Chaos in hydrology: A case study in Konya Basin, Turkey

Conference Paper · September 2018

DOI: 10.1142/9789813273238_0168

3 authors:

Didem Odabaşı Cingi
Istanbul Aydin University

Ergun Eray Akkaya
Beykoz University

Dilek Eren Akyüz
Istanbul University-Cerrahpasa

CHAOS IN HYDROLOGY: A CASE STUDY IN KONYA BASIN, TURKEY

DIDEM ODABASI CINGI

Department of Construction Technologies, Istanbul Aydin University, Istanbul, Besyol/Kucukcekmece 34295, Turkey

ERGUN ERAY AKKAYA

Mechatronics Engineering Department, Istanbul Gelisim University, Istanbul, Avcilar 34315, Turkey

DILEK EREN AKYUZ†

Civil Engineering Department, Istanbul University, Istanbul, Avcilar 34320, Turkey

Every natural behaviour is non-linear, but not always is chaotic. This paper aims to investigate low dimensional chaotic behaviour of study area: Konya Basin by using non-linear time series techniques with three stages: i) Mutual Information, ii) False Nearest Neighbour (FNN) algorithm, iii) Stretching Exponential. These techniques calculate the delay time, the embedding dimensions and the maximal positive Lyapunov exponent respectively. The data set consists of daily average flow rates of three stations in Konya Basin through the study period between 1968 and 2014. Analysed data implied that these time series have shown chaos. This information helps catchment manager to forecast future and the extreme flow rates such as droughts and floods.

1. Introduction

Konya Basin is one of the 25 river basins in Turkey. It has approximately 7% of Turkey with an area of 5.5 million hectares. 2% of the available surface water resources in Turkey is Konya Basin. On the other hand, it has about 17% of the groundwater potential of our country due to its large closed basin [1]. In this study, we analysed the observed data from three stations located in the east part of Konya Basin (Figure 1). The 1968-2015 Konya Basin daily average flow rates time series reveals significant inter-annual and inter-decadal fluctuations. Such wet or dry periods are associated with regional climatic variability and are vital for understanding or predicting drought and the long-term accessibility of water. In this case, Konya Basin is known as semi-arid. Therefore, water management concepts such as water treatment and pollution control are important for water quality and sustainability of the basin [2].

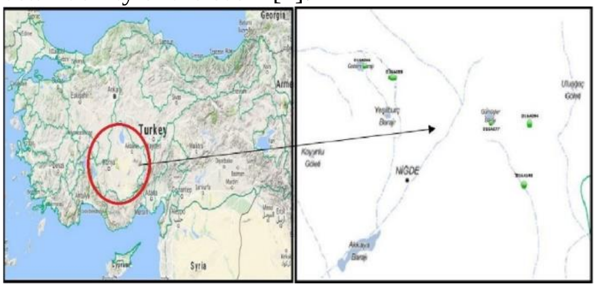

Figure 1. Location of Konya Basin and stations.

A multitude of factors indicate the climatic state and hence the Konya Basin flow rate. The integrating effect of the basin may lead to its fluctuations being determined largely by a few, unknown dynamical variables that may be complex, nonlinear functions of the physical variables. In this paper we test the hypothesis that the dynamics of Konya Basin can be described by a small set of variables, that is, it is low-dimensional. Such evidence may be useful for developing low-order models that explain at least a part of Konya Basin variability and perhaps can be ultimately related to low-frequency climatic variability.

Some current studies have demonstrated that low-dimensional deterministic strategies can be connected as an elective technique for modelling and the results are promising [3]. A sophisticated behaviour in nature can be distinguished as deterministic and disorderly or non-deterministic and randomized, subject to its base dynamics. Therefore, it is quite difficult to distinguish a chaotic deterministic system from a purely random system. Since both of them can produce irregular and apparently unpredictable temporal and/or spatial variability. The theory of chaos in hydrological sciences was introduced in the late 1980s [4, 5]. Since then, applications of chaos theory in hydrological sciences have been significantly advanced [6, 7, 8]. As things stand, chaotic and low dimensional deterministic time series techniques verifies the obtained results in the field of research, modelling and predicting possible river flow dynamics.

All the time an adequate measure of care is practiced in applying low-dimensional deterministic methods and in evaluating the results, such methods can be valuable in studying dynamics of river flow [6]. Furthermore, late articles which make prediction on river flow by using chaos theory, can reveal the number of factors that impact the river flow dynamics [3, 9]. This situation shows that; non-linear analysis is picking up significance and convenience for river flow analysis.

