



ELSEVIER

Contents lists available at ScienceDirect

## Data in Brief

journal homepage: [www.elsevier.com/locate/dib](http://www.elsevier.com/locate/dib)

## Data Article

# Neuro-fuzzy inference system Prediction of stability indices and Sodium absorption ratio in Lordegan rural drinking water resources in west Iran

Afshin Takdastan<sup>a,g</sup>, Majid Mirzabeygi (Radfard)<sup>b,c,d,\*</sup>,  
 Mahmood Yousefi<sup>b</sup>, Abbas Abbasnia<sup>b</sup>,  
 Rouhollah Khodadadia<sup>a,b,c,d,e</sup>, Hamed Soleimani<sup>b</sup>,  
 Amir Hossein Mahvi<sup>b,e</sup>, Davood Jalili Naghan<sup>a,f,\*\*</sup>

<sup>a</sup> Department of Environmental Health Engineering, School of Health, Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran

<sup>b</sup> Department of Environmental Health Engineering, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

<sup>c</sup> Student's Scientific Research Center, Tehran University of Medical Sciences, Tehran, Iran

<sup>d</sup> Ava Salamat Entrepreneurs Institute, Tehran, Iran

<sup>e</sup> Center for Solid Waste Research, Institute for Environmental Research, Tehran University of Medical Science, Tehran, Iran

<sup>f</sup> Deputy Health Sharekord University of Medical Science Health, Sharekord, Iran

<sup>g</sup> Environmental Technologies Research Center, Ahvaz Jundishapur University of Medical Science, Ahvaz, Iran

## ARTICLE INFO

## Article history:

Received 17 January 2018

Accepted 28 February 2018

Available online 13 March 2018

## Keywords:

Drinking water

Villages of Lordegan city

Stability index

## ABSTRACT

According to World Health Organization guidelines, corrosion control is an important aspect of safe drinking-water supplies. Water always includes ingredients, dissolved gases and suspended materials. Although some of these water ingredients is indispensable for human beings, these elements more than permissible limits, could be endanger human health. The aim of this study is to assess physical and chemical parameters of drinking water in the rural areas of Lordegan city, also to determine corrosion indices. This cross-sectional study has carried out with 141 taken samples during 2017 with 13 parameters, which has been analyzed based on standard method and to estimate the water quality indices from

\* Corresponding author at: Department of Environmental Health Engineering, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran.

\*\* Corresponding author.

E-mail addresses: [Radfard.tums.ac.ir@gmail.com](mailto:Radfard.tums.ac.ir@gmail.com) (M. Mirzabeygi (Radfard)), [d.jalili225@gmail.com](mailto:d.jalili225@gmail.com) (D.J. Naghan).

groundwater using ANFIS. Also with regard to standard conditions, results of this paper are compared with Environmental Protection Agency and Iran national standards. Five indices, Ryznar Stability Index (RSI), Langlier Saturation Index (LSI), Larson-Skold Index (LS), Puckorius Scaling Index (PSI), and Aggressive Index (AI) programmed by using Microsoft Excel software. Owing to its simplicity, the program, can easily be used by researchers and operators. Parameters included Sulfate, Sodium, Chloride, and Electrical Conductivity respectively were 13.5, 28, 10.5, and 15% more than standard level. The amount of Nitrate, in 98% of cases were in permissible limits and about 2% were more than standard level. Result of presented research indicate that water is corrosive at 10.6%,89.4%,87.2%,59.6% and 14.9% of drinking water supply reservoirs, according to LSI, RSI, PSI, LS and AI, respectively.

© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

### Specifications Table

Subject area	Chemistry
More specific subject area	Describe narrower subject area
Type of data	Tables, Figures
How data was acquired	This cross-sectional study has carried out with 141 taken samples during 2017 with 13 parameters, which has been analyzed based on standard method and to estimate the water quality indices from groundwater using ANFIS
Data format	Raw, Analyzed
Experimental factors	The mentioned parameters above, in abstract section, were analyzed according to the standards for water and wastewater treatment handbook.
Experimental features	The levels of physical and chemical parameters were determined.
Data source location	Chaharmahal and Bakhtiari province, Lordegan city Iran
Data accessibility	Data are included in this article

### Value of the data

- Selecting the most realistic appropriate index, is one of the most sophisticated stages of water stability evaluations.
- The results of this study clearly indicated that with appropriate selection of input variables, ANFIS as a soft computing approach can estimate water quality indices properly and reliably.
- ANFIS approach is also extensible to solve, the other problems related to manage of the groundwater quality in agriculture and drinking water sectors such as nitrate, sulfate and etc.
- Considering present study, many of rural drinking water supply reservoirs, need to pay attention to achieve Iran national water quality standards.

