Journal of Civil Engineering and Urbanism

Volume 2, Issue 6: 220-225 (2012) ISSN-2252-0430

ORIGINAL ARTICLE

Evidence of Low-Dimensional Determinism in Monthly Water Level Time Series in Lake Van, Turkey

Rasoul Daneshfaraz[1], Hakimeh Asadi[2*] and Seyed Mahdi Saghebian[3]

[1]*Department of Civil Engineering, University of Maragheh, Maragheh, Iran, E-mail: daneshfaraz@yahoo.com*

[2]*Department of Water Engineering, University of Tabriz, Tabriz, Iran, E-mail: tu.i2006@yahoo.com*

[3]*Department of Civil Engineering, Ahar Branch, Islamic Azad University, Ahar, Iran, E-mail smsaghebian@iau-ahar.ac.ir*

*Corresponding author's Email: tu.i2006@yahoo.com

ABSTRACT: Adequate knowledge of lake water level variations is important for proper planning and management of water resources and design of environment. In this study chaotic behavior of monthly water level variations in the lake Van during January 1944 - April 2002 is investigated. The lake Van is the largest lake in Turkey. The methods and indicators of chaos theory (power spectrum, average mutual information, false nearest neighbours, correlation dimension and largest Lyapunov exponents) were applied. The value of power spectrum (4.1) indicate that chaotic (fractal) behavior to the lake water level time series. The optimal delay time (5 month) and embedding dimension (5) are obtained from average mutual information and false nearest neighbours techniques, respectively. Optimal values are then used for the estimation of the correlation dimension and the largest Lyapunov exponent for inspecting possible signatures of chaotic dynamics. The low correlation dimension (3.2) suggest the presence of low-dimensional chaos; also imply that the water level dynamics are dominantly governed by four variables. The positive largest Lyapunov exponent value (0.0085) indicated a signature of chaos. These results give a positive indication towards considering lake water level as a chaotic system.

Keywords: Chaos theory, Correlation dimension, Lyapunov exponent, Power spectrum, Time series, Van Lake.

INTRODUCTION

Lake water level studying plays a significant role in management of fresh water supply, designing and planning of lakeshore structures and environment. It is necessary to develop models for simulation of the level variations in order to control future lake level changes (Sen et all, 2000).

Recently different methods widely have been used for lake water level forecasting, including Artificial Neural Networks (Altunkaynak 2007; Ondimu and Murase 2007), Triple diagram model (Altunkaynak et al. 2003), Markov model (Sen et al. 2000), Cluster regression (Sen et al. 1999), Fuzzy Logic (Altunkaynak and Sen 2007), Support vector machines (Çimen and Kisi 2009), Adaptive Neuro-fuzzy Inference system (Mpallas et al. 2011). During the past two decades, chaos theory showed its applicability in solving a wide class of problems in many areas of natural sciences and, in particular, of civil engineering and water-related applications: these studies are devoted to model and forecast natural phenomena and require, in each case, a deeper comprehension of the underlying dynamics. A preliminary study on the application of nonlinear time series analysis using delay coordinate embedding on the tidal data from the Venice Lagoon from 1980 to 1984, was carried out by Vittori (1992). Koçak (1997) successfully applied nonlinear prediction to water level time series. Zaldivar et al. (1998) employed nonlinear time series analysis for the detection of high water levels in Venice (Italy), whereas nonlinear dynamic analysis of coastal waters and the comparison with other existing methods have been reported by Frison et al. (1999). Solomatine et al. (2000) suggested the possibility of accurate predictions of the surge water level in the North sea with similar techniques. Rahmstorf (2003) used a semi-empirical approach to study sea level fluctuations based on earth temperature changes. Khatibi et al. (2011) successfully employed chaos theory to hourly water level at Hillarys Boat Harbour, Western Australia.

Lake Van is the largest lake in Turkey and for no studies have been found about the application of chaos theory in modeling of the lake water level, so the present study can be considered as a pioneering study presenting the usage of this method. This paper aims at investigating the possible presence of chaotic signals in the Van lake water level time series during January 1944- April 2002. The remainder of the paper is organized as follows. Section 2 presents the methodology in this study. In Section 3, the data used, case study and results obtained are explained. The conclusions of this study are presented in Section 4.

To cite this paper: Daneshfaraz R, Asadi H, Saghebian SM. 2013. Evidence of Low-Dimensional Determinism in Monthly Water Level Time Series in Lake Van, Turkey. *J.Civil Eng. Urban.*, 2(4): 220-225.

