Journal of Chemical Health Risks

www.jchr.org

ORIGINAL ARTICLE

Modeling of Groundwater Resources Heavy Metals Concentration Using Soft Computing Methods: Application of Different Types of Artificial Neural Networks

Meysam Alizamir[1], Soheil Sobhanardakani[*2], Lobat Taghavi[3]

[1]Young Researchers & Elite Club, Hamedan Branch, Islamic Azad University, Hamedan, Iran

[2]Department of the Environment, College of Basic Sciences, Hamedan Branch, Islamic Azad University, Hamedan, Iran

[3]Department of the Environmental Pollution, College of Environment and Energy, Science and Research Branch, Islamic Azad University, Tehran, Iran

(Received: 7 December 2016 Accepted: 6 February 2017)

KEYWORDS

Heavy metals concentration;
Multi-layer perceptron;
Radial basis function;
Groundwater resources;
Asadabad Plain

ABSTRACT: Nowadays, groundwater resources play a vital role as a source of drinking water in arid and semiarid regions and forecasting of pollutants content in these resources is very important. Therefore, this study aimed to compare two soft computing methods for modeling Cd, Pb and Zn concentration in groundwater resources of Asadabad Plain, Western Iran. The relative accuracy of several soft computing models, namely multi-layer perceptron (MLP) and radial basis function (RBF) for forecasting of heavy metals concentration have been investigated. In addition, Levenberg-Marquardt, gradient descent and conjugate gradient training algorithms were utilized for the MLP models. The ANN models for this study were developed using MATLAB R 2014 Software program. The MLP performs better than the other models for heavy metals concentration estimation. The simulation results revealed that MLP model was able to model heavy metals concentration in groundwater resources favorably. It generally is effectively utilized in environmental applications and in the water quality estimations. In addition, out of three algorithms, Levenberg-Marquardt was better than the others were. This study proposed soft computing modeling techniques for the prediction and estimation of heavy metals concentration in groundwater resources of Asadabad Plain. Based on collected data from the plain, MLP and RBF models were developed for each heavy metal. MLP can be utilized effectively in applications of prediction of heavy metals

* Corresponding author: s_sobhan@iauh.ac.ir (S. Sobhanardakani).

concentration in groundwater resources of Asadabad Plain.

INTRODUCTION

Water is one of the most important agents for every living organism on earth. The 3% of global fresh water is large enough to meet the requirements of man for millions of years [1]. Various sources especially industrialization and urbanization can cause discharge of kinds of pollutants such as heavy metals, pesticides, POPs, etc. into the water resources worldwide [1-3].

Heavy metals are one of the most poisonous and serious groups of pollutants due to their high toxicity, abundance, and ease of accumulation from human and other various species. “The behavior of these elements in the environment depends on some characteristics especially their inherent chemical properties” [1].

Cadmium is very toxic metal with a natural occurrence in soil, but it is also spread in the environment due to anthropogenic origins [4]. It can finally enter the human body via the food chains [5]. Cadmium is present in virtually all food and ingestion of foods is the primary way for human's exposure to this metal [6, 7]. Long-term exposure to lower contents of Cd leads to a build up in the kidneys and possible kidney disease, lung damage, and frangible bones. Hypertension, arthritis, diabetes, anemia, cardiovascular disease, cirrhosis, hypoglycemia, headaches, osteoporosis, strokes and cancer are its some odd long-term results. Cadmium also affects the female reproduction system and severely affects the feminine endocrine system [5, 8-13].

Lead is a widespread pollutant and highly toxic metal, is bioaccumulative and nor easily metabolized, therefore, does not degrade in the environment. Lead usually can enter to the environment via human activities such as mining and industrial activities, fuel combustion, etc. and due to the Pb. has no known functions in biological systems, recognized as a major environmental health risk throughout the world and cause adverse effects to human health [14]. The main source of human exposure to Pb is food, believed to provide about 80%-90% of daily doses. Although, Pb affects humans and animals of all ages, but the effects of this element especially organic compound are most serious in young children [8]. The excessive intake of Pb can damage the central nervous system, skeletal, enzymatic, endocrine, circulatory, and immune systems, disturb hemoglobin synthesis, kidneys and blood system in adults and delays in mental development and physical in children [15-17].

Zinc is known as an essential functional and structural element in biological systems. It is one of the important metals for normal growth and development in human beings, and often has an important role in catalyzing reactions such as the redox. Of course, zinc harms some physiological processes like breathing [18, 19].

