Chaotic Characteristic Analysis of Air Traffic System

Cong Wei(丛玮), Hu Minghua(胡明华)*

College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing, 210016, P. R. China

(Received 22 October 2013; revised 16 March 2014; accepted 30 March 2014)

Abstract: Chaotic characteristics of traffic flow time series is analyzed to further investigate nonlinear characteristics of air traffic system. Phase space is reconstructed both by time delay which is built through mutual information, and by embedding dimension which is based on false nearest neighbors method. In order to analyze chaotic characteristics of time series, correlation dimensions and the largest Lyapunov exponents are calculated through Grassberger-Procaccia (G-P) algorithm and small-data method. Five-day radar data from the control center in Guangzhou area are analyzed and the results show that saturated correlation dimensions with self-similar structures exist in time series, and the largest Lyapunov exponents are all equal to zero and not sensitive to initial conditions. Air traffic system is affected by multiple factors, containing inherent randomness, which lead to chaos. Only grasping chaotic characteristics can air traffic be predicted and controlled accurately.

Key words: air traffic; chaos; phase space reconstruction; correlation dimension; the largest Lyapunov exponent **CLC number:** U8 **Document code:** A **Article ID:** 1005-1120(2014)06-0636-07

1 Introduction

Air traffic system is a nonlinear dynamic system involving aircraft, airspace structure, pilots and controllers. Traffic situations involving different traffic flows are full of nonlinear characteristics such as certainty and randomness, order and disorder, contingency and inevitability, quantitative change and qualitative change. Recently, considerable progresses have been made in air traffic system complexity. Many concrete concepts and evaluation methods have been put forward such as air traffic complexity, airspace complexity, dynamic density, air traffic control complexity and cognitive complexity. Although chaos is an important theme in nonlinear science, researches on air traffic chaos have not been seen obviously.

Generally, chaos is a seemingly irregular and stochastic phenomenon in a deterministic system. Chaos is not a simple disorder, but an ordered structure with rich interior arrangements and no obvious period and symmetry, which is a new existence in nonlinear system^[1-3]. In the field of air

traffic system, Li Shanmei and Xu Xiaohao used the largest Lyapunov exponent to prove that flight conflict time series were chaotic^[4]. In the field of ground transportation, many scholars have made significant progress in chaos research. There are massive in-depth studies about chaos concepts, basic characteristics, identification methods, predictions and other aspects^[1-2,5-8]. For example, the ground traffic flow is chaotic, and a short term prediction for the flow based on chaotic characteristics can be achieved^[1,5-8]. In order to improve and perfect the analyses of air traffic system nonlinear characteristics and manage air traffic more effectively, it is necessary to distinguish and analyze air traffic system chaos.

In this paper, time delay and embedding dimension are determined using mutual information method and false nearest neighbors method respectively. Then a phase space is reconstructed for a traffic flow time series based on Takens embedding theorem. The correlation dimension is calculated through Grassberger-Procaccia (G-P) method and the largest Lyapunov exponent is calculated through the small-data method. The chaculated

otic characteristics of air traffic flows are analyzed on the basis of these two kinds of characteristics.

2 Chaos Identification Method

The commonly used quantitative analysis methods for chaos include: surrogate data method, the largest Lyapunov exponent method, correlation dimension method, Kolmogorov entropy method, Poincare section method, power spectrum exponent, etc^[3]. Attractor is a point set or subspace in a phase space, whose tracks tend to approach attractor as time passes by. Chaotic attractors exist in chaos system, known as strange attractors. The dimension of chaos system is fraction, and their tracks are regular. The existence of chaos in the system can be determined by two basic characteristics: One is whether an attractor in phase space is of a self-similar fractal structure, the other is whether the system is sensitive to the initial condition. Based on the reconstructed phase space, correlation dimension and the largest Lyapunov exponent are calculated and used to analyze chaotic characteristics of attractors.

Air Traffic has nonlinear characteristics, including uncertainty, universality, invisiblity and unexpectedness, due to different controllers' actions, unpredictable weather conditions, complex airspace structure and many other factors. The uncertainty reflects the randomness of chaotic phenomenon. The conductivity shows that the results sensitively rely on initial conditions. Therefore, air traffic evolution cannot be described by determined mathematical equations. However, it is a good choice to study on data of observable variables or metrics. There are many kinds of meaningful metrics in the field of air traffic and the number of aircraft is the most common and important one. Therefore, air traffic flow time series consist of a set of metric values (number of aircraft) which are sorted according to the time.

