Application of an improved adaptive chaos prediction model in aero-engine performance parameters

CHUNXIAO ZHANG
College of Science
Civil Aviation University of China
Tianjin
CHINA
cxzhang@cauc.edu.cn

JUNJIE YUE
College of Aeronautical Engineering
Civil Aviation University of China
Tianjin
CHINA
jjyue_yjs09@cauc.edu.cn

Abstract: Based on the research of complexity and non-linearity of aero-engine exhaust gas temperature (EGT) system, a regularization adaptive chaotic prediction model applied in short time forecasting of EGT was proposed. In this research, we develop a new hybrid particle swarm optimization (HPSO) arithmetic in order to improve the accuracy of the forecasting model. This arithmetic enhanced the ability of dealing with integer variables and constraints by adding and changing some manipulations to fit in with optimizing continuous and integer variables. The test results are based on QAR data supplied by a civil airline company, and show that the proposed framework performs better than the traditional chaotic forecasting model on prediction accuracy. Therefore, this arithmetic is efficient and feasible for a short-term prediction of aero-engine exhaust gas temperature.

Key–Words: Exhaust gas temperature (EGT); Regularization; Adaptive chaos prediction ; Hybrid particle swarm optimization(HPSO); Principal component regression(PCR);Aero-engine.

1 Introduction

Advanced monitor and prognostic schemes to determine engine condition is important for modern civil aircraft in order to reduce unnecessary maintenance action and improve aircraft safety. Thus, the current estimates of engine condition are necessary before upcoming fights to avoid delays. In fact, abnormality of engine can be identified by monitoring performance parameters of engine. Therefore, we can accurately realize the engine condition monitoring and fault prognostic by predicting various performance parameters.The prediction on engine performance parameters can provide adequate time and decision-making basis for the prevention and troubleshooting, which is important for the effective implementation of aero-engine condition monitoring and condition-based maintenance.

At present, many domestic and foreign experts and scholars in the research of on-wing engine short-term forecast have researched several methods and made vary degrees of results. In [1-3], time series model based on performance parameters trend prediction of aero-engine were proposed, but those traditional time series analysis models were a linear model for smooth, zero-mean, normally distributed random sequence, so it is not suitable for non-linear forecasting. In [4], a chaos arithmetic for civil aero-engine forecasting model was built and solved by ordinary least squares (OLS), the model existed the serious multicollinearity that will increase the predicting error. Some researchers and scholars forecasted engine systems reliability by neural network mode[5], they improved BP neural network to dynamically forecast read-time, but the over learning and unstable training of neural network were insurmountable problems. In [6], the Kalman filter method was proposed to estimate the aero-engine health parameters. Kalman filter shows its superior ability in solving forecasting problem and improving the prediction accuracy of time series model. In fact , the efficiency of Kalman filter is difficult to evaluated because of promiscuous noise.

Aero-engine system is a time-varying, nonlinear, non-stationary, stochastic system, so this paper will adopt a short-term chaotic prediction model of engine performance parameters EGT. With the shorten prediction time interval, the short-term requirements of prediction accuracy of the performance parameters would be higher. The parameters of a chaotic prediction model include integer variables and continuous variables, so we need to select the appropriate mixed-integer optimization algorithm [7][8] for optimal estimation of nonlinear on-wing engine performance parameters. The hybrid particle swarm optimization algorithm is a kind of the optimization algorithms, which adaptively selects the embedding dimension, delay time, the number of adjacent points and the regularization parameter to obtain better predicted outcome than fixed values of these parameters.

Modern heuristic algorithms are considered as effective tools for nonlinear optimization problems [9]. The objective function is not necessary to be differentiable and continuous in these algorithms. A particle swarm optimization (PSO) is one of the modern heuristic algorithms [10][11] and can be applied to nonlinear and non-continuous optimization problems with continuous variables. It has been developed through simulation of simplified social models. A hybrid PSO (HPSO) adds a selection mechanism of evolutionary computation (EC) to PSO and it can generate a high-quality solution within short calculation time [12].

In this research,we propose on-wing engine performance parameters regularization adaptive prediction model using an HPSO. The proposed method can greatly improve on-wing engine short-term forecast accuracy and efficiency, and effectively overcome the problem of strong non-linear, random and difficult prediction of the aero-engine exhaust temperature system.

This paper is organized as follows. Firstly, we take the EGT time series analyzing in section 2. Section 3 presents the operation of the adaptive chaos prediction algorithm based on regularization. We introduce the evaluation function and algorithm's detail. A improved hybrid particle swarm optimization are given in section 4. Flow chart of algorithms is discussed in section 5. Section 6 discusses and compares result of prediction. A conclusion will be drawn in section 7.

2 Time series analyzing

2.1 EGT's modification

High quality of data is necessary before processed by any arithmetic theories. But actual data are always disturbed by all kinds of factors and can not meet requirement of arithmetic processing , therefore, actual data must be processed before modeling to decrease chaos as much as possible.

Conditions of the environment (such as outside overall temperature) of an airplane in different flights are different. On principle of the relative theory of aero-engine, the important performance parameters of the same motor under different working conditions, are different greatly, thereby actual performance parameter data acquiring from different flights cannot be compared directly. According to this problem, similar theory offers a good method to eliminate influence on performance parameters of aero-engine caused by different working conditions.

