

Nonlinear dynamics, delay times, and embedding windows

H.S. Kim^{1,a}, R. Eykholt^{b,*}, J.D. Salas^c

^a Department of Civil Engineering, Colorado State University, Fort Collins, CO 80523, USA

^b Department of Physics, Colorado State University, Fort Collins, CO 80523, USA

^c Hydrologic Science and Engineering Program, Department of Civil Engineering, Colorado State University, Fort Collins, CO 80523, USA

Received 11 September 1996; received in revised form 10 January 1998; accepted 12 August 1998

Communicated by A.M. Albano

Abstract

In order to construct an embedding of a nonlinear time series, one must choose an appropriate delay time τ_d . Often, τ_d is estimated using the autocorrelation function; however, this does not treat the nonlinearity appropriately, and it may yield an incorrect value for τ_d . On the other hand, the correct value of τ_d can be found from the mutual information, but this process is rather cumbersome computationally. Here, we suggest a simpler method for estimating τ_d using the correlation integral. We call this the C–C method, and we test it on several nonlinear time series, obtaining estimates of τ_d in agreement with those obtained using the mutual information. Furthermore, some researchers have suggested that one should not choose a fixed delay time τ_d , independent of the embedding dimension m , but, rather, one should choose an appropriate value for the delay time window $\tau_w = (m - 1)\tau$, which is the total time spanned by the components of each embedded point. Unfortunately, τ_w cannot be estimated using the autocorrelation function or the mutual information, and no standard procedure for estimating τ_w has emerged. However, we show that the C–C method can also be used to estimate τ_w . Basically τ_w is the optimal time for independence of the data, while τ_d is the first locally optimal time. As tests, we apply the C–C method to the Lorenz system, a three-dimensional irrational torus, the Rossler system, and the Rabinovich–Fabrikant system. We also demonstrate the robustness of this method to the presence of noise. ©1999 Elsevier Science B.V. All rights reserved.

Keywords: Delay time; Correlation integral; Embedding; Time series

PACS: 05.45.+b; 47.52.+j

1. Introduction

Analysis of chaotic time series is common in many fields of science and engineering, and the method of delays has become popular for attractor reconstruction from scalar time series. From the attractor dynamics, one can estimate the correlation dimension and other quantities to see whether the scalar time series is chaotic or stochastic. Therefore, attractor reconstruction is the first stage in chaotic time series analyses. Since the choice of the delay

* Corresponding author Tel.: +1-970-491-7366; fax: +1-970-491-7947; e-mail: eykholt@lamar.colostate.edu

¹ Present Address: Department of Construction Engineering, Sun Moon University, Korea

time τ_d for attractor reconstruction using the method of delays has not been fully developed, many researchers use the autocorrelation function, which is computationally convenient and does not require large data sets. However, it has been pointed out [1] that the autocorrelation function is not appropriate for nonlinear systems, and, instead, τ_d should be chosen as the first local minimum of the mutual information. Unfortunately, this approach is cumbersome computationally and requires large data sets [2].

According to Packard et al. [3] and Takens [4], the method of delays can be used to embed a scalar time series $\{x_i\}, i = 1, 2, \dots$, into an m -dimensional space as follows:

$$\mathbf{x}_i = (x_i, x_{i+1}, \dots, x_{i+(m-1)t}), \quad \mathbf{x}_i \in \mathbb{R}^m, \quad (1)$$

where t is the index lag. If the sampling time is τ_s , the delay time is $\tau_d = t\tau_s$. Takens' theorem assumes that we have an infinite noise-free data set, in which case, we can choose the delay time almost arbitrarily. However, since real data sets are finite and noisy, the choice of the delay time plays an important role in the reconstruction of the attractor from the scalar time series. If τ_d is too small, the reconstructed attractor is compressed along the identity line, and this is called *redundance*. If τ_d is too large, the attractor dynamics may become causally disconnected, and this is called *irrelevance* [5].

In common practice, the delay time τ_d is chosen so as to ensure that the components of \mathbf{x}_i are independent, and the same delay time is used for all embedding dimensions m . However, in recent years, it has been suggested that the delay time window $\tau_w = (m-1)\tau$, which is the entire time spanned by the components of \mathbf{x}_i , should be independent of m instead [6–11]. In this case, the delay time τ varies with the embedding dimension m . Rosenstein et al. [6] compared several approaches for estimating τ_d and τ_w and indicated the critical disadvantages of each, such as inconsistencies and long computation times. The geometrical concepts of *redundance* and *irrelevance* were used to evaluate the quality of an attractor's reconstruction. It has been suggested [11] that the delay time window may be set to $\tau_w \geq \tau_p$, where τ_p is the mean orbital period, which can be approximated by examining the oscillations of the time series. Martinerie et al. [7] examined the delay time window and compared it with the delay times estimated using the autocorrelation function and the mutual information, but they concluded that τ_w cannot be estimated using either of these methods. Basically, τ_w is the optimal time for independence of the data, but these methods estimate only the first locally optimal time, which is τ_d .

In this work, we develop a technique for choosing either the delay time τ_d or the delay time window τ_w using the correlation integral. Although τ_w may be a more useful quantity for the estimation of the dimension than τ_d , many researchers continue to use τ_d for attractor reconstruction using the method of delays. Also, the estimation of both τ_d and τ_w using this technique helps to illuminate the difference between these two quantities. Brock et al. [12,13], in their development of a test for nonlinearity in a time series, used the statistic $S(m, N, r) = C(m, N, r) - C^m(1, N, r)$, where $C(m, N, r)$ is the correlation integral (see Eq. (2)). Here, we use the similar statistic $S(m, N, r, t) = C(m, N, r, t) - C^m(1, N, r, t)$, and we examine its dependence on the index lag t . Hence, we call our technique the C–C method. Compared with the use of the mutual information, the C–C method is easier to implement, useful for smaller data sets, and less demanding computationally. Also, its estimate of τ_d agrees with the first local minimum of the mutual information.

