Vol 18 No 8, August 2009
1674-1056/2009/18(08)/3287-08

Chinese Physics B

© 2009 Chin. Phys. Soc.
and IOP Publishing Ltd

A method to improve the precision of chaotic time series prediction by using a non-trajectory

Yan Hua(闫 华), Wei Ping(魏 平), and Xiao Xian-Ci(肖先赐)

School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 610054, *China*

(Received 3 September 2008; revised manuscript received 23 February 2009)

Due to the error in the measured value of the initial state and the sensitive dependence on initial conditions of chaotic dynamical systems, the error of chaotic time series prediction increases with the prediction step. This paper provides a method to improve the prediction precision by adjusting the predicted value in the course of iteration according to the evolution information of small intervals on the left and right sides of the predicted value. The adjusted predicted result is a non-trajectory which can provide a better prediction performance than the usual result based on the trajectory. Numerical simulations of two typical chaotic maps demonstrate its effectiveness. When the prediction step gets relatively larger, the effect is more pronounced.

Keywords: non-trajectory, chaotic time series, prediction
PACC: 0545

1. Introduction

The existence of chaos has been demonstrated in many different kinds of physical systems and phenomena. Some typical ones are, for example, the ocean-atmosphere system,[1] the solar system,[2,3] earthquakes[4] and the like. Meteorologists research chaotic phenomena related to the weather with the hope of forecasting weather more accurately, especially for some weather disasters such as hurricanes, floods and droughts. Astronomers study chaotic motion in the solar system aiming to predict the motion of bodies and investigate the stability and longevity of planetary systems. Seismologists use chaotic theory to help them understand the formation of earthquakes and improve the prediction of it. Because of the extensive existence of chaos in nature, the prediction of chaotic time series plays an important role in the scientific studies of related fields.

To predict chaotic time series, we first need to build a model and then take the noisy measured values as initial conditions of the model for iteration. The central problem of the first step is how to build the model appropriately. This is an active research area all the time[5−9] and many different methods have been developed for this purpose. Some representative methods are various types of neural networks, such as multilayer perceptions,[10−13] radial basis function neural networks[14−22] and recurrent neural networks,[23−26] support vector machines,[27−32] local prediction[33−37] and polynomial models.[38,39] There is relatively less research about the second step. Usually, some noise reduction methods are firstly applied to the measured values in order to reduce the influence of noises.[39−42] Then the processed values are iterated with the built model to make multi-step-ahead predictions. Due to the sensitive dependence on initial conditions for chaotic dynamical systems, the prediction error increases with the prediction step. In recent years, some researchers have devoted themselves to the improvement of prediction effectiveness by adjusting the predicted values appropriately in the course of iteration. Strelioff and Hübler[43] suggested a method to extend prediction to chaotic systems with fold dynamics by using probability density function dynamics. Their critical idea is to recognize cases when the probability density near the critical point is sufficient to create a new peak first, and then to use the image of the critical point to replace the image of the current most probable state as the most likely point for the next step. This method can improve the prediction effectiveness of the most likely trajectory. Danforth and Yorke[44] studied the case when an ensemble of initial conditions is used in the course of iteration for specifying the probability associated with each possible outcome. They showed how to inflate or systematically perturb the ensemble of solutions so that the effectiveness of the forecast can be improved. While they considered the problem of improving the statisti-

cally predicted results during the iterative procedure, we mainly discuss how to improve an exact trajectory prediction during the iterative process. Usually, a trajectory (we will give its definition later) of the model built is used as the predicted result of a true trajectory, but in the method we propose, a non-trajectory is used as the predicted result for improving the prediction effectiveness. In order to concentrate on the course of iteration, we assume that the perfect model which produces the chaotic time series is available. This means that the model has been well built and can be used for iteration. Therefore, this paper will discuss mainly the following problem: when the perfect model is available, the noises in measured values will lead to the degradation of prediction precision as the prediction step increases. We will give an alternative approach to improve the precision of chaotic time series prediction for this case.

This paper is organised as follows. In Section 2, we give a one-dimensional mathematical model which we use to explain our method. Then a probabilistic model of initial conditions is introduced as a mathematical description of measurement. In Section 3, we explain why the prediction precision degrades with the increase of the prediction step and this directly motivates our proposed method. Then in Sections 4 and 5, the proposed method is specified and numerical experiments of two typical one-dimensional chaotic maps are reported to show its effectiveness. Finally, a short conclusion is drawn in Section 6.

2. A model of a one-dimensional chaotic map and a mathematical description of measurement

We use the following one-dimensional map to specify our idea

$$x_{k+1} = f(x_k), \tag{1}$$

whose domain is a closed interval $[x_{\min}, x_{\max}]$. For our method, we require that f should have a continuous derivative in its domain.

