Application of neural network techniques to predict the heavy metals in acid mine drainage from South African mines

John Kabuba* and Andani Valentia Maliehe
Department of Chemical Engineering, Faculty of Engineering and Technology, Vaal University of Technology, Vanderbijlpark, Private Bag X021, South Africa
*Corresponding author. E-mail: johnka@vut.ac.za

ABSTRACT

Acid Mine Drainage (AMD) is the formation and movement of highly acid water rich in heavy metals. Prediction of heavy metals in the AMD is important in developing any appropriate remediation strategy. This paper attempts to predict heavy metals in the AMD (Zn, Fe, Mn, Si and Ni) from South African mines using Neural Network (NN) techniques. The Backpropagation (BP) neural network model has three layers with the input layer (pH, SO4\textsuperscript{2-} and total dissolved solids (TDS)) and output layer (Cu, Fe, Mn and Zn). After BP training, the NN techniques were able to predict heavy metals in AMD with a tangent sigmoid transfer function (\textit{tansig}) at hidden layer with five neurons and linear transfer function (\textit{purelin}) at output layer. The Levenberg-Marquardt back-propagation (\textit{trainlm}) algorithm was found as the best of 10 BP algorithms with mean-squared error (MSE) value of 0.00041 and coefficient of determination (R) for all (training, validation and test) value of 0.99984. The results indicate that NN can be considered as an easy and cost-effective technique to predict heavy metals in the AMD.

Key words: acid mine drainage, heavy metals, neural networks, prediction

HIGHLIGHTS

- NNs are applied for prediction of the heavy metals in AMD by considering pH, SO4\textsuperscript{2-} and TDS.
- Sampling is time consuming and highly costly for analysis.
- Levenberg-Marquardt BP training algorithm was found the best with \( R = 0.00041 \).
- NN is a viable method to predict heavy metals rapidly and cost-effectively in the AMD.
Optimal NN structure using the *trainlm* BP algorithm

Training, validation, and test MSE for the LMA for the Levenberg-Marquardt back-propagation (*trainlm*) algorithm
INTRODUCTION

Mines’ activities are the major contributors to acid mine drainage (AMD), which poses considerable risk to the environment (Fosso-Kankeu 2018). AMD environmental impacts are provoked by the low pH and high dissolved metals concentrations, which are toxic (Rodríguez-Galán et al. 2019). The most toxic dissolved metal in AMD is Fe, presented in Table 1.

AMD from mines’ activities has been reported to be derived mostly from the oxidation of pyrite. The oxidation occurs on the pyrite in two stages, of which the initial stage produces sulphuric acid and ferrous sulphate (FeSO4) and the next one produces orange-red ferric hydroxide (Fe(OH)3) and additional sulphuric acid (McCarthy 2011). The combination of water, oxygen and oxidizing bacteria causes the pyrite and sulphide minerals in mine wastes to oxidize and form AMD. Pyrite oxidation is quite complex; therefore different reactions can represent it under different conditions (Kefeni et al. 2017). Some examples of reactions for the major and common pyrite oxidation processes are as follows:

The initial reaction involves sulphide mineral oxidation into dissolved Fe2+, SO42− and H+, which Equation (1) illustrates:

\[
\text{FeS}_2 + \frac{7}{2} \text{O}_2 + \text{H}_2\text{O} \rightarrow \text{Fe}^{2+} + 2\text{SO}_4^{2-} + 2\text{H}^+ \tag{1}
\]

An increase in total dissolved solids (TDS) and acidity in the water is represented by the dissolved Fe2+, SO42− and H+. If they are not neutralized, they cause pH decrease. If the environment contains enough oxygen, pH and bacteria activity which are the oxidizing requirements, then a lot of the ferrous iron (FeO) will be oxidized to ferric iron (Fe3+) as shown in Equation (2):

\[
\text{Fe}^{2+} + \frac{1}{4} \text{O}_2 + \text{H}^+ \rightarrow \text{Fe}^{3+} + \frac{1}{2} \text{H}_2\text{O} \tag{2}
\]

