Shahid Beheshti University

Sustainable Earth Review

Journal homepage: http://sustainearth.sbu.ac.ir

Chaotic nature of river flow time series: upstream to downstream

Leila Malekani[a*], Mohammad Ali Ghorbani[b], Hakan Tongal[c], Habibeh Abbasi[a]

[a] Faculty of Technical and Engineering of Marand, Department of Civil Engineering, Tabriz University, Tabriz, Iran
[b] Faculty of Agriculture, Department of Water Engineering, Tabriz University, Tabriz, Iran
[c] Faculty of Engineering, Department of Civil Engineering, Süleyman Demirel University, Turkey

A B S T R A C T

Study the dynamics of runoff in the watershed and how influence of the morphological characteristics is very important in solving water engineering problems. In this paper, the dynamic behavior of the runoff of the four basins located in the Iran, on basin in Turkey and two basins located in United States using chaos theory is studied. There is several stream gauge station in all selected basin. For all basin time series of runoff extraction and then chaotic parameters are calculated.

The delay time is determined using the mutual information method, the embedding dimension and correlation dimension is estimated using the false nearest neighbor algorithm and D2 algorithm, respectively.

Analysis of the results showed that the correlation dimension has increased from upstream to downstream which constitutes the increasing complexity and number of required variables for modeling. Delay times and embedding dimension do not have relationship with position of the station and values are fairly similar to each other. It is clear that we need more studies to reach a general conclusion on this point.

A R T I C L E I N F O

Keywords:
Correlation dimension
Delay times
Dynamic
Embedding dimension
Runoff

Article history:
Received: 24 Sep 2020
Accepted: 28 Oct 2020

*corresponding author
E-mail address:
lmalekani@tabrizu.ac.ir
(L. Malekani)

1. Introduction

Adequate knowledge of river flow plays a significant role in hydrology, for proper planning and management of water resources, and environment. River flow is a dynamic, nonlinear, complex system in nature and often influenced by the physical characteristics of the watershed. Many nonlinear techniques include chaos theory, artificial intelligence systems, and catastrophe theory are used for the river flow modeling (see for details, Kumar and Foufoula-Georgiou (1997), ASCE Task Committee (2000a, 2000b), Sivakumar (2000), Labat et al. (2005), Ghorbani et al. (2010), Sivakumar and Berndtsson (2010), and Khatibi et al. (2012)). The fundamental concept in chaos theory is that complex and random-looking behaviors are not necessarily the result of actual random systems but can also be from simple nonlinear deterministic systems with sensitive dependence on initial conditions (Lorenz, 1963; Wilks, 1991).

Chaos theory has found specific applications in investigating the presence of chaotic behavior (low or high dimensional deterministic) in the river flow and related processes because such complex systems can be represented in a simplified form through chaotic approach. Also this approach determines the complexity level of a system that provides the parameters and required information for subsequent predictive analyses. (Jayawardena and Lai, 1994; Porporato and Ridolfi, 1997; Liu et al. (1998); Krasovskaia et al. (1999); Sivakumar et al. (2001); Ghorbani et al. (2010), Lisi and Villi (2001); Tang and Hu (2010); Xu et al. (2013); Li et al. (2013)). The research by Ng et al. (2007) focused on the application of chaotic analytical techniques to daily stream flow data of the Saugeen River in Ontario, Canada and hydrologic series comprised of outliers. Different techniques and concepts of chaotic theory were adopted to enhance the understanding of the phenomena of outliers.

