CrossMark

ORIGINAL PAPER

# Utilization of PSO algorithm in estimation of water level change of Lake Beysehir

**Meral Buyukyildiz**[1] · **Gulay Tezel**[2]

**Abstract** In this study, unlike backpropagation algorithm which gets local best solutions, the usefulness of particle swarm optimization (PSO) algorithm, a population-based optimization technique with a global search feature, inspired by the behavior of bird flocks, in determination of parameters of support vector machines (SVM) and adaptive network-based fuzzy inference system (ANFIS) methods was investigated. For this purpose, the performances of hybrid PSO-$\varepsilon$ support vector regression (PSO-$\varepsilon$SVR) and PSO-ANFIS models were studied to estimate water level change of Lake Beysehir in Turkey. The change in water level was also estimated using generalized regression neural network (GRNN) method, an iterative training procedure. Root mean square error (RMSE), mean absolute error (MAE), and coefficient of determination ($R^2$) were used to compare the obtained results. Efforts were made to estimate water level change (L) using different input combinations of monthly inflow-lost flow (I), precipitation (P), evaporation (E), and outflow (O). According to the obtained results, the other methods except PSO-ANN generally showed significantly similar performances to each other. PSO-$\varepsilon$SVR method with the values of minMAE = 0.0052 m, maxMAE = 0.04 m, and medianMAE = 0.0198 m; minRMSE = 0.0070 m, maxRMSE = 0.0518 m, and medianRMSE = 0.0241 m; min$R^2$ = 0.9169, max$R^2$ = 0.9995, median$R^2$ = 0.9909 for the I-P-E-O combination in testing period became superior in forecasting water level change of Lake Beysehir than the other methods. PSO-ANN models were the least successful models in all combinations.

## 1 Introduction

One of the most important components of the water resources is a lake, which is highly vulnerable to external influences, and basins surrounding it. Lake water levels are affected by climate change and human activity. Water used for precipitation, flow, evaporation, supply of drinking water, irrigation and energy leads to an increase or decrease in lake water level. Increased or decreased lake water level causes huge economic losses in the lake environments and even sometimes irreversible changes in nature (Cengiz and Kahya 2006).

Recognition of changes in lake level is of vital importance for planning and construction works, design works in lakes and the areas surrounding them, ensuring the management and productivity of agricultural lands, sustaining existing or acquired aquatic life, and the operation of wetlands which are of special significance and protected by various international treaties and drinking water facilities.

Lake level changes can be observed regularly by monitoring stations to be built at certain places or estimated using mathematical methods, such as a water balance model. Nevertheless, a variety of artificial intelligence methods, including artificial neural networks (ANN), support vector machines (SVM), adaptive network-based fuzzy inference system (ANFIS), fuzzy logic (FL), and wavelet transform (WT), which have been recently employed to estimate many hydrological parameters, such as precipitation (Ramana et al. 2013), temperature (Chithra et al. 2014), sediment (Goyal 2014), flow (Kisi et al. 2012a), are also commonly used to estimate

✉ Meral Buyukyildiz
meralbyildiz@selcuk.edu.tr

Gulay Tezel
gtezel@selcuk.edu.tr

1 Department of Civil Engineering, Engineering Faculty, Selçuk University, Konya, Turkey

2 Department of Computer Engineering, Engineering Faculty, Selçuk University, Konya, Turkey

🰀 Springer

lake water level (Altunkaynak 2007, 2014; Ondimu and Murase 2007; Chang et al. 2014; Karimi et al. 2013; Seo et al. 2015).

De Domenico et al. (2013) compared the success of chaos theory with that of the auto-regressive integrated moving average (ARIMA) model in estimating sea water level for daily, weekly, 10-day, and monthly time scales at the Cocos (Keeling) Islands. In conclusion, an acceptable performance was achieved in both models; however, the chaos theory model was more successful than the ARIMA model in forecasting. Altunkaynak (2007) used ANN and auto-regressive moving average (ARMA) models to estimate water level and determined that ANN models performed better than ARMA models. Khan and Coulibaly (2006) investigated the ability of SVM method in estimating the water level of Lake Erie and compared the results with those of multilayer perceptron (MLP) and seasonal auto-regressive (SAR) models. Kavehkar et al. (2011) used genetic programming (GP) model to estimate the change in the water level of Lake Urmia in Iran and compared with those of the ANN. Cimen and Kisi (2009) studied the potential of SVM and ANN techniques in modeling lake level fluctuations and found that the SVM model performed better than the ANN. Wei (2012) investigated the performance of wavelet SVMs and classical Gaussian SVMs for forecasting the hourly water levels. He found that wavelet SVMs outperform the Gaussian SVMs for water level predictions. Khatibi et al. (2014) investigated the performance of the models, including seasonal auto-regressive integrated moving average (SARIMA), ANN, gene expression programming (GEP), multiple linear regression (MLR), and nonlinear local prediction (NLP), to estimate water levels of six lakes with different physical characteristics. Kisi et al. (2012b) studied the usability of GEP, ANN, ARMA, and ANFIS techniques to estimate the daily water level of Lake Iznik, located in west of Turkey.

Other than the methods such as ANN, ANFIS, SVM, wavelet, fuzzy logic, scientists have recently developed the evolutionary optimization methods inspired by nature which include genetic algorithm (GA), particle warm optimization (PSO), ant colony optimization (ACO), artificial bee colony (ABC), memetic algorithm (MA), differential evolution (DE), artificial immune systems (AIS). These optimization algorithms developed on the basis of events in nature are called heuristic methods (Karaboga 2004). Disadvantages of conventional techniques such as early convergence, local minima, and high computational complexity have led to the use of heuristic methods in training artificial neural networks (Kennedy and Eberhart 1995). Hybrid models have been developed by using heuristic methods also in the determination of the parameters of methods such as SVM, ANFIS, etc. PSO, one of these heuristic methods, is widely used to estimate hydrological parameters, particularly to train ANN; however, the number of studies on determination of SVM parameters is

more limited. In the literature, there are a few examples of the use of PSO in determining ANFIS parameters for the prediction of hydrological data (Pousinho et al. 2011).

Chau (2006) developed PSO-based artificial neural network (PSO-ANN) approach to estimate the water level of Shing Mun River in Hong Kong using PSO to train multilayer perceptrons. Mohandes (2012) employed PSO-ANN model to predict global solar radiation using data from 41 stations in Saudi Arabia and compared with those of the BP algorithm. Sedki and Ouazar (2010) used PSO-ANN algorithm to estimate daily flows of a basin in a semiarid region of Morocco, while Tapoglou et al. (2014) used it to simulate the hydraulic head changes in an observation well in the region of Agia, Chania, Greece.

