Hybrid ANN with Invasive Weed Optimization Algorithm, a New Technique for Prediction of Gold and Silver in Zarshuran Gold Deposit, Iran

Habibollah Bazdar[1], Hadi Fattahi[1, *], Feridon Ghadimi[1]

1- Department of Mining Engineering, Arak University of Technology, Arak, Iran.

* Corresponding Author: H.fattahi@arakut.ac.ir

Abstract

Because of high cost of drilling and analysis of samples, it needs to predict gold and silvers based on pathfinder such as As, Sb, Cd, Pb and Zn and decrease the cost and time exploration project implementation. In this paper, the model based on a multilayer perceptron artificial neural network (MLP-ANN) optimized by invasive weed optimization algorithm (IWO) to predict of gold and silver in Zarshuran gold deposit, Iran is proposed. The IWO is used to decide the initial weights of the ANN. The ANN-IWO model is trained using an experimental data set to approximate the relation between Sb, Cd, Pb and Zn as inputs and gold and silver as output. Furthermore, the performance of the ANN-IWO model is compared with multiple linear regression (MLR). The results obtained indicate that the ANN-IWO model has strong potential to prediction of gold and silver with high degree of accuracy and robustness.

Keywords: Gold and silver; Artificial neural network; Invasive weed optimization algorithm; Multiple linear regression.

1- Introduction

Gold is the main mineral resource in most countries because of its role in economic, industries and public culture. So, most exploration project in gold thread defined. The most important test that defines the presence or absence of mineralization is drilling. High cost of drilling and analysis of samples provide a model that has good agreement with reality. Sometimes it needs to decrease the cost and time exploration project implementation. Mineral prospectively mapping (MPM) and prediction of ore grade is main contributor to this work. Statistical methods such as kriging interpolation, weight of evidence and logistic regression has been applied in this field. Soft computing is also used as modern modeling in gold exploration projects. Skabar applied MPM to mapping gold mineralization potential in the Castlemaine region of Victoria, Australia, and results are compared with a method based on estimating probability density functions and using a range of geological, geophysical and geochemical input variables. Harris and Pan used probabilistic neural network (PNN) to classify mineralized and nonmineralized cells using eight geological, geochemical, and geophysical variables. This investigation demonstrates that a probabilistic neural network can be a useful tool for mineral favorability mapping. A PNN trained and calibrated well on training data, indicating that it could distinguish mineralized cells from those that are barren by a set of eight geological, geochemical, and geophysical variables (Harris and Pan 1999; Harris *et al.*, 2003). Singer and Kouda used the ability of a probabilistic neural network to classify deposits into types based on a simple representation of mineralogy and six broad rock types. They concluded the comparison of the spatial distribution of the neural network's estimated deposit classes and permissive tracts determined by experts shows that the probabilistic neural network is able is perform well at generalization (Singer and Kouda 1996)

Leite *et al*. presented application of probabilistic neural networks to map the potential for platinum group elements (PGE) mineralization sites in the northeast portion of the caraja´s mineral province (CMP), Brazilian amazon. They concluded mineral potential maps provided by this method can be used as reconnaissance guides for future detailed ground surveys of possible new PGE occurrences, which is of critical importance to shorten exploration time and costs in such densely forested Amazonian terrains (Leite and de Souza Filho 2009a; Leite and de Souza Filho 2009b). Samanta *et al*. was performed comparative evaluation of various local and global learning algorithms in neural network modeling for ore grade estimation in three deposits: gold, bauxite, and iron ore. Four local learning algorithms, standard back-propagation, back-propagation with momentum, quick prop back-propagation, and Levenberg–Marquardt back-propagation, along with two global learning algorithms, novel and simulated annealing, were investigated for this purpose. The study revealed it is better to apply global learning algorithms in neural network training since many real-life applications of neural network modeling show local minima problems in error surface (Samanta *et al.,* 2002; Samanta *et al.,* 2004). Another intelligence method such as support vector regression (SVR) (Li *et al.,* 2013; Li *et al.,* 2010), neuro-fuzzy (Asadi *et al.,* 2014; Brown *et al.,* 2003), also used for MPM and grade estimation of gold and relevant elements.

During the last years, artificial neural network (ANN) modeling was used often in different separation and technological applications, mainly due to its powerfulness for solving complex multiple regression problems. ANN's ability for mapping non-linear relationships between the inputs and outputs of a system has extended the field of applications of ANN modeling (Simpson 1990).The main goal of this paper is to extend the ANN's approach in order to finding appropriate initial values for optimization main parameters such as weighting and Biases. Some evolutionary algorithms such as particle swarm optimization (Price and Bauer 1985) and genetic algorithm (Hajihassani *et al.,* 2014) can be used for this determination. Recently, a novel numerical stochastic optimization method inspired by colonizing weed has been introduced for optimizing antenna problems. This algorithm has been first used by Mehrabian and Lucas (2006) in dynamic and control systems theory. They have named the algorithm invasive weed optimization (IWO). In the present work, we propose IWO for optimizing the weights of feed-forward neural network. Then simulation results demonstrate the effectiveness and potential of the new proposed network for gold and silver prediction compared with multiple linear regression (MLR) using the same data.

2- Materials and Methods

2.1- Geological setting

Zarshuran deposit is located in northwestern Iran, and 8 km northeast of Zarshuran village (Fig. 1).The Zarshuran area consists mainly of Precambrian rocks (Samimi 1992). The oldest unit, the Iman Khan unit, forms the core of the Iman Khan anticline, symmetrically plunging both NW and SE over some 7 km (Mehrabi *et al.,* 1999).The Iman Khan unit mainly consists of chlorite–amphibole– schist and locally serpentinite. The Iman Khan schist is followed stratigraphically upwards by the Chaldagh limestone, the Zarshuran black shale with silica and carbonate intercalations and the Qaradash shale, tuff and sandstone (Mohajer *et al.,* 1989). An Oligocene to Miocene granitoid, intruded into the mineralized Precambrian formations and is highly altered, mylonitized and weakly mineralized (Fig. 1).The main faults present trend northeast-southwest and northwest-southeast in the Precambrian formation. Gold occurs mainly as disseminations in carbonaceous, siliceous, and calcareous beds

within the Zarshuran black shale. Gold is also found in hydrothermal veins of massive quartz (jasperoid) and quartz veinlets formed by carbonate replacement along high-angle faults in Chaldagh limestone.