Journal homepage http://www.ojceu.ir/main/

METHODOLOGY

Several methods have used for the investigation of chaotic signal in time series. In this study the concepts and methods of chaos theory such as power spectrum, average mutual information, false nearest neighbours, correlation dimension and largest Lyapunov exponent are employed to analyse the chaotic signal of the water level time series.

Power Spectrum Analysis

Since chaotic systems are aperiodic, a power spectrum analysis can indicate the presence of periodic regimes (Ng et al. 2007). If the power spectrum, $E(f)$, obeys a power law form

$$E(f) \propto f^{-\beta} \tag{1}$$

where f is the frequency and β is the spectral exponent, this is an indication of the absence of characteristic time scale in the range of the power law. In such a case, fractal behaviour may be assumed to hold (Sivakumar, 2006).

Phase space reconstruction

The concept of phase space is a powerful tool for characterizing dynamical systems. The delay embedding is one of the most popular methods for reconstructing phase space from a univariate or multivariate time series (Takens, 1981), assumed to be generated by a deterministic dynamical system with D degrees of freedom. The Takens theorem stated that the underlying (unknown) dynamics can be fully recovered by building a m-dimensional space wherein the components of each state vector $\vec{Y}_t$ are defined through the delay coordinates

$$\vec{Y}_t = \left(X_t, X_{t-\tau}, X_{t-2\tau}, ..., X_{t-(m-1)\tau}\right) \tag{2}$$

where $m > 2D_2$ is called *embedding dimension* and τ is referred to as *delay time*. If the dynamics of the system can be reduced to a set of deterministic laws, trajectories converge towards a subset of the phase space with fractional dimension, called attractor.

Many methods have been proposed for the estimation of optimal values of the embedding parameters: within this work we adopted the minimization of the False Nearest Neighbours (FNN) for m and of the Average Mutual Information (AMI) for τ, as suggested by Cellucci et al. (2003).

For a given time series sequence $\{x_0, x_1, x_2, ..., x_i, ..., x_n\}$ the mutual information indicates the amount of information about the state $x_{i+\tau}$ if the state of x_i is known. The average mutual information is defined by:

$$I(\tau) = -\sum_{ij}^{N} p_{ij}(\tau) \ln \frac{p_{ij}(\tau)}{p_i(\tau)p_j(\tau)} \tag{3}$$

where $p_i(\tau)$ and $p_j(\tau)$ correspond to the probability of finding x_i in the *ith* and $x_{i+\tau}$ in the *jth* interval, and $p_{ij}(\tau)$ is their joint probability. The first local minimum of $I(\tau)$ estimates the optimal selection for the delay time required for phase space reconstruction. In practice, it provides the maximum delay time such that $x_{i+\tau}$ adds the largest amount of information about x_i.

Successively, false nearest neighbours (FNN) search is used to determine the optimal embedding dimension m. In fact, a small value of m may not be sufficient to reconstruct the phase space, whereas a large value of m causes a large unfolding of the attractor and high computational cost (Cellucci et al. 2003). The method is as follows. For a fixed embedding dimenion m, and for each point in the phase space, we identify the K nearest neighbours. Then, we repeat our procedure in the $m+1$-dimensional phase space: if the reconstruction is not optimal, the nearest neighbours are different, i.e. they were false nearest neighbours.

FNN method employs the search of false neighbours in phase space: when the ratio between the number of false neighbours at the dimension $m+1$ and m is below a given threshold, generally smaller than 5%, each $m' > m+1$ is an optimal embedding. However, if m' is too large, a poor reconstruction of few embedding states with several components is obtained and the next analyses should not be performed (Kennel et al. 1992).

Correlation dimension

Correlation dimension is the most widely used as fractal dimension quantifier, and is based on the correlation integral (Grassberger & Procaccia, 1983b).

For an m-dimensional phase space, the correlation function $C_m(r)$ is defined as the fraction of states closer than r (Fraser and Swinney 1986):

$$C_m(r) = \lim_{N\to\infty} \frac{2}{(N-w)(N-w-1)} \sum_{i=1}^{N} \sum_{j=i+1+w}^{N} H(r - |\vec{Y}_i - \vec{Y}_j|) \tag{4}$$

where H is the Heaviside step function, $\vec{Y}_i$ is the i-th state vector, N is the number of points on the reconstructed attractor and r is the radius of the sphere centered on Y_i or Y_j.

