

# Analysis of Chaotic Characters for the Monthly Runoff Series at Fudedian Station in Liaohe Basin

Haiying Hu<sup>1</sup>, Huamao Huang<sup>2</sup>

<sup>1</sup>School of Civil and Transportation Engineering, South China University of Technology, Guangzhou, P. R. China

<sup>2</sup>School of Science, South China University of Technology, Guangzhou, P. R. China

Email: cthyhu@scut.edu.cn, schhm@scut.edu.cn

Received September, 2012

## ABSTRACT

The evolution of monthly runoff is affected both by climate environment and human activities, and its characteristics play an important role in runoff prediction and simulation. In this paper, the G-P and the principal component analysis method, which are both based on the reconstruction theory of the phase space, are used to study the chaos characteristics of the monthly runoff series at Fudedian station in Liaohe basin. The results show that the monthly runoff series have a large probability of chaos.

**Keywords:** Chaos Analysis; Saturated Correlation Dimension; Principal Component Analysis; Monthly Runoff Series

## 1. Introduction

Over the past two decades, the theory of chaos has shown its applicability to a wide class of scientific problems in various fields, including astronomy, biology, chemistry, ecology, engineering, and physics. The discovery of chaos, starting with Lorenz [1], provides a helpful discussion. Casdagli [2] presents an effective comparative study between chaotic and stochastic models. The discovery that even very simple deterministic systems can produce seemingly irregular outputs (e.g. time series) provided an impetus to seek possible chaotic patterns in real systems, through applications of a host of nonlinear dynamic and chaos methods.

Chaotic studies are necessary for a complex or nonlinear system. Based on the chaotic theory, it is well known that a random-like series can be attributed to deterministic rules. In fact, with the development of the chaotic theory, chaotic study in geosciences has not only been restricted to the basic but it has also come into the practical study stage, such as the characteristics of chaotic evolution, dynamical system reconstruction and nonlinear prediction. The discovery and continuing research of fractals and deterministic chaos gives hydrology research a new theorem, methods and opportunities. Therefore, the application of the concept of chaos theory in hydrology has been gaining considerable interest in recent times [3-5].

Runoff change is very complex, which is the combined result of climate environment and human activities. This paper investigates the existence of chaotic behavior in

the monthly runoff time series at the Fudedian station in Liaohe basin, China. Monthly data observed over a period of 39 years (1964-2004) are studied. Four nonlinear dynamic methods, with varying levels of complexity, are employed: (1) phase space reconstruction; (2) autocorrelation function; (3) correlation dimension method; and (4) principal component analysis method. These methods provide either direct or indirect identification of chaotic behaviors.

## 2. Study Area

Liaohe River located in the southwest of northeast China, is one of China's seven major rivers. The Liaohe River Basin lies between 40°30'-45°10'N and 117°00'-125°30'E. The climate of the Liaohe Basin is chiefly characterized by monsoonal weather conditions with the annual average temperature of 5°C to 11°C. The region of this study lies between Fudedian and Tieling Station in Liaohe basin, with an area of 13839 km<sup>2</sup>. Affected by the atmospheric circulation, precipitation in the study region is very unevenly distributed in time and space, which is mainly concentrated in the summer season with 60-70% of the annual precipitation. Rainstorm lasts shorter and rainfall concentrates. Consequently, floods in the study area are frequent and serious.

## 3. Methodologies

### 3.1. Phase Space Reconstruction

The first step in the process of chaotic analysis is that of

attempting to reconstruct the dynamics in phase-space. Monthly runoff series can be decided by many factors, and a number of variables will be required when we use equations to describe the process of its evolution. Generally, it is believed that the time series of a variable is the overall performance of the system, so the chaotic behavior of the system can be studied by investigating the single variable time series which can decide the long-term evolution of the system.

