Modified correlation entropy estimation for a noisy chaotic time series

A. W. Jayawardena,[1,a)] Pengcheng Xu,[2] and W. K. Li[3]

[1]*International Centre for Water Hazard and Risk Management,*
Public Works Research Institute, 305-8516 Tsukuba, Japan
[2]*Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, China*
[3]*Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong, Hong Kong*

A method of estimating the Kolmogorov–Sinai (KS) entropy, herein referred to as the modified correlation entropy, is presented. The method can be applied to both noise-free and noisy chaotic time series. It has been applied to some clean and noisy data sets and the numerical results show that the modified correlation entropy is closer to the KS entropy of the nonlinear system calculated by the Lyapunov spectrum than the general correlation entropy. Moreover, the modified correlation entropy is more robust to noise than the correlation entropy. © *2010 American Institute of Physics*. [doi:10.1063/1.3382013]

Correlation entropy (CE), which is an approximation of the Kolmogorov–Sinai (KS) entropy, is an important quantity that can be used to identify chaos in a nonlinear dynamical system. It gives an indication of the predictability of the nonlinear time series since the inverse of the KS entropy evaluates the maximum time step for prediction. Estimating the CE accurately is important for the simulation and forecasting of a chaotic time series. However, this is made difficult by the presence of noise which blurs the signal of the time series, thereby introducing errors in the calculation of the correlation sum. To overcome this difficulty, we introduce a new quantity, herein referred to as the modified correlation entropy (MCE), as an approximation of the KS entropy. For noise-free data, the MCE is equivalent to the general CE. For noisy chaotic data, the MCE is closer to the KS entropy than the general CE.

I. INTRODUCTION

Study of chaos in time series has seen a rapid growth during the past two decades or so. Many types of time series which appear to be stochastic superficially are in fact of a deterministic nature of low dimension. If such types can be identified, short-term predictions can be made using local models rather than stochastic global models. The identification of chaos involves the determination of certain invariants such as the correlation dimension, Lyapunov exponent, entropies, among others. They are usually carried out in a reconstructed phase space of the original time series. By Takens's embedding theorem (Takens, 1981), a nonlinear time series can be embedded into a delay phase space by a special embedding dimension and an arbitrary choice of the time delay.

Important characteristics of a chaotic time series are that it is initial value sensitive, consists of a topological mixture, has dense periodic orbits, and is fractal. The sensitivity of

initial value can be evaluated by the largest Lyapunov exponent while fractal structure can be validated by estimating the fractal dimension. The topological mixture rate of a nonlinear time series can be evaluated by the dynamical entropy.

The concept of dynamical entropy, which was first introduced by Kolmogorov (1958) and Sinai (1959), provides a method to describe the mixing rate of a deterministic system. Hence it is called the "KS" entropy, which is based on an invariant measure of the attractor and it is a "measure of theoretic entropy" of a nonlinear system. Later, the concept of "topological entropy" (T-Entropy) that does not employ a metric for the system was introduced by Adler *et al.* (1965). Given the evolution function of the dynamical system, the T-Entropy or the KS entropy can be calculated. However, for nonlinear time series, the evolution function is unknown, and therefore the T-Entropy or the KS entropy cannot be calculated from the definitions. Several methods have been developed to estimate the KS entropy for a chaotic time series. Grassberger and Procaccia (1983) introduced a method to approximate the KS entropy for a chaotic time series using the correlation sum (correlation integral). The estimation of the entropy by this method is then called the CE. Another entropy measure for a time series is its "information entropy," which is also called the "Shannon entropy" since it was first introduced by Shannon (1948) and Weaver and Shannon (1949). A generalized entropy for a nonlinear system is also introduced according to the *q*-deformed algebra and special functions developed by Renyi (1971), and it is called the "Renyi entropy." The information entropy and the CE have been proved to be special cases of the Renyi entropy. Methods have been developed to calculate various entropy measures for a chaotic time series. Grassberger and Procaccia (1983) provided a simple formula to estimate the CE for a chaotic time series, which has been applied widely. Kantz and Schürmann (1996) introduced a method to calculate a finite order entropy to approximate the KS entropy. Schürmann and Grassberger (1996) provided a method to estimate the information entropy for a symbolic sequence, which can also be applied for a chaotic time series. Recently,

---
a)Author to whom correspondence should be addressed. Electronic mail: hrecjaw@hkucc.hku.hk.

the above methods have been employed to estimate the entropy measure for some special cases, such as chaotic time series with small data size (Bonachela *et al.*, 2008).

However, all chaotic time series are blurred by noise. The presence of additive noise strongly affects the evolution of the correlation sum, which in turn affects the estimation of the CE. The correlation sum for a chaotic time series with additive Gaussian noise has been discussed by several authors. Most of the correlation integral equations depend nonlinearly on the KS entropy. Hence, it is difficult to estimate the KS entropy by these correlation integral equations.

In this study, by employing the correlation integral equation obtained by Diks (1999) (p. 123) and Oltmans and Verheijen (1997) for a chaotic time series with additive Gaussian noise, we provide a method to estimate the KS entropy, herein referred to as the "MCE," which has the advantage that it is similar to the CE and that it can be calculated by the correlation sum directly without any statistical methods. It is also robust to noise, i.e., the presence of additive Gaussian noise does not affect the estimation of the KS entropy. Hence, the MCE is suitable for all chaotic time series regardless of whether they are clean or noisy.

The application of the method is demonstrated with four chaotic time series in this study: two artificial and two real world. The numerical results show that the MCE is a robust estimation for the KS entropy. The KS entropy is also estimated by using the Lyapunov spectrum, and numerical results for the four data sets used show that the MCE is closer to the KS entropy than the CE.

The flow of this paper is arranged as follows. In Sec. II, some definitions of various entropies are introduced. In Sec. III, the correlation sum is introduced for a clean and a noisy chaotic time series. In Sec. IV, a method to estimate the KS entropy for a clean or a noisy chaotic time series is developed. In Sec. V, the estimation method for the Lyapunov spectrum is introduced to estimate the KS entropy. In Sec. VI, the proposed estimation method is applied to several data sets, and the results and discussion are presented in Sec. VII.

II. ENTROPY OF A DYNAMICAL SYSTEM

There are several types of entropies defined in the literature. In the context of thermodynamics, entropy refers to the amount of "disorder" in the system—the higher the entropy, the higher the amount of disorder. For a closed thermodynamic system, it is the measure of the amount of thermal energy not available to do mechanical work. In statistical mechanics, it refers to the amount of uncertainty in the system. In information theory, it is a measure of the uncertainty associated with a random variable. Shannon entropy (Shannon, 1948) refers to a measure of the average information content that is missing by not knowing the value of the random variable. In statistical thermodynamics, it measures the degree to which the probability of the system is spread out over all possible substates.

In a time series, entropy is an important entity as its inverse gives the time scale relevant for the predictability of the system. It also provides the topological information about the folding process. Several definitions of entropies can be found in the literature [e.g., Kantz and Schreiber (2003)], and some of them are briefly given here.

In information theory, the entropy of a random variable is defined as

$$K(Z) = -\sum_i p_i \ln p_i, \tag{1}$$

where p_i is the probability of the random variable Z taking a specified value z_i [i.e., $\Pr(Z=z_i)=p_i$]. In thermodynamics, a similar expression is

$$K(Z) = -k_B \sum_i p_i \ln p_i, \tag{2}$$

where k_B is the Boltzmann constant.

