**Journal of Biodiversity and Environmental Sciences (JBES)**
ISSN: 2220-6663 (Print) 2222-3045 (Online)
Vol. 5, No. 1, p. 396-404, 2014
http://www.innspub.net

RESEARCH PAPER OPEN ACCESS

# Forecast the suspended load baron chay river by chaos theory

**Ahmad Pour Mohammad Aghdam[1], Edris Merufinia[2*], Hazhar Hadad[1]**

[1]*Department of Civil Engineering, College Water Engineering, Mahabad Branch , Islamic Azad University , Mahabad, Iran*

[2]*Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran*

[3]*Department of Civil Engineering, Mahabad Branch, Islamic Azad University, Mahabad, Iran*

Article published on July 14, 2014

**Key words:** Baron Chay River, Chaos theory, Correlation dimension, Suspended sediment.

**Abstract**

Estimation of rivers suspended load is one of the major issues of topics related to river engineering, reservoirs management, schemes and hydrologic projects. 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. Correlation dimension method is applied for investigating chaotic behavior of the daily suspended sediment statistics, which is the resultant of correlation dimensions, expresses chaotic behavior in the time series to illustrate efficiency of chaos theory for predicting suspended sediment, daily suspended sediment statistics of Maku Baron Chay River investigated for 5years. The delay time and optimum embedding dimension were obtained 8 and 5 respectively. The low amount of correlation dimension (d = 3) represents the chaotic behavior of sediment time series.

***Corresponding Author:** Edris Merufinia ✉ edris.marufynia@yahoo.com

## 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.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.

In order to investigate dynamic precipitation in various time scales, Sivakumar(2001) carried out a study by chaos theory. To accomplish this, the participation data of Leef River in Mississippi in 4 time scales of daily, second, fourth and eighth day is analyzed during a 25 years period and used the method of correlation dimension to demonstrate the dynamic behavior of participation. Limited correlation dimension for each 4 scales is 82.4, 26.5, 42.6, and 84.8 respectively which suggests the possibility of chaotic behavior in each 4 scales.

Rigonda *et al.* (2004) discussed and investigated the data of three rivers in daily, third, and fifth day basis in regard to the ability of being chaotic. A number of data series showed chaotic behavior and investigate random behavior.

Khan *et al.* (2005) studied the possibility of presence of chaotic signals of conceived time series from hydrological systems. In this study the amounts of daily discharge of Arkansas and Colorado Rivers during 4 years, are investigated which none of raw time series showed chaos. This study illustrates that hydrological data may or may not possess a certain chaos feature.

Ng *et al.* (2007) studied the application of chaos analysis techniques on daily noised flow series. In this study, they investigated the effect of disorder on the complexity of a system by using the concepts of chaos visually and considering the quantity, and showed that the presence of disorder adds on the complexity of time series analysis.

Ghorbani *et al.*(2011), by studying the data of suspending sediments of Lighvan River during 21 statistical years, by using chaos theory, calculated the delay time (62 days) and based on the conceived delay time and for surrounding dimension of 34, the correlation dimension is calculated equal to 1.6, which is a reason to chaotic behavior of the Lighvan River's sediment amount.

The methods available thus far are the correlation dimension method (Grassberger, 1983), the nonlinear prediction method (Farmer, 1987) including deterministic versus stochastic diagram (Casdagli, 1991), the Lyapunov exponent method (Wolf, 1985), the surrogate data method (Theiler, 1992), and the linear and nonlinear redundancies (Palus, 1995; Prichard, 1995). Among these the correlation dimension method has been the most widely used one for the investigation of deterministic chaos in hydrological phenomena (Hense, 1987; Puente, 1996; Sangoyomi, 1996; Sivakumar, 2000). In the present study, the correlation dimension method is employed, and the presence of a low-dimensional attractor (a geometric object which characterizes the long-term behavior of a system in the phase space) is taken as an indication of chaos.

