

# A practical low-cost model for prediction of the groundwater quality using artificial neural networks

Nima Heidarzadeh

## ABSTRACT

The quality monitoring of water wells is always a costly and time-consuming process. To avoid the unnecessary cost and time of future sampling, applying some powerful and well known prediction models can be a suitable alternative. In this research, the groundwater quality of Amol-Babol aquifer was predicted using the artificial neural network (ANN) model with a data set from 1987 to 2010. Sodium (Na) was considered as the response variable in the ANN model due to its high concentration for irrigation. Also, to select the studied wells in the neural network, a geographic information system (GIS)-based zoning of Na was conducted for 20 years. Afterwards, the sensitive area was detected. Based on pre-modeling, the three properties of pH, electrical conductivity and total hardness were the best input variables. The results indicated that the Na concentration in three wells can be estimated by training six monitoring wells with a high accuracy. The best network is a two-layer network of the Logsig-Tansig transfer functions with four and three neurons in the first and second layers, respectively. In the best model, the coefficients of determination ( $R^2$ ) were 0.99 and 0.98 for the training and the validation periods, respectively, with a root mean square error of 0.08.

**Key words** | artificial neural networks, ANN, feed-forward networks, groundwater, sodium

**Nima Heidarzadeh**

Civil Engineering Group, Technical and Engineering  
Department,  
Kharazmi University,  
Tehran,  
Iran  
E-mail: [n.heidarzadeh@khu.ac.ir](mailto:n.heidarzadeh@khu.ac.ir)

## INTRODUCTION

Water resources quality management is one of the most important concerns, particularly for groundwater resources. To evaluate groundwater quality, the operators of the wells apply a monitoring plan in aquifers. During the monitoring, several samples are collected at a pre-set schedule and many quality characteristics are analyzed in laboratories. This process is costly and time-consuming. To avoid the unnecessary cost and time, applying some powerful and well-known prediction models, worked based on the historical recorded data as their input, can be a suitable alternative. One of the appropriate methods to predict the quality of water wells is an artificial neural networks (ANNs) model. ANN is able to find the relationship between the set of inputs and outputs to predict each output corresponding to the desired inputs, without considering any initial assumptions

and prior knowledge of the relationships between the studied variables (Kohzadi *et al.* 1996). Thus, to avoid field surveys and reduce the cost of sampling and laboratory analyses, this method can be used to estimate the groundwater quality. There are many investigations to apply this model for predicting water resources quality or other characteristics.

Karamouz & Araghi Nezhad (2004) used two approaches of ANN to predict floods and changes in groundwater quality and level in an aquifer. They prepared the initial data for training the ANN model by considering multiple inputs and outputs for the aquifer. Noorani *et al.* (2008) predicted temporal and spatial changes in the groundwater level in the Tabriz plain using ANN. They used six types of combinations and training algorithms to determine

the best structure of the ANN. The results of the research showed that a feedback ANN with Levenberg-Marquardt training algorithm provides the best predictions. Furthermore, in a similar study, Ray & Klindworth (2000) confirmed the ability of ANN to predict the concentration of nitrate in wells. Also, Momeni *et al.* (2011) evaluated the application of ANN to predict the groundwater level at Dasht-E-Naz plain in the city of Sari, Mazandaran province, Iran. In this study, the input variables were precipitation and evapotranspiration data during the desired period and output of the model was the groundwater depth in the same period. Additionally, Abbasi *et al.* (2013) used a feed-forward ANN model (feed-forward back propagation) to predict the amount of total dissolved solids (TDS) in the groundwater to decide whether it is suitable for irrigation. They used pH as the input of the model. The model presents  $R^2$  of 0.9 and 0.64 for training and prediction periods, respectively. Mehrdadi *et al.* (2012) predicted the TDS of effluent with the neural network in Fajr industrial wastewater treatment plant in the south of Iran. Zare *et al.* (2011) estimated the nitrate concentration in groundwater using ANN and linear regression (LR) models. Results showed an acceptable and appropriate accuracy for both ANN and LR methods. Results also declared that ANN requires fewer parameters with more accuracy in comparison to LR models. Singh & Datta (2010) applied a feed-forward ANN to estimate temporal and spatial changes of unknown pollution sources, unknown flow, and transport parameters using a back-propagation (BP) algorithm for training.

