State-space prediction model for chaotic time series

A. K. Alparslan, M. Sayar, and A. R. Atilgan*
Polymer Research Center, TUBITAK Advanced Polymeric Materials Research Center and School of Civil Engineering, Boğaziçi University, Bebek 80815, Istanbul, Turkey
(Received 14 July 1997; revised manuscript received 1 April 1998)

A simple method for predicting the continuation of scalar chaotic time series ahead in time is proposed. The false nearest neighbors technique in connection with the time-delayed embedding is employed so as to reconstruct the state space. A local forecasting model based upon the time evolution of the topological neighboring in the reconstructed phase space is suggested. A moving root-mean-square error is utilized in order to monitor the error along the prediction horizon. The model is tested for the convection amplitude of the Lorenz model. The results indicate that for approximately 100 cycles of the training data, the prediction follows the actual continuation very closely about six cycles. The proposed model, like other state-space forecasting models, captures the long-term behavior of the system due to the use of spatial neighbors in the state space. [S1063-651X(98)10707-9]

PACS number(s): 05.45.+b

Measured time series are usually the basis for characterizing a dynamical system. In practical cases, however, it is not possible to observe all relevant dynamical variables pertaining to the system. The most common case is limited to a scalar time evolution of a variable for a finite duration of time. One of the most challenging endeavors is to predict the continuation of the time evolution of the scalar variable monitored. A finite-dimensional linear system produces a signal that can be characterized by a finite number of frequencies. Based upon this fact, either in frequency or in time domain there are methods for time series prediction [1]. For nonlinear processes, however, these methods become inappropriate since a global model cannot be applied to the entire state space where the signal lives [2,3].

Eckmann and Ruelle [4] suggested first the idea of finding the relation between the delay coordinates [5–7] of a point and the points that appear at some time later in the state space. This idea was followed immediately [8–11] and also found an application in the solution of a so-called inverse problem in iterated function systems [12]. A competition was also arranged to test the success of prediction algorithms proposed until 1993 [13]. Among those registered for the competition, two methods prove to be the most successful [14,15]. One uses a connectionist neural network [14] and the other utilizes the delay coordinate embedding based methodology [15] based upon the Eckmann-Ruelle proposition. More recently, wavelets [16] and genetic algorithms [17] have also been suggested for nonlinear predictions. And, a methodology based upon a nonlinear prediction technique [18] has been proposed to probe dynamical coupling [19] among nonlinear systems. In what follows we propose a state-space prediction model whose success will be shown to be comparable with those of Sauer's [15] and Wan's [14], although present methodology is simpler and easier to implement.

The prediction problem may be formulated as follows: Suppose that the time evolution of the system behavior is reconstructed in the state space [7,20,21]. A total of n different points on the attractor located in the state space are known. These points are $P(1),P(2),\ldots,P(n)$. With respect to a fixed reference frame, a point is represented by m numbers, m being the dimension of the state space. There are two questions: (i) Can one determine the point $P(n+1)$, and (ii) if (i) is achieved, how far can the consecutive points $P(n+1),P(n+2),\ldots,P(n+n^*)$ be found, that is, what is the maximum value for n^*?

We first postulate that for any point $P(n)$, the succeeding point $P(n+1)$ can be found using the preceding points of $P(n+1)$ and the time evolution information of the spatial neighbors of $P(n)$ that are located within a certain cutoff distance r (Fig. 1). Note that $P(n)$ is a column vector with m rows. The preceding points of $P(n+1)$ may be expressed in a (dm) dimensional vector $P_-(n+1)$ as

$$P_-(n+1)=[P^T(n)P^T(n-1)P^T(n-2)\cdots P^T(n-d+1)], \tag{1}$$

where d is referred to as the model dimension. According to our conjecture, the following equation may be written:

$$P(n+1)=CP_-(n+1), \tag{2}$$

where C is an $[m\times(dm)]$ coefficient matrix that contains an average time evolution information of the spatial neighbors of $P(n)$. In Fig. 1, $P(k_i)$ denotes these spatial neighbors. Here i can take values from 1 to p, which is the number of neighbors that fall into the sphere whose center is at $P(n)$ and radius is the cutoff distance r. The entries of the coefficient matrix C, therefore, depend on the location of $P(n)$ and its topological neighborhood. The C matrix can be obtained from the following equation:

$$C=BA^{-1}, \tag{3}$$

where A is a $[(dm)\times p]$ matrix whose ith column consists of $P_-(k_i+1)$ and B is an $(m\times p)$ vector whose ith column is made of $P(k_i+1)$. The generalized inverse of A can be found using the singular value decomposition technique [22].

*Author to whom correspondence should be addressed.

