Erciyes Üniversitesi
Fen Bilimleri Enstitusu Dergisi
Cilt 38, Sayı 3, 2022

Erciyes University
Journal of Institue Of Science and Technology
Volume 38, Issue 3, 2022

Investigation Of Lake Water Level Forecasting Performances Of Subband Decomposition Techniques

Levent Latifoğlu[*1], **Tefaruk Haktanır**[2]

[*1] Erciyes Üniversitesi Mühendislik Fakültesi İnşaat Mühendisliği, KAYSERİ
[2] Nuh Naci Yazgan Üniversitesi Mühendislik Fakültesi İnşaat Mühendisliği, KAYSERİ

(Alınış / Received: 10.08.2022, Kabul / Accepted: 29.11.2022, Online Yayınlanma / Published Online: 30.12.2022)

Keywords
Forecasting of lake water level,
Artificial Neural Networks,
Empirical Mode Decomposition,
Singular Spectrum Analysis,
Discrete Wavelet Transform

Abstract: In this study, hybrid methods have been developed for the estimation of the monthly average water level of a natural lake in the coming months from the next one to the sixth month ahead. Lake water level data were preprocessed using Discrete Wavelet Transform (DWT), Empirical Mode Decomposition (EMD), Singular Spectral Analysis (SSA) techniques and these subband signals were applied to the input data of Artificial Neural Networks (ANN). Thus, three different hybrid models were obtained and the prediction performance of these models was analyzed. According to obtained results, it was observed that the hybrid approaches obtained with the preprocessing methods applied to the water level data improved the model performance and EMD-ANN and SSA-ANN hybrid models were found to better predict average monthly lake water levels for a forecast period of one to six months than the ANN and DWT-ANN model.

Göl Su Seviyesi Tahmininde Alt Bant Ayrıştırma Tekniklerinin Performanslarının İncelenmesi

Anahtar Kelimeler
Göl su seviyesi tahmini,
Yapay Sinir Ağları,
Ampirik Kip Ayrıştırma,
Tekil Spektrum Analizi,
Ayrık Dalgacık Dönüşümü

Öz: Bu çalışmada, doğal bir gölde ortalama su seviyesinin bir aydan altıncı aya kadar olan aylık ileri tahmini için hibrit yöntemler geliştirilmiştir. Göl su seviyesi verileri Ayrık Dalgacık Dönüşümü (DWT), Ampirik Kip Ayrıştırma (EMD), Tekil Spektrum Analiz (SSA) teknikleri kullanılarak ön işleme tabi tutulmuştur ve elde edilen bu alt bant sinyalleri Yapay Sinir Ağlarına (YSA) giriş verileri olarak uygulanmıştır. Böylece üç farklı hibrit model elde edilmiş olup bu modellerin tahmin performansı analiz edilmiştir. Elde edilen sonuçlara göre, su seviyesi verilerine uygulanan ön işleme yöntemleri ile elde edilen hibrit yaklaşımların model performansını iyileştirdiği gözlemlenmiştir ve EMD-ANN ve SSA-ANN hibrit modellerinin bir ila altı aylık bir tahmin dönemi için ortalama aylık göl suyu seviyelerini ANN ve DWT-ANN modeline göre daha iyi tahmin ettiği görülmüştür.

*Corresponding Author, email: leventlatifoglu@gmail.com

1.Introduction

Information about the water-related events acquired from gauged data is important because these data are used to estimate their future magnitudes. Quantitative estimation of future hydrological phenomena is important for analyzing and predicting the behavior of the nature and also for the optimal utilization of water resources leading to a more productive water management. Lakes are one of the most important resource on the world. Therefore, accurate water level forecasting of natural lakes is crucial. Lake water level estimations are needed for planning of water supplies, hydropower plants, commercial navigation, fishing industry, design of hydraulic and other structures along the coasts, evaluation of the aquatic ecosystem, and monitoring and regulation of water quality [1, 2].
As an example for the significance of forecasting the lake water surface elevations in a couple of months ahead, we can mention the present situation of Lake Beyşehir and the man-made Beyşehir canal conveying irrigation water diverted from it. Lake Beyşehir is the largest freshwater lake in Türkiye having clear water suitable for irrigation and even for municipal purpose. At a northern edge of this lake there is a man-made canal known as Beyşehir Creek which discharges the lake with a flow rate as much as 30 m^3/s into the land-locked Konya plain

provided that the lake water surface elevation is higher than 1123.35 m above mean sea level [3]. With the objective of releasing as much water as needed during the cultivation season, the Beyşehir Weir, which has 15 bays and 15 sluice gates, was contracted by the Ottoman Empire to a German company in 1908, and the construction was completed in 1914. Conveyance of water through Beyşehir Creek with the purpose of irrigating cultivated fields along the 60 km length until the canal joins the Suğla Lake has been provided in a controlled manner since 1914, and this practice is still operational today with further smaller diversions and subsidiary canals having been added along the way [3]. This irrigation activity has significantly increased the production of mainly chickpea and lintel in that portion of the fertile Konya plain. In short, the accurate prediction of the water surface elevation of Lake Beyşehir one, two, or three months ahead of the summer season would improve irrigation planning and management. As a second example of the significant effects of lake water level changes, the case of Lake Van can be mentioned. The excessive rising of Lake Van in 1996, the largest lake in Türkiye, which lies in the Eastern Anatolia Region, caused disruption of train services between Iran and Türkiye during that period, among other problematic issues. While the average water level of Lake Van over the period from 1944 through 2007 was around 1648.0 m above MSL, the water level reached an elevation of 1950.2 m in 1996 [4-6]. Because the train station was built along Lake Van's coast to transfer the Ro-Ro train from the ferry boat to the railway at the station, the unexpected rise of Lake Van that year caused inundation of the train station. There were problems with bus services, too. And, quite a few commercial and residential buildings and some agricultural fields nearby the coast of the lake were also inundated, causing considerable losses [4-6]. If the lake surface elevation could have been forecasted three months prior to the flooding, the transportation management to and from Iran would have been planned in another way, and precautionary measures in other sectors also could have been taken beforehand. Lake water level data, like other hydrological data, are inherently nonlinear and non-stationary. Many studies have been done to estimate and simulate these data. The early studies had been based on linear approaches such as parametric autoregressive and autoregressive moving average models, which had been introduced in the 1970s to analyze time series [7, 8]. These techniques assume that the time series is stationary and linear [9, 10]. The wavelet transform (WT), which is frequently used in the analysis of time series, enables the analysis of signals in both the time and frequency domains. WT is used in the estimation of the hydrological data as well because it makes the estimation easier by decomposing the signal at different resolution levels. Sifüzzaman analyzed the WT and Fourier Transform (FT) methods by detailed comparative studies and concluded that WT has several advantages over FT [11].

Wavelet analysis and short-time Fourier analysis have been designed for linear but non-stationary data analysis [12, 13]. Using wavelet analysis, Smith et al. carried out studies on the characterization and estimation of stream flows [14].

Yet, nonparametric models have also been used for hydrological time series analysis and estimation. Yakowitz, for example, presented a stochastic model for river flows [15]. In recent years, artificial intelligence (AI) techniques have been widely used in nonlinear hydrological applications and noisy data sets [16]. Artificial intelligence methods include mathematical optimization algorithms as well as logic, classification, statistical learning, and probability-based methods. The Artificial Neural Networks (ANN) approach, based on the principle of parallel processing of information by simulating the neuron structure in the human brain, has been used in many studies on the estimation of hydrological data. In the study conducted by the ASCE Task Committee, the place of ANN in hydrology has been investigated, and the strengths and limitations of the ANN methods have been compared to those of other approaches [17]. In general, an ANN model consists of three layers: (1) the input layer, into which the input data are loaded, (2) the hidden layer, to which the activation function is applied, and (3) the output layer, which produces the expected outcome. Hornik et al. have applied the forward feed-back propagation (FFBP) algorithm to ANN for the first time and used it for the modeling of hydrological data. This algorithm finds the error between the real and the estimated values by investigating the best link weights [18]. In relevant literature, many studies on flow estimation using different ANN structures and learning algorithms can be found [19]. Furthermore, hybrid models, Mamdani and Tagai Sugeno fuzzy system approaches, and ANN models for flow estimation have been introduced by combining Support Vector Machines (SVM), Adaptive Neuro Fuzzy Inference System (ANFIS) models, and ANN models [20-22]. Partal (2008) developed a hybrid model for flow estimation by combining different ANN structures with WT [23]. The flow data were decomposed into periodic components in his model. The ANN structure was used to estimate the decomposed data. The result was a successful estimation obtained by the developed hybrid structure. As presented in the related literature, approaches based on artificial intelligence techniques such as ANN, SVM, and ANFIS, which are nonlinear models, have been used in the estimation of non-linear and non-stationary data. Hybrid approaches have been developed to improve the performance of these approaches. The most widely used hybrid approach is also the WT-ANN model [24]. The Empirical Mode Decomposition (EMD) technique is another approach for analyzing non-linear and non-stationary data [25]. Kişi et al (2014) and Rezaie-Balf et al. (2019) revealed in their studies the success of a hybrid approach using the EMD method in estimating flow data [26, 27]. Another important method that differentiates the signals with respect to time is the single-spectrum analysis (SSA) technique, which is one of the most powerful approaches among multivariate analysis schemes. This method has been widely used recently in the analysis of stationary and non-stationary data, because it is a non-parametric model and does not require priority of the data as in the EMD technique [28, 29]. In another study, Latifoğlu et al. (2015) demonstrated the success of a hybrid approach using the SSA method in the estimation problem [30, 31]. According to relevant literature research, hybrid models have been used at an increasing rate in forecasting studies in recent years. In this study, for estimating the monthly water surface elevations of the natural lake of Eğirdir in Türkiye, a pre-processing treatment has been applied to the raw gauged data by separating them into subbands by each one of the EMD, SSA, and wavelet methods separately. Next, the estimation performance of the subband data has been analyzed by ANN using the pre-processed data. The lake water level data have been estimated by each of the hybrid models EMD-ANN, SSA-ANN, and WT-ANN. The performances of these hybrid models and of the conventional ANN model have been compared with each other. To the best of our knowledge, there is not a study in the literature analyzing and comparing the performances of the EMD-ANN, SSA-ANN, and WT-ANN hybrid models for forecasting lake water levels. Some meteorological factors, such as evaporation from the lake surface, wind speed, humidity, and temperature, affect the lake water level. Naturally, these happenings vary spatially from one region to another. Instead of inserting many inputs into the forecasting model, some of which may not even be gauged, forming a forecasting model that uses the recorded previous lake levels is obviously more advantageous. In this study, the lake water levels for the near future for periods from one month up to six months have been forecasted using past recorded data only by hybrid models, which have not been used before for lake level predictions using proposed models.