This paper illustrated the use of the autocorrelation functions, mutual information, power spectrum analysis, phase space reconstruction, correlation dimension, surrogate tests, and Hurst coefficients for the analysis of chaotic systems. Based on the results of analyses, the analyzed series exhibited random-like fluctuations. The existence of outliers was found to increase the complexity of the analyzed series. The differentiation of a highly complex system from a random process, and the impact of outliers on the complexity of a system were quantitatively as well as visually presented from a chaotic perspective. Khatibi et al. (2012) investigated the existence of chaotic behavior in the river stage and discharge time series observed at the Sogutluhan hydrometric station, Turkey. Five nonlinear dynamic methods were employed: (1) phase space reconstruction; (2) false nearest neighbor (FNN) algorithm; (3) correlation dimension method; (4) Lyapunov exponent method; and (5) local approximation method. As the discharge data used in this study are produced by rating directly gauged stage time series, it becomes feasible to investigate any interference triggered by chaotic signals with the rating. The results indicated the existence of low-dimensional chaos in the two time series. They also suggested that the rating of the stage time series to obtain the discharge time series amplified significantly the fluctuations in the latter in the presence of chaotic signals. Kędra (2014) used several independent methods, techniques and tools for studying daily discharge from two selected gauging stations of the mountain river in southern Poland, the Raba River to determine the underlying dynamics of river flow. The outcomes indicated that the investigated discharge dynamics is not random but deterministic. Moreover, the results completely confirm the nonlinear deterministic chaotic nature of the studied process. All mentioned studies demonstrated the chaotic behavior of river flow at the certain stations on the river. In the case of a river network, the author is not aware of any attempts to study the nature of the river flow upstream to downstream. So the main aim of this study is to detect the possibility of presence of chaotic behavior in the consecutive gauge stations of seven basins from different areas of the world. Different estimators are adopted to capture the presence of deterministic and chaotic dynamics in all gauge stations, namely average mutual information, false nearest neighbors, correlation dimension. The phase space corresponding to the underlying dynamical system is reconstructed by means of optimal delay time and embedding dimension, obtained from first local minimum of the average mutual information and from the false nearest neighbors techniques, respectively. The main signatures of chaotic dynamics will be investigated by means of the correlation dimension.

2. Material and Methods

2.1. Methodology

Chaos theory or nonlinear time series analysis involves a host of methods, essentially based on the phase space reconstruction of the process, from scalar or multivariate measurements of its physical observables (Manlio et al., 2013).

2.2. Reconstruction of phase space

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 (Ghorbani et al., 2012). A method for reconstructing phase-space from an observed time series has been presented by Takens (1981). Using a single variable X_i ($i=1, 2, \ldots, N$) the phase space (multi-dimensional) can be reconstructed using the method of delays (Takens, 1981) as following:

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

Where $j = 1,2,\ldots,N-(m-1)\,\tau/\Delta t$; τ is delay time; m is the dimension of the Y_j or embedding dimension and Δt is sampling time. 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 (Domenico et al., 2012).

The time delay τ can be defined by the average mutual information method (Fraser and Swinney, 1986). This method defines how the measurements *X(t)* at time *t* are connected in an information theoretic fashion to measurements $X(t+\tau)$ at time $t+\tau$ (Abarbanel, 1996). The Average Mutual Information (AMI) is defined as:

$$I(\tau)=\sum_{i,i+\tau} P(X_i, X_{i+\tau}) \log_2\left[\frac{P(X_i, X_{i+\tau})}{P(X_i)\,P(X_{i+\tau})}\right] \quad (2)$$

Where *i* is total number of samples. $P(X_i)$ and $P(X_{i+\tau})$ are individual probabilities for the measurements of X_i and $X_{i+\tau}$.$P(X_i, X_{i+\tau})$ are the joint probability density for measurements $P(X_i)$ and $P(X_{i+\tau})$. The appropriate time delay τ is defined as the first minimum of the average mutual information $I(\tau)$(Ghorbani et al., 2012).

2.3. Correlation dimension

Correlation dimension method is one of the most efficient methods to determine the presence of chaos (Jani et al., 2013). This method is used as a fractal dimension quantifier and is based on the correlation integral (Grassberger & Procaccia, 1983).

