Prediction of mine water quality by physical parameters

Manoj Khandelwal and T N Singh*

Department of Earth Sciences, Indian Institute of Technology Bombay, Powai, Mumbai 400 076
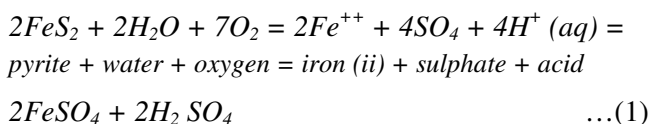
Present paper is an attempt to predict the chemical parameters like sulphate, chlorine, chemical oxygen demand, total dissolved solids and total suspended solids in mine water using artificial neural network (ANN) by incorporating the pH, temperature and hardness. The prediction by ANN is also compared with Multivariate Regression Analysis (MVRA). For prediction of chemical parameters of mine water, 30 data set were taken for the training of the network while testing and validation of network was done by 10 data set with 923 epochs. The predicted results of chemical parameters of mine water by ANN are very satisfactory and acceptable as compared to MVRA, and seem to be a good alternative for pollutants prediction.
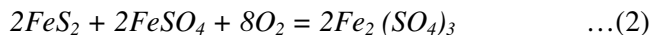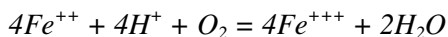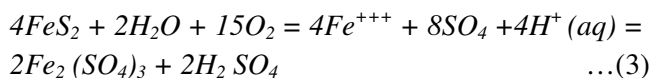
Introduction

Different mining processes (excavation, drainage pattern, geological structures, ground water conditions of the mining area) deteriorate the quality of mine water, which concerns to mine workers and population residing near mining area[1]. Acid Mine Drainage (AMD) is one of the frightening problems in mines, particularly in some of the coalmines, where coal contains pyritic material. When water passes through the pyritic strata, it gets acidic and adversely affects the performance of the machines due to its higher corrosive nature and deteriorates the water quality[2]. The pyrite is ubiquitous in most metal sulphide deposits[3]. The released acids may be partially consumed by the second stage of weathering and increase pH of water[4]. This may release more base cations and metals, which potentially increase total dissolved solids (TDS) of mine water[5]. There are different approaches to treat the acid mine water[6]:

$$2FeS_2 + 2H_2O + 7O_2 = 2Fe^{++} + 4SO_4 + 4H^+ \ (aq) =$$

pyrite + water + oxygen = iron (ii) + sulphate + acid

$$2FeSO_4 + 2H_2 SO_4 \qquad \qquad \ldots(1)$$

Further oxidation of iron (II) consumes some protons:

$$4Fe^{++} + 4H^+ + O_2 = 4Fe^{+++} + 2H_2O$$

$$2FeS_2 + 2FeSO_4 + 8O_2 = 2Fe_2 \ (SO_4)_3 \qquad \ldots(2)$$

But the overall reaction is still acid producing:

$$4FeS_2 + 2H_2O + 15O_2 = 4Fe^{+++} + 8SO_4 +4H^+ \ (aq) = 2Fe_2 \ (SO_4)_3 + 2H_2 \ SO_4 \qquad \ldots(3)$$

The pH indicates the acidic nature of mine supplemented by variations in temperature and hardness due to presence of carbonates. To study quality of mine water, chemical compositions like sulphate, chlorine (Cl), chemical oxygen demand (COD), TDS and total suspended solids (TSS) are analysed by atomic absorption, water analyser kit etc[7]. Feng *et al*[8] tried to neutralise acid mine water with $Ca(OH)_2$ or with $CaCO_3$ to get maximum precipitation of metals. Voleski[9] used bio-sorption for heavy metal removal by biomass based ion exchange concept.

