JAEHR Journal of Advances in Environmental Health Research
Publishing House

Forecasting of heavy metals concentration in groundwater resources of Asadabad plain using artificial neural network approach

Meysam Alizamir[1], **Soheil Sobhanardakani**[2]

1 Department of Civil Engineering, Young Researchers and Elite Club, Hamedan Branch, Islamic Azad University, Hamedan, Iran
2 Department of the Environment, School of Basic Sciences, Hamedan Branch, Islamic Azad University, Hamedan, Iran

Original Article

Abstract

Nowadays 90% of the required water of Iran is secured with groundwater resources and forecasting of pollutants content in these resources is vital. Therefore, this research aimed to develop and employ the feedforward artificial neural network (ANN) to forecast the arsenic (As), lead (Pb), and zinc (Zn) concentration in groundwater resources of Asadabad plain. In this research, the ANN models were developed using MATLAB R2014 software program. The artificial intelligence models were trained with the data collected from field and then utilized as prediction tool. Levenberg-Marquardt (LM) and Bayesian regularization (BR) algorithms were employed as ANN training algorithms and their performance was evaluated using determination coefficient and the root mean square error. The results showed that the ANN models could potentially forecast heavy metals concentration in groundwater resources of the studied area. Coefficients of determination for ANN models for As, Pb and Zn in testing phase were 0.9288, 0.9823 and 0.8876, respectively. Finally, based on the simulation results, it was demonstrated that ANN could be applied effectively in forecasting the heavy metals concentration in groundwater resources of Asadabad plain.
KEYWORDS: Neural Networks, Heavy Metals, Groundwater, Forecasting, Risk

Date of submission: 10 Oct 2015, ***Date of acceptance:*** 12 Jan 2016

Citation: Alizamir M, Sobhanardakani S. **Forecasting of heavy metals concentration in groundwater resources of Asadabad plain using artificial neural network approach.** J Adv Environ Health Res 2016; 4(2): 68-77.

Introduction

Urbanization, industrialization, agriculture and exploitation of natural resources are basic activities associated with living in contemporary societies that have imposed pollutant loads especially toxic metals into natural cycles such as soil, water and air cycles.[1,2] Nowadays, throughout the world, heavy metals have been taken into consideration due to their ability to accumulate in the biota, toxicity and adverse health effects even at low concentrations.[3-5] In this regard, heavy metals pollution of the groundwater is one of the serious environmental problems. Some of the heavy metals considered as micronutrients can cause adverse health effects when their content exceed the permissible limit in drinking water.[6-8] Thus, assessment of groundwater resources pollutants especially heavy metals are very important with respect to human health.

Arsenic (As) is a widely distributed metalloid, occurring in soil, rock, water and air. The production of energy from fossil fuel and smelting of non-ferrous metals are the two significant industrial processes that lead to As contamination of environment, especially polluting the atmosphere. Other

Corresponding Author:
Soheil Sobhanardakani
Email: s_sobhan@iauh.ac.ir

sources of contamination are the use of arsenical pesticides, manufacture and wood preservatives.[9]

Zinc (Zn) on the other hand, is an essential functional and structural element in biological systems often catalyzing reactions by binding to substrates by favoring various reactions, such as the mediation of redox reactions or oxidation-reduction reactions, through reversible changes in the oxidation state of the metal ions. Of course, Zn harms several physiological processes like breathing.[10,11]

Some of the toxic metals, such as lead (Pb), may damage the kidney and cause symptoms of chronic toxicity, including poor reproductive capacity, impaired organ function, tumors, hypertension, and hepatic dysfunction.[12] Moreover, Pb can also affect brain function by interfering with synapse formation and neurotransmitter release. Exposure to Pb has been associated with reduced intelligence quotient (IQ), learning disabilities, hyperactivity, slow growth, impaired hearing and antisocial behaviors.[13] The most common routes of human Pb exposure are inadvertent ingestion of Pb paint, inhalation of traffic exhaust fumes and consumption of Pb-contaminated foods.[14,15]

In the last decade, artificial neural networks (ANNs), as massively parallel distributed processing systems that have similar performance to biological neural networks of human brain, has been widely used as a powerful prediction and management tool in solving environmental problems such as water resources modeling and management problems.[16,17] ANNs have proven to be an efficient approach for modeling qualitative water resource variables. Daliakopoulos et al.[18] employed ANNs in groundwater level forecasting. They used back-propagation (BP) and Levenberg-Marquardt (LM) algorithms for training of ANN models. Antar et al.[19] utilized ANN models with back-propagation (BP) algorithm in daily runoff estimation.

The ANNs is a useful computational method for predicting and modeling complex relationships among parameters, especially when there is no clear relation between them.[20] In this study, the structure of the ANNs consisted of three layers, i.e. the input, hidden and output layers. The developed computational intelligence model was applied for prediction of As, Pb and Zn concentration in groundwater resources of Asadabad plain, western Iran.