For an *m*-dimensional phase space the correlation function $C_m(r)$ is given by Theiler (1986) as:

$$C_m(r)=\lim_{N\to\infty}\frac{2}{N(N-1)}\sum_{i,j=1}^{N} H(r-|Y_i-Y_j|) \quad (3)$$

where H is the Heaviside step function, with $H(u)=1$ for $u\geq 0$, and $H(u)=0$ for $u\leq 0$, where $u=r-|Y_i-Y_j|$, N is the number of points on the reconstructed attractor, r is the radius of the sphere centered on Y_i or Y_j (Jani et al., 2013). If the time series is characterized by an attractor, then for positive values of r the correlation function $C_m(r)$ is related to the radius r. For stochastic time series $C_m(r)\propto r^m$ holds, whereas for chaotic time series the correlation function scales with r as:

$$C_m(r)\propto r^{D_2} \quad (4)$$

Where D_2, called *correlation exponent.* The correlation exponent is defined by:

$$D_2=\lim_{r\to 0}\frac{\ln C_m(r)}{\ln r} \quad (5)$$

And can reliably be estimated as the slope in the $\ln C_m(r)$ vs. $\ln(r)$ plot. The slope can be computed by the least-squares fit of a straight line (i.e. scaling region) over a length scales of r. According to Grassberger-Procassia algorithm (1983), in case of deterministic data set the plot of ‘m’ versus ‘D_2’ should be a straight line parallel to embedding dimension, in case of stochastic data set, it should be straight line sloping 45 degrees to *x* and *y* axis. For a chaotic system, the correlation exponent initially increases but finally saturates after an especial embedding dimension. The saturation value of the correlation exponent is defined as the correlation dimension. If the value of correlation dimension is relatively small and fractal, the system can be considered as low-dimensional deterministic chaotic dynamic. Sivakumar and Singh (2012) classify systems with correlation dimension to the four groups of low, medium, high-dimensional and unidentifiable. The associated dimensionalities are as follows: (1) low-dimensional, with d ≤ 3.0; (2) medium-dimensional, with 3.0<d ≤ 6.0; (3) high-dimensional, with d >6.0; and (4) unidentifiable.

2.4. River flow data

In this study, historical time series at the five basins from different regions of the world were used. Table 1 presents detailed spatial and temporal information about these basins. The statistical parameters corresponding to each time series are given in Table 2, while Figure 1(a)-(g) shows the observed river flow time series variations in the selected stations versus time.

Table 1. Spatial and temporal detail of the case studies

Basin	Station ID	Country	Date (y-m)		Location		Elevation (m)	Drainage Area (km^2)
			Start	End	Latitude	Longitude		
AjiChai	31005	Iran	1988-10	2010-10	47°28'48"	37°52'48"	1,660	1,679
	31014		1981-10	2010-10	47°13'59"	38°00'00"	1,600	2,451
	31117		1991-10	2010-10	46°48'58"	38°09'58"	1,518	5,635
	31015		2001-10	2010-10	46°23'60"	38°06'36"	1,450	7,280
	31045		1976-10	1996-10	46°02'60"	38°00'36"	1,310	9,670
BalikhliChai	19049	Iran	1981-10	2010-10	38°07'00"	47°54'03"	2,068	18
	19051		1981-10	2010-10	38°02'02"	47°59'38"	1,624	168
	19910		1999-10	2010-10	38°02'05"	48°02'11"	1,584	567
	19053		1981-10	2010-10	38°08'16"	48°11'19"	1,440	1,044
Kizilirmak	1535	Turkey	1996-10	2002-12	39°56'18"	32°49'44"	1,243	6,607
	1501		1996-10	2002-12	39°55'0"	32°33'56"	995	15,582
Bear	10011500	U.S.	1962-01	2012-12	40°57'55"	110°51'10"	2,428	445
	10020100				41°26'04"	111°01'01"	1,967	1,955
	10038000				40°43'53"	110°51'10"	1,871	6,338
Eel	11472800	U.S.	1990-01	2005-10	39°49'45"	123°04'11"	457	528
	11473900		1990-01	2005-12	39°42'23"	123°19'27"	275	1930
	11475000		1990-01	2005-12	40°13'06"	123°37'53"	66	5457
	11478000		1990-01	2005-12	40°29'30"	124°05'55"	11	8063

