Data on Estimation for Sodium Absorption Ratio: Using Artificial Neural Network and Multiple Linear Regressions

7 authors, including:

Majid Radfarad
Shiraz University of Medical Sciences
44 PUBLICATIONS   **413** CITATIONS

SEE PROFILE

Hamed Soleimani
Tehran University of Medical Sciences
26 PUBLICATIONS   **276** CITATIONS

SEE PROFILE

Hamed Akbari
Islamic Azad University Tehran Science and Research Branch
29 PUBLICATIONS   **180** CITATIONS

SEE PROFILE

Amir Adibzadeh
Health Research Center
29 PUBLICATIONS   **69** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Evaluation of corrosion and scaling tendency indices in water distribution system: a case study of Torbat Heydariye, Iran View project

Impacts of Climate Change and Drought Phenomenon on the Chemical Quality of Groundwater Resources View project

Title: Data on Estimation for Sodium Absorption Ratio: Using Artificial Neural Network and Multiple Linear Regressions

Article Type: Data Article

Corresponding Author: Dr. Amirhossin mahvi,

Corresponding Author's Institution:

First Author: Majid Mirzabeygi (Radfarad)

Order of Authors: Majid Mirzabeygi (Radfarad); Amirhossin mahvi; Hamed Soleimani; Amir Adibzadeh; Hesam Akbari; hamed Akbari; samira nabavi

Abstract: The sodium absorption ratio (SAR) is one of the parameters used in assessing groundwater quality for irrigation purposes. The specified parameter is very important in terms of problems related to water permeability and soil ventilation. In this study, the data of the groundwater quality of Aras catchment area in the years 2010-2014 were investigated for estimating the sodium absorption ratio(SAR). The neural network is defined as a system of processor elements, called neurons, which create a network by a set of weights. In this study, a 3-layer MLP neural network including a hidden layer, an input layer and an output layer was designed. The number of neurons in the input and output layers of this network was considered 4 and 1, respectively, due to having four input variables (pH, sulfate, chloride and electrical conductivity) and one output variable (sodium absorption ratio). The effect of pH, sulfate, chloride and electrical conductivity were estimated to be 11.34%, 72.22%, 94 % and 91%, respectively. According to these results, it can be said that in estimating the sodium absorption ratio to the chloride variable, the designed neural network is more sensitive than other variables. In this research, artificial neural network and multiple linear regression methods were used to estimate the rate of sodium absorption ratio of groundwater resources of the Aras catchment area. The results of these two methods were compared with the model's performance evaluation criteria, namely root mean square error(RMSE), mean absolute error (%) and correlation coefficient. The results of this study showed that artificial neural network is a helpful and exact tool for predicting the amount of sodium absorption ratio(SAR) in groundwater resources of Aras catchment area and these results are not comparable with the results of multiple linear regressions.

Detailed Response to Reviewers

All comments were corrected

Data article

Data on Estimation for Sodium Absorption Ratio: Using Artificial Neural Network and Multiple Linear Regressions

Majid Radfard[a], Hamed Soleimani[b], Samira Nabavi[b], Bayram Hashemzadeh [c] , Hesam Akbari[a**], Hamed Akbari[a],Amir Adibzadeh*

[a] Health Research Center, Life Stayle institute, Baqiyatallah University of Medical Sciences, Tehran.

[b] Department of Environmental Health, School of public Health, Tehran University of Medical Sciences, Tehran, Iran.

[c] Department of Environmental Health, School of public Health, Khoy University of Medical Sciences, Khoy, Iran.

[*]Corresponding author. Health Research Center, Life Style institute, Baqiyatallah University of Medical Sciences, Tehran. Email: Radfard.tums.ac.ir@gmail.com (Amir Adibzadeh)

[**]Corresponding author. Health Research Center, Life Style institute, Baqiyatallah University of Medical Sciences, Tehran. Email:Radfard.tums.ac.ir@gmail.com (Hesam Akbari)

Abstract:

In this article the data of the groundwater quality of Aras catchment area were investigated for estimating the sodium absorption ratio (SAR) in the years 2010-2014. The artificial neural network (ANN) is defined as a system of processor elements, called neurons, which create a network by a set of weights. In the present data article, a 3-layer MLP neural network including a hidden layer, an input layer and an output layer had been designed. The number of neurons in the input and output layers of network was considered to be 4 and 1, respectively, due to having four input variables (including: pH, sulfate, chloride and electrical conductivity (EC)) and only one output variable (sodium absorption ratio). The impact of pH, sulfate, chloride and EC were estimated to be 11.34%, 72.22%, 94 % and 91%, respectively. ANN and multiple linear regression methods were used to estimate the rate of sodium absorption ratio of groundwater resources of the Aras catchment area. The data of both methods were compared with the model's performance evaluation criteria, namely root mean square error (RMSE), mean absolute error (%) and correlation coefficient. The data showed that ANN is a helpful and exact tool for predicting the amount SAR in groundwater resources of Aras catchment area and these results are not comparable with the results of multiple linear regressions.

