

Estimation of Monthly Total Dissolved Solids Using ANN and LS-SVM Techniques in the Aji Chay River, Iran

Mahdieh Jannatkah¹, Abolghasem Akbari^{2*}, Aida Bagheri Basmanji¹, Ebrahim Rahmani¹, Jonathan Peter Cox³

1. Department of Civil Engineering, University of Tabriz, Tabriz, Iran

2. Department of Civil Engineering, Binaloud Institute of Higher Education, Mashhad, Iran

3. Caribbean Institute for Meteorology and Hydrology, B23006, Barbados

Email: akbariinbox@yahoo.com (Corresponding author)

Received:30 September 2020; Accepted: 24 November 2020; Available online: 30 December 2020

Abstract: This research follows on from diverse international efforts to safeguard one of the largest natural lakes in the world, Urmia lake in North West Iran. In this research two new numerical packages based on Artificial Neural Networks (ANN) and the Least Square Support Vector Machine (LS-SVM) models were developed to estimate monthly Total Dissolved Solid (TDS) in the Aji Chay River, one the main tributaries of Urmia lake, Iran. A feed forward back propagation (FFB) model was used to obtain a set of coefficients for a linear model, and the radial basis function (RBF) kernel was employed for the LS-SVM model. The input data sets of both the ANN and LS-SVM models consists of six water quality parameters: TDS, Mg^{2+} , Na^+ , Ca^{2+} , Cl^- , and SO_4^{2-} , all collected on a monthly time scale over a period of 30 years from the Vanyar and Zarnagh stations, in the Aji Chay watershed. The research demonstrated that both models can effectively predict the variability of TDS, but for the Vanyar station with the ANN model (giving an R^2 value of 0.913 and RMSE of 0.0032, a Nash-Sutcliffe Efficiency (NSE) coefficient 0.812 and as such has a more efficient and accurate estimation when compared to the LS-SVM model with $R^2=0.871$ and RMSE =0.097 and NSE=0.86. The analysis of Zarnagh station data shows $R^2=0.853$ and RMSE=0.0162, NSE= 0.854 for SVM and $R^2=0.903$ and RMSE =0.0091 and NSE=0.85 for ANN.

Keywords: Urmia Lake; Aji Chay River; TDS; ANN; LS-SVM.

1. Introduction

The evaluation and prediction of surface water quality is one of the central challenges in the water resource industries today. Due to the parametric complexity, the high cost of phenomenological water examinations [1-6] in both field and laboratory situations, and also the lack of experimental water quality data, many researchers have utilized data-driven techniques for water quality data retrieval [7-12].

Due to reasonable accuracy and relatively low cost, data-driven modelling techniques, known also as black box models, have become widespread in recent decades [12-16]. The ANN is one such black box model with a high potential for prediction in complicated non-linear systems

This technique requires a training or calibration phase, and generally estimates the amount of qualitative and quantitative parameters. It is relatively accurate in determining the standard deviation of data, and furthermore, has the capability of modelling the fundamental relationship between the inputs and outputs with a generalization potential [17,18]. The ANN models can be set with a limited number of input variables; conversely, a comprehensive number of records is needed to provide quality training data. This is essential as data-driven methods have a limited capability to provide accurate forecasts of events that are outside the range of the training dataset. Furthermore, when excessive numbers of variables are exploited as inputs, the most correlated variables logically dominate the model and, consequently, it is not possible to utilize all the physical knowledge or available measurements. Nonetheless, this can be solved by pre-processing techniques which select the most sensitive variables and, thus, reduce the input space [19, 20].

Another advanced soft method is the LS-SVM, proposed by Vapnik. It is based on the theory of statistical modeling which utilizes quadratic programming techniques. The LS-SVM has been used for time series estimation with acceptable levels of accuracy [21-24]. Nevertheless, this technique is more time consuming and has high computational requirements as a result of the required limited optimization programming. The quality parameters of surface water are usually defined as pH, ions, BOD, COD, total dissolved solids (TDS). The latter is the combined amount of inorganic and organic substances contained in water [25]. In the present study, ANN and LS-

SVM soft methods were developed to estimate the contamination of water resources with high accuracy and low cost. The case study was performed in the Aji Chay river, a water source that is mainly used for agriculture purposes and ecological stability. This study links to the comprehensive international efforts to protect Urmia lake. The Aji Chay River discharges into this highly endangered lake and potentially play a main role in preventing its desiccation [26]. As such, the developed ANN and LS-SVM techniques were used to model the variation of TDS of the Aji Chay River which has escalating salinity due to the increases in pollution levels.

The main objective of this research is the evaluation of Monthly Total Dissolved Solid using ANN and LS-SVM techniques in the Aji Chay River [26-28]. Previously numerous researchers have studied the behavior of intelligent techniques in the prediction of water quality parameters, but there are few studies that have compared both ANN and LS-SVM models to evaluate the total dissolved solids of the AjiChay river, and as such forms the principal novelty of this research.