Similar theory indicates that, on the precondition of geometrical similarity, if ratios of cognominal physical measures at corresponding sections are equal to each other respectively, working states of aero-engine are similar without considering the all kinds of different working environment aero-engine undergoing. Plenary and necessary condition of the similarity of one motor's working state is conservation of the flight Mach number and the conservation rotational speed.

According to the similar theory of aero-engine, EGT can be modified to the same atmospheric condition to eliminate the influence of different working conditions on aero-engine performance parameters and enhance the comparable of different flight data[13].

$$EGT\_cor = \frac{EGT\_raw}{\theta T2} \tag{1}$$

$$\theta T2 = \frac{273.15 + TAT}{288.15} \tag{2}$$

Where EGT_cor is the standardized EGT data, EGT_raw is the original data; TAT is the total air temperature.

2.2 Outliers detection

Outlier is defined as an observation that deviates too much from other observations that it arouses suspicions that it was generated by a different mechanism from other observations [15]. The identification of outliers can lead to the discovery of useful and meaningful knowledge and has a number of practical applications in areas such as transportation, ecology, public safety, public health, climatology, and location based services. Recently, a few studies have been conducted on outlier detection for large data set [16]. However, most existing research focuses on the algorithm based on special background, compared with outlier detection approach is still rare. This paper adopts one technique which is based on traditional outlier detection techniques.

S is regarded as the standard deviation of a measurement [14].

$$S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{N-1}} \tag{3}$$

Where $x_1, x_2, , x_N$ is the observed time series, N is the length of this series, $\bar{x}$ is the average of this series. If $x_i - \bar{x} \geqslant 3S$, the data x_i should be detected ; then $\bar{x}$ and S should compute again.

3 Adaptive chaos prediction algorithm based on regularization

Nonlinear analysis methods developed in nonlinear dynamics systems are applied to time series data in order to capture as much of the underlying structure as possible. One key question is whether the dynamics of the system can be reconstructed on the basis of the data given. Based on the case in which the considered system is deterministic, the dynamics unambiguously evolves one state in phase space into another. Therefore, the reconstruction of phase space is a fundamental problem that plays the key role in many applications: as long as the state space reconstruction remains unjustified ,all of the consequent analysis may be wrong. From this point of view, we discuss the phase space embedding of realistic signals in the example of the aero-engine performance parameters. An introduction to nonlinear dynamics is found in [17][18]. An account of nonlinear time series methods is given by [19].

Let the observed chaotic time series be denoted by$\{x_n\}_{n=1}^{N}$ .Following the approach introduced by Packard et al. [20], and a firm mathematical basis put on by Taken [21], we reconstruct a phase space as below. The phase space is

$$S_n = (x_n, x_{n-\tau}, \cdots, x_{n-(m-1)\tau}) \in R^m \tag{4}$$

$n = (m-1)\tau + 1, (m-1)\tau + 2, \cdots, N$, Where τ is the delay time, m is the embedding dimension and N is the length of the signal.

Indeed, the reconstructed phase space can be converted into a multi-input single-output (MISO) model.

3.1 Regularized estimation

For forecasting of chaotic time series, the most commonly used approach is the phase space local approximation method [22] which predicts by interpolation based on neighbouring data points in the reconstructed phase space. Since this method of relying on the neighbouring data points ignores the dynamic information included in the entire system, the forecasted value may suffer a great error when some neighbouring data points are unusual.

Assume thats_T is the state vector occurred in time T. we will predict the value of x_{T+1},which occurred after one step. Prediction model is expressed as.

$$\begin{aligned} \hat{x_{T+1}} &= \hat{GS_T} = c_0 + c_1 x_T + c_2 x_{T-\tau} + \\ &\cdots + c_m x_{T-(m-1)\tau} + e \end{aligned} \tag{5}$$

Here $c_0, c_1, \cdots, c_m$ are undetermined coefficients and e denotes random error. $E(e) = 0$ and $Var(e)$ =constant. Calculate the Euclidean distance between the target points s_T and various reconstruction vectors. Let $S_{\alpha_k}(k = 1, 2, \cdots, K)$ be the nearest neighbouring points of the target points s_T.

$$S_{\alpha_k} = (x_{\alpha_k}, x_{\alpha_{k-\tau}}, \cdots, x_{\alpha_k-(m-1)\tau}) \tag{6}$$

The problem is transformed to sake the $c_0, c_1, \cdots, c_m$, that meet the following matrix equation

$$y = Xc \tag{7}$$

Where

$$y = [x_{\alpha_1+p}, x_{\alpha_2+p}, \cdots, x_{\alpha_K+p}] \in R^K \tag{8}$$

$$X = \begin{pmatrix} 1 & x_{\alpha_1} & \cdots & x_{\alpha_1-(m-1)\tau} \\ 1 & x_{\alpha_2} & \cdots & x_{\alpha_2-(m-1)\tau} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{\alpha_K} & \cdots & x_{\alpha_K-(m-1)\tau} \end{pmatrix} \tag{9}$$

If the matrix XX^T is invertible, the above equation can be expressed as

$$c_{OLS} = (X^T X)^{-1} X^T y \tag{10}$$

c_{OLS} is regarded as the ordinary least square estimation, abbreviated as OLS.