The nature of dynamics of a real-world system may be stochastic, deterministic or in between. This can be identified, at least as a preliminary indicator, by using the phase space concept. A method for reconstructing a phase-space from a single time series has been presented by Takens [6]. The Takens theorem states that the underlying dynamics can be fully recovered by building an  $m$ -dimensional space. The dynamics of a scalar time series  $(x_1, x_2, x_3, \dots, x_n)$  are embedded in the  $m$ -dimensional phase-space. The phase-space is defined by:

$$Y(i) = \{x(i), x(i + \tau), \dots, x(i + (m-1)\tau)\} \quad i = 1, 2, \dots, N \tag{1}$$

where  $N = n - (m-1)\tau$ ,  $\tau$  is the time delay,  $m$  is the embedding dimension, and  $x_i = \{x_1, x_2, x_3, \dots, x_n\}$  with  $N$ -observed values.

To determine the fractal dimension of a dynamical system, one should first decide the time delay and the embedding dimension for the correlation dimension analysis. Usually, the time delay is made with the help of the autocorrelation function and the embedding dimension can be obtained by the G-P method [7]. Correlation dimension can be used to estimate the sufficient embedding dimension and the estimated fractal dimension. A strange attractor could also be revealed in a chaotic system under a phase space reconstruction environment.

For a monthly runoff time series  $x_1, x_2, x_3, \dots, x_n$ , the reconstruction of the phase-space is defined by:

$$\begin{cases} Y_1 = \{x_1, x_{1+\tau}, \dots, x_{1+(m-1)\tau}\} \\ Y_2 = \{x_2, x_{2+\tau}, \dots, x_{2+(m-1)\tau}\} \\ \dots \\ Y_l = \{x_l, x_{l+\tau}, \dots, x_{l+(m-1)\tau}\} \end{cases} \tag{2}$$

where,  $l = n - (m-1)\tau$ .

### 3.2. Autocorrelation Function

The purpose of using autocorrelation function in chaotic analysis is to help select a proper time delay. The time delay can't be too small, or too large. The choice of a too small time delay causes information overlap among subsequent datum. On the other hand, the choice of a too

large time delay can lead to loss of all relevant information in phase space reconstruction because neighboring trajectories diverge. Therefore, the selection of a suitable time delay would allow unfolding of the attractor in the phase space while the components of any state vector must be as uncorrelated as possible.

A good choice of the time delay is essential for geometrical and numerical analysis of the phase space reconstruction. For the time series  $x_1, x_2, x_3, \dots, x_n$ , the Autocorrelation function is defined as follows:

$$C(\tau) = \frac{1}{n - \tau} \frac{\sum_{i=1}^{n-\tau} (x_i - \mu)(x_{i+\tau} - \mu)}{\sigma^2(x)} \tag{3}$$

where,  $\tau$  is the time delay,  $\mu$  and  $\sigma$  is the mean and standard deviation of the time series, respectively. It may be chosen as the lag time at which the auto-correlation becomes zero. However, considering various values of  $\tau$  demonstrates that the results do not show a strong dependence on the actual value chosen.

### 3.3. G-P Method

There are few distinct methods for computing fractal dimensions: relative dispersion analysis, correlation analysis, Fourier analysis and rescaled range analysis. To estimate the fractal dimension of a time series, the concept of correction dimension is useful. Correlation dimension is a nonlinear measure of the correlation between pairs lying on the attractor. Correlation dimension estimation is related to the relative frequency with which the attractor visits each covering element. Correlation dimension is generally a lower bound measure of the fractal dimension.

The correlation dimension can be measured by the G-P method suggested by Grassberger and Procaccia [7]. The main steps of the G-P method are:

1) For a time series  $x_1, x_2, x_3, \dots, x_n$ , selecting a proper time delay  $\tau$  and embedding dimension  $m$  and reconstructing the phase space.