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 (1991)). During the past two decades, the theory of chaos showed its applicability in solving a wide class of problems in many areas of natural sciences. The discovery that very simple deterministic systems can produce seemingly irregular time series pushed researchers to try identifying such systems and apply chaos theory in order to predict their behavior. However, chaotic signal analysis is still a novel approach in many areas related to civil engineering and to water-related problems in particular. In literature, many researchers have investigated the stream flow modeling with chaos theory. The papers by Jayawardena& Lai (1994); Porporato&Ridolfi (1997); Stehlik (1999) have shown the presence of low dimensional deterministic behavior in the stream flow process. Islam & Sivakumar (2002), Lisi& Villi (2001), have suggested the possibility of accurate stream flow predictions using nonlinear deterministic approaches. Elshorbagy*et al.*(2002) has performed noise reduction and missing data estimation Qingfang&Yuhua (2007) has developed a new local linear prediction model for chaotic stream flow series.

The goals of the study includes determining the chaotic potential of discharge data in daily scale, modeling the flow by chaos theory, and predicting the amount of flow discharge using chaos theory. this study is focused on the chaotic behavior of river flow and then by using chaos theory it estimates the river's flow.

## Material & methods

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) 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 Takens (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:

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

Where $\tau$ 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 $j = 1, 2, \ldots, N - (m-1)\tau/\Delta t$, $m$ is the dimension of the vector $Y_t$, also called the embedding dimension, and $\tau$ is a delay time taken to be some suitable multiple of the sampling time $\Delta t$(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&Procaccia (1983) defined the correlation sum*C(r)* as:

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

where H is the Heaviside step function with H (u) = 1 for u$> 0$, *H(u)* = 0 for *u* ≤ 0; *N* is the number of points in the vector time series *{Y(t)}*; $N_{ref}$ ($\leq N$) 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 $\|Y_i - Y_j\|$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:

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

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 $X_t$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:

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

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 $d$th order polynomial F($Xt$) as follows:

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

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:

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

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

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

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

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

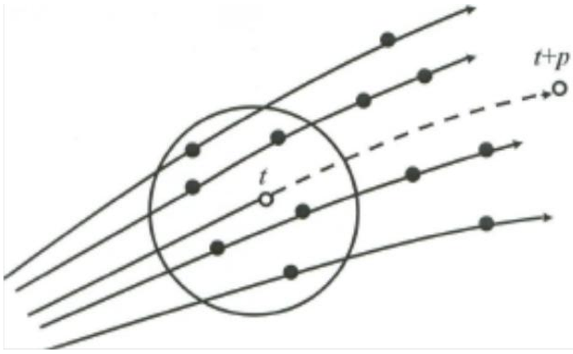


**Fig. 1.** Local Prediction Mechanism and Model.

*Study Area and Data Used*

Maku is a city in the West Azerbaijan Province, Iran. At the 2006 census, its population was 41,865, in 10,428 families. It is situated 22 kilometers (14 mi) from the Turkish border in a mountain gorge at an altitude of 1634 meters. The Zangmar River cuts through the city. Maku Free Trade and Industrial Zone is Iran's largest and the world's second largest free trade zone and will encompass an area of 5000 square km when it will open in 2011.

Plantation Hydrometery Station was established on the Baron Chay River in 1989 Years.Its height from sea level is 1712 meters. Surface catchment basin of 654 square kilometers and is home to the Caspian Sea.


**Fig. 2.** Location of Baron Chay Maku Basin.

Fig. 2 shows the variations of daily river flow time series and Table 1 presents some of the important statistics of the time series.

## Results

In order to reconstruct the original phase space, we first estimate reconstruction parameters, the delay times $\tau$ and embedding dimension $m$. The method used for the determination of the sufficient embedding dimension is based on the calculation of the percentage of false nearest-neighbors for the time series.

