



Prediction of suspended sediment loading by means of hybrid artificial intelligence approaches

Banu Yilmaz¹ · Egemen Aras² · Murat Kankal³ · Sinan Nacar⁴

Received: 24 April 2019 / Revised: 27 August 2019 / Accepted: 1 October 2019 / Published online: 11 October 2019
© Institute of Geophysics, Polish Academy of Sciences & Polish Academy of Sciences 2019

Abstract

The main aim of the research is to use the artificial neural network (ANN) model with the artificial bee colony (ABC) and teaching–learning-based optimization (TLBO) algorithms for estimating suspended sediment loading. The stream flow per month and SSL data obtained from two stations, İnanlı and Altınsu, in Çoruh River Basin of Turkey were taken as precedent. While stream flow and previous SSL were used as input parameters, only SSL data were used as output parameters for all models. The successes of the ANN-ABC and ANN-TLBO models that were developed in the research were contrasted with performance of conventional ANN model trained by BP (back-propagation). In addition to these algorithms, linear regression method was applied and compared with others. Root-mean-square and mean absolute error were used as success assessing criteria for model accuracy. When the overall situation is evaluated according to errors of the testing datasets, it was found that ANN-ABC and ANN-TLBO algorithms are more outstanding than conventional ANN model trained by BP.

Keywords Artificial bee colony · Çoruh river basin · Estimation · Suspended sediment loading · Teaching–learning-based optimization

Introduction

Prediction of suspended sediment loading (SSL) conveyed by a stream is needed for most of the water supply project plannings and managements. Sediment is a major problem in the design of the dead storage, sediment transport in the

stream, reservoir filling and channel navigability. In addition, if the granules also carry pollutants such as nutrients, pesticides and other chemicals, the estimation of SSL has an additional importance for environmental engineering. Mc Bean and Al-Nassri (1988) examined the vagueness in suspended sediment rating curves and found that sediment loading against discharge is fallacious for indicating goodness of fit.

Alternatively, they suggested to set up regression between discharge and sediment concentration as well.

It is presumed that there are linear relationships between variables in most of the present techniques for time-series analysis. However, in reality, temporary variations in data never show simplistic regularities and it is hard to analyze and estimate the data correctly. Linear repetition and combination of them have often found to be insufficient to describe the behavior of such data. It seems necessary to apply complex nonlinear models such as artificial neural network (ANN) to analyze real-world temporal data (Cigizoglu 2001). Recently, researchers have looked for easier, cheaper and more reliable methods to estimate the SSL and decided to use this model (Ardichoglu et al. 2007; Cigizoglu 2001; Tayfur and Guldal 2006). The ANN practices have been used in many branches of water supply applications successfully

✉ Egemen Aras
egemen.aras@btu.edu.tr

Banu Yilmaz
banuyilmaz@ktu.edu.tr

Murat Kankal
mkankal@uludag.edu.tr

Sinan Nacar
sinannacar@ktu.edu.tr

¹ Department of Civil Engineering, Faculty of Technology, Karadeniz Technical University, 61830 Trabzon, Turkey

² Department of Civil Engineering, Faculty of Engineering and Natural Sciences, Bursa Technical University, 16330 Bursa, Turkey

³ Department of Civil Engineering, Faculty of Engineering, Bursa Uludağ University, 16059 Bursa, Turkey

⁴ Department of Civil Engineering, Faculty of Engineering, Karadeniz Technical University, 61080 Trabzon, Turkey

(Cigizoglu 2001; Cobaner 2011; Jain and Jha 2005; Kisi 2006; Schulze et al. 2005), and it has been thought to give encouraging outcomes, especially in modeling the suspended sediment (Alp and Cigizoglu 2007; Ardicioglu et al. 2007; Cigizoglu 2004; Jain 2001; Kisi 2008; Tayfur 2002; Tayfur and Guldal 2006; Zhu et al. 2007). Tayfur and Guldal (2006) found an ANN model to estimate the daily amount of SSL in rivers. The efficiency of the model was also tried versus the nonlinear black-box model built on two-dimensional unit sediment graph theory (2D-USGT), and when the results were compared, ANN was seen to be much more successful than the 2D-USGT. Ardicioglu et al. (2007) investigated the capacity of two distinct FFBP neural network algorithms, namely the Levenberg–Marquardt and gradient descent in suspended sediment prediction. Alp and Cigizoglu (2007) employed two distinct ANN algorithms, which are the feed-forward backpropagation (FFBP) method and radial basis functions, for predicting the daily amount of SSL in rivers. The monthly stream flow and suspended sediment data of two centrals, Palu and Çayağzı, in the Fırat Basin in Turkey were taken as precedents in the study. Cigizoglu (2004) examined the performance of multilayer perceptrons (MLPs) in daily SSL prediction and foreseeing. As seen in the graphs and statistics of the study, MLPs were found to show a complex nonlinear behavior of the sediment series better than the conventional models. Jain (2001) applied an ANN model to show an integrated stage-discharge-sediment concentration relation and determined that the ANN results were superior to conventional sediment rating curve. Tayfur (2002) found ANN approach from the physically based model for sheet sediment transportation and noted that ANN gives superior results in contrast to the physically based model in some cases. Zhu et al. (2007) used ANN model for monthly suspended sediment flux. According to the results, ANN showed superior compliance under the identical data necessity compared to multiple linear regression and power relation models. Kisi (2008) also designed ANN models for SSL prediction by means of statistical preprocessing of the data. The methodology was assessed with the help of the flow and sediment data from two centrals, Quebrada Blanca and Rio Valenciano, in the USA. Additionally, three distinct training algorithms were contrasted with each other for the chosen input vector.

In recent years, distinct artificial intelligence methods such as support vector machine (SVM) and adaptive neuron fuzzy inference system have been accomplished to resolve prediction issues in sediment engineering applications (Cobaner et al. 2009; Kisi et al. 2009; Kisi and Shiri 2012; Kisi 2012; Lafdani et al. 2013; Rajae et al. 2009). When considering studies that have been done, the various forms of SVM models have shown promising results (Kisi 2012; Lafdani et al. 2013). In the same way, genetic algorithm has been applied to indicate the relationship between sediment

loading and stream flow discharge (Pour 2011). Altunkaynak (2009) and Lafdani et al. (2013) benefited from this method for modeling sediment loading.

As seen in the literature review, it is seen that artificial bee colony (ABC) method has been used successfully in many various areas such as design (Karaboga 2009), cluster (Zhang et al. 2010), programming (Pan et al. 2011) and prediction (Kisi et al. 2012; Uzlu et al. 2014a, b; Yilmaz et al. 2018). Along the same line, teaching–learning–based optimization (TLBO) algorithm has been used in some areas such as electricity (Niknam et al. 2012), computer (Sathapathy and Naik 2011), civil engineering (Dede 2013) and energy (Kankal and Uzlu 2017; Uzlu et al. 2014c).

ABC and TLBO algorithms which are novel, simple and robust optimization algorithms were used in this study during training ANN to predict SSL. Linear regression which is a simple technique was also used in comparison. Even though ABC is a novel technique, it is used in various ranges of engineering. Yet in terms of suspended sediment modeling, there has been just one study (Kisi et al. 2012). In accordance with the literature review, ANN-TLBO technique has not been used in the estimation of SSL and has not been compared with ANN-ABC and ANN-BP algorithms.