Iran is located within the arid and semi-arid areas. Therefore, groundwater resources are vital role in supply of the required water of this country [20] and forecasting of pollutants content in these resources is very important. The purpose of this study was to investigate the performance of two types of ANN models for prediction of heavy metals (Cd, Pb and Zn) concentration in groundwater resources of Asadabad Plain.

MATERIALS AND METHODS

Study area

Asadabad Plain with an aquifer area about 962 km^2 is located in Southwest of Hamedan Township in the west part of Iran at an altitude of 1650 m above the sea level [21].

Statistical properties of data used in this study are presented in Table 1.

Table 1. Statistical properties of dataset.

Heavy metal	Min.	Max.	Mean	Standard deviation
Cd	2.01	10.05	4.55	**1.22**
Pb	2.5	51.52	13.5	**7.39**
Zn	0.97	267.35	27.32	**33.7**

MLP model

Artificial neural networks (ANNs) are inspired from nervous system of living organisms; these networks consist of layers parallel processing elements, called neurons Haykin. In recent years, multi-layer feed forward networks have been applied extensively to solve various engineering problems [22, 23]. A three- layer ANN model using backpropagation (BP) algorithm proposed to predict the removal efficiency of As (III) and As(V) of *Botryococcus braunii* used as phycoremediator materials [24]. A back propagation artificial neural network developed based on experimental results for prediction of As (III) removal efficiency from water. Based on the results of this study, ANN can predict removal process of As(III) efficiently [25]. An ANN model in estimating nitrate and sulfate contamination employed in water sources. The results of this study showed good capability of the developed model to relate inputs and outputs [26].

Neurons as the elements of ANN models are organized in layers [27]. The neurons in a layer receive input parameters from the previous layer and give their output to the next layer [28]. The ANNs employed in the current research consist of three layers: the input, one hidden and output. The MLP was trained using different algorithms, Levenberg-Marquardt, gradient descent and conjugate gradient.

Figure 1 shows the A three-layer MLP neural network for this study, having one hidden layer with several nodes between the input and output layers. The code of ANN modeling was written using MATLAB software. In ANN modeling, choosing the hidden nodes number is an important task. Here the ANN with one hidden layer was utilized and hidden nodes number was determined using trial and error method. The sigmoid and linear activation functions were used for the hidden and output nodes, respectively (Table 2).

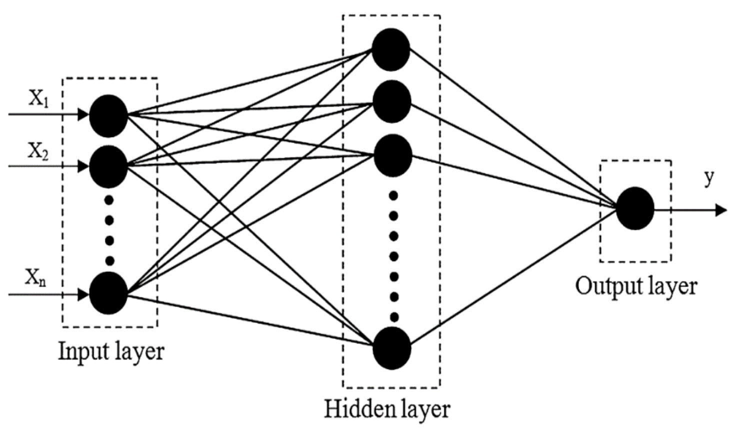

Figure 1. A three-layer MLP neural network model

Table 2. Chosen MLP model functions.

Training Algorithms	**1) Levenberg-Marquardt** **2) Gradient descent** **3) Conjugate gradient**
Activation function (hidden layer)	Sigmoid
Activation function (output layer)	Linear
Performance functions	R^2 and Root mean square error

Radial Basis Function

Radial basis function (RBF) networks are non-linear hybrid networks that they always have a single hidden layer and this layer uses Gaussian transfer functions [29]. Figure 2 shows a typical RBF network by a hidden layer with four nodes, four inputs and four outputs. A more detailed description of RBF models can be found in Haykin [30].

