

Prediction of heavy metals in acid mine drainage using artificial neural network from the Shur River of the Sarcheshmeh porphyry copper mine, Southeast Iran

R. Rooki · F. Doulati Ardejani · A. Aryafar ·
A. Bani Asadi

Received: 16 March 2010 / Accepted: 28 January 2011 / Published online: 17 February 2011
© Springer-Verlag 2011

Abstract Water is one of the basic and fundamental requirements for the survival of human beings. Mining of the sulphide mines usually produce a significant amount of acid mine drainage (AMD) contributing to huge amounts of chemical components and heavy metals in the receiving waters. Prediction of the heavy metals in the AMD is important in developing any appropriate remediation strategy. This paper attempts to predict heavy metals (Cu, Fe, Mn, Zn) from the AMD using backpropagation neural network (BPNN), general regression neural network (GRNN) and multiple linear regression (MLR), by taking pH, sulphate (SO₄) and magnesium (Mg) concentrations in the AMD into account in Shur River, Sarcheshmeh porphyry copper deposit, southeast Iran. The comparison between the predicted concentrations and the measured data resulted in the correlation coefficients, *R*, 0.92, 0.22, 0.92 and 0.92 for Cu, Fe, Mn and Zn ions using BPNN method. Moreover, the *R* values were 0.89, 0.37, 0.9 and 0.91 for Cu, Fe, Mn, and Zn taking the GRNN method into consideration. However, the correlation coefficients were

low for the results predicted by MLR method (0.83, 0.14, 0.9 and 0.85 for Cu, Fe, Mn and Zn ions, respectively). The results further indicate that the ANN can be used as a viable method to rapidly and cost-effectively predict heavy metals in the AMD. The results obtained from this paper can be considered as an easy and cost-effective method to monitor groundwater and surface water affected by AMD.

Keywords Acid mine drainage (AMD) · Heavy metals · Artificial neural network (ANN) · Sarcheshmeh copper mine

Introduction

Copper exploitation is a major water quality problem due to acid mine drainage (AMD) generation in Sarcheshmeh mine, Kerman Province, southeast Iran. The oxidation of sulphide minerals, in particular pyrite exposed to atmospheric oxygen during or after mining activities generates acidic waters with high concentrations of dissolved iron (Fe), sulphate (SO₄) and heavy metals (Williams 1975; Moncur et al. 2005). The low pH of AMD may cause further dissolution and the leaching of additional metals (Mn, Zn, Cu, Cd, and Pb) into aqueous system (Zhao et al. 2007). AMD containing heavy metals have detrimental impact on aquatic life and the surrounding environment. Shur River in the Sarcheshmeh copper mine polluted by AMD with pH values ranging between 2 and 4.5 and high concentrations of heavy metals. The prediction of heavy metals in Shur River is useful in developing proper remediation and monitoring methods.

The Sarcheshmeh copper deposit recognised to be the fourth largest mine in the world contains 1 billion tonnes averaging 0.9% copper and 0.03% molybdenum (Banisi and

R. Rooki
Mining Engineering, Shahrood University of Technology,
Shahrood, Iran

F. Doulati Ardejani
Faculty of Mining, Petroleum and Geophysics,
Shahrood University of Technology, Shahrood, Iran

A. Aryafar (✉)
Department of Mining Engineering, Faculty of Engineering,
Birjand University, Birjand, Iran
e-mail: ahariafar@yahoo.com

A. Bani Asadi
Environmental Geology, Faculty of Geosciences,
Shahrood University of Technology, Shahrood, Iran

Finch 2001). This ore body is located at southeast of Iran, Kerman Province. Mining operation has placed many low-grade waste dumps and has posed many environmental problems. Environmental problems of sulphide minerals oxidation and AMD generation in the Sarcheshmeh copper mine and its impact on the Shur River have been investigated in the past (Marandi et al. 2007; Shahabpour and Doorandish 2008; Doulati Ardejani et al. 2008; Bani Assadi et al. 2008).

Many investigations have been carried out on the behaviour of the heavy metals in AMD and their impact on the receiving water bodies (Govil et al. 1999; Merrington and Alloway 1993; Hammack et al. 1998; Herbert 1994; Moncur et al. 2005; Smuda et al. 2007; Wilson et al. 2005; Lee and Chon 2006; Dinelli et al. 2001; Canovas et al. 2007). The conventional method of measuring the heavy metals is involved in the sampling and a time-consuming and expensive laboratory analysis. Less study has been carried out for the prediction of heavy metals in AMD. Therefore, investigation of a method that can predict the concentrations of heavy metals in water affected by AMD is necessary to develop an appropriate remediation and monitoring method for comprehensive assessment of the potential environmental impacts of AMD.

Artificial neural networks (ANN) have gained an increasing popularity in different fields of engineering in the past few decades, because of their capability of extracting complex and nonlinear relationships. Kemper and Sommer (2002) estimated the heavy metal concentration in soils from reflectance spectroscopy using back propagation network and multiple linear regression. Almasri and Kaluarachchi (2005) applied the modular neural networks to predict the nitrate distribution in ground water using the on-ground nitrogen loading and recharge data. Khandelwal and Singh (2005) predicted the mine

water quality by the physical parameters using back propagation neural network (BPNN) and multiple linear regression. Singh et al. (2009) modelled the backpropagation neural network to predict water quality in the Gomti River (India). Erzin and Yukselen (2009) used the back propagation neural network for the prediction of zeta potential of kaolinite.

The literature review has shown that despite many research works having been conducted related to the application of the ANN method in mining and relevant environmental problems, the ANN method has not been directly used to predict heavy metals in AMD. In this paper, attention has been focused on the prediction of the heavy metals in the Shur River impacted by AMD. The results obtained from the predictions using ANN and MLR are compared with the measured concentrations of major heavy metals sampled and analysed in Shur River of Sarcheshmeh copper mine, southeast Iran.

Site description

Sarcheshmeh copper mine is located 160 km to southwest of Kerman and 50 km to southwest of Rafsanjan in Kerman province, Iran. The main access road to the study area is Kerman–Rafsanjan–Shahr Babak road. This mine belongs to Band Mamazar–Pariz Mountains. The average elevation of the mine is 1,600 m. The mean annual precipitation of the site varies from 300 to 550 mm. The temperature varies from +35°C in summer to –20°C in winter. The area is covered with snow for about 3–4 months per year. The wind speed sometimes exceeds 100 km/h. A rough topography is predominant at the mining area; Fig. 1 shows the geographical position of the Sarcheshmeh copper mine.

Fig. 1 Location of the Sarcheshmeh mine and Shur River (modified after Shahabpour and Doorandish 2008; Derakhshandeh and Alipour 2010)

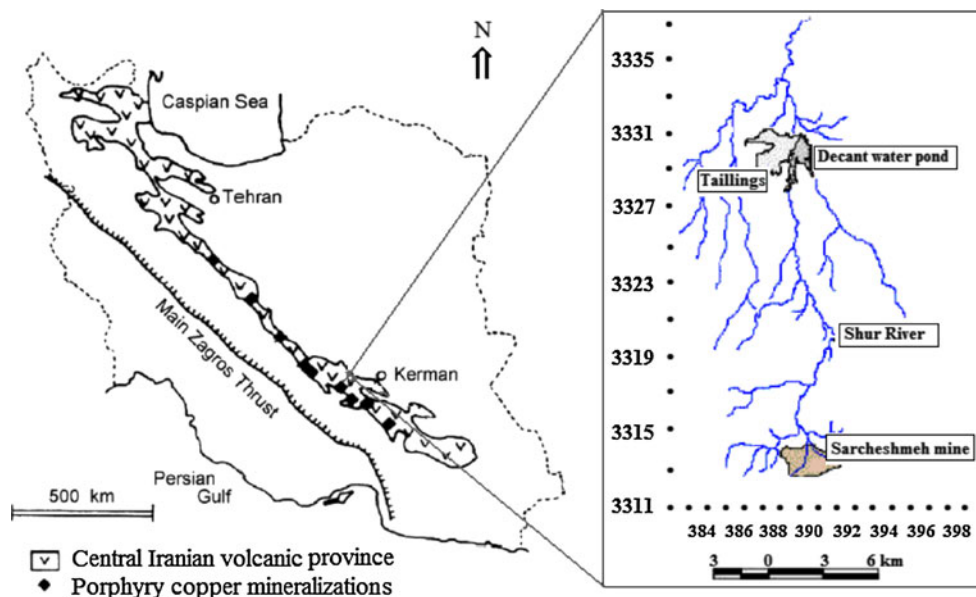


Table 1 Maximum, minimum and mean physical and chemical constituents including heavy metals of the Shur River

	pH	SO ₄ ²⁻	Cl ⁻	HCO ₃ ⁻	Ca ²⁺	Mg ²⁺	Cu	Fe	Mn	Zn	TDS	EC (μS/cm)
Min	3.3	27	0	0	92	13	0	0.01	0.04	0	446	870
Max	7.20	1,526	230	628	460	123	158	23	52	31.48	2,080.68	2,260
Mean	5.27	778.45	23.56	34.01	182.78	56.70	20.29	4.60	16.05	6.33	1,009.90	1,306.52

Concentrations of elements are given in ppm

The orebody in Sarcheshmeh is oval shaped with a long dimension of about 2,300 m and a width of about 1,200 m. This deposit is associated with the late Tertiary Sarcheshmeh granodiorite porphyry stock (Waterman and Hamilton 1975). The porphyry is a member of a complex series of magmatically related intrusives emplaced in the Tertiary volcanics at a short distance from the edge of an older near-batholith-sized granodiorite mass. Open pit mining method is used to extract copper deposit in Sarcheshmeh. A total of 40,000 tons of ore (average grades 0.9% Cu and 0.03% molybdenum) is approximately extracted per day in Sarcheshmeh mine (Banisi and Finch 2001).