The artificial neural network (ANN) has ability to learn from the pattern acquainted before. Once the network has been trained with sufficient number of sample data sets, it can make predictions on the basis of its previous learning about the output related to new input data set of similar pattern[10]. In the present investigation, an attempt has been made to predict the mine water quality by simple field observations and analysis like pH, temperature and hardness, to get the information about sulphur, chlorine, COD, TDS and TSS using ANN. The results were also compared with multivariate regression analysis (MVRA) to find the superiority of ANN over MVRA.

———————
*Author for correspondence
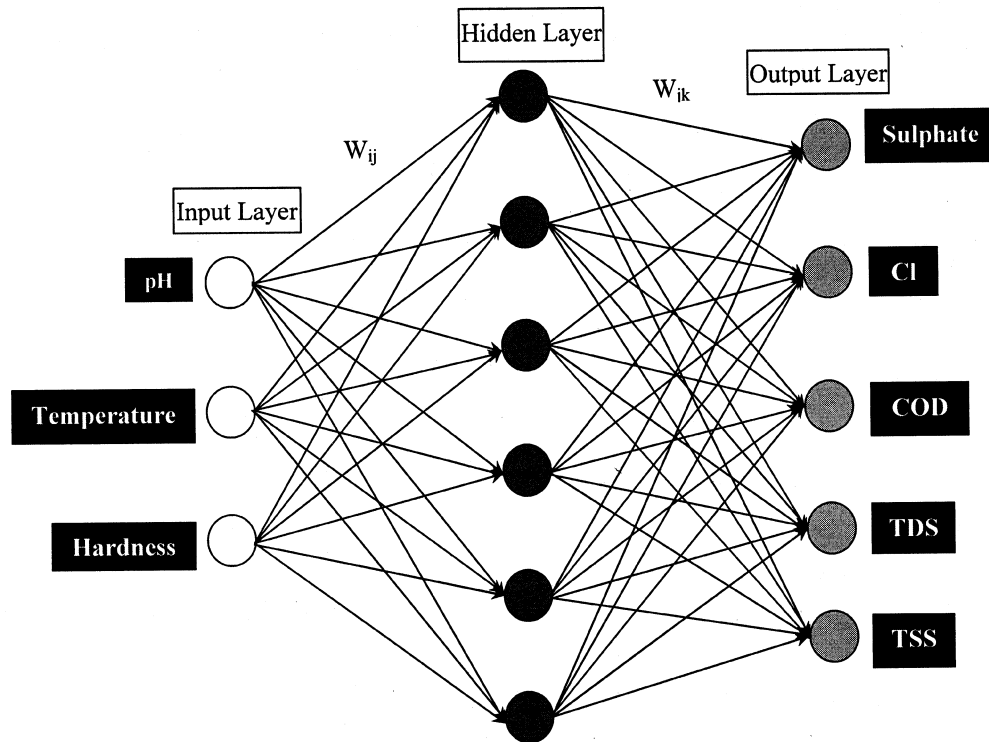Tel: 91-22-2572-2545; Fax: 91-22-2572-3480
E-mail: tnsingh@iitb.ac.in

Fig. 1—Back propagation neural network

The paradigms in this field are based on direct modelling of the human neuronal system[11]. Particular network can be defined using three fundamental components: transfer function, network architecture and learning law[12]. A network first needs to be trained before interpreting new information. Back-propagation algorithm is the most versatile and robust technique, which provides the most efficient learning procedure for multilayer neural networks. The feed forward back propagation neural network (BPNN) always consists of at least three layers: input layer, hidden layer and output layer (Fig. 1). Each layer consists neurons, and each neuron is connected to the next layer through weights i.e. neurons in the input layer will send its output as input for neurons in the hidden layer and similar is the connection between hidden and output layer. Number of hidden layers and number of neurons in the hidden layer change according to the problem to be solved. The number of input and output neuron is same as the number of input and output variables.