According to singular value decomposition (SVD) theorem:

$$X = U\Sigma V^T \tag{11}$$

Where,

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{pmatrix} \tag{12}$$

$$U = [u_1, u_2, \cdots, u_r] \in R^{k \times r},$$
$$V = [v_1, v_2, \cdots, v_r] \in R^{m \times r}$$

To take (11) into (10), we get

$$\hat{c_{OLS}} = \sum_{i=1}^{r} \frac{(u_i, y)}{\sigma_i} v_i \tag{13}$$

Because the smallest singular values will make the larger variance of the least squares estimation. In order to reduce this effect, we remove the latter q items expression of the least-squares estimation, which is called principal component regression estimation, abbreviated as PCR. Above formula (13) becomes

$$\hat{c_{PCR}} = \sum_{i=1}^{q} \frac{(u_i, y)}{\sigma_i} v_i \tag{14}$$

where q is called the number of factors.

3.2 Method of selecting parameters K

The main steps of selecting parameters K[22] are as following.

Step1:Take $K = 2m+1, 2m+2, \cdots, 2m+10$.

Step2:Each of the above K calculate in accordance with the following formula.

$$H = X(XX^T)^{-1}X^T \tag{15}$$

$$D(K) = tr(H) \tag{16}$$

$$s(K) = \frac{[Y - X(XX^T)^{-1}X^TY]^T[Y - X(XX^T)X^Y]}{K - D(K)} \tag{17}$$

Step 3: Choose the smallest K ,which makes the minimum of $s(K)$. Denoted $\overset{*}{K}$;

Step 4: Construct a local prediction model according to $\overset{*}{K}$.

3.3 Evaluation function

In the practical application, a part of the time series is usually used for training, the other part for verifying. If the difference that between predictive value and actual value is fewer, the model is more ideal. This paper adopts the normalized mean squared error (NMSE) value as the evaluation function of the prediction accuracy, it can be defined as:

$$NMSE = \frac{MSE}{VAR} \tag{18}$$

Here VAR is the variance of the sequences; MSE is the mean square error of time series.

3.4 Chaotic forecasting arithmetic for aeroengine

When the matrix X is not full rank or its condition number is too large in the equation (11), the estimated parameters are unreliable. So it requires that improved X which is full rank and whose condition number is not greater than a given positive number M_1 to ensure the robustness of parameter estimating. To further reduce the PCR error, regularization parameter q should be a different value when the different target is predicted. Specific steps are as follows:

Step 1: Select the embedding dimension m, the delay time τ, threshold of the condition M_1 and threshold of determining the number of factors M_2, which are determined according to a new (HPSO) algorithm, which will be seen in section 4.They are greater than zero;

Step 2: Using the method of section 3.2, adaptively determine K-the number of the nearest neighbouring points and construct X, then X do SVD processing;

Step 3: If $\sigma_1/\sigma_r \langle M_1$, then is taken. Otherwise, perform step 4;

Step 4: $r = r - 1$; $\sigma_r = \sigma_{r-1}$;

Step 5: Repeat step 3 and step4 until $\sigma_1/\sigma_r \langle M_1$;

Step 6: Determine the appropriate embedding dimension. Using the singular value of X from 3.1 content, find out the value of q, which meet $\sigma_1/\sigma_r \langle M_2(i = 1, 2, \cdots, r)$ and the maximum subscript;

Step 7: If $m_0 = m$, compute $c_0, c_1, \cdots, c_m$ according to equation (14). Then use equation (5) to predict;

Step 8: For the new target point, we repeat the above steps until the end.

4 Hybrid particle swarm optimization arithmetic

4.1 Objective function and constraints

The objective function and constrains are constructed according to sector 3.

$$\begin{aligned} &\min NMSE(m, \tau, M_1, M_2) \\ &s.t. M_1 \geq M_2 \\ &M_1^L \leq M_1 \leq M_1^U \\ &M_2^L \leq M_2 \leq M_2^U \\ &m \in \{m^L, \cdots, m^U\} \\ &\tau \in \{\tau^L, \cdots, \tau^U\} \end{aligned} \tag{19}$$

where L represents the lower property value and U represents the upper property value.

4.2 Definition

PSO algorithm defines each particle as a potential solution to a problem in D-dimensional space. At each iteration, the particle's position is updated by a stochastic velocity which depends on its previous velocity, its own previous best solution and that of its neighbourhood. For standard PSO, the i-th particle's position and velocity updating formulas are given by:

$$\begin{aligned} v_{ij}(t+1) \quad = \quad & wv_{ij}(t) + c_1r_1(pbest_{i,j}(t) - z_{ij}(t)) + \\ & c_2r_2(gbest_j(t) - z_{ij}(t)) \\ & z_{ij}(t+1) = z_{ij}(t) + v_{ij}(t+1) \end{aligned} \tag{20}$$

Where $i = 1, \cdots, N, j = 1, \cdots., D$; N is the population size of particle swarm; $z_i(t)$ and $v_i(t)$ are the position and velocity of particle i at time step t; respectively$pbest_{i,j}(t)$ is the personal best position of particle i; $gbest_j(t)$is the best position found by the neighbourhood of particle i; w is inertia factor; c_1 and c_2 are acceleration coefficients; $r_1$and r_2 are random numbers uniformly distributed in [0, 1].