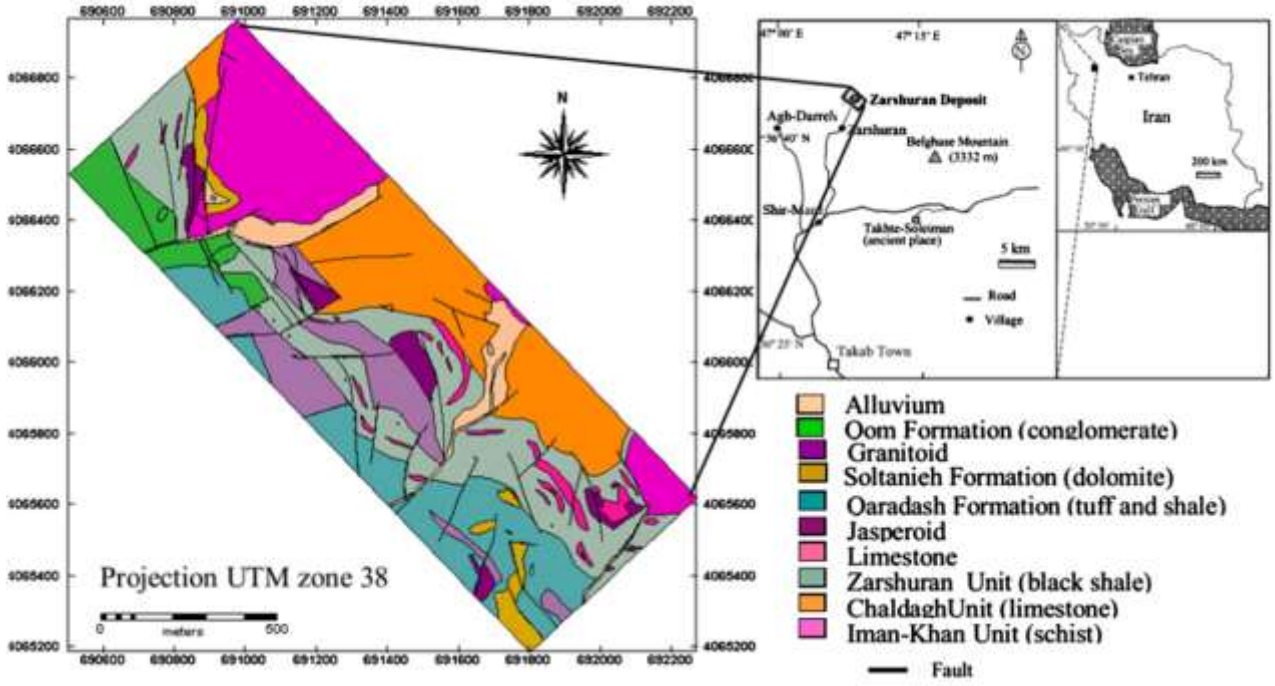

Figure 1) Location and detailed geological map of the Zarshuran gold mining district, NW Iran (Asadi 2000).

The Chaldagh limestone is also mineralized at its contacts with the Iman Khan schist and with the overlying Zarshuran unit. Quartz, calcite, dolomite and clays are the principal minerals of the host rocks. Decalcification, silicification, and argillization characterize mineralized rocks. Decalcification increased the porosity and permeability of the host rocks and thus provided favorable sites for hydrothermal mineralization. The ore consists mainly of orpiment and pyrite, and to a lesser extent sphalerite, galena, realgar and stibnite, with subordinate cinnabar, HgS, lorandite, $TlAsS_2$, christite, $TlHgAsS_3$, coloradoite, HgTe, getchellite, $AsSbS_3$, aktashite, $Cu_6Hg_3As_4S_{12}$, baumhuerite, $Pb_3As_4S_9$, boulangerite, $Pb_5Sb_4S_{11}$, geochronite, $Pb_{14}(Sb,As)_6S_{23}$, plagionite, $Pb_5Sb_8S_{17}$ and twinnite, $Pb(Sb,As)_2S_4$. Sulfide oxidation is mainly confined to veins, veinlets, and fracture zones. Quartz, calcite, fluorite, hematite and barite are the main gangue minerals. Accessory minerals include apatite, $Ca_5(PO_4)_3$ (OH, F, Cl), rutile, TiO_2, zircon, $ZrSiO_2$, and xenotime, YPO_4 (Asadi, 2000). Gold is rarely visible but occurs invisibly in arsenian pyrite and sphalerite (Asadi *et al.*, 1999). The petrography, mineralogy and trace element geochemistry of the Zarshuran gold deposit show that it is a Carlin-like sediment-hosted disseminated gold deposit. The petrography, mineralogy and trace element geochemistry of the Zarshuran gold deposit show that it is a Carlin-like sediment-hosted disseminated gold deposit. The association of mineralization at Zarshuran with a magmatic intrusion, and the presence of tellurium in concentrations sufficient to precipitate telluride, suggest a greater magmatic component in the mineralizing hydrothermal solution than is typical for most Carlin-type deposits, best known in the western United States (Asadi 2000).

2.2-Sample Collection and Analysis

Sampling was carried out on two original trenches that conducted in the square form sides of each square was 2 meters and approximately 3 kg per square was removed as representative sample. The total number of samples was 108

samples that 56 samples from the first trench and the 52 samples from second trench were taken. Length of first trench approximately 125 m and height of it varies from 3 to 6 meters. The color of the trench is yellow and black and orange color. Contact of schist and limestone is visible (Fig. 2).

Figure 2) First trench that contact of schist and limestone is visible.

Figure 3) Second trench in sampling area.

Second trench is about 105 meters long and its height varied from 5 to 10 m and containing of limestone, silica, fluorite crystals and jasperoid and its color combination of yellow and orange and black have been less (Fig. 3).

