ELSEVIER

2012 AASRI Conference on Computational Intelligence and Bioinformatics

Chaotic Time Series Prediction Based On Binary Particle Swarm Optimization

Xiaoxiao Cui, Mingyan Jiang*

School of Information Science and Engineering, Shandong University, Jinan 250100, P.R. China

Abstract

Prediction of chaotic time series based on the phase space reconstruction theory has been applied in many research fields. Local linear model is widely used in chaos prediction due to its versatility and small computation amount. The embedding dimension and time delay parameters of the local linear prediction model can take different values with those of the phase space reconstruction. The Binary Particle Swarm Optimization (BPSO) is applied to choose the optimal parameters of the new local linear prediction model for its strong search ability. The main objective of this approach is to increase the predictive accuracy of the local linear model. In this paper the local linear one-step and multi-step predictive model predicts the chaotic time series respectively. Simulation results show the feasibility and effectiveness of the proposed method.

1. Introduction

Chaotic phenomenon is an irregular motion produced by determinate nonlinear dynamical system. It is widely encountered in various fields ranging from physics and mathematics to biology and others. Though

---

* Corresponding author. Tel.:13361073798.
E-mail address: jiangmingyan@sdu.edu.cn.

chaos seemed complicated and stochastic, a certain rule indicating that it can be predicted in the short term exists internally-[1]. Modeling and prediction of chaotic time series has already been an important branch of nonlinear information procession, such as weather forecasting, economic analysis and biomedical signal procession, etc.-[2]. With the development of chaos theory and research on its application, many methods have been proposed to predict chaotic time series, mainly including global and local prediction methods. The global methods use the whole time series data to fit the true function of the chaos attractor-[3]. It is obvious that if new information is taken into account, all parameters of the model have to be changed. Besides, the equation of global dynamic system is rather difficult to fit because of the complex attractor's construction and computation-[4]. There are a few prediction tools about the global model, such as artificial neural network and Least Square Support Vector Machine (LS-SVM). Data normalization is necessary for artificial neural network, which makes calculation more complicated-[5]. Parameters are extremely crucial to LS-SVM learning results and generalization ability, but in most cases they are empirical values-[6]. The local methods can overcome the disadvantage above because it builds model only on the current state and utilizes part of the past information. On the basis of phase space reconstruction, parameters of the local linear method are valued the same with those of reconstruction traditionally-[7]. But parameters which are optimal for the state space reconstruction may not be suitable for the prediction. The improvement is that one parameter is optimized by minimizing the standard prediction errors, when other parameter is empirical value-[8]. The obtained parameter value, in theory, is optimal on the condition that the other parameter is fixed. But both of the parameters value may not be optimal for the prediction model. In this paper BPSO is introduced to search the optimal parameters of the new local linear prediction model by minimizing the standard prediction errors. For the local linear model there are one-step and multi-step prediction.

This paper is organized as follows: The local linear prediction model, BPSO and adaptive weighted fusion algorithm are described in Section 2. The new local linear prediction model is illustrated in Section 3. Simulation and results are shown in Section 4 and discussion is described in Section 5. The conclusion is given in Section 6.

2. Background theory

2.1. The local linear prediction model

Phase space reconstruction from a scale time series is the foundation of chaotic time series analysis. Takens has proved that state space can be reconstructed based on Whitney's topological embedding theorem–[9-10]. For observed scalar time series $\{x(n), n = 1, 2, \cdots, N\}$, the state space can be reconstructed as $X(n) = [x(n), x(n-\tau), \cdots, x(n-(m-1)\tau)]^T$, $n = 1 + (m-1)\tau, 2 + (m-1)\tau, \cdots, N$. according to Takens's delay embedding theorem. Where m is the embedding dimension and τ is the time delay. With the optimal parameters the state evolution of chaotic systems can be reflected by the state space.

