Pol. J. Environ. Stud. Vol. 28, No. 5 (2019), 3451-3462
DOI: *10.15244/pjoes/93923* ONLINE PUBLICATION DATE: 2019-04-28

Original Research

Estimating Dam Reservoir Level Fluctuations Using Data-Driven Techniques

Fatih Üneş[1]*, Mustafa Demirci[1], Bestami Taşar[1], Yunus Ziya Kaya[2], Hakan Varçin[1]

[1]Iskenderun Technical University, Civil Engineering Department, Hydraulics Division, İskenderun, Hatay, Turkey
[2]Osmaniye Korkut Ata University, Civil Engineering Department, Hydraulics Division, Osmaniye-Turkey

Received: 26 February 2018
Accepted: 2 August 2018

Abstract

Estimating dam reservoir level is very important in terms of the operation of a dam, the safety of transport in the river, the design of hydraulic structures, and determining pollution, the salinity of the river flow fluctuations and the change of water quality in the dam reservoir. In this study, an adaptive network-based fuzzy inference system (ANFIS), support vector machines (SVM), radial basis neural networks (RBNN) and generalized regression neural networks (GRNN) approaches were used for the prediction and estimation of daily reservoir levels of Millers Ferry Dam on the Alabama River in the USA. Particularly, the feasibility of ANFIS as a prediction model for the reservoir level has been investigated. The Millers Ferry Dam on the Alabama River in the USA was selected as a case study area to demonstrate the feasibility and capacity of ANFIS, SVM, RBNN, and GRNN. The model results are compared with conventional auto-regressive models (AR), auto-regressive moving average (ARMA), multi-linear regression (MLR) models, and artificial intelligence models for the best-input combinations. The comparison results show that ANFIS models give better results than classical and other artificial intelligence models in estimating reservoir level.

Keywords: reservoir level, prediction, adaptive network-based fuzzy inference system, support vector machines, radial basis neural networks, generalized regression neural networks

Introduction

Water-resource investigations should embrace and deal with effective reservoir water storage, which can be employed for water supply, irrigation, providing flood and drought control, hydropower and energy security of local generation. Dam reservoirs are used to provide storage for water and are important elements in planning and managing water resources. If the dam reservoir level is estimated correctly, the most suitable dam operation performance can be obtained. Reservoir level parameters are inflow to reservoir, water storage in reservoir, water release from reservoir, evaporation, soil moisture and infiltration. They represent uncertainties and must be considered in water resource operation. Reservoir water level is also important in the analysis and design of several water resource projects such as dam construction, irrigation needs and flood control. During a flood, the opening of the dam weir passage should be sufficient to ensure that reservoir

*e-mail: fatih.unes@iste.edu.tr

capacity does not exceed the limits and that discharges do not flow down. If drought occurs, the reservoir needs to collect water and be released adequately to fulfill its aims. Reservoir water release modeling is significant to make fast and accurate decisions when dealing with flood and drought. Statistical prediction is important in determining dam reservoir level.

Recently, artificial intelligence techniques such as artificial neural networks (ANN), adaptive network-based fuzzy inference system (ANFIS), and radial basis neural network (RBNN) methods have been widely applied in the modeling complex nonlinear phenomena in hydrology and water resource systems.

The artificial intelligence approach is a technique widely used in several engineering problems such as estimating suspended sediment, evaporation, groundwater level and nearshore bar volumes [1-5]. Zaji and Bonakdari [6] applied genetic algorithm artificial neural network and genetic programming methods to estimate lake level. Nitsure et al. [7] used ANN and fuzzy techniques to estimate changes in seawater level. Unes [8] and Unes et al. [9] developed an ANN model for predicting reservoir level fluctuation. Parmar and Bhardwaj [10] used neural networks, fuzzy and wavelet coupled model to predict river water. Yanping et al. [11] used a radial basis neural network model to establish the surrogate model of groundwater flow numerical simulation in Jinquan Industrial Park, Inner Mongolia, China. Wiktor et al. [12] used a neural network model based on salinity, since salinity reduces the amount of nutrients and soil temperature. They investigated spatial analysis of soil properties and identification of factors affecting their diversity in agricultural areas of mountain catchments. Shiri et al. [13] developed an extreme learning machine, genetic programming and artificial neural network models for predicting water-levels in Urmia Lake.

Emamgholizadeh et al. [14] used the ANN model for forecasting groundwater levels of Bastam Plain. Sun and Trevor [15] developed multiple model combination methods for annual maximum water level prediction during river ice breakup. Unes [16] predicted plunging depth of density flow in a dam reservoir using the ANN technique. Himan et al. [17] developed a neuron-fuzzy model for identifying the most suitable location for a water reservoir in the area of Batu Pahat, Johor, Malaysia. Demirci and Baltacı [18] estimated suspended sediment of the Sacremento River in the USA using fuzzy logic. Seo et al. [19] used artificial intelligence techniques for daily water level forecasting. Piri and Kahkha [20] predicted the water level fluctuations of Chahnimeh Reservoirs in Zabol using ANN, ANFIS and the Cuckoo optimization algorithm.

In this study, ANFIS, RBNN, SVM and conventional auto regressive (AR), autoregressive moving average (ARMA), and multi linear regression (MLR) techniques are investigated to estimate daily reservoir levels. The later sections of this article are arranged as follows. In the second part, the applied dataset as well as the ANFIS, RBNN and SVM techniques are discussed. The third section describes the statistical proposals applied for model analysis, and the fourth section contains cues and discussions. Finally, the last section gives the conclusions of this study.

Materials and Methods

Case Study

Daily reservoir level records of Millers Ferry Dam on the Alabama River in the USA were used in the present study. The Millers Ferry Lock and Dam are located near a community on the Alabama River, Alabama. The

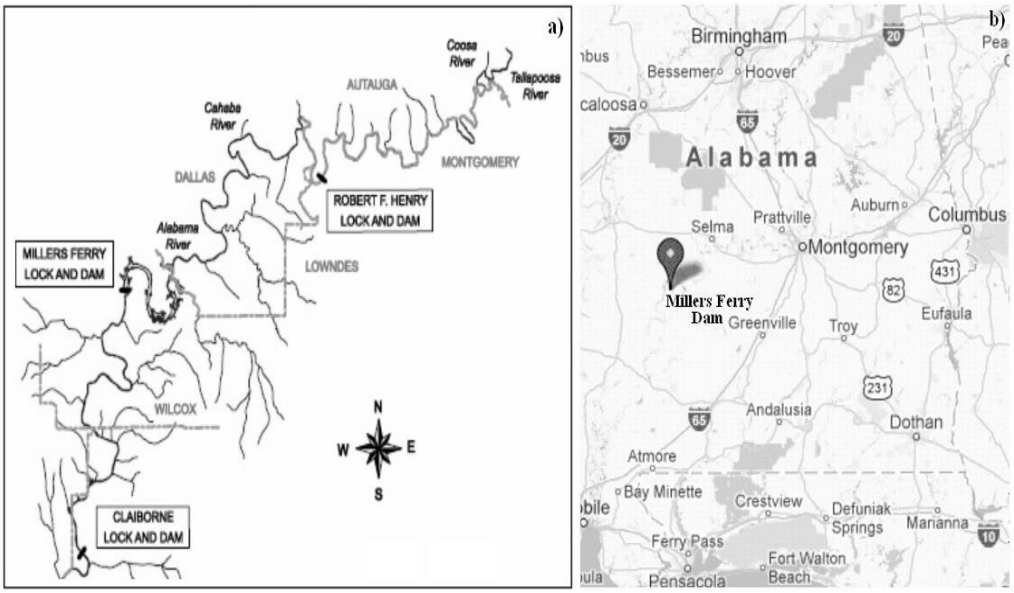

Fig. 1. a) Millers Ferry Dam Map and b) location of Millers Ferry Dam in USA.