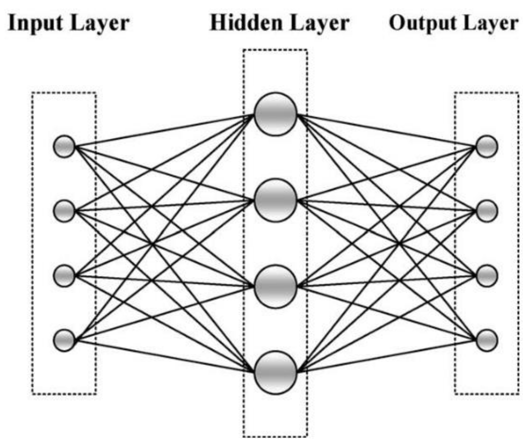

Figure 2. Structure of a typical RBF neural network model [29]

Model Development

The collected field data served to develop the two soft computing models (MLP and RBF) to determine the three heavy metals concentration (Cd, Pb and Zn) in groundwater resources of Asadabad Plain. These soft computing models were developed and were coded using MATLAB software. The collected data was divided into training and testing parts (80% and 20%, respectively). The input layers were groundwater heavy metals concentration in past time and the output layer was in the current time. In the ANN modeling, the selection of input parameters is an essential step. Different statistical methods were suggested for appropriate input vectors for a model [31, 32]. Different previous lags were con sidered as input candidates to the model in this study.

The inputs denote the previous groundwater heavy metals concentration in the Asadabad Plain (t, t-1, t-2), and the output layer corresponds at time t+1. In the current study, a partial autocorrelation function (PACF) was employed to find the best-input combinations.

Models Performance Evaluation

The performance of artificial intelligence methods in training and testing periods is evaluated via two common statistical indicators such as determination coefficient (R^2) and root mean square error (RMSE), which are expressed as follows:

$$R^2 = \frac{\sum_{i=1}^{n}(O_i^{Predicted} - \overline{O}_i^{Predicted})(O_i^{Observed} - \overline{O}_i^{Observed})}{\sqrt{\sum_{i=1}^{n}(O_i^{Predicted} - \overline{O}_i^{Predicted})^2 (O_i^{Observed} - \overline{O}_i^{Observed})^2}} \quad (1$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(O_i^{Predicted} - \overline{O}_i^{Predicted})^2}{n}} \quad (2$$

Where $O_i^{Observed}$ and $O_i^{Predicted}$ respectively account for the observed and predicted values at time *i*, while the terms $\overline{O}_i^{Observed}$ and $\overline{O}_i^{Predicted}$ imply the mean of the observed and predicted values; and *n* shows the number of data points.

RESULTS AND DISCUSSION

During the past decade, artificial intelligence techniques (AIT) have been successfully applied in different fields of environmental problems. In this study, the same training and testing data set employed for the development of MLP and RBF models. For the MLP model, three different training of the algorithms were used to train the network. During the development of ANN models, different numbers of hidden layer neurons can be used. By applying Hecht-Nielsen method [33] via a trial-and-error procedure, the optimum number of nodes in the hidden layer is found. For all heavy metals concentration, all the ANN models were first trained using the data in the training sets (using the first 80% of the data) to obtain the optimized set learning coefficients, and then tested (using the 20% of the data). The ANN models were then compared using statistical indicators of determination coefficient (R^2) and root mean square error (RMSE). Table 3 shows the ANN models perfor mance statistics (RMSE and R^2) for Cd, Pb and Zn concentration. For all heavy metals, the MLP model performed better than the RBF model in testing and training phases. Figures 3 to 8 illustrate the results with the performance indices between observed and predicted data for the training and testing data sets.

The results in Figure 3 show good performance of MLP in Cd concentration prediction in both training and testing periods (R^2=0.9721 and RMSE=0.2109 µg/l in training phase and R^2=0.9465 and RMSE=0.2601 µg/l in testing phase). The MLP model yields better performance in comparison with RBF model in prediction of Pb concentration (R^2=0.9945 and RMSE=0.5739 µg/l in training phase and R^2=0.9821 and RMSE=0.7627 µg/l in testing phase) (Figure 5). Figure 7 demonstrates priority of MLP model in the prediction of Zn concentration (R^2=0.9931 and RMSE=3.0040 µg/l in training phase and R^2=0.9373 and RMSE=5.0030 µg/l in testing phase).

Table 3. Comparative performance of ANNs for Cd, Pb and Zn concentration.

