6th Intercontinental Geoinformation Days

igd.mersin.edu.tr

Lake level forecasting with radial based neural networks

Mehmet Fehmi Yıldız [1], Vahdettin Demir*[1]

[1]KTO Karatay University, Civil Engineering Department, Konya, Türkiye

Keywords
Michigan-Huron
Radial Based Neural Networks
Lake levels
Forecast

Abstract

Many water resources across the world have increasing and decreasing water levels. The change of water level in lakes, which is one of the water resources, is associated with climate change and the effects of climate change can be seen at lake levels the fastest. Lake Michigan-Huron studied in this study is an 8 km wide body of water formed by the confluence of Lake Michigan and Huron. Lake Michigan-Huron is the largest freshwater lake in the world. The aim of this study is to estimate the water level of Lake Michigan-Huron in the USA. For this purpose, radial-based artificial neural networks were used. In the forecast model, lake levels in the past months and periodicity number were used as input data. The lake water level (m) data used has a record length of 104 years (1918-2021). All data is divided into 4 parts (M1, M2, M3 and M4). 75% of all data was used for the training phase (M1+M2+M3) and 25% for the testing phase (M4). The test sections were changed from M1 to M4 so that the training and testing rates remained constant. Mean absolute error (MAE), root-mean-square error (RMSE) and coefficient of determination (R^2) were used as evaluation criteria. As a result, it is seen that the models make very good predictions in all data sets and in the training-test phases. However, according to the test results, the data set that gives the most successful results is the M1 package and the input set using data that has been lag time for 7 months.

1. Introduction

Humans have been interested in water since its inception, trying to study water movements, recognize features, identify detection hazards, and make the most of the water outside. The branch of science that manages the distribution and properties of water on Earth is called hydrology. The science of hydrology, which provides its relationship with the internal environment and efforts to control its environment, began to gain more importance. As a result of the hydrological operation, the basic structures that maintain their water consumption and attitudes can be identified (Koca, 2014).

Due to the changes in water bodies, long research has been started. Water is an indispensable resource for human life. Therefore, research on the quality and quantity of existing water resources has intensified, and the storage facility of closed water basins such as lakes has gained importance (Teltik et al., 2008).

The water level of many lakes in the world is observed to rise and fall due to various reasons. In the studies, it is thought that the reasons for the change in the lake level are meteorological and hydrological features (evaporation, precipitation, flow, etc.), tectonic movements, changes in the ozone layer and climate change (Teltik et al. 2008). In addition, the use of water resources to provide more water than normal in order to meet the water needs of agricultural activities and cities also causes the capacity of water reserves such as lakes to decrease (Albek et al., 2017).

Some studies on the use of RBANN in the literature, In the study of Desmukh and Tanty (2015), a comprehensive review was made on the artificial neural network (ANN) used in the field of hydrology-related problems. They stated that it can be well exemplified by artificial intelligence with precipitation-flow modeling, stream flow modeling, water quality modeling and its applications in groundwater (Desmukh and Tanty 2015).

In 2018, Arslan et al. A study was carried out to examine the seasonal variation of Adana Seyhan Dam Lake area. In their study, they achieved highly accurate results in the classification of water structures with the artificial neural network method (Arslan et al. 2018).

In a study conducted by Dikbaş and Fırat in 2005, a comparison of POM and ANN models was made in three-dimensional hydrodynamic modeling in lakes. When the

* **Corresponding Author**

(mehmetfehmi.yildiz@ogrenci.karatay.edu.tr) ORCID ID 0000-0001-6550-8401
*(vahdettin.demir@karatay.edu.tr) ORCID ID 0000-0002-6590-5658

Cite this study

Yıldız, M. F., & Demir, V. (2023). Lake level forecasting with radial based neural networks. Intercontinental Geoinformation Days (IGD), 6, 211-215, Baku, Azerbaijan

results obtained with both models were compared, it was seen that the methods had advantages and disadvantages compared to each other. ANN requires previously obtained observation results and calculations, while POM does not. In addition, ANN achieves results in a much shorter time than POM. As a result, it can be recommended to use the artificial neural network method in certain sections and studies that require many detailed calculations (Dikbaş and Fırat, 2005).