Measurement is the precondition of prediction. We assume that the measured values have a normal distribution $N(\bar{x}_0, \sigma^2)$, where $\bar{x}_0$ represents the true initial value and σ^2 is a small variance reflecting the non-systematic noises in the measurement. Since a single measured value is not reliable enough for the existence of noises, we carry out measurements N times for $\bar{x}_0$, and then take the mean of all the measured values as the estimated value which will be substituted into Eq.(1) for prediction by iteration. Since σ^2 is small, we can regard the true value and the estimated value of the initial state as two neighbors in the domain of Eq.(1).

For the convenience of the following description, we define a trajectory of Eq.(1) as $\{x_0, x_1, x_2, \ldots\}$, where each pair of adjacent points x_k and x_{k+1} satisfies Eq.(1).

3. An explanation of the degradation of prediction precision

We have already mentioned that the prediction precision degrades as the prediction step increases. This is due to the effect of the positive Lyapunov exponent. The Lyapunov exponent is a very important parameter of a dynamical system which describes the rate of divergence or convergence of nearby trajectories. A dynamical system with a positive Lyapunov exponent has the property of sensitive dependence on initial conditions, which means that no matter how close two initial states are, the two trajectories starting from them will diverge exponentially fast from each other. Since the true value and the estimated value of an initial state are regarded as two neighbors in the domain, the trajectories starting from them will diverge from each other. This is the reason why the prediction precision degrades as the prediction step increases.

For specifying our critical idea, we continue to consider how nearby trajectories starting from two neighbors in the domain of Eq.(1) diverge from each other. This problem refers to images of two neighbors under Eq.(1). Instead of only considering images of two neighbors, we consider the image of a small closed interval $[a_0, b_0]$ whose length is $\delta_0 = b_0 - a_0$, which is a small positive value. According to the intermediate value theorem and the maximum value theorem,[45] it is easy to prove that the image of a closed interval under a continuous map is still a closed interval. So we can use $[a_1, b_1]$ to denote $f([a_0, b_0])$, where $a_1 = \min f([a_0, b_0])$ and $b_1 = \max f([a_0, b_0])$. Let $\delta_1 = b_1 - a_1$ be the length of the image. Then, when f is a monotonic function in $[a_0, b_0]$, it is obvious that the following relation

$$|f'(x)| \approx \delta_1/\delta_0 \tag{2}$$

is satisfied for any x that belongs to $[a_0, b_0]$ since f has a continuous derivative and δ_0 is small. In this instance, if $\delta_1/\delta_0 > 1$, the effect of f is expansion. Otherwise, for $\delta_1/\delta_0 < 1$, the effect of f is contraction. When f is not a monotonic function in $[a_0, b_0]$, the effect is to fold. For this case, relation (2) is not satisfied and we deal with this case differently from the monotonic case in our method. We will specify this later. The two trajectories starting from two neighbors diverge from each other due to the expansion effect of map f. Although the contraction effect makes two points closer, the positive Lyapunov exponent of map f determines that the expansion effect predominates over the contraction effect. Because of this, the true value and the estimated value, as neighbors in the domain, will finally move apart under iteration.

4. Using a non-trajectory to improve the prediction precision

The above-mentioned cause of degradation of prediction precision directly motivates the method we proposed. In fact, the critical idea of this method is to adjust the predicted value when it diverges from the true value to a certain extent under iteration. The adjustment makes the predicted value move closer to the true value so that the prediction precision can be improved. After the adjustment, since the predicted result no longer satisfies the definition of the trajectory (that is to say, the predicted values at time k and $k+1$ may not satisfy Eq.(1)), it is a non-trajectory. In the following paragraphs, we expound how to adjust the predicted value properly in detail.

Suppose the current time is k. Besides making a prediction by applying f to the current predicted value x_k only, we also apply f to small intervals $I_{k,L} = [x_k - d^b_{k,L}, x_k]$ and $I_{k,R} = [x_k, x_k + d^b_{k,R}]$ on the left and right sides of x_k respectively, where $d^b_{k,L}$ and $d^b_{k,R}$ are small positive numbers denoting the lengths of intervals at time k before iteration. The subscripts L and R mean left and right, respectively, and the superscript b means before iteration. Usually, we choose $d^b_{k,L} = d^b_{k,R} = d$ as long as $I_{k,L}$ and $I_{k,R}$ are in the domain $[x_{\min}, x_{\max}]$ of f, where d is a small positive number. Otherwise let $d^b_{k,L} = x_k - x_{\min}$ and $d^b_{k,R} = x_{\max} - x_k$. The purpose of obtaining $d^b_{k,L}$ and $d^b_{k,R}$ in this way is to ensure that the intervals $I_{k,L}$ and $I_{k,R}$ are in the domain $[x_{\min}, x_{\max}]$ of f. Applying f to $I_{k,L}$ and $I_{k,R}$, we obtain $f(I_{k,L})$ and $f(I_{k,R})$. Let $d^a_{k,L}$ and $d^a_{k,R}$ be their lengths respectively, where the superscript a means after iteration.