2.1 Phase space reconstruction

Takens theorem proves that a one-dimension time series can be reconstructed into a phase space which is equivalent to the original dynamic system in topological sense, so properties and regularities of time series can be analyzed through reconstruction phase space. For this reason, phase space reconstruction is very important for the research on chaotic time series $^{[9]}$. Since time series of actual problems do have a limited length and contain noisy data, confirmation of embedding dimension m and time delay τ is crucial for phase space reconstruction.

2.1.1 Time delay

If time delay τ gets smaller value, the values of each two components x $(t+j\tau)$ and $x(t+(j+1)\tau)$ of X $(t)=\{x(t),x(t+\tau),x(t+2\tau),\cdots,x(t+(m-1)\tau)\}$ are too close to be distinguished. If time delay τ gets bigger value, $x(t+j\tau)$ and $x(t+(j+1)\tau)$ are likely to be completely independent from each other. That is to say, there is no relevance between projections of two chaotic attractors' tracks. Mutual information function can measure linear or nonlinear relevance of random variables and it is used to confirm time delay^[3].

Assume that $S = \{s_1, s_2, \dots, s_n\}$ and $Q = \{q_1, q_2, \dots, q_m\}$ are discrete information series. Based on Shannon's theory of communication, information entropy of S and Q can be expressed as

$$H(S) = -\sum_{i=1}^{n} P_{s}(s_{i}) \log_{2} P_{s}(s_{i})$$
 (1)

$$H(\mathbf{Q}) = -\sum_{i=1}^{m} P_{q}(q_{i}) \log_{2} P_{q}(q_{i})$$
 (2)

where P_s (s_i) is the probability of s_i in S, and $P_q(q_j)$ the probability of q_j in Q.

H(S,Q) is the joint entropy of (S,Q)

$$H(S,Q) = -\sum_{i,j} P_{sq}(s_i,q_j) \log_2 P_{sq}(s_i,q_j)$$
 (3)

where $P_{sq}(s_i, q_j)$ is the joint distribution probability of s_i and q_i .

The mutual information of S and Q is

$$I(S,Q) = H(S) + H(Q) - H(S,Q)$$
(4)

Define $[s,q] = [x(t),x(t+\tau)]$, s represents the time series x(t), q the time series $x(t+\tau)$ with time delay τ . Then $I(\mathbf{S},\mathbf{Q}) = I(x(t),x(t+\tau))$ can represent the certainty of $x(t+\tau)$ when x(t) is known. When $I(\mathbf{S},\mathbf{Q}) = 0$, $x(t+\tau)$ is completely unpredictable, which means x(t) and $x(t+\tau)$ are fully unrelated. When $I(\mathbf{S},\mathbf{Q})$ reaches the minimum, x(t) and $x(t+\tau)$ are un-

related in the biggest possibility. The τ , which makes $I(\mathbf{S},\mathbf{Q})$ get the first minimum, can be taken as optimal time delay^[3]. A key step in mutual information method is how to determine $P_{s_l}(s_i,q_i)$, and in Refs. [10-11] space cells method divided by equal distance are used to calculate distribution probability.

Assume that S and Q represent one-dimensional coordinate axis respectively, then (S,Q) forms a two-dimensional space. Each coordinate axis is divided into 2^n intervals by equal distance. Intervals can be indicated with $1,2,\dots,i,\dots,2^n$ or $1,2,\dots,j,\dots,2^n$. This two-dimensional space is divided into 2^{2n} regions and every region can be indicated with (1,1), (1,2), \dots , (2,1), \dots , (i,j), \dots , $(2^n,2^n)$. Calculate the number N(i,j) of two-dimensional variables within the (i,j) region which are composed of S and Q time series, so $P_{s_l}(s_i,q_j)$ can approximate to be

$$P_{sq}(s_i,q_i) = N(i,j)/N \tag{5}$$

$$P_{sq}(s_i) = \sum N(i,j)/N = N(i)/N$$
 (6)

$$P_{s_l}(q_j) = \sum_{i} N(i,j) / N = N(j) / N$$
 (7)

where $N = \sum_{i,j} N(i,j)$ is the total number of variables within the whole region. This is a simple explanation of distribution probability calculation method and detailed descriptions are illustrated in Refs. [10-11].