A study was conducted by Aksoy et al. in 2020 on the estimation of the water level in Yalova Gökçe Dam using ANN. According to the data they obtained as a result of the analyzes, the dam water level for 2019 was 73.77, while the actual water level of the dam was measured as 72.13 meters. As a result, it is thought that the use of ANN algorithms will be beneficial in estimating the water level of Gökçe Dam (Aksoy et al. 2020).

In 2012, Okkan and Dalkılıç conducted a study on the modeling of monthly flows of the Kemer Dam using radial-based neural networks. When they evaluated the results of their study in terms of minimum and maximum currents, the results of the RBANN model were successful for most months. In addition, it is thought that the problems encountered in other artificial neural network models can be overcome with RBANN (Dalkılıç and Okkan, 2012).

The purpose of this study was to analyze the water level of Lake Michigan-Huron in the United States and to determine changes in the lake's water level. For this purpose, monthly lake water levels in Lake Michigan-Huron between 1918-2021 were estimated with radial-based artificial neural networks and the predictions in various data sets (the training set is 3 parts and the test set is one part, and the test set is constantly changing.) were compared with the observed data.

2. Material and Method

2.1. Material

In the study, monthly water levels year from January 1918 to December 2021 were used. Data obtained from "https://www.lre.usace.army.mil/Missions/Great-Lakes-Information/Great-Lakes-Information.aspx#ICG_ETH_22302". Statistical information of the data used is given in Table 1.

The data are observed as monthly average (m) and there are no gaps in the data. In addition, station information in Excel format is available for all researchers free of charge.

Table 1. Statistical information for water levels (m)

Michigan-Huron	
Average	176.44
Standard error	0.0116
Median	176.45
Standard Deviation	0.410
Sample Variance	0.168
Kurtosis	-0.787
Skewness	0.101
Smallest	175.57
Largest	177.5
Number of Data	1248

2.2. Study Area

Lake Michigan is the third largest of the five great lakes in the northern United States and is connected to Lake Huron by the Strait of Mackinac (Demir, 2022). It is located 176 meters above sea level and its deepest point is 281 meters. Lake Huron is also located in North America and is the 4th largest lake in the world.

Lake Huron is connected to Lake Michigan by the Straits of Mackinac and to Lake Superior by a series of straits. Huron and Michigan are hydrologically a single lake because the flow of water through the straits keeps water levels in overall balance. Although the flow is generally eastward, water moves in both directions depending on local conditions. Combined, Lake Michigan-Huron is the world's largest freshwater lake by area (Michigan-Huron, 2023). The study area is given in Figure 1.

Figure 1. Lake Michigan-Huron (Demir and Yaseen, 2022)

The most important factor in choosing this study area is that when Michigan-Huron Lake is considered as a whole, it is the largest freshwater lake in the world in terms of surface area and the data are continuous.

2.3. Method

Artificial neural networks are a method based on the biological nervous system in humans. Artificial neural networks consist of elements called neurons, which are connected in parallel and have a non-linear structure. It is used in object recognition, system modeling, signal processing and solving complex engineering problems. Artificial neural networks realize the learning process with examples. In other words, it can be defined as the machine-transferred version of the human learning mechanism (D'Addona, 2014).

2.4. Radial based neural network

The radial-based neural network model can be considered as a combination of a data modeling technique for a high-dimensional space and a schema such as an ANN network. In the RBANN model, three layers are defined as input layer, hidden layer and output layer, but unlike the classical ANN structure, a nonlinear clustering analysis and radial based activation functions are used in the transition from the input layer to the hidden layer in the radial-based neural network model (Dalkılıç and Okkan, 2012).