2. ANNs and Least Squares Support Vector Machines

2.1 ANNs

ANNs can be used in water quality estimation and modeling. In this technique, feed-forward (FF) and back-propagation (BP) network patterns can be used for the present TDS simulation. It has been shown that the BP network pattern with a three-layered structure is desired for predicting and evaluating water resources problems. As indicated in Fig. 1, in order to estimate water quality time series, a three-layered feed-forward neural network (FFNN) provides a general framework for expressing nonlinear functional mapping between a set of input and output data. In this figure, i , j and k indicate input layer, hidden layer and output layer neurons, respectively, and the w variable is the used weight by the operating neuron. The term “feed-forward” refers to neuron connectivity being defined from a neuron in the input layer to other neurons in the hidden layer, or from a neuron in the hidden layer to neurons in the output layer. i.e. the input and output layers are not connected to each other. The clear explanation for an output value of a three-layered FFNN is given by Eq. 1 [29-32].

$$\hat{y}_k = f_{0[\sum_{j=1}^{MN} W_{kj} \cdot f_h(\sum_{i=1}^{NN} W_{ji} \cdot x_i + w_{j0}) + w_{ko}]} \quad (1)$$

where NN and MN are the number of the neurons in the input and hidden layers, $W_{ji}W_{ji}$ is a weight in the hidden layer connecting the i^{th} neuron in the input layer to the j^{th} neuron in the hidden layer, $w_{j0}w_{j0}$ is the bias for the j^{th} hidden neuron, f_h is the activation function of the hidden neuron, $W_{kj}W_{kj}$ is a weight in the output layer connecting the j^{th} neuron in the hidden layer to the k^{th} neuron in the output layer, $w_{ko}w_{ko}$ is the bias for the k^{th} output neuron, f_o is the activation function for the output neuron, x_i is the i^{th} input variable of the input layer and y_k are analysed and observed output variables, respectively. The weights aren't constant in the hidden and output layers, and their amounts can be varied during the network training.

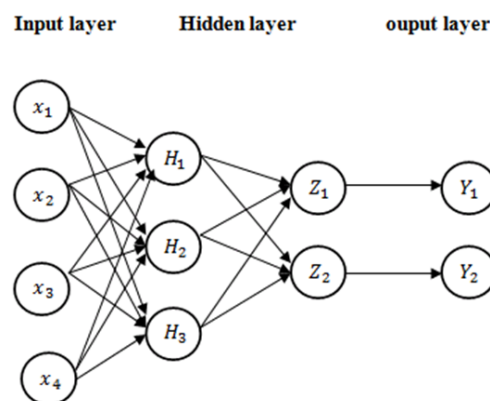


Fig. 1. A schematic diagram of a three-layer ANN.

2.2 Least Squares Support Vector Machines

For the available training data sets, the dimension (D) is defined as equation (2). Where x_k is the input data, y_k is the output data, and R , k and N are real numbers, kernel function and natural numbers respectively. Based on the Mercer principle, the kernel function $K(0, 0)$ is defined according to the mapping function $\phi(0)$:

$$D = (x_k, y_k) \quad x_k \in \mathbb{R}^n \quad D = (x_k, y_k) \quad x_k \in \mathbb{R}^n, y_k \in \mathbb{R}^n, k = 1, 2, \dots, N \quad (2)$$

$$\omega = \sum_{k=1}^N a_k \phi(x_k)$$

$$K(x_k, y_k) = \phi(x_i)^T \phi(x_i) \quad (3)$$

LS-SVM is solved in the main space; therefore, the following optimization problem is shown in Eq. 4 and Eq.5:

$$\min J(\omega, b, \xi) = \frac{1}{2} \omega^T \omega + \frac{1}{2} \gamma \sum_{k=1}^N \xi_k^2 \quad (4)$$

$$\text{sty}_k = \omega^T \cdot \phi(x_k) + b + \xi_k \quad (k = 1, 2, \dots, N) \quad (5)$$

The γ is an acceptable adjustment parameter, it is also the compromise between the training error and pattern complexity. The favourable function has a wide extension and generalization capability. The larger the value of the γ addresses, the smaller the regression error for the process. The variable ω is the weight matrix. LS-SVM describes a distinct loss function compared with the standard SVM, and it transfers the limit of the difference into the limit of the equation.

By using the Lagrange function:

$$L(\omega, b, \xi, a) = J(\omega, b, \xi) - \sum_{k=1}^N a_k [\omega^T \phi(x_k) + b + \xi_k - y_k] \quad (6)$$

$a_k \in R$ indicates Lagrange multipliers, the optimal a, b can be obtained from the following (Karush-Kuhn-Tucker) KKT conditions:

$$\frac{\partial L}{\partial \omega} = 0, \quad \frac{\partial L}{\partial b} = 0, \quad \frac{\partial L}{\partial \xi} = 0, \quad \frac{\partial L}{\partial a} = 0 \quad (7)$$

The following equations are obtained after implying the above conditions:

$$\omega = \sum_{k=1}^N a_k \phi(x_k) \quad (8)$$

$$\sum_{k=1}^N a_k = 0, \quad a_k = \gamma \xi_k \quad \text{and} \quad \omega^T \cdot \phi(x_k) + b + \xi_k - y_k = 0$$

By eliminating the variables ξ and ω , A linear system of functions is therefore obtained:

$$\begin{bmatrix} 0 & \theta^T \\ \theta & K + \gamma^{-1} I_n \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (9)$$

where $y = [y_1, y_2, \dots, y_k]^T$, $\theta = [1, \dots, 1]$, $a = [a_1, \dots, a_n]^T$ and $K_{ij} = \phi(x_i)^T \cdot \phi(x_j)$ ($i, j = 1, \dots, N$).