2) The correlation integral  $C(r)$  is then calculated, which is given by

$$C(r) = \frac{1}{N^2} \sum_{\substack{i,j=1 \\ i \neq j}}^N \theta(r - |Y_i - Y_j|) \tag{4}$$

where  $\theta$  is the Heaviside step function,  $u = r - |Y_i - Y_j|$  with  $\theta(u) = 1$  for  $u \geq 0$ , and  $\theta(u) = 0$ , for  $u < 0$ .  $N$  is the number of points on the reconstructed attractor,  $r$  is the radius of the sphere centered on  $Y_i$  or  $Y_j$ .  $|Y_i - Y_j|$  is the Euclidean distance between the elements of  $Y_i$  and  $Y_j$ .

3) For a large number of points, the correlation integral follows the power law:

$$C(r) \propto \alpha r^D \tag{5}$$

where  $\alpha$  is a constant; and  $D$  is the correlation dimension, which represents generally good estimate of the fractal dimension of the attractor. The slope of  $\ln C(r)$  versus  $\ln r$  is given by:

$$D = \lim_{r \rightarrow 0} \frac{\ln C(r)}{\ln r} \tag{6}$$

The behavior of  $D$  provides one technique for determining the presence of chaos in a time series, such that (i) for stochastic processes,  $D$ , varies linearly with increasing  $m$ , without reaching a saturation value; (ii) for deterministic processes the value of  $D$  saturates after a certain value of  $m$ . The saturation value  $D$  is the fractal dimension of the attractor or the time series.

4) Gradually increasing the phase space dimension  $m$  and repeating the first three steps, calculating the estimation of systematic dimension  $D(m)$  until the  $D(m)$  is not obviously changed while increasing the phase dimension  $m$ . That means the  $D(m)$  has tended to the saturation.

### 3.4. Principal Component Analysis Method

The principal component analysis method is able to effectively identify the chaos and noise. In this method, the covariance of the trajectory line matrix  $Y_{lm}$  formed in the reconstruction phase space is calculated first, and then the eigenvalues  $\lambda_i (i = 1, 2, \dots, m)$  and the corresponding eigenvectors of the covariance matrix  $U_i (i = 1, 2, \dots, m)$  are calculated later. Eigenvalue can be arranged by size as follows:  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ .

The eigenvalues  $\lambda_i$  and the eigenvectors  $U_i$  are principal components. The sum of all the eigenvalues  $\gamma$  is obtained as:

$$\gamma = \sum_{i=1}^m \lambda_i \tag{7}$$

The map with the axes of the index  $i$  and  $\ln(\lambda_i/\gamma)$  is referred to the primary component spectrum, and the primary component spectrum of the chaotic sequence is a straight line or containing a linear portion, with negative slope and though the vertex. Thus the chaotic sequence can be identified by use of this property.

## 4. Analysis and Results

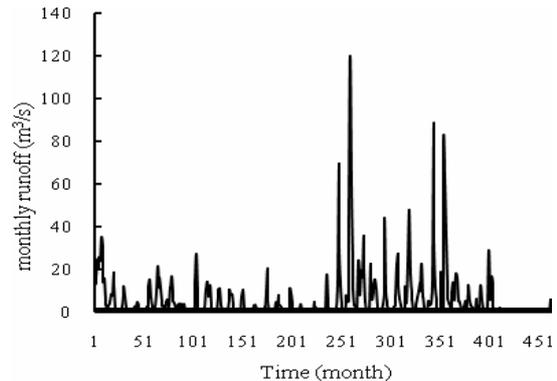
In this paper, the monthly runoff time series at Fudedian station in Liaohe basin, China is used and analyzed to investigate the possible existence of chaotic behavior. The average monthly runoff is  $6.0 \text{ m}^3/\text{s}$ . **Figure 1** shows the variation of monthly runoff obtained from Fudedian station.