**Table 1.** Statistics of Daily River flow data from Baron Chay River

| Data Type | Minimum | Maximum | Mean | Variance | Standard deviation | The coefficient of skewness | Elongation factor |
|---|---|---|---|---|---|---|---|
| $Q(\frac{m^3}{Sec})$ | 0.018 | 14.56 | 2.84 | 8.40 | 2.89 | 1.46 | 1.84 |
| Suspended sediment $(\frac{ton}{day})$ | 0.00 | 7901.41 | 2132.68 | $1.7 \times 10^6$ | 1305.78 | 0.98 | 1.77 |

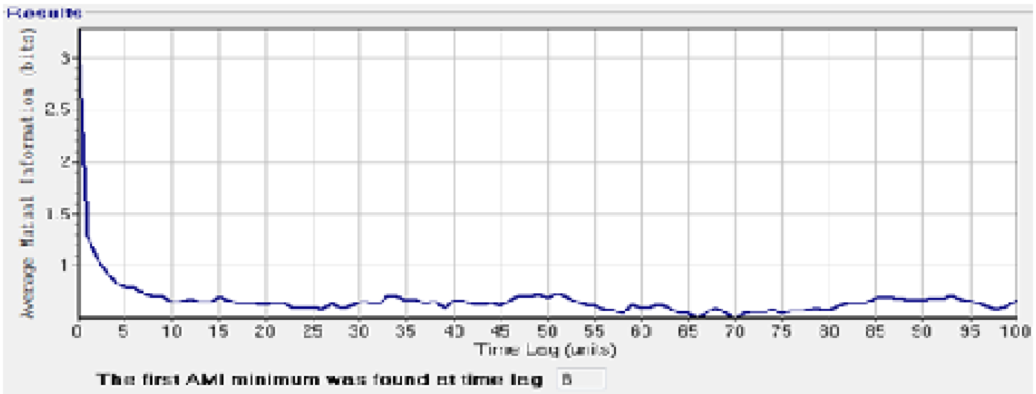


**Fig. 4.** Mutual information function of suspended sediment in the from Baron Chay River daily time scale.

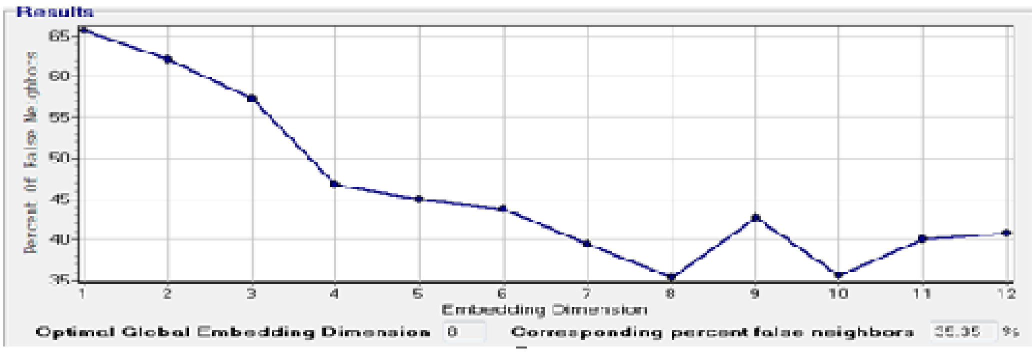


**Fig. 5.** False nearest neighbors for different values offrom Baron Chay River.

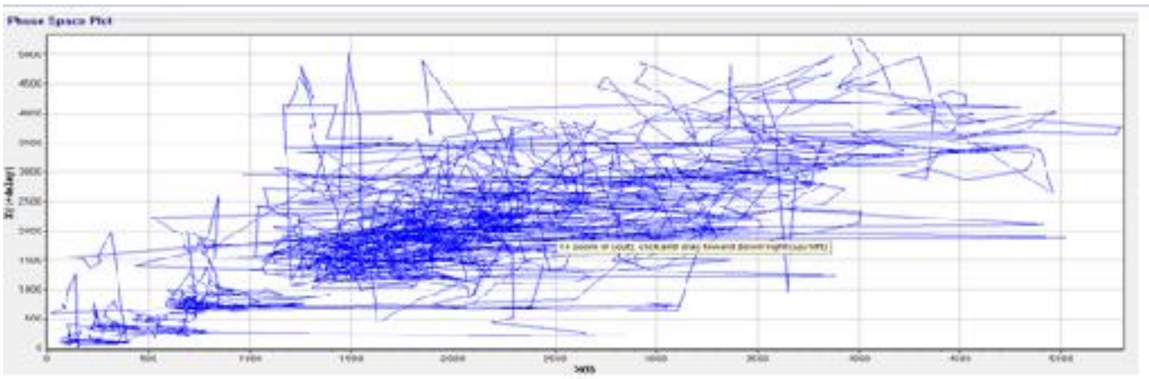


**Fig. 6.** The daily suspended sedimentfrom Baron Chay River.

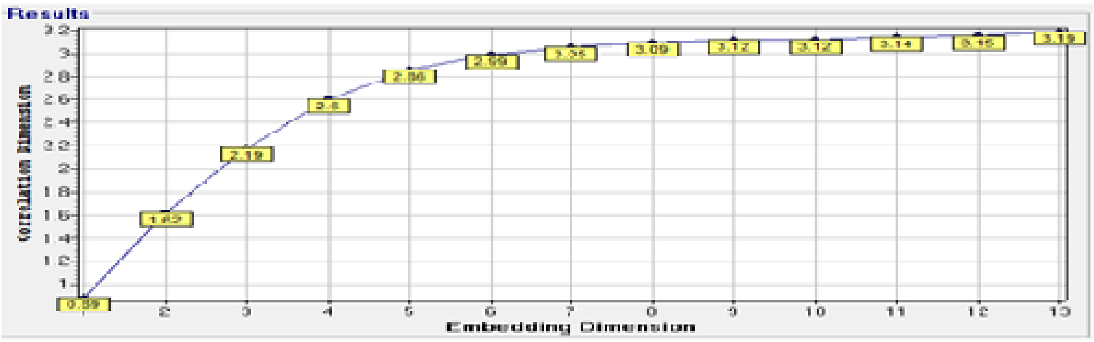