García Nieto et al. (2014) developed hybrid models using PSO and SVM methods for long-term estimation of turbidity in the Nalon river basin in the north of Spain. Sudheer et al. (2014) studied estimation performance of PSO-SVM model using monthly streamflow data from St. Regis River in the nearby Clark Fork and Swan River in the nearby Bigfork and compared with those of the ANN and ARMA models. Zhao and Wang (2010) used the PSO algorithm to determine the parameters of the SVR model to rainfall forecasting model using monthly rainfall data in Guangxi, China, during 1954–2008. They compared the results of PSO-SVR model with conventional SVR method.

Pousinho et al. (2011) employed the hybrid PSO-ANFIS approach to predict short-term wind power in Portugal and compared with those of the ARIMA, neural network (NN), NN combination with wavelet transform (NNWT), and wavelet neuro fuzzy (WNF).

Considering that lake water is usually fed by precipitation and streamflow and water loss is dependent on the amount of water drawn from lakes and evaporation, it would be an important step to establish the correlation between the change in lake water level and such hydrometeorological variables (Yurtcu 2006).

The aim of this study was to determine the usefulness of fast convergent PSO in the solution of nonlinear problems in prediction of optimal values of the parameters of SVM and ANFIS methods. The change in water level of Lake Beysehir, the largest freshwater lake in Turkey, was estimated using hybrid PSO models. The performance of the hybrid PSO models was compared with the GRNN. All models with various input structures were constructed, and the best model was determined using different performance criteria.

## 2 Study area and data.

Lake Beysehir (37° 40′ 54″ N, 31° 43′ 22″ E) is the largest freshwater lake and the third largest lake by

surface area in Turkey. It is situated on the west end of Konya-Çumra closed sub-basin, which comprises the southwest of Konya closed basin, the largest closed basin in Anatolia, and it is 90 km away from the province of Konya. Lake Beysehir is within the borders of Konya and Isparta provinces. It is currently 1121.93 m above sea level; its surface area is approximately 650 km$^2$, and rainfall area is 4086.4 km$^2$. Although a variety of numbers was reported, the lake is approximately 8.5 m deep, 45 km long, and 10–25 km wide and has a circumference of 120 km. Lake Beysehir was opened for operation in 1914, and it is currently utilized for the purposes of supplying drinking water, irrigation, fishing, various commercial business, and tourism. Figure 1 shows Lake Beysehir and its location in Turkey.

In this study, monthly inflow-lost flow (I), precipitation (P), evaporation (E), and outflow (O) values determined by the 4th Regional Directorate of State Hydraulic Works (DSI) were used as inputs in order to estimate the monthly change in Lake Beysehir water level (L). Three hundred forty-eight-month data in total covering the period from 1962 to 1990 were used, and the first 276-month data (1962–1984) were used in training process, and the remaining 72-month data (1985–1990), in testing process. Here, level values were given as elevation so the values of change in level per month were calculated by taking the differences in elevation compared to the previous month.

## 3 Methodology.

### 3.1 Generalized regression neural network

GRNN, developed by Specht (1991) as an alternative to feed forward backpropagation algorithms, does not require an iterative training procedure.

GRNN consists of four layers, including the input layer, the hidden layer, the summation layer, and the output layer. If $f(x,y)$, the common probability intensity

function, is known, regression of dependent $y$ variable with respect to independent $x$ variable is expressed by Eq. (1).

$$E[y|X] = \frac{\int_{-\infty}^{+\infty} yf(X,y)dy}{\int_{-\infty}^{+\infty} f(X,y)dy} \tag{1}$$

Otherwise, $f(x,y)$ function is estimated using Eq. (2) by utilizing observed $X^i$ and $Y^i$ values.

$$\widehat{f}(X,Y) = \frac{1}{(2\pi)^{(p+1)/2}s^{(p+1)}} \tag{2}$$

$$\times \frac{1}{n}\sum_{i=1}^{n}\exp\left[-\frac{(X-X^i)^T(X-X^i)}{2s^2}\right] \times \exp\left[-\frac{(Y-Y^i)^2}{2s^2}\right]$$

where $n$ indicates the number of observed data; $s$, the spread value; and $p$, the dimension of vector $x$. $D_i^2$ is a scalar function and defined as

$$D_i^2 = (X-X^i)^T(X-X^i) \tag{3}$$

Equation (4) is obtained by solving the integrals in Eq. 1 (Alp and Cigizoglu 2004).

$$\widehat{Y}(X) = \frac{\sum_{i=1}^{n} Y^i \exp\left(-\frac{D_i^2}{2s^2}\right)}{\sum_{i=1}^{n} \exp\left(-\frac{D_i^2}{2s^2}\right)} \tag{4}$$

### 3.2 Support vector regression

SVM was proposed by Vapnik for the first time to solve classification and regression type problems (Vapnik 1995). When SVM was used for regression, it is called as support vector

Fig. 1 Lake Beysehir and its location in Turkey
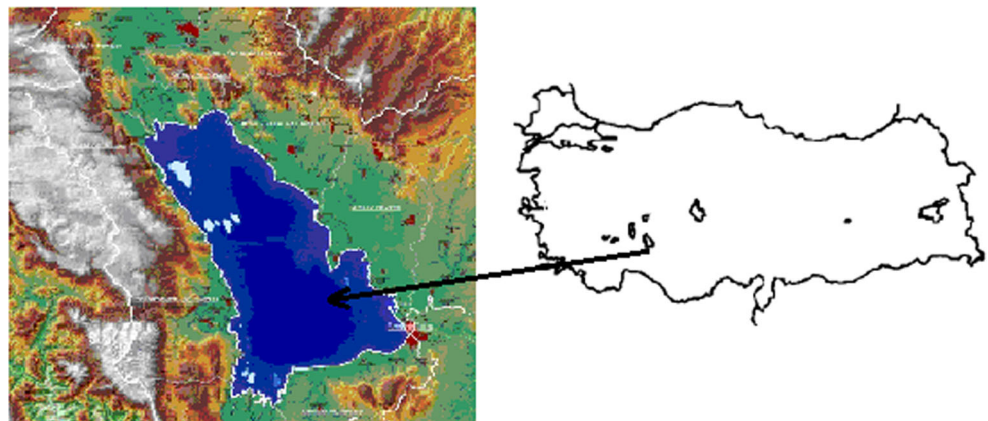
regression (SVR). Traditional machine learning methods involve several problems, including the demand for numerous training data, low convergence rate, local minima, and overfitting/underfitting (Lu et al. 2002). SVM overcomes these problems by operating on the basis of structural risk minimization (Shen et al. 2004). SVM's output function is obtained as Salat and Osowski (2004).

$$f(x) = \sum_{i=1}^{N} \left(\alpha_i - \alpha_i^*\right) K(x,z) + b_i \qquad (5)$$

where $\left(\alpha^i - \alpha_i^*\right)$ indicates Lagrange factors, $K(x,z)$, the inner product core function defined in accordance with the Mercer theorem and $b_i$, the bias value. A detailed description of SVR is available in Buyukyildiz et al. (2013).