Specifications Table

Subject area	Chemistry
More specific subject area	Water quality and monitoring
Type of data	Tables , Figures
How data was acquired	Data on groundwater resources quality in the Aras catchment area was obtained from West Azerbaijan Water and Wastewater Company during the years 2010-2014 and was studied for estimation of sodium absorption ratio (SAR).
Data format	Raw, Analyzed
Experimental factors	The sodium absorption ratio (SAR), 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	Aras, West Azerbaijan province, Iran.
Data accessibility	Data are included in this article
Related research article	A.Takdastana, M. Mirzabeygi (Radfard), M.Yousefi, A. Abbasnia, R. Khodadadia, A H.Mahvi, D.Jalili Naghan, Neuro-fuzzy inference system Prediction of stability indices and Sodium absorption ratio in Lordegan rural drinking water resources in west Iran, Data in Breif 18(2018)255-261.

Value of the data

- The data of this article can be used to environmental management and better exploitation of groundwater resources.
- Considering the present data, many of the sampling drinking water supply reservoirs need to pay attention to achieve Iran national water quality standards.

.

- The results clearly indicate that with appropriate selection of input variables, Artificial Neural Network and Multiple Linear Regressions as a soft computing approach can be used to estimate water quality indices properly and reliability

1. Data

Two algorithms, including seven Back-propagation algorithms and Lewenberg-Markow, have been used in this data article. The Comparison of the performance of seven Back-propagation algorithms in estimating the sodium absorption ratio with the number of neurons 10 in the hidden layer has been shown in Table 1 and Table 2, indicating the comparison of the different neurons performance in the hidden layer in estimating the sodium absorption ratio using the Lewenberg-Markow algorithm. The optimized output of neural network and data performance criteria for has been shown in Figure 1 and 2. Also, Figure 3 indicates the actual SAR values in groundwater resources and their predicted values via multiple linear regression.

Table 1. Comparison of the performance of seven Back-propagation algorithms in estimating the sodium absorption ratio with the number of neurons 10 in the hidden layer

Back-propagation algorithms	Neural network process	Evaluation of model performance			Repeat number
		R	MAE	MSE	
	Training	0.883	1.34	5	132
Trainbfg (BFGS quasi-Newton)	Test	0.819	1.33	4.54	132
	Training + Test + Validation	0.866	1.36	4.7	132
	Training	0.916	1.15	3.07	39
Traincgp (Polak–Ribie´re	Test	0.855	1.2	3.16	39
conjugate gradient)	Training + Test + Validation	0.893	1.2	3.78	39
	Training	0.726	13.09	197.07	52
Traingd (gradient descent)	Test	0.736	13.89	221.88	52
	Training + Test + Validation	0.762	13.17	193.17	52
	Training	0.806	1.71	7.25	500
Traingda (adaptive learning	Test	0.85	1.54	4.58	500
rate back-propagation)	Training + Test + Validation	0.811	1.68	6.55	500
	Training	0.898	1.23	4.13	143
Trainscg (scaled conjugate	Test	0.86	1.16	2.55	143
gradient)	Training + Test + Validation	0.89	1.25	3.85	143
	Training	0.868	1.4	4.9	34
Traincgf (Fletcher–Powell	Test	0.0804	1.52	4.49	34

conjugate gradient)	Training + Test + Validation	0.85	1.43	5.16	34
	Training	0.906	1.05	2.77	29
Trainlm (Levenberg–Marquardt)	Test	0.9	.92	1.9	29
	Training + Test + Validation	0.0901	1.08	3.52	29

Table 2. Comparison of the different neurons performance in the hidden layer in estimating the sodium absorption ratio using the Lewenberg-Markow algorithm