Heavy Metal Concentration	Methods	Training		Testing	
		RMSE	R^2	RMSE	R^2
Cd	MLP	0.2109	0.9721	0.2601	0.9465
	RBF	0.2135	0.9714	0.2904	0.9402
Pb	MLP	0.5739	0.9945	0.7627	0.9821
	RBF	0.5877	0.9942	0.8237	0.979
Zn	MLP	3.0040	0.9931	5.0030	0.9373
	RBF	6.3529	0.9693	7.6509	0.8742

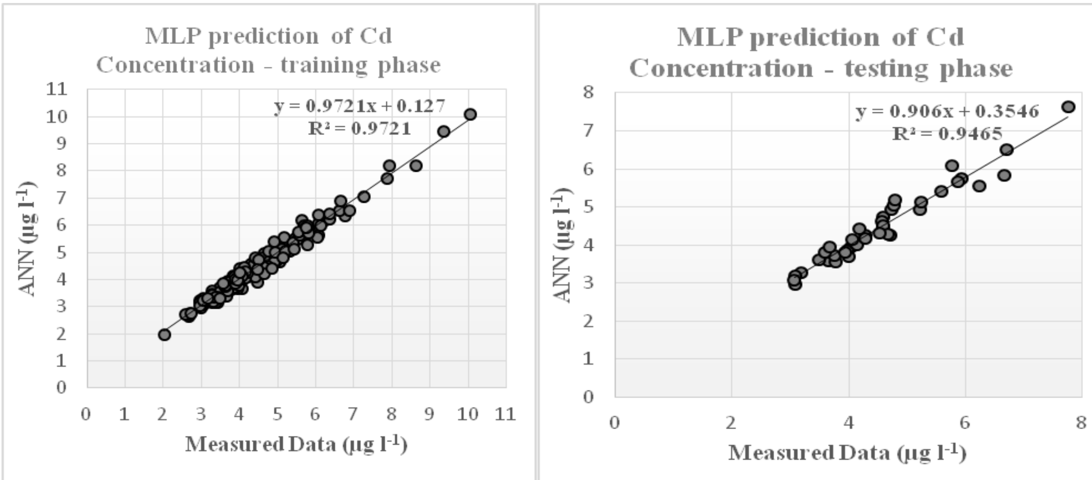

Figure 3. Observed and simulated Cd concentration by MLP model during the training and testing phases.

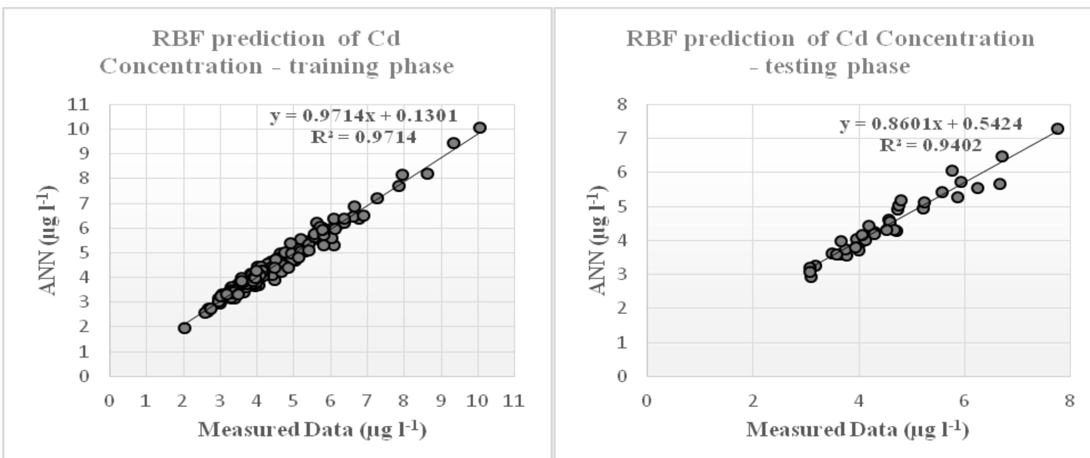

Figure 4. Observed and simulated Cd concentration by RBF model during the training and testing phases.