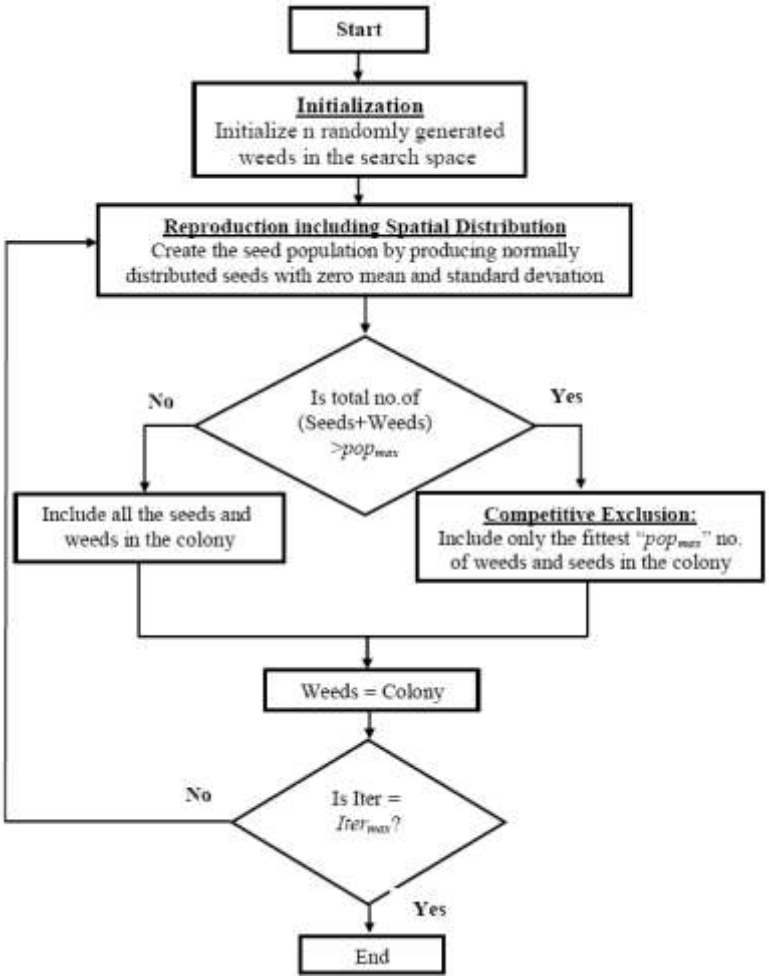

Figure 4) A Flow-Chart Representation of the IWO Algorithm.

Samples after preparation were sent to laboratory for analysis. Samples were analyzed by inductively coupled plasma optical emission spectrometry (ICP-OES) for Au, Ag, Al, As, Ba, Be, Bi, Ca, Cd, Ce, Co, Cr, Cu, Fe, K, Li, La, Mg, Mn, Mo, Na, Ni, P, Pb, S, Sb, Sc, Sr, Th, Ti, U, V, Y, Yb, Zn and Zr. These elements were selected as direct and pathfinder indicator for gold mineralization in an epithermal setting.

3- Hybrid ANN- Invasive Weed Optimization Algorithm

3.1- Artificial Neural Network

ANN firstly was introduced by McCulloch and Pitts (1943) who presented ability of this technique to calculate any logical functions. Processing of the information is executed with the help of many interconnected simple elements known as neurons which are placed in distinct layers of the network. Multi-layer perceptron, the most famous type of ANNs, consists of at least three layers: input, intermediate or hidden layers and output. Difficulty level of the problem determines the number of the hidden layers and neurons (Simpson 1990). The neurons are linked from a layer to the next one, but this connection is not within the same layer. Once a series of inputs presents to the network, the input values are transmitted through the links to the second layer. In every link, the transmitted value is multiplied to the weight of the link. The weighted values are come together at a node in the hidden layer and a bias is summed to the weighted values in that particular node. Consequently, the achieved value transfer to an activation function and a signal is created. Using the departing links of hidden nodes, the results are transmitted to the output layer. Similar to hidden nodes, the input values of the output nodes are weighted, biased, summed and transferred to the activation function. The created values of activation functions in output layers are the outputs of the network. Performance of an ANN is dependent on architecture of the network which is the pattern of the connections existing between the neurons. The network should be trained with sufficient input–output patterns that are known as the training data (Meulenkamp and Grima 1999). As the error reached specified error goal, training is finished and the optimum model is determined.

3.2-Invasive Weed Optimization Algorithm

In 2006, Mehrabian and Lucas proposed the IWO (Mehrabian and Lucas 2006), a derivative-free, metaheuristic algorithm, mimicking the ecological behavior of colonizing weeds. The algorithm for IWO may be summarized as follows:

a) Initialization: A population of initial solutions is being dispread over the d dimensional problem space with random positions.

b) Reproduction: Each member of the population is allowed to produce seeds depending on its own, as well as the colony's lowest and highest fitness, such that, the number of seeds produced by a weed increases linearly from lowest possible seed for a weed with worst fitness to the maximum no. of seeds for a plant with best fitness.

c) Spatial distribution: The generated seeds are being randomly distributed over the d dimensional search space by normally distributed random numbers with mean equal to zero; but varying variance. This step ensures that the produced seeds will be generated around the parent weed, leading to a local search around each plant. However, the standard deviation (SD) of the random function is made to decrease over the iterations.

If sd_{max} and sd_{min} be the maximum and minimum standard deviation and if *pow*be a real no. , then the standard deviation for a particular iteration may be given as in equation (1):

$$sd_{ITER} = \left(\frac{iter_{\max} - iter}{iter_{\max}} \right)^{pow} \left(sd_{\max} - sd_{\min} \right) + sd_{\min} \quad \text{Eq.1}$$

This step ensures that the probability of dropping a seed in a distant area decreases nonlinearly with iterations, which results in grouping fitter plants and elimination of inappropriate plants. Therefore, this is a selection mechanism of IWO.

d) Competitive Exclusion: If a plant leaves no offspring then it would go extinct, otherwise they would take over the world. Thus, there is a need of some kind of competition between plants for limiting maximum number of plants in a colony. Initially, the plants in a colony will reproduce fast and all the produced plants will be included in the colony, until the number of plants in the colony reaches a maximum value pop_{max}. However, it is expected that by this timethe fitter plants have reproduced more than undesirable plants. From then on, only the fittest plants, among the existing ones and the reproduced ones; are taken in the colony and the steps 1 to 4 are repeated until the maximum number of iterations has been reached, *i.e.* the colony size is fixed from thereon to pop_{max}. This method is known as competitive exclusion and is also a selection procedure of IWO. A flowchart of the whole optimization process has been shown in Fig 4. For more details about the IWO see (Mehrabian and Lucas 2006; Rad and Lucas 2007; Roy *et al.,* 2011;Xing and Gao 2014).

In this paper, IWO is used to decide the initial weights of the ANN.

4- Prediction of Gold and Silver Using Hybrid ANN-IWO Model

To prediction of gold and silver using hybrid ANN-IWO model, all relevant parameters should be determined, due to the fact that ANNs work based on given data and do not have previous knowledge about the subject of prediction. Following sections describe the inputs and output parameters and prediction of gold and silver using hybrid ANN-IWO model.

4.1- Inputs and Outputs Data

In ANN-IWO modeling, any type of input can be used as long as they have effects on output results. A dataset that includes 108 experimental data sets was employed in current study, while 86 data points (80%) were utilized for constructing the model and the remainder data points (22 data points) were utilized for model performance evaluation. According to the correlation matrix (Table 1), five input parameters including As, Sb, Cd, Pb and Zn output including gold and silver were used. Partial dataset used in this study for constructing the in ANN-IWO model are shown in Table 2. Also, descriptive statistics of the data sets used for modeling are shown in Table 3.

Table 1) Correlation matrix between gold, silver and independent variables.

