of Lyapunov exponents to predict epileptic seizures'," *Physical Review Letters*, vol. 94, no. 1, Article ID 019801, 2005.
- [17] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, "Determining Lyapunov exponents from a time series," *Physica D*, vol. 16, no. 3, pp. 285–317, 1985.
- [18] Z. Wang, *Of Time Series*, China Statistics Press, Beijing, China, 2005.
- [19] J. Zeyu, Y. Yue, and Z. Wenzhong, "Application of improved partial least square regressive model in power load forecasting," *Power DSM*, vol. 13, no. 1, pp. 10–14, 2011.
- [20] D. V. Pisarenko and V. F. Pisarenko, "Statistical estimation of the correlation dimension," *Physics Letters A*, vol. 197, no. 1, pp. 31–39, 1995.