Materials and methods

Definition of the research area

The Çoruh River arises from Civilikaya Hill which is located in the Mescit Mountains at the North of the Erzurum Plateau. It runs over East Anatolia and the East Black Sea Regions of Turkey and lastly flows into the Black Sea near Batumi in Georgia (DSI 2006). Although the catchments constitute approximately % 22 of Turkey's woodland, Çoruh River catchment is not a thoroughly developed area and is deprived of arable. Rural earning in the region is one-third of the average of Turkey. A large-scale development program has been designed by the Turkey's government to remove inequality between the catchment societies of Çoruh River and those in other districts of Turkey. Several of these far-reaching investments have been fulfilled, and there are already many other investments which are at the building and planning phase (Akpınar et al. 2011).

For these stations, the data were provided from the report of Turkey General Directorate of Renewable Energy (YEGM). İnanlı Station is at a distance of about 48 km from Artvin-Yusufeli Highway and is located between 41°42'58"E and 40°53'20"N coordinates. Since most of the Çoruh River Basin in the study area consists of narrow valleys with very steep slopes, the highways providing access in the basin often follow the river. The SSL measurement was taken between the years of 2005 and 2011 at the station. Altınsu

Station which is by the roadside near Altınsu with a distance of 5 km from Artvin-Yusufeli Highway is in between $41^{\circ}53'36''\text{E}$ and $41^{\circ}09'47''\text{N}$ coordinates. A SSL measurement was also taken at the station between the years of 1984 and 2001. Because of technical reasons emerging for a period of few months, the data are not continuous for both stations.

The Çoruh River is totally 427 km in length. 400 km part of the river and 19,872 km² of basin covering total of 21,962 km² area are within the borders of Turkey, and the remainder is within Georgia (DSI 2009) (Fig. 1). Çoruh is the fastest running river of Turkey and arises from the western side of the Mescit Mountain (3225 m). After arising from these mountains, Çoruh River flows to the westward over Bayburt and İspir and afterward it reaches Artvin provincial border at Yokuşlu Village of Yusufeli by forming arch. Lastly, it follows a tectonic hollow and flows into Karadeniz in Batumi (Yılmaz et al. 2016).

The Çoruh River is one of the major water supplies of Turkey, particularly as of hydroelectric potency. Approximately 85% of the total annual flow of the Çoruh River is seen between May and July (DSI 2009). Additionally, the

river which has relatively large and variable flow rates conveys high amount of sediment and deposits (estimated at 5 million cubic meters/year) stemming from the erosion in mountains (Berkun 2010). Due to severe erosion in the basin, the fact that the completed, under-construction and planned dam reservoirs get full in a short while and decrease their economic life, poses a risk.

Available data

The monthly stream flow and SSL data of İnanlı Station and Altınsu Station on Çoruh River in Turkey were taken for the study. The location of the Çoruh River catchment together with river gaging stations is shown in Fig. 1. When selecting the stations used in the study, it was preferred that the length of the dataset to be modeled was sufficient and far from human and dam impacts.

For these stations, the data were provided from the report of Turkey General Directorate of Renewable Energy (YEGM). İnanlı Station is in between $41^{\circ}42'58''\text{E}$ and $40^{\circ}53'20''\text{N}$ coordinates. The SSL measurement was taken



Fig. 1 Site location map of Çoruh Stream catchment and place of stations

at the station between the years of 2005 and 2011. Altınsu Station is by the roadside and near between 41°53'36"E and 41°09'47"N coordinates. A SSL measurement was taken at the station between the years of 1984 and 2001. Because of technical reasons emerging for a period of few months, the data are not continuous for both stations.

The process applied to both stations before the analysis phase is the determination of data anomalies. The values which are not fit to the dataset in comparison with other values are called outliers. The presence of outliers can cause the deviation of the dataset from the normal distribution, and as a result of this, the analysis can be negatively affected. These values as well as increase in the standard deviation of the data may also change the shape of the distribution and result in erroneous decisions as a result of the statistical decision process. Therefore, identifying and removing these values is one of the significant procedures to be done before analysis (Uckardes et al. 2010). Outliers were calculated for flow and SSL data at two stations in this study and were ignored for averting incorrect decision making in the following prediction methods. The values which do not meet these conditions are ignored. Although this method is not recommended for all statistical analyses, the standard deviation method was found appropriate because the extreme values of the data in our study were clearly seen. Detailed information regarding the method was also given in several studies in the past (Miller 1991; Howell 1998).

Whole regulated data taken from stations were classified into three in all models: 60% of which were training (118 data for Altınsu and 45 data for İnanlı station), 20% were validation (38 data for Altınsu and 15 data for İnanlı station) and 20% were test set (38 data for Altınsu and 15 data for İnanlı station). The first one was used to train ANN models and named as the training set. After completing training period, rest of the data, namely validation data, were used to

minimize overfitting. Finally, test set was used to determine how the ANN works on new data.

The statistical analyses of training, validation and testing sets are presented in Table 1. This table shows the mean (x_{mean}), standard deviation (S_d), variation coefficient (C_v), skewness coefficient (C_s), maximum (x_{max}), minimum (x_{min}) and maximum mean ratio ($x_{\text{max}}/x_{\text{mean}}$), respectively. As it is understood from the skewness coefficients in the sixth column of the table, the distribution ratio of stream flow and sediment is low. It is also seen that $x_{\text{max}}/x_{\text{mean}}$ values in the last column of Table 1 are also suitable for simulations. From the literature review on sediment prediction, these values were seen to be higher than the values of the study (Kisi et al. 2009; Kisi and Shiri 2012).

ANN (artificial neural network)

Artificial neural neurons which are the main processing element of ANN were obtained by simulating real nerve cells. ANN is modeled in the brain where neurons are connected with complex models. The neurons basically consist of input which are multiplied by weights and then computed by a mathematical function determining the activation of the neuron. Another function (which may be the identity) computes the output of the artificial neuron (sometimes in dependence of a certain threshold). ANNs combine artificial neurons in order to process information (Gershenson 2013).