	Ag	Al	As	Ba	Ca	Cd	Cr	Pb	S	Zn	Sb	Au
Ag	1											
Al	-0.321	1										
As	0.764	-0.221	1									
Ba	0.503	0.353	0.445	1								
Ca	-0.156	0.375	0.008	-0.468	1							
Cd	0.765	0.221	0.999	0.446	0.01	1						
Cr	0.083	0.695	0.075	0.483	-0.549	0.074	1					
Pb	0.874	-0.124	0.781	0.593	/0.141	0.783	0.164	1				
S	0.302	0.337	0.136	0.421	-0.353	0.136	0.297	0.315	1			
Zn	0.589	0.1766	0.68	0.614	-0.218	0.683	0.475	0.71	0.099	1		
Sb	0.792	-0.081	0.864	0.546	-0.093	0.866	0.162	0.883	0.355	0.719	1	
Au	0.655	-0.115	0.774	0.518	-0.094	0.748	0.12	0.689	0.143	0.707	0.752	1

4.2- Pre-Processing of Data

In data-driven system modeling methods, some pre-processing steps are usually implemented prior to any calculations, to eliminate any outliers, missing values or bad data. This step confirms that the raw data retrieved from database is perfectly proper for modeling. In order to softening the training procedure and improving the accuracy of prediction, all data samples are normalized to adapt to the interval [-1,1] according to the following linear mapping function:

$$x_M = 2\left(\frac{x - x_{\min}}{x_{\max} - x_{\min}}\right) - 1 \qquad \text{Eq.2}$$

Where x is the original value from the dataset, x_M is the mapped value, and x_{min} (x_{max}) denotes the minimum (maximum) raw input values, respectively. It is to be noted that model outputs will be remapped to their corresponding real values by the inverse mapping function ahead of calculating any performance criterion.

Table 2) Partial dataset used in this study for constructing the in ANN-IWO model

Sample No.	Input parameters				Output parameters	
	Cd (ppm)	Sb (ppm)	Pb (ppm)	Zn (ppm)	Au (ppb)	Ag (ppm)
1	6.4	1.12	30	53	6	0.22
2	12.8	1.07	28	24	7	0.22
3	10.8	1.08	25	34	7	0.2
4	9	1.22	38	51	8	0.23
5	2.8	1.14	86	48	23	0.21
6	0.7	1.13	80	20	5	0.21
7	16.4	11.6	41	23	5	0.22
8	10.9	1.19	27	9	691	0.2
9	44.3	8	29	0.75	37	0.21
10	7.3	1.1	13	0.75	20	0.21

Table 2) 3Descriptive statistics of the data sets.

Parameter	Min	Max	Mean
Au(ppb)	4	1458	121.75
Ag(ppm)	0.2	8.5	1.7906
As(ppm)	29.2	10975.7	2291.472
Cd(ppm)	0.7	366.6	76.5407
Sb(ppm)	0.95	207.9	56.5476
Pb(ppm)	7	2675	664.3519
Zn(ppm)	0.75	4100	820.456

4.3- Tuning Parameters for the IWO

To develop an accurate ANN model, the training, and validation processes are the important steps. In the training process, a set of input-output patterns is repeated to the ANN. From that, weights of all the interconnections between neurons are adjusted until the specified input yields the desired output. Through these activities, the ANN learns the correct input-output response behavior. The model training stage includes choosing a criterion of fit (Root mean squared error) and an iterative search algorithm to find the network parameters that minimizes the criterion. The IWO was used in an effort to formalize a systematic approach to training the ANN, and to insure creation of a valid model. It was used to perform global search algorithms to update the weights and biases of neural network. The control parameters used for running the IWO shown in Table 4.

Table 4) The control parameters used for running the IWO.

Parameter	Value
Number of initial seeds (N_{Weed})	100
Maximum of algorithm iteration ($iter_{max}$)	100
Maximum of weeds (P_{max})	100
Maximum of seeds around each weed (S_{max})	5
Minimum of seeds around each weed (S_{min})	1
Maximum standard deviation (sd_{max})	1
Minimum standard deviation (sd_{min})	0.001
Nonlinear ratio (pow)	2

4.4-Network Design

In this paper, after building several ANN-IWO models based on trial and error, the best result of each model, listed in Tables 5 and 6, are compared and the one with maximum of R^2 and minimum of MSE is chosen. The best ANN-IWO architecture for prediction ofgold and silver was: 5-5-2 (5 input units, 5 hidden neurons, 2 output neuron). IWO is used as neural network optimization algorithm and the RMSE used as a cost function in this algorithm. The goal in proposed algorithm is minimizing this cost function. Figure 5 shows architecture of best ANN-IWO model for prediction of gold and silver in Zarshuran gold deposit, Iran. Also, in this study, LogSig and TanSig were used as transfer function between input and hidden layer, as well as were used as transfer function between hidden and output layer (Tables 5 and 6), shown by the following equations:

$$\mathrm{TanSig} = \frac{2}{(1+\exp(-2x))} - 1 \qquad \text{Eq.3}$$

$$Log\,\mathrm{Sig} = \frac{1}{1+\exp(-x)} \qquad \text{Eq.4}$$

Table 5) Comparison between the best results of some ANN-IWO models for training data set.

Structure	Transfer Functions	Au		Ag	
		MSE	R^2	MSE	R^2
5-6-6-6-2	TanSig- TanSig- TanSig- LogSig	0.5255	0.0005	0.5233	0.0007
5-6-5-6-2	LogSig- TanSig- LogSig- Purelin	0.4212	0.0295	0.3516	0.3050
5-5-4-2	LogSig- LogSig- LogSig	0.4436	0.5536	0.3256	0.8456
5-3-7-2	LogSig- LogSig- TanSig	0.1554	0.6554	0.0752	0.9014
5-2-2	LogSig- TanSig	0.1141	0.6175	0.0358	0.9372
5-3-2	LogSig- TanSig	0.1021	0.6309	0.0256	0.9567
5-4-2	LogSig- TanSig	0.1177	0.6659	0.0191	0.9681
5-6-2	LogSig- TanSig	0.1193	0.6605	0.0280	0.9571
5-7-2	LogSig- TanSig	0.1555	0.6729	0.0285	0.9573
5-5-2	TanSig- TanSig	0.1271	0.6621	0.0468	0.6463
5-5-2	**LogSig- TanSig**	**0.0915**	**0.6920**	**0.0268**	**0.9553**

Table 6) Comparison between the best results of some ANN-IWO models for testing data set

