Proc. of the Int. Conf. on Water Resour. &
Environ. Res. (Volume I)
Oct. 29-31, 1996 Kyoto, Japan

PREDICTION OF MONTHLY TEMPERATURE TIME SERIES USING RECONSTRUCTED CHAOTIC SYSTEM EQUATIONS

M. MATSUMOTO[1], A. KAWAMURA[1], K. JINNO[1], R. BERNDTSSON[2], and S. XU[1]

abstract

Prediction of long-term changes in temperature is of vital importance for estimation of future available water resources. We investigate an observed 238-year monthly time series of temperature by dynamical systems theory and use results from the analysis to make short-term predictions in real-time. The methodology for this is (1) based on the time series behavior in phase-space and dimension of attractor, determine a reference system of equations, (2) assume a general structure of governing system equations by Taylor series expansion, (3) use the reference system of equations as initial state for the chaotic system, and (4) identify parameters and structure of governing system equations by a parameter updating procedure. We use the Lorenz equations as a reference system of equations and the extended kalman filter to identify the structure of the governing system and to make updated predictions of the chaotic monthly temperature. We show that predictions can be made for filtered monthly temperature time series if the prediction lead time is short. The results indicate that parts of monthly temperature variation at a point may follow a chaotic time trajectory as influenced by large-scale atmospheric flow.

1. INTRODUCTION

Prediction of climatic variables such as temperature is important since variation of these has profound impacts on the availability of water resources. However, temperature is highly variable and unpredictable in the long-term. This is due to complex interactions between large-scale atmospheric flow and local physiographical conditions with many independent and irreducible degrees of freedom. A common procedure, however, is to reduce this complex reality into a set of partial differential equations for atmospheric mass and heat transport, e.g., in general circulation models (GCMs). During recent years, an alternative approach has emerged, namely, to analyze and build models directly from available observations (Farmer and Sidorowitch, 1987; Sugihara and May, 1990; Mundt et al., 1991; Jinno et al., 1995). The general idea for this is the application of dynamical systems theory to geophysical processes.

The first and most basic step when analyzing a time series within the framework of dynamical systems theory, is to perform a phase space reconstruction. The assumption behind a phase space reconstruction is that the past and future of the time series contain information about unobserved variables that may be used to define a state of the process at the present time (Casdagli et al., 1991). The procedure of a phase space reconstruction is motivated due to unknown properties of the dynamical system such as relevant variables and their total number. Their total number may be determined by estimating the dimension of the time series (dimension of the attractor).

We use observed long-term time series (238 years; 1753-1990) of monthly temperature to investigate if they can be predicted by the use of dynamical systems theory. The methodology for this is: (1) based on the time series behavior in phase space and dimension of attractor, determine

[1]: Department of Civil Engineering, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka, 812, Japan
[2]: Department of Water Resources Engineering, Lund University, Box 118, S-221 00 Lund, Sweden

a reference system of equations, (2) assume a general structure of governing system equations by Taylor series expansion, (3) use the reference system of equations as initial state for the chaotic system, and (4) identify parameters and structure of governing system equations by a parameter updating procedure.

We investigate two types of temperature variation, nonlinear long-term trends and short-term variations. Dimension estimates are calculated using both these types of temperature variability. After this, the Lorenz equations are assumed as a reference system of equations as initial state for short-term monthly temperature prediction. By applying the extended kalman filter technique and a general structure of governing system equations by Taylor series expansion, parameters are updated in real-time and identified. We close with a summary and discussion on how the results may be practically used for climate predictions.

2. THEORY

We use the methodology of Jinno et al. (1995) to reconstruct the chaotic system for monthly temperature and use this system of equations to predict the future behavior of temperature. The methodology for this is (1) based on the time series behavior in phase-space and dimension of attractor, determine a reference system of equations, (2) assume a general structure of governing system equations by Taylor series expansion, (3) use the reference system of equations as initial state for the chaotic system, and (4) identify parameters and structure of governing system equations by a parameter updating procedure. Figure 1 gives an outline of this methodology. The figure shows the hidden original system and the true time series (area within broken line in Fig. 1) which will never be known. Instead, we are forced to deal with distorted and noisy observations.

