An Investigation on Changes and Prediction of Urmia Lake water Surface Evaporation by Chaos Theory

Farzin, S.1*, Ifaei, P.1, Farzin, N.2, Hassanzadeh, Y.1 and Aalami, M.T.1

1Faculty of Water Engineering, Department of Civil Engineering, University of Tabriz, Tabriz, Iran
2Faculty of Electrical and Computer Engineering, University of Tehran, Tehran, Iran

ABSTRACT: Chaos theory discusses unstable and non-periodic behavior of non-linear fluctuating dynamic systems. Since evaporation level variations of lakes have a dynamic identity, chaos theory can have a unique role in collecting data of these variations. Therefore it is an important and challenging problem in the field of planning and qualitative and quantitative management of Urmia Lake to verify whether the evaporation time series are stochastic or chaotic that is discussed in the present study. In addition to introduction of different utilities of chaos theory, the monthly evaporation amounts of Urmia Lake in the past 40 years have been studied and predicted in the present research. So after calculating the delay time ($T=7.5$) by using average mutual information method (AMI) and embedding dimension ($d=3$) by using false nearest neighbor algorithm (FNN), the slope of correlation dimension diagram has been computed. The non-integer amount of the slope (2.47) represents that the system is chaotic. Lyapunov exponent and broad band in Fourier power spectrum are other indexes reported in the present study and their provided results ensures that the system is chaotic. Thus the amount of Urmia Lake evaporation is predictable. Therefore the amount of evaporation in the recent 10 years (1997-2007) have been predicted by means of false nearest neighbor algorithm and verified with the observed data. The results agree with the high accuracy of chaos theory predictions so the amount of evaporation of the Lake is predicted for 10 following years (2007-2017).

Key words: Urmia Lake, evaporation, chaos theory, correlation dimension, Lyapunov exponent, Fourier power spectrum

INTRODUCTION
The importance and identity of Urmia Lake as an international natural heritage on one side and the comprehensive water sources management on the other side have attracted the attention of various researchers around the world (Ahmadzadeh Kokya et al., 2011; Ahmadi et al., 2011). Urmia Lake is the vastest domestic lake and the second saltiest lake in the world. The lake has been the habitat of many animals. It ranks high importance of different economic, social and environmental points of view in country and district. The water elevation of the lake as a strategic environment has decreased 6 meters in the recent years for various reasons.

There are numerous unknown natural phenomena which seem completely random and irregular in a short scale of time but it’s probable to achieve a regularity and period of frequency by changing the scale and make them predictable in the future. In other words, there is a regularity hidden in any irregularities. This idea is the main basis of chaos theory which studies the unstable and non-periodic behavior in non-linear fluctuating dynamic systems and is too sensitive to initial conditions and the situation at the beginning of the behavior. Chaotic processes are essentially certain. Thus instability, non-periodic behavior, certain systems and non-linearity simultaneously define a chaotic system (Kocak et al., 2000). Since the changes of evaporation level of lakes have a dynamic nature and the management of such sensitive ecologic environments needs precise information at different intervals, chaos theory can play a unique role in collecting information of the phenomena. The theory was first used in 1965 by a scientist called Edward Lorenz in meteorology and developed in all experimental, mathematical, managerial, behavioral and social sciences and topics later that date. It has
Farzin, S. et al.

provided fundamental changes in science especially in meteorology, astronomy, mechanics, physics, mathematics, biology, economy and management (Kocak et al., 2000). Noting that the theory possesses a great capacity to analyze various phenomena, many researchers have already developed extensive researches benefiting from it. Some are noted below: Solomatine et al., (2001) predicted North Sea water levels by using chaos theory and artificial neural network method. The results prove the priority of non-linear methods of univariate time series over linear models like autocorrelation and ARIMA in prediction of water levels on offshore (Stehlik, 1999). In that research, two series of discharge data of ariver in the Czech Republic with different time delays were studied in order to verify the possibility of the chaotic behavior of the data. Regonda et al., (2004) studied the data of three rivers in different time scales of 5 and 7 days to see if they are chaotic. Some of data series appeared to be chaotic and some stochastic. Khan et al., (2005) 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., (2007), 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. Ng et al., (2007) studied the application of analytical chaotic techniques for daily noisy flow series. Wu et al., (2009) used chaos and geometric fractal models with low-use data to predict water quality time series and compared the results with Grey and AR time series models and concluded that the accuracy of chaos model is higher than Grey model and AR time series model. Shang et al., (2009) used non-linear time series modeling techniques to analyze suspended sediments data. The results showed that there are chaotic properties in sediment transport.