**Fig. 7.** Changes correlated with an increase in the surrounding plot of suspended sediment.

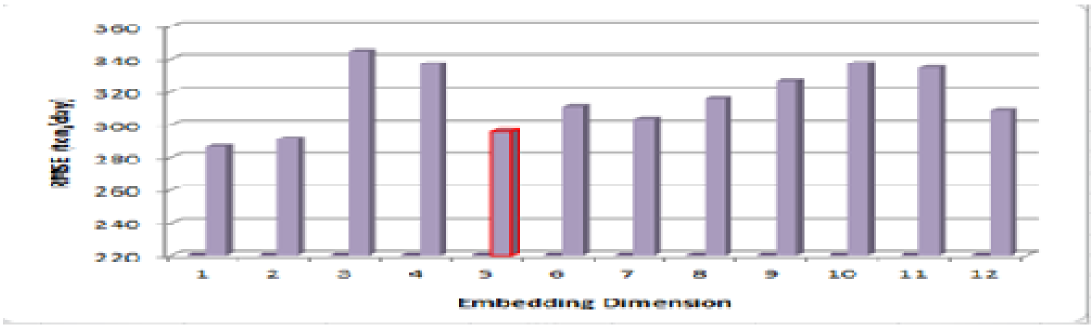


**Fig. 8.** RMSE graphs for various dimensions of the inscribed.

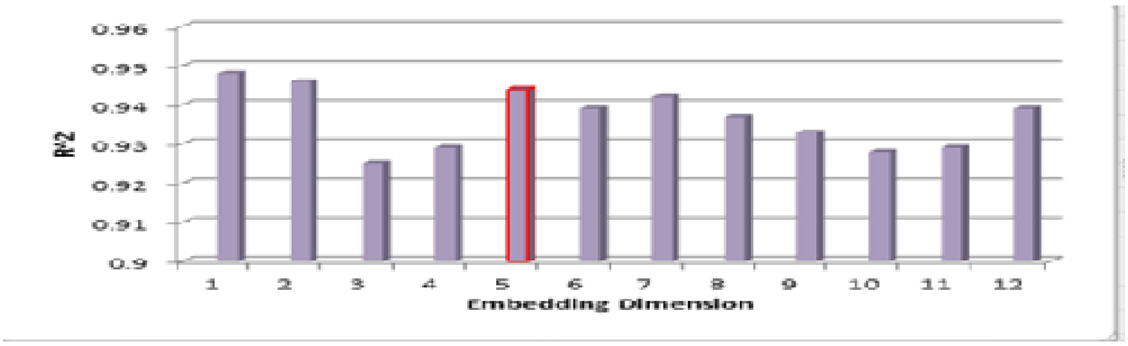


**Fig. 9.** $R^2$graphs for various dimensions of the inscribed

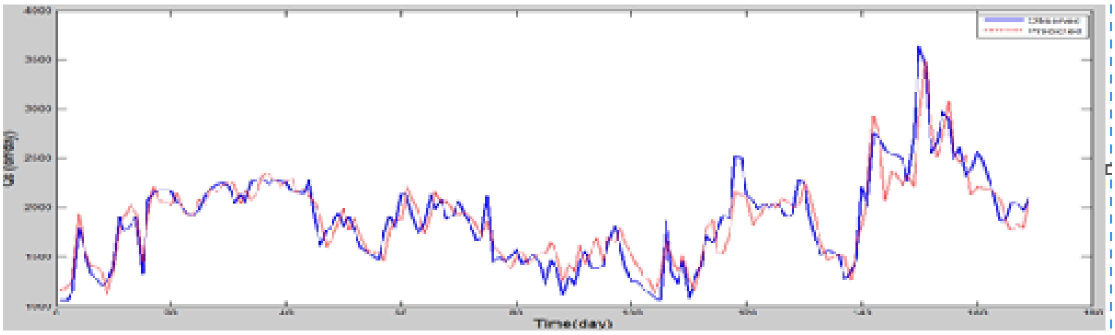


**Fig. 10.** Comparison of the calculated and observed sediment graph from Baron Chay River with chaos theory model.

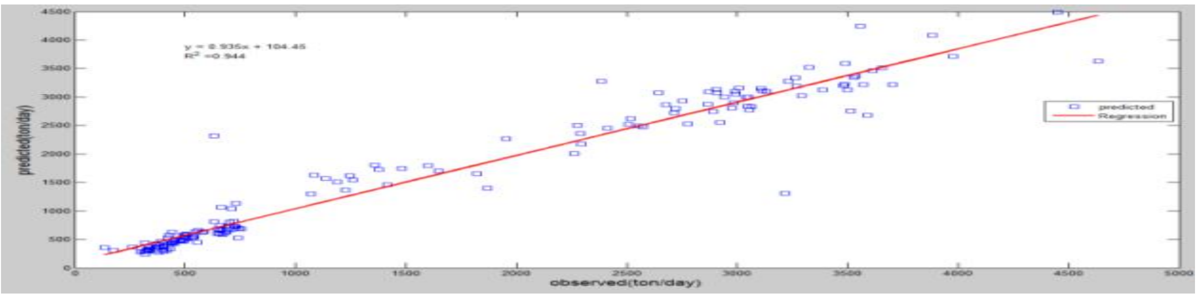


**Fig. 11.** The scatter plot of estimated and observed precipitation data and the regression line data.

**Table 2.** Statistical forecasting of daily suspended sediment component dimensions, inscribed with the from Baron Chay River.

| daily suspended sediment | | | | | |
|---|---|---|---|---|---|
| $R^2$ | RMSE | ED | $R^2$ | RMSE | ED |
| 0.942 | 303.2 | 7 | 0.948 | 286.88 | 1 |
| 0.937 | 316.21 | 8 | 0.946 | 291.57 | 2 |
| 0.933 | 326.42 | 9 | 0.925 | 345.11 | 3 |
| 0.928 | 337.19 | 10 | 0.929 | 336.85 | 4 |
| 0.929 | 334.54 | 11 | 0.944 | 295.84 | 5 |
| 0.939 | 308.92 | 12 | 0.939 | 310.87 | 6 |

## Conclusion and discussion

Delay time obtained by average mutual information method for baroon chay river of Maku. The best appropriate embedding dimension determined by false nearest neighbor method equal to 5. Correlation dimension for time series was 3.1 it is mean that a number of the necessary variable for description the system is 3. A little of obtained correlation dimension at daily time scale display existence of chaos on suspended sediment time series of the Baron chay river.