The order-q Renyi entropy is defined as (Renyi, 1971)

$$K_q(P_\varepsilon) = \frac{1}{1-q} \ln \sum p_i^q, \tag{3}$$

where p_i is the fraction of the measure contained in the partition $P_i(\varepsilon)$ of side length $\leq \varepsilon$. The Shannon entropy, evaluated by the l'Hospital rule for $q=1$, is

$$K_1(P_\varepsilon) = -\sum_i p_i \ln p_i. \tag{4}$$

Block entropies of block size n are defined as

$$K_q(n, P_\varepsilon) = \frac{1}{1-q} \ln \sum_{i_1, i_2, \ldots, i_m} p_{i_1, i_2, \ldots, i_n}^q, \tag{5}$$

where $p_{i_1, i_2, \ldots, i_n}$ are the joint probabilities that at an arbitrary time n, the observable falls into the interval I_{i_1}, at time $n+1$, to the interval I_{i_2}, and so on, where I_t is the interval. Then, the order-q entropies are

$$k_q = \sup_P \lim_{n \to \infty} \frac{1}{n} K_q(n, \varepsilon)$$
$$= \sup_P \lim_{n \to \infty} (K_q(n+1, P_\varepsilon) - K_q(n, P_\varepsilon)), \tag{6}$$

where the supremum $\sup_P$ indicates that one has to maximize over all possible partitions P_ε and usually implies the limit $\varepsilon \to 0$.

In Eq. (6), k_1 (when $q=1$) is the KS entropy and k_0 (when $q=0$) is the T-Entropy. KS entropy is also sometimes called the metric entropy or simply Kolmogorov entropy. KS entropy has a numerical value of zero for nonchaotic systems and positive values for chaotic systems. It can also be thought of as the additional information obtained by observing the state of the system at a certain time given *a priori* knowledge of the entire past. Numerically, entropies are computed for a finite order q, which in the limit as $q \to \infty$ converges to the KS entropy. However, there are problems in estimating KS entropy by using Eq. (6), as it would require box counting which would be difficult in a high dimensional phase space. There are also difficulties arising from the limitations of the data set as it is nontrivial to obtain the limit process as $\varepsilon \to 0$.

III. CORRELATION SUM

Since it is difficult to calculate the probability density by counting boxes, a substitute method, which is called the relative integral method, is usually employed to estimate the dynamic entropy for a chaotic time series. Considering a chaotic time series $\{x_n\}$, $1 \leq n \leq N_s$, where N_s is the length of the time series embedded into a delay phase space with embedding dimension m and time delay τ, a point in the reconstructed phase space can be denoted by

$$X_i = (x_i, x_{i+\tau}, \ldots, x_{i+(m-1)\tau}). \tag{7}$$

The correlation sum of the time series is then given by

$$C_m(r) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j \neq i}^{N} H(r - \|X_i - X_j\|), \tag{8}$$

where X_i and X_j are two different points in the reconstructed phase space, N is the number of sample points in the reconstructed phase space, and $H(x)$ is the Heaviside function

$$H(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0. \end{cases}$$

In this study, $\|\bullet\|$ is the Euclidean norm.

The correlation sum of a chaotic time series is a very important entity. It can be used to estimate many related parameters. It is therefore necessary to find a "good" description of the correlation sum in order to estimate such parameters accurately. For a noise-free chaotic time series, Frank *et al.* (1993) proved that the correlation sum behaves as

$$C_m(r) = \phi[\exp(-m\tau K)](r/\sqrt{m})^D \quad \text{for} \ \ r \to 0, m \to \infty, \tag{9}$$

where ϕ is a constant, D is the correlation dimension, K is the KS entropy (order 2 entropy), m is the embedding dimension, and τ is the time delay.

For a chaotic time series with additive Gaussian noise, Diks (1999) (p. 123) and Oltmans and Verheijen (1997) proved that the correlation integral satisfies

$$C_m(r) = \frac{\phi e^{-m\tau K} m^{-D/2} \sigma^{D-m} 2^{-m} r^m}{\Gamma(m/2+1)}$$
$$\times M\left(\frac{m-D}{2}, \frac{m}{2}+1, -\frac{r^2}{4\sigma^2}\right), \tag{10}$$

where ϕ is a constant [same as in Eq. (9)], σ is the Gaussian noise level, and $M(a,b,z)$ is the Kummer's confluent hypergeometric function, which has the following integral representation:

$$M(a,b,z) = \frac{\Gamma(b)}{\Gamma(a)\Gamma(b-a)} \int_0^1 e^{zt} t^{a-1} (1-t)^{b-a-1} dt, \tag{11}$$

where

$$a = \frac{m-D}{2},$$

$$b = \frac{m}{2}+1, \tag{12}$$

$$z = -\frac{r^2}{4\sigma^2}.$$

The Kummer's confluent hypergeometric function also satisfies the following equation:

$$\frac{d}{dz} M(a,b,z) = \frac{a}{b} M(a+1, b+1, z), \tag{13}$$

where

$$a+1 = \frac{m-D}{2}+1 = \frac{(m+2)-D}{2}$$

and

$$b+1 = \frac{m}{2}+2 = \frac{m+2}{2}+1.$$

The correlation dimension and KS entropy, together with the noise level of the chaotic time series, can be estimated theoretically by applying Eqs. (9) and (10). Diks (1999) suggested that for fixed embedding dimension, Eq. (10) can be applied to obtain a nonlinear maximum likelihood estimate of the correlation dimension D and the noise level σ. Furthermore, for large embedding dimensions, the estimation of the KS entropy for noisy time series is given by (Diks, 1999)

$$K = \frac{1}{\tau} \ln\left[\frac{C_m(r)}{C_{m+1}(r)}\right] - \frac{1}{\tau} \ln\left[\frac{r}{2\sqrt{m}\sigma}\right]. \tag{14}$$

The estimation of the KS entropy by Eq. (14) strongly depends upon the estimation of the noise level, which is very difficult for noisy chaotic time series. Hence, the use of Eq. (14) cannot provide a better estimation of the KS entropy for a noisy chaotic time series.

Equation (14) suggests a relationship between the correlation sums $C_m(r)$ and $C_{m+2}(r)$, which will then be used to estimate the KS entropy for a noisy chaotic time series.

IV. MODIFIED CORRELATION ENTROPY

By comparing the correlation sum [Eq. (9)] for two different values of the embedding dimension, m and $m+2$, we have, when $r \approx 0$,

$$\frac{C_m(r)}{C_{m+2}(r)} = \exp(2\tau K)\left(1 + \frac{2}{m}\right)^{D/2}, \tag{15}$$

which can be written as

$$\ln \frac{C_m(r)}{C_{m+2}(r)} = 2\tau K + \ln \left(1 + \frac{2}{m} \right)^{D/2}. \qquad (16)$$

The CE K_2 as defined by Diks (1999) (p. 111) is

$$K_2 = \frac{1}{2\tau} \ln \left[\frac{C_m(r)}{C_{m+2}(r)} \right]. \qquad (17)$$