All previous studies show the ability of ANN to predict several kinds of groundwater characteristics. However, they do not propose a practical management procedure by using such a modeling tool. The assumption of the present study is that by applying an ANN model, we can predict the quality of some wells of an aquifer with knowing the characteristics of the other wells and with fewer inputs of the model, consequently lowering the cost of sampling and analysis. For example, based on a quality monitoring plan established in Iran, physico-chemical properties of the wells have been periodically analyzed over the past 20 years by the regional water authorities. The conventional plan includes analyzing several physico-chemical properties of major anions ( $\text{Cl}^-$ ,  $\text{SO}_4^{2-}$ ,  $\text{HCO}_3^-$ ,  $\text{CO}_3^{2-}$ ), major cations ( $\text{Na}^+$ ,  $\text{K}^+$ ,

$\text{Mg}^{+2}$ ,  $\text{Ca}^{+2}$ ), pH, TDS, electrical conductivity (EC), and total hardness (TH). This study suggests a procedure to reduce the number of physico-chemical properties mentioned above and the number of wells to be monitored. To achieve the goal, the quality of some wells of the Amol-Babol aquifer will be predicted by ANN model and a suitable model will be provided to reduce laboratory costs resulting from unnecessary analyses of the properties.

## METHODS

### Study area

The study area is the Amol-Babol plain, located in the Mazandaran province, Iran and at the southern coast of the Caspian Sea. The plain is located in  $52^\circ 10' - 52^\circ 50'$  E and  $36^\circ 22' - 36^\circ 44'$  N, with an area of 1,571 km<sup>2</sup>. The area is mainly used for rice cultivation. The area has a Mediterranean climate with a mild and wet condition. The average annual rainfall of the area is 1,000 (mm). The mean annual temperature is between 15 and 18 °C. Figure 1 shows the location of each of the monitoring wells and the surrounding land uses.

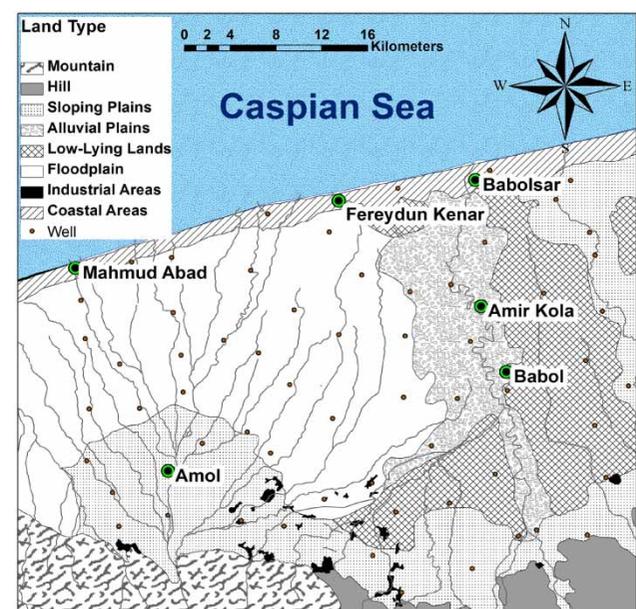


Figure 1 | Land uses and location of monitoring wells.

## Water quality assessment and geographic information system (GIS)-based zoning

The main consumption of the studied wells is for agricultural activities, particularly rice cultivation. Hence, the important model variables were determined by comparing the physico-chemical properties of the wells with the Food and Agriculture Organization (FAO) irrigation standards (FAO 2010). The required quality data were obtained by Mazandaran Regional Water Company for all Amol-Babol plain monitoring wells. Initially 2,132 data were evaluated during the period 1987 to 2010. Then the desired chemical, the ion(s) with a concentration of more than the FAO standard, was selected to be zoning. In this regard, the point-based concentration data of the desired chemical were transferred into spatial raster data in ArcGIS 10.1 software. Then, an interpolated map was created by applying inverse distance weighting (IDW), Kriging, Co-Kriging, and radial basis function (RBF) interpolation techniques. These techniques apply the point data, for example the Na concentration of each well, to interpolate them by different algorithms (ESRI 2008). After that, the best interpolated map was classified into appropriate intervals of the desired chemical concentration. In this study, the zoning was conducted for sodium (Na) with the available data of 77 wells for 20 years. In the next step, nine nearby wells were selected in the critical region to be modeled in the ANN. Moreover, EC, TH, and pH were used as input variables, because these properties are easily measurable by portable devices of water quality assessment or by simple and low cost laboratory methods.

## ANNs

ANN is a tool for estimating complex linear or nonlinear mapping when the mapping cannot be expressed using conventional mathematical equations. ANNs will estimate these relations during their training. ANN has a layered structure and it is composed of an input layer, an output layer, and one or more intermediate layers. Each layer consists of a number of nodes or neurons that are connected to networks with different weights. Based on the connection of the nodes to each other, neural networks were divided into feed-forward networks and feed-back networks (Menhaj 2000).

## The used ANN model

There are several algorithms for training feed-forward networks. In the simplest implementation feed-forward networks, weights and biases are updated in a way that reduces the goal function (i.e. the opposite slope). Equation (1) shows a single epoch of the algorithm (Rumelhart & McClelland 1986).

$$X_{k+1} = X_k - \alpha_k g_k \quad (1)$$

in which,  $X_k$  is the current vector of weights and biases,  $g_k$  is the current slope of activation function and  $\alpha_k$  is the learning rate.