Table 1. Statistical parameters of the applied data set during the study period.

	X_{max}	X_{min}	X_{mean}	S_x	C_{sx}
Training period	24.56	23.52	24.26	0.247	-1.472
Testing period	24.57	23.77	24.34	0.126	-0.937
Whole period	24.57	23.52	24.29	0.215	-1.739

X_{mean} X_{max}, X_{min}, S_X and C_{SX} denote the mean, maximum, minimum, standard deviation and skewness coefficient, respectively.

dam is a lock and a hydroelectric dam on the Alabama River that was built and operated by the United States Army Corps of Engineers. The construction of the complex began in 1963 and was completed in 1974. The Millers Ferry Station was commissioned in 1970 and has a production capacity of 90 megawatts. The reservoir covers 70 km^2 and has approximately 800 km of shoreline (Fig. 1).

The gauging station datum is 40 m above sea level. The time-series data of Millers Ferry Dam station (USGS Station No. 02427505 operated by the US Geological Survey (USGS)) were used in the study. Data sample consists of 6 years (from 1 October 2006 to 19 December 2012) of daily reservoir level records. The first 1461 daily levels data were used for training, the remaining 811 daily levels data were used for testing each model. Table 1 indicates the statistical parameters of used data during the study period.

Adaptive Network Fuzzy Inference System (ANFIS)

Jang [21] first introduced NF. It approximates any real continuous function on a compact set to any degree of accuracy. The NF is functionally equivalent to fuzzy inference systems (Jang et al., [22]). In this study, a fuzzy inference system with two and more input variables such as x, y, z .., and one output variable f was used. The resulting Sugeno fuzzy reasoning system is shown in Fig. 2. Here the output f = f(x,y,z) is the weighted average of the individual rules outputs and is a crisp consecutive function. In this study, x, y and z might be considered as previously recorded dam reservoir levels $L_{(t)}$, $L_{(t-1)}$,… and $L_{(t-n)}$, while the output f would represent reservoir level at the following time step, $L_{(t+1)}$. The corresponding general ANFIS architecture for a Sugeno model is illustrated in Fig. 3. The ANFIS is a network structure consisting of a number of nodes connected through directional links. The basic learning rule of ANFIS is a multilayer feed-forward network that uses neural network learning algorithms and fuzzy reasoning.

Each node has a node function that can be set or fixed parameters. The network learning or training phase is a process used to determine the parameter values sufficiently to match the training data.

A typical rule set with two fuzzy *If/Then* rules can be expressed as:

Rule 1: *If x is* A_1 *,and y is* A_3 *,and z is* A_5*; then* $f_1 = p_1x + q_1y + r_1z + c_1$

Rule 2: *If x is* A_2 *and y is* A_4 *,and z is* A_6*; then* $f_2 = p_2x + q_2y + r_2z + c_2$

…where A_1, A_2, A_3, A_4, A_5, and A_6 are linguistic labels (such as "low," "medium" or "high"), f_1 and f_2 indicate the output functions of rule 1 and rule 2, respectively, and $\{p_i, q_i, r_i, c_i\}$ are the parameters referred to as consequent parameters.

Layer 1: Each node (i) in this layer is an adaptive node that represents subscription functions defined by generalized bell functions, e.g.:

$$Z_{l,i} = \mu A_i(x) \quad \text{for } i = 1, 2, 3 \text{ or}$$
$$Z_{l,i} = \mu A_{i-3}(x) \quad \text{for } i = 4, 5, 6$$

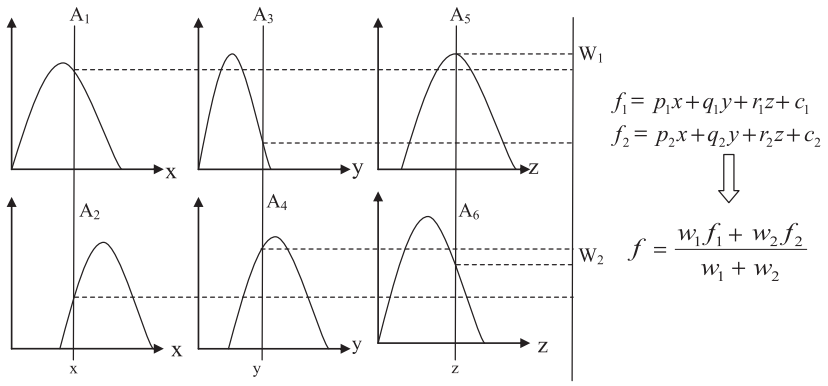

Fig. 2. Typical Sugeno fuzzy model for two inputs and with two rules.

…where x (or y) is the i^{th} node input and A_i (or A_{i-2}) is a linguistic tag associated with this node. Z_1 in the expressions indicates how quickly the blurred conglomerate will meet the quantifier of entry level *A (= A_1, A_2, A_4, A_5...A_n)* and input *x (or y, z)*. Membership functions for A_1,..,A_n are generally defined by the generalized bell functions, e.g.:

$$Z_{1,i} = \mu_1(x) = \frac{1}{1+\left|(x-c_1)/a_1\right|^{2b_1}} \quad (1)$$

…where x (or y) and {a_1, b_1 and c_1} are adaptive variables known as pre-parameters. These parameters were calibrated using the gradient descent technique during the training phase. The output of this layer is the membership value of the premise part.

Layer 2: This layer consists of nodes that multiply the incoming signals and send out the product. This output represents the most effective resistance of the rule.

For example, in Fig. 3:

$$Z_{2,1} = w_1 = \mu_1(x)\mu_3(y) \quad (2)$$

Layer 3: In this layer, the nodes calculate the ratio of the ignition power rules to the sum of the fire power of all rules.

$$Z_{3,1} = \overline{w}_1 = \frac{w_1}{w_1 + w_2 + w_3 + w_4} \quad (3)$$

Layer 4: This layer's nodes are adaptive with node functions:

$$Z_{4,1} = \overline{w}_1 f_1 = \overline{w}_1(p_1 x + q_1 y + r_1 z + c_1) \quad (4)$$

…where $\overline{w}_1$ is the output of Layer 3 and {pi,qi, ri} are the parameter sets. The parameters of this layer are called the result parameters.

Layer 5: The last fixed output of this layer is calculated as the sum of all incoming signals:

$$Z_{5,1} = f = \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 \quad (5)$$

General form:

$$Z_{5,i} = f = \sum_{i=1}^{n} \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (6)$$

An adaptive network is created that is also functionally equivalent to a Sugeno first-order fuzzy inference system. Further information about ANFIS can be found in Jang et al. [22].