### 4.1. Determination of the Delay Time

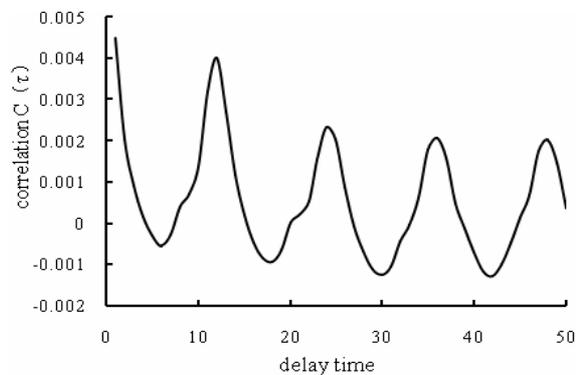
The quality of a phase space portrait depends on the time delay and therefore, a reasonable value of the time delay is desired. According to the formula of autocorrelation function method, monthly runoff time series at Fudedian station are analyzed and the correlation function value with the change of the delay time can be obtained as shown in **Figure 2**. It is seen from **Figure 2** that the correlation function value of monthly runoff time series at Fudedian station becomes zero for the first time while the delay time is 4. Therefore, the delay time in phase space reconstruction can be determined as  $\tau = 4$ .

### 4.2. Calculation of the Saturation Correlation Dimension

Saturation correlation dimension is calculated by the G-P method. The embedding dimension  $m$  is set to 4, 5, 6, 7, 8, 9, 10, 11, 12, and 13. The result can be seen in **Figures 3 and 4**. **Figure 3** shows the behavior of correlation function  $\ln C(r)$  against  $\ln r$  for values of increasing  $m$ . **Figure 4** shows that the correlation dimension values  $D(m)$  increase with increasing the embedding dimension values,



**Figure 1. The variation of monthly runoff at Fudedian station.**



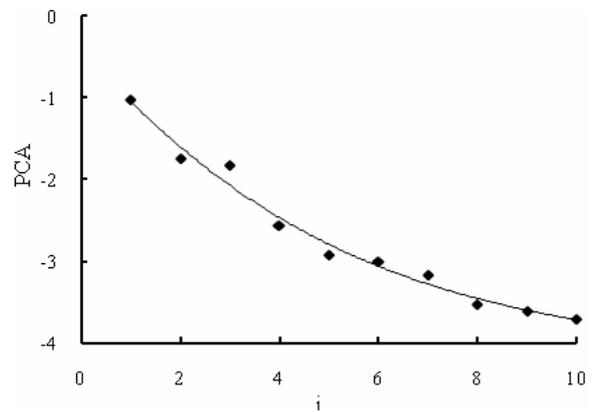
**Figure 2. Autocorrelation function plot for monthly runoff series of Fudedian Station.**

$m$ , up to a certain value and then reaches a plateau, where its value saturates. This is an indication of deterministic dynamics and the saturated correlation dimension is 2.66 for monthly runoff data at Fudedian Station. The nearest integer above the correlation dimension value ( $D = 3$ ) is taken as the minimum dimension of the phase space that can embed the attractor. The value of  $m$  at the saturation point ( $m = 10$ ) is supposed to provide the sufficient number of variables to describe the dynamics of the attractor.

### 4.3. Principal Component Analysis

The monthly runoff time series at Fudedian station are analyzed by the principal component analysis method and the relation of dimension  $i$  and  $PCA$  for monthly runoff series of Fudedian station can be seen in **Figure 5**.

It is seen from **Figure 5** that there is a straight line with negative slope in the relation diagram of embedded dimension  $i$  and  $PCA$  for monthly runoff series of Fudedian station. The result shows that monthly runoff series



**Figure 5.** The relation of dimension  $i$  and  $PCA$  for monthly runoff series of Fudedian station.

have chaotic characteristics. This further confirmed that the monthly runoff series is a chaotic sequence.