2.1.2 Embedding dimension

The geometrical significance of restructured phase space is to recover its motion track in high dimensional phase space according to the original time series. As the track of original time series is distorted when projected to a one-dimensional space, the previously non-adjacent points may turn adjacent after being projected. As a result, only the embedding of dimension reconstruction time series can reestablish the track of high-dimensional chaotic motion and the track of expansion, in order to judge the authenticity of the adjacent points, which can be taken as the method of embedding dimensions $m^{\lceil 12-13 \rceil}$.

In the *m*-dimensional phase space, phase points can be expressed as $X(t) = \{x(t), x(t+\tau), x(t+2\tau), \dots, x(t+(m-1)\tau)\}.$

There exists X_F which is the nearest neighbor point. Their distance is $R_m(t) = \| \mathbf{X}(t) - \mathbf{X}_F(t) \|$.

The distance between two points will change when the dimension of phase space increase from m to m+1. The new distance is described as follows

$$R_{m+1}^{2}(t) = R_{m}^{2}(t) + \| \mathbf{X}(t+m_{\tau}) - \mathbf{X}_{F}(t+m_{\tau}) \|^{2}$$
(8)

If R_{m+1} (t) changes largely compared to $R_m(t)$, $X_F(t)$ can be considered as a false neighbor point. Set

$$S_{m} = \frac{\|X(t+m_{\tau}) - X_{F}(t+m_{\tau})\|}{R_{m}(t)}$$
(9)

If $S_m > S_T$, $X_F(t)$ is a false nearest neighbor point of X(t). The range of threshold S_T is [10,50].

Calculate the proportion of false nearest neighbors repeatedly with the increase of m. When the proportion does not vary with the increase of m, optimal embedding dimension can be confirmed.

2, 2 Chaotic characteristic calculation

Reconstruct phase space based on time delay and embedding dimension. Correlation dimension and the largest Lyapunov exponent are used to analyze whether time series is chaotic or not.

2. 2. 1 Correlation dimension

The movement track of chaotic attractors in phase space forms a special curve which is of a self-similar structure. Fractal dimension can describe such a self-similar structure. Grassberger and Procaccia put forward correlation dimension method, i. e. G-P method. Saturated correlation dimension is approximate to fractal dimension and commonly used to represent fractal dimension, so it can well reflect geometrical features of chaotic attractors^[14-15]. G-P method is particularly applicable to observed data and easy for implementation, so it has been widely used in this field.

Assume x(1), x(2), x(3), ..., x(t), ..., t=1,2,...,N is the given time series. Reconstruct the phase space and get the new vector. That is $\mathbf{X}(t) = \{ x(t), x(t+\tau), x(t+2\tau), ..., x(t+(m-1)\tau) \}$, t=1,2,...,M, $M=N-(m-1)\tau$.

The distance of M vectors can be calculated by the means of Euclidean norm. Based on the given positive number r, correlation vectors are those with a less vector distance than r. The proportion of total pairs of correlation vectors in M^2 is correlation integral

$$C_{n}(r) = \frac{1}{M^{2}} \sum_{i=1}^{M} \sum_{j=1}^{M} \theta(r - \|\mathbf{X}(i) - \mathbf{X}(j)\|)$$

$$= \frac{1}{M^{2}} \sum_{i=1}^{M} \sum_{j=1}^{M} \theta(r - r_{ij})$$
(10)

where $\theta(\cdot)$ is Heaviside function

$$\theta(x) = \begin{cases} 0 & x \leqslant 0 \\ 1 & x > 0 \end{cases} \tag{11}$$

Correlation integral $C_n(r)$ is closely related to the positive number r. If r gets the larger value, all r_{ij} will be less than r which makes $C_n(r) = 1$. If r gets the less value, all r_{ij} will be larger than r which makes $C_n(r) = 0$. According to actual conditions, choose appropriate r to avoid extremes. When $r \rightarrow 0$, the relationship between $C_n(r)$ and r is shown as follows

$$\lim_{r \to 0} C_n(r) \propto r^D \tag{12}$$

where D is correlation dimension. D can be expressed as

$$D = \lim_{r \to 0} \ln C_n(r) / \ln r \tag{13}$$

To estimate the correlation dimension D of the actual time series, plot $\ln C_n(r) - \ln r$ curves for every embedding dimension based on the selected r values and their corresponding $C_n(r)$ values. The correlation exponent is the slope of the linear portions on the curves except for portions whose slopes are 0 or ∞ . Least square method is usually used to calculate the correlation exponent. Correlation exponents of random time series will never become saturated as embedding dimension increases. However, correlation exponents increase monotonically with increasing embedding dimensions. When correlation exponent become saturated, the saturated correlation exponent is correlation dimension $\mathbb{C}^{[3]}$.