2. Material and Method

Recently, artificial intelligence techniques have been used as an alternative to classical approaches for estimation of hydrological data. Among them, ANN is quite common. In addition to the traditional ANN method, the relevant literature reveals that data pre-processing has recently been used for estimation calculations. The wavelet transform has been used in earlier attempts at pre-processing the hydrological data, which decomposed the data according to its frequency characteristics [23, 32, 33]. Because the hydrological data have a non-linear and non-stationary structure, the EMD scheme turns out to be one of the methods that can be used appropriately for the time-domain analysis of these data [26, 34]. Another suitable method of analysis for stationary/non-stationary and linear/nonlinear data is the SSA approach. This method is a multivariate analysis method that processes signals based on their time domain properties. Because there is no need for prior data, the EMD methods are widely used [28-31, 35]. In this study, the performance of a hybrid approach for estimation of lake water levels has been investigated. For this purpose, the gauged lake water level data have been divided into two segments as training and test data and further divided into subbands by the EMD, SSA, and DWT methods [30, 31]. The subband signals have been estimated using ANN. The flow chart of the study is shown below in Figure 1.

2.1 Lake water level data description

The monthly average water surface elevations of Lake Eğirdir had been continuously gauged from 1953 through 1999. Hence, this data consisting of 523 successive numbers has been used as the material for this study (Figure 2). 310 of these elements have been used for the training phase, and the rest, 213 values, have been used for the testing stage [31]. These data are normalized as shown in Figure 3.

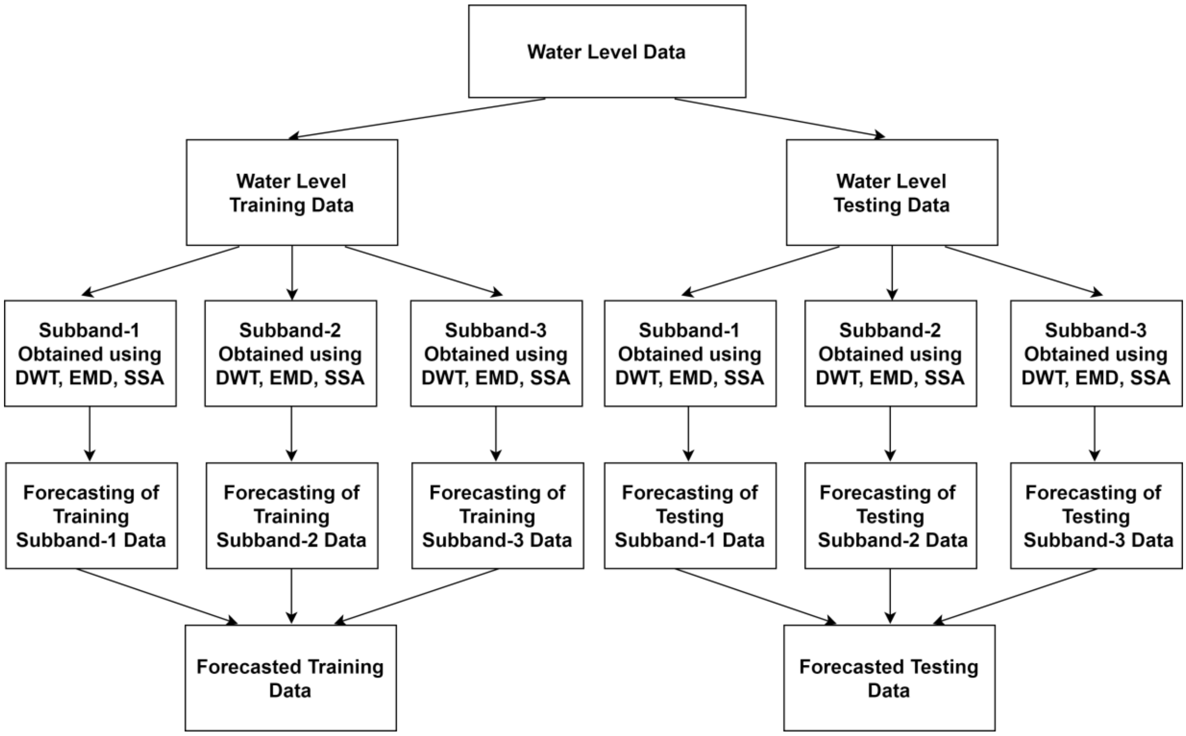


Figure 1. Flow chart for the proposed forecasting model

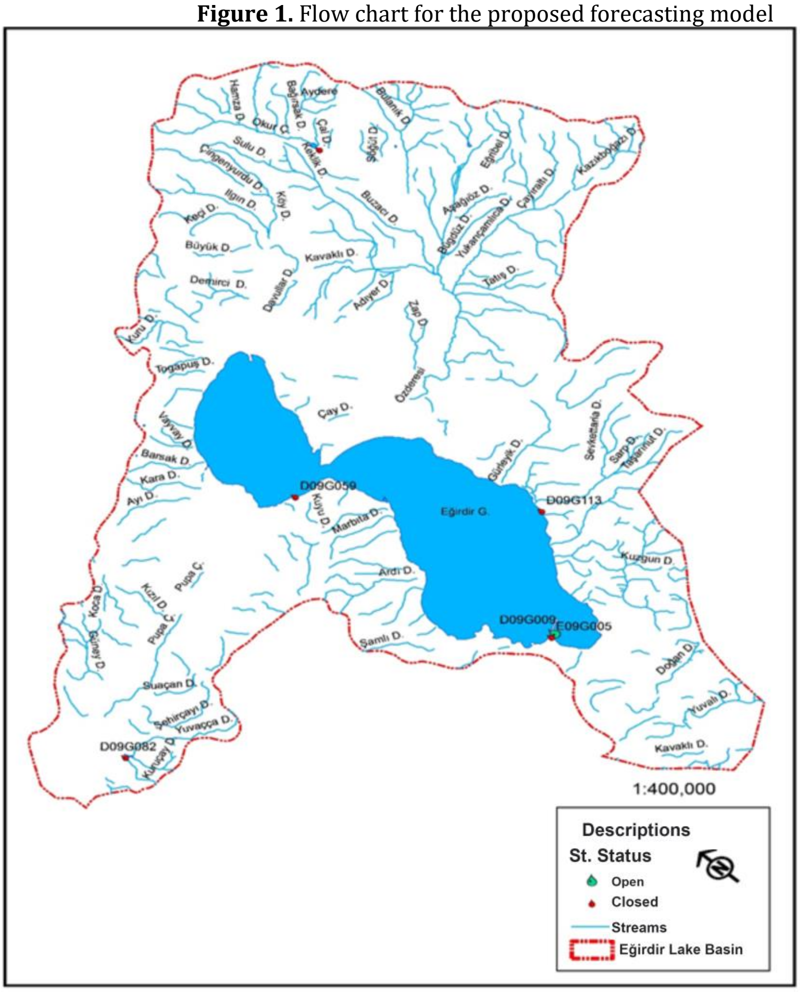


Figure 2. Basin of the Lake Eğirdir in Türkiye

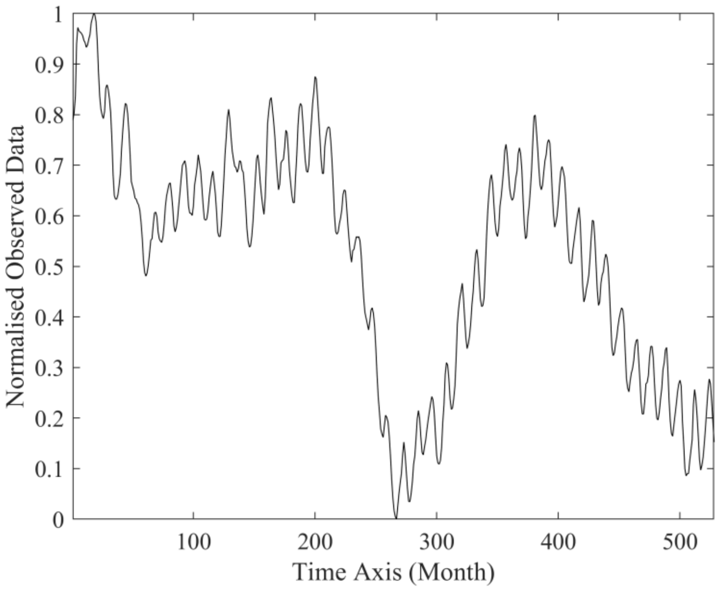
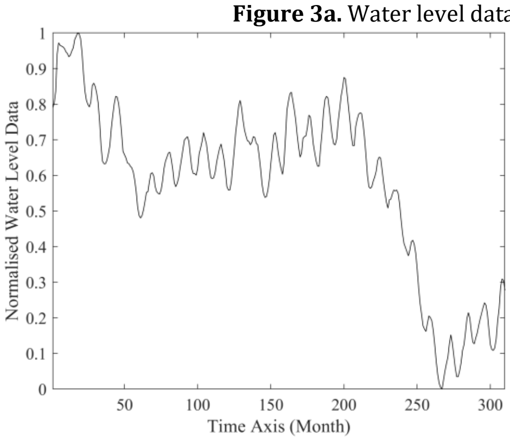


Figure 3a. Water level data measured on Lake Eğirdir



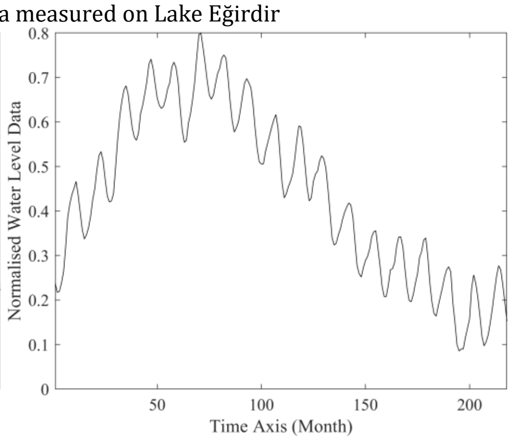


Figure3b. Train data for forecasting model

Figure3c. Test data for forecasting model

2.2 Empirical Mode Decomposition method

The Empirical Mode Decomposition (EMD) method has decomposed a signal into components called Intrinsic Mode Functions (IMF). The signal is defined as the sum of the IMF components and the residual signal, as depicted by Equation 1 below [25].

$$x(t) = \sum_{i=1}^{k} IMF_i + r(t) \tag{1}$$

The signal is subjected to the constraints defined below.

