



VICTORIA UNIVERSITY
MELBOURNE AUSTRALIA

Investigation of quantitative and qualitative changes in groundwater of Ardebil plain using ensemble artificial intelligence-based modeling

This is the Published version of the following publication

Sarreshtedar, Ayda, Sharghi, Elnaz, Afkhaminia, Amin, Nourani, Vahid and Ng, Anne WM (2022) Investigation of quantitative and qualitative changes in groundwater of Ardebil plain using ensemble artificial intelligence-based modeling. *Water Supply*, 22 (9). pp. 7140-7157. ISSN 1606-9749

The publisher's official version can be found at
<https://iwaponline.com/ws/article/22/9/7140/89833/Investigation-of-quantitative-and-qualitative>

Note that access to this version may require subscription.

Downloaded from VU Research Repository <https://vuir.vu.edu.au/47279/>

Investigation of quantitative and qualitative changes in groundwater of Ardebil plain using ensemble artificial intelligence-based modeling

Ayda Sarreshtedar^a, Elnaz Sharghi^b, Amin Afkhaminia^{ib a,*}, Vahid Nourani^{a,c} and Anne Ng^c

^a Center of Excellence in Hydroinformatics and Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran

^b Department of Water Resources Engineering, Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran

^c College of Engineering, Information Technology and Environment, Charles Darwin University, Ellengowan, Brinkin, NT 0810, Australia

*Corresponding author. E-mail: afkhaminia74@gmail.com

 AA, 0000-0002-0857-694X

ABSTRACT

Groundwater is an essential source to supply water for various sectors. This paper aimed to predict the quantitative and qualitative changes in groundwater over time and to evaluate the efficiency of different modeling methods. This study is based on three steps. In the first step, quantitative and qualitative piezometers were clustered by the Growing Neural Gas Network (GNG) method, and the central piezometer of each cluster was used on behalf of each cluster. In the second step, four different Artificial Intelligence (AI) models were applied, namely Artificial Neural Network (ANN), Adaptive Neuro-Fuzzy Inference System (ANFIS), Support Vector Machine (SVM), and Emotional Artificial Neural Network (EANN). As a post-processing approach three different ensemble methods were used: simple average ensemble (SAE), weighted average ensemble (WAE), and nonlinear neural network ensemble (NNE). In the third step, the outputs of single AI models were used to enhance the evaluation results. Therefore, the results demonstrate that the NNE led to reach the better performance for three GWL, TDS, and TH parameters up to 37, 29, and 23% on average, respectively. Study results will lead to the improvement of AI applications in groundwater research and will benefit groundwater development plans.

Key words: black-box, clustering, groundwater parameters, model combination

HIGHLIGHTS

- Artificial intelligence methods were used to predict quantitative and qualitative changes in groundwater.
- The GNG method was used for clustering.
- Ensemble artificial intelligence-based modeling was employed to enhance the individual modeling results.
- The results show better performance when using ensemble artificial intelligence-based modeling.

1. INTRODUCTION

Throughout the world, groundwater is a major source of fresh water for many uses like drinking, agricultural and industrial purposes. According to [Gintamo *et al.* \(2022\)](#), population growth and increased groundwater consumption pose considerable threats to aquifers. Today, the development of communities living in semi-arid areas, where food production is tightly controlled by the amount and distribution of rainfall and groundwater, depends on people's effort to manage limited water resources. Thus, it can be seen that in the near future, the interaction between the growth of urbanization and the application of an efficient management method to conserve water resources will continue. Due to costly and time-consuming solutions to groundwater problems and even in many cases the impossibility of a complete solution, analysis of these valuable resources is vital.

Neural networks have also been applied successfully for studies of quantitative and qualitative changes in groundwater. Groundwater modeling is complicated by the influence of natural and/or human factors such as complexity, nonlinearity, multi-scales, and randomness, all of which affect quantitative and qualitative assumptions ([Nourani *et al.* 2014](#)). These specific models cannot provide explicit equations and physical relationships which are not easily predictable. So in complex phenomena, black-box models are generally used. The efficiency of the model is strongly influenced by the quality and

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

quantity of input and output data. Semi-distributive or conceptual models are the interface between the above two models (Nourani 2017). The use of Black-box AI (Artificial intelligence) approaches has gained popularity in current years because of its good performance, specifically [i.e. Artificial Neural Networks (ANN), Adaptive Neuro-Fuzzy Inference Systems (ANFIS), Support Vector Machines (SVM), and, Emotional Artificial Neural Network (EANN)] which are capable of handling complex phenomena have become prominent in water resource issues. The artificial neural network has been used as a different approach to estimating aquifer water quality (Maier & Dandy 1996). Moreover, Derbela & Nouriri (2020), employed ANN models to predict the dynamic changes in piezometric levels in Nebhana aquifers. Correlation analysis demonstrated that the piezometric levels were toughly influenced by monthly rainfall, evapotranspiration, and initial water table level, Also this study revealed that ANN models can be adopted for future groundwater prediction to estimate trends in piezometric levels. Dehghani & Torabi Poudeh (2022) concluded that ANN's performance can be enhanced effectively through combination with other models. During another study analyzing groundwater quality, Shwetank & Chaudhary (2022) used four ANFIS models to identify the best one. The results, however, determined that all four models were dependable. SVM is another AI-based model for reliable estimation of groundwater level and quality. This method tries to minimize the operation risk. Nordin *et al.* (2021) in a comprehensive study, compared four widely used AI methods [ANN, ANFIS, EA (evolutionary algorithm), and SVM] for modeling and forecasting groundwater quality. As a result of their generalization and faster optimization, EA, and SVM have surpassed ANN and ANFIS in predictive modeling applications. Hydrological studies have been successfully implemented using EANN, a novel version of the classic ANN model. As part of their research in the area of modeling rainfall-runoff in watersheds, Nourani (2017) applied an emotional Artificial Neural network method. He predicted the runoff level by two methods of modeling named EANN and FFNN using the parameters of two catchments with different climate conditions and the results showed that EANN can have 13 and 34% higher efficiency in terms of training and validation, respectively. In addition, in the field of GWL, another study was conducted to assess the efficacy of the EANN model for forecasting one month in the field of GWL.

Today, soft computing tools and artificial intelligence, by accepting that the results can be inaccurate in some parts and by focusing on the human mind and nature, are used to model nonlinear hydraulic and hydrological processes. But using different models to solve a specific issue may lead to distinct results. It is obvious that some models will show higher performance than other models. In other words, no single model is better than another for the analysis and study of hydrological processes (Shamseldin *et al.* 1997). Using a collection of individual models in combination, the ensemble method can provide results with high predictive accuracy and low errors using the advantages of all models at the same time. Sharghi *et al.* (2019) used the ensemble method and a 19% increase was gained by using these models. Nadiri *et al.* (2015) incorporated four models including a feed-forward neural network (FFNN), backpropagation neural network, ANFIS, and, SVM to predict GWL in Meshgin plain. Despite the ability of all models, the results show the superiority of the ensemble method over any of the individual AI models.

Undoubtedly, the best way to understand the behavior of an aquifer system is through long-term research for each region. Therefore, with data mining and clustering tools, indicators can be obtained with acceptable accuracy, based on which necessary decisions can be made regarding aquifer management (Al-adamat *et al.* 2003). Classifying different objects has always been one of the most important human activities and this is part of the learning process that explains why cluster analysis in many cases is considered a branch of cognitive pattern and artificial intelligence. One of the problems in studying an aquifer through AI methods is the number of observation wells in the aquifer, the analysis of each one is time-consuming alone. One of the statistical methods for presenting the distribution map and quantitative and qualitative distribution of several parameters in groundwater aquifers is the use of cluster analysis. The idea of clustering is to collect objects that are similar with the least possible difference between them in a set of specific groups (Kim *et al.* 2020). Clustering is a way of categorizing multidimensional inputs into homogenous groups, which can be extracted from unlabeled inputs to form clusters with the greatest degree of similarity within groups and the greatest degree of dissimilarity between groups (Nourani & Kalantari 2010). There are several clustering methods commonly used, e.g. k-means (Hsu *et al.* 2002) and C-means (Ayvaz *et al.* 2007) which are typically based on the number of clusters (Nourani *et al.* 2013). Recent works on unsupervised clustering using self-organizing (SOM) maps have been able to find many applications in hydrology (e.g., Hsu *et al.* 2002; Kalteh *et al.* 2008; Toth 2009; Nourani & Parhizkar 2013; Nourani *et al.* 2013). SOM-based Growing Neural Gas Network (GNG) algorithms allow for the learning of complex relationships without knowing any prior knowledge. A topology is created by GNG using the competitive Hebbian tutorial. There is a limit to how many data points can be clustered using most existing techniques, but SOM and GNG provide a solution by quantifying the data, and despite the similarities between the two methods, GNG has been proven to be the best option thus far (Ventocilla & Riverio 2020; Ventocilla *et al.* 2021). This

framework has been used in a variety of fields as a tool for multipurpose analysis specifically because of its high level of flexibility for recognizing complex patterns (Shi *et al.* 2014; Viejo *et al.* 2014; Santos & Nascimento 2016; Abdi *et al.* 2017). Considering the research background and the ability of the GNG method in clustering and pattern recognition also, being aware of no groundwater investigations done with this method at this time, in this article, the mentioned method has been hired for clustering groundwater piezometers. Due to the successes achieved in hydrological studies using black-box models, the focus of this research is on investigating the efficiency of AI methods to predict and investigate changes in groundwater parameters. On the other hand, the vacancy of a deep and comprehensive study that can evaluate the ability of mentioned and especially ensemble methods, on the groundwater of Ardabil plain is felt. Groundwater studies have not previously employed the GNG clustering method. Also, this report is the first to combine four AAN, ANFIS, SVM, and EANN techniques in the field of groundwater.