Prediction of chaotic time series is to forecast the dynamic trajectories of reconstructed state space. That is to find the mapping function $F : \hat{x}(N+1) = F(X_N)$, where X_N is the current state vector and $\hat{x}(N+1)$ is the next predictive data. The local linear prediction expression is:

$$\hat{x}(N+n) = a_0 + \sum_{i=1}^{m} a_i x(N-(i-1)\tau) = AY(N). \tag{1}$$

Where $A = [a_0, a_1, a_2, \cdots, a_m]$, $Y(N) = [1, x(N), x(N-\tau), \cdots, x(N-(m-1)\tau)]^T$. And n is the predictive step. Clearly, the embedding dimension of the reconstructed space becomes the order of the prediction model. In

order to compute A, we need to find out the nearest neighboring points of the current state $X(N)$ in the reconstructed state space. These nearest neighboring points are obtained by the formula:

$$dist(i) = \|X_N - X_i\|, \ i = 1, \ 2, \ \cdots, N-1. \tag{2}$$

Form formula (2) we can see that the *dist* can reflect the similarity of state evolution of the neighbouring and current state vectors. A can be obtained by computing the equation $AB = D$ and k is the number of the neighbouring points. B is an $(m+1)*k$ matrix whose ith column is made up of Y and $D = [x(N_1 + step), x(N_2 + step), \cdots, x(N_k + step)]$. Number of the neighbouring points is usually no smaller than $m+1$ to make sure that the matrix is non-singular. For local linear one-step prediction the predictive step is one each time, and then the obtained prediction data is added into the training set to predict the next data with the same procedure. The local multi-step prediction is achieved by changing prediction step, and the obtained prediction value is also added into the training set.

The optimal value of parameters m and τ in phase space reconstruction may not be optimal for prediction in the local linear prediction model. They should be more suitable for the prediction.

2.2. BPSO

Kenney and Eberhart presented the Particle Swarm Optimization (PSO) algorithm as an optimization technique to solve problems of real-number spaces in 1995-[11]. It is described as:

$$v_{ij}(t+1) = wv_{ij}(t) + c_1 r_{1j}(t)[p_{best}(t) - x_{ij}(t)] + c_2 r_2(t)[g_{best}(t) - x_{ij}(t)],$$
$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1). \tag{3}$$

Where $v_{ij}(t)$ is the velocity of particle i and t is the number of iterations. $x_{ij}(t)$ represents the particle position. j is dimension of velocity or position of particles. c_1 and c_2 are the positive constant parameters. r_1 and r_2 are random number between 0 and 1. p_{best} is the best position of the particle itself and g_{best} is the best position among all the particle swarm. And w is the inertia weight to influence the convergence behavior of PSO. In most cases

$$w = w_{max} - (w_{max} - w_{min})k / k_{max}. \tag{4}$$

It is used to speed up the convergence. But typical problems are set in a variable space with discrete and qualitative features, such as scheduling and routing problems. Then Kenney and Eberhart proposed BPSO to settle zero-one integer programming problems in 1997-[12]. In BPSO algorithm $v_{ij}(t)$ is no longer the velocity of particles but the possibility of $x_{ij}(t)$ equaling to 1. Sigmoid function is used to confine the possibilities to $[0 \ 1]$. Values of v are assigned $[V_{min} \ V_{max}]$ in case of Equation (5) reaching saturation-[13].

$$sigmoid(v) = \frac{1}{1 + e^{-v}} \tag{5}$$

3. Modeling method based on BPSO

Parameters of the traditional local linear are valued the same with that of phase space reconstruction during the traditional predictive methods. Ref [8] points out that parameters of the model can be assigned different to improve the predictive result. For these two parameters, the authors estimate the optimal value of one parameter with the other parameter fixed by minimizing the standard prediction error in Ref [8]. This method improves the predictive error indeed. But the problem is that there always a corresponding value for one parameter with the other parameter settled, so the obtained value could not be optimal in theory. In this paper the BPSO is used to evaluate the both parameters simultaneously.

The local linear prediction model uses the BPSO algorithm to choose the parameters m and τ for improving the precision of prediction. Then m and τ become the particle swarms, so the dimension of the particle swarm is two. The process of obtaining the optimal parameters is illustrated as follows:

Step 1: Setting parameters

Initialize the swarm parameters. In this paper swarm size is 30 and number of iteration is 30, length of the binary number is 4. Other parameters are set as: $w_{max} = 1.2, w_{min} = 0.4, c_1 = c_2 = 2$.

Step 2: Learning and Computing

*1).*Initialize particles with random positions and velocities;

*2).*Input the training set and search for nearest neighboring points. Local linear model begins to predict with the training set;

*3).*Compute the fitness values of each particle. To evaluate the performance of the prediction model, the standard prediction errors is chosen as the fitness function.

$$e(m,\tau) = \sqrt{\frac{\sum_{i=1}^{l}[x(n+i) - \hat{x}(n+i)]^2}{\sum_{i=1}^{l}[x(n+i) - \bar{x}]^2}}, i = 1, 2, \cdots, l. \qquad (6)$$

Where l is the length of the prediction data and

$$\bar{x} = \frac{1}{l}\sum_{i=1}^{l}x(n+i) \qquad (7)$$

$x(n+i)$ is the real data and $\hat{x}(n+i)$ is the output data of the model. Value of parameters could be acquired by minimizing formula (6).

