PHYSICAL REVIEW A VOLUME 38, NUMBER 3 AUGUST 1, 1988

Noise reduction in dynamical systems

Eric J. Kostelich
Center for Nonlinear Dynamics and Department of Physics, University of Texas, Austin, Texas 78712

James A. Yorke
Institute for Physical Science and Technology and Department of Mathematics, University of Maryland, College Park, Maryland 20742
(Received 29 October 1987)

A method is described for reducing noise levels in certain experimental time series. An attractor is reconstructed from the data using the time-delay embedding method. The method produces a new, slightly altered time series which is more consistent with the dynamics on the corresponding phase-space attractor. Numerical experiments with the two-dimensional Ikeda laser map and power spectra from weakly turbulent Couette-Taylor flow suggest that the method can reduce noise levels up to a factor of 10.

The ability to extract information from time-varying signals is limited by the presence of noise. Methods of noise reduction are a subject of widespread interest in communication,[1] physical systems,[2] and experimental measurements.[3] Recent experiments to study the transition to turbulence in systems far from equilibrium, like those by Fenstermacher *et al.*,[4] Behringer and Ahlers,[5] and Libchaber *et al.*,[6] succeeded largely because of instrumentation that enabled them to quantify and reduce the noise.

In recent years, traditional methods of time series analysis like power spectra have been augmented by new methods. In many cases, the time series can be viewed as a dynamical system with a low-dimensional attractor that can be reconstructed from the time series using time delays.[7] Because the dynamics of the phase-space attractor are not localized in a time or frequency domain, traditional noise-reduction methods like Wiener[8] and Kalman[9] filters are not applicable. In this paper we describe a noise-reduction procedure that works by taking many nearby points in phase space (corresponding to widely varying times in the original signal) to find a local approximation of the dynamics. These approximations can be used collectively to produce a new time series whose dynamics are more consistent with those on the phase-space attractor. We demonstrate the efficacy of the method using chaotic attractors obtained from the Ikeda laser map[10] and a Couette-Taylor fluid flow experiment.[11]

The discrete sampling of the original signal means that the points on the reconstructed attractor can be treated as iterates of a nonlinear map f whose exact form is unknown. However, we assume that f is nearly linear in a small neighborhood about each attractor point x_i and write

$$x_{i+1}=f(x_i)\approx A_i x_i+b_i \equiv L(x_i)$$

for some matrix A_i and vector b_i. (The matrix A_i is the Jacobian of f at x_i.) This can be done with least-squares procedures similar to those described in Ref. 12. Let $\{x_j\}_{j=1}^n$ be a collection of points in a small ball around the ith reference point, and let $y_i=f(x_j)$ denote the observed image of x_j. The kth row $a_i^{(k)}$ of A_i and kth component $b_i^{(k)}$ of b_i are given by the least-squares solution of the equation

$$y_j^{(k)}=b_i^{(k)}+(a_i^{(k)}\,|\,x_j)\ , \tag{1}$$

where $y_j^{(k)}$ is the kth component of y_j and $(\ |\ )$ denotes the dot product. (Farmer and Sidorowich[13] have generated similar approximations for the different purpose of forecasting chaotic time series.)

We remark that Eq. (1) can be ill-conditioned, for example, when the unstable manifold at x_i is nearly one dimensional and A_i is 2×2. We detect this situation by computing the singular values and right singular vectors[14] of the matrix X whose jth row is x_j to find the condition number of X, which is defined as the ratio of the largest to the smallest singular value. When the condition number is sufficiently large, we solve Eq. (1) using the components of x_j contained in the subspace spanned by the singular vectors corresponding to the largest singular values. (For instance, we find a one-dimensional linear approximation of f wherever the points x_j fall nearly along a single line.) Moreover, because error exists both in the points x_j and their observed images y_j, a modified least-squares procedure as described in Ref. 15 often gives better estimates of A_i and b_i.