- The total number of maximum or minimum points and zero crossing points in the data set must be equal or differ by one element only.
- The average of the envelopes formed by local maximas at any position and the envelopes defined by local minimas should be zero.

This time-domain separation approach is simple to implement and incredibly efficient. The signals in the time axis are handled as a mix of repeating and original oscillations with a local mean of zero that are dispersed symmetrically around it in this technique.

The initial step is to identify the maxima and minima, as well as the zero crossing locations.

The upper and lower envelopes of the signals are obtained **in the second** step by treating both peaks and minima with cubic interpolations.

In the third stage, Equation 2 below computes the average of the upper and lower envelopes for the whole time period.

$$m1(t) = [z_{max}(t) + z_{min}(t)] / 2 \tag{2}$$

In the fourth step, the average signal is subtracted from the actual signal as given by Equation 3 below to obtain the detail signal, d(t).

$$d(t) = x(t) - m(t) \tag{3}$$

The detail signal is processed as if it were a new original signal in **the fifth stage**, and the process steps are repeated until the terminate requirement is met.

The two-level decomposition approach was used in this study to obtain good estimation accuracy on both the training and testing parts of the lake water level data. As shown in Figure 4, three subsets were found, including two subsets of IMF and R [31]. In all data sets, the last value has to be neither maxima nor minima. The reason for this is; if the final values in the data are determined as maxima or minima, a limitation of the so-called boundary effect of the EMD method arises for estimation studies and makes the model dependent on the future [35, 36]. In order to prevent this, while the IMF values in the training data have been determined, the final values in the data have been chosen not including the extremum points [36, 37].

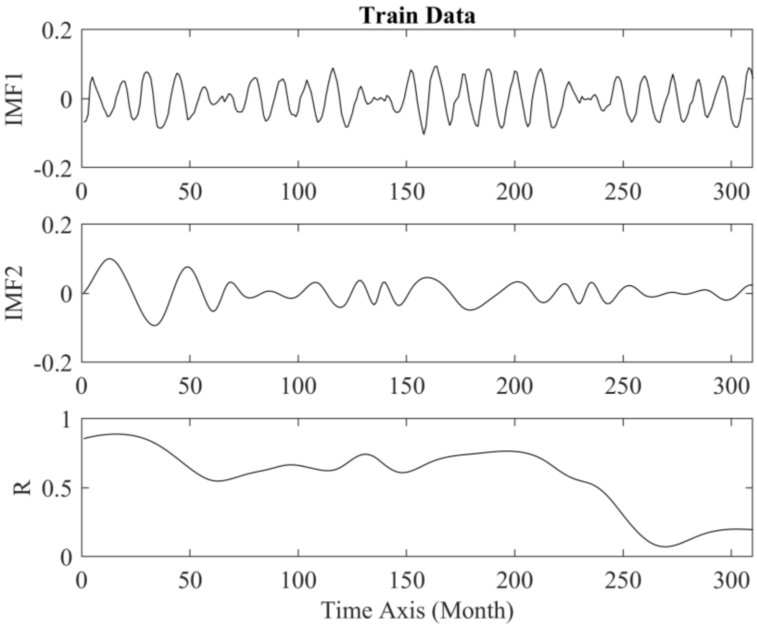


Figure 4a. Training subband signal decomposed by the EMD method

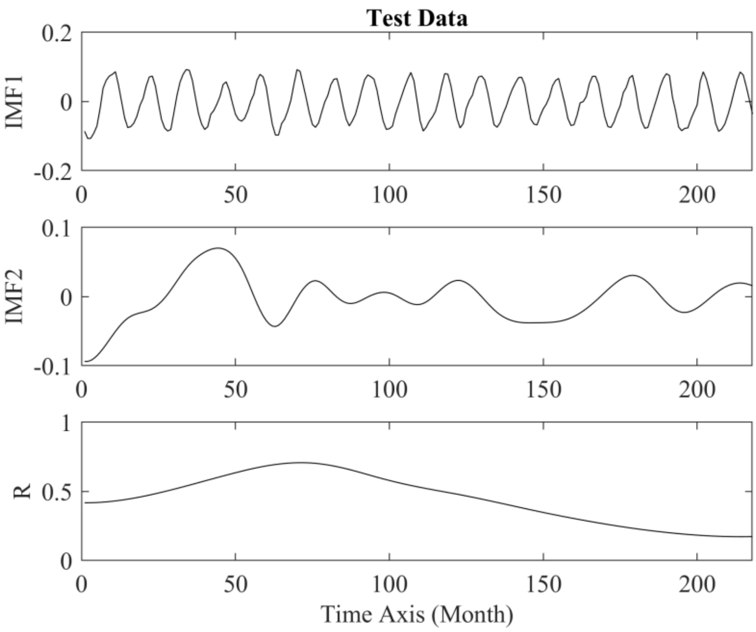


Figure 4b. Testing subband signal decomposed by the EMD method

2.3 Singular Spectrum Analysis method

Singular-spectrum analysis (SSA) is a decomposition method that uses the time information of the signal. SSA does not require any parametric model and does not call for a prerequisite. This method involves two main stages: decomposition and reconstruction. In the decomposition step, the embedding process and singular value decomposition (SVD) are applied, which express the signal as a matrix. In the reconstruction phase, there are grouping and diagonal averaging processes [28].

In the embedding process, which is the first step of the decomposition stage, the time series: X = [X_1, X_2, X_3,……,X_N] is expressed as a matrix of K rows and L columns (K×L) instead of a one-dimensional single row by taking the specified window length (L). The embedding process, expressed as the trajectory matrix of the signal, is also referred to as the Hankel matrix. The window length L is in the range of 1<L<N, N is the length of the signal and K = N − L + 1. In this case, the trajectory matrix Tx is defined as follows:

$$T_X = \begin{bmatrix} x_1 & x_2 & \cdots & x_L \\ x_2 & x_3 & \cdots & x_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_K & x_{K+1} & \cdots & x_N \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_K \end{bmatrix} = X_1 + X_2 + \ldots + X_K \quad (4)$$

The SVD process is carried out after the embedding process as a second step of the decomposition stage. Multiplications of Tx and TxT matrices yield three kinds of matrices called the eigen triple (U orthonormal unit matrix of K×K size and S diagonal matrix of K×L size from the Tx_i*Tx_iT operation, and V orthonormal unit matrix of L×L size from the Tx_iT*Tx_i operation). And, X_i's are determined as follows:

$$X_i = U_i * S_i * V_i \quad (5)$$

Next, the second stage, the reconstruction stage including the grouping and averaging steps is executed. At the grouping step, the singular values obtained from SVD (S_i) are plotted, categorized according to their components such as monthly or seasonal components, white noise, trend. Equation 4 is rewritten with the number of categories determined based on the representation rate of eigenvalues.

(6)

$\dot{I}1$ $\dot{I}2$ $\dot{I}m$

At the last step of this stage called averaging, the regrouped matrix Tx is transformed into a time serie by averaging the X_{ij} elements diagonally. The average of those X_{ij} elements for which the sum of the i and j indices is a constant yields the kth element of the new time series component, called the reconstructive component (RC) satisfying i+j = k+2. This procedure is said to be dehankelization of the matrix. At the end of this process, the new subband time series components (RC) are obtained.

Since the monthly lake water level data used in this study should reflect the nature's periodic structure, the window length is determined to be 12 in the embedding process. The data are divided into subbands according to the eigenvalues (1.), (2.) and (3-12.). These are the subband signals, which are expected to lead to the best estimates as seen in Figure 5 [31].

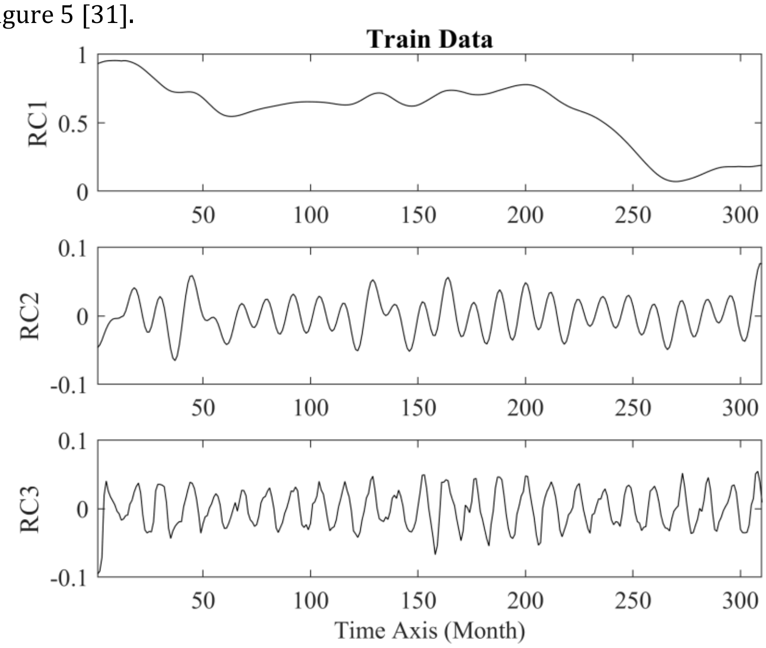


Figure 5a. Training subband signals decomposed by the SSA method

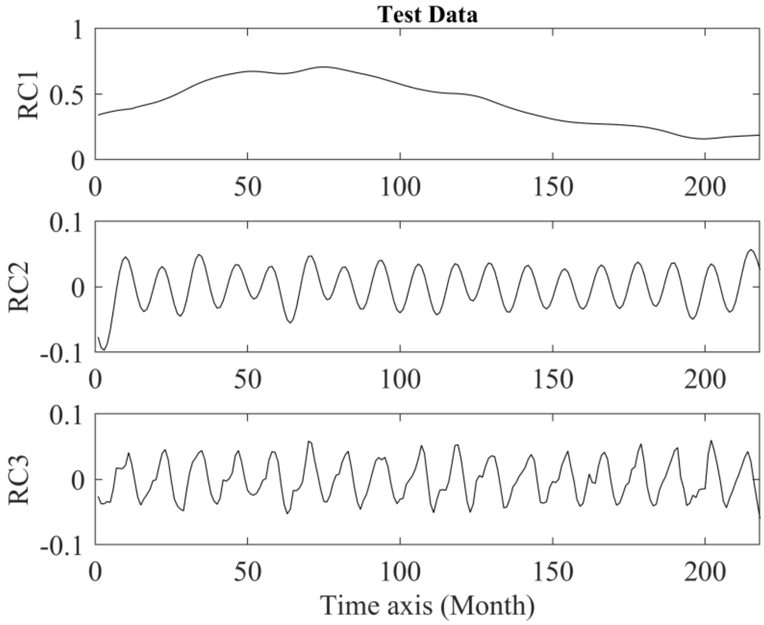


Figure 5b. Testing subband signals decomposed by the SSA method

2.4 Discrete Wavelet Transform method

The Discrete Wavelet Transform (DWT) method runs like a filter bank that decomposes the signals according to the frequency bands. This method, which is based on sub-band coding, has been introduced for the fast calculation of the Continuous Wavelet Transform method, is used to analyze the signal at different scales using different cut-off frequency filters. In order to handle the low frequencies in it, the signal is passed through the low pass filter, and to analyze the high frequencies in the signal, it is passed through the high-pass filter. The resolution of the signal, indicating the amount of detailed information in the signal, is changed by the filtering process, and the scale is replaced by downsampling and upsampling operations. x[n], where n is an integer, represents the signal (Fig. 6). While the low pass filter is indicated by G0, the high-pass filter is indicated by H0 [36].