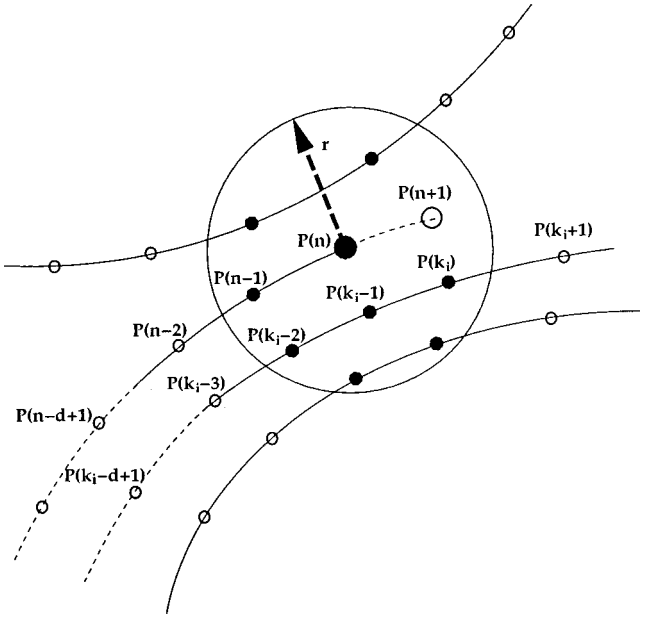

FIG. 1. A schematic view illustrating the collection of spatial neighbors and their time evolution in the state space. Here r denotes the radius of the cutoff sphere, d the number of consecutive points preceding in the time history of a trajectory. $P(n)$ represents the last point in the training data and $P(k_i)$ (any of the filled circles) is a spatial neighbor within the cutoff sphere of the last point.

The same scheme may be followed to calculate $P(n+2)$, which follows the last point, which is now $P(n+1)$.

We need to define an error between the predicted and the actual trajectory of the system so as to monitor the accuracy of the prediction model within the range of locality assumption. Root-mean-square (rms) error, which calculates the root mean square of the differences between the predicted and the actual data at all points, gives a single number about the error. Instead, we need the evolution of the error along the prediction horizon [15]. Therefore, plotting the rms errors of neighboring predictions within a window length of l_w of each point would be more meaningful. This is referred to as the moving rms error. The error of this nature may be expressed as

$$e(i)=\left(\frac{\sum_{k=i-l_w/2+1}^{k=i+l_w/2}[x(k)-\overline{x}(k)]^2}{l_w}\right)^{1/2}, \tag{4}$$

where $i=l_w/2,...,n^*-l_w/2$. Here n^* is the total number points predicted, and $x(k)$ and $\overline{x}(k)$ are the actual and predicted data, respectively. We now would like to test the proposed methodology for a benchmark case.

Complex signals were commonly assumed to be the output of a complicated system with a large number of active degrees of freedom. However, realization of nonlinear systems with a relatively small number of degrees of freedom, while deterministic in principle, can create output signals that look complex and mimic stochastic signals, such as the Lorenz model [23,24]. This model describes the Rayleigh-Bénard convection arising from the two-dimensional Navier-Stokes equation, which is formulated for a fluid slab of finite thickness subjected to gravity loading, heated from below, and between the top cold and the bottom hot surface the temperature is held constant. The partial differential equations were transformed to a set of three ordinary differential equations [23]. Later Lorenz described the way he derived the model both physically and mathematically [24]. The three ordinary differential equations are

$$\begin{aligned}\dot{x}&=-\sigma x+\sigma y,\\ \dot{y}&=-xz+rx-y,\\ \dot{z}&=xy-bz,\end{aligned} \tag{5}$$

where x is the amplitude of the convection motion, y is the temperature difference between the ascending and descending currents, z is the distortion of the vertical temperature profile from linearity, and σ, r, and b are dimensionless parameters. The Lorenz system has been studied extensively and Eqs. (5) was recently shown to be isomorphic to a Doppler broadened optically pumped laser model [25]. The parameters in Eqs. (5) are most commonly selected to be $\sigma=10$, $r=28$, and $b=8/3$ for a rich dynamical behavior [26].

For generating the data, the Lorenz equations displayed in Eqs. (5) are integrated for 3×10^4 time steps with a step size of 5×10^{-3}. The Bulirsch-Stoer method [27] is used to obtain the time evolution of the parameters x, y, and z whose starting values are $x(0)=y(0)=z(0)=10$. The time-delayed embedding together with the false nearest neighbor technique [21] is utilized to reconstruct the state space with $m=3$ using the convection amplitude x. The objective is to capture the actual time continuation of the amplitude x.

