

*Received May 15, 2012; reviewed; accepted July 13, 2012*

## COMPARISON BETWEEN NEURAL NETWORKS AND MULTIPLE REGRESSION METHODS IN METALLURGICAL PERFORMANCE MODELING OF FLOTATION COLUMN

**Fardis NAKHAEI, Mehdi IRANNAJAD**

Department of Mining & Metallurgical Engineering, Amirkabir University of Technology, Tehran, Iran  
Iranajad@aut.ac.ir

**Abstract:** Artificial neural networks are relatively new computational tools which their inherent ability to learn and recognize highly non-linear and complex relationships makes them ideally suited in solving a wide range of complex real-world problems. In this research, different techniques (Linear regression, Non-linear regression, Back propagation neural network, Radial Basis Function for the estimation of Cu grade and recovery values in flotation column concentrate are studied. Modeling is performed based on 90 datasets at different operating conditions at Sarcheshmeh pilot plant, a copper concentrator in Iran, which include chemical reagents dosage, froth height, air and wash water flow rates, gas holdup and Cu grade in the rougher feed and flotation column feed, column tail and final concentrate streams. The results of models were also expressed and analyzed by intuitive graphics. The results indicated that a four-layer BP network gave the most accurate metallurgical performance prediction and all of the neural network models outperformed non-linear regression in the estimation process for the same set of data.

**Keywords:** *metallurgical performance; separation; neural networks; non-linear regression; prediction; flotation column*

### Introduction

The flotation column was invented by the Canadian Pierre Boutin in the 1960s. The flotation column has continued to attract more and more attention with its wide use in the metal and non-metal mineral processing fields. It has many advantages including high separation efficiency, large production capacity, and low cost. The ability to operate columns with deep froth beds and to wash the froth was the main reasons cited for the improved metallurgical performance. Concentrate grade and recovery modeling is usually commenced at the mineral processing stage of a mining project and is continued throughout the life of a mine. The major focus of this task lies in recovery

and grade estimations, which are useful for mine investment decisions, plant design and control planning.

In recent years, many researchers have focused on the relationship between metallurgical performance and operation properties. These research results show that efficiency of process depends on its operation conditions and most effective parameters are feed grade, grain size, froth height and wash water and air rates and etc. Thus, some researchers by using classic statistical methods and recently by developing intelligent techniques have established models based on experimental conditions to estimate grade and recovery.

Zhang et al. (2007) implemented artificial neural networks (ANNs) for coal mining information fusion. They studied the parameters affecting the volume of a gas burst in Chinese coal mines. A total of 20 gas bursts and six different factors (working depth, seam thickness, gas content, mining intensity, adjacent layer spacing and adjacent layer gas content) were used to train and validate the network. Al Thyabat (2008) used neural networks (NN) for the optimization of froth flotation. A multi-layered feed-forward NN was used to study the effect of feed mean size, collector dosage and impeller speed on flotation recovery and grade. Gupta et al. (1999) also worked on this topic, focusing on phosphate flotation. In order to find some relationships between rate constants and operating variables, they suggested a hybrid model combining first principles and NNs. Once calibrated, the prediction of the effect of frother concentration, particle size, air flow rate and bubble diameter on phosphate recovery was made possible. Ozbayoglu et al. (2008) applied different techniques for the estimation of coal Hardgrove Grindability Index (HGI) values. Nonlinear regression and NN techniques are used for predicting the HGI values for the specified coal parameters. Results indicate that a hybrid network which is a combination of four separate NNs gave the most accurate HGI prediction and all of the NN models outperformed non-linear regression in the estimation process. Jorjani et al. (2007) investigated the application of NNs in organic and inorganic sulfur reduction from coal. This work is an attempt to solve the important question: with the use of experimental data resulted from laboratory level, can we predict directly the organic and inorganic sulfur reduction by means of NNs? In this context the present study, ANN and multivariate statistical models are presented which have potential of predicting with acceptable accuracy using some simple parameters. The aim of this work is to evaluate two models described in the literature to estimate Cu grade and recovery in the column flotation concentrate, and to identify which models and method give the best predictions.

The remainder of the paper is organized as follows. Section 2 provides a brief description of the pilot plant and the selection of data. Sections 3 introduce the architecture and learning algorithm of BPNN and RBFNN. Section 4 describes the result and discussion and section 5 concludes the presentation.