Since Iran is located within the dry and semi dry regions, about 90% of the required water is secured with groundwater resources.[21] Therefore forecasting pollutants content in these resources is vital. This research aimed to predict the heavy metals concentration (As, Pb and Zn) in groundwater resources of Asadabad plain using ANNs.

Materials and Methods

Asadabad plain with aquifer area about 962 km^2 and 1650 m above the sea level is located in southwest of Hamadan township in the west part of Iran. Agricultural practices are done in a large scale in this area making discharge of chemical pollutants to the groundwater resources very likely. Since the most required water of study area is secured with groundwater resources,[22] the implementation of this study seems necessary.

Groundwater samples were collected from 30 different locations including open and tube wells in 3 replicates to evaluate the heavy metal contamination in 2012. Figure 1 shows the sampling stations in the study area. The sampling locations were selected based on different land use patterns, including agricultural and residential areas. The samples were taken in pre-cleaned, acid-soaked 200 ml polyethylene bottles to avoid unpredictable changes in characteristic as per standard procedures.[23,24] The collected samples were filtered (Whatman no. 42), preserved with 6N HNO_3 (Suprapur Merck, Germany) and kept at 4 °C for further analysis.[23,25] Concentrations of heavy metals [As, Zn, Pb, cadmium (Cd) and copper (Cu)]

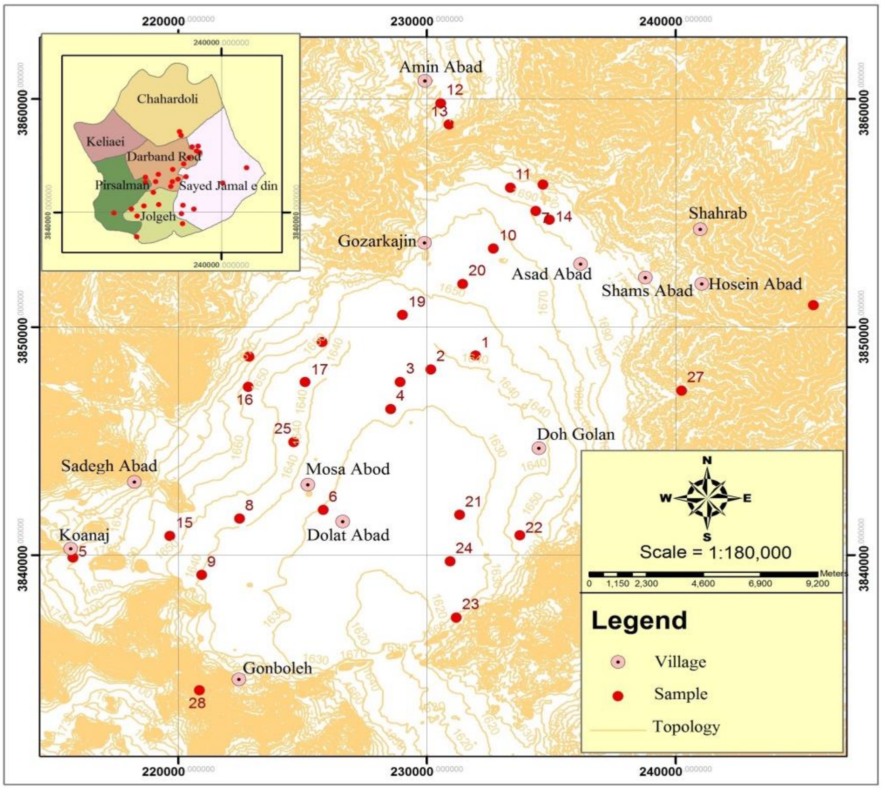

Figure 1. Map of sampling stations

in samples were determined using inductively coupled plasma-optical emission spectrometry (ICP-OES) (Varian, 710-ES, Australia). Table 1 gives statistical properties of data used in this study.

Artificial neural networks can indicate nonlinear and complex relationships using a part of input-output training patterns from the data. In fact ANNs provide a non-linear mapping between inputs and outputs by its intrinsic ability.[26] ANNs characterization is based on architecture that demonstrates the connection pattern between nodes, connection weights determination method and the activation function.[27,28] Due to the ability of ANNs in the learning of system's behavior from representative data, it would be possible for them to solve large-scale complex problems.[16,29] The most commonly used ANN is the feedforward neural networks. These networks have been successfully applied in different studies in environmental problems.[20,30-34] Architecture of these networks have input layer, hidden layer and output layer. A typical multi layered feedforward neural network is included of multiple elements called nodes and connection pathways that link them.[16]

In the developed ANN models, number of hidden layers, number of neurons and choice of training and stopped training algorithms were carefully chosen for evading the overtraining problem. In this study the optimal number of layers and neurons in the hidden layer were identified using trial and error procedure. Training process is employed for generating an output vector from ANNs which is as close as possible to target vector.[16]