Suppose that the true value $\bar{x}_k$ and the predicted value x_k at time k are neighbors in the domain and then we have $\bar{x}_k \in I_{k,L}$ or $\bar{x}_k \in I_{k,R}$. For the convenience of specification, we only discuss the first situation $\bar{x}_k \in I_{k,L}$ and the discussion for the second one is quite similar. If f is monotonically increasing in $I_{k,L}$, then $\bar{x}_{k+1} \le x_{k+1}$, where $\bar{x}_{k+1} = f(\bar{x}_k)$ and $x_{k+1} = f(x_k)$ denote, respectively, the true value and the predicted one at time $k+1$. Usually, strict equality is rare, so we can deem $\bar{x}_{k+1} < x_{k+1}$. If f is monotonically decreasing in $I_{k,L}$, then $\bar{x}_{k+1} > x_{k+1}$. If f does not change monotonically, the fold phenomenon we have mentioned before occurs. For this case, since we cannot judge which one of $\bar{x}_{k+1}$ and x_{k+1} is bigger, we use the following value to replace x_{k+1}:

$$\tilde{x}_{k+1} = \min f(I_{k,L}). \tag{3}$$

Because $\bar{x}_k \in I_{k,L}$, we have $\bar{x}_{k+1} \in f(I_{k,L})$. So we can judge $\bar{x}_{k+1} > \tilde{x}_{k+1}$. After obtaining the order relation of $\bar{x}_{k+1}$ and x_{k+1}, the relation between $\bar{x}_{k+2}$ and x_{k+2} can be judged by the same method. Through the above method, we can judge the order relation of true values and prediction values at time $k+1, k+2, \ldots$, if the order relation at time k has been known.

From the above analysis, we see that the true and predicted trajectories diverge gradually. Hereinafter, we will quantitatively analyze this separation process. Suppose the current time is k and the true value is $\bar{x}_k$. Then the true value at time $k+n$ is $\bar{x}_{k+n} = f^n(\bar{x}_k)$, where f^n is the composition of f with itself n times. According to the chain rule of differentiation, we have

$$\left.\frac{\mathrm{d}f^n(x)}{\mathrm{d}x}\right|_{x=\bar{x}_k} = \prod_{i=0}^{n-1} f'(\bar{x}_{k+i}), \tag{4}$$

where $\bar{x}_{k+i} = f^i(\bar{x}_k), i = 0, 1, 2, \ldots, n-1$. Here, $f^0(\bar{x}_k) = \bar{x}_k$. Since $\bar{x}_k$ and x_k are neighbors, the following relation is satisfied:

$$\left|\left.\frac{\mathrm{d}f^n(x)}{\mathrm{d}x}\right|_{x=\bar{x}_k}\right| \approx \left|\frac{f^n(\bar{x}_k) - f^n(x_k)}{\bar{x}_k - x_k}\right| = \left|\frac{\bar{x}_{k+n} - x_{k+n}}{\bar{x}_k - x_k}\right|. \tag{5}$$

If there is no fold phenomenon during the course of iteration of intervals at time $k+i$ $(i = 0, 1, 2, \ldots, n-1)$, according to Eq.(2), we have

$$|f'(\bar{x}_{k+i})| \approx \frac{\mathrm{d}^a_{k+i,C_{k+i}}}{\mathrm{d}^b_{k+i,C_{k+i}}}, \quad i = 0, 1, 2, \ldots, n-1, \tag{6}$$

where the subscript C_{k+i} can take either L or R in terms of $\bar{x}_{k+i} \in I_{k+i,L}$ or $\bar{x}_{k+i} \in I_{k+i,R}$ which can be determined by the previous method.

Substitute Eqs.(5) and (6) into the left and right sides of Eq.(4) respectively to yield

$$\left|\frac{\bar{x}_{k+n} - x_{k+n}}{\bar{x}_k - x_k}\right| \approx \prod_{i=0}^{n-1} \frac{\mathrm{d}^a_{k+i,C_{k+i}}}{\mathrm{d}^b_{k+i,C_{k+i}}}. \tag{7}$$

Let $r_{k,k+n-1} = \prod_{i=0}^{n-1} \frac{\mathrm{d}^a_{k+i,C_{k+i}}}{\mathrm{d}^b_{k+i,C_{k+i}}}$. It denotes the total expansion/contraction ratio from time k to $k+n-1$. Then we have

$$\bar{x}_{k+n} \approx x_{k+n} \pm r_{k,k+n-1} \left|\bar{x}_k - x_k\right|. \tag{8}$$