The number w is called *Theiler window* and it is the correction needed to avoid spurious results due to temporal correlations instead of dynamical ones. For stochastic time series $C_m(r) \propto r^m$ holds, whereas for chaotic time series the correlation function scales with r as

$$C_m(r) \propto r^{D_2} \tag{5}$$

where D_2, called *correlation exponent.* The correlation exponent is defined by

$$D_2 = \lim_{r\to 0} \frac{\ln C_m(r)}{\ln r} \tag{6}$$

and can be reliably 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 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

To cite this paper: Daneshfaraz R, Asadi H, Saghebian SM. 2013. Evidence of Low-Dimensional Determinism in Monthly Water Level Time Series in Lake Van, Turkey. *J.Civil Eng. Urban.*, 2(4): 220-225.
Journal homepage: http://www.ojceu.ir/main/

. For chaotic system, the correlation exponent initially increases but finally saturates after a especial embedding dimension. The saturation value of the correlation exponent is defined as the correlation dimension. Figure 1 shows the above three cases (Stochastic, Deterministic, Chaotic) characteristics.

If the value of correlation dimension is small and fractal, in this case the system is a low-dimensional deterministic chaotic dynamic.

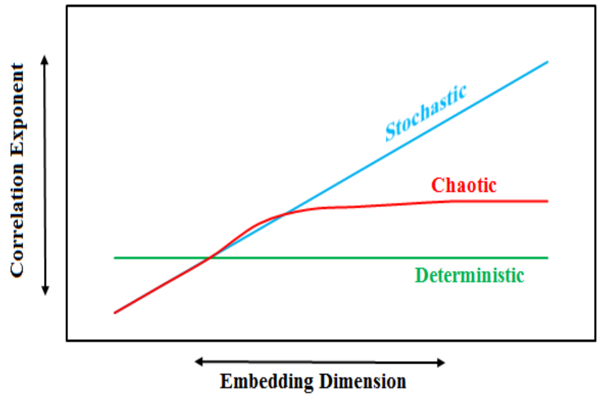

Figure 1. Plot differentiating deterministic, stochastic and chaotic system

Lyapunov exponents:

The unpredictability is one of the important characteristics of a chaotic system, because the sensitive dependence on initial conditions. The largest Lyapunov exponent need only be considered, as it determines the total predictability of the system (Nagesh Kumar and Dhanya 2011). In general, Lyapunov exponent (λ) is a quantitative measure of the sensitive dependence on the initial conditions and to discriminate between chaotic dynamics and periodic signals are often used. Lyapunov exponents quantify the divergence of nearby trajectories in the phase space, along a given direction. Given two nearby states and their Euclidean distance $d(t_0)$ at time t_0, the largest Lyapunov exponent $\lambda_{\max}$, corresponding to the dominant divergence direction, is defined as:

$$\lambda_{\max} = \lim_{t\to\infty} \frac{1}{t - t_0} \log \frac{d(t)}{d(t_0)} \quad (7)$$

In the present work we adopt the method proposed by Rosenstein et al. for the estimation of $\lambda_{\max}$ (Rosenstein et al.1993): it makes use of the stretching factor

$$S(t) = \frac{\Delta t}{t} \sum_{i=1}^{t/\Delta t} \log\left[\frac{1}{|\Omega_i|} \sum_{j\in\Omega_i}^{N} \left|\vec{Y}_i - \vec{Y}_j\right| \right] \quad (8)$$

along an orbit of $\Delta t / t$ time steps, where $|\Omega_i|$ is the number of neighbours in the neighbourhood Ω_i of the reference state $\vec{Y}_i$, and Δt is the sampling time of measurements. For a chaotic dynamics, the stretching factor $S(t)$ is expected to be proportional to time, the largest Lyapunov exponent $\lambda_{\max}$ being the proportionality constant. In other word in the case of chaotic dynamics, a plot of the stretching factor S against the number of points N (or time $t = N. \Delta t$) will yield a curve with a linear increase at the beginning, followed by an almost flat region. The slope of the linear portion of the first part of this curve gives an estimate of $\lambda_{\max}$ (Rosenstein et al., 1993). To be chaotic, $\lambda_{\max}$ must exceed zero. Only for systems with $\lambda_{\max}$ between zero and one are chaotic predictions of any practical use. Usually in practice, one is interested in the maximal Layapunov exponent that can be used to categorise the type of the motion of the system as presented in Table 1.