The mathematical representation of radial basis neural networks is as follows.

$$\emptyset_j = exp\lfloor -\|x - c_j\| / \sigma_j^2 \rfloor \quad (1)$$

Here is x the input vector, c_j j. It is the center of the Gaussian function and σ_j is the standard deviation. Equation $\|x - c_j\|$ indicates the Euclidean distance between vectors x and c_j. j. the activation level of the intermediate node is equal to $\emptyset_j$.

Interlayer outputs;

$$y_{kj} = \prod \emptyset_j(x, c, \sigma) \quad (2)$$

k. the output of the node is given by Equation 2.

$$o_k = \sum_{j=1}^{J} w_{kj}\, y_{kj} \quad (3)$$

Here w_{kj} k. With the exit node j. is the weight between the middleware node (Kılıç et al., 2012). The basic structure of RBANN is given in Figure 2.

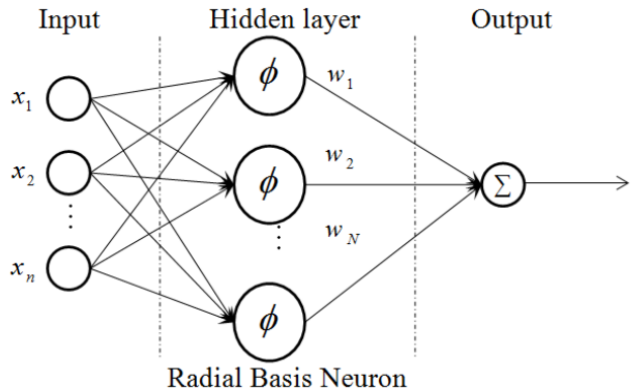

Figure 2. Radial-based neural network structure (Chen et al. 2019)

In Figure 2, the RBANN structure basically consists of three parts and the output data is obtained by multiplying the input data with the weights after reaching the hidden layer.

3. Results

In the modeling phase, the data were first shifted by lag time, and then the estimation results were obtained by separating them into training and test sets. The estimated data with the observed data were evaluated by considering the comparison criteria. Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and coefficient of determination (R^2) were used as comparison criteria. The formulas of the comparison criteria are given in Equation 4-6.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_p - Y_o)^2} \quad (4)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|Y_p - Y_o| \quad (5)$$

$$R^2 = \left(\frac{N*(\sum Y_o*Y_p) - (\sum Y_o)*(\sum Y_p)}{\sqrt{(N*\sum Y_o^2) - (\sum Y_o)^2 * (N*\sum Y_p^2) - (\sum Y_p)^2}}\right)^2 \quad (6)$$

N is the number of data used in the equations, Yp represents the estimated value in the model, Yo represents the observed value. Since Equation (4-5) has error results for our comparison criteria, the unit of error results in the model is m^3/s.

High error results show that the model is far from real data, that is, it gives unsuccessful results. If the error values are close to zero, it indicates that the results of the model are close to the real data.

The coefficient of determination R^2 can take a value between 0 and 1. It is interpreted that the closer the value is to 1, the higher the model's fitness and accuracy.

In the study, data sets with 12 inputs were created. These; 1 month lag (T-1), T-2, T-3, T-4 and T-5 with 1 year lag (Y-1), Y-2, Y-3, Y-4 and Y-5.

Month numbers representing the periodicity of the data were used recursively as the 12th input set. 75% of all data were used in the training phase and 25% in the testing phase. At this stage, four different combinations were tried, namely training (75%) and testing (25%). These are M1 (the part with the oldest data), M2, M3 and M4 (the part with the most recent data). Model results are given in Table 2 according to the training and testing phases.

In the table, the most successful result in the test phase was obtained in 7 inputs. The scatter plot of the best method is given in **Figure 3**.

