International Journal of the Physical Sciences Vol. 6(10), pp. 2565-2571, 18 May, 2011
Available online at http://www.academicjournals.org/IJPS
DOI: 10.5897/IJPS11.320
ISSN 1992 - 1950 ©2011 Academic Journals

Full Length Research Paper

A new parameters joint optimization method of chaotic time series prediction

Changsheng Xiang[1], Wei Zhou[2], Zheming Yuan[3]*, Yuan Chen[2] and Xingyao Xiong[3]

[1]Orient Science and Technology College, Hunan Agricultural University, Changsha, 410128, China.
[2]College of Bio-safety Science and Technology, Hunan Agricultural University, Changsha, 410128, China.
[3]Hunan Provincial Key Laboratory for Crop Germplasm Innovation and Utilization, Changsha, 410128, China.

Accepted 18 April, 2011

To improve the prediction performance of chaotic time series, a new method is proposed for parameters joint optimization of phase space reconstruction and support vector machine (SVM). The main idea of the joint optimization method is that the parameters from phase space reconstruction and SVM are designed jointly using uniform design firstly, and then the parameters are optimized jointly based on self-calling SVM. The results tested by chaotic time series indicate that the proposed method has more advantages than traditional methods, such as better prediction accuracy and lower computational complexity.

Key words: Uniform design, support vector machine, parameters joint optimization, chaotic time series.

INTRODUCTION

The chaotic time series prediction has been widely applied in many fields, such as economics, signal processing, communication, biology and control, etc. It is a key problem to determine the optimal delay time (τ) and embedding dimension (m) in phase space reconstruction (Chen et al., 2004; Liu et al.,2005) . τ and m were selected independently in previous studies and some relative methods were developed. The methods determining τ mainly include autocorrelation function method (Albano et al., 2002) and average mutual information method (Fraser, 1989); while the methods determining m mainly include trial method and false neighbor point method (Maguire et al., 1998). Although Takens theory has proved τ and m of an infinite and noiseless time series are mutually independent, all time series are finite and influenced inevitably by various noises in real world, so now many researchers address that τ and m are interrelated. Correspondingly, C-C (Kim, 1999) and τ-m automatic method (Zhang et al., 2010) are emerged, However, C-C method calculation results are under-stable and τ-m automatic method evaluation criteria are subjective (Ma et al., 2004). Developing a new and effective joint optimization method of τ-m is a hotspot in the current chaotic time series analysis.

After determining τ and m, the support vector machine (SVM) was adopted to model in the most current researches because of nonlinear characteristics of chaotic time series. SVM is a machine learning method based on the structural risk minimization principle. It eliminates the problems of small sample, nonlinearity, over fitting, dimensionality curse, local minimum and generalization ability is very good. So SVM has been widely used in nonlinear time series prediction (Wu et al., 2007). SVM prediction precision is related to kernel function and its parameters. In most cases, radial basis kernel function (RBF) has a good prediction performance. The representative software of SVM—Libsvm adopts grid search to obtain SVM parameters, which is very time-consuming and not suitable for large samples. If default parameters not optimized are adopted, prediction precision will decrease sharply (Huang et al., 2007). The least squares support vector machine (LSSVM) is reformulations to standard SVM (Pan et al., 2009). LSSVM is suitable for solving large-scale problems and its prediction precision is slightly less than Libsvm result. When the RBF is selected for LSSVM, LSSVM needs to optimize two parameters: regularization parameter γ and kernel width σ. In previous studies, phase space reconstruction involving τ and m and LSSVM optimization

*Corresponding author: E-mail: zhmyuan@sina.com.