Values a and b are obtained by solving Eq. 10. This equation indicates the LS-SVM process function which is utilized for water quality investigation. [33].

$$f(x) = \sum_{i=1}^N a_i K(x, x_i) + b \quad (10)$$

The option and creation of kernel functions is an important stage which affects the implementation of LS-SVM. This can prepare an essential process to extend LS-SVM from a linear situation to nonlinear phase. There are different general kernel functions such as linear kernel, q-order polynomial kernel function, RBF function and sigmoid kernel function [34]. The (Gaussian) radial basis function kernel, or RBF kernel, is a popular kernel function used in various kernelized learning algorithms. The RBF kernel on two samples x and x' , represented as feature vectors in some input space and define as:

$$K(x, x') = \exp \left(-\frac{\|x - x'\|^2}{2\sigma^2} \right) \quad (11)$$

$\|x - x'\|^2$ may be recognized as the squared Euclidean distance between the two feature vectors. σ is a free parameter. An equivalent, but simpler, definition involves a parameter $\gamma = (1/2\sigma^2)$:

$$K(x, x') = \exp (-\gamma \|x - x'\|^2) \quad (12)$$

Since the value of the RBF kernel decreases with distance and ranges between zero (in the limit) and one (when $x = x'$), it has a ready interpretation as a similarity measure. The feature space of the kernel has an infinite number of dimensions; for $\sigma = 1$, its expansion is shown in eq.13.

$$\exp\left(-\gamma\|x-x'\|^2\right) = \sum_{j=0}^{\infty} \frac{x^t x'}{j!} \exp\left(-\left(\frac{1}{2}\right)\|x\|^2\right) \exp\left(-\left(\frac{1}{2}\right)\|x'\|^2\right) \quad (13)$$

3. Study area

This study was carried out in one of the largest watersheds in North-Western Iran, the AjiChay catchment with a drainage area of 51,876 km² and principal channel length of 265 km. This river originates from the southern regions of the Sabalan and the Qusheh-dagh Mountains and eventually drains to the basin of Urmia Lake. It is important to study of water quality in this river due to the fact that Ajichay river is one of the main available resources with average discharge of 0.6 m³/s as drinking water for the west Azarbaijan province. It also important in terms of environmental aspects because this river is drained into the Urmia lake. Therefore, the quality of Ajichay river water, affects drinking water of West Azarbaijan province area as well as quality of Urmia lake. To evaluate the accuracy of the developed ANN and SVM methods, 30-years of TDS data from the AjiChay River were collected for the Vanyar station. This sampling point is located at 38°07'00"N and 46°24'18"E and Zarnagh station located at 47°14'N 38'00E coordinates with an elevation of 1470 m above mean sea level (See Figure 2) [35].

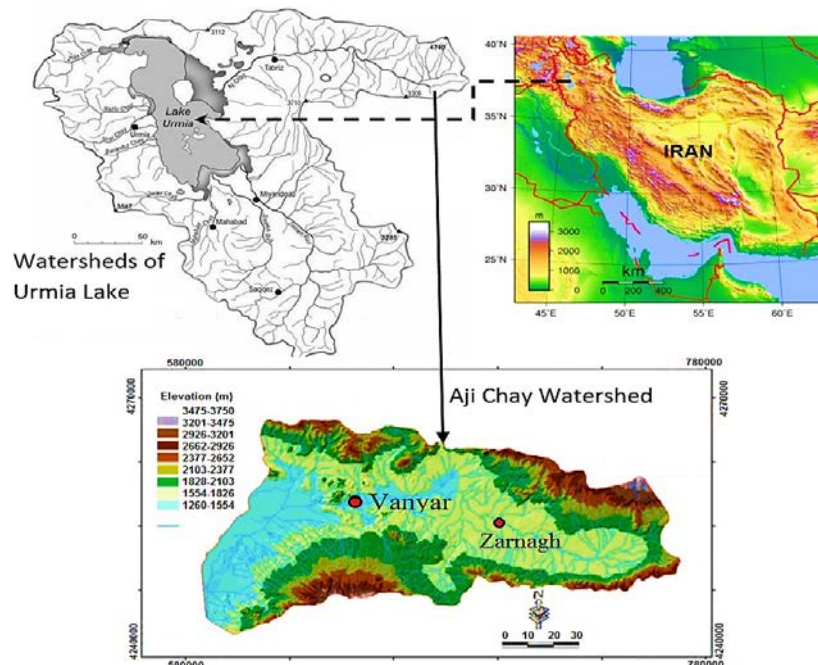


Fig. 2. Layout of the study area. Aji Chay River, Urmia Lake and Vanyar and Zarnagh stations

4. Results and discussion

The inputs of the models were the monthly Mg^{2+} (Magnesium), Na^+ (Sodium), Ca^{2+} (Calcium), Cl^- (Chloride), and SO_4^{2-} (Sulfate) with the output being the TDS. The time series of total observed TDS data (experimental data collected over 30 years from Vanyar and Zarnagh stations) were randomly classified into two separate parts, i.e., 20% testing, and 80% training, i.e., 20% testing and 80% training, to operate the applied methods. Table 1 shows the statistical specification of the collected TDS data from Vanyar and Zarnagh stations, and Table 2 represents the utilized input data classification of the identified models (M). A trial and error process is utilized to acquire the optimum percent of data for training and testing assortment. The purpose of this study is to determine the best value of performance criteria. Accordingly, three classifications of data were evaluated which includes 30–70 (i.e. 30% of data for testing and 70% of them for training data) 20–80 and 40–60 modes. Between these three classifications, state of 20-80 had optimum result in the criteria performance. Hence, in the both ANN and SVM models, 80% of observed data were used for training and 20% were used for the test issue.