In this study, $\varepsilon$-SVR model was used as SVR model. Vapnik proposed $\varepsilon$-SVR by introducing an alternative $\varepsilon$-insensitive loss function. This loss function allows the concept of margin to be used for regression problems. The purpose of SVR is to find a function having maximum $\varepsilon$ deviation from actual target vectors for all training data given, which should be at the same level as much as possible. In other words, error on any training data should be smaller than $\varepsilon$ (Pal and Arun 2006).

The most important thing to consider in SVR is the selection of $\varepsilon$, C, and kernel function parameters. $\varepsilon$-Insensitive error term constant indicates the range in which error is neglected. As this constant gets smaller, the number of support vectors to be found by SVM increases. C—the regularization constant maintains order between system complexity characterized by weight vector and prediction errors measured by insensitive variables. Using these variables in the most appropriate values is highly effective in the success of SVM (Ekici 2007). However, there is not any decisive criterion for selecting both kernel function and model parameters (Lin 2006). In this study, radial basis kernel function, one of kernel functions widely used in SVR applications, was used and it is defined as follows.

$$K(x_i, x) = \exp\left(-\gamma \|x_i - x_j\|^2\right) \qquad (6)$$

$\gamma$ is the kernel parameter that ensures radius control. SVM architecture is shown in Fig. 2.

### 3.3 Adaptive network-based fuzzy inference system

ANFIS is a hybrid intelligence method that utilizes parallel calculation and learning ability of artificial neural networks and inference feature of fuzzy logic. ANFIS model, developed by Jang (1993), uses Sugeno-type fuzzy inference system and hybrid learning algorithm. Learning algorithm of ANFIS is a hybrid learning algorithm consisting of combined use of the
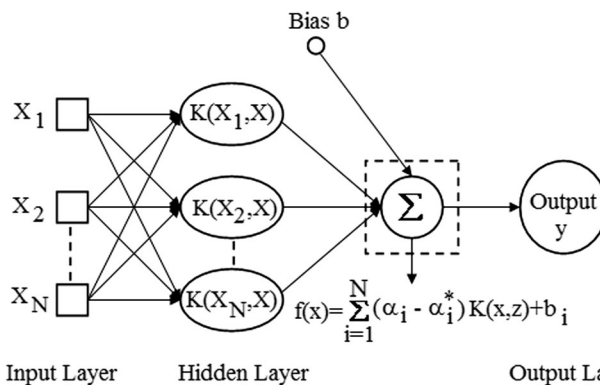


**Fig. 2** General structure of support vector machines

least squares method and backpropagation learning algorithm. Network-based inference systems have two important adjustments: structural adjustment and variable adjustment. Structural adjustment involves the number of variables to be calculated, the number of rules, expression of definition spaces of each input-output variable using fuzzy sets, and creation of the structure of rules, whereas variable adjustment, the centers of membership functions, gradients, widths, and calculation of weights of fuzzy logic rules. It is more advantageous than other systems due to the simplicity of optimization of its parameters (Jang 1993). A general ANFIS architecture consisting of two inputs such as $x_1$ and $x_2$, a single output, $y$ and four rules is as that shown in Fig. 3.

The layers in Fig. 3 comprise the input, the fuzzification, the rule, the normalization, the defuzzification, and overall layers. Detailed information on the functioning of these layers is available in Jang (1993), Jang et al. (1997).

### 3.4 Particle swarm optimization

PSO is a population-based stochastic search algorithm developed by Kennedy and Eberhart (1995), inspired by the behavior of bird flocks. It was designed to solve nonlinear problems. It is used to provide a solution to multi-parameter and multivariable optimization problems. The system is initiated by a
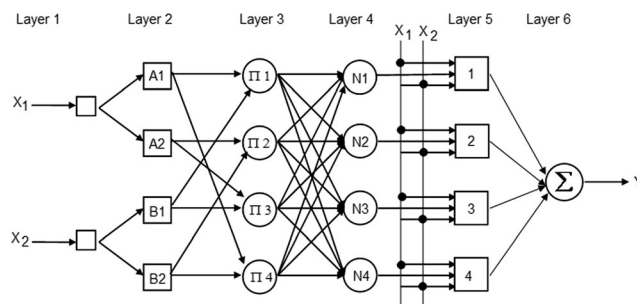


**Fig. 3** Adaptive network-based fuzzy inference system (Jang 1993)

population containing random solutions and searches for the most optimum solution by updating generations. Potential possible solutions, called a particle in PSO, follow the optimum particle on that particular moment and travel around the problem space. The most important difference of PSO from conventional optimization techniques is that it does not require derivative knowledge. This property takes away the burden of complex operations required for solution of many problems. The algorithm of PSO has a small number of parameters, making it significantly easy to apply PSO. PSO can be successfully applied in many areas, including function optimization, fuzzy system control, and artificial neural network training (Zhao et al. 2005).

### 3.4.1 PSO-based SVR and ANFIS application

In this study, PSO was used to optimize the values of the parameters of SVR and ANFIS methods in PSO-SVR and PSO-ANFIS hybrid models. The algorithm of these hybrid models is as follows.

(1)   Selecting parameters for SVR/ANFIS

a.   SVR: type of Kernel function and the range of SVR parameters ($C$, $\varepsilon$, and $\gamma$) are set.
b.   ANFIS: the number and type of membership function for input are set.

(2)   PSO: the number of particles, acceleration constants $c_1$, $c_2$ of the particles, $w$ (inertia weight) value, and the number of iterations are set.
(3)   Training

a1.   Training with SVR: The particles' SVR parameters considered as position ($C$, epsilon, and gamma) are randomly assigned within the limits set for each particle. These parameters are used in SVR.
a2.   Training with ANFIS: The particles' output membership function parameters considered as position are randomly assigned within the limits set for each particle. These parameters are used in the constants (A) of "constant" function specified as ANFIS output membership function
b.   SVR/ANFIS is trained for each particle using training data and the corresponding output values.

(4)   Calculation of fitness values using fitness function for each particle
(5)   The stop criterion is compared. Step 3–step 8 is repeated until stop criterion is met.
(6)   For all particles, the fitness value of each particle is compared with the best position of the particle ($p_{\text{best}}$). If the current value is better than $p_{\text{best}}$, $p_{\text{best}}$ value is replaced with the current value.

(7)   Best fitness value is compared with all previous best values ($g_{\text{best}}$) of the population and if the current value is the best, array index of the particle is reset to the current value.
(8)   Speed and position values for each particle are updated using following equations.

$$v_{ij}^{t} = wv_{ij}^{t-1} + c_1 r_1 \left( p\text{best}_{ij}^{t-1} - x_{ij}^{t-1} \right) + c_2 r_2 \left( g\text{best}_{j}^{t-1} - x_{ij}^{t-1} \right) \quad (7)$$

$$x_{ij}^{t} = x_{ij}^{t-1} + v_{ij}^{t} \quad (8)$$

(9)   Testing data is applied to the model developed according to SVR/ANFIS parameters by which the best fitness value has been obtained.

