A Method for Predicting Chaotic Time Series based on a Local Approximation Technique

Ken-ichi ITOH and Yoko MAEMURA

SUMMARY An improved method based on a local approximation technique is presented for predicting chaotic time series. In the local approximation technique, a state space is reconstructed from a time series using delay coordinates and then a local predictor is constructed on the basis of the motion of the nearest neighbors in the state space. To increase the prediction accuracy of the local approximation, a new method is proposed for selecting the nearest neighbors in the state space. The efficacy of the proposed method is demonstrated using chaotic time series generated by the Ikeda map and the Hénon map.

key words: *chaos, time series, prediction*

1. Introduction

Short-term prediction methods for chaotic time series are being intensively studied [1]-[4]. The local approximation technique [1], [2], one of the prediction methods for chaotic time series, is outlined below. The first step is to reconstruct a state space from a univariate time series using delay coordinates [5]. The next step is to assume a functional relationship between the current point and a future point in the state space. To predict this future point, the nearest neighbors of the current point are found in the state space and then a local predictor is constructed on the basis of the motion of the nearest neighbors. The relationship between state space reconstruction and prediction accuracy was discussed in detail by Casdagli et al. [6].

In the state space, the direction of the trajectory of the nearest neighbor sometimes differs from the direction of the trajectory of the current point. In such cases, the accuracy of the prediction obtained using the local approximation method decreases because the motion of the nearest neighbors incorrectly approximates that of the current point.

To increase the prediction accuracy in such cases, this paper proposes a new method for selecting the nearest neighbors in a state space. The efficacy of the proposed method is confirmed through numerical experiments using chaotic time series generated using the Ikeda map [7] and the Hénon map [8].

Section 2 reviews the local approximation method. Section 3 defines the proposed method for selecting the nearest neighbors. Section 4 describes results of the experiments on the prediction accuracy. Section 5 concludes the paper.

2. Local Approximation Method

This section reviews the local approximation method for predicting chaotic time series [1]. Given a time series, x_i, the attractor can be reconstructed in an m-dimensional state space by forming the delay vector

$$X_i = (x_i, x_{i-\tau}, \cdots, x_{i-(m-1)\tau}), \tag{1}$$

where τ is the time delay [5]. Assume that the box-counting dimension of the attractor is D_0. If $m > 2D_0$, then m-dimensional delay vectors generically form an embedding of the original state space [9]. Namely, the structure of the attractor in the original state space is preserved in the m-dimensional state space. The prediction is executed by estimating the change of the trajectory with time in the m-dimensional state space.

Assume that the observed values x_1 to x_t up to time t are given and the value x_{t+p} at time p in the future is to be predicted. Using Eq. (1), the attractor is reconstructed from the observed values x_1 to x_t. The relationship between the current point, X_t, and the future point, X_{t+p}, on the attractor is approximated by function F,

$$X_{t+p} \cong F(X_t). \tag{2}$$

In the local approximation method, the nearest n neighbors X_{T_h} $(h = 1, 2, \cdots, n)$ of X_t are selected and then function F is estimated on the basis of the relationship between X_{T_h} and X_{T_h+p}. To predict x_{t+p}, it is sufficient to determine only the first element of X_{t+p}; it is not necessary to determine all m elements of X_{t+p}. Accordingly, the value x_{t+p} is estimated using linear polynomial f,

$$x_{t+p} \cong f(X_t) = a_0 + \sum_{k=1}^{m} a_k x_{t-(k-1)\tau}. \tag{3}$$

The coefficients $a_0, a_1, \cdots, a_m$ are determined as follows. For each point X_i in the m-dimensional state space, except X_t, the Euclidean distance r from X_t is calculated:

$$r = \| X_i - X_t \| = \Big(\sum_{k=0}^{m-1} (x_{i-k\tau} - x_{t-k\tau})^2\Big)^{1/2}. \tag{4}$$

By comparing r, the nearest n neighbors X_{T_h} of X_t are selected from among all points in the m-dimensional state space. The coefficients $a_0, a_1, \cdots, a_m$ can be calculated by using a least-squares fit:

$$\sum_{h=1}^{n} (x_{T_h+p} - f(X_{T_h}))^2 = \min. \qquad (5)$$

The minimum number of n needed to give a unique solution to the least-squares problem is $m + 1$; however, choosing an n larger than the minimum number decreases the prediction error.