As discussed; evaporation study and consequently the water level have a great importance in the fields of lake ecologic values, the amount of run-off, agriculture, water sources management and generally a lot of daily issues and various experimental, semi-experimental and intelligent models have been developed for that. Noting the effects of various factors on evaporation in a district, it is likely to introduce evaporation changes system as a dynamic system. Since chaos theory studies dynamic systems, it can be a modern approach to study this process by the use of the theory which has not been developed yet.

MATERIALS & METHODS

Location and geographic characteristic: Urmia Lake is about 140 kilometers long, 55 kilometers wide and at most 18 meters deep in its extreme extended limit. The lake is located in northwest of Iran. Its area has fluctuated between 4000 to 6000 square kilometers in the past years but the average area is estimated about 5000 square kilometers. The Lake is the vastest wetland in Iran plateau and the twentieth most extensive lake in the world. The average salinity is about 220 to 300 grams per liter, depending on various space and time conditions, that made it the second over saturated salt lake in the world. The lake is one of the most important natural habitats in the district and it is known as a national park by UNESCO because of its unique specialties. Fig. 1 displays the geographic location of the district and Urmia Lake (ME-MWP, 2006).

![Fig.1. Geographic location of Urmia Lake in northwest of Iran](image)

A variety of changes have occurred in Urmia Lake water level in the recent years. Fig. 2 displays the water elevation changes while the study period of time (1967-2007) in meters. The highest and lowest water elevations of Urmia Lake have been 1277.67 and 1272.18 meters respectively in 1995 and 2007 (Ministry of Energy, 2006).

Chaos theory: Chaos is a non-linear behavior in a range between periodic and random behavior. These kinds of systems are sensitive to initial conditions and the situation at the beginning of behavior. Chaotic systems sometimes seem regular or periodic or even random at first glance. Both assumptions (being periodic or random) are incorrect in any cases. Chaotic systems are definitely certain systems that mean the system can be determined in any moment as:

$$X(t) = (x(t), x(t - \tau), x(t - 2\tau), ..., x(t - (k - 1)\tau))$$

Eq. (1)
Fig. 2. Urmia Lake water elevation changes in a 40 years period

Where \( t \) = a scalar index corresponding to time in order to represent data and \( \tau \) = intervals between observations. Now suppose \( F \) = non-linear function dominating the system. Since the system is determined for any \( \tau + t \), it is written as:

\[
F : \mathbb{R}^k \rightarrow \mathbb{R}^k
\]

\[
x(t + \tau) = F(X(t)) + p(t)
\]

Probability variable \( p(t) \) is supposed to be equal to zero average because of rounding, adjustment or measurement errors. The present error (the difference between two adjacent states) in determined systems either remains too small (sustainable systems) or grow exponentially (chaotic systems) whereas the error and difference distributes randomly and stochastically in random systems (Frazier et al., 2004).

Embedding dimensions: As the chaotic systems behavior depends on various parameters in any moment, \( k \) has a great amount. Additionally, as the effective parameters in a system increase, data volume and calculations increase greatly. So it will be better to use random variable models to describe a system with amounts greater than 10 for \( k \). There are various methods to decrease system dimensions and to choose parameters which explain system behavior with acceptable errors. One of the most prominent methods is finite dimensions theory. Vector \( y(t) \) is supposed instead of vector \( x(t) \) in the theory after decreasing system dimensions from \( k \) to \( d \). the theory is written as:

\[
s(t) = h(x(t)) \quad R^k \rightarrow R
\]

\[
y(t) = [s(t), s(t + T), s(t + 2T), \ldots, s(t + (d-1)T)]
\]

Where \( T = n\tau \quad (n = 1, 2, 3, \ldots) \) = time lag. If the quantity \( d \) is great enough vector \( y(t) \) rebuild most of important dynamic properties of \( x(t) \). So there is no need for to explain many system properties. Two steps should be taken in order to apply the theory effectively: in the first step, time lag \( (T) \) and in second step the finite dimension \( (d) \) is obtained (Abarbanel, 1996).