## 4 Application and results

In order to estimate the monthly change in water level of Lake Beysehir using PSO-$\varepsilon$SVR, PSO-ANFIS, and GRNN methods; before applying the methods, the following equation was used to eliminate unit differences between parameters used and the data was made non-dimensional between 0 and 1.

$$X_{\text{norm}} = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (9)$$

where $X_{\text{norm}}$, $X_i$, $X_{\min}$, and $X_{\max}$ represent normalized, observed, minimum, and maximum values for all parameters, respectively.

Models were constructed using five different input combinations of the variables I, P, E, and O, namely (i) I-P-E-O, (ii) I-P-E, (iii) P-E, (iv) P, and (v) E, and the effects of individual variables on water level change were investigated. Mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination ($R^2$) criteria, given in the following equations, were used to evaluate the success of the generated models.

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} \left| Yi_{\text{observed}} - Yi_{\text{estimate}} \right| \quad (10)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left( Yi_{\text{observed}} - Yi_{\text{estimate}} \right)^2} \quad (11)$$

$$R^2 = \frac{\left[ \sum_{i=1}^{N} \left( Yi_{\text{observed}} - \overline{Y}_{\text{observed}} \right) \left( Yi_{\text{estimate}} - \overline{Y}_{\text{estimate}} \right) \right]^2}{\sum_{i=1}^{N} \left( Yi_{\text{observed}} - \overline{Y}_{\text{observed}} \right)^2 \sum_{i=1}^{N} \left( Yi_{\text{estimate}} - \overline{Y}_{\text{estimate}} \right)^2} \quad (12)$$
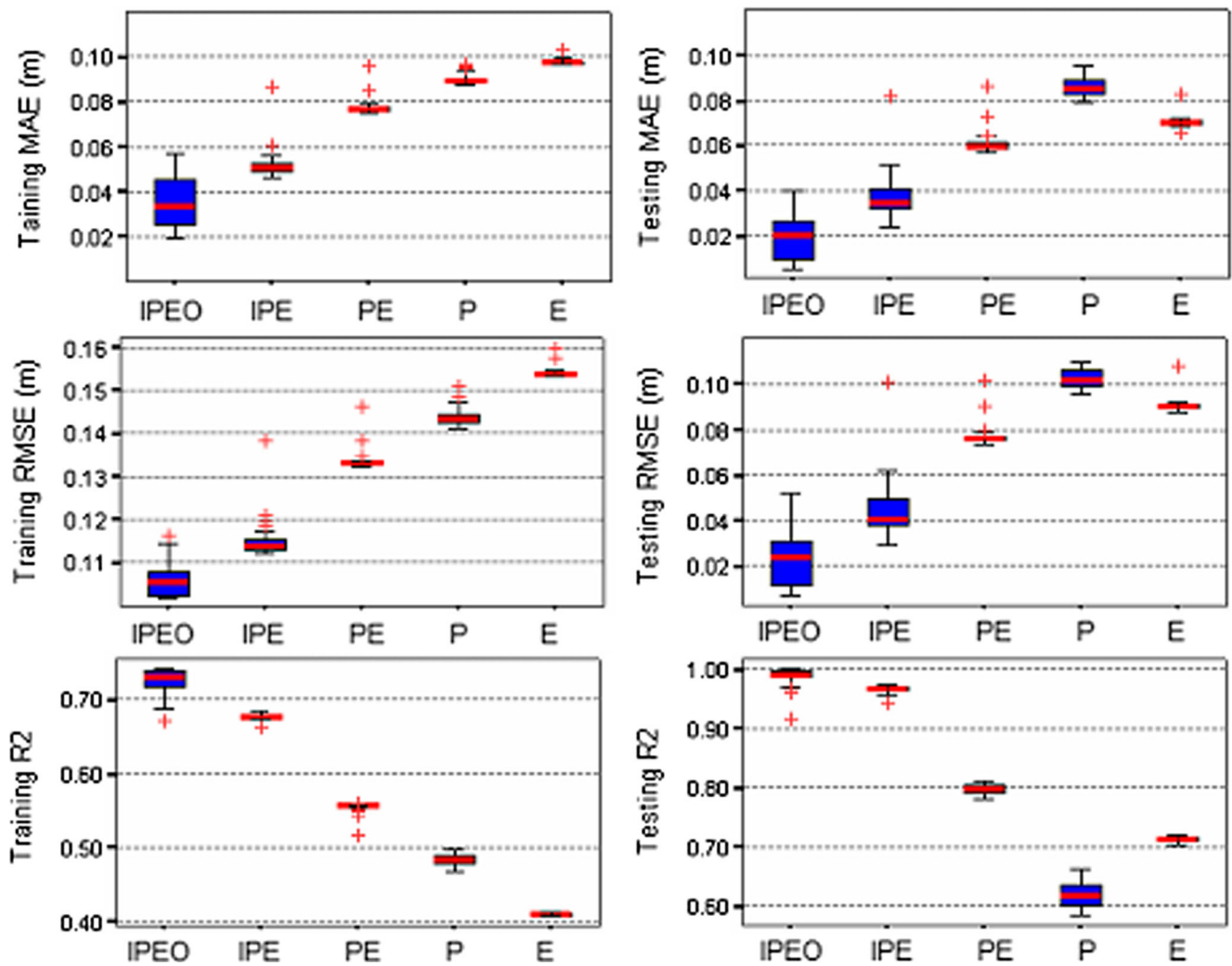
**Fig. 4** Box plots for training- and testing-MAE, RMSE, and $R^2$ values obtained using PSO-$\varepsilon$SVR method for five different inputs at the end of 30 runs

For PSO-$\varepsilon$SVR method, a software package known as LIBSVM Software (Chang and Lin 2001) was employed and the applications were performed on MATLAB 2010b.

While estimating monthly water level change, PSO-$\varepsilon$SVR and PSO-ANFIS hybrid models were developed using PSO algorithm, a heuristic algorithm, for the determination of both $\varepsilon$SVR and ANFIS-method parameters. While constructing both

models, the following were selected: the number of particles in PSO = 30, initial inertia weight $w_{max} = 0.9$, and perminate inertia weight $w_{min} = 0.4$. The range of maximum speed values that a particle can have was [−1, 1], the number of iterations for each input combination was 10, and $c_1$ and $c_2$ learning factors were 2. $c_1$ and $c_2$ are constants representing acceleration terms, which draw each particle to its best position ($p_{best}$) and to the
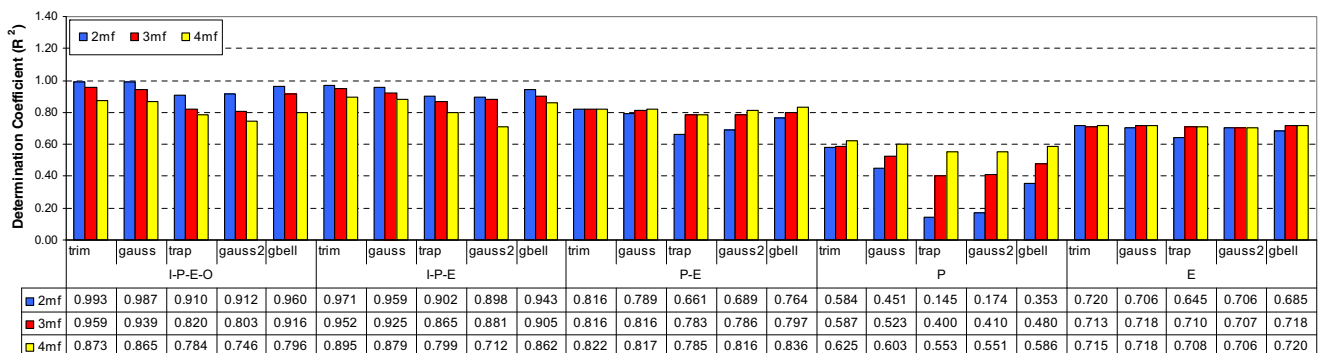


