



Investigation of recent level changes in Lake Van using water balance, LSTM and ANN approaches

Mehmet Cihan Aydın¹ · Gamze Gelberi² · Ali Emre Ulu¹

Received: 10 March 2022 / Accepted: 26 December 2023 / Published online: 8 February 2024
© The Author(s) 2024

Abstract

Lake Van, the greatest soda water lake in the world located in the east of Turkey, has always attracted the attention of researchers due to its significant water level changes. Identifying the causes for these level changes is very important with respect to the scientific world and the sustainability of the ecological balance. Although extensive research has been carried out on the water level changes in Lake Van in the past years, no any study exists which analyzes the recent level changes. In this study, recent water level changes in Van Lake were analyzed using two different methods, and the results were evaluated. First, the water level changes during the period 2010–2020 were examined through the meteorological and hydrological data collected by the water budget method. Second, the water level changes for 2000–2020 were estimated using the LSTM and NAR neural networks. In the light of the results, reasons for the recent level changes in Lake Van were discussed. It was concluded that the changes, especially those occurred after 2014, may be hydrometeorological. However, the unusual level changes between 2010 and 2013 cannot be explained by water balance, may be tectonic or volcanic origin. The findings of this study indicate that these changes may continue in the future, and therefore, further research with different disciplines is needed on this subject.

Keywords Lake Van · Level change · Water budget · LSTM-RNN · NAR · ANN

Introduction

As the world's population grows, the increase in industrial activity on the planet is negatively affecting the ecological balance on a daily basis, leading in the long term to global climate change. These interactions are most evident in enclosed lake basins, which are natural water resources. Many of the world's natural lakes have experienced significant changes in water levels in recent years, and some are even at risk of drying up. In addition to hydrometeorological variables, tectonic, geological and seismic factors can be

the main causes of lake level changes. Therefore, predicting the level changes in lakes is an important issue in order to monitor and control their environmental effects (Şener and Morova 2011). However, changes in lake levels are the complex result of many environmental factors, including precipitation, direct and indirect runoff, evaporation from the free water surface, air and water temperature, and interactions between the lake and shallow aquifers (Kişi et al. 2012). Therefore, the estimation of lake water levels is very difficult due to nonlinear hydrological processes. It is more economical to create models that predict level oscillations using level data from previous years, although it is possible to create complex models using hydrological and meteorological data such as precipitation, temperature, evaporation, and flow to predict lake level oscillations (Aytek et al. 2014).

Lake Van, which is the world's greatest soda lake, has undergone significant level and volume changes under cold and humid climatic conditions during glacial and Holocene periods (Landmann et al. 1996). The most notable indicator to level change in the past may be the formation of terrace deposits around the lake (Aydın and Doğu 2018; Üner et al. 2012). Tomonaga et al. (2017) found that the salinity

✉ Mehmet Cihan Aydın
mcaydin@gmail.com

✉ Ali Emre Ulu
aliemreulu@gmail.com

Gamze Gelberi
gelberi.gamze@gmail.com

¹ Department of Civil Engineering, Bitlis Eren University, Bitlis, Turkey

² State Hydraulic Works 17th Regional Directorate, Van, Turkey

changes in the pore water of the sediment in the lake indicate major lake level changes in the past. According to previous studies on lake terraces and with seismic and sedimentological surveys, it was recognized two major transgressions of up to +105 m from the current lake level. Çukur et al. (2014) investigated potential causes related to changes in climate, volcanism and regional tectonics since the formation of Lake Van. The study, which refers the dramatic changes in the water level from the formation of the lake, showed that these changes in Lake Van are more likely climatic, but the changes could also be tectonic and volcanic due to a tectonically active environment.

The prominent past studies focused on water level changes in Van Lake are shown in Table 1. While some of these studies suggest that the changes in water levels in Lake Van are hydrometeorological, others argue that factors such as tectonic or seismic activity may also be at work. These studies generally cover the water level changes of Lake Van in the period 1944–2016, for which historical data are available. As can be seen in Table 1, the water budget approach is generally used to analyze the water levels of Lake Van, but in recent years, artificial neural networks have been preferred

together with some hybrid soft computing methods. Difficulties in deterministic modeling of multiple and highly uncertain nonlinear events have led most researchers to use black box models such as stochastic and artificial intelligence (e.g., Altunkaynak 2007; Teltik 2008; Kişi et al. 2012; Aytek et al. 2014; Seo et al. 2015; Shafaei and Kişi 2016; Doğan et al. 2016; Liang et al. 2018; Piasecki et al. 2018; Hrnjica and Bonacci 2019; Baek et al. 2020; Aydin et al. 2021). Some of these studies showed that the long short-term memory (LSTM) method is quite successful for estimating lake levels (Liang et al. 2018; Hrnjica and Bonacci 2019; Baek et al. 2020). In addition, to provide hydrologists and water resource planners with a new vision for the sustainable management of lakes, Zhu et al. (2020) presented a comprehensive review of the applications of machine learning models for predicting lake water level dynamics. Fijani et al. (2021) developed a hybrid algorithms for forecasting multistep ahead (up to three months) two lake water levels. For the generation of robust multistage lake water level forecasts, this algorithm was shown to have great potential.

Although many studies have been conducted on the level measurements of Lake Van in the past, there is no study yet

Table 1 Literature review on the study of water levels changes in Van Lake

References	Method	Period	Highlighting
Batur (1996)	Water budget	1987–1994	The change in lake level is hydrometeorological
Kadioğlu et al. (1997)	Historical data analysis with monthly climate polygons	1944–1994	Climate change plays a crucial role in the fluctuation of water levels in Lake Van
Kadioğlu et al. (1999)	Dynamic water balance model	1944–1994	The rise in the lake level after 1968 is hydrometeorological
Yıldız and Deniz (2005)	Observational and measurement	1950–2000	The rapid rise in water levels between 1993–1994 caused serious ecological
Utkucu (2006)	Historical data evaluations	1869–2003	The lake level oscillations correlate well with the $M \geq 4.0$ earthquakes in the lake
Batur (2009a)	Water budget	1987–1996	The reason for the change in lake water level can be explained by the water budget
Batur (2009b)	Multiple regression analysis	1944–2007	The volume change in the lake is related to hygienic variables
Düzen (2011)	Hydrogeological approach, water budget	1960–2009	The water budget showed that the lake may be fed by groundwater
Kişi et al. (2012)	Artificial neural network (ANN)	1961–1982	GEP is superior to all of the applied models
Aksoy et al. (2013)	Stochastic models	1944–2007	The multi-trend model is suitable for planning future projects around the lake
Aytek et al. (2014)	Gene expression programming (GEP)	1944–2002	GEP model is better than neuro-fuzzy models
Utkucu and Arman (2015)	Historical data evaluations	1944–2010	$M \geq 5.0$ earthquakes may triggered 1 m water level fluctuations
Doğan et al. (2016)	ANN	2001–2012	The lake level fluctuations have a long-term decreasing trend
Jalili et al. (2016)	Trend and coherency analyses	1965–2013	Significant common inter-annual frequencies were found between the two lake levels
Altunkaynak (2019)	The hybrid additive season algorithm–multilayer perceptron (ASA-MP) model	1960–2016	The ASA-MP model showed a very good performance for prediction the lake level series
Yaseen et al. (2020)	Multilayer Perceptron (MLP) and Whale Optimization Algorithm (WOA)	1943–2016	The MLP_WOA model is an applicable method for predicting the lake levels

on the level changes in the recent years (2010–2020). In this study, the changes in Lake Van for this period were analyzed using the water budget together with some soft computing methods, which are the heuristic time series methods used in recent years. Thus, the time series of the lake level were estimated considering the hydrological data such as lake level measurements, stream discharges flowing into the lake, precipitation and evaporation data.

Lake Van

Located in the east of Turkey, Lake Van is Turkey's largest and also the world's largest soda and salty lake. The surface area of the lake varied between 3550 and 3620 km² in 1973 to 2020. The total drainage of Lake Van is ~15,500 km² that includes all lakes in the basin. Its surface area was between ~3550 and ~3620 km² in 1943 to 2016, with an average 3580 km². Its volume also varied between ~560 and ~575 km³ with an average of 568 km³ in same period (Düzen 2011). The deepest point of Lake Van was estimated as 452.9 m for the water level of 1648.33 m from evaluation data of bathymetric maps (Degens and Kurtman 1978; SHODB 1983; Düzen 2011; Aydin and Karakuş 2016; Khatibi et al. 2020).

The lake was formed by the separation of the Nemrut lava wall, which was formed because of the activity of the Nemrut volcano, and the Muş-Van Depression and the accumulation of water behind the volcanic wall. There are mountains and hills around the lake, and the peaks of these elevations can reach 4000 m in some points (Kadioğlu et al. 1997; Aytek et al. 2014). The lake water cannot be used for drinking and agricultural irrigation due to its high salinity. The precipitation falling on the basin and the lake, the streams flowing into and the snowmelt feed the lake. The streams feeding the lake are shown in Fig. 1. A large amount of precipitation falls as snow in the region during the winter months, and the lowest water level in the lake is usually measured in this season. The lake has a drainage area of 12,500 km² excluding drainage areas of the other lake in the basin and has no natural outlet (Erinç 1953; Aytek et al. 2014). The main streams feeding the lake and their characteristic features are given in Table 2. In general, it is reported that 4.2 km³ of water evaporates annually from the lake surface and these losses are balanced by the long-term averages of annual precipitation and runoff amounts of 1.7 and 2.5 km³, respectively (Aytek et al. 2014).