4.2.1 Constraint matrix

Create a $N \times 1$ dimensional constraint matrix C to reflect whether the solutions meet the constraints of each situation or not. The value of the matrix elements will compute according to the following formula.

$$c_i = \begin{cases} 0, & M_1 \geq M_2, \quad (21a) \\ M_2 - M_1, & otherwise. \quad (21b) \end{cases}$$

If $c_i = 0$,z_i is the feasible solution. Otherwise, it isn't the one. We calculate the degree of constraint violations according to the following formula.

$$A(z_i) = c_i \tag{22}$$

4.2.2 Pareto dominance

All the dominant individual z_i of the current population $pop(t)$ are denoted by

$$DS_i = \{z_l \in pop(t) : z_l < z_i\} \tag{23}$$

Then the degree of being dominated ofz_i can be recursively expressed as

$$D(z_i) = 1 + \sum_{l \in DS_i} D(z_l) \tag{24}$$

4.3 Creation and transformation of several operating

In order to adapt to solve mixed-integer nonlinear programming (MINLP), it requires the algorithm to enhance the processing capacity of constraints and integer variables. In this paper, we present some creation and transformation related with operations of particle swarm optimization (PSO) algorithm.

4.3.1 Velocity and position update

The flight speed of particle gradually become smaller as the iteration, which make species diversity gradually disappeared ,and thus species fall into local optimum. To this end, we can use local minima of the particle information to adjust the update formula of the speed of each component in order to maintain the diversity of the population [24][25]. In the four components, we randomly select $N1 = ceil(4{*}Pm)$ components. Then update their velocity according to the following formula (25).

$$v_{ij}(t+1) = wv_{ij}(t) + r(gbest_j(t) - z_{ij}(t)) \tag{25}$$

From the remaining 4-$N1$ components we randomly selected $N2 = ceil((4 - N1) \times Pc)$ component. Then according to their velocity,update the following formula(26).

$$v_{ij}(t+1) = wv_{ij}(t) + r(pbest_{ij}(t) - z_{ij}(t)) \tag{26}$$

Finally, the remaining 4-$N1$-$N2$ components are updated in accordance with the following formula (27).

$$\begin{aligned} v_{ij}(t{+}1) \quad = \quad & wv_{ij}(t) + r_1(pbest_{u(i),j}(t) - z_{ij}(t)) \\ + \quad & r_2(gbest_j(t) - z_{ij}(t)) \end{aligned} \tag{27}$$

Where $v_{ij}(t)$ denotes the j-th component of the speed of the particle i at time step t; $ceil$ is a function of rounding up; Pm and Pc are the pre-set probability, here they are selected as 0.5; w is the inertia weight; r,$r_1 and r_2$ are random numbers uniformly distributed in [0,1]. $pbest_{ij}(t)$ can represent the j-th component of local optimum of the particle i at time step t, $gbest_j(t)$ is the j-th component of the global optimum at time step t; $pbest_{u(i),j}(t)$ is the j-th component of local optimum of the particle at time step t and $u(i)$ is a random number uniformly distributed in [0, N]; N is the number of populations.

Particle's position is updated according to the following formula (28).

$$z_{ij}(t+1) = z_{ij}(t) + v_{ij}(t+1) \tag{28}$$

Where z_{ij} is the j-th component of position of the particle i at time step t .

4.3.2 Rounded to the integer variable accordance with the probability

A particle z_i updates its velocity and position according to equations (25) $\sim$ (28). In order to ensure that integer variables are integer-values,we need to take the rounding operation. The distance function[7] is defined as:

$$d_{ij}{}^{(k)} = |z_{i,n+j} - y_{jk}|; \quad i = 1, \cdots, N, j = 1, 2, k = 1, \cdots., mj \tag{29}$$

Where $d_{ij}{}^{(k)}$ is the absolute distance between integer variables $z_{i,n+j}$ and the k-th value y_{jk} of the range; mj is the number of integers, which are contained in the j-th variable. If $d_{ij}{}^{(k)} \neq 0$, $z_{i,n+j}$ is not an integer. Then, we should take the rounding operation. We select the k-th value of the range by the following formula (30) to calculate the probability of $p_{ij}{}^{(k)}$.

$$p_{ij}{}^{(k)} = \frac{1}{d_{ij}{}^{(k)}} / \sum_{l=1}^{mj} \frac{1}{d_{ij}{}^{(l)}}; k = 1, \cdots., mj \tag{30}$$

4.3.3 Determining the local and global best positions

For MINLP, this constraint-based matrix, this paper uses the concept of Pareto dominance [7][23] based on constraint matrix to evaluate the merits of the particle solution and determine the local optimum and global optimum .determining the local optimum through the following method. If $pbest_i(t-1) < z_i(t)$, $pbest_i(t)$ is equal to $pbest_i(t-1)$; if $z_i(t) < pbest_i(t-1)$ $pbest_i(t)$ is equal to $z_i(t)$; if $pbest_i(t-1)$ and $z_i(t)$ do not Pareto dominance, we discuss the points in three cases: a) $pbest_i(t-1)$ and $z_i(t)$ are feasible solutions.If $f(pbest_i(t-1)) < f(z_i(t))$,$pbest_i(t)$ is equal to $pbest_i(t-1)$.otherwise, $pbest_i(t)$ is equal to $z_i(t)$. b) $pbest_i(t-1)$ and $z_i(t)$ aren't feasible solutions. We randomly select one as the local optimum. c) If $pbest_i(t-1)$ and $z_i(t)$ are only a feasible solution, $pbest_i(t)$ is equal to that feasible solution.