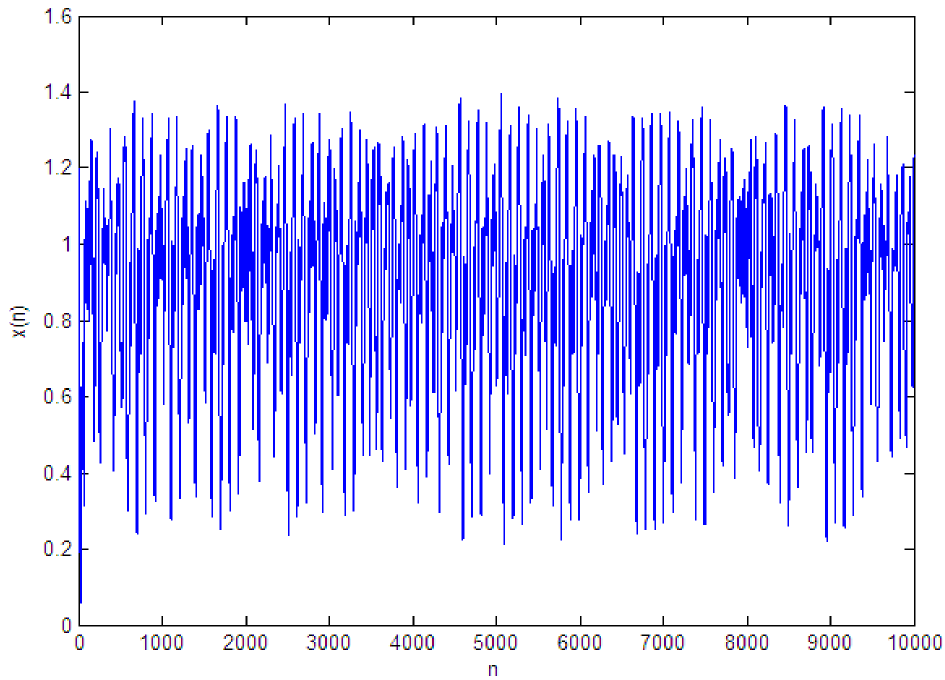

Figure 1. The Mackey-Glass time series when Δ=30.

involving γ and σ are determined independently, so the determined τ and m are not always ensure LSSVM has the optimal prediction precision. Therefore, parameters joint optimization of τ, m, γ and σ purely from data driven not based on any priori knowledge, is a very attractive choice. Multilevel exhaustive search optimization is very time-consuming, so uniform design (UD) will be introduced because it can effectively reduce the number of experiments (Yuan et al., 2009).

Aimed at the joint optimization of τ, m, γ and σ of chaotic time series prediction, a new optimization method is proposed based on UD and SVM (UD-SVM) in this paper. Simulation experiments are carried out based on two chaotic time series data to test the validity of UD-SVM.

MACKEY-GLASS TIME SERIES PREDICTION

Mackey-Glass time series

Mackey-glass time series is often used as a standard to test the prediction performance of nonlinear system models. The definition of Mackey-glass time series is as follows:

$$\frac{dx(n)}{dn}=\frac{0.2x(n-\Delta)}{1+x^{20}(n-\Delta)}-0.1x(n) \quad (1)$$

Where Δ is parameter. When Δ≥17, chaos is presented. The higher the Δ, the higher the chaotic degree. In this paper, Δ=30, the fourth-order five-step Runge-Kutta method to carry out numerical integration and get chaotic time series data (Figure 1).

In order to compare the results of literature (Cui et al., 2004), γ is fixed to 50, the train set is the first 3000 data points, the test set is the data points from 3001 to 6000 and the valid set is the data points from 6001 to 9000 of the time series. Data preprocessing is as follows:

$$X(k,i)=\frac{x(k,i)-mean\_x(i)}{std\_x(i)}$$

$$y(k)=\frac{y(k)-mean\_y}{std\_y} \quad (2)$$

where $mean\_x(i)$ and $std\_x(i)$ are arithmetic mean value and standard deviation of the ith column of input vector $x(i)$ respectively; $mean\_y$ and std_y are arithmetic mean value and standard deviation of output vector $y(k)$ respectively.

UD and LSSVM prediction

The predetermination upper and lower limits of 3 factors (τ, m and σ) are listed in Table 1 according to the relevant studies of Mackey-Glass time series. 24 treatments (parameters combination) are produced by 3 factors and 8 levels UD based on DPS 6.0. For each treatment, the train set is trained while the valid set is predicted by

Table 1. The predetermination upper and lower limits of the parameters for Mackey-Glass time series.

Factors	τ	m	σ
Upper limit	8	8	4
Lower limit	1	1	0.5
Step	1	1	0.5

Table 2. 24 treatments designed by UD and RMSEs calculated by LSSVM for Mackey-Glass time series.