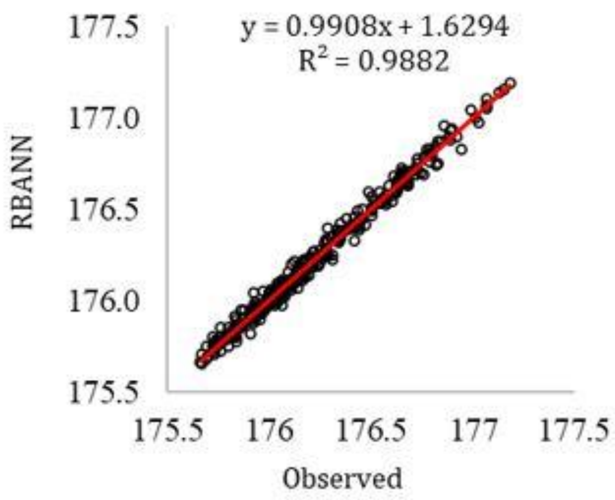

Figure 3. Scatter plot

When the values in the graph are examined, it is seen that the results obtained using the radial-based artificial neural network model are compatible with the water level data of the observed lake, and the graph equation approaches the y=x line, and the R^2 value is 0.9882. The variation of these estimates in the time series is shown in Figure 4.

Table 2. RBANN training and test results (Model: RBANN)

Part	Input	Training			Test		
		RMSE	MAE	R^2	RMSE	MAE	R^2
M1	1	0.069	0.055	0.97	0.066	0.054	0.964
	2	0.046	0.037	0.986	0.044	0.035	0.984
	3	0.044	0.034	0.987	0.042	0.033	0.986
	4	0.043	0.033	0.988	0.041	0.031	0.987
	5	0.042	0.033	0.989	0.039	0.03	0.987
	6	0.042	0.032	0.989	0.039	0.03	0.988
	7	0.041	0.032	0.989	0.038	0.03	0.988
	8	0.041	0.032	0.989	0.038	0.03	0.988
	9	0.042	0.032	0.989	0.039	0.031	0.987
	10	0.042	0.033	0.988	0.04	0.031	0.987
	11	0.043	0.033	0.988	0.04	0.031	0.987
	12	0.039	0.03	0.99	0.038	0.03	0.988
M2	1	0.067	0.054	0.973	0.071	0.057	0.962
	2	0.045	0.036	0.988	0.048	0.037	0.983
	3	0.043	0.034	0.989	0.045	0.035	0.985
	4	0.042	0.033	0.989	0.044	0.034	0.986
	5	0.042	0.033	0.989	0.043	0.033	0.986
	6	0.041	0.031	0.99	0.042	0.032	0.987
	7	0.041	0.032	0.99	0.041	0.031	0.988
	8	0.04	0.032	0.99	0.041	0.031	0.988
	9	0.041	0.032	0.99	0.042	0.031	0.987
	10	0.042	0.033	0.989	0.043	0.032	0.987
	11	0.042	0.033	0.989	0.042	0.032	0.987
	12	0.039	0.031	0.991	0.04	0.031	0.988
M3	1	0.069	0.055	0.964	0.066	0.054	0.945
	2	0.045	0.036	0.985	0.049	0.039	0.973
	3	0.043	0.034	0.986	0.046	0.036	0.975
	4	0.042	0.032	0.987	0.045	0.035	0.976
	5	0.041	0.032	0.987	0.044	0.034	0.977
	6	0.04	0.031	0.988	0.042	0.033	0.978
	7	0.04	0.031	0.988	0.043	0.033	0.978
	8	0.04	0.031	0.988	0.044	0.034	0.977
	9	0.04	0.031	0.988	0.044	0.035	0.977
	10	0.041	0.032	0.987	0.044	0.035	0.976
	11	0.042	0.032	0.987	0.047	0.037	0.975
	12	0.04	0.031	0.988	0.046	0.036	0.976
M4	1	0.068	0.055	0.971	0.068	0.055	0.963
	2	0.048	0.038	0.985	0.046	0.037	0.983
	3	0.044	0.034	0.988	0.044	0.035	0.984
	4	0.042	0.033	0.989	0.042	0.033	0.986
	5	0.041	0.032	0.989	0.041	0.033	0.986
	6	0.04	0.031	0.99	0.042	0.033	0.986
	7	0.04	0.031	0.99	0.041	0.032	0.987
	8	0.04	0.031	0.99	0.042	0.033	0.986
	9	0.041	0.032	0.99	0.042	0.033	0.986
	10	0.041	0.032	0.989	0.042	0.033	0.986
	11	0.042	0.032	0.989	0.043	0.034	0.985
	12	0.039	0.03	0.99	0.04	0.032	0.987