When the time series is noise-free, Eq. (15) holds, and Eqs. (16) and (17) then simplify to a relationship between the CE K_2 and the KS entropy K as

$$K_2 = K + \frac{1}{2\tau} \ln \left(\left[1 + \frac{2}{m} \right]^{D/2} \right), \qquad (18)$$

which converges to K as $m \to \infty$, i.e.,

$$\lim_{m \to \infty} K_2 = K.$$

For a noisy chaotic time series, by comparing the correlation sum [Eq. (10)] for two different values of the embedding dimension, m and $m+2$, we have

$$\frac{C_m(r)}{C_{m+2}(r)} = \frac{m^{-D/2}}{e^{-2\tau K}(m+2)^{-D/2}2^{-2}\sigma^{-2}r^2} \frac{\Gamma(m/2+2)}{\Gamma(m/2+1)} \frac{M(a,b,z)}{M(a+1,b+1,z)}$$

$$= \exp(2\tau K) \left(\frac{m+2}{m} \right)^{D/2} \frac{2\sigma^2(m-D)M(a,b,z)}{r^2 dM(a,b,z)/dz} = \exp(2\tau K) \left(\frac{m+2}{m} \right)^{D/2} \frac{2\sigma^2(m-D)}{r^2 d \ln[M(a,b,z)]/dz}, \qquad (19)$$

where $z = -r^2/(2\sigma^2)$, which has been defined in Eq. (12). Hence, we obtain

$$\exp(2\tau K) \left(\frac{m+2}{m} \right)^{D/2} (m-D)$$

$$= \frac{C_m(r)}{C_{m+2}(r)} \frac{r^2}{2\sigma^2} \frac{d \ln[M(a,b,z)]}{dz}. \qquad (20)$$

Taking natural logarithms for both sides of Eq. (10) gives

$$\ln(C_m(r)) = \ln(\phi e^{-m\tau K} m^{-D/2} 2^{-m} \sigma^{D-m}) + m \ln(r)$$
$$- \ln(\Gamma(m/2+1))$$
$$+ \ln \left(M \left(\frac{m-D}{2}, \frac{m}{2}+1, -\frac{r^2}{4\sigma^2} \right) \right). \qquad (21)$$

Differentiating Eq. (21) with respect to r, we obtain

$$\frac{d}{dr} \ln[C_m(r)] = \frac{m}{r} + \frac{d}{dz} \ln \left[M \left(\frac{m-D}{2}, \frac{m}{2}+1, -\frac{r^2}{4\sigma^2} \right) \right]$$
$$\times \left(-\frac{r}{2\sigma^2} \right),$$

from which we obtain

$$\frac{r^2}{2\sigma^2} \frac{d}{dz} \ln[M(a,b,z)] = m - r \frac{d}{dr} \ln[C_m(r)]. \qquad (22)$$

Substituting Eq. (22) into Eq. (20) gives

TABLE I. Statistics of data sets used in this study. Note: Under the noise level column, C means that the data set is clean; A/B indicates that A is the added noise level and B is the actual noise level; R indicates that the data set is raw and NR-1 indicates that the data set has been noise reduced by method 1, while NR-2 indicates that the data set has been noise reduced by the method 2. (The actual noise level $\sigma^2 = 1/N \Sigma_{n=1}^{N} (x_n - \bar{x}_n)^2$, where x_n and $\bar{x}_n$ are the clean and noisy time series, respectively.)

| Data	Length	Maximum	Minimum	Mean	Variance	Noise level
Lorenz A	5000	18.02	−18	0.3074	7.8977	C
Lorenz B	5000	18.21	−19	0.3091	7.911	0.5/0.5020
Lorenz C	5000	19.27	−20.12	0.3109	7.9561	1.0/1.004
Lorenz D	5000	20.33	−21.19	0.3127	8.0325	1.5/1.5060
Rössler A	5000	17.29	−14.59	0.13	8.2	C
Rössler B	5000	17.98	−15.28	0.132	8.21	0.5/0.5020
Rössler C	5000	18.97	−16.18	0.134	8.245	1.0/1.004
Rössler D	5000	20.17	−17.1	0.136	8.313	1.5/1.5060
Mekong A	4292	21 000	891	4302	2624.8	R
Mekong B	4292	19 447	951	4297	2609	NR-1
Mekong C	4292	19 091	903	4737	3177	NR-2
Chao A	5844	4320	71	627.4	470.5	R
Chao B	5844	4200	232	622	440	NR-1
Chao C	5844	3927	95	626.9	647	NR-2
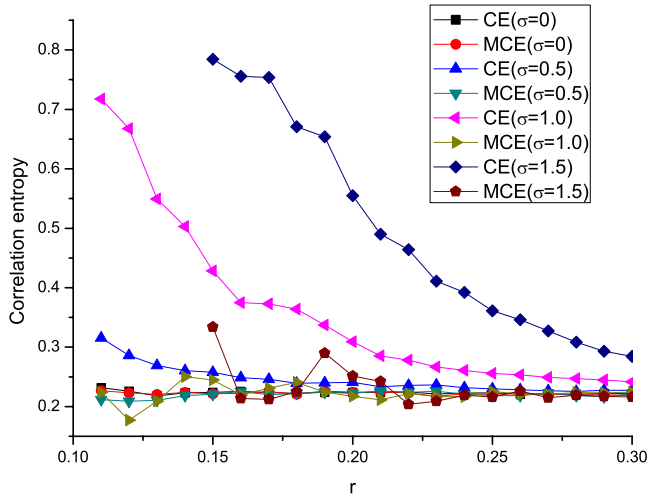
FIG. 1. (Color online) CE and MCE vs radius r for Lorenz data (embedding dimension $m=20$, time delay $\tau=1$ for $0.1<r<0.3$; σ indicates the noise level).

FIG. 3. (Color online) CE and MCE vs radius r for Mekong data (embedding dimension $m=20$, time delay $\tau=1$ for $0.1<r<0.3$; NR-1: noise reduced by method 1, NR-2: noise reduced by method 2).

$$\exp(2\tau K)\left(\frac{m+2}{m}\right)^{D/2}(m-D)$$
$$=\frac{C_m(r)}{C_{m+2}(r)}\left(m-r\frac{d}{dr}\ln[C_m(r)]\right). \tag{23a}$$

Taking natural logarithms for both sides of the above equation gives

$$2\tau K+\ln\left(\frac{m+2}{m}\right)^{D/2}=\ln\left(\frac{C_m(r)}{C_{m+2}(r)}\left(m-r\frac{d}{dr}\ln[C_m(r)]\right)\right)$$
$$-\ln(m-D). \tag{23b}$$

Note that the left-hand side of Eq. (23b) is just the right-hand side of Eq. (16). Since the right-hand side of Eq. (23b) does not depend on K, we can define a new CE for the noisy time series by
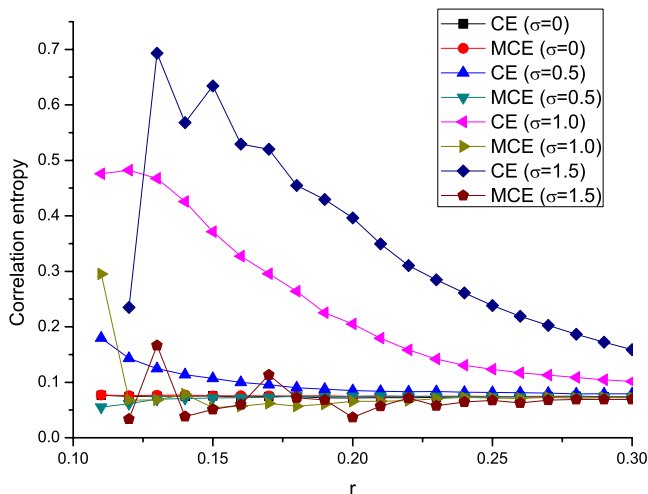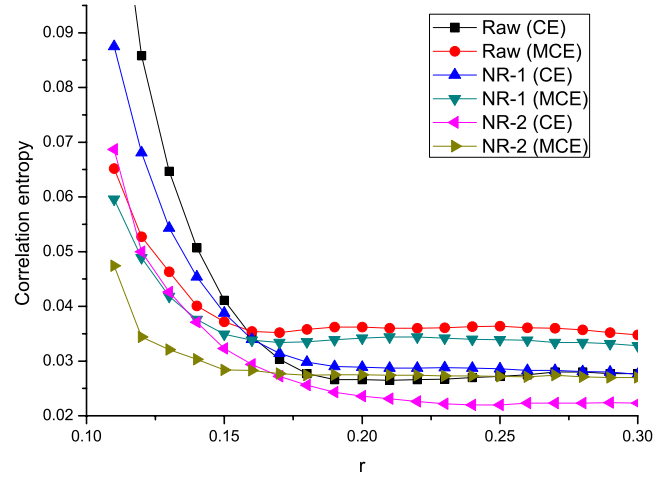
$$\bar{K}_2=\frac{1}{2\tau}\ln\left[\frac{C_m(r)}{C_{m+2}(r)}\right]$$
$$+\frac{1}{2\tau}\ln\left[\left(m-\frac{d\ln(C_m(r))}{d\ln r}\right)\Bigg/(m-D)\right]. \tag{24}$$