Structure	Transfer Functions	Au		Ag	
		MSE	R^2	MSE	R^2
5-6-6-6-2	TanSig- TanSig- TanSig- LogSig	0.5326	0.0008	0.5112	0.0010
5-6-5-6-2	LogSig- TanSig- LogSig- Purelin	0.3256	0.0518	0.2445	0.0336
5-5-4-2	LogSig- LogSig- LogSig	0.4256	0.5456	0.3125	0.8354
5-3-7-2	LogSig- LogSig- TanSig	0.1259	0.6825	0.0578	0.9214
5-2-2	LogSig- TanSig	0.1247	0.6442	0.0260	0.6567
5-3-2	LogSig- TanSig	0.0695	0.6912	0.0195	0.9688
5-4-2	LogSig- TanSig	0.1540	0.6962	0.0179	0.9720
5-6-2	LogSig- TanSig	0.0961	0.7578	0.0223	0.9657
5-7-2	LogSig- TanSig	0.1258	0.7208	0.0184	0.9734
5-5-2	TanSig- TanSig	0.1593	0.6370	0.0582	0.9413
5-5-2	**LogSig- TanSig**	**0.0879**	**0.7368**	**0.0144**	**0.9815**

The simulation performance of the ANN-IWO model was evaluated on the basis of mean squared error (MSE), variance account for (VAF), root mean squared error (RMSE), squared correlation coefficient (R^2) and mean absolute percentage error (MAPE), shown by the following equations:

$$MSE = \frac{1}{n}\sum_{k=1}^{n}(t_k - \hat{t}_k)^2 \quad \text{Eq.5}$$

$$VAF = \left(1 - \frac{\text{var}(t_k - \hat{t}_k)}{\text{var}(t_k)}\right) \quad \text{Eq.6}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(t_k - \hat{t}_k)^2} \quad \text{Eq.7}$$

$$R^2 = 1 - \frac{\sum_{k=1}^{n}(t_k - \hat{t}_k)^2}{\sum_{k=1}^{n} t_k^2 - \frac{\sum_{i=1}^{n}\hat{t}_k^2}{n}} \quad \text{Eq.8}$$

$$MAPE = \frac{1}{n}\sum_{k=1}^{n}\left|\frac{t_k - \hat{t}_k}{t_k}\right| \times 100 \quad \text{Eq.9}$$

where t_k be the actual value and $\hat{t}_k$ be the predicted value of the k^{th} observation, n be the number of observations.

A comparison between predicted values of gold and silver by the ANN-IWO model and measured values for 108 data sets at training and testing phases is shown in Figs. 6 and 7. As shown in Figs. 6 and 7, the results of the ANN-IWO model in comparison with actual data show a good precision of the ANN-IWO model (see Table 7).

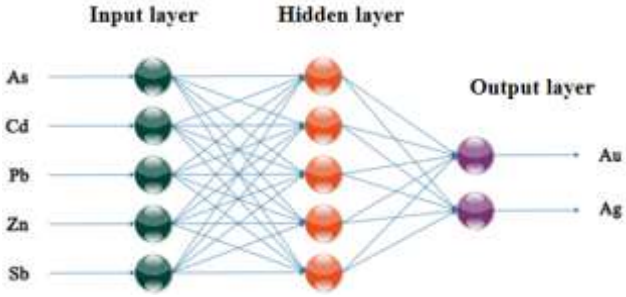

Figure 5) Architecture of best ANN-IWO model for prediction of gold and silver.

In addition, performance analysis of the ANN-IWO model for predicting gold and silveris shown in Table 7. The performance indices obtained in Table 7 indicate the high performance of the ANN-IWO model that can be used successfully for the prediction of gold and silver. Furthermore, correlation between measured and predicted values of gold and silver for training and testing phases are shown in Figures 8 and 9.

Table 7) Performance of the model for predicting gold and silver.

Data Set		R^2	MSE	RMSE	VAF	MAPE
Au	Train Data Set	0.6920	0.0915	0.3024	0.6658	549.0
	Test Data Set	0.7368	0.0879	0.2965	0.5255	125.9
Ag	Train Data Set	0.9553	0.0268	0.1636	0.9537	377.5
	Test Data Set	0.9815	0.0144	0.1199	0.9814	60.1

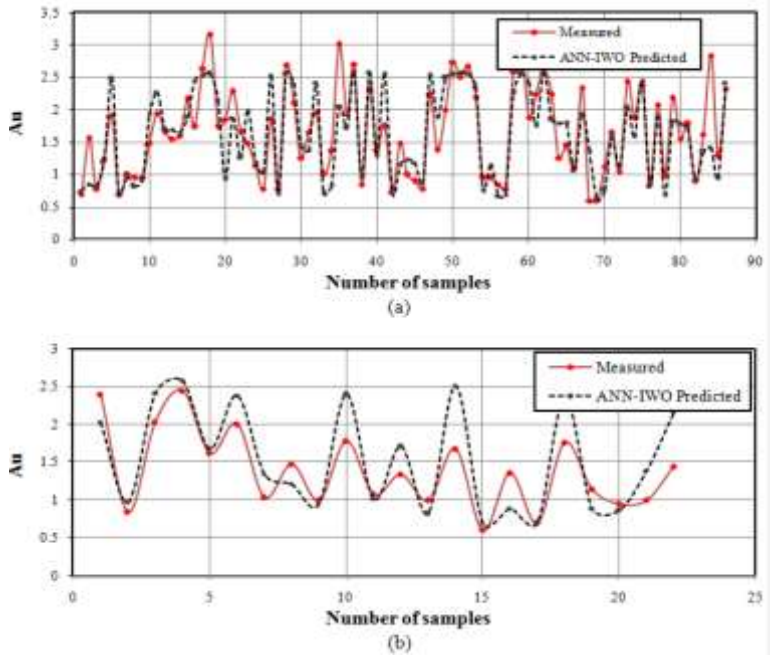

Figure 6) Comparison between measured and predicted gold a) trainingdatasets b) testing datasets.

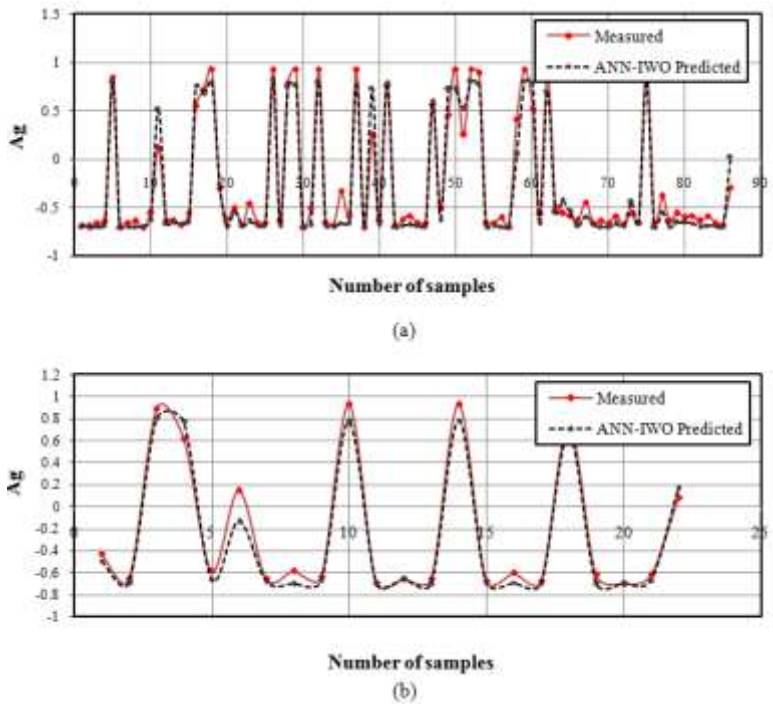

Figure 7) Comparison between measured and predicted silver a) trainingdatasets b) testing datasets.