**Fig. 5** Maximum testing-$R^2$ values obtained according to the number and type of membership functions of five different input combinations for the most successful models.
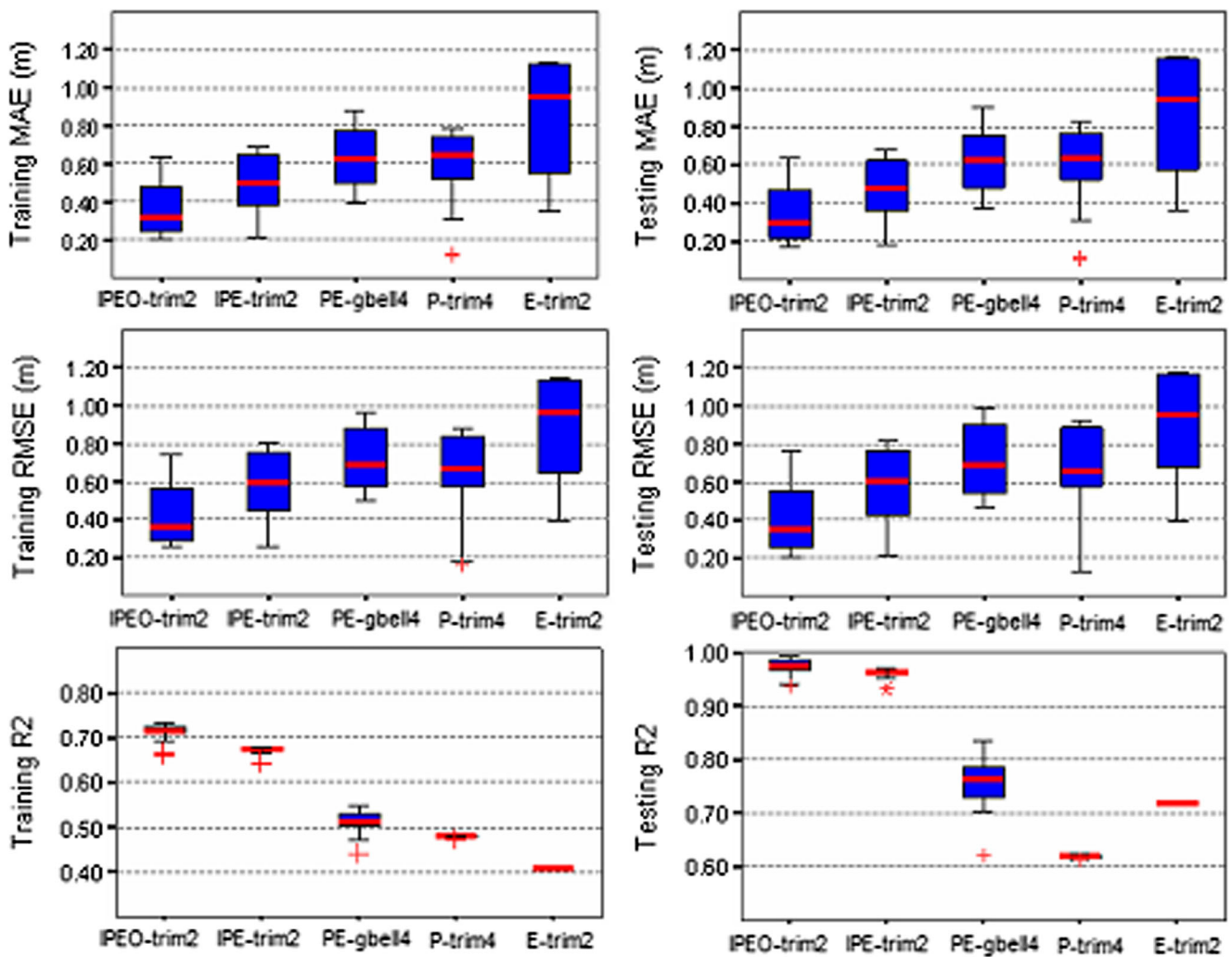
**Fig. 6** Box plots for training- and testing-MAE, RMSE, and $R^2$ values obtained using PSO-ANFIS-GP method for five different inputs at the end of 30 runs

best of the population ($g_{best}$). $c_1$ helps the particle move according to its experiences, while $c_2$ helps it move according to the experiences of the other particles in the population.

### 4.1 PSO-εSVR application

For the estimation of monthly water level, error term ($\varepsilon$), regulatory factor ($C$) of $\varepsilon$SVR method, and $\gamma$ parameters of radial basis kernel function were determined using PSO algorithm. The number of runs while estimating $\varepsilon$SVR parameters using PSO was 30. Search spaces for $C$, $\varepsilon$, $\gamma$ parameters were [1 100], [0.01 0.5], [0.1 8], respectively. Box plots for training- and testing-MAE, RMSE, and $R^2$ values obtained for five different inputs at the end of 30 runs are shown in Fig. 4. According to the box plots in Fig. 4, the most successful PSO-$\varepsilon$SVR model was the model comprised of I-P-E-O inputs, by which the lowest training- and testing-MAE and RMSE values and the highest training and testing $R^2$ values were obtained. The least successful model was the PSO-$\varepsilon$SVR model with E input which gave the highest training MAE and

RMSE values and the lowest training $R^2$ value and the model with $P$ input for testing values. Testing data of 30 runs for I-P-E-O, which was the most successful input combination in estimating water level change according to box plots were as follows: minMAE = 0.0052 m, maxMAE = 0.04 m, and medianMAE = 0.0198 m; minRMSE = 0.0070 m, maxRMSE = 0.0518 m and medianRMSE = 0.0241 m; min$R^2$ = 0.9169, max$R^2$ = 0.9995, median$R^2$ = 0.9909. Standard deviation values for MAE, RMSE, and $R^2$ of 30 runs for I-P-E-O input were 0.0111, 0.0278, and 0.0171, respectively.