On the right-hand side of Eq.(8), x_{k+n} (the predicted value at time $k+n$) and $r_{k,k+n-1}$ can be obtained through computation. $|\bar{x}_k - x_k|$ is the absolute value of prediction error at time k and it cannot be obtained during the course of prediction because the true value $\bar{x}_k$ is not known, but we can make a rough estimate of it. On the left-hand side of Eq.(8), $\bar{x}_{k+n}$ is the true value we intend to predict. According to Eqs.(3) and (8), we propose the following method for improving the prediction precision: when the fold phenomenon occurs at $k+n-1$, x_{k+n} should be adjusted according to Eq.(3); when no fold phenomenon is observed from time k to $k+n-1$, since $r_{k,k+n-1}$ reflects the total expansion/contraction extent during this period of time, we choose a critical value λ and propose that if $r_{k,k+n-1} \le \lambda$, we make no adjustment to x_{k+n}. Otherwise, as long as $r_{k,k+n-1} > \lambda$ (that is to say, the expansion reaches a certain extent), x_{k+n} should be adjusted according to Eq.(8). That is to say, x_{k+n} is adjusted as follows:

$$\tilde{x}_{k+n} = \begin{cases} \min f\left(I_{k+n-1,C_{k+n-1}}\right), & \text{if } A, \\ x_{k+n}, & \text{if } B \text{ and } r_{k,k+n-1} \le \lambda, \\ x_{k+n} \pm r_{k,k+n-1}\Delta_k, & \text{if } B \text{ and } r_{k,k+n-1} > \lambda, \end{cases} \tag{9}$$

where the symbol A denotes the fact that the fold phenomenon occurs at time $k+n-1$; the symbol B denotes the fact that no fold phenomenon occurs from time k to $k+n-1$; the subscript C_{k+n-1} has the same meaning as in Eq.(6) and Δ_k denotes an estimate of $|\bar{x}_k - x_k|$. Usually, Δ_k is chosen as a small positive number. The sign $\pm$ in Eq.(9) is chosen as follows: if $\bar{x}_{k+n} > x_{k+n}$, then the plus sign is taken; if $\bar{x}_{k+n} < x_{k+n}$, then the minus sign is taken. Here the order relation of $\bar{x}_{k+n}$ and x_{k+n} can be obtained by the previous method. After making an adjustment according to the first or third equation in Eq.(9), the starting time point of computing the ratio $r_{p,q}$ should be reset to $p = k+n$. The reason is as follows. If an adjustment is made according to the first equation, then the fold phenomenon occurs. Since we cannot estimate the derivative by using Eq.(6) for this case, the starting time point should be reset. If an adjustment is made according to the third equation, then from Eq.(8), we can find that the predicted value newly adjusted becomes closer to the true value than the one without an adjustment. That is to say, the expansion effect from the time point k to $k+n-1$ has been counteracted to a certain extent, so we should reset the starting time point p to $k+n$.

Because the intervals $I_{j,L}$ and $I_{j,R}$ at every time step j are very short, usually map f varies monotonically in these intervals. This means that very few fold phenomena occur during the course of iteration. Thus, when the parameters λ and Δ_k are chosen properly, an adjustment will be made every few time steps according to the third equation in Eq.(9). This makes the predicted value after being adjusted closer to the true value, namely, the rate of separation slows down, so the prediction precision is improved.

When an adjustment is made according to the first or third equation in Eq.(9), it is obvious that $\tilde{x}_{k+n}$ and x_{k+n-1} no longer satisfy Eq.(1), so the predicted result is a non-trajectory. Hereinafter, in order to avoid confusion, we will call the predicted result a predicted sequence, instead of the usual name — a predicted trajectory.

When using the above method, we should know the order relation of the estimated initial value x_0 and the true initial value $\bar{x}_0$, but without other information besides x_0, we cannot determine this. Therefore, there

will be two different predicted sequences corresponding to two situations $\bar{x}_0 < x_0$ and $\bar{x}_0 > x_0$. Since one of the two situations must be wrong, the corresponding predicted sequence may not improve the prediction precision. We choose the sequence with a better prediction effectiveness as the final predicted result in the following simulations.

5. Simulation experiments and discussion

To validate the effectiveness of the proposed method, two one-dimensional chaotic maps are used for the numerical simulations.