In the second stage of the method, we use the linear (more precisely, linear + constant) maps L to correct errors in the observed trajectories as follows. Given a "window" of consecutive points $\{x_i, x_{i+1}, \ldots, x_{i+p}\}$ on the observed trajectory, we find the collection of points $\{\hat{x}_i, \hat{x}_{i+1}, \ldots, \hat{x}_{i+p}\}$ closest to the observed ones which also best satisfy the corresponding linear maps. More precisely, the new trajectory $\{\hat{x}_{i+k}\}_{k=0}^{p}$ minimizes the sum of squares

 ©1988 The American Physical Society

$$\sum_{k=0}^{p} \|\hat{x}_{i+k} - L(\hat{x}_{i+k-1})\|^2 + \|\hat{x}_{i+k} - x_{i+k}\|^2 + \|\hat{x}_{i+k+1} - L(\hat{x}_{i+k})\|^2 \quad (2)$$

(terms with subscripts outside $[i,i+p]$ are omitted). This procedure can be iterated by replacing the original trajectory $\{x_i\}$ with the most recent least-squares trajectory $\{\hat{x}_i\}$, then finding a new solution to Eq. (2).[16] Moreover, the windows can overlap; for instance, the second window can begin in the middle of the first.

We have conducted numerical experiments using the attractor produced by the Ikeda map $f(z)=\rho + c_2 z \exp\{i[c_1 - c_3/(1+|z|^2)]\}$, which models the dynamics of a bistable laser cavity.[10] We consider the attractor for the mapping $z_{j+1}=f(z_j)$, where $\rho=1$, $c_1=0.4$, $c_2=0.9$, $c_3=6$. (The complex number z_j is identified with the 2-vector x_j.) Numerical evidence suggests that initial conditions in $[0.5,1.8]\times[-2,1]$ are in the basin of a chaotic attractor whose numerically calculated Lyapunov exponents are 0.7296, −1.034 (logarithms base 2) and whose Lyapunov dimension is 1.71.

We measure the noise level in terms of the *pointwise error* $e_j = \|x_{j+1} - L(x_j)\|$, i.e., the distance between the observed image and the predicted one [using the linear maps from Eq. (1)]. The *mean error* is $E=(\sum_j e_j^2/N)^{1/2}$, the root-mean-square value of the pointwise error over all N points on the attractor. We define the *noise reduction* $R = 1 - E_{\text{fitted}}/E_{\text{noisy}}$, where the mean errors are computed for the adjusted and original noisy attractor, respectively. (R is a measure of the self-consistency of the time series, assuming that the linear maps are accurate approximations of the true dynamics.)

The numerical experiments on the Ikeda attractor use 65 536 iterates, to which 0.1% uniformly distributed random noise is added. The noise is independent of the dynamics, i.e., the input to the computer program is the series $\{z_j + \eta_j : \eta_j \text{ random}\}$, for which $E_{\text{noisy}} = 7.588\times 10^{-4}$. The linear maps L are computed using at least 50 points about each attractor point. Points are collected until the condition number of Eq. (1) is less than ten.[17] Trajectory adjustment is done in windows of 24 points, and the windows overlap by two points. After noise reduction, $E_{\text{fitted}} = 1.178\times 10^{-4}$, so that the total noise reduction R is 84%. When 1% noise is added, we find $R=83\%$.

