

Chaotic Analysis on Series of Groundwater Depth in Hetao Irrigation District

Weiguang Wang^{1,2} Shizhang Peng¹

¹State Key Laboratory of Hydrology-Water Resource and Hydraulic Engineering, Hohai University, Nanjing 210098, P.R. China

²School of Hydrology-Water Resource, Hohai University, Nanjing 210098, P.R. China

Abstract

The chaotic behavior of groundwater depth time series is investigated using correlation dimension method and the CAO method based on the reconstruction of phase-space. A groundwater depth time series, each five days series of 25 years (with a total of 1793 values) observed in Hetao irrigation district of Yellow River basin, China, is studied. Results show that the chaotic character of groundwater level may exist under the natural condition, chaotic analysis on groundwater depth time series offer theoretical basis and gist for groundwater depth forecasting.

Keywords: Correlation dimension, CAO method, Chaotic behavior, Groundwater depth

1. Introduction

The tremendous spatial and temporal variability of hydrological processes has been believed, recently, to be due to the influence of a large number of variables. Consequently, the majority of the previous investigations on modeling hydrological processes have essentially employed the concept of a stochastic process. However, recent studies have indicated that even simple deterministic systems, influenced by a few nonlinear interdependent variables, might give rise to very complicated structures (i.e. deterministic chaos). Therefore, it is now believed that the dynamic structures of the seemingly complex hydrological processes, such as rainfall and runoff, might be better understood using nonlinear deterministic chaotic models than the stochastic ones.

The investigation of the existence of chaos in hydrological processes has been of much interest lately. The outcomes of the investigations are very encouraging as they provided evidence regarding the existence of low-dimensional chaos implying the possibility of accurate short-term predictions.

The science of chaos is a burgeoning field, and the available methods to investigate the existence of chaos

in time series are still in a state of infancy. However, the considerable attention that the theory has received in almost all fields of natural and physical sciences has motivated improvements in existing methods for the diagnosis of chaos and the proposal of new ones. The methods available thus far are the correlation dimension method (Grassberger, 1983)[1], the nonlinear prediction method (Farmer, 1987)[2] including deterministic versus stochastic diagram (Casdagli, 1991)[3], the Lyapunov exponent method (Wolf, 1985)[4], the surrogate data method (Theiler, 1992)[5], and the linear and nonlinear redundancies (Palus, 1995)[6]. Among these the correlation dimension method has been the most widely used one for the investigation of deterministic chaos in hydrological phenomena.

It is relevant to note that the application of chaos identification methods, particularly the correlation dimension method, to hydrological time series and the reported results have very often been questioned because of the fundamental assumptions with which the methods have been developed, that is, that the time series is infinite and noise-free. Important issues, in the application of chaos identification methods to hydrological data, for example, data size, noise, delay time, etc., and the validity of chaos theory in hydrology have been discussed in detail by Sivakumar(2000)[7]. At the same time, the studies by Sivakumar reveals that we should employed more than one method in the investigation of the existence of chaos in hydrological processes and to verify or confirm the results. Thus, In this thesis, we identify the existence of chaos in the groundwater depth time series in Hetao irrigation district of Yellow River basin, from 1981 to 2005 (Figure 1) by correlation dimension method and CAO method.

2. Correlation dimension method

For a scalar time series x_t , where $t = 1, 2, \dots, N$, the phase space can be reconstructed using the method of delays(Takens,1980)[8]. The basic idea in the method

of delays is that the evolution of any single variable of a system is determined by the other variables with which it interacts. Information about the relevant variables is thus implicitly contained in the history of any single variable. On the basis of this an “equivalent” phase space can be reconstructed by assigning an element of the time series x_t and its successive delay as coordinates of a new vector time series

$$Y_j = (X_j, X_{j+\tau}, X_{j+2\tau}, \dots, X_{j+(m-1)\tau}) \quad (1)$$

Where $j = 1, 2, \dots, N - (m-1)\tau / \Delta t$, m is the dimension of the vector Y_j , also called the embedding dimension, and τ is a delay time taken to be some suitable multiple of the sampling time Δt .