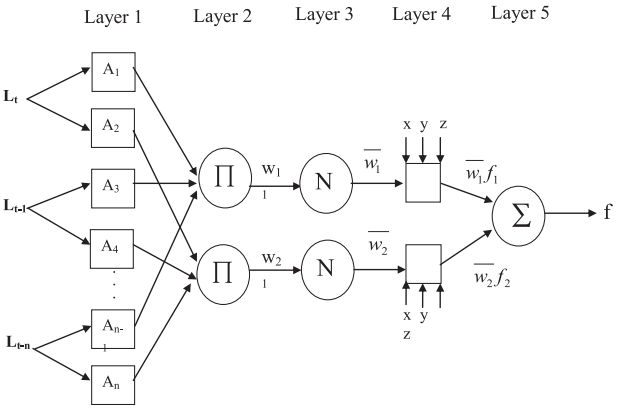

Fig. 3. ANFIS network structure.

Support Vector Machine (SVM)

SVM is a learning approach found by Cortes and Vapnik [23] for solving the classification and regression problems. Classification of variables on a plane can be conducted by drawing a boundary between them. The boundary that is drawn between variables must be as far from each variable as possible. SVM provides the opportunity to define how this boundary between a variables group can be drawn. It is based on statistical learning theory. When the training data *{(x_1, y_1),...., (x_n, y_n)}* set is considered, it specifies the input field of every $x_i \subset R^n$ samples and has a corresponding target value $y_i \subset R$ for i = 1,..., n, where n corresponds to the size of the training data.

The SVM estimating function can be expressed as:

$$y = (K_{xi} \cdot W_{jk}) + b \quad (7)$$

…where the kernel function is K_{xi}, b is bias term of SVM network and W_{jk} is called the Lagrange multipliers that determine the relative significance of the training data sets for final output. The kernel function of non-linear radial basis is:

$$K_{xi} = e^{-\gamma(x_i - y_i)} \quad (8)$$

…where γ is a user-defined parameter. The network architecture of SVM is given in Fig. 4.

Radial Basis Neural Networks (RBNN)

The RBNN were first formulated into the neural network literature by Broomhead and Lowe [24]. A radial basis function network is defined as an artificial neural network. This network uses radial basic functions as functions to enable it.

An RBNN contains a hidden layer of radial units and each hidden layer models a Gaussian response

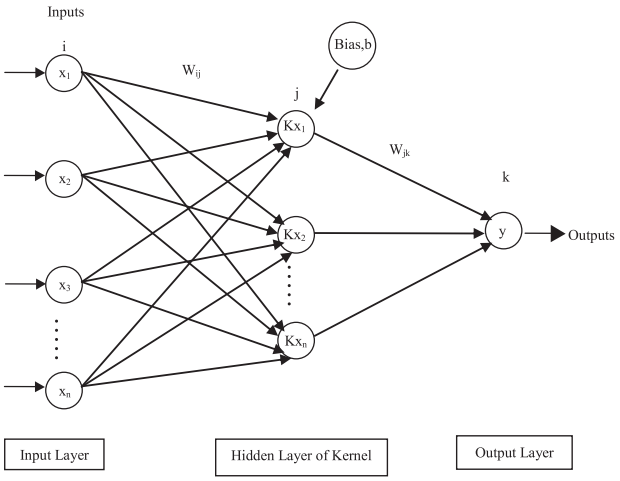

Fig. 4. Network architecture of SVM.

surface. In order to model any function in any shape, it is not compulsory to have a multiple hidden layer for the functions that are nonlinear, since modeling any function is possible when we have adequate radial units. The standard RBF has an output layer with identity activation function (see Haykin [25]). The inspiration of RBNN model came from the observation made in biological neurons having locally tuned responses. In particular, some parts of the nervous system have neurons that illustrate locally tuned response characteristics, for example, visual cortical cells sensitive to rods guided in a specific direction within a small region of the visual field (Poggio and Girosi [26]).

RBNN model theory is based on the field of interpolation of multivariate functions. The RBNN model input layer contains input data. The hidden layer transforms the data from the input field into a hidden field using a non-linear function. The hidden layer transforms the data from the input field into a hidden field using a non-linear function. The output layer is linear and responsive to the network.

$$R_d \rightarrow R \text{ with } F(x_s) = y_s \text{ for all } s = 1,..., N$$

...where F is a function of a linear space. In the RBNN approach, the interpolating function F is a linear combination of basis functions:

$$F(x) = \sum_{s=1}^{N} w_s \phi\left(\left\| x - x_s \right\|\right) + p(x) \tag{9}$$

…where $\|\cdot\|$ denotes Euclidean norm, $w_1,..., w_N$ are real numbers, ϕ a real valued function, and p is a polynomial of degree.

The interpolation problem is to reveal the real coefficients $w_1,..., w_N$ and polynomial mean $p = \Sigma^D_{l=1} a_l p_j$, where $p_1,..., p_D$ are standard stiffness of p_D and $a_1,..., a_D$ are the actual coefficients. If the interpolation problem has a unique solution for selecting any data point, this function ϕ is called the radial basis function. In some cases, the polynomial in Equation (10) can be neglected by combining this equation with Equation (11):

$$\phi w = y \tag{10}$$

…where $w = (w_1,..., w_N)$, $y = (y_1,...,y_N)$, and ϕ is an NxN matrix defined by

$$\phi = \left(\phi\left(\left\| x_k - x_s \right\|\right)\right)_{k,s=1,....,N} \tag{11}$$

The solution w of the interpolation problem can be explicitly calculated and found in the form of: $w = \phi^{-1}y$. The most popular and widely used radial basis unction is the Gaussian basis function:

$$\phi = e^{-\left(\left\| x - c_j \right\| / 2\sigma^2\right)} \tag{12}$$

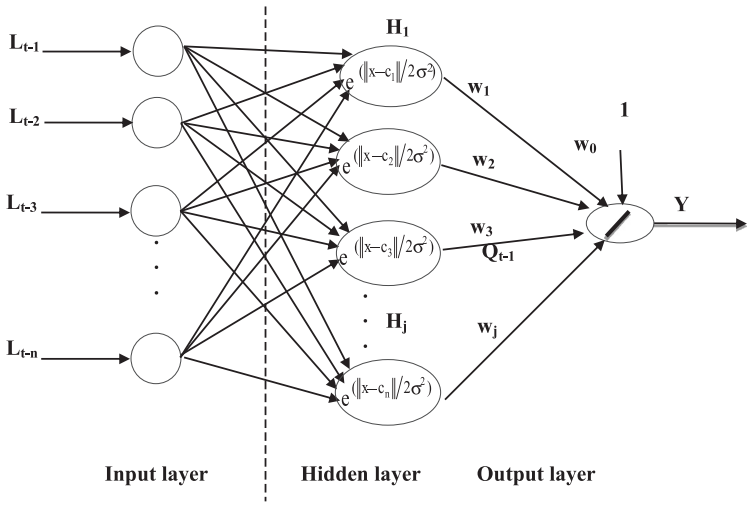

Fig. 5. Radial Basis Function Network.