Sampling and field methods

Sampling of waters in the Shur River downstream from the Sarcheshmeh mine was carried out in February 2006. Water samples consist of water from Shur River (Fig. 1) originating from Sarcheshmeh mine, acidic leachates of heap structure, run-off of leaching solution into the River and samples affected by tailings along the Shur River. The water samples were immediately acidified by adding HNO₃ (10 cc acid/1,000 cc sample) and stored under cool conditions. The equipments used in this study were sample container, GPS, oven, autoclave, pH meter, atomic adsorption and ICP analysers. The pH of the water was measured using a portable pH meter in the field. Other physical parameters were total dissolved solids (TDS), electric conductivity (EC) and temperature. Analyses for dissolved metals were performed using atomic adsorption spectrometer (AA220) in water lab of the National Iranian Copper Industries Company (NICIC). Although not given here, ICP (model 6000) was also used to analyse the concentrations of those heavy metals that are detected in the range of ppb. Table 1 gives the minimum, maximum and the mean values of some physical and chemical parameters.

Method

Back propagation neural network design

Artificial neural networks (ANN) are generally defined as information processing representation of the biological

neural networks. ANN has gained an increasing popularity in different fields of engineering in the past few decades, because of their ability of deriving complex and nonlinear relationships. The mechanism of the ANN is based on the following four major assumptions (Hagan et al. 1996):

- Information processing occurs in many simple elements that are called neurons (processing elements).
- Signals are passed between neurons over connection links.
- Each connection link has an associated weight, which, in a typical neural network, multiplies the signal being transmitted.
- Each neuron applies an activation function (usually nonlinear) to its net input in order to determine its output signal.

Figure 2 shows a typical neuron. Inputs (*P*) coming from another neuron are multiplied by their corresponding weights (*W_{1,R}*), and summed up (*n*). An activation function (*f*) is then applied to the summation, and the output (*a*) of that neuron is now calculated and ready to be transferred to another neuron. Many types of neural network architectures and algorithms are available. In this study, a generalised regression neural network (GRNN) is used.

In this network, each element of the input vector *p* is connected to each neuron input through the weight matrix *W*. The *i*th neuron has a summer that gathers its weighted inputs and bias to form its own scalar output *n* (*i*). The various *n* (*i*) taken together form an *S*-element net input vector *n*. Finally, the neuron layer outputs form a column vector *a* (Eqs. 1, 2).

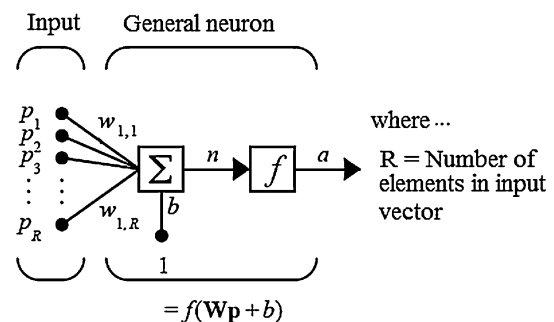


Fig. 2 A typical neuron (Demuth and Beale 2002)

$$n_j = \sum_{i=1}^R (p_i w_{ij} + b_j), \quad j = 1, 2, \dots, S \tag{1}$$

where

$$b = \begin{bmatrix} b_1 \\ b_{2S} \\ b_S \end{bmatrix}, \quad P = \begin{bmatrix} P_1 \\ P_2 \\ P_R \end{bmatrix}, \quad W = \begin{bmatrix} w_{1,1}w_{1,2} \dots w_{1,R} \\ w_{2,1}w_{2,2} \dots w_{2,R} \\ \dots \\ w_{S,1}w_{S,2} \dots w_{S,R} \end{bmatrix}$$

Then, final output of network is calculated by:

$$a_S = f(n_S) \tag{2}$$

Here, f is an activation function, typically a step function or a sigmoid function, which takes the argument n and produces the output a . Figure 3 shows examples of various activation functions.

Backpropagation neural networks (BPNN) are recognised for their prediction capabilities and ability to generalise well on a wide variety of problems. These models are supervised type of networks; in other words, trained with both inputs and target outputs. During training the network tries to match the outputs with the desired target values. Learning starts with the assignment of random weights. The output is then calculated and the error is estimated. This error is used to update the weights until the stopping criterion is reached. It should be noted that the stopping criteria is usually the average error or epoch.

Network training: the over fitting problem

One of the most common problems in the training process is the over fitting phenomenon. This happens when the error on the training set is driven to a very small value, but when new data is presented to the network the error is large. This problem occurs mostly in case of large networks with only few available data. Demuth and Beale (2002) have shown that there are a number of ways to avoid over fitting problem. Early stopping and automated Bayesian regularization methods are most common. However, with immediate fixing of the error and the number of epochs to an adequate level (not too low/not too high) and dividing

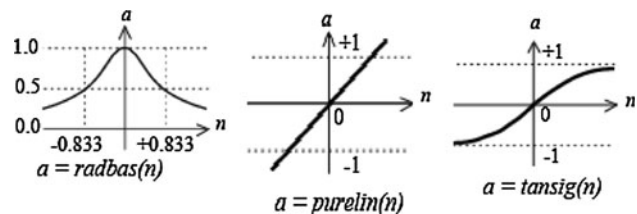


Fig. 3 Three examples of transfer functions (Demuth and Beale 2002)

the data into two sets, training and testing, one can avoid such problem by making several realizations and selecting the best of them. In this paper, the necessary coding was added through MATLAB multi-purpose commercial software to implement the automated Bayesian regularization for training BPNN. In this technique, the available data is divided into two subsets. The first subset is the training set, which is used for computing the gradient and updating the network weights and biases. The second subset is the test set. This method works by modifying the performance function, which is normally chosen to be the sum of squares of the network errors on the training set. The typical performance function that is used for training feed forward neural networks is the mean sum of squares of the network errors according to the following equation:

$$mse = \frac{1}{N} \sum_{i=1}^N (e_i)^2 = \frac{1}{N} \sum_{i=1}^N (t_i - a_i)^2 \tag{3}$$

where, N represents the number of samples, a is the predicted value, t denotes the measured value and e is the error.

It is possible to improve generalisation if we modify the performance function by adding a term that consists of the mean of the sum of squares of the network weights and biases which is given by:

$$msereg = \gamma mse + (1 - \gamma) msw \tag{4}$$

Where, $msereg$ is the modified error, γ is the performance ratio, and msw can be written as:

$$msw = \frac{1}{N} \sum_{i=1}^N w_i \tag{5}$$

Performance function will cause the network to have smaller weights and biases, and this will force the network response to be smoother and less likely to over fit (Demuth and Beale 2002).

