Journal of Civil Engineering and Urbanism

Volume 4, Issue 3: 293-297 (2014) ISSN-2252-0430

Estimation of Daily Discharge of Baranduz River via Chaos Theory

Ahmad Pour Mohammad Aghdam[1], Reza Asheghi[2], Edris Merufinia[3*]

[1] *Department of Civil Engineering, College Water Engineering, Mahabad Branch, Islamic Azad University, Mahabad, Iran*
[2] *Department of Civil Engineering, Roudehen Branch, Islamic Azad University, Roudehen, Iran*
[3] *Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran*

*Corresponding author's E-mail: edris.marufynia@yahoo.com

ORIGINAL ARTICLE
Received 05 May. 2014
Accepted 13 May. 2014

ABSTRACT: The chaotic behavior of monthly precipitation time series is investigated using the phase-space reconstruction technique and the principal component analysis method. To reconstruct phase space, the time delay and embedding dimension are needed and for this purpose, average mutual information and algorithm of false nearest neighbors are used. The delay time for Baranduz River is calculated via the average mutual information method which is equal to 66. The most suitable inscribed dimension, by use of false nearest neighbors approach, is about 28. The correlation in time series of water flow is equal to 3.1 which require at least 3 variables to describe the system. The low value of correlation in daily scale is an indication of the existence of chaos in the water flow of Baranduz Chay River.

Keywords: Baranduz Chay River, Chaos Theory, Correlation Dimension, Delay Time

INTRODUCTION

Study of river flow is important for designing, exploitation and study of water supply systems. River flow processes is dynamic, nonlinear, extremely complex, and are affected by several interconnected physical variables, so that different methods including hydrologic modeling, time series analysis, artificial neural networks, fuzzy logic, neuro-fuzzy, genetic programming and recently chaos theory are used for river flow modeling (Alizade, 2013).

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 and Procaccia ,1983) the nonlinear prediction method (Sugihara and May, 1990) including deterministic versus stochastic diagram (Wolf et al,1985), the Lyapunov exponent method (Theiler et al., 1992), the Kolmogorov entropy method, the surrogate data method (Casdagli, 1991), and the linear and nonlinear redundancies (Prichard and Theiler, 1995).

A chaotic system is defined as a deterministic system in which small changes in the initial conditions may lead to completely different behavior in the future. Signal from the chaotic system is often, at first sight, indistinguishable from a random process, despite being sensitive to initial conditions) behavior of many systems was observed by many researchers for a number of decades, but was first described as such by Lorenz Wilks (Wilks, 1991). Sivakumar (2000) revealed that the presence of noise in the data does not significantly influence the correlation dimension estimates (though it significantly influence the prediction accuracy estimates). This suggests that the correlation dimension may be used as a preliminary indicator to identify the existence of chaos in the monthly precipitation time series.

Elshorbagy et al. (2004) has performed noise reduction and missing data estimation. Qingfang and Yuhua (2009) has developed a new local linear prediction model for chaotic stream flow series. Khan et al. (2006) verified the probability of chaotic signals existence in finite time series and showed that finite hydrologic data can also have chaotic behavior. Kocak et al. (2008), studied the prediction of monthly flow in Yamla dam using the local prediction of chaos approach in which the short time predictions showed better results in comparison with other methods. Damle and Yalcin (2007), began to predict the volume of floods using chaos theory and showed that the predicted amounts by chaos theory have a considerable accuracy in comparison with those predicted with time series models.

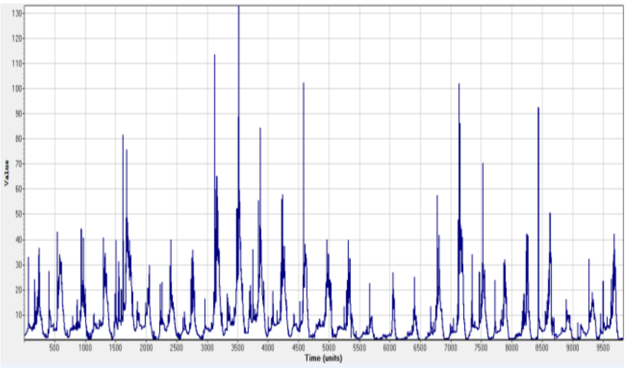

Figure 1. Time Series Diagram of water flow in Baranduz River

MATERIAL AND METHODS

It is relevant to note that the application of chaos identification methods, particularly the correlation