## Structure of the used network

The first step in using a neural network is to determine its architecture. The architecture of a neural network usually is defined as the number of hidden layers, transfer functions of each layer and the number of neurons in each layer. Each of these parameters is very important in the performance of neural networks. Some of these parameters were determined by the problem state and some others were determined by trial and error method. In the present study, the model was a dual layer model. Also, the combination of transfer functions and the number of neurons in the hidden layer was changed from 1 to 10 using a trial and error procedure. The simplicity of the network should be considered in selecting the neurons in the hidden layer. In other words, to avoid the complexity of the network between two similar options, the option with fewer neurons should be selected. The functions of Tansig, Logsig, and Purelin were used as transfer functions of the different layers.

## Selection of training function

Training functions are responsible for adjusting and modifying the weights and bias values in order to have a better training. The Levenberg-Marquardt (Trainlm) propagation algorithm is a multivariate algorithm in which the root mean square error (RMSE) decreases at each epoch. This feature caused the algorithm to have the fastest performance

(Pham & Sagiroglu 2001). Sharma & Venugopalan (2014) showed that the Trainlm function converges faster with a lower number of epochs compared with the other training functions. For this reason, the Trainlm function is used as the training function in this study.

### Design parameters of a feed-forward network

Training rate should be introduced in the training algorithm of a feed-forward network. This parameter, which is shown by  $\alpha$ , determines the speed of convergence of the algorithm in the network. Training rate is multiplied by the value of the slope, and it is used to update weights and bias. If the rate chosen is too large, the training process will not have enough stability and if the rate chosen is very small, the algorithm needs too much time to converge. Determining the appropriate training rate is one of the most critical stages of the BP algorithm (Razavi 2006).

In this study, this parameter is considered to be a constant value to avoid the risk of divergence.

Another important parameter in the design of feed-forward networks is momentum ratio (MC). This parameter receives a number between 0 and 1. When the MC is zero, weight changes are only from the slope of the activation function, and when it is 1, weight changes are based on previous weight changes and the slope will be ignored. Other factors that are effective in the network design are Show, Goal and Epoch. The Show parameter indicates the number of rounds that are shown after training, Goal is the error rate and Epoch is the iterations required for the training process. Training stops whenever it reaches the determined epoch or the amount of the effectiveness function becomes less than the goal parameter. Table 1 demonstrates the considered values for network design parameters in this study.

**Table 1** | Characteristics of ANN training parameters

Parameter	Value/description
Show	100
Lr	0.05
Goal	$e^{-5}$
MC	0.9
Epochs	100
Function	Newff

In this study, zoning helped to select more appropriate wells for modeling in the ANN. As a result, nine wells were selected in the critical region. Six and three of the wells were selected for training and the test, with 160 and 85 recorded data, respectively. Afterwards, the modeling was conducted to find the optimal model by the following steps.

First step: a fixed model was considered for different input variables. This means that the type of transfer functions and layer structure were fixed and only inputs were different in the models. The selected variables for this step were the coordinates of each well (as UTM longitude (x) and latitude (y) for UTM zone 39), the date of the analysis (year and month), and the desired properties (TH, EC, and pH). The best subset of variables can be achieved by comparing eight models' performances which resulted in higher  $R^2$  and lower RMSE.

Second step: After specifying the most suitable input data in the ANN, they were examined for a specific layering of the combination of different transfer functions to find the best one by the model performance. Nine different models were evaluated in this step.

Third step: After detecting the best combination of transfer functions, layers with different neuronal arrangements were selected to optimize RMSE and  $R^2$  as the model performance indicators. In this step, 12 different neuronal arrangements were considered.

The single output of all models in each step is Na concentration. The whole procedure conducted in the study as mentioned above is summarized in a flowchart in Figure 2.

## RESULTS AND DISCUSSION

### Results of the water quality assessment

The most important physico-chemical properties as input variables were selected by evaluating the deviation of each property from the maximum allowed concentration proposed by FAO standards for irrigation use over 20 years. The results of this comparison showed that Na, at 120 times more than FAO standards, is one of the most important properties (Figure 3). In this regard, zoning of the Na concentration was conducted during these years