When the pH values lie between 2.3 and 3.5, Fe3+ precipitates as Fe(OH)3 and jarosite, which then leaves a little Fe3+ in solution while the pH is simultaneously lowered as shown in Equation (3):

\[
\text{Fe}^{3+} + 3\text{H}_2\text{O} \rightarrow \text{Fe(OH)}_3 + 3\text{H}^+ \tag{3}
\]

Any Fe3+ in the second equation which does not precipitate from solution through Equation (3) can be used to oxidize additional pyrite as shown in Equation (4):

\[
\text{FeS}_2 + \frac{15}{2} \text{O}_2 + \frac{7}{2} \text{H}_2\text{O} \rightarrow \text{Fe(OH)}_3 + 2\text{SO}_4^{2-} + 4\text{H}^+ \tag{4}
\]

Table 1 | Composition of AMD sample in literature

| Element | Concentration (mg/L) |
|---------|----------------------|
|         | Edraki et al. (2005) | Nieto et al. (2013) | Ayora et al. (2016) | Vital et al. (2018) | Ryu et al. (2019) |
| Al      | 76–3,073             | 163                 | 25.1               | 293                 | 150               |
| Ca      | 464–514              | 81.6                | 441                | 313                 | 170               |
| Fe      | 13–1,487             | 153                 | 744                | 134                 | 340               |
| Mg      | 1,280–1,664          | 199                 | 1,104              | 436                 | 220               |
| SO4     | 8,390–22,700         | 2,122               | 11,700             | 8,250               | 4,300             |
| Cu      | 3–138                | 11.7                | 165                | 615                 | 90                |
| Mn      | 22–230               | 21.9                | 467                | 203                 |                   |
| Zn      | 2–81                 | 51.2                | 976                | 68.5                | 120               |
A combination of Equations (1)–(3) represents acid generation that produces Fe\(^{2+}\), which eventually precipitates as Fe(OH)\(_3\):

\[
\text{FeS}_2 + 14\text{Fe}^{3+} + 8\text{H}_2\text{O} \rightarrow 16\text{Fe}^{2+} + 2\text{SO}_4^{2-} + 16\text{H}^+ \tag{5}
\]

The overall equation for stable Fe\(^{3+}\) that is used to oxidize additional pyrite is a combination of Equation (1)–(3):

\[
\text{FeS}_2 + \frac{15}{2} \text{O}_2 + \frac{1}{2} \text{H}_2\text{O} \rightarrow \text{Fe}^{3+} + 2\text{SO}_4^{2-} + \text{H}^+ \tag{6}
\]

All the equations, except for Equations (2) and (3), have assumed that the mineral being oxidized is pyrite and oxygen is the oxidant. There are other sulphide minerals, like pyrrhotite (FeS) and chalcocite (Cu\(_2\)S), which have other ratios of metal sulphide and metals other than Fe. There are different reaction pathways, rates and stoichiometries for additional oxidants and sulphide minerals; however, with limited research on those variations (Akcil & Koldas 2005).

In South Africa and other parts of the world, there is interconnectivity in the ground and surface water systems, which makes it possible for high volumes of contaminated water to affect the water supply (Ghadimi 2015). Aluminium (Al) could cause a number of problems within the distribution network. Aluminium precipitates can increase the final water turbidity and my enmesh micro-organisms to obstruct the disinfection process (Arendze & Sibiya 2017). The presence of iron (Fe) and manganese (Mn) results in staining as well as offensive tastes and appearances, an aesthetic and operational concern. In moderate doses, it is considered as an essential nutrient for human health (Demiral et al. 2021). Other heavy metals that carry problematic effects include zinc (Zn). High concentration of Zn in human beings causes eminent health problems such as nausea, vomiting, anaemia, stomach cramps, skin irritation and cholesterol problems (Fu & Wang 2011). Nickel (Ni) exceeding its critical level might cause serious lung and kidney problems, gastrointestinal distress, pulmonary fibrosis and skin dermatitis (Kabuba & Banza 2020). High concentration of copper (Cu) brings illnesses in brain, kidneys, immunological system and haematological system. Copper consumption may also lead to serious toxicological concerns, since it could be deposited in the brain, skin and liver (Moharbi et al. 2020). These toxic heavy metals are becoming serious environmental problems and cause severe health problems.