Determine the global optimum: Check whether the set contains a feasible solution. If there are some feasible solutions, we take the minimum value of objective function from the set of feasible solutions as $gbest(t)$ .Otherwise, calculate the dominated degree of the particles in the set $\{pbest_i(t)\}_{i=1}^{N}$. All the particles whose dominated degree is equal to 1.0 constitute non-inferior set $NDS(t)$, in which the number of particles is $N_{NDS(t)}$ .we randomly select a particle from the above set as $gbest(t)$.

4.3.4 Stochastic mutation based solution repair strategy

The MINLP exists both constraints and integer variables, which leads to the PSO algorithm is often time-consuming. To speed up the search excellent speed, the algorithm should be added the solution of the repair operation. Repair operation is only the continuous variables of unfeasible solution $z_i(t)$ as follows:

$$z_{ij}(t) = z_{ij}(t) + \delta(t)(x_{ij}^U - x_{ij}^L)(0.5 - r) \tag{31}$$

If $z_{ij}(t) < x_j^L$,do $z_{ij}(t) = x_j^L$; if $z_{ij}(t) > x_j^U$,do $z_{ij}(t) = x_j^U$;if $z_{ij}(t)$ becomes a feasible solution after mutation, the repair is to be completed. Otherwise, determine the local optimum of the evaluation principles according to section 4.3.3 and compare the pros and cons between before variation and after variation. If there is no improvement, we eliminate the variation. Then the execution of the next variation of continuous variables continues.When all continuous variables are mutated, the solution is still unfeasible. The repair process can be repeated many times and repetitions recorded as SRN. $\delta(t)$ are random numbers uniformly distributed in [0,1].

4.3.5 Multi-particle swarm strategy

To further enhance the diversity of the population and improve the ability of global optimization algorithms, this paper adopts a simple multi-particle swarm strategy. SN particle groups are established and each particle group runs independently in parallel. After NT iteration, the global optimum of the whole population is determined, which comes from their respective global optimum from the global optimum of SN subgroups to in and pass it to each sub-groups.

4.4 Hybrid particle swarm optimization steps

The main steps of the modified hyhrid particle swarm optimization are as following.

Step 1:Set the parameters: $N, SN, Lmt, w,$ SRN and NT;

Step 2: Let $t = 0$ and initialize the various subgroups. Rounded to the integer variable according to sector 4.3.2 and local optimum of the particle are themselves;

Step 3:Set $t = t+ 1$, each sub-group independently carry out 4)$\sim$6) operation;

Step 4:Create the constraint matrix of the population to determine the local optimum and global optimum according to sector 4.3.3;

Step 5: Perform a stochastic mutation based solution repair strategy and update velocity and position according to sector 4.3.1 and 4.3.4;

Step 6: Rounded to the integer variable accordance with the probability according to 4.3.2;

Step7: After each round of iteration NT, the global optimum of the whole population is determined and the global optimum of the sub-group is passed to the sub-group;

Step 8: Determine whether the termination condition has been reached or not. If the conditions are met, output the global optimal and stop. Otherwise, go to step 3 .

5 Flow chart of algorithms

As the paper used a combination of the two algorithms in this paper,the logical structure of algorithms is more complex. In order to explain the calculation, this paper presents the flow chart of algorithms. The detailed steps of the improved adaptive chaos prediction model based on are as following in the Fig 1.

6 Simulation results

The most work hours of aeroengine are in the cruise, so the performance parameters of engine at this flight condition are critical. Therefore, the forecasted EGT value in the cruise can provide data support for aeroengine to determine whether it is an exception or not. The EGT data series in this article is from an Air China Boeing 777 passenger aircraft engines for nearly one year.The actual measurement EGT data is shown in Fig 2.

Firstly, we proposed a method to amend the time series of aero-engine exhaust temperature according to section 2.1, and the result is shown in Fig 3.

Secondly, we removed abnormal data of above set by the Pauta standard according to section 2.2. And the result of outliers detection is shown in Fig 4.

Comparing Fig 3 and Fig 4, we can see that the original EGT data have some outliers, so we should detect these outliers in order to better reflect the true model.

Then, set the initial parameters of hybrid particle swarm optimization algorithm, where N=20, w=0.5, Lmt=1000, SN=2, NT=20, SRN=5. Through these methods of the section 2∼4, use the former EGT data to train the chaos prediction model, thus get different $c_0, c_1, \cdots, c_m$, then use the trained model predict the latter 20 points and get some prediction results of different methods.When $M_1 = 124.4336$,$M_2 = 52.0352$,$m = 24$ and τ =2, we find the optimal NMSE which is equal to 0.2585. The results of some chaotic prediction models are shown in Table 1 and Table 2.

As can be seen from Table 1, different forecasting methods obtain different chaotic characteristics (embedding dimension and delay time) of optimal NMSE. Secondly, this paper proposed adaptive chaotic prediction algorithm based on hybrid particle swarm optimization, which can make the prediction accuracy to 0.2585.This prediction accuracy is much better than the prediction accuracy of OLS and PCR.Because the data information is stored in the variance,the proposed evaluation function NMSE is more reasonable than the relative error (the deviation value divided by the original value).