3. Proposed Method

In the local approximation method, it is important to select the nearest neighbors correctly. Through reconstructing the attractor using delay coordinates, the nearest neighbors in a reconstructed space may differ from the nearest neighbors in the original space [6]. In such cases, even if the condition for embedding, i.e., $m > 2D_0$ is satisfied, the direction of the trajectory of the nearest neighbor sometimes differs from the direction of the trajectory of the current point in the m-dimensional state space. These incorrect nearest neighbors decrease the prediction accuracy.

To increase the prediction accuracy in the above cases, this paper proposes a new method for selecting the nearest neighbors. To predict accurately future point X_{t+p}, it is desirable to select the nearest neighbors, X_{T_h}, of current point X_t so that future points X_{T_h+p} of X_{T_h} are also as close as possible to X_{t+p}. However, it is difficult to select such nearest neighbors because the value of X_{t+p} is unknown. Hence, the motions of nearest neighbors are investigated retroactively to the past in this paper. Namely, the nearest neighbors of X_t are selected taking into consideration the distance between $X_{i-\tau}$ and $X_{t-\tau}$ as well as the distance between X_i and X_t. For that purpose, the following equation is introduced to replace Eq. (4):

$$r = w \cdot \left\| X_{i-\tau} - X_{t-\tau} \right\| + (1-w) \cdot \left\| X_i - X_t \right\| \quad , \qquad (6)$$

where w is a weighting factor. When $w = 0$, Eq. (6) is equal to Eq. (4). By using this new strategy for selecting the nearest neighbors, it is expected that the motion of the nearest neighbors better approximates that of the current point. As a result, the prediction accuracy should be increased.

In Fig. 1, it is desirable for predicting future point X_{t+p} to select X_b instead of X_a as the nearest neighbor of X_t. However, when the conventional method mentioned in Sec. 2 is applied, X_a is selected as the nearest neighbor because X_a is closer to X_t than X_b. By using the proposed method, X_b can be selected as the nearest neighbor.

4. Experimental Results

Experiments were conducted on the prediction accuracy of the proposed method. In the experiments, chaotic time series generated by the Ikeda map [7] and the Hénon map [8] were used.

The Ikeda map is a two-dimensional map,

$$\begin{aligned} x_{n+1} &= 1 + \mu\,(x_n \cos t - y_n \sin t), \\ y_{n+1} &= \mu\,(x_n \sin t + y_n \cos t), \end{aligned} \qquad (7)$$

where $t = 0.4 - 6.0 / (1 + x_n^2 + y_n^2)$ and the parameter value is $\mu = 0.7$. The time series taken was the x coordinate starting at $x_0 = y_0 = 0.3$. Figure 2(a) shows the phase plot of the Ikeda map.

The Hénon map is a two-dimensional map,

$$\begin{aligned} x_{n+1} &= y_n + 1 - Ax_n^2, \\ y_{n+1} &= Bx_n, \end{aligned} \qquad (8)$$

where the parameter values are $A = 1.4$ and $B = 0.3$. The time series taken was the x coordinate starting at $x_0 = y_0 = 0.3$. Figure 2(b) shows the phase plot of the Hénon map.

Assume that observed values v_i $(i = 1, 2, \cdots, N)$ for N steps, as well as their predicted values z_i $(i = 1, 2, \cdots, N)$ are given. The accuracy of the prediction is evaluated in terms of relative error E:

$$E = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(z_i - v_i)^2}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(v_i - \bar{v})^2}}, \qquad (9)$$

where $\bar{v} = \frac{1}{N}\sum_{i=1}^{N} v_i$. When $E = 0$, the prediction is perfect.