ANN models require that input data sets be normalized between 0.05 and 0.95. This was accomplished applying equation 14 at the model initialization process.

$$x_{new} = 0.8 \frac{x - x_{min}}{x_{max} - x_{min}} + 0.1 \tag{14}$$

The variables x_{new} , x , x_{min} and x_{max} are the normalized values of the original parameter, the original data, and the minimum and maximum values in the data set, respectively. The standardization provides a better TDS variation as an acceptable result for testing regression correlation.

Table 1. Statistical characteristics of the input data used for testing and training

Statistical characteristics	TDS _{Vanyar}		TDS _{Zarnagh}	
	Test	Train	Test	Train
Data number	320	80	266	66
Max value	0.903	0.925	0.9	0.9
Min value	0.124	0.098	0.125	0.099
Average	0.298	0.017	0.305	0.221

4.1 The ANN model in preliminary test

For training and testing data settings for the ANN method, the observed data are selected to be under an 80% training vs. 20% testing pattern, and the ANN numerical evaluation is determined based on R² and RMSE. The applied ANN was trained using MATLAB version 7.8. The variables' types and their amounts utilized in the present ANN technique are listed in Table 2. In this table, Epochs addresses the number of training steps, MF is a membership function and “Trimf” is the abbreviation for triangular membership function (see Figure 3).

Table 2. Parameter types and values used in the developed ANN model

Gradient	Value
Transfer function	TANSIG
Network type	Feed-forward back prop
Number of layers	2
Adaption learning function	learn GDM
Performance function	MSE
Training function	Train LM

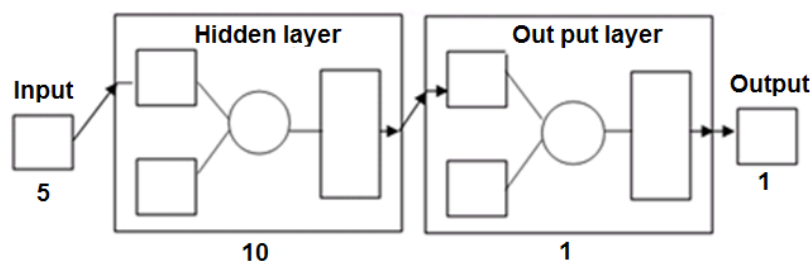


Fig. 3. The architecture of the developed ANN technique

4.2 LS-SVM model in preliminary test

The training and testing data set for the developed LS-SVM technique is similar to the ANN model (80%-20%), and the LS-SVM evaluation process is also taken as R² and RMSE and NSE. The kernel used in the LS-SVM numerical analysis was selected to be (RBF) with the regularization parameters.

4.3 Outputs of the analysis

The numerical TDS results in Tables 4 and 5 showed that the ANN method has R²=0.913, RMSE=0.0032, and NSE=0.812 for the M₂ pattern, and R²=0.871, RMSE=0.09 and NSE=0.86 for the M₂ pattern using the LS-SVM model in evaluation of Vanyar station dataset. For Zarnagh station: R²= 0.853, RMSE= 0.016 and NSE=0.853 is obtained for LS-SVM analysis. The out puts of ANN modelling for Zarnagh station are R²= 0.903, RMSE= 0.009,

and NSE= 0.806 (Table 6,7). These facts determine the capability and workability of ANN and LS-SVM as efficient soft-methods in estimating TDS. The lower RMSE and higher R^2 of the ANN technique compared to the corresponding results from the utilized LS-SVM indicates greater precision of the ANN in predicting the TDS. The relationship between observed and predicted TDS based on the two investigated approach were presented in form of scatter plot and statistical indicators as shown in Figure 4, Figure 5, Figure 6. Figure 7. Sensitivity analysis of TDS were performed based on more important parameters such as Mg^{2+} and Ca^{2+} as shown in Figure 8 and Figure 9. Statistical evaluation of the models to predict TDS using developed ANN method at Zarnagh and Vanyar stations are presented in Table 4 and Table 5 respectively. The same Statistical indices were calculated to evaluate the power prediction of TDS using developed LS-SVM method at Zarnagh and Vanyar stations as shown in Table 6 and Table 7. Statistical indicators for Sensitivity analysis of TDS for Mg^{2+} and Ca^{2+} illustrated in Table 8 and Table 9.

4.4 Sensitivity analysis

In order to find out which parameter had the most important influence on ANN outputs, sensitivity analysis was conducted on data set. Based on the configurations illustrated in Table 3, to be more specific, the effect of each parameter on the amount of output changes was investigated by keeping the other parameters constant. The results show that in Vanyar station, the parameter Ca^{2+} is the most important parameter in analysis due to more accuracy and the minimum error with the following values: $R^2=0.9431$, RMSE= 0.0043, NSE= 0.86. Table 8 shows the result of sensitivity analysis on Ca^{2+} .