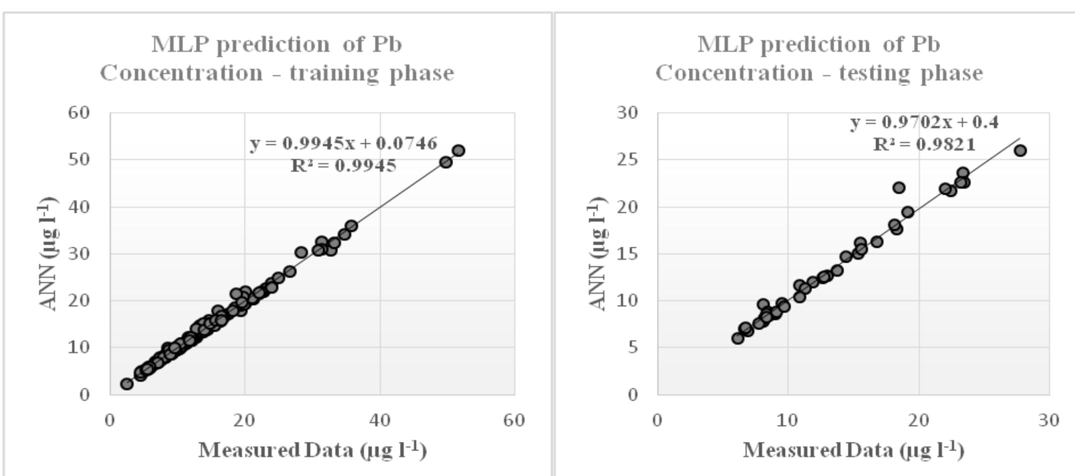

Figure 5. Observed and simulated Pb concentration by MLP model during the training and testing phases.

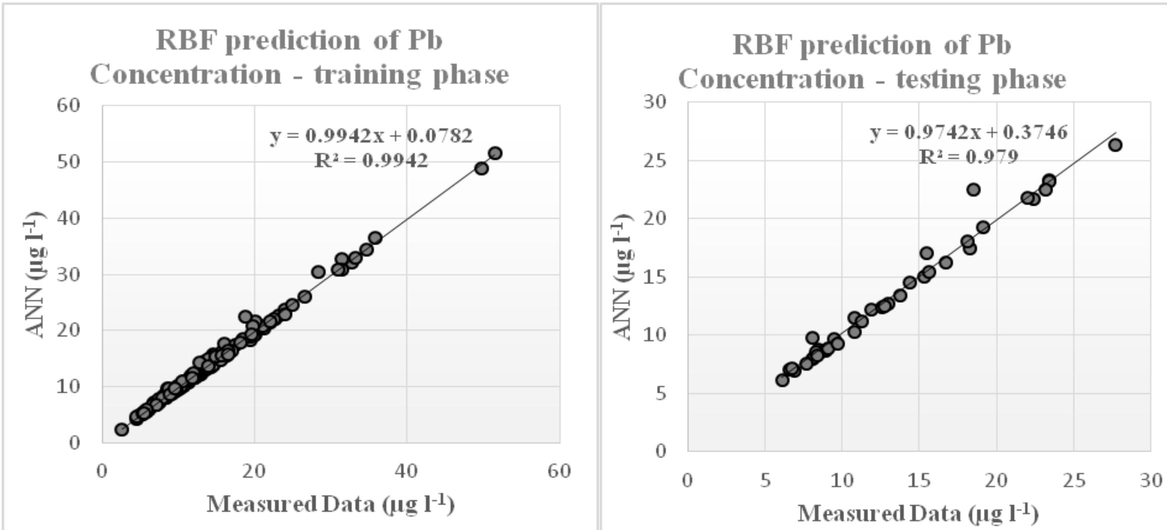

Figure 6. Observed and simulated Pb concentration by RBF model during the training and testing phases.