Heavy metals prediction using BPNN

According to the correlation matrix (Table 2), pH, SO₄ and Mg that have most dependent on heavy metals (Cu, Mn and Zn) concentrations were selected as inputs of the network. The outputs of network were heavy metals concentrations including Cu, Fe, Mn and Zn. In view of the requirements of the neural computation algorithm, the data of both the independent and dependent variables were normalised to an interval by transformation process. In this study, normalisation of data (inputs and outputs) was done for the range of (-1, 1) using Eq. 6 and the number of training data (44) and test data (12) were then selected randomly.

$$p_n = 2 \frac{P - P_{min}}{P_{max} - P_{min}} - 1 \tag{6}$$

Table 2 Correlation matrix between heavy metals concentrations and independent variables

	pH	SO ₄	Cl	HCO ₃	Ca	Mg	Cu	Fe	Mn	Zn	TDS	EC
PH	1											
SO ₄	-0.713	1										
Cl	-0.046	-0.035	1									
HCO ₃	0.310	-0.507	0.534	1								
Ca	-0.308	0.645	0.088	-0.014	1							
Mg	-0.625	0.694	-0.004	-0.238	0.121	1						
Cu	-0.697	0.663	-0.101	-0.199	0.238	0.764	1					
Fe	0.021	0.078	-0.195	-0.255	-0.164	-0.008	-0.022	1				
Mn	-0.758	0.668	-0.160	-0.305	0.173	0.709	0.809	0.068	1			
Zn	-0.710	0.624	-0.186	-0.306	0.204	0.659	0.720	0.156	0.910	1		
TDS	-0.407	0.649	-0.235	-0.334	0.509	0.366	0.390	0.099	0.563	0.613	1	
EC	-0.779	0.767	0.212	-0.096	0.475	0.727	0.757	-0.081	0.739	0.628	0.334	1

Table 3 R and RMS between predicted and measured heavy metals concentrations in training and test data

n	R ₁	R ₂	RMS ₁	RMS ₂	R ₁	R ₂	RMS ₁	RMS ₂	R ₁	R ₂	RMS ₁	RMS ₂	R ₁	R ₂	RMS ₁	RMS ₂
	Cu	Cu	Cu	Cu	Fe	Fe	Fe	Fe	Mn	Mn	Mn	Mn	Zn	Zn	Zn	Zn
5	0.97	0.94	7.77	6.48	0.37	0.03	4.15	3.72	0.97	0.92	2.91	5.12	0.95	0.92	1.89	1.34
6	0.97	0.92	7.68	6.50	0.35	0.22	4.10	3.45	0.97	0.92	2.90	5.18	0.95	0.92	1.91	1.31
7	0.97	0.88	7.85	9.28	0.26	0.12	4.31	3.56	0.98	0.93	2.48	4.82	0.95	0.92	1.78	1.68
8	0.98	0.86	6.63	10.88	0.29	0.11	4.28	3.50	0.98	0.91	2.39	4.84	0.96	0.91	1.59	1.57
9	0.98	0.81	6.64	13.78	0.49	-0.12	3.89	5.79	0.98	0.86	2.28	5.56	0.97	0.81	1.45	3.97
10	0.98	0.75	6.29	15.80	0.36	-0.12	4.17	3.83	0.98	0.83	2.15	6.20	0.97	0.88	1.33	3.32

where, P_n is the normalised parameter, p denotes the actual parameter, p_{\min} represents a minimum of the actual parameters and p_{\max} stands for a maximum of the actual parameters.

In this research, several architectures (varied numbers of neurons in hidden layer) with Automated Bayesian Regularization algorithm for ANN model with the default parameter values for algorithm (Demuth and Beale 2002) are used to predict heavy metals concentrations using BPNN. Two criteria were used to evaluate the effectiveness of each network and its ability to make accurate predictions. The root mean square error (RMS) can be calculated as follows:

$$RMS = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \tag{7}$$

where, y_i is the measured value, \hat{y}_i denotes the predicted value, and n stands for the number of samples. RMS indicates the discrepancy between the measured and predicted values. The lower the RMS, the more accurate the prediction is. Furthermore, the efficiency criterion, R^2 , is given by:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n y_i^2 - \frac{\sum_{i=1}^n \hat{y}_i^2}{n}} \tag{8}$$

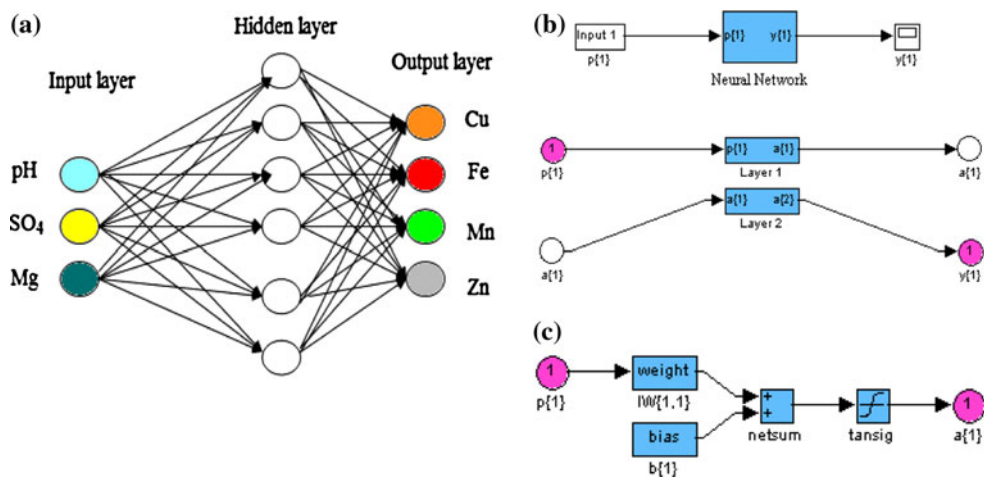
where R^2 efficiency criterion represents the percentage of the initial uncertainty explained by the model. The best fitting between measured and predicted values, which is unlikely to occur, would have $RMS = 0$ and $R^2 = 1$. Table 3 gives the correlation coefficient (R) and RMS between predicted and measured concentrations in training and test from any architect.

The indices 1 and 2 for R and RMS in Table 3 are related to the training and test data, respectively and n is the number of neurons in the hidden layer.

The optimal network for this study is a feed forward multilayer perceptron (Cybenko 1989; Hornik et al. 1989; Haykin 1994; Noori et al. 2009, 2010), having one input layer with three inputs (pH, SO₄, Mg), one hidden layer with six neurons that each neuron has a bias and is fully connected to all inputs and utilises sigmoid hyperbolic tangent (tansig) activation function (Fig. 4). The output layer has four neurons (Cu, Fe, Mn and Zn) with a linear activation function (purelin) without bias. Linear activation function can provide any range of data in output without any limitation for output values.

Bayesian regularization algorithm (trainbr) was used as training function to prevent overtraining of the ANN models. Figure 4a shows the backpropagation neural

Fig. 4 **a** Backpropagation neural network architecture, **b** general schematic diagram of network and its layers, **c** structure of hidden layer (Layer 1)



network architecture. In Fig. 4b, Layer 1 is hidden layer and Layer 2 is output layer. Figure 4c shows the structure of the hidden layer.

Figure 5 shows the training process of the network. In this figure, SSE is the sum square error for training data. One feature of this algorithm is that it provides a measure of how many network parameters (weights and biases) are being effectively used by the network. The final trained network employs approximately 44 parameters out of the 52 total weights and biases in the 3-6-4 network. The training may stop with the message “Maximum MU reached”. This is typical, and is a good indication that the algorithm has truly converged. In the present case, the algorithm was stopped in 171 epochs. In this network learning rate was 0.5.

The selected BPNN (3 neurons in input layer, 6 neurons in hidden layer and 4 neurons in output layer) provided a good-fit model for two data sets of Cu, Mn and Zn concentrations and poor fit for Fe concentration. The correlation coefficient (R) values for the training and test data and the respective values of RMS for the two data sets are highlighted in Table 3. Figure 6a–h compares the network predictions versus measured concentrations for training and test data.

GRNN model

General regression neural network has been proposed by Specht (1991). GRNN is a type of supervised network and also trains quickly on sparse data sets but, rather than categorising it. GRNN applications are able to produce continuous valued outputs. GRNN is a three-layer network where there must be one hidden neuron for each training pattern.

GRNN is a memory-based network that provides estimates of continuous variables and converges to the underlying regression surface. GRNNs are based on the estimation of probability density functions, having a

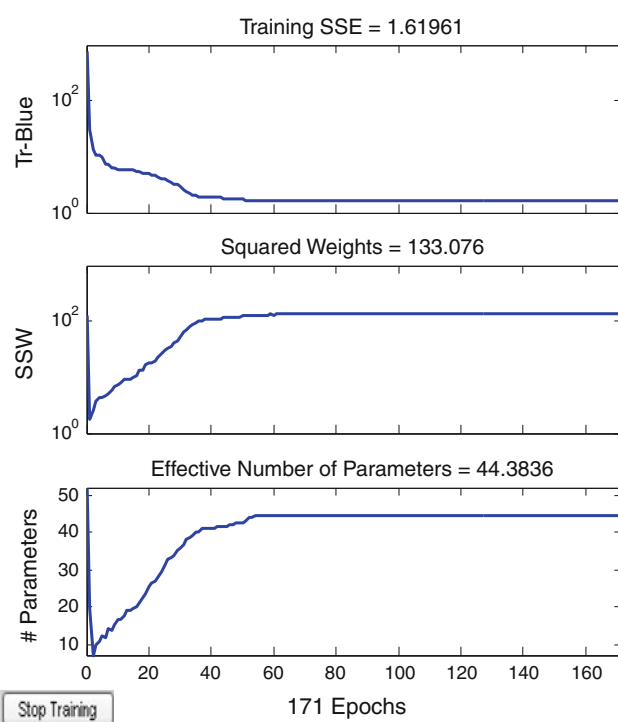


Fig. 5 Sum squared error, sum squared weights and effective number of parameters versus the epoch in training step

feature of fast training times and can model nonlinear functions. GRNN is an one-pass learning algorithm with a highly parallel structure. GRNN algorithm provides smooth transitions from one observed value to another even with sparse data in a multidimensional measurement space. The algorithmic form can be used for any regression problem in which an assumption of linearity is not justified. GRNN can be thought as a normalised radial basis functions (RBF) network in which there is a hidden unit centred at every training case. These RBF units are usually probability density functions such as the Gaussian. The only weights that need to be learned are the widths of the RBF

units. These widths are called “smoothing parameters”. The main drawback of GRNN is that it suffers badly from the curse of dimensionality. GRNN cannot ignore irrelevant inputs without major modifications to the basic algorithm. So GRNN is not likely to be the top choice if there are more than 5 or 6 non-redundant inputs. The regression of a dependent variable, Y , on an independent variable, X , is the computation of the most probable value of Y for each value of X based on a finite number of possibly noisy measurements of X and the associated values of Y . The variables X and Y are usually vectors. To implement system identification, it is usually necessary to assume some functional form. In the case of linear regression, for example, the output Y is assumed to be a linear function of the input, and the unknown parameters, a_i , are linear coefficients.