Table 3. Applied input configurations for ANN, LS-SVM and Bayesian Network methods (applied on both Vanyar and Zarnagh stations data)

Model	Model input (s)
M ₁	Mg^{2+} - Na^+ - Cl^-
M ₂	Mg^{2+} - Na^+ - Ca^{2+} - Cl^- - SO_4^{2-}
-M ₃	Mg^{2+} - Na^+ - SO_4^{2-} - Ca^{2+}
M ₄	Ca^{2+} - Cl^-
M ₅	Na^+ - Ca^{2+}

Table 4. Statistical evaluation of the models to predict TDS using the developed ANN method- Vanyar station

Model	Train			Test		
	NSE	RMSE	R^2	NSE	RMSE	R^2
M ₂	0.871	0.016	0.931	0.812	0.003	0.913
M ₁	0.869	0.067	0.906	0.804	0.053	0.880
M ₃	0.822	0.091	0.827	0.782	0.077	0.814
M ₄	0.803	0.010	0.836	0.745	0.090	0.825
M ₅	0.790	0.054	0.794	0.713	0.117	0.730

Table 5. Statistical evaluation of the models to predict TDS using the developed ANN method -Zarnagh station

Model	Train			Test		
	NSE	RMSE	R^2	NSE	RMSE	R^2
M ₂	0.885	0.021	0.925	0.806	0.0091	0.903
M ₁	0.863	0.016	0.912	0.830	0.054	0.901
M ₃	0.852	0.013	0.922	0.803	0.080	0.892
M ₄	0.801	0.190	0.913	0.781	0.143	0.880
M ₅	0.868	0.290	0.894	0.864	0.201	0.853
M ₅	0.790	0.398	0.794	0.713	0.311	0.730

Table 6. Statistical evaluation of the models to predict TDS using the developed LS-SVM method -Vanyar station

Model	Training			Test		
	NSE	RMSE	R^2	NSE	RMSE	R^2
M ₂	0.850	0.09	0.895	0.860	0.097	0.871
M ₁	0.847	0.050	0.89	0.818	0.084	0.842
M ₃	0.839	0.093	0.81	0.790	0.160	0.783
M ₄	0.786	0.158	0.87	0.720	0.270	0.759
M ₅	0.825	0.290	0.80	0.805	0.330	0.704

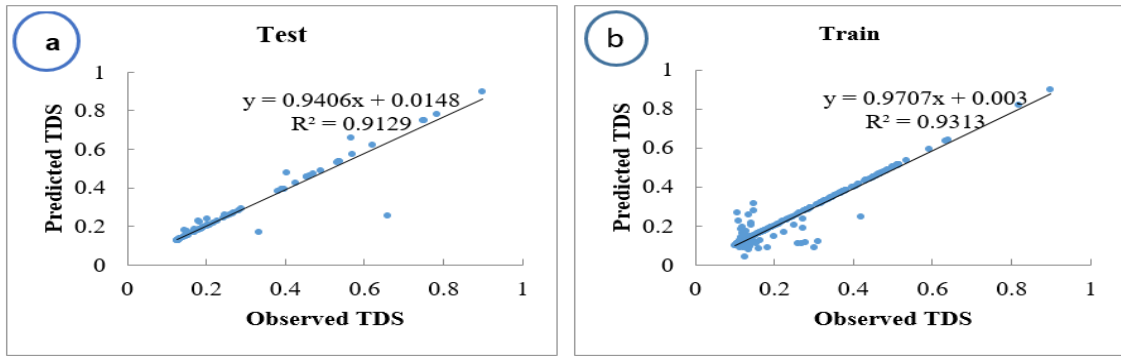


Fig. 4. Comparison of observed TDS with calculated via ANN a) test, b) train at Vanyar station

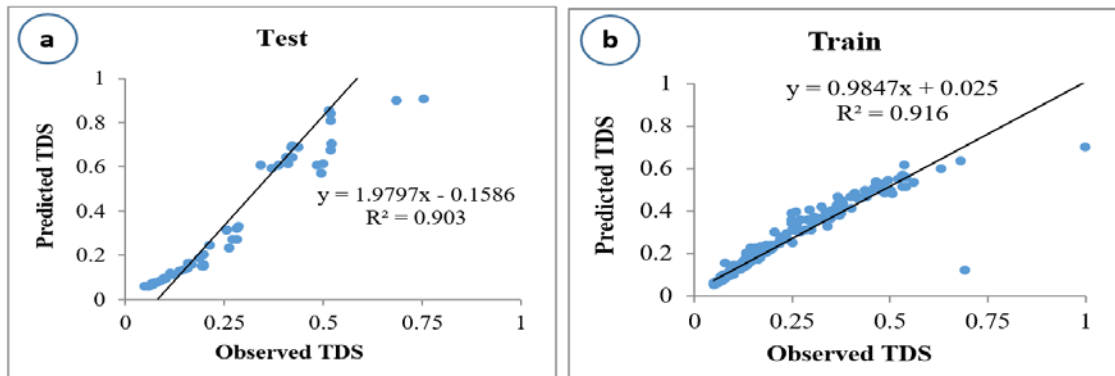


Fig. 5. Comparison of observed TDS with predicted via ANN. a) test, b) train at Zarnagh station