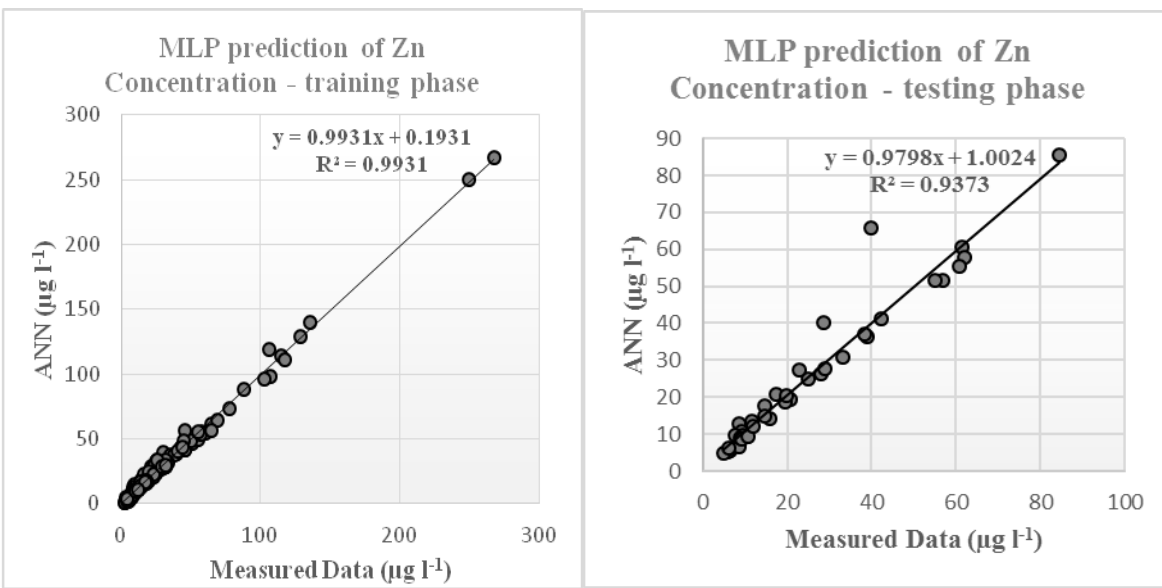

Figure 7. Observed and simulated Zn concentration by MLP model during the training and testing phases.

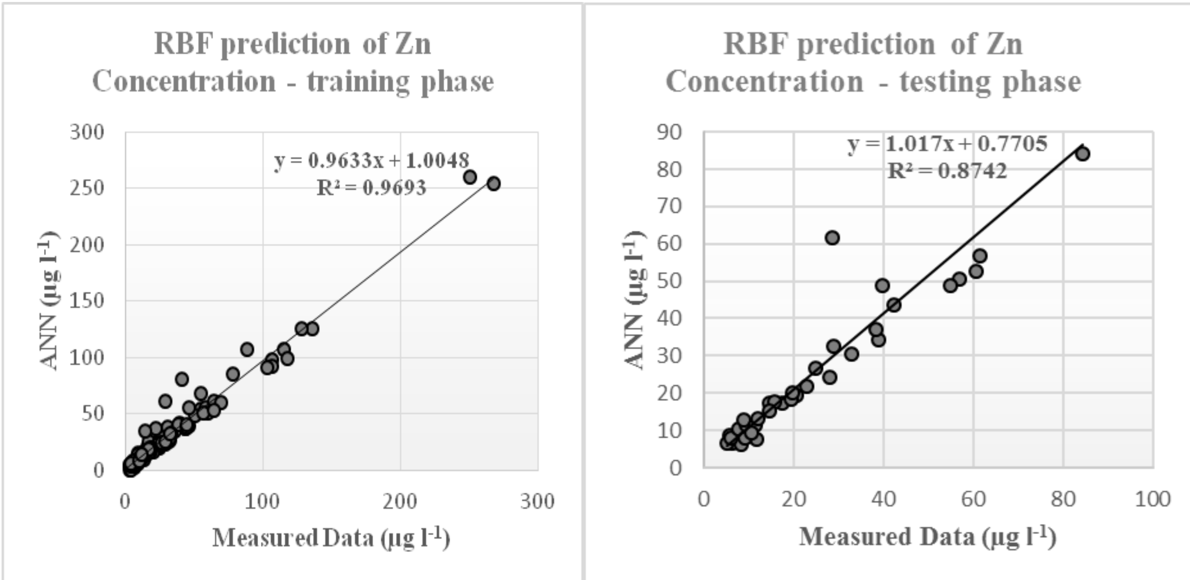

Figure 8. Observed and simulated Zn concentration by RBF model during the training and testing phases.

As shown in Table 3, based on results of statistical parameters, the MLP model with R^2= 0.9721 in the training phase and with R^2=0.9465 in testing phase has better accuracy in comparison with the RBF in prediction of Cd concentration. In addition, in prediction of Pb concentration, the MLP model with R^2= 0.9945 in the training phase and with R^2=0.9821 in testing phase yields better performance. The comparison clearly shows that the RBF model has failed to produce accurate results compared with the proposed MLP model. Moreover, the MLP model with R^2= 0.9931 in the training phase and with R^2=0.9373 in testing phase, revealed reliability of the MLP model was better than the RBF model for forecasting groundwater Zn concentration.