No.	RMSE	τ	M	σ
N_1	0.003189	1	4	4
N_2	0.003275	1	8	1
N_3	0.005110	1	2	2
N_4	0.008495	5	8	3
N_5	0.011255	7	6	4
N_6	0.010084	8	8	2.5
N_7	0.010604	6	7	2
N_8	0.008328	2	5	1.5
N_9	0.009395	2	7	3.5
N_{10}	0.011790	6	6	1.5
N_{11}	0.020350	3	6	2.5
N_{12}	0.026522	8	4	3
N_{13}	0.024580	7	5	0.5
N_{14}	0.029389	3	3	0.5
N_{15}	0.029241	3	3	3
N_{16}	0.040853	4	5	3.5
N_{17}	0.036683	4	7	0.5
N_{18}	0.065221	5	4	1
N_{19}	0.057433	2	1	2.5
N_{20}	0.063453	4	2	1.5
N_{21}	0.111680	6	2	3.5
N_{22}	0.126060	7	3	2
N_{23}	0.133870	5	1	4
N_{24}	0.191660	8	1	1

LSSVM, and the predicted root mean square errors (RMSEs) are listed in Table 2.

Parameters optimization based on Libsvm

Taking RMSE values as dependent variable and τ, m and σ as independent variables from 24 treatments to form a new train set (Table 2), while taking τ, m and σ of all combinations to form a new test set. The new test set has 512 (8×8×8) samples and the dependent variable (RSMEs) values are unknown. Because traditional empirical risk minimization models, such as multiple linear regressions, response surface method and artificial neural network have many defects, especially poor generalization ability under the condition of small samples. The prediction precision of LSSVM is a little less than that of SVM, so SVM is used as the regression model here. Libsvm2.8 is a simple and easy SVM software package, which contains four programs, namely svmscale used for normalizing original data, svmtrain used for training, svmpredict used for predicting and gridregression.py used for automatic searching the optimal c, g, p. The usage and parameters setting of various programs refer to literature (Lin et al., 2007).

Based on Libsvm, the new train set is trained and the new test set is predicted. The optimal parameters combination (τ=1, m=8, σ=2) is confirmed according to the smallest RMSE value. Based on the optimal parameters combination, the train set (the first 3000 data of Mackey-Glass) is trained and the valid set (from 6001 to 9000 data of Mackey-Glass) is predicted using LSSVM, and the RMSE value is equal to 0.002326. The results show the prediction precision is very high, so the optimal parameters are reasonable based on the proposed joint optimization method.

Model verification

Based on the optimal parameters combination (τ=1, m=8, σ=2), the train set (the first 3000 data of Mackey-Glass) is trained and the test set (from 3001 to 6000 data of Mackey-Glass) is predicted using LSSVM. The results are shown as Figure 2 and the RMSE value is

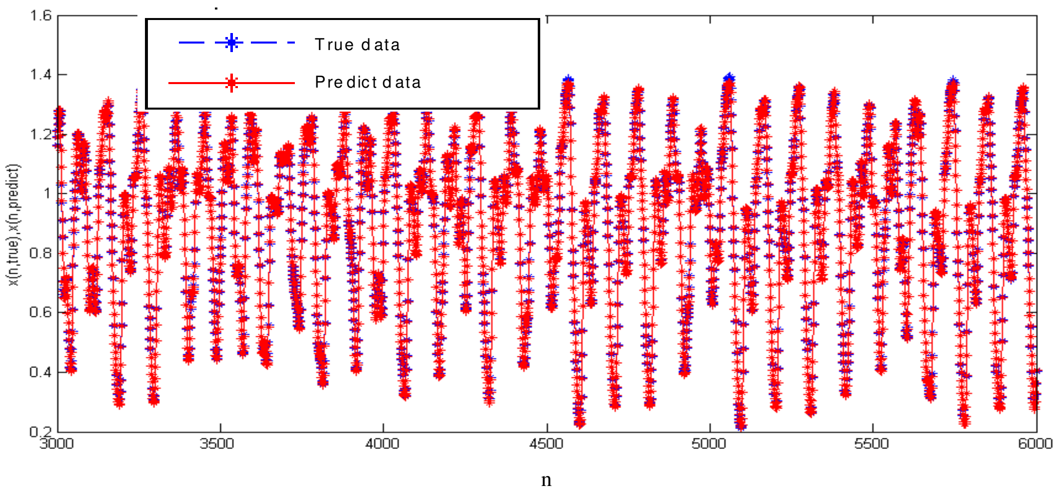

Figure 2. Comparison of the true values and corresponding predicted values using UD-SVM of Mackey-Glass time series.