*4).*Iterate and update the particle swarm. Compare the fitness value of each iteration particle swarm and update the g_{best} and p_{best}.

*5).*Obtain the optimal parameters when iteration ends.

Step 3: Predicting

Local linear model predicts with the parameters obtained from Step 2.

4. Simulation and results

In this paper the classical Lorenz time series is introduced to verify the feasibility and practicability of the proposed method.

Based on Galerkin's approximation to the partial differential equations of thermal convection in the lower atmosphere derived by Salzman, Lorenz proposed three ordinary differential equations [14]. It is given by the following three Ordinary Differential Equations (ODEs):

$$\begin{cases} \dot{x} = \alpha(y - x) \\ \dot{y} = (\beta - z)x - y. \\ \dot{z} = xy - \gamma z \end{cases} \tag{8}$$

Where x is the amplitude of the convection motion, y is the temperature difference between ascending and descending currents and z is distortion of the vertical temperature profile from linearity. α, β and γ are dimensionless parameters. In this paper these parameters are set to be $\alpha = 10$, $\beta = 28$ and $\gamma = 8/3$ for a rich dynamical behavior. The standard fourth-order Runge–Kutta method is used to solve the equations. The time step is $h = 0.01$ and the x value are used to simulate and test.

In this paper, 2000 data is chosen to build the local linear one-step model of Lorenz time series and the following 50 data is to be predicted. For comparison, the fitness value of proposed one-step predictive model and traditional model are shown in Table 1.

Table 1. Comparison results of Lorenz time series

Method	Fitness value
Traditional method	0.00491
Proposed method	0.00215

In order to compare the method of Ref [8] and the new model, 3000 data is chosen to build the model of Lorenz time series. The following 600 data is to be predicted. With the method of Ref [8] we set the parameter $m = 3$ and search for the optimal value of τ that minimizes $e(m, \tau)$. Then we get that the optimal value of τ is 1. Comparison results are shown separately in Fig.1 and Fig.2. Simulation results show that the optimal method is better than the method proposed in Ref [8].
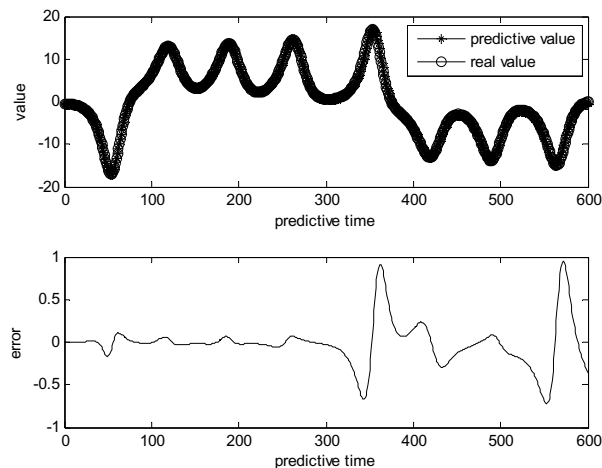

Fig.1. the local linear multi-step prediction output of Lorenz time series with $m = 2, \tau = 2$
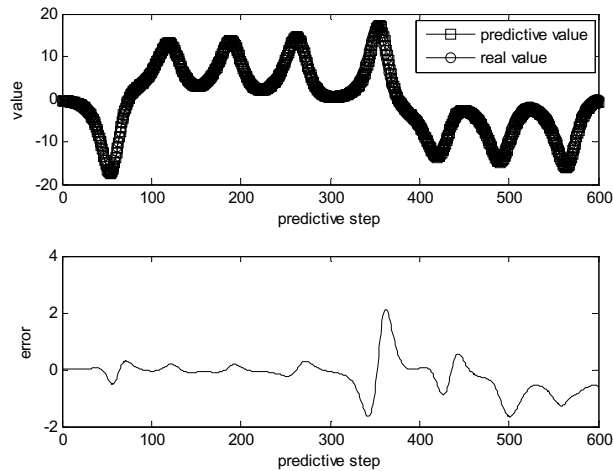
Fig.2. the local linear multi-step prediction output of Lorenz time series with $m = 3, \tau = 1$