2. Correlation integral and BDS statistic

The correlation dimension introduced by Grassberger and Procaccia [14,15] is widely used in many fields for the characterization of strange attractors. The correlation integral for the embedded time series is the following function:

$$C(m, N, r, t) = \frac{2}{M(M-1)} \sum_{1 \leq i < j \leq M} \Theta(r - \|\mathbf{x}_i - \mathbf{x}_j\|), \quad r > 0, \quad (2)$$

where

$$\Theta(a) = 0, \quad \text{if } a < 0, \quad \Theta(a) = 1, \quad \text{if } a \geq 0,$$

N is the size of the data set, t is the index lag, $M = N - (m - 1)t$ is the number of embedded points in m -dimensional space, and $\|\cdots\|$ denotes the sup-norm. $C(m, N, r, t)$ measures the fraction of the pairs of points $\mathbf{x}_i, i = 1, 2, \dots, M$, whose sup-norm separation is no greater than r . If the limit of $C(m, N, r, t)$ as $N \rightarrow \infty$ exists for each r , then the fraction of all state vector points that are within r of each other is denoted by $C(m, r, t) = \lim_{N \rightarrow \infty} C(m, N, r, t)$, and the correlation dimension is defined as $D_2(m, t) = \lim_{r \rightarrow 0} [\log C(m, r, t) / \log r]$. Since N remains finite for real data sets, then we cannot let r go to 0; instead, we look for a linear region of slope $D_2(m, t)$ in the plot of $\log C(m, N, r, t)$ vs. $\log r$.

Brock et al. [12,13] studied the BDS statistic, which is based on the correlation integral, to test the null hypothesis that a given data set is independently and identically distributed (iid). The test has been particularly useful for chaotic systems and nonlinear stochastic systems. For completeness, we briefly review the BDS statistic. Let F be the invariant distribution of a variable X in the state space, and define the spatial correlation integral as follows:

$$C(m, r) = \int \int \Theta(r - \|\mathbf{x} - \mathbf{y}\|) dF(\mathbf{x}) dF(\mathbf{y}), \quad r > 0. \quad (3)$$

If X is iid, then using $\Theta(r - \|\mathbf{x} - \mathbf{y}\|) = \prod_{k=1}^m \Theta(r - |x_k - y_k|)$ yields $C(m, r) = C^m(1, r)$, where

$$C(1, r) = \int [F(x + r) - F(x - r)] dF(x) \equiv C. \quad (4)$$

Denker and Keller [16] showed that $C(m, N, r)$ is a U-statistic estimator. Using the U-statistics theory for an absolutely regular process, Brock et al. [12,13] proved that, as $N \rightarrow \infty$, $\sqrt{N}[C(m, N, r) - C^m(1, r)]$ approaches a normal distribution with mean zero and variance

$$\sigma^2(m, r) = 4 \left[K^m - C^{2m} + 2 \sum_{i=1}^{m-1} (K^{m-i} C^{2i} - C^{2m}) \right], \quad (5)$$

where

$$K \equiv \int [(F(x + r) - F(x - r))]^2 dF(x) \quad (6)$$

(assuming that $K > C^2$). Thus, the BDS statistic defined by

$$\text{BDS}(m, N, r) = \frac{\sqrt{N}}{\sigma(m, r)} [C(m, N, r) - C^m(1, r)] \quad (7)$$

approaches a standard normal distribution. However, since the distribution F is generally unknown, then we cannot obtain the values of C and K and the variance $\sigma^2(m, r)$ from the above definitions. Instead, the correlation integral $C(1, r)$ and the variance $\sigma^2(m, r)$ must be estimated from the sample data. Thus, $C(1, r)$ is estimated by $C(1, N, r, t)$, and $\sigma^2(m, r)$ is estimated by

$$\hat{\sigma}^2 = 4 \left\{ m(m-1) \hat{C}^{2(m-1)} (\hat{K} - \hat{C}^2) + \hat{K}^m - \hat{C}^{2m} + 2 \sum_{i=1}^{m-1} [\hat{C}^{2i} (\hat{K}^{m-i} - \hat{C}^{2(m-i)}) - m \hat{C}^{2(m-i)} (\hat{K} - \hat{C}^2)] \right\}, \quad (8)$$

where $\hat{C} = C(m, N, r, t)$ and

$$\hat{K} = \frac{6}{M(M-1)(M-2)} \sum_{1 \leq i < j < k \leq M} \Theta(r - \|\mathbf{x}_i - \mathbf{x}_j\|) \Theta(r - \|\mathbf{x}_j - \mathbf{x}_k\|), \quad (9)$$

with $M = N - (m - l)t$. Then, under the iid hypothesis, if $\hat{K} > \hat{C}^2$ and $m > 1$, the BDS statistic becomes

$$\text{BDS}(m, N, r) = \frac{\sqrt{N}}{\hat{\sigma}} [C(m, N, r, t) - C^m(1, N, r, t)], \quad (10)$$

and this converges to a standard normal distribution as $N \rightarrow \infty$.

The BDS statistic originates from the statistical properties of the correlation integral, and it measures the statistical significance of calculations of the correlation dimension. Even though the BDS statistic cannot be used to distinguish between a nonlinear deterministic system and a nonlinear stochastic system, it is a powerful tool for distinguishing random time series from chaotic or nonlinear stochastic time series. Further statistical properties and proofs can be found in [12,13].