### 1. Data

The Table 1 shows Pearson's correlation factors between chemical water quality parameters and water indices and Table 2 shows predicting performance in different steps of ANFIS, Fig. 1. Checking and training errors (PSI index prediction) for optimization of epochs, Fig. 2 also shows estimated indexes versus real indexes for checking dataset, Fig. 3 shows comparison of first and second steps

optimization in estimating of indexes for checking dataset and finally Fig. 4 shows location of water sampling sites in Lordegan city, Chaharmahal and Bakhtiari province.

## 2. Experimental design, materials and methods

### 2.1. Study area description

Chaharmahal and Bakhtiari province is one of the 31 provinces of Iran. Lordegan is a city in Chaharmahal and Bakhtiari Province, Iran that coordinates is: 31°30'37"N 50°49'46"E.

### 2.2. Materials and methods

#### 2.2.1. Sample collection and analytical procedures

This cross-sectional study has carried out with 141 taken samples during 2017 with 13 parameters, which has been analyzed based on standard method [1–8].

#### 2.2.2. Description of ANFIS

With introduction of fuzzy inference systems to the world of research, because of the great features of these methods in the control systems, as well as simple expression of the variables in terms of linguistic

**Table 1**

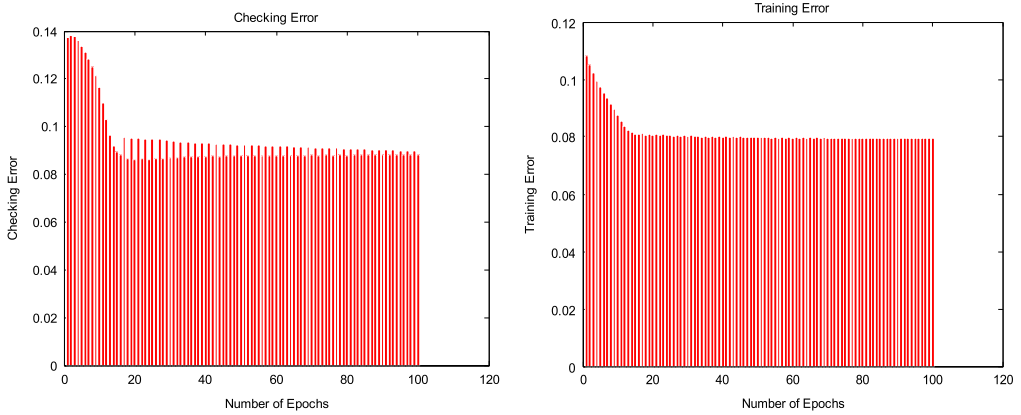
Pearson's correlation factors between chemical water quality parameters and water indices.

	TDS	TA	CH	PSI	RSI	AI	SAR
TDS	1						
TA	0.254467	1					
CH	0.413105	0.150231	1				
PSI	-0.38863	-0.7011	-0.74581	1			
RSI	-0.41141	-0.52671	-0.78821	0.941166	1		
AI	0.444762	0.478963	0.705398	-0.83255	-0.9527	1	
SAR	0.720472	0.111289	-0.03624	0.010696	-0.01206	0.10911	1