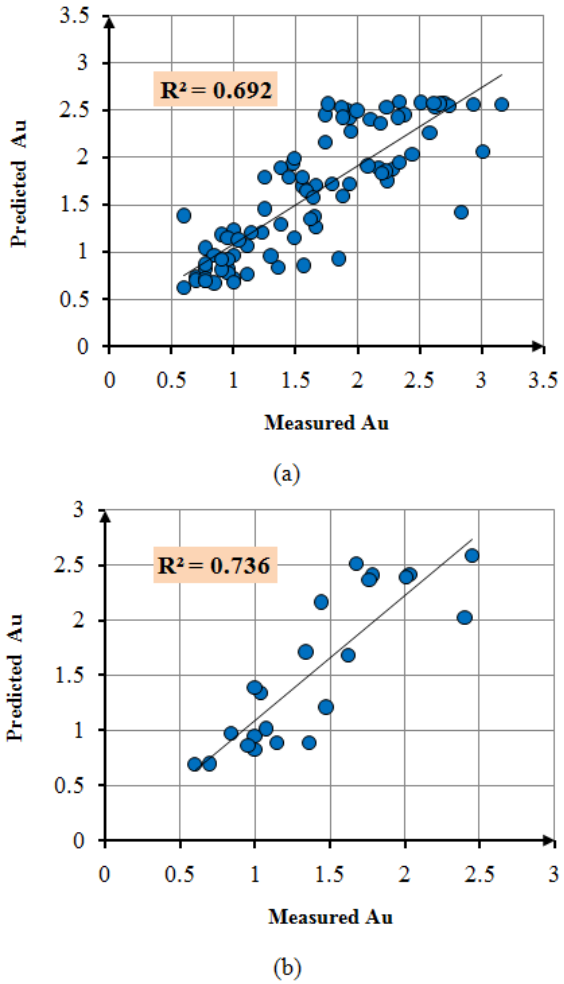

Figure 8) Correlation between measured and predicted values of gold a) training datasets b) testing datasets.

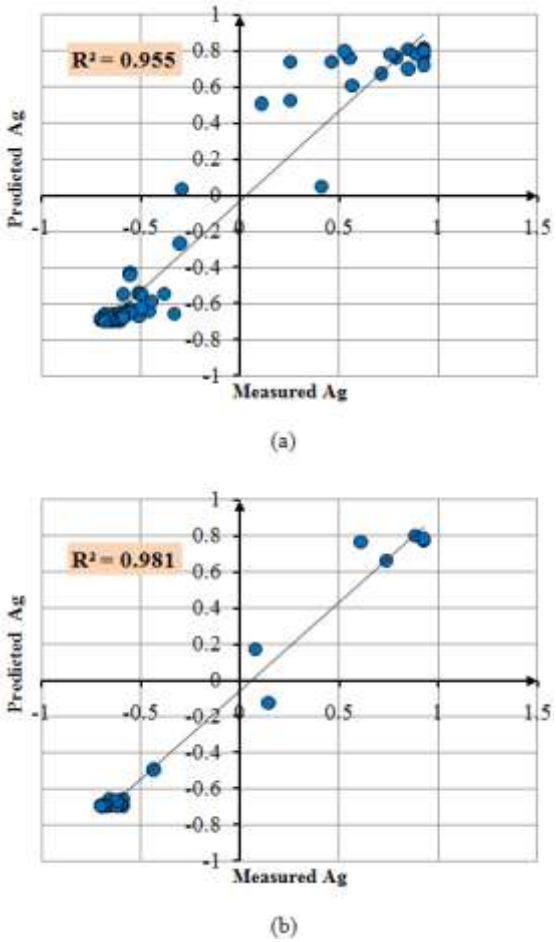

Figure 9) Correlation between measured and predicted values of silver a) training datasets b) testing datasets.

5- 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:

$$y = C_1 + C_2 x_2 + ... + C_n x_n + e \quad \text{Eq.10}$$

Where *y* is the dependent variable, *x* is are the independent random variables and *e* is a random error which is the amount of variation in *y* not accounted for by the linear relationship. The parameters C is, 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.

Table 8) Statistical characteristics of the multiple regression models.

Model	Method	Independent variables	Coefficient	Standard error	Standard error of estimate	t value	F ratio	Sig. level	Determination coefficient (R^2)
Eq. a	Enter	Constant	5.424	4.0033	0.41217	1.345	29.228	0.182	0.646
		As	-3.297	2.669		-1.235		0.220	
		Sb	0.173	0.147		1.181		0.241	
		Cd	3.656	2.659		1.375		0.173	
		Pb	-0.020	0.116		-0.171		0.856	
		Zn	0.186	0.064		2.917		0.005	
Eq. b	Enter	Constant	-3.445	3.074	0.31417	-1.096	47.328	0.227	0.764
		As	1.097	2.034		0.542		0.589	
		Sb	-0.056	0.112		-0.149		0.882	
		Cd	-0.787	2.027		-0.383		0.703	
		Pb	0.600	0.089		6.056		0.000	
		Zn	-0.066	0.049		-1.479		0.143	

Table 9) The comparison of the results (R2, MSE) of two methods in training and testing data.

Data set		R^2		MSE	
		ANN-IWO	MLR	ANN-IWO	MLR
Au	Train data set	0.6916	0.6463	0.0915	0.1580
	Test data set	0.7369	0.7602	0.0879	0.1122
Ag	Train data set	0.9553	0.7646	0.0268	0.0855
	Test data set	0.9815	0.8859	0.0144	0.0479

In this paper, regression analysis was performed using the training and test data employed in ANN-IWO model. Gold and silver was considered as the dependent variable and As,Sb, Cd, Pb and Zn 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 gold and silver are given as below:

$$\text{Au} = 5.424 - 3.297\ \text{As} + 0.173\ \text{Sb} + 3.656\ \text{Cd} - 0.020\ \text{Pb} + 0.186\ \text{Zn} \quad \text{(a)}$$

$$\text{Ag} = -3.445 + 1.097\ \text{As} - 0.056\ \text{Sb} - 0.787\ \text{Cd} + 0.600\ \text{Pb} - 0.066\ \text{Zn} \quad \text{(b)}$$

The statistical results of the model are given in Table 8. Gold and silver were estimated according to the Eqs. a and b. Figures 10 and 11 shows the correlation between measured gold and silver and those predicted using MLR with five inputs.