The method does not need to assume a specific functional form. A Euclidean distance (D_i^2) is estimated between an input vector and the weights, which are then

rescaled by the spreading factor. The radial basis output is then the exponential of the negatively weighted distance. The GRNN equation can be written as:

$$D_i^2 = (X - X^i)^T (X - X^i) \tag{9}$$

$$Y(X) = \frac{\sum_{i=1}^n Y_i \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^n \exp\left(-\frac{D_i^2}{2\sigma^2}\right)} \tag{10}$$

where σ is the smoothing factor (SF).

The estimate $Y(X)$ can be visualised as a weighted average of all of the observed values, Y_i , where each observed value is weighted exponentially according to its Euclidian distance from X . $Y(X)$ is simply the sum of Gaussian distributions centred at each training sample. However, the sum is not limited to being Gaussian. In this theory, the optimum smoothing factor is determined after several runs according to the mean squared error of the

Fig. 6 Comparison of the network predictions and measured concentrations for training and test data using BPNN model. **a** Correlation between BPNN Cu versus measured Cu (training data). **b** Correlation between BPNN Cu versus measured Cu (test data). **c** Correlation between BPNN Fe versus measured Fe (training data). **d** Correlation between BPNN Fe versus measured Fe (test data). **e** Correlation between BPNN Mn versus measured Mn (training data). **f** Correlation between BPNN Mn versus measured Mn (test data). **g** Correlation between BPNN Zn versus measured Zn (training data). **h** Correlation between BPNN Zn versus measured Zn (test data)

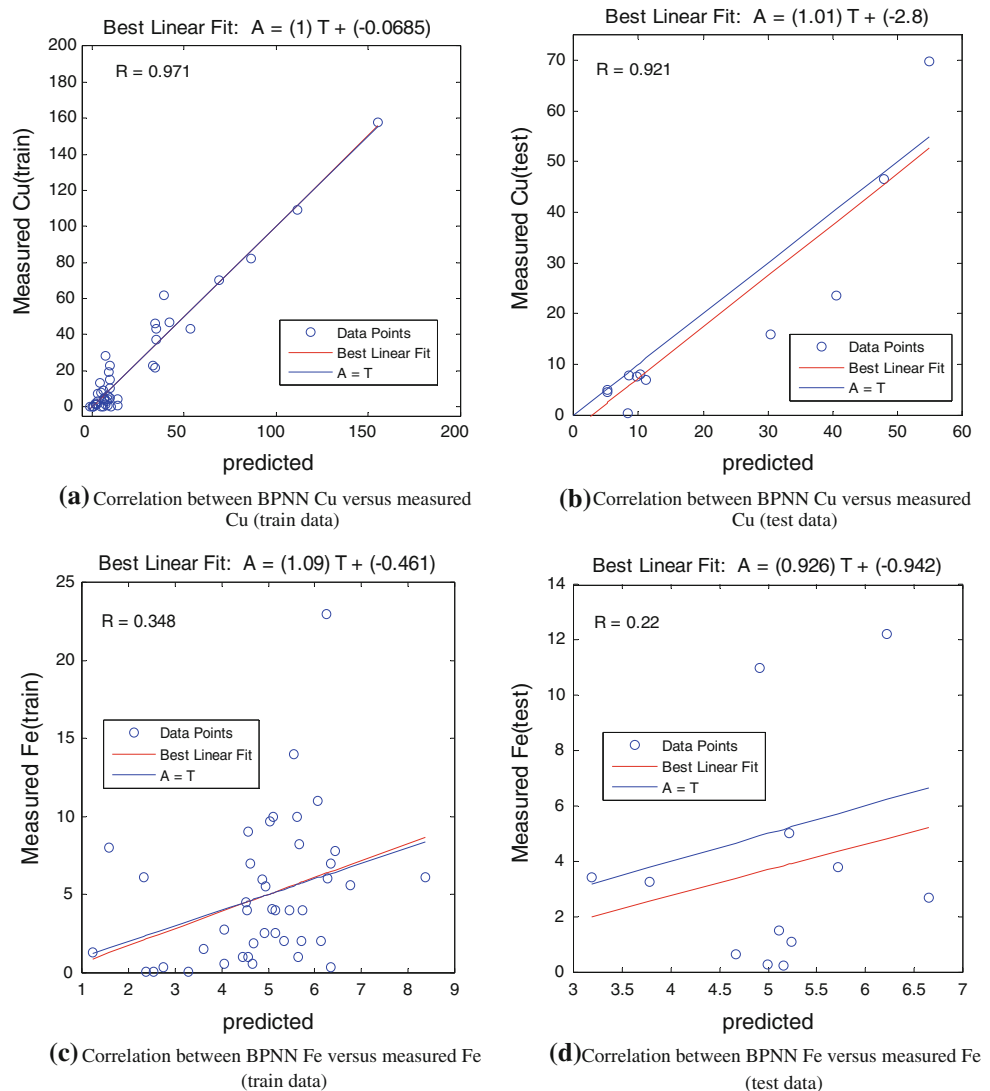
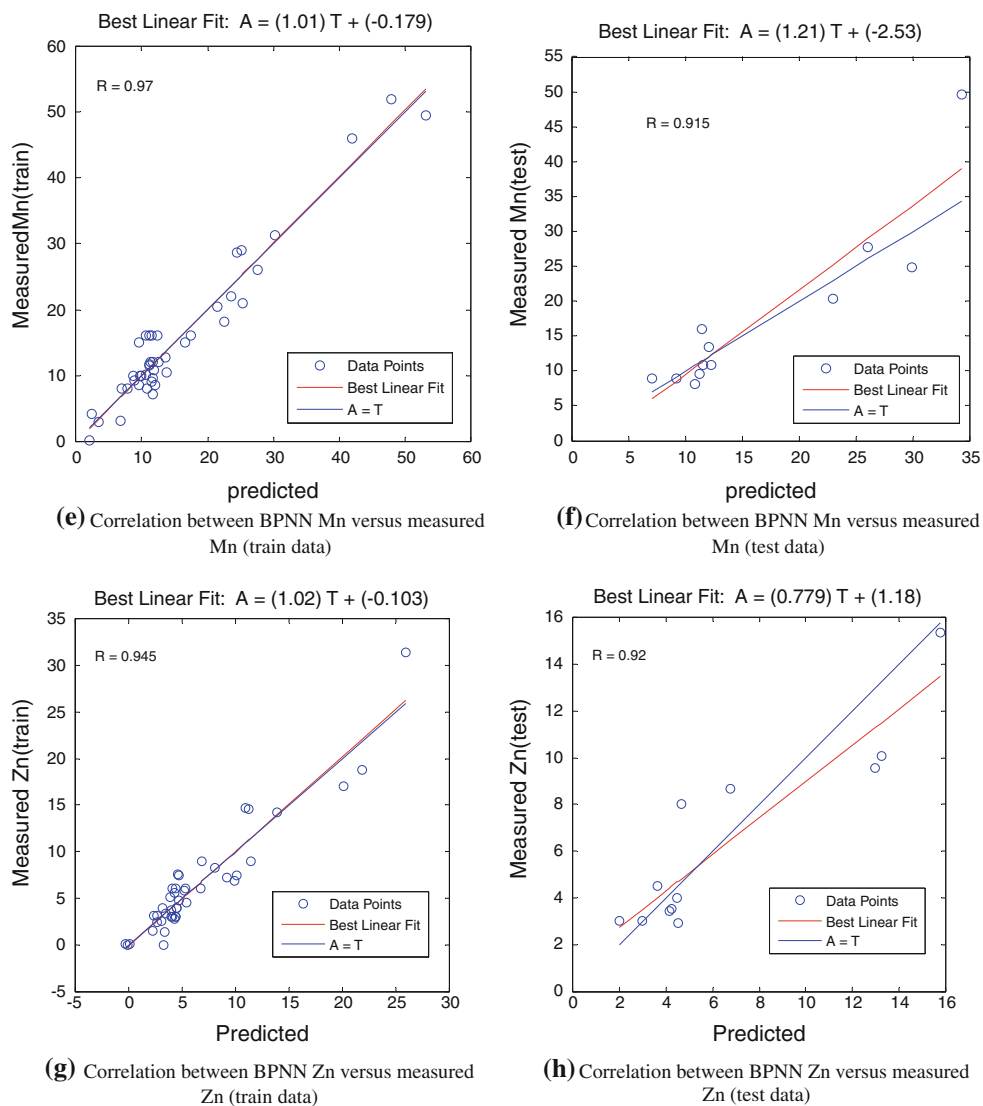