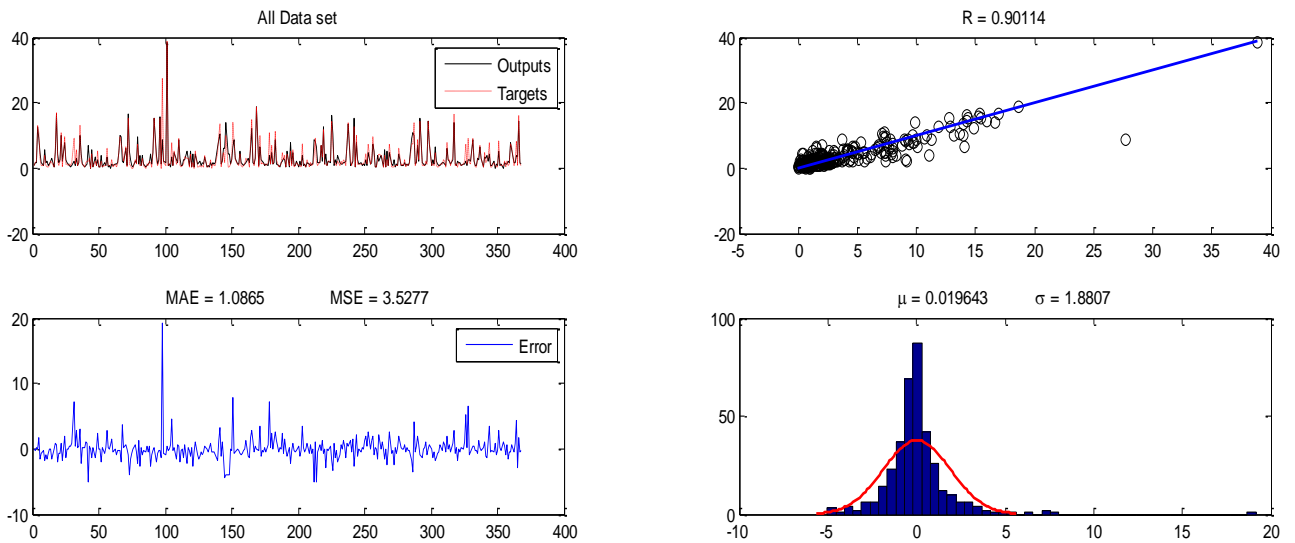| Number of neurons	R	MAE	MSE	Repeat number
3	0.879	1.23	4.59	25
5	0.86	1.24	4.81	32
7	0.91	1.22	3.61	25
10	0.9	0.923	1.9	29
12	0.89	1.14	4.03	33
17	0.88	1.09	2.37	29
20	0.895	1.08	3.76	30
30	0.892	1.09	3.92	34

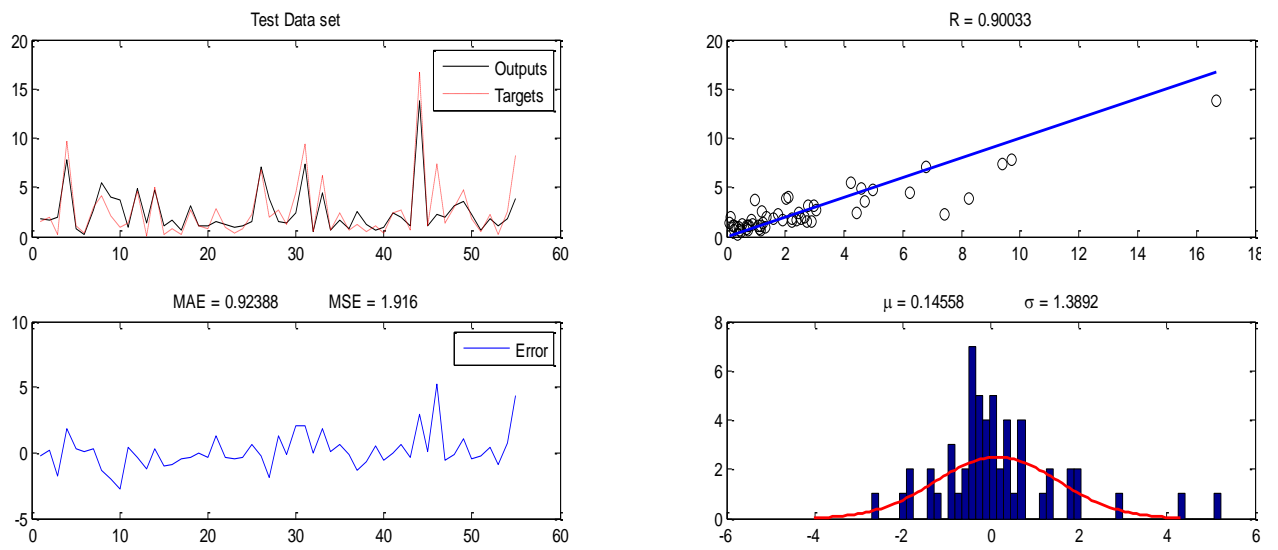Fig1. Optimized neural network output and model performance criteria for all data.