…where cj (c Є Rd) is the center of the neuron in the hidden layer and ϕ is the activation function, which is a non-linear function and has many types, such as Gaussian, multiquadric, thinspline and exponential functions. The basic architecture for an RBF network is shown in Fig. 5. The input layer is simply a fan-out layer and nothing happens in this layer. The hidden layer applies a non-linear mapping from the input part to a (generally) higher dimensional part where the models become linearly separable.

Autoregressive (AR), Autoregressive Moving Average (ARMA) and Multi-Linear Regression (MLR) Models

In the prediction of the time series, good results have been obtained by researchers in the last few decades. Recently, the use of traditional methods in time series analysis has been increased by researchers. In this study, traditional methods such as Box-Jenkins (AR), autoregressive moving average (ARMA) and the autoregressive method of the multilinear regression (MLR) model have been applied for time series analysis. The time series modeling methods used have also been effective for a long time, but they are still experiencing fixed and linearity problems and only give reasonable accuracy.

An order p as an autoregressive model that can be written as AR(p) can be described as

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} \ldots\ldots\ldots + \varphi_{t-p} y_{t-p} + z_t \quad (13)$$

…where z_t is a completely random process; and $E(z_t) = 0$, $Var(z_t) = \sigma_z^2$. The parameters $\varphi_1, \ldots, \varphi_p$ are called AR coefficients. The term "autoregressive" is defined by the degeneracy of X_t over past values. In this study, AR (1) ... .AR (5) models are applied to the reservoir level data using MATLAB. The Yule-Walker equation is used to estimate the AR coefficients.

ARMA [p, q] (autoregressive moving average) models (Box and Jenkins [27]) use a weighted linear combination of previous values and shocks, which can be written as:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} \ldots\ldots\ldots + \varphi_{t-p} y_{t-p} + \ldots + \\ + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} \ldots\ldots\ldots + \theta_{t-p} a_{t-p} \quad (14)$$

…where y_t is the predicted value, a_i's are the shocks or residuals, and φ_i's and θ_i's are the associated weights. The notation [*p, q*] specifies the number of autoregressive and moving average terms in the model. The Box-Jenkins method is a method by which statistical significance can be tested. It provides a systematic recursive approach to determine the optimal number of terms and changes the weights until an optimum set of weights is found.

In this paper, models ARMA (1, 1), ARMA (3, 3) and ARMA (5, 5) have been applied to Miller Dam reservoir level data using MATLAB programming.

MLR models are used as a weighted linear combination of previous reservoir level values. Assuming that the Y-dependent variable is affected by m independent variables X_1, X_2,…,X_m, and that a linear equation is chosen for the relationship between them, the Y regression equation can be written as:

$$y = a + b_1 x_1 + b_2 x_2 + \ldots + b_m x_m + e \quad (15)$$

…where y is a linear combination of the parameters, (b_m)s are the constants of regression equation, e is the residual term, and the subscript *i* indexes a particular observation.

The training and testing data groups used for the models are also used for MLR models. For each model, mean square error (MSE), mean absolute error (MAE) between model estimations and the observed values are computed as follows:

$$MSE = \frac{1}{N}\left(\sum_{i=1}^{N} Yi_{observed} - Yi_{forecast}\right)^2 \quad (16)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|Yi_{observed} - Yi_{forecast}\right| \quad (17)$$

…where N and Y_i denote the number of data sets and dam reservoir level, respectively.

Results and Discussion

Predicting reservoir level variations by ANFIS, RBNN, SVM, MLR, AR and ARMA models is the aim of this study. Based on the partial autocorrelation function analysis, several input combinations were used to estimate reservoir levels (Fig. 6). L_t represents the lake level at time interval t. Input combinations show the last recorded daily delayed lake levels (L_{t-5}, L_{t-4}, L_{t-3}, L_{t-2}, and L_{t-1}).

With autocorrelation and partial autocorrelation statistics, we estimate daily reservoir level time series at lag 0 to lag 14 intervals and 95% confidence interval (Fig. 6). The autocorrelation function shows a correlation between the pairs of all points of the reservoir-level time series.

As shown from Fig. 6, the lake-level values are moderately auto correlated. The slower water exchange in a dam reservoir can lead to inertia of the water levels. This produces a low-frequency oscillating component. A partial auto-correlation refers to autocorrelation of a series with itself under steady-state conditions, while controlling the effect of intermittent delays. The autocorrelation of a series is spontaneously indicated by partial autocorrelation without the mixed effects of delayed autocorrelation (Mutlu et al., [28]). The partial autocorrelation function

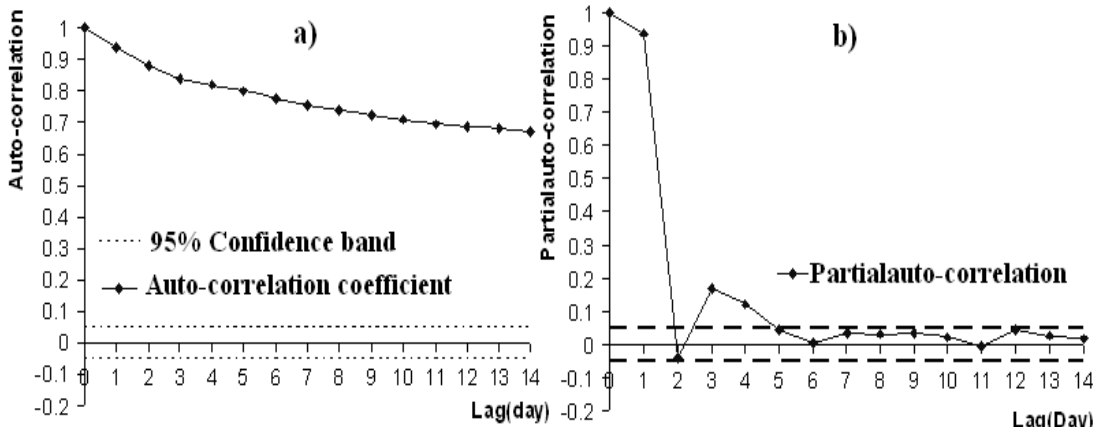

Fig. 6. a) Autocorrelation and b) partial auto-correlation functions of daily reservoir levels.

describes a significant correlation up to delay 5, and falls within confidence bounds for this reason. The pattern in partial autocorrelation decreases rapidly, confirming autoregressive process dominance. Partial autocorrelation coefficients suggest that the delay level should be included for up to 5 days.

The partial autocorrelation coefficients suggest the incorporation of lake-level values of up to 5 days lag in input vector to the employed models (Fig. 6). The correlation analysis was implemented to evaluate the degree of effect of each variable and to select the most effective input vectors. The following combinations of input data were evaluated: (i) L_{t-1}, *(ii)* $L_{t-1,}$ L_{t-2}, and L_{t-3}, (iii) L_{t-1} , L_{t-3}, and L_{t-5}. The output layer has only one neuron, and the reservoir level L_t represents the current day in all models.

Predicting Reservoir Level Fluctuations of Millers Dam

According to the partial autocorrelation function, the 5 lag times have a significant influence on predicting levels for the following time step (L_{t-1}). Therefore, it seems that up to 5 lag-time reservoir levels are necessary

Table 2. Comparison of different models with MSE, MAE and R for testing period..