Table 1. Possible types of dynamics systems and the corresponding maximum lyapunov exponents (Siek 2011)

Maximum Lyapunov exponent	$\lambda_{\max} \prec 0$	$\lambda_{\max} = 0$	$0 \prec \lambda_{\max} \prec \infty$	$\lambda_{\max} = \infty$
Type of dynamics system	Stable fixed point	Stable Limit cycle	Deterministic chaos	Noise (Random motion)

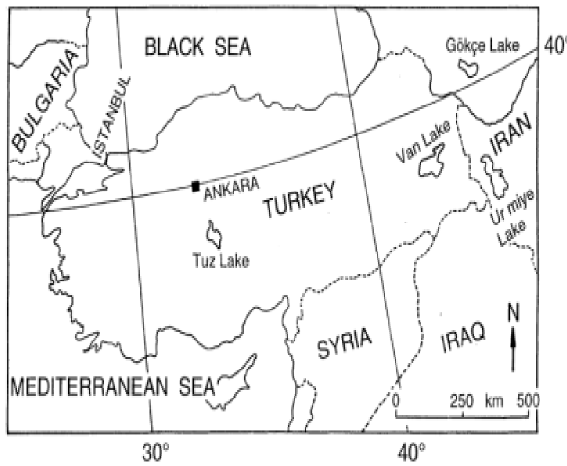

Figure 2. The location of Lake Van in Turkey (Sen et al. 2000)

Table 2. Statistics of monthly lake water level time series at the Van lake

Statistic	**Value**
Number of data	700
Mean(m)	1648.20
Standard deviation(m)	1.14
Maximum value(m)	1650.53
Minimum value(m)	1646.68
Variance(m^2)	1.32
Skewness	0.32
Kurtosis	-1.96

RESULTS AND DISCUSSION

Study area and data used

The monthly lake level data during January 1944 - April 2002 of Lake Van, the largest lake in Turkey, the biggest soda lake in the world, and the world's fourth closed basin lake with a volume of about 600 km^3 is used in the study. The lake is located on the Anatolian high plateau in eastern Turkey (38.5°N and 43°E) (Figure 2). Lake Van has a large drainage basin of 12500 km^2. The lake surface averages about 3600 km^2 , the surface is approximately 1650 m above sea level and the deepest point is 457 meters (Sen et al., 2000). The lake is fed by small rivers, rainfall and melts water of ice. During winter months the lake has the lowest level, and rise after the spring with melting of snow from surrounding mountains. (Çimen and Kisi, 2009). The statistical parameters of the water level data for the Van lake are given in Table 2 and Figure 3 shows the variations of monthly data series.

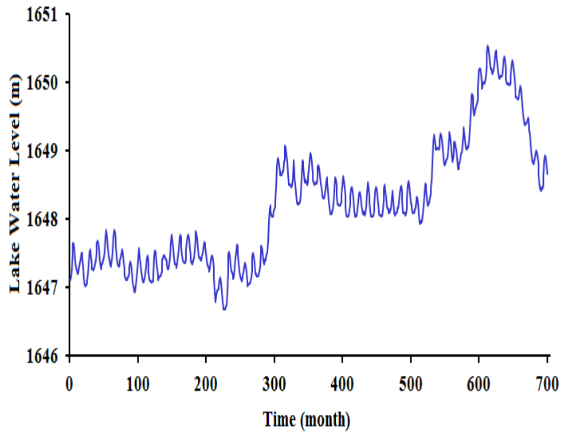

Figure 3. Monthly lake water level time series (January 1944 - April 2002)

Power Spectrum

Figure 4 shows the power spectrum of the monthly water level series observed in the Van lake. The spectrum has been averaged over logarithmically spaced frequency intervals.The value of the β computed from the slope of the solid line, is approximately 4.1, and this an indication that chaotic (fractal) behavior to the lake water level time series during this time interval.