Table 2. Statistics of the river flow time series for the case studies considered

Basin	Station	Statistics						
		Number of Data	Mean (m^3/s)	Min. (m^3/s)	Max. (m^3/s)	Std. Deviation (m^3/s)	Skew	Kurtosis
AjiChai	31005	240	4.5	0	32.0	6.1	2.4	6.1
	31014	120	2.6	0	32.0	5.4	3.6	14.3
	31117	240	6.0	0	50.29	10.3	2.43	5.8
	31015	360	9.5	0	107.70	16.4	2.97	10.49
	31045	288	8.6	0	107.31	15.62	3.61	15.78
BalikhliChai	19049	360	0.2	0.01	2.01	0.23	5.37	33.04
	19051	360	1.2	0.12	8.66	0.73	2.79	23.28
	19910	144	2.0	0.36	8.47	1.51	1.98	4.86
	19053	360	3.45	0.03	18.33	2.87	1.60	4.09
Kizilirmak (daily data)	1535	72	33.21	3.03	232.26	46.99	2.61	7.14
	1501	72	57.79	3.96	349.32	73.34	2.60	6.87
Bear	10011500	612	5.5	0.6	56.3	8.25	2.59	7.24
	10020100	612	6.2	0.01	72.5	10.04	2.91	10.48
	10038000	612	12.1	0.0	105.1	14.66	2.84	9.51
Eel	11472800	192	157.36	0.88	1317.55	246.81	2.33	5.54
	11473900	192	95.81	0.15	877.42	155.23	2.47	6.53
	11475000	192	30.79	0.02	219.86	44.7	2.14	4.86
	11472800	84	17.13	0.09	103.88	22.99	1.86	3.75