Models	Input combination	MSE(m^2)	MAE(m)	R
ANFIS1	L_{t-1}	0.0031	0.0404	0.898
ANFIS 3	L_{t-3}, L_{t-2}, and L_{t-1}	0.0030	0.0397	0.903
ANFIS 5	L_{t-5}, L_{t-4}, L_{t-3}, L_{t-2}, and L_{t-1}	0.0028	0.0396	0.909
RBNN1	L_{t-1}	0.0039	0.0453	0.868
RBNN 3	L_{t-3}, L_{t-2}, and L_{t-1}	0.0036	0.0433	0.879
RBNN 5	L_{t-5}, L_{t-4}, L_{t-3}, L_{t-2}, and L_{t-1}	0.0035	0.0431	0.880
SVM 5	L_{t-5}, L_{t-4}, L_{t-3}, L_{t-2}, and L_{t-1}	0.0784	0.0564	0.810
GRNN 5	L_{t-5}, L_{t-4}, L_{t-3}, L_{t-2}, and L_{t-1}	0.0040	0.0467	0.865
ANN5	L_{t-5}, L_{t-4}, L_{t-3}, L_{t-2}, and L_{t-1}	0.0032	0.0415	0.893
MLR1	L_{t-1}	0.0068	0.0602	0.792
MLR3	L_{t-3}, L_{t-2}, and L_{t-1}	0.0061	0.0564	0.797
MLR5	L_{t-5}, L_{t-4}, L_{t-3}, L_{t-2}, and L_{t-1}	0.0060	0.0567	0.800
AR(1)	L_{t-1}	0.0069	0.0602	0.792
AR(3)	L_{t-3}, L_{t-2}, and L_{t-1}	0.0063	0.0573	0.796
AR(5)	L_{t-5}, L_{t-4}, L_{t-3}, L_{t-2}, and L_{t-1}	0.0060	0.0564	0.799
ARMA(1-1)	L_{t-1}	0.0066	0.0587	0.792
ARMA(3-3)	L_{t-3}, L_{t-2}, and L_{t-1}	0.0061	0.0565	0.798
ARMA(5-5)	L_{t-5}, L_{t-4}, L_{t-3}, L_{t-2}, and L_{t-1}	0.0060	0.0566	0.799

for the task at hand. The ANFIS, RBNN, SVM, GRNN, ANN (Üneş at al. 2015), ARMA, AR and MLR model results of daily dam reservoir level predictions with various and optimal input combination for the testing periods can be seen in Table 2. Each input combination was applied for introducing the models to produce good predictions. The table clearly shows that the ANFIS 5 model with the input combination L_{t-5}, L_{t-4}, $L_{t-3,}$ L_{t-2} and L_{t-1} gives the best results. Table 2 also shows that the ANFIS 5 model performs better than the other models in terms of the R, MSE and MAE statistics in the test period.

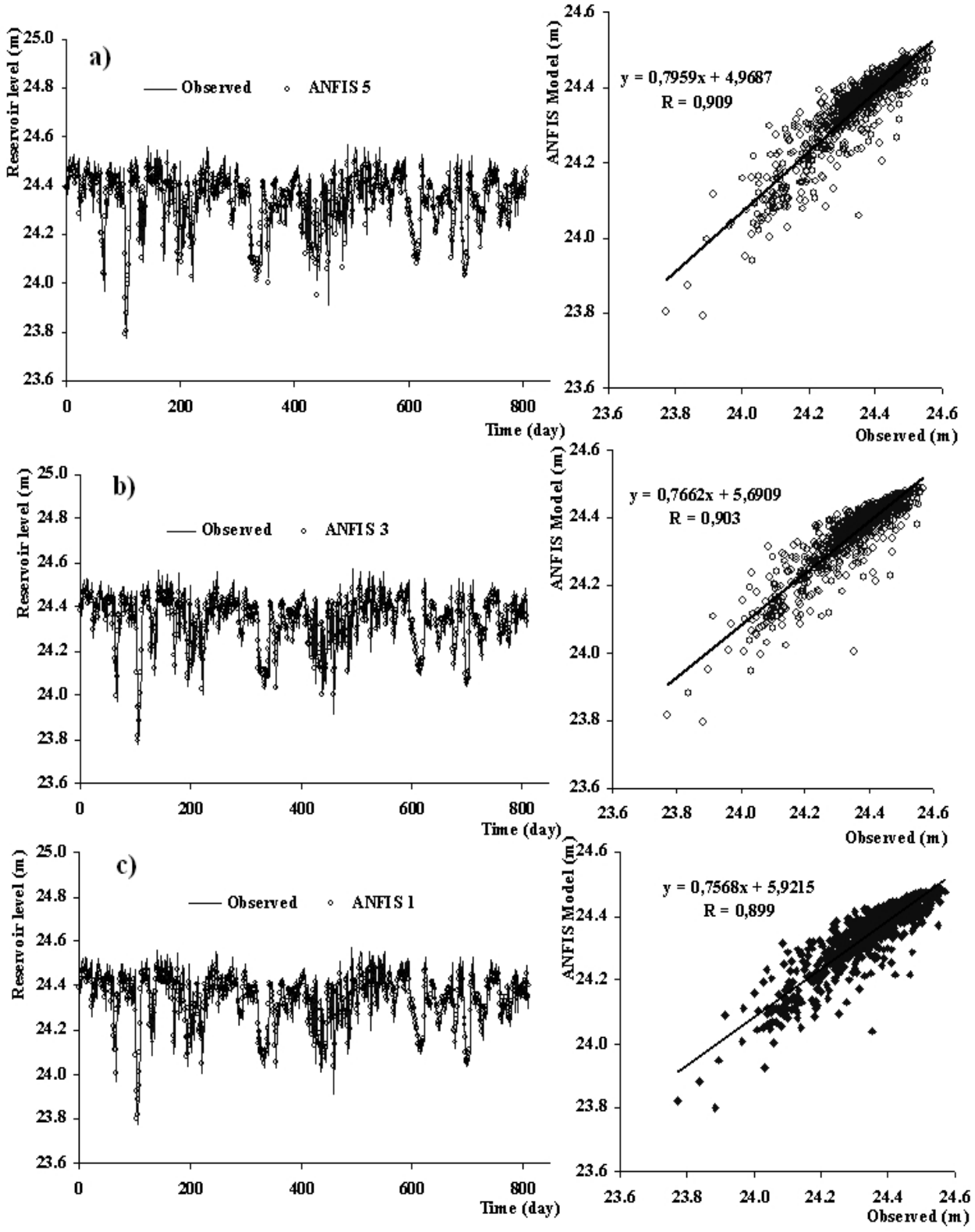

Fig. 7. Observed and predicted reservoir levels for Millers Dam in the test period: a) ANFIS5, b) ANFIS3, and c) ANFIS1.

Increasing the accuracy of the model while adding more input variables does not significantly affect predictive accuracy.