Simulation experiment 1 The first simulation experiment is carried out by using the logistic map $x_{k+1} = 4x_k(1-x_k), x_k \in [0,1], k = 0,1,\ldots$. This is a typical chaotic dynamical system[46] and $[0,1]$ is its domain. To test the average performance of this method for the whole domain, we divide $[0,1]$ into $N_1 = 10^3$ subintervals $[(i-1)/N_1, i/N_1], i = 1,2,\ldots,N_1$ of uniform length. Then we choose N_1 true initial values $\bar{x}_0^i, i = 1,2,\ldots,N_1$ as follows: $\bar{x}_0^i$ is chosen from the i-th subinterval according to uniform distribution. Let $\bar{T}^i = \{\bar{x}_k^i, k = 0,1,\ldots,N_2\}$ be the true trajectory starting from $\bar{x}_0^i$, where $\bar{x}_{k+1}^i = 4\bar{x}_k^i(1-\bar{x}_k^i), k = 0,1,\ldots,N_2-1$ and $N_2 = 10$. To assess the performance of the method in the sense of statistical average, we compute $N_3 = 10^4$ predicted sequences for each $\bar{T}^i$. The q-th sequence is obtained as follows. Firstly, measure the initial value $\bar{x}_0^i$. As we have mentioned before, the measured values have a normal distribution $N(\bar{x}_0^i, \sigma^2)$, where $\sigma^2 = 10^{-5}$. Produce $N_4 = 5 \times 10^3$ measured values for each $\bar{x}_0^i$ and take their mean value $x_0^{i,q}$ as the estimated value of $\bar{x}_0^i$, where $q = 1,2,\ldots,N_3$. Secondly, apply the logistic map to $x_0^{i,q}$ for N_2 times and adjust the predicted values by using the proposed method in the course of iteration. Since applying f to the whole intervals $I_{k,L}$ and $I_{k,R}$ is impossible in actual computation, we use N_L and N_R equidistant-spacing points in these two intervals separately for iteration. Here, $N_L = N_R = 10^3$. The other parameters of the method are taken to be $d = 10^{-4}$ and $\lambda = 6$. Suppose that adjustments in terms of the third equation of Eq.(9) occur at time points $k_i, i = 1,2,\ldots$ ($k_i < k_{i+1}$ for each i) and the first six corresponding values of Δ_k are $\Delta_{k_1} = 10^{-6}, \Delta_{k_2} = 10^{-4}, \Delta_{k_3} = 10^{-3}, \Delta_{k_4} = 10^{-3}, \Delta_{k_5} = 10^{-3}$ and $\Delta_{k_6} = 10^{-3}$ (in fact, during the N_2 times iteration, the number of adjustments is less than six, so giving six values of Δ_k is enough). When the first adjustment occurs, the error between the predicted value and the true value is very small, so Δ_{k_1} is taken as a small value. Since the error increases with the prediction step, we also increase the value of Δ_{k_i} gradually. Let $\{x_k^{i,q}, k = 0,1,\ldots N_2\}$ be the q-th predicted sequence of $\bar{T}^i$.

We validate the improvement of the prediction precision through comparison with the direct iteration method. The direct iteration method means applying the logistic map to $x_0^{i,q}$ for N_2 times directly without any adjustment and let $\hat{x}_k^{i,q}$ be its results. Here parameters i, q and k have the same meaning as above.

For $\bar{T}^i$, the absolute value of the estimation error of $\bar{x}_0^i$ of the q-th predicted sequence is $e_0^{i,q} = |x_0^{i,q} - \bar{x}_0^i|$ and the absolute values of the prediction errors of the two methods are $e_k^{i,q} = |x_k^{i,q} - \bar{x}_k^i|$ and $\hat{e}_k^{i,q} = |\hat{x}_k^{i,q} - \bar{x}_k^i|, k = 1,2,\ldots,N_2$ respectively. So the mean values with respect to N_3 sequences are $E_0^i = \left(\sum_{q=1}^{N_3} e_0^{i,q}\right)/N_3$, $E_k^i = \left(\sum_{q=1}^{N_3} e_k^{i,q}\right)/N_3$ and $\hat{E}_k^i = \left(\sum_{q=1}^{N_3} \hat{e}_k^{i,q}\right)/N_3, k = 1,2,\ldots,N_2$. They reflect the average estimation and prediction errors. Since the prediction errors can also be considered as the distances between the predicted and true values, we call E_k^i and $\hat{E}_k^i$ ($k = 1,2,\ldots,N_2$) the average distances between the predicted sequence and the true trajectory. The fact that E_k^i and $\hat{E}_k^i$ increase with k reflects the divergence of the predicted sequence and true trajectory. So we can give the following definition: for $\bar{T}^i$, the average divergent speed from time 0 to k ($k > 0$) between the predicted sequence and the true trajectory $\bar{T}^i$ is $v_{0,k}^i = \left(E_k^i - E_0^i\right)/k$ for the method we proposed and is $\hat{v}_{0,k}^i = \left(\hat{E}_k^i - E_0^i\right)/k$ for the direct iteration method. Then, for $\bar{T}^i$, the relative decreasing percentage of the average divergent speed from time 0 to k is $p_{0,k}^i = \left(\hat{v}_{0,k}^i - v_{0,k}^i\right)/\hat{v}_{0,k}^i$ which reflects the improvement of performance at $\bar{x}_0^i \in [(i-1)/N_1, i/N_1]$ by using the method we proposed. Since the length of $[(i-1)/N_1, i/N_1]$ is small, we can consider that $p_{0,k}^i$ represents the performance gain in the i-th subinterval. So the mean $P_{0,k} = \left(\sum_{i=1}^{N_1} p_{0,k}^i\right)/N_1$ represents the average performance gain till time k ($k > 0$) in the whole domain and we call it the mean of relative decreasing percentage of average divergent speed. It is obvious that $P_{0,k} > 0$ means improvement of the prediction precision while $P_{0,k} < 0$ means degradation

after applying our method, and the bigger the absolute value of $P_{0,k}$ is, the stronger the corresponding effect is.