MLP neural network is effective tool for predicting heavy metals concentration. The optimum MLP model proposed in this study shows very promising results for improving planning environmental management in Asadabad Plain. RBF model convergence speed is faster but it has higher levels of prediction errors. ANN model developed by 5 neurons in hidden layer for predicting of water pollution sources in different areas in Turkey and ANN can be used as a powerful analytical tool in water quality modeling [34]. Besides, a three-layer MLP was used by 25 hidden neurons to classify different combinations and groups of water contamination (nitrate and sulfate). The results showed that the good accuracy and high validity of ANN models in predicting of water contaminations [35].

CONCLUSIONS

The MLP offered the most promising results compared to RBF model. For the MLP models, three different algorithms, Levenberg-Marquardt, conjugate gradient and gradient descent, were utilized, Levenberg-Marquardt algorithm was found to be better than the others were. Successful application of MLP technique in the present study shows that MLP was a robust tool for prediction of heavy metals concentration in groundwater resources and suggested that the use of this method is an appropriate practical approach for managers and researchers in other fields of environmental issues.

ACKNOWLEDGMENTS

The authors are grateful to the Hamedan Branch, Islamic Azad University for providing facilities to conduct and complete this study. The authors declare that there is no conflict of interest.

REFERENCES

1. Sarala Thambavani D., Uma Mageswari T.S.R., 2013. Metal pollution assessment in ground water. Bull Environ Pharm Life Sci. 2(12), 122-129.

2. Srinivas J., Purushotham A.V., Murali Krishna K.V.S.G., 2013. Determination of water quality index in industrial areas of Kakinada, Andhra Pradesh, India. Int Res J Environ Sci. 2(5), 37-45.

3. Yari A.R., Siboni M.S., Hashemi S., Alizadeh M., 2013. Removal of heavy metals from aqueous solutions by natural adsorbents (A review). Arch Hyg Sci. 2(3), 114-124.

4. Zhu F., Fan W., Wang X., Qu L., Yao S., 2011. Health risk assessment of eight heavy metals in nine varieties of edible vegetable oils consumed in China. Food Chem Toxicol. 49, 3081-3085.

5. Chen X., Wang K., Wang Z., Gan C., He P., Liang Y., Jin T., Zhu G., 2014. Effects of lead and cadmium co-exposure on bone mineral density in a Chinese population. Bone. 63, 76-80.

6. Hellström L., Persson B., Brudin L., Petersson Grawé K., Öborn I., Järup L., 2007. Cadmium exposure pathways in a population living near a battery plant. Sci Total Environ. 373, 447-455.

7. Pérez A.L., Anderson K.A., 2009. DGT estimates cadmium accumulation in wheat and potato from phosphate fertilizer applications. Sci Total Environ. 407, 5096-5103.

8. Krejpcio Z., Sionkowski S., Bartela J., 2005. Safety of fresh fruits and juices available on thePolish market as determined by heavy metal residues. Polish J Environ Stud. 14(6), 877-881.

9. Peralta-Videa J.R., Lopez M.L., Narayan M., Saupe G., Gardea-Torresdey J., 2009. The biochemistry of environmental heavy metal uptake by plants: Implications for the food chain. Int J Biochem Cell Biol. 41, 1665-1677.

10. Minh N.D., Hough R.L., Thuy L.T., Nyberg Y., Mai L.B., Vinh N.C., Khaim N.M., Öborn I., 2012. Assessing dietary exposure to cadmium in a metal recycling community in Vietnam: Age and gender aspects. Sci Total Environ. 416, 164-171.

11. Zhang W.L., Du Y., Zhai M.M., Shang Q., 2014. Cadmium exposure and its health effects: A 19-year follow-up study of a polluted area in China. Sci Total Environ. 470-471, 224-228.

12. Chen H., Teng Y., Lu S., Wang Y., Wang J., 2015. Contamination features and health risk of soil heavy metals in China. Sci Total Environ. 512-513, 143-153.

13. Lin L.Q., Cong L., Yun W.H., Yang J., Ming H., Wan Z.B., Kai Ch., Lei H., 2015. Association of soil cadmium contamination with ceramic industry: A case study in a Chinese town. Sci Total Environ. 514, 26-32.

14. Matovic V., Buha A., Đukić-Ćosić D., Bulat Z., 2015. Insight into the oxidative stress induced by lead and/or cadmium in blood, liver and kidneys. Food Chem Toxicol. 78, 130-140.

15. Cherfi A., Abdoun S., Gaci O., 2014. Food survey: Levels and potential health risks of chromium, lead, zinc and copper content in fruits and vegetables consumed in Algeria. Food Chem Toxicol. 70, 48-53.