In this study, the GNG clustering method was used to investigate fluctuations in groundwater level and quality parameters such as TDS and TH to reduce the volume of inputs and classify piezometric stations. A primary objective of the study is to examine how the ANN, ANFIS, SVM, and, EANN techniques perform in groundwater studies. In this regard, four single AI-based models, which are the feed-forward neural network (FFNN), adaptive neuro-fuzzy inference system (ANFIS), support vector machine (SVM), and emotional artificial neural network (EANN) were used for groundwater parameters modeling. Thereafter, the results of individual models were combined using three different ensemble methods named Simple Average Ensemble (SAE), Weighted Average Ensemble (WAE), and Neural Network ensemble (NNE) to improve the predictive performance. In general, the effectiveness of combined models was compared to individual models.

2. MATERIALS AND METHODS

2.1. Study area and data Set

Ardabil plain (38°03'–38°27'N and 47°55'–48°20'E) is situated in north-western Iran and covers an area of 990 km². In Figure 1 the study area is shown geographically. The wettest and driest months are May and August, respectively. The

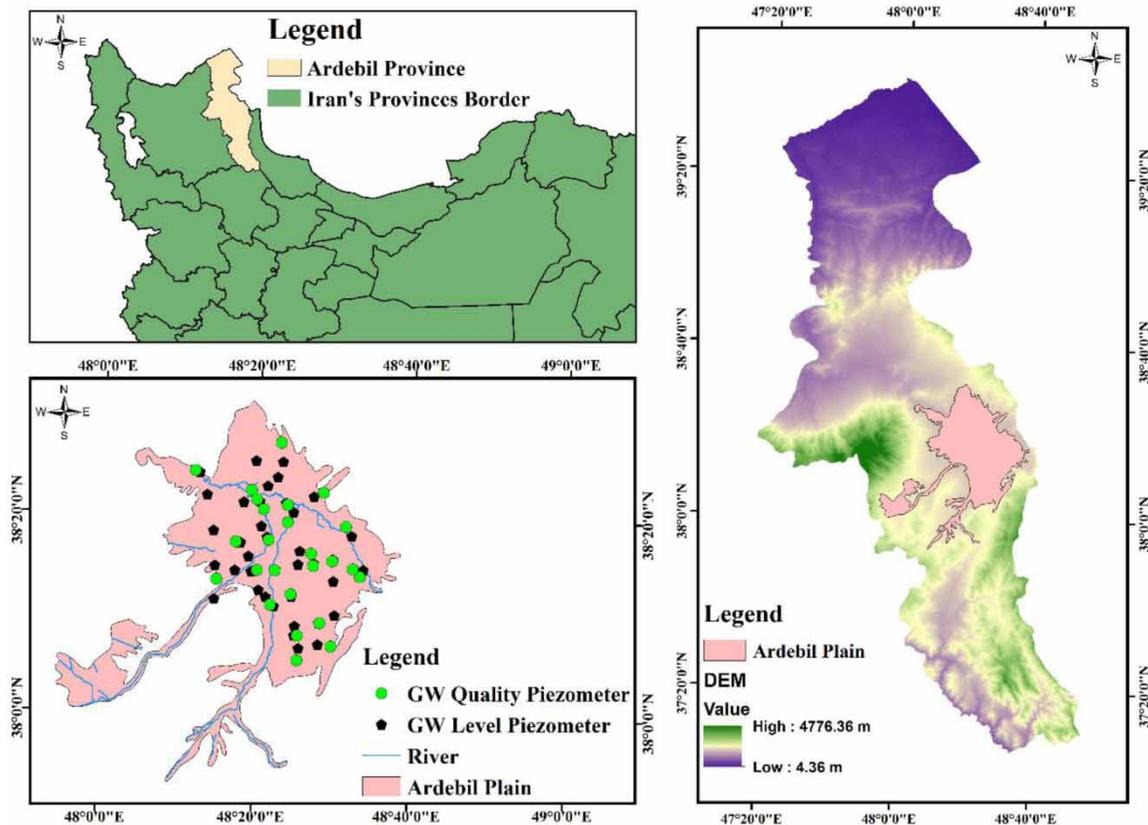


Figure 1 | Study area.

GWLs are analyzed with 39 piezometers on the plain. The monthly GWLs information is available for quantitative analysis. The chemical properties of the available data include the sum of the major cations and anions present in the water, the electrical conductivity (EC), the sodium adsorption ratio (SAR), and the total hardness (TH), and the total solid dissolved in water (TDS). Groundwater quality variables are measured twice a year in wet and dry months. Data used for the application covers 20 years (1996–2019) of observations, from which 70 percent are used for training and the rest for validation randomly. All data used in this study were obtained from Ardabil Regional department.

AI-based modeling relies on the careful selection of inputs to produce optimal results. Thus, to increase modeling efficiency and optimize the input layer, quantitative and qualitative data of groundwater were estimated and, clustered using the GNG method. Lately, studies proved that the highest correlation between the present groundwater level values was with its previous values. According to this explanation based on a trial and error procedure, previous time steps ($t - 1$, $t - 2$, ...) until lag 6 were considered as inputs, while the current groundwater level (t) was the output of the model. The proposed methodology takes three steps into account. The spatial clustering method GNG was employed, first, to classify quantitative piezometers using the GWL parameter, and then to classify qualitative piezometers based on the GWL and quality parameters. It is noteworthy that the input data were mapped to a value between -1 and 1 before clustering. In step 2, first, in the qualitative field, each cluster was fitted with data from the central piezometer. Comparing different input candidates and the output of selecting dominant inputs and lags of the model using Mutual Information (MI) showed that groundwater quality parameters (TDS and TH) were primarily influenced by variations in (So₄, Cl, Ca, Mg, TDS, TH) and GWL fluctuations. Then, the dominant inputs for quantitative and qualitative analysis of groundwater in each cluster are imposed into the ANN, ANFIS, SVR, and EANN, respectively. In the third step, to improve the obtained results of modeling, the outputs of individual models were considered as the dominant input of the combined model. The following sections contain a summary of the results obtained. As a result, in this study, three ensemble methods were used to boost single model accuracy. Figure 2 depicts the ensemble approach schematically.

2.2. Growing neural gas network-GNG

The GNG algorithm can be considered as one of the most powerful ways to train unsupervised and incremental neural networks as it can build and update the network's structure without depending on size and shape information. This method determines the structure of the data by creating a topological map. The formation of the algorithm begins with the creation of a diagram consisting of two nodes, each node is created using a random instance. During this process, the position of the nodes evolves so that the node position vectors are updated to model the data topology, and then large structures are subdivided into smaller structures, each of which forms a cluster shows (Subba Rao 2012). This network is primarily for adding new nodes to a small primary network (Cirrincione *et al.* 2012). The GNG neurons compete for alignment with the input data set to choose the most similar ones (Morell *et al.* 2014).

2.3. ANN

Part of the inspiration for this network comes from the way the biological neural system processes data and information, to learn and generate knowledge. One of the outstanding features of the artificial neural network is its ability to be used in any

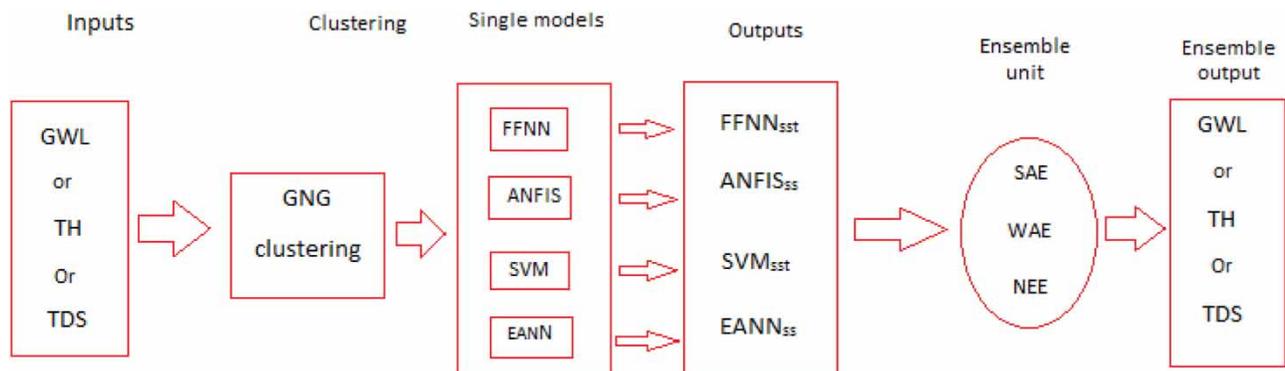


Figure 2 | Schematic view of the processes.

case that requires learning linear and nonlinear mapping. ANN must first know what it does not know, that is being taught first, and then expect to be able to solve the problem in this way. Generally, artificial neural networks are comprised of three layers, the input, the middle and, the output layers. Several processing elements are found in each layer known as neurons. As each neuron receives an input signal, it converts it into an activated value by using an activation function called a sigmoid. The observational data are given to the input layer and then the input data are processed in the middle layer, and finally, the values predicted by the neural network are obtained from the output layer (Nourani *et al.* 2011). In comparison with all other methods, feed-forward and back-forward propagation neural networks trained using Levenberg–Marquardt (FFNN-LMB) are the most practical and useful. In terms of predicting groundwater levels, these two approaches were more effective and, successful (Daliakopoulos *et al.* 2005; Sujatha & Kumar 2010; Chitsazan *et al.* 2013).