Mutual Information function can be used to measure time lag \( (T) \) in the first step (equation 4). As reported by (Frazier et al., 2004) an appropriate amount for \( T \) would be the first minimum amount of the function.

\[
I(\tau) = \sum_{i(n), j(n+\tau)} \left\{ \frac{\text{Pr}(s(t), s(t + T))}{\text{Pr}(s(t)) \text{Pr}(s(t + T))} \log_2 \frac{\text{Pr}(s(t), s(t + T))}{\text{Pr}(s(t)) \text{Pr}(s(t + T))} \right\}
\]

Where \( \text{Pr} \) = probability function amount that can be extracted from the corresponding data histogram. To determine the finite dimension \( (d) \) in the second step a method developed by Cao is usually used; as reported by (Cao, 1997):

\[
E1(d) = \frac{E(d + 1)}{E(d)}
\]

\[
E(d) = \frac{1}{N - dT} \sum_{t=0}^{N-dT-1} \left\| y_{d+1}(t) - y_{d+1}^{NN}(t) \right\|
\]

Where \( N \) = data series length, \( d \) = finite dimension and \( NN \) = closest adjacent vector to another vector that is determined as:

\[
\left\| y_{d}(t) - y_{d}^{NN}(t) \right\| = \max_{0 \leq j \leq d} \left| s(t + jT) - s^{NN}(t + jT) \right|
\]

\( E1(d) \) approaches one as the amount of \( d \) increases. The most appropriate amount of \( d \) is an amount where the changes of \( E1(d) \) are stopped and an approximately constant procedure is followed. The following quantities are also defined for systems with random variables:

\[
E2(d) = \frac{E^*(d + 1)}{E^*(d)}
\]

\[
E^*(d) = \frac{1}{N - dT} \sum_{t=0}^{N-dT-1} \left\| s(t + dT) - s^{NN}(t + dT) \right\|
\]
Lake water Surface Evaporation

The index is spike like for periodic data at frequencies representing the system and close to zero at other frequencies but the index has a broadband in chaotic systems. It is very difficult to distinguish between chaotic systems and the systems with intense random behavior by using this index (Ng, 2007). The index is plotted in fig. 3 for a chaotic behavior (on the right side) and a periodic behavior (on the left side):

Fractal dimension concept: Fractal is originated from the Latin word “Fractus” that means broken and irregular pieces and is used to explain fractal dimensions concept. The concept was developed by Hausdorff in 1919 specifically in mathematics and physics. The concept was expanded in 1967 by Mandelbrot to describe random and irregular natural phenomena. Thus the French mathematician, Mandelbrot, attracted attentions by introducing fractal geometry as geometry of the nature in 1983 (Turner, 1998). Mandelbrot observed that the coastline does not approach L by lessoning the measurement unit but increases instead. A relationship as

\[ L \propto \varepsilon^{1-D} \]

is obtained while the increasing amount of L is plotted on an axis with the decreasing amount of \( \varepsilon \) in a logarithm-logarithm scale. It is easy to understand that the real number D is the fractal dimension of the coastline regarding fig. 4 (Mart, 1992).

That was the beginning point of explaining several concepts originating from fractal geometry concerning self-similarity. So many explanations and descriptions of chaotic and irregular phenomena in science and engineering were born.

The main property of a fractal model description is its dimension. A dimension is a simple concept in Euclidean shapes described by ordinary integers. For instance the dimension amount is 2 on plane. But in a fractal model, structure repeat causes a line to occupy the surface. The amount of dimension is varied from 1 to 2 in this case. The amount approaches 2 by the increase of complexity and the richness of repeating structure.

Fractal dimension measurement methods: There have been a noticeable number of algorithms to obtain fractal dimension in the recent fifty years which can be classified in two general categories:
1-size-measure relationships: based on the recursive length or area measurement of a curve or surface using various measurement scales.
2-application of relationships: based on approximation or fitting a curve or surface to a known fractal function or a statistical property like variance.

A. Box counting method: A current algorithm in measuring signals fractal dimension of photos is box counting method developed by Wu for the first time. Faster and more accurate algorithms of the method were reported later (Hilborn, 2000).

The fractal body is covered by a network of squares or cubes with side lengths of R and a few not-empty
squares $N(R)$ in box counting method. The body fractal dimension, $D_b$, will be determined as:

$$D_b = - \lim_{R \to 0} \frac{\log[N(R)\!]}{\log R}$$  \hspace{1cm} \text{Eq. (10)}

Slope of logarithm curve of $R$-$N(R)$, $D_b$, is obtained by graphical method considering the linear part which helps obtaining the fractal dimension note that:

$$D_b = - b$$  \hspace{1cm} \text{Eq. (11)}

The dimension is called box dimension or Minkowski dimension.