To cite this paper: Pour Mohammad Aghdam A, Asheghi R, Merufinia E. 2014. Estimation of Daily Discharge of Baranduz River via Chaos Theory. *J. Civil Eng. Urban.*, 4 (3): 293-297.
Journal homepage http://www.ojceu.ir/main/

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) and therefore are not reported herein. It is relevant to note, however, that the studies by Sivakumar reveal that the presence of noise in the data does not significantly influence the correlation dimension estimates (though it significantly influence the prediction accuracy estimates). This suggests that the correlation dimension may be used as a preliminary indicator to identify the existence of chaos in the monthly precipitation time series.

Reconstruction of phase space

While for stochastic systems there is no specific rule for phase-space reconstruction except some physical and/or statistical considerations, the optimal phase-space reconstruction of a deterministic uni/multivariate nonlinear system is obtained by "embedding" the dynamics of the process utilizing the so-called delay time method. The first step in the process of chaos theory is reconstructing the dynamics in phase space. The concept of phase-space is a powerful tool for characterizing dynamic system, because with a model and a set of appropriate variables, dynamics can represent a real world system as the geometry of a single moving point .A method for reconstructing phase-space from a sight time series has been presented by Taken (1981). The time series is assumed to be generated by a nonlinear dynamic system with m degrees of freedom. It is therefore necessary to construct an appropriate series of state vectors Y_t with delay coordinates in the m-dimensional phase space:

$$Y_t = \{x_t, x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-(m-1)\tau}\} \qquad (1)$$

Where τ is referred to as the delay time and for a digitized time series is a multiple of the sampling interval used, while m is termed the embedding dimension. If the dynamics of the system can be reduced to a set of deterministic laws, the trajectories of the system converge towards the subset of the phase space, called the attractor. For a scalar time series t x_t , where t = 1,2,…,N, the phase space can be reconstructed using the method of delays (Takens, 1980). where $j = 1, 2, \dots, N-(m-1)\tau/\Delta t$, m is the dimension of the vector Y_t , also called the embedding dimension, and τ is a delay time taken to be some suitable multiple of the sampling time Δt (Packard, 1980). Take a scalar time series $x_1, x_2, \dots, x_n$ in system phase-space as an example. Supposing its dimension d is 1, its dimension of embedding phase-space should be 3. If here $m = 4$, $x_1, x_2, \dots, x_4$ forms the first vector Y_1 of a four-dimensional state space and then moving right one step,$x_1, x_2, \dots, x_4, x_5$ forms the second vector Y_2. Just do in the same way, $Y_1, Y_2, Y_3, \dots, Y_l$ forms the time series of reconstruction phase-space.

Correlation dimension

Grassberger and Procaccia (1983) defined the correlation sum *C(r)* as:

$$c(r) = \frac{1}{N_{ref}} \sum_j^{N_{ref}} \frac{1}{N} \sum_i^N H\left(r - \left\|Y_i - Y_j\right\|\right); \qquad i \neq j \qquad (2)$$

where H is the Heaviside step function with H(u) = 1 for u> 0 , $H(u) = 0$ for $u \le 0$; N is the number of points in the vector time series *{**Y(t)**}*; N_{ref} $(\le N)$ is the number of reference points taken from the vector time series ***Y****(t); r* is the radius of sphere centered on either of the points *{**Yi** }*or *{**Yj**}*. The norm $\left\|Y_i - Y_j\right\|$ may be any of the three usual norms, the maximum norm, the diamond norm, or the Euclidean norm, of which the Euclidean norm is widely used. Correlation sums are calculated for a series of embedding dimensions. If an attractor for the system exists, then, for small *r*, it can be shown that:

$$c(r) \cong r^d \qquad (3)$$

Where d is the correlation exponent. It may be estimated by the slope of a straight line in the plot of log *(C(r))* vs. log *(r)* for each value of *m*. For random processes, *d* varies linearly with increasing *m* without reaching a saturation value, whereas for deterministic processes, the value of *d* levels off after a certain *m*. The saturation value of *d* is defined as the correlation dimension *D* of the attractor or the time series. The nearest integer above the saturation value of *d* provides the minimum number of embedding dimensions of the phase space necessary to model the dynamics of the attractor. When the optimal embedding dimension is not known, such as for example in a real time series, the correlation dimension is calculated for increasing embedding dimensions until it reaches a saturation value. It should be noted that in the plots of log *(C(r))* vs. log *r*, there are large statistical errors for small and large values of *r*. In between, however, there is a region in which the value of *d* remains reasonably constant. This region is called the scaling region.

The slope is generally estimated by a least-squares fit of a straight line over a certain range of *r*, called the scaling region. The presence/absence of chaos can be identified using the correlation exponent versus the embedding dimension plot. If the correlation exponent saturates and the saturation value is low, then the system is generally considered to exhibit low-dimensional 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 variables necessary to model the dynamics of the attractor. On the other hand, if the correlation exponent increases without limit with increase in the embedding dimension, the system under investigation is generally considered as stochastic.

Local prediction

A correct phase-space reconstruction in a dimension m facilities an interpretation of the underlying dynamics in the form of an m-dimensional map f_T, according to:

$$Y_{j+T} = f_T(Y_j) \qquad (4)$$

Where Y_j and Y_{j+T} are vectors of dimension m, describing the state of the system at times *j* (e.g. current state) and *j*+*T* (e.g. future state), respectively. The problem then is to find an appropriate expression for f_T (i.e. F_T). Local approximation entails the subdivision of the f_T domain into many subsets (neighborhoods), each of which identifies some approximations F_T, valid only in that same subset. In other words, the dynamics of the system is described step by step locally in the phase space. By considering a time series of a single variable, it is possible to reconstruct the phase space. Before applying reconstruction procedure it is necessary to have some information, embedding dimension, delay time, etc., concerning the attractor. One of the independent coordinates mentioned above is taken as the time series itself. The remaining coordinates are formed by its (m-1) lagged time series shifted by (m-1) multiples of the correlation time τ, at which correlation between coordinates become zero. It is assumed that the time series data are generated from a chaotic dynamical system in the v-dimensional space (v is dimension of attractor). In this m-dimensional space, prediction is performed by estimating the change of X_i with time. Considering the relation between the points X_t and X_{t+p} at time p later on the attractor is approximated by function F as:

$$X_{t+p} \cong F(X_t) \qquad (5)$$

In this prediction method, the change of X_t with time on the attractor is assumed to be the same as those of nearby points,(X_{t_h}, h=1,2,3,…,n). Here in, X_{t+p} is determined by the dth order polynomial F (X t) as follows:

$$x_{t+p} \cong f_0 + \sum_{k_1=0}^{m-1} f_{1k_1} X_{t-k_1\tau} + \sum_{\substack{k_2=k_1\\k_1=0}}^{m-1} f_{2k_1k_2} X_{t-k_1\tau} X_{t-k_2\tau} + \ldots + \sum_{\substack{k_d=k_{d-1}\\ \vdots \\ k_2=k_1\\k_1=0}}^{m-1} f_{dk_1k_2\ldots k_d} X_{t-k_1\tau} X_{t-k_2\tau} \ldots X_{t-k_d\tau} \qquad (6)$$

Using n of X_{Th} and $X_{T_{h+p}}$ for which the values are already known, the coefficients f are determined by solution of the following equation,

Where:

$$X \cong Af \qquad (7)$$

$$x = (x_{T_{1+p}} + x_{T_{2+p}}, \ldots, x_{T_{n+p}}) \qquad (8)$$

$$f = (f_0, f_{10}, f_{11}, \ldots, f_{1(m-1)}, f_{200}, \ldots, f_{d(m-1)(m-1)\ldots(m-1)}) \qquad (9)$$