Table 1. Statistical properties of dataset

Heavy metal	Min.	Max.	Mean	Standard deviation
As	21.27	100.24	55.2	16.5
Pb	2.50	51.52	13.5	7.4
Zn	0.97	267.35	27.3	33.7

As: Arsenic; Pb: Lead; Zn: Zink

Training of a network is based on the optimization process for weights in the nodes links and bias terms;[16] this process continues until the values of the output layer are as close as possible to the actual outputs. As far as there are different types of training algorithms, the selection of an algorithm that provides the best fit to the data is required. Therefore, in this study the Levenberg-Marquardt (LM) training algorithm [35,36] was used. The sigmoid and linear activation functions were utilized for the hidden and output node, respectively. Figure 2 shows the schematic feed-forward network for this study, having one hidden layer (HL) with several nodes between the input and output layers. The code of ANN modelling was written using MATLAB software.

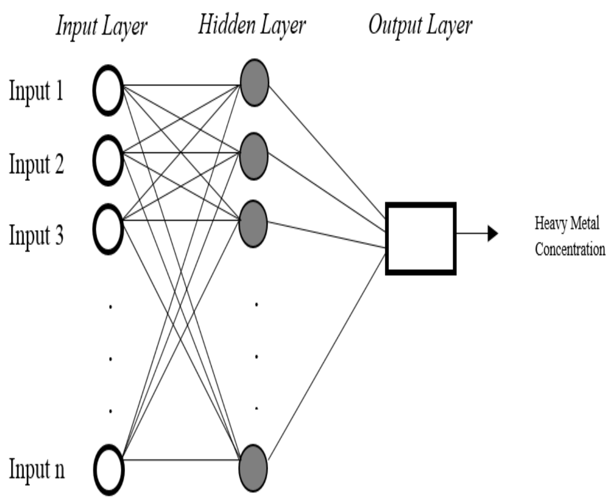

Figure 2. A schematic architecture of the artificial neural network (ANN) model used in this study

An ANN model is developed in this study to predict groundwater heavy metals concentration in Asadabad plain for a period using the observed time series of different heavy metals concentration. Different numbers of hidden layer neurons were tried for ANN models in the study. Simple trial and error showed that the number of hidden layer neurons that gave the minimum root mean square error (RMSE) was 5 for different input combinations. The neural network is first trained to perform predictions of the groundwater heavy metals concentration using previously observed data. In this study for training and testing of ANN models, collected field data have been used. The input layers were groundwater heavy metals concentration in the past and the output layer was current groundwater heavy metals concentration. In the ANN modeling, the selection of input parameters is a very essential step. Different statistical methods were suggested for appropriate input vectors for a model.[37,38] Different previous lags were considered as input candidates to the model in this study. The inputs denotes the previous groundwater heavy metals concentration in the Asadabad plain (t, t-1, t-2), and the output layer corresponds to the groundwater heavy metals concentration at time *t*+1. In the current study, a partial autocorrelation function (PACF) was used to find the best input combinations.

The efficiencies of the models developed in this study were surveyed using two standard statistical performance evaluation criteria. The statistical criteria considered were RMSE, and determination coefficient (R^2).

1. The correlation between predicted and actual values (R^2) that computed according to equation 1:

$$R^2 = \frac{\left[\sum_{i=1}^{n}(O_i-\overline{O_i})\times(P_i-\overline{P_i})\right]^2}{\sum_{i=1}^{n}(O_i-\overline{O_i})\times\sum_{i=1}^{n}(P_i-\overline{P_i})} \quad (1)$$

where the predicted and observed values are P_i and O_i respectively, and the total number of test data is given by n.

2. RMSE was obtained through the equation 2:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i-O_i)^2}{n}} \quad (2)$$

The smaller the value of this statistic, the better the function of the model.

Table 2. Comparative performance of artificial Neural Network (ANN) for different heavy metal concentration

Metal Concentration	Training		Testing	
	RMSE	R^2	RMSE	R^2
As	3.3173	0.9585	4.0303	0.9288
Pb	0.1698	0.9995	0.7638	0.9823
Zn	5.0885	0.9805	7.2230	0.8876

RMSE: Root mean square error; As: Arsenic; Pb: Lead; Zn: Zink

Results and Discussion

The performance of ANN models with Levenberg-Marquardt algorithm were compared in predicting three heavy metals concentration in groundwater resources of Asadabad plain. The output of the model was a prediction of groundwater heavy metal concentration one day in advance (t + 1). The best architecture for the network was selected by trial and error. The results of this study have shown that ANN was effective for predicting groundwater heavy metals concentration. The performance of this model in terms of determination coefficient and root mean square error statics during the training and testing period is shown in table 2. It is apparent from this table that the model performance was good, and the models have forecasted metals concentration with reasonable accuracy.