MSE and MAE values of the MLR model for various lead times during the test period are given in Table 2. Using L_{t-1} as the only input variable resulted in less accurate simulations with MSE = 0.0068 m^2 and MAE = 0.0602 m, while the input combination of L_{t-5}, L_{t-3}, and L_{t-1} give the best results with MSE = 0.0060 m^2 and MAE = 0.0567 m compared to the others. Thus, the MLR application shows that reservoir water levels up to L_{t-5} significantly increase

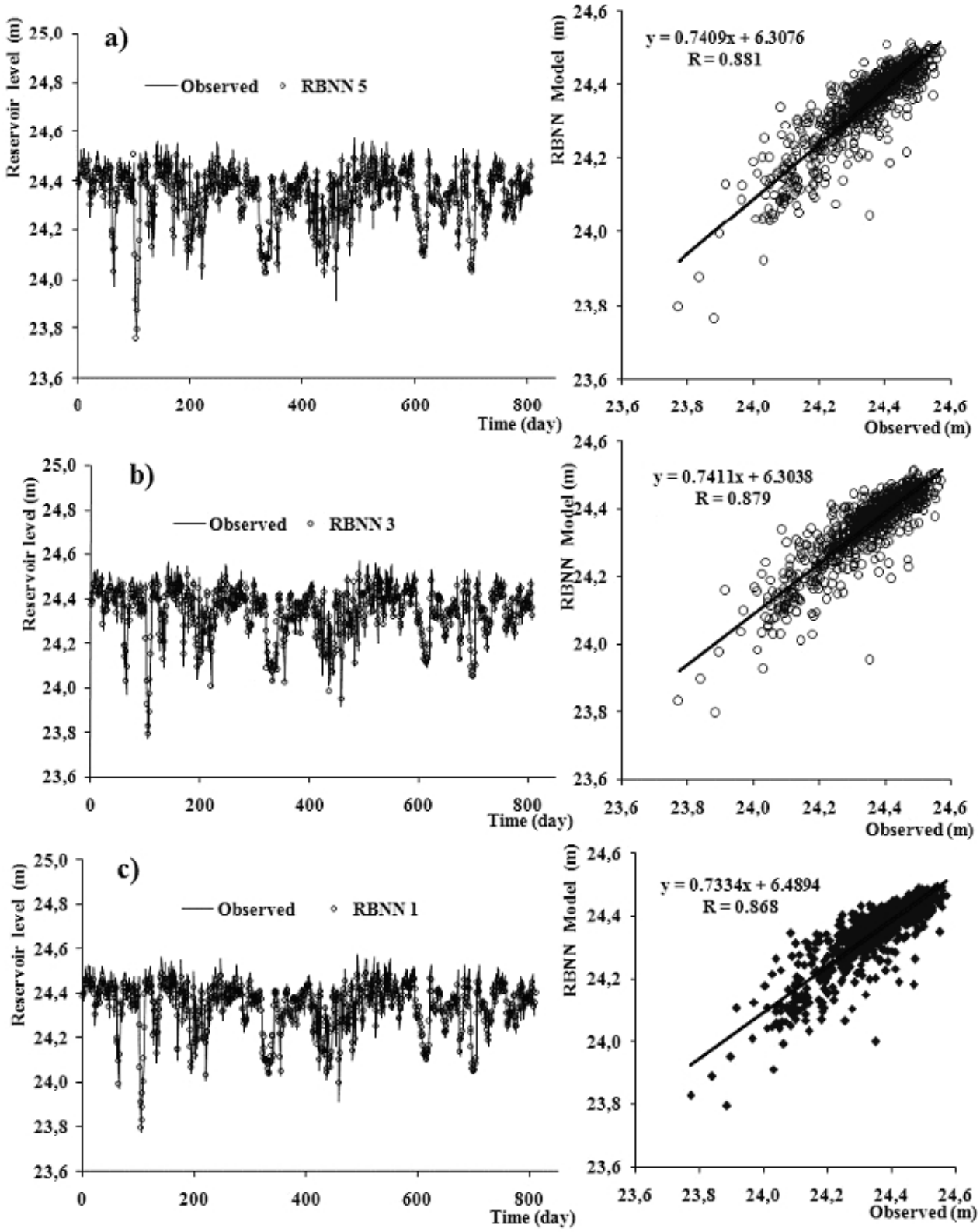

Fig. 8. Observed and predicted reservoir levels for Millers Dam in the test period: a) RBNN5, b) RBNN3, and c) RBNN1.

modeling accuracy, while adding more variables does not noticeably affect prediction accuracy. The overall evaluation of the other three linear methods reveals that the ARMA5, AR5 and MLR5 models give similar accuracy. The similar accuracy of the ARMA5, AR5 and MLR5 models can be clearly seen from Table 2.

The back-propagation ANFIS network described in the preceding section was applied in MATLAB code for forecasting reservoir levels using the previously recorded daily reservoir level values. Similarly, a different number of input combinations is applied for real observed data sets to find out how well the models perform. The ANFIS network's training was stopped after 500 epochs. The testing statistics of ANFIS models in terms of R, MAE and MSE are presented in Table 2. Similar to the ANFIS model, the RBNN model was solved by writing a MATLAB code. In the model, mean squared error goal and maximum number of neurons were taken as 0.005 and 15, respectively. The testing statistics of RBNN models in terms of R, MAE and MSE are also given in Table 2.

Fig. 7 shows model performance of the ANFIS. As seen in Fig. 7, the ANFIS 5 model performs better than the other ANFIS models in terms of R in the test period. The ANFIS 5 model also has less scattered predictions than the other models and provides the highest R coefficient (0.909) for the L_{t-5}, L_{t-4}, L_{t-3}, L_{t-2} and L_{t-1} input combination. As seen from Table 2, ANFIS models perform better than the other models with the evaluation of all methods. The test results also show that increasing input combination leads to a little increase in model accuracy. The R increases from 0.898 to 0.909 for the ANFIS and from 0.792 to 0.881 for the other models (Fig. 7).

Fig. 8 shows that the RBNN 5 model performs better than the other RBNN models in terms of the R, MSE

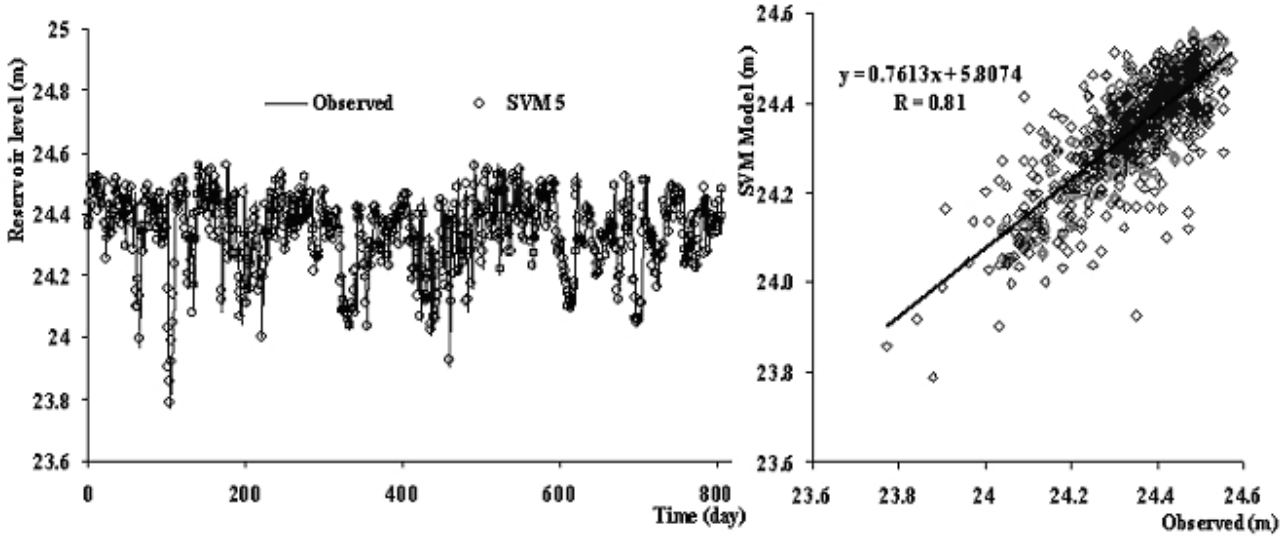

Fig. 9. Observed and predicted reservoir levels for Millers Dam in the test period – SVM 5.