And A is the $\frac{n\times(m+d)!}{m!d!}$ Jacobian matrix which in its explicit form is:

$$A=\begin{bmatrix} 1x_{T_1} & x_{T_{1-\tau}} & \ldots & x_{T_{1-(m-1)\tau}} & x_{T_1}^2 & \ldots & x_{T_{1-(m-1)\tau}}^d \\ 1x_{T_2} & x_{T_{2-\tau}} & \ldots & x_{T_{2-(m-1)\tau}} & x_{T_2}^2 & \ldots & x_{T_{2-(m-1)\tau}}^d \\ \vdots & & \vdots & & \vdots & & \vdots \\ 1x_{T_n} & x_{T_{n-\tau}} & \ldots & x_{T_{n-(m-1)\tau}} & x_{T_n}^2 & \ldots & x_{T_{n-(m-1)\tau}}^d \end{bmatrix} \qquad (10)$$

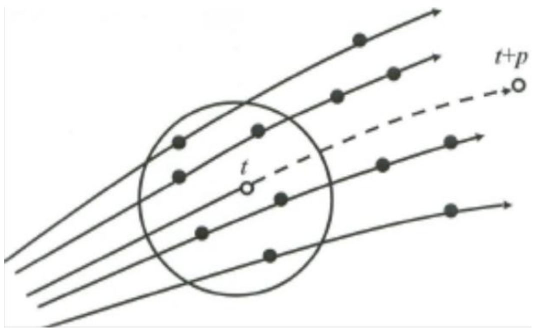

Figure 2. Local Prediction Mechanism and Model

Study Area and Data Used

In this survey, we have utilized water flow data of Baranduz Chay River. This river whose area is equal to 1203 $Km^3$is located in the northwest, between Urmia Lake and mutual border between Iran, Turkey and Iraq. The geographical extent of the Basin is about 44 45 'to 45 14' East and 37 06 'to 37 29' North latitude. The length of main water streams is about 75 Km. The maximum value of altitude and minimum latitude in exited section is, respectively, equal to 3500 and 1250 meters from free water's surface. This basin has four Hydrometer stations which are located in Babarud, Dizaj, Gasemlu and Bikaran. In figure 3 the conditions of sub-basin of Baranduz chay is displayed.

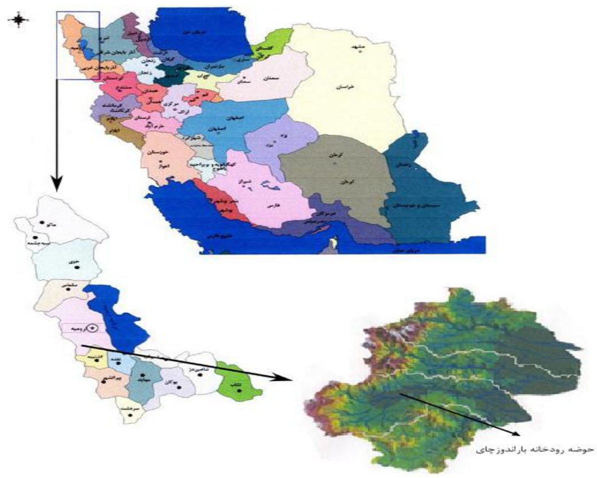

Figure 3. The Location of Baranduz river Basin in West Azarbaijan Province

Table 1. Statistical Characteristics of Baranduz chay River data (1746 data)

Data Type	Minimum	Maximum	Mean	Variance	Standard deviation	Coefficient of skewness	Elongation coefficient
Flow rate (S / m)	0.001	132.96	7.8751	92.294	9.607	3.0103	15.240

To cite this paper: Pour Mohammad Aghdam A, Asheghi R, Merufinia E. 2014. Estimation of Daily Discharge of Baranduz River via Chaos Theory. *J. Civil Eng. Urban.*, 4 (3): 293-297.
Journal homepage http://www.ojceu.ir/main/

RESULTS AND DICUSSION

The statistic and information about Baranduz Chay River is collected from the Regional Water Authority of Urmia which include a 5 year statistical period (from 2005 to 2010). In these data and statistics the sampling of water flow and daily suspended sediment is measured simultaneously (given the number of 1764 data). After removal of aback data and homogeneity test, for evaluation and assessment of time series of suspended sediment we used Visual Recurrence Analysis (VRA) software. For this purpose, we load the file linked the suspended sediment to VRA software.