B. Correlation dimension: Since box counting method is too complicated or sometimes impossible for real data with many dependent variables, the correlation dimension was first invented by Grassberger and Procaccia in 1983 to facilitate algorithm calculations (Hilborn, 2000). The algorithm represents the dynamicity of a system according to its corresponding time group. It is not necessary to have correlation between variables to plot cause and effect diagram but the existence of a causality relationship is essential. A simulated model has to reconstruct the system structure and imitate its behavior somehow. Behavior imitation is not limited to old experiences rebuilding but it has to be able to respond quietly inexperienced phenomena and policies. The correlation between variables represents situation in the past (Sterman, 2000). Correlation dimension is one of the current methods to determine system chaos and chaotic dimension. The quantity is not an integer for chaotic systems. If a sphere is imagined around specific points of data, the average of points’ number in the sphere except the center of the sphere can be written as (Hilborn, 2000):

$$C(R) = \frac{1}{N(N-1)} \sum_{i \neq j} \sum_{j \neq j'} \theta(R - |x(i)-x(j)|)$$  \hspace{1cm} \text{Eq. (12)}

Where $N$=data amplitude length, $\theta(x)$= Heaviside step function:

$$\theta(x) = \begin{cases} 
0 & x < 0 \\
1 & x \geq 0 
\end{cases}$$  \hspace{1cm} \text{Eq. (13)}

As $R$ approaches zero, $C(R)$ changes rate will be:

$$C(R) = \lim_{R \to 0} k R^{D_c}$$  \hspace{1cm} \text{Eq. (14)}

$D_c$ is determined by solving equation 17 as:

$$D_c = \lim_{R \to 0} \frac{\log C(R)}{\log R}$$  \hspace{1cm} \text{Eq. (15)}

$D_c$ represents correlation dimension in the equation. Since data series are not continuous, $R$ cannot approach zero too much (there will be no points in the sphere found). Three diagrams can be used to verify chaos and estimate correlation dimension:

- $\log C(R)/\log R$ versus $\log R$ diagram. Scaling part can be determined by this diagram that is the part where for the continuous amounts of $\log(R), \log C(R)/\log R$ ratio reaches a constant amount for various embedding dimensions and is saturated that is the fractal dimension.

- Correlation power changes versus various embedding dimensions diagram. Chaotic behavior and the appropriate correlation dimension can be determined by this diagram.

- LogC(R) versus Log(R) diagram. Diagram slope for various embedding dimensions can be determined using minimum squares method in scaling part. $\ln C(R)$ versus $\ln R$ diagram can be drawn in this case (see fig. 11); the slope of the linear limb of diagram can approximate $D_c$. If $D_c$ is not an integer, one of the important properties of chaotic systems appear (Grassberger and Procaccia, 1983).

Prediction: All dynamic systems affected by changes will inevitably encounter prediction while deciding about future. Prediction is a tool to help deciding, so the answers of decisive questions should be found using prediction. Prediction value depends on its benefit in decision time. Any predictions improving decision quality will be useful ignoring what happens in the future.

Little changes will cause great consequences in non-linear systems discussed by chaos theory but this problem needs time to appear. So the short time behavior of such a system may be logically predictable. Thus, designing near future process and the success of programming lies on the predictability of the system. The problem of evaporation changes of Urmia Lake is undoubtedly an important problem and a prediction based on precise criteria can be helpful in its effective control, programming and offering managerial strategies for coming years. This aim has been attained using the false nearest neighborhood algorithm.

RESULTS & DISCUSSION

The evaporation of Lake Surface is the main outflow of the lake. The monthly amounts of evaporation are verified and estimated in the study period of time. The monthly amounts time series of lake surface evaporation in 40 years are illustrated in fig. 5.
The best amount of $d$ is that the changes of $d$ are stopped and there is an almost unchanging procedure followed. The evaporation amounts statistical properties used in the present research are calculated and represented in Table 1.