2.4. ANFIS

This system was first introduced in 1993 by Jang (Jang 1993). To form this system, two structures of neural networks and a fuzzy system have been combined. This method uses each method's advantage to increase decision-making power in conditions of uncertainties (Abraham 2005). The analysis operation is performed by the fuzzy inference operator. Various fuzzy inference operators that can be used for this purpose, among these operators, Sugeno and Mamdani are very common (Lin & Chen 2005; Nourani *et al.* 2011). This method is done by Fuzzy the inputs through the membership functions, in which the input values are mapped by a relation to the range of [0–1]. In the ANFIS structure, each approach has two or more membership functions, to which these system parameters must be sent. Figure 3 shows the three trapezoidal, Gaussian, and triangular membership functions. An example of a simple membership function is a triangular function with straight lines. The Trapezoidal membership function is a triangular function cut at the top, these two membership functions have the advantage of simplicity. The most popular membership function is the Gaussian function, which has the advantage of softness and non-zero value at all points.

2.5. SVM

One of the relatively new machine learning methods is the support vector machine method, which was introduced in 1995 by Kurtz and Vapnik. The mentioned network can be considered a method of supervised learning. As opposed to other neural networks, a support vector machine evaluates the operational risk of the objective function, instead of minimizing its modeling error. In SVM, instead of using a line to separate, a confidence margin is used. In Figure 4 H_1 line does not properly separate the two categories of data shown. Contrariwise H_2 and H_3 lines do this operation correctly. The difference between H_2 and H_3 lines is the operational risk or the risk of not being properly classified. Perceptron neural networks do not distinguish between these two dividing lines, but for a support vector machine network, H_3 is better than H_2 . Also, the main goal of this network is to reach H_3 . Another version of SVM for regression was provided in 1997 by Vapnik (Vapnik *et al.* 1996). SVM, the most common form of which is the support vector regression (SVR) is the method used in this study (Basak *et al.* 2007). The general objective of this procedure is to perform a linear regression and then apply a kernel to result to create a nonlinear regression.

2.6. EANN

The EANN model is a new generation of classical ANN model. A component of this model involves an artificial emotion unit that release hormones to regulate neuron function, as well as modifying hormonal weights based on values generated by neuron, input and output. Figure 5, illustrates how each hidden layer of EANN information undergoes constant transformation between input and output layers. In addition, these nodes generate the dynamic hemorrhages H_a , H_b , and H_c .

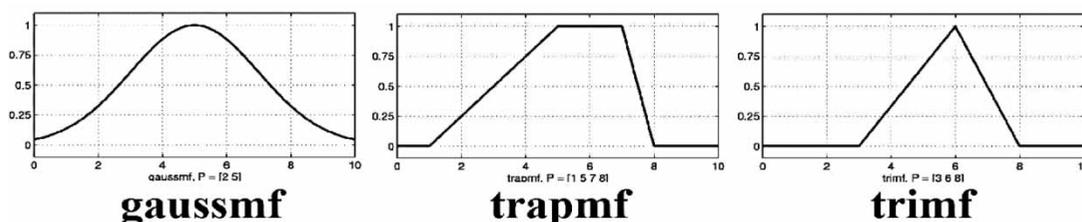


Figure 3 | Trapezoidal, Gaussian and triangular membership functions.

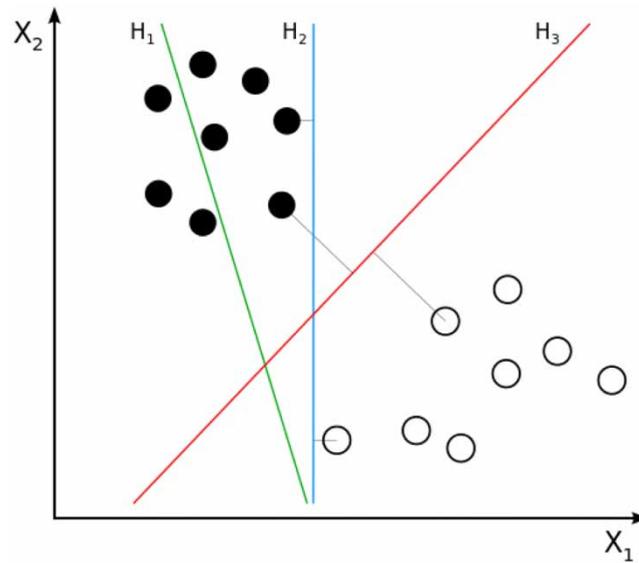


Figure 4 | Types of separators.

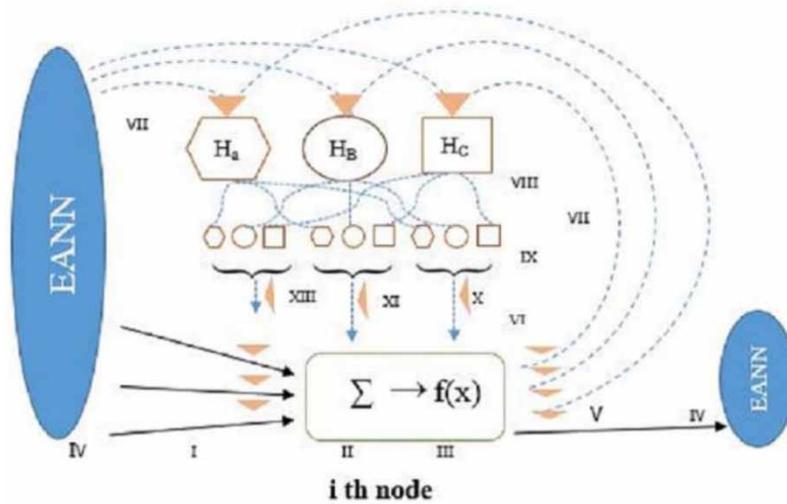


Figure 5 | EANN structure.

Iterating in the training process improves the coefficients that are assumed in this model using the pattern between the input and output data. Education process affects other node units (neurons) through hormonal coefficients, such as activation, weight and, net function. In Figure 5, the dashed and plotted lines show the hormonal and neural pathways of information (Nourani 2017).

2.7. Ensemble

One of the important advances in artificial neural networks during the last decade is the model combination method. Essentially, it is a way to enhance final model performance by merging the outputs of various models (Sharghi et al. 2018). In this multi-model method, the main objective is to maximize the estimation accuracy by using the specific capability of each model. By combining different models, the prediction performance with this method can be increased compared to the performance of individual models. Combining various methods can significantly improve modeling performance, as demonstrated by theoretical and experimental evidence (Zhang 2003). In this research, the model combination method has been performed

using three methods: a simple linear averaging (SAE), a linear-weighted averaging (WAE) and, a nonlinear neural ensemble technique (NEE). The formula used for the simple averaging model is:

$$\bar{f}(t) = \frac{1}{N} \sum_{i=1}^N f_i(t)$$

where $f(t)$ depicts the output of the simple ensemble model, N indicates the number of single methods, and f_i shows the results of every single model.

Combination is performed using the following formula according to linear-weighted averaging:

$$\bar{f}(t) = \sum_{i=1}^N w_i f_i(t)$$

In this formula, w_i is the weight applied to the result of i th model and it is calculated as follows:

$$w_i = \frac{DC_i}{\sum_{i=1}^N DC_i}$$

where DC_i is the i th method efficiency.

With a neural ensemble model, the outputs from individual models are imported into another FFNN for training as inputs and then a nonlinear ensemble is produced. A trial-and-error process was then employed for determining the ideal number of epochs and hidden layers.

2.8. Efficiency criteria

During the spatial clustering phase, the silhouette coefficient was used to measure the validity of the clustering. In a cluster, this coefficient measures the degree of similarity between the members, as shown in the equation below:

$$S_{(i)} = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

$S_{(i)}$ is the silhouette value of member i . Higher values of $S_{(i)}$ indicate more similarity of members in the same cluster. The quality of clustering can be determined by measuring the average width of the silhouette across the entire data set. Distances between clusters are measured as Euclidian distances by using the mean dissimilarity of the clusters $a(i)$. $b(i)$, which expresses the least average dissimilarities between member i and the members of other clusters.

To measure the accuracy of the models, two criteria of determination coefficient (DC) and root mean square error (RMSE) have been used. The results in both stages of training and validations should be in the desired range, which for DC is close to 1 and for RMSE close to zero. The values of the mentioned statistics were obtained using the below functions:

$$DC = 1 - \frac{\sum_{i=1}^n (y_i - x_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}}$$

In the above formulas, y_i is the data related to the model calculation, x_i is the observational data, n is the number of data points and \bar{y} and \bar{x} are the average of the mentioned data.

3. RESULTS

This section discusses the results obtained from each part separately.