We have performed similar numerical trials with the Hénon attractor,[18] for which the $(j+1)$st time series value is given by

$$x_{j+1} = f(x_j, x_{j-1}) = 1 - 1.4x_j^2 + 0.3x_{j-1} .$$

In this case the pointwise error can be measured exactly by replacing L with f (the mean error E then becomes a "correctness index"). When 1% noise is added to the in-

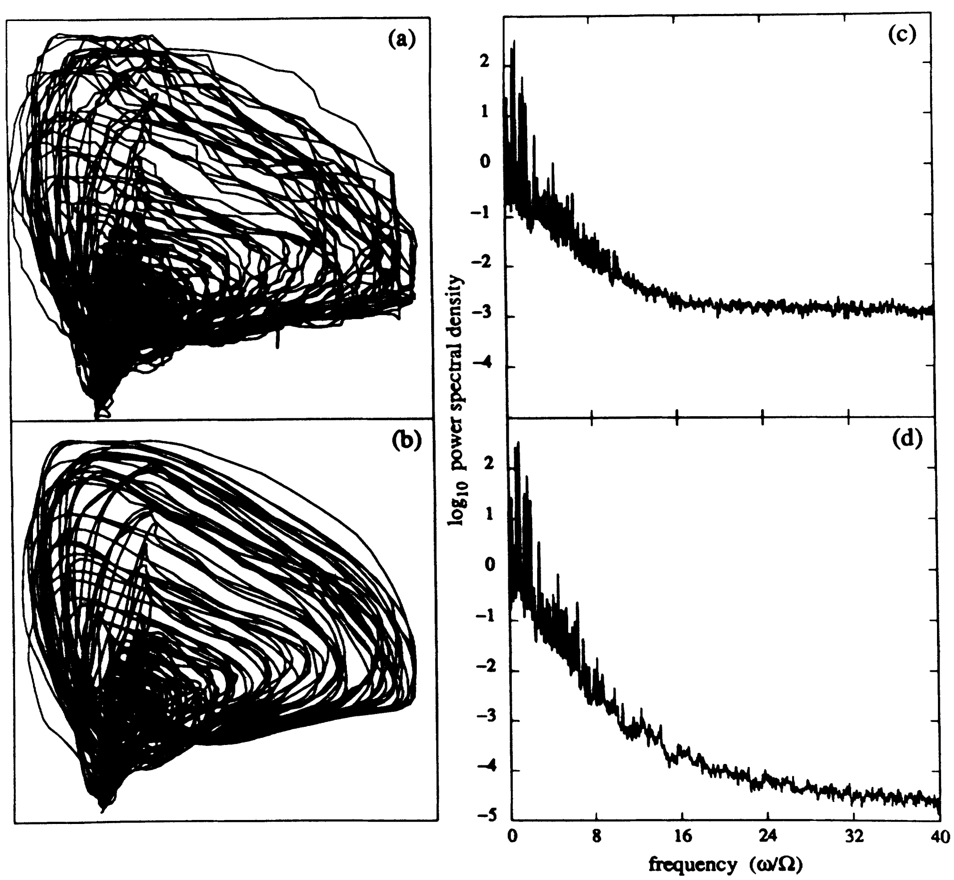

FIG. 1. Chaotic attractors from the Couette-Taylor fluid flow experiment described in Ref. 11 at $R/R_c = 12.9$. (a) Raw data. (b) Attrator after the noise reduction procedure described in the text. (c) and (d) Power spectra corresponding to (a) and (b), respectively. The units of frequency and power spectral density are as described in Ref. 11.

put as described above, the noise reduction (measured with the actual map) is 79%.[19] In addition, noise levels can be reduced almost as much in cases where the noise is added to the dynamics, i.e., where the input is of the form

$$\{x_{j+1}:x_{j+1}=f(x_j+\eta_j,x_{j-1}+\eta_{j-1}),\eta_j,\eta_{j-1}\text{ random}\}\ .$$

Next we consider the application of the method to data from the Couette-Taylor fluid flow experiment described in Ref. 11. Figure 1(a) shows a two-dimensional phase portrait of the raw time series at a Reynolds number $R/R_c=12.9$, which corresponds to weakly chaotic flow.[11] The corresponding phase portrait from the filtered time series[20] is shown in Fig. 1(b). The noise reduction, using the above criterion with the linear maps, is 63%.