### 4.2 PSO-ANFIS application

In this study, Grid Partition (GP) technique was used to set the rules in the case of PSO-ANFIS models. In PSO-ANFIS-GP models, types of input and output membership functions and the numbers of membership functions were determined for each input combination. Trimf, gaussmf, trapmf, gauss2mf, and gbellmf functions were used for input membership functions. The numbers 2-3-4 were considered as the number of
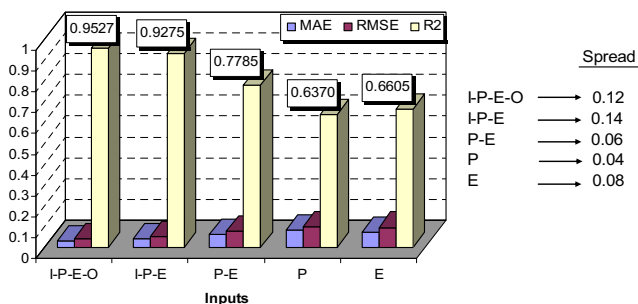
**Fig. 7** Testing-MAE, RMSE, $R^2$, and spread values for the most successful models obtained by GRNN method

membership functions for each input combination. Constant function was used as the output membership function. Optimum values of a constant of output membership functions were determined using PSO. According to the type and number of different membership functions, each input combination was run 30 times, and the most successful models were identified by considering the run result with the highest $R^2$ value. Maximum testing-$R^2$ values obtained according to the number and type of membership functions of five different input combinations for the most successful models are given in Fig. 5.

According to $R^2$ results in Fig. 5, the most successful PSO-ANFIS-GP models were obtained with trimf for I-P-E-O, I-P-E, P and E combinations and with gbellmf for P-E combination. The obtained number of membership functions for the most successful models was two in the cases of I-P-E-O, I-P-E, and E inputs and four in the cases of P-E and P inputs. Box plots for MAE, RMSE and $R^2$ obtained using the results of 30 runs of these models are shown in Fig. 6.

According to the box plots in Fig. 6, the most successful PSO-ANFIS-GP model was the model comprised of I-P-E-O input combination for which minimum training- and testing-

MAE and RMSE values and maximum training- and testing-$R^2$ values were obtained. The least successful model by $R^2$ value was the PSO-ANFIS model for which E input was used for training values and P input was used for testing values. Testing data of 30 runs for I-P-E-O, which was the most successful input combination in estimating water level change according to box plots were as follows: minMAE = 0.1755 m, maxMAE = 0.6333 m and medianMAE = 0.2998 m; minRMSE = 0.2062 m, maxRMSE = 0.7677 m and medianRMSE = 0.3507 m; minR$^2$ = 0.9359, maxR$^2$ = 0.9932, medianR$^2$ = 0.9744. Standard deviation values for MAE, RMSE and $R^2$ of 30 runs for I-P-E-O input were 0.1436, 0.1747, and 0.0155, respectively.

### 4.3 GRNN application

Spread value ($n$) was tested with an increment of 0.01 in the range of [0.01 5] for 5 different input combinations in order to estimate water level change using GRNN technique. The results obtained for testing data at the end of the analysis are given in Fig. 7. According to these results, I-P-E-O input combination was the most successful GRNN model with a spread value of $n = 0.12$ and $R^2 = 0.9527$, whereas the least successful model was the GRNN model obtained in the case of P input combination with a spread value of $n = 0.04$ and $R^2 = 0.6370$. The highest and lowest MAE and RMSE criteria were obtained for P and I-P-E-O combinations, respectively.

The results obtained from the models constructed using PSO-εSVR, PSO-ANFIS-GP, and GRNN methods employed in this study in order to estimate water level change in Lake Beysehir are listed in Table 1. Table 1 also shows the results of the study by Buyukyildiz et al. (2014).

**Table 1** Results of PSO-εSVR, PSO-ANFIS-GP hybrid models, and GRNN models

| Model inputs | PSO-ANN (Buyukyildiz et al. 2014) | | PSO-ANFIS-GP | | | PSO-εSVR | | GRNN | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $(i, j, k)$ | $R^2$ | Input MF type | Number of MFs | $R^2$ | $(C, \varepsilon, \gamma)$ | $R^2$ | Spread value | | $R^2$ |
| I-P-E-O | (4,5,1) | 0.931 | trimf | 2 2 2 2 | 0.9932 | (1.146, 0.273, 1.738) | *0.9995* | 0.12 | | 0.9527 |
| I-P-E | (3,11,1) | 0.687 | trimf | 2 2 2 | 0.9709 | (1.446, 0.270, 0.982) | 0.9715 | 0.14 | | 0.9275 |
| P-E | (2,3,1) | 0.765 | gbellmf | 4 4 | 0.8355 | (1.867, 0.186, 1.793) | 0.8171 | 0.06 | | 0.7785 |
| P | (1,7,1) | 0.560 | trimf | 4 | 0.6248 | (5.773, 0.215, 2.702) | 0.6609 | 0.04 | | 0.6370 |
| E | (1,3,1) | 0.543 | trimf | 2 | 0.7200 | (1.269, 0.270, 1.369) | 0.7203 | 0.08 | | 0.6605 |
| | MLP (Buyukyildiz et al. 2014) | | ANFIS-GP (Buyukyildiz et al. 2014) | | | εSVR (Buyukyildiz et al. 2014) | | RBNN (Buyukyildiz et al. 2014) | | |
| | $(i, j, k)$ | $R^2$ | Input MF type | Number of MFs | $R^2$ | $(C, \varepsilon, \gamma)$ | $R^2$ | Number of neurons | Spread value | $R^2$ |
| I-P-E-O | (4,2,1) | 0.991 | trimf | 2 2 2 2 | 0.9932 | (1, 0.01, 0.1) | 0.9988 | 6 | 4.26 | 0.9986 |
| I-P-E | (3,2,1) | 0.970 | trimf | 2 2 2 | 0.9714 | (4, 0.18, 0.2) | 0.971 | 5 | 1.59 | 0.9710 |
| P-E | (2,5,1) | 0.825 | trimf | 4 4 | 0.8188 | (45, 0.04, 0.2) | 0.816 | 8 | 1.47 | 0.8219 |
| P | (1,5,1) | 0.673 | gaussmf | 3 | 0.6647 | (27, 0.02, 2.3) | 0.666 | 11 | 0.02 | 0.6891 |
| E | (1,7,1) | 0.701 | trimf | 4 | 0.7206 | (6, 0.26, 0.1) | 0.720 | 2 | 0.32 | 0.7266 |

The best value is presented in italic

(a)



(b)