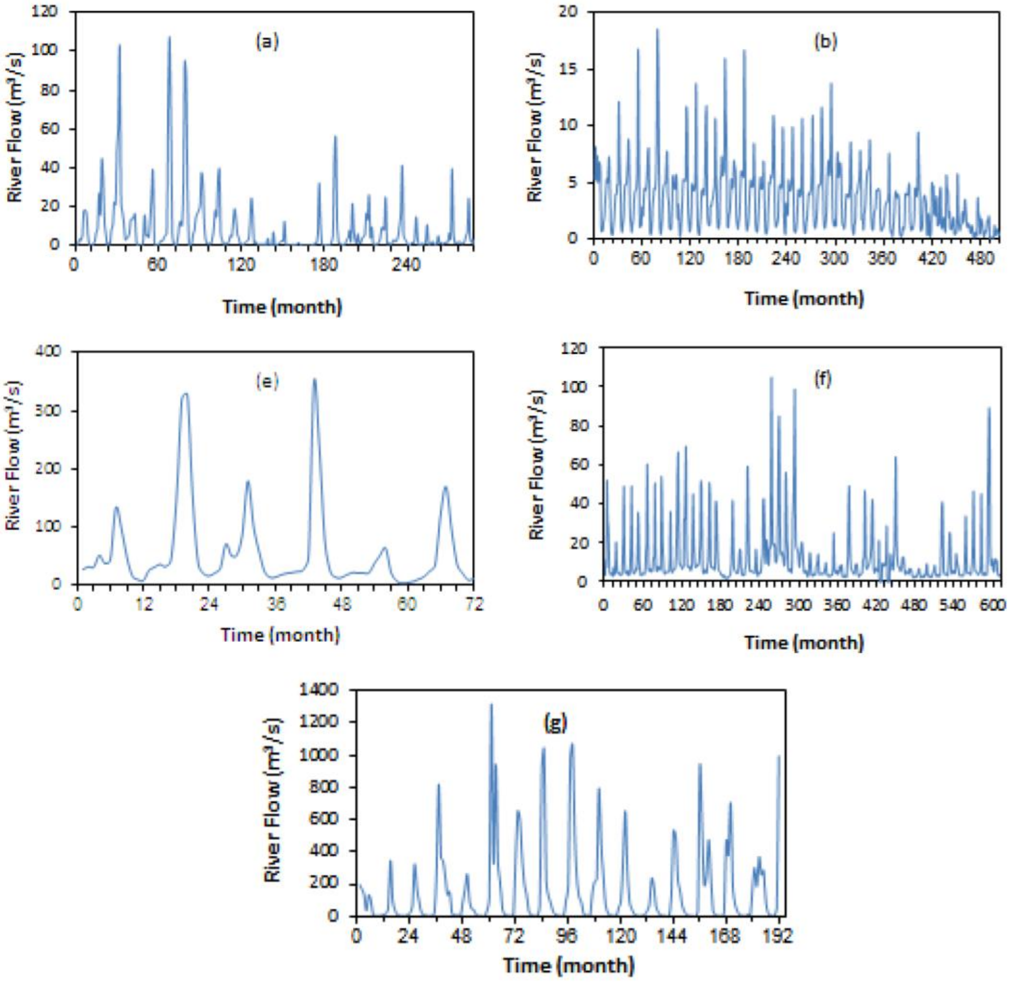

Fig. 1. River flow time series at the stations located in basin outlets: a) 31045 in the AjiChai basin; b) 19053 in the BalikhliChai basin; c) 1501 in the Kizilirmak basin; d) 10038000 in the bear basin; e) 11477000 in the Eel basin

3. Results and discussion

To determine the chaotic dynamics for each of the stations, all of the river flow time series are used for phase-space reconstruction. Fig. 2(a)–(g) shows the phase-space reconstruction of time series at stations 31045, 19053, 21425, 17035, 1535, 38000 and 7000 respectively. The plots show the projection of the attractor on the plane with delay time equal 1. The delay time (τ) is estimated by using the average mutual information (AMI) method. The first minimum in the mutual information function can be considered as the optimal delay time (Regonda et al., 2004). The mutual information function of the stations are plotted and shown in Figure 3. Hence, the optimal delay time is chosen as 3, 4, 4, 8, 7, 6 and 4 for 31045, 19053, 21425, 17035, 1535, 10038000 and 11477000 stations, respectively (table 3). The initial exponential decay of mutual information functions indicates that the river flow series may be of chaotic nature (Li et al., 2013). After determining the delay time values, the correlation integrals were computed by the Grassberger-Procaccia algorithm for different embedding dimensions (m). Figure 4 shows the relationship between the correlation function, C(r), and the radius r (i.e. log$C(r)$ versus log r) for embedding dimensions, different value m. The relationship between the correlation exponent values (D_2) and the embedding dimension values m is shown in Figure 5.