Most neural network structures use some (certain) type of neuron. Figure 2 shows the structure of a single artificial neuron that receives input from one or more sources being other neurons or data input. The inputs can come from the neuron to another neuron or directly from the outer world. The node or artificial neuron multiplies each of these inputs

Table 1 Statistics analyses of İnanlı and Altınsu stations

Data Set	Data type	x_{mean}	S_d	$C_v (S_d/x_{\text{mean}})$	C_s	x_{max}	x_{min}	$x_{\text{max}}/x_{\text{mean}}$
<i>İnanlı station</i>								
Training	Q (m ³ /s)	102.50	90.79	0.88	2.02	414.69	32.02	4.04
	SSL (ton/day)	7201	13,963	1.94	2.72	61.67	124.66	8.56
Validation	Q (m ³ /s)	154.64	124.96	0.8087	1.17	409.65	50.12	2.64
	SSL (ton/day)	18,999	23,818	1.25	1.25	73,386	312.81	3.86
Testing	Q (m ³ /s)	93.18	101.90	1.09	2.37	395.78	38.88	4.25
	SSL (ton/day)	3410	6,847	2.01	2.86	25,841	124.43	7.58
<i>Altınsu station</i>								
Training	Q (m ³ /s)	158.2	156.54	0.99	1.57	641.5	25.59	4.05
	SSL (ton/day)	15,554	34,932	2.24	3.27	206,993	48.60	13.31
Validation	Q (m ³ /s)	117.1	95.53	0.82	1.71	437.7	36	3.74
	SSL (ton/day)	7705	16,625	2.15	3.17	73,609	101.5	9.55
Testing	Q (m ³ /s)	116.69	126.30	1.08	1.93	471.6	35.46	4.04
	SSL (ton/day)	8014	17,915	2.23	2.66	74,189	62.50	9.26

by a weight. Then, it adds the multiplications and passes the sum to an activation function. Some neural networks do not use an activation function when their principle is different. The following equation summarizes the calculated output:

$$f(x, w) = \phi(x \cdot w) = \phi\left(\sum_{i=1}^p (x_i \cdot w_i)\right) \quad (1)$$

In the equation, variables x and w represent the input vector and weight vector of the neuron when there are p inputs into the neuron. Greek letter ϕ (phi) denotes an activation function. The process results in a single output from a neuron (Anderson et al. 1988).

The ANN is a black-box model that converts input data to output data by means of a particular set of nonlinear basis functions. ANN is defined as complex systems that connect to each other with various connection geometries of artificial nerve cells generated by the nerve cells in the human brain. The network consists of input, output and great numbers of hidden layers of parallel operation components called neurons or nodes, and each layer is entirely bonded to the next layer with interconnection weights. Firstly, detected weight values are gradually altered at each of the iterations during a training process. Afterward, predicted outputs compare with known outputs. In the back-propagation process, the weight adjustments required to minimize the mistakes are determined by inverting any error. Information about MLP is given in detail by Kankal et al. (2011).

The generalization ability of FFBP is strong. If it is trained correctly, it means that it can give correct results even in situations that have never been seen before. Additionally, the most widely utilized network is the three-layer feed-forward ANN (ASCE 2000). The sequence of

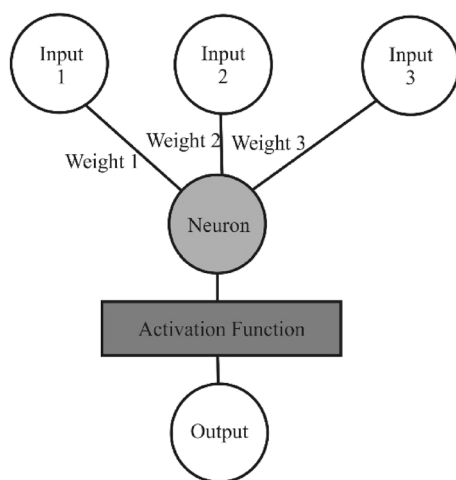


Fig. 2 An artificial neuron (Anderson et al. 1988)

variables for the input layer was flow, flow with one lag and SSL with one lag, respectively. SSL (t time) estimations were used for the network output layer.

Subsequently selecting the input and output variables, the optimum number of neurons in the hidden layer was found by trial and error. The number of hidden neurons was chosen as 5 and 10. The maximum period number was adjusted to 10,000. There was no transfer function at the input layer neurons. For activation of the hidden and output layers of the network, hyperbolic tangent sigmoid transfer function (tansig) and linear transfer function (purelin) were used, respectively.

Among the applied ANNs, BP algorithm is most commonly employed for training MLP in solving various engineering problems. The weights were updated based on their contribution to the error function in this algorithm. The reason for updating the weights is to decrease the error. The Delta rule is a weight update algorithm in the training of neural networks. The algorithm progresses sequentially layer by layer and updates weights as it goes. The weights in the network are brought up to date by the quantity of Δw_{ij} given below:

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} \quad (2)$$

where E is the error between actual and desired network outputs, w_{ij} is the j th weight factor applied to the i th input to the network and η is the learning rate indicating the size of change realized in the learning parameter (Sajwan and Rajesh 2011). In this study, the learning rate and momentum coefficients were tested in three different ways to achieve the optimum result: 0.1, 0.5 and 1, respectively.

Basically, an objective function, such as minimization of sum square error, mean square error or mean absolute error, is chosen for optimizing the neural network. Mean square error criteria were chosen in this study. The objective of the algorithm is to find the optimal weights that generate an output vector close to the target output vector within a selected accuracy (Nourani et al. 2012). The error between the computed and actual outputs is calculated and back-propagated to the layers. Then, the weights are updated depending on their contribution to the error function. The MSE, which can be used as an error value, is given as follows (Uzlu et al. 2014a; Khoshnevisan et al. 2013):

$$MSE = \frac{1}{N} \sum_{i=1}^N (SSL_{observed} - SSL_{predicted})^2 \quad (3)$$

where N is the number of observations, $SSL_{observed}$ is the observed value and $SSL_{predicted}$ is the estimated value belonging to SSL.

ABC (artificial bee colony) algorithm

ABC algorithm that was developed by Karaboga (2005) is an optimization algorithm and one of the most popular swarm optimization algorithms stimulating intelligent foraging behavior of honey bee swarms (Karaboga et al. 2012). It has been practiced in several areas for resolving different problems until today.

In the ABC algorithm, the state of a nutrition resource symbolizes a probable resolution to the optimization problem, and the nectar quantity of a nutrition source complied with the quality (conformity) of related solution (Kisi et al. 2012).

The foraging process comprises three groups of bees which are onlooker, scout and worker bee. It can be referred to Uzlu et al. (2014a) for more detailed information regarding the ABC algorithm.

TLBO (teaching–learning-based optimization) algorithm

TLBO algorithm is a new meta-heuristic optimization algorithm which gained inspiration by teaching–learning process in the class. It was developed by Rao in 2011 and consists of two fundamental phases: (i) by the help of teacher (known as teacher phase) and (ii) interacting with the other learners (known as learner phase) (Rao et al. 2011). The supremacy of TLBO in comparison with the other evolutionary methods is that it simplifies numerical structure and that it is independent on a number of control parameters to define the algorithm's performance (Toğan 2012).

In the teaching phase, the student having minimum objective function (f) value in the whole population is found and imitated as a teacher. The other students in the current population are modified as neighborhood of the teacher.

During learning step of the algorithm, the learning of the students (i.e., learners) is enhanced by means of interaction among themselves. They can also have knowledge by means

of interacting with others. In this way, a learner randomly interacts with the other learner in the population for learning something novel (Patel and Savsani 2016). The flowchart for TLBO algorithm is given in Fig. 3. Detailed Information regarding TLBO algorithm and its application is given by Rao et al. (2011).

ANN training with ABC algorithm

In the research, it is suggested to apply the ABC algorithm to overcome the negativity with the BP algorithm. The ABC algorithm was chosen as the optimized tool as it has capability to reach the best resolutions by relatively moderate computational necessities. In that vein, the ABC algorithm was applied to ANNs in the training process for acquiring more satisfying parameters comprising of weight and bias values in order to decrease the error functions. Parameters were continually updated till it reached the convergence criterion (Yeh and Hsieh 2012).