Table 1: Chaotic characteristics of prediction

	Adaptive chaotic prediction	OLS prediction	PCR prediction
m	24	30	9
τ	2	16	3
NMSE	0.2585	0.7081	0.5662

Form Table 2 we can see that, thought the NMSE of adaptive chaotic prediction model is the least, not all of the predictive values are optimal. the comparison result is shown in Fig 5.

Table 2: Actual data and forecasting result of EGT ^0C

	Original signals	Adaptive chaotic prediction	OLS prediction	PCR prediction
1	444.89	442.94	428.55	439.62
2	431.28	427.38	432.11	445.78
3	418.81	426.92	418.99	430.86
4	422.16	427.75	419.62	418.46
5	446.80	445.67	436.49	434.51
6	446.80	439.01	443.13	433.93
7	430.00	433.06	442.96	427.17
8	453.59	441.57	431.71	442.39
9	451.38	445.06	443.20	448.50
10	460.34	457.10	452.18	451.21
11	460.34	464.76	462.60	451.58
12	442.30	443.21	442.81	453.00
13	438.01	434.50	439.10	436.49
14	449.43	441.52	431.25	445.47
15	454.78	441.14	439.53	450.32
16	426.03	422.19	438.55	448.05
17	426.14	420.41	408.19	428.37
18	435.08	438.88	430.47	426.76
19	435.46	440.42	426.53	435.46
20	418.88	430.15	427.76	432.84

In Fig 5, through comparing with OLS prediction and PCR prediction model, prediction result of adaptive chaotic prediction model is closer to the actual developing trend, prediction error is smaller and forecast accuracy is higher.

Calculate the deviation of predicted values and original values. Forecast error as shown in Fig 6.

Fig 6 shows the results of the experiment error using many chaotic prediction models. These results are remarkably different. We can see that the result of adaptive chaotic prediction model has the smallest fluctuation. So the model in this paper is a good method to obtain better prediction results.

We make a simple statistical analysis for forecast errors to analyze the number and percentage of errors within a certain range, and the specific data is shown in Table 3.

Table 3: The number and percentage of errors

Deviation range	Adaptive chaotic prediction	OLS prediction	PCR prediction
0∼5	11(55%)	8(40%)	8(40%)
5∼10	6(30%)	4(20%)	4(20%)
above 10	3(15%)	8(40%)	8(40%)

Form Table 3 we can see that the number of the optimal prediction error of adaptive chaotic prediction model falling within a far smaller range is more than the other two prediction methods. If we use the method in this paper, 85% of forecast error of the engine exhaust temperature will fall with in 10^oC. It provides good protection to better diagnose engine condition.

7 Conclusion

In this research, we developed an adaptive regularization chaotic prediction method incorporated with a hybrid particle swarm optimization algorithms, to solve the prediction problem of on-wing engine performance parameters. The framework included three components, the adaptive chaotic model, the hybrid algorithms, and Numerical tests, utilizing QAR data from a civil airline company's operations, were performed to evaluate the predicted NMSE. During the testing process, several chaotic prediction algorithms were also performed, and we compared with the NMSE of different algorithms, to demonstrate the advantages of our method.

In practice, aero-engine system is a time-varying, nonlinear, non-stationary, stochastic system, the monitoring process of engine condition is usually performed based on a flight data by the prediction of engine condition parameters, making integrated problems difficult to solve in a closed form. Since it is difficult to extend the current optimization methods to deal with the integrated problem, we aim to develop a heuristic approach embedded particle swarm algorithms to explore the solution space more effectively on reaching the solution and enhance the prediction accuracy. The test results show that 55% of forecast error of the engine exhaust temperature fall within 5^oC and 85% of forecast error of the engine exhaust temperature fall within 10^oC.

In addition, the comparison result of different algorithms shows that both the solution quality and computational efficiency of our method perform well. The results also show that proposed forecasting model has a clear advantage in multicollinearity of variables and the stability problem of regression equation.

Finally, in future we hope to explore an integrated model or algorithm for predicting various performance parameters, so that related optimization theories or issues can be investigated. Nevertheless, the heuristic approach and the test results discussed in this research should be useful as a reference for the development of the engine condition monitoring and fault diagnosis.

Acknowledgements: This work is supported by the Fundamental Research Funds for the Central Universities under Grant No.ZXH 2009B003 and Research Funds in Civil Aviation University of China No.06kys06.

References:

[1] Shirong Ge, Hua Zhu, Complicate tribological systems and quantitative study methods of their problems, *Tribology.* 22, 2002, pp. 405–408.

[2] Minliang Gan, Hongfu Zuo, Zhong Yang, et al, The method of time series modeling and its application in the spetral analysis of lubricating oil, *Spectroscope and Spectral Analysis.* 20, 2000, pp. 64–67.

[3] Yunxue Song, Kexing Zhang, Yongsheng Shi, Research on aeroengine performance parameters forecast based on multiple linear regression forecasting method, *Journal of Aerospace Power.* 24, 2009, pp. 427–431.

[4] Yunxue Song, Peng Peng, Yongsheng Shi, Chaos arithmetic for civil aeroengine forecasting, *Journal of Aerospace Power.* 26, 2010, pp. 675–679.