## Pilot plant description

The Sarcheshmeh pilot plant has an identical circuit configuration compared to the plant and can process 1.6 Mg/h (tone/h) of ore. The rougher flotation bank consists of 14 cells (35 dm<sup>3</sup> each) in three units (6cell–4cell–4cell) and the regrind mill is a 76.2 cm by 137.2 cm ball mill. The scavenger banks have 6 cells (30 dm<sup>3</sup>). The single stage flotation column operation which is employed in the cleaner circuit is composed of a column with 26 cm internal diameter and 540 cm height. Figure 1 shows the flotation circuit of the pilot plant. The column is fed by the mixture of rougher stage and scavenger stage concentrates, previously classified in cyclones. Column tailing is used as scavenger feed, and column concentrate is the plant final product.

The pilot flotation column is instrumented with flow meters for feed, wash water and air as well as with a conductivity profile. Local control loops are implemented to regulate feed, tails, wash water and air flow rates. Two 23 cm long spargers are used, made of PVC tubes. The holes of 1.5 mm in diameter in a grid with dimensions of 2.5 cm × 2 cm were drilled.

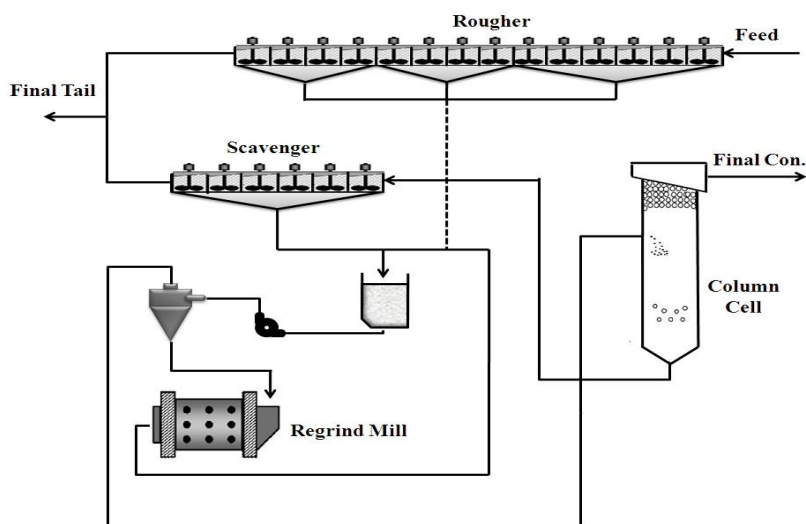


Fig. 1. The flow sheet of flotation circuit of the Sarcheshmeh pilot plant

The pulp–froth interface position is measured using a semi-analytical method based on the conductivity profile along the column. The non-floated flow rate is also controlled by a variable speed peristaltic pump driven by a frequency inverter. The pressure measurements are used to calculate the values of the air holdup and of the froth layer height. The data acquisition system is also connected via a port to a microcomputer.

## Collecting data

The basic idea behind the current research is statement of ANN (artificial neural network) and MNLR (multi non-linear regression) capability to estimate the Cu grade and recovery in flotation column concentrate. The most important step in developing an ANN and MNLR is to assemble data that can be used for training and testing the model. Therefore, a series of reliable experimental data was collected in 4 tests. A total of 90 data pairs have been selected from the experimental database. Pilot plant data was collected, over a period of 13 min based on RTD (residence time distribution), in order to cover fluctuations in all the measured variables related to the concentrate grade prediction (Nakhaei, et al. 2012).

The simultaneous measured variables are chemical reagents dosage, froth height, air and wash water flow rates, gas holdup, Cu and grade in the rougher feed, flotation column feed, column tail and final concentrate streams. The sampling period was chosen as 13 min. The maximum, minimum, mean and standard deviation of variables are given in Table 1.

The pH was adjusted to 11.8 with lime. In all tests, the rougher feed flow rate was kept 1.6 tone/h. The particle size characterization and solid percent data are given in Table 2. The chemical reagents used for the flotation process include collector and frother. Frother and collector were added in rougher cells and ball mill (before flotation circuit), respectively. The reagent type and dosage (collector and frother) used in the flotation circuit are presented in Table 3. The chemical analysis and mineralogical composition of the samples show that in all samples, the ore consists of mainly 1.78%  $\text{CuFeS}_2$ , 0.27%  $\text{Cu}_2\text{S}$  and 0.083%  $\text{MoS}_2$ .