This paper aimed to predict the groundwater heavy metals concentration in the Asadabad plain using ANN models. In the first step, field data were analyzed and various input combinations were tried and the best input combination for each heavy metal concentration was selected afterwards. In the second step the data was divided for training and testing phases and finally, the groundwater heavy metals concentration was computed by the optimal ANN model. R^2 and RMSE statics of the optimal model for As, Pb and Zn are given in table 2 for the training and testing periods. Generally, lowest value of root mean square error and highest value of coefficient of determination is desirable for choosing best model. Training data proportion is a very important factor for efficiency of an ANN model; low proportion of the training data can lead to insufficient neural network training, therefore presenting a sufficient portion of data can improve the accuracy of the model in the training and testing periods.

The predicted metal concentration plots during training and testing period are shown in figures 3 to 5.

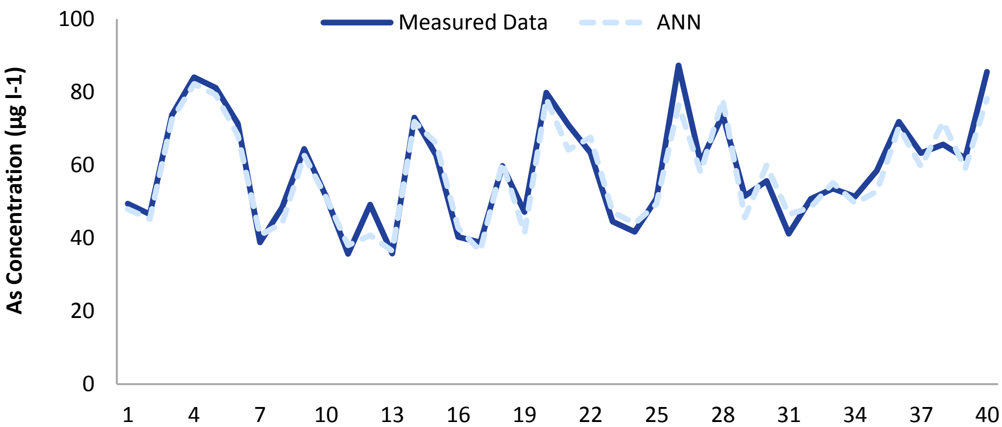

Figure 3. Comparison of the actual groundwater heavy metal concentration values and those predicted for As in testing phase
ANN: Artificial neural network

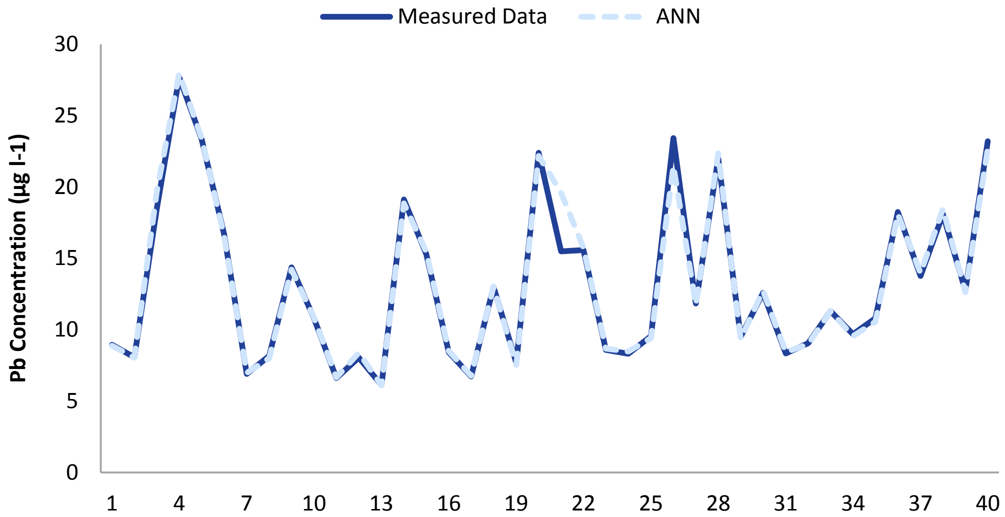

Figure 4. Comparison of the actual groundwater heavy metal concentration values and those predicted for Pb in testing phase
ANN: Artificial neural network

In general, the results indicated the potential of ANN technique in predicting metals concentration at Asadabad plain. Also extent of the match between the field data and predicted heavy metals concentration of ANN models are shown in figures 6 to 8 for As, Pb and Zn, respectively. As displayed in figures 6 to 8, during the training phase, the intelligent model estimated metals concentration better than the testing phase and in all the cases, R^2 was higher than 0.85 which indicated that the metals concentration values estimated by the ANN forecasting models were closely matched with the analyzed values. On the other hand, the developed ANN models had acceptable accuracy and precision in predicting of heavy metal concentration in Asadabad plain groundwater resources. RMSE and R^2 values in training and testing periods are reported in table 2.