3.1. Results of clustering

To study the groundwater conditions on the Ardabil plain, it is crucial to categorize quantitative and qualitative piezometers and distinguish the dominant piezometer (central) on the plain. Piezometers that exhibit similar and dominant behavior were identified using GNG. A trial-and-error method was employed for determining the optimal number and the quality of the formed clusters. The clustering results of quantitative piezometers into three clusters and qualitative piezometers for TDS and TH parameter studies into three and four clusters are shown in Table 1, respectively. In addition, clusters are shown in Figure 6. According to the table, central piezometers have the highest value of silhouette coefficient.

The central piezometers are given in the fifth column of Table 1. Among the central piezometers, the third cluster with the highest coefficient values is the best clustering, and also among the qualitative piezometers, the second cluster is the best cluster for the TDS and TH parameters. This paper revealed, the GNG method scores best in comparison with the others. Thus, due to the obtained results, the GNG can be considered as a reliable method for groundwater level and its chemical parameters clustering.

3.2. Result of single AI models

In the second stage, the best results of each model have been presented in the following sections, which use only one AI method to train and verify. In this case, the current time step is determined by its previous time steps until lag 6. The time steps $h(t - 1)$, $h(t - 2)$, ..., $h(t - 6)$ were thus fed into four different methods (ANN, ANFIS, SOM, and EANN) for estimating GWL value $h(t)$ at the current time step.

The value of the SO_4 , Ca, Mg, Cl parameters, and GWL, which were more proportional to TDS and TH based on MI processing, was fed as input into the mentioned individual methods for the qualitative process. It was essential to determine the number of middle neurons and calibrate the network at the right number of iterations in order to prevent overfitting in the FFNN model. It is important in order to avoid too small neurons, which may capture inconsistent information, and too many neurons, which may cause overfitting.

To select the most efficient model, training iterations of 2–100 were conducted using the Levenberg–Marquardt algorithm and log-sigmoid function. Here, 2–15 middle neurons were also tested to select the optimal network. Table 2 presents the optimal results of FFNN for both qualitative and quantitative studies.

Another AI-based model is ANFIS, which is characterized for its ability to be used for handling nonlinear processes with uncertainty by employing a fuzzy concept. The Sugeno operation was utilized for the calibration of membership function (MF) parameters. Additionally, MFs of triangular, Gaussian, and trapezoidal shapes showed a good ability for predicting

Table 1 | Results of clustering

Parameter	Cluster no.	Piezometers	Silhouette coefficient	Central piezometer
GWL	1	P ₂₆ ,P ₇ ,P ₂ ,P ₃₁ ,P ₃₅ ,P ₂₃ ,P ₁₄ ,P ₃₉ ,P ₂₄ , P ₁₆ ,P ₁₇ ,P ₃₆ ,P ₃₈ ,P ₃₂	0.13,0.24,0.31,0.41,0.52,0.56,0.59,0.64,0.67,0.68,0.69,0.72,0.74,0.78	P ₃₂
	2	P ₁₀ ,P ₆ ,P ₁₈ ,P ₁₃ ,P ₁₁ ,P ₁ ,P ₂₅ ,P ₃₃ ,P ₁₂ , P ₁₉ ,P ₉ ,P ₄ ,P ₂₇ ,P ₂₀	0.15,0.22,0.26,0.37,0.38,0.41,0.43,0.53,0.55,0.64,0.68,0.69,0.73,0.75	P ₂₀
	3	P ₅ ,P ₃₀ ,P ₁₅ ,P ₃₄ ,P ₈ ,P ₂₂ ,P ₂₈ ,P ₃ ,P ₂₉ , P ₂₁ ,P ₃₇	0.21,0.24,0.44,0.46,0.53,0.7,0.7,0.71,0.74,0.81,0.82	P ₃₇
TDS	1	P ₂ ,P ₄ ,P ₁₄ ,P ₂₃ ,P ₈ ,P ₁ ,P ₁₉ ,P ₃ ,P ₁₆	0.01,0.13,0.53,0.61,0.62,0.77,0.82,0.83,0.84	P ₁₆
	2	P ₁₅ ,P ₁₃ ,P ₇ ,P ₁₂ ,P ₁₇ ,P ₅ ,P ₁₀ ,P ₂₄ ,P ₁₁ , P ₁₈ ,P ₂₂ ,P ₂₁	0.45,0.51,0.52,0.72,0.82,0.83,0.83,0.84,0.85,0.85,0.86,0.87	P ₂₁
	3	P ₂₀ ,P ₂₆ ,P ₉ ,P ₂₅	0.37,0.46,0.66,0.67	P ₂₅
TH	1	P ₃ ,P ₁ ,P ₈ ,P ₁₉	0.32,0.64,0.67,0.69	P ₁₉
	2	P ₇ ,P ₁₂ ,P ₂₂ ,P ₅ ,P ₁₀ ,P ₁₁ ,P ₁₇ ,P ₂₄ ,P ₁₈ , P ₂₁	0.4,0.79,0.87,0.87,0.88,0.89,0.9,0.91,0.92,0.93	P ₁₈
	3	P ₂₃ ,P ₂ ,P ₄ ,P ₁₆ ,P ₁₅ ,P ₁₃ ,P ₁₄	0.45,0.47,0.63,0.68,0.71,0.76,0.8	P ₁₄
	4	P ₂₀ ,P ₂₆ ,P ₉ ,P ₂₅	0.4,0.57,0.69,0.74	P ₂₅

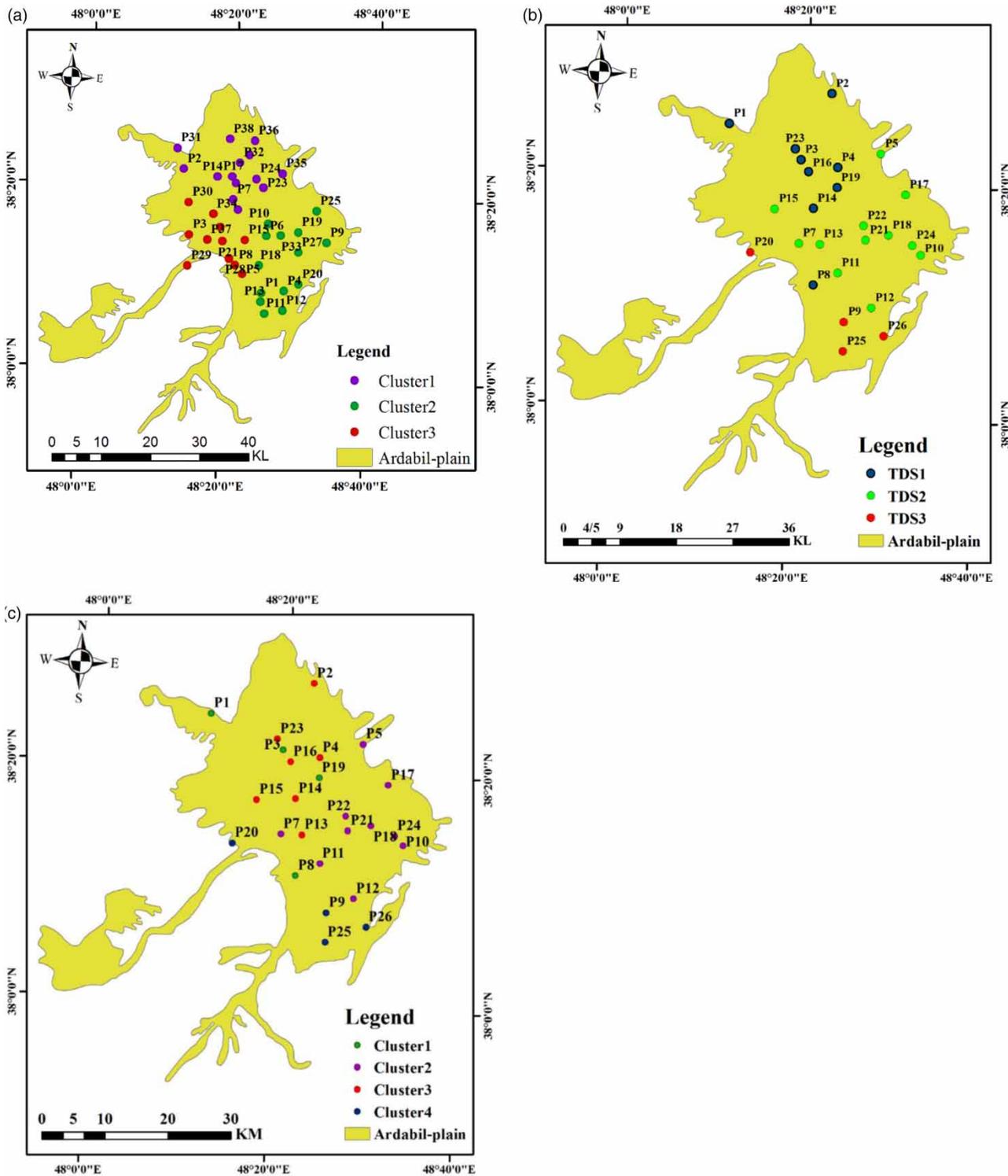


Figure 6 | GNG clustering (a): GWL, (b): TDS, (c): TH.

and analyzing groundwater as well as a consistent MF in the output layer. The best model was determined by comparing two and three MFs and the number of 5–500 iterations using a trial-and-error method. The outcomes of optimum ANFIS method are tabulated in [Table 2](#).