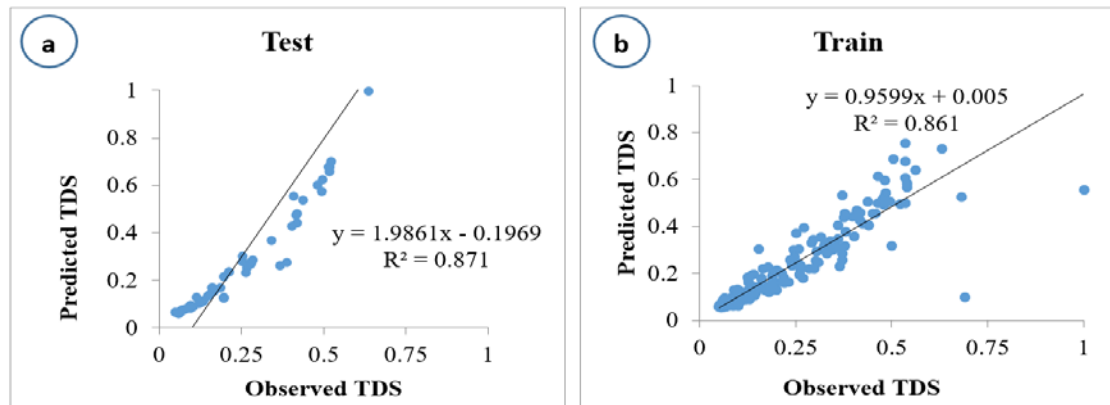


Fig.6. Comparison of observed TDS with calculated via LS-SVM. a) Test, b) Train Vanyar

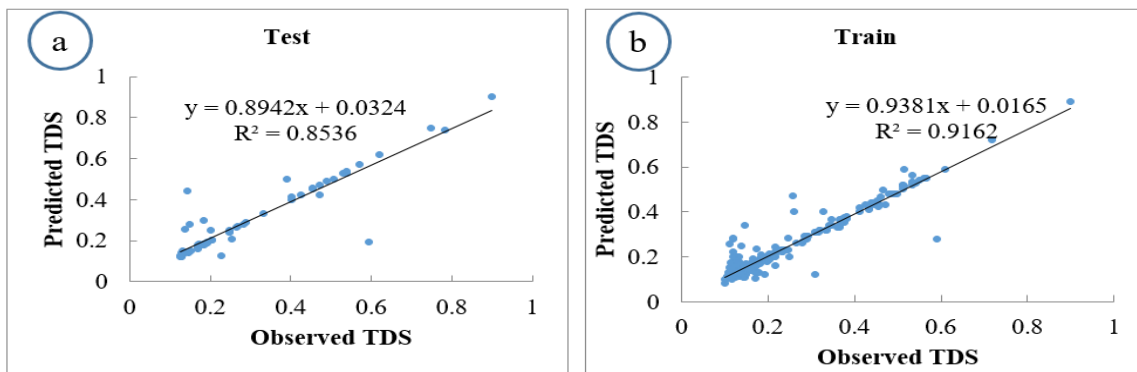


Fig. 7. Comparison of observed TDS with predicted via LS-SVM a) test, b) train at Zarnagh station

Table 7. Statistical evaluation of the models to predict TDS using the developed LS-SVM method-Zarnagh station

Model	Train			Test		
	NSE	RMSE	R ²	NSE	RMSE	R ²
M ₂	0.879	0.01	0.916	0.85	0.016	0.8535
M ₁	0.874	0.017	0.897	0.830	0.029	0.8472
M ₃	0.849	0.021	0.902	0.803	0.038	0.8295
M ₄	0.825	0.153	0.844	0.781	0.250	0.8017
M ₅	0.819	0.285	0.802	0.864	0.325	0.760

Table 8. The Sensitivity Analysis of water quality parameters (Vanyar station)

Test	R ²	Ca-TDS	Cl-TDS	Mg-	Na-TDS	SO ₄ -
		RMSE	0.004	0.011	0.012	0.028
	NSE	0.890	0.823	0.780	0.730	0.700
Train	R ²	0.963	0.873	0.819	0.798	0.870
	RMSE	0.001	0.008	0.011	0.017	0.024
	NSE	0.910	0.900	0.874	0.768	0.810

Table 9. The Sensitivity Analysis of water quality parameters (Zarnagh station)

Test	R ²	Ca-TDS	Cl-TDS	Mg-	Na-TDS	SO ₄ -
		RMSE	0.0234	0.08	0.0026	0.037
	NSE	0.885	0.823	0.900	0.740	0.723
Train	R ²	0.940	0.927	0.932	0.893	0.842
	RMSE	0.012	0.023	0.004	0.024	0.059
	NSE	0.891	0.821	0.915	0.719	0.863

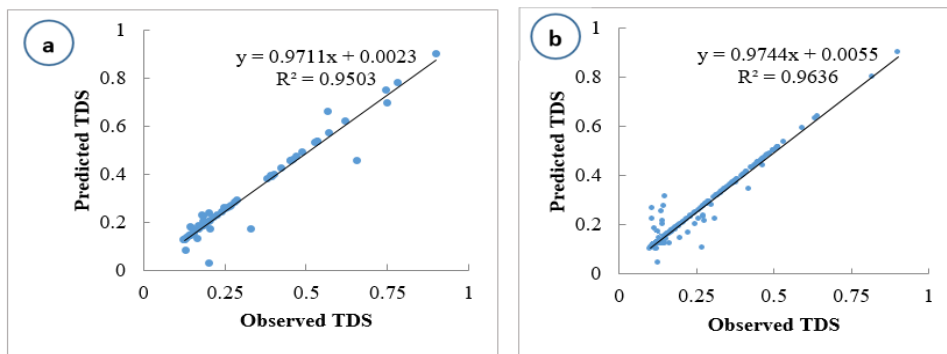


Fig. 8. Sensitivity analysis of the outputs of ANN (Vanyar station)-The most effective parameter(Ca²⁺)