Figure 1(a) shows the simulation result under the above conditions. From Fig.1(a), we can see that $P_{0,k} \geq 0$ for every prediction step k. This indicates that our method improves the prediction precision. From Fig.1(a), we can also note that $P_{0,1} = 0\%$. This is because the maximum of the absolute value of the derivative of the logistic map in $[0, 1]$ is 4, which is smaller than $\lambda = 6$. So at the first prediction step, our method does not make any adjustment to the predicted value. Subsequently, as the prediction step increases from 2 to 10, the approach makes the divergent speed between the predicted sequence and the true trajectory decrease, and $P_{0,k}$ increases from 0.41% to 33.27%. This reflects that the approach brings more performance gain for relatively larger prediction steps. The reason is as follows: first, an adjustment at a certain prediction step makes the divergent speed at some subsequent steps also decrease; second, as the prediction step increases, the number of adjustments also increases. So we obtain more performance gain for relatively larger prediction steps. But this does not mean that the larger the prediction step is, the more gain we can obtain. Actually, although the algorithm can decrease the divergent speed, it cannot stop the divergence of the predicted sequence and true trajectory. So when the prediction step is large enough, they will eventually diverge.

Figure 1(a) shows the average performance gain in the whole domain. In Fig.1(b), we give the E_k^i and $\hat{E}_k^i$ $(k = 1, 2, \ldots, N_2)$ for a certain $\bar{x}_0^i$ in order to reflect the change of the distances between the predicted sequence and the true trajectory before and after the application of our algorithm. From this figure, we can see that by applying our approach, the distances certainly become closer and as the prediction step increases, the effect becomes more pronounced. This is also consistent with the above analysis.

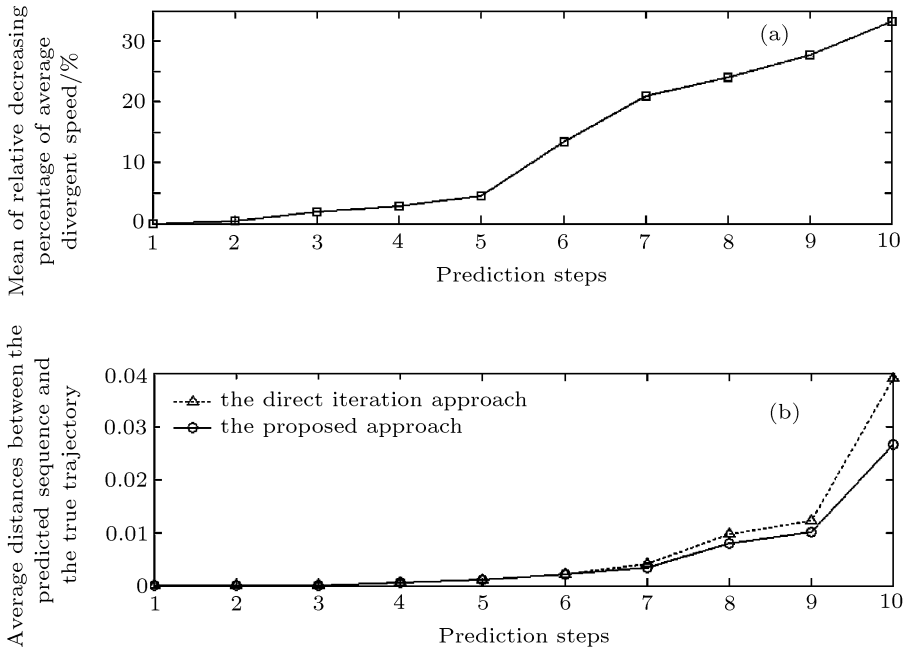


Fig.1. (a) The mean of the relative decreasing percentage of the average divergent speed $P_{0,k}$. It measures the average performance gain of the algorithm in the whole domain. The fact that $P_{0,k} \geq 0$ for every prediction step k indicates the improvement of the prediction precision. For relatively larger prediction steps, the proposed approach can provide a better performance gain. (b) A comparison of average distances between the predicted sequence and the true trajectory before and after the application of our approach for a certain true initial value. This result is consistent with Fig.1(a).