Fig. 6 continued



estimate, which must be kept at minimum. This process is referred to as the training of the network. If a number of iterations pass with no improvement in the mean squared error, that smoothing factor is determined as the optimum one for that data set. While applying the network to a new set of data, increasing the smoothing factor would result in decreasing the range of output values (Specht 1991). In this network, there are no training parameters such as the learning rate, momentum, optimum number of neurons in hidden layer and learning algorithms as in backpropagation network but there is a smoothing factor that its optimum is gained as trial and error. The smoothing factor must be greater than 0 and can usually range from 0.1 to 1 with good results. The number of neurons in the input layer is the number of inputs in the proposed problem, and the number of neurons in the output layer corresponds to the number of outputs. Because GRNN networks evaluate each output independently of the other outputs, GRNN networks

may be more accurate than backpropagation networks when there are multiple outputs. GRNN works by measuring how far the given sample pattern is from the patterns in the training set. The output that is predicted by the network is a proportional amount of all the output in the training set. The proportion is based upon how far the new pattern is from the given patterns in the training set.

Heavy metals prediction using GRNN

In this method, the training and test data in BPNN were used. To obtain the best network, the GRNN was trained by different smooth factors to gain the optimum smooth factor according to correlation coefficient and RMS error between measured and predicted values in training and test data. The results are given in Table 4.

In Table 4, the indices 1 and 2 for R and RMS are related to training and test data, respectively and SF stands

Table 4 *R* and RMS with different smooth factors in training and test data

SF	<i>R</i> ₁ Cu	<i>R</i> ₂ Cu	RMS ₁ Cu	RMS ₂ Cu	<i>R</i> ₁ Fe	<i>R</i> ₂ Fe	RMS ₁ Fe	RMS ₂ Fe	<i>R</i> ₁ Mn	<i>R</i> ₂ Mn	RMS ₁ Mn	RMS ₂ Mn	<i>R</i> ₁ Zn	<i>R</i> ₂ Zn	RMS ₁ Zn	RMS ₂ Zn
0.02	1.00	0.79	1.03	12.55	1.00	0.35	0.39	3.65	1.00	0.79	0.40	7.95	1.00	0.84	0.20	2.22
0.04	1.00	0.81	2.31	11.99	0.99	0.34	0.75	3.63	1.00	0.82	0.81	7.43	1.00	0.84	0.47	2.20
0.06	0.99	0.83	3.68	11.46	0.97	0.37	1.10	3.54	0.99	0.85	1.27	7.05	0.99	0.86	0.77	2.07
0.08	0.99	0.86	4.67	10.63	0.95	0.41	1.65	3.47	0.99	0.88	1.62	6.77	0.99	0.89	1.02	1.81
0.10	0.99	0.89	5.47	9.89	0.90	0.37	2.20	3.55	0.99	0.90	1.89	6.60	0.98	0.91	1.31	1.56
0.12	0.98	0.91	6.18	9.43	0.86	0.26	2.62	3.69	0.98	0.92	2.23	6.58	0.96	0.93	1.66	1.42
0.14	0.98	0.92	6.88	9.31	0.81	0.13	2.95	3.82	0.97	0.92	2.69	6.69	0.94	0.93	2.02	1.39
0.16	0.97	0.92	7.58	9.34	0.77	0.00	3.21	3.91	0.96	0.92	3.19	6.84	0.92	0.93	2.36	1.41
0.18	0.97	0.93	8.28	9.33	0.73	-0.10	3.44	3.98	0.95	0.92	3.65	6.99	0.90	0.93	2.64	1.46
0.20	0.96	0.93	8.94	9.23	0.68	-0.19	3.64	4.03	0.94	0.93	4.03	7.09	0.87	0.93	2.86	1.50

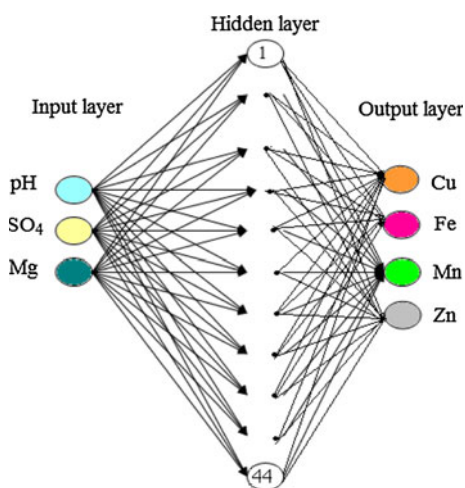


Fig. 7 Schematic diagram of GRNN network

for smooth factor. The optimum smooth factor (SF) was selected 0.10 according to evaluated criteria in training and test data in Table 2. Taking this SF into consideration, Fig. 7 shows the schematic diagram of GRNN network. The general diagram of network, the associated layers and the structure of hidden layer are shown in Fig. 8a and b, respectively. This network (Fig. 7) has three layers; input layer with 3 neurons (pH, SO₄ and Mg), hidden layer incorporating 44 neurons (number of training samples) with radbas activation function in all neurons and output layer with 4 neurons (Cu, Fe, Mn and Zn) with linear activation function. In Fig. 8a, Layer 1 is hidden layer and Layer 2 is output layer and as mentioned above, Fig. 8b is the structure of hidden layer.

Figure 9 compares the measured and predicted concentrations of heavy metals in training and test data. The selected GRNN (3 nodes in input layer, 44 nodes in hidden layer, and 4 nodes in output layer) provided a good-fit model for the two data sets of Cu, Mn and Zn concentrations and poor fit for Fe concentration. The correlation coefficients (*R*) for the

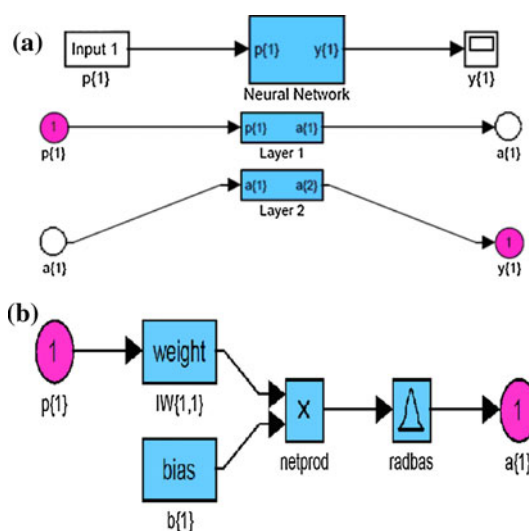


Fig. 8 **a** General diagram of network and its layers, **b** structure of hidden layer (Layer 1)

training and test data and the respective values of RMS for two data sets are shown in Table 4. A closely followed pattern of variation by the measured and predicted heavy metals, *R* and RMS values suggest a good-fit of the heavy metals (Cu, Mn and Zn) model to the data set. The poor-fit model for Fe ion is a result of low correlation between Fe and independent variables in Table 2.

Multiple linear regression

Multiple linear regression (MLR) is an extension of the regression analysis that incorporates additional independent variables in the predictive equation. Here, the model to be fitted is:

$$y = B_1 + B_2x_2 + \dots + B_nx_n + e \tag{11}$$

where *y* is the dependent variable, *x_i*s are the independent random variables and *e* is a random error (or residual)

which is the amount of variation in y not accounted for by the linear relationship. The parameters B_i s, stand for the regression coefficients, are unknown and are to be estimated. However, there is usually substantial variation of the observed points around the fitted regression line. The deviation of a particular point from the regression line (its predicted value) is called the residual value. The smaller the variability of the residual values around the regression line, the better is model prediction.