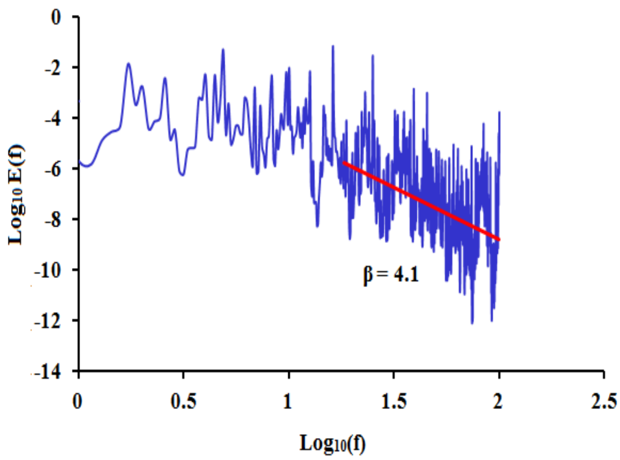

Figure 4. Power spectrum for monthly water level in the Van lake

Time lag and Embedding dimension

In this study τ is computed using the AMI method using time lags of 1-200 month. The AMI shows well-defined first minima at time lag 5 month (Figure 5).

Hence, the optimal embedding delay τ_{opt} is chosen as 5 for our analysis.

The method used for the determination of the sufficient embedding dimension is based on the calculation of the percentage of false nearest-neighbours for the time series. In Figure 6 we show the density of FNN *vs*. the embedding dimension *m* for monthly time series: the optimal embedding is chosen to be *m* = 5.

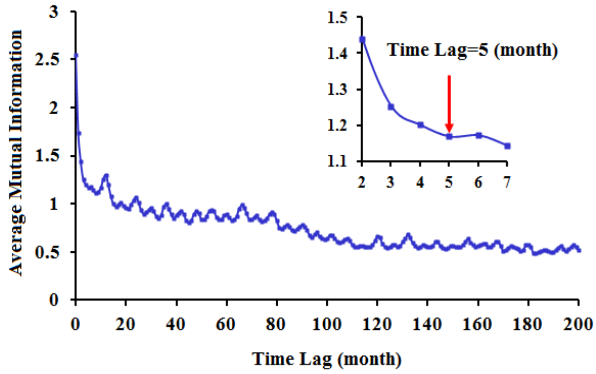

Figure 5. Average mutual information(AMI) function of the lake water level time series dimension

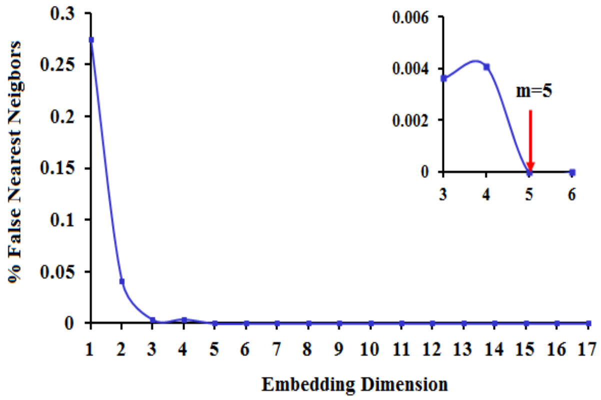

Figure 6. Percentage of false nearest neighbour (FNN) of the lake water level time series in embedding

Correlation dimension

The correlation function calculated for the dataset using the delay times (τ =5), determined by the AMI method in the previous section, and for embedding dimensions, *m*, from 1 to 20. Figure 7 shows the relationship between the correlation function *C(r)* and the radius *r* (i.e. *lnC(r)* versus *lnr*) for increasing *m.*

The relationship between the correlation dimension values $D_2(m)$ and the embedding dimension values *m* is shown in Figure 8.

It can be seen that the value of correlation exponent increases with the embedding dimension up to a certain value and then saturates beyond it. The saturation of the correlation exponent beyond a certain embedding dimension value is the indication of an existence of deterministic dynamics. The saturated correlation dimension is ~ 3.2, ($D_2 = 3.2$). The value of correlation dimension suggests the possible presence of chaotic behaviour and fractal characteristics in the lake water level time series. Because correlation dimension for the monthly water level is above 3, at least four independent variables are needed to describe the dynamics of the lake level variation of the Van lake. This also is taken as the minimum dimension of the phase space that can embedded the attractor.