Table 9 compares the correlation coefficient R^2 and mean square error (MSE) associated with two methods for both training and test data. It is well illustrated in Table 9 that a close agreement can be seen between the predicted and measured data when the ANN-IWO method is used. Low

correlation values between the model predictions and measured data using MLR method describes its low capability in prediction gold and silver.

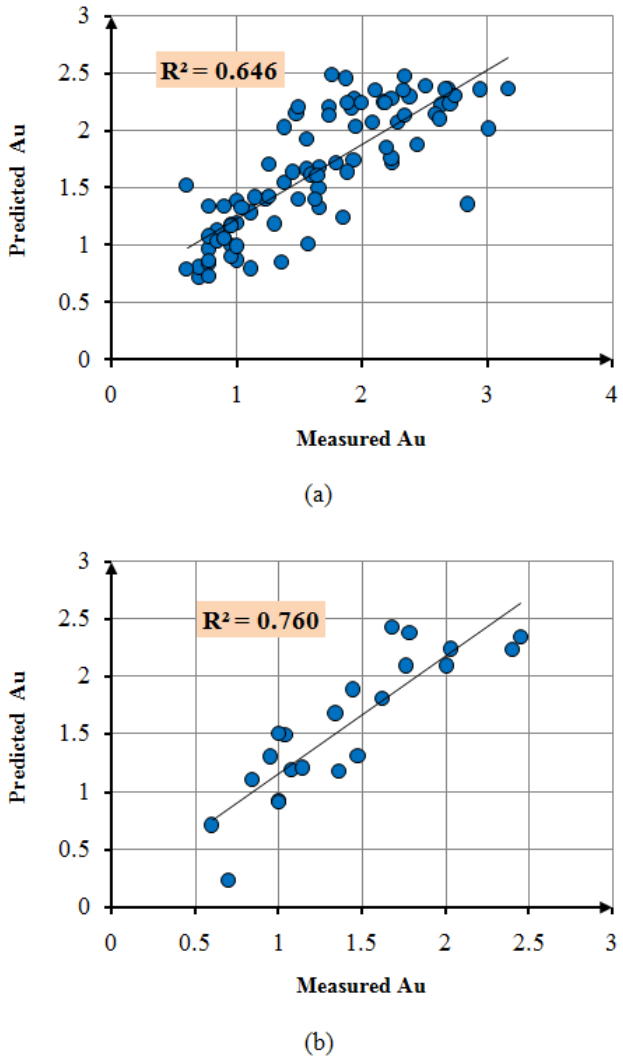

Figure 10) Comparison of the predicted using MLR and measured gold a) training data, b) testing data.

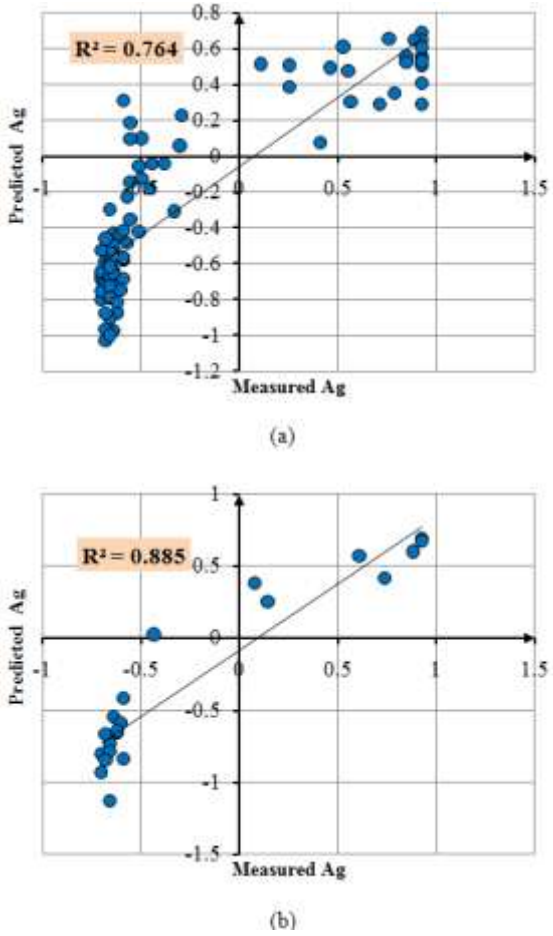

Figure 11) Comparison of the predicted using MLR and measured silver a) training data, b) testing data

6- Conclusions

A new method to predict gold and silver by experimental data has been presented using ANN-IWO method. The predictions for gold and silver using ANN-IWO method with MLR method are presented and compared with the measured data. The input data for the ANN-ICA and MLR models have been selected based on the high values of the correlation coefficients between gold and silver. It was found that a close agreement was achieved between the predicted and measured concentrations for gold and silver when the ANN-IWO method was used. Low correlation values between the model predictions and measured data using MLR method describes its low capability in prediction gold and silver.

Acknowledgments

We acknowledge the Arak University of Technology, Iran and especially Sarya Mohammadi for providing the data and permission to publish them in this study.

References:

Asadi, H. 2000. The Zarshuran gold deposit model applied in a mineral exploration GIS in Iran. PhD Thesis, University of Delft and ITC, The Netherlands, 127–165.

Asadi, H., Voncken, J., Hale, M. 1999. Invisible gold at Zarshuran, Iran. Economic Geology: 94, 1367–1374.

Asadi, H. H., Porwal, A., Fatehi, M., Kianpouryan, S., Lu, Y.-J. 2014. Exploration feature selection applied to hybrid data integration modeling: Targeting copper-gold potential in central Iran. Ore Geology Reviews: 71, 819–838.

Brown, W., Groves, D., Gedeon, T. 2003. Use of fuzzy membership input layers to combine subjective geological knowledge and empirical data in a neural network method for mineral-potential mapping. Natural Resources Research: 12, 183–200.

Hajihassani, M., Jahed Armaghani, D., Sohaei, H., Tonnizam Mohamad, E., Marto, A. 2014. Prediction of airblast-overpressure induced by blasting using a hybrid artificial neural network and particle swarm optimization. Applied Acoustics: 80, 57–67.

Harris, D., Pan, G. 1999. Mineral favorability mapping: a comparison of artificial neural networks, logistic regression, and discriminant analysis. Natural Resources Research: 8, 93–109.