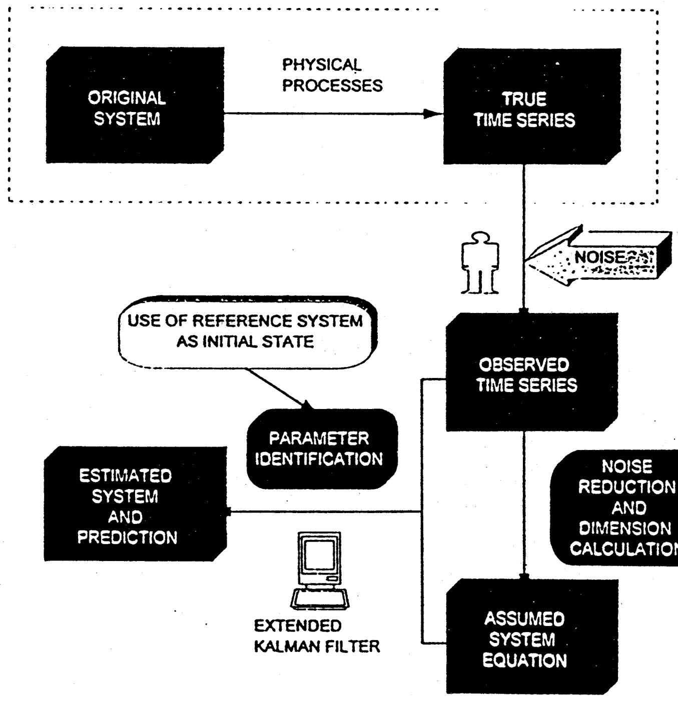

Figure 1. Outline of the methodology to predict monthly temperature (Matsumoto, 1996).

The first step in our analysis is to reduce the noise without changing the underlying deterministic chaotic signal and to perform a dimension calculation. The dimension calculation is started with a phase space reconstruction (e.g., Henderson and Wells, 1988; Tsonis and Elsner, 1990). This simply means that the time series is plotted against itself with a proper time lag. By performing this analysis properties of the attractor can be investigated. Occurrence of an attractor means that the future time behavior is not random but instead settles on a pattern close to that of the attractor. This in turn, indicates the type of nonlinearity and if it is possible to make predictions into the future for the time series. By determining the dimension of the attractor it is possible to evaluate if a low-dimensional equation system can be used to describe the nonlinear time behavior for use in simple forecasting schemes (e.g., Grassberger and Procaccia, 1983a; 1983b). The correlation integral $C(r)$ of the attractor defines the density of points around a specific coordinate with radius r of the time series and can be used to determine the dimension. 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. In this paper, the algorithm according to Grassberger (1990) was used to estimate correlation integrals. Values of d that are not integers indicate a fractal and thus a chaotic attractor. The dimension d of the attractor is given by the slope of log $C(r)$ for the slope of log r. For deriving the dimension d of the attractor from observations $x(t)$ it is sufficient to embed it in an m-dimensional space ($d < m$):

$$\mathbf{x}(t) = [x(t), x'(t), \ldots, x^{(m-1)}(t)] \tag{1}$$

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; Ruelle, 1981). According to this and introducing a time lag τ one gets (e.g., Grassberger and Procaccia, 1983a, 1983b):

$$x(t),\ x(t+\tau),\ x(t+2\tau), \ldots, x(t+(m-1)\tau) \tag{2}$$

However, known methods for dimension calculations are noise sensitive and, therefore, a noise-reduction scheme has to be employed. In this paper, we use the algorithm of Schreiber (1993) which was especially developed for dimension estimations. The idea of the method is to replace each coordinate in $\mathbf{x}_i$ by an average value over a suitable neighborhood in the phase space.

Preliminary investigations have shown that the Lorenz equations (Lorenz, 1963) may be used as a reference system of equations to be used as initial state for the predictions (Matsumoto, 1996). These equations are physically relevant and simplification of large-scale atmospheric flow. The Lorenz equations are (Lorenz, 1963):

$$\begin{cases} \dot{x} = \sigma(y-x) \\ \dot{y} = Rx - y - xz \\ \dot{z} = xy - bz \end{cases} \tag{3}$$

where the parameters σ, R, and b are constants. In order to use the Lorenz system as an initial state for the system to be identified, a linear transformation $x^{*} = \gamma x$ and $t^{*} = Tt$ is necessary. This is done to adjust the amplitude of $x(t)$ and to synchronize the time series. Consequently, the modified Lorenz system becomes:

$$\begin{cases} \dfrac{dx^*}{dt^*} = \dfrac{\sigma}{T}(\gamma y - x^*) \\ \dfrac{dy}{dt^*} = \dfrac{1}{T}(\dfrac{R}{\gamma}x^* - y - \dfrac{1}{\gamma}x^* z) \\ \dfrac{dz}{dt^*} = \dfrac{1}{T}(\dfrac{1}{\gamma}x^* y - bz) \end{cases} \tag{4}$$