The annual water level values between 1944 and 2020 of the Tatvan Lake observation station are given in Fig. 2 (DSI 2021). According to these data, there has been a continuous increase in the level of Lake Van over a 76-year period. The maximum level increase in this period is remarkable with about 4 m. The highest water elevation was measured as

1650.53 m in March 1989 and June 1995, while the lowest value was measured as 1646.67 m in September 1956. There was a 3.85 m level difference between the maximum and minimum water levels in this historical period. The most significant increases occurred between 1961–1964, 1981–1989, 1991–1996 and 2002–2008. Due to the level rise in these periods, some houses, roads and agricultural areas around the lake were flooded (Aydin and Doğu 2018).

Water budget method

The main cause of lake level fluctuations may be seasonal and climatological changes. These changes may be revealed by the water budget (water balance) method. Therefore, some hydrological and meteorological data are needed to be able to analyze the water budget. However, the achievement of this method depends on the quantity and quality of the measured data. On the other hand, it is possible to find a logical relationship between the lake level oscillations and the changes in the hydrometeorological data. For the water budget analysis, two data set were considered as inputs to Lake Van: (i) The monthly average values of the precipitation falling on the lake surface, and (ii) The flow values obtained from the flow observation stations of the main streams (DSI 2021) entering the lake (Table 2). For the output, only evaporation values were taken in accordance with the assumptions that there is no groundwater and surface water output.

The water budget can be determined by the continuity equation given in Eq. (1) for a closed catchment. Unlike black box techniques, the water budget method is based on basic deterministic relationship between inputs and outputs. Thus, the accuracy of the method often depends on the nonlinear input and output data and the logical relationship between them. In addition, the method allows a simple physical model to be established between its inputs and outputs. The water budget method in this study allows a simple physical model to be established between hydrometeorological inputs and outputs.

$$Q_i - Q_o = \Delta V / \Delta t \quad (1)$$

In Eq. (1), Q_i and Q_o are inlet and outlet, respectively, while ΔV is change in the water volume and Δt is time interval. Some assumptions need to be made to setup a simple mathematical model in which only the active parameters are taken into account. The water budget chart for Lake Van is given in Fig. 3. In this system, the runoff directly entering the lake, the river flows and the precipitation falling directly on the lake surface can be considered as the total inlet to the lake. The runoff coming from around the lake and the sub-basin areas without a measuring station were calculated

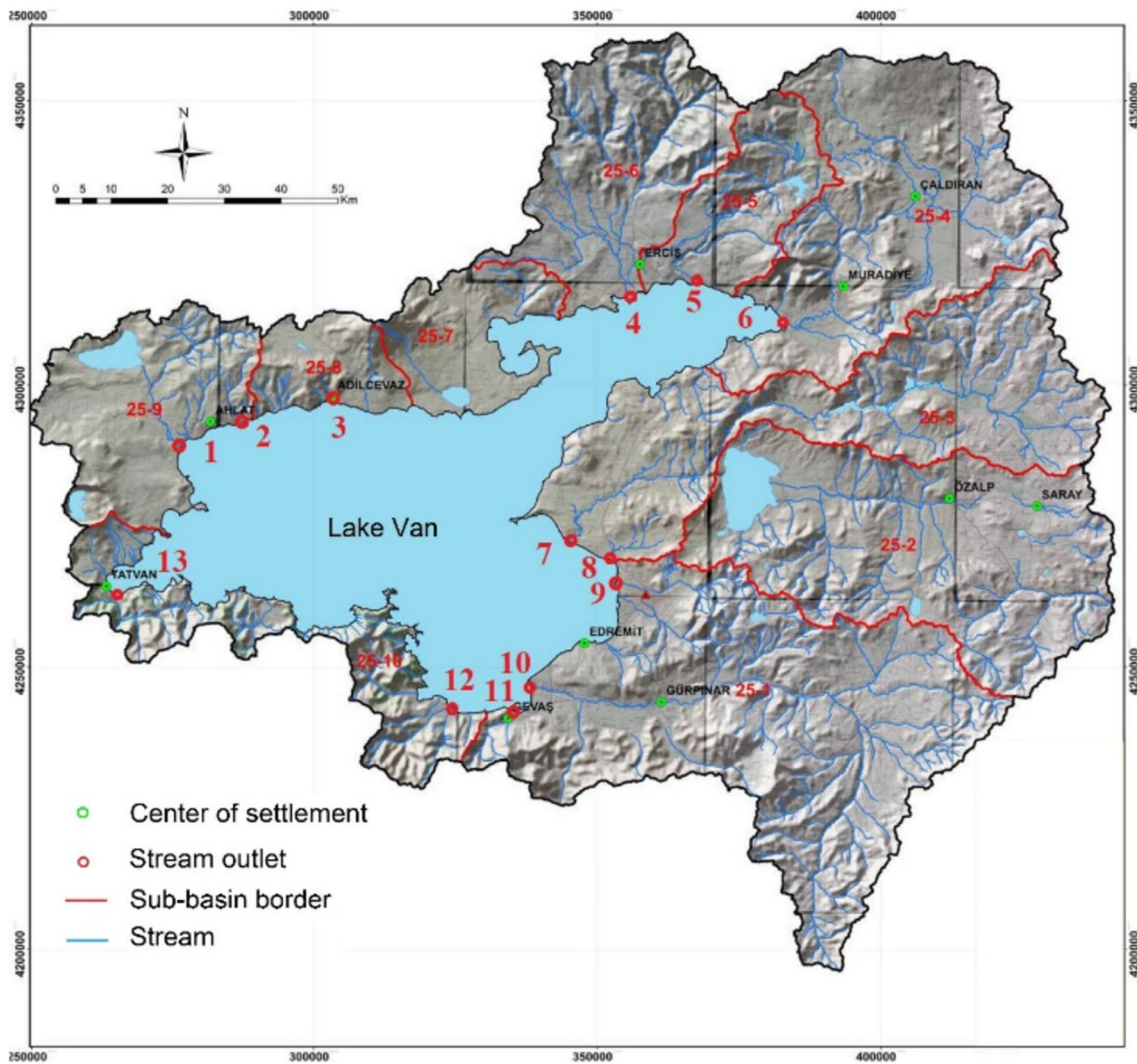


Fig. 1 Lake Van sub-basins and stream outlets (modified from DSI 2015)

Table 2 Specifications of the main streams discharging into Lake Van (DSI 2018)

Stream	Location	Flow Meas. Station	Altitude (m)	Mean flow rate (m ³ /s)	Annual yield (L/s/km ²)	Total annual flow (Million m ³)	Annual rainfall depth (mm)
Kasap S. (Adilcevaz)	Adicevaz	D25A045	1654	0.33	N/A	10.55	N/A
Delicay S	Ercis, Kadirasker	D25A013	1992	2.75	8.30	86.72	263
Bendimahi S (Muradiye S.)	Van, Karahan	E25A005	1915	11.15	7.90	342	249
Karasu S	Van, Topaktas	D25A039	1730	2.78	1.60	87.85	51
Engil Suyu (Dönemec S, Guzelsu S.)	Van, Cicekli	E25A027	1652	2.49	1.00	78.55	31
Gevas S	Van, Gevas	D25A021	1652	0.66	11.10	20.69	350
Buyukdere S	Gevas, Akdamar	D25A038	1658	0.27	2.60	8.60	82
Kotum S. (Kucuksu)	Tatvan, Kucuksu	D25A040	1750	1.65	24.60	52.20	777

Fig. 2 Annual max. and min. water level measurements between 1944 and 2020 in Lake Van (DSI 2021)