(c)
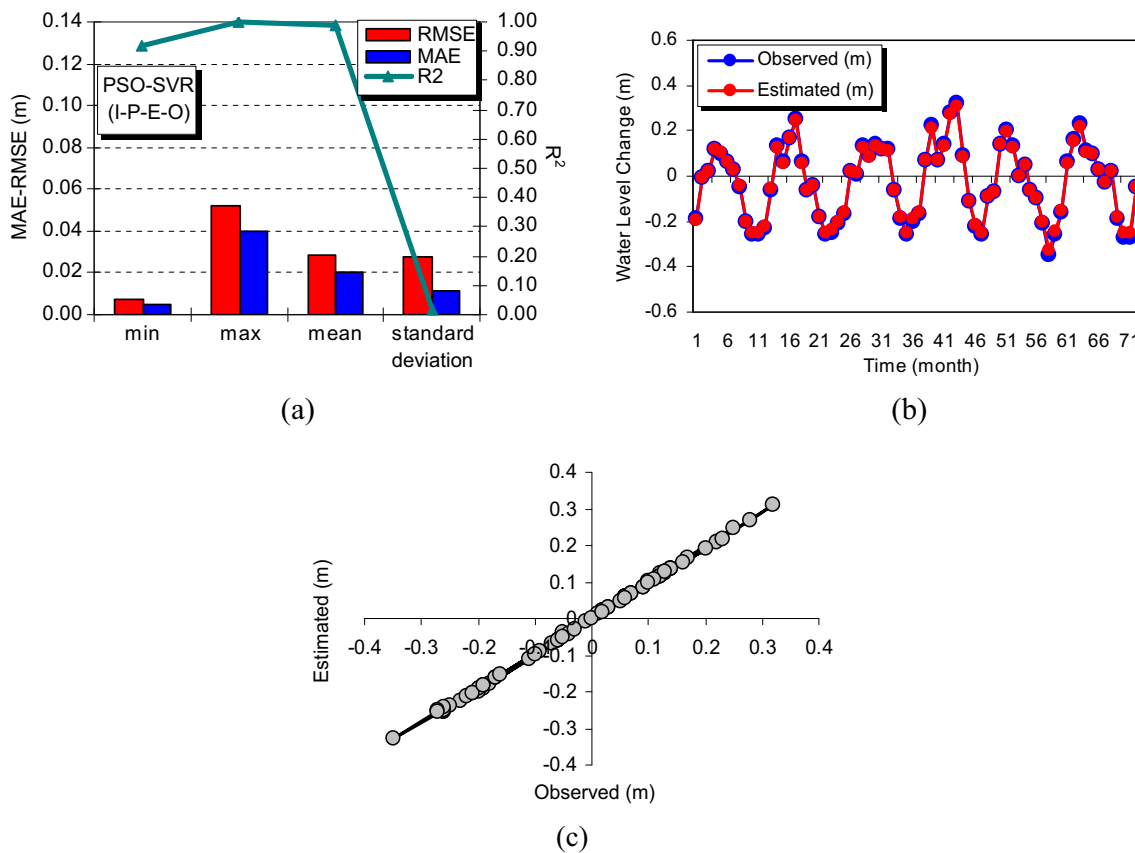
**Fig. 8** **a**) performance criteria **b**)time series **c**) scatter diagram for testing data of the most successful PSO- eSVR (1.146, 0.273, 1.738) model

In view of Table 1, when the results obtained by parameter optimization using PSO are compared with the results obtained using only SVR and ANFIS-GP (Buyukyildiz et al. 2014), it is clear that parameter optimization was effective on the success. In PSO-ANN hybrid model, PSO was used to update weights, rather than parameter optimization, and it was found to have no positive effect on the success of ANN.

The input combination for which the most successful results was obtained in all methods given in Table 1 was I-P-E-O combination. Among the employed methods, PSO-εSVR method was the method with the highest estimation success with an $R^2$ value of 0.9995 in I-P-E-O combination. PSO-ANN models were the least successful model in all combinations, while the other methods generally showed performances significantly similar to each other. The least successful method among the models employed in this study to estimate water level change was GRNN method. The obtained $C$, $\varepsilon$, and $\gamma$ parameters for I-P-E-O combination of PSO-εSVR method, the most successful method, were 1.146, 0.273, and 1.738, respectively.

Minimum, maximum, mean, and standard deviation values for testing data performance criteria, scatter diagram, and time series graph of PSO-εSVR (1.146, 0.273, 1.738) model obtained in the combination of I-P-E-O for 30 runs are shown in Fig. 8a, b, c, respectively. The small standard deviation values

in Fig. 8a demonstrate that PSO-εSVR model achieved a significant success in terms of estimation. Because all $R^2$ values were above 90 %. Scatter diagram and time series graph of monthly water level in Fig. 8b, c, respectively, clearly indicate that the constructed PSO-εSVR (1.146, 0.273, 1.738) model was successful in estimating monthly water level change of Lake Beysehir.

## 5 Conclusions

The objective of this study was to optimize the most appropriate parameter values of SVR and ANFIS methods using PSO algorithm in order to estimate water level change of Lake Beysehir. ε-SVR method was employed in SVR applications, while GP method was used to set the rules in the case of ANFIS models. Additionally, GRNN method was used to predict water level change and the results of the model were compared with the results obtained from hybrid PSO-εSVR and PSO-ANFIS models. While constructing the models, five different input combinations of inflow-lost flow (I), precipitation (P), evaporation (E), and outflow (O) data were used, with monthly water level change as the output. According to the results of the models, the models developed using I-P-E-O combination were the most successful models in all three methods in

estimating monthly water level change. The models with the least success were the models in which only precipitation was used as the input in all three methods. The most successful model among hybrid PSO-$\varepsilon$SVR, PSO-ANFIS-GP, and GRNN models was PSO-$\varepsilon$SVR with an $R^2$ value of 0.9995 obtained using I-P-E-O inputs. When the results of PSO-$\varepsilon$SVR and PSO-ANFIS-GP models obtained in this study were compared with $\varepsilon$SVR and ANFIS-GP results, hybrid PSO models were found to be slightly more successful.

These results suggest that the PSO algorithm, which does not require any derivative information, has a small number of parameters requiring computation and provides a fast solution by taking away the burden of complex operations in solving nonlinear problems and can be successfully used to optimize parameters of the estimator used to estimate lake water level change.

Estimation of water level change with satisfactory accuracy would shed light to further studies on water resources management such as flood control, management of local water supply, recreation activities, ecosystem sustainability, etc.