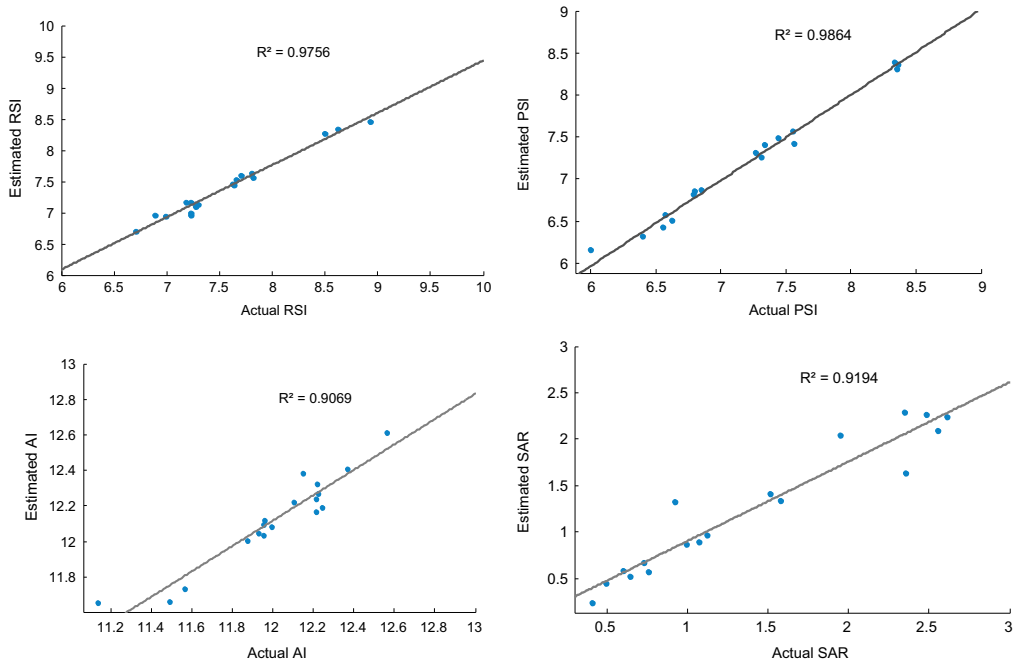
**Table 2**

Predicting performance in different steps of ANFIS.

Indexes	RMSE		2nd step optimization	Radii	Epoch	CC
	Initial	1st step optimization				
PSI						
- Train	0.1082	0.0875	0.0806	0.9	16	
- Check	0.1375	0.1164	0.0861			0.993
RSI						
- Train	0.2240	0.1839	0.1553	0.75	20	
- Check	0.2887	0.2808	0.2163			0.984
AI						
- Train	0.1363	0.1159	0.0969	0.6	100	
- Check	0.2284	0.2028	0.1673			0.952
SAR						
- Train	0.6474	0.6060	0.5473	0.55	20	
- Check	0.3284	0.3154	0.2713			0.9588