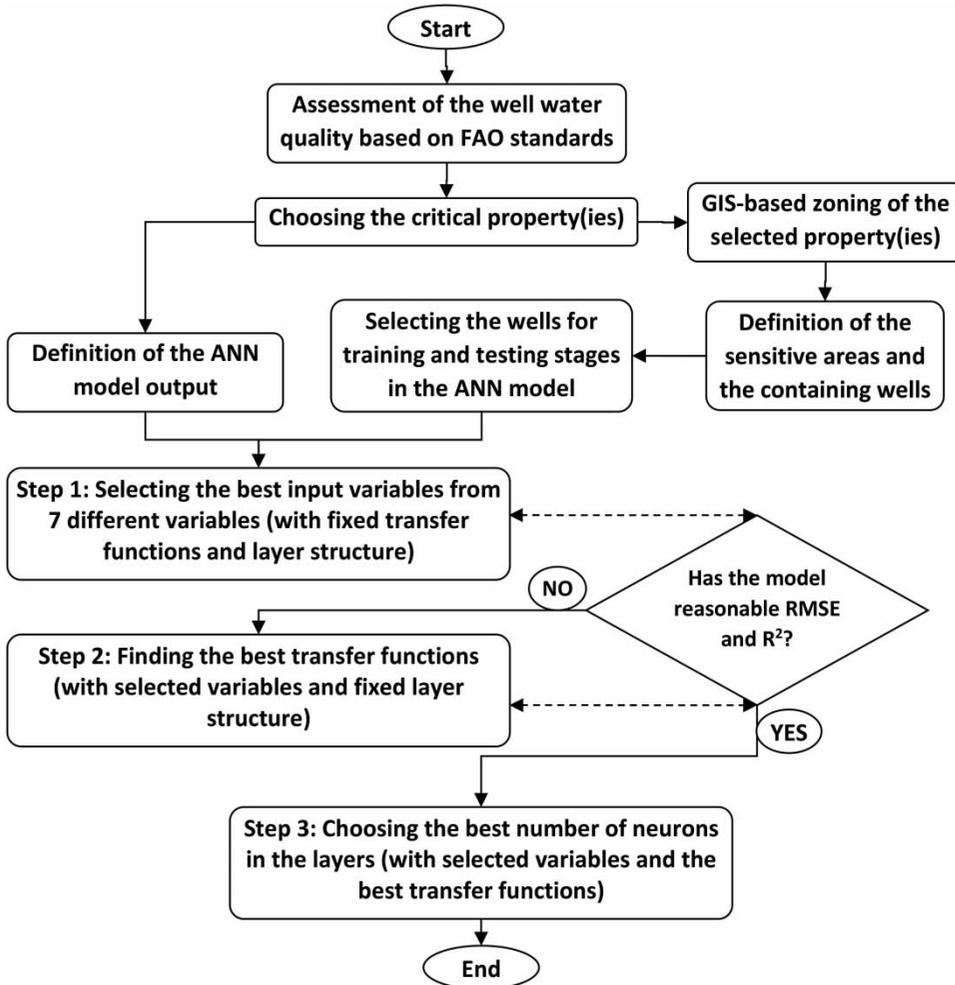


Figure 2 | Flowchart of the several steps conducted in the study.

using ArcGIS 10.1 software, as shown in Figure 4. Furthermore, as suggested by Mohammadi & Soltani (2011), TH, pH, and EC were the other important properties. Since

these properties can be easily measured by portable devices, Na was considered to be predicted by them in the ANN model.

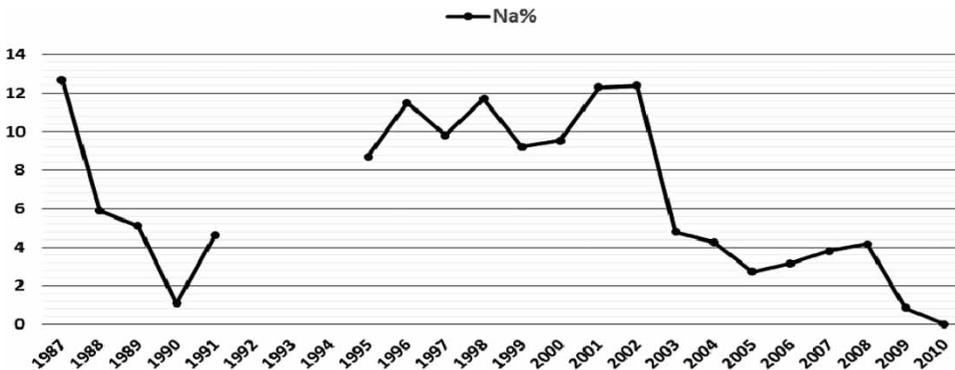


Figure 3 | The times where Na concentration is more than the suggested maximum value in the FAO standard in different years.