4. Discussion

In a study conducted by Demir in 2021, the water level changes of Lake Michigan were examined using MARS, M5-tree and LSSVR methods. These three models have gone through training and testing phases. RMSE, MAE and R^2 were used as evaluation criteria. In the study, 80% of the data was used in the training phase and the remaining 20% in the testing phase. The data period of the study is between 1918 and 2020. Data deferred up to 8 months were used as the input set. When the results were examined, it was seen that better results were obtained with the MARS method (RMSE=0.0359, MAE=0.0288, R^2=0.9922). In addition, it was stated that the periodicity effect increased the model performance. (Demir, 2022).

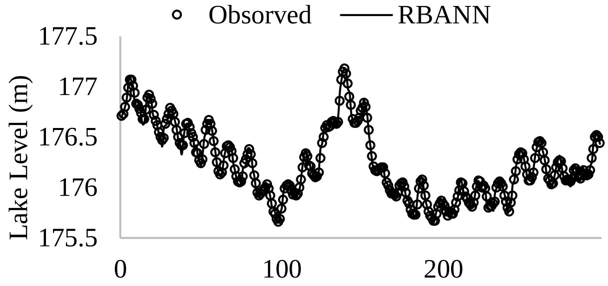

Figure 4. Timeline graph of test data

Figure 4 shows the estimates of test data for the M1 package. RBANN appears to capture the highs and lows of the test data well.

A study was conducted by Demir and Yaseen on estimating the water level in 5 large lakes in the USA by 2021. In this study, they aimed to find a reliable model for lake level estimation. In their study, they used three different models: MARS, M5-tree and LSSVR. In the study, the data set was divided into 4 parts and 75% of the data was used in the training and the remaining 25% in the testing phase. Data deferred up to 3 months were used as the input set. When the results of these three models are compared with MAE, RMSE and R^2 (RMSE=0.344, MAE=0.287, R^2=0.426), it is seen that the LSSVR model gives better results (Demir and Yaseen 2022).

In a study conducted by Çubukçu et al. in 2021 on the estimation of the monthly average water levels of Lake Michigan, data between 1981 and 2021 were used and studied with three different artificial neural network models. These models are multilayer ANN, radial basis ANN and generalized ANN models. RMSE, MAE and R2 were used as comparison criteria. In general, it was seen that all models gave good results, but according to the test results, the best training algorithm was seen as multilayer ANN, giving the best results in 12 inputs. (MAE= 0.0342, RMSE= 0.0435, R^2= 0.9906). The best method was found to be MANN, RBANN and GRNN, respectively (Çubukçu et al. 2021).

In our study, unlike the above study, when we analyze the data between 1918-2021 with the RBANN model, we see that it gives the best results in 7 entries (MAE=0.03, RMSE=0.038, R^2=0.988).

5. Conclusion

Using the monthly average lake water level data of Lake Michigan-Huron between 1918-2021, which revealed this result, estimates were made with the radial-based artificial neural network model and the results were compared.

In the study, 75% of all data were trained and 25% were tested. At this stage, four different combinations were tried, namely training (75%) and testing (25%). Evaluation criteria include absolute mean error (MAE), root mean square error (RMSE), and determination value (R^2).

It is seen that the results obtained make very good predictions in all inputs sets and training-test tests. However, according to his tests, the most successful result was the M1 data set package and the 7 inputs.

When compared with the results of previous studies, the RBANN model has acceptable accuracy.

As a result, the use of RBANN to estimate the Michigan-Huron lakes water level would be beneficial.