Various investigations have been conducted on the behaviour of the heavy metals in AMD (Cánovas et al. 2007; Rooki et al. 2011; Fosso-Kankeu 2018). The conventional technique of measuring the heavy metals involves sampling and a time-consuming and high-cost laboratory analysis. Therefore, to predict the concentrations of heavy metals in water affected by AMD, it is mandatory to develop an appropriate technique to remediate and monitor for comprehensive assessment of the potential environmental impact of AMD. Few studies have been conducted for prediction of heavy metals in AMD using Neural Networks (NNs).

NNs are information processing based on the structure of a biological neural system (Rooki et al. 2011; Kabuba et al. 2014). Due to complexity of the technique, it is difficult to predict using conventional mathematical modelling. NNs is considered a promising method because of their simplicity toward prediction and modeling (Elmolla et al. 2010). In many diverse fields, NNs have been applied over the years, and such applications have been found successful and satisfactory. Some of the applications include the recognition of patterns, processing of signals and images, system identification and modelling as well as predictions of stock market. The reason for such success can be attributed to the fact that since NNs are parallel methods of processing information, they are able to extract relationships that are nonlinear and complex (Ghadimi 2015).

The main objective of the study is to develop an NN model to predict heavy metals in acid mine drainage from South African mines.

**METHODOLOGY**

**Collection of raw and neutralized AMD samples**

The raw samples were collected from the pipeline where acid water enters the treatment plant (influent). The raw sample is the acid water prior to any chemical addition. The point of sample collection was opened and water was allowed to run for 10 min. The samples were then taken using 2 L bottles and immediately stored in a dark space.

The water sample looked like pure water until it was allowed to stand; in the process of storage, a further oxidation reaction of dissolved Fe\(^{3+}\) ions to form ferric ions in the sample takes place. The original colour of wastewater samples changes to...
brownish and reddish. Thereafter, hydrolysis of Fe$^{3+}$ with water forms the solid Fe(OH)$_3$ (ferrihydrite), which is orange-red in colour, and the release of additional acidity. This reaction is pH dependent and under very acidic conditions of less than about pH 3.5, the solid mineral does not form and Fe$^{3+}$ remains in solution. However, the sample pH was about 4, which is a higher value, and a precipitate was formed, commonly referred to as ‘yellow boy’ (Fosso-Kankeu 2018). The neutralized sample was taken from an area in the plant where treatment has already taken place by adding 95% Ca(OH)$_2$ for neutralization. The solid precipitations were reincorporated in the water sample via acidification by adding HNO$_3$ prior to analysis.

**Analysis for AMD properties and analytical techniques**

- **Inductively Coupled Plasma – Optical Emission Spectrometry (ICP-OES) analysis**

The acid mine water samples were submitted on the same day after collection to Setpoint laboratories in Johannesburg (South Africa) for the ICP-OES analysis, using a model number Varian 700-ES. The water samples were characterized by ICP-OES analysis to determine the concentrations of inorganic elements contained in these samples. The prepared standards of inorganic elements to be detected were used to calibrate the ICP-OES equipment for the ICP-OES analysis of the selected inorganic elements (transition elements, lighter inorganic elements and hazardous heavy elements) in the leachate samples (Olesik et al. 1994). In this process 0.2 g from reference material for digestion was used. Wet digestion of samples was performed using 5 mL mixtures of acids: HNO$_3$:HClO$_4$ (3:1). Each sample was heated up to 180 °C for 3 h on the heating digestion block. Then the acid digest sample was allowed to cool and filtered into a 25 mL volumetric flask, using Whatmann filter paper, and made up to mark with de-ionized water. The blank digest sample was similarly processed. Calibration curves for elements were created from mixtures of high-purity element standard solutions (Merck, Germany) in a 4% (v/v) HNO$_3$ matrix.