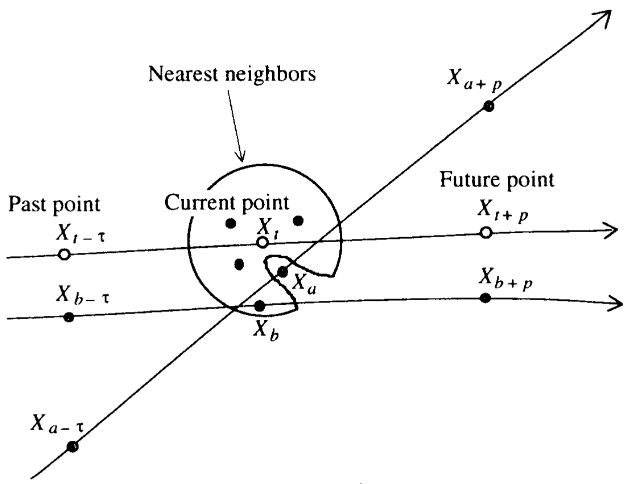


Fig. 1 A new method for selecting the nearest neighbors in the local approximation technique.

L steps starting from point s of the time series are used as the training data for reconstructing the attractor. The following N steps are used to predict p steps ahead. In other words, based on past observed values $v_s, v_{s+1}, \cdots, v_{s+L-1}$ and observed value v_i at the current point, predicted value z_{i+p} at p steps ahead is estimated and compared to observed value v_{i+p}. This procedure is iterated for N times for $i = s + L$ to $i = s + L + N - 1$. Using these N predicted values and observed values, the relative error E is determined.

The parameters are set as follows. Training data length L is varied from 200 to 6400. Prediction data length N is set to L. The time delay τ is set to 1. Time series data of 40,000 points each were prepared for the Ikeda map and the Hénon map. For each value of L, 10 sets of test data were constructed randomly selecting starting point s from the time series data.

Figure 3(a) shows an example of the prediction results for the Ikeda map with $L = 400$, $m = 3$, $p = 1$, and $n = 8$, when the conventional method is applied. The predicted value is nearly equal to the observed value for most prediction points. However, there are a few points with large prediction error. It is surmised that these prediction errors are caused by incorrectly selecting the nearest neighbors as noted in Sec. 3. By using the proposed method, these points with a large prediction error can be eliminated as shown in Fig. 3(b).

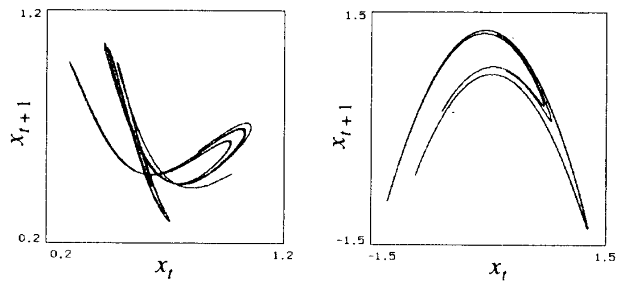
Figure 4 shows relative error E as a function of weighting factor w for the Ikeda map. The parameters are set as follows: $m = 3$, $p = 1$. The value of each point in Fig. 4 is the average for the 10 test data sets, as well as in the following figures. As shown in Fig.4, the relative error can be decreased by using weighting factor w. The reason why the prediction error for $w = 1$ is smaller than that for $w = 0$ remains to be clarified.

Figure 5 shows relative error E as a function of training data length L for the Ikeda map. The parameters are also set as follows: $m = 3$, $p = 1$. In Fig. 5, two cases of $w = 0.3$ and $w = 0.5$ are denoted for the proposed method. Similar results are obtained for the two cases of $n = 8$ and $n = 16$. When the conventional method, namely $w = 0$, is applied for the Ikeda map, points with large prediction error as shown in Fig. 3(a) are generated for every value of L. Therefore, the relative error is not significantly decreased when L is increased. By using the proposed method, the nearest neighbors can be correctly selected which substantially decreases the relative error.

Figure 6 shows relative error E as a function of prediction steps p for the Ikeda map. The parameters are set as follows: $m = 3$, $n = 8$, $w = 0.3$. In Fig. 6, the two cases of $L = 400$ and $L = 3200$ are denoted. When $p < 4$, the relative error is decreased by using the proposed method. When $p \geq 4$, the prediction accuracy of the proposed method is roughly equal to that of the conventional method. The results show that the proposed method is effective for the short-term prediction of time series of the Ikeda map.