To differentiate between the different processing units, values called biases are introduced in the transfer functions. These biases are referred to as the temperature of a neuron. Except for the input layer, all neurons in BPNN are associated with a bias neuron and a transfer function. The bias is much like a weight, except that it has a constant input of 1, while the transfer function filters the summed signals received from this neuron. These transfer functions are designed to map a neurons or layers net output to its actual output and they are simple step functions either linear or non-linear functions. The application of transfer functions depends on the purpose of the neural network. Output layer produces computed output vectors corresponding to the solution.

Network Training

During training of the network, data is processed through the input layer to hidden layer, until it reaches the output layer (forward pass). Output is compared to the measured values ("true" output). The difference or error between both is processed back through the network (backward pass) updating the individual weights of the connections and the biases of the individual neurons. The input and output data are mostly represented as vectors called training pairs. The process is repeated for all the training pairs in the data set, until the network error converged to a threshold minimum defined by a corresponding cost

function; usually the root mean squared error (RMS) or summed squared error (SSE).

In Fig. 1, the jth neuron is connected with a number of inputs

$$x_i = (x_1, x_2\ x_3 \ldots x_n) \qquad \ldots(4)$$

The net input values in the hidden layer will be

$$Net_j = \sum^n_{i=1} x_i w_{ij} + \theta_j \qquad \ldots(5)$$

where, x_i = input units, w_{ij} = weight on the connection of ith input and j^{th} neuron; θ_j = bias neuron (optional); and n = number of input units.

So, the net output from hidden layer is calculated using a logarithmic sigmoid function

$$O_j = f(Net_j) = 1/1 + e^{-(Net_j + \theta_j)} \qquad \ldots(6)$$

The total input to the kth unit is

$$Net_k = \sum^n_{j=1} w_{jk}\ O_j + \theta_k \qquad \ldots(7)$$

where, θ_k = bias neuron; and w_{jk} = weight between j^{th} neuron and k^{th} output.

So, the total output from l^{th} unit will be,

$$O_k = f(Net_k)$$

In the learning process, the network is presented with a pair of patterns, an input pattern and a corresponding desired output pattern. The network computes its own output pattern using its (mostly incorrect) weights and thresholds. Now, the actual output is compared with the desired output. Hence, the error at any output in layer k is

$$e_l = t_k - O_k \qquad \ldots(8)$$

where, t_k = desired output, and O_k = actual output.

The total error function is given by

$$E = 0.5\ \sum^n_{k=1}\ (t_k - O_k)^2 \qquad \ldots(9)$$

Training of the network is basically a process of arriving at an optimum weight space of the network. The descent down error surface is made using the following rule:

$$\nabla W_{jk} = -\eta\ (\delta E / \delta W_{jk}) \qquad \ldots(10)$$

where, η is the learning rate parameter, and E is the error function.

The update of weights for the $(n+1)^{th}$ pattern is given as:

$$W_{jk}(n+1) = W_{jk}(n) + \nabla W_{jk}(n) \qquad \ldots(11)$$

Similar logic applies to the connections between the hidden and output layers[13]. This procedure is repeated with each pattern pair of training exemplar assigned for training the network. Each pass through all the training patterns is called a cycle or epoch. The process is then repeated as many epochs as needed until the error is within the user specified goal is reached successfully. This quantity is the measure of how the network has learned.

Dataset

The values of different input parameters ranges as follows: pH, 6.5-9.2; temperature, 26.0-32.9°C; and hardness, 600.5-711.4 mg/l. The value for different output parameters ranges as follows: sulphate, 206.1-401.2; Cl, 24.5-1009.1; COD, 21.0-235.2; TDS, 459-796; and TSS, 240.6-1180.0 mg/l.

All the input and output parameters were scaled between 0 and 1. This was done to utilize the most sensitive part of neuron and since output neuron being sigmoid can only give output between 0 and 1, the scaling of output parameter was necessary.

Scaled value = (max. value – unscaled value)/(max. value – min. value)

Network Architecture

Feed forward network, adopted here as architecture, is reported to be suitable for problem identification. Pattern matching is basically an input/output mapping problem. Closer the mapping, better will be the performance of the network. The architecture of the network has one input layer with 3 neurons, one hidden layer with 6 neurons and one output layer with 5 neurons. The training of the network has been carried out up to 923 epochs with 0.002 error goal.