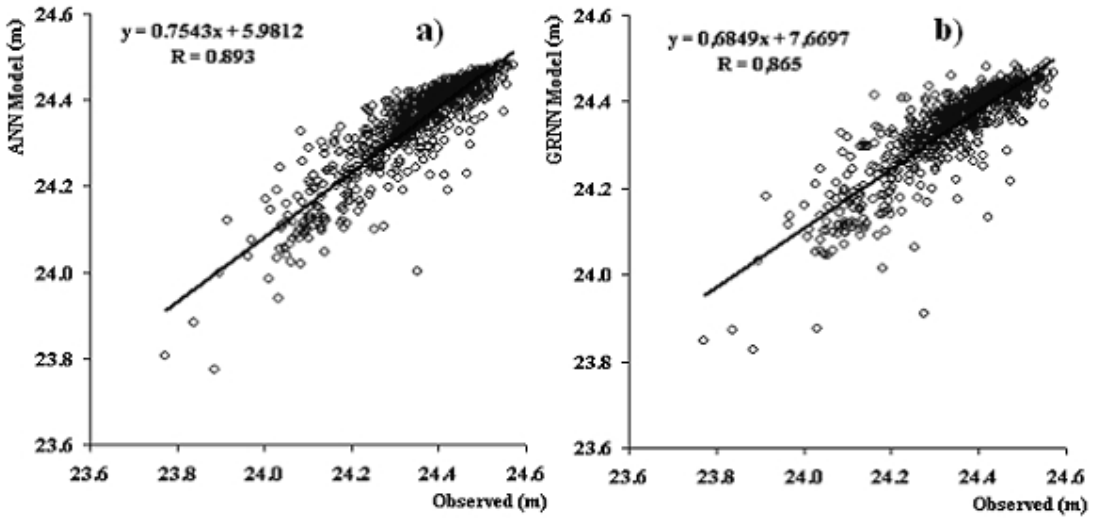

Fig. 10. Scatter graphs for Millers Dam in the test period for: a) ANN (Üneş at all 2015) and b) GRNN.

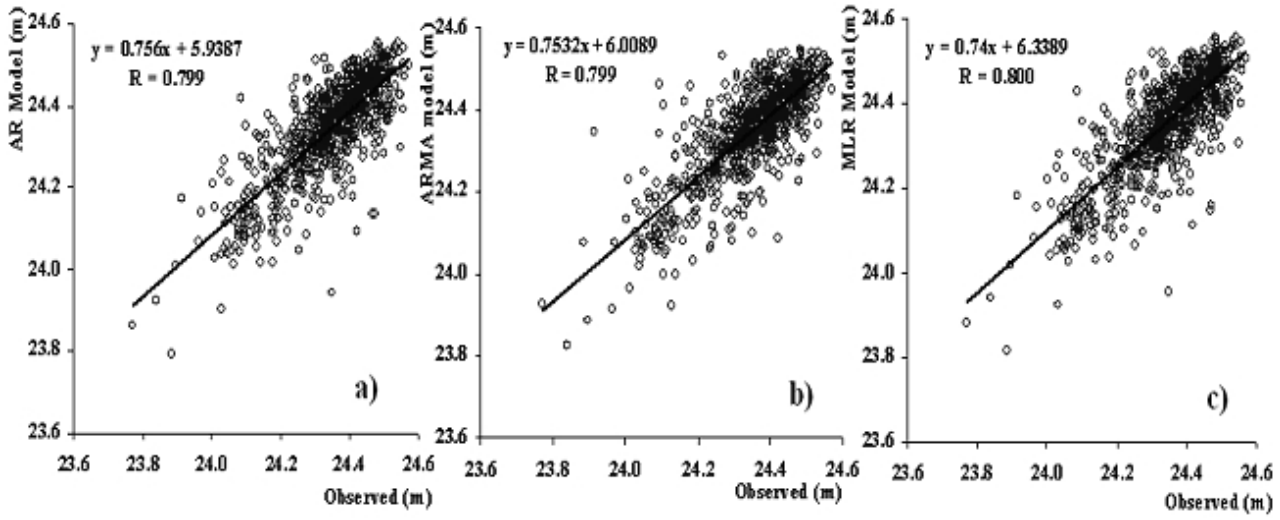

Fig. 11. Scatter graphs for Millers Dam in the test period for: a) AR5, b) ARMA5, and c) MLR5 model.

and MAE statistics in the test period. Fig. 8 also shows that the RBNN 5 model has less scattered predictions than the other RBNN models and provides the highest R coefficient (0.880) for the L_{t-5}, L_{t-4}, $L_{t-3,}$ L_{t-2} and L_{t-1} input combination.

Correlation coefficient R = 0.81 was obtained for the test from the SVM 5-5 (L_{t-5}, L_{t-4}, $L_{t-3,}$ L_{t-2} and L_{t-1} input combination) model. It is shown in Fig. 9 that the SVM estimates at the test phase show worse results than the ANFIS and RBNN model results, and they show better results than the AR, ARMA and MLR model results for the observed daily lake levels.

AR and ARMA models are implemented in this study and the application of the AR(5) and ARMA(5-5) models gave similar low R coefficient (0.792-0.799) for the test period (Table 2). The previous ANN and GRNN model investigations for the test period are also compared with the observed reservoir levels using the scatter plots for Miller Dam Reservoir. Such scatter graphs are given for ANN (Ünes at all. [4]) and GRNN models in Fig. 10 and AR(5), ARMA(5-5) and MLR(5) models in Fig. 11. These scatter plots of autoregressive models provide a low agreement between observation and prediction reservoir levels. However, AR and ARMA models have the smallest R coefficient (0.792-0.799) compared to the other model results, as seen in Table 2.

The forecasting of different input combinations for 1-day ahead displays that the model clearly increases with different input combinations. Fig. 7 shows the scatter plots of the observed and simulated dam reservoir levels during the ANFIS 5 test periods. As seen from Table 2, the ANFIS 5 model has the smallest MSE (0.0028 m^2), MAE (0.0396 m) and the highest R (0.909) for five-input combination during the test period. According to all model results, ANFIS picks out as having very small MSE, MAE and high R values for the same input combination.

Conclusions

In this study, dam reservoir level variations are predicted using artificial and statistical techniques for Millers Ferry Dam in the USA. Neuro fuzzy and radial base models were developed for predicting reservoir levels and then compared with SVM, GRNN, ANN, models and conventional models ARMA, AR and MLR. The following conclusions can be reached from the study.