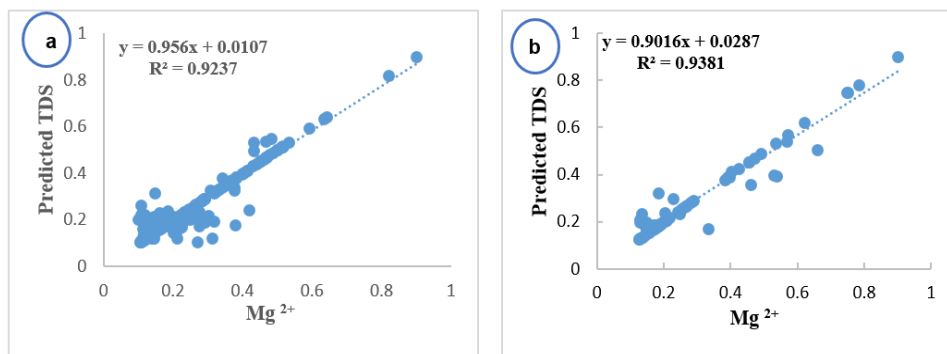


Fig. 9. Sensitivity analysis of the outputs of ANN (Zarnagh station)-The most effective parameter (Mg²⁺)

5. Conclusion

This study developed ANN and LS-SVM techniques, to predict the TDS of rivers, i.e., Aji Chay River. To assess the performance of used methods, Root mean square error (RMSE), correlation coefficient (R^2) and NSE were applied as performance indicators for the analysis. The obtained results demonstrated that the ANN technique has a lower RMSE (0.0032) and higher R^2 (0.913) with a reasonable NSE coefficient (0.812) for the testing data of Vanyar station, compared to that obtained from the LS-SVM method with ($R^2=0.871$ and $RMSE=0.097$ and $NSE=0.86$). The observed results have a high correlation with the estimated analytical data, and a high accuracy for the ANN meta-method in estimating TDS compared to the results from the LS-SVM method. The analysis of Zarnagh station data shows $R^2=0.903$ and $RMSE=0.0091$, $NSE=0.806$ for ANN and $R^2=0.853$, $RMSE=0.016$ and $NSE=0.85$ for SVM. The novelty of the developed methods can be used for predicting the TDS of other surface water bodies.

6. Acknowledgements

The authors would like to express their great appreciation to the Water and Wastewater Company of East Azerbaijan Province, Iran, for their valuable support and providing data for this study.

7. References

- [1] Chang FJ, Tsai YH, Chen PA, Coynel A, Vachaud G. Modeling water quality in an urban river using hydrological factors—Data driven approaches. *Journal of Environmental Management*. 2015;151:87-96.
- [2] Verma AK, Singh TN. Prediction of water quality from simple field parameters. *Environmental Earth Sciences*. 2013;69(3):821-829.
- [3] Pinho JL, Vieira JP, Do Carmo JA. Hydroinformatic environment for coastal waters hydrodynamics and water quality modelling. *Advances in Engineering Software*. 2004;35(3-4):205-222.
- [4] Zhang G, Li X, Wu S, Gu P. Effect of source water quality on arsenic (V) removal from drinking water by coagulation/microfiltration. *Environmental Earth Sciences*. 2012;66(4):1269-1277.
- [5] Sarala Thambavani D, Uma Mageswari TS. Water quality indices as indicators for potable water. *Desalination and Water Treatment*. 2014;52(25-27):4772-4782.
- [6] Sharifinia M, Ramezanpour Z, Imanpour J, Mahmoudifard A, Rahmani T. Water quality assessment of the Zarivar Lake using physico-chemical parameters and NSF-WQI indicator, Kurdistan Province-Iran. *International Journal of Advanced Biological and Biomedical Research*. 2013;1(3):302-312.
- [7] Khataee AR, Vafaei F, Jannatkah M. Biosorption of three textile dyes from contaminated water by filamentous green algal *Spirogyra* sp.: Kinetic, isotherm and thermodynamic studies. *International Biodeterioration & Biodegradation*. 2013;83:33-40.
- [8] Kobya M, Soltani RD, Omwene PI, Khataee A. A review on decontamination of arsenic-contained water by electrocoagulation: Reactor configurations and operating cost along with removal mechanisms. *Environmental Technology & Innovation*. 2020;17:100519.
- [9] Khataee A, Azamat J, Bayat G. Separation of nitrate ion from water using silicon carbide nanotubes as a membrane: Insights from molecular dynamics simulation. *Computational Materials Science*. 2016;119:74-81.
- [10] Saboe D, Ghasemi H, Gao MM, Samardzic M, Hristovski KD, Boscovic D, Burge SR, Burge RG, Hoffman DA. Real-time monitoring and prediction of water quality parameters and algae concentrations using microbial potentiometric sensor signals and machine learning tools. *Science of the Total Environment*. 2020;10:142876.
- [11] García-Alba J, Bárcena JF, Ugarteburu C, García A. Artificial neural networks as emulators of process-based models to analyse bathing water quality in estuaries. *Water Research*. 2019;150:283-295.
- [12] Thair SK, MJ AH, Ayad S. Prediction of water quality of Euphrates river by using artificial neural network model (spatial and temporal study). *International Research Journal of Natural Sciences*. 2014;2(3):25-38.
- [13] Maedeh AP, Mehrdadi N, Bidhendi GN, Abyaneh HZ. Application of artificial neural network to predict total dissolved solids variations in groundwater of Tehran plain. Iran. *Int. J. Environ. Sustain*. 2013;2(1):10-20.
- [14] Nasr M, Zahran HF. Using of pH as a tool to predict salinity of groundwater for irrigation purpose using artificial neural network. *The Egyptian Journal of Aquatic Research*. 2014;40(2):111-115.
- [15] Gazzaz NM, Yusoff MK, Aris AZ, Juahir H, Ramli MF. Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors. *Marine pollution bulletin*. 2012;64(11):2409-2420.
- [16] Akhtar MK, Corzo GA, Van Andel SJ, Jonoski A. River flow forecasting with artificial neural networks using satellite observed precipitation pre-processed with flow length and travel time information: case study of the Ganges river basin. *Hydrology and Earth System Sciences*. 2009;13(9):1607.