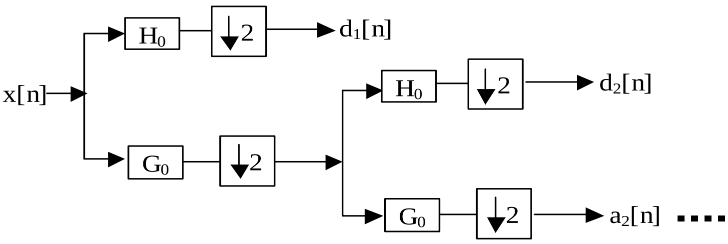


Figure 6. Subband decomposition operation during DWT

For each level, while the high pass filter produces the information of the detail (d[n]), the low pass filter produces the information of the approximation (a[n]) by the scale function. At each decomposition level, the semi-band filters produce a signal in half the frequency band. The filtering process continues until it reaches the desired level. This process is also called "multi-resolution analysis". The maximum number of levels depends on the length of the data. This operation is the opposite of the decomposition process (Fig. 7). The detail and the approximation coefficients at each level are sampled with 2 and passed through the low- and high-pass filters, then combined. Then, this process continues to the level of the decomposition process to obtain the original signal [36].

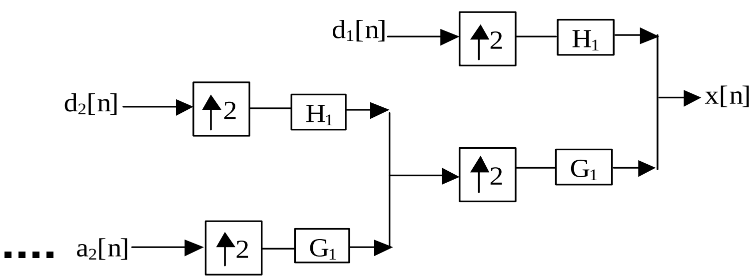


Figure 7. Reconstruction operation during DWT

In this study, the lake water level data have been treated by a two-level decomposition process. Thus, the lake water level signal has been decomposed into the approximation (a2) and detail (d2, d1) subbands (Fig. 8). We have used Daubechies' 4th wavelet function in the decomposition process. The following figure shows the lake water level data used in the training and testing phases, which are decomposed into subbands by using the Discrete Wavelet Transform (DWT) method.

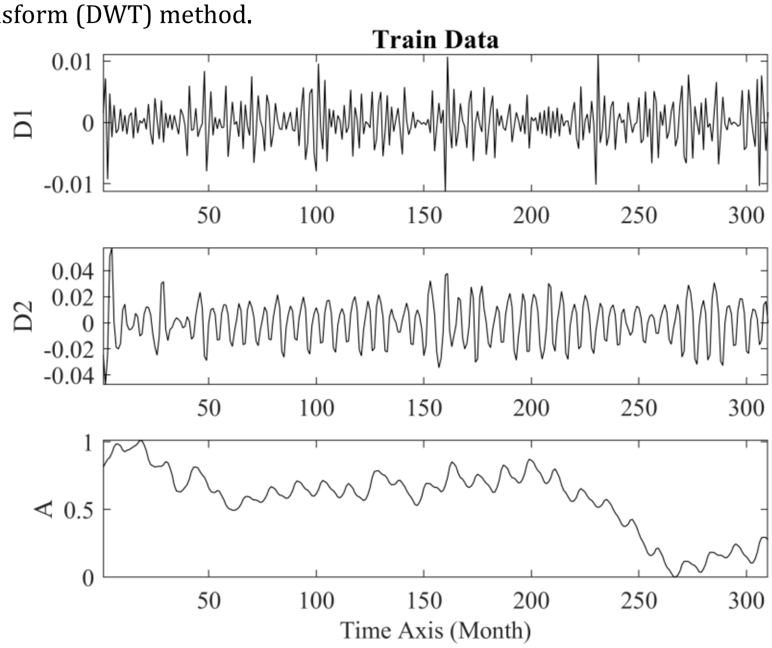


Figure 8.a. Training subband signal decomposed by the DWT method

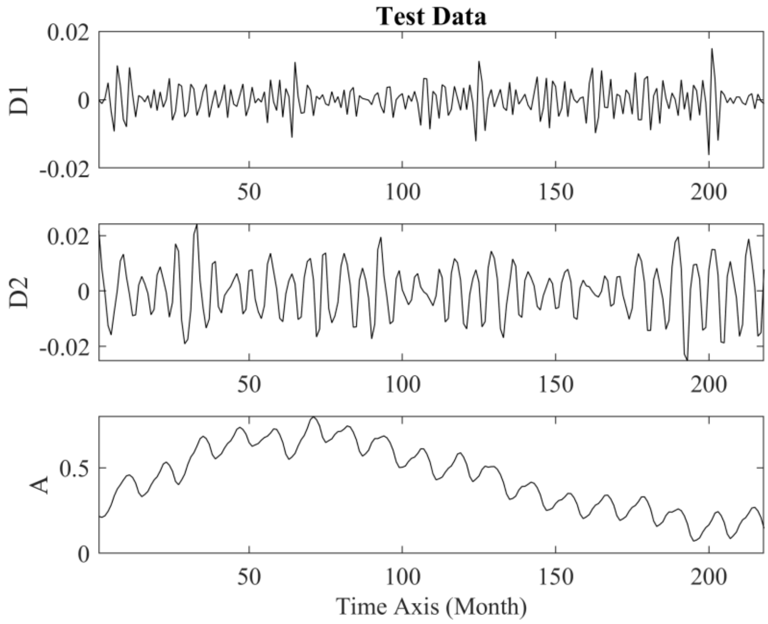


Figure 8.b. Testing subband signal decomposed by the DWT method

2.4 The method of Artificial Neural Networks

Because of its features, such as the capacity to handle nonlinear structures and parallel and serial processing capabilities, the ANN has recently become widely employed in data estimation. The sum and activation functions, as well as the learning strategy of the processor components and the learning rule, are employed in an ANN structure, and the topology determined by the connection of process elements produces the network model. Artificial neural cells (neurons) unite to create their ANN. Neurons are not collected at random. Cells, in general, comprise a three-layer network and are arranged in parallel inside each layer. These are the input, inner (hidden), and output layers [37, 38].

The most widely used version of ANN is the multi-layer perceptron (MLP) model. An MLP model consists of the input layer, one or more hidden layers, and the output layer, and each one has at least one neuron layer. The input layer processor components act as a buffer, distributing input signals to the hidden layer processing units. Each of these parts incorporates the input data into the sum function by multiplying it by the weight coefficients that represent the effectiveness of the input over the hidden neuron. These sum functions are then transferred via a transfer function to compute the output value of that neuron. These procedures are carried out for each processor in this tier. The processor units of the output layer also serve as hidden layer elements, calculating network output values. Because the information flow in the MLP model is in the forward direction, it is also known as "feed forward ANN". One advantage of this approach is that it may train the network using a variety of learning techniques. The weights of the network are modified according to the training procedure until the error between the network's output and the desired output is reduced. Back Propagation Neural Network (BPNN) is a feed forward network and is the most commonly used ANN model in time series estimation. In our study, the MLP model and back propagation neural network (BPNN) feed forward network structure were used. In the ANN, different learning algorithms are used to train the network. MLP-ANN is based on a supervised learning strategy in which both input and output (that has to be produced corresponding to the input) are applied to the network during training. The task of the network is to produce an output for each input. MLP-ANN learning occurs in two steps. The output of the network is computed in the first stage, the forward calculation stage. The weights are determined in the second stage, the backward calculation stage, based on the difference between the predicted output and the network output. The MLP-ANN learning rule is known as the back-propagated MLP learning rule because it is carried out in this manner via back-propagating. The Levenberg-Marquardt (LM) learning method, which outperforms the ANN in terms of computing speed, is utilized here to estimate lake water level data. The LM algorithm is similar to the Quasi-Newton method based on the least squares calculation and the maximum neighborhood approaching the second-degree training speed without the use of a Hessian matrix [39].

2.5 Parameters Used For Performance Criteria

The statistical criteria commonly used for comparing the estimation accuracies of the hydrological data by various modeling techniques are the mean absolute error (MAE), the mean square error (MSE), the determination coefficient (R^2), and the correlation coefficient (R). In the lake water level estimation study, we have used these parameters to evaluate the performances of the models.

The Mean Square Error (MSE): MSE is the arithmetic average of the squares of the differences of the observed values in the series from the ones estimated by the model used, and is defined by Equation 7 below.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(X_{observed,i} - X_{estimated,i})^2 \tag{7}$$

The Mean Absolute Error (MAE): MAE is the arithmetic average of the absolute differences of the observed values in the series from the ones estimated by the model used, and is defined by Equation 8 below.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|X_{observed,i} - X_{estimated,i}| \tag{8}$$

Determination Coefficient (R^2): R^2 is a measure which quantitatively reflects the accuracy of the estimated values given by the model defined by Equation 9 below. The second term in Equation 9 approaches zero for a powerful model. Therefore, R^2 approaches 1 for a good model.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}[X_{observed,i} - X_{estimated,i}]^2}{\sum_{i=1}^{N}[X_{observed,i} - mean\ of\ X_{observed,i}]^2} \tag{9}$$

Correlation Coefficient (R): It indicates the degree and trend of whether a linear relationship exists between the observed and estimated series. R accepts values ranging from -1 to +1. There is no association between the two data sets if R is near to zero, a significant positive relationship if R is close to +1, and a negative relationship if R is close to -1. Equation 10 below defines R.

$$R = \frac{1}{N-1}\sum_{i=1}^{N}\left(\frac{X_i - \mu_X}{\sigma_X}\right)\left(\frac{Y_i - \mu_Y}{\sigma_Y}\right) \tag{10}$$

Here, μ_X and σ_X represent the mean and standard deviation of the X data set, respectively, whereas μ_Y and σ_Y represent the mean and standard deviation of the Y data set.