Table 1. The maximum, minimum, mean and standard deviation of variables in different operating conditions

Variable	Index	Maximum	Minimum	Mean	Standard deviation
Froth height (cm)	$H_F$	120	35	83.61	20.52
Air holdup (%)	$\epsilon_g$	92	71	82.36	4.1
Cu grade in the rougher feed (%)	RF	0.93	0.77	0.82	0.04
Cu recovery in the flotation column (%)	Re	91.27	83.34	87.33	1.75
Cu grade in the flotation column feed (%)	CoF	11.96	6.95	8.89	1.22
Cu grade in the flotation column tail (%)	CoT	2.68	1.05	1.81	0.45
Cu grade in the flotation column concentrate (%)	FC	25.21	15.93	21.13	2.12
Frother dosage (g/Mg)	Fr	36	32	34	1.64
Collector dosage (g/Mg)	C	40	36	38	1.64
Wash water flow rate (cm/s)	$Q_w$	0.34	0.11	0.27	0.08
Air flow rate (cm/s)	$Q_a$	1.72	0.63	1.1	0.25

Table 2. Flotation conditions used in the experiments (pH=11.8)

Parameter	Rougher feed	Column feed	Final concentrate	Final tail
Solid (%)	27	14	14.5	28
# -325 (%)	48	85	74	54

Table 3. Reagents type was used in the flotation circuit

Reagents	Type and commercial name
Collectors	Z <sub>11</sub> , Nascol 451
Frothers	MIBC, Dowfroth 250

## Methods

### Back propagation neural network architectures

Back propagation is the most commonly used neural network type due to its simplicity in implementation and its successful generalization capabilities for complex data sets. Back propagation neural network is a three-layered feed forward architecture. The three layers are input layer, hidden layer and output layer. Functioning of back propagation proceeds in three stages, namely learning or training, testing or inferences and validation. Figure 2 shows the p-q-1 (p input neurons, q hidden neurons, and 1 output neuron) architecture of a back propagation neural network model. Input layer receives information from the external sources and passes this information to the network for processing. Hidden layer receives information from the input layer, and does all the information processing, and output layer receives processed information from the network, and sends the results out to an external receptor. The input signals are modified

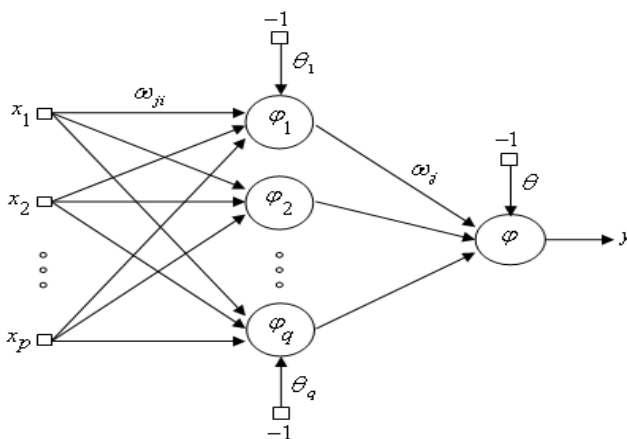


Fig. 2. A schematic diagram of multilayer feed forward neural network

by interconnection weight, known as weight factor  $w_{ji}$ , which represents the interconnection of  $i^{\text{th}}$  node of the first layer to  $j^{\text{th}}$  node of the second layer. The sum of modified signals (total activation) is then modified by a sigmoid transfer function ( $\varphi$ ). Similarly, outputs signal of hidden layer are modified by interconnection weight ( $w_{ki}$ ) of  $k^{\text{th}}$  node of output layer to  $j^{\text{th}}$  node of hidden layer. The sum of the modified signal is then modified by sigmoid transfer function and output is collected at output layer (Mosavi, 2011).