5. Discussion

All the figures above demonstrate the short-term predictive of chaotic time series. The predictive validity depends on not only the dynamic characteristics of time series, but the predictive step as well. From error diagrams we can see that the error of prediction is increasing basically with the prediction step in the local linear prediction. The reason is that each prediction value obtained by the model is used to predict the next data. Therefore the shorter the predictive step is, the more accurate the predictive result will be. Furthermore, the local linear multi-step prediction predicts chaotic time series with the training data by changing the step. The neighboring points do not need to be recalculated during the predictive procedure, which reduces the complexity of computation and is more suitable for the long-term chaos prediction.

To get better predictive value, we can also use some other methods such as the adapted weighted fusion algorithm to fusion the obtained data. The adapted weighted fusion algorithm allocates the weighted factors automatically according to the adaptive criteria-[15]. It merges the output obtained by local one-step and multi-step prediction model with weighted factors and improves predictive accuracy and the algorithm can minimize the predictive error theoretically.

6. Conclusion

A new method of predicting chaotic time series based on BPSO is proposed in this paper. Parameters of prediction model can be set to different value with those of phase space reconstruction. Because the parameters have important impact on the prediction accuracy, optimal values are chosen for the prediction model. BPSO is applied to search the optimal parameters due to its strong global search capability. The local linear one-step prediction updates the training set to search for new neighboring points and increases the complexity of computation. The multi-step prediction can overcome the disadvantage by changing the predictive step. The adaptive weighted fused algorithm also can be employed to improve the predictive accuracy by fusing the models with weighted factors. Simulation results demonstrate that the proposed method is superior to the traditional local linear model and more suitable for high precision short-term prediction.

Acknowledgements

This work is supported by the Natural Science Foundation of Shandong Province (No. ZR2010FM040).

References

[1] H. Kantz and T. Schreiber. Nonlinear Time Series Analysis, 2nd ed. Cambridge University Press 2003.
[2] M. Casdagli. Nonlinear prediction of chaotic time-series. Physica D 1989:35: 335-356.
[3] A.J. Lichtenberg, M.A. Lieberman. Regular and Stochastic Motion. Springer 1983.
[4] Ansgar Freking, Wolfgang Kinzel. Learning and predicting time series by neural networks. Phys. Rev. E, 2002:65, 050903.
[5] D. S. K. Karunasinghe, S. Y. Liong. Chaotic time series prediction with a global model: Artificial neural network. Journal of Hydrology, 2006: 323 (1): 92-105.
[6] Ping Liu, Jian Yao. Application of Least Square Support Vector Machine based on Particle Swarm Optimization to Chaotic Time Series Prediction. IEEE International Conference on Intelligent Computing and Intelligent Systems, 2009: 4: 458-462.
[7] D. Kugiumtzis, O. C. Lingjerde, N. Christophersen. Regularized local linear prediction of chaotic time series. Physica D, 1998:112: 344-360.
[8] Qingfang Meng, Yuhua Peng. A new local linear prediction model for chaotic time series. Physics Letters A, 2007: 370: 465–470.
[9] Kennel M B, Brown R, Abarbanel H D I. Determining embedding dimension for phase-space reconstruction using a geometrical construction. Physical Review A, 1992: 45: 3403-3411.
[10] F. Takens. On the numerical determination of the dimension of an attractor. Dynamical System and Turbulence, Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1981: 898:230-241.
[11] J.Kennedy, R.C.Eberhart. Particle swarm optimization. IEEE Service Center, 1995: IV: 1942-1948.
[12] James Kennedy, Russell C. Eberhart. A Discrete Binary Version Of The Particle Swarm Algorithm. IEEE Service Center, 1997:4104-4108.
[13] Xuyi Chun, Xiaoran Bin. An Improved Binary Particle Swarm Optimizer. Pattern Reorganization & Artificial Intelligent, 2007: 20(6): 788-793,.(in Chinese)
[14] E.N. Lorenz. Deterministic nonperiodic flow. Journal of the Atmospheric Sciences, 1963: 20: 131-141.
[15] Yu Zhang, Jinhe Ran. Dynamic Weighted Track Fusion Algorithm Based on Track Comparability Degree. IEEE, 2010: 710-713.