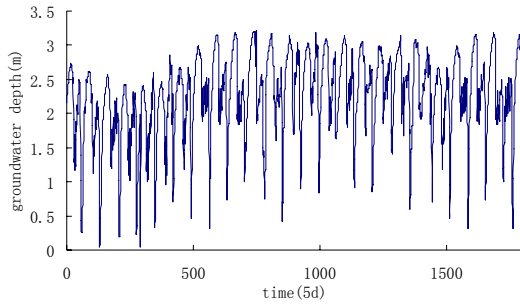


Fig. 1: Groundwater depth of Hetao irrigation district in Yellow River basin of China.

The goal of determining the dimension of an attractor is that the dimensionality of an attractor furnishes information on the number of dominant variables present in the evolution of the corresponding dynamical system. Dimension analysis will also reveal the extent to which the variations in the time series are concentrated on a subset of the space of all possible variations. The central idea behind the application of the dimension approach is that systems whose dynamics are governed by stochastic processes are thought to have an infinite value for the dimension. A finite, noninteger value of the dimension is considered to be an indication of the presence of chaos.

Correlation dimension is a measure of the extent to which the presence of a data point affects the position of the other points lying on the attractor. The correlation dimension method uses the correlation integral for distinguishing between chaotic and stochastic behaviors. The concept of the correlation integral is that an irregular-looking process arising from deterministic dynamics will have a limited number of degrees of freedom equal to the smallest number of first-order differential equations that capture the most important features of the dynamics. Thus, when one constructs phase spaces of increasing dimension for an infinite data set, a point will be reached where the dimension equals the number of

degrees of freedom and beyond which increasing the dimension of the representation will not have any significant effect on the correlation dimension.

According to the embedding theorem of Takens[1980][8], to characterize a dynamic system with an attractor dimension d , an $(m = 2d + 1)$ -dimensional phase space is required. However, Abarbanel et al.[1990][9] suggested that $m > d$ would be sufficient. For an m -dimensional phase space the correlation function $C(r)$ is given by

$$C(r) = \lim_{N \rightarrow \infty} \frac{2}{N(N-1)} \sum_{i,j} H(r - |Y_i - Y_j|) \quad (2)$$

$1 \leq i < j \leq N$

Where H is the Heaviside step function, with $H(u) = 1$ for $u > 0$, and $H(u) = 0$ for $u \leq 0$, where $u = r - |Y_i - Y_j|$, N is the number of point on the reconstructed attractor, r is the radius of the sphere centered on Y_i or Y_j .

If the time series is characterized by an attractor, then for positive values of r the correlation function $C(r)$ is related to the radius r by the following relation:

$$C(r) \approx \alpha r^v \quad (3)$$

$r \rightarrow 0$
 $N \rightarrow \infty$

where α is a constant; and v is the correlation exponent or the slope of the $\ln C(r)$ versus $\ln r$ plot is given by:

$$v = \lim_{\substack{r \rightarrow 0 \\ N \rightarrow \infty}} \frac{\ln C(r)}{\ln r} \quad (4)$$

The slope is generally estimated by a least squares fit of a straight line over a certain range of r , called the scaling region. For a finite data set, such as the precipitation data series, it is clear that there is a separation r below which there are no pairs of point; that is, it is “depopulated.” At the other extreme, when the value of r exceeds the set diameter, the correlation function increases no further; that is “saturated.” Therefore, for a finite data set, the region sandwiched between the depopulation region and the saturation region is considered as the scaling region. A region is to estimate the local slope given by $d[\ln C(r)] / d[\ln r]$.

To observe whether chaos exist, the correlation exponent(or local slope) values are plotted against the value of the correlation exponent is finite, low, and noninteger, the system is considered to exhibit low-dimension chaos. The saturation value of the correlation exponent is defined as the correlation dimension of the attractor. The nearest integer above

the saturation value provides the minimum number of phase spaces or variables necessary to model the dynamics of the attractor. On the contrary, if the correlation exponent increases without bound with increase in the embedding dimension, the system under investigation is considered as stochastic .