Testing and Validation of ANN Model

To test and validate ANN model, the new data sets have been chosen. The results are presented in this section to demonstrate the network performance, which is the coefficient of correlation between predicted and observed values. Training of the network was done using 1 hidden layer with 6 hidden neurons. Using Bayesian regulation[14], there was no danger of over-fitting problems; hence, the network was trained with 923 training epochs.

From the graphs (Figs 2-6), it can be said that value of correlation coefficient (R^2) is coming high enough (0.7138 – 0.8835).

Multivariate Regression Analysis (MVRA)

The purpose of MVRA is to learn more about the relationship between several independent or predictor
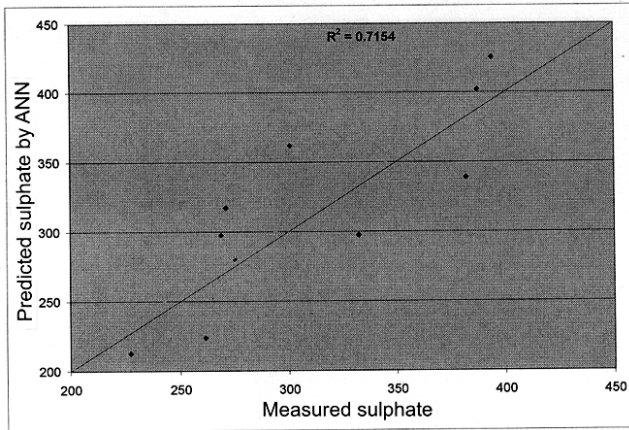
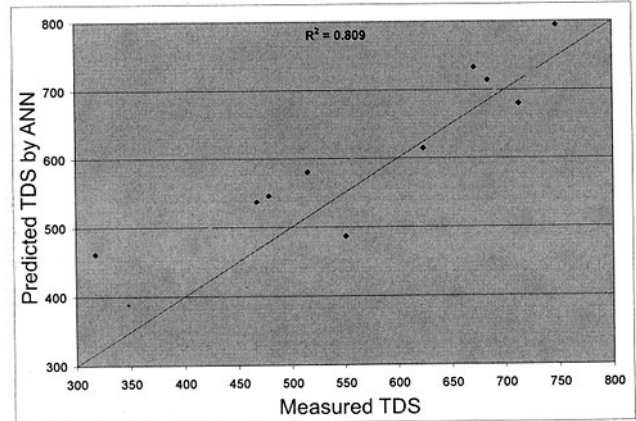Fig. 2 — Graph between measured and predicted sulphate

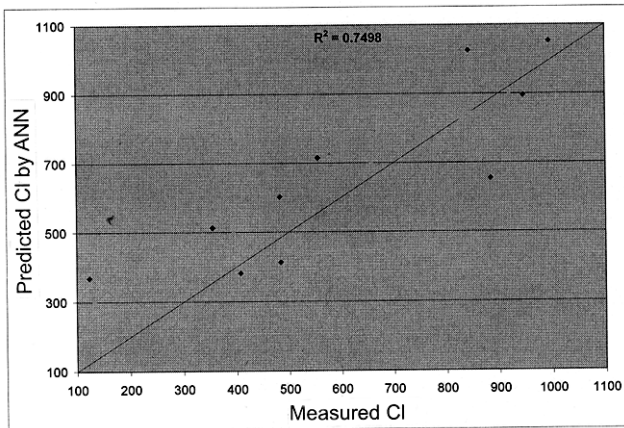Fig. 5 — Graph between measured and predicted TDS