ABC algorithm includes the following three control parameters: the number of nutrition sources equal to the number of onlookers or employed bees (SN); the limit value; and the maximum cycle number (MCN) (Yeh and Hsieh 2012). The ABC algorithm parameters were set as follows: colony size (NP) = 50 and 100; SN = 25 and 50; parameter range = $[-1, 1]$; and MCN = 2000. Control parameter values of ANN-ABC are shown in Table 2.

ANN training with TLBO algorithm

As in ABC algorithm, the TLBO algorithm is also implemented in ANN training to resolve the BP algorithm's drawbacks. This algorithm is used for the same purpose as ABC. In other words, the TLBO algorithm was used to specify weight and bias values that would reduce the error values to the minimum in the training process. The proposed training

Fig. 3 Suggested flowchart for TLBO algorithm (Kankal and Uzlu 2017)

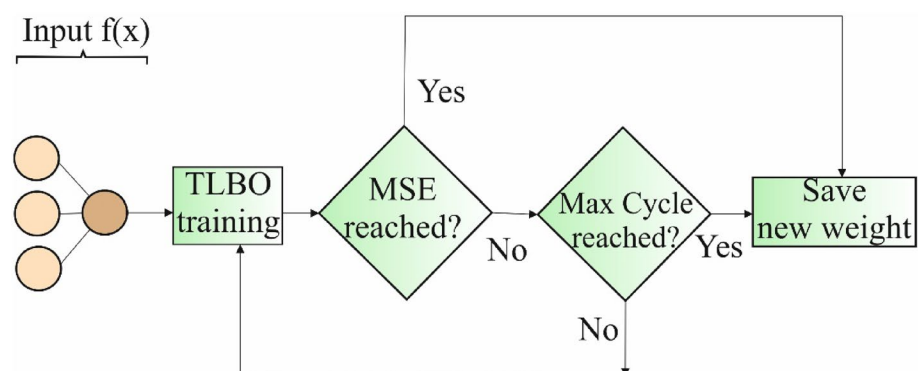


Table 2 Control parameters values of ANN-ABC and ANN-TLBO models

ANN-ABC				ANN-TLBO				
NP	SN (NP/2)		MCN	Range of parameters		NMI	SP	
50	100	25	50	2000	[- 1.1]	2000	50	100

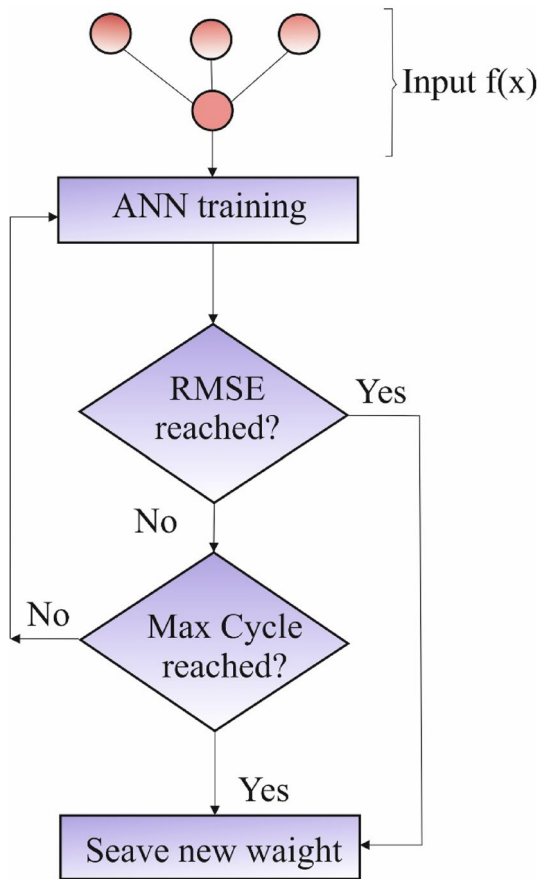


Fig. 4 Suggested training plan for ANN-ABC and ANN-TLBO (Uzlu et al. 2014a, b, c)

scheme for ANN-ABC and ANN-TLBO is shown in Fig. 4 (Hassim and Ghazali 2012).

The control parameters of TLBO were set as below: number of maximum iterations (NMI)=2000 and size of

population (SP)= 50 and 100, respectively. The control parameter values of ANN-TLBO are shown in Table 2.

Linear regression

Regression analysis is one of the commonly used techniques in multifactorial data analysis. It is assumed that the y dependent variable is affected by independent variables such as x_m , and if a linear equation is selected for the relationship between them, the regression equation of y is written as follows:

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_mx_m \tag{4}$$

In this equation, y shows the estimated value when the independent variables are x_1, x_2, \dots, x_m and regression coefficients are b_0, b_1, \dots, b_m . This simple technique was used to compare the results of the other methods of the study.

Results and discussion

Determination of the input variable is of great importance in computational analyses such as ANN. It has been observed that input variables affect the results, particularly when the relationship of complex phenomena such as natural phenomenon is modeled. In Table 3, there are correlation coefficients indicating the relation between inputs and outputs of the İnanlı and Altınsu stations, respectively. These coefficients are calculated separately for training, validation and test sets. In other words, these coefficients show to what extent the input data have effect on the output data. As seen in Table 3, the correlation coefficients decreased markedly after the second lag; hence, only one lag input scenario combination was established.

Table 3 Correlation coefficients between input parameters and SSL obtained from İnanlı and Altınsu stations

Time series	İnanlı station				Altınsu station			
	Training set	Test set	Validation set	All dataset	Training set	Test set	Validation set	All dataset
Q_t	0.710	0.963	0.553	0.666	0.785	0.902	0.7262	0.789
Q_{t-1}	0.472	0.619	0.187	0.364	0.3534	0.437	0.160	0.357
Q_{t-2}	0.148	0.067	-0.322	0.005	-0.016	0.048	-0.142	0.001
SSL_{t-1}	0.332	0.432	0.242	0.330	0.2534	0.368	0.2382	0.273
SSL_{t-2}	0.248	0.087	-0.243	0.069	-0.086	0.002	-0.061	-0.060

These effects in the tables were taken into account, and (Q_t) current days flow data and (Q_{t-1}) previous days flow data and (SSL_{t-1}) previous (generally previous month) SSL data were selected as the input variables for both stations. The output data are SSL value which was measured once or twice a month. Only one lag input scenario combination was constructed in composed scenarios of the simulation experiments in the majority of previous studies because correlations became weaker by the second lag.

That the input parameters are within a very wide range makes it difficult to train the ANN. The reason is that there are too small and too large values in the dataset. For preventing this negative effect, input values were normalized as follows:

$$\text{Normalized value} = \left[\frac{\text{Raw value} - \text{Minimum value}}{\text{Maximum value} - \text{Minimum value}} \right] \times (0.9 - 0.1) + 0.1 \quad (5)$$

More than one evaluation criterion was used to make a better comparison between the models and evaluate the results more accurately. Correlation between real values and output values and accuracy of the estimation results were evaluated by the root-mean-square error (RMSE) and the mean absolute error (MAE). Formulations regarding these criteria are given below:

Root-mean-square error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (SSL_{\text{observed}} - SSL_{\text{predicted}})^2} \quad (6)$$

Mean absolute error (MAE):

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |SSL_{\text{observed}} - SSL_{\text{predicted}}| \quad (7)$$

where N is the number of observations, SSL_{observed} is the observed value and $SSL_{\text{predicted}}$ is the estimated value belonging to SSL.