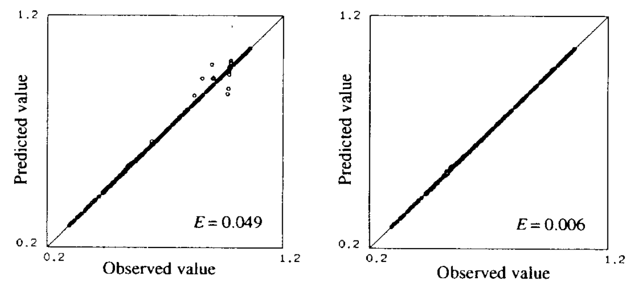
Figure 7 shows relative error E as a function of embedding dimension m for the Ikeda map. The parameters are set as follows: $p = 1$, $n = 8$, $w = 0.3$. In Fig. 7, two cases of $L = 400$ and $L = 3200$ are also denoted. When $m < 4$, the proposed method is very effective compared to the conventional method. The minimum embedding dimension of the Ikeda map is 3 because the condition for embedding is $m > 2D_0$ and D_0 of the Ikeda map is 1.32. However, when the conventinal method is applied, the relative error for $m = 3$ is considerably larger than the minimum value of the relative error obtained by changing m. By using the proposed method, the relative error for $m = 3$ is approximately held to the minimum value. This implies that when the proposed method is applied, it is not necessary to seek the optimal embedding dimension for prediction because a sufficient level of prediction accuracy can be obtained at the minimum embedding dimension.

Figures 8-11 show some test results for the Hénon map. The parameters of Figs. 8-11 are equal to the parameters of Figs. 4-7. When the conventional method is applied for the Hénon map, points with large prediction error such as those shown in Fig. 3(a) are not generated because the nearest neighbors are correctly selected. Accordingly, the relative



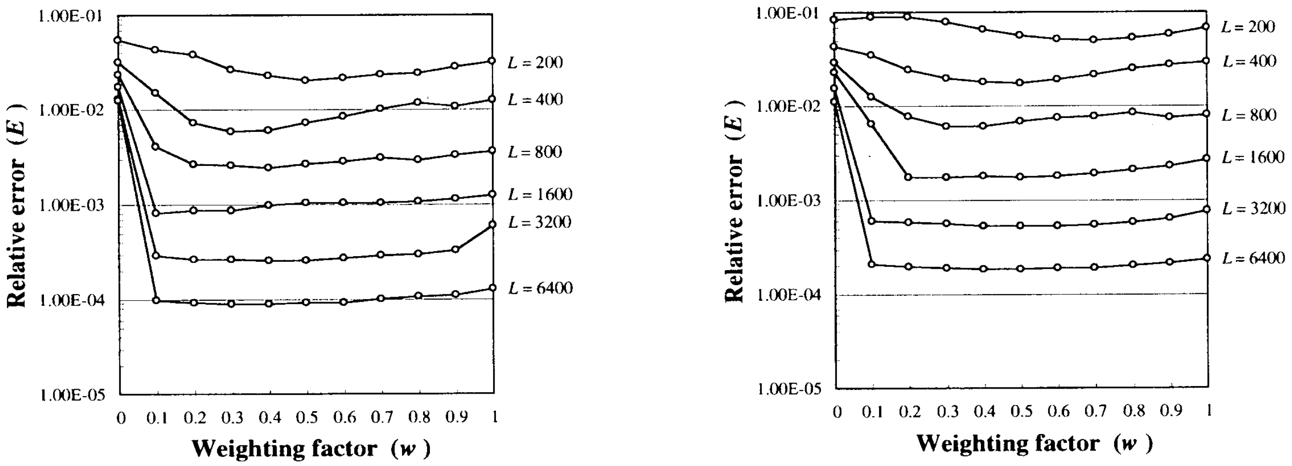
(a) Ikeda map (b) Hénon map

Fig. 2 Phase plots of the Ikeda map and the Hénon map.