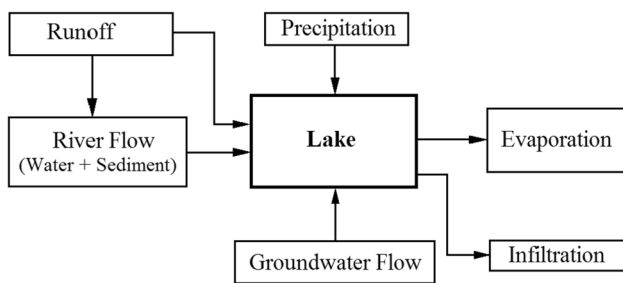
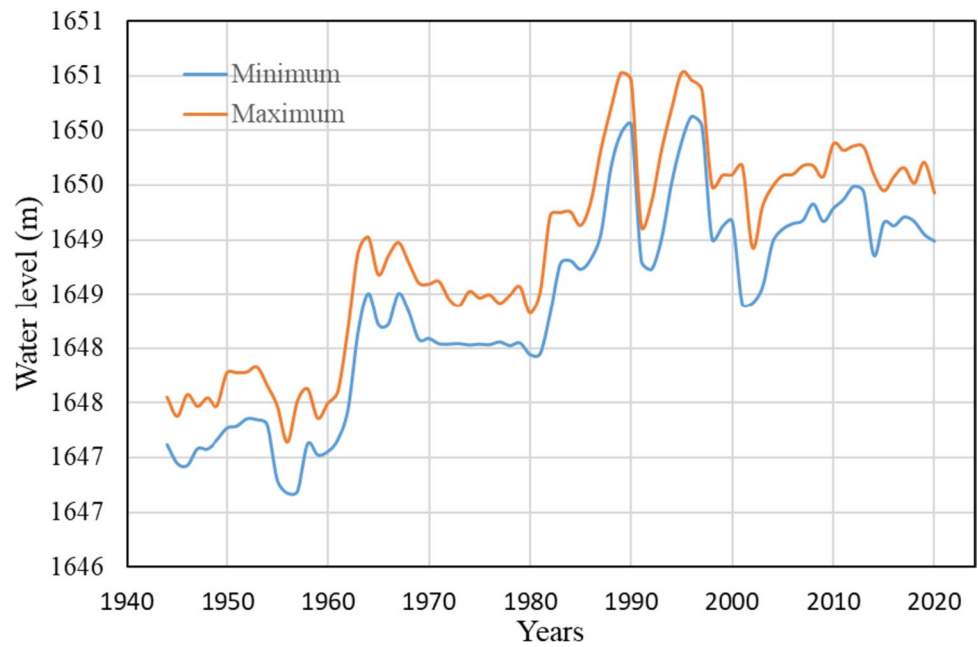


Fig. 3 Water budget scheme for a closed basin, Lake Van (modified from Ozer et al. 2022)

that depend on the precipitation, basin area and runoff coefficients, and added to the inflow. System outputs, on the other hand, are more complex. Evaporation can be considered as the only outlet, assuming that the ground below the lake level is completely saturated, and there is no infiltration. Since the lake water is salty, it can be assumed that there will be no losses such as irrigation and drinking water. The level changes to be obtained by this method will represent the hydrometeorological level changes in the lake.

Estimation of evaporation

According to this water budget model, it is important to determine the evaporation since it is assumed the only outlet of the lake. However, there is no system to measure evaporation on the lake surface. Instead, the evaporation pans employed by meteorology stations around the lake are used to detect evaporation. Batur et al. (2009a) calculated

the annual total evaporation values measured by the evaporation pans at the Lake Van meteorology station, as well as the evaporation values using the Penman method. Drawing from the results of their study, Batur et al. (2009a) suggested that the ceiling coefficient should be 1.0 for Lake Van, instead of 0.70 suggested by the World Meteorological Organization (WMO) for average climatic conditions. Therefore, in this study, the data obtained from the station were used for the evaporation of Lake Van. On the other hand, evaporation pans can measure evaporation in warm seasons and cannot measure evaporation when the temperature drops below zero. However, even in the winter months when the temperature is low, evaporation from the lake surface continues. In addition, the change in heat storage affects the amount of evaporation in lakes of high depth such as Lake Van. Therefore, it may be expected that the evaporation values measured from Lake Van are incomplete and inaccurate.

In the case of unlimited water, all the losses that may occur by evaporation from the surface are called potential evaporation (PE). The two most used methods for potential evaporation calculation are Penman (1948) and Thornthwaite (1948) methods. Although it gives better results, the Penman (1948) method is very difficult to apply in many cases because it requires very comprehensive data such as net radiation, average air temperature and wind speed, saturated vapor pressure, and real vapor pressure. The monthly potential evapotranspiration can be estimated by the method originally presented by Thornthwaite (1948) which requires limited data (Doğdu 2011). It should be noted that although the Thornthwaite method is a method for measuring potential evapotranspiration (PEt), this temperature-dependent method was assumed using for the average monthly

evaporation estimate in the lake in this study. The method requires some simple input parameters such as the latitude of the study area (LAT, decimal degrees) and the average monthly temperature (T , degrees Celsius). Beginning with temperature and latitude, the potential evapotranspiration (millimeters) is calculated as follows (Davie 2008):

$$PEt = 16k \left(\frac{10T}{I} \right)^a \tag{2}$$

$$I = \sum_{j=1}^{12} i \tag{3}$$

$$i = \left(\frac{T}{5} \right)^{1.514} \tag{4}$$

, where T stands for average temperature ($^{\circ}\text{C}$) for a month, while I is annual heat index, and i is monthly heat index. Term k is a correction factor to account for unequal day length between months, which can be taken from the tables (De Marsily 1986) based on the latitude of the study site. Term a can be calibrated by the following cubic function of I (Davie 2008):

$$a = 6.75 \times 10^{-7}I^3 - 7.71 \times 10^{-5}I^2 + 1.79 \times 10^{-2}I + 0.49239 \tag{5}$$

The data provided by NASA and DLR (German Aerospace Center) obtained through the Advance Very High Resolution Radiometer (AVHRR) satellite were used for the water surface temperature values of Lake Van (Kavak and Karaoglan 2012). In addition, monthly average temperature

data around the lake were taken into account. Figure 4a shows the monthly water surface temperature averages of several years obtained from different sources, and the changes in the meteorological observation temperature average values throughout the year. A nonlinear relationship between the water surface temperature and the evaporation values obtained by the Thornthwaite method was determined with $R^2 = 96\%$ (Fig. 4b). Batur et al. (2009b) found the average annual evaporation amount as 1135 mm using the measuring pan and 1177 mm by the Penman method between 1956 and 2007. In the study, the average annual evaporation was measured as 1004 mm when the pan coefficient was considered as 1.0 and 704 mm when the pan coefficient was 0.70. When Thornthwaite method was used, it was measured as 746 mm. According to Thornthwaite’s method based on temperature change, the 10-year total evaporation change was calculated as ± 41 mm, and the standard deviation was 22 mm, which is not large enough to explain the annual level changes.

Water budget analysis

Figure 5 presents the actual (observed) water levels in Lake Van and the annual average water level changes in the water budget calculated in relation to the measured evaporation, precipitation and flow values. According to these data, the changes of lake level cannot be explained by the water budget, except for the changes occurred between 2017 and 2019. However, one-to-one matching of the water level changes may not be expected due to the

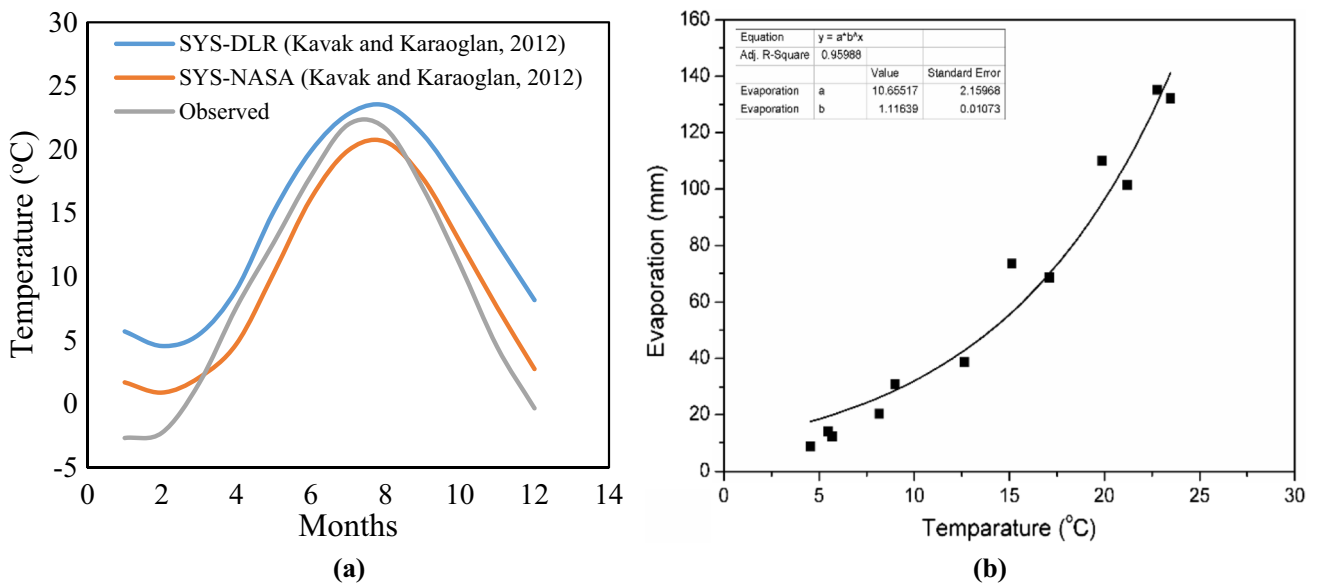


Fig.4 Temperature and evaporation changes in Lake Van: **a** Distribution of average monthly temperatures in a year, **b** The relationship between monthly average temperature and evaporation