16. Moriarity R.S., Harris J.T., Cox R.D., 2014. Lead toxicity as an etiology for abdominal pain in the emergency department. J Emerg Med. 46, e35-e38.

17. Rahman M.A., Rahman M.M., Reichman S.M., Lim R.P., Naidu R., 2014. Heavy metals in Australian grown and imported rice and vegetables on sale in Australia: Health hazard. Ecotoxicol Environ Safe. 100, 53-60.

18. Narin I., Tuzen M., Sari H., Soylak M., 2005. Heavy metal content of potato and corn chips from Turkey. Bull Environ Contam Toxicol. 74, 1072-1077.

19. Öztürk E., Atsan E., Polat T., Kara K., 2011. Variation in heavy metal concentrations of potato

(*Solanum tuberosum* L.) cultivars. J Animal Plant Sci. 21(2), 235-239.

20. Sobhanardakani S., Jamali M., Maànijou M., 2014. Evaluation of As, Zn, Cr and Mn concentrations in groundwater resources of Razan Plain and preparing the zoning map using GIS. J Environ Sci Technol. 16(2), 25-38. (In Persian)

21. Fazel Tavasol S., Vusuq B.P., Manshuri M., 2010. The investigation of heavy metal (Sn-Pb) concentration in ground water resources and their environmental effects, Case study: North Chardoly Plain, West of Iran. Proceedings of the 1st International Applied Geological Congress; Apr 26-28; Mashhad, Iran. p. 26-28.

22. Hattab N., Hambli R., Motelica-Heino M., Bourrat X., Mench M., 2013. Application of neural network model for the prediction of chromium concentration in phytoremediated contaminated soils. J Geochem Explor. 128, 25-34.

23. Keskin T.E., Dugenci M., Kacaroglu F., 2015. Prediction of water pollution sources using artificial neural networks in the study areas of Sivas, Karabuk and Bartın (Turkey). Environ Earth Sci. 73(9), 5333-5347.

24. Podder M.S., Majumder C.B., 2016. The use of artificial neural network for modelling of phycoremediation of toxic elements As(III) and As(V) from wastewater using *Botryococcus braunii*. Spectrochim Acta A Mol Biomol Spectrosc. 155, 130-145.

25. Mandal S., Mahapatra S.S., Sahu M.K., Patel R.K., 2015. Artificial neural network modeling of As(III) removal from water by novel hybrid material. Process Safety Environ Protect. 93, 249-264.

26. Mohamad Nor A.S., Faramarzi M., Md Yunus M.D., Ibrahim S., 2014. Nitrate and sulfate estimations in water sources using a planar electromagnetic sensor array and artificial neural network method. IEEE Sensor J. 15(1), 497-504.

27. Hornik K., Stinchcombe M., White H., 1990. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. Neur Net. 3(5), 551-560.

28. Govindaraju R.S., 2000. Artificial neural networks in hydrology. I: Preliminary concepts. J Hydrol Eng. 5(2), 115-123.

29. Daliakopoulos I.N., Coulibaly P., Tsanis I.K., 2005. Groundwater level forecasting using artificial neural networks. J Hydrol. 309(1-4), 229-240.

30. Haykin S., 1998. Neural Networks: A Comprehensive Foundation. New Jersey: Prentice-Hall, Upper Saddle River. 88-92.

31. Sudheer K.P., Gosain A.K., Ramasastri K.S., 2002. A data-driven algorithm for constructing artificial neural network rainfall-runoff models. Hydrol Process. 16(6), 1325-1330.

32. Lin J.Y., Cheng C.T., Chau K.W., 2006. Using support vector machines for long-term discharge prediction. Hydrol Sci J. 51(4), 599-612.

33. Hecht-Nielsen R., 1987. Kolmogorov's mapping neural network existence theorem. Proceedings of the International Symposium on Neural Networks; New York, USA.

34. Keskin T.E., Dugenci M., Kacaroglu F., 2015. Prediction of water pollution sources using artificial neural networks in the study areas of Sivas, Karabuk and Bartın (Turkey). Environ Earth Sci. 73(9), 5333-5347.

35. Nor A.S.M., Faramarzi M., Yunus M.A.M., Ibrahim S., 2015. Nitrate and sulfate estimations in water sources using a planar electromagnetic sensor array and artificial neural network method. IEEE Sensors J. 15(1), 497-504.