3. CAO method[10]

Cao think that the problem of the false neighbor method[11] is how to choose the threshold value, which is often used to determine the embedding dimension. Different time series data may have different threshold values. These imply that it is very difficult and even impossible to give an appropriate and reasonable threshold value which is independent of the dimension d and each trajectory's point, as well as the considered time series data.

To avoid the problem, Cao method instead define the following quantity

$$E(m) = \frac{1}{N - m\tau} \sum_{i=1}^{N-m\tau} a(i, m) \quad (5)$$

$E(m)$ is dependent only on the dimension m and the lag τ . To investigate its variation from m to $m+1$, we define

$$E1(m) = E(m+1) / E(m) \quad (6)$$

Cao method found that $E1(m)$ stop changing when m is greater than some value m_0 if the time series comes from an attractor. Then $m_0 + 1$ is the minimum embedding dimension we look for. It is necessary to define another quantity which is useful to distinguish deterministic signals from stochastic signals. Let

$$E2(m) = E^*(m+1) / E^*(m) \quad (7)$$

Where:

$$E^*(m) = \frac{1}{N - m\tau} \sum_{i=1}^{N-m\tau} |x_{i+m\tau} - x_{n(i,m)+m\tau}| \quad (8)$$

For time series data from a random set of numbers, $E1(m)$, in principle, will never attain a saturation value as m increases. But in practical computations, it is difficult to resolve whether the $E1(m)$ is slowly increasing or has stopped changing if m is sufficiently large. In fact, since available observed data samples are limited, it may happen that $E1(m)$ stops changing at some m although the time series is random. To solve this problem, Cao method consider the quantity $E2(m)$. For random data, since the future values are independent of the past values, $E2(m)$ will be equal to 1 for any m in this case. However, for deterministic

data, $E2(m)$ is certainly for all m ; in other words, there must exist some m 's such that $E2(m) \neq 1$.

Cao method recommend calculating both $E1(m)$ and $E2(m)$ for determining the minimum embedding dimension of a scalar time series, and to distinguish deterministic data from random data.

4. Result and analysis

$\ln C(r)$ versus $\ln r$ is show in Figure 2. Figure 2 is obtained by using formula (2) and (4). From the Figure 2, the scaling region is existed; herein the groundwater series has chaotic character. The slope of the line in scaling region is the correlation dimension. The relationship between the correlation dimension values and the embedding dimension values is shown in Figure 3. It can be seen that the correlation dimension values increases with the embedding dimension up to a certain value and then saturates beyond that value. The saturation of the correlation dimension beyond a certain embedding dimension value is an indication of the existence of deterministic dynamics. The saturated correlation dimension is about 3.469 ($v = 3.469$), and the embedding dimension $m = 11$. The finite and low correlation dimension is an indication that the groundwater depth series exhibit chaotic behavior.

We test the groundwater depth time series with Cao method and show the results in Figure 4. Very clearly the minimum embedding dimension is 11, which is the same as the results tested by correlation dimension method, furthermore, $E2$ of the groundwater series is always not equal to 1, which indicated that the groundwater depth series exhibit chaotic behavior. As a comparison, we show the results from random colored noise by Cao method in figure 5. Here the $E2$ values approximately equal 1 for any m and have certainly no relation to the $E1$ values. One can see that the Cao method can distinguish random time series from chaotic time series.

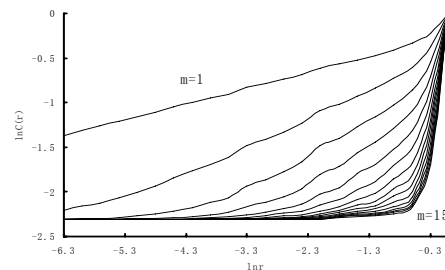


Fig. 2: $\ln C(r)$ versus $\ln r$ plot.