3. Results

3.1 Results for forecasting the lake water levels using the ANN Model

In the first stage of this study, in order to evaluate the effect of pre-processing by the DWT, SSA, and EMD methods on the estimation of the water level data of Lake Eğirdir, the data estimation process has been performed without any pre-processing on these data by using the ANN model. For the estimation of the data, the MLP-ANN structure consists of the input, an hidden layer, and an output layer. The number of hidden layer neurons has been increased from 1 to 9 stepwise and the number of neurons in the hidden layer has been determined according to the lowest mean square error. In order to determine the best ANN model; MSE and MAE values have been used as the network performance criteria and trained by using the Levenberg Marquardt back propagation algorithm. Successively one-month through six-months ahead forecasted lake levels have been computed using such formed ANN model and the error values that occured during the ANN training and testing stages are given in Table 1. Figure 9 shows the graphs of one-month ahead and six-months ahead forecasted values of the test phase by the ANN model [31].

Table 1. The performance values of the estimated data obtained from one to six ahead forecasting results using the ANN model for the monthly water levels of Lake Eğirdir

ONE AHEAD FORECASTING WITH ANN MODEL							
TEST				**TRAIN**			
MSE	**MAE**	**R**	**R^2**	**MSE**	**MAE**	**R**	**R^2**
0.0009	0.0264	0.9870	0.9742	0.0008	0.0227	0.9860	0.9855
TWO AHEAD FORECASTING WITH ANN MODEL							
0.0033	0.0501	0.9541	0.9103	0.0027	0.0436	0.9542	0.9522
THREE AHEAD FORECASTING WITH ANN MODEL							
0.0064	0.0696	0.9120	0.8318	0.0051	0.0610	0.9120	0.9098
FOUR AHEAD FORECASTING WITH ANN MODEL							
0.0093	0.0846	0.8726	0.7614	0.0073	0.0731	0.8726	0.8617
FIVE AHEAD FORECASTING WITH ANN MODEL							
0.0117	0.0935	0.8443	0.7128	0.0090	0.0806	0.8443	0.8472
SIX AHEAD FORECASTING WITH ANN MODEL							
0.0136	0.0982	0.8317	0.6897	0.0096	0.0821	0.8317	0.8293

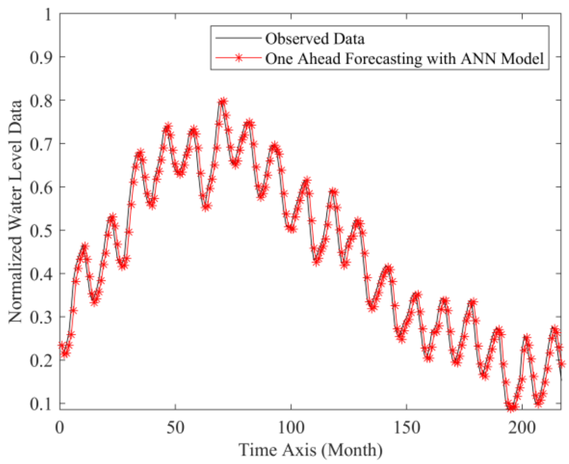


Figure 9a- One month ahead forecasted water level data obtained from the ANN model

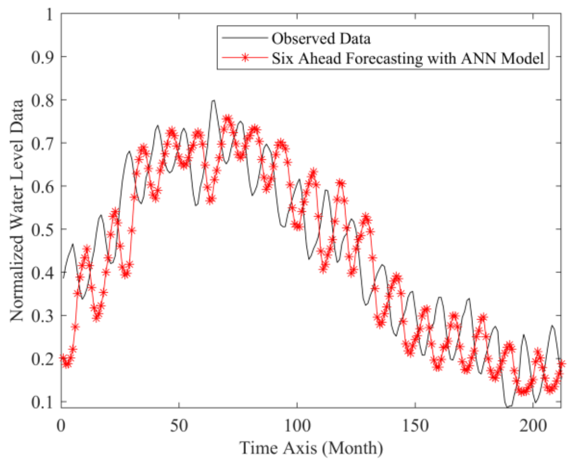


Figure 9b- Six month ahead forecasted water level data obtained from the ANN model

3.2 Results for forecasting the lake water levels using the EMD-ANN model

First, the training and test data have been decomposed into two IMF components by the EMD method. In this way, hydrological training and test data have been defined by the three-component equation of: The ANN model has been used to estimate the data of Lake Eğirdir, which has been divided into three components, two IMF and one residual data.

For the estimation study, the same structure and design procedures have been carried out in the EMD-ANN model as with the ANN. For each subband, ANNs have been trained using the Levenberg Marquardt back propagation algorithm. In order to determine the best ANN model, the MSE and MAE values have been used as performance criteria. Using the EMD-ANN model, one-month through six-months ahead forecasted values have been obtained, and the error values that occur during the EMD-ANN training and testing phases have turned out to be as given in Table 2. Figure 10 shows the graphs of one-month ahead and six-months ahead forecasted values of the test phase by the EMD-ANN model [31].

Table 2a. The performance values of the subband estimated data obtained for one-step to six-step ahead forecasting results of subbands by using EMD-ANN model for monthly water levels of Lake Eğirdir

	TEST				TRAIN			
	MSE	**MAE**	**R**	**R^2**	**MSE**	**MAE**	**R**	**R^2**
				ONE AHEAD FORECASTING				
IMF1	0.0006	0.0198	0.8599	0.7378	0.0008	0.0246	0.8698	0.7442
IMF2	0.00001	0.0065	0.9816	0.9831	0.000001	0.0036	0.9913	0.9292
R	0.00001	0.0073	0.9995	0.999	0.000001	0.0043	0.9997	0.9992
				TWO AHEAD FORECASTING				
IMF1	0.0017	0.0339	0.5063	0.2596	0.0023	0.0417	0.4919	0.2398
IMF2	0.0002	0.0122	0.8644	0.8073	0.0001	0.0064	0.9697	0.9442
R	0.0002	0.0116	0.9984	0.9967	0.0001	0.0068	0.9989	0.9976
				THREE AHEAD FORECASTING				
IMF1	0.0022	0.0395	0.0505	0.0193	0.0030	0.0481	0.0172	0.0118
IMF2	0.0004	0.0169	0.8837	0.8761	0.0001	0.0092	0.9360	0.8142
R	0.0004	0.0170	0.9963	0.9945	0.0003	0.0152	0.9973	0.9927
				FOUR AHEAD FORECASTING				
IMF1	0.0018	0.0357	0.4144	0.2433	0.0024	0.0427	0.4380	0.1639
IMF2	0.0007	0.0208	0.6428	0.7841	0.0002	0.0126	0.8855	0.4132
R	0.0007	0.0226	0.9936	0.9899	0.0006	0.0208	0.9950	0.9872
				FIVE AHEAD FORECASTING				
IMF1	0.0010	0.0264	0.7332	0.6246	0.0011	0.0273	0.7954	0.5480
IMF2	0.0009	0.0233	0.7027	0.6658	0.0003	0.0153	0.7940	0.2411
R	0.0011	0.0281	0.9901	0.9837	0.0010	0.0266	0.9918	0.9802
				SIX AHEAD FORECASTING				
IMF1	0.0005	0.0192	0.9042	0.8220	0.0005	0.0184	0.8722	0.7586
IMF2	0.0005	0.0183	0.6907	0.4309	0.0010	0.0237	0.4103	0.2796
R	0.0015	0.0327	0.9878	0.9757	0.0015	0.0334	0.9859	0.9720

Table 2b The magnitudes of the parameters used as performance criteria for estimation of monthly water levels of Lake Eğirdir by the EMD-ANN model for one-step to six-step ahead forecastings

TEST				TRAIN			
MSE	**MAE**	**R**	**R^2**	**MSE**	**MAE**	**R**	**R^2**
			ONE AHEAD FORECASTING WITH EMD-ANN MODEL				
0.0009	0.0252	0.9879	0.9758	0.0007	0.0226	0.9939	0.9876
			TWO AHEAD FORECASTING WITH EMD-ANN MODEL				
0.0025	0.0430	0.9649	0.9306	0.0021	0.0392	0.9817	0.9640
			THREE AHEAD FORECASTING WITH EMD-ANN MODEL				
0.0035	0.0503	0.9515	0.9076	0.0032	0.0474	0.9729	0.9468
			FOUR AHEAD FORECASTING WITH EMD-ANN MODEL				
0.0034	0.0471	0.9563	0.9175	0.0033	0.0468	0.9716	0.9433
			FIVE AHEAD FORECASTING WITH EMD-ANN MODEL				
0.0027	0.0405	0.9698	0.9405	0.0030	0.0445	0.9743	0.9483
			SIX AHEAD FORECASTING WITH EMD-ANN MODEL				
0.0027	0.0498	0.9717	0.9429	0.0020	0.0443	0.9742	0.9493

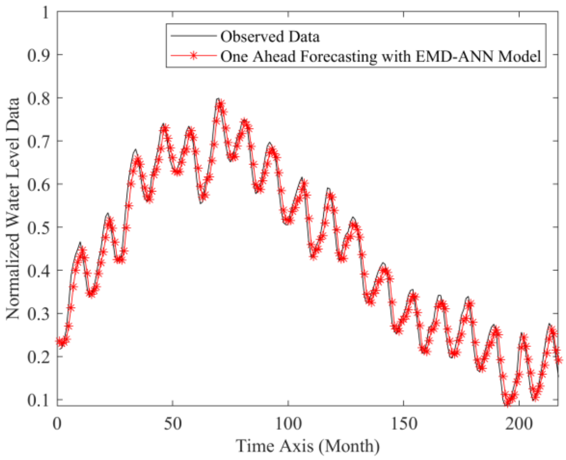


Figure 10.a. One month ahead forecasted water level data obtained from EMD-ANN model