Figure 1(c) and 1(d) show the power spectra for the corresponding time series. We emphasize that the dynamical information used to adjust the trajectories (viz., the motion of ensembles of points which are close together in phase space) corresponds to portions of the original signal that are widely and irregularly spaced in time. One question therefore is whether reducing the high-frequency noise corresponds to discovering the true dynamics which have been masked by noise. We believe that the answer is yes, based on those cases where there is an underlying low-dimensional dynamical system. However, in chaotic process some high-frequency components remain, because they are appropriate to the dynamics.

The method is particularly useful in calculations of dynamical quantities such as metric entropy and attractor dimension from experimental data. As an example, we consider the correlation dimension.[21] Let $C(x_i,\epsilon)$ denote the fraction of points on the attractor that fall within a distance ϵ of a randomly chosen (with respect to the natural measure) reference point x_i. Let $C(\epsilon)$ be the average values of $C(x_i,\epsilon)$ over the reference points x_i. Then $C(\epsilon)\sim\epsilon^d$ for small ϵ, where d is the correlation dimension.[21]

The dimension calculation illustrated in Fig. 2 is for the Ikeda attractor described above. The value of $C(\epsilon)$ is estimated from 1000 reference points using 48 values of ϵ, equally spaced on a logarithmic scale from 2^{-10} to 2^{-2} (only the range $2^{-6}\le\epsilon\le2^{-3}$ is shown). Distances are normalized so that the total attractor extent is 1. The dimension, which is estimated as the derivative of $\log C(\epsilon)$ with respect to $\log\epsilon$, is taken as the slope of the regression line through six consecutive $(\log\epsilon,\log C(\epsilon))$ pairs. Although the noise level in the input is only 1%, noise inflates the dimension estimate even at ball sizes which are 3% of the attractor extent (top curve). However, the dimension estimate for the fitted attractor (middle thin curve) compares favorably to that obtained from the noiseless attractor (bold line).

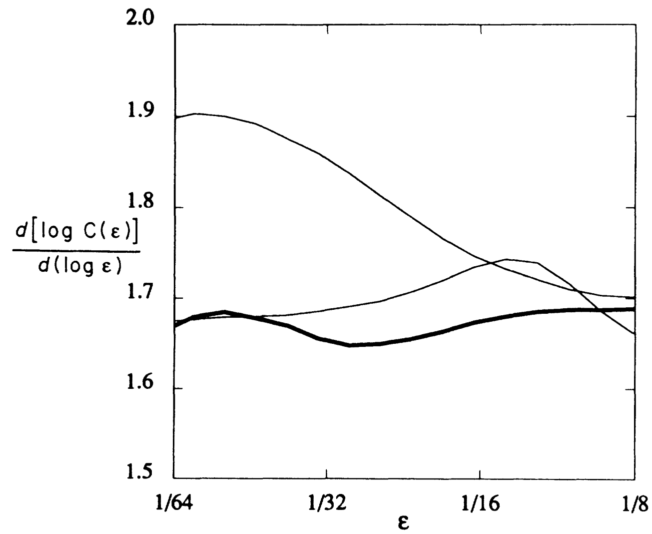

FIG. 2. Grassberger-Procaccia correlation dimension d for the Ikeda attractor, using a data set with 1% uniformly distributed random noise (top curve), the same data set after the noise-reduction procedure (middle curve), and the original noiseless attractor (bold curve). The Lyapunov dimension of the attractor is 1.71.

Since accurate linear approximations are essential for the success of the method, there must be an ample number of points in a small neighborhood about each point on the attractor. Thus, the data requirements depend on the dimension of the attractor. This method is best suited to situations where large amounts of data can be collected but the measurement precision is limited. This method promises to be of considerable value in the analysis of experimental data when the time series can be viewed as arising from a dynamical system with a low-dimensional attractor.