2.2.2 The largest Lyapunov exponent

Chaotic systems have a sensibility for initial conditions, i. e. tracks which are initially adjacent will diverge with exponential rate in phase space^[2-3]. The Lyapunov exponent is used to

measure the exponential rate of convergence or divergence between adjacent tracks which are in different initial conditions. The existence of chaos in the system can be determined by whether the largest Lyapunov exponent λ_{max} is bigger than zero.

Wolf method, Jocobian method, small-data method and P-norm method^[16-18] are usually used to identify the largest Lyapunov exponent. Among them, small-data method exhibits a good robustness to the noisy data of time series and is easy to calculate, especially for small amount of data. Consequently, small-data method is used to calculate the largest Lyapunov exponent.

Transform the time series by fast Fourier transformation (FFT) and the average period is $P^{[19]}$. Reconstruct the phase space, and then the new vector is \mathbf{X} $(t) = \{x(t), x(t+\tau), x(t+2\tau), \cdots, x(t+(m-1)\tau)\}, t=1,2,\cdots,M,M=N-(m-1)\tau$. Find the nearest neighbors of each point in the phase space and limit short-term separation. $X(\hat{t})$ is the nearest neighbor point. The minimum distance is described as follows

$$d_{t}(0) = \min_{\hat{t}} \| \boldsymbol{X}(t) - \boldsymbol{X}(\hat{t}) \|$$

$$|t - \hat{t}| > P$$
(14)

where $\hat{t}=1,2,\cdots,M$ and $t\neq\hat{t}$. $d_t(0)$ is the initial distance between the tth point of phase space and the nearest neighbor. $\|\cdot\|$ is the Euclidean norm.

Calculate the distance between $m{X}(t)$ and $m{X}(\hat{t})$ after the ith discrete time step

$$d_{t}(i) = \| \mathbf{X}(t+i) - \mathbf{X}(\hat{t}+i) \|$$

$$i = 1, 2, \dots, \min(M-t, M-\hat{t})$$
(15)

For each i, calculate the average value of $\ln d_i(i)$, that is

$$x(i) = \frac{1}{q\Delta t} \sum_{j=1}^{q} \ln d_{j}(i)$$
 (16)

where q is the number of non-zero $d_t(i)$.

Select a linear portion of the curve to calculate the largest Lyapunov exponent [19-20] and make the least squares regression line of the linear portion of the curve, the slope of the line is the largest Lyapunov exponent λ_{max} . If $\lambda_{max} > 0$, time series are chaotic.

3 Result Analysis

Select flight radar data from the 16th sector of control center in Guangzhou area and confirm 7:30 to 24:00 as a valid period of one day. Matlab is used to process data and calculate chaotic characteristic values. The sampling interval is taken to be 1 min. A time series of 990 actual data about air traffic flow is obtained, as shown in Fig. 1.

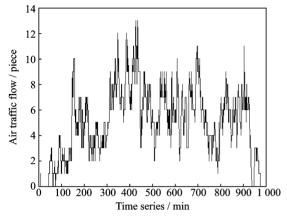


Fig. 1 Time series of air traffic flow

According to Figs. 2, 3, mutual information gets the first minimum when time delay τ is 11 and the rate of false nearest neighbors drops to 0 when embedding dimension is 8. Therefore, time delay and embedding dimension are confirmed.

Increase embedding dimension from m=2 to m=15 one by one. As embedding dimension increases, the linear portions of curves are dense and tend to be parallel to each other. The slope of linear portions tends to be a stable value which means correlation dimension reaches saturated, as shown in Fig. 4.

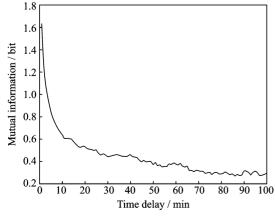


Fig. 2 Relationship curve between mutual information and time delay

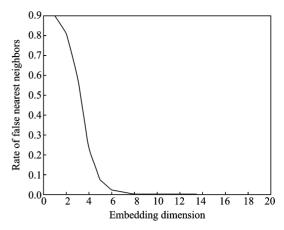


Fig. 3 Relationship curve between rate of false nearest neighbors and embedding dimension

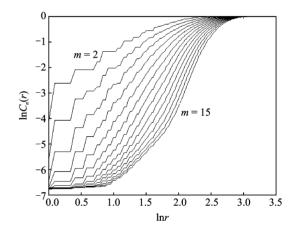


Fig. 4 Results of the G-P method

When embedding dimension increases, these correlation dimensions become more stable as shown in Table 1. The non-integer correlation dimension proves that the attractor of air traffic flow time series in phase space is of a self-similar fractal structure.