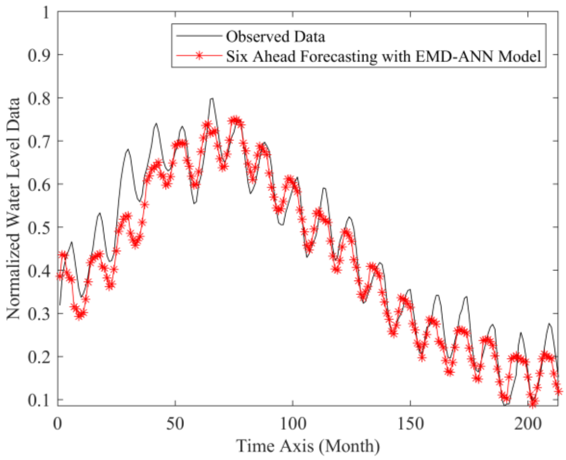


Figure 10.b. Six month ahead forecasted water level data obtained from EMD-ANN model

3.3 Results for forecasting the lake water levels using the SSA-ANN model

The signals, which are the lake water levels in this case, are decomposed into related and non-related components symbolized by RCi, using the SSA method, and they have been recovered as RC1 + RC2 + RC3. In the SSA process, the signals are subjected to the hankelization process with a 12 length of window size and transformed into the trajectory matrix. The component with the highest eigenvalue is denoted by RC1, the component with the second highest eigenvalue by RC2, and the component with the other eigenvalues by RC3.
For the estimation study, the same structure and design procedures have been carried out in the SSA-ANN model as in the ANN and EMD-ANN models. Using the SSA-ANN model also, one-month through six-months ahead forecasted values have been obtained; and the resulting error values that occurred with this model are presented in Table 3. Figure 11 shows the graphs of one-month-ahead and three-months-ahead forecasted values of the test phase by the SSA-ANN model [31].

Table 3.a The performance values of the subband estimation data obtained from one to six ahead forecasting results of subbands using the SSA-ANN model for monthly water levels of Lake Eğirdir

				ONE AHEAD FORECASTING				
	TEST				TRAIN			
	MSE	MAE	R	R^2	MSE	MAE	R	R^2
RC1	0.000001	0.0031	0.9999	0.9954	0.000001	0.0024	0.9999	0.9987
RC2	0.0003	0.0070	0.9889	0.8185	0.000001	0.0042	0.9983	0.8243
RC3	0.0008	0.0247	0.8709	0.7416	0.0007	0.0216	0.9232	0.7574
				TWO AHEAD FORECASTING				
RC1	0.000001	0.0031	0.9998	0.9963	0.000001	0.0029	0.9998	0.9968
RC2	0.0003	0.0079	0.9801	0.7515	0.0001	0.0072	0.9948	0.7857
RC3	0.0025	0.0428	0.5477	0.3369	0.0022	0.0394	0.7457	0.3577
				THREE AHEAD FORECASTING				
RC1	0.0001	0.0055	0.9994	0.9914	0.0001	0.0074	0.9996	0.9929
RC2	0.0001	0.0068	0.9912	0.8942	0.0001	0.0077	0.9929	0.9655
RC3	0.0037	0.0526	0.1315	0.098	0.0036	0.0499	0.5053	0.0897
				FOUR AHEAD FORECASTING				
RC1	0.0001	0.0060	0.9991	0.9846	0.000001	0.0057	0.9993	0.9877
RC2	0.0001	0.0079	0.9917	0.8142	0.0001	0.0097	0.9880	0.9537
RC3	0.0034	0.0499	0.1001	0.0606	0.0049	0.0863	0.1806	0.0842
				FIVE AHEAD FORECASTING				
RC1	0.0001	0.0075	0.9986	0.9757	0.0001	0.0053	0.9991	0.9814
RC2	0.0002	0.0094	0.9862	0.9478	0.0002	0.0119	0.9815	0.9431
RC3	0.0034	0.0500	0.0829	0.0823	0.0048	0.0553	0.1652	0.1563
				SIX AHEAD FORECASTING				
RC1	0.0002	0.0090	0.9979	0.9642	0.0001	0.0084	0.9984	0.9741
RC2	0.0003	0.0110	0.9785	0.9696	0.0003	0.0141	0.9737	0.8081
RC3	0.0029	0.0446	0.4435	0.4081	0.0046	0.0543	0.2383	0.2315

Table 3b The performance values of the estimation data obtained from one to six six ahead forecasting results using the SSA-ANN model for the monthly water levels of Lake Eğirdir

SSA-ANN MODEL ONE AHEAD FORECASTING							
TEST				TRAIN			
MSE	MAE	R	R²	MSE	MAE	R	R²
0.0010	0.0257	0.9869	0.9758	0.0008	0.02180	0.9939	0.9776
TWO AHEAD FORECASTING WITH SSA-ANN MODEL							
0.0028	0.0451	0.9611	0.9540	0.0024	0.0357	0.9623	0.9589
THREE AHEAD FORECASTING SSA-ANN MODEL							
0.0037	0.0545	0.9430	0.9214	0.0030	0.0526	0.9659	0.9533
FOUR AHEAD FORECASTING SSA-ANN MODEL							
0.0038	0.0552	0.9456	0.9325	0.0024	0.0511	0.9528	0.9489
FIVE AHEAD FORECASTING SSA-ANN MODEL							
0.0039	0.0568	0.9446	0.9502	0.0036	0.0517	0.9514	0.9605
SIX AHEAD FORECASTING SSA-ANN MODEL							
0.0053	0.0583	0.9529	0.9520	0.0045	0.0467	0.9532	0.9529

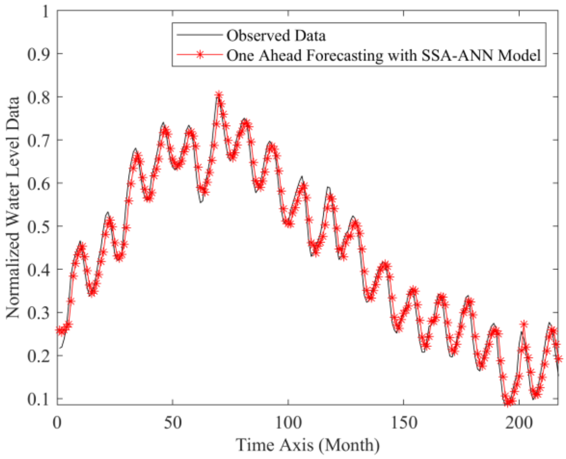


Figure 11.a. One month ahead forecasted water level data computed by the SSA-ANN model

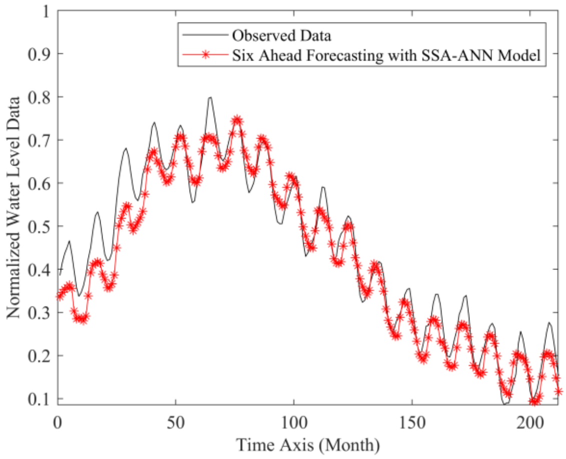


Figure 11.b. Six month ahead forecasted water level data computed by the SSA-ANN model

3.4 Results For Forecasting The Lake Water Levels Using The DWT-ANN Model

For the monthly lake water level data, the DWT method also has been decomposed into related and non-related components. The two levels of the signals that are decomposed by DWT are approximation (A) and detail (D1, D2) signals. In this way, both hydrological training and test data are expressed by the equation: x (t) = A + D1 + D2. For the estimation study, the same structure and design procedures have been carried out in the DWT-ANN model as in the ANN, EMD-ANN, SSA-ANN models. Using the DWT-ANN model also, one-month through six-months ahead forecasted values have been computed and the error values that occurred during the training and test phases turned out to be as given Table 4. Figure 12 shows the graphs of one-month ahead and three-months ahead forecasted values of the test phase by the DWT-ANN model.

Table 4.a The performance values of the subband estimation data obtained from one to six ahead forecasting results of the subbands using the DWT-ANN model for the monthly water levels of Lake Eğirdir

	ONE AHEAD FORECASTING							
	TEST				TRAIN			
	MSE	MAE	R	R²	MSE	MAE	R	R²
D1	0.000001	0.0034	0.4273	0.1826	0.000001	0.0028	0.4302	0.1851
D2	0.000007	0.0068	0.4379	0.1917	0.00002	0.0122	0.4494	0.2019
A	0.00009	0.0253	0.9882	0.9766	0.00007	0.0215	0.9934	0.9868
	TWO AHEAD FORECASTING							
D1	0.000002	0.0035	0.3841	0.1475	0.00001	0.0030	0.3161	0.0999
D2	0.000009	0.0078	0.3254	0.1059	0.00002	0.0121	0.4974	0.2474
A	0.0034	0.0486	0.9578	0.9174	0.0024	0.0396	0.9777	0.9559
	THREE AHEAD FORECASTING							

D1	0.000002	0.0035	0.4710	0.2219	0.000001	0.0029	0.2150	0.0462
D2	0.000003	0.0051	0.7602	0.5779	0.00001	0.0085	0.7731	0.5976
A	0.0064	0.0681	0.9219	0.8500	0.0044	0.0534	0.9594	0.9204
FOUR AHEAD FORECASTING								
D1	0.000002	0.0038	0.1326	0.0176	0.00001	0.0030	0.1369	0.0187
D2	0.000008	0.0077	0.3314	0.1098	0.00002	0.0120	0.4913	0.2414
A	0.0089	0.0801	0.8914	0.7947	0.0063	0.0638	0.9413	0.8860
FIVE AHEAD FORECASTING								
D1	$2.4456x10^{-5}$	0.0039	0.1479	0.0219	0.000001	0.0032	0.0093	0.00008
D2	$1.0356x10^{-4}$	0.0081	0.1755	0.0308	0.00002	0.0126	0.3453	0.1192
A	0.0106	0.0867	0.8714	0.7593	0.0076	0.0702	0.9280	0.8612
SIX AHEAD FORECASTING								
D1	0.00002	0.0042	0.2515	0.0633	0.000001	0.0030	0.2234	0.0499
D2	0.000005	0.0058	0.7448	0.5548	0.000009	0.0077	0.7879	0.6207
A	0.0107	0.0867	0.8688	0.7548	0.0085	0.0748	0.9186	0.8438