Five different time series with different end point indices, 16 000, 18 000, 20 000, 22 000, and 24 000 are prepared as training data for prediction. Continuation of the amplitude data is predicted for 10^3 time steps for each initial condition. Results are then compared with the actual continuation, which is calculated by integrating the Lorenz equations further in time with the Bulirsch-Stoer method. The parameter set of the algorithm is identical for all five runs. The model dimension d, which represents the number of consecutive points in a trajectory whose linear combinations determine the succeeding point, is set to 3. The radius of the sphere r, which controls the number of spatial neighbors collected from the adjacent trajectories is set to 0.2. Comparison studies are placed in Fig. 2. The moving rms error obtained by averaging over five sets is displayed in Fig. 3.

The results indicate that predictions with different starting points contribute to the accumulation of error differently. However, the averaged error demonstrates that the error starts growing considerably after around 600 time steps, then fluctuates within an interval about 200 steps. The later steps however, diverge significantly from the actual continuation. It is observed that the predictions follow the actual trajectory for a certain period of time. The predicted results then start diverging from the actual one. This is observed whenever the trajectory approaches zero where the number of spatial neighbors increases enormously. Therein the sensitivity to the initial conditions is so critical that the error accumulated by the prediction algorithm also grows exponentially. The predicted trajectory gets away slightly from the true continuation first, and as it comes near zero once more, this slight

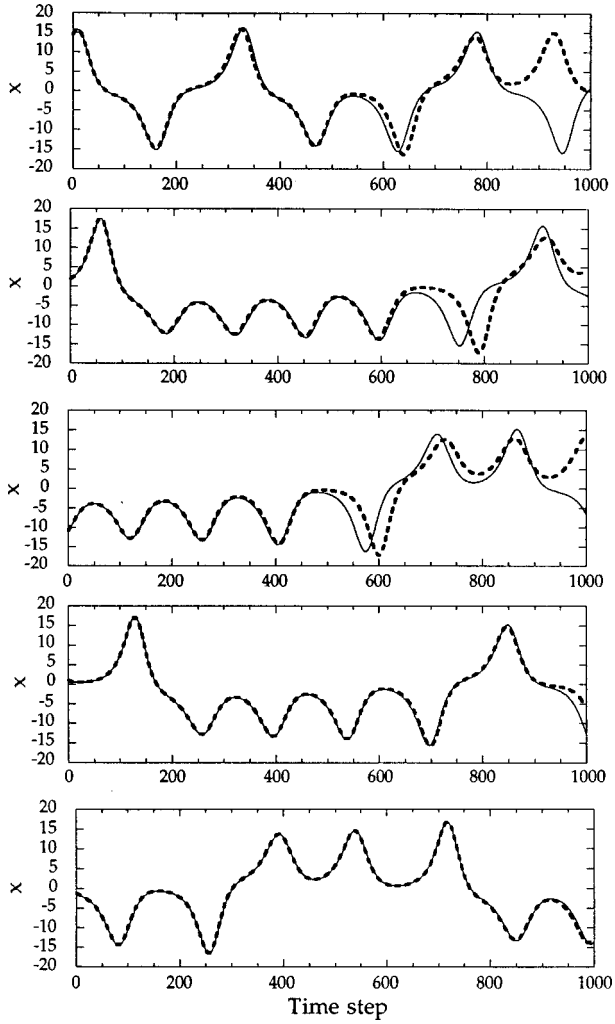

FIG. 2. Predicted (dashed lines) and the actual continuation (solid lines) of the time evolution for the x coordinate of the Lorenz model for five different initial conditions. These figures illustrate the dependence of prediction performance on the starting points.

difference results in a shift to a completely different region of the attractor. However, the predicted trajectory catches the actual continuation with a time lag, as the method utilizes the state space neighbors for predictions. This fact is an advantage of the state-space prediction methods over neural networks [13] and would serve a good starting point to create a model for long-term predictions.

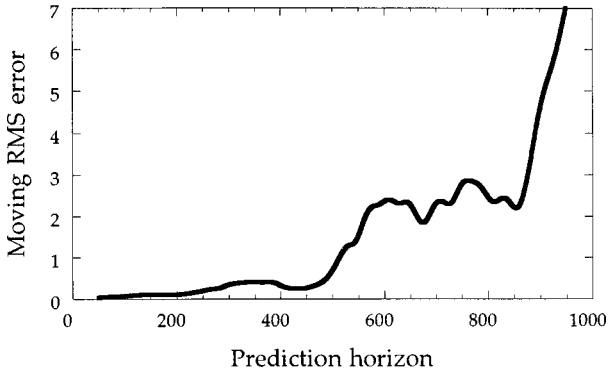

FIG. 3. The moving rms error of the prediction for the x coordinate of the Lorenz model. The error is obtained by averaging over five different predictions exhibited in Fig. 2.

One can enlarge the cutoff radius r of the sphere in order to include the far neighbors. This may invalidate the local linear model assumption. Using a sphere with smaller radius, on the other hand, may increase the effect of noise [15]. Therefore, one needs to find an optimum radius in the spectrum.