References

Aksoy, A. S., Aksoy, Y. R., Keskin, M. E., & Yılmazkoç, B., (2017). Göl Seviye Tahmini: Eğirdir Gölü: Mühendislik Bilimleri ve Tasarım Dergisi, 5(3), 601-608.

Albek, E. A., Albek, M. & Göncü, S. (2017). Trend Analysis of Burdur, Eğirdir, Sapanca and Tuz Lake Water Levels Using Nonparametric Statistical Methods. Afyon Kocatepe University Journal of Sciences and Engineering, 17(2), 555–570.

Arslan, N., Çiçekli, S., Dönmez, C., Şekertekin, A., (2018). Adana Seyhan Baraj Gölü Alanının Mevsimsel Değişiminin Yapay Sinir Ağları ile Analizi. Uzaktan Algılama ve Coğrafi Bilgi Sistemi Sempozumu, Ankara.

Chen, T., Cheng, P., He, H., Yan, Y. (2019). Tree Height Estimation of Forest Plantation in Mountainous Terrain from Bare-Earth Points Using a DoG-Coupled Radial Basis Function Neural Network, 11,1271.

Çubukçu, E. A., Yılmaz, C. B., Demir, V., & Sevimli, M. F. (2021). Forecasting of monthly average lake levels of Lake Michigan with artificial neural networks. *Advanced Engineering Days (AED)*, *1*, 4-7.

D'Addona, D. M. (2014). Neural Network, in: CIRP Encycl. Prod. Eng., Springer Berlin Heidelberg, Berlin, Heidelberg, 911–918. https://doi.org/10.1007/978-3-642-20617-7_6563.

Demir, V. (2022). Enhancing monthly lake levels forecasting using heuristic regression techniques with periodicity data component: application of Lake Michigan. *Theoretical and Applied Climatology*, *148*(3-4), 915-929. https://doi.org/10.1007/s00704-022-03982-0

Demir, V., & Yaseen, Z. M. (2023). Neurocomputing intelligence models for lakes water level forecasting: a comprehensive review. *Neural Computing and Applications*, *35*(1), 303-343.

Desmukh, T., Tanty, R., (2015). Application of Artificial Neural Network in Hydrology- A Review. International Journal of Engineering Research and Technology, 4(06), 184-188.

Dikbaş, F., Fırat, M., (2006). Göllerde Üç Boyutlu Hidrodinamik Modellemede POM ve Yapay Sinir Ağları Yöntemlerinin Kullanılması: Gökpınar Baraj Gölü Örneği: Mühendislik Bilimleri Dergisi, 1(12), 43-50.

Kılıç, E., Özbalcı, Ü., & Özçalık, H. R. (2012). Comparison of MLP and RBF Structures in Modeling of Nonlinear Dynamic Systems with Artificial Neural Networks, Elektrik - Elektronik ve Bilgisayar Mühendisliği Sempozyumu, 29 Kasım - 01 Aralık 2012, Bursa.

Koca, Y. (2014). Rize İyidere Alt Havzası İkİzdere Kesiti İçin Birim Hidrografın Belirlenmesi. Uzmanlık tezi, Orman ve Su İşleri Bakanlığı, Ankara.

Michigan-Huron (2023), Lake Michigan–Huron https://en.wikipedia.org/wiki/Lake_Michigan%E2%80%93Huron, Access date: 25.05.2023

Okkan, U., & Yıldırım Dalkiliç, H. (2012). Radyal tabanlı yapay sinir ağları ile Kemer Barajı aylık akımlarının modellenmesi. *Teknik Dergi*, *23*(112), 5957-5966.

Teltik, İ., Aksoy, H., Ünal, N. E. (2008). Van Gölü Su Seviyesi Stokastik Modelleri, Van Gölü Hidrolojisi ve Kirliliği Konferansı, 21-22 Ağustos 2008, DSİ XVII. Bölge Müdürlüğü, Van, 74-81.