Output from a neuron in the input layer is

$$O_{pi} = x_{pi}, \quad i = 1, 2, \dots, p. \quad (1)$$

Output from a neuron in the hidden layer is

$$O_{pj} = \varphi(\sum_{i=0}^p w_{ji} O_{pi} + \theta_j), \quad j = 1, 2, \dots, q. \quad (2)$$

Output from a neuron in the output layer is

$$O_{pk} = \varphi(\sum_{i=0}^q w_{ki} O_{pj} + \theta) = y, \quad k = 1, 2, \dots, n. \quad (3)$$

### Learning or training in back propagation neural network

Batch model type of supervised learning has been used in the present case, where interconnection weights are adjusted using delta rule algorithm after sending the entire training sample to the network. During training, the predicted output is compared with the desired output, and the mean square error is calculated. If the mean square error is more than a prescribed limiting value, it is back propagated from output to input, and weights are further modified till the error or number of iterations is within a prescribed limit. Mean square error,  $E_p$  for pattern  $p$  is defined as (Samanta et al., 2009):

$$E_p = \sum_{i=1}^n \frac{1}{2} (D_{pi} - O_{pi})^2 \quad (4)$$

where  $D_{pi}$  is the target output, and  $O_{pi}$  is the computed output for the  $i^{\text{th}}$  pattern. Weight change at any time  $t$ , is given by:

$$\Delta w(t) = -\eta E_p(t) + \alpha \Delta w(t-1) \quad (5)$$

$\eta$  = learning rate ( $0 < \eta < 1$ );  $\alpha$  = momentum coefficient ( $0 < \alpha < 1$ ).

### Radial basis function network

#### Architecture of radial basis function network

Basically radial basis function network is composed of large number of simple and highly interconnected artificial neurons and can be organized into several layer, i.e.

input layer, hidden layer, and output layer by Haykin (2004). These inputs are the non-linearly transformed by a set of RBFs, creating a new space,  $j$ , in the hidden layer. Through this non-linear transformation only, the network acquires the capability of non-linear function al mapping. The most commonly used non-linear RBF is the Gaussian function. Other types of non- linear functions such as multi quadratic and the thin-plate spline are alternately used (Haykin, 2004). The general consensus is that the type of the basis function does not have much impact on the general performance of the RBF network. The Gaussian RBFs of the network are characterized by two sets of parameters: radial basis centers,  $\mu$  , and widths,  $\sigma$  . Hence, the different basis functions in the hidden layer are recognized by these parameters. The  $i^{th}$  basis function of the network can be expressed by the following mathematical formula

$$\Phi_i(x) = \exp\left(-\frac{\|x - \mu_i\|}{\sigma_i}\right) \tag{6}$$

where  $\Phi_i(x)$  is the output from the  $i^{th}$  basis function,  $\mu_i$  the center of the  $i^{th}$  basis function,  $\sigma_i$  the width of the  $i^{th}$  basis function and  $\| \cdot \|$  the Euclidian distance of the input from the center. The outputs from the hidden layer are then linearly combined to produce the final output ( $y$ ) of the network, which in the study is the Cu grade and recovery in flotation column concentration. This can be expressed as:

$$y = \sum_{i=1}^n w_i \cdot \Phi_i(x) \tag{7}$$

where  $y$  is the output of the network,  $w_i$  the weight associated with the  $i^{th}$  basis function and  $n$  the number of the basis functions. While the weight parameters of a feed-forward network are determined using a complex non-linear optimization algorithm, demanding expensive computational time, the weight parameters of a RBF network can be fixed by using a least square algorithm. This is the point where a RBF network gains substantial computational advantage over a feed-forward network.

**Preparation of data sets**

The normalized data sets are used for training the network. In the present case, the data are normalized between 0 and 1 to ensure that each input is represented in the network training as well as different kinds of input quantities are normalized in the same scale. The data sets are normalized in the range of 0 and 1 using (Nakhaei et al., 2010):

$$N_p = \frac{A_p - meanA_{ps}}{stdA_p} \tag{8}$$

where,  $A_p$  is actual parameter,  $meanA_{ps}$  is mean of actual parameters,  $stdA_p$  is standard deviation of actual parameter and  $N_p$  is normalized parameter.

### Testing and validation of neural network

Entire experimental data set is divided into training set, testing set. The error on the testing set is monitored during the training process. The testing error will normally decrease during the initial phase of training, as does the training set error. However, when the network begins to overfit the data, the error on the testing set will typically begin to rise. When the testing error starts increasing for a specified number of iterations, the training is stopped; and the weights at the minimum value of the testing error are returned.