2. Theory and Data

In this study, three gauging stations were selected for investigation of chaotic behaviour within the closed basin. The data set contains the daily average flow rates of the stations between years 1968 and 2014. The methods shown schematically in Figure 2 were used to analyse time series data.

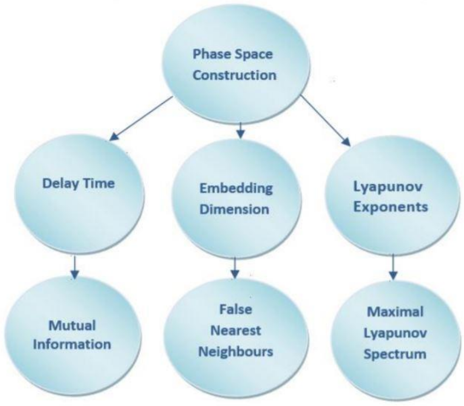

Figure 2. Data analysis steps in time series.

2.1. *Phase Space Reconstruction*

The reconstruction of a d-dimensional phase space is a vital step to analyse the underlying dynamics of a scalar time series. From the phase-space configuration of a system, we can obtain some information about the asymptotic properties (*e.g.* positive Lyapunov exponents) of the studied system, indicating how chaotic a system is, and the topological dimension of an attractor. Since attractors contain geometrical and dynamical properties of the original phase space, study properties of attractors signify study properties of the system. So, phase space reconstruction enables studying unobserved variables. In reconstructed phase space, we can determine the relation between current and future situation. For time series of scalar flow rate $s_{nn}(1), s_{nn}(2), s_{nn}(3), \ldots, s_{nn}(n)$ the delay of the reconstructed phase space $\vec{y}_{nn}(k)$ is defined by,

$$\vec{y}_{nn}(k) = [\, s_{nn}(k), s_{nn}(k+\tau), s_{nn}(k+2\tau), \ldots, s_{nn}(k+(d-1)\tau)] \quad (1)$$

Where, $k = 1, 2, \ldots, n$; τ and d are the time delay and the embedding dimension respectively. Thus, the two parameters, time delay τ and embedding dimension d, need to be selected before reconstructing the phase space. Time delay can be measured from the first zero of correlation function (linear criterion) or first minimum of average mutual information [10].

2.2. *Mutual Information*

Some of the past studies recommend the use of the autocorrelation function to choose the time delay τ [5, 11]. But in fact, autocorrelation function is a linear statistic and does not agree the nonlinear correlations. However, in 1986, Fraser and Swinney suggested using the mutual information method for the determination of the optimum choice of the time delay for the phase space reconstruction [10]. At this point, it is crucial to state that, for some attractors, the method choice does not matter. However, for others, the prediction of τ might depend exceedingly on the approach applied. If nonlinear correlation is considered, the mutual information method is more accurate regarding the proper choice of τ [12].

$$S = -\sum_{ij} P_{ij}(\Delta t) \ln \frac{P_{ij}(\Delta t)}{P_i P_j} \tag{2}$$

where; P_i and P_j are the probabilities to determine a time series value into the i^{th} and j^{th} intervals of a section of the available space, and $P_{ij}(\Delta t)$ is the probability that an observation value takes part in the i^{th} interval and, after some observational delay time Δt, later takes part in the j^{th} interval [13]. The first minimum of the mutual information is widely used for its quality to capture the nonlinear correlation in time series data [10]. And, it can be calculated by hand.

2.3. *False Nearest Neighbours*

The False Nearest Neighbours (FNN) method is the most popular tool for choosing the minimum embedding dimension. In the phase space, the attractor trajectory is directly influenced by embedding dimension. Thus, it is important for the neighbourhood of the points. FNN is based on the hypothesis that nearly located two points in a sufficient embedding dimension should remain close when dimension increases. If the embedding dimension is too small, selected points can be seen as neighbours but in fact they are distant to each other. If the embedding dimension is too large, this situation causes undesirable consequences such as redundancy on chaotic data and decreasing performance of many algorithms. When these cases are taken into consideration, the false neighbours should be corrected. In this case, FNN percentage should drop to zero when the optimum embedding dimension value is reached. This process is calculated as follows:

$$R_i = \frac{|R_{i+1} - R_{j+1}|}{\|\vec{R_i} - \vec{R_j}\|} \tag{3}$$

where; $\vec{R_i}$ is a constructed vector by using delay time, $\vec{R_j}$ is the nearest neighbour of $\vec{R_i}$ vector in a random dimension. This process should be iterated for all consecutive vectors to obtain R_i value. If the point of data is selected as a false neighbour in distance, R_i crosses above the certain threshold [14].