Table 1 Correlation dimensions for different embedding dimensions

Embedding dimension	9	10	11	12	13	14	15
Correlation	1 26	1 72	4 04	5 25	6.06	6 14	6 18
dimension	4. 20	4.73	4.94	J. 2J	0.00	0.14	0.10

Analyze air traffic flow time series with small-data method. The relationship curve of discrete time steps and average values is shown in Fig. 5. There exists no obvious linear portion in this curve, so the largest Lyapunov exponent is unable to be calculated through least square method. According to the theory in Ref. [20],

assume the largest Lyapunov exponent equals to 0 in this case. The system is not sensitive to the initial condition. In order to verify the result of the largest Lyapunov exponent, select Wolf method is seleted to recalculate the largest Lyapunov exponent. The result still equals to 0.

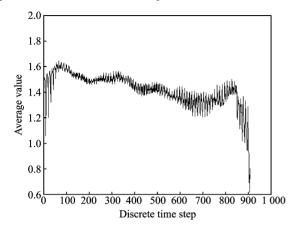


Fig. 5 Analytical curve for the largest Lyapunov exponent

Select radar data of November 14 to November 18. The air traffic flow sampling interval is taken to be 1 min. Results of five days are shown in Table 2. Based on the observation of five days' radar data, there are always no obvious linear portions in the analytical curves for the largest Lyapunov exponent. The largest Lyapunov exponent equals to 0 which means this kind of flow time series is not sensitive to initial conditions. However, air traffic flow time series always produce saturated correlation dimension as embedding dimension increases, which are of self-similar fractal structures. In order to demonstrate the universality of this phenomenon, select another ten sectors to test whether they are in accordance with the 16th sector results on the same day. Analytical results of ten sectors are shown in Table 3 and they turn out to be in accordance with the same regularity. Generally speaking, air traffic system is chaotic. Air traffic is affected by a number of factors. Operations of air traffic flow are orderly in general with the command of controllers, but strong random characteristics still exist in the system. Uncertain disturbances that air traffic system suffers from are caused by aircraft performance, pilots' abilities, severe weather, controllers and so on. If air traffic flow is completely random, it is difficult to analyze the operation regulation. If air traffic flow is clearly chaotic or periodic, the operation of traffic is regular. Only when nonlinear characteristics of air traffic are fully and correctly analyzed, can air traffic be predicted and managed accurately.

Table 2 Analytical results of five day data

	D .	Time Embedding		Correlation	The largest	
Date	Date	delay	dimension	dimension	Lyapunov exponent	
	14	11	8	6.14	0	
	15	14	9	6.35	0	
	16	12	11	5.67	0	
	17	8	14	7.98	0	
	18	11	9	6.26	0	

Table 3 Analytical results of ten sectors on the same day

<u> </u>	Time	Embedding	Correlation	The largest	
Sector	delay	dimension	dimension	Lyapunov exponent	
AR01	12	15	5.22	0	
AR02	9	11	1.85	0	
AR03	13	12	5.64	0	
AR04	13	10	6.26	0	
AR05	10	15	5.18	0	
AR06	14	10	4.98	0	
AR07	11	11	5.87	0	
AR08	14	8	5.98	0	
AR09	16	12	4.62	0	
AR10	15	15	6.38	0	

4 Conclusions

Air traffic system is complex and nonlinear. The operation of air traffic system is influenced by many uncertain factors. Based on the five-day radar data from the control center of Guangzhou area, air traffic flow time series are calculated. After reconstruction of phase space, correlation dimensions and the largest Lyapunov exponents are deduced. Although the largest Lyapunov exponents are all equal to 0 which means they are not sensitive to initial conditions, the saturated correlation dimensions show that the time series are of self-similar structures. As a result, the system is chaotic in terms of flow time series. For a long time, air traffic managers have thought flow distribution was disorder and irregu-

lar. However, research results of this paper provide a new idea of analyzing air traffic operation data. With the development of air transportation, air traffic distribution is becoming more and more intensive and the chaotic characteristics will be more obvious.