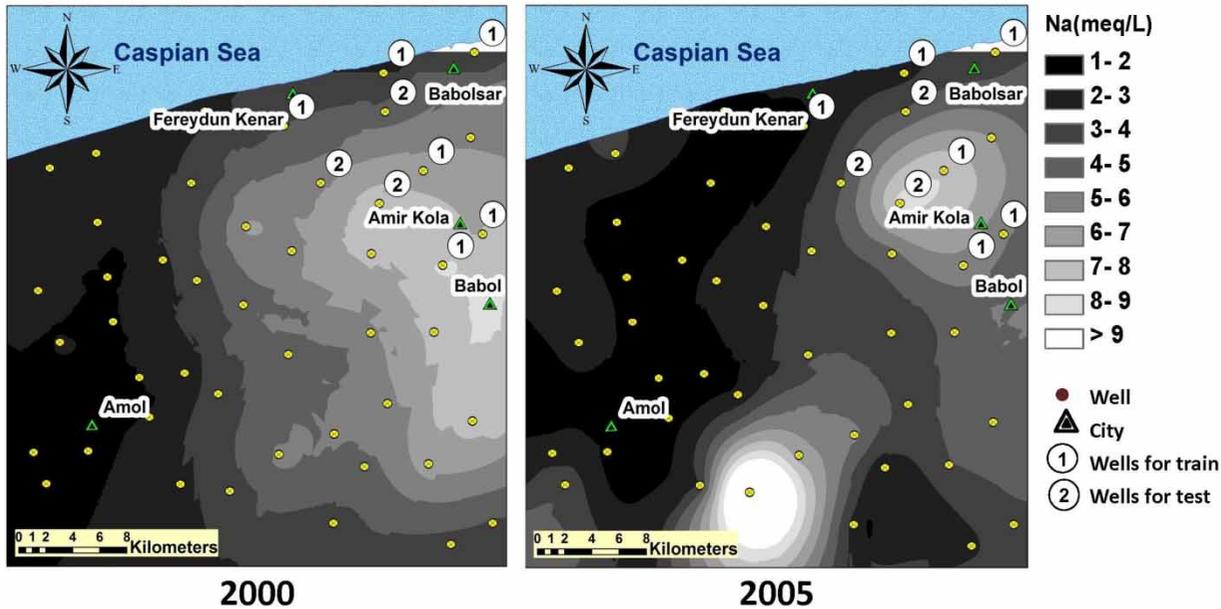


Figure 4 | Na zoning in selected years and location of training and testing wells used in the ANN.

There are 77 monitoring wells in the study area. Table 2 shows the mean, minimum, maximum, and standard deviation of the physico-chemical properties of all wells during 20 years based on 2,141 recorded data. The sampling frequency from wells during the data period was every six months, the middle of spring and autumn.

**Zoning results**

Table 3 shows the results of evaluating the GIS-based interpolation methods for Na. Due to the lower RMSE, the Kriging interpolation method is the best one in the region. This interpolation method is more efficient for many plains (Gallichand *et al.* 1992), however it may not be an appropriate method in some areas (Nakhaei & Mahmoodi 2012).

Table 2 | Water quality characteristics of the monitoring wells from 1987 to 2010 based on 2141 recorded data

Property	pH	Na <sup>+</sup> (meq/L)	HCO <sub>3</sub> <sup>-</sup> (meq/L)	Cl <sup>-</sup> (meq/L)	EC (μS/cm)
Average	7.7	3.35	6.86	2.32	1152
Standard deviation	0.28	3.57	1.89	3.2	518.6
Minimum	6.4	0.13	1.5	0.2	318
Maximum	8.6	44	25.1	45	5,530

Table 3 | Results of evaluating the interpolation methods for Na zoning

Interpolation method	RMSE
Kriging	1.70
Co-Kriging	1.95
RBF	3.65
IDW	7.25

The results of Na zoning showed that the more polluted area was dispersed between the cities of Amir Kola and Babol (especially in industrial parts). Hence, nine wells near the mentioned cities were investigated, of which six wells (near the city) and three wells (out of the city) were selected for training and testing of the ANN model, respectively (Figure 4). The lower distance of the training wells from the cities can be useful to reduce the cost of sampling due to fuel and time saving.

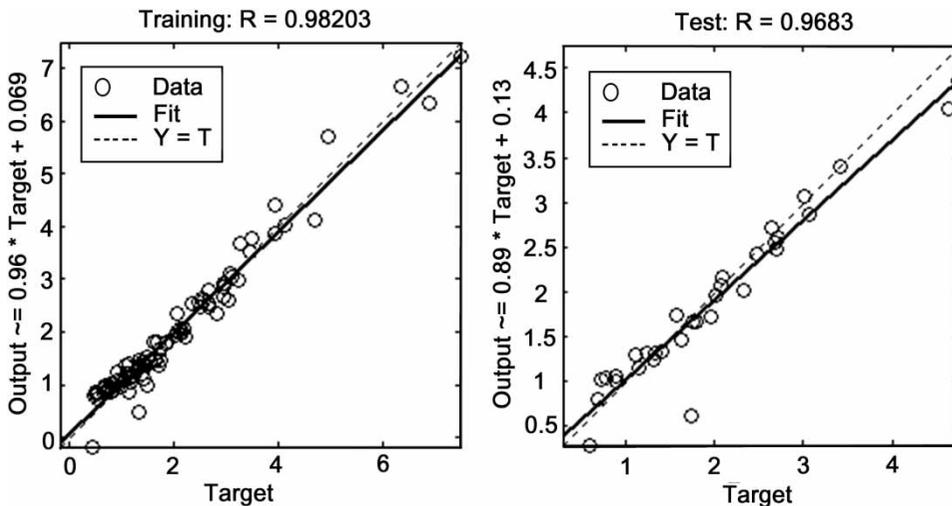
**ANN model results**

**Detection of the best possible inputs**

In the first step, seven input variables were introduced into the model including UTM(x) and UTM(y), sampling year and month, TH, EC, and pH. Then, eight different models were created to detect the best combination of the variables as shown in Table 4. In all models, the combination of