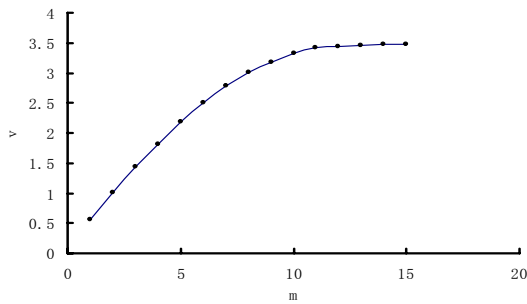


Fig. 3: Relationship between embedding dimension and correlation dimension.

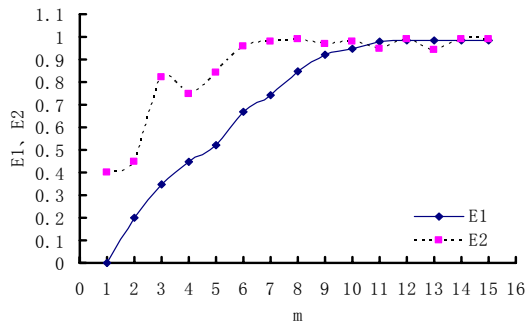


Fig. 4: The values E1 and E2 for the groundwater depth time series.

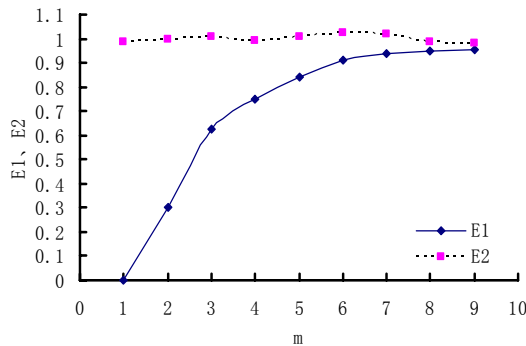


Fig.5: The values E1 and E2 for the random colored noise.

5. Conclusions

An attempt was made in this study to investigate the chaotic feature of the groundwater depth time series of Hetao irrigation district in China. Correlation dimension method and Cao method were employed to achieve the target. E2 value of the groundwater series, which is always not equal to 1, and the finite and low correlation dimension of series are an indication that the groundwater depth series exhibit chaotic behavior. This results provides scientific gist

for predicting the groundwater depth with chaotic approach.

Acknowledgement

This project is supported by National Basic Research Program of China (Grant No. 2006CB403402).

References

- [1] P. Grassberger, I. Procaccia, Measuring the strangeness of strange attractors. *Physica D*, pp. 189-208,1983.
- [2] D. J. Farmer, J. J. Sidorowich, Predicting chaotic time series. *Phys Rev Lett*, pp.845-849,1987.
- [3] M. Casdagli, Nonlinear prediction of chaotic time series. *Physica D*, pp.335-341,1989.
- [4] A. Wolf, J. B. Swift, H. L. Swinney and J. A. Vastano, Determining Lyapunov exponents from a time series, *Physica D*, pp.285-317,1985.
- [5] J. Theiler, Generating surrogate data for time series with several simultaneously measured variables. *Phya.Rev.Lett*, pp.951-954,1992.
- [6] M. Palus, Testing for nonlinearity using redundancies: Quantitative and qualitative aspects. *Physica D*, pp. 186-205,1995.
- [7] B. Sivakumar, Chaos theory in hydrology: important issues and interpretations, *J. Hydrol*, pp.1-20,2000.
- [8] F. Takens, Detecting strange attractors in turbulence, In: D A Rand, L S Yong, Dynamical systems and turbulence. *Warwick, Lecture Notes in Mathematics*, pp.366 -381,1980.
- [9] Abarbanel HDI, R. Brown, J.B. Kadtke, Prediction in chaotic nonlinear systems: Methods for time series with broadband Fourier spectra, *Phya.Rev.Lett*, pp.783-807,1990.
- [10] L.Y. Cao, Practical method for determining the minimum embedding dimension of a scalar time series. *Physica D*, pp.43-50,1997.
- [11] M B Kennel, R Brown, and H D I Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction", *Phys.Rev.A*, pp.34-45,1992.