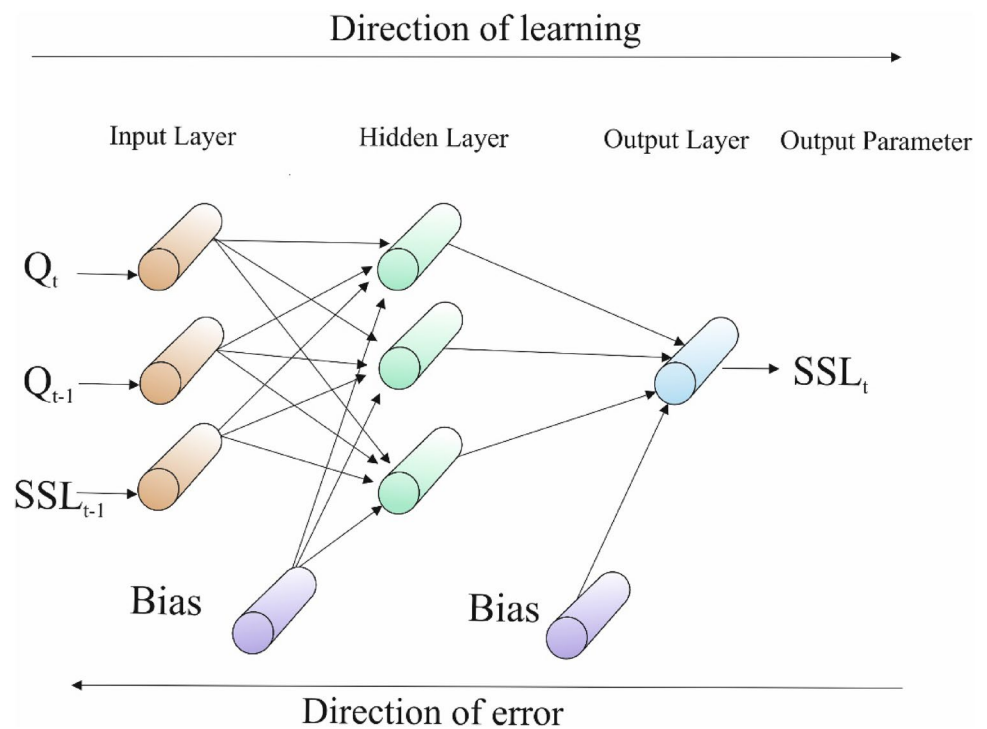
One of the problems that arises in ANN training is the overfitting of the network. During overfitting, the error on the training set decreases to a too small value, while for the test data offered to the network the error is large. Accordingly, the network can memorize instances of education but cannot learn to generalize for new situations. The period number, number of hidden layers and node number of hidden layer must be selected by trial and error for avoiding overfitting during training. The number of nodes in hidden layers affects the sensitivity of the network. While few nodes can cause inadequacy, too many nodes may be concluded as overfitting. Two different node counts as 5 and 10 were tested in this study to reach the optimum number of nodes in

the hidden layer. A three-layer network that is composed of input, hidden and output layers was selected in the present study for all generated neural networks (Fig. 5). The fundamental configurations regarding ANN-ABC, ANN-TLBO and ANN-BP methods are shown in Table 4.

While the models used in the study are evaluated, the RMSE value was interpreted firstly. RMSE values may be different from each other while the correlations of any two datasets may be the same. Accordingly, RMSE and MAE values may not be zero when the linear dependency between the two series is 100%. Therefore, the fact that a model has a high correlation value does not mean that it gives very good estimates. Scatter diagrams and time series of prediction performances for the testing period using ANN-BP, ANN-ABC, ANN-TLBO and linear regression are shown in Fig. 6 for Altınsu station and Fig. 7 for İnanlı station. Although the ANN-BP model seems to be successful in reflecting peak values, the closest values to the observed values are obtained with the ANN-ABC model in Fig. 6 and with this model was obtained the highest success clearly in Altınsu station. Briefly, during overfitting, the error on the training set decreases to a too small value, while for the test data offered to the network the error is large; it was seen that ANN-ABC and ANN-TLBO models were superior than ANN-BP at İnanlı station. It was found that ANN-ABC model was better than ANN-BP model in the Altınsu station. The performance of ANN-BP in both stations was increased with the use of ABC or TLBO algorithms in ANN training.

Error values of BP, ABC, TLBO algorithms for ANN and linear regression in testing set are shown in Table 5. According to the comparisons in RMSE and MAE criteria, the most acceptable results for each criteria is marked in bold. In accordance with Table 5, the model with RMSE value of 2097 and MAE value of 1191 in ANN-TLBO (3-5-1) was the best in İnanlı station for the testing set. The model with RMSE value of 2352 and MAE value of 1362 in ANN-ABC (3-10-1), and RMSE value of 2595 and MAE value of 1344 in ANN-BP (3-10-1) was the best. In addition, RMSE value is 5403 and MAE value is 2701 in linear regression. When an evaluation was made among all the methods used, it was seen that the smallest RMSE value was reached with the ANN-TLBO method in İnanlı station.

As seen in Table 5, the model with RMSE value of 7296 and MAE value of 3755 in ANN-BP (3-5-1) is the best for the testing set. The model with RMSE value of 7269 and MAE value of 4206 in ANN-ABC (3-10-1), and RMSE value of 7745 and MAE value of 4585 in ANN-TLBO (3-10-1) is the best. RMSE value is 11,208 and MAE value is 6668 in linear regression. It can be seen that the lowest RMSE value was obtained by the ANN-ABC method among all the methods applied in Altınsu station.

Fig. 5 Suggested ANN model for SSL estimation**Table 4** Basic configurations regarding ANN-ABC, ANN-TLBO and ANN-BP methods

Number of hidden layer	Single hidden layer
Network structure (input–hidden–output neurons)	3-5-1 (trial and error) 3-10-1 (trial and error)
Evaluation of training process	Root-mean-square error (RMSE)
Transfer function (input to hidden layer)	Tangent sigmoid function
Transfer function (hidden to output layer)	Linear function
Training epochs	10,000 (for ANN-BP) 5000 (for ANN-ABC and ANN-TLBO)