Table 4b The performance values of one to six ahead forecasting data obtained by using the DWT-ANN model for monthly water levels of Lake Eğirdir

ONE AHEAD FORECASTING WITH WAVELET-ANN MODEL							
TEST				**TRAIN**			
MSE	**MAE**	**R**	**R²**	**MSE**	**MAE**	**R**	**R²**
0.0010	0.0267	0.9874	0.9750	0.00007	0.0224	0.9932	0.9865
TWO AHEAD FORECASTING WITH WAVELET-ANN MODEL							
0.0038	0.0516	0.9529	0.9081	0.0024	0.0400	0.9779	0.9563
THREE AHEAD FORECASTING WITH WAVELET-ANN MODEL							
0.0067	0.0697	0.9189	0.8443	0.0044	0.0540	0.9592	0.9200
FOUR AHEAD FORECASTING WITH WAVELET-ANN MODEL							
0.0090	0.0816	0.8899	0.7920	0.0066	0.0662	0.9381	0.8800
FIVE AHEAD FORECASTING WITH WAVELET-ANN MODEL							
0.0109	0.0887	0.8663	0.7505	0.0083	0.0738	0.9213	0.8488
SIX AHEAD FORECASTING WITH WAVELET-ANN MODEL							
0.0113	0.0897	0.8613	0.7418	0.0094	0.0796	0.9094	0.8270

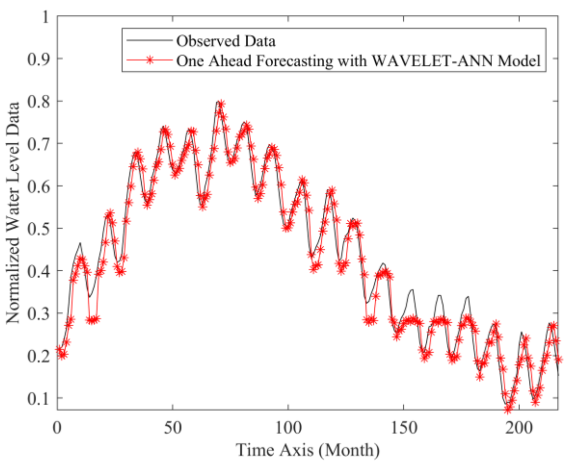


Figure 12.a. One month ahead forecasted water level data obtained by the DWT-ANN model

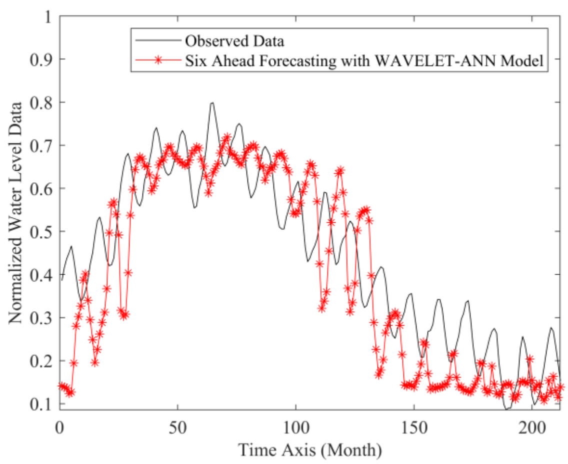
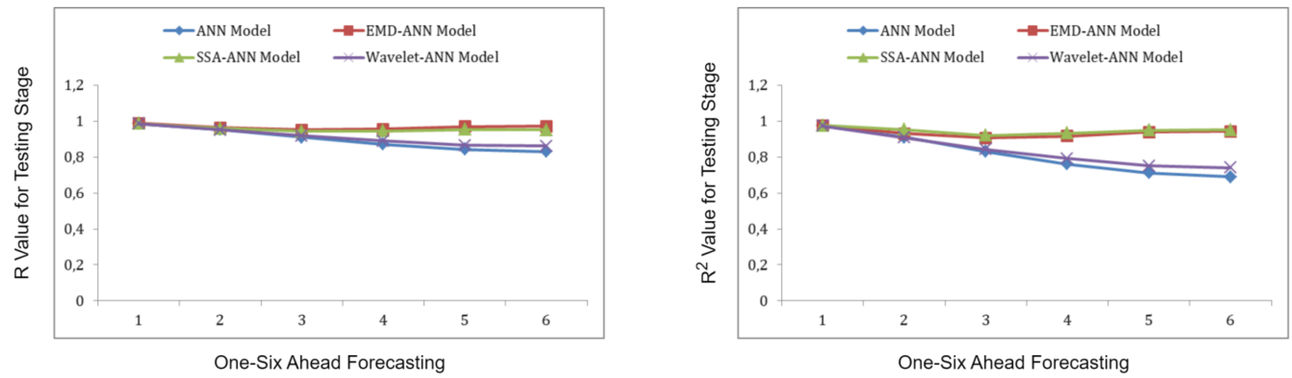


Figure 12.b. Six month ahead forecasted water level data obtained by the DWT-ANN model

4. Discussion and Conclusion

When the literature is examined, there are studies that include linear and nonlinear approaches for estimating the lake water level. In a study carried out in the literature, lake water levels were estimated up to 3-day time intervals by using the autoregressive moving average (ARMA), ANN, adaptive-neuro-fuzzy inference system (ANFIS), and gene expression programming (GEP). According to the results obtained, the superiority of GEP, ANFIS, and ANN models over ARMA models has been demonstrated [40]. In another study, a recurrent neural network (RNN) and ANFIS models were constructed, and the most suitable model was searched. In addition,

autoregressive (AR) and autoregressive moving average (ARMA) models, which are classical stochastic models, were obtained and compared with RNN and ANFIS models. The results show that RNN and ANFIS can be applied successfully and provide higher performance in lake level estimation than AR and ARMA models [41]. In a review article study, it was seen that the main machine learning (ML) models ANN, SVM, ANFIS, hybrid models, evolutionary ML models, ELM, and deep learning models were used in lake water level estimation.
Also, in another review article, the performance of M5-Tree, multivariate adaptive regression spline (MARS) and least square support vector regression (LSSVR) models in lake water level prediction was analyzed [42]. Demir has shown that the MARS model performs better than LSSVR and M5-tree [43]. Although great progress has been made in the application of ML models to predict the water level in lakes, there is still much space for further research [44].
In this study, the lake water level was estimated using the ANN, EMD-ANN, SSA-ANN, and DWT-ANN methods, and the estimation performance of these four methods with the effect of preprocessing the time series data in the forecasting study was analyzed. The difference between this study and the literature studies is that it analyzes the performance of the preprocessing methods in lake water level estimation.
All of the parameters of each of the four models are computed using the available recorded series, and in our study, this is the 310-month portion of the approximately 44-year-long gauged data. Once a model is formed in the optimized way summarized in the relevant sub-sections above, based on the single lake water value of the present month, the water level of either one-month or two-months, …, or six-months ahead is computed (estimated) by that model. For example, for estimating the 311st water level by any model, the water level measured on the 310th day is given as the input value. Similarly, for any one-month ahead estimated water level, the actual water level measured one month before, for each 213-element test series, is used. Also, the actual water level of the present month is used for estimation of water levels more than one month ahead. Figures 9, 10, 11, and 12 reveal the same information for the models of ANN, EMD-ANN, SSA-ANN, and DWT-ANN, respectively. The black lines in all of these eight figures are exactly the same because they exhibit the actual lake water levels measured over the 213-month test period on normalized scales. The red points represent the estimated values in the 213-element test segment. In any case, the closer the red points are to the solid black line, the better the estimation accuracy. These eight figures are provided here to allow the reader to make a visual assessment of the goodness of estimation accuracy of any model. As it can be observed in these figures, although the one-month ahead forecasts by all four methods look pretty good, the six-months ahead water levels computed by the EMD-ANN and SSA-ANN models are closer to the black line than the other two. The two-months, three-months, four-months, and five-months ahead levels are not given with the purpose of not elongating the paper. They have appearances that are relatively similar to the six-month-ahead figures, given here as Figures 9.b through 12.b.
Figures 13a and 13b show the magnitudes of the performance parameters of R, and R^2 obtained for all four models over the 213-month test segment, respectively. As it is seen in these four figures, one-month ahead prediction abilities of all of the four models seem to be good and close. However, the values begin deviating from each other appreciably with increasing forecast periods up to six months. The calculated performance parameters for the models of ANN, and DWT-ANN deteriorate while those for the models of EMD-ANN and SSA-ANN stay fairly good and much better than the former with increasing forecast periods.



c d

Figure 13. The performance parameters of the EMD-ANN, SSA-ANN, DWT-ANN and ANN models for the estimation of monthly water levels of Lake Eğirdir
a. R values, **b.** R^2 values