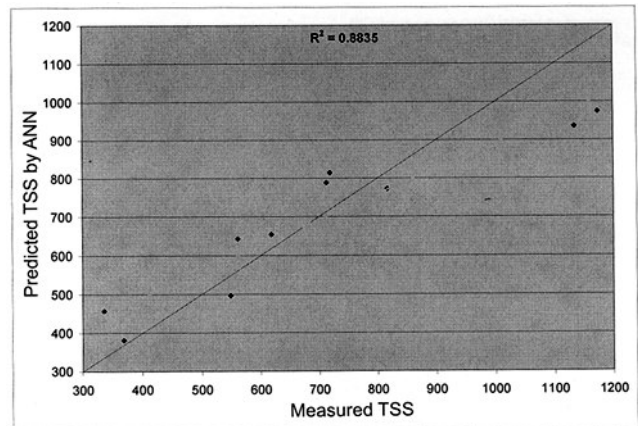Fig. 3 — Graph between measured and predicted Cl

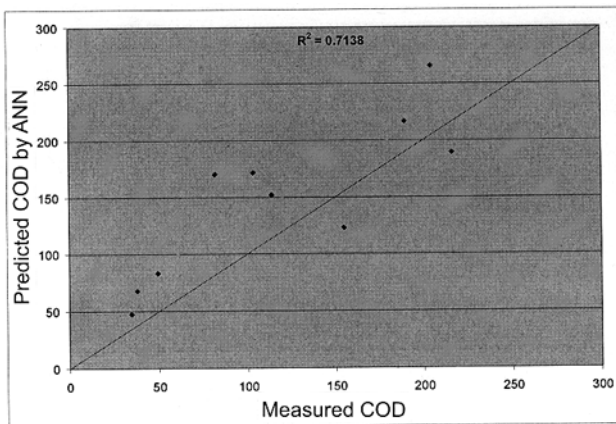Fig. 6 — Graph between measured and predicted TSS

Fig. 4 — Graph between measured and predicted COD

variables and a dependent or criterion variable. The goal of regression analysis is to determine the values of parameters for a function that cause the function to best fit a set of data observations provided. In linear regression, the function is a linear equation. When there is more than one independent variable, then MVRA is used to get best-fit equation. MVRA solves data sets by performing least squares fit. It constructs and solves simultaneous equations by forming the regression matrix and solving for the co-efficient using the backslash operator. MVRA has been done by same data sets and same input parameters used in ANN. From the graphs (Figs 7-11) it can be easily said that value of R^2 is coming low (0.0146-0.2793) and MVRA is not able to estimate the values of sulphate, Cl, COD, TDS and TSS upto an acceptable limit.

Comparison of ANN Prediction with MVRA

The graphs (Figs 12-16) between measured and predicted values of sulphate, Cl, COD, TDS and TSS show that prediction by ANN is very closer with the measured one whereas prediction by MVRA has very high errors (Table 1).
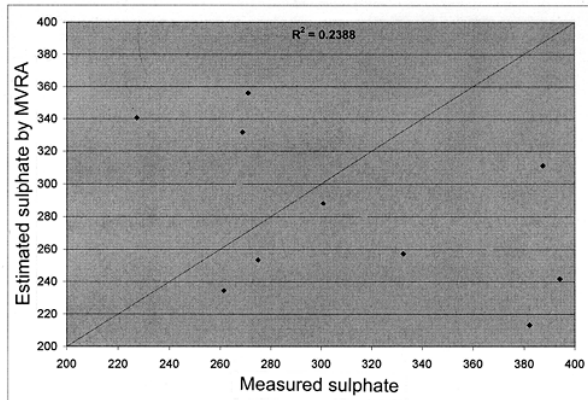
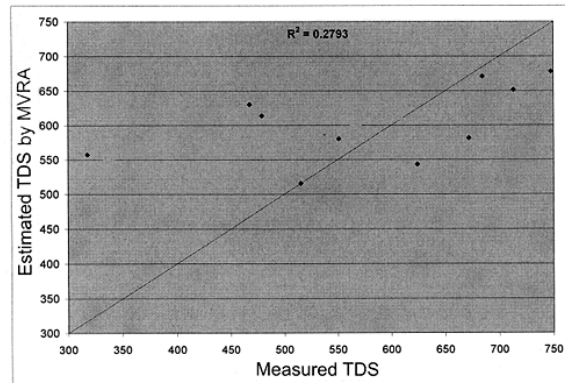Fig. 7 — Graph between measured and estimated sulphate
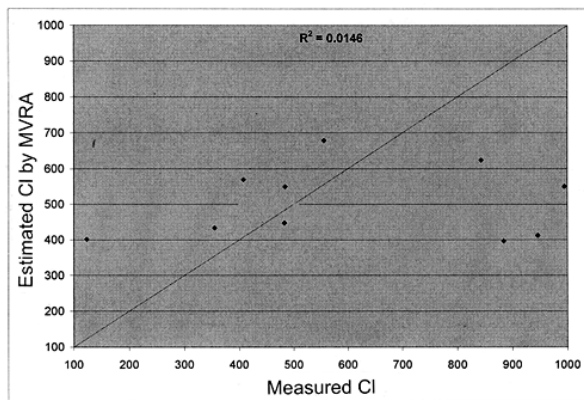