3. Measure of nonlinear dependence

3.1. C–C method

This study is concerned with the properties of $S(m, N, r, t) = C(m, N, r, t) - C^m(1, N, r, t)$. We refer to a comment by Brock et al. [12]: “If the stochastic process $\{x_i\}$ is iid, it will be shown that $C(m, r) = C^m(1, r)$ for all m and r . That is to say, the correlation integral behaves much like the characteristic function of a serial string in that the correlation integral of a serial string of independent random variables is the product of the correlation integrals of component substrings.” This leads us to interpret the statistic $S(m, N, r, t)$ as the serial correlation of a nonlinear time series. Therefore, it can be regarded as a dimensionless measure of nonlinear dependence. For fixed m, N , and r the plot of $S(m, N, r, t)$ vs. t is a nonlinear analog of the plot of the autocorrelation function vs. t .

In order to study the nonlinear dependence and eliminate spurious temporal correlations, we must subdivide the time series $\{x_i\}$, $i = 1, 2, \dots, N$, into t disjoint time series, where t is the index lag. $S(m, N, r, t)$ is then computed from these disjoint time series as follows:

For $t = 1$, we have the single time series $\{x_1, x_2, \dots, x_N\}$, and

$$S(m, N, r, 1) = C(m, N, r, 1) - C^m(1, N, r, 1). \quad (11)$$

For $t = 2$, we have the two disjoint time series $\{x_1, x_3, \dots, x_{N-1}\}$ and $\{x_2, x_4, \dots, x_N\}$, each of length $N/2$, and we average the values of $S(m, N/2, r, 2)$ for these two series:

$$S(m, N, r, 2) = \frac{1}{2} \{ [C_1(m, N/2, r, 2) - C_1^m(1, N/2, r, 2)] + [C_2(m, N/2, r, 2) - C_2^m(1, N/2, r, 2)] \}. \quad (12)$$

For general t , this becomes

$$S(m, N, r, t) = \frac{1}{t} \sum_{s=1}^t [C_s(m, N/t, r, t) - C_s^m(1, N/t, r, t)]. \quad (13)$$

Finally, as $N \rightarrow \infty$, we can write

$$S(m, r, t) = \frac{1}{t} \sum_{s=1}^t [C_s(m, r, t) - C_s^m(1, r, t)], \quad m = 2, 3, \dots \quad (14)$$

For fixed m and t , $S(m, r, t)$ will be identically equal to 0 for all r if the data are iid and $N \rightarrow \infty$. However, real data sets are finite, and the data may be serially correlated; so, in general, we will have $S(m, r, t) \neq 0$. Thus, the locally optimal times may be either the zero crossings of $S(m, r, t)$ or the times at which $S(m, r, t)$ shows the least

variation with r , since this indicates a nearly uniform distribution of points. Hence, we select several representative values r_j , and we define the quantity

$$\Delta S(m, t) = \max\{S(m, r_j, t)\} - \min\{S(m, r_j, t)\}, \quad (15)$$

which is a measure of the variation of $S(m, r, t)$ with r . The locally optimal times t are then the zero crossings of $S(m, r, t)$ and the minima of $\Delta S(m, t)$. The zero crossings of $S(m, r, t)$ should be nearly the same for all m and r , and the minima of $\Delta S(m, t)$ should be nearly the same for all m (otherwise, the time is not locally optimal). The delay time τ_d will correspond to the first of these locally optimal times.

3.2. Choosing m and r for finite sample sizes

In determining the nonlinear dependence of a finite time series by using the statistic $S(m, N, r, t)$, one must have criteria for selecting the values of m and r . In addition, one must know the role of the sample size N . For a fixed value of N , as m becomes large, the data become very sparse, so that $C(m, N, r, t)$ becomes vanishingly small. Also, if r exceeds the size of the attractor, then $C(m, N, r, t)$ saturates, since most pairs of points are within the distance r . Thus, neither m nor r should be too large.

Appropriate choices for m , N , and r may be found by examining the BDS statistic. Brock et al. [12] carried out such an investigation using time series generated from a variety of asymptotic distributions. Time series with three sample sizes, $N = 100, 500$ and 1000 , were generated by Monte Carlo simulations from six asymptotic distributions: a standard normal distribution, a student- t distribution with three degrees of freedom, a double exponential distribution, a chi-square distribution with four degrees of freedom, a uniform distribution, and a bimodal mixture of normal distributions. These studies led to the conclusion that m should be between 2 and 5 and r should be between $\sigma/2$ and 2σ . In addition, the asymptotic distributions were well approximated by finite time series when $N \geq 500$.

As an example, we illustrate the dependence of the statistic $S(m, N, r, 1)$ on m , N , and r for the Lorenz system of three coupled differential equations [17]:

$$\dot{x} = -a(x - y), \quad \dot{y} = -xz + cx - y, \quad \dot{z} = xy - bz, \quad (16)$$

where a , b , and c are constants. We solve this system of equations for $a = 16.0$, $b = 4.0$, and $c = 45.92$ to generate a time series of the variable x with $\tau_s = 0.01$. The values of $S(m, N, r, 1)$ for various sample sizes N and various values of m and r are shown in Fig. 1. For $N > 100$, $S(m, N, r, 1)$ is positive for all r between 0 and 4σ , where σ is the standard deviation of the data set, and it is nearly 0 for $r > 4\sigma$. For $N > 1000$, the behavior of $S(m, N, r, 1)$ as a function of r within the range $0 < r < 4\sigma$ is nearly the same as for $N = 1000$, and this represents the true correlation of the time series.