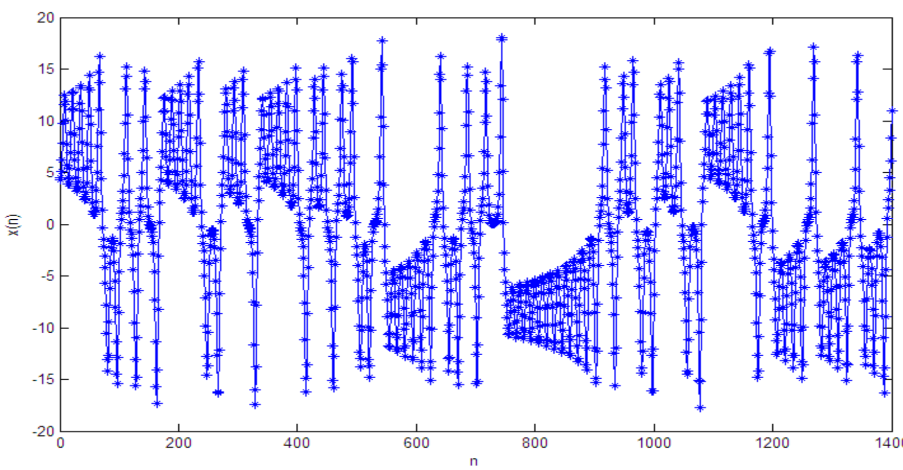

Figure 3. The Lorenz time series.

0.0028. The optimal results reported by Cui (2004) are as follows: τ=1, m=3, σ=2, RMSE=0.0034. The results show UD-SVM prediction precision is superior to that reported by Cui (2003). So UD-SVM is a fast and efficient τ-m-SVM parameters joint optimization method.

Lorenz chaotic time series prediction

Lorenz system is a deterministic nonlinear dissipative system and has disordered and non-periodic phenomena. Its studies run through the whole chaos science development and therefore it is an important chaotic time series.

$$\begin{aligned} \dot{x} &= 10(y-x) \\ \dot{y} &= 28x - y - xz \\ \dot{z} &= -\frac{8}{3}z + xy \end{aligned} \tag{3}$$

Under initial conditions (x_0=5 , y_0=5 , z_0=15), using the fourth-order five-step Runge-Kutta algorithm to get its numerical solution and sampling interval time is 0.05, the first 1400 data points of chaotic time series about x component are produced (Figure 3), and the time series have complex nonlinear chaotic characteristics obviously.

In order to compare with literature (Cui et al., 2005), the first 600 data points are selected as the train set, 400 data points from 601 to 1000 as the test set and γ is fixed to 10. In addition, 400 data points from 1001 and 1400 are selected as the valid set. The predetermination upper and lower limits of 3 factors are list in Table 3. 24 treatments are produced by 3 factors and 8 levels UD based on DPS 6.0. For each treatment, the train set is trained the valid set is predicted using LSSVM and the predicted RMSEs are show in Table 4.

The data in Table 4 are taken as a new train set, the 3 parameters and 8 levels complete combinations which are 512 samples as a new test set. The training and predicting are carried out based on Libsvm and get the optimal parameters combination (τ=1, m=5, σ=2). The train set of Lorenz time series are trained and the

Table 3. The predetermination upper and lower limits of the parameters for Lorenz time series.

Factors	τ	M	σ
Median	4	4	4
Upper limit	8	8	8
Lower limit	1	1	1
Step	1	1	1

Table 4. 24 treatments designed by UD and RMSEs calculated by LSSVM for Lorenz time series.

No.	τ	m	σ	RMSE
N_1	1	4	8	0.34108
N_2	1	8	2	0.34956
N_3	1	2	4	0.2906
N_4	5	8	6	4.9375
N_5	7	6	8	6.6285
N_6	8	8	5	7.7915
N_7	6	7	4	6.2560
N_8	2	5	3	0.4664
N_9	2	7	7	0.4702
N_{10}	6	6	3	4.8262
N_{11}	3	6	5	1.1097
N_{12}	8	4	6	5.5303
N_{13}	7	5	1	4.6805
N_{14}	3	3	1	0.5873
N_{15}	3	3	6	0.9594
N_{16}	4	5	7	2.4362
N_{17}	4	7	1	2.3891
N_{18}	5	4	2	2.5264
N_{19}	2	1	5	3.4680
N_{20}	4	2	3	1.7052
N_{21}	6	2	7	5.3989
N_{22}	7	3	4	4.8446
N_{23}	5	1	8	5.2356
N_{24}	8	1	2	6.1827

the valid set of Lorenz time series are predicted based on LSSVM with the optimal parameters combination, and the RMSE is 0.25188. The prediction precision has been improved significantly compared with the prediction precision before optimization (Table 4).