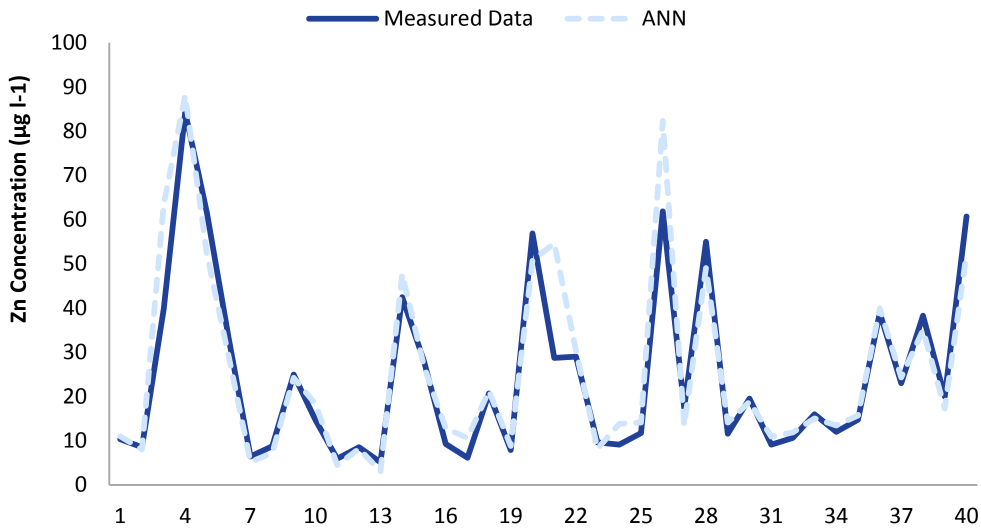

Figure 5. Comparison of the actual groundwater heavy metal concentration values and those predicted for Zn in testing phase
ANN: Artificial neural network

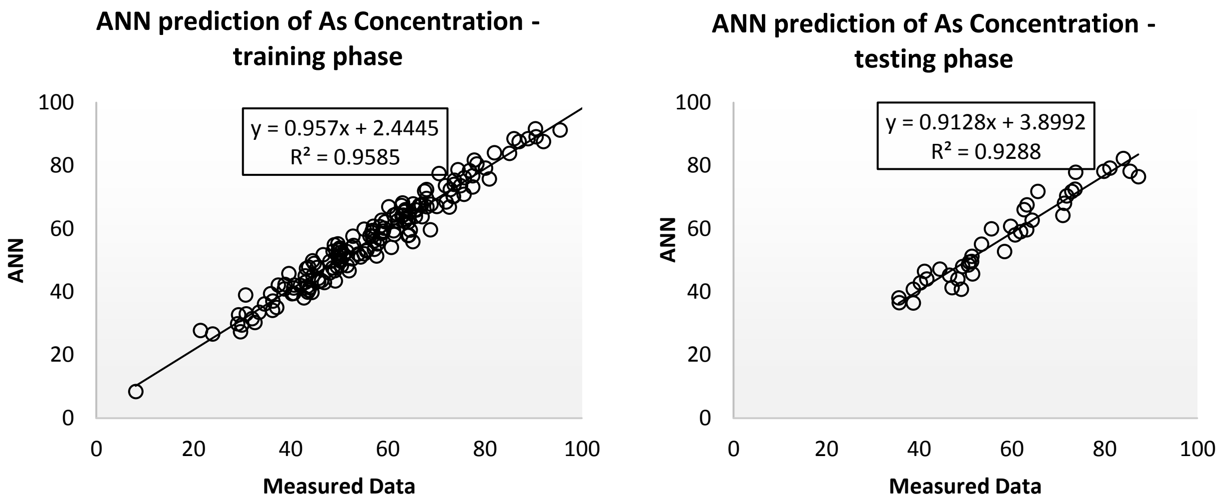

Figure 6. Scattered plot of predicted with actual values for As in training (left) and testing (right) phases
ANN: Artificial neural network

The obtained results of the proposed ANN model for the prediction of As concentration are: R^2 = 0.9585 and RMSE = 3.3173 µg l^{-1} in training period and R^2 = 0.9288 and RMSE = 4.0303 µg l^{-1} in testing period. It is clear from the obtained results that the developed ANN models for As concentration yielded acceptable performance. Also, the results of model for the prediction of Pb concentration are: R^2 = 0.9995 and RMSE = 0.1698 µg l^{-1} in training period and R^2 = 0.9823 and RMSE = 0.7638 µg l^{-1} in testing period. Results of the model for Zn concentration revealed that the model yielded good accuracy, in this case: R^2 = 0.9805 and RMSE = 5.0885 µg l^{-1} in training period and R^2 = 0.8876 and RMSE = 7.2230 µg l^{-1} in testing period. The results of this study indicated that environmental managers can improve their designs and evaluations via ANN for forecasting heavy metals concentration in Asadabad plain. Compared to the results of this study, Keskin et al. developed ANN model by 5 neurons in hidden layer for predicting water pollution sources in different areas in Turkey and reported that that ANN can be used as powerful analytical tool in water quality modeling.[39]