2.4. *Maximal Lyapunov Exponent*

Lyapunov exponent measures the exponential divergence and convergence of initially close state-space trajectories and quantities the amount of chaos in the system. So, Lyapunov exponent shows the chaotic nature of attractor. If exponential growth rate of nearby trajectories is positive, this state signs a chaos. This exponential growth rate is called as maximal Lyapunov exponent. In that case, maximal Lyapunov exponent physically can be considered as the rate of divergence of close trajectories. The stretching of the trajectories is calculated as follows:

$$S(\epsilon, m, t) =< \ln \frac{1}{u_n} \sigma s_n \epsilon u_n | s_{n+t} - s_{n'+t} | > \qquad (4)$$

Where; $s_{n'}$ is a neighbouring point, ϵ is the box size, after t time, the distance between the selacted points will be $s_{n+t} - s_{n'+t}$. This equation calculates the growth of this distance in time. If S (ϵ,m,t) is linear, the maximal Lyapunov exponent can be calculated from the slope.

2.5. *Evaluation of Data*

This study includes the dynamics of the daily average flow rates of 3 gauging stations located in Konya Basin. Data were obtained from Directorate General for State Hydraulic Works (DSI). We tried to select convenient combinations of time delay and embedding dimension of reconstructed phase space. In this case, we calculated statistical properties of the time series such as minimum, average, maximum, skewness and standard deviation values (Table 1).

Table 1. Statistical properties of time series.

Station Number	Length (Day)	Observation Years	Average (m^3/s)	Max (m^3/s)	Min (m^3/s)	Skewness	Std. Dev.
D16A066	13128	1969-2014	0.061	0.640	0.000	1.182	0.100
D16A077	13877	1968-2014	0.107	0.620	0.000	1.120	0.161
D16A089	13150	1971-2014	0.116	3.700	0.000	3.676	0.179

By the reason of identifying and interpreting the relationships in data, we generated time series plots for each station (Figure 3). And Table 2 demonstrates the calculated values such as mutual information, embedding dimension and Lyapunov k (maximal Lyapunov exponent of a given scalar data set) [15].

Table 2. Nonlinear time series analysis results of Konya Basin.

Station Number	Time Delay (Days)	Embedding Dimension (m)	Lyap_k (values)
D16A066	20	3	0.0106
D16A077	17	3	0.0360
D16A089	18	3	0.0286

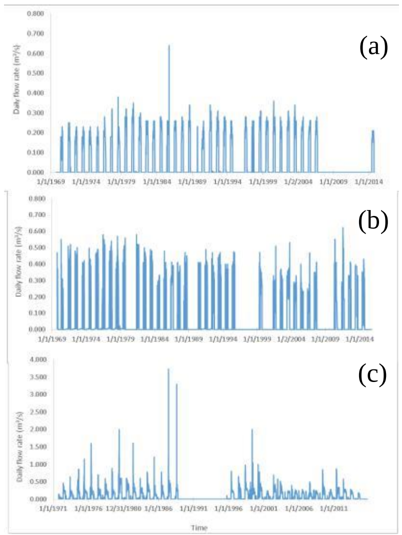

Figure 3. Flow rates time series for a) D16A066 station, b) D16A077 station, c) D16A089 station.

3. Application

First, we analysed the raw data as depicted in Figure 3. And we saw the variation of autocorrelation coefficient of the original data series of daily average flow rates. It also indicates seasonal variation with some quick decay in some lags. Then we marked zero-crossing point. This value gives information about downward tendency of the linear correlation. Thereafter the result obtained from the first minimum of the mutual information output which is shown in Figure 4(a) was compared with the zero-crossing point. It is sign of the nonlinear information of descent value of the data at lag 18. This state shows that the nonlinear correlation is more acceptable than the linear correlation in the daily average flow rates data series. Then, we decided to apply the chaotic time series prediction based on non-linear time series techniques to the daily average flow rates of three stations in Konya Basin.

This study consists of three steps. In the first step, mutual information is calculated by using TISEAN package [13] to determine time delay. In the second step, we computed false nearest neighbours to find the embedding dimension. In the last step, we used these values to obtain maximal positive Lyapunov exponent for the stations (Table 2).

As seen in Figure 4, slope of the maximal positive Lyapunov exponent of D16A066 station is increasing. This situation indicates the chaotic behaviour. Figure 4 also shows tendency is chaotic. In Table 2; time delay, embedding dimension and maximal Lyapunov exponents of three stations are presented as computed values. In this case, other stations sign nearly same results. They also have chaotic behaviour.