We call the CE estimated by Eq. (24) the "MCE" for a chaotic time series. When the chaotic time series is noise-free, by using Eq. (9), or using Eqs. (10) and (11) (Jayawardena *et al.*, 2008), it can be shown that

$$D=\frac{d\ln[C_m(r)]}{d\ln(r)}. \tag{25}$$

Equation (24) ensures that when the chaotic time series is noise-free, the MCE and the CE [Eq. (17)] are equivalent, i.e.,

FIG. 2. (Color online) CE and MCE vs radius r for Rössler data (embedding dimension $m=20$, time delay $\tau=1$ for $0.1<r<0.3$; σ indicates the noise level).
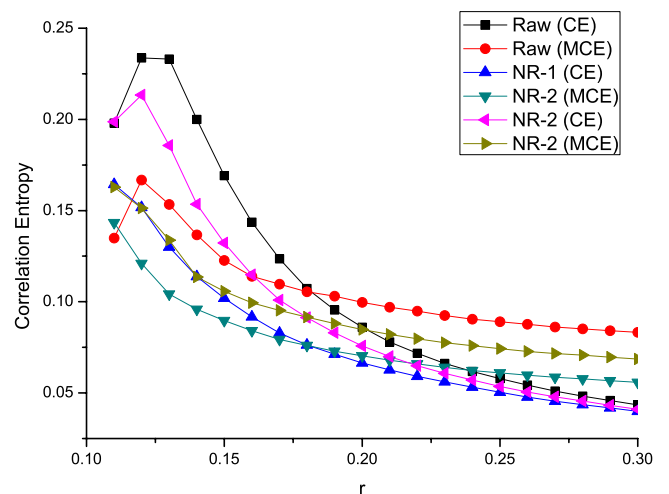

FIG. 4. (Color online) CE and MCE vs radius r for Chao Phraya data (embedding dimension $m=20$, time delay $\tau=1$ for $0.1<r<0.3$; NR-1: noise reduced by method 1, NR-2: noise reduced by method 2).
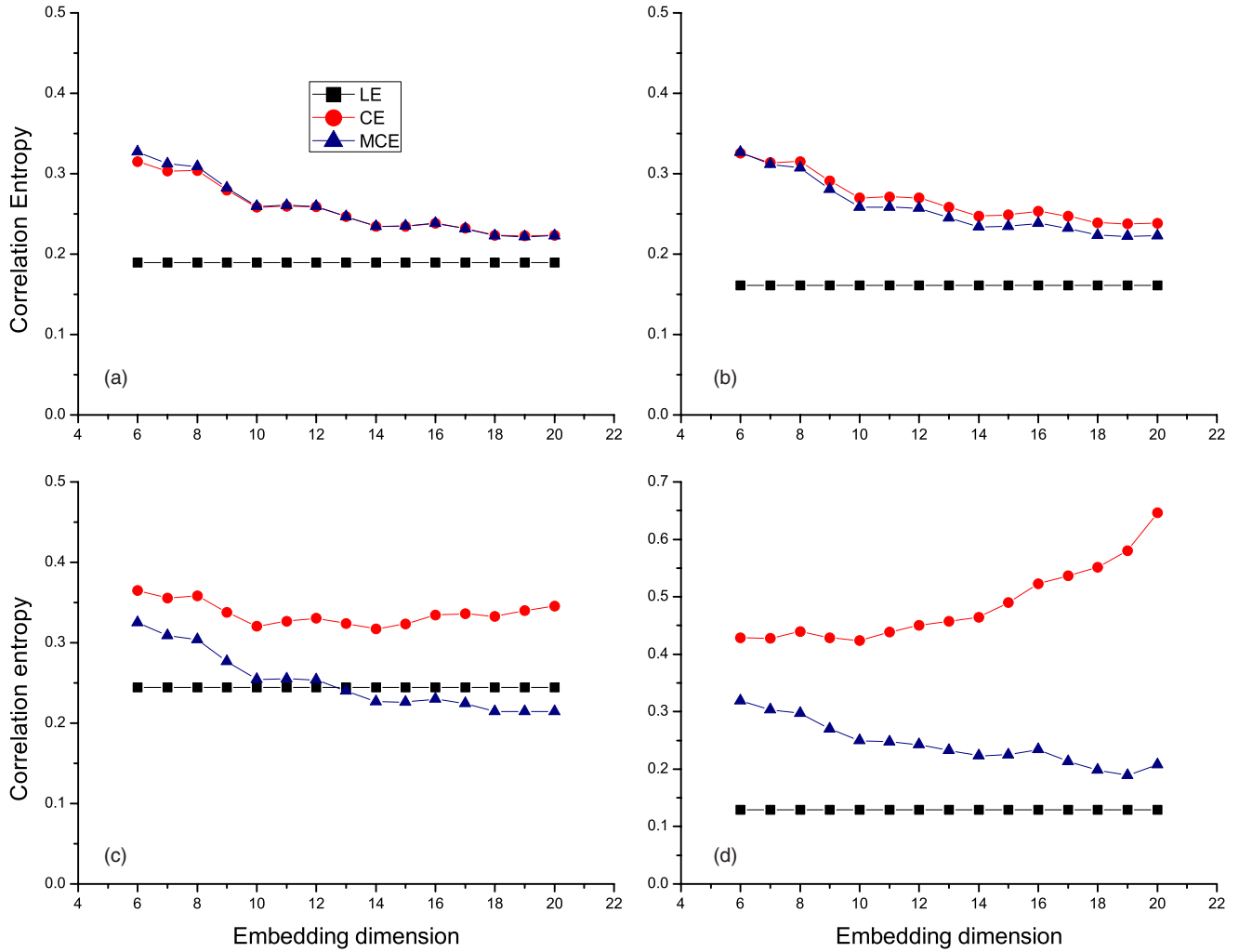
FIG. 5. (Color online) CE, MCE, and LE vs embedding dimension for Lorenz data with four different noise levels (a for clean data; b for σ=0.5; c for σ =1.0; and d for σ=1.5).

$$\bar{K}_2 = K_2. \tag{26}$$

The estimation of the MCE by Eq. (24) needs the correlation dimension of the chaotic time series, which can be given *a priori*, or be estimated using the correlation sum. In the former case it can be given approximately as

$$D \approx m_d - 1, \tag{27}$$

where m_d is the minimum embedding dimension of the chaotic time series, which can be obtained by the false nearest neighbor (FNN) method (Kantz and Schreiber, 2003). In the latter case, it can be estimated by employing the correlation sum using Eq. (8). Given several values $r_1, r_2, \ldots, r_L$, the corresponding correlation sums are then obtained as $C_m(r_1), C_m(r_2), \ldots, C_m(r_L)$. Then the correlation dimension is approximated by

$$D \approx \frac{1}{L} \sum_{i=1}^{L} \frac{d \ln[C_m(r_i)]}{d \ln(r_i)}. \tag{28}$$

In this study, the first method [Eq. (27)] is used.