We thank Anke Brandstäter for providing data from her Couette-Taylor experiment. E. K. thanks Bill Jefferys, Andrew Fraser, Randy Tagg, and Harry Swinney for helpful discussions. This research is supported by the U.S. Department of Energy Office of Basic Energy Sciences, U.S. Air Force Office of Scientific Research, U.S. Defense Advanced Research Projects Agency–Applied and Computational Mathematics Program.

[1]For instance, see R. G. Gallager, *Information Theory and Reliable Communication* (Wiley, New York 1968).

[2]For instance, see *Sixth International Conference on Noise in Physical Systems,* edited by P.H.E. Meijer, R. D. Mountain, and R. J. Soulen (U.S. GPO, Washington, D.C., 1981), and similar volumes.

[3]For instance, see D. R. Brillinger, *Time Series: Data Analysis and Theory,* (Holden-Day, San Francisco, 1980); M. Priestly, *Spectral Analysis and Time Series* (Academic, London, 1981), two volumes.

[4]P. R. Fenstermacher, H. L. Swinney, and J. P. Gollub, J. Fluid Mech. **94**, 103 (1979).

[5]R. P. Behringer and G. Ahlers, J. Fluid Mech. **125**, 219 (1982); G. Ahlers and R. P. Behringer, Phys. Rev. Lett. **40**, 712 (1978).

[6]A. Libchaber, S. Fauve, and C. Laroche, Physica D **7**, 73 (1983).

[7]Examples of the use of this method for the analysis of experimental data are given in *Dimensions and Entropies in Chaotic Systems,* edited by G. Mayer-Kress (Springer-Verlag, Berlin, 1986).

[8]For example, see L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, N.J., 1975).

[9]For example, see A. E. Bryson and Y. C. Ho, *Applied Optimal Control* (Ginn, Waltham, Mass., 1969).

[10]S. M. Hammel, C. K. R. T. Jones, and J. V. Moloney, J. Opt. Soc. Am. B **2**, 552 (1985).

[11]A. Brandstäter and H. L. Swinney, Phys. Rev. A **35**, 2207 (1987).

[12]J.-P. Eckmann and D. Ruelle, Rev. Mod. Phys. **54**, 617 (1985); J.-P. Eckmann, S. O. Kamphorst, D. Ruelle, and S. Ciliberto, Phys. Rev. A **34**, 4971 (1986); M. Sano and Y. Sawada, Phys. Rev. Lett. **55**, 1082 (1985).

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

[14]J. J. Dongarra, C. B. Moler, J. R. Bunch, and G. W. Stewart, *LINPACK User's Guide* (Society of Industrial and Applied Mathematics, Philadelphia, 1979).

[15]W. H. Jefferys, Astron. J. **85**, 177 (1980); *ibid.* **86**, 149 (1981); M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics* (Griffin, London, 1961), Vol. 2, p. 375.

[16]Because the input is a scalar time series, we constrain the trajectory adjustment to yield a scalar time as output. For instance, if $x_1=(\xi_i,\xi_2)$ consists of the first two values in the time series, then the output begins with values $\hat{\xi}_1,\hat{\xi}_2,\hat{\xi}_3$ such that $\hat{x}_1=(\hat{\xi}_1,\hat{\xi}_2),\hat{x}_2=(\hat{\xi}_2,\hat{\xi}_3)$.

[17]To save CPU time, a maximum of 200 points is used to compute the linear maps. Despite these constraints, it is never necessary to compute a one-dimensional map approximation for this attractor as described in the text.

[18]M. Hénon, Commun. Math. Phys. **50**, 69 (1976).

[19]When the program is run on noiseless input, the mean error in the output is 0.025% of the attractor extent, which suggests that errors arising from small nonlinearities are negligible when the input contains enough points.

[20]In the noise reduction procedure, the attractor is reconstructed in four dimensions from a time series containing 32 768 values. Linear maps are computed using at least 50 points in each ball. Trajectories are fitted using windows of 24 attractor points which overlap by 6 points.

[21]P. Grassberger and I. Procaccia, Physica D **9**, 189 (1983).