The delay time in daily time scale is equal to be about 8, which is calculated by use of the average mutual information method. The value of the inscribed dimension with regard to figure 4 in daily time scale is equal to 8.

Table 2. Statistical forecasting of daily suspended sediment of Baranduz chay River components inscribed with various inscribed dimensions.

Daily Flow Discharge					
Inscribed Dimension ED	**RMSE**	**R^2**	**Inscribed Dimension ED**	**RMSE**	**R^2**
1			7	1.0982	0.982
2	0.9275	0.987	8	2.1623	0.937
3	0.8100	0.991	9	1.4924	0.967
4	0.8034	0.986	10	1.6060	0.962
5	0.9492	0.986	11	1.9826	0.945
6	1.3101	0.974	12	1.0611	0.983

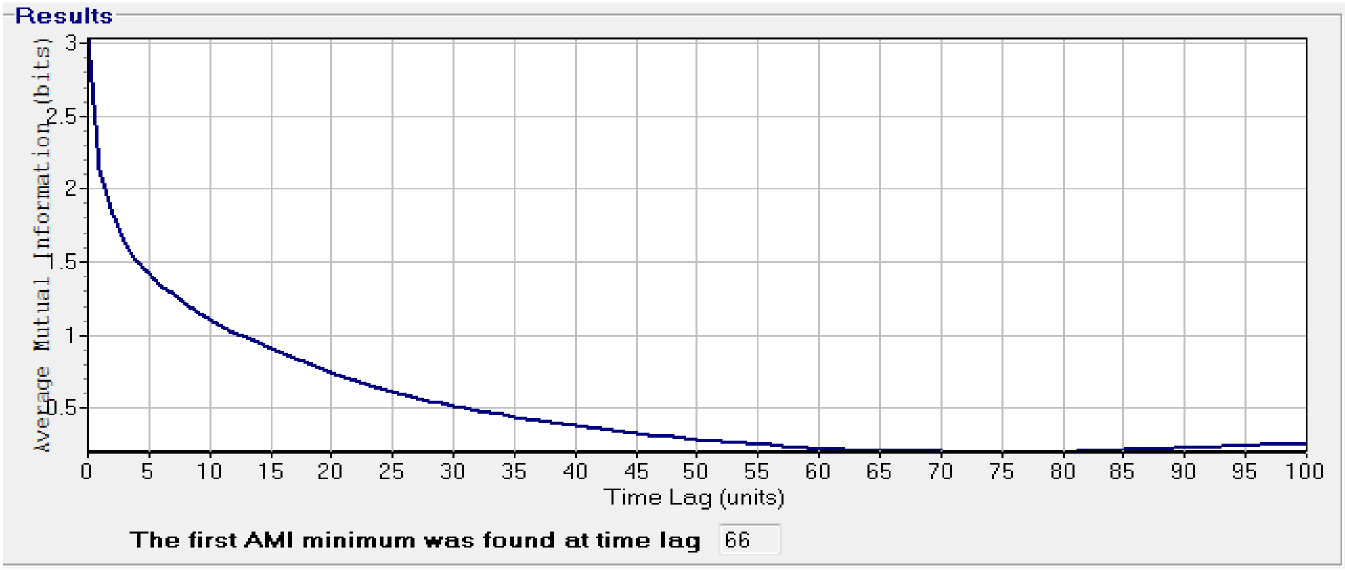

Figure 4. Mutual information function of flow of Baranduz chay in daily time scale

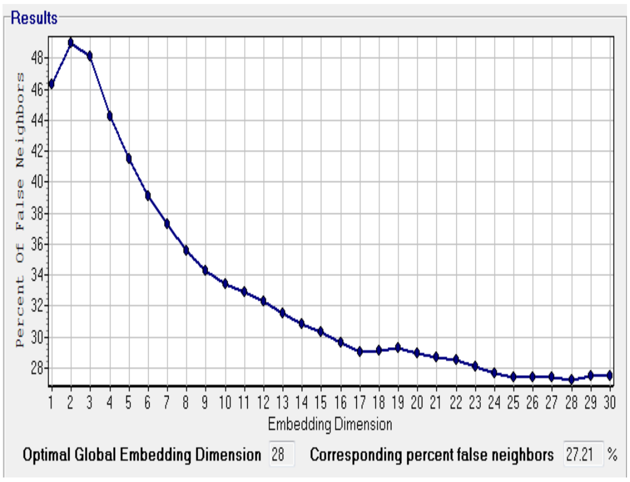

Figure 5. Percentage values of false nearest neighbors for different aspects of Baranduz Chay River