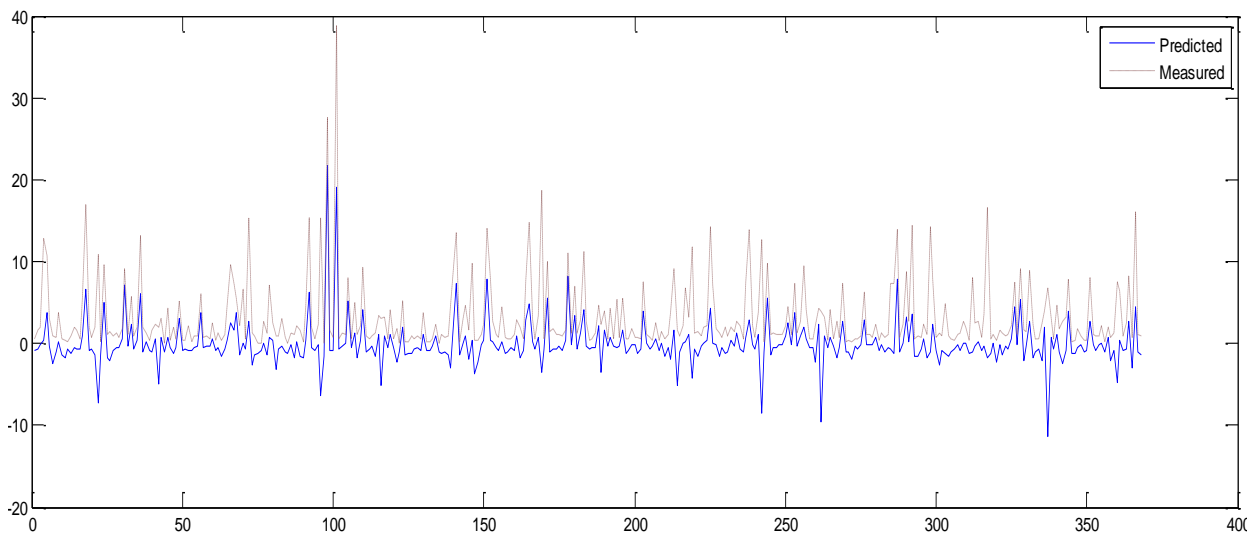**Fig 2.** Optimized neural network output and model performance criteria for test data.

Fig 3. Actual SAR values in groundwater resources and their predicted values with multiple linear regression.