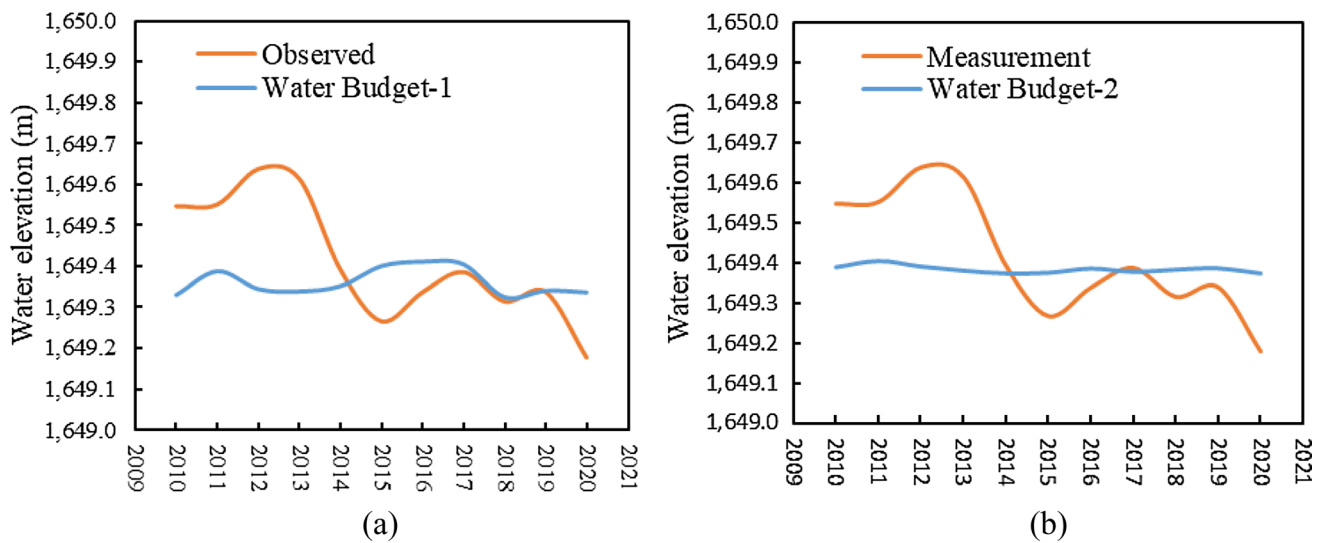


Fig. 5 Average level changes of Lake Van in long-term period: **a** According to the measured evaporation, **b** According to the average evaporation by Thornthwaite method

assumptions in the water budget method and the errors and deficiencies in the measurement values. Instead, comparison of standardized model and observation data will be more useful in revealing the relationship between the events.

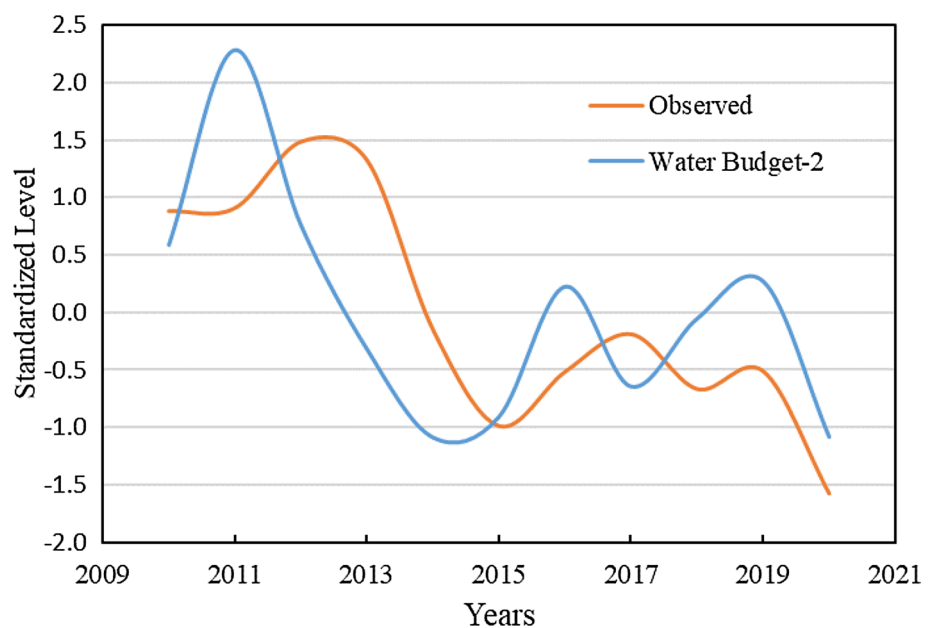
The following relation was used to standardize the data:

$$Z = \frac{X - \mu}{\sigma} \tag{6}$$

Here, Z is the standardized value, X is the raw data value, and μ and σ are the mean and standard deviation

of the data, respectively. The annual standardized level changes obtained according to Thornthwaite evaporation values compared the standardized observed values in the graph in Fig. 6. Since the annual average evaporation values are taken as constant in this graph, the effect of evaporation on the change is neglected as in Fig. 5b. Thus, it is aimed to reveal the effect of precipitation and flow on the level changes. Figure 6 also shows that the level change relates to the water budget of the lake, but the effect of hydrologic variables on the lake level change happened with a delay of about 1 year. Batur et al. (2009b) state that there is a 1- to 2-year delayed relationship between

Fig. 6 Comparison of standardized annual average Lake Van levels



the water level of Lake Van and the flows entering the lake, and that the lake level can react to climatic events in 2 months to 3 years in lakes with similar problems around the world. In this study, this delay is to be less since precipitation falling on the lake surface and stream flows are taken into account.

When the monthly changes are examined, it is understood that the delay may be seasonal (3–4 months). Therefore, the standardized monthly water budget values have been moved forward by 4 months to reveal the monthly variation relationship (Fig. 7). In Figs. 7a–d, 1-, 3-, 6- and 12-month moving averages of standardized values were plotted in order to monitor the monthly, seasonal, semi-annual and annual changes, respectively. Simple moving average (SMA) values were calculated according to the following formula:

$$\text{SMA}(n) = \frac{P_m + P_{m-1} + \dots + P_{m-(n-1)}}{n} \quad (7)$$

Here, n is the number of data included in the average, and m is the order indices of the values.

When the observed and monthly water budget level changes are examined (Fig. 7a), the general trends are in agreement with each other, but particularly the monthly changes in the water budget approach fluctuate more and differ from the observed values. When considered the three-month moving averages (SMA (3)) in Fig. 7b, the seasonal changes are related to the water level. It should be noted that this overlap occurs with a delay of 4 months. Furthermore, the graph in Fig. 7c reveals that the changes in the lake water level also match seasonally (dry and wet periods) in semi-annual periods. All these graphs in Fig. 7 show that the water level of Lake Van changes seasonally with the water budget in annual periods. In other words, seasonal changes in Lake Van level can be explained by hydrometeorological variables. However, the chart in Fig. 7d comparing the 12-month moving averages (similar to Fig. 5) shows that the annual changes cannot be related to the water budget alone. It is seen that the lake water levels were above normal levels, especially between the years 2010 and 2014, while the lake level decreased between 2013 and 2015 and approached the normal level after 2015. According to these results, the significant increases especially between the years 2010–2013 show that they cannot be explained by hydrometeorological events (water budget). Considering, the major natural events during these years, it can be suggested that they may be related to the 7.2 magnitude earthquake that occurred in Van on October 23, 2011. Activities such as crustal movements, water and gas outflows that may occur on the lake floor before and after such a big earthquake may be the main reasons for these elevations in the lake level. Utkucu (2006) showed that alternating rises and decreases

of the both change of the annual mean of the lake level and the annual lake level oscillations correlate quite good with the occurrence of $M \geq 4.0$ earthquakes across and near the vicinity of Lake Van basin. Utkucu and Arman (2015) also suggested that the episodic water level fluctuations of Lake Van in the order of 1-m trigger $M \geq 5.0$ earthquakes along the lake basin-related faults.