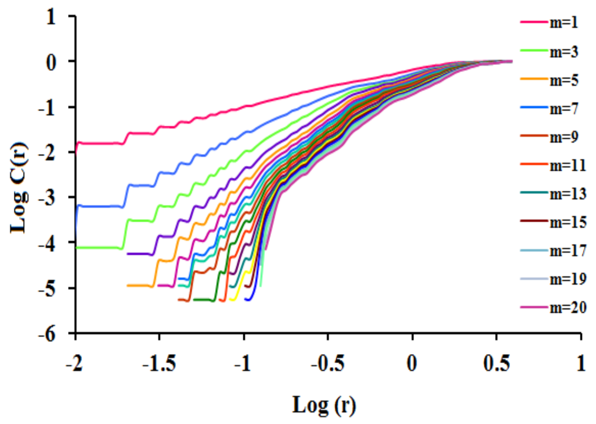

Figure 7. *lnC(r)* versus *ln r* plots for water level data from The Van lake

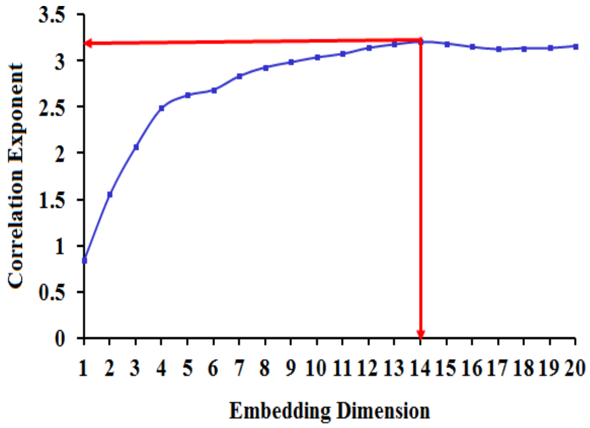

Figure 8. Relation between correlation exponent D2 and embedding dimension *m*

Largest Lyapunov exponent

We apply the method of estimation of the largest Lyapunov exponent described above to the water level time series for the Van lake data, using the same delay time and embedding dimension as before. Figure 9 shows the curve for the stretching factor *S* versus the number of points *N*. The slope value corresponding to the largest Lyapunov exponent is obtained after the least-squares line fit for the water level time series and is found to be 0.0085. This positive value indicate a strong signature of chaos.

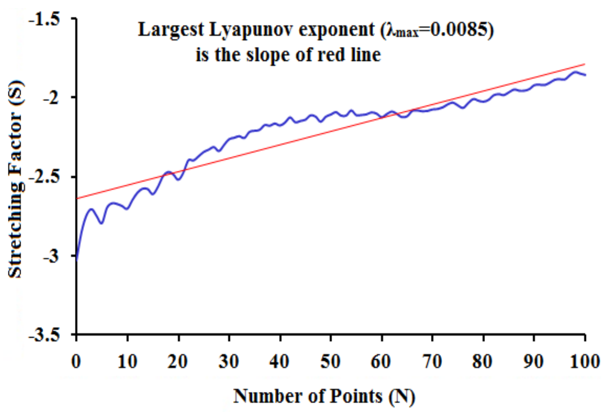

Figure 9. Estimation of the largest Lyapunov exponent using the method of Rosenstein et al. (1993)

CONCLUSIONS

This study investigated the existence of chaotic signals in the monthly water level data in lake Van, Turkey from January 1944 - April 2002. The power spectrum, average mutual information approach, the false nearest neighbor algorithm, the correlation integral analysis and the Lyapunov exponents analysis were used in the research by TISEAN package (Hegger et al. 1999) . The value of the power spectrum indicate that chaotic (fractal) behavior to the lake water level time series. The mutual information approach and the false nearest neighbor algorithm provided a time lag and embedding dimension which is needed to reconstruct phase space. The correlation dimension method provided a low fractal-dimensional attractor thus suggesting a possibility of the existence of chaotic signals. Based on the attractor dimensions, the minimum number of variables essential to model the monthly water level dynamics was identified as 4. Finally the positive largest Lyapunov exponent indicated a existence of chaos behavior, therefore , short-term reliable predictions are possible The authors are not aware of any study exploring possible presence of chaotic behaviors in Van lake level, so that comparsion of the results couldn't possible with another reference.

ACKNOWLEDGMENTS

The authors would like to acknowledge and thank the Prof. Ozgur Kisi from Keyseri university, Turkey for providing the data used in this study. The authors wish to express their special thanks to the anonymous reviewers who have given vital input for improving the quality of the manuscript.