Table 2 | Outputs of single models

Piezometers	Models	Model architecture	DC		RMSE	
			train	verify	train	verify
P ₃₂	ANN	(3-5-1)	0.84	0.73	0.08	0.38
	ANFIS	Triangular shaped-2	0.73	0.52	0.33	0.41
	SVR	(50-0.3-0.1)	0.82	0.79	0.26	0.28
	EANN	(10-10)	0.86	0.78	0.22	0.21
P ₂₀	ANN	(3-11-1)	0.89	0.87	0.31	1.6
	ANFIS	Gaussian shaped-2	0.84	0.78	1.6	2.17
	SVR	(60-0.2-0.5)	0.89	0.86	0.58	1.4
	EANN	(10-8)	0.82	0.8	0.21	0.18
P ₃₇	ANN	(3-7-1)	0.85	0.78	0.28	0.34
	ANFIS	Gaussian shaped-2	0.76	0.70	0.12	0.15
	SVR	(50-0.1-0.33)	0.87	0.68	0.25	0.38
	EANN	(10-10)	0.75	0.64	0.14	0.15
TDS ₁	ANN	(5-12-1)	0.75	0.71	0.22	0.24
	ANFIS	Triangular shaped-2	0.74	0.73	0.22	0.25
	SVR	(20-1-1)	0.83	0.7	0.17	0.26
	EANN	(10-12)	0.64	0.61	0.28	0.24
TDS ₂	ANN	(5-6-1)	0.89	0.78	0.09	0.11
	ANFIS	Gaussian shaped-2	0.82	0.78	0.14	0.17
	SVR	(50-1-0.2)	0.86	0.74	0.12	0.19
	EANN	(10-10)	0.89	0.76	0.10	0.17
TDS ₃	ANN	(5-6-1)	0.79	0.64	0.17	0.27
	ANFIS	Triangular shaped-2	0.70	0.68	0.19	0.22
	SVR	(20-0.1-0.33)	0.69	0.56	0.22	0.22
	EANN	(10-10)	0.53	0.51	0.27	0.24
TH ₁	ANN	(5-8-1)	0.84	0.69	0.05	0.29
	ANFIS	TrimF-2	0.72	0.70	0.28	0.3
	SVR	(50-1-1)	0.86	0.78	0.08	0.28
	EANN	(10-6)	0.81	0.67	0.23	0.32
TH ₂	ANN	(5-14-1)	0.85	0.79	0.16	0.2
	ANFIS	TrimF-2	0.84	0.79	0.17	0.18
	SVR	(50-0.2-0.33)	0.9	0.82	0.08	0.2
	EANN	(10-10)	0.82	0.78	0.18	0.22
TH ₃	ANN	(5-9-1)	0.79	0.66	0.26	0.3
	ANFIS	Gaussian shaped-2	0.80	0.77	0.20	0.27
	SVR	(60-0.2-0.5)	0.82	0.78	0.24	0.28
	EANN	(10-8)	0.81	0.7	0.26	0.28
TH ₄	ANN	(5-10-1)	0.65	0.61	0.28	0.31
	ANFIS	Triangular shaped-2	0.86	0.73	0.09	0.25
	SVR	(60-0.01-0.33)	0.75	0.65	0.23	0.28
	EANN	(10-10)	0.80	0.73	0.24	0.21

The third AI-based method exploited in this investigation was SVR. To model SVR, the radial basis function (RBF) kernel can produce more accurate results because of its smoothness assumption, even though it includes fewer tuning parameters than a model using other kernels (Noori *et al.* 2011). The results of optimal SVM models are tabulated in Table 2.

The last AI applied in this study was EANN. This method can model through hormonal glands and gives accurate outcomes in learning application (Lotfi *et al.* 2014). The EANN was designed utilizing the Sigmoid activation function and 1–10 hormone parameters, 1–10 neurons, and 1–2 training iterations were used to select the network with high productivity. Table 2 indicates the ideal results of EANN models.

Examining the results of quantitative studies, it can be seen that the DC of training for all three central piezometers has a higher value than the DC for validation, which may have originated from a large number of data at this stage. The results of

validation's DC in P_{20} are 16% higher than other piezometers on average. The reason for the low value in the other two piezometers can be the presence of illegal wells and lots of agricultural land in this area. Also, P_{20} is located in the south of this plain. Considering that the highest amount of nutrition of the Ardebil plain aquifer is from the southern parts where the river enters it, so it can be considered as a reason for the high DC of piezometer in this area. Regarding the observed outcomes from qualitative modeling, it can be seen that the DC of training data is more than the DC of validation data. Conversely, modeling the amount of the total solid dissolved in water (TDS) in cluster 2 and the total hardness (TH) in cluster 2 have shown better evaluation results up to 20 and 14% on average in contrast with other clusters, respectively. According to Figure 6, it can be seen that these two clusters are almost scattered in the central and eastern parts of Ardebil plain which can be a reason to justify the relative highness of this criterion according to Daneshvar Vousoughi *et al.* (2013). This study stated that the water quality situation in the northern and southwestern of this regions is unfavorable compared to other areas. Also, from clusters in which prediction and evaluation have been done with high accuracy, can be seen that the most correlation is seen between the selected inputs and groundwater quality. Qualitative study results reveal the fact that the FFNN model outperformed other methods in averaging DC of GWL validation data set modeling. The average DC of TDS validation results for the ANFIS method was higher than other methods and for the TH study, SVM had a higher result compared with other models. Also in a few results, EANN outperformed the other methods. The estimation performance of the models employed in this study was best visualized through scatter plots. Figure 7 illustrates the scatter plots of the single model result compared to the observed result. It is evident in this figure that, in some models, some methods overestimate observed values while others underestimate them. As shown in Figure 7. A, point 1 of the EANN model provided a closer estimation of the record value (i.e. the difference between the observed and computational data in the EANN case is only 2 units, while this difference in the SVR case increases to 10), which was countered by the SVR model at point 2 (4 units

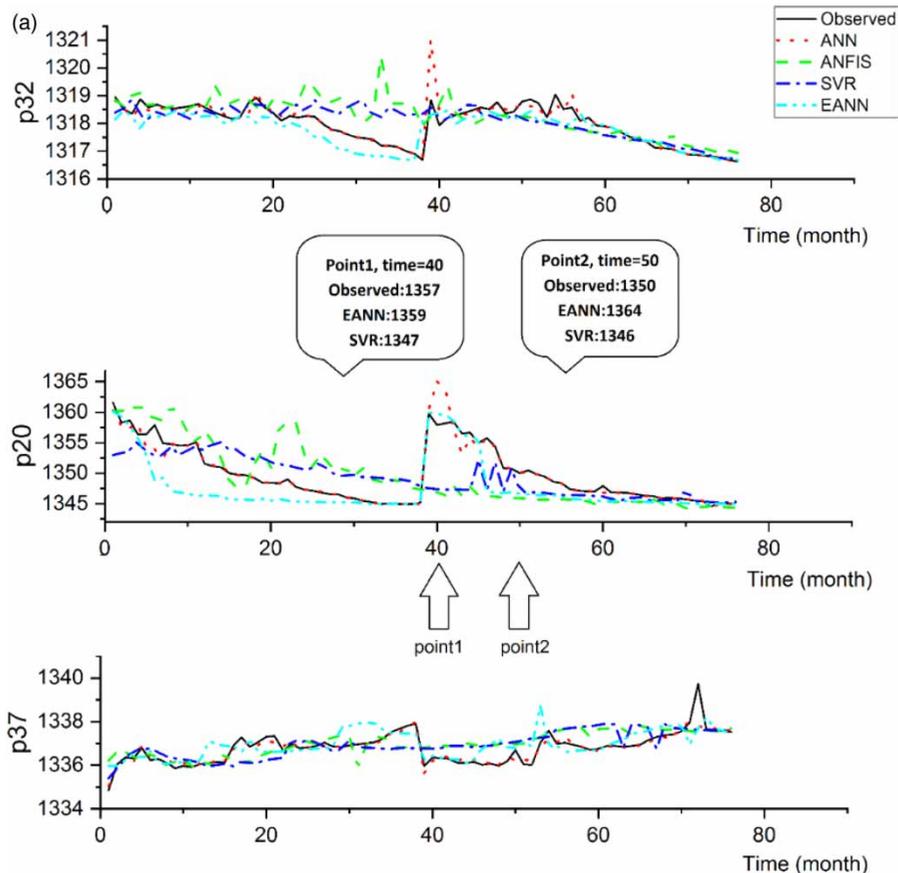


Figure 7 | Recorded vs. estimated precipitation values obtained by ANN, ANFIS, SOR and EANN methods in the testing phase for (a): GWL, (b): TDS, (c): TH parameters. (Continued).