(a) Conventional method (b) Proposed method ($w = 0.3$)

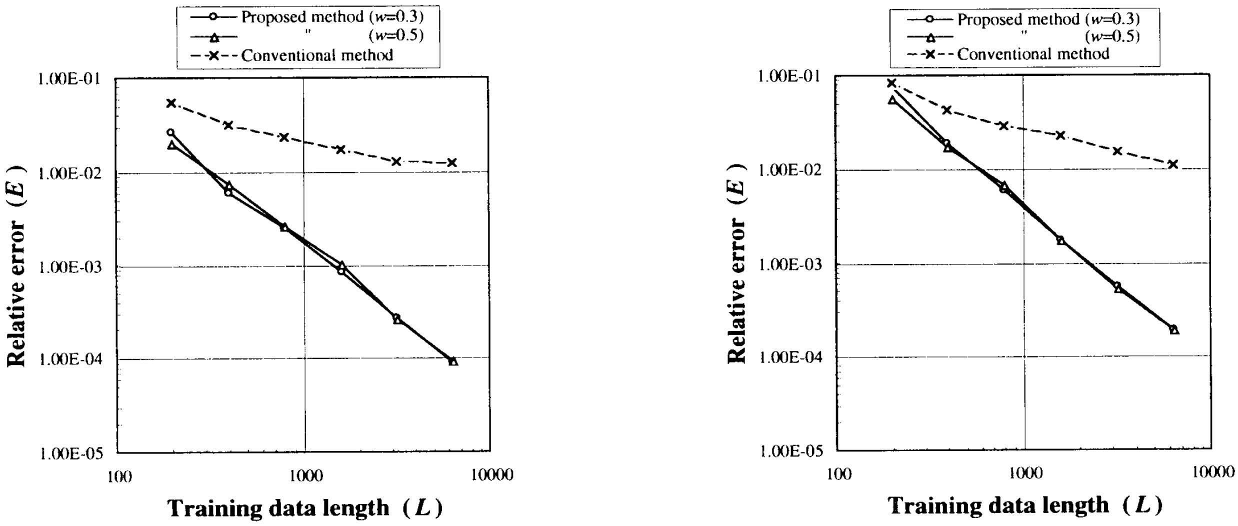
Fig. 3 Predicted value vs. observed value for the Ikeda map.



(a) $n = 8$

(b) $n = 16$

Fig. 4 Relative error as a function of weighting factor for the Ikeda map.



(a) $n = 8$

(b) $n = 16$

Fig. 5 Relative error as a function of training data length for the Ikeda map.

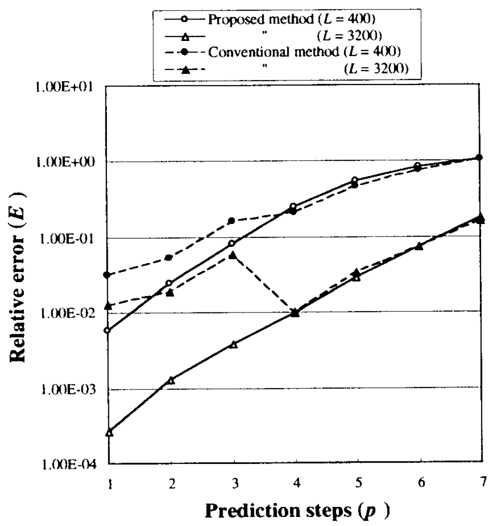


Fig. 6 Relative error as a function of prediction steps for the Ikeda map.