As in Jinno et al. (1995) we assume that the system dynamics can be expressed by a general first-order simultaneous nonlinear differential equation system. The nonlinear differential equations $f(x,y,z)$ are expanded into Taylor series and terms up to second order are included. The system equation then becomes:

$$\begin{cases} \dot{x} = f_1(x,y,z) = a_{10}+a_{11}x+a_{12}y+a_{13}z+a_{14}xy+a_{15}xz+a_{16}yz+a_{17}x^2+a_{18}y^2+a_{19}z^2 \\ \dot{y} = f_2(x,y,z) = a_{20}+a_{21}x+a_{22}y+a_{23}z+a_{24}xy+a_{25}xz+a_{26}yz+a_{27}x^2+a_{28}y^2+a_{29}z^2 \\ \dot{z} = f_3(x,y,z) = a_{30}+a_{31}x+a_{32}y+a_{33}z+a_{34}xy+a_{35}xz+a_{36}yz+a_{37}x^2+a_{38}y^2+a_{39}z^2 \end{cases} \tag{5}$$

where a_{ij} (i = 1,2,3, j = 0, 1, ... , 9) are parameters. Then, x, y, z, and a_{ij} in Eq. (5) are used in the system vector X according to:

$$X = [x_1, x_2, \cdots, x_{32}, x_{33}]^T = [x, y, z, a_{10}, \cdots, a_{19}, a_{20}, \cdots, a_{29}, a_{30}, \cdots, a_{39}]^T \tag{6}$$

Accordingly, the system equation becomes:

$$\begin{cases} \dot{x}_1 = f_1(X) = x_4+x_5x_1+x_6x_2+x_7x_3+x_8x_1x_2+x_9x_1x_3+x_{10}x_2x_3+x_{11}x_1^2+x_{12}x_2^2+x_{13}x_3^2 \\ \dot{x}_2 = f_2(X) = x_{14}+x_{15}x_1+x_{16}x_2+x_{17}x_3+x_{18}x_1x_2+x_{19}x_1x_3+x_{20}x_2x_3+x_{21}x_1^2+x_{22}x_2^2+x_{23}x_3^2 \\ \dot{x}_3 = f_3(X) = x_{24}+x_{25}x_1+x_{26}x_2+x_{27}x_3+x_{28}x_1x_2+x_{29}x_1x_3+x_{30}x_2x_3+x_{31}x_1^2+x_{32}x_2^2+x_{33}x_3^2 \\ \dot{x}_i = f_i(X) = 0 \qquad (4 \le i \le 33) \end{cases} \tag{7}$$

In the extended kalman filter (EKF), the nonlinear function $f(X)$ is expanded into a Taylor series at a point X^*. The observation equation in the EKF is a vector function according to:

$$Y = g(X) \tag{8}$$

It is further assumed that only one of the three variables, x, is observed, and the observation vector Y becomes scalar with element y_1 according to:

$$y_1 = g_1(X) = x_1 \tag{9}$$

The EKF is consequently used to identify the structure of the governing system and to make updated predictions of monthly temperature.

2. CHAOTIC CHARACTERISTICS OF TEMPERATURE

Figure 2 shows an example of long-term temperature trends for the data used. The monthly temperature data were averaged over approximately 11-year periods (see Berndtsson et al., 1996) after a linear trend corresponding to an increase of 0.84°C per 100 years was removed (see further Kawamura et al., 1993). Spline functions were then used to interpolate between these averaged values. Figure 2, consequently, shows how temperature may vary on an approximately decade scale. Figure 3 shows the corresponding phase space portrait with a time lag τ of 2 years. As seen from the figure, the behavior of long-term temperature in time appears not to be random. Instead, it seems as if a type of attractor with both smaller and larger excursions from the mean emerges. A possible explanation for this type of attractor may be the correlation with sunspot cycle length as shown by Berndtsson et al. (1996).