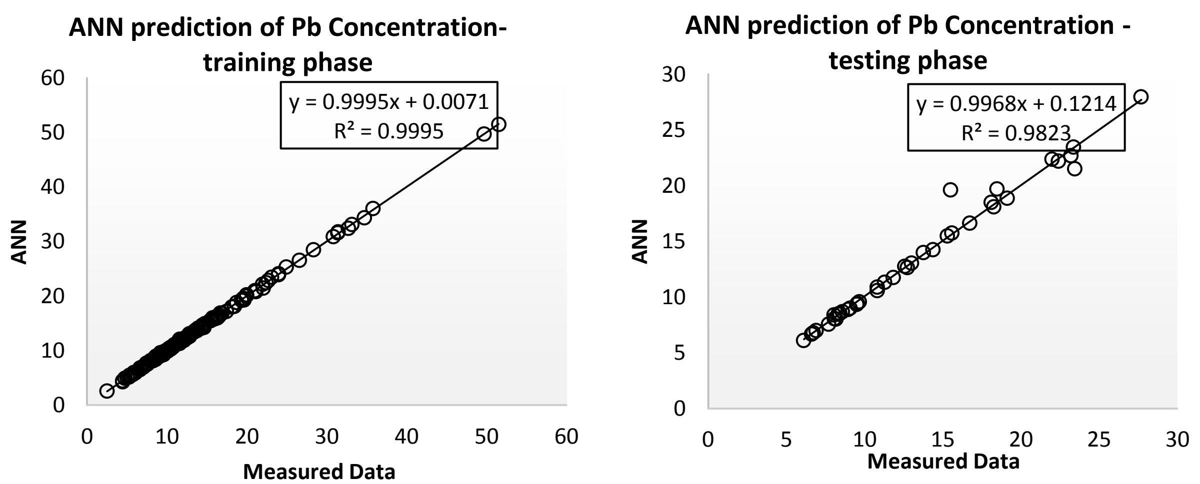

Figure 7. Scattered plot of predicted with actual values for Pb in training (left) and testing (right) phases
ANN: Artificial neural network

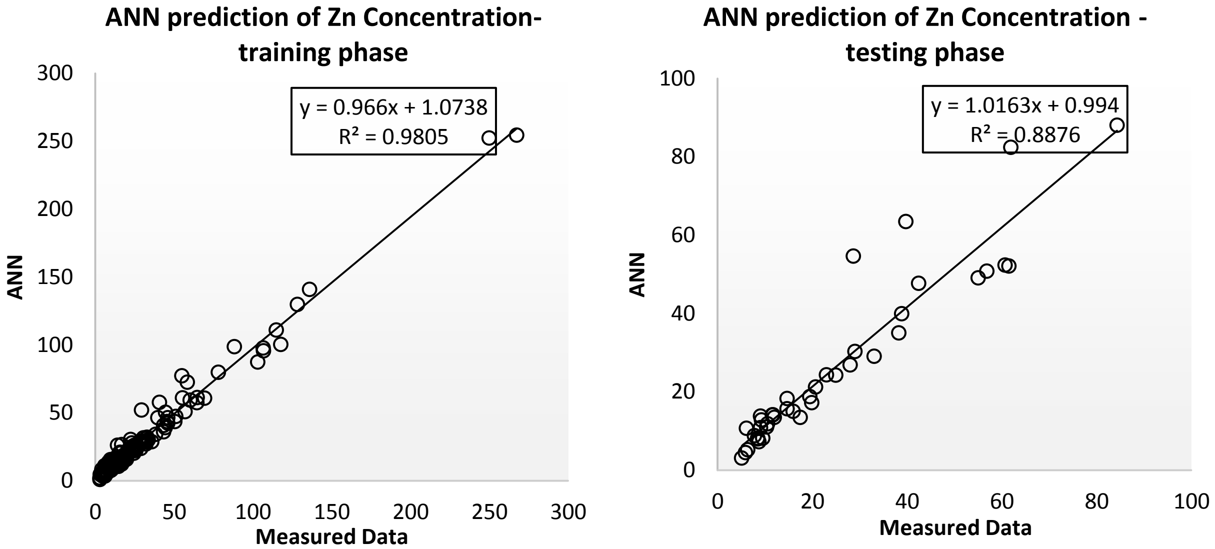

Figure 8. Scattered plot of predicted with actual values for Zn in training (left) and testing (right) phases
ANN: Artificial neural network