Simulation experiment 2 The second simulation experiment is performed using the Kent map[46]

$$x_{k+1} = \begin{cases} x_k/a, & \text{if } x_k \in [0, a) \\ (1 - x_k)/(1 - a), & \text{if } x_k \in [a, 1] \end{cases}, \ k = 0, 1, \ldots, \tag{10}$$

where $a = 0.4$. Its domain is $[0, 1]$. All the simulation conditions and parameters are the same as those for the first simulation experiment. The result is shown in Fig.2. Figure 2(a), which plots $P_{0,k}$ versus the prediction

step k, indicates that, by applying the proposed method, the divergent speed between the predicted sequence and the true trajectory decreases and the prediction precision is improved. The monotonically increasing curve in Fig.2(a) also indicates that the improvement is more significant for relatively larger prediction steps. Figure 2(b) plots E_k^i and $\hat{E}_k^i$ versus the prediction step k for a certain $\bar{x}_0^i$. By comparing the two curves in Fig.2(b), it can be seen that the proposed approach certainly makes the rate of divergence of the predicted sequence and true trajectory decrease. This also confirms the improvement of the prediction precision.

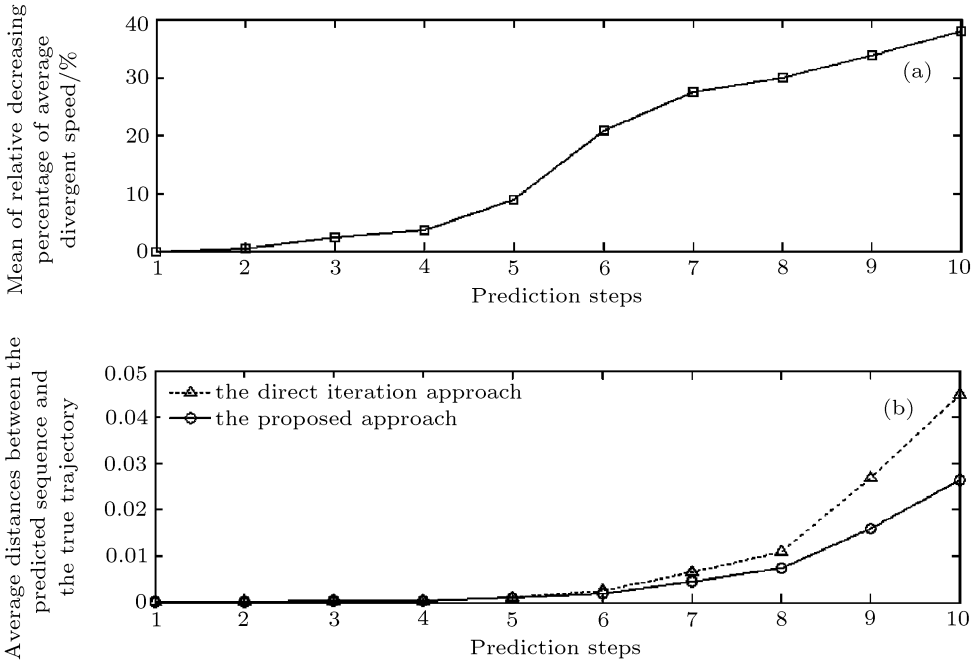


Fig.2. (a) The mean of relative decreasing percentage of average divergent speed $P_{0,k}$. (b) A comparison of average distances between the predicted sequence and the true trajectory before and after the application of our approach for a certain true initial value.

Here we point out that the Kent map does not satisfy the condition of having a continuous derivative in its domain (see the requirement for the function f in Eq.(1)) since it is not differentiable at $x = a$. Even so, the simulation result shows that our approach remains valid. This indicates that, in practice, the requirement for the function f in Eq.(1) need not to be strictly satisfied, and only needs to be approximately satisfied.

The above two simulation experiments demonstrate that the non-trajectory can certainly provide better prediction performance than the usual method (i.e., the direct iteration method) using the trajectory for prediction, so the algorithm we proposed can indeed improve the precision of chaotic time series prediction. Especially when the prediction step is relatively larger, the effect becomes more marked.

6. Conclusion

This paper has provided a method which can improve the precision of chaotic time series prediction by adjusting the predicted value properly according to the evolution of small intervals on the left and right sides of the predicted value. Compared with the usual predicted result based on the trajectory, our predicted result which is a non-trajectory provides better prediction performance. We have demonstrated its effectiveness by applying it to the logistic map and the Kent map. The reason why this approach can bring a performance gain is that: we consider not only the iteration of the predicted value itself, but also the evolution information of small intervals in the neighborhood of the predicted value. In this paper, we always assume that the perfect model is available because we want to focus on the problem of adjusting predicted values properly during iteration. But this is not a limitation of our method. In fact, for the model built according to actual data, the proposed method is still applicable. But the difference between the built model and the perfect one may influence the

effect to a certain extent. We hope that the approach we proposed can be applied to some related areas of chaotic time series prediction.