# References

Alp M, Cigizoglu HK (2004) Farklı yapay sinir ağı metodları ile yağış-akış ilişkisinin modellenmesi. ITU Dergisi 3(1):80–88(in Turkish)

Altunkaynak A (2007) Forecasting surface water level fluctuations of lake van by artificial neural networks. Water Resour Manag 21: 399–408

Altunkaynak A (2014) Predicting water level fluctuations in lake Michigan-Huron using wavelet-expert system methods. Water Resour Manag 28:2293–2314

Buyukyildiz M, Tezel G, Yilmaz V (2013) Water level prediction models based on soft computing techniques (SCTs) for Lake Beysehir, Turkey. Proceedings of the Second International Conference on Water, Energy and the Environment, Kusadası, Turkey September 21–24

Buyukyildiz M, Tezel G, Yilmaz V (2014) Estimation of the change in lake water level by artificial intelligence methods. Water Resour Manag 28:4747–4763

Cengiz, TM, Kahya E (2006) Türkiye göl su seviyelerinin eğilim ve harmonik analizi. itüdergisi/dmühendislik 5 (3): 215–224 (in Turkish)

Chang, CC, Lin CJ (2001) LIBSVM: a library for support vector machines. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm

Chang FJ, Chen PA, Lu YR, Huang E, Chang KY (2014) Real-time multi-step-ahead water level forecasting by recurrent neural networks for urban flood control. J Hydrol 517:836–846

Chau KW (2006) Particle swarm optimization training algorithm for ANNs in stage prediction of Shing Mun river. J Hydrol 329(3–4): 363–367

Chithra NR, Thampi SG, Surapaneni S, Nannapaneni R, Kumar Reddy AA, Kumar JD (2014) Prediction of the likely impact of climate change on monthly mean maximum and minimum temperature in the chaliyar river basin, India, using ANN-based models. Theor Appl Climatol. doi:10.1007/S0070-014-1257-1

Cimen M, Kisi O (2009) Comparison of two different data driven techniques in modeling lake level fluctuations in Turkey. J Hydrol 378(3–4):253–262

De Domenico M, Ghorbani MA, Makarynskyy O, Makarynska D, Asadi H (2013) Chaos and reproduction in sea level. Appl. Math. Mod. 37(6):3687–3697

Ekici S (2007) Elektrik Güç Sistemlerinde Akıllı Sistemler Yardımıyla Arıza Tipi ve Yerinin Belirlenmesi. PhD Thesis. Fırat Üniversitesi. Fen Bilimleri Enstitüsü (in Turkish)

García Nieto PJ, García-Gonzalo E, Alonso Fernández JR, Díaz Muñiz C (2014) Hybrid PSO–SVM-based method for long-term forecasting of turbidity in the Nalón river basin: a case study in northern Spain. Ecological Engineering. 73:192–200

Goyal MK (2014) Modeling of sediment yield prediction using M5 model tree algorithm and wavelet regression. Water Resour Manag 28: 1991–2003

Jang JSR (1993) ANFIS adaptive-network-based-fuzzy inference systems. IEEE Transactions on Systems, Man, and Cybernetics 23(3): 665–685

Jang JSR, Sun CT, Mizutani E (1997) Neuro fuzzy and soft computing a computational approach to learning and machine intelligence. Prentice Hall, USA

Karaboga D (2004) Yapay zeka optimizasyon algoritmaları. Atlas Yayınları(in Turkish)

Karimi S, Kisi O, Shiri J, Makarynskyy O (2013) Neuro-fuzzy and neural network techniques for forecasting sea level in Darwin harbor. Australia Comp Geosci 52:50–59

Kavehkar S, Ghorbani MA, Khokhlov V, Ashrafzadeh A, Darbandi S (2011) Exploiting two intelligent models to predict water level: a field study of Urmia lake. Iran Int J Civil Environ Eng 3(3):162–166

Kennedy J, Eberhart R(1995) Particle swarm optimization. In: Proceedings of the International Conference on Neural Networks. IV: 1942–1948

Khan MS, Coulibaly P (2006) Application of support vector machine in lake water level prediction. Hydr. Eng. 11:199–205

Khatibi R, Ghorbani MA, Naghipour L, Jothiprakash V, Fathima TA, Fazelifard MH (2014) Inter-comparison of time series models of lake levels predicted by several modeling strategies. J Hydrol 511:530–545

Kisi O, Nia AM, Gosheh MG, Tajabadi MRJ, Ahmadi A (2012a) Intermittent streamflow forecasting by using several data driven techniques. Water Resour Manag 26:457–474

Kisi O, Shiri J, Nikoofar B (2012b) Forecasting daily lake levels using artificial intelligence approaches. Comput Geosci 41: 169–180

Lin H (2006) Support vector machines for regression and its application for prediction of machine degradation based on vibration signals. Thesis, Master of Science. University of Alberta. Edmonton, Alberta

Lu W, Wang W, Leung A, Lo S, Yuen R, Xu Z, Fan H (2002) Air pollutant parameter forecasting using support vector machines. IJCNN ‹02, Proceedings of the 2002 International Joint Conference on Neural Networks.1: 630–635

Mohandes MA (2012) Modeling global solar radiation using particle swarm optimization (PSO). Sol Energy 86:3137–3145

Ondimu S, Murase H (2007) Reservoir level forecasting using neural networks: Lake Naivasha. Biosyst Eng 96(1):135–138

Pal M, Arun G (2006) Prediction of the end-depth ratio and discharge in semi-circular and circular shaped channels using support vector machines. Flow Meas Instrum 17(1):49–57

Pousinho HMI, Mendes VMF, Catalão JPS (2011) A hybrid PSO-ANFIS approach for short-term wind power prediction in Portugal. Energy Convers Manag 52:397–402

Ramana RV, Krishna B, Kumar SR, Pandey NG (2013) Monthly rainfall prediction using wavelet neural network analysis. Water Resour Manag 27:3697–3711

Salat R, Osowski S (2004) Accurate fault location in the power transmission line using support vector machine approach, power systems. IEEE Transactions on 19:879–886

Sedki A, Ouazar D (2010) Hybrid particle swarm and neural network approach for stream flow forecasting. Math Model Nat Phenom 5(7):132–138

Seo Y, Kim S, Kisi O, Singh VP (2015) Daily water level forecasting using wavelet decomposition and artificial intelligence techniques. J Hydrol 520:224–243

Shen J, Pei ZJ, Lee ES (2004) Support vector regression in the analysis of soft-pad grinding of wire-sawn silicon wafers. CITSA 2004/ISAS 2004: 19–24

Specht DF (1991) A general regression neural networks. IEEE Transactions on Neural Networks 2(6):568–576

Sudheer C, Maheswaran R, Panigrahi BK, Mathur S (2014) A hybrid SVM-PSO model for forecasting monthly streamflow. Neural. Comput. And Applic. 24:1381–1389

Tapoglou E, Trichakis IC, Dokou Z, Nikolos IK, Karatzas GP (2014) Groundwater-level forecasting under climate change scenarios using an artificial neural network trained with particle swarm optimization. Hydrol Sci J 59(6):1225–1239

Vapnik V (1995) The nature of statistical learningtheory. Springer, New York

Wei CC (2012) Wavelet kernel support vector machines forecasting techniques: case study on water-level predictions during typhoons. Expert Syst Appl 39:5189–5199

Yurtcu S (2006) Eber gölü su seviye değişiminin bulanık mantıkla modellenmesi. Teknoloji 9(1):67–77(in Turkish)

Zhao S, Wang L (2010) Support vector regression based on particle swarm optimization for rainfall forecasting. Third International Joint Conference on Computational Science and Optimization IEEE

Zhao F, Ren Z, Yu D, Yang Y (2005) Application of an improved particle swarm optimization algorithm for neural network training. 0–7803-9422-4/05/2005 IEEE