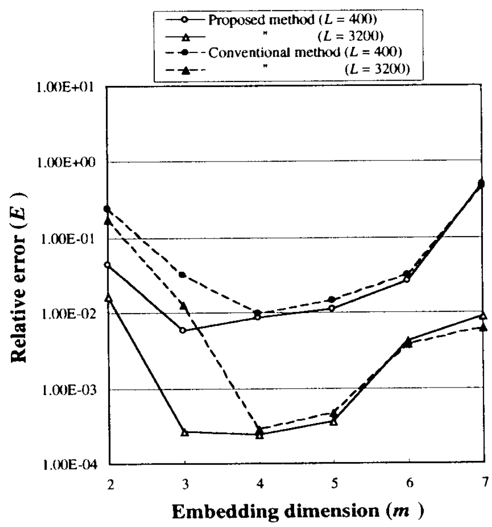
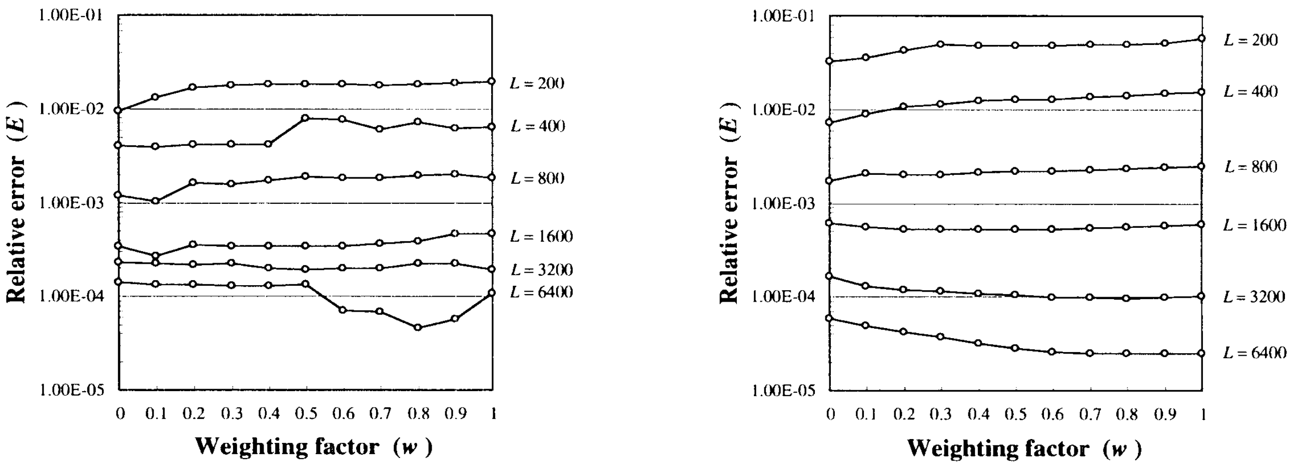
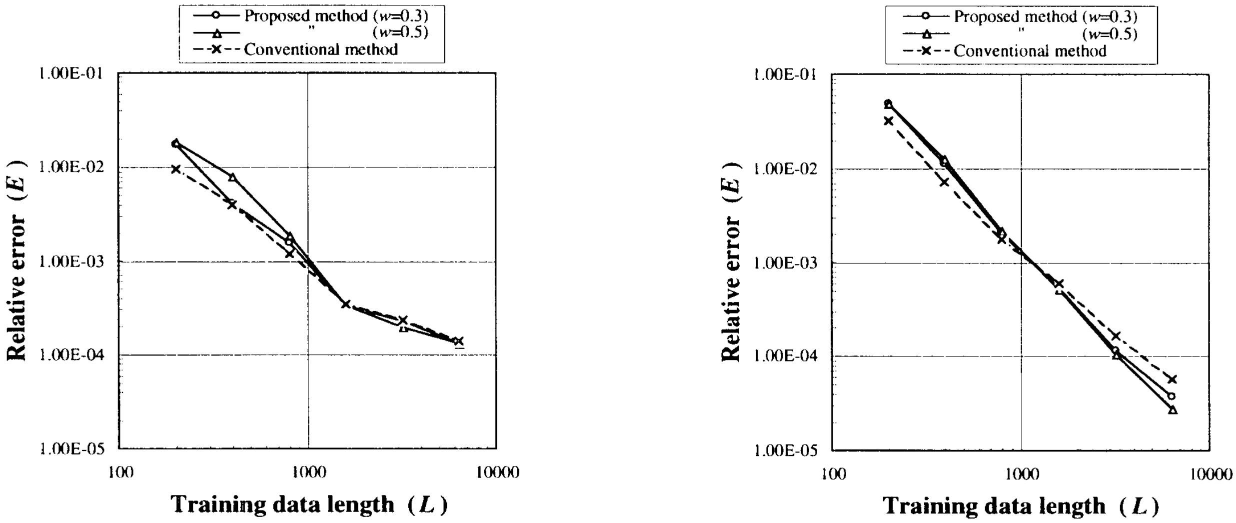


Fig. 7 Relative error as a function of embedding dimension for the Ikeda map.



(a) $n = 8$ (b) $n = 16$

Fig. 8 Relative error as a function of weighting factor for the Hénon map.



(a) $n = 8$ (b) $n = 16$

Fig. 9 Relative error as a function of training data length for the Hénon map.