Long short-term memory (LSTM)

The long short term memory (LSTM) network is a different type of recurrent neural network. In practice, simple recurrent neural networks (RNNs) are limited in their capacity to learn longer-term dependencies. Therefore, Hochreiter and Schmidhuber (1997) developed a special structure of RNN called LSTM networks. When requiring the network to learn long-term relationships, the RNN model cannot use information from old periods. LSTM networks, on the other hand, implement additional gates to control what information in the hidden cell is transmitted as output and move on to the next hidden state. These additional gates thus allow the network to learn long-term relationships in the data more effectively. Compared to simple RNNs, less sensitivity to time gap makes LSTM networks a better choice for analyzing sequential data. Unlike conventional RNNs, there are typically additional units such as a memory cell, input gate, output gate and forget gate in a LSTM block architecture. The diagram in Fig. 8 illustrates how the gates forget, update and output the cell and hidden states. The LSTM architecture uses the following formulae to describe their components at time step (t) which control the cell state and hidden layer state (Hochreiter and Schmidhuber 1997; MathWorks 2021):

$$i_t = \sigma_g(W_f x_t + R_i h_{t-1} + b_i) \quad (8)$$

$$f_t = \sigma_g(W_f x_t + R_f h_{t-1} + b_f) \quad (9)$$

$$g_t = \sigma_c(W_g x_t + R_g h_{t-1} + b_g) \quad (10)$$

$$o_t = \sigma_g(W_o x_t + R_o h_{t-1} + b_o) \quad (11)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (12)$$

$$h_t = o_t \odot \sigma_c(c_t) \quad (13)$$

In Eqs. (8–13), h_t and c_t denote the output (also known as the hidden state) and the cell state at time step t , respectively. W is the input weights, R is the recurrent weights, b is the bias, while indices of i , f , g , and o denote the input gate,

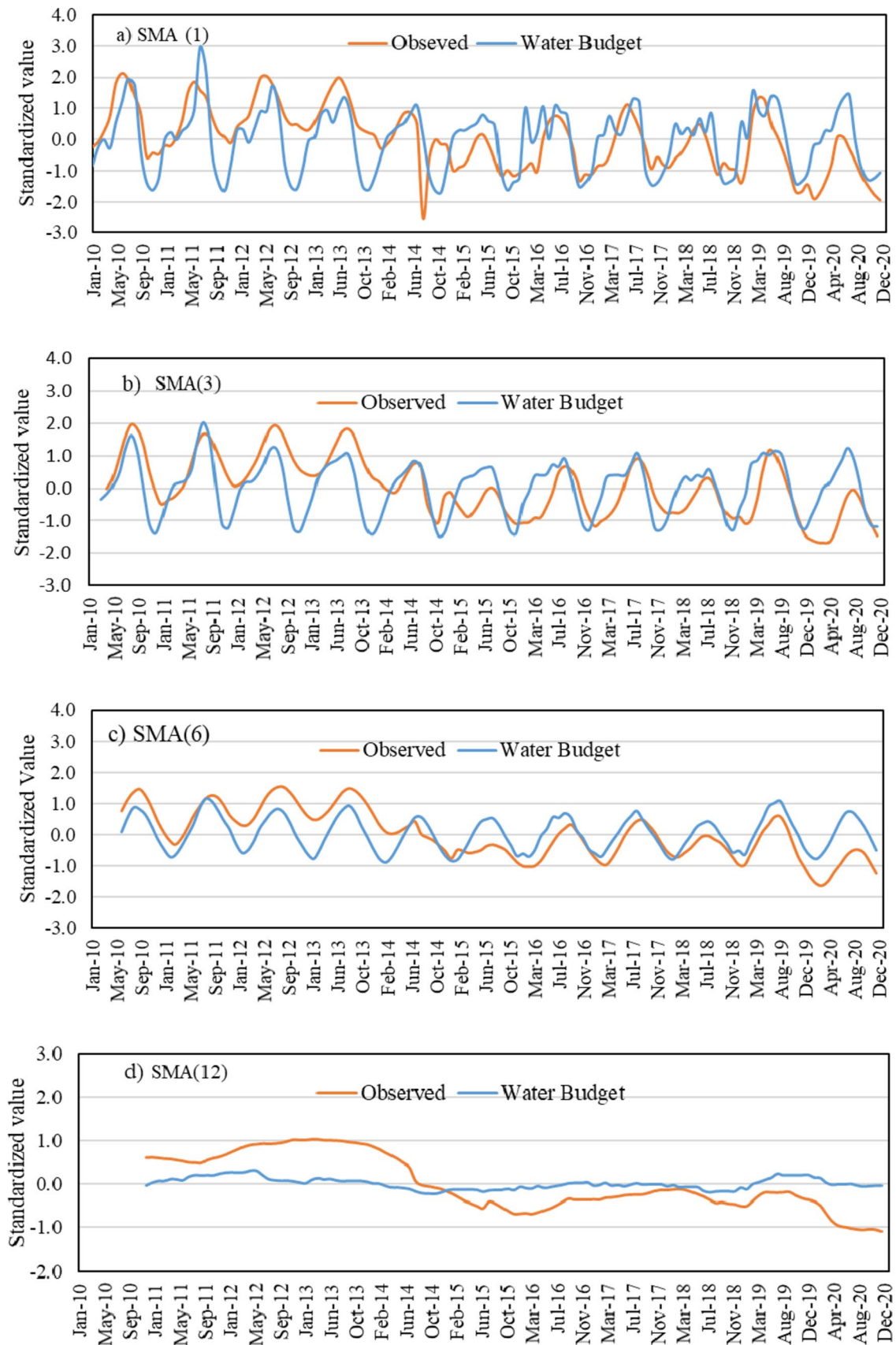
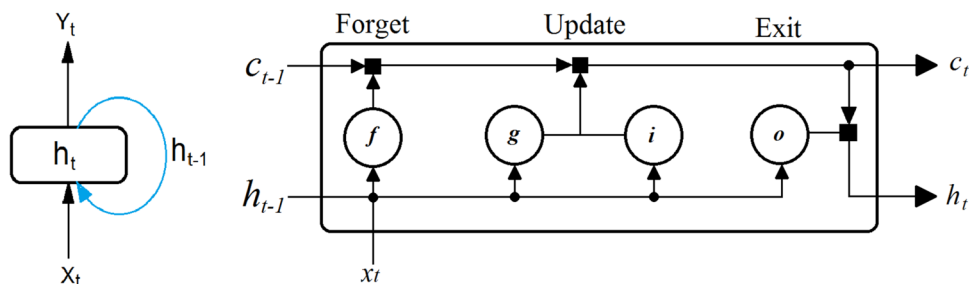


Fig. 7 Comparison of monthly standardized level changes in Lake Van

Fig. 8 Architecture of recurrent neural network (RNN) and LSTM block (MathWorks 2021)



forget gate, cell candidate and output gate, respectively. σ_c represents the state activation function which uses the hyperbolic tangent function (*tanh*) by default, and σ_g represents the gate activation function which uses the sigmoid function given by $\sigma(x) = (1 + e^{-x})^{-1}$.

The standard deviation of the prediction errors used to test the accuracy of the model, i.e., root-mean-square-error (RSME) is calculated as follows:

$$RSME = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2} \tag{14}$$

In this part of the study, water levels of Lake Van were simulated with LSTM network model using MATLAB R2019b software. The monthly average water elevations of Lake Van during period of 2000–2020, which were obtained from the lake level measurement station were considered as time series (Fig. 9).

70% of the 252 data were used to train the LSTM network, and the remaining 30% of data were used for testing. Training data and test data at prediction time were standardized for better program performance. The LSTM network model was built with 200 hidden layers. Training was performed for 250 epochs, and the gradient threshold value was taken as 1 to solve the gradient disappearance problem in LSTM networks. The initial learning rate was set as 0.005, and the learning rate was reduced by multiplying a factor of 0.2 after 125 epochs.

Having trained the model, the values of the forward time steps were estimated, with the value of each time step being the input of the next time step. RMSE error value was found as 0.257 (Fig. 10). Then, the predicted values were compared with the test values.

The error value was reduced by updating the predicted values in the LSTM network model with their actual values. Thus, the new RMSE error value was reduced to 0.145 (Fig. 11). As

Fig. 9 Monthly average water levels of Van Lake (2000–2020)