REFERENCES

1. Altunkaynak A (2007) Forecasting surface water level fluctuations of Lake Van by artificial neural networks. Water Res Manag 21(2): 399–408.
2. Altunkaynak A, Ozger M, Sen Z (2003) Triple diagram model of level fluctuations in Lake Van, Turkey. Hydrol Earth Sys Sci 7(2): 235–244.
3. Altunkaynak A, Sen Z (2007) Fuzzy logic model of lake water level fluctuations in Lake Van, Turkey. Theor Appl Climatol 90, 227–233.
4. Cellucci CJ, Albano AM, Rapp PE (2003) Comparative study of embedding methods. Phys Review E 67(6): 66210.
5. Çimen M, Kisi O (2009) Comparison of two different data-driven techniques in modeling lake level fluctuations in Turkey. J Hydro 378: 253–262.
6. Fraser AM, Swinney HL (1986) Independent coordinates for strange attractors from mutual information. Phys Review A 33(2):1134-1140.
7. Frison TW, Abarbanel HDI, Earle MD, Schultz JRW (1999) Chaos and predictability in ocean water levels. J Geophysical Res 104 (4):7935-7951.

8. Grassberger P, Procaccia I (1983) Characterization of strange attractors. Phys Rev Lett 50: 346-349.
9. Hegger R, Kantz H, Schreiber T (1999) Practical implementation of nonlinear time series methods: The TISEAN package. Chaos. 9:413-435.
10. Kennel M, Brown R, Abarbanel HDI (1992) Determining embedding dimension for phase-space reconstruction using a geometrical construction. Phys Rev A 45: 3403–11.
11. Khatibi R, Ghorbani MA, Aalami MT, Kocak K, Makarynskyy O, Makarynska D, Aalinezhad M (2011) Dynamics of hourly sea level at Hillarys Boat Harbour,Western Australia: a chaos theory perspective. Ocean Dynamics 61:1797–1807.
12. 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.
13. Mpallas L, Tzimopoulos C, Evangelides C (2011) Comparison between Neural Networks and Adaptive Neuro-fuzzy Inference System in Modeling Lake Kerkini Water Level Fluctuation Lake Manegement using Artificial Intelligence. J Environmental Science and Technology 4(4): 366-376.
14. Nagesh Kumar D, Dhanya CT (2011) Predictability and Chaotic Nature of Daily Streamflow. 34th IAHR World Congress - Balance and Uncertainty. 26 June - 1 July, Brisbane, Australia.
15. Ng WW, Panu US , Lennox WC (2007) Chaos based Analytical techniques for daily extreme hydrological observations. J Hydro 342:17– 41.
16. Ondimu S, Murase H (2007) Reservoir level forecasting using neural networks: Lake Naivasha. Biosystems Engineering 96(1): 135–138.
17. Rahmstorf S (2003) A semi empirical approach to projecting future sea level rise. Science 315: (5810) 368-370.
18. Rosenstein MT, Collins JJ, Deluca CJ (1993) A practical method for the calculating largest Lyapunov exponents from small datasets. Phys D 65:117-34.
19. Sen Z, Kadioglu M, Batur E (1999) Cluster regression model and level fluctuation features of Van Lake, Turkey. Ann Geophysicae 17: 273–279.
20. Sen Z, Kadioglu M, Batur E (2000) Stochastic modelling of the Van Lake monthly level fluctuations in Turkey. Theoretical and Applied Climatology 65: 99–110.
21. Siek M (2011) Predicting Storm Surges Chaos, Computational Intelligence, Data Assimilation, Ensembles. DISSERTATION Submitted in fulfillment of the requirements of the Board for the Doctorate of Delft University of Technology. PP:207.
22. Sivakumar B (2006) Suspended sediment load estimation and the problem of inadequate data sampling: a fractal view. Earth Surf Process Landforms 31:414–427.
23. Solomatine DP, Rojas CJ, Velichov S, Wust JC (2000) Chaos theory in predicting surge water levels in the Norh Sea. 4-th International Conference on Hydroinformatics, Iowa, USA.
24. 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.
25. Vittori G (1992) On the chaotic features of tide elevation in the lagoon of Venice. In Proceedings of the ICCE’92, 23rd International Conference on Coastal Engineering, 4–9 October 1992, Venice, Italy.361–362.
26. Zaldivar JM, Strozzi F, Gutierrez E, Shepherd IM (1998) Early detection of high water at Venice Lagoon using chaos theory techniques. European report 17317. ISPRA: E.C.