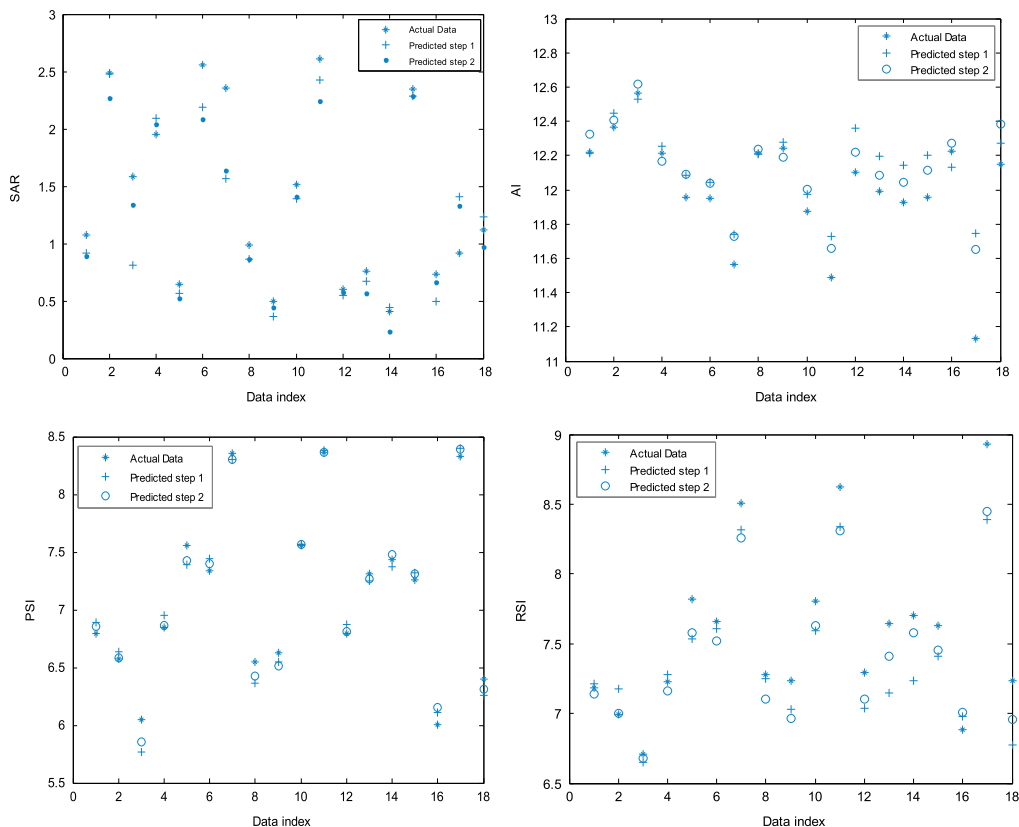


**Fig. 1.** Checking and training errors (PSI index prediction) for optimization of epochs.



**Fig. 2.** Estimated indexes versus real indexes for checking dataset.

variables, use of this method in many applications was developed and expanded quickly. But designing a powerful fuzzy inference system, which needs to determine appropriate membership functions and fuzzy rules, is not easy, and mainly will be possible with a lot of trial and error and with the involvement of experts' comments in order to achieve the best performance, and in some cases, identification of all the rules is impossible. For this reason, the idea of using learning algorithms of the artificial neural network was introduced for fuzzy systems. ANFIS is an adaptive neural fuzzy inference system which was first proposed in 1993 by Jung. The system is an intelligent hybrid method, which uses features of the Takagi-Sugeno fuzzy inference system and adaptive neural network [9–11]. ANFIS, using artificial neural network learning algorithms, including gradient descent algorithm and back-propagation algorithm leads to improve fuzzy system rules. ANFIS, because of having the adaptive fuzzy system parameters compared



**Fig. 3.** Comparison of first and second steps optimization in estimating of indexes for checking dataset.

with artificial neural network, is trained faster, and usually provides better results. Schematic structure of ANFIS, with two input variables and an output variable, which lead to the Sugeno model, is shown in Fig. 1. Two rules (Eqs. (1), (2)) in the Sugeno fuzzy model can be written as follows:

$$\text{If } x \text{ is } A1 \text{ and } y \text{ is } B1 \text{ then } f1 = p1x + q1 + r1 \tag{1}$$

$$\text{If } x \text{ is } A2 \text{ and } y \text{ is } B2 \text{ then } f2 = p2x + q2 + r2 \tag{2}$$

Generally, the structure of an ANFIS model has 5 distinct layers, and each of these layers plays a special role. In the first layer, every node  $i$ , according to Eq. (3), is known as a comparative node, and the outputs of this layer represent the degree of membership assigned to inputs in the fuzzy form.

$$O^3_i = \mu_{A_i}(x) \tag{3}$$

In the above equation,  $x$  is an input variable to node  $i$  and  $A_i$  is small, large... linguistic labels, that are associated with the node function. In other words,  $O^3$  is membership function of fuzzy set  $A$ , which the input variable  $x$  has certain circumstances in it. Membership function  $A$ , can be Gaussian function or triangular membership function and so on. The membership function parameters that are used in this layer are known as the first parameter. In MATLAB, using `genfis1` and `genfis2` functions, you can create a fuzzy inference system, which covers the first layer of ANFIS [9,12]. In the `genfis1` function, to create a fuzzy system, there is to need to specify the number and type of membership function well. However, `genfis2` function creates this fuzzy system using Subtractive clustering. In this study, `genfis2` function is used with Gaussian membership function and Subtractive clustering to create the fuzzy system. In the Subtractive clustering, it is assumed that all data have candidate for clustering center according to the density of the