However, as N decreases below 100, $S(m, N, r, 1)$ fails to represent the true correlation. For example, when $N = 15$ and $m = 10$, $S(m, N, r, 1)$ vanishes at $r = (0.5)\sigma$, and it departs substantially from the curves for $N = 1000$. On the other hand, for $m = 2$, $S(m, N, r, 1)$ has roughly the correct behavior even for N as small as 15. These results are in agreement with the conclusions of Brock et al. Within the ranges $2 \leq m \leq 5$ and $\sigma/2 \leq r \leq 2\sigma$, the quantity $S(m, N, r, 1)$ represents the serial correlation for the wide range of values of N considered in Fig. 1. Finally, note that, while the asymptotic distributions are well approximated by finite time series of size $N \geq 500$, it is not necessary for N to be this large in order to use $S(m, N, r, 1)$ to study the nonlinear dependence of the time series.

For the range $\sigma/2 \leq r \leq 2\sigma$, we choose four representative values $r_1 = (0.5)\sigma$, $r_2 = (1.0)\sigma$, $r_3 = (1.5)\sigma$, and $r_4 = (2.0)\sigma$. We then define the following averages of the quantities given by Eqs. (14) and (15):

$$\bar{S}(t) = \frac{1}{16} \sum_{m=2}^5 \sum_{j=1}^4 S(m, r_j, t), \quad (17)$$

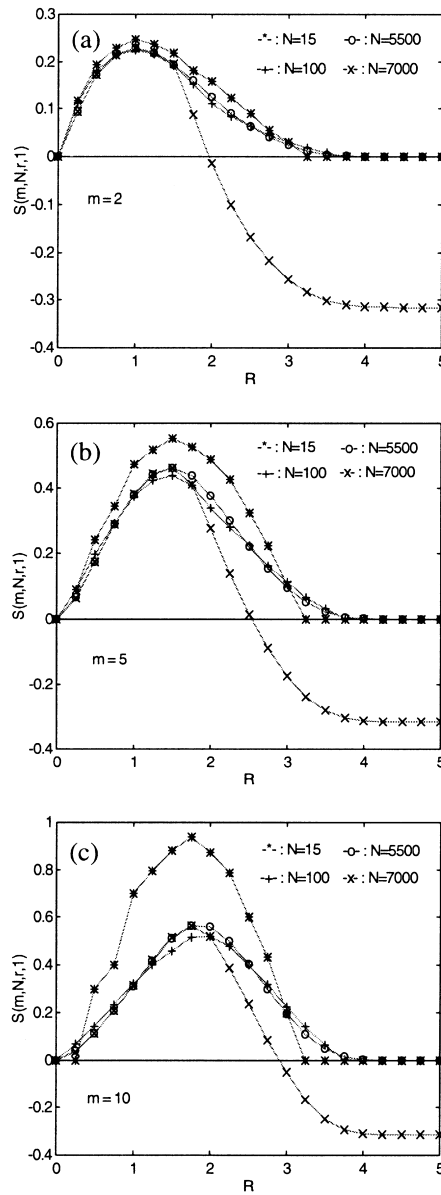


Fig. 1. $S(m, N, r, 1)$ vs. $R = r/\sigma$ for various sample sizes N and for $m = 2, 5,$ and 10 using time series of the variable x from the Lorenz system of Eq. (16) with $a = 16.0, b = 4.0, c = 45.92,$ and $\tau_s = 0.01$.

$$\Delta \bar{S}(t) = \frac{1}{4} \sum_{m=2}^5 \Delta S(m, t), \tag{18}$$

and we look for the first zero crossing of $\bar{S}(t)$ or the first local minimum of $\Delta \bar{S}(t)$ to find the first locally optimal time for independence of the data, which gives the delay time $\tau_d = t\tau_s$. The optimal time is the index lag t for which $\bar{S}(t)$ and $\Delta \bar{S}(t)$ are both closest to 0. If we assign equal importance to these two quantities, then we may simply