Harris, D., Zurcher, L., Stanley, M., Marlow, J., Pan, G. 2003. A comparative analysis of favorability mappings by weights of evidence, probabilistic neural networks, discriminant analysis, and logistic regression. Natural Resources Research: 12, 241–255.

Leite, E. P., de Souza Filho, C. R. 2009a. Artificial neural networks applied to mineral potential mapping for copper-gold mineralizations in the Carajás Mineral Province, Brazil. Geophysical Prospecting: 57, 1049–1065.

Leite, E. P., de Souza Filho, C. R. 2009b. Probabilistic neural networks applied to mineral potential mapping for platinum group elements in the Serra Leste region, Carajás Mineral Province, Brazil. Computers and Geosciences: 35, 675–687.

Li, X.-l., Li, L.-h., Zhang, B.-l., Guo, Q.-j. 2013. Hybrid self-adaptive learning based particle swarm optimization and support vector regression model for grade estimation. Neurocomputing: 118, 179–190.

Li, X., Xie, Y., Guo, Q. 2010. A new intelligent prediction method for grade estimation. In: Advances in Neural Networks-ISNN 2010. Springer, pp 507–515.

McCulloch, W. S., Pitts, W. 1943. A logical calculus of the ideas immanent in nervous activity. The Bulletin of Mathematical Biophysics: 5, 115–133.

Mehrabi, B., Yardley, B., Cann, J. 1999. Sediment-hosted disseminated gold mineralisation at Zarshuran, NW Iran. Mineralium Deposita: 34, 673–696.

Mehrabian, A. R., Lucas, C. 2006. A novel numerical optimization algorithm inspired from weed colonization. Ecological informatics: 1, 355–366.

Meulenkamp, F., Grima, M. A. 1999. Application of neural networks for the prediction of the unconfined compressive strength (UCS) from Equotip hardness. International Journal of Rock Mechanics and Mining Sciences: 36, 29–39.

Mohajer, G., Parsaie, H., Fallah, N., Ma'dani, F. 1989. Mercury exploration in the Saein Dez-Takab area. Ministry of Mines and Metals of Iran, Tehran, 227 p (in Persian).

Price, R. H., Bauer, S. J. 1985. Analysis of the elastic and strength properties of Yucca Mountain tuff, Nevada. Paper presented at the 26th US Symposium on Rock Mechanics, United States, December 1985.

Rad, H. S., Lucas, C. A recommender system based on invasive weed optimization algorithm. In: Evolutionary Computation, 2007. CEC 2007. IEEE Congress on, 2007. IEEE, pp 4297–4304.

Roy, G. G., Das, S., Chakraborty, P., Suganthan, P. N. 2011. Design of non-uniform circular antenna arrays using a modified invasive weed optimization algorithm. Antennas and Propagation, IEEE Transactions on: 59, 110–118.

Samanta, B., Bandopadhyay, S., Ganguli, R. 2002. Data segmentation and genetic algorithms for sparse data division in Nome placer gold grade estimation using neural network and geostatistics. Exploration and mining geology: 11, 69–76.

Samanta, B., Bandopadhyay, S., Ganguli, R., Dutta, S. 2004. Sparse data division using data segmentation and Kohonen network for neural network and geostatistical ore grade modeling in Nome offshore placer deposit. Natural Resources Research: 13, 189–200.

Samimi, M. 1992. Reconnaissance and preliminary exploration in the Zarshuran area. Kavoshgaran Engineering Consultant, Tehran: 47.

Simpson, P. K. 1990. Artificial neural system—foundation, paradigm, application and implementations. Pergamon Press, New York.

Singer, D. A., Kouda, R. 1996. Application of a feedforward neural network in the search for Kuroko deposits in the Hokuroku district, Japan. Mathematical Geology: 28, 1017–1023.

Xing, B., Gao, W.-J. 2014. Invasive Weed Optimization Algorithm. In: Innovative Computational Intelligence: A Rough Guide to 134 Clever Algorithms. Springer, pp 177–181.

Received: 22 October 2015 / Accepted: 04 December 2015 / Published online: 12 December 2015

EDITOR–IN–CHIEF:

Dr. Vahid Ahadnejad*:*
Payame Noor University, Department of Geology.PO BOX 13395–3697, Tehran, Iran.
E–Mail: edchief@jtethys.org

EDITORIAL BOARD:

Dr. Jessica Kind:
ETH Zürich Institut für Geophysik, NO H11.3, Sonneggstrasse 5, 8092 Zürich, Switzerland
E–Mail: jessica.kind@erdw.ethz.ch

Prof. David Lentz
University of New Brunswick, Department of Earth Sciences, Box 4400, 2 Bailey Drive Fredericton, NB E3B 5A3, Canada
E–Mail: dlentz@unb.ca

Dr. Anita Parbhakar–Fox
School of Earth Sciences, University of Tasmania, Private Bag 126, Hobart 7001, Australia
E–Mail: anitap1@utas.edu.au

Prof. Roberto Barbieri
Dipartimento di Scienzedella Terra e Geoambientali, Università di Bologna, Via Zamboni 67 – 40126, Bologna, Italy
E–Mail: roberto.barbieri@unibo.it

Dr. Anne–Sophie Bouvier
Faculty of Geosciences and Environment, Institut des science de la Terre, UniversitédeLausanne, Office: 4145.4, CH–1015 Lausann, Switzerland
E–Mail: Anne–Sophie.Bouvier@unil.ch

Dr. Matthieu Angeli
The Faculty of Mathematics and Natural Sciences, Department of Geosciences, University of Oslo
Postboks 1047 Blindern, 0316 OSLO, Norway
E–Mail: matthieu.angeli@geo.uio.no

Dr. Miloš Gregor
Geological Institute of DionysStur, MlynskaDolina, Podjavorinskej 597/15 DubnicanadVahom, 01841, Slovak Republic
E–Mail: milos.gregor@hydrooffice.org

Dr. Alexander K. Stewart
Department of Geology, St. Lawrence University, Canton, NY, USA
E–mail: astewart@stlawu.edu

Dr. Cristina C. Bicalho
Environmental Geochemistry, Universidade Federal Fluminense – UFF, Niteroi–RJ, Brazil
E–mail: ccbicalho@gmail.com

Dr. Lenka Findoráková
Institute of Geotechnics, Slovak Academy of Sciences, Watsonova 45,043 53 Košice, Slovak Republic
E–Mail: findorakova@saske.sk

Dr. Mohamed Omran M. Khalifa
Geology Department, Faculty of Science, South Valley, Qena, 83523, Egypt
E–Mail: mokhalifa@svu.edu.eg

Prof. A. K. Sinha
D.Sc. (Moscow), FGS (London). B 602, VigyanVihar, Sector 56, GURGAON 122011, NCR DELHI, Haryana, India
E–Mail: anshuksinha@gmail.com