Also, Nor et al. used a three-layer multi-layer perceptron (MLP) by 25 hidden neurons to classify different combinations and groups of water contamination by nitrate and sulfate and reported that ANN can yield reliable results.[40]

Conclusion

The importance of water quality on human health is well known and in recent years have attracted a great deal of interest. The use of groundwater for the water supply needs of many urban and rural communities especially in arid and semi-arid climate have increased in the last decade and, in Iran, groundwater continues to play an important role in the socio-economic development of the country. With regard to the importance of prediction of groundwater qualitative situation in the Asadabad plain, ANN model has been utilized to predict groundwater As, Pb and Zn concentration. The best input combinations were selected by correlation analysis from previous values of heavy metals concentration for the ANN models to estimate groundwater heavy metals concentration in the Asadabad plain. The different algorithms used for training the models were Levenberg-Marquardt algorithm and Bayesian regularization. The performance of these algorithms was evaluated and it was found that Levenberg-Marquardt was a good choice for forecasting heavy metals concentration in the selected monitoring wells. The results indicated that the Levenberg-Marquardt algorithm took less time for training of the network compared to Bayesian regularization algorithm. Also, Levenberg-Marquardt algorithm gave the best metals concentration forecasts compared to Bayesian regularization algorithm. The performance evaluation criteria (RMSE) and determination coefficient were very good and consistent in the all metals concentration predictions.

The optimum ANN model proposed in the current research showed very promising results for improving environmental management planning in Asadabad plain and it can be effectively used in the future studies in this region.

Conflict of Interests

Authors have no conflict of interests.

Acknowledgements

Authors are grateful to the Hamadan Branch, Islamic Azad University for providing

facilities to conduct and complete this study.

References

1. Gribble GW. The natural production of chlorinated compounds. Environ Sci Technol 1994; 28(7): 310A-9A.
2. Nasrabadi T. An IndexApproach to Metallic Pollution in RiverWaters. Int J Environ Res 2015; 9(1): 385-94.
3. Morillo J, Usero J, Gracia I. Partitioning of metals in sediments from the Odiel River (Spain). Environ Int 2002; 28(4): 263-71.
4. Baghvand A, Nasrabadi T, Bidhendi GN, Vosoogh A, Karbassi A, Mehrdadi N. Groundwater quality degradation of an aquifer in Iran central desert. Desalination 2010; 260(̃13): 264-75.
5. Hossein Pour M, Lashkaripour G, Dehghan P. Assessing the effect of heavy metal concentrations (Fe, Pb, Zn, Ni, Cd, As, Cu,Cr) on the quality of adjacent groundwater resources of Khorasan Steel Complex). Int J Pl An and Env Sci 2014; 4(2): 511-8.
6. Prasanna MY, Chidambaram S, Hameed AS, Srinivasamoorthy K. Hydrogeochemical analysis and evaluation of groundwater quality in the Gadilam river basin, Tamil Nadu, India. J Earth Syst Sci 2011; 120(1): 85-98.
7. Prasad B, Kumari P, Bano S, Kumari S. Ground water quality evaluation near mining area and development of heavy metal pollution index. Appl Water Sci 2014; 4(1): 11-7.
8. Sobhanardakani S. Evaluation of the water quality pollution indices for groundwater resources of Ghahavand plain, Hamadan province, western Iran. Iran J Toxicol 2016; 10(3): 35-40.
9. Jarup L. Hazards of heavy metal contamination. Br Med Bull 2003; 68: 167-82.
10. Tahsin N, Yankov B. Research on accumulation of zinc (Zn) and cadmium (Cd) in sunflower oil. Journal of Tekirdag Agricultural Faculty 2007; 4(1): 109-11.
11. Sobhanardakani S, Jamshidi K. Assessment of Metals (Co, Ni, and Zn) Content in the Sediments of Mighan Wetland Using Geo-Accumulation Index. Iran J Toxicol 2015; 9(30): 1386-90.
12. Abou-Arab AAK, Ayesh AM, Amra HA, Naguib K. Characteristic levels of some pesticides and heavy metals in imported fish. Food Chem 1996; 57(4): 487-92.
13. Dahiya S, Karpe R, Hegde AG, Sharma RM. Lead, cadmium and nickel in chocolates and candies from suburban areas of Mumbai, India. J Food Comp Anal 2005; 18(6): 517-22.
14. Hosseini SV, Sobhanardakani S, Tahergorabi R, Delfieh P. Selected heavy metals analysis of Persian sturgeon's (Acipenser persicus) caviar from Southern Caspian Sea. Biol Trace Elem Res 2013; 154(3): 357-62.
15. Adekunle IM, Akinyemi MF. Lead levels of certain consumer products in Nigeria: a case study of smoked fish foods from Abeokuta. Food Chem Toxicol 2004; 42(9): 1463-8.
16. Haykin SS. Neural Networks: A Comprehensive Foundation. 2nd ed. Upper Saddle River, NJ: Prentice Hall; 1999.
17. Nikolos IK, Stergiadi M, Papadopoulou MP, Karatzas GP. Artificial neural networks as an alternative approach to groundwater numerical modelling and environmental design. Hydrol Process 2008; 22(17): 3337-48.
18. Daliakopoulos IN, Coulibaly P, Tsanis IK. Groundwater level forecasting using artificial neural networks. J Hydrol 2005; 309(̃14): 229-40.
19. Antar MA, Elassiouti I, Allam MN. Rainfall-runoff modelling using artificial neural networks technique: a Blue Nile catchment case study. Hydrolog Process 2006; 20(5): 1201-16.
20. Shakeri Abdolmaleki A, Gholamalizadeh Ahangar A, Soltani J. Artificial neural network (ann) approach for predicting cu concentration in drinking water of chahnimeh1 reservoir in Sistan-Balochistan, Iran. Health Scope 2013; 2(1): 31-8.
21. Sobhanardakani S, Jamali M, Maanijou M. Evaluation of as, Zn, Cr and Mn concentrations in groundwater resources of razan plain and preparation of zoning map using gis. J Environ Sci Technol 2014; 16(2): 25-38.
22. Fazel Tavasol S, Vusuq BP, Manshuri M. The investigation of heavy metal (Sn-Pb) concentration in ground water resources and their environmental effects, Case study: North Chardoly Plain. Proceedings of the 1st International Applied Geological Congress; 2010 Apr 26-28; Mashhad, Iran.
23. Clesceri LS, Greenberg AE, Eaton AD. Standard methods for the examination of water and waste water, quality assurance. 20th ed. Washington DC: American Public Health Association; 1999.
24. Eaton AD, Franson MA. Standard methods for the examination of water & wastewater. Washington, DC: American Public Health Association; 2005.
25. Edet A, Offiong O. Evaluation of water quality pollution indices for heavy metal contamination monitoring. A study case from Akpabuyo-Odukpani area, Lower Cross River Basin (southeastern Nigeria). Geo J 2002; 57(4): 295-304.
26. Hornik K, Stinchcombe M, White H. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. Neural Net 1990; 3(5): 551-60.
27. Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. Neural Net 1989; 2(5): 359-66.
28. Aziz AR, Vincent Wong KF. A Neural-Network