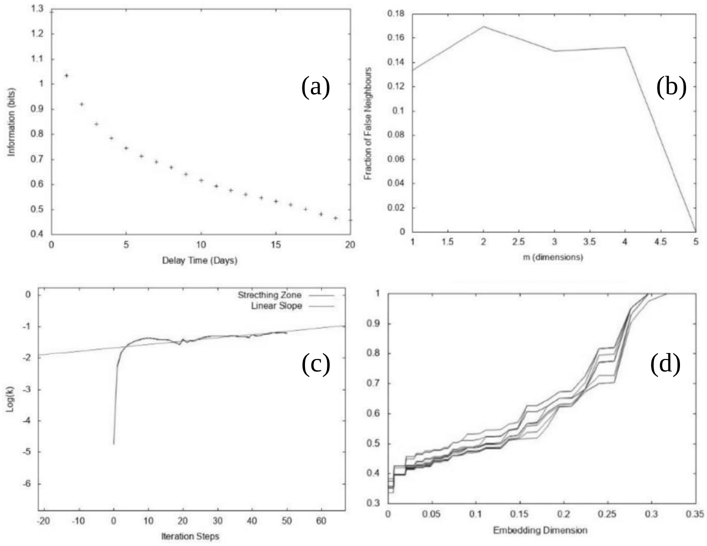

Figure 4. Chaotic time series analysis for D16A066 station a) Mutual information, b) False nearest neighbour, c) Lyapunov exponent by Kantz algorithm, d) Dimensional correlation.

4. Conclusion

Low-dimensional chaos is defined as a dynamical system indicating one positive Lyapunov Exponential when the embedding dimension of the system is greater than or equal to 3. Maximal Lyapunov Exponent is defined as the rate of divergence of close trajectories and it is counterbalanced by the other effects so the system is considered as stable. In this case, the attractor is limited and for small perturbation of the trajectories we remain in the attractor.

There is another concept called as hyper-chaos. Since it would be impossible to counterbalance more than one positive Lyapunov exponent, hyper-chaos cannot be seen in a low-dimensional dynamical system. If embedding dimension of a dynamical system is greater than 3 and there is one positive Lyapunov exponent in this dynamical system, this is called as Higher Dimensional Chaos. When a Higher Dimensional Dynamical System (with embedding dimension is greater than 3) contents more than one positive Lyapunov Exponent, it is called as hyper-chaos.

In the light of this information, we analysed chaotic behaviour for three stations. For each station, there is only one positive Lyapunov exponent was calculated. The embedding dimension of each station was obtained as 3. But embedding dimension of the station D16A077 was calculated at the limit value (3). This means that our study area could be Higher Dimensional Dynamical System. In this case, low-dimensional non-linear techniques would be useless for this station. We would use other methods that appropriate to HD chaos such as HD Lorenz, Duffing, Rössler and Van der Pol oscillators, modified canonical Chua's circuits.

This study is the first step of chaotic time series estimation. By defining the chaotic behaviour, making forward and backward estimates will be easier and the missing data of each station can be defined. In this way, extreme values can be calculated such as flood and drought estimation. These results will help us to understand and evaluate the real world much better in mathematical terms.

References

1. wwf.org, (2018).
2. D.E. Mercan, S. Kabdasli, *PhD thesis* (2006).
3. R. Khatibi, B. Sivakumar, M.A. Ghorbani, O. Kisi, K. Koçak, D.F. Zadeh, *Journal of Hydrology* **414-415**, 108–117 (2012).
4. A. Hense *Beitr Phys Atmos* **60**(1):34–47 (1987).
5. I. Rodriguez-Iturbe, F.B. De Power, M.B. Sharifi, K.P. Georgakakos. *Water Resour Res* **25**(7):1667–1675 (1989).
6. B. Sivakumar, (2003).
7. H. Tongal, M.C. Demirel, M.J. Booij, *Stoch Environ Res Risk Assess* **27**:489–503 (2013).
8. M.S. Kyoung, H.S. Kim, B. Sivakumar, V.P. Singh, K.S. Ahn, *Stoch Environ Res Risk Assess* **25**(4):613–625 (2011).
9. N.H. Adenan and M.S. Noorani, *Journal of Korean Society of Civil Engineering*, **18**, 2268-2274 (2014).
10. A.M. Fraser and H.L. Swinney, *Physical Review*, **A33**, 1134 (1986).
11. Q. Wang and T.Y. Gan, *Water Resour. Res.* **34**(9), 2329–2339 (1998).
13. A.A Tsonis and J.B. Elsner, *Nature* **333**, 545–547 (1988).
14. R. Hegger, H. Kantz and T. Schreiber, *Chaos*, **9**, 413 (1999).
15. E.E. Akkaya, A.S. Hacinliyan, *PhD thesis*, (2016).
16. H. Kantz, *Physics Letters*, **A185**, 77 (1994).