Table 1: The corresponding statistics of the studied evaporation data

<table>
<thead>
<tr>
<th>Data number</th>
<th>480</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average (mm)</td>
<td>127.27</td>
</tr>
<tr>
<td>Standard deviation (mm)</td>
<td>90.74</td>
</tr>
<tr>
<td>Maximum amount (mm)</td>
<td>329.67</td>
</tr>
<tr>
<td>Minimum amount (mm)</td>
<td>2.58</td>
</tr>
<tr>
<td>Skewness (mm)</td>
<td>0.415</td>
</tr>
</tbody>
</table>

The amount of the time lag is presumed to be equal to the first minimum amount of mutual information function when finite dimensions theory is used. Mutual information function is illustrated in Fig. 6 for this purpose. Since appropriate $T$ is equal to the first minimum amount of the function, data time lag will be 7.5 according to the figure.

$E(\hat{d})$ and $E(\hat{d})$ are also calculated (see Figs 7 and 8). It is raised to some extent according to the fig. and has a loop at $d=3$. Thus the embedding dimension will be equal to 3.

According to Fig. 8, it is obvious that this diagram remains less than one for little amounts of $d$ that represents the chaotic behavior of the system. Fourier power spectrum, Lyapunov exponent and correlation dimension are respectively illustrated in Figs 9, 10 and 11 to verify chaos in the system.

It is clear that the broadband in Fourier power spectrum represents chaotic behavior of the system. Referring to the preceding figure, it is clear that the Lyapunov exponent is positive that agrees with chaotic behavior of the system and its sensitivity to the initial conditions.

$logC(R)$ versus $log(R)$ is illustrated in Fig. 11 to determine the correlation dimension. The slope of the linear part of the diagram approximates $D_c$. Getting profitable from plotting versus $d$, varies according to $d$ increasing and never reaches a saturated amount in a periodic process whereas the amount of is saturated after a certain $d$ in a definite process. The saturated amount is attractor fractal dimension (correlation dimension) or time series (Elshorbagy et al., 2002). Concluded from the preceding fig. non-integer amount of correlation dimension diagram slope (approximately 2.47) represents the chaotic behavior of the system. Eventually, a brief review of chaos indexes; broadband of Fourier power spectrum, positive Lyapunov exponent and non-integer amount of correlation dimension diagram slope (approximately 2.47) all confirm the chaotic behavior of the system.

In order to predict the evaporation amount by chaos theory using false nearest neighborhood algorithm, evaporation level is predicted in last 10 years (1997-2007) at first. The results are compared with observed data to verify the accuracy (see Fig. 12). The calculated correlation coefficients (about 0.9) and square root of the average of error squares (0.23) represent an acceptable accuracy of chaos theory in predicting the evaporation level of Urmia Lake. Since the verification of accuracy is satisfying, the evaporation amount is predicted for following 10 years (2007-2017). Fig. 13 illustrates the results.
Lake water Surface Evaporation

Fig. 9. Broadband in Fourier power spectrum of studied period

Fig. 10. Positive Lyapunov exponent

Fig. 11. Non-integer slope of correlation dimension diagram (approximately 2.47)

Fig. 12. Comparison between evaporation time series and observed data
CONCLUSION

The calculation of Urmia Lake water surface evaporation amount and subsequently verification of water level changes in order to protect this unique natural heritage is an important issue with a dynamic nature. Such an ever-changing phenomenon must absolutely get analyzed by modern and proper methods. Noting the fact, chaos theory seems to be a powerful tool to analyze, model and control such a phenomenon. The present research is an effort to verify the function of chaos theory in modeling the complicated non-linear behavior of Urmia Lake water surface evaporation amount in a 40 years period of time. It appears to be correct to assume the system non-linear according to parameters like human factors, weather situation and other physical reasons. The amounts of delay time $T=7.5$ and $d=3$ were obtained using auto-correlation and counting false nearest neighborhood methods. The results obtained from the amounts of Fourier power spectrum, Lyapunov exponent and correlation dimension were used to distinguish between chaotic and stochastic behavior. The non-integer amount of correlation dimension equal to 2.47, broadband of Fourier power spectrum and positive Lyapunov exponent all confirmed the chaotic behavior of water evaporation changes. The $E^2(\hat{d})$ diagram is less than one for small amounts of $d$ that confirms the chaotic behavior of studied time series too. The false nearest neighborhood algorithm was used to predict the evaporation amount by chaos theory. For that, the evaporation amounts in the last 10 years (1997-2007) were predicted at first and then compared with observed data to verify the accuracy. The verification results confirmed high accuracy of chaos prediction. Finally Urmia Lake water surface evaporation amounts were predicted for following 10 years (2007-2017). The results are important for future programming and proposing managerial strategies.

REFERENCES