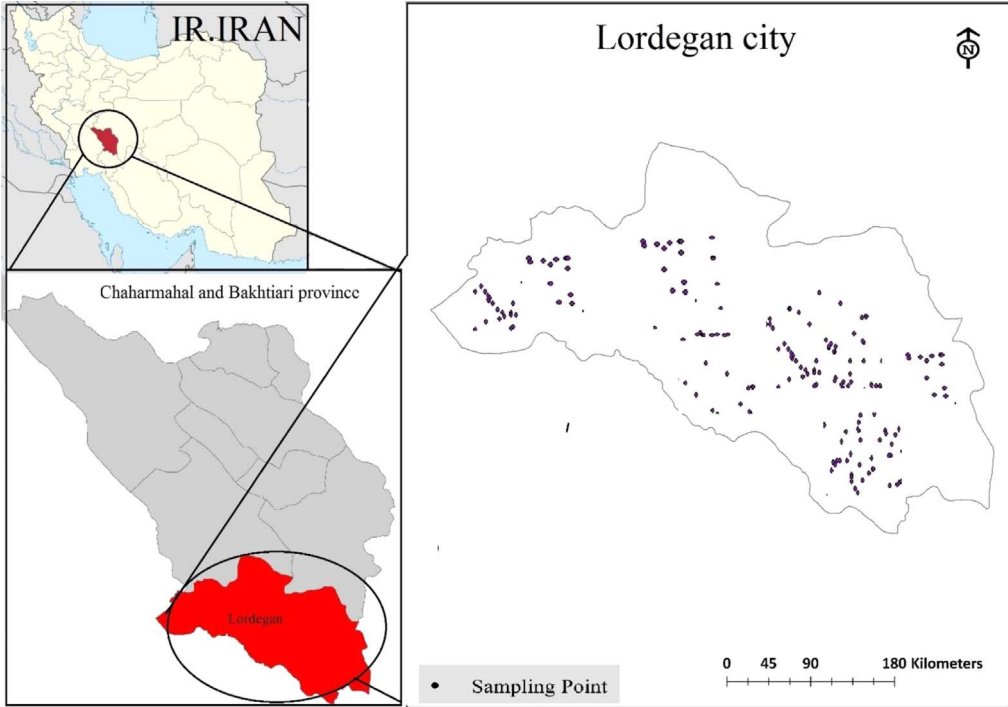


Fig. 4. Location of water sampling sites in Lordegan city, Chaharmahal and Bakhtiari province.

surrounding data. So, using the following Eq. (4), we can have a measure of density of data  $x_i$ .

$$D_i = \sum_{j=1}^n \exp \left[ \frac{\|X_i - X_j\|^2}{(ra/2)^2} \right] \tag{4}$$

Here,  $ra$  is a positive constant, which indicates the radius of the neighbored radius. Therefore, a data with the greatest amount of density ( $D_i$ ), has more the number of data points which are located in the surrounding neighborhood. First cluster center  $X_{c1}$  is dedicated to data that has the highest density  $D_{c1}$ . The density value of each data point  $x_i$  is modified as follows Eq. (5) [9–12].

$$D_i = D_i - D_{c1} \exp \left[ \frac{-\|X_i - X_{c1}\|^2}{(\frac{rb}{2})^2} \right] \tag{5}$$

Here,  $rb$  is a positive constant, and defines a neighborhood, which has a significant reduction in the amount of density. Therefore, the numbers of data points that are located near the center of the first cluster  $x_{c1}$  significantly reduce the amount of density. After correcting the density function, the next cluster center is belonged to a data which has maximum density. This process continues until a sufficient number of clusters are reached. In the second layer of ANFIS, the outputs of the first layer are combined. As a result, the output of this layer can be written according to Eq. (6).

$$O_1^6 = W_1 = \mu_{A_i}(X) \mu_{B_i}(Y) \tag{6}$$

Here, each output node burning represents the firing strength of a rule. The next layer is the second layer, which normalizes the output of before layers output, and calculates the firing strength

of  $i$ -th rule than the firing strength of all the rules based on Eq. (7).

$$O_1^7 = W_i = \frac{W_i}{W_1 + W_2} \quad (7)$$

Nodes of the fourth layer have the adaptive nature, and the parameters of this node ( $r_i$ ,  $q_i$ ,  $p_i$ ) are called as consequent parameters. Eq. (8) can be written to describe the layer 4. In this equation,  $f$  is a first order function of the Sugeno fuzzy system.

