Trends in Hydrology, 1 (1994)

Dynamical systems theory applied to long-term temperature and precipitation time series

Ronny Berndtsson*, Kenji Jinno, Akira Kawamura**, Jonas Olsson* and Shiguo Xu****
*Department of Water Resources Engineering, Lund University, Lund, Sweden, **Department of Civil Engineering (SUIKO),Kyushu University, Fukuoka, Japan

ABSTRACT

We investigate 238-year monthly series of sunspots, temperature, and precipitation by dynamical systems theory. Raw time series of these variables do not show any chaotic deterministic properties. However, after noise reduction, all three variables display a low-dimensional chaotic behavior. Thus, we view this as indicative, though not conclusive, of chaos. The results can be used to better understand sudden jumps and changes in climatological data. We delineate a prediction technique that can be used to combine both deterministic and stochastic components of the time series.

INTRODUCTION

The recent progress in dissipative nonlinear dynamical systems which display chaotic behavior has framed a new view on time series analyses [e.g., *Schuster*, 1984]. While white or colored noise in many ways may be indistinguishable from chaotic trajectories [e.g., *Osborne and Provenzale*, 1989; *Provenzale et al.*, 1991], observed time series may at least tentatively be regarded as deterministically chaotic if prediction methods based on underlying deterministic properties are significantly better as compared to autoregressive linear models based on stochastic theory [see further *Farmer and Sidorowitch*, 1987; 1988; *Sugihara and May*, 1990; *Mundt et al.*, 1991].

There appears to be no universal method at present to discriminate between colored noise with power-law spectra and underlying dynamical processes in data [e.g., *Theiler et al.*, 1992]. This is especially true for finite and noisy experimental data. Instead, several methods have been delineated to make probable or at least indicate the presence of deterministic chaos in observed time series [*Provenzale et al.*, 1992; *Theiler et al.*, 1992]. Therefore, we prefer to view results in this paper as indicative of chaos only and leave strict definitions of underlying chaotic behavior for future analyses.

As indicated above, our intention of this paper is not to prove the existence of nonlinear dynamical systems in climatological data or not. Instead, we aim at employing any apparent deterministic property to make short-term predictions. In recent years, a number of studies have used techniques to exploit chaotic dynamics to provide improved forecasting capabilities [*Farmer and Sidorowich*, 1987; *Sugihara and May*, 1990; *Mundt et al.*, 1991; *Elsner and Tsonis*, 1992; *Tsonis and Elsner*, 1990]. Climatological data are highly variable and unpredictable in the long term. This is caused by complex interactions with many independent and irreducible degrees of freedom. Instead of solving a set of partial differential equations for atmospheric flow, an alternative approach is to build models directly from available observations [*Elsner and Tsonis*, 1992].

An appealing property of a class of low-dynamical dissipative systems is that its existence can provide knowledge of the number of equations needed to describe the system. The trajectories of these systems are not confined to periodic or quasi-periodic cycles but instead to aperiodic, never repeatable, however, completely deterministic evolutions in time. Because of the aperiodic behavior, these systems exhibit frequency spectra that are similar to those of colored noise (broadband spectra). These systems, are extremely sensitive to initial conditions and thus they are called chaotic. Although the chaotic system is unpredictable in the long term, the system's settling on a fractal trajectory (strange attractor) may be used for short-term predictions.

In this paper, we analyze observed monthly time series of sunspots, temperature and precipitation by dynamical systems theory. The analyses on sunspot data are made to compare the applied techniques with results from other studies (e.g., *Mundt et al.*, 1991). The objective is to show that at least parts of the observed time series can be viewed as deterministically chaotic. We delineate methods for possible short-term predictions.

MATERIALS AND METHODS

Theory

Lorenz [1963] was the first to display possible chaotic properties of the atmosphere. He showed that a dynamic system may be described by:

$$\frac{dx}{dt} = F[x] \qquad (1)$$

where t denotes time that is the only dependent variable. The vector $x = (x_1, x_2, x_3, \dots, x_n)$ represents a state of the system and a set of n ordinary differential equations and can be thought of as points along a time axis in phase space where the vector $F(x)$ is a nonlinear operator acting on x. For some initial conditions the vector x can be shown to have a chaotic evolution, i.e., x approaches a strange attractor. At small changes of the initial conditions, x will have a very different evolution.