2. Experimental Design, Materials and Methods

2.1. Study area description

Aras Catchment area is a plain located in the northern half of the East Azerbaijan province, West Azerbaijan province and Ardabil province. Extensive precipitation in this region, in addition to its impact on climate moderation, has created numerous rivers [18].
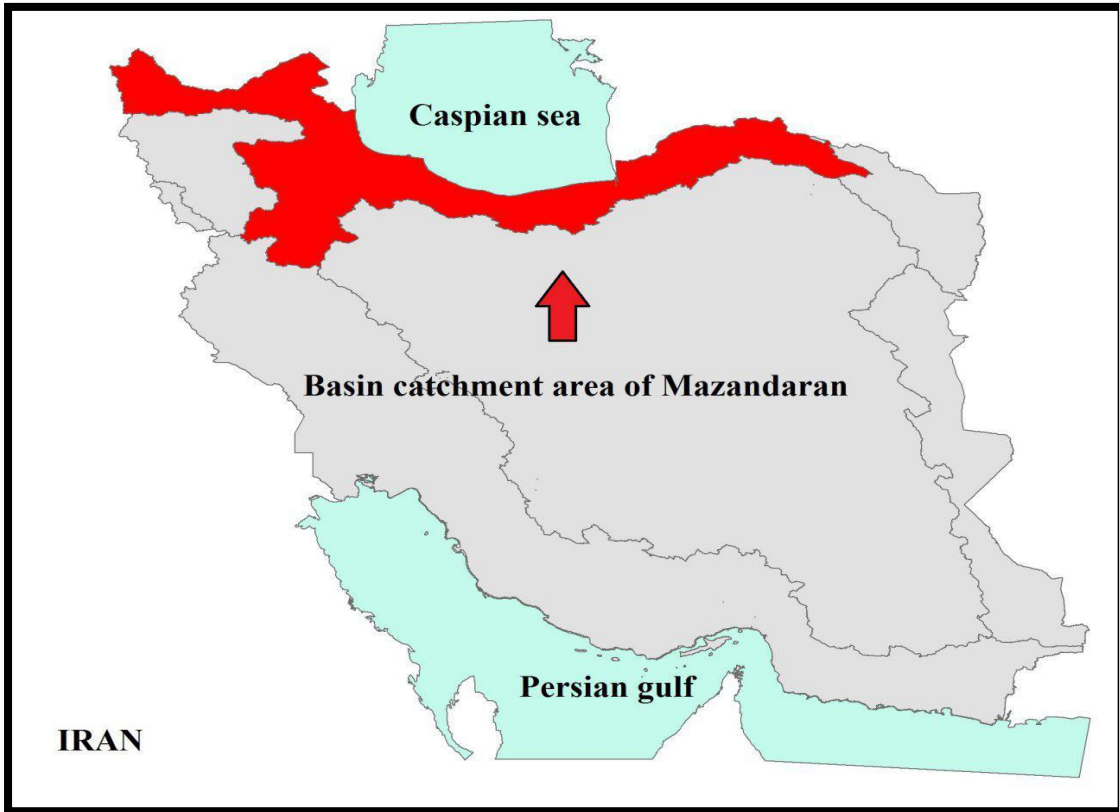
Fig 4. Study area

2.2. Material and methods:

Data on groundwater resources were collected a during the years 2010-2014 and water samples were analyzed following the standard methods for examination of water and waste water [1-10] in terms of estimation of sodium absorption ratio (SAR). A two-layer neural network with a tangent-sigmoid transfer function for the hidden layer and a linear transfer function for the output layer was used. The input parameters of the neural network included sulfate, chloride, electrical conductivity (EC) and pH, and the sodium absorption ratio (SAR) were considered as the network output parameter. The data on these parameters were divided into training, testing and data validation. 70% of these data was used for training, 15% of data for validation and other 15% for testing. Considering that today BP neural networks have become a common tool for modeling environmental systems, so in this study, 8 BP algorithms were selected and their results were tested to obtain the best algorithm. For all algorithms, a dual layer network with a tan-sigmoid transfer function on the hidden layer and

a linear transfer function in the output layer was used. In choosing the best BP algorithm, the number of neurons was considered 10. The results of the model's performance with their BP algorithm are presented in Table 1. The performance of the BP algorithm was evaluated with mean squared error (MSE), mean absolute error (MAE) and correlation coefficient (R) between the output of the models and the actual data set. The algorithm with the least training error and the maximum correlation coefficient was selected as the most suitable algorithm. The Langenberg-Marquard algorithm (trainlm) was chosen as the best algorithm to predict the sodium adsorption ratio (SAR). To optimize the number of neurons after selecting the best BP algorithm, namely Levenberg-Marquard, the number of neurons was optimized by keeping other parameters intact. As shown in Table 2, in the number of neurons greater than the optimal number of neurons, 10, the mean square error (MSE) was not significantly altered. Therefore, all the modeling steps were done based on the number of neurons 10 and the Lewenberg-Markow algorithm to predict the sodium absorption ratio [1-20].

Acknowledgements:

Conflict of Interest: The authors of this article declare that they have no conflict of interests.

References:

[1] Yousefi, M., et al., Modification of pumice with HCl and NaOH enhancing its fluoride adsorption capacity: Kinetic and isotherm studies. Human and Ecological Risk Assessment: An International Journal, 2018: p. 1-13.

[2] Babaeia, A., et al., Data on groundwater quality, scaling potential and corrosiveness of water samples in Torbat-e-Heydariyeh rural drinking water resources, Khorasan-e-Razavi province, Iran. Data in Brief, 2018. 19: p. 2260-2266.

[3] Heidarinejad, Z., et al., Data on Quality Indices of Groundwater resource for Agricultural Use in the Jolfa, East Azerbaijan, Iran. Data in Brief, 2018.

[4] Jalili, D., et al., Data on Nitrate-Nitrite pollution in the groundwater resources a Sonqor plain in Iran. Data in Brief, 2018.

[5] Akbari, H., et al., Data on investigating the Nitrate concentration levels and quality of bottled water in Torbat-e Heydarieh, Khorasan razavi province, Iran. Data in Brief, 2018.