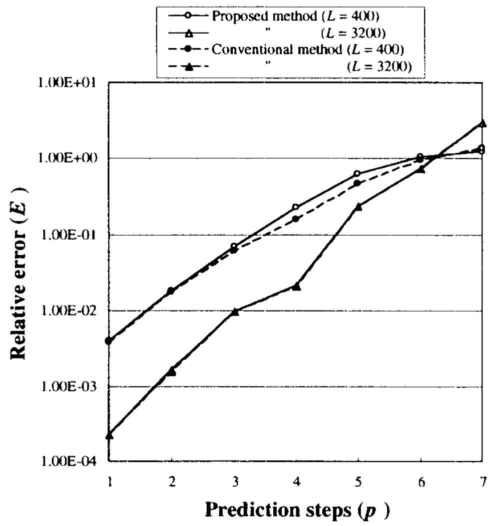


Fig. 10 Relative error as a function of prediction steps for the Hénon map.

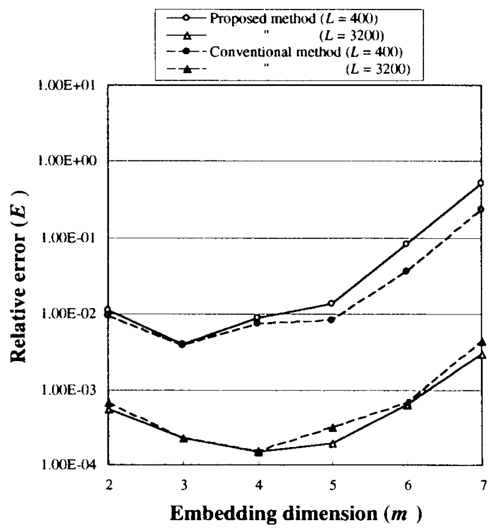


Fig. 11 Relative error as a function of embedding dimension for the Hénon map.

error is sufficiently decreased as L is increased as shown in Fig. 9. The minimum embedding dimension of the Hénon map is 3 because the condition for embedding is $m > 2D_0$ and D_0 of the Hénon map is 1.26. The relative error for $m = 3$ is roughly equivalent to the minimum value of the relative error obtained by changing m. When the proposed method is applied for the Hénon map, the relative error is of the same order as that obtained by the conventional method.

Some trial tests were conducted on continuous time systems such as the Lorenz model [10] and the Rössler model [11]. The results show that the prediction accuracy of the proposed method nearly equals the prediction accuracy of the conventional method. The proposed method is also usable for continuous time systems.

5. Conclusion

This paper proposed an improved method based on a local approximation technique for predicting chaotic time series. To increase the prediction accuracy of the local approximation, a new strategy was employed for selecting the nearest neighbors in a state space. By using chaotic time series generated by the Ikeda map and the Hénon map, numerical prediction experiments were conducted. As a result, the efficacy of the proposed method was confirmed. In future work, the proposed method will be applied to real-world data.

References

[1] J. D. Farmer and J. J. Sidorowich, "Predicting chaotic time series," Phys. Rev. Lett., vol.59, no.8, pp.845-848, Aug. 1987.

[2] G. Sugihara and R. M. May, "Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series," Nature, vol.344, no.19, pp.734-741, April 1990.

[3] M. Casdagli, "Nonlinear prediction of chaotic time series," Physica D, vol.35, pp.335-356, 1989.

[4] A. Lapedes and R. Farber, "Nonlinear signal processing using neural networks: Prediction and system modeling," Los Alamos National Laboratory Report, no.LA-UR-87-2662, 1987.

[5] N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, "Geometry from a time series," Phys. Rev. Lett., vol.45, no.9, pp.712-716, Sept. 1980.

[6] M. Casdagli, S. Eubank, J. D. Farmer, and J. Gibson, "State space reconstruction in the presence of noise," Physica D, vol.51, pp.52-98, 1991.

[7] K. Ikeda, "Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system," Opt. Commun., vol.30, no.2, pp.257-261, 1979.

[8] M. Hénon, "A two-dimensional mapping with a strange attractor," Commun. Math. Phys., vol.50, pp.69-77, 1976.

[9] T. Sauer, J. A. Yorke, and M. Casdagli, "Embedology," J. Stat. Phys., vol.65, no.3,4, pp.579-616, 1991.

[10] E. N. Lorenz, "Deterministic non-periodic flow," J. Atmos. Sci., vol.20, pp.130-141, 1963.

[11] O. E. Rössler, "An equation for continuous chaos," Phys. Lett., vol.57A, pp.397-398, 1976.