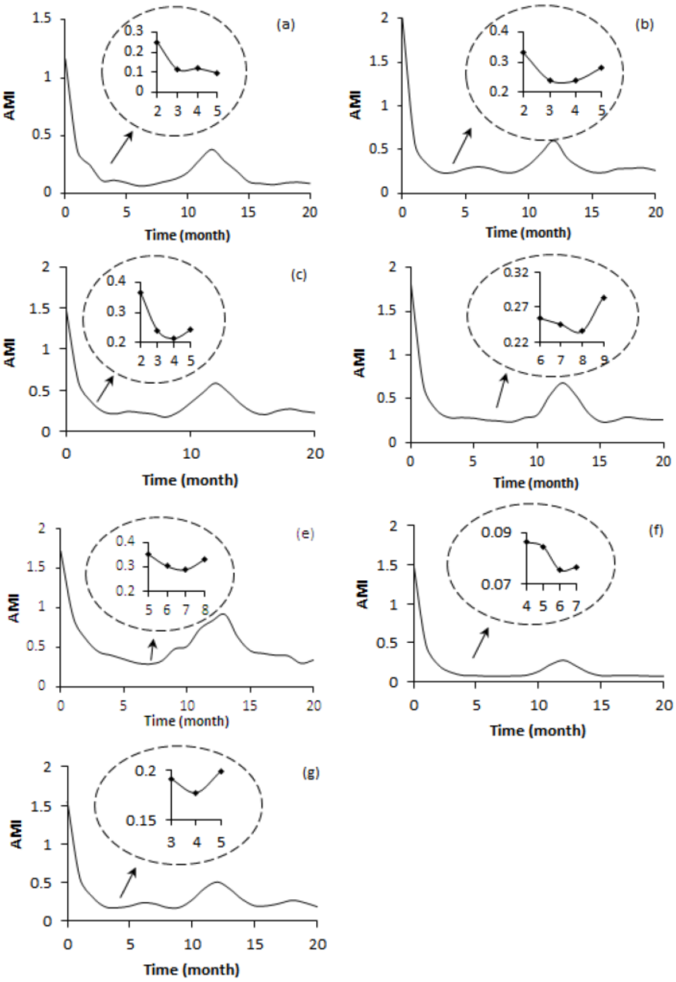

Fig. 2. Average mutual information function and its relative change with lag time at the stations that located in outlet basin: a) 31045 in the Ajichai basin; b) 19053 in the Balikhlichai basin; e) 1501 in the Kizilirmak basin; f) 10038000 in the bear basin; g) 11477000 in the Eel basin

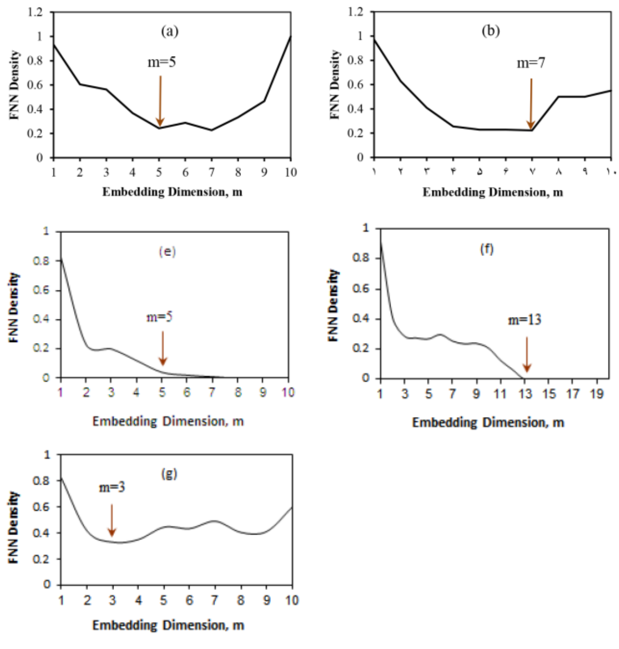

Fig. 3. Percentage of false nearest neighbor of the monthly river flow time series in embedding dimension at the stations: a) 31005 in the Ajichai basin; b) 19053 in the Balikhlichai basin; e) 1501 in the Kizilirmak basin; f) 10038000 in the bear basin; g) 11477000 in the Eel basin