[5] K. Xu , M. Xie , C. Tang, Application of neural net works in forecasting engine systems reliability, *Applied Soft Computing.* 2, 2003, pp. 255–268.

[6] Yonghua Wang, Tao Sun, Xiuxia Wang, Research on the model of aeroengine based on Kalman filter, *Gas Turbine Technology.* 24, 2011, pp. 31–34.

[7] Yijun He, Dezhao Chen, Hybrid particle swarm optimization algorithm for mixed-integer nonlinear programming , *Journal of Zhejiang University (Engineering Science).* 42, 2008, pp. 747–7519.

[8] Zhao Liu, Lishan Kang , Liangxiao Jiang, et al, New PSO algorithm for MINLP problems, *Minimicro Systems.* 26, 2005, pp. 991–993.

[9] J. Kennedy, R. Eberhart, Particle swarm optimization, *IEEE International Conference on Neural Networks.* 4, 1995, pp. 1942–1948.

[10] H. Yoshida, K. Kawata, Y. Fukuyama, et al, A particle swarm optimization for reactive power and voltage control considering voltage security assessment, *IEEE Transactions on Power Systems.* 15, 2000, pp. 1232–1239.

[11] P. Angeline, Using selection to improve particle swarm optimization, *Evolutionary Computation Proceedings,IEEE World Congress on Computational Intelligence.* 1998. pp. 84–89

[12] Yun Wang, *Aircraft Engine Principles,* Beihang University Press, Beijing 2009.

[13] Jinwen Liang, *Error Theory and Data Processing,* China Metrology Press, Beijing 2001.

[14] D. M. Hawkins, *Identification of Outliers,* Chapman and Hall, London–New York 1980.

[15] C. C. Aggarwal, P. S. Yu , An effective and efficient algorithm for high-dimensional outlier detection, *The VLDB Journal.* 14, 2005, pp. 211–220.

[16] D. Kaplan , L. Glass, *Understanding Nonlinear Dynamics,* Springer–Verlag, New York–Berlin–Heidelerg 1995.

[17] E. Ott, *Chaos in Dynamical Systems,* Cambridge University Press, England 1993.

[18] H. Kantz, T. Schreiber, *Nonlinear Time Series Analysis,* Cambridge University Press, England 1997.

[19] Packard. N. H, Crutchfield. J. P, Farmer. J. D, Geometry from a time series, *Physical Review Letters.* 45, 1980, pp. 712–715.

[20] F. Takens, Detecting strange attractors in turbulence, *Dynamical Systems and Turbulence.* 898, 1981, pp. 366–381.

[21] Tien-Yien Li, James A.Yorke, Period three implies chaos. *Physical Review Letters.* 59, 1987, pp.845–848.

[22] JianMing Yi, On measuring and correcting the effects of data mining and model selection, *Journal of the American Statistical Association.* 93, 1998, pp. 120–131.

[23] Xiaofeng Xie, Wenjun Zhang, Zhilian Yang, Hybrid particle swarm optimizer with mass extinction, *International Conference on Communications, Circuits and Systems and West Sino Expositions.* 2, 2002, pp. 1170–1173.

[24] J. J. Liang, A. K. Qin, S. Baskar, et al , Evaluation of comprehensive learning particle swarm optimizer , *Lecture Notes in Computer Science.* 3316, 2004, pp. 230–235.

[25] Haiyan Wang , Shan Lu, *Nonlinear Time Series Analys and its Application,* Science Press, Beijing 2006.