Fig. 10 — Graph between measured and estimated TDS

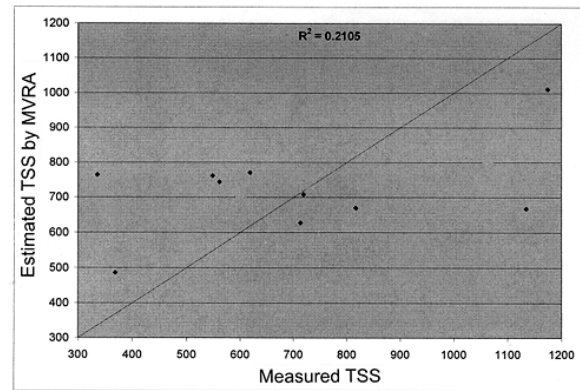
Fig. 8 — Graph between measured and estimated Cl

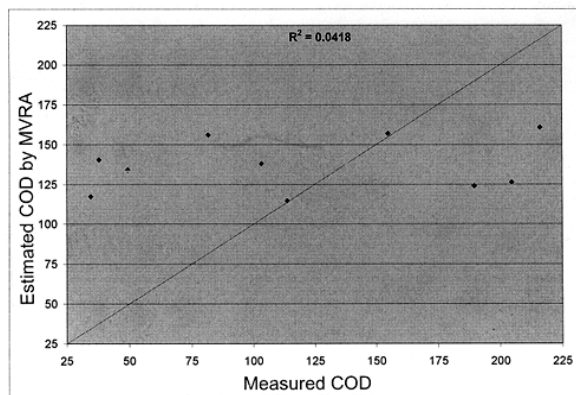Fig. 11 — Graph between measured and estimated TSS

Fig. 9— Graph between measured and estimated COD
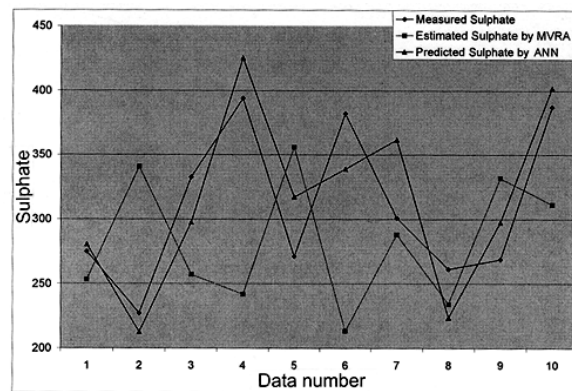

Fig. 12 — Graph between measured and predicted values of sulphate by ANN and MVRA