The ANFIS model gives better estimates of the reservoir level fluctuations than the RBNN, ANN, SVM, and GRNN models and conventional models. AR, ARMA and MLR models did not reach the desired accuracy for the Millers Ferry Dam case study (in the problem) and it cannot provide a good prediction for reservoir levels. These models could not reach the intended accuracy due to the problem of nonlinearity in the dam reservoir level.

However, the ANFIS model adapts to the changing input conditions. The advantages of the model are that fluctuations in the reservoir level can be explained by specifying that the model structures include nonlinear dynamics of the entire data set.

The models and methods have no limitations in the form of fixed assumptions or formal constraints.

It is seen from the results that ANFIS is a beneficial alternative method for dam reservoir level prediction. The reservoir level estimations can be quite informative for determining the periodic water supply strategies, the hydroelectric energy computations and flood management studies.

Therefore, an ANFIS model can be developed for a specific region, and can be quite useful in the water resource management studies, such as dam reservoir level, habitat in dam reservoirs, water quality modeling, management and studies.

Acknowledgements

The data used in this study were downloaded from the web server of the USGS. The authors wish to thank the staff of the USGS who are associated with data observation, processing, and management of USGS web sites.

Conflict of Interest

The authors declare no conflict of interest.

References

1. TAŞAR B., KAYA Y.Z., VARÇIN H., ÜNEŞ F., DEMIRCI M. Forecasting of suspended sediment in rivers using artificial neural networks approach. Int. Eng. J. Sci. J. **4**, 79, **2017**.
2. DEMIRCI M., ÜNEŞ F., SAYDEMIR S. Suspended sediment estimation using an artificial intelligence approach. Sediment matters. Eds. P. Heininger, J. Cullmann. Springer International Publishing, 83, **2015**.
3. TAŞAR B., ÜNEŞ F., DEMIRCI. M., KAYA Y.Z. Forecasting of Daily Evaporation Amounts Using Artificial Neural Networks. Journal of Dicle University Engineering **9** (1), 543, **2018.**
4. DEMIRCI M., UNES F., AKOZ M.S. Prediction of cross-shore sandbar volumes using neural network approach. Journal of Marine Science and Technology. **20** (1), 171, **2015**. doi:10.1007/s00773-014-0279-9
5. ÜNEŞ F., DEMIRCI M., ISPIR E., KAYA Y. Z., MAMAK M., TASAR B. Estimation of Groundwater Level Using Artificial Neural Networks: a Case Study of Hatay-Turkey. Environmental Engineering 10th International Conference, **2017** DOI: https://doi.org/10.3846/enviro.2017.092
6. ZAJI A.H., BONAKDARI H. Robustness lake water level prediction using the search heuristic-based artificial intelligence methods. ISH Journal of Hydraulic Engineering, **2018**, doi.org/10.1080/09715010.2018.1424568.
7. NITSURE S.P., LONDHE S.N., KHARE K.C. Prediction of sea water levels using wind information and soft computing techniques. Applied Ocean Research, **47**, 344, **2014**.
8. UNES F. Dam reservoir level modeling by neural network approach: A case study. Neural Network World **4** (10), 461, **2010**.
9. UNES F., DEMIRCI M., KISI O. Prediction of Millers Ferry dam reservoir level in USA using artificial neural network. Periodica Polytechnica Civil Engineering, **59** (3), 309, **2015**.
10. PARMAR K.S., BHARDWAJ R. River water prediction modeling using neural networks, fuzzy and wavelet coupled model. Water Resources Management, **29**, 17, **2015**.
11. YANPING Y., WENXI L., DEFA H., HONGCHAO L., LEI Z. Application of dual-response surface methodology and radial basis function artificial neural network on surrogate model of the groundwater flow numerical simulation. Polish Journal of Environmental Studies, **26** (4), 1835, **2017**.
12. WIKTOR H., DARIUSZ M., MAREK R., EDYTA K., ARTUR R.P. Applying an artificial neural network (ANN) to assess soil salinity and temperature variability in agricultural areas of a mountain catchment. Polish Journal of Environmental Studies, **26** (6), 2545, **2017**.
13. SHIRI J., SHAMSHIRBAND S., KISI Ö., KARIMI S., BATENI S. M., NEZHAD S.H.H., HASHEMI A. Prediction of water-level in the Urmia Lake using the extreme learning machine approach. Water Resources Management, **30**, 5217, **2016**.
14. EMAMGHOLIZADEH S., MOSLEMI K., KARAMI G. Prediction the groundwater level of Bastam Plain (Iran) by artificial neural network (ANN) and adaptive neuro-fuzzy ınference system (ANFIS). Water Resources Management, **28**, 5433, **2014**.
15. SUN W., TREVOR B. Multiple model combination methods for annual maximum water level prediction during river ice breakup. Hydrological Processes, **3** (2), 421, **2018**.
16. UNES F. Prediction of density flow plunging depth in dam reservoir: An artificial neural network approach. CLEAN - Soil Air Water, **38** (3), 296, **2010**.
17. HIMAN S., BAKHTYAR A.A., BAHARIN B.A., MOHAMMAD J.T.A., SOROUSH K., SAEED E. Assessment of WLC and Fuzzy Logic methods for site selection of water reservoirs in Malaysia. Polish Journal of Environmental Studies, **25** (3), 1223, **2016**.
18. DEMIRCI M., BALTACI A. Prediction of suspended sediment in river using fuzzy logic and multilinear regression approaches. Neural Computing and Applications, **23**(1), 145, **2013**.
19. SEO Y., KIM S., KISI O., SİNGH V.P. Daily water level forecasting using wavelet decomposition and artificial intelligence techniques. Journal of Hydrology **520**, 224, **2015**.
20. PIRI J., KAHKHA M.R.R. Prediction of water level fluctuations of Chahnimeh reservoirs in Zabol using ANN, ANFIS and Cuckoo optimization algorithm. Iranian Journal of Health, Safety and Environment **4** (2), 706, **2016**.
21. JANG J.S.R. ANFIS: Adaptive-Network-Based Fuzzy Inference System, IEEE Transactıons On Systems, Man, And Cybernetics **23** (3), 665, **1993**.
22. JANG J.S.R., SUN C.T., MIZUTAN E. Neuro-Fuzzy and Soft Computing : A Computational Approach to Learning and Machine Intelligence, New Jersey, **1997**.
23. CORTES C., VAPNIK V. Support vector networks. Machine Learning **20**, 1, **1995**.
24. BROOMHEAD D.S., LOWE D. Radial Basis Functions, Multi-Variable Functional Interpolation and Adaptive Networks, London, **1988**.
25. HAYKIN S. Neural Networks. A Comprehensive Foundation, Prentice-Hall, Upper Saddle River, New Jersey, **1999**.
26. POGGIO T., GIROSI F. Networks for approximation and learning. Proceedings of the IEEE **78** (9), 1481, **1990**.
27. BOX G.P., JENKINS G. Time Series Analysis, Forecasting and Control, Holden Day, San Francisco, **1970**.
28. MUTLU E., CHAUBEY I., HEXMOOR H., BAJWA S.G. Comparison of artificial neural network models for hydrologic predictions at multiple gauging stations in an agricultural watershed. Hydrological Processes **22** (26), 5097, **2008**.