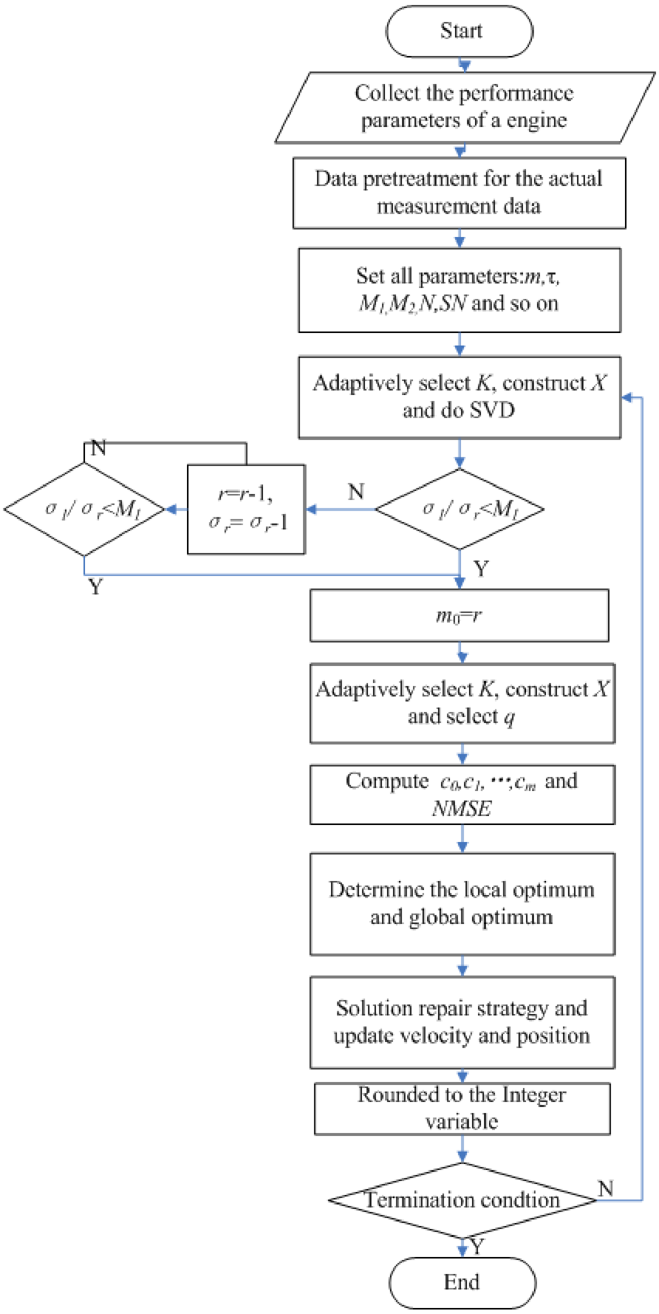

Fig. 1: Flow chart

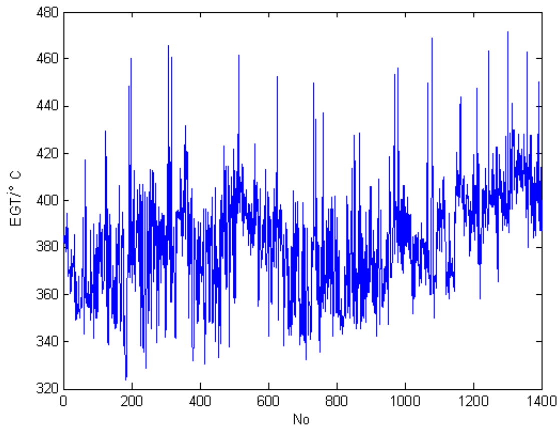

Fig. 2: The original values of unstandardized EGT

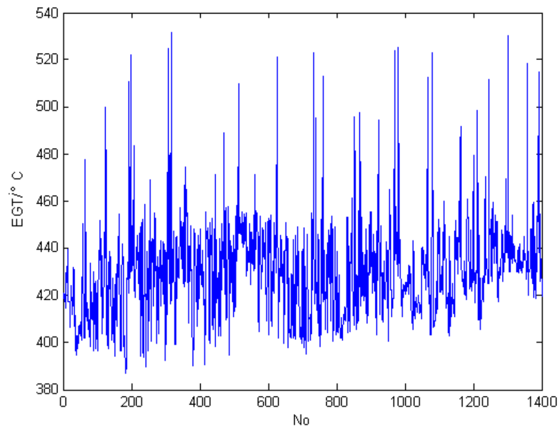

Fig. 3: The original values of standardized EGT

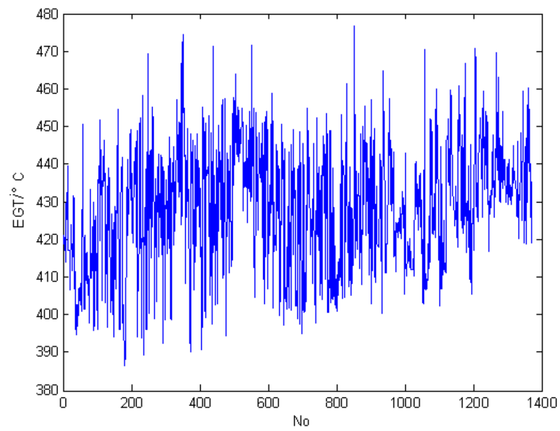

Fig. 4: The original EGT values detecting outliers

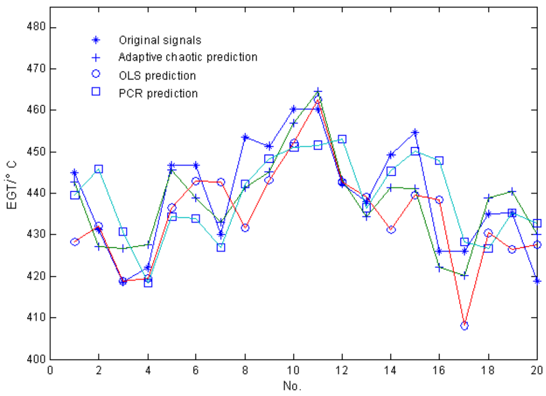

Fig. 5: The original and predicted values of EGT

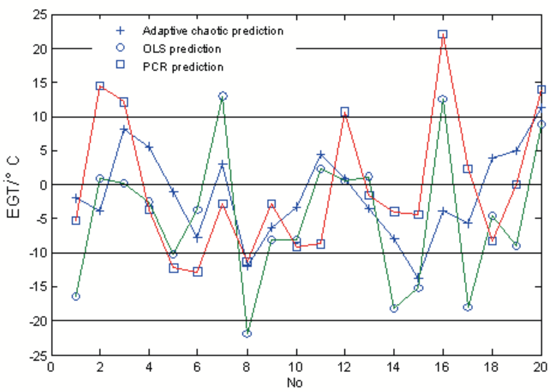

Fig. 6: Predicted errors of EGT's forecasting result