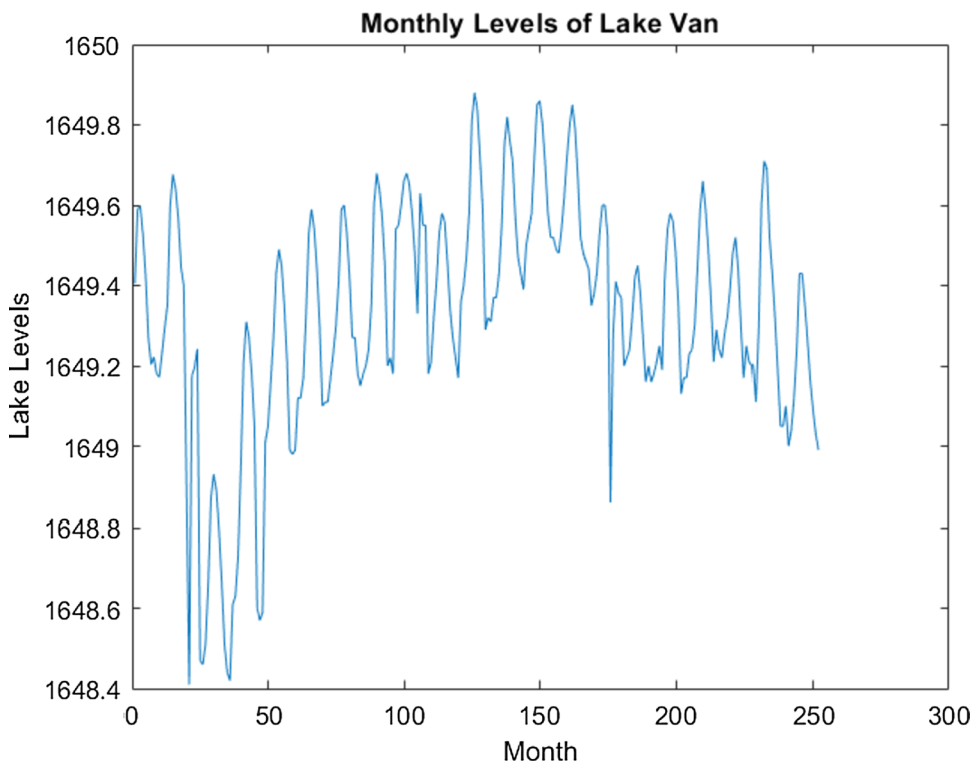


Fig. 10 Measured and predicted water levels with RMSE errors distribution

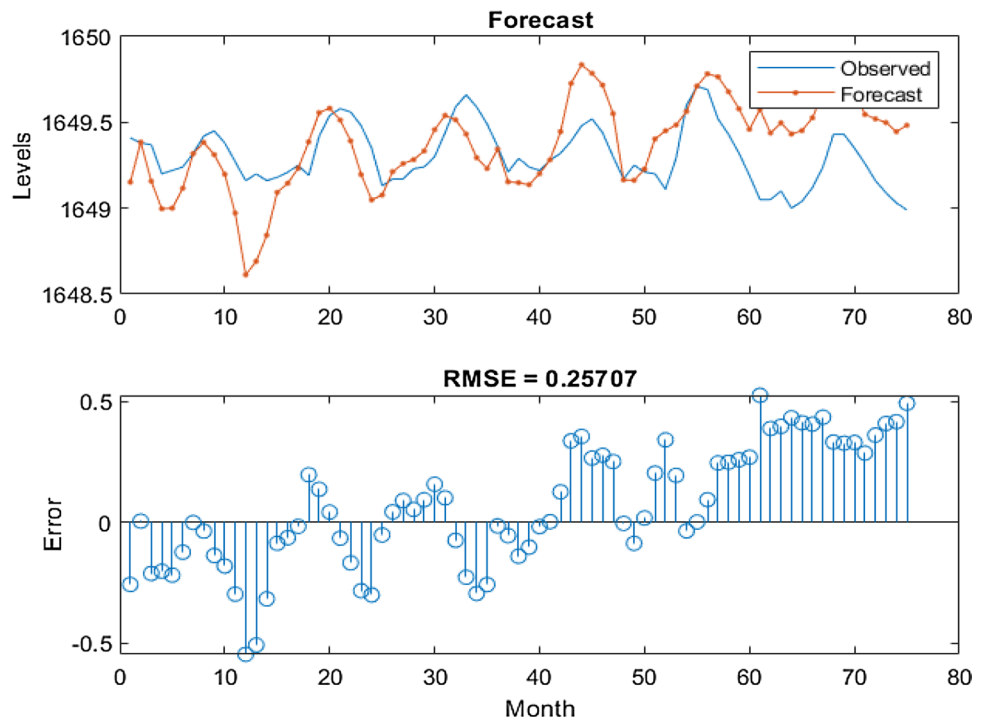
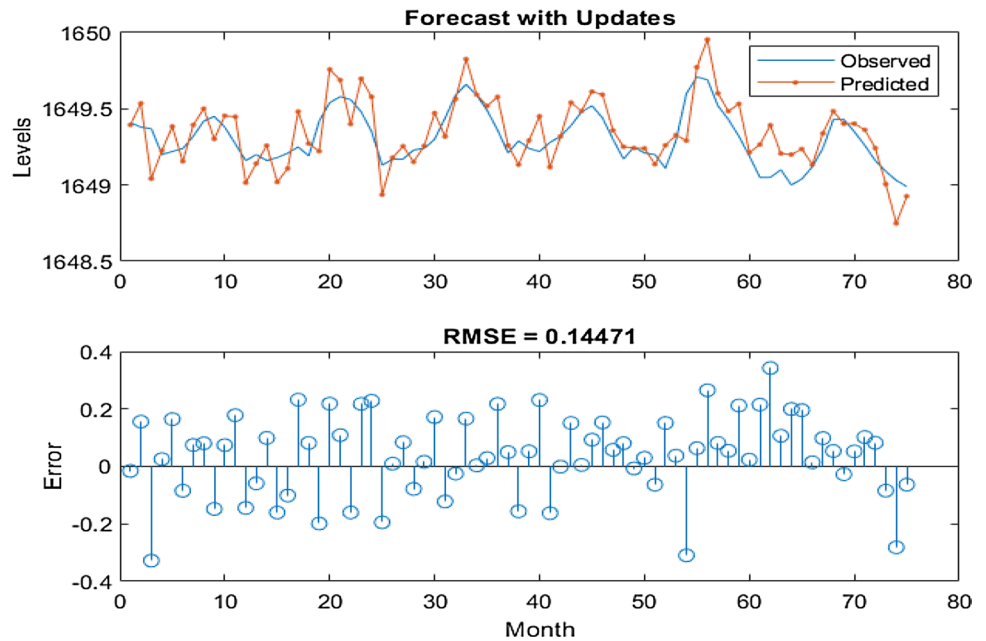


Fig. 11 Model retrained with real values and RMSE error distribution



a result of this update, a better estimation has been obtained. Figure 11 shows the time series estimated as the continuation of the level changes according to the last model trained are given in Fig. 12.

Neural network time series

Hydrological time series, such as water level changes, can be modeled using a nonlinear autoregressive (NAR) neural network. Predicting sequential time series requires a multi-step process. Closed-loop networks can perform multistep prediction. Closed-loop networks can continue forecasting using internal feedback in the absence of external feedback. NAR predicts future values of a time series only from past values (MathWorks 2023). The neural network diagram of the created in MATLAB 2018b for 72 months delayed is given in Fig. 13. Series of monthly water levels in the lake were modeled using the NAR neural network with ten hidden layer with delays of 2, 6, 12, 24, 36 and 72 months. The analysis used 70% data for training, 15% for validation and 15% for testing. The performance of

each NAR neural network is shown in Table 3. NAR(72A) and NAR(72B) of 72 months delay indicate the use of *Levenberg–Marquardt* and *Bayesian regularization* schemes, respectively, as training algorithms. According to Table 1, although the best test results were obtained for NAR(36), the results of NAR(72) with the longest delay were taken into account for comparison with LSTM. While the test correlations for NAR(72A) and NAR(72B) were 74% and 88% respectively, the RSME values obtained were 0.204 and 0.099. The performance indicators and prediction time series of the NAR(2A) and NAR(2B) models are shown in Figs. 14 and 15, respectively. It should be noted that the learning performance (correlation) of both models is above 90%. According to these results, NAR(2B) performed slightly better than NAR(2A) and LSTM in the long term.

Fig. 12 Extended forward forecasted lake level changes with observed values

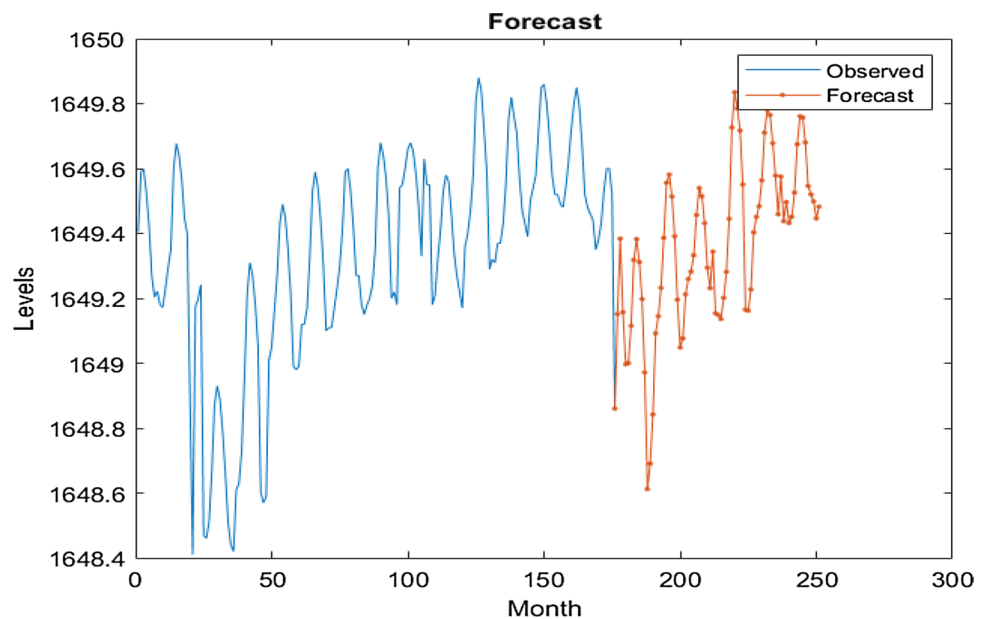


Fig. 13 Architecture of NAR neural network (MathWorks 2023)