Conclusions

The study presents prediction of mine water quality by simple field tests. The correlation co-efficient is 0.7154, 0.7498, 0.7138, 0.809 and 0.8835 for sulphate, Cl, COD, TDS and TSS respectively by ANN, however; the correlation co-efficient is 0.2388,

0.0146, 0.0418, 0.2793 and 0.2105 by MVRA. It is observed that ANN has given very sound results as compared to MVRA, and may be used to predict for future analysis of mine water. This method will be helpful to cope-up with cumbersome techniques of mine water chemical analysis.   The results given are

Table — 1 Input and output data set parameters

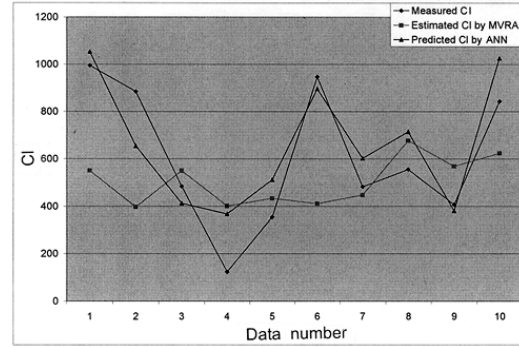| S No.	pH	Temp	Hardness mg/l	Sulphate mg/l			Cl mg/l			COD mg/l			TDS mg/l			TSS mg/l		
				Measured	MVRA	ANN	Measured	MVRA	ANN	Measured	MVRA	ANN	Measured	MVRA	ANN	Measured	MVRA	ANN
1.	6.92	27.76	612.40	274.90	253.41	280.4	994.70	550.25	1052.3	113.80	114.70	152.1	478.94	613.73	545.2	368.74	486.45	381.2
2.	8.45	26.84	701.15	227.30	340.67	212.6	883.90	396.54	654.8	154.84	156.97	122.8	684.25	670.82	714.3	548.61	761.64	496.8
3.	7.51	29.58	645.51	332.57	257.16	297.5	484.27	549.61	412.8	204.70	126.33	265.6	551.01	580.06	486.7	1134.20	667.88	935.9
4.	8.81	28.24	618.70	394.10	241.81	425.4	123.58	401.58	368.2	37.84	140.54	67.8	671.80	581.36	732.8	718.80	708.12	814.6
5.	8.12	27.16	719.84	271.06	356.10	317.2	354.80	433.85	512.7	81.73	156.06	170.4	747.80	678.13	794.6	618.43	770.92	654.7
6.	9.02	29.31	604.80	382.17	213.19	338.9	946.15	412.14	895.7	103.29	138.08	171.7	624.07	543.28	614.5	561.17	744.02	643.1
7.	9.22	31.17	711.02	300.84	288.03	361.8	482.90	447.66	602.6	216.16	160.71	189.3	317.82	557.42	461.3	1174.10	1011.14	974.4
8.	7.03	32.81	663.78	261.40	234.48	223.7	555.71	678.24	715.1	34.58	117.38	47.5	515.40	515.39	580.1	335.84	764.36	457.6
9.	6.94	28.45	704.80	268.80	331.73	297.4	407.90	569.29	381.4	49.40	134.21	83.2	713.06	651.46	679.9	816.90	670.08	772.5
10.	6.52	29.16	690.48	387.49	311.23	402.1	843.31	624.06	1025.4	189.64	123.92	216.6	467.57	630.24	536.4	713.02	627.85	788.9

Fig. 13 — Graph between measured and predicted values of Cl by ANN and MVRA
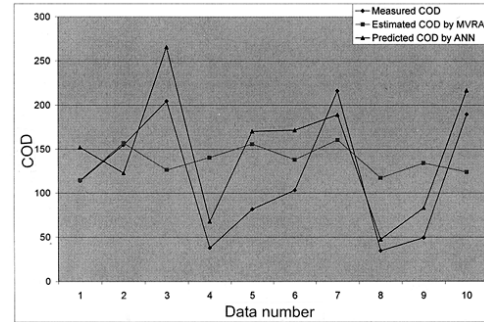