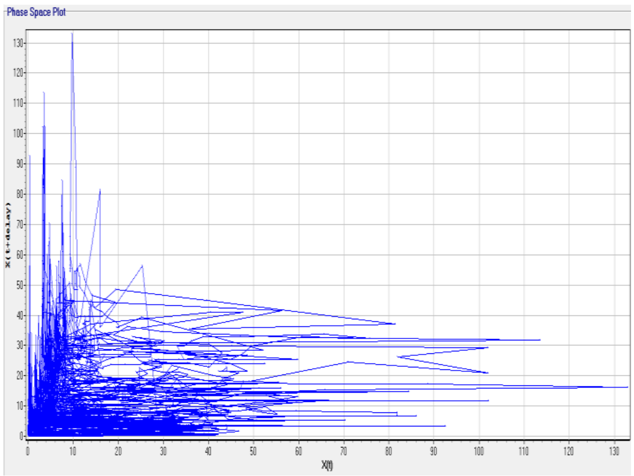

Figure 6. Space of Daily river flow of Baranduz chay River

To cite this paper: Pour Mohammad Aghdam A, Asheghi R, Merufinia E. 2014. Estimation of Daily Discharge of Baranduz River via Chaos Theory. *J. Civil Eng. Urban.*, 4 (3): 293-297.
Journal homepage http://www.ojceu.ir/main/

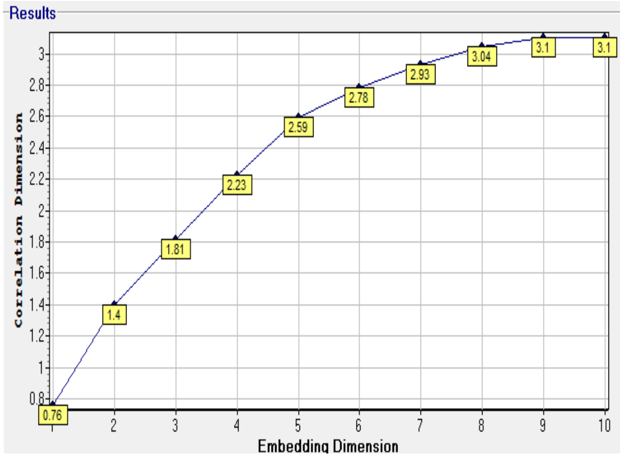

Figure 7. The variation diagram of correlation dimension in terms of increase in inscribed dimension of river's daily discharge

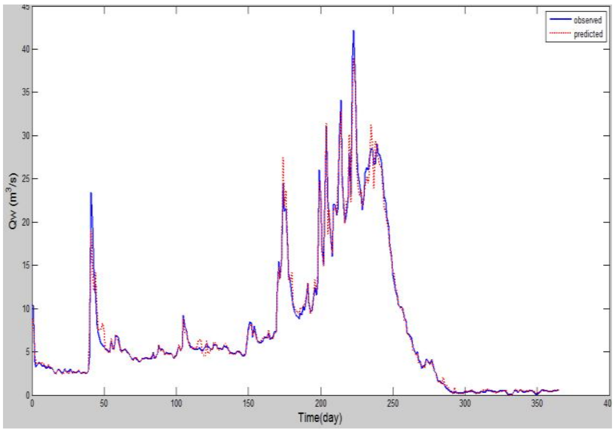

Figure 8. Comparison of the computational values with the observed daily flow rate of the Baranduz Chay River by use of chaos theory model