### 5. Conclusions

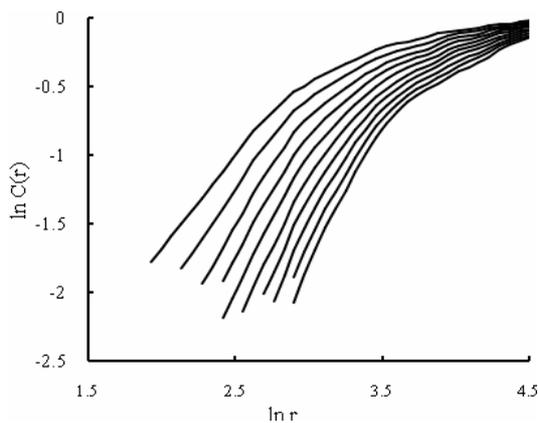
This paper investigates possible chaotic behaviors in the monthly runoff at the Fudedian station in Liaohe basin, China. The autocorrelation function, the correlation dimension and the principal component analysis method are used in the analysis. The results from these methods provide convincing indication and confirmation of the existence of a mild low-dimensional chaos in monthly runoff series for the data used. On the basis of the attractor dimension, the minimum number of variables essential to model the dynamics of the monthly runoff of the Fudedian station in Liaohe basin is identified as three and the number of variables sufficient is ten.

### 6. Acknowledgements

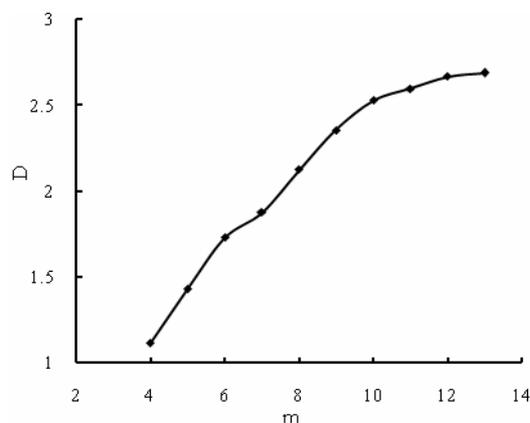
This paper is supported by the National Natural Sciences Foundation of China (grant number: 51209096), the Open Funding of the State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering of Hohai University (grant number: 2011490411), and the Fundamental Research Funds for the Central Universities, SCUT(grant number: 2011ZZ0017).

### REFERENCES

- [1] E. N. Lorenz, "Deterministic Nonperiodic Flow," *Journal of the Atmospheric Sciences*, Vol. 20, No. 2, 1963, pp. 130-141.  
[doi:10.1175/1520-0469\(1963\)020<0130:DNF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2)
- [2] M. Casdagli, "Chaos and Deterministic Versus Stochastic Nonlinear Modeling," *Journal of the Royal Statistical Society Series B*, Vol. 54, No. 2, 1991, pp. 303-328.
- [3] A. Hense, "On the Possible Existence of a Strange Attractor for the Southern Oscillation," *Beiträge zur Physik der*



**Figure 3.** Relational curves of  $\ln C(r) \sim \ln r$  for monthly runoff series of Fudedian Station.



**Figure 4.** Relational curves of  $D(m) \sim m$  for monthly runoff series of Fudedian Station.

- Atmosphäre*, Vol. 60, No. 1, 1987, pp. 34-47.
- [4] B. Sivakumar, "Chaos Theory in Hydrology," *Journal of Hydrology*, Vol. 227, 2000, pp. 1-20.
- [5] M. N. Islam and B. Sivakumar, "Characterization and Prediction of Runoff Dynamics:a Nonlinear Dynamical View," *Advances in Water Resources*, Vol. 25, No. 2, 2002, pp. 179-190. [doi:10.1016/S0309-1708\(01\)00053-7](https://doi.org/10.1016/S0309-1708(01)00053-7)
- [6] F. Takens, "Detecting Strange Attractors in Turbulence," In: Rand, D.A., Young, L.S. (Eds.), *Lectures Notes in Mathematics*, Vol. 898, Springer-Verlag, New York, 1981, pp. 366-381.
- [7] P. Grassberger and I. Procaccia, "Characterization of Strange Attractors," *Physical Review Letters*, Vol. 50, No. 5, 1983, pp. 346-349. [doi:10.1103/PhysRevLett.50.346](https://doi.org/10.1103/PhysRevLett.50.346)