In this study, the configurations of these four models have been formed the way summarized here and their model parameters have been determined out of the 310-month sequentially gauged data of water levels of Lake Eğirdir as summarized above. We have used the remaining 213-month segment of sequentially gauged water levels in Lake Eğirdir as testing for the developed models. The measured water levels of Lake Eğirdir over this 213-month period are the actual naturally occurring water levels, and we have computed the estimated water levels for one-month ahead to six-month-ahead periods using these four models. Therefore, Figures 10 through 12, and especially Figure 13, are tangible validations of these models, and yet they are comparisons of them for betterness in estimation accuracies. Visual observations of graphs of actual water levels and estimated water levels are in the same figures as those in Figures 9 through 12. The second quantification of the R and R^2performance of these forecasted values is shown in Figure 13. The ANN, DWT-ANN, EMD-ANN, and SSA-ANN models are used to forecast the monthly average water levels of Lake Eğirdir in Turkey one, two, and six months ahead. In the study of the lake water level by the ANN model, one input layer, one hidden layer, and one output layer structure is used. In the estimation study carried out by the EMD-ANN model, the hydrological data are divided into three subbands. Two of these are the IMF components, and the other is the data called the residual component. These subbands are named IMF1, IMF2, and R, and an input-output ANN model is used for the estimation of IMF1, IMF2, and R subband data. In the estimation study performed by the SSA-ANN model, the hydrological data are divided into three subbands called reconstruction components (RCs). These subband data consist of the subband data RC1 having the first highest eigenvalue, the subband data RC2 having the second highest eigenvalue, and the subband data RC3 having the sum of the other residual eigenvalues. In the estimation study performed by the DWT-ANN model, the hydrological data are divided into three subbands. Two of these are the detail subbands, and the third one is the approximation data subband. These subbands are named D1, D2, and A. As for the other models, one input, one output ANN is used for the estimation of D1, D2, and A subband data. In all of the models developed for the lake water level estimation, we have performed estimation studies from one month ahead up to six months ahead. The worst estimation performance is obtained in three-month-ahead forecasting. The performances are compared in three-months ahead forecasting, which is used to determine the best results under the worst conditions. For the three-months ahead forecasting, the magnitudes of the MSE, MAE, R, and R^2 performance parameters have turned out to be; for the ANN model, 0.0064, 0.0696, 0.9120, and 0.8318; for the EMD-ANN model, 0.0035, 0.0503, 0.9515, and 0.9076; for the SSA-ANN model, 0.0037, 0.0545, 0.9430, and 0.9214; and for the DWT-ANN model, they are: 0.0067, 0.0697, 0.9189, and 0.8443, respectively. In the study of the hydrological data on temporal variations of the water surface level of a natural lake, it is seen that the performances of the estimation studies carried out by the EMD-ANN and SSA-ANN models are better than the estimation performances of the ANN and DWT-ANN models. Also, when compared to other models, the EMD-ANN and SSA-ANN models perform quite well, particularly in two-month, three-month, and six-month ahead estimation studies. The overall result of this study is that the preprocessing performed by the SSA and EMD procedures for forecasting the water level of a natural lake one to six months in the future by such a hybrid ANN method appreciably improves the estimation performance. Hence, the conclusion and recommendation for natural lake management administrations is that a hybrid ANN model similar to either the EMD-ANN or the SSA-ANN models developed in this study will enable them to forecast the average lake water surface elevations in the coming couple of months up to six months with a pretty good accuracy. In this study, we have used two-thirds of the gauged lake water data for obtaining the relevant parameters of these models, and we have used the remaining one third for testing the goodness of the forecasts. The goodness of estimations has been checked quantitatively by those four comparison criteria. The prediction ability of the recommended models is indeed good because the errors are very small and the R^2s are very close to 1.0. The proposed approach should be used for forecasting the future water levels of the other lakes reliably.

References

[1] Chow, V.T., Maidment, D.R., Mays, L.W. 1988. Applied Hydrology. McGraw-Hill. NY

[2] Ebtehaj, I., Bonakdari, H., Gharabaghi, B. 2019. A reliable linear method for modeling lake level fluctuations. Journal of Hydrology, 570, 236–250.

[3] Muşmal, H. 2015. Beyşehir Regülâtörü (Taş Köprü), Tarih Okulu Dergisi (TOD) Journal of History School (JOHS). 357-373

[4] Degens, E. T., Wong, H. K., Kempe, S., & Kurtman, F. (1984). A geological study of Lake Van, eastern Turkey. Geologische Rundschau, 73(2), 701–734.

[5] Batur, E., Kadıoğlu, M., Özkaya, M., Saban, M., Akın, İ., Kaya, Y. 2008. Water Level Modelling of Lake Van and Estimation of Extrem Levels. Van Lake Hydrology and Pollution Conference, 10-25.

[6] Kılınçaslan, T. 2000. The rising water level in Lake Van: environmental features of the Van basin which increase the destructive effect of the disaster. Water Science and Technology, 42(1–2), 173–177.

[7] Box, G.E.P., Jenkins, G.M. 1970. Time Series Analysis: Forecasting and Control. Holden-Day. San Francisco.

[8] Carlson, R.F., MacCormick, A.J.A., Watts, D. G. 1970. Application of linear random models to four annual streamflow series. Water Resources Research. 6, 1070-1078.

[9] Grimaldi, S. 2004. Linear parametric models applied to daily hydrological series. Journal of Hydrologic Engineering. 9, 383-391.

[10] Noakes, D.J., McLeod, A.I., Hipel, K.W. 1985. Forecasting monthly riverflow time series. International Journal of Forecasting. 1:179-190.

[11] Sifuzzaman, M., Islam, M.R., Ali, M.Z.., 2009. Application of wavelet transform and its advantages compared to fourier transform. J. Phys. Sci.. 13, 121–134.

[12] Cahill, A.T. 2002. Determination of changes in streamflow variance by means of a wavelet-based test. Water Resources Research, 38, 1065-1078.

[13] Padmanabhan, G., Rao. A.R. 1988. Maximum entropy spectral analysis of hydrologic data. Water Resources Research. 24, 1519-1533.

[14] Smith, L. C., Turcotte, D. L., Isacks, B. L. 1998. Stream flow characterization and feature detection using a discrete wavelet transform. Hydrological processes. 12, 233-249.

[15] Yakowitz, S. J. 1973. A stochastic model for daily river flows in an arid region. Water Resources Research. 9, 1271-1285.

[16] Kentel, E. 2009. Estimation of river flow by artificial neural networks and identification of input vectors susceptible to producing unreliable flow estimates. Journal of hydrology. 375, 481-488.

[17] Hydrology, A. T. C. on A. of A. N. N. in. 2000. Artificial neural networks in hydrology. I: Preliminary concepts. Journal of Hydrologic Engineering, 5(2), 115–123.

[18] Hornik, K., Stinchcombe, M., White, H. 1989, Multilayer feedforward networks are universal approximators. Neural Networks. 2, 359-266.

[19] Veintimilla-Reyes, J., Cisneros, F., Vanegas, P. 2016. Artificial Neural Networks applied to flow prediction: A use case for the Tomebamba river. Procedia Engineering. 162, 153-161.

[20] El-Shafie, A., Taha, M. R., & Noureldin, A. 2007. A neuro-fuzzy model for inflow forecasting of the Nile river at Aswan high dam. Water resources management. 21, 533-556.

[21] He, Z., Wen, X., Liu, H., & Du, J. 2014. A comparative study of artificial neural network. adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. Journal of Hydrology. 509, 379-386.

[22] Noori, R., Karbassi, A. R., Moghaddamnia, A., Han, D., Zokaei-Ashtiani, M. H., Farokhnia, A., & Gousheh, M. G. 2011. Assessment of input variables determination on the SVM model performance using PCA. Gamma test. and forward selection techniques for monthly stream flow prediction. Journal of Hydrology. 401, 177-189.

[23] Partal, T. 2008. River flow forecasting using different artificial neural network algorithms and wavelet transform. Canadian Journal of Civil Engineering. 36, 26-38.

[24] De Macedo Machado Freire P.K., Santos C.A.G., Da Silva G.B.L. 2019. Analysis of the use of discrete wavelet transforms coupled with ANN for short-term streamflow forecasting. Applied Soft Computing. 80:494-505, 2019.

[25] Huang, N. E., Shen, Z., & Long, S. R., A new view of nonlinear water waves: the Hilbert spectrum. Annual review of fluid mechanics. 31, 417-457.

[26] Kişi O., Latifoğlu L., Latifoglu F. 2014. Investigation of Empirical Mode Decomposition in Forecasting of Hydrological Time Series. Water Resources Management. 28, 4045-4057.

[27] Rezaie-Balf M., Kim S., Fallah H., Alaghmand S. 2019. Daily river flow forecasting using ensemble empirical mode decomposition based heuristic regression models: Application on the perennial rivers in Iran and South Korea. Journal of Hydrology. 572, 470-485.

[28] Hassani, H., Soofi, A. S., & Zhigljavsky, A. A. 2010. Predicting daily exchange rate with singular spectrum analysis. Nonlinear Analysis: Real World Applications. 11, 2023-2034.

[29] Marques, C. A. F., Ferreira, J. A., Rocha, A., Castanheira, J. M., Melo-Goncalves, P., Vaz, N., Dias, J. M. 2006. Singular spectrum analysis and forecasting of hydrological time series. Physics and Chemistry of the Earth. Parts A/B/C. 31:1172-1179.

[30] Latifoğlu, L., Kis,i O., Latifoglu, F. 2015. Importance of hybrid models for forecasting of hydrological variable. Neural Computing & Applications. 26, 1669-1680.

[31] Latifoğlu, L. July 2017. Forecasting of Hyrological Variables Using New Hybrid Methods, Erciyes University, Graduate School of Natural and Applied Sciences Ph.D. Thesis.

[32] Mehr, D. A., Kahya, E., Bagheri, F., Deliktas, E. 2014. Successive-station monthly streamflow prediction using neuro-wavelet technique. Earth Sci. Informatics. 7, 217–229.

[33] Sahay, R. R., Srivastava, A. 2014. Predicting monsoon floods in rivers embedding wavelet transform. genetic algorithm and neural network. Water resources management. 28, 301-317.

[34] Karthikeyan, L., Kumar, D.N. 2013. Predictability of nonstationary time series using wavelet and EMD based ARMA models. Journal of Hydrology. 502,103–119.

[35] Wang, X., Wu, J., Liu, C., Wang, S., & Niu, W. 2016. A Hybrid Model Based on Singular Spectrum Analysis and Support Vector Machines Regression for Failure Time Series Prediction. Quality and Reliability Engineering International. 32, 2717-2738.

[36] Broughton, S.A., Bryan, K. 2018. Discrete Fourier Analysis and Wavelets: Applications to Signal and Image Processing. Wiley.

[37] Zhang, Z. 201). Artificial neural network. In Multivariate time series analysis in climate and environmental research, 1–35.

[38] Wang, S.-C. 2003. Artificial neural network. In Interdisciplinary computing in java programming. 81–100.

[39] Marquardt, DW. 1963. An algorithm for least-squares estimation of nonlinear parameters. Journal of the society for Industrial and Applied Mathematics. 11, 431-441.

[40] Kişi, O., Shiri, J., Nikoofar, B. 2012. Forecasting daily lake levels using artificial intelligence approaches. Computers & Geosciences. 41, 169-180.

[41] Güldal, V., Tongal, H. 2010. Comparison of recurrent neural network, adaptive neuro-fuzzy inference system and stochastic models in Eğirdir Lake level forecasting. Water resources management. 24, 105-128.

[42] Demir, V., Yaseen, Z. M., 2022. Neurocomputing intelligence models for lakes water level forecasting: a comprehensive review. Neural Computing and Applications. 1-41.

[43] 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), 915-929.

[44] Zhu, S., Lu, H., Ptak, M., Dai, J., Ji, Q. 2020. Lake water-level fluctuation forecasting using machine learning models: a systematic review. Environmental Science and Pollution Research. 27, 44807-44819.