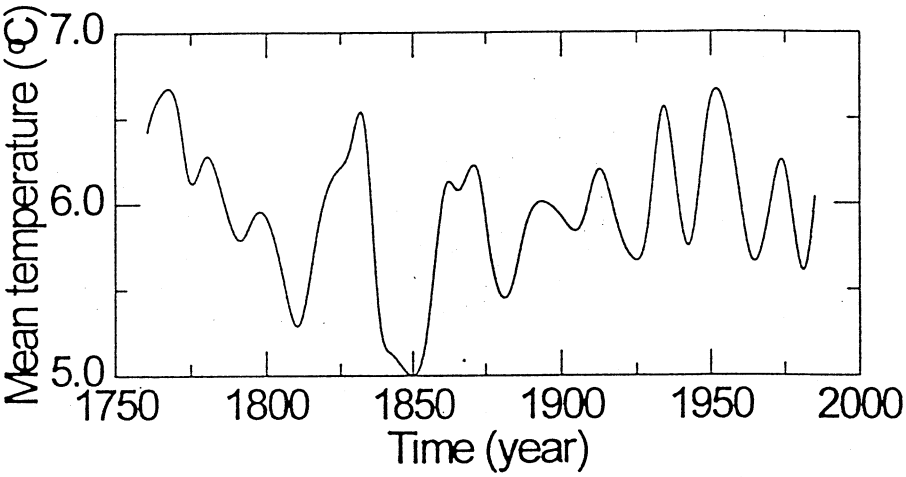

Figure 2. Monthly temperature averaged over approximately 11-year periods after removal of linear trend and interpolated by spline functions.

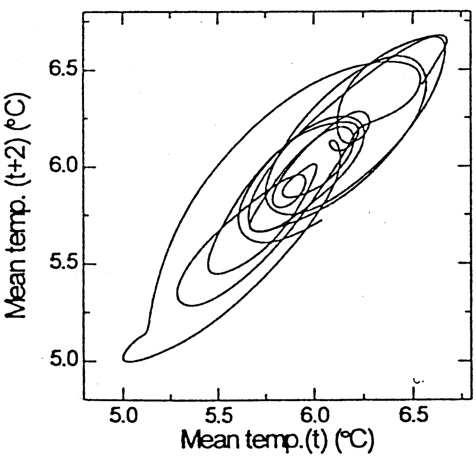

Figure 3. Attractor for interpolated mean temperature (data as in Fig. 2). Time lag is 2 years.

The occurrence of an attractor means that the future time behavior is not random but instead settles on a pattern close to that of the attractor. This in turn, may indicate the type and degree of nonlinearity and if it is possible to make predictions into the future for the time series. Of course, interpolated data like this have to be interpreted cautiously and may only indicate properties of the actual process. Even so, it is believed that the phase-space portrait in Fig. 3 embraces some of the general and long-term behavior of temperature.

A similar analysis was made for short-term (monthly) temperature variation. Figure 4 shows the attractors for raw (annual cyclic component removed) and noise-reduced using the algorithm of Schreiber (1993). As seen from the figure, an attractor with similarity to the well-known Lorenz attractor emerges for the noise-reduced data. Because of this similarity, the Lorenz equations were chosen as a reference system for initial conditions in the below prediction scheme of temperature.

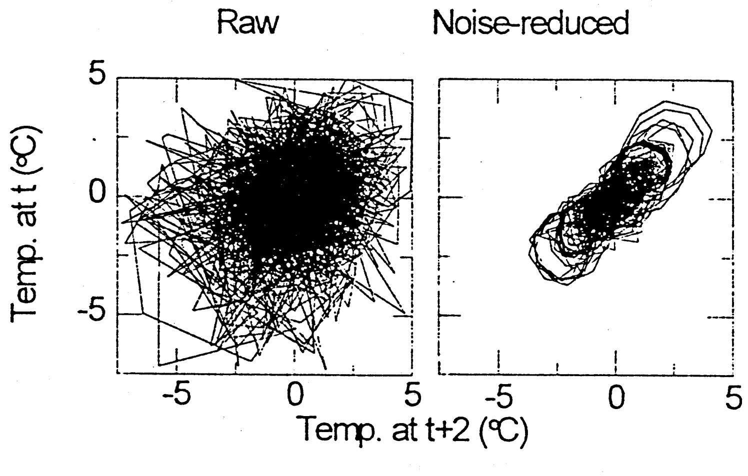

Figure 4. Comparison of attractors for raw (annual cyclic component removed) and noise-reduced monthly temperature.