- **Ion chromatography (IC analysis)**

About 20 μL sample was injected into a Metrosep A Supp 4–250/4 anion-exchange column (stationary phase), which was held at 25 °C, with a pressure of 5.83 Mpa and a flow rate of 1.00 mL/min. Before analysis, the liquid samples were filtered through 0.22 μm Millipore filter paper. The IC analysis of the sample was carried out under isocratic conditions using disodium carbonate (Na$_2$CO$_3$) (1.8 mmol/L) and sodium hydrogen carbonate (NaHCO$_3$) (1.7 mmol/L) as mobile phase with a pH 10.30. Different standards were used during the IC analysis including fluoride 2.0 mg/L, chloride 2.0 mg/L, nitrite 5.0 mg/L, bromide 10.0 mg/L and nitrate 10.0 mg/L.

- **EC, TDS and pH values measurements**

The electrical conductivity (EC), which indicates the levels of salinity of the water, TDS and pH of the water samples were recorded using a Hana HI 991301 pH meter with portable EC/pH/TDS/temperature probe.

- **NN model**

In this work, data sets of 28 samples were collected to develop the NN model and were divided into input and target matrix. The input variables for the NN model were selected based on the physical and chemical characteristics that have greater impact on AMD. These characteristics appear mostly in water that contains heavy metals; they are considered to have most dependence on heavy metals. The targets were chosen based on the heavy metals that are high in concentration in the sample taken and are considered unhealthy for service water and need to be removed before AMD is released to the public. The input variables identified were decided to be pH, SO$_4^{2-}$ and TDS while the targets were chosen to be Zn, Fe, Mn, Si and Ni.

The reason for selecting Zn, Fe, Mn, Si and Ni was based on the fact that they were found to be in high concentrations in the Western Basin AMD plant samples and according to literature.

**Neural network procedure**

**Data processing**

A database is critical when modelling an NN. In the first part, the database was generated by collecting a large number of data points from the experimental data. After evaluating all the experimental results, the collected data were arranged in a set of input vectors as a column in a matrix. Then another set of target vectors was arranged (the correct output vectors for each of the input vectors) to a second matrix in an MS Excel sheet. The input variables were pH, SO$_4^{2-}$ and TDS. The corresponding Zn, Fe, Mn, Si and Ni were used as a target. To ensure that all variables in the input data are important, principal component
Analysis (PCA) was performed as an effective procedure for the determination of input parameters. In this research, a multiple-layered perceptron (MLP) type BPNN was used for modeling.

**NN design procedure**

Five important aspects that must be determined in the design procedure of NNs are as follows: (1) selection of the BP training algorithm, (2) data distribution, (3) selection of the NN structure, (4) selection of the initial weight and (5) sensitivity analysis.

**Selection of inputs, targets and functions**

The data points, which are the results of the laboratory analyses, are exported to the Neural Network Toolbox V4.0 MATLAB, 2020a mathematical software. The inputs (three) and targets (five) were defined from the exported data. Selections of training, adaptation learning and performance functions were done.

**Train NN using information and parameters**

The training was done using the input and target as training information. The parameters include information such as epochs, goal, time and showing of command line.

**Trial and error**

Testing of the NN was done using parameters mentioned above as well as the gradient, the validation check parameter was used to check the validity of the NN and the overall data regression as presented in the plots of the NN. Fitting was observed; if the fitting is over or under, then there is re-initializing and editing of weights. The cycle of training, testing, validation and regression was done until the fitting was good. When the fitting was good, the output and error results were retrieved, and comparison was done between the targets and outputs.