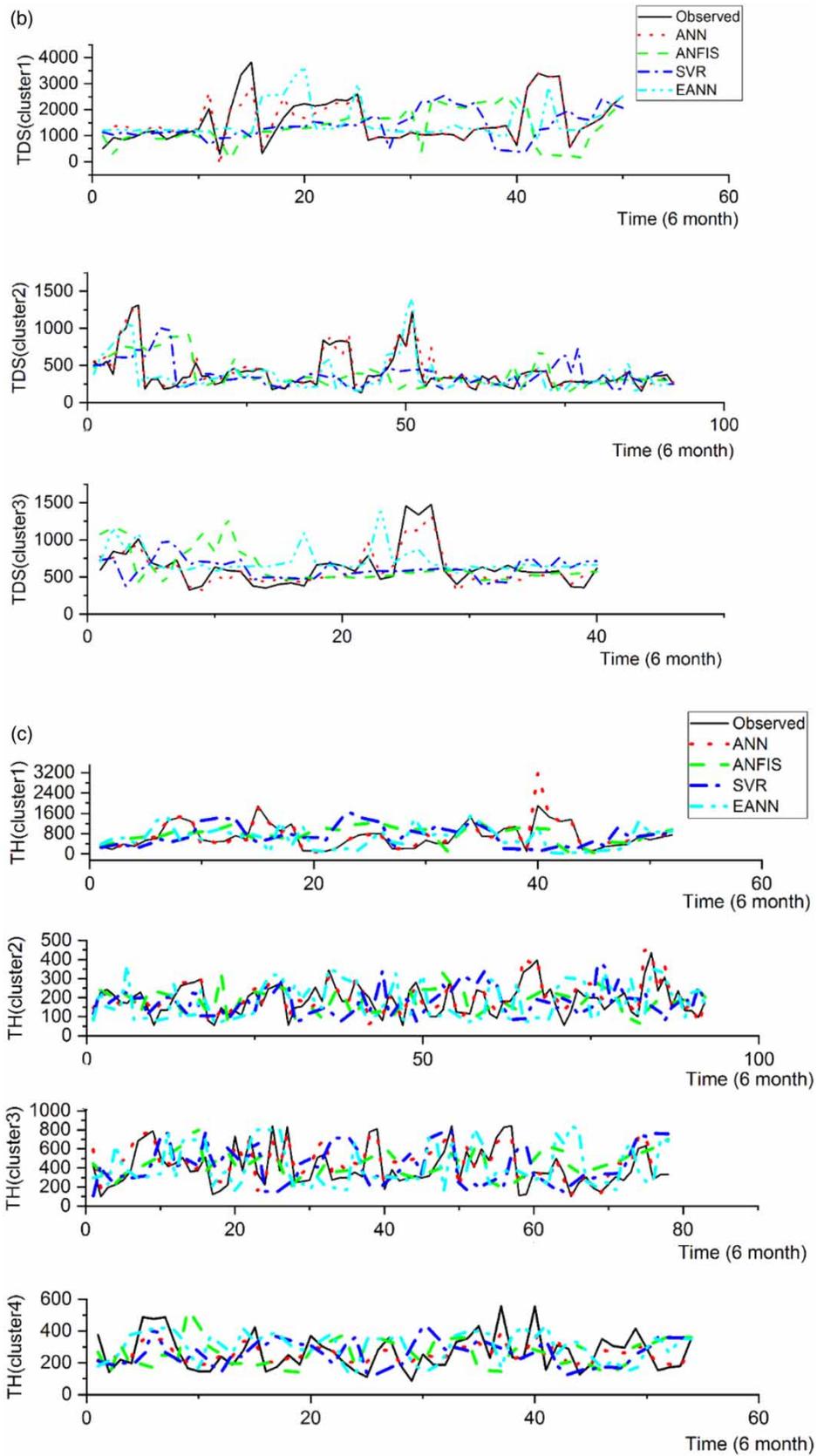


Figure 7 | Continued.

versus 14), even when the two dates are in close proximity. There are also some times that none of the models could simulate the observed data with good accuracy. Therefore, different models at different time points can lead to different performances, according to the results from the selected points. Furthermore, a study was developed to examine the effectiveness of combined (GA-EANN) models over the use of single models (EANN, FFNN, GRNN) for predicting one-month ahead groundwater levels. In comparison to the developed single AI-based models, the GA-EANN model offered a marked improvement in performance (Roshni *et al.* 2019). Consequently, ensemble methods can provide more accurate groundwater field estimations. On this subject, a possible method for improving the efficiency of models is to combine the outputs of multiple methods.

3.3. Results of ensemble modeling

The third step in the modeling process involved combining the results from four independent artificial intelligence models to enhance prediction efficiency. The linear SAE model outperformed developed AI single models in some piezometers and clusters. It is a fact that the linear averaging of the data set often yields a value lower than the highest and a value higher than the lowest (Nourani *et al.* 2020). It is possible that WAE is slightly better than SAE as a second ensemble model. This might be the case due to the weighting assigned to the parameters based on their importance. As with FFNN, NNE was also trained using the Levenberg–Marquardt algorithm. In addition, both the output and hidden layers were activated by the sigmoid activation function.

Table 3 | Outputs obtained by linear, weighted and non-linear ensemble methods

Piezometer	Iteration	Model architecture	DC		RMSE	
			Train	Verify	Train	Verify
P32	Simple linear averaging	(4,13,1)	0.81	0.7	0.22	0.32
	Weighted averaging		0.81	0.72	0.26	0.27
	Non-linear averaging		0.85	0.82	0.23	0.21
P20	Simple linear averaging	(4,22,1)	0.86	0.83	0.67	1.33
	Weighted averaging		0.86	0.83	1.12	1.73
	Non-linear averaging		0.96	0.93	0.19	0.14
P37	Simple linear averaging	(4,10,1)	0.8	0.7	0.19	0.25
	Weighted averaging		0.81	0.7	0.22	0.3
	Non-linear averaging		0.89	0.8	0.1	0.14
TDS1	Simple linear averaging	(4,6,1)	0.74	0.68	0.22	0.24
	Weighted averaging		0.74	0.69	0.23	0.24
	Non-linear averaging		0.85	0.79	0.16	0.21
TDS2	Simple linear averaging	(4,11,1)	0.86	0.76	0.11	0.16
	Weighted averaging		0.86	0.76	0.22	0.16
	Non-linear averaging		0.95	0.88	0.07	0.12
TDS3	Simple linear averaging	(4,10,1)	0.67	0.59	0.21	0.23
	Weighted averaging		0.69	0.6	0.21	0.23
	Non-linear averaging		0.85	0.74	0.16	0.12
TH1	Simple linear averaging	(4,20,1)	0.8	0.71	0.16	0.29
	Weighted averaging		0.81	0.71	0.22	0.29
	Non-linear averaging		0.92	0.82	0.15	0.2
TH2	Simple linear averaging	(4,12,1)	0.85	0.79	0.14	0.2
	Weighted averaging		0.85	0.79	0.11	0.2
	Non-linear averaging		0.95	0.9	0.08	0.12
TH3	Simple linear averaging	(4,11,1)	0.8	0.72	0.24	0.28
	Weighted averaging		0.8	0.73	0.24	0.28
	Non-linear averaging		0.87	0.83	0.2	0.23
TH4	Simple linear averaging	(4,22,1)	0.76	0.68	0.21	0.26
	Weighted averaging		0.77	0.68	0.23	0.26
	Non-linear averaging		0.82	0.78	0.19	0.24

In a trial and error process, the number of middle neurons and the number of iterations for training varied between 2 and 25, and between 5 and 300, respectively. As a combination technique in hydrological modeling, non-linear neural networks have been successfully used (Elkiran *et al.* 2018; Sharghi *et al.* 2018; Nourani *et al.* 2021). The obtained results of ensemble models are presented in Table 3. As shown in this table, all ensemble methods can be used to enhance the performance of single models in GW modeling.

As revealed by the results, all three methods of the model combination were able to outperform the individual models in the majority of cases, and in Figure 8 an appropriate correspondence is apparent between the observational and computational data in this method. On the scatter plots, this figure illustrates how the NNE model performs in the estimations of GWL, TDS, and TH. Also, it shows less spread of points for observed and estimated GWL, TDS, and TH parameters. As mentioned, all models and techniques have their own merits and demerits. Some methods offered higher estimates, while others offered lower estimates. Therefore, the mentioned combination methods show more preferable results compared