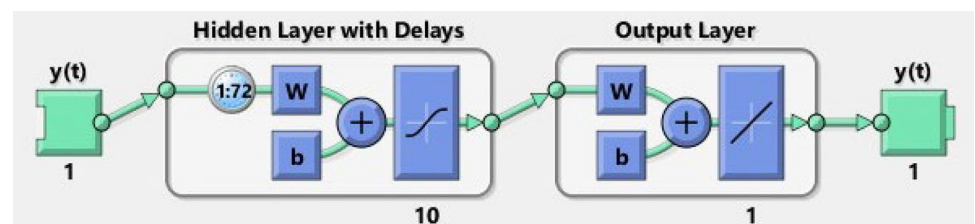


Table 3 Testing performance of NAR models with various delayed months

Model	NAR(2)	NAR(6)	NAR(12)	NAR(24)	NAR(36)	NAR(72A)	NAR(72B)
RSME	0.0952	0.1167	0.1518	0.1718	0.0995	0.2038	0.0991
R	0.941	0.895	0.852	0.824	0.946	0.737	0.881

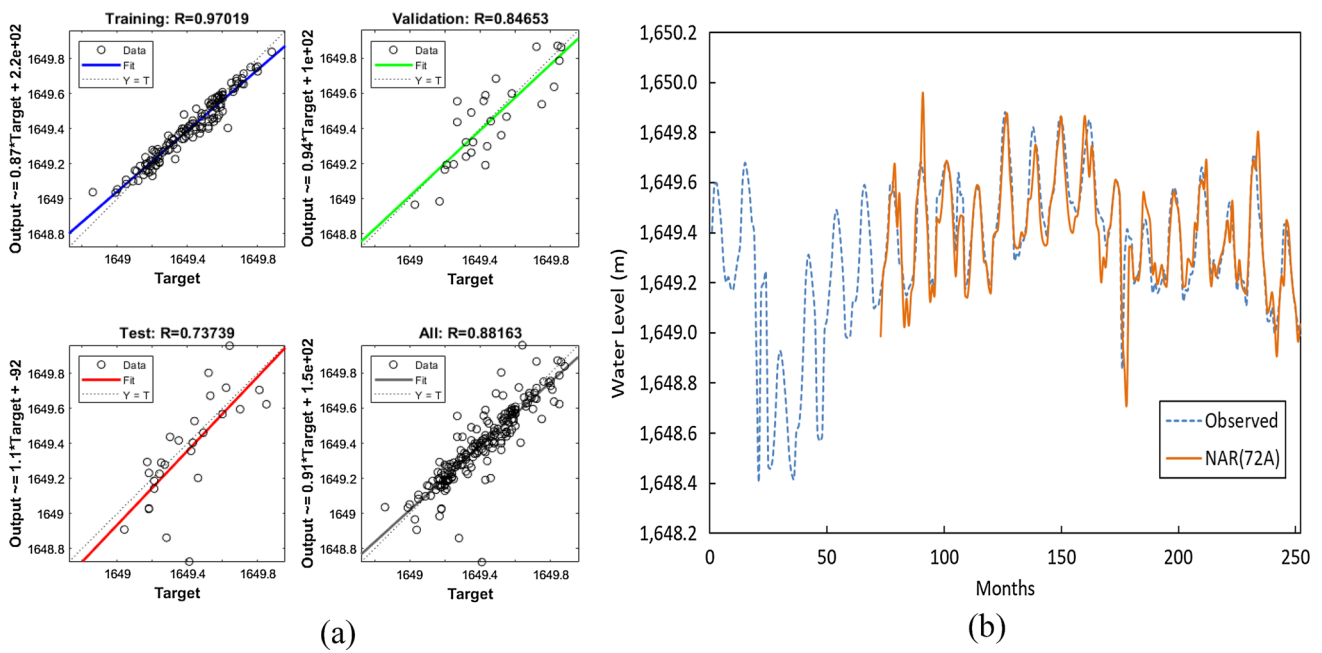


Fig. 14 Model performance of NAR(72A) neural network; **a** Regression indicators, **b** Predicted and observed time series

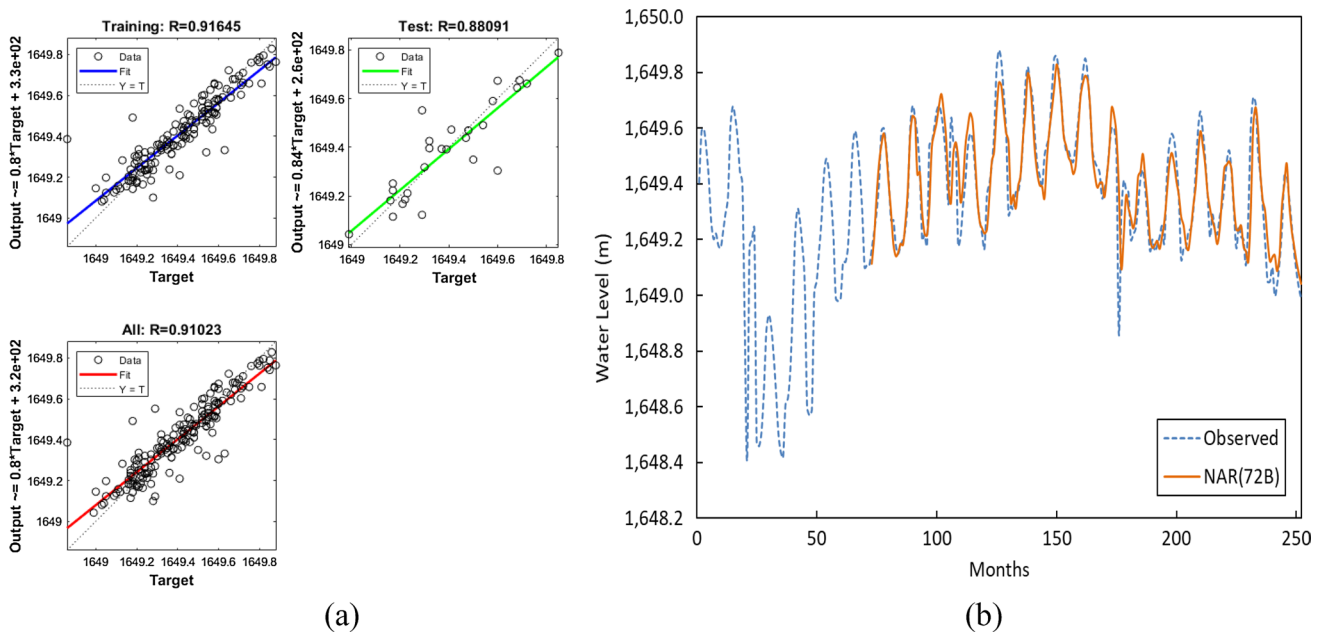


Fig. 15 Model performance of NAR(72B) neural network; **a** Regression indicators, **b** Predicted and observed time series

Conclusion

The water level fluctuations in Lake Van, the world’s largest soda lake, have always attracted the attention of researchers in the recent past. Although these level changes are generally considered hydrometeorological,

they have been attributed to different reasons in each decades period, namely various ideas have been put forward so far. In this study, the changes in the water levels of Lake Van in recent years were analyzed using water budget, LSTM and NAR neural network methods. According to the results of water budget analysis, it can be said that the main reason for the level fluctuations in Van Lake in

recent years is seasonal hydrometeorological variables. The effect of hydrometeorological variables on lake level fluctuations overlaps with a delay of a few months, i.e., seasonally. On the other hand, the water budget approach cannot explain the annual water level changes in the period of 2010–2013, which points to other possible reasons. The 7.2 magnitude earthquake that occurred in Van, Turkey, in 2011, may be the main reason for the significant level changes in the period. However, further research is therefore needed on this issue in order to reach clearer scientific results. Another reason for the uncertainties in level changes may be related to the inability to accurately measure or estimate the amount of evaporation, which is considered the only output of the lake in the water budget approach. In addition, time series analyses conducted with LSTM and NAR networks predicted the water level successfully. The RSME error in predicting lake levels was found to be 0.145 for LSTM and 0.099 for NAR models. It can be stated in general that such the neural network methods were found to be particularly successful in modeling periodically time series.

Acknowledgements This study is derived from the Master's Thesis of Gelberi (2021). In addition, we would like to thank the Turkish General Directorate of State Hydraulic Works (DSI), and the Turkish State Meteorological Service for their support.

Funding The authors received no specific funding for this work.

Data Availability All data are available from the corresponding author.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standard The study complied with all ethical standards and guidelines.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aksoy H, Unal NE, Eris E, Yuce MI (2013) Stochastic modeling of Lake Van water level time series with jumps and multiple trends. *Hydrol Earth Syst Sci* 17(6):2297–2303
- Altunkaynak A (2007) Forecasting surface water level fluctuations of Lake Van by artificial neural networks. *Water Resour Manag* 21:399–408
- Altunkaynak A (2019) Predicting water level fluctuations in Lake Van using hybrid season-neuro approach. *J Hydrol Eng* 24(8):04019021
- Aydın FA, Doğu AF (2018) Level changes in lakes and their causes: the example of Lake Van. *Yuzuncu Yıl University Soc. Sci Inst J* 41:183–208 (in Turkish)
- Aydın MC, Gelberi G, Budak U (2021) Estimation of the van lake levels using long short term memory (LSTM) network, Middle East International Conference On Contemporary, March 27–28, 2021, Ankara, Turkey
- Aydın H, Karakaş H (2016) Estimation of evaporation for Lake Van. *Environ Earth Sci* 75:1275
- Aytek A, Kisi O, Guven A (2014) A genetic programming technique for lake level modeling. *Hydrol Res* 45:529–539
- Baek SS, Pyo J, Chun JA (2020) Prediction of water level and water quality using a CNN-LSTM combined deep learning approach. *Water* 12:3399
- Batur E (1996) Water budget and basin climate of Lake Van. Istanbul Technical University. Institute of Science and Technology, Istanbul (in Turkish)
- Batur E, Kadioğlu M, İbrahim A, Özkaya M, Saban M, Elkatmış MN, İlikçi A (2009a) The water budget of Lake Van and the relationship of the lake water level with areal precipitation and flows. *Water Resour* 2:12–26 (in Turkish)
- Batur E, Kadioğlu M, Özkaya M, Saban M, İbrahim A (2009b) Van lake water level modeling. *Water Resour* 2:27–40 (in Turkish)
- Çukur D, Krastel S, Schmincke HU, Sumita M, Tomonaga Y, Namık Çağatay M (2014) Water level changes in Lake Van, Turkey, during the past ca. 60 0 ka: climatic, volcanic and tectonic controls. *J Paleolimnol* 52(3):201–214
- Davie T (2008) Fundamentals of hydrology, second edition, routledge fundamentals of physical geography. Taylor & Francis e-Library, New York
- De Marsily G (1986) Quantitative hydrogeology. Paris School of Mines, Fontainebleau
- Degens ET, Kurtman F (1978) The geology of Lake Van. MTA Press, Publication No. 169. Ankara
- Doğan E, Kocamaz UE, Utkucu M, Yıldırım E (2016) Modeling daily water level fluctuations of Lake Van (Eastern Turkey) using Artificial Neural Networks. *Fundam Appl Limnol* 187:177–189
- Doğdu MŞ (2011) Computer program for calculation of water budget components used in hydrological studies. *Techn Bull General Directorate State Hydraulic Works* 112:19–32 (in Turkish)
- DSI (2015) Van closed basin master plan report. Ministry of forestry and water affairs, general directorate of state hydraulic works, Department of Survey, Planning and Allocation Prepared by Eser Proje ve Mühendislik A. Ş., Ankara (in Turkish)
- DSI (2018) 2015 Akım Gözlem Yıllığı, Cilt-1, Orman Ve Su İşleri Bakanlığı Devlet Su İşleri Genel Müdürlüğü, DSI Teknoloji Dairesi Başkanlığı Basım ve Foto-Film Şube Müdürlüğü, Ankara 2018.
- DSI (2021) General Directorate of State Hydraulic Works, 17th Regional Directorate, Flow and Lake Level Observation Station Data, Van
- Düzen H (2011) Hydrogeological approach to Van Lake Water level changes. *Yüzcüncü Yıl University. Institute of Natural and Applied Sciences, Jeol. Engineering Department, Master Thesis, Van* (in Turkish)
- Eriñç S (1953) Eastern Anatolian Geography. Istanbul University, Istanbul (in Turkish)
- Fijani E, Khosravi K, Barzegar R, Quilty J, Adamowski J, Kheyrollah-Pour H (2021) Advanced hybrid machine learning algorithms for

- multistep lake water level forecasting. DOI: <https://doi.org/10.21203/rs.3.rs-912383/v1>
- Gelberi G (2021) Analysis of Van Lake level changes in recent years. Bitlis Eren University, Graduate Institute. Bitlis, Turkey (**in Turkish**)
- Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780
- Hrnjica B, Bonacci O (2019) Lake level prediction using long short-term memory recurrent neural networks. In: 12th International scientific conference on production engineering, pp 1–6
- Jalili S, Hamidi SA, Morid S, Namdar Ghanbari R (2016) Comparative analysis of Lake Urmia and Lake Van water level time series. *Arab J Geosci* 9(14):1–10
- Kadioğlu M, Şen Z, Batur E (1997) The greatest soda-water lake in the world and how it is influenced by climatic change. In *Annales Geophysica* 15(11):1489–1497
- Kadioğlu M, Şen Z, Batur E (1999) Cumulative departures model for lake-water fluctuations. *J Hydrol Eng* 4(3):245–250
- Kavak MT, Karadoğan S (2012) Investigation of water surface temperature change in van lake with AVHRR satellite data. IV. Remote sensing and geographic information systems Symposium (UZAL-CBS 2012), 16–19 October 2012, Zonguldak, Turkey (**in Turkish**).
- Khatibi R, Ghorbani MA, Naghshara S, Aydın H, Karimi V (2020) A framework for 'Inclusive Multiple Modelling' with critical views on modelling practices—Applications to modelling water levels of Caspian Sea and Lakes Urmia and Van. *J Hydrol* 587:124923
- Kisi O, Shiri J, Nikoofar B (2012) Forecasting daily lake levels using artificial intelligence approaches. *Comput Geosci* 41:169–180
- Landmann G, Reimer A, Kempe S (1996) Climatically induced Lake level changes at Lake Van, Turkey, during the Pleistocene/Holocene transition. *Global Biogeochem Cycles* 10(4):797–808
- Liang C, Li H, Lei M, Du Q (2018) Dongting lake water level forecast and its relationship with the three gorges dam based on a long short-term memory network. *Water* 10:1389
- MathWorks (2021) Long Short-Term Memory (LSTM), Help Center, MATLAB Documentation, USA, The MathWorks, Inc
- MathWorks (2023) Nonlinear Autoregressive Neural Network, Help Center, MATLAB Documentation, USA, The MathWorks, Inc
- Ozer C, Aydın MC, Budak U, Ulu AE (2022) Investigation of pollution carried to van lake by streams within bitlis province boundaries, BEU Scientific Research Project, No: BEBAP 2022.02.
- Penman HL (1948) Natural evaporation from open water, bare soil and grass. *Proc R Soc London Ser A Math Phys Sci* 193:120–145
- Piasecki A, Jurasz J, Adamowski JF (2018) Forecasting surface water-level fluctuations of a small glacial lake in Poland using a wavelet-based artificial intelligence method. *Acta Geophys* 66:1093–1107
- Şener E, Morova N (2011) Determination of water level changing of burdur lake with fuzzy logic and linear regression analysis. *J Nat Appl Sci* 15
- Seo Y, Kim S, Kisi O, Singh VP (2015) Daily water level forecasting using wavelet decomposition and artificial intelligence techniques. *J Hydrol* 520:224–243
- Shafaei M, Kisi O (2016) Lake level forecasting using wavelet-SVR, wavelet-ANFIS and wavelet-ARMA conjunction models. *Water Resour Manag* 30:79–97
- SHODB (1983) Bathymetric map of Lake Van. Turkish Naval Forces, Department of Navigation, Hydrography and Oceanography, Istanbul
- Teltik İ (2008) Stochastic modeling of van lake level. Graduate School of Natural and Applied Sciences. Istanbul Technical University, Turkey (**in Turkish**)
- Thornthwaite CW (1948) An approach toward a rational classification of climate. *Geogr Rev* 38:55–94
- Tomonaga Y, Brennwald MS, Livingstone DM et al (2017) Porewater salinity reveals past lake-level changes in Lake Van, the Earth's largest soda lake. *Sci Rep* 7:313
- Üner S, Yeşilova Ç, Yakupoğlu T (2012) The traces of earthquake (Seismites): Examples from Lake Van Deposits (Turkey). In: Damico S (ed) *Earthquake Research and Analysis Seismology. G and Earthquake Geology*. IntechOpen, Croatia, pp 21–31
- Utkucu M (2006) Implications for the water level change triggered moderate ($M \geq 4.0$) earthquakes in Lake Van basin. *Eastern Turkey J Seismol* 10:105–117
- Utkucu M, Arman H (2015) Water level variations and earthquakes in the Lake Van Area, Eastern Turkey. In: International conference on engineering geophysics, Al Ain, United Arab Emirates, 15–18 November 2015. Society of Exploration Geophysicists, pp 132–135
- Yaseen ZM, Naghshara S, Salih SQ, Kim S, Malik A, Ghorbani MA (2020) Lake water level modeling using newly developed hybrid data intelligence model. *Theoret Appl Climatol* 141(3):1285–1300
- Yıldız MZ, Deniz O (2005) The impacts of the level changes in closed basin lakes on the coastal settlements: the Lake Van example. *J Soc Sci* 15(1):15–31
- Zhu S, Lu H, Ptak M, Dai J, Ji Q (2020) Lake water-level fluctuation forecasting using machine learning models: a systematic review. *Environ Sci Pollut Res* 27:44807–44819

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.