Fig. 14 — Graph between measured and predicted values of sulphate by ANN and COD
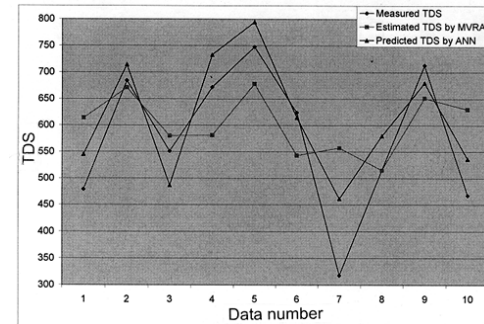

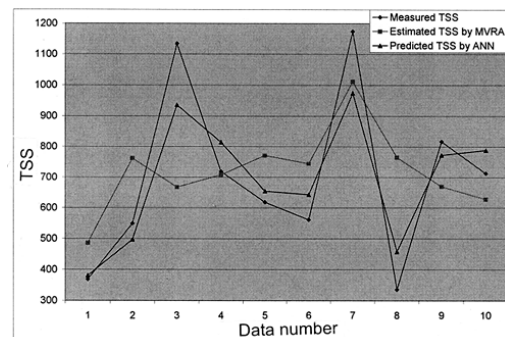Fig. 15 — Graph between measured and predicted values of TDS by ANN and MVRA

Fig. 16 — Graph between measured and predicted values of TSS by ANN and MVRA

based on limited data set. The more number of data may provide better correlation coefficient by ANN due to non-bias and non-linearity nature of simulation.

References

1 Haddon M, Mines and misdemeanours, *Water Bulletin*, (1994) 598.

2 Dubey R K & Singh T N, Influence of pH of water on mechanical properties of sandstone-An experimental approach, *J Sci Ind. Res,* **59** (2000) 583-586.

3 NRA, Abandoned mines and the water environment, National Rivers Authority, Water Quality Series Report, HMSO, 1994.

4 Mitchell P, Prediction, prevention, control and treatment of acid rock drainage, in *Environmental Policy in Mining, Corporate Strategy and Planning for Closure*, edited by A Warhurst & L Noronha, (Lewis Publisher,) 1999, 117-143.

5 Banks D, The abandonment of the Killingdal sulphide mine, Norway: A saga of acid mine drainage and radioactive waste disposal, *Mine Water & Environment*, **13** (1994) 35-48.

6 Santos S, Machado R, Correia M J N & Cavalho J R, Treatment of mining waters, *Minerals Engg,* **17** (2004) 225-232.

7 Ghosh S K Singh T P N & Tiwari R K, Quality of mine waters in Jharia coalfields, *IAWPC TECH Annual,* **11** (1984) 25-28.

8 Feng D, Aldrich C & Tan H, Treatment of acid mine water by use of heavy metal precipitation and ion exchange, *Minerals Engg,* **13** (2000) 623-642.

9 Voleski B, Detoxification of metal-bearing effluents: Biosorption for the next century, *Hydrometallurgy*, **59** (2001) 203-216.

10 Khandelwal M & Singh T N, Prediction of waste dump stability by an intelligent approach, *Nat Symp New Equipment – New Technology, Management and Safety*, 2002, 38-45.

11 Kosko B, Neural networks and fuzzy systems: A dynamical systems approach to machine intelligence (Prentice Hall of India, New Delhi) 1994, 12-17.

12 Simpson P K, Artificial neural system – Foundation, paradigm, application and implementations (Pergamon Press, New York) 1990.

13 Khandelwal M, Application of neural network for the prediction of triaxial constants from uniaxial compressive strength, M Tech Thesis, Deptt Mining Engg, Banaras Hindu University, Varanasi, 2002.

14 MacKay D J C, Bayesian interpolation, *Neural Computation,* **4** (1992) 415-447.