### Results and discussion

All data analysis methods used 60 training and 30 testing data points. Different combinations of the data set are used during the process, so all the data points have eventually been tested. Eight input parameters were setup as network inputs to the input layer. These parameters are chemical reagents dosage (frother and collector), froth height, air and wash water flow rates, gas holdup, Cu grade in the rougher feed, flotation column feed streams are considered as input which have the influence on cu grade and recovery in flotation column concentrate are considered as output parameter.

#### Multi linear regression

The same approach is used in linear and non-linear regressions as it was used in the neural network system training and testing. A statistical model of regression, the data of which is similar to that of NNs, had been employed to predict Cu grade and recovery. It should be mentioned that 60 sets of data were selected for simulating the regression model. Also, the 30 data were used to test the performance of the model where inputs are referred to independent variables which are the same as inputs used in NN. The MLR equations forecasted the Cu grade and recovery with correlation coefficients of 0.87 and 0.85 in testing stage respectively. The MSE between the estimated Cu grade and recovery and the desired data was 1 and 0.96 respectively. Results show that the average error for linear regression was considerably higher than the other models.

#### Multi non-linear regression

Two multivariable regression equations were developed for the prediction of the Cu grade and recovery, which were deduced for these results as the follows:

$$\begin{aligned}
 Fc = & 147.2951 - 0.0047H_f + 3.448 \cdot 10^{-4}H_f^2 - 0.0009C^2 + 0.1235Fr \\
 & - 1.7435\varepsilon_g + 0.01087\varepsilon_g^2 + 10.3393Q_a - 4.328Q_a^2 + 8.1804Q_w \\
 & + 6.434Q_w^2 - 184.757RF + 109.8199RF^2 + 1.3818CoF - 0.0596CoF^2
 \end{aligned} \tag{9}$$



$$\begin{aligned}
 Re = & 25.3116 - 0.0267H_f - 3.05 \cdot 10^{-4}H_f^2 - 0.0524C - 0.0027Fr^2 \\
 & + 0.6013\varepsilon_g - 0.0028\varepsilon_g^2 + 11.2774Q_a - 3.3297Q_a^2 - 5.9272Q_w \\
 & + 17.171Q_w^2 + 106.0232RF - 64.9079RF^2 - 2.3751CoF + 0.1408CoF^2
 \end{aligned}
 \tag{10}$$

All symbols used in equations were explained in Table 1.

The statistical significance of the regression coefficient is a test which indicates whether there is a relationship between independent variables effect and the dependent variable. To assess the significance of regression coefficients, one needs to set a risk level (called the P-value level). As the rule of thumb, in most cases the P-value is set at level 0.05 (for 95% confidence).

The regression coefficient obtained from equations 9 and 10 showed that the majority of independent variables and their interactions had a significant effect on the grades and recoveries in a way that the P-values were less than 0.05. The MNLN equations forecasted the Cu grade and recovery with correlation coefficients of 0.9 and 0.86, respectively

### BP neural network

Best network architecture (i.e. number of hidden layers, number of neurons in the hidden layers, learning rate and momentum coefficient) has been obtained by trial and error based on mean square error MSE in training, MSE in testing, and the number of iterations. A feed-forward ANN was trained with the BP algorithm, the model of which was designed by software package MATLAB 2009. The best results were obtained with Feed-forward NN with 8-12-8-2 arrangement that was capable to estimate