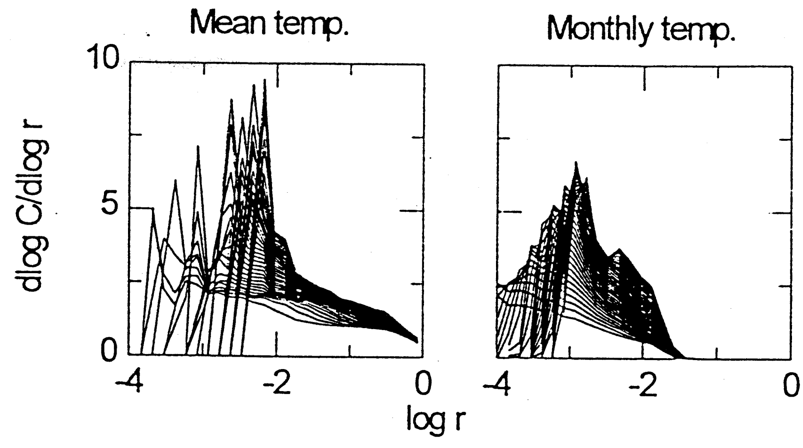

Figure 5. Slopes *d*log *C*(*r*)/*d*log *r* versus log *r* for temperature attractors in Figs. 3 and 4.

Figure 5 shows the resulting $d\log C(r)/d\log r$ vs. $\log r$ for the long- and short-term temperature attractors of Figs. 3 and 4. The m-embedding was chosen between $2 \leq m \leq 40$. As seen from the figure, there is a clear scaling region for $-1.5 < \log r < 0$ for both types of temperature variation. Both types of temperature variation display saturation at a correlation dimension $d < 2.5$ in this range. Consequently, a nonlinear equation system with three independent variables seems appropriate to describe the evolution in time for both types of temperature variation (Takens, 1981; Ruelle,1981).

3. REAL-TIME PREDICTION OF MONTHLY TEMPERATURE

Following the above analysis, a prediction system with three independent variables was assumed as system equation. Because of the similarity between the empirical temperature attractor and Lorenz' well known attractor, the Lorenz model was chosen as reference system for the initial state. Consequently, initial parameter values in Eq. (6) were chosen as: $a_{11} = -\sigma/T = -10/0.13$, $a_{12} = \sigma^{*}\gamma/T = 10 \cdot 0.20/0.13$, $a_{21} = R/(\gamma^{*}T) = 28/(0.20 \cdot 0.13)$, $a_{22} = -1/T = -1/0.13$, $a_{26} = -1/(\gamma^{*}T) = -1/(0.20 \cdot 0.13)$, $a_{33} = -b/T = (-8/3)/0.13$, $a_{34} = 1/(\gamma^{*}T) = 1/(0.20 \cdot 0.13)$, and all other terms were set to 0.

Figure 6 shows an example of one- and three-month ahead predictions for the temperature time series. It is seen that one-month ahead predictions can be made with small errors while three-month ahead predictions generate quite large errors. A possible explanation for this may be that the chaotic trajectories in time lose their original information content in a very rapid way.