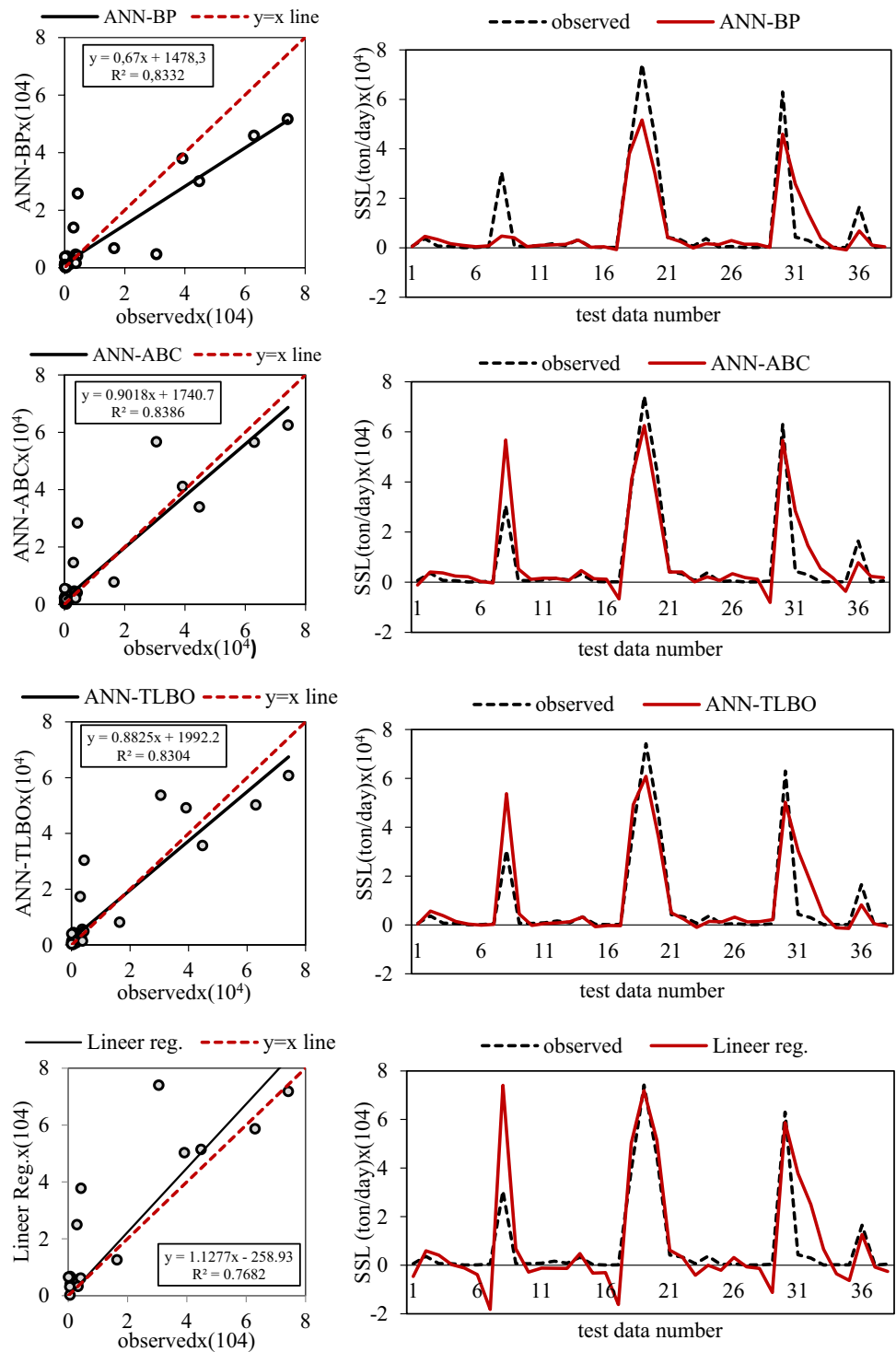
TLBO is a new meta-heuristic optimization algorithm which is based on the teacher's natural phenomenon and has been applied on a large number of fields. In this study, the hybrid version of this algorithm with ANN was used. The results belonging to TLBO were significantly more successful than other methods for one station. However, when using this algorithm, results should be taken into consideration in terms of secrets of TLBO's dominance. According to these, there are several mistakes or misconceptions regarding TLBO in the original papers. Comparisons of TLBO algorithm with other methods were not fair because of experiments not repeated in the same setting. Moreover, the stopping criteria for the TLBO experiments were different from the other compared algorithms. Additionally, there are misconceptions about parameter-less control. In short, it was provided reminders for meta-heuristic researches and practitioners in order to avoid similar

mistakes regarding both the qualitative and quantitative aspects (Črepinšek et al. 2012).

Conclusions

The applicability of ANN-ABC and ANN-TLBO techniques in monthly SSL estimation was demonstrated in the study. The predictions of these models were compared to ANN-BP and linear regression model. Based on the results of testing dataset, the proposed ANN-ABC and ANN-TLBO models were more robust than classical ANN-BP model. As expected, the results obtained from linear regression were found to be the most unsuccessful according to the performance evaluation criteria. The analyses showed that the amount of sediment carried at any given time by the flow was not only dependent on the existing hydrological

Fig. 6 Scatter diagrams and time series of prediction efficiency for the testing period by means of ANN-BP, ANN-ABC, ANN-TLBO and linear regression—Altınsu Station