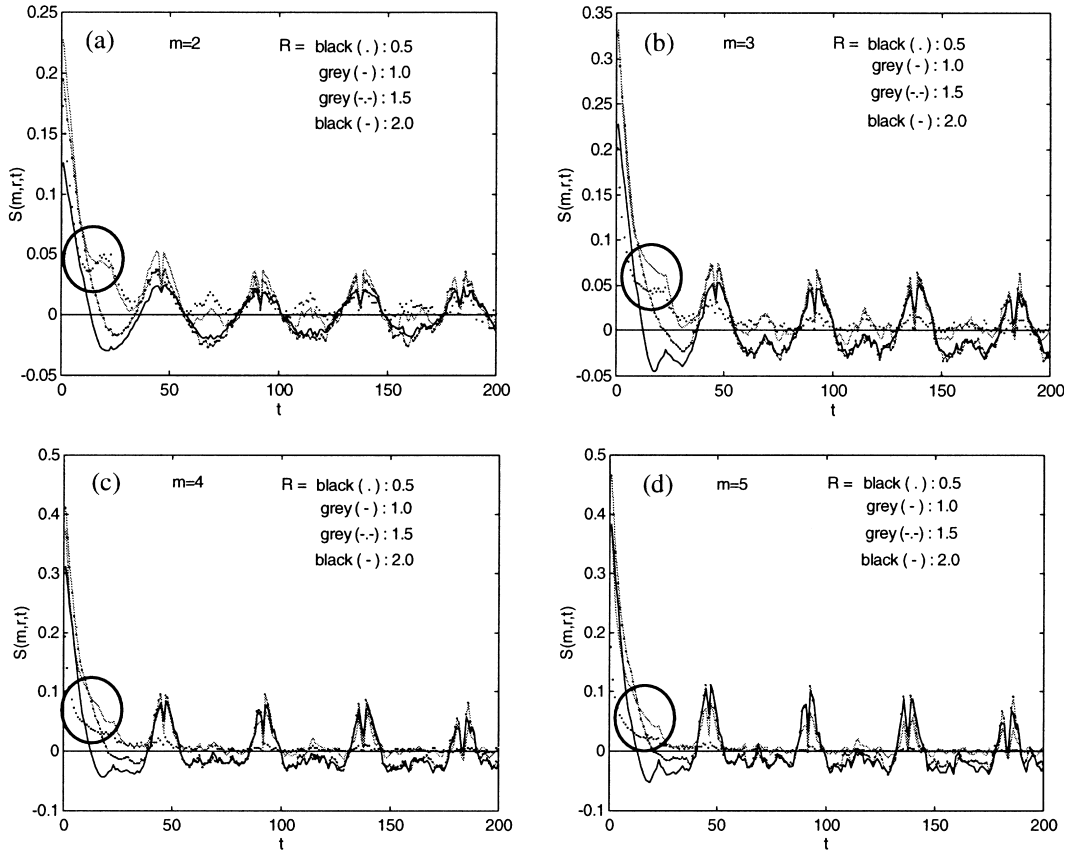


Fig. 2. $S(m, r, t)$ for the variable x from the Lorenz system of Eq. (16) with $a = 16.0$, $b = 4.0$, $c = 45.92$, and $\tau_s = 0.01$ using 3000 data points. The circles indicate the vicinity of τ_d , where the first local minimum occurs in the variation of $S(m, r, t)$ with r .

look for the minimum of the quantity

$$S_{\text{cor}}(t) = \Delta \bar{S}(t) + |\bar{S}(t)|, \quad (19)$$

and this optimal time gives the delay time window $\tau_w = t\tau_s$.

4. Numerical examples

The following examples show that the C–C method is relatively easy to implement, that the delay times τ_d agree with those obtained using mutual information, that the delay time windows τ_w agree with those obtained by Martinerie et al. [7], and that the method is robust to the presence of noise.

4.1. Applications to dynamical systems

We solve the Lorenz system of Eq. (16) for $a = 16.0$, $b = 4.0$, and $c = 45.92$ to generate a time series of 3000 data points with $\tau_s = 0.01$. We then compute $S(m, r, t)$ from Eq. (14), and the results are shown in Fig. 2. The circles in Fig. 2 indicate the index lag t where the variation of $S(m, r, t)$ with r is at its first local minimum, and Fig. 3(a) shows this first local minimum of $\Delta S(m, t)$ more clearly. We choose the delay time at this point, which gives

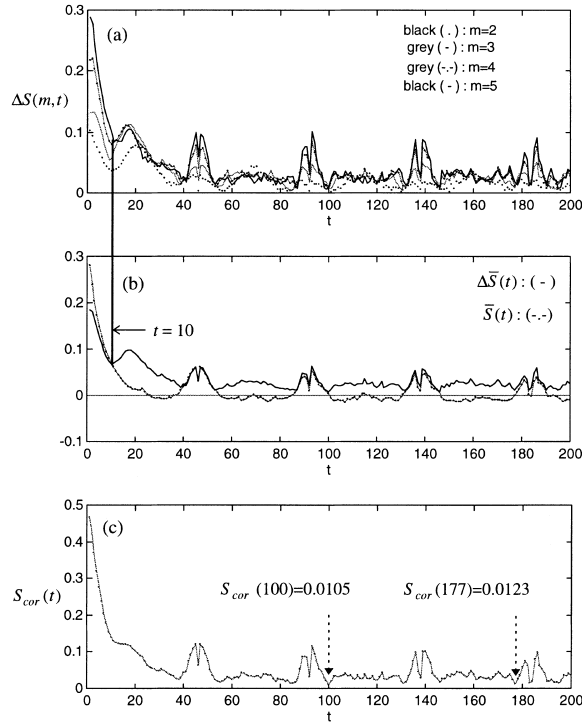


Fig. 3. $\Delta S(m, t)$, $\Delta \bar{S}(t)$, $\bar{S}(t)$, and $S_{cor}(t)$ for the variable x from the Lorenz system of Fig. 2. The solid line locates $\tau_d = 10\tau_s$, and the minimum of $S_{cor}(t)$ yields $\tau_w = 100\tau_s$.

$\tau_d = 10\tau_s = 0.10$. This agrees with the delay time found by Abarbanel et al. [18] using the first local minimum of the mutual information. Also, from the minimum of $S_{cor}(t)$ in Fig. 3(c), we choose the delay time window as $\tau_w = 100\tau_s = 1.0$.

If we instead choose $a = 10$, $b = 28$, and $c = 8/3$, as used by Martinerie et al. [7], to generate a time series of 3000 data points with $\tau_s = 0.01$, we then obtain a delay time $\tau_d = 0.18$ and delay time window $\tau_w = 1.23$. For this case, Martinerie et al. [7] obtained $\tau_d = 0.17$ from the first minimum of the mutual information and $\tau_w = 1.0$ from their empirical procedure, and our values are in agreement with theirs.

Next, we generate 3000 data points for the variable x from each of the following three systems: the three-torus [7]

$$x_i = \sin \left[\frac{3i}{500} \right] + \sin \left[\frac{3\sqrt{2}i}{250} \right] + \sin \left[\frac{9\sqrt{3}i}{500} \right], \quad (20)$$

the Rabinovich–Fabrikant system [19]

$$\dot{x} = y(z - 1 + x^2) + \gamma x, \quad \dot{y} = x(3z + 1 - x^2) + \gamma y, \quad \dot{z} = -2z(\alpha + xy), \quad (21)$$

and the Rossler system [20]

$$\dot{x} = -(y + z), \quad \dot{y} = x + dy, \quad \dot{z} = e + z(x - f). \quad (22)$$

Table 1 summarizes our results for the four systems studied.