Note that the estimation process for the MCE does not involve the noise level of the noisy chaotic time series.

Hence, the above estimation method would be easier for a noisy chaotic time series than the estimation method based on Eq. (14). Furthermore, the MCE is more robust to noise than the general CE obtained by Eq. (18).

V. KS ENTROPY AND THE LYAPUNOV SPECTRUM

In a dynamical system, the Lyapunov exponent characterizes the rate of separation of initially nearby trajectories. The rate of separation can be different for different initial separations, thus leading to a whole spectrum of Lyapunov exponents, with the number of such exponents being equal to the dimension of the phase space. The largest of them determines the predictability of the system and a positive largest exponent gives an indication that the system is chaotic. For an evolutionary system, the spectrum of Lyapunov exponents $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ depends upon the starting point x_0 and is defined from the Jacobian matrix as follows:

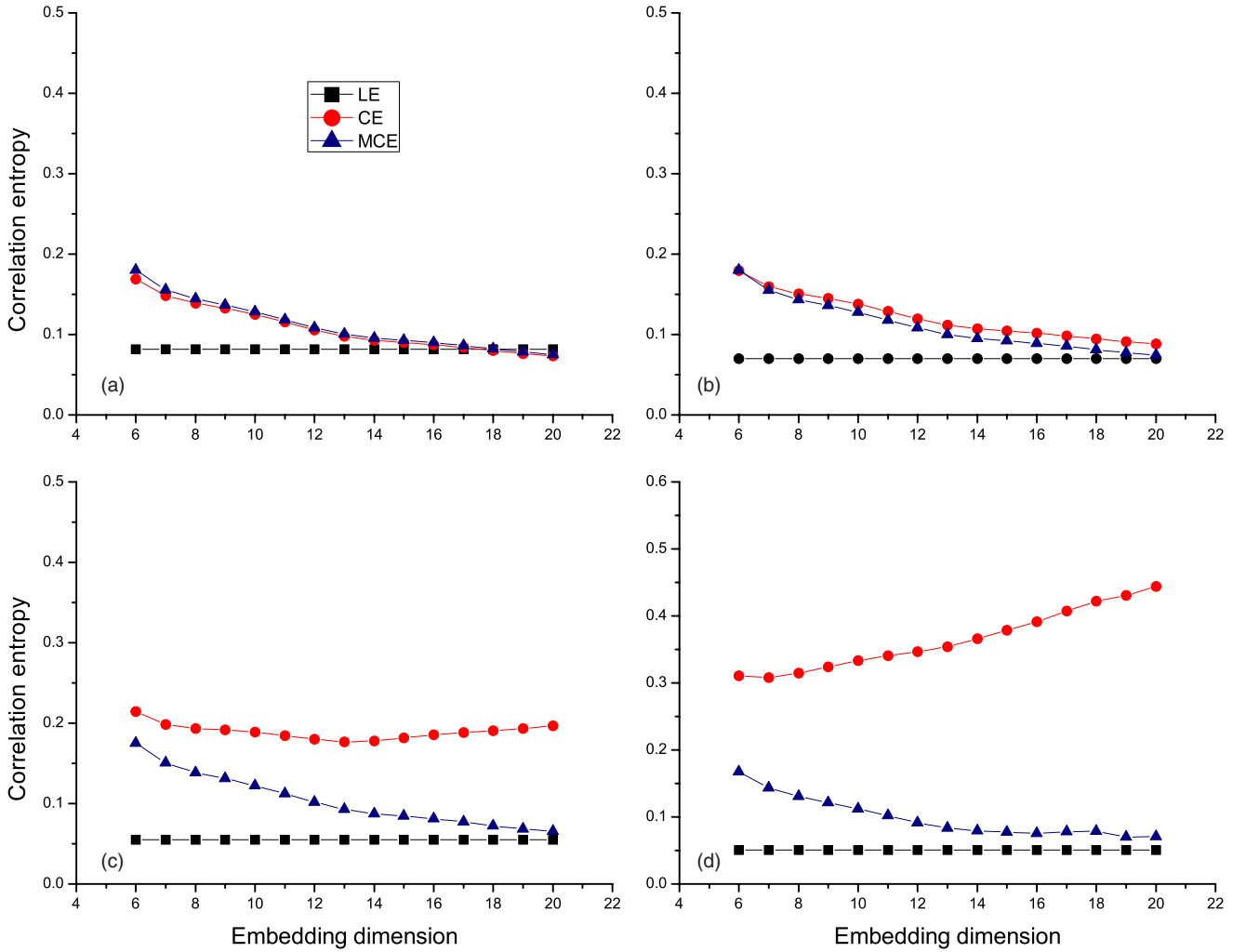$$\lambda_j = \lim_{N \to \infty} \frac{1}{N} \|J[F^{(N)}(x_0)]e_j\|, \tag{29}$$

FIG. 6. (Color online) CE, MCE, and LE vs embedding dimension for Rössler data with four different noise levels (a for clean data; b for $\sigma=0.5$; c for $\sigma=1.0$; and d for $\sigma=1.5$).

$$J[F^{(N)}(x_0)] = \left. \frac{dF^{(N)}(x)}{dx} \right|_{x_0}, \qquad (30)$$

where $J(F)$ is the Jacobian matrix of the nonlinear function F, $F^{(n)}$ is the nth iteration of the nonlinear evolution function, and $\{e_j\}, j=1,2,\ldots,m$ is a set of orthogonal basis of the tangent space at $\mathbf{x}_0$. The evolution function relates the future value to the present value according to the equation

$$\mathbf{x}_{n+1} = F(\mathbf{x}_n), \qquad (31)$$

where $\{\mathbf{x}_n\}_{n=1}^N$ is the trajectory in the reconstructed phase space of the chaotic time series.

For a conservative system, the sum of all Lyapunov exponents is zero. For a dissipative system, it is negative. According to the theorem of Pesin (1977), the sum of all positive Lyapunov exponents gives an estimate of the KS entropy. The Lyapunov spectrum estimated as described below is used to estimate the KS entropy to compare with the MCE.