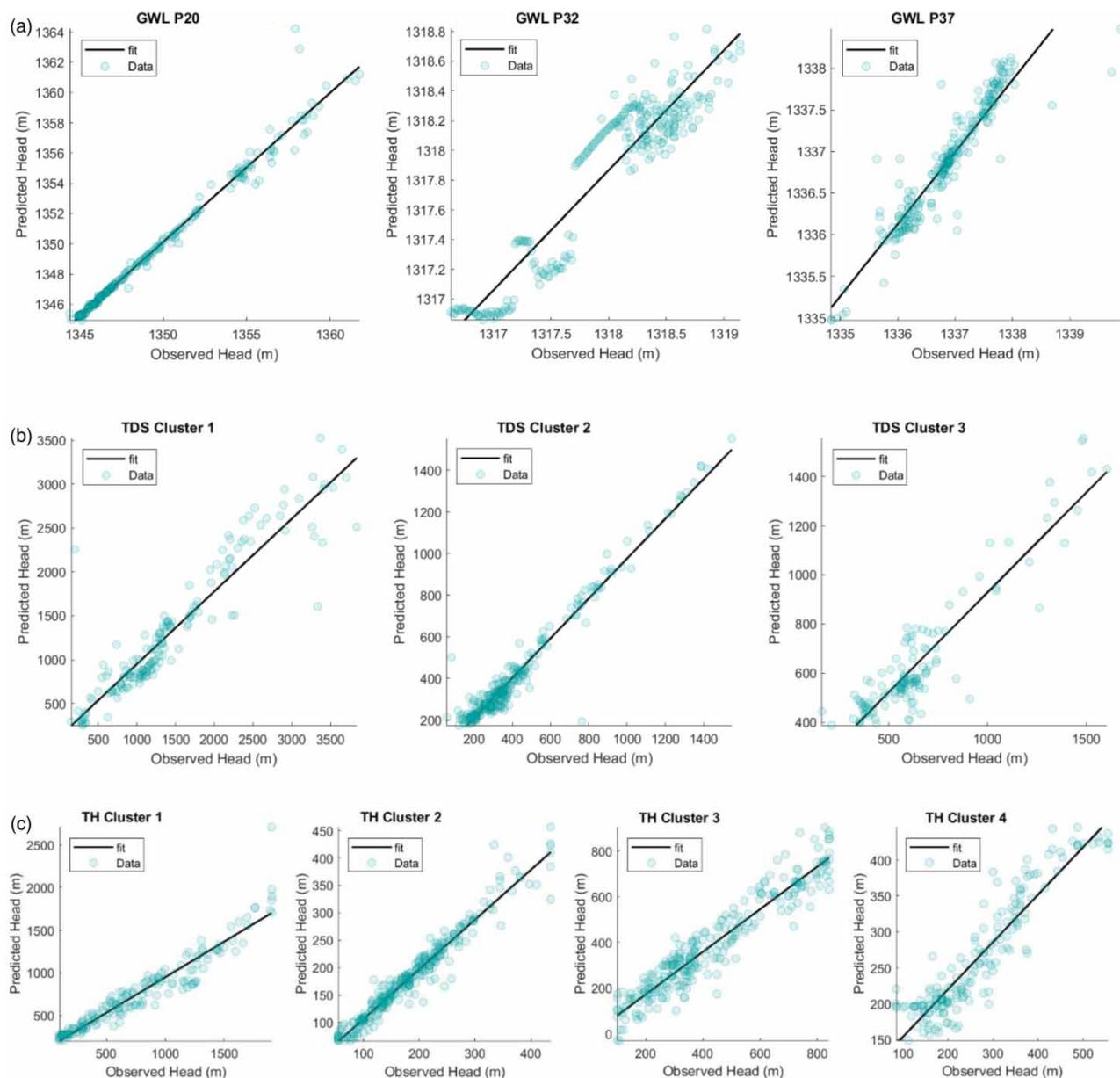


Figure 8 | correspondence of the observational and computational data for (a): GWL, (b): TDS, (c): TH parameters.

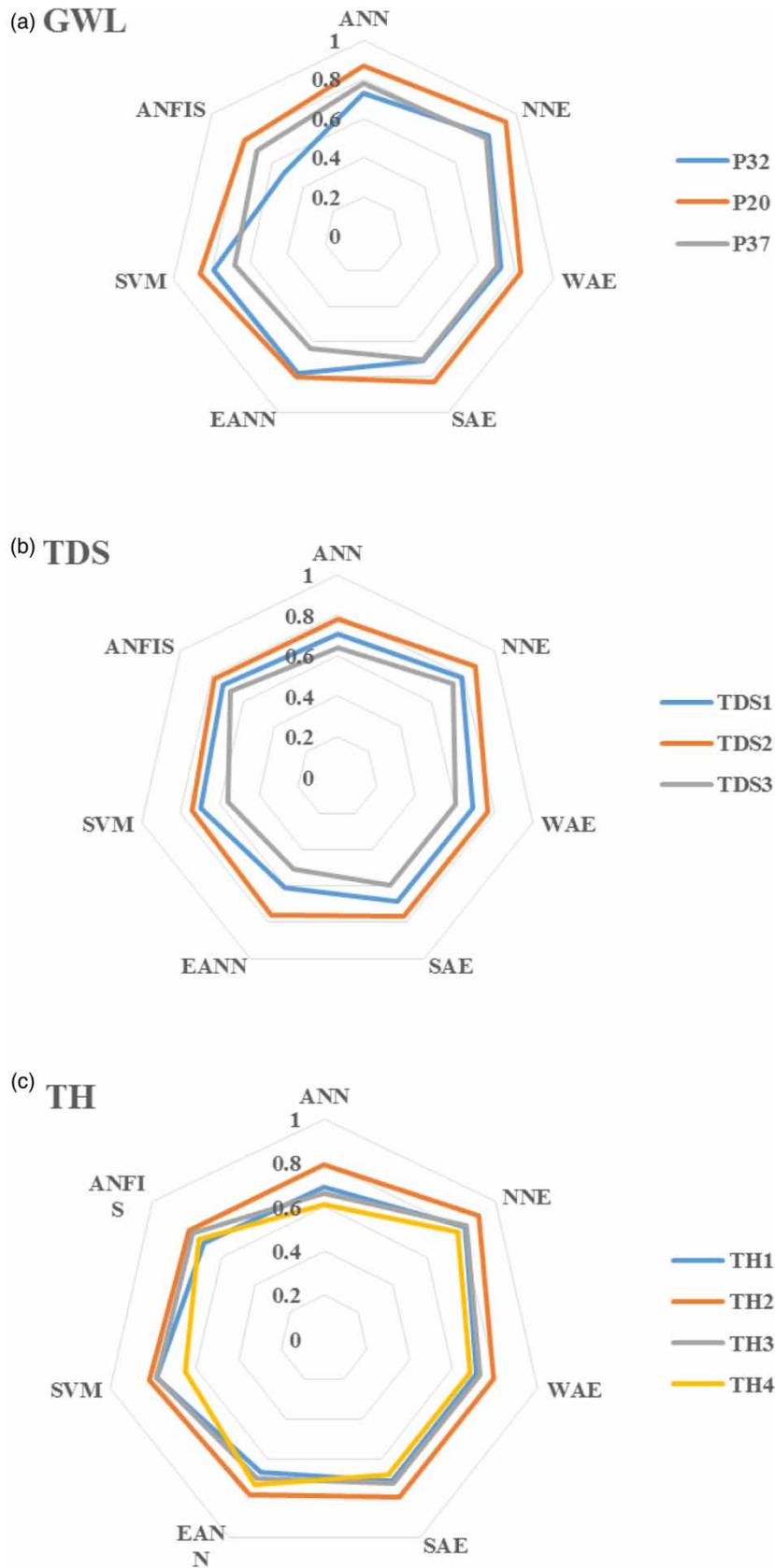


Figure 9 | Radar diagrams of (a): GWL, (b): TDS, (c): TH modeling.

to individual models due to the unique capabilities of each model. In general, the outputs of the simple and weighted ensemble were quite the same because the efficiency of the simple and weighted ensemble techniques followed the same path. As can be seen in Table 3, nonlinear averaging is capable in all cases of being more powerful than linear ensemble methods. In terms of DC, it is necessary to mention that the nonlinear neural ensemble technique came up with better validation outputs up to 57, 19, and 25% for P_{32} , P_{20} , and P_{37} piezometers and 30, 12, and 45% for TDS_1 , TDS_2 , and TDS_3 and finally 22, 15, 26 and 28% for TH_1 , TH_2 , TH_3 and TH_4 index, respectively. All in all, the ensemble models in general, and single AI models specifically, perform well in GW estimations. In Figure 9 for all GWL, TDS, and TH models, radar diagrams are used to create a visual representation of the comparison between single and ensemble methods. The clustering method has been used previously to classify the observed piezometers of the Ardebil plain. The SVM and ANN methods were combined to predict the GWL. The results showed that the best values of DC in training and verification were 0.94 and 0.89 respectively. In this study, however, GNG clustering method was applied and two other AI methods (ANFIS and EANN) were added to the prediction stage. These steps improved the accuracy of prediction up to 0.96 for training and 0.93 for verification stages.

4. CONCLUSION

In this study, changes in groundwater levels and their quality were studied and predicted using artificial intelligence networks. In this regard, the GNG clustering technique was used to divide the whole study area into several groups. Furthermore, the piezometers in the quantitative and qualitative wells were modeled. In the next step, to increase the efficiency and accuracy of modeling, the model combination method was performed using simple linear averaging, linear-weighted averaging, and nonlinear neural ensemble procedures. In the end, the obtained results were compared. Based on the outputs, the ensemble model could result in a promising improvement in groundwater parameters modeling, conversely, the efficiency of simple linear averaging and linear-weighted averaging is directly related to individual models, therefore the poor results of each model affect these ensemble models. Among the three different methods of ensemble modeling, the nonlinear neural technique is more efficient. Comparing the results of the third ensemble method with the best result obtained from individual ones, TDS_2 showed the highest improvement of 12%, while P_{37} indicated the lowest progress of 2%. Considering that in this study only static averaging of the results of individual methods was used, it is suggested that in future studies, by a dynamic and comparative selection of the results of individual methods in terms of minimizing estimation error, the efficiency of the model combination method will be improved. In this study, an ensemble unit is developed using only black-box models. Thus, physical-based models should be included in GW studies, alongside AI models, to investigate and combine their performances.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

- Abdi, A., Hassanzadeh, Y. & Ouarda, T. B. 2017 *Regional frequency analysis using growing neural gas network*. *Journal of Hydrology* **550**, 92–102.
- Abraham, A. 2005 Adaptation of fuzzy inference system using neural learning. In: *Fuzzy Systems Engineering: Theory and Practice*, Vol. 181 (N. Nedjah & L. M. Mourelle, eds.). Springer, Berlin, pp. 53–83.
- Al-adamat, R. A. N., Foster, I. D. L. & Baban, S. M. J. 2003 *Groundwater vulnerability and risk mapping for the Basaltic aquifer of the Azraq basin of Jordan using GIS and remote sensing and DRASTIC*. *Applied Geography* **23** (4), 303–324.
- Ayvaz, M. T., Karahan, H. & Aral, M. M. 2007 *Aquifer parameter and zone structure estimation using kernel-based fuzzy c-means clustering and genetic algorithm*. *Journal of Hydrology* **343** (3–4), 240–253.
- Basak, D., Pal, S. & Patranabis, D. C. 2007 Support vector regression. *Neural Information Processing-Letters and Reviews* **11** (10), 203–224.
- Chitsazan, M., Rahmani, G. & Neyamadpour, A. 2013 *Groundwater level simulation using artificial neural network: a case study from Aghili plain, urban area of Gotvand, south-west Iran*. *Geopersia* **3** (1), 35–46.
- Cirrincione, M., Pucci, M. & Vitale, G. 2012 *Growing neural gas-based MPPT of variable maps pitch wind generators with induction machines*. *IEEE Transactions on Industry Applications* **48** (3), 1006–1016.
- Daliakopoulos, I. N., Coulibaly, P. & Tsanis, I. K. 2005 *Groundwater level forecasting using artificial neural networks*. *Journal of Hydrology* **309** (1), 229–240.