Analyzing chaotic characteristics of air traffic system is significant to improve and perfect non-linear characteristics analyses of traffic flow, predict the traffic accurately and manage air traffic effectively. Flow is just one of the metrics describing multiple dimensional air traffic. Future research is to select other useful and meaningful metrics and analyze whether these metrics' time series are chaotic. Moreover, chaos characteristics auto-extracting technique is also an immediate focus of further research.

References:

- [1] Wang D S, He G G. Summary and prospects of the study on traffic chaos [J]. China Civil Engineering Journal, 2003, 36(1): 68-74. (in Chinese)
- [2] Briggs K. An improved method for estimating Liapunov exponents of chaotic time series[J]. Physics Letters A, 1990, 151(1): 27-32.
- [3] Han M. Prediction theory and method of chaotic time series [M]. Beijing: China Water & Power Press, 2007:21-43. (in Chinese)
- [4] Li S M, Xu X H, Meng L H. Flight conflict forecasting based on chaotic time series[J]. Transactions of Nanjing University of Aeronautics & Astronautics, 2012, 29(4):388-394.
- [5] Safonov L A, Tomer E, Strygin V V, et al. Multifractal chaotic attractors in a system of delay-differential equations modeling road traffic[J]. Chaos: An Interdisciplinary Journal of Nonlinear Science, 2002, 12(4): 1006-1014.
- [6] Blank M. Dynamics of traffic jams: order and chaos [J]. Mosc Math J, 2001, 1(1): 1-26.
- [7] Lu Y, Chen Y H, He G G. The computing of maximum Lyapunov exponent in traffic flow applying the improved small-data method[J]. Systems Engineering-Theory & Practice, 2007, 28(1): 85-90. (in Chinese)

- [8] Dong C J, Shao C F, Ma Z L, et al. Temporal-spacial characteristic of urban expressway under jam flow condition [J]. Journal of Traffic and Transportation Engineering, 2012, 12(3): 73-79. (in Chinese)
- [9] Packard N H, Crutchfield J P, Farmer J D, et al. Geometry from a time series [J]. Physical Review Letters, 1980, 45(9): 712.
- [10] Fraser A M, Swinney H L. Independent coordinates for strange attractors from mutual information [J]. Physical Review A, 1986, 33(2): 1134-1140.
- [11] Rosenstein M T, Collins J J, de Luca C J. Reconstruction expansion as a geometry-based framework for choosing proper delay times[J]. Physica D: Nonlinear Phenomena, 1994, 73(1): 82-98.
- [12] Wang H Y, Sheng Z H. Choice of the parameters for the phase space reconstruction of chaotic time series [J]. Journal of Southeast University: Natural Science Edition, 2000, 30(5):113-117. (in Chinese)
- [13] Kennel M B, Brown R, Abarbanel H D I. Determining embedding dimension for phase space reconstruction using a geometrical reconstruction[J]. Phys Rev A, 1992, 45(6): 3403-3411.
- [14] Grassbergr P, Procaccia I. Measuring the strangeness of strange attractors[J]. Physica D, 1983, 9(1-2):189-208.
- [15] Wang J, Yang AB, Liu SY, et al. Improved fast algorithm to calculate correlation dimension of time series[J]. Journal of Huazhong University of Science and Technology: Natural Science Edition, 2012, 40 (10):101-104. (in Chinese)
- [16] Wolf A, Swift J B, Swinney H L, et al. Determining Lyapunov exponents from a time series[J]. Physica D: Nonlinear Phenomena, 1985, 16(3): 285-317.
- [17] Barana G, Tsuda I. A new method for computing Lyapunov exponents [J]. Phys Lett A, 1993, 175 (6): 421-427.
- [18] Rosenstein M T, Collins J J, Deluca C J. A practical method for calculating largest Lyapunov exponents from small data[J]. Physica D, 1993, 65(1/2): 117-134.
- [19] Yang Y F, Wu M J, Gao Z, et al. Parameters selection for calculating largest Lyapunov exponent from small data sets[J]. Journal of Vibration, Measurement & Diagnosis, 2012, 32(3):371-374. (in Chinese)
- [20] Chen G. Dynamic chaos features auto-extracting technique of non-linear time series[J]. Journal of Aerospace Power, 2007,22(1):1-7. (in Chinese)