**RESULTS AND DISCUSSION**

**Analysis of raw AMD and treated AMD samples**

The water characteristics results of the raw samples and samples treated with Ca(OH)$_2$ are shown in Table 2. The results are of the two representative samples.

The results obtained indicated that the lowest recorded pH of the raw AMD was 4.17 and the sulphate content was found to be between 1,627 mg/L (minimum) and 1,634 mg/L (maximum), which is very high. 600 mg/L is the maximum point where human health will not be affected by sulphate concentrations (Penner et al. 2020). This confirms that the samples analysed were indeed AMD. According to Moodley et al. (2017), AMD is generally characterized by low pH, high heavy metal content and high salinity, although the sulphate and metal concentrations in the water vary based on the mine.

Samples treated with Ca(OH)$_2$ were of cleaner standards. The data shows an increase in pH from 2.57 to a maximum of 8.26 due to the addition of Ca(OH)$_2$. This also leads to the total alkalinity of the treated samples increasing due to the addition

| Water characteristics | Sample numbers | 1 | 2 |
|-----------------------|---------------|---|---|
|                       | Raw           | Treated | Raw | Treated |
| Acidity (mg/L CaCO$_3$) | 258           | ≤0.01   | 250 | ≤0.01   |
| Conductivity (mS/m 25 °C) | 362           | 331     | 364 | 327     |
| Dissolved Organic (mg/L) | 2.50          | 2.50    | 2.50 | 2.50    |
| pH                    | 4.17          | 7.90    | 4.17 | 8.26    |
| Sulphate (mg/L)       | 1,627         | 1,510   | 1,634 | 1,520   |
| Total Alkalinity (mg/L CaCO$_3$) | <0.10        | 25.8    | <0.10 | 20.2    |
| TDS (mg/L @180 °C)    | 2,357         | 2,153   | 2,370 | 2,127   |
| Total hardness (mg/L) | 2,000         | 200     | 2,000 | 200     |
| Total Organic Carbon  | <0.10         | <0.10   | <0.10 | <0.10   |
of Ca. An insignificant reduction in TDS is observed in the neutralized solutions when compared to the raw solution on the samples, which means that treatment of AMD has very little impact in reducing TDS. This can be attributed to the fact that ions such as Al, Ag, Cr, Mo, Fe, and Pb react with OH\textsuperscript{-} from the added Ca(OH)\textsubscript{2} at high pH of 8.2 to form metal hydroxide precipitates (Penner et al. 2020). Conductivity indicates the level of salinity of water, and AMD conductivity is high due to high salinity. The addition of Ca(OH)\textsubscript{2} increases pH and lowers the salinity of the AMD, and therefore the conductivity of the treated solution also decreased, as seen in Table 2.

Heavy metals were identified, with Zn, Fe, Mn, Si and Ni being the ones with high concentrations. Heavy metals analysis was done on the raw AMD and treated AMD samples to determine if the heavy metals were removed by the neutralization process in order get the water to cleaner standards. Table 3 shows that in the presence of oxygen, ferrous iron was oxidized to ferric iron, which precipitated at a pH of lower than 4 in other samples.

Ferric hydroxide formed a yellowish-orange solid known as yellow boy, which usually precipitates at pH greater than 3.5. Therefore, when the pH increased to 8.26, ferric hydroxide was precipitated. In all samples, Fe went from high concentrations to <10.0 mg/L, which is below detection. This can be seen in sample 1, where the initial concentration of Fe went from 289 mg/L to <10.0 mg/L.