So, the optimal parameters combination (τ=1 , m=5 , σ=2) are used to train the train set and predict the test set of Lorenz data, The results are shown in Figure 4. The RMSE is 0.25884 and the correlation coefficient between predicted values and true values is 0.9994. The optimal parameters reported by Cui (2005) are τ=1 , m=6 , σ=5, the correlation coefficient is 0.9991 while the RMSE reported has not given. It can be seen that the UD-SVM computational complexity is lower, its predicted values fit its true values very well and its prediction performance is superior to the method which is proposed by Cui (2005). Thus the results prove that UD-SVM is a fast and efficient method for τ-m-SVM parameters joint optimization.

DISCUSSIONS

Because the time series have noise inevitably and its lengths are finite, the existing methods for determining τ and m are not suitable for the practical application. Even though the τ and m can be got by different methods, they are contradictory sometimes. For example, for the monthly mean sunspots, Wu et al. (2004) think its m should be less than 8 (Wu et al. 2004; Gu et al. 1999) , but the opinion of Ma et al. (2009) is opposite Because phase space reconstruction and prediction model are interdependent relationship, they should be optimized jointly, but they are optimized independently in the current researches. Therefore, the τ-m-SVM parameters joint optimization based on data driven proposed in this

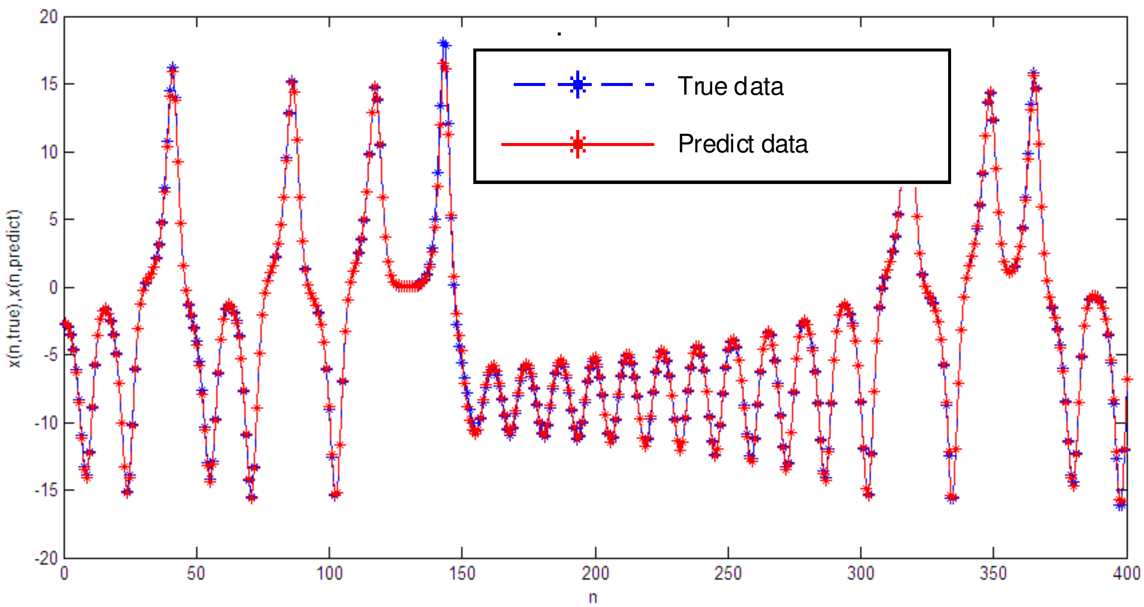

Figure 4. Comparison of the true values and corresponding predicted values using UD-SVM of Lorenz time series.

paper provides a new idea for solving the problem.

Multiparameter and multilevel joint optimization is very time consuming for large train samples, so UD and self-calling SVM are used to reduce the number of experiments and convert large sample search into small sample search in this paper, thereby it greatly reduce computational complexity under the condition of ensuring prediction precision. Empirical risk minimization regression models, such as multiple linear regression and partial least square regression, etc, assume that the sample is sufficiently or infinitely large, but the parameters joint optimization after converting is a typical small sample and there are often complex nonlinear relations among various parameters and dependent variable. Therefore these models generalization ability has a serious shortage, and the prediction precision of the optimal parameters combination deduced by empirical risk minimization regression models is often unsatisfactory. When optimizing parameters by response surface method, the selection of reconstruction function is a key problem and the adopted usually quadratic polynomial has approximation limited the nonlinear ability for nonlinear functions (Gupta and Manohar, 2004). Artificial neural network has very good nonlinear prediction ability, but it has many defects such as undetermined model structure, over fitting and local minimum, etc (Zhang and Hu, 2008). This paper solves the problem by nonlinear SVM which is based on minimum structural risk principle and has strong generalization ability and is suitable for small samples. If the optimal parameters combination exceeds the predetermination upper or lower limit of parameters, a new UD needs to be carried out.