The Jacobian of Eqs. (29) and (30), by the chain rule, can be written as

$$\begin{aligned}
JF^{(N)}(x_0) &= [JF(x_{N-1})] \cdot [JF^{(N-1)}(x_0)] \\
&= [JF(x_{N-1})] \cdot [JF(x_{N-2})] \cdot [JF^{(N-2)}(x_0)] = \cdots \\
&= [JF(x_{N-1})] \cdot [JF(x_{N-2})] \cdots [JF(x_1)] \cdot [JF(x_0)] \\
&= A_n \cdot A_{n-1} \cdots A_2 \cdot A_1, \qquad (32)
\end{aligned}$$

where $A_n = \mathbf{J}F(x_{n-1}), 1 \leq n \leq N$. Using the QR decomposition, matrix $\mathbf{A}$ can be decomposed into two matrices: an orthogonal matrix $\mathbf{Q}$ and an upper triangular matrix $\mathbf{R}$. We define matrices $\mathbf{Q}_n, \mathbf{R}_n$ via

$$\mathbf{A}_{n+1}\mathbf{Q}_n = \mathbf{Q}_{n+1}\mathbf{R}_{n+1}, \qquad (33)$$

where $\mathbf{Q}_0 = I$, the identity matrix. Then, the Lyapunov spectrum of the system is given by
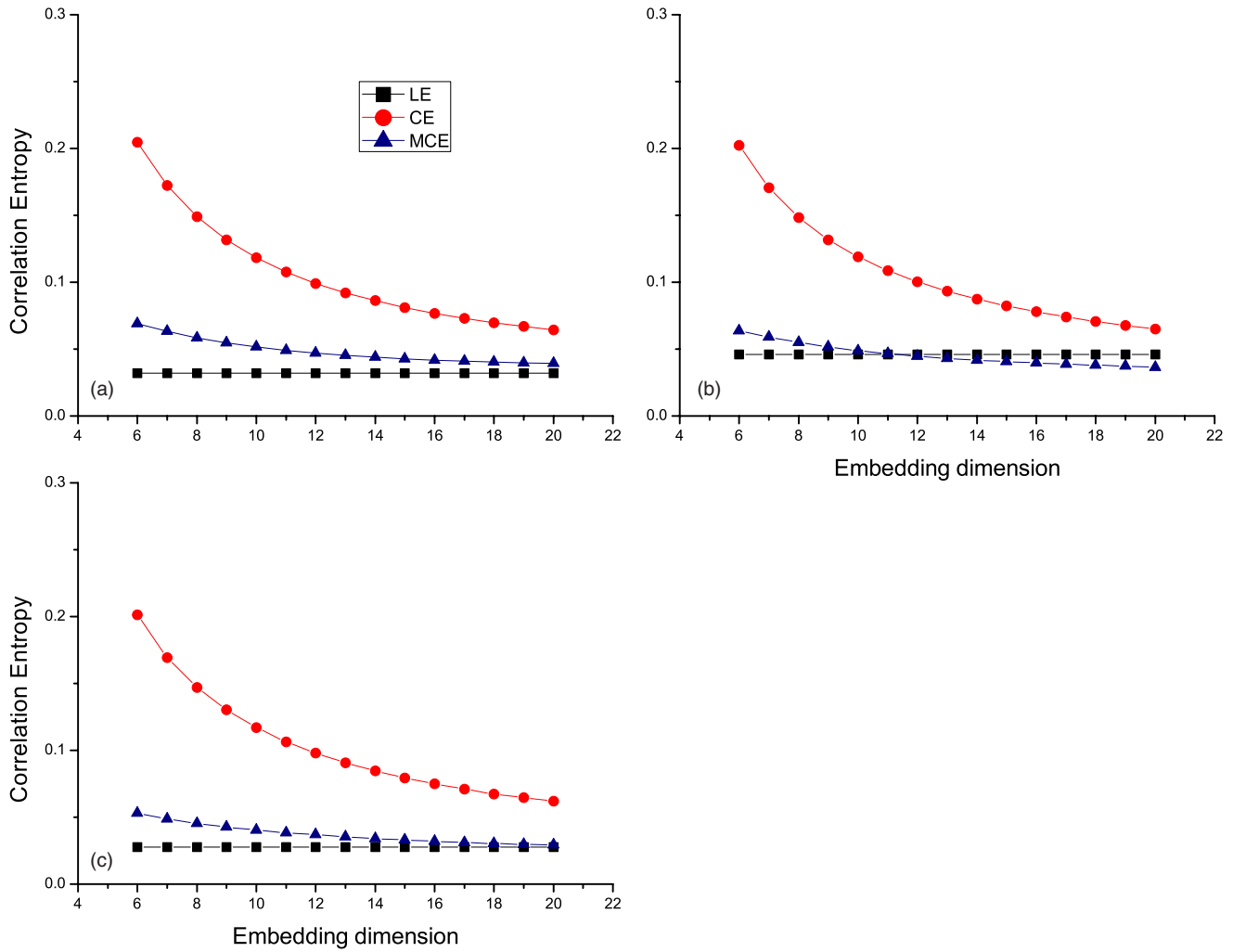
FIG. 7. (Color online) CE, MCE, and LE vs embedding dimension for Mekong data with three different noise levels (a for raw data; b for noise reduced by method 1; c for noise reduced by method 2).

$$\lambda_j = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \ln(\|R_n(j,j)\|), \qquad (34)$$

where $R_n(j,j)$ is the element of matrix $\mathbf{R}_n$ at jth row and jth column. The Jacobian matrix at each time point in the trajectory of the nonlinear system is evaluated by employing the least-squares method to some nearest neighbors of the corresponding points in the reconstructed phase space. The above procedure follows that developed by Brown *et al.* (1991).

VI. APPLICATION

The proposed method is then applied to four data sets to estimate the MCE: two artificial chaotic time series and two real-world chaotic time series. The artificial chaotic times series are generated by the Lorenz and Rössler equations. The Lorenz map is (Lorenz, 1963) defined by the equations

$$\begin{cases} \dot{x} = & \sigma(y - x) \\ \dot{y} = & -xz + rx - y \\ \dot{z} = & xy - bz, \end{cases} \qquad (35)$$

and it becomes chaotic for $\sigma = 10$, $r = 28$, and $b = 8/3$. Equation (35) is a three dimensional differential system, which can be solved numerically using the fourth order Runge–Kutta method with a time step of 0.005 and initial conditions of

$$x(0) = 12.5, y(0) = 2.5, z(0) = 1.5.$$

The values of the x-coordinate are then recorded as the realized time series. To ensure that the values of the time series are in the chaotic attractor, the first 4000 values are discarded. Then, a time series of length 5000 is generated for Eq. (35) by

$$x_n = x(100 + 0.2n), \quad 1 \le n \le 5000.$$

Gaussian noise with noise levels (σ) 0.5, 1.0, and 1.5 are then added to the series.

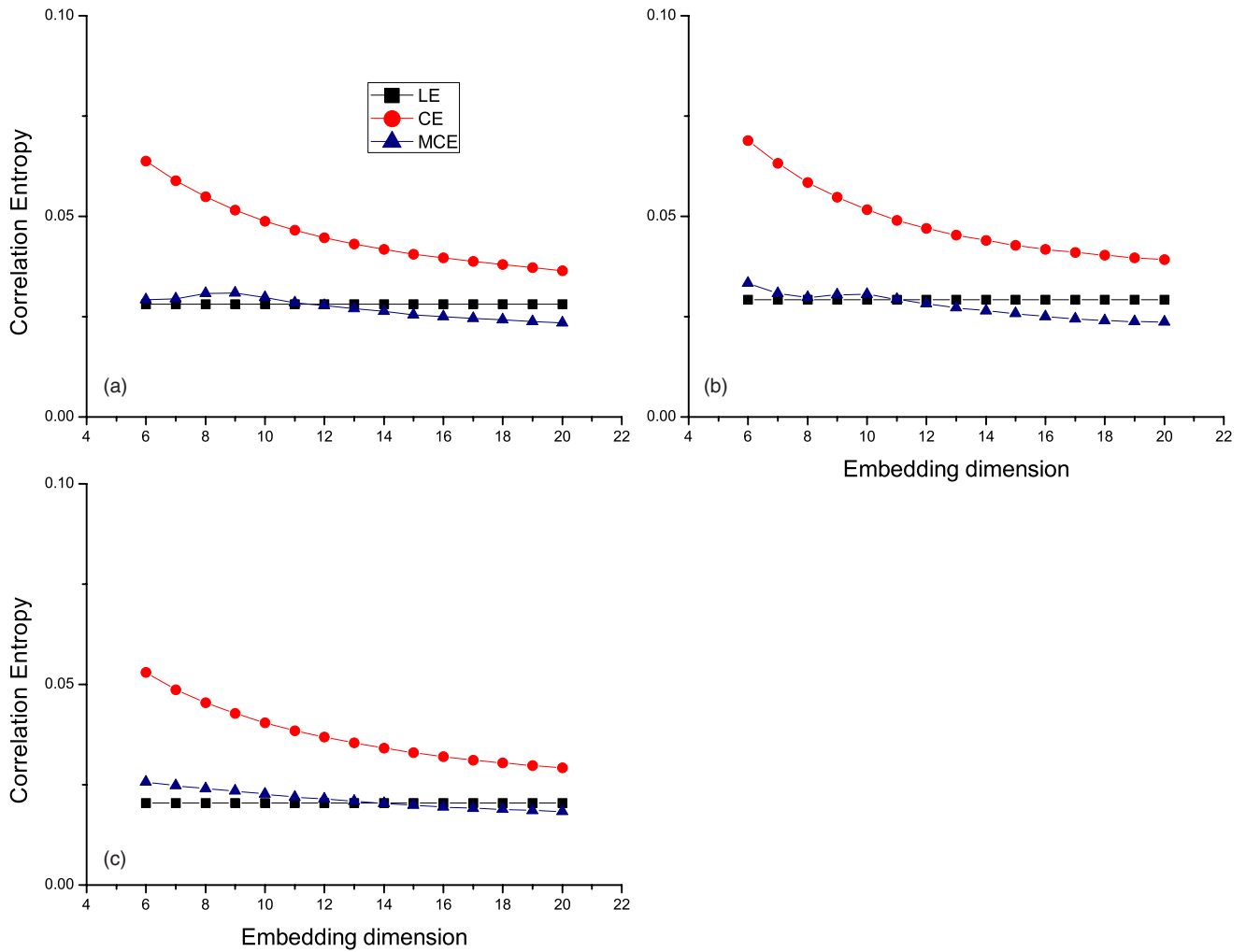The Rössler system is defined by the equations (Rössler, 1976)

FIG. 8. (Color online) CE, MCE, and LE vs embedding dimension for Chao Phraya data with three different noise levels (a for raw data; b for noise reduced by method 1; c for noise reduced by method 2).

$$\begin{cases} \dot{x} = & -y - z \\ \dot{y} = & x + ay \\ \dot{z} = & b - z(x - c), \end{cases} \tag{36}$$

and it becomes chaotic for $a=0.15$, $b=0.2$, and $c=10$. Equation (36) is also a three dimensional differential system which can be solved numerically using the fourth order Runge–Kutta method with a time step of 0.02 and initial conditions

$$x(0) = 0.1, y(0) = 0.1, z(0) = 0.1.$$

The values of the x-coordinate are then recorded as the realized time series. To ensure that the values of the time series are in the chaotic attractor, the first 5000 points are discarded. Then, a time series of length 5000 is generated for Eq. (36) by

$$x_n = x(100 + 0.5n), \quad 1 \le n \le 5000.$$

Gaussian noise with noise levels (σ) 0.5, 1.0, and 1.5 are then added to the series.

The two real-world time series come from the field of hydrology. They are the flow measurements made at two gauging stations across two major rivers in Asia, namely, the Mekong, which is a transboundary river that runs through six countries (China, Myanmar, Lao, Thailand, Cambodia, and Vietnam), and the Chao Phraya that runs through Thailand. For Mekong, daily discharges measured at Nong Khai (location: 17.87° North, 102.72° East; basin area: 302 000 km^2; GRDC Reference No. 2969090) for the period of April 1980–December 1991, and for Chao Phraya, daily discharges measured at Nakhon Sawan (location: 15.67° North, 100.2° East; basin area: 110 569 km^2; GRDC Reference No. 2964100) for the period of April 1978–March 1994 were used in the study.

Two noise reduction methods, described elsewhere (Jayawardena and Gurung, 2000), are applied to the two raw data sets. The first noise removal method (noise reduced by method 1) follows that of Grassberger *et al.* (1993), while the second method (noise reduced by method 2) follows that of

Schreiber and Grassberger (1991). Some statistics of these four data sets are given in Table I. However, for convenience of selecting the correlation radius r in calculating the correlation sum using Eq. (8), all of the above data sets have been rescaled into the interval [0,1].

VII. RESULTS AND DISCUSSION

In this study, a new approximating quantity for the KS entropy, the MCE, is introduced for a chaotic time series. This quantity is more robust to noise than the general CE and is estimated by employing the correlation sum. Hence, the estimation process strongly depends on the calculation of the correlation sum. The correlation sum depends on several parameters: the amount of the sample points N, the embedding dimension m, the time delay τ, and the correlation radius r. The effects of these parameters on the estimation of the correlation sum are discussed by Ramsey and Yuan (1990), and they note that the sample size of the chaotic time series should not be smaller than 2000. However, they also pointed out that when the embedding dimension is large, the sample size should also be large to ensure that the correlation sum results in a small bias. In this study, 5000 sample points each for Lorenz and Rössler data, 4292 sample points for the Mekong data, and 5844 sample points for the Chao Phraya data were used to estimate their respective correlation sums.

The embedding dimension for the Lorenz data and the Rössler data is 3, since they are generated from a three dimensional chaotic attractor. Based on a previous study (Jayawardena *et al.*, 2002) that uses the false nearest neighborhood method, each of the two real-world data sets has embedding dimension of 3. By Takens' embedding theorem (Takens, 1981), the time delay for a chaotic time series can be chosen arbitrarily. In this study, it is set as 1.

The correlation radius r is also important for the correlation sum calculation. Too small values of r would lead to no neighbors for a point while a large value of r would miss the microstructure of the chaotic attractor. In this study, the correlation sums were calculated for $0.1 < r < 0.3$ using rescaled data.

The numerical results of the comparison between the CE and the MCE, both calculated by using correlation sums, for all the data sets used in this study are shown in Figs. 1–4. In the simulation, the embedding dimension is 20 and the time delay is 1 for all the data sets. Figure 1 shows the results for the Lorenz data for four noise levels: clean data and noisy data with noise levels 0.5, 1.0, and 1.5. They show that the curves for the MCE and the CE are almost identical for the clean Lorenz data set. For noisy data, the curves of the MCE are still close to the curve for the clean data, while the curves of the CE are far apart from the curve of clean data especially for small values of r. These results demonstrate that the MCE is more robust to noise than the CE. A similar conclusion can be made for the Rössler data as well, which are shown in Fig. 2. The results for the Mekong and Chao Phraya data, which are shown in Figs. 3 and 4, also indicate the same pattern. It is also seen that the MCE is more stable and that it varies in a smaller range compared with the CE.

Equations (18) and (24) show that the MCE and the CE converge to the KS entropy of the system when the embedding dimension is large enough. For comparison, the KS entropy is also estimated by the Lyapunov spectrum method, as described in Sec. V. This is referred to as Lyapunov entropy in this context. Figures 5 and 6 show the variation in the different entropies [CE, MCE, and Lyapunov entropy (LE)] as functions of the embedding dimension for different noise levels for the Lorenz and Rössler data. The corresponding results for the Mekong and Chao Phraya data are shown in Figs. 7 and 8. All the four data sets show that the MCEs are closer to the LE than the CE especially for large value of the embedding dimension. The three sets of graphs for the Mekong and Chao Phraya data correspond to raw data, noise reduced-1, and noise reduced-2.

VIII. CONCLUSION

In this study, a new method of estimating the KS entropy for a chaotic time series is presented. It is applicable to clean as well as noisy data. Both theoretical analysis and numerical results, as demonstrated by the application of the method to Lorenz data, Rössler data, Mekong data, and Chao Phraya data, show that the MCE is more robust to noise than the CE. The numerical results show that the MCE is closer to the KS entropy for both noisy and noise-free time series data.

APPENDIX A: ALGORITHM 1 (FOR MODIFIED CORRELATION DIMENSION)

(1) Re-scale the time series $\{x_n\}_{n=1}^{N}$ to [0, 1], and obtain a new time series, $\{\bar{x}_n\}_{n=1}^{N}$.

(2) Estimate the minimum embedding dimension m_d for the re-scaled time series by FNN method.

(3) Given an embedding dimension m ($m \geq m_d$) and a time delay τ, embed the chaotic time series into the delay phase space. Points in the delay phase space are $\{\mathbf{X}_n\}, 1 \leq n \leq N-(m-1)\tau$.

(4) Give several values of r, $0.1 = r_1 < r_2 < \cdots r_i < \cdots r_L = 0.3$, calculate the correlation sum $C_m(r_i)$ for the time series by Eq. (8).

(5) Increase the embedding dimension to $m+2$, embed the time series $\{\bar{x}_n\}_{n=1}^{N}$ into a new delay phase space. Points in the new delay phase space are $\{\mathbf{Y}_n\}, 1 \leq n \leq N-(m+2-1)\tau$.

(6) Give several values of r, $0.1 = r_1 < r_2 < \cdots r_i < \cdots r_L = 0.3$, calculate the correlation sum $C_{m+2}(r_i)$ for the time series by Eq. (8),

(7) Calculate $d \ln(C_m(r))/d \ln(r)$ by

$$\frac{d \ln(C_m(r_i))}{d \ln(r_i)} = \frac{\ln(C_m(r_i)) - \ln(C_m(r_{i-1}))}{\ln(r_i) - \ln(r_{i-1})}, \quad 1 < i \leq L$$

(8) Calculate

$$K_2(i) = \frac{1}{2\tau} \ln \frac{C_{m+2}(r_i)}{C_m(r_i)}, \quad 1 \leq i \leq L$$

(9) Calculate

$$\bar{K}_2(i) = \frac{1}{2\tau} \ln \frac{C_{m+2}(r_i)}{C_m(r_i)} + \frac{1}{2\tau} \ln \left(\frac{m - d \ln(C_m(r_i))/d \ln(r_i)}{m - m_d + 1} \right),$$

$$1 < i \leq L$$

APPENDIX B: ALGORITHM 2 (FOR LYAPUNOV SPECTRUM)

(1) Given a trajectory $\{x_n\}, 1 \le n \le N$.
(2) Let $Q_0 = I$, identity matrix.
(3) For every $\mathbf{x}_n$, find its N_B (assumed to be 20 for all data sets) nearest neighbors in the reconstructed phase space, $\{\mathbf{x}_n^k\}, 1 \le k \le N_B$.
(4) Estimate parameters for local linear model

$$x_{n+m\tau} = a + \sum_{i=1}^{m} b_i x_{n+(i-1)\tau}$$

by the least square method using the N_B nearest neighbors.
(5) Set Matrix

$$\mathbf{A}_n = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_1 & b_2 & b_3 & \cdots & b_m \end{bmatrix}$$

(6) QR decomposition: $\mathbf{Q}_k \mathbf{R}_k = \mathbf{A}_k \mathbf{Q}_{k-1}$
(7) $\lambda_j = 1/N \Sigma \ln(R_n(j,j)), \quad j = 1, 2, \ldots \ldots N$

Adler, R. L., Konheim, A. G., and McAndrew, M. H., "Topological entropy," Trans. Am. Math. Soc. **114**, 309–319 (1965).

Bonachela, J. A., Hinrichsen, H., and Munoz, M., "Entropy estimations of small data set," J. Phys. A: Math. Theor. **41**, 202001 (2008).

Brown, R., Bryant, P., and Abarbanel, H. D., "Computing the Lyapunov spectrum of a dynamical system from an observed time series," Phys. Rev. A **43**, 2787–2806 (1991).

Diks, C., Nonlinear Time Series Analysis, Methods and Applications (World Scientific, Singapore, 1999).

Frank, M., Blank, H.-R., Heindl, J., Kaltenhauser, M., Kochner, H., Kreissche, W., Muller, N., Pocher, S., Sporer, R., and Wagner, T., "Improvement of K_2 entropy calculations by means of dimension scaled distances," Physica D **65**, 359–364 (1993).

Grassberger, P., Hegger, R., Kantz, H., Schaffrath, C., and Schreiber, T., "On noise reduction methods for chaotic data," Chaos **3**, 127–141 (1993).

Grassberger, P. and Procaccia, I., "Estimation of the Kolmogorov entropy from a chaotic signal," Phys. Rev. A **28**, 2591–2593 (1983).

Jayawardena, A. W. and Gurung, A. B., "Noise reduction and prediction of hydrometeorological time series: Dynamical systems approach vs. stochastic approach," J. Hydrol. **228**, 242 (2000).

Jayawardena, A. W., Li, W. K., and Xu, P., "Neighbourhood selection for local modeling and prediction of hydrological time series," J. Hydrol. **258**, 40–57 (2002).

Jayawardena, A. W., Xu, P. C., and Li, W. K., "A method of estimating the noise level in a chaotic time series," Chaos **18**, 023115 (2008).

Kantz, H. and Schreiber, T., Nonlinear Time Series Analysis (Cambridge University Press, Cambridge, England, 2003).

Kantz, H. and Schürmann, T., "Enlarger scaling ranges for the KS-entropy and information dimension," Chaos **6**, 167–171 (1996).

Kolmogorov, A. N., "A new invariant of transitive dynamical systems," Dokl. Akad. Nauk SSSR **119**, 861–864 (1958).

Lorenz, E. N., "Deterministic non-periodic flow," J. Atmos. Sci. **20**, 130–141 (1963).

Oltmans, H. and Verheijen, P. J. T., "Influence of noise on power-law scaling functions and an algorithm for dimension estimations," Phys. Rev. E **56**, 1160–1170 (1997).

Pesin, Y. B., "Characteristic Lyapunov exponents and smooth ergodic theory," Russ. Math. Surveys **32**, 55–112 (1977).

Ramsey, J. B. and Yuan, H. J., "The statistical properties of dimension calculations using small data sets," Nonlinearity **3**, 155–176 (1990).

Renyi, A., Probability Theory (North-Holland, Amsterdam, 1971).

Rössler, O. E., "An equation for continuous chaos," Phys. Lett. **57A**, 397–398 (1976).

Schreiber, T. and Grassberger, P., "A simple noise-reduction method for real data," Phys. Lett. A **160**, 411–418 (1991).

Schürmann, T. and Grassberger, P., "Entropy estimation of symbol sequences," Chaos **6**, 414–427 (1996).

Shannon, C. E., "A mathematical theory of communication," Bell Syst. Tech. J. **27**, 379–423 and 623–656 (1948).

Sinai, A. G., "On the concept of entropy of a dynamical system," Dokl. Akad. Nauk SSSR **142**, 768–776 (1959).

Takens, F., "Detecting strange attractors in turbulence," Lect. Notes Math. **898**, 366–381 (1981).

Weaver, W. and Shannon, C. E., The Mathematical Theory of Communication (University of Illinois Press, Urbana, IL, 1949).