In this study, regression analysis was performed using the training and test data employed in neural network data. Heavy metal concentrations were considered as the dependent variables and pH, SO_4 and Mg were considered as the independent variables. A computer-based package called SPSS (Statistical Package for the Social Sciences) was used to carry out the regression analysis. The estimated regression relationships for heavy metals are given as below:

$$Cu = 29.639 - 10.423 \times pH + 0.01354 \times SO_4 + 0.649 \times Mg \tag{12}$$

$$Fe = 7.796 - 0.452 \times pH + 0.001609 \times SO_4 - 0.03118 \times Mg \tag{13}$$

$$Mn = 34.451 - 5.708 \times pH + 0.004639 \times SO_4 + 0.145 \times Mg \tag{14}$$

$$Zn = 15.927 - 2.78 \times pH + 0.002376 \times SO_4 + 0.06109 \times Mg \tag{15}$$

The statistical results of the model are given in Table 5. Heavy metal concentrations were estimated according to the Eqs. 12–15. Figure 10 shows the correlation between measured heavy metal concentrations and those predicted using MLR with three inputs.

As can be seen in Fig. 10, the inappropriate predictions of the heavy metals shown by negative values is the most

Fig. 9 Comparison of the network predictions and measured concentrations for training and test data using GRNN model. **a** Correlation between GRNN Cu versus measured Cu (training data). **b** Correlation between GRNN Cu versus measured Cu (test data). **c** Correlation between GRNN Fe versus measured Fe (training data). **d** Correlation between GRNN Fe versus measured Fe (test data). **e** Correlation between GRNN Mn versus measured Mn (training data). **f** Correlation between GRNN Mn versus measured Mn (test data). **g** Correlation between GRNN Zn versus measured Zn (training data). **h** Correlation between GRNN Zn versus measured Zn (test data)

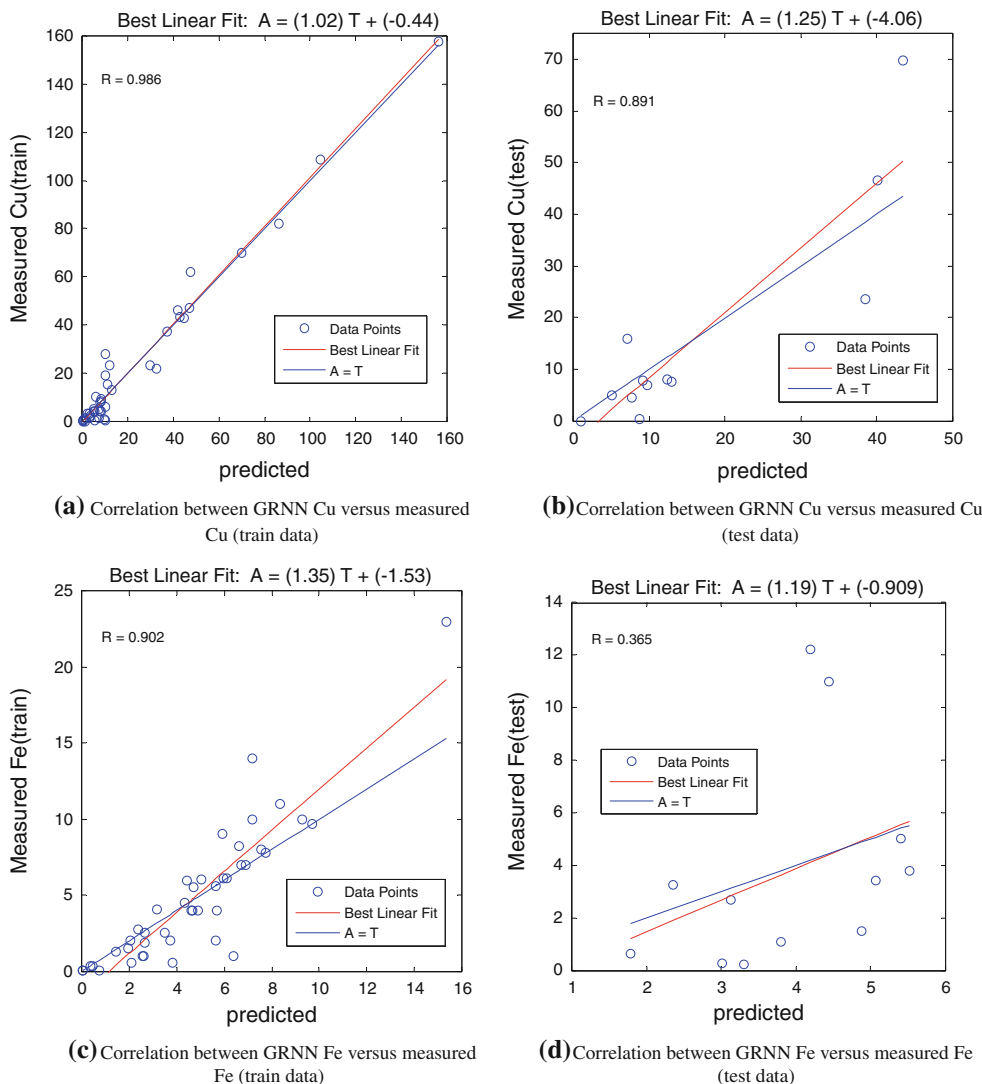
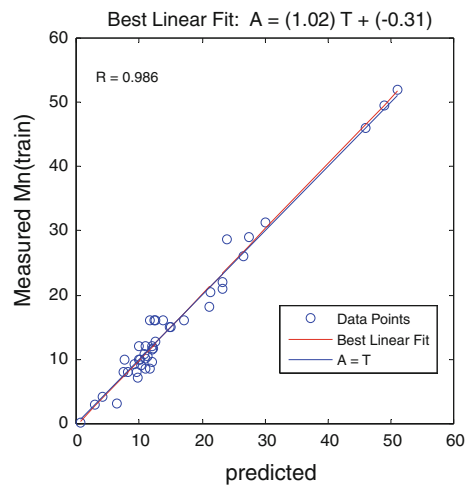
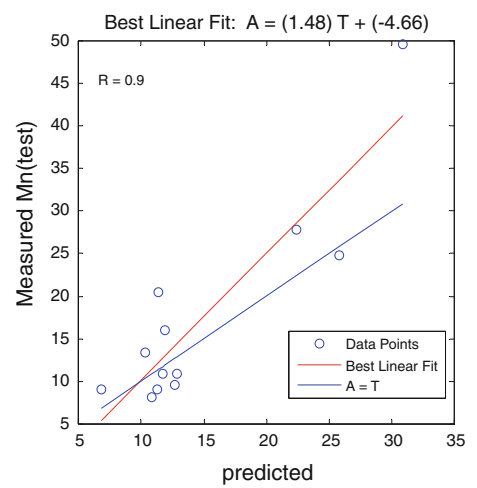


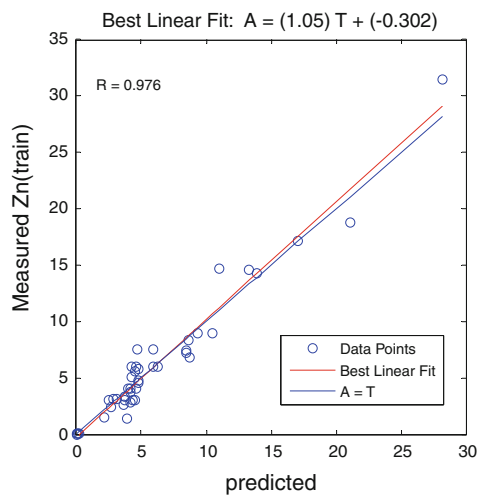
Fig. 9 continued



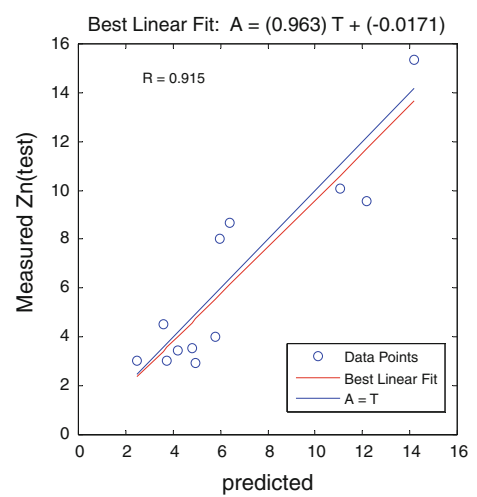
(e) Correlation between GRNN Mn versus measured Mn (train data)



(f) Correlation between GRNN Mn versus measured Mn (test data)



(g) Correlation between GRNN Zn versus measured Zn (train data)