Conclusions

SVM parameters and phase space reconstruction are optimized jointly because they have interdependent relationships. A parameters joint optimization method is proposed based on UD-SVM in this paper. The simulation results show that UD-SVM is a fast and efficient τ-m-SVM parameters joint optimization method based on data driven for chaotic time series. But for noise and time-varying parameters hyper chaotic time series, both regularization parameter γ and the train set window length need to be optimized too and it will be explored in the further research.

ACKNOWLEDGMENTS

I would like to express my special thanks to referees for their precious advice to my paper. This work was supported by Hunan Provincial Natural Science Foundation of China (Grant No. 10JJ1005), Innovative Research Team in University (Grant No. IRT0963), the Ph.D. Programs Foundation of Ministry Education (Grant No.200805370002), the scientific research projects of Hunan Province (Grant NO. 10C0803).

REFERENCES

Albano AM, Muench J, Schwartz C (2002). Correlation-dimension and autocorrelation fluctuations in epileptic seizure dynamics. Phys. Rev. Lett., 65(3): 1921-1926.

Chen GH, Li Y, Sheng ZH (2004). The chaotic behavior and control of a duopoly model based on heterogeneous expectations. Syst. Eng. Theor. Pract., 24(5): 84-90.

Cui WZ, Zhu CC, Bao WX, Liu JH (2004). Prediction of the chaotic time series using support vector machines. Acta Phys. Sin., 53(10): 3303-3311.

Cui WZ, Zhu CC, Bao WX, Liu JH (2005). Prediction of the chaotic time series using support vector machines for fuzzy rule-based modeling. Acta Phys. Sin., 54(7): 3009-3018.

Fraser AM (1989). Information, entropy in strange attractors. IEEE Trans. Info. Theor., 35(2): 245-262.

Gupta S, Manohar CS (2004). An improved response surface method for the determination of failure probability and importance measures. Struct. Saf., 26(2): 123-139.

Gu SS, Wang ZQ, Cheng JT (1999). The fractal research and predicating on the time series of sunspot relative number. Appl. Math. Mech., 20(1): 79-85.

Huang CM, Lee YJ, Lin DK, Huang SY (2007). Model selection for support vector machines via uniform design. Comput. Stat. Data Anal., 52(1): 335-346.

Kim H, Eykholt SR, Salas JD (1999). Nonlinear dynamics, delay times and embedding windows. Phys. D., 27: 48-60.

Lin CJ (2007). Projected gradient methods for non-negative matrix factorization. Neural Comput., 19: 2756-2779.

Liu HD, Li Q (2005). Chaotic time series nonlinear prediction based on support vector machines. Syst. Eng. Theor. Pract., 25(9): 94-99.

Maguire LP, Roche B, Mcginnity TM (1998). Predicting a chaotic time series using a fuzzy neural network. Info. Sci., 112: 125-136.

Ma HG, Li XH, Wang GH (2004). Selection of embedding dimension and delay time in phase space reconstruction. J. Xian Jiao Tong Univ., 38(4): 335-338.

Ma QL, Zheng QL, Peng H, Qin JW (2009). Chaotic time series prediction based on fuzzy boundary modular neural networks. Acta. Phys. Sin., 58(3): 1410-1419.

Pan JS, Hong, Zhou QF (2009). Integrated application of uniform design and least-squares support vector machines to transfection optimization. BMC Biotechnol.. 9(1): 52-57.

Wu Q, Liu WY, Yang YH (2007). Time series online prediction algorithm based on least squares support vector machine. J. Central South Uni. Technol., 14(3): 442-446.

Wu YL, Dong FA, Li SF (2004). Predicting on the time series of sunspot number based on chaos theory. J. Shaanxi Normal Uni., 32(1): 81-88.

Yuan ZM, Zuo B, Tan SQ (2009). Experimental design and analysis of the optimal fermentation medium based on uniform design and support vector regression. Chinese J. Process Eng., 9(1): 148-152.

Zhang SQ, Jia J, Gao M (2010). Study on the spatial parameters selection of chaotic time series reconstruction. Acta. Phys. Sinica., 59(3): 1576-1582.

Zhang JF, Hu SS (2008). Chaotic time series prediction based on multi-kernel learning support vector regression. Acta. Phys. Sin., 57(5): 2708-273.