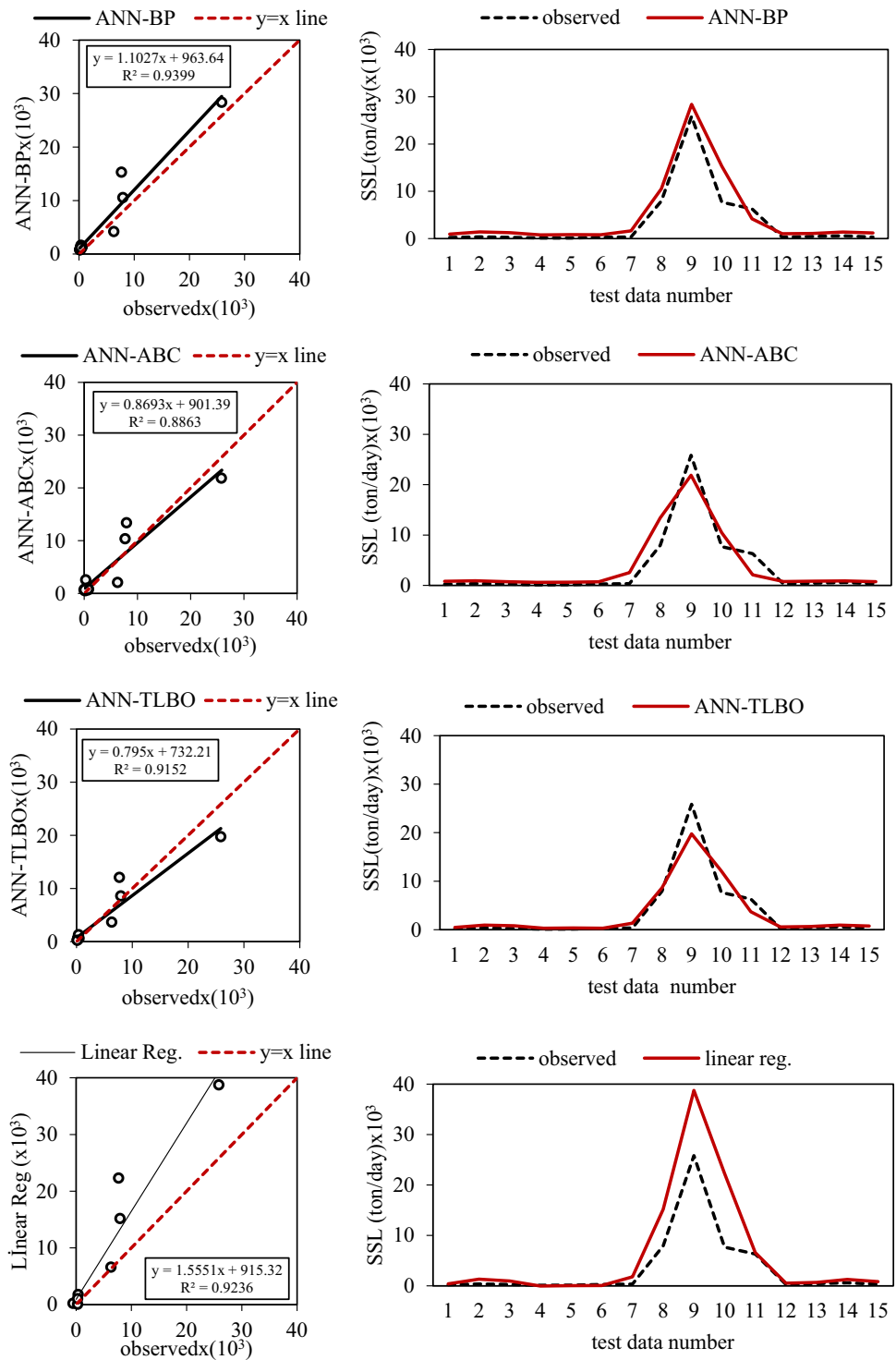


conditions but also on delayed effects. Since these effects were taken into account in all methods used, stronger predictions could be made between input and output data.

RMSE was used to compute the errors and similarities in prediction of the observed responses in comparing the performances of the four methods. For İnanlı station, TLBO algorithm 62%, ABC algorithm 56% and

BP algorithm 52% revealed better prediction than linear regression. The RMSE with the value of 3.7% in the ANN-ABC method and that with the value of 12.7% in the ANN-TLBO method were lower than the ANN-BP model for testing set. For Altınsu station, TLBO algorithm 31%, ABC algorithm 35% and BP algorithm 0.35% revealed better prediction level than linear regression. In

Fig. 7 Scatter diagrams and time series of prediction efficiency for the testing period by means of ANN-BP, ANN-ABC, ANN-TLBO and linear regression—Iranlı Station



the ANN-ABC method, the RMSE value of 0.38% was lower than the value in ANN-BP model for testing set. The proposed ANN approaches for estimating SSL are numerically more sensitive than the classical ANN model performed in the present study. As a result, the recommended

model has capacity to provide efficiently correct results at least for the resolved numerical examples in this study.

Even if two stations were used in the study, the results obtained can be consolidated with the results of other datasets. In addition, the models used in these studies (ANN-ABC and ANN-TLBO) can be compared with different

Table 5 Error values of BP, ABC, TLBO algorithms for ANN and linear regression in testing set

ANN architecture	BP algorithm		ABC algorithm		TLBO algorithm		Linear reg	
	RMSE (ton/day)	MAE (ton/day)	RMSE (ton/day)	MAE (ton/day)	RMSE (ton/day)	MAE (ton/day)	RMSE (ton/day)	MAE (ton/day)
<i>İnanlı station</i>								
3-5-1	2761	1756	2571	2073	2097	1191	5403	2701
3-10-1	2595	1344	2352	1362	2215	1041		
<i>Altınsu station</i>								
3-5-1	7296	3755	7560	4236	7863	5024	11,208	6668
3-10-1	8079	4.237	7269	4206	7745	4585		

meta-heuristic methods. Since the sediment in the monthly SSL is difficult to estimate, it should be supported and developed by other works.

Performances belonging to TLBO, ABC, BP algorithms and linear regression were compared with each other in this study. Other than that, different methods such as PSO (particle swarm optimization) and DE (differential evolution) can be used to make comparisons. Besides, comparing TLBO algorithm with other methods should be fair in terms of repeat in the same setting, especially in experimental studies.

Acknowledgments The authors thank YEGM (General Directorate of Renewable Energy) for the hydrological data of the research. This study is dedicated in memory of the late Assoc. Prof. Dr. Murat İhsan KÖMÜRÇÜ, who died in February 2013.

Compliance with ethical standards

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

References

- Akpınar A, Kömürçü Mİ, Kankal M (2011) Development of hydro-power energy in Turkey: the case of Coruh river basin. *Renew Sustain Energy Rev* 15(2):1201–1209
- Alp M, Cigizoglu HK (2007) Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. *Environ Model Softw* 22(1):2–13
- Altunkaynak A (2009) Sediment load prediction by genetic algorithms. *Adv Eng Softw* 40(9):928–934
- Anderson JA., Rosenfeld E, Pellionisz A (eds) (1988). *Neurocomputing* (vol 2). MIT Press, Cambridge
- Ardıçlıoğlu M, Kişi Ö, Haktanır T (2007) Suspended sediment prediction using two different feed-forward back-propagation algorithms. *Can J Civ Eng* 34(1):120–125
- ASCE Task Committee (2000) Artificial neural networks in hydrology. I: preliminary concepts. *J Hydrol Eng* 5(2):115–123
- Berkun M (2010) Hydroelectric potential and environmental effects of multidam hydropower projects in Turkey. *Energy Sustain Dev* 14(4):320–329
- Cigizoglu HK (2001) Suspended sediment estimation for rivers using artificial neural networks and sediment rating curves. *Turk J Eng Environ Sci* 26(1):27–36
- Cigizoglu HK (2004) Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons. *Adv Water Resour* 27(2):185–195
- Cobaner M (2011) Evapotranspiration estimation by two different neuro-fuzzy inference systems. *J Hydrol* 398(3):292–302
- Cobaner M, Unal B, Kisi O (2009) Suspended sediment concentration estimation by an adaptive neuro-fuzzy and neural network approaches using hydro-meteorological data. *J Hydrol* 367(1):52–61
- Črepinšek M, Liu SH, Mernik L (2012) A note on teaching-learning-based optimization algorithm. *Inf Sci* 212:79–93
- Dede T (2013) Optimum design of grillage structures to LRFD-AISC with teaching-learning based optimization. *Struct Multidisc Optim* 48(5):955–964
- DSI (General Directorate of State Hydraulic Works) (2006) Yusufeli dam and hydropower plant project, Chapter I: introduction. Environmental Impact Assessment, Draft Final Report, Ankara
- DSI (General Directorate of State Hydraulic Works) (2009) Coruh River development plan. In: International workshop on trans-boundary water resources management; Tbilisi, Georgia
- Gershenson C (2003) Artificial neural networks for beginners. arXiv preprint cs/0308031.
- Hassim YMM, Ghazali R (2012) Training a functional link neural network using an artificial bee colony for solving a classification problems. arXiv preprint arXiv:1212.6922.
- Howell DC (1998) *Statistical methods in human sciences*. Wadsworth, New York
- Jain SK (2001) Development of integrated sediment rating curves using ANNs. *J Hydraul Eng* 127(1):30–37
- Jain SK, Jha R (2005) Comparing the stream re-aeration coefficient estimated from ANN and empirical models/Comparaison d'estimations par un RNA et par des modèles empiriques du coefficient de réaération en cours d'eau. *Hydrol Sci J* 50(6):1037–1052
- Kankal M, Uzlu E (2017) Neural network approach with teaching-learning-based optimization for modeling and forecasting long-term electric energy demand in Turkey. *Neural Comput Appl* 28:737–747
- Kankal M, Akpınar A, Kömürçü Mİ, Özşahin TŞ (2011) Modeling and forecasting of Turkey's energy consumption using socio-economic and demographic variables. *Appl Energy* 88(5):1927–1939
- Karaboga D (2005) An idea based on honey bee swarm for numerical optimization. Technical report-tr06, Erciyes University, Engineering Faculty, Computer Engineering Department
- Karaboga N (2009) A new design method based on artificial bee colony algorithm for digital IIR filters. *J Frankl Inst* 346(4):328–348
- Karaboga D, Ozturk C, Karaboga N, Gorkemli B (2012) Artificial bee colony programming for symbolic regression. *Inf Sci* 209:1–15
- Khoshnevisan B, Rafiee S, Omid M, Yousefi M, Movahedi M (2013) Modeling of energy consumption and GHG (greenhouse gas)