- Daneshvar Vousoughi, F., Dinpashoh, Y., Aalami, M. T. & Jhajharia, D. 2013 Trend analysis of groundwater using non-parametric methods (case study: Ardabil plain). *Stochastic Environmental Research Risk Assessment* **27** (2), 547–559.
- Dehghani, R. & Torabi Poudeh, H. 2022 Application of novel hybrid artificial intelligence algorithms to groundwater simulation. *International Journal of Environmental Science and Technology* **19** (5), 4351–4368.
- Derbela, M. & Nouri, I. 2020 Intelligent approach to predict future groundwater level based on artificial neural networks (ANN). *Euro-Mediterranean Journal for Environmental Integration* **5** (3), 1–11.
- Elkiran, G., Nourani, V., Abba, S. I. & Abdullahi, J. 2018 Artificial intelligence-based approaches for multi-station modeling of dissolved oxygen in river. *Global Journal of Environmental Science and Management* **4** (4), 439–450.
- Gintamo, T. T., Mengistu, H., Xu, Y. & Kanyerere, T. H. 2022 'Using GIS-Based Modified DRASTIC Modelling of the Cape Flats Aquifer to Assess Coastal Aquifer Vulnerability in Urban Hydrogeology'. South Africa. Available at SSRN.
- Hsu, K., Gupta, H. V., Gao, X., Sorooshian, S. & Imam, B. 2002 Self-organizing linear output map (SOLO). 'an artificial neural network suitable for hydrologic modeling and analysis'. *Water Resource Research* **38** (12), 1–38.
- Jang, J. S. R. 1995 ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems Man and Cybernetics* **23** (3), 665–685.
- Kalteh, A. M., Hjorth, P. & Berndtsson, R. 2008 Review of selforganizing map (SOM) in water resources: analysis, modeling, and application. *Environmental Modelling & Software* **23** (7), 835–845.
- Kim, H., Kim, H. K. & Cho, S. 2020 Improving spherical k-means for document clustering: 471 fast initializations, sparse centroid projection, and efficient cluster labeling. *Expert Systems with Applications* **150**, 113288.
- Lin, G. F. & Chen, G. R. 2005 Determination of aquifer parameters using radial basis function network approach. *Journal of the Chinese Institute of Engineers* **28** (2), 241–249.
- Lotfi, E., Khosravi, A. & Nahavandi, S. 2014 Wind power forecasting using emotional neural networks. In *2014 IEEE International Conference on Systems, man, and Cybernetics*. pp. 311–316.
- Maier, H. R. & Dandy, G. C. 1996 The use of artificial neural networks for the prediction of water quality parameters. *Water Resources Research* **32** (4), 1013–1022.
- Morell, V., Cazorla, M., Orts-Escolano, S. & Garcia-Rodriguez, J. 2014 3D maps representation using GNG. *Mathematical Problems in Engineering* **2014**, 972304.
- Nadiri, A., Vahedi, F., Asghari Moghadam, A. & Kadkhodae, A. 2015 Use of artificial intelligence model supervised to predict groundwater level. *Civil Engineering and Environmental Engineering Journal of Tabriz University* **46**, 101–112. [In Persian].
- Noori, R., Karbassi, A. R., Moghaddamnia, A., Han, D., Zokaei-Ashtiani, M. H., Farokhnia, A. & Gousheh, M. G. 2011 Assessment of input variables determination on the SVM model performance using PCA, Gamma test, and forward selection techniques for monthly streamflow prediction. *Journal of Hydrology* **401** (3–4), 177–189.
- Nordin, N. F. C., Mohd, N. S., Koting, S., Isamail, Z., Sherif, M. & EL-Shafie, A. 2021 Groundwater quality forecasting modelling using artificial intelligence: a review. *Groundwater for Sustainable Development* **14**, 100643.
- Nourani, V. 2017 An Emotional ANN (EANN) approach to modeling rainfall-runoff process. *Journal of Hydrology* **544**, 267–277.
- Nourani, V. & Kalantari, O. 2010 Integrated artificial neural network for spatiotemporal modeling of rainfall-runoff-sediment process. *Environmental Engineering Science* **27** (5), 411–422.
- Nourani, V. & Parhizkar, M. 2013 Conjunction of SOM-based feature extraction method and hybrid wavelet-ANN approach for rainfall-runoff modeling. *Journal of Hydroinformatics* **15** (3), 829–848.
- Nourani, V., Kisi, Ö. & Komasi, M. 2011 Two-hybrid artificial intelligence approaches for modeling rainfall-runoff process. *Journal of Hydrology* **402**, 41–59.
- Nourani, V., Hosseini Baghanam, A., Adamowski, J. & Gebremichael, M. 2013 Using self-organizing maps and wavelet transforms for space-time pre-processing of satellite precipitation and runoff data in neural network-based rainfall-runoff modeling. *Journal of Hydrology* **476**, 228–243.
- Nourani, V., Hosseini Baghanam, A., Adamowski, J. & Kisi, O. 2014 Applications of hybrid wavelet-artificial intelligence models in hydrology: a review. *Journal of Hydrology* **514**, 358–377.
- Nourani, V., Gökçekuş, H. & Umar, I. K. 2020 Artificial intelligence-based ensemble model for prediction of vehicular traffic noise.
- Nourani, V., Gokcekus, H. & Gelete, G. 2021 Estimation of suspended sediment load using artificial intelligence-based ensemble model. *Complexity* **2021**, 6633760.
- Roshni, T., Jha, M. K., Deo, R. C. & Vandana, A. 2019 Development and evaluation of hybrid artificial neural network architectures for modeling Spatio-temporal groundwater fluctuations in a complex aquifer system. *Water Resources Management* **33** (7), 2381–2239.
- Santos, C. P. & Nascimento, M. C. V. 2016 Growing neural gas as a memory mechanism of a heuristic to solve a community detection problem in networks. *Procedia Computer Science* **96**, 485–494.
- Shamseldin, A. Y., O'Connor, K. M. & Liang, G. C. 1997 Methods for combining the outputs of different rainfall-run models. *Journal of Hydrology* **197** (1–4), 203–229.
- Sharghi, E., Nourani, V. & Behfar, N. 2018 Earth fill dam seepage analysis using ensemble artificial intelligence based modeling. *Journal of Hydroinformatics* **20** (5), 1071–1108.
- Sharghi, E., Nourani, V., Najafi, H. & Gokcekus, H. 2019 A conjunction of newly proposed emotional ANN(EANN) and Wavelet transform for suspended sediment load modeling. *Water Supply* **19** (6), 1727–1734, IWA publishing.
- Shi, J., Chen, C. & Zhong, S. 2014 Privacy preserving growing neural gas over arbitrarily partitioned data. *Neurocomputing* **144**, 427–435.

- Shwetank, S. & Chaudhary, J. K. 2022 Hybridization of ANFIS and fuzzy logic for groundwater quality assessment. *Journal Groundwater for Sustainable Development* **18**, 100777.
- Subba Rao, N. 2012 PIG: a numerical index for dissemination of groundwater contamination zones. *Hydrological Processes* **26**, 3344–3350.
- Sujatha, P. & Kumar, G. N. P. 2010 Prediction of groundwater levels using different artificial neural network architectures and algorithms. *ISH Journal of Hydraulic Engineering* **16** (1), 20–30.
- Toth, E. 2009 Classification of hydro-meteorological conditions and multiple artificial neural networks for streamflow forecasting. *Hydrology and Earth System Sciences* **13** (9), 1555–1566.
- Vapnik, V., Golowich, S. E. & Smola, A. 1996 Support vector method for function approximation, regression estimation, and signal processing. In: *Advances in Neural Information Processing Systems 9* (M. C. Mozer, M. Jordan & T. Petsche, eds.). MIT Press, Cambridge.
- Ventocilla, E. & Riverio, M. 2020 A comparative user study of visualization techniques for cluster analysis of multidimensional data sets. *Information Visualization* **19** (4), 318–333.
- Ventocilla, E., Martins, R. M., Paulovich, F. & Riveiro, M. 2021 Scaling the growing neural gas for visual cluster analysis. *Big Data Research* **26**, 100254.
- Viejo, D., Garcia-Rodriguez, J. & Cazorla, M. 2014 Combining visual features and growing neural gas networks for robotic 3D SLAM. *Information Sciences* **276**, 174–185.
- Zhang, G. P. 2003 Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* **50**, 159–175.

First received 30 April 2022; accepted in revised form 18 July 2022. Available online 22 July 2022