The simple prediction algorithm proposed herein uses both spatial and temporal information about the trajectories in the state space. The relative weight of these two components depends on the parameters selected. The model dimension d and the cutoff radius r are the free parameters that determine the weight. For example, decreasing the model dimension and increasing the cutoff radius, one can formulate an algorithm in which spatial information is more dominant.

An additional measure to scale the relative contributions of the neighbors with respect to their locations in the cutoff sphere may be considered [15]. In the algorithm proposed here, we do not limit the number of spatial neighbors coming from a trajectory. Thus, the nearest trajectories can contribute to the neighboring set with more than one point. This results in an increased weighting of the contribution coming from the nearest trajectories.

The authors gratefully acknowledge partial support from Bogazici University Research Funds, Project No. 97A0401, and from The Boeing Company, Project No. PO FR-513570-07LLN.

[1] P. Brockwell and R. Davis, *Time Series: Theory and Methods* (Springer-Verlag, New York, 1987).

[2] H. Tong, *Nonlinear Time Series* (Clarendon Press, Oxford, 1990).

[3] H. Tong, *A Personal Overview of Nonlinear Time Series Analysis from a Chaos Perspective in Nonlinear Dynamics and Time Series: Building a Bridge between the Natural and Statistical Sciences*, Fields Institute Communications Vol. 11, edited by C. D. Cutler and D. T. Kaplan (American Mathematical Society, Providence, RI, 1996).

[4] J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985).

[5] N. Packard *et al.*, Phys. Rev. Lett. **45**, 712 (1980).

[6] F. Takens, in *Dynamical Systems and Turbulence, Warwick, 1980*, edited by D. Rand and L.-S. Young (Springer, Berlin, 1981), p. 366.

[7] T. Sauer, J. A. Yorke, and M. Casdagli, J. Stat. Phys. **65**, 579 (1991).

[8] J. D. Farmer and J. J. Sidorowich, Phys. Rev. Lett. **59**, 845 (1987).

[9] J. D. Farmer and J. Sidorowich, in *Evolution, Learning and Cognition*, edited by Y. C. Lee (World Scientific, Singapore, 1988).

[10] M. Casdagli, Physica D **35**, 335 (1989).

[11] M. Giona, F. Lentini, and V. Cimagalli, Phys. Rev. A **44**, 3496 (1991).
[12] G. Mantica and B. G. Giraud, Chaos **2**, 225 (1992).
[13] N. A. Gershenfeld and A. S. Weigend, *The Future of Time Series: Learning and Understanding, in Time Series Prediction: Forecasting the Future and Understanding the Past*, edited by A. S. Weigend and N. A. Gershenfeld (Addison-Wesley, Redwood City, CA, 1993), p. 1.
[14] E. A. Wan, in *Time Series Prediction: Forecasting the Future and Understanding the Past*, edited by A. S. Weigend and N. A. Gershenfeld (Addison-Wesley, Redwood City, CA, 1993), p. 195.
[15] T. Sauer, in *Time Series Prediction: Forecasting the Future and Understanding the Past*, edited by A. S. Weigend and N. A. Gershenfeld (Addison-Wesley, Redwood City, CA, 1993), p. 241.
[16] U. Parlitz and G. Mayer-Kress, Phys. Rev. E **51**, 2709 (1995).
[17] G. G. Szpiro, Phys. Rev. E **55**, 2557 (1997).
[18] S. J. Schiff *et al.*, Phys. Rev. E **54**, 6708 (1996).
[19] L. M. Pecora, T. L. Carrol, and J. F. Heagy, Phys. Rev. E **52**, 3420 (1995).
[20] H. D. I. Abarbanel *et al.*, Rev. Mod. Phys. **65**, 1331 (1993).
[21] H. D. I. Abarbanel, in *Nonlinear Dynamics and Time Series: Building a Bridge between the Natural and Statistical Sciences*, Fields Institute Communications Vol. 11, edited by C. D. Cutler and D. T. Kaplan (American Mathematical Society, Providence, RI, 1996).
[22] G. H. Golub and C. F. Van Loan, *Matrix Computations* (Johns Hopkins University Press, 1989).
[23] E. N. Lorenz, J. Atmos. Sci. **20**, 130 (1963).
[24] E. N. Lorenz, Tellus, Ser. A **36**, 98 (1984).
[25] R. Gilmore *et al.*, Phys. Rev. E **55**, 2479 (1997).
[26] A. J. Lichtenberg and M. A. Lieberman, *Regular and Stochastic Motion* (Springer, New York, 1983).
[27] W. H. Press *et al.*, *Numerical Recipes* (Cambridge University Press, Cambridge, England, 1992).