$$O_i^8 = W_i f_i = W_i (p_i x + q_i y + r_i) \quad (8)$$

And finally, for the output of layer 5, which is the final output of the ANFIS system, we can write Eq. (9) [9–12].

$$O_i^5 = z = \sum_i \bar{w}_i f_i \quad (9)$$

## Acknowledgement

The authors expressed their appreciate to the health deputy of Lordegan Medical Sciences University and Rural Water and Wastewater Office, about providing information and their Cooperation in preparation laboratory facilities.

## Transparency document. Supporting information

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2018.02.075>.

## References

- [1] M. Mirzabeygi, N. Yousefi, A. Abbasnia, H. Youzi, M. Alikhani, A.H. Mahvi, Evaluation of groundwater quality and assessment of scaling potential and corrosiveness of water supply networks, Iran, *J. Water Supply: Res. Technol.-Aqua* (2017) (jws2).
- [2] M. Mirzabeygi, M. Naji, N. Yousefi, M. Shams, H. Biglari, A.H. Mahvi, Evaluation of corrosion and scaling tendency indices in water distribution system: a case study of Torbat Heydariye, Iran, *Desalination Water Treat.* 57 (2016) 25918–25926.
- [3] A. Abbasnia, M. Alimohammadi, A.H. Mahvi, R. Nabizadeh, M. Yousefi, A.A. Mohammadi, H. Pasalari, M. Mirzabeygi, Assessment of groundwater quality and evaluation of scaling and corrosiveness potential of drinking water samples in villages of Chabahr city, Sistan and Baluchistan province in Iran, *Data Brief* 16 (2018) 182–192.
- [4] A.A. Mohammadi, M. Yousefi, A.H. Mahvi, Fluoride concentration level in rural area in Poldasht city and daily fluoride intake based on drinking water consumption with temperature, *Data Brief* 13 (2017) 312–315.
- [5] M. Yousefi, H. Najafi Saleh, A.A. Mohammad, A.H. Mahvi, M. Ghadrpoori, H. Suleimani, Data on water quality index for the groundwater in rural area Neyshabur County, Razavi province, Iran, *Data Brief* 15 (2017) 901–907.
- [6] H. Soleimani, A. Abbasnia, M. Yousefi, A.A. Mohammadi, Changani Khorasgani Fazlollah. Data on assessment of groundwater quality for drinking and irrigation in rural area Sarpol-e Zahab city, Kermanshah province, Iran, *Data Brief* 17 (2018) 148–156.
- [7] F.B. Asghari, A.A. Mohammadi, Z. Aboosaedi, M. Yaseri, M. Yousefi, Data on fluoride concentration levels in cold and warm season in rural area of Shout (West Azerbaijan, Iran), *Data Brief* 15 (2017) 528–531.
- [8] A.A. Mohammadi, H. Najafi Saleh, A.H. Mahvi, M. Alimohammadi, R. Nabizadeh, M. Yousefi, Data on corrosion and scaling potential of drinking water resources using stability indices in Jolfa, East Azerbaijan, Iran, *Data Brief* 16 (2018) 724–731.
- [9] V. Ranković, J. Radulović, I. Radojević, A. Ostojčić, L. Čović, Neural network modeling of dissolved oxygen in the Gruža reservoir, Serbia, *Ecol. Model.* 221 (2010) 1239–1244.
- [10] L. Boriwal, R. Sarviya, M. Mahapatra, Modeling the resistance spot welding of galvanized steel sheets using Neuro-Fuzzy Method, *Int. Proc. Adv. Soft Comput., Intell. Syst. Appl.* (2018) (pp.37–5).
- [11] M.I. Yesilnacar, E. Shahinkaya, Artificial neural network prediction of sulfate and SAR in an unconfined aquifer in southeastern Turkey, *Environ. Earth Sci.* 67 (2012) 1111–1119.
- [12] M. Ay, Ö. Kişi, Estimation of dissolved oxygen by using neural networks and neuro fuzzy computing techniques, *KSCE J. Civil. Eng.* 21 (2017) 163–169.