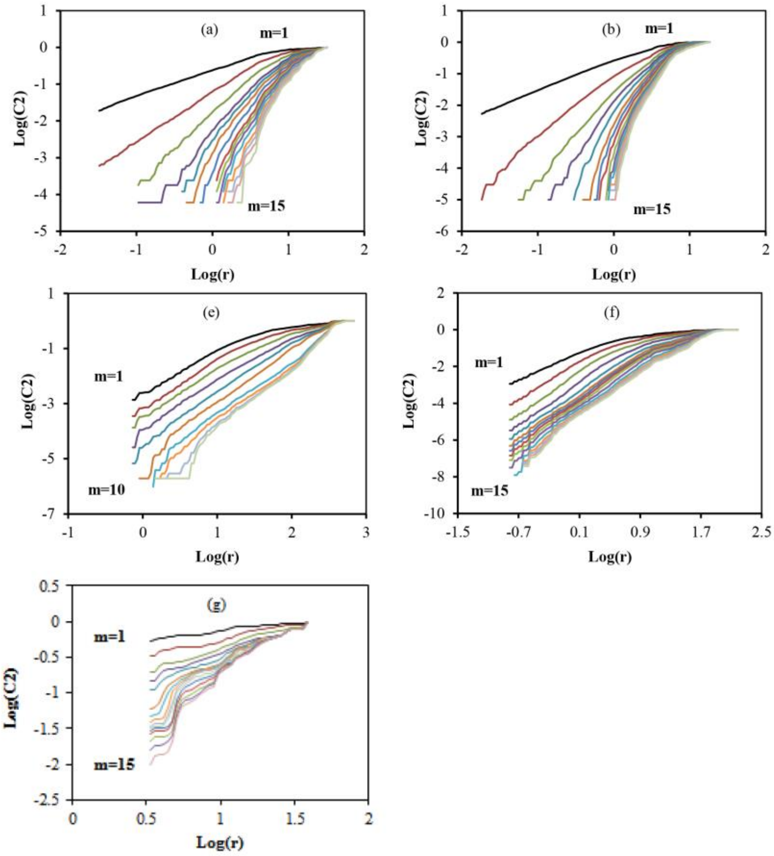

Fig. 4. Log C(r) versus log (r) plots for monthly river flow data: a) 31045 in the Ajichai basin; b) 19053 in the Balikhlichai basin; c) 21425 in the Beheshabad basin; d) 17035 in the Taleghan basin; e) 1501 in the Kizilirmak basin; f) 10038000 in the Bear basin; g) 11477000 in the Eel basin

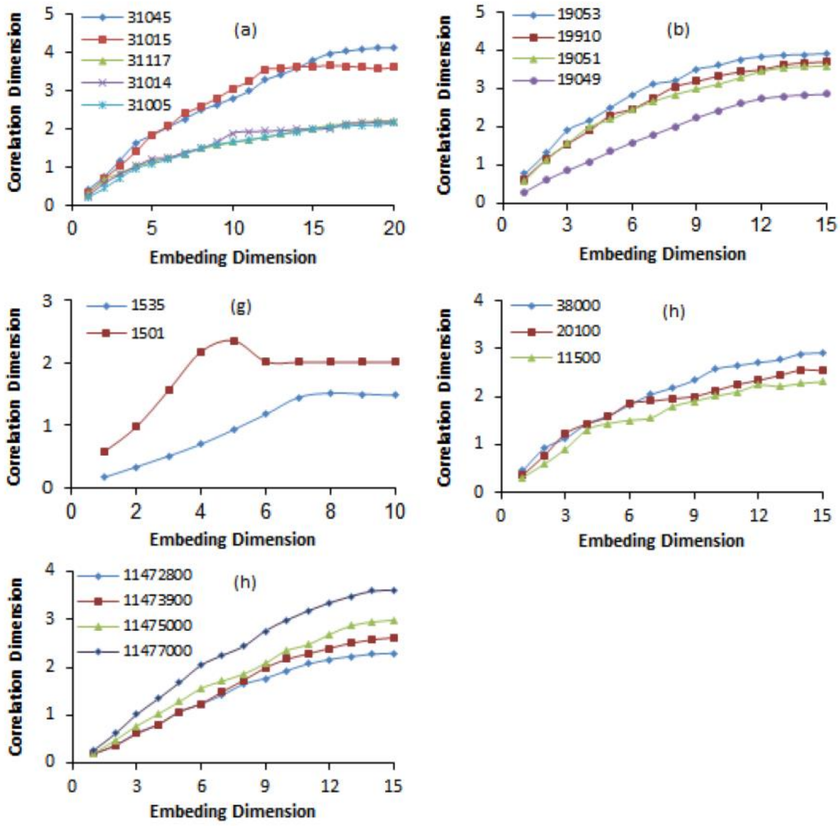

Fig. 5. Relationship between embedding dimension and correlation exponent for different station in: a) Ajichai basin; b) Balikhlichai basin; g) Kizilirmak basin; h) Bear basin; i) Eel basin

4. Conclusion

4.1. Conclusions and possibility for future study

Correlation dimension (D2) in all of monthly runoff time series is a positive value and limited there for all of them are chaotic. For delay times and embedding dimension not seen relationship between them values and position of the station in basins. It seems that they do not significantly increase or decrease but in all of basin correlation dimension value increase from upstream to downstream. The rate of increase depends on the conditions of each basin such as changes in slope, land use, soil and climatic conditions. This study is among the first investigation it is clear that we need more studies to reach a general conclusion on this point. In homogeneous basin changes of D2 is less. Obtain a relation between correlation dimension and position of gauge station, Area or other physical characteristics of basin is important because D2 expressed as the number of relevant parameters and can determined type time series in terms of complexity.