For the three-torus, τ_d corresponds to the first zero crossing of $S(m, r, t)$, as indicated by the circles in Fig. 4. For each value of m , this occurs for $r = (0.5)\sigma$, and averaging these four values of t gives $t = 55$ and $\tau_d = 55\tau_s = 55$. Also, as shown in Fig. 5(c), the minimum of $S_{cor}(t)$ occurs for $t = 101$, which gives $\tau_w = 101\tau_s = 101$.

Table 1
Summary of results for four dynamical systems

System	Parameters	τ_s	C–C method		MI ^b		Ref.
			$\tau_d(t)$	$\tau_w(t)$	τ_d	τ_w	
Lorenz	$a = 16, b = 4.0, c = 45.92$	0.01	0.1(10)	1.0(100)	0.1		[18]
	$a = 10, b = 28, c = 8/3$	0.01	0.18(18)	1.23(123)	0.17	1.0	[7]
Three-torus		1.0	55(55)	101(101)	54	100	[7]
R–F ^a	$\gamma = 0.87, \alpha = 1.1$	0.01	0.52(52)	1.28(128)			
Rossler	$d = 0.2, e = 0.4, f = 5.7$	0.05	0.85(17)	9.55(191)	1.0	3.3	[7]
	$d = 0.2, e = 0.2, f = 5.0$	0.05	0.80(16)	5.85(117)		$6 \leq \tau_w < 14$	[11]

^aR–F: Rabinovich–Fabrikant system.

^bMI: The first minimum of the mutual information.

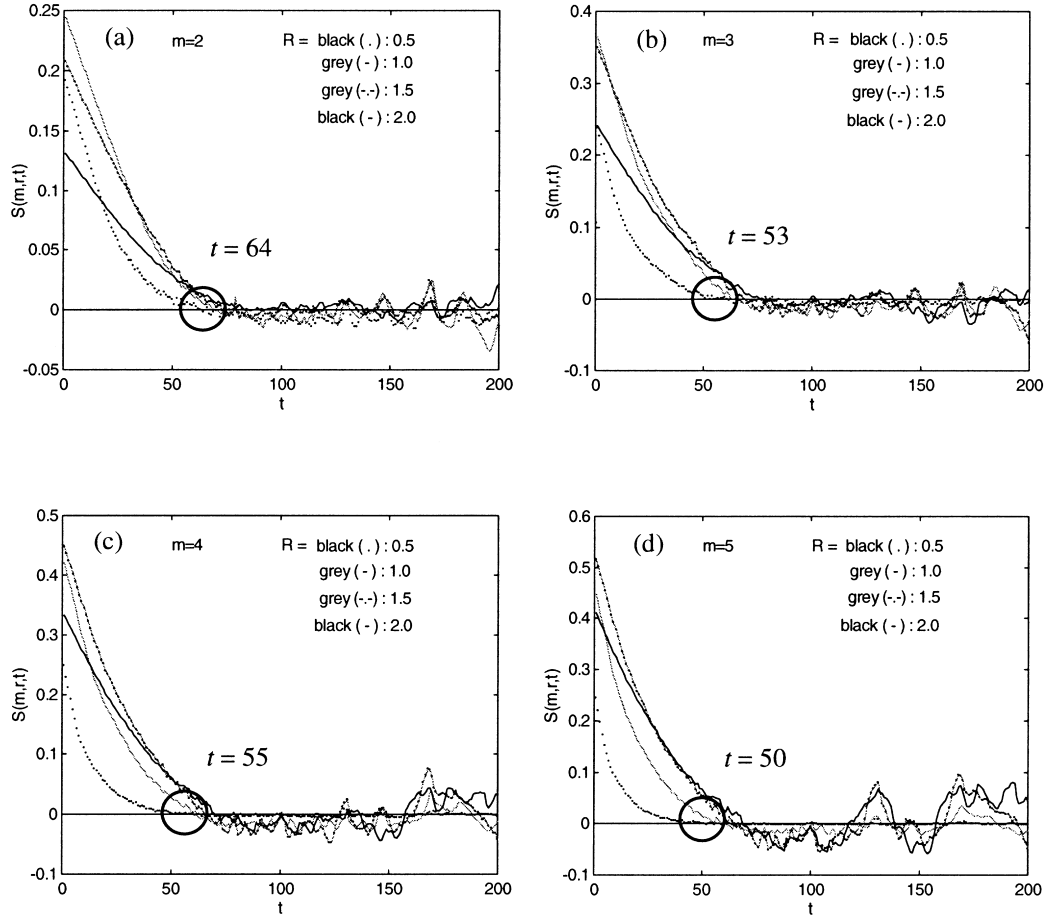


Fig. 4. $S(m, r, t)$ for the three-torus of Eq. (20) using 3000 data points with $\tau_s = 1$. The circles indicate the vicinities of τ_d , which is where the first zero crossing occurs, and $\tau_d = 55\tau_s$ is obtained by averaging these four zero crossings.

The other cases are handled similarly, and Table 1 shows that our results are generally in agreement with those of previous works. The only disagreement occurs for the determination of τ_w for one case of the Rossler system, so we show the plot of $S_{\text{cor}}(t)$ for this system in Fig. 6. Based on these plots, we believe that the empirical result $\tau_w = 3.3$ from [7] is incorrect (this would correspond to $t = 66$ in Fig. 6).