**Table 4** | Combination of different inputs in the ANN with fixed functions and layers and the model performance

Model performance	Input variables included in the model							Model name		
	UTM (x)	UTM (Y)	Year	Month	TH	pH	EC	Train	Test	RMSE
VR-Model 1	✓	✓	✓	✓	✓	✓	✓	0.98	0.95	0.21
VR-Model 2	✓	✓	-	-	✓	✓	✓	0.98	0.96	0.25
VR-Model 3	-	-	✓	✓	✓	✓	✓	0.98	0.96	0.22
VR-Model 4	-	-	✓	-	✓	✓	✓	0.98	0.96	0.21
VR-Model 5	-	-	-	-	✓	✓	✓	<b>0.98</b>	<b>0.96</b>	<b>0.20</b>
VR-Model 6	-	-	-	-	✓	✓	-	0.38	0.14	3.1
VR-Model 7	-	-	-	-	-	✓	✓	0.71	0.75	1.13
VR-Model 8	-	-	-	-	✓	-	✓	0.98	0.96	0.23



**Figure 5** | Results of training and testing stages with EC, pH and TH as inputs of VR-Model 5.

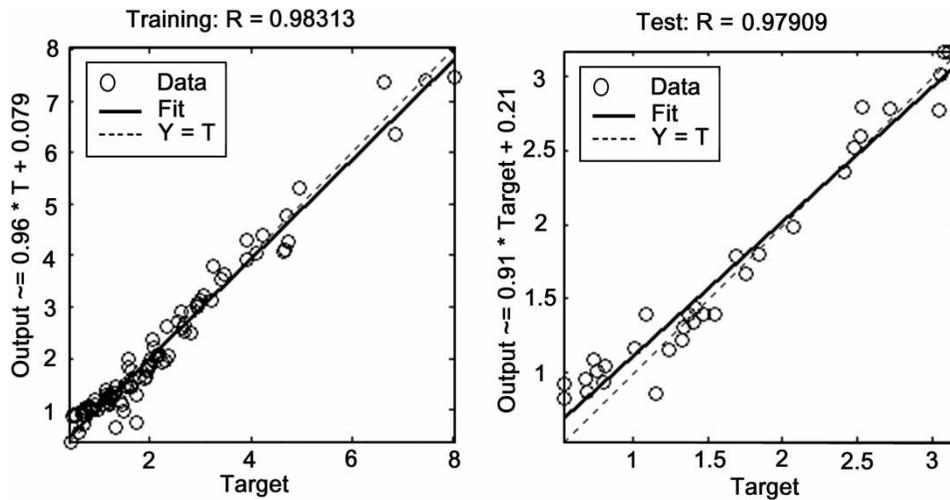
functions and the number of neurons were fixed (Tansig-Purelin, 1-1-2-1). The model with the highest  $R^2$  and lowest RMSE was likely to be the best one. As the model performance indicated in Table 4 and Figure 5, the VR-Model 5 is the best model. Thus, the best possible subsets of input variables are TH, pH and EC. If fewer input variables are considered, EC and TH will be the best subset.

**Selection of the best combination of functions**

In this step, different combinations of the transfer functions were evaluated. This evaluation was performed by a fixed layering of (1-1-2-1) as well as the fixed input variables of

**Table 5** | Comparison of different combinations of training functions with fixed layers

Model name	F(1)	F(2)	$R^2$		RMSE
			Train	Test	
TF-Model 1	tansig	purelin	0.98	0.96	0.20
TF-Model 2	purelin	purelin	0.98	0.94	0.29
TF-Model 3	purelin	tansig	0.98	0.96	0.23
TF-Model 4	tansig	tansig	0.98	0.96	0.20
TF-Model 5	tansig	logsig	0.98	0.97	0.20
<b>TF-Model 6</b>	<b>logsig</b>	<b>tansig</b>	<b>0.98</b>	<b>0.97</b>	<b>0.17</b>
TF-Model 7	logsig	logsig	0.98	0.96	0.19
TF-Model 8	logsig	purelin	0.98	0.95	0.23
TF-Model 9	purelin	logsig	0.98	0.96	0.22



**Figure 6** | Results of training and testing stages with EC, pH and TH as inputs and the combination of Logsig-Tansig function (TF-Model 6).