Manganese (Mn) precipitation is variable due to its many oxidation states, but it generally precipitates at a pH of 9.0 to 9.5. In samples that reached a pH of 9.0, it was precipitated. In samples that did not reach a pH of 9.0, Fe precipitation largely removed Mn from the water at pH 8 due to co-precipitation, because the iron concentration in the water was much greater than the manganese content. Some of the Mn remains as dissolved Mn and in very small concentrations.

Complete precipitation of Zn occurs at pH 10.1; however, the pH of the samples was only increased to 8.26. At these pH values, most of the Zn forms zinc hydroxide and precipitates out of solution, while some of the Zn remains as dissolved Zn. In sample 1, for instance, the concentration of Zn was reduced from 0.33 mg/L to 0.07 mg/L. The same concept applies for Ni, which completely precipitates at pH 10.8 to 11. At pH between 7.0 and 8.0, most of the Ni forms nickel hydroxide and precipitates out of solution, leaving the rest as dissolved Ni. In sample 1, for instance, the concentration of Ni was reduced from 0.521 mg/L to 0.0272 mg/L.

**Application of NN**

According to Toma et al. (2004), to design NN, the important aspects must be determined: (1) selection of the BP training algorithm, (2) data distribution, (3) selection of the NN structure and (4) selection of the initial weight.

- Selection of the BP training algorithm

Ten BP training algorithms were compared, as illustrated in Table 4, in order to select the best suitable BP training algorithm.

The NNToolbox was used, where the laboratory results of the samples were imported. In all the BP training algorithms, a three-layer NN with a tangent sigmoid transfer function (tansig) at the hidden layer and a linear transfer function (purelin) at the output layer were used. The chosen training algorithm was the Levenberg-Marquardt back-propagation (trainlm) due to the fact that it had the smallest mean squared error (MSE) of 0.00041, meaning the error of this algorithm is very low. For

**Table 3** | Heavy metal analysis results for AMD (Raw and Treated)

| Heavy metals (mg/L) | Sample numbers |    |    |
|---------------------|----------------|----|----|
|                     | 1 Raw          | 1 Treated | 2 Raw | 2 Treated |
| Zn                  | 0.33           | 0.07 | 0.36 | 0.08 |
| Fe                  | 289            | <0.10 | 289 | <0.10 |
| Mn                  | 45             | 0.0533 | 46.5 | 0.0502 |
| Ni                  | 0.521          | 0.0272 | 0.490 | 0.0203 |
| Si                  | 8.56           | 0.37 | 8.58 | 0.38 |
cross validation of NN, coefficient of determination ($R$) and $MSE$ were evaluated. The $MSE$ is elaborated in Equation (7):

$$MSE = \frac{\sum_{N} (SP_{\text{cal}} - SP_{\text{exp}})^2}{N}$$

(7)

### Table 4 | Comparison of 10 BP algorithms

| BP algorithms                        | Function | MSE    | BLE               |
|--------------------------------------|----------|--------|-------------------|
| 1. Batch gradient descent            | traingd  | 0.54628| $y = 0.94x - 11$  |
| 2. Batch gradient descent with momentum | traingdm | 0.60734| $y = 0.94x - 11$  |
| 3. BFGS quasi-Newton back-propagation | trainbfg | 0.07108| $y = x + 4.5$     |
| 4. Fletcher-Reeves conjugate gradient back-propagation | traincfg | 0.15623| $y = x + 2.4$     |
| 5. Levenberg-Marquardt back-propagation | trainlm  | 0.00041| $y = x + 1.4$     |
| 6. One step secant back-propagation  | trainoss | 0.17672| $y = x - 1.9$     |
| 7. Polak-Ribiere conjugate gradient back-propagation | traincgp | 0.12782| $y = x + 2.5$     |
| 8. Powell-Beale conjugate gradient back-propagation | traincgb | 0.15873| $y = 0.99x + 0.86$|
| 9. Scaled conjugate gradient back-propagation | trainscg | 0.13185| $y = x - 1.3$     |
| 10. Variable learning rate