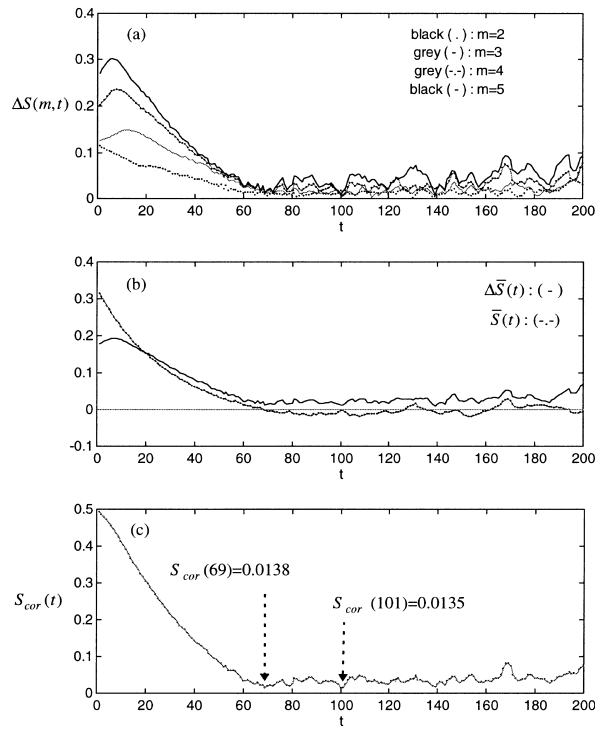


Fig. 5. $\Delta S(m, t)$, $\Delta \bar{S}(t)$, $\bar{S}(t)$, and $S_{cor}(t)$ for the three-torus of Fig. 4. The minimum of $S_{cor}(t)$ gives $\tau_w = 101\tau_s$.

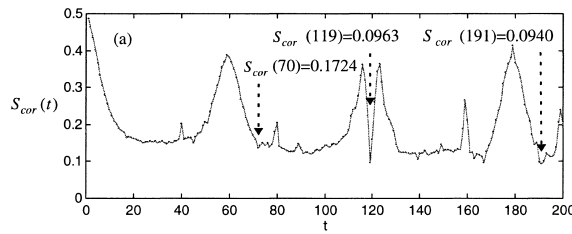


Fig. 6. $S_{cor}(t)$ for the variable x from the Rossler system of Eq. (22) with $d = 0.2$, $e = 0.4$, $f = 5.7$, and $\tau_s = 0.05$. The minimum gives $\tau_w = 191\tau_s$.

4.2. Noise effects

To study the effects noise on the C–C method, we added Gaussian noise to the Lorenz time series. Specifically, we examined the time series $x_i = L_i + \eta\sigma\varepsilon_i$, where L_i is the noise-free Lorenz time series, σ is its standard deviation, ε_i is a Gaussian iid random variable with zero mean and a standard deviation of 1, and η is the strength of the noise. Noise levels of 10%, 30%, 50%, and 100% ($\eta = 0.1, 0.3, 0.5$ and 1.0) were added to the Lorenz time series, and the C–C method was performed for each of these noise levels. The results are shown in Figs. 7 and 8, and we observe that the estimates of τ_d and τ_w remain unchanged for $\eta = 0.1$ and 0.3 , but not for $\eta = 0.5$ and 1.0 . However, even in these cases, the estimates of τ_d and τ_w remain within a factor of 2 of the correct values. This robustness to noise is due to the fact that Eqs. (17) and (18) involve averages, and the noise tends to average to 0.

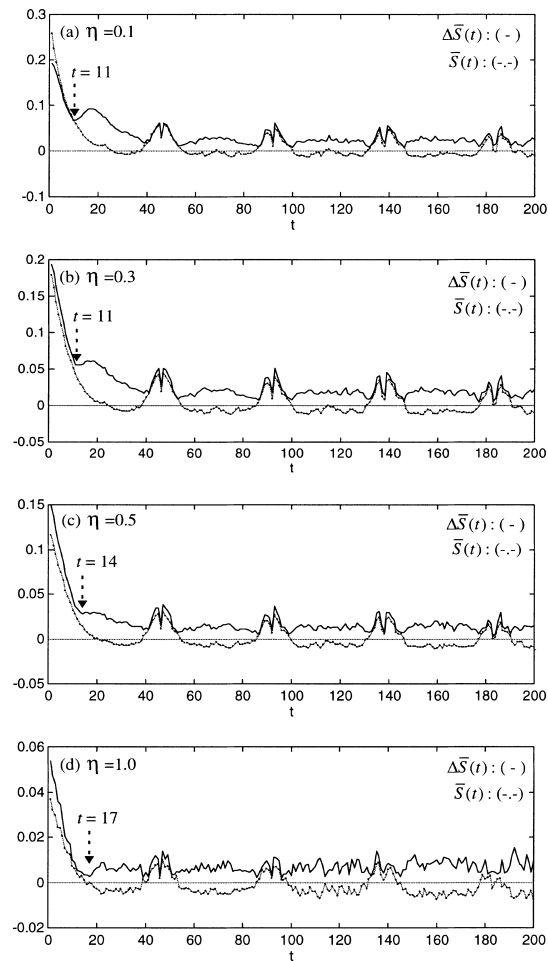


Fig. 7. $\Delta\bar{S}(t)$ and $\bar{S}(t)$ for the variable x from the Lorenz system of Fig. 2 with added noise levels of: (a) 10%, (b) 30%, (c) 50%, and (d) 100%. The resulting estimates of the delay time are: (a) $\tau_d = 11\tau_s$, (b) $\tau_d = 11\tau_s$, (c) $\tau_d = 14\tau_s$, and (d) $\tau_d = 17\tau_s$.

5. Concluding remarks

Martinerie et al. [7] compared the delay time windows τ_w for the Lorenz system, a three-torus, and the Rossler system with the delay times τ_d found from the autocorrelation function and the mutual information. They concluded that the autocorrelation function and the mutual information could not give the value of τ_w . We have introduced a new method, called the C–C method, and shown that it can be used to find both τ_d and τ_w . The values found for τ_d agree well with those found from the first minimum of the mutual information, and the values found for τ_w agree well with the empirical estimates made by Martinerie et al. [7] and by Kugiumtzis [11]. It should be noted that the C–C method makes use of a statistic within the reconstructed phase space, rather than analyzing the temporal evolution of the time series.