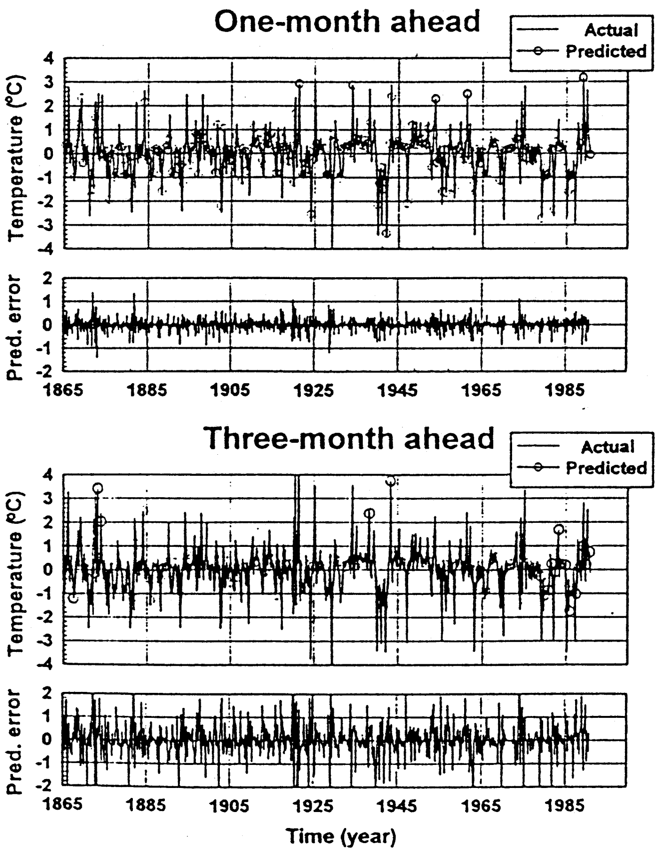

Figure 6. One- and three-month ahead predictions of monthly temperature with errors.

4. SUMMARY AND CONCLUSIONS

We have outlined a methodology to use dynamical systems theory to make updated predictions of monthly temperature in real-time. It was shown that good predictions could be made at least for one-month ahead predictions. Three-month ahead predictions were less successful. A possible way to improve the predictions may be to include more observables in the system equation. For example, if a physical cause-effect relationship could be included in Eq. (5) for the y and/or z terms, this may improve the predictions significantly. An example of such a relationship is the significant dependence between sunspot cycle length and long-term temperature trends.

ACKNOWLEDGMENTS

The study was sponsored by the Swedish Natural Sciences Research Council and the Swedish Water Regulation Authority (Vaso/Huva). This support is gratefully acknowledged.

REFERENCES

Berndtsson, R., M. Matsumoto, K. Jinno, A. Kawamura, and S. Xu, (1996), Relationship between sunspot cycle length and local land air temperature, J. Geophys. Res., (subm.).

Casdagli, M., S. Eubank, J. D. Farmer, and J. Gibson, (1991), State space reconstruction in the presence of noise, Physica D, 51, 52-98.

Farmer, J. D., and J. J. Sidorowich, (1987), Predicting chaotic time series, Phys. Rev. Lett., 59, 845-848.

Grassberger, P., (1990), An optimized box-assisted algorithm for fractal dimensions, Phys. Lett. A, 148, 63-68.

Grassberger, P., and I. Procaccia, (1983a), Characterization of strange attractors, Phys. Rev. Let., 50, 346-349.

Grassberger, P., and I. Procaccia, (1983b), Measuring the strangeness of strange attractors, Physica, 9D, 189-208.

Henderson, H. W., and R. Wells, (1988), Obtaining attractor dimensions from meteorological time series, Adv. Geophys., 30, 205-237.

Jinno, K., S. Xu, R. Berndtsson, A. Kawamura, and M. Matsumoto, (1995), Prediction of sunspots using reconstructed chaotic system equations, J. Geophys. Res., 100, 14773-14781.

Kawamura, A., K. Jinno and R. Berndtsson, (1993), Analyses of a 238-year long monthly precipitation time sequence by use of filtering techniques, IAHS Publ. No. 213, 103-112.

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

Matsumoto, M., (1996), Prediction of chaotic time series, M. Sc. Thesis, Dep. Civil Eng., Kyushu Univ., Fukuoka, pp. 1-110.

Mundt, M. D., W. B. Maguire II, and R. R. P. Chase, (1991), Chaos in the sunspot cycle: analysis and prediction, J. Geophys. Res., 96, 1705-1716.

Ruelle, D., (1981), In: Nonlinear phenomena in chemical dynamics, Eds. A. Pacault and C. J. Vidal, Springer Verlag.

Schreiber, T., (1993), An extremely simple nonlinear noise reduction method, Phys. Rev. E. Stat. Physics, 47, 2401-2404.

Sugihara, G., and R. M. May, (1990), Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series, Nature, 344, 734-741.

Takens, F., (1981), Detecting strange attractors in fluid turbulence, In: Dynamical systems and turbulence, Ed. by D. Rand and L.-S. Young, Springer Verlag, pp. 366-381.

Tsonis, A. A., and J. B. Elsner, (1990), Multiple attractors, fractal basins and longterm climate, Beitr. Phys. Atmosph., 63, 171-176.