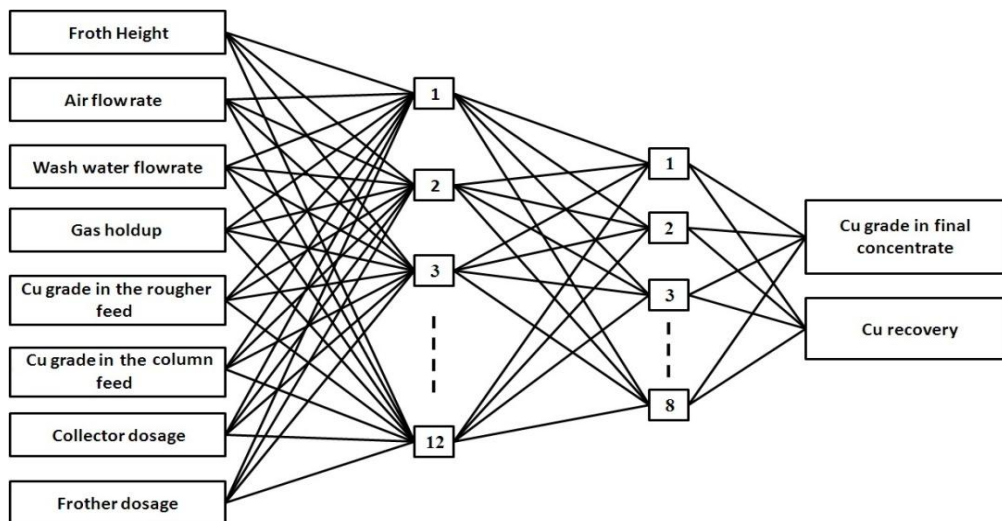


Fig. 3. Optimum structures of the proposed BPNN for estimating metallurgical performance of Cu in flotation column

metallurgical performance of Cu in flotation column (Fig. 3). Log-sigmoidal and linear were used as transfer functions in hidden layers and output layer, respectively. In NN modeling approach, a Levenberg-Marquardt training algorithm was used as a learning rule. The training process was stopped after 500 epochs for metallurgical performance of Cu. In this study, the associated ANNs analyses were carried out with an optimal value of learning rate of 0.2 and the momentum of 0.8. The R (correlation coefficient) values of testing set for Cu grade and recovery are 0.92 and 0.92, respectively.

### RBF neural network

In this section an attempt has been made in selecting the methodology of the network. The present methodology is used with an advantage that hidden layer weights or centre vectors are optimized first then it forward the processes to the output layer. Hidden-output layer weights are further optimized using gradient descent technique. So there was no backflow at the input-hidden layer which reduces the simulation time of the code. Best network architecture (i.e. number of centre vectors in the hidden layers, learning rate and momentum coefficient) has been obtained by trial and error based on mean square error in training, testing, and the number of iterations.

Best network architecture (i.e. number of centre vectors in the hidden layers, learning rate and momentum coefficient) has been obtained by trial and error based on mean square error in training, testing, and the number of iterations. Number of centre vectors varies from 10 to 30 and range of  $\eta$  and  $\alpha$  varies from 0.1 to 0.9. The best results were obtained with 8-30-2 arrangement that was capable to estimate metallurgical performance of Cu in flotation column. Also, the associated ANNs analyses were carried out with an optimal value of learning rate of 0.7 and the momentum of 0.9 with 1000 iterations.

Table 4. Metallurgical performance estimation results by various models

Std error	MIN error	MAX error	MSE	R	Index	Model
0.99	-2.07	1.63	1.00	0.87	FC	MLR
0.92	-2.44	2.09	0.96	0.85	Re	
0.84	-1.37	2.31	0.98	0.9	FC	MNLr
0.91	-2.41	1.26	1.09	0.86	Re	
0.73	-1.07	1.56	0.71	0.92	FC	BPNN
0.66	-1.3	0.82	0.48	0.92	Re	
0.83	-1.36	1.82	0.75	0.91	FC	RBFNN
0.73	-1.7	1.38	0.66	0.90	Re	

### Comparison of Cu grade and recovery prediction by NN and statistical methods

Apparently, it has been observed from Table 4 as well as Figs 4 and 5 that Cu grade and recovery predicted by both BPNN and RBFNN is within a very good tolerance limit compared with the traditional regression method. Both the networks are well fitted with the data collected and show the least variation error in predicting the Cu grade and recovery. It has also been observed that BPNN shows better prediction accuracy compared to RBFNN but RBFNN converges faster compared to BPNN.

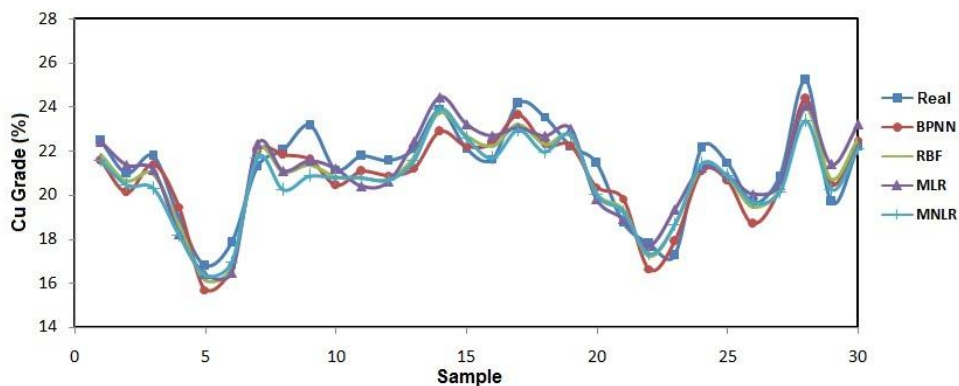


Fig. 4. NN values and regression values against observed (experimental) values in testing stage