References

[1] Tziperman E, Stone L, Cane M A and Jarosh H 1994 *Science* **264** 72
[2] Lissauer J J 1999 *Rev. Mod. Phys.* **71** 835
[3] Sussman G J and Wisdom J 1992 *Science* **257** 56
[4] de Sousa Vieira M 1999 *Phys. Rev. Lett.* **82** 201
[5] Chng E S, Chen S and Mulgrew B 1996 *IEEE Trans. Neural Netw.* **7** 190
[6] Karunasinghe D S K and Liong S Y 2006 *Journal of Hydrology* **323** 92
[7] Small M and Tse C K 2002 *Phys. Rev.* E **66** 066701
[8] Abel M, Ahnert K, Kurths J and Mandelj S 2005 *Phys. Rev.* E **71** 015203(R)
[9] Shi Z W and Han M 2007 *IEEE Trans. Neural Netw.* **18** 359
[10] Cybenko G 1989 *Mathematics of Control, Signals, and Systems* **2** 303
[11] Hornik K, Stinchcombe M and White H 1989 *Neural Networks* **2** 359
[12] Hornik K, Stinchcombe M and White H 1990 *Neural Networks* **3** 551
[13] Funahashi K 1989 *Neural Networks* **2** 183
[14] Broomhead D S and Lowe D 1988 *Complex Systems* **2** 321
[15] Hartman E, Keeler J D and Kowalski J M 1990 *Neural Computation* **2** 210
[16] Park J and Sandberg I W 1991 *Neural Computation* **3** 246
[17] Park J and Sandberg I W 1993 *Neural Computation* **5** 305
[18] Poggio T and Girosi F 1990 *Proceedings of the IEEE* **78** 1481
[19] Leung H, Lo T and Wang S C 2001 *IEEE Trans. Neural Netw.* **12** 1163
[20] Li J and Liu J H 2005 *Acta Phys. Sin.* **54** 4569 (in Chinese)
[21] Harpham C and Dawson C W 2006 *Neurocomputing* **69** 2161
[22] Zhang J F and Hu S S 2007 *Acta Phys. Sin.* **56** 713 (in Chinese)
[23] Zhang J S and Xiao X C 2000 *Chin. Phys. Lett.* **17** 88
[24] Han M, Xi J H, Xu S G and Yin F L 2004 *IEEE Trans. Signal Process.* **52** 3409
[25] Han M, Shi Z W and Guo W 2007 *Acta Phys. Sin.* **56** 43 (in Chinese)
[26] Ma Q L, Zheng Q L, Peng H, Zhong T W and Qin J W 2008 *Chin. Phys.* B **17** 536
[27] Thissen U, Van Brakel R, de Weijer A P, Melssen W J and Buydens L M C 2003 *Chemometrics and Intelligent Laboratory Systems* **69** 35
[28] Mukherjee S, Osuna E and Girosi F 1997 *Neural Networks for Signal Processing* [1997] VII. *Proceedings of the 1997 IEEE Workshop* 511
[29] Müller K R, Smola A J, Rätsch G, Schölkopf B, Kohlmorgen J and Vapnik V 1997 *Proceedings of the 7th International Conference on Artificial Neural Networks* 999
[30] Liu H, Liu D and Deng L F 2006 *Chin. Phys.* **15** 1196
[31] Zhang J S, Dang J L and Li H C 2007 *Acta Phys. Sin.* **56** 67 (in Chinese)
[32] Zhang J F and Hu S S 2008 *Acta Phys. Sin.* **57** 2708 (in Chinese)
[33] Giona M, Lentini F and Cimagalli V 1991 *Phys. Rev.* A **44** 3496
[34] Gong X F and Lai C H 1999 *Phys. Rev.* E **60** 5463
[35] Kugiumtzis D, Lingjærde O C and Christophersen N 1998 *Physica* D: *Nonlinear Phenomena* **112** 344
[36] Alparslan A K, Sayar M and Atilgan A R 1998 *Phys. Rev.* E **58** 2640
[37] Meng Q F and Peng Y H 2007 *Phys. Lett.* A **370** 465
[38] Yan H, Wei P and Xiao X C 2007 *Acta Phys. Sin.* **56** 5111 (in Chinese)
[39] Kantz H and Schreiber T 1997 *Nonlinear Time Series Analysis* (Cambridge: Cambridge University Press) pp51, 154, 211
[40] Kostelich E J and Yorke J A 1988 *Phys. Rev.* A **38** 1649
[41] Kostelich E J and Schreiber T 1993 *Phys. Rev.* E **48** 1752
[42] Schreiber T and Kantz H 1995 *Chaos* **5** 133
[43] Strelioff C C and Hübler A W 2006 *Phys. Rev. Lett.* **96** 044101
[44] Danforth C M and Yorke J A 2006 *Phys. Rev. Lett.* **96** 144102
[45] Munkres J R 2004 *Topology, Second Edition* (Beijing: China Machine Press) p147 (in Chinese)
[46] Huang R S and Huang H 2005 *Chaos and Its Applications, Second Edition* (Wuhan: Wuhan University Press) pp134, 141, 402 (in Chinese)