References

Abarbanel, H.D.I., 1996. *Analysis of observed chaotic data*. Springer – Verlag, New York.

De Domenico, M., Ghorbani, M.A., Makarynskyy, O., Makarynska, D. & Asadi, H., 2013. Chaos and reproduction in sea level. A*pplied Mathematical Modelling*, 37(6), 3687-3697.

Ghorbani, M.A., Kisi, O. & Alinezhad, M., 2010. A probe into the chaotic nature of daily streamflow time series by correlation dimension and largest Lyapunov methods. *Applied Mathematical Modelling*, 34, 4050-4057

Grassberger, P. & Procaccia, I., 1983. Characterization of strange attractors. *Physical review letters*, 50(5), 346 p.

Jayawardena, A.W. & Lai, F., 1994. Analysis and prediction of chaos in rainfall and stream flow time series. *Journal of hydrology*, 153(1-4), 23-52.

Kędra, M., 2014. Deterministic chaotic dynamics of Raba River flow (Polish Carpathian Mountains). *Journal of Hydrology*. 509(13), 474-503.

Khatibi, R., Sivakumar, B., Ghorbani, M.A., Kisi, O., Koçak, K. & FarsadiZadeh, D., 2012. Investigating chaos in river stage and discharge time series. *Journal of Hydrology*, 414-415, 108-117.

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, 185-193.

Krasovskaia, I., Gottschalk, L. & Kundzewicz, Z.W., 1999. Dimensionality of Scandinavian river flow regimes. *Hydrological Sciences Journal*, 44(5), 705-723.

Li, X., Gao, G., Hu, T., Ma, H. & Li, T., 2014. Multiple time scales analysis of runoff series based on the chaos theory. *Desalination and Water Treatment*, 52(13-15), 2741-2749.

Ng, W.W., Panua, U.S. & Lennox, W.C., 2007. Chaos based Analytical techniques for daily extreme hydrological observations. *Journal of Hydrology*, 342(1-2), 17-41.

Ondimu, S. & Murase, H., 2007. Reservoir level forecasting using neural networks. *Lake Naivasha. Biosystems Engineering*, 96(1), 135-138.

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

Regonda, S., Sivakumar, B. & Jain, A., 2004. Temporal scaling in river flow: Can it be chaotic? *Journal of Hydrological Science*, 49(3), 373-385.

Sivakumar, B., 2001. Rainfall dynamics at different temporal scales: A chaotic perspective. *Hydrol Earth Sys Sci*, 5(4), 645-651.

Sivakumar, B., 2001. Rainfall dynamics at different temporal scales: A chaotic perspective. *Hydrology and Earth System Sciences*, 5(4), 645-652.

Sivakumar, B. & Singh, V.P., 2012. Hydrologic system complexity and nonlinear dynamic concepts for a catchment classification framework. *Hydrology and Earth System Sciences*, 16(11), 4119-4131.

Sivakumar, B., Berndtsson, R. & Persson, M., 2001. Monthly runoff prediction using phase space reconstruction. *Hydrological sciences journal*, 46(3), 377-387.

Takens, F., 1981. *Detecting strange attractors in turbulence, in Lectures Notes in Mathematics*,edited by D.A.Rand and L.S.Young, Springer-Verlag, New York. 898, 366-381.

Tang, Q. & Hu, T., 2010. Testing for Chaos of Yangtze River Streamflow at Different Timescales Seventh *International Conference on Fuzzy Systems and Knowledge Discovery*, 2151-2154.

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