As at the this paper outcome embedding dimension is somewhat high, to do prediction process use from embedding dimension d≤ m<d+1 (from 4 to 7) that result the best prediction on the appropriate embedding dimension. RMSE and $R^2$calculated 295.84 and 0.944 respectively at Table (2) to the optimum embedding dimension (=5) by Tisean software.

## References

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

**Elshorbagy A**. Simonovic, S.P. &Panu, U.S. 2002.".Noise reduction in chaotic hydrologic time series".facts and doubts, Journal of Hydrology. **256,** 147-165

**Grassberger P, Procaccia I.** 1983. Measuring the strangeness of strange attractors.Physica ;**9,** 189-208.

**Ghorbani MA, KisiO, Aalinezhad M**. 2011. A Probe into the Chaotic Nature of Daily Stream Flow time Series by Correlation Dimension and

Largest Lyapanov Methods", Applied Mathematical Modelling, **34,** 4050-4057.

**Farmer DJ, Sidorowich JJ.** 1987. Predicting chaotic time serie. Phys Rev Lett; **59,** 845-8.

**Hense A**. 1987. On the possible existence of a strange attractor for the southern oscillation. BeitrPhys Atmos; **60(1)**, 34-47.

**Jayawardena AW, Lai F.** 1994. Analysis and prediction of chaos in rainfall and stream flow time series. Journal of Hydrology. **153,** 23-52.

**Islam MN, Sivakumar B**. 2002. Characterization and prediction of runoff dynamics: a nonlinear dynamical view .Advances in Water Resources. **25,** 179-190.

**Koçak K.** 1997. Application of Local Prediction Model to Water Level Data. A satellite Conference to the 51st ISI Session in Istanbul, Turkey: Water and Statistics, Ankara-Turkey, pp: 185-193

**Lisi F, Villi V.** 2001. Chaotic forecasting of discharge time series: A case study, Journal of the American Water Resources Association. **37(2)**, 271- 279.

**Ng WW, Panu US, Lennox WC.** 2007. "Chaosbased Analytical techniques for daily extreme hydrological observations", Journal ofHydrology, **342,** 17–41.

**Palus M.** 1995. "Testing for nonlinearity using redundancies: Quantitative and qualitative aspects". Physica, **80**, 186-205.

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

**Puente CE, Obregon N**. 1996. A deterministic geometric representation of temporal rainfall: Results for a storm in Boston. Water Resources, **32(9)**, 2825-39.

**Porporato A, Ridolfi L**. 1997. Nonlinear analysis of river flow time sequences, Water Resources Research. **33(6),** 1353-1367.

**Qingfang M, Yuhua P**. 2007. A new local linear prediction model for chaotic time series, Physics Letters A. **370**, 465-470.

**Regonda SK, Sivakumar B, Jain A.** 2004. Temporal scaling in river flow: can it bechaotic?, Hydrological Sciences Journal- desSciences Hydrologiques, **49(3)**, 373-385.

**Stehlík J.** 1999. Deterministic chaos in runoff series. Journal of Hydrology and Hydrodynamics. **47(4),** 271-287.

**Sangoyomi TB, Lall U, Abarbanel HD**. 1996. Nonlinear dynamics of the Great Salt Lake: Dimension estimation. Water Resources; **32(1)**, 149-59.

**Sivakumar B**. 2000. Chaos theory in hydrology: Important issues and interpretations. Journal of Hydrology; **227**, 1-20.

**Takens F.** 1980. Detecting strange attractors in turbulence, in Dynamical Systems and Turbulence, Lecture Notes in Mathematics, 898, edited by Rand DA, Young LS. Springer-Verlag, New York. 366-81.

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

**Wolf A, Swift JB, Swinney HL, Vastao A.** 1985.Determining Lyapunov exponents from d time series. Physics **16**, 285-317.

**Wilks DS.** 1991. Representing serial correlation of meteorological events and forecasts in dynamic decision-analytic models. Monthly Weather Review **119**, 1640-1662.