Approach to the Determination of Aquifer Parameters. Ground Water 1992; 30(2): 164-66.
29. ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. Artificial Neural Networks in Hydrology. I: Preliminary Concepts. J Hydrol Eng 2000; 5(2): 115-23.
30. Amin S, Farjoud Mr, Shabani A. Groundwater contamination by heavy metals in water resources of Shiraz area. Iran Agric Res 2011; 30(1-2): 21-32.
31. Taghipour H, Mosaferi M, Pourakbar M, Armanfar F. Heavy metals concentrations in groundwater used for irrigation. Health Promot Perspect 2012; 2(2): 205-10.
32. Ghosh A, Das P, Sinha K. Modeling of biosorption of Cu(II) by alkali-modified spent tea leaves using response surface methodology (RSM) and artificial neural network (ANN). Appl Water Sci (2015) 5: 191 2015; 5(2): 191-9.
33. Gnanasangeetha D, SaralaThambavani D. Modelling of As3+ adsorption from aqueous solution using Azadirachta indica by artificial neural network. Desalin Water Treat 2015; 56(7): 1839-54.
34. Yurtsever U, Yurtsever M, Sengil A, Yilmazçoban NK. Fast artificial neural network (FANN) modeling of Cd(II) ions removal by valonia resin. Desalin Water Treat 2014; 56(1): 83-96.
35. Marquardt DW. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. J Soc Ind Appl Math 1963; 11(2): 431-41.
36. Levenberg K. A method for the solution of certain non-linear problems in least squares. Quart Appl Math 1944; 2(2): 164-8.
37. Sudheer KP, Gosain AK, Ramasastri KS. A data-driven algorithm for constructing artificial neural network rainfall-runoff models. Hydrol Process 2002; 16(6): 1325-30.
38. Lin J, CHENG CT, Chau KW. Using Support Vector Machines for Long-Term Discharge Prediction. Hydrol Sci J 2006; 51(4): 599-612.
39. Keskin TE, Dügenci M, Kaçaroglu F. Prediction of water pollution sources using artificial neural networks in the study areas of Sivas, Karabük and Bartin (Turkey). Environ Earth Sci 2015; 73(9): 5333-47.
40. Nor AS, Faramarzi M, Yunus MA, Ibrahim S. Nitrate and sulfate estimations in water sources using a planar electromagnetic sensor array and artificial neural network method. IEEE Sens J 2015; 15(1): 497-504.