TH, EC, and pH. Performance of nine selected combinations is presented in Tables 5 and Figure 6. As a result, the combination of Logsig-Tansig (TF-Model 6) was the best one due to its lower RMSE and higher  $R^2$ .

### Finding the best layering

Based on the previous steps' results, the Logsig-Tansig function together with the three input variables of TH, EC, and pH were selected for change in the layering of the model, and after that the results were evaluated for the layers

**Table 6** | Comparison of different combinations of number of neurons in each layer with fixed transfer functions

Model name	Layer 1	Layer 2	$R^2$		RMSE
			Train	Test	
LY-Model 1	1	2	0.98	0.96	0.18
LY-Model 2	2	2	0.98	0.97	0.16
LY-Model 3	2	1	0.99	0.96	0.19
LY-Model 4	2	3	0.99	0.96	0.16
LY-Model 5	3	4	0.99	0.97	0.12
<b>LY-Model 6</b>	<b>4</b>	<b>3</b>	<b>0.99</b>	<b>0.98</b>	<b>0.08</b>
LY-Model 7	5	3	0.99	0.97	0.10
LY-Model 8	4	5	0.99	0.97	0.11
LY-Model 9	4	6	0.99	0.98	0.14
LY-Model 10	6	8	0.98	0.96	0.22
LY-Model 11	7	10	0.97	0.97	0.19
LY-Model 12	10	8	0.98	0.96	0.16

with different neurons. As shown in Table 6, 12 different neuron arrangements were selected to improve the performance of the model. The results illustrated that LY-Model 6 is the best model due to its lower RMSE and higher  $R^2$ .

According to these three steps, the analysis showed that the best ANN model should have the following specifications: input variables of TH, EC, pH (VR-Model 5, Table 4); transfer functions logsig and tansig (TF-Model 6, Table 5); and three neurons in the first layer and four neurons in the second (LY-Model 6, Table 6).

This model provides a good approximation to predict the concentration of Na in unknown wells with the  $R^2$  and the RMSE of 0.98 and 0.08, respectively (Figure 7).

## CONCLUSION

Based on the results of this study, the Na concentration of unknown wells was accurately predicted by the physico-chemical properties of the other nearby wells. Two subsets of the physico-chemical properties can be appropriate for the ANN model: (1) two input variable subsets, TH and EC; and (2) three input variable subsets, TH, pH, and EC. Moreover, the best structure for the ANN model should be Logsig-Tansig as transfer functions and 3 and 4 neurons in the first and second layer, respectively.

The main finding of the study is that unknown physico-chemical properties of some wells can be predicted by using the long-term data of other wells. This helps us to avoid

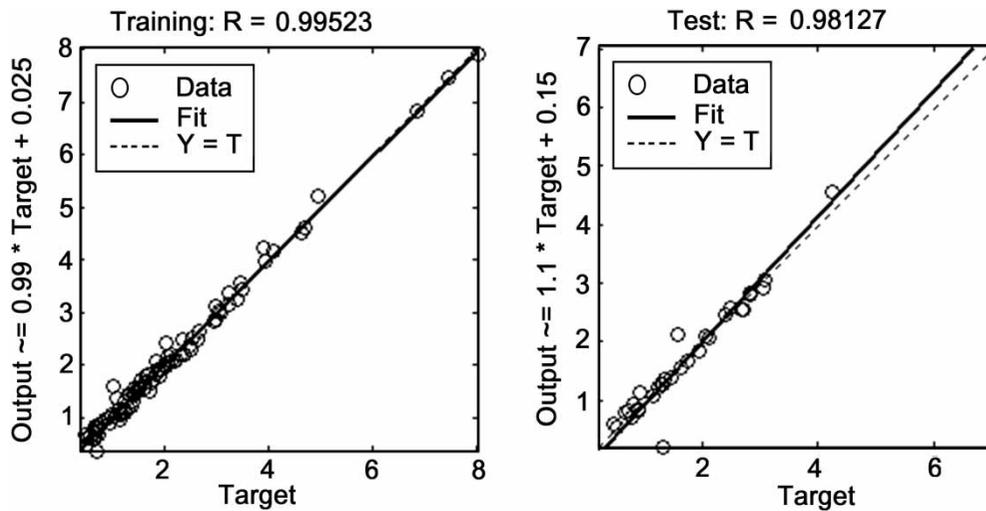


Figure 7 | Results of the training and testing stages of LY-Model 6 model with EC, pH and TH as inputs and the Logsig-Tansig function.

direct sampling or laboratory analysis of the unknown wells, especially when the wells are located far from the cities and direct sampling could be costly. For example, in the case of this study, we can continue analyzing six wells near the cities in the future rather than nine wells. Moreover, it would be possible to analyze fewer physico-chemical properties, in the case of this study, TH, pH and EC, in the wells near the cities. Also, some of the properties (pH and EC) can be analyzed by low-cost portable devices. Then, we can predict other desirable properties, in this study Na. Consequently, similar ANN models are recommended for other plains. The regional water authorities can save their laboratory and staff costs by such models, because they can analyze fewer physico-chemical properties for fewer wells without becoming concerned about the lack of data. Also, it is recommended that a periodic analysis of the unknown wells (for example at 3–5 year intervals) be undertaken to assess the validity of the model in the future.