[6] Mirzabeygi, M., et al., Heavy metal contamination and health risk assessment in drinking water of Sistan and Baluchistan, Southeastern Iran. Human and Ecological Risk Assessment: An International Journal, 2017. 23(8): p. 1893-1905.

[7] A. Abbasnia, N.Yousefi , AH.Mahvi ,R.Nabizadeh ,M.Radfard, M.Yousefi & M.Alimohammadi , Evaluation of groundwater quality using water quality index and its suitability for assessing water for drinking and irrigation purposes; case study of Sistan and Baluchistan province (Iran) .Human and Ecological Risk Assessment: An International Journal (2018) (DOI:10.1080/10807039.2018.1458596)

[8] M.Radfard , M.Yunesian , R. Nabizadeh Nodehi ,H. Biglari , M. Hadi , N.Yosefi , M.Yousefi ,A. Abbasnia , AH. Mahvi . Drinking water quality and Arsenic health risk assessment in Sistan-and-Baluchestan, Southeastern province Iran. Human and Ecological Risk Assessment: An International Journal (2018) (DOI:10.1080/10807039.2018.1458210).

[9] Soleimani, H., et al., Data on drinking water quality using water quality index (WQI) and assessment of groundwater quality for irrigation purposes in Qorveh&Dehgolan, Kurdistan, Iran. Data in Brief, 2018.

[10] M. Mirzabeygi, N. Yousefi, A.Abbasnia, H. Youzi., M. Alikhani, AH. Mahvi, Evaluation of groundwater quality and assessment of scaling potential and corrosiveness of water supply networks, Iran, Journal of Water Supply: Research and Technology-Aqua. jws2(2017)

[11] A.Takdastana, M. Mirzabeygi (Radfard), M.Yousefi, A. Abbasnia, R. Khodadadia, A H.Mahvi, D.Jalili Naghan, Neuro-fuzzy inference system Prediction of stability indices and Sodium absorption ratio in Lordegan rural drinking water resources in west Iran, Data in Breif 18(2018)255-261.

[12] M .Mirzabeygi, M.Yousefi, H. Soleimani, AA. Mohammadi, AH.Mahvi, A. Abbasnia. The concentration data of fluoride and health risk assessment in drinking water in the Ardakan city of Yazd province, Iran. Data in Brief. 15(2018)40-46.

 [13] A. Abbasnia, M. Alimohammadi,AH. Mahvi, R. Nabizadeh,M. Yousefi,AA. Mohammadi ,H. Pasalari H ,M. Mirzabeigi, 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.

[14] Abbasnia A, Radfard M, Mahvi AH, Nabizadeh R, Yousefi M, Soleimani H, et al. Groundwater quality assessment for irrigation purposes based on irrigation water quality index and its zoning with GIS in the villages of Chabahar, Sistan and Baluchistan, Iran. Data in brief. 2018;19:623-31.

 [15] Abbasnia A, Ghoochani M, Yousefi N, Nazmara S, Radfard M, Soleimani H, et al. Prediction of human exposure and health risk assessment to trihalomethanes in indoor swimming pools and risk reduction strategy. Human and Ecological Risk Assessment: An International Journal. 2018:1-18 (DOI:10.1080/10807039.2018.1488582)

[16] Akazem Neisia, M.M.b., Ghader Zeydunib, Asghar Hamzezadehc, Davoud Jalilid, Abbas Abbasniab, Mahmood Yousefie, Rouhollah Khodadadif,, Data on fluoride concentration levels in cold and warm season in City area of Sistan and Baluchistan Province, Iran. Data in Brief, 2018.

[17] HR.Maier,GC. Dandy , The use of artificial neural networks for the prediction of water quality parameters, Water resources research, 32(1996) 1013-22.

[18] https://en.wikipedia.org/wiki/East_Azerbaijan_Province

[19] Gharibi, H., et al., A novel approach in water quality assessment based on fuzzy logic. Journal of Environmental Management, 2012. 112: p. 87-95.

[20] Sowlat, M.H., et al., A novel, fuzzy-based air quality index (FAQI) for air quality assessment. Atmospheric Environment, 2011. 45(12): p. 2050-2059.

*Conflict of Interest Form

[Click here to download Conflict of Interest Form: Conflict of Interest.docx](#)

*Conflict of Interest Form

[Click here to download Conflict of Interest Form: Conflict of Interest.docx](#)