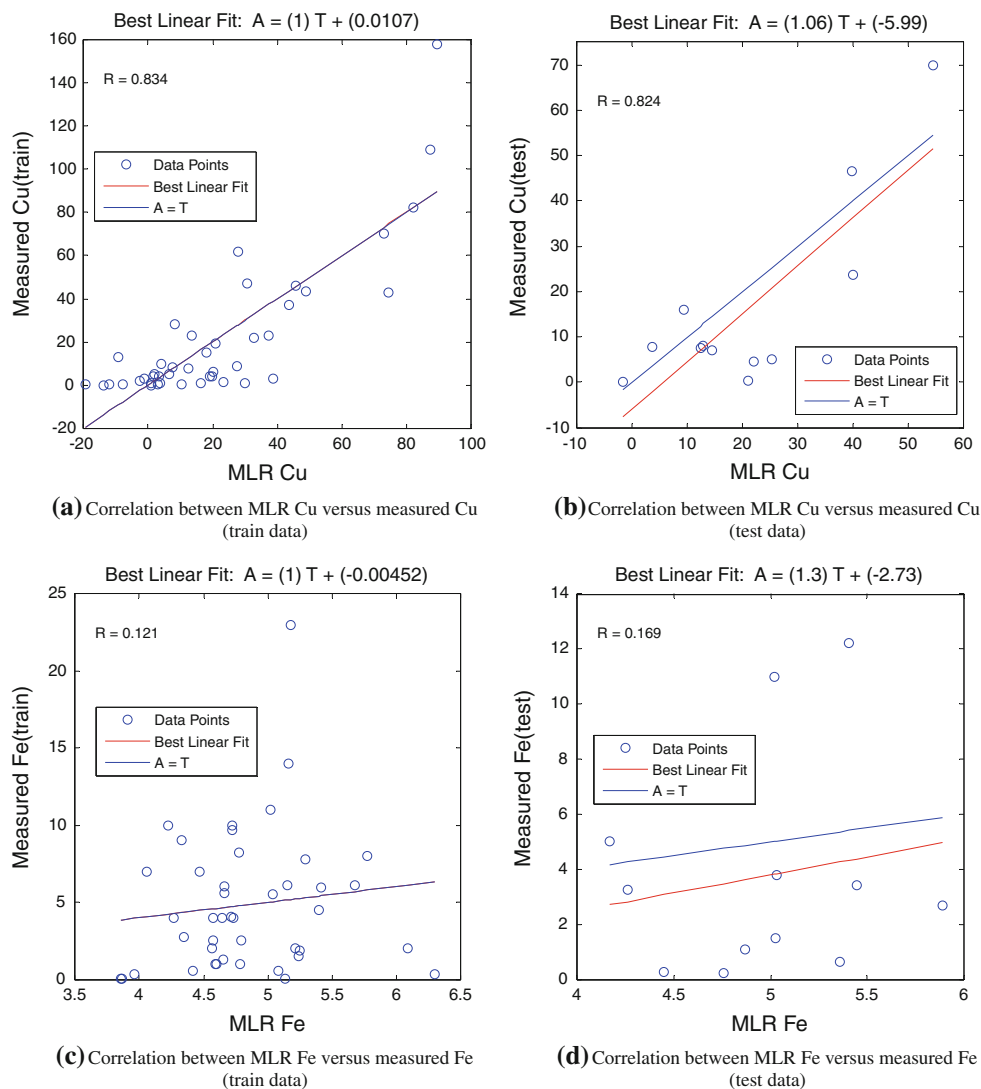


(h) Correlation between GRNN Zn versus measured Zn (test data)

Table 5 Statistical characteristics of the multiple regression models

Model	Method	Independent variables	Coefficient	Standard error	Standard error of estimate	<i>t</i> value	<i>F</i> ratio	Sig. level	Determination coefficient (<i>R</i> ²)
Eq. 12	Enter	Constant	29.639	32.646	18.87	0.908	29.718	0.370	0.696
		pH	-10.423	4.150		-2.512		0.016	
		SO ₄	0.01354	0.014		0.956		0.345	
		Mg	0.649	0.202		3.213		0.003	
Eq. 13	Enter	Constant	7.796	8.058	4.66	0.967	0.192	0.339	0.015
		pH	-0.452	1.024		-0.442		0.661	
		SO ₄	0.001609	0.003		0.460		0.648	
		Mg	-0.03118	0.050		-0.625		0.536	
Eq. 14	Enter	Constant	34.451	10.775	6.229	3.197	35.337	0.003	0.731
		pH	-5.708	1.370		-4.167		0.000	
		SO ₄	0.004639	0.005		0.992		0.327	
		Mg	0.145	0.067		2.172		0.036	
Eq. 15	Enter	Constant	15.927	6.633	3.83	2.401	20.617	0.021	0.613

Fig. 10 Comparison of the predicted concentrations using MLR and measured concentrations for training and test data. **a** Correlation between MLR Cu versus measured Cu (training data). **b** Correlation between MLR Cu versus measured Cu (test data). **c** Correlation between MLR Fe versus measured Fe (training data). **d** Correlation between MLR Fe versus measured Fe (test data). **e** Correlation between MLR Mn versus measured Mn (training data). **f** Correlation between MLR Mn versus measured Mn (test data). **g** Correlation between MLR Cu versus measured Cu (training data). **h** Correlation between MLR Zn versus measured Zn (test data)



important disadvantage of the MLR method compared to ANN method.

Table 6 compares the correlation coefficient R and root mean square error (RMS) associated with three methods for both training and test data. It is well illustrated in Table 6 that the BPNN and GRNN methods predicted some what similar results. Furthermore, a close agreement can be seen between the predicted concentrations and measured data when the ANN method (BPNN and GRNN) is used. Low correlation values between the model predictions and measured data using MLR method describes its low capability in prediction heavy metals.

In Table 6, the indices 1 and 2 for R and RMS are related to training a test data, respectively.

Conclusions

A new method to predict major heavy metals in Shur River impacted by AMD has been presented using ANN method.

The predictions for heavy metals (Cu, Fe, Mn and Zn) using ANN method incorporating BPNN and GRNN approaches together with MLR method are presented and compared with the measured data. The input data for the ANN and MLR models have been selected based on the high values of the correlation coefficients between heavy metals and pH, SO_4 and Mg^{2+} concentrations. In this paper, the BPNN model has three layers including input layer (pH, SO_4 and Mg^{2+}), hidden layer (6 neurons) with tansig activation function and output layer (Cu, Fe, Mn and Zn) with linear activation function. Whereas, the GRNN consists of three layers, i.e. input layer (pH, SO_4 and Mg^{2+}), hidden layer (44 neurons) with radbas activation function and output layer (Cu, Fe, Mn and Zn) with linear activation function. An optimal smoothing factor of 0.10 was obtained for GRNN model by a trial and error process. It was found that the BPNN and GRNN methods predicted some what similar results. Furthermore, a close agreement was achieved between the predicted and measured concentrations for heavy metals