In this study, we worked a prediction of a conservative chemical property of Na. Future works can focus on other important properties, weather conservative or non-conservative, such as nitrate.

## REFERENCES

Abbasi, M. P., Mehrdadi, N., Nabi Bidhendi, G. R. & Zare Abya, H. 2013 Application of artificial neural network to predict

total dissolved solids variations in groundwater of Tehran plain, Iran. *International Journal of Environment and Sustainability* 2, 10–20.

ESRI 2008 ArcGIS Spatial Analyst Tutorial, California, USA.

Gallichand, J., Bouckland, D., Marcotte, D. & Hendry, M. J. 1992 Spatial interpolation of soil salinity and sodicity for a saline soil in southern Alberta. *Canadian Journal of Soil Science* 72, 503–516.

Food and Agriculture Organization (FAO) Water quality for agriculture 2010. <http://www.fao.org/docrep/003/T0234E/T0234E01.htm#tab1> (accessed 11 July 2010).

Karamouz, M. & Araghi Nezhad, S. 2004 Application of artificial neural networks in flood estimation. In: *4th International Conference on Decision Making in Urban & Civil Engineering, Porto*, October 29–31.

Kohzadi, N., Boyd, M. S., Kermanshahi, B. & Kaastra, L. 1996 A comparison of artificial neural networks and time series model for forecasting commodity price. *Neurocomputing* 10, 169–181.

Mehrdadi, N., Hasanlou, H., Jafarzadeh, M. T., Hasanlou, H. & Abodolabadi, H. 2012 Simulation of low TDS and biological units of Fajr industrial wastewater treatment plant using artificial neural network and principal component analysis hybrid method. *Journal of Water Resource and Protection* 4, 370–376.

Menhaj, M. B. 2000 *Foundations of Artificial Neural Network*. Amirkabir University, Tehran, Iran (in Persian).

Mohammadi, A. & Soltani, J. 2011 Study of effect of drought on change in the quantity and quality of groundwater for agricultural usage on Jiroft plain. In: *Eleventh Conference on Irrigation and Reduced Evaporation, Shahid Bahonar University, Kerman, Iran* (in Persian).

Momeni, B., Palangi, J. A. & Aghdasi, B. 2011 Evaluation of groundwater level using artificial neural networks. In: *4th*

- National Conference on Water Resources Management, Tehran, Iran* (in Persian).
- Nakhaei, M. & Mahmoodi, D. 2012 Hydrochemical optimization interpolation methods in GIS. In: *16th Conference of the Association of Geologists, Shiraz University, Iran* (in Persian).
- Noorani, V., Hasanzadeh, Y., Komasi, M. & Sharghi, E. 2008 Modeling of rainfall – runoff using a combined wavy – fuzzy neural network. In: *Proceedings of 4th National Congress on Civil Engineering, Tehran, Iran* (in Persian).
- Pham, D. & Sagioglu, S. 2001 Training multilayered perceptrons for pattern recognition: a comparative study of four training algorithms. *International Journal of Machine Tools and Manufacture* **41**, 419–430.
- Ray, C. & Klindworth, K. K. 2000 Neural networks for agrichemical vulnerability assessment of rural private wells. *Journal of Hydrologic Engineering* **5**, 162171.
- Razavi, F. 2006 Rain prediction applying artificial neural network. MS Thesis, Amirkabir University, Tehran, Iran (in Persian).
- Rumelhart, D. E. & McClelland, J. L. 1986 *Parallel Distribution Processing: Exploration in the Microstructure of Cognition*. MIT Press, Cambridge, MA.
- Sharma, B. & Venugopalan, K. 2014 Comparison of neural network training functions for hematoma classification in Brain CT images. *IOSR Journal of Computer Engineering (IOSR-JCE)* **16**, 31–35.
- Singh, R. M. & Datta, B. 2010 Groundwater pollution source identification and simultaneous parameter estimation using pattern matching by artificial neural network. *Environmental Forensics* **5**, 143–153.
- Zare, A. H., Bayat, V. M. & Daneshkare, A. P. 2011 Forecasting nitrate concentration in groundwater using artificial neural network and linear regression models. *International Agrophysics* **25**, 187–192.

First received 14 May 2016; accepted in revised form 26 October 2016. Available online 30 January 2017