- [17] Tanty R, Desmukh TS. Application of artificial neural network in hydrology—A review. *Int. J. Eng. Technol. Res.* 2015;4:184-188.
- [18] Gebler D, Wiegleb G, Szoszkiewicz K. Integrating river hydromorphology and water quality into ecological status modelling by artificial neural networks. *Water research.* 2018;139:395-405.
- [19] Malik A, Kumar A, Singh RP. Application of heuristic approaches for prediction of hydrological drought using multi-scalar streamflow drought index. *Water Resources Management.* 2019;33(11):3985-4006.
- [20] Adnan RM, Yuan X, Kisi O, Yuan Y. Streamflow forecasting using artificial neural network and support vector machine models. *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS).* 2017;29(1):286-294.
- [21] Kahani M, Ahmadi MH, Tatar A, Sadeghzadeh M. Development of multilayer perceptron artificial neural network (MLP-ANN) and least square support vector machine (LSSVM) models to predict Nusselt number and pressure drop of TiO₂/water nanofluid flows through non-straight pathways. *Numerical Heat Transfer, Part A: Applications.* 2018;74(4):1190-1206.
- [22] Guo T, He W, Jiang Z, Chu X, Malekian R, Li Z. An improved LSSVM model for intelligent prediction of the daily water level. *Energies.* 2019;12(1):112.
- [23] Rostami S, Rashidi F, Safari H. Prediction of oil-water relative permeability in sandstone and carbonate reservoir rocks using the CSA-LSSVM algorithm. *Journal of Petroleum Science and Engineering.* 2019;173:170-186.
- [24] Kardani MN, Baghban A. Utilization of LSSVM strategy to predict water content of sweet natural gas. *Petroleum Science and Technology.* 2017;35(8):761-767.
- [25] Mudgal LK. Comparative studies of physico-chemical parameters of two reservoirs of Narmada River, MP, India. *Current World Environment.* 2013;8(3):473.
- [26] Niazi F, Mofid H, Modares NF. Trend analysis of temporal changes of discharge and water quality parameters of Ajichay River in four recent decades. *Water Quality, Exposure and Health.* 2014;6(1-2):89-95.
- [27] Damadipour M, Nazarpour M, Alami MT. Evaluation of particle size distribution using an efficient approach based on image processing techniques. *Iranian Journal of Science and Technology, Transactions of Civil Engineering.* 2019;43(1):429-441.
- [28] Danesh-Yazdi M, Ataie-Ashtiani B. Lake Urmia crisis and restoration plan: Planning without appropriate data and model is gambling. *Journal of Hydrology.* 2019;576:639-651.
- [29] Akhgar A, Toghraie D, Sina N, Afrand M. Developing dissimilar artificial neural networks (ANNs) to prediction the thermal conductivity of MWCNT-TiO₂/Water-ethylene glycol hybrid nanofluid. *Powder Technology.* 2019;355:602-610.
- [30] Ozel HU, Gemici BT, Gemici E, Ozel HB, Cetin M, Sevik H. Application of artificial neural networks to predict the heavy metal contamination in the Bartin River. *Environmental Science and Pollution Research.* 2020:1-8.
- [31] Asadi H, Shahedi K, Jarihani B, Sidle RC. Rainfall-runoff modelling using hydrological connectivity index and artificial neural network approach. *Water.* 2019;11(2):212.
- [32] Kim HG, Hong S, Jeong KS, Kim DK, Joo GJ. Determination of sensitive variables regardless of hydrological alteration in artificial neural network model of chlorophyll a: case study of Nakdong River. *Ecological Modelling.* 2019;398:67-76.
- [33] Lu Y, Yin Q, Li H, Sun H, Yang Y, Hou M. The LS-SVM algorithms for boundary value problems of high-order ordinary differential equations. *Lu et al. Advances in Difference Equations.* 2019; 2019:195
- [34] Chang Q, Chen Q, Wang X. Scaling Gaussian RBF kernel width to improve SVM classification. In: 2005 International Conference on Neural Networks and Brain. 2005. IEEE. (Vol. 1, pp. 19-22).
- [35] Mansouri H, Jorkesh Z, Ajalloeian R, Sadeghpour AH. Investigating effects of water salinity on geotechnical properties of fine-grained soil and quartz in a sandstone case study: Ajichay project in northwest Iran. *Bulletin of Engineering Geology and the Environment.* 2017;76(3):1117-1128.



© 2020 by the author(s). This work is licensed under a [Creative Commons Attribution 4.0 International License](http://creativecommons.org/licenses/by/4.0/) (<http://creativecommons.org/licenses/by/4.0/>). Authors retain copyright of their work, with first publication rights granted to Tech Reviews Ltd.