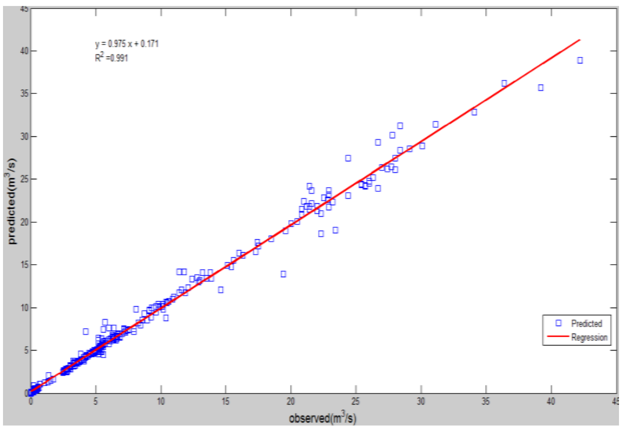

Figure 9. Scatter diagram of estimated and observed precipitation data and the regression line data

Summary of results

As it was mentioned, the value of delay time, which was calculated through the average mutual information method, is equal to 66 and the most suitable inscribed dimension is 28. The correlation dimension of time series of daily discharge was 3.1 which implies that at least three variables is needed to give an account of the system. The low value of correlation dimension is an indication of the existence of chaos in time series of daily discharge f Baranduz River, in terms of daily scale. Since the range of calculated inscribed dimension via false nearest neighbors in this survey is somehow high, in order to estimate we will use inscribed dimension values of $d<m \leq 2d+1$ (that is from 4 to 7 values of inscribed dimension). So, the result of such estimation will be the best case of inscribed dimension. During the estimation process of TIEASEN software in terms of table 2, for optimized inscribed dimension value of 4, the resulting values for RMSE and R^2 is equal to 0.8034 and 0.991, respectively.

REFERENCES

Alizade. A, (2013), Applied Hydrology, Ferdosi Mashhad Publication, Iran

Casdagli M. (1991), Chaos and deterministic versus stochastic non-linear modeling. J R Stat Soc, Ser B; 54(2):303-28.

Damle, C. and Yalcin, A. (2007). Flood Prediction Using Time Series Data Mining. Journal of Hydrology, 333: 305- 316.

Elshorbagy, A., Simonovic, S.P. & Panu, U.S. (2004). Estimation of missing stream flow data using principles of chaos theory, Journal of Hydrology, 255, 123-133.

Grassberger P, Procaccia I. (1983), Estimation of the Kolmogorov entropy from a chaotic signal. Phys Rev A;28:2591-3.

Khan, S., Ganguly, A. R. and Saigal, S., (2006). Detection And Predictive Modeling Of Chaos In Finite Hydrologycal Time Series. Nonlinear Processes in Geophysics, 12, 41- 53.

Kocak, K., Bali, A. and Bektasoglu, B., (2008). Prediction of Monthly Flows by Using Chaotic Approach. International Congress on River Basin Management, 22-24 March, Antalya, Turkey, 4 (117), 553-559.

Prichard D, Theiler J. (1995), Generalized redundancies for time series analysis. Physica D; 84:476-93.

Qingfang, M. & Yuhua, P. (2009). A new local linear prediction model for chaotic time series", Physics Letters A. 370: 465-470.

Sivakumar B. (2000), Chaos theory in hydrology: Important issues and interpretations. J Hydrol;227:1-20.

Sugihara G, May RM. (1990), Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. Nature; 344:734-41.

Theiler J, Eubank S, Longtin A, Galdrikian B, Farmer JD. (1992), Testing for nonlinearity in time series: The method of surrogate data. Physica D; 58:77-94.

Wilks, D.S. (1991). Representing serial correlation of meteorological events and forecasts in dynamic decision-analytic models. Monthly Weather Review .119:1640-1662.

Wolf A, Swift JB, Swinney HL, Vastao A. (1985), Determining Lyapunov exponents from d time series. Physica D; 16:285-317.

To cite this paper: Pour Mohammad Aghdam A, Asheghi R, Merufinia E. 2014. Estimation of Daily Discharge of Baranduz River via Chaos Theory. *J. Civil Eng. Urban.*, 4 (3): 293-297.
Journal homepage: http://www.ojceu.ir/main/