Fig. 10 continued

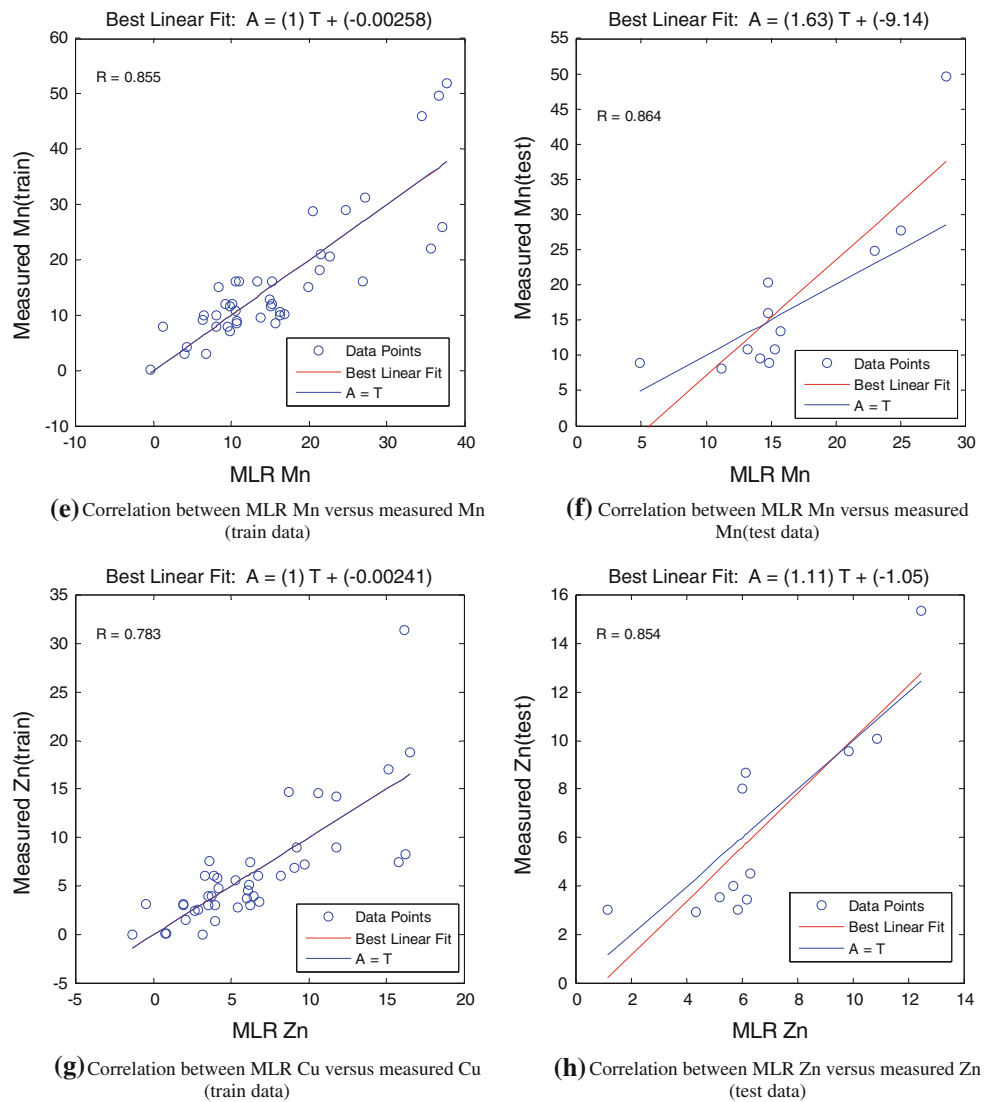


Table 6 The comparison of the results (R , RMS) of three methods in training and test data

Method	R_1 Cu	R_2 Cu	RMS ₁ Cu	RMS ₂ Cu	R_1 Fe	R_2 Fe	RMS ₁ Fe	RMS ₂ Fe	R_1 Mn	R_2 Mn	RMS ₁ Mn	RMS ₂ Mn	R_1 Zn	R_2 Zn	RMS ₁ Zn	RMS ₂ Zn
BPNN	0.97	0.92	7.68	6.50	0.35	0.22	4.10	3.45	0.97	0.92	2.90	5.18	0.95	0.92	1.91	1.31
GRNN	0.99	0.89	5.47	9.89	0.90	0.37	2.20	3.55	0.99	0.90	1.89	6.60	0.98	0.91	1.31	1.56
MLR	0.83	0.80	17.77	12.47	0.12	0.16	4.39	3.94	0.85	0.86	5.86	7.09	0.78	0.85	3.61	2.03

(Cu, Mn and Zn) when the ANN method (BPNN and GRNN) was used. However, the correlation factor was low between predicted Fe concentration and its associated measured data. Low correlation values between the model predictions and measured data using MLR method describes its low capability in prediction heavy metals.

Acknowledgments The authors thank Shahrood University of Technology and National Iranian Copper Industries Company (NIC-ICO) for their continuous support.

References

Almasri MN, Kaluarachchi JJ (2005) Modular neural networks to predict the nitrate distribution in ground water using the on-ground nitrogen loading and recharge data. Environ Model Softw 20:851–871

Bani Assadi A, Doulati Ardejani F, Karami GH, Dahr Azma B, Atash Dehghan R, Alipour M (2008) Heavy metal pollution problems in the vicinity of heap leaching of Sarcheshmeh porphyry copper mine. In: 10th international mine water association congress, 2–5 June 2008, Karlovy Vary, Czech Republic, pp 355–358

- Banisi S, Finch JA (2001) Testing a floatation column at the Sarcheshmeh copper mine. *Miner Eng* 14(7):785–789
- Canovas CR, Olias M, Nieto JM, Sarmiento AM, Ceron JC (2007) Hydrogeochemical characteristics of the Tinto and Odiel Rivers (SW Spain). Factors controlling metal contents. *Sci Total Environ* 373:363–382
- Cybenko G (1989) Approximation by superposition of a sigmoidal function. *Math Control Signals Syst* 2:303–314
- Demuth H, Beale M (2002) Neural network toolbox for use with MATLAB, user's guide version 4
- Derakhshandeh R, Alipour M (2010) Remediation of acid mine drainage by using tailings decant water as a neutralization agent in Sarcheshmeh copper mine. *Res J Environ Sci* 4(3):250–260
- Dinelli E, Lucchhini F, Fabbri M, Cortecchi G (2001) Metal distribution and environmental problems related to sulfide oxidation in the Libiola copper mine area (Ligurian Apennines, Italy). *J Geochem Explor* 74:141–152
- Doulati Ardejani F, Karami GH, Bani Assadi A, Atash Dehghan R (2008) Hydrogeochemical investigations of the Shour River and groundwater affected by acid mine drainage in Sarcheshmeh porphyry copper mine. In: 10th international mine water association congress, 2–5 June 2008, Karlovy Vary, Czech Republic, pp 235–238
- Erzin Y, Yukselen Y (2009) The use of neural networks for the prediction of zeta potential of kaolinite. *Math Geosci* 41:779–797
- Govil PK, Reddy GLN, Gnanaswara RT (1999) Environmental pollution in India: heavy metals and radiogenic elements in Nacharam Lake. *J Environ Health* 61(8):23–28
- Hagan MT, Demuth HB, Beale MH (1996) Neural neural network design. PWS Publishing, Boston, MA
- Hammack RW, de Vegt AL, Schoeneman AL (1998) The removal of sulfate and metals from mine waters using bacterial sulfate reduction: pilot plant results. *J Mine Water Environ* 17(1):8–27
- Haykin S (1994) Neural networks: a comprehensive foundation, 2nd edn. Macmillan, New York
- Herbert RB Jr (1994) Metal transport in groundwater contaminated by acid mine drainage. *Nord Hydrol* 25:193–212
- Hornik K, Stinchcombe M, White H (1989) Multilayer feed forward networks are universal approximators. *Neural Netw* 2:359–366
- Kemper T, Sommer S (2002) Estimate of heavy metal contamination in soils after a mining accident using reflectance spectroscopy. *Environ Sci Technol* (36):2742–2747
- Khandelwal M, Singh TN (2005) Prediction of mine water quality by physical parameters. *J Sci Ind Res* 64:564–570
- Lee JS, Chon HT (2006) Hydrogeochemical characteristics of acid mine drainage in the vicinity of an abandoned mine, Daduk Creek, Korea. *J Geochem Explor* 88:37–40
- Marandi R, Doulati Ardejani F, Marandi A (2007) Biotreatment of acid mine drainage using sequencing batch reactors (SBRs) in the Sarcheshmeh porphyry copper mine. In: Cidu R, Frau F (eds) IMWA symposium 2007: water in mining environments, 27–31 May 2007, Cagliari, Italy, pp 221–225
- Merrington G, Alloway BJ (1993) Leaching characteristics of heavy metals from three historical Pb–Zn mine tailings heaps in the United Kingdom. *Min Ind* 102:71–152
- Moncur MC, Ptacek CJ, Blowes DW, Jambor JL (2005) Release, transport and attenuation of metals from an old tailings impoundment. *Appl Geochem* 20:639–659
- Noori R, Abdoli MA, Jalili Ghazizade M, Samieifard R (2009) Comparison of neural network and principal component-regression analysis to predict the solid waste generation in Tehran. *Iran J Public Health* v38:74–84
- Noori R, Khakpour A, Omidvar B, Farokhnia A (2010) Comparison of ANN and principal component analysis-multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistic. *Expert Syst Appl* 37:5856–5862
- Shahabpour J, Doorandish M (2008) Mine drainage water from the Sarcheshmeh porphyry copper mine, Kerman, IR Iran. *Environ Monit Assess* 141:105–120
- Singh KP, Basant A, Malik A, Jain G (2009) Artificial neural network modeling of the river water quality—a case study. *Ecol Model* 220:888–895
- Smuda J, Dold B, Friese K, Morgenstern P, Glaesser W (2007) Mineralogical and geochemical study of element mobility at the sulfide-rich Excelsior waste rock dump from the polymetallic Zn–Pb–(Ag–Bi–Cu) deposit, Cerro de Pasco, Peru. *J Geochem Explor* 92:97–110
- Specht DF (1991) A general regression neural network. *IEEE Trans Neural Netw* 2(6):568–576
- Waterman GC, Hamilton RL (1975) The Sarcheshmeh porphyry copper deposit. *Econ Geol* 70:568–576
- Williams RE (1975) Waste production and disposal in mining, milling, and metallurgical industries. Miller-Freeman, San Francisco, p 489
- Wilson B, Lang B, Pyatt FB (2005) The dispersion of heavy metals in the vicinity of Britannia Mine, British Columbia, Canada. *Ecotoxicol Environ Saf* 60:269–276
- Zhao F, Cong Z, Sun H, Ren D (2007) The geochemistry of rare earth elements (REE) in acid mine drainage from the Sitai coal mine, Shanxi Province, North China. *Int J Coal Geol* 70:184–192