For the choice of m and r in the C–C method, there is no firm theoretical basis [12]. However, when we use the recommendation of Brock et al. [12], the C–C method works well. It is also important to consider the data sample size N . In particular, for the Lorenz system, if $N > 7000$, then $S(m, N, r, t)$ will always cross zero at $t = 1$. However, if this first zero crossing is ignored, we have found that the C–C method still works well. Also, the C–C method shows robustness to the presence of noise, especially for noise levels below 30%.

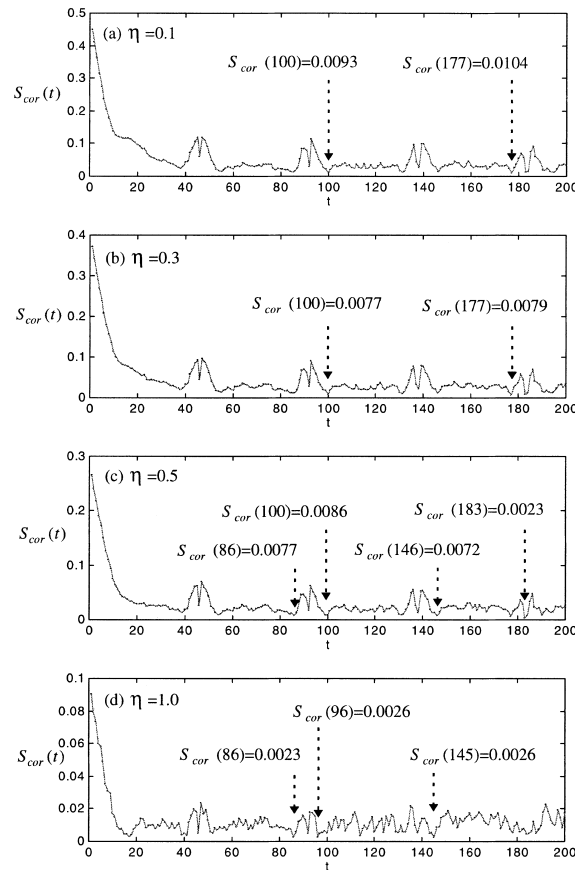


Fig. 8. $S_{cor}(t)$ for the variable x from the Lorenz system of Fig. 2 with added noise levels of: (a) 10%, (b) 30%, (c) 50%, and (d) 100%. The resulting estimates of the delay time window are: (a) $\tau_w = 100\tau_s$, (b) $\tau_w = 100\tau_s$, (c) $\tau_w = 183\tau_s$, and (d) $\tau_w = 86\tau_s$.

In summary, the C–C method is a relatively simple method for estimating both τ_d and τ_w , it requires relatively small data sets, and it is not computationally demanding.

Acknowledgements

This research has been partially supported by NSF Grant CMS-9625685 on “Uncertainty and Risk Analysis Under Extreme Hydrologic Events.”

References

- [1] A.M. Fraser, H.L. Swinney, Phys. Rev., A 33 (1986) 1134.
- [2] A.A. Tsonis, Chaos: From Theory to Applications, Plenum Press, New York, 1992.
- [3] N.H. Packard, J.P. Crutchfield, J.D. Farmer, R.S. Shaw, Phys. Rev. Lett. 45 (1980) 712.
- [4] F. Takens, in: D.A. Rand, L.S. Young (Eds.), Dynamical Systems and Turbulence, Lecture Notes in Mathematics, vol. 898 Springer, Berlin, 1981, p. 336.
- [5] M. Casdagli, S. Eubank, J.D. Farmer, J. Gibson, Physica D 51 (1991) 52.
- [6] M.T. Rosenstein, J.J. Collins, C.J. De Luca, Physica D 73 (1994) 82.
- [7] J.M. Martinerie, A.M. Albano, A.I. Mees, P.E. Rapp, Phys. Rev. A 45 (1992) 7058.

- [8] D.S. Broomhead, G.P. King, *Physica D* 20 (1986) 217.
- [9] A.I. Mees, P.E. Rapp, L.S. Jennings, *Phys. Rev. A* 36 (1987) 340.
- [10] A.M. Albano, J. Muench, C. Schwartz, A.I. Mees, P.E. Rapp, *Phys. Rev. A* 38 (1988) 3017.
- [11] D. Kugiumtzis, *Physica D* 95 (1996) 13.
- [12] W.A. Brock, D.A. Hsieh, B. LeBaron, *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence*, MIT Press, Cambridge, MA, 1991.
- [13] W.A. Brock, W.D. Dechert, J.A. Scheinkman, B. LeBaron, *Econ. Rev.* 15 (1996) 197.
- [14] P. Grassberger, I. Procaccia, *Phys. Rev. Lett.* 50 (1983) 345.
- [15] P. Grassberger, I. Procaccia, *Physica D* 7 (1983) 153.
- [16] M. Denker, G. Keller, *J. Stat. Phys.* 44 (1986) 67.
- [17] E.N. Lorenz, *J. Atmos. Sci.* 20 (1963) 130.
- [18] H.D.I. Abarbanel, R. Brown, J.J. Sidorowich, L.S. Tsimring, *Rev. Mod. Phys.* 65 (1993) 1331.
- [19] M.I. Rabinovich, A.L. Fabrikant, *Sov. Phys. JETP* 50 (1979) 311.
- [20] O.E. Rossler, *Phys. Lett. A* 57 (1976) 397.