- emissions in wheat production in Esfahan province of Iran using artificial neural networks. *Energy* 52:333e8
- Kisi Ö (2006) Daily pan evaporation modelling using a neuro-fuzzy computing technique. *J Hydrol* 329(3):636–646
- Kisi Ö (2008) Constructing neural network sediment estimation models using a data-driven algorithm. *Math Comput Simul* 79(1):94–103
- Kisi O (2012) Modeling discharge-suspended sediment relationship using least square support vector machine. *J Hydrol* 456:110–120
- Kisi O, Shiri J (2012) River suspended sediment estimation by climatic variables implication: comparative study among soft computing techniques. *Comput Geosci* 43:73–82
- Kisi O, Haktanir T, Ardicioglu M, Ozturk O, Yalcin E, Uludag S (2009) Adaptive neuro-fuzzy computing technique for suspended sediment estimation. *Adv Eng Softw* 40(6):438–444
- Kisi O, Ozkan C, Akay B (2012) Modeling discharge-sediment relationship using neural networks with artificial bee colony algorithm. *J Hydrol* 428:94–103
- Lafdani EK, Nia AM, Ahmadi A (2013) Daily suspended sediment load prediction using artificial neural networks and support vector machines. *J Hydrol* 478:50–62
- McBean EA, Al-Nassri S (1988) Uncertainty in suspended sediment transport curves. *J Hydraul Eng* 114(1):63–74
- Miller J (1991) Reaction time analysis with outlier exclusion: bias varies with sample size. *Q J Exp Psychol* 43(4):907–912
- Niknam T, Azizipanah-Abarghooee R, Narimani MR (2012) A new multi objective optimization approach based on TLBO for location of automatic voltage regulators in distribution systems. *Eng Appl Artif Intell* 25(8):1577–1588
- Nourani V, Sharghi E, Aminfar MH (2012) Integrated ANN model for earthfill dams seepage analysis: Sattarkhan dam in Iran. *Artif Intell Res* 1:22e37
- Öcal O (2007) Yapay sinir ağları algoritması kullanılarak akarsu havzalarında yağış-akış-katı madde ilişkisinin belirlenmesi (in Turkish), Dissertation, Pamukkale University
- Pan QK, Tasgetiren MF, Suganthan PN, Chua TJ (2011) A discrete artificial bee colony algorithm for the lot-streaming flow shop scheduling problem. *Inf Sci* 181(12):2455–2468
- Patel VK, Savsani VJ (2016) A multi-objective improved teaching-learning based optimization algorithm (MO-ITLBO). *Inf Sci* 357:182–200
- Pour OMR, Shui LT, Dehghani AA (2011) Genetic algorithm model for the relation between flow discharge and suspended sediment load (Gorgan river in Iran). *Electron J Geotech Eng* 16:539–553
- Rajaei T, Mirbagheri SA, Zounemat-Kermani M, Nourani V (2009) Daily suspended sediment concentration simulation using ANN and neuro-fuzzy models. *Sci Total Environ* 407(17):4916–4927
- Rao RV, Savsani VJ, Vakharia DP (2011) Teaching-learning-based optimization: a novel method for constrained mechanical design optimization problems. *Comput Aided Des* 43(3):303–315
- Sajwan N, Rajesh K (2011) Designing aspects of artificial neural network controller. *Int J Sci Eng Res* 2–4
- Satapathy SC, Naik A (2011) Data clustering based on teaching-learning-based optimization. In: International conference on swarm, evolutionary, and memetic computing (pp 148–156). Springer, Berlin
- Schulze FH, Wolf H, Jansen HW, Van der Veer P (2005) Applications of artificial neural networks in integrated water management: fiction or future? *Water Sci Technol* 52(9):21–31
- Tayfur G (2002) Artificial neural networks for sheet sediment transport. *Hydrol Sci J* 47(6):879–892
- Tayfur G, Guldal V (2006) Artificial neural networks for estimating daily total suspended sediment in natural streams. *Hydrol Res* 37(1):69–79
- Toğan V (2012) Design of planar steel frames using teaching-learning based optimization. *Eng Struct* 34:225–232
- Uckardes F, Sahinler S, Efe E (2010) Aykırı gözlemlerin belirlenmesinde kullanılan bazı istatistikler(in Turkish). *KSU J Nat Sci* 13(1):42–45
- Uzlu E, Akpınar A, Öztürk HT, Nacar S, Kankal M (2014a) Estimates of hydroelectric generation using neural networks with the artificial bee colony algorithm for Turkey. *Energy* 69:638–647
- Uzlu E, Kömürcü Mİ, Kankal M, Dede T, Öztürk HT (2014b) Prediction of berm geometry using a set of laboratory tests combined with teaching-learning-based optimization and artificial bee colony algorithms. *Appl Ocean Res* 48:103–113
- Uzlu E, Kankal M, Akpınar A, Dede T (2014c) Estimates of energy consumption in Turkey using neural networks with the teaching-learning-based optimization algorithm. *Energy* 75:295–303
- Yeh WC, Hsieh TJ (2012) Artificial bee colony algorithm-neural networks for S-system models of biochemical networks approximation. *Neural Comput Appl* 21(2):365–375
- Yilmaz B, Aras E, Nacar S (2016) Estimation of daily suspended sediment load with an artificial neural network. In: 1st International Black Sea congress on environmental sciences (IBCESS), pp 708–720
- Yilmaz B, Aras E, Nacar S, Kankal M (2018) Estimating suspended sediment load with multivariate adaptive regression spline, teaching-learning based optimization, and artificial bee colony models. *Sci Total Environ* 639:826–840
- Zhang C, Ouyang D, Ning J (2010) An artificial bee colony approach for clustering. *Expert Syst Appl* 37(7):4761–4767
- Zhu YM, Lu XX, Zhou Y (2007) Suspended sediment flux modeling with artificial neural network: an example of the Longchuanjiang River in the Upper Yangtze Catchment. *China Geomorphol* 84(1):111–125