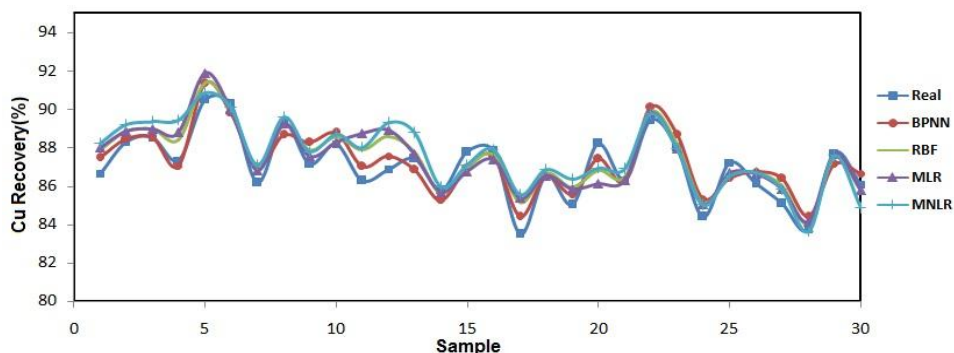


Fig. 5. NN values and regression values against observed (experimental) values in testing stage

### Conclusion

This paper uses two techniques – ANN and statistical methods – to estimate Cu grade and recovery values in flotation column concentrate. This study has shown that BPNN is effective for predicting metallurgical performance of flotation column. Significant advantage of this model is that it can provide satisfactory predictions with short as well as large data. The experimental results suggest that BPNN models can offer reliable frameworks for modeling Cu grade and recovery in flotation column. Similarly it

has also been observed that RBFNN (radial basis function neural network) based prediction systems achieve faster convergence compared to BPNN (back propagation neural network) based system but with higher levels of prediction errors. Therefore, it can be a very powerful tool for treating the experimental data in other similar processes. Also, the model performance may be improved by considering additional program such as genetic algorithms (GA) and fuzzy systems.

### Acknowledgments

The authors would like to acknowledge the support of Department of Research and Development of Sarcheshmeh Copper Plants for this research.

### References

- AL THYABAT S., 2008. *On the Optimization of Froth Flotation by the Use of an Artificial Neural Network*. Journal of China University of Mining and Technology, 18, 418-426.
- HAYKIN S., 2004. *Neural Networks, A Comprehensive Foundation*, second ed. Prentice Hall, NJ, USA.
- JORJANI E.S., CHELGANI C., MESROGHLI Sh., 2007. *Prediction of Microbial Desulfurization of Coal using Artificial Neural Network*. Journal of Minerals Engineering 20, 1285-1292.
- MOSAVI M.R., 2011. *Error reduction for GPS accurate timing in power systems using Kalman filters and neural networks*. J. Electr. Eng. 161-168 R. 87 NR 12.
- NAKHAEI F., SAM A. MOSAVI M.R., VAGHEI Y. 2012. *Recovery and Grade Accurate Prediction of Pilot Plant Flotation Column Concentrate: Neural Network and Statistical Techniques*, International Journal of Mineral Processing, In press.
- NAKHAEI F., SAM A. MOSAVI M.R., ZEIDABADI S., 2010. *Prediction of Copper Grade at Flotation Column Concentrate using Artificial Neural Network*. IEEE Conference on Signal Processing.
- OZBAYOGLU G., OZBAYOGLU A., OZBAYOGLU M., 2008. *Estimation of Hardgrove Grindability Index of Turkish coals by neural networks*. Int. J. Miner. Process. 85 (4), 93-100.
- SAMANTA B., BANDOPADHYAY S., 2009. *Construction of a radial basis function network using an evolutionary algorithm for grade estimation in a placer gold deposit*, Computers & Geosciences 35, 1592-1602.
- ZHANG X.Q., WANG H.B., YU H.Z., 2007. *Neural Network Based Algorithm and Simulation of Information Fusion in the Coal Mine*. Journal of China University of Mining and Technology, 17(4), 595-598.