Equation (1) can be expressed as a single nonlinear differential equation according to:

$$x^{(n)} = f[x, x', \dots, x^{(n-1)}] \qquad (2)$$

This is in turn equivalent to:

$$x(t) = [x(t), x'(t), \dots, x^{(n-1)}(t)] \qquad (3)$$

In the climatological reality, however, $F(x)$ and initial conditions are unknown. Instead, one often has observations of $x(t)$, e.g., temperature, precipitation records, etc. According to a theorem by *Takens* [1981, see also *Ruelle*, 1981 and *Packard et al.*, 1980] it is possible to use the observations $x(t)$ to evaluate the dimension of the attractor.

The general procedure to evaluate the attractor dimension is to perform a phase space (sometimes called state space) reconstruction. The basic idea behind a phase space reconstruction is that the past and future of the time series contain information about unobserved state variables that may be used to define a state at the present time [*Casdagli et al.*, 1991]. The procedure of phase space reconstruction is motivated due to unknown properties of the dynamical system such as relevant variables and their total number. Phase space reconstruction was introduced in dynamical systems by *Packard et al.* [1980], *Ruelle* [1981], and *Takens* [1981], even though the basic idea goes as long back as *Yule* [1927].

For deriving the dimension d of the attractor from observations $x(t)$ it is sufficient to embed it in an m-dimensional space ($d < m, n$):

$$x(t) = [x(t), x'(t), \dots, x^{(m-1)}(t)] \qquad (4)$$

Consequently, it is not necessary to know the original system's dimension n or state variables as long as m is chosen large enough [$m = 2d + 1$; *Takens*, 1981]. According to this and introducing a time lag τ one gets [*Grassberger and Procaccia*, 1983a, 1983b]:

$$x(t),\ x(t+\tau),\ x(t+2\tau), \dots, x(t+(m-1)\tau) \qquad (5)$$

Following Eq. (5), new time series are generated according to:

$$\begin{aligned}&x(t_1),\ x(t_2), \dots, x(t_N)\\&x(t_1+\tau),\ x(t_2+\tau), \dots, x(t_N+\tau)\\&x(t_1+(m-1)\tau),\ x(t_2+(m-1)\tau) \dots, x(t_N+(m-1)\tau)\end{aligned} \qquad (6)$$

where N is a set of points on the attractor embedded in the m-dimensional phase-space. In the vector x with the coordinates $\langle x(t_1), \dots, x(t_i + (m-1)\tau\rangle$, a point can be chosen x_i so that all distances $|x_i - x_j|$ for m-1 points can be calculated. By repeating this for all i one gets:

$$C(r) = \frac{1}{N^2} \sum_{\substack{i,j=1\\ i \neq j}}^{N} \theta(r - |x_i - x_j|) \qquad (7)$$

where θ is the Heaviside function defined by $\theta(x) =$

0 if $x < 0$ and $\theta(x) = 1$ if $x > 0$. The entity $C(r)$ is called the correlation integral for the strange attractor and defines the density of points around a specific coordinate x_i. In this paper, the algorithm according to *Grassberger* [1990] was used to estimate correlation integrals. The correlation integral $C(r)$ is used to describe the dimension d of the attractor, i.e., if the attractor is a line, surface or volume. If the attractor can be described by a line one expects that the number of points within a distance r from a coordinate is proportional to r/ϵ, where ϵ is a point in the middle of the attractor. If, on the other hand, the attractor is a surface $C(r)$ is proportional to $(r/\epsilon)^2$, and similarly if the attractor is a volume $C(r)$ should be proportional to $(r/\epsilon)^3$. Consequently, we find that for small r, $C(r)$ should relate as:

$$C(r) \sim r^d \tag{8}$$

Values of d that are not integers indicate a fractal and thus chaotic attractor. The dimension d of the attractor is given by the slope of log $C(r)$ for the slope of log r according to:

$$\log C(r) = d|\log r| \tag{9}$$

The dimension of the attractor indicates according to above how many variables that are necessary to describe the evolution in time. For example, if $d = 2.5$, this indicates that the time series can be described by an equation system containing 3 variables. It is, however, difficult to estimate the structure of the equation system. This is subject to intensive research [e.g., *Rössler*, 1976; *Gousbet*, 1991a; 1991b; *Xu et al.*, 1993]. Generally, one may assume a common structure of a third-order equation system (after Taylor series expansion):

$$\begin{aligned}
\frac{dx}{dt} &= a_{11} + a_{12}x + a_{13}y + a_{14}z + a_{15}xy + a_{16}xz + \\
&\quad a_{17}yz + a_{18}x^2 + a_{19}y^2 + a_{110}z^2 \\
\frac{dy}{dt} &= a_{21} + a_{22}x + a_{23}y + a_{24}z + a_{25}xy + a_{26}xz + \\
&\quad a_{27}yz + a_{28}x^2 + a_{29}y^2 + a_{210}z^2 \\
\frac{dz}{dt} &= a_{31} + a_{32}x + a_{33}y + a_{34}z + a_{35}xy + a_{36}xz + \\
&\quad a_{37}yz + a_{38}x^2 + a_{39}y^2 + a_{310}z^2
\end{aligned} \tag{10}$$

After assuming initial conditions, the parameters (a-terms) in Eq. (10) can be determined by a parameter identification technique, e.g., Kalman filtering. At the same time, forecasting of the system can be made [e.g., *Xu et al.*, 1993].

It should be mentioned, that there are many difficulties involved in the above procedure and that all parts of the referred technique are object of intensive research. One major difficulty is involved in the fact that empirically observed geophysical time series contain noise. The technique to determine the attractor dimension according to above is very noise sensitive. Therefore, time series need firstly to be cleaned by a noise reduction algorithm. *Schreiber* [1993] gives a simple method for noise reduction especially developed for dimension estimations. The idea of the method is to replace each coordinate in x_i by an average value over a suitable neighborhood in the phase space. A radius η is chosen and for each coordinate in x_i the set Ω^{η}_i for all neighbors x_j so that:

$$supp\{|x_{j-k} - x_{i-k}|, \ldots, |x_{j+l} - x_{i+l}|\} \equiv \|x_j - x_i\|_{supp} \tag{11}$$

where k and l denote past and future coordinates. Consequently, coordinates in x_i are replaced by mean values in Ω^{η}_i:

$$x_i^{corr} = \frac{1}{|\Omega_i^{\eta}|} \sum_{\Omega_i^{\eta}} x_j \tag{12}$$

The question whether time varying climatological phenomena have low-dimensional properties or not is a much discussed issue at present. In fact, the complexity of climatological systems and the large number of degrees of freedom makes it unlikely that, e.g., precipitation is governed by a system of few variables [e.g., *Lorenz*, 1991]. However, if elements of observed time series can be shown to obey deterministic chaos, e.g., underlying trends, then a more complete understanding of the system and possibly better prediction techniques can be achieved.

Data base

The time series used in this paper consist of monthly sunspots, temperature, and precipitation observations. The temperature and precipitation data are in many respects unique and they merit some further description. Observations of temperature and precipitation have been done regularly at the city of Lund in the south of Sweden since 1741 [*Tidblom*, 1876]. Because of some gaps in the observations during the early years of measurements we use monthly series from 1753-1990 (238 years). The location for observations has changed over this period about four times for the rain gage and three times for the temperature gage. Also the type of gage has varied and the use of wind shield. The largest horizontal distance change for the gages has been about two kilometers. Simultaneous observations over a 15 years period for precipitation and a 6 years period for temperature at these locations indicated an absolute average difference of about 23 mm/year for precipitation and 0.16 °C for temperature. Uncertainties and errors are inherent in such long records. However, changes in location of the observations are expected to mainly result in an abrupt change in the mean value of the time series and not affecting other statistical properties.

RESULTS AND DISCUSSION

Trend removal and noise reduction

Figure 1 shows monthly mean Wolf sunspot data published by *Chernosky and Hagan* [1958] and consecutive issues of *J. Geophys. Res.* Monthly sunspot numbers were used in this paper to give reference values to other studies for the noise reduction techniques and correlation integral estimations used in this paper. The figure shows raw and smoothed sunspot numbers after five iterations in the noise reduction algorithm (Eq. (11) and (12)). It is seen that the algorithm keeps the main features of the time series even after several iterations. Figure 2 shows the spectral density of the two graphs in Figure 1. It is clearly seen that the noise algorithm mainly affects the high frequencies (short-term properties) of the time series and leave low frequencies more or less unaffected.

Figures 3 and 4 show corresponding time series for temperature and precipitation [see also *Kawamura et al.*, 1993]. Spectral analyses showed that temperature exhibited a 1-year and 6-month and

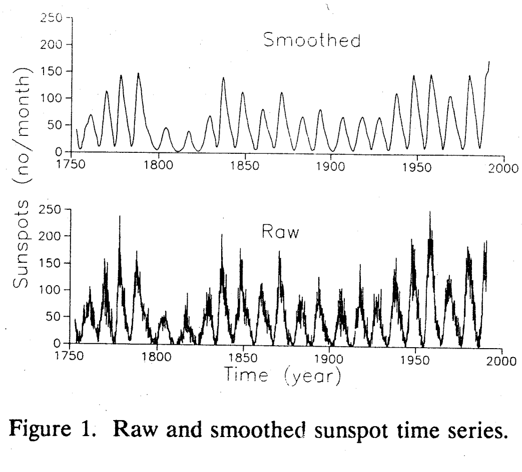

Figure 1. Raw and smoothed sunspot time series.

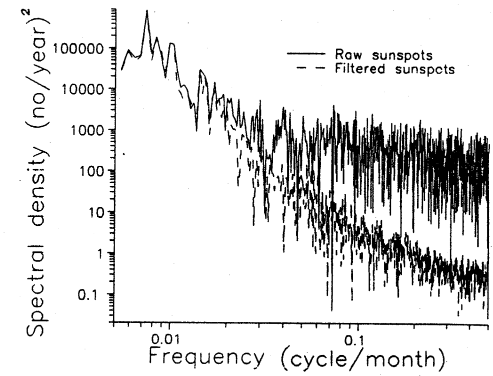

Figure 2. Fft spectral density for raw and smoothed sunspot time series in Figure 1.

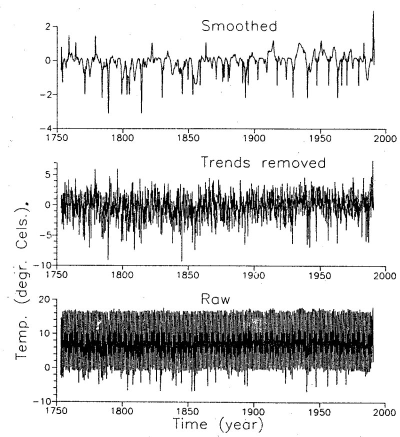

Figure 3. Time series of temperature (raw, after removal of cyclic trends, and smoothed; note different y-axis scales).

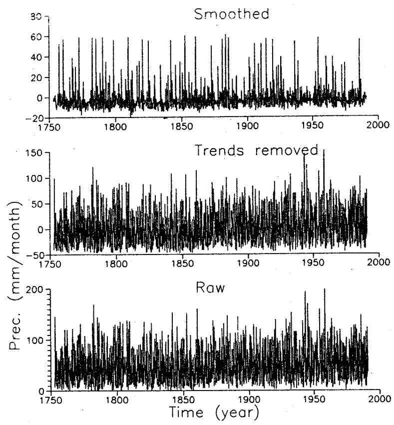

Figure 4. Time series of precipitation (raw, after removal of cyclic trend, and smoothed; note different y-axis scales).

precipitation a 1-year cyclic component that was statistically significant. Thus, these cyclic components were regarded as trend and therefore removed from the respective time series (the graphs labelled "Trends removed" in Figures 3 and 4). Figure 5 shows an example of autocorrelation for the temperature time series and it is seen that the correlation is close to zero already after 2-3 months.

After trend removal according to above, the time series were treated for noise in the same manner as sunspots (Figures 3 and 4). Five iterations were used for temperature and two for precipitation. The degree of smoothing has by necessity to be done in a subjective way since the true noise level is not known beforehand. In general, however, the main features of the time series (low frequency components) were sought to be kept while high frequency components, that can be considered as random noise were as much as possible removed.

From Figures 3 and 4 it is seen that a large portion of the variance is removed when using the noise reduction algorithm. This depends probably on a high degree of noise inherently buried in the time series used, as in all geophysical data. However, even if a significant portion of the variance is removed in the noise reduction, it is worthwhile to analyze the remaining low-frequency part of the time series. It is possible that this part of the time series can explain basic elements of the variation and thus contribute to a greater knowledge of factors that govern time

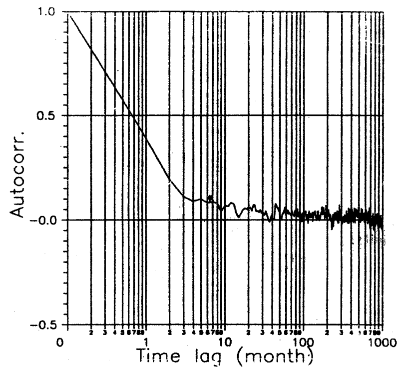

Figure 5. Autocorrelation of raw temperature residuals.

variation and also to increase forecasting capabilities.

Correlation dimensions

Correlation dimensions were calculated according to the algorithm given by *Grassberger* [1990]. Before this, phase space portraits were investigated for the different time series. Figures 6 and 7 show examples of this for sunspots and temperature, respectively. The figures display the effects of noise reduction in a remarkable way. For the raw (cyclic trends removed for temperature) time series no relationship is discernible in phase space. However, after noise reduction a very clear pattern can be seen. It therefore appears possible that strange attractors may be hidden in noisy geophysical time series.

Figure 8 shows the corresponding plots of $C(r)$ versus r for raw and smoothed sunspot time series (embedding dimensions $m = 2, 40$) and Figure 9 the local slopes $d\log C(r)/d\log r$ for the same data. A general time lag $\tau = 20$ was used throughout the analyses. This value appeared to be much larger than the dynamic correlation of the time series as shown in Figure 5.

The effects of noise reduction is clear in Figures 8 and 9. The raw data do not display any clear scaling region, while smoothed sunspots show saturation at about $d < 2$ for $0 < \log r < -1.5$. *Mundt et al.* [1991] found $d \approx 2.3$ for the same kind of sunspot data. However, they used a different noise reduction scheme for their study. *Theiler et al.* [1992] investigated sunspot data and similarly as in this study found no underlying deterministic chaotic component for the raw data. Consequently, it can be stated that

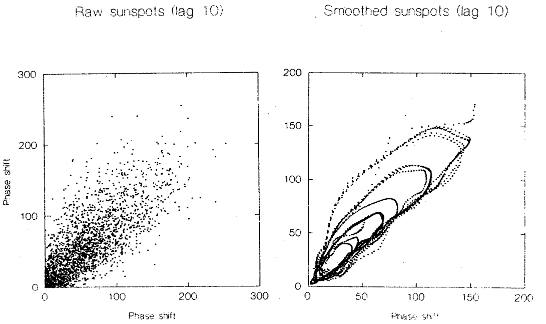

Figure 6. Strange attractor for raw and smoothed sunspots (note different scales).

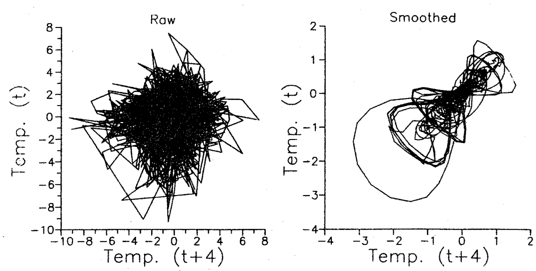

Figure 7. Strange attractor for raw (cyclic components removed) and smoothed temperature (note different scales).

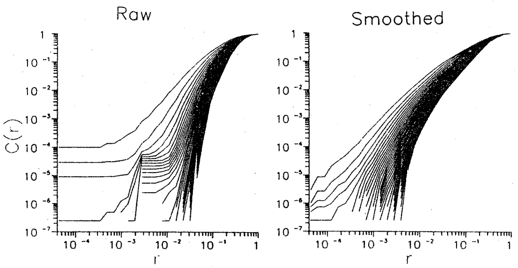

Figure 8. Plot of $C(r)$ versus r for raw and smoothed sunspot time series (embedding dimensions $m = 2$, 40).

noise has a significant effect on the outcome of the correlation dimension analyses.

In a similar way, Figure 10 shows local slopes $d\log C(r)/d\log r$ for raw (cyclic trends removed) and smoothed temperature. Here, smoothed temperature shows saturation at about $d < 4$ over the entire

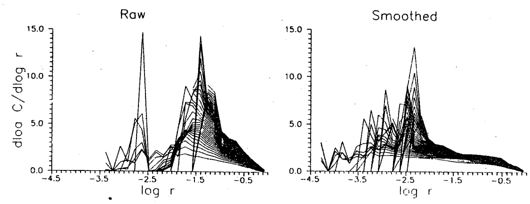

Figure 9. Slopes $d\log C(r)/d\log r$ for raw and smoothed sunspots as in Figure 7 (embedding dimensions $m = 2$, 40).

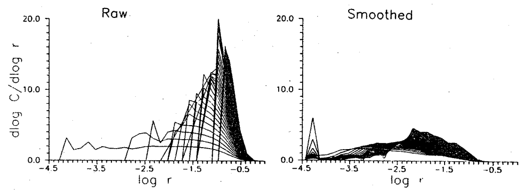

Figure 10. Slopes $d\log C(r)/d\log r$ for raw and smoothed temperature time series (embedding dimensions $m = 2$, 40).

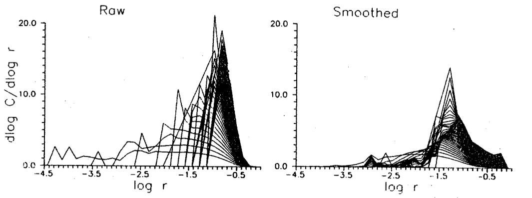

Figure 11. Slopes $d\log C(r)/d\log r$ for raw and smoothed precipitation time series (embedding dimensions $m = 2$, 40).

range of r while raw data do not show any similar characteristics.

In Figure 11, precipitation similarly, displays a small scaling region of r. However, in this case, only two iterations in the noise reduction were made. Increasing number of iterations would lead to an increasing scaling region.

SUMMARY

This study shows that low-dimensional chaotic components appear to be present in monthly time series of sunspots, temperature, and precipitation. However, the chaotic components can be distinguished only after a noise reduction scheme. The noise reduction scheme tends to reduce the original variance of the data significantly. This was especially true for the rather complex temperature and precipitation time series. There are thus obvious needs to continue to investigate effects of noise reduction on the estimated correlation dimension for geophysical time series.

Even if a large portion of the original variation is lost due to noise reduction, studies like in this paper may help to explain underlying low-frequency deterministic components. If deterministic chaotic components can be identified, it may help to explain conspicuous jumps and other changes that cannot be resolved by standard linear or autoregressive methods.

We prefer to view the results of this study as indicative of chaos only. As *Provenzale et al.* [1992] recently pointed out, there is still no unique method to distinguish between low-dimensional chaos and correlated noise. Instead, several data analyses techniques need to be combined to ascertain a chaotic behavior. Anyhow, the dynamical systems theory is a recent scientific progress which will expand quickly in the future. We view this theory as a new and promising way to characterize hydrological and geophysical time series.

Acknowledgment: This cooperative study was supported by the Japanese-German Center Berlin Special Exchange Program. The research was partly also supported by the Swedish Natural Science Research Council and the OK Environmental Foundation.

REFERENCES

Casdagli, M., S. Eubank, J. D. Farmer, and J. Gibson, 1991. *Physica D, 51*, 52-98.

Chernosky, E. J. and M. P. Hagan, 1958. *J. Geophys. Res., 63*, 755-788.

Elsner, J. B. and A. A. Tsonis, 1992. *Bull. Am. Meteorol. Soc., 73*, 49-60.

Farmer, J. D. and J. J. Sidorowich, 1987. *Phys. Rev. Lett., 59*, 845-848.

Farmer, J. D. and J. J. Sidorowich, 1988. *Rep. LA-UR-88-901, Los Alamos Natl. Lab.*, Los Alamos, N.M.

Gousbet, G. 1991a. *Phys. Rev. A, 43*, 5321-5331.

Gousbet, G. 1991b. *Eighth Symp. on Turbulent Shear Flows, Techn. Univ. of Munich*, Sep. 9-11, 1991.

Grassberger, P. 1990. *Phys. Lett. A, 148*, 63-68.

Grassberger, P. and I. Procaccia, 1983a. *Phys. Rev. Let., 50*, 346-349.

Grassberger, P. and I. Procaccia, 1983b. *Physica, 9 D*, 189-208.

Grassberger, P., T. Schreiber, and C. Schaffrath, 1991. *Int. J. Bifurc. Chaos, 1*, 521-547.

Grassberger, P., R. Hegger, H. Kantz, C. Schaffrath, and T. Schreiber, 1993. *Chaos, 3*, 127-143.

Kawamura, A., K. Jinno and R. Berndtsson, 1993. Proc. of Intern. Symp. on Extreme Hydrological Events: Precipitation, Floods, and Droughts, July 20-23, 1993, Yokohama, *IAHS Publ. No. 213*, 103-112.

Lorenz, E. N. 1963. *J. Atmos. Sci., 20*, 130-141.

Lorenz, E. N. 1991. *Nature, 353*, 241-244.

Mundt, M. D., W. B. Maguire II, and R. R. P. Chase, 1991. *J. Geophys. Res., 96*, 1705-1716.

Osborne, A. R. and A. Provenzale, 1989. *Physica D, 35*, 357-381.

Packard, N. H., J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, 1980. *Phys. Rev. Lett., 45*, 712-716.

Provenzale, A., A. R. Osborne, and R. Soi, 1991. *Physica D, 47*, 361.

Provenzale, A., L. A. Smith, R. Vio, and G. Murante, 1992. *Physica D, 58*, 31-49.

Ruelle, D. 1981. In: *Nonlinear phenomena in chemi cal dynamics*, Eds. A. Pacault and C. J. Vidal, Springer Verlag, 1981.

Rössler, O. E. 1976. *Z. Naturforsch., 31*, 1664-1670.

Schreiber, T. 1993. *Phys. Rev. E. Stat. Physics, 47*, 2401-2404.

Schuster, H. G. 1984. *Deterministic chaos, an introduction*, Physikverlag, Weinheim, 220 pp.

Sugihara, G. and R. M. May, 1990. *Nature, 344*, 734-741.

Takens, F. 1981. In: *Dynamical systems and turbulence*, Ed. by D. Rand and L.-S. Young, Springer Verlag, pp. 366-381.

Theiler, J., S. Eubank, A. Longtin, B. Galdrakin, and J. D. Farmer, 1992. *Physica D, 58*, 77-94.

Tidblom, A. V. 1876. *Lunds Universitets Årsskrift, Tom, XII, Afd. Mathem. Naturvetenskap*, 1-78, Lund, Sweden.

Tsonis, A. A. and J. B. Elsner, 1990. *Beitr. Phys. At mosph., 63*, 171-176.

Xu, S., K. Jinno, A. Kawamura, R. Berndtsson, and J. Olsson, 1993. *Proc. of Annual Conf. of Japan Soc. of Civ. Eng., 37*, pp. 853-856.

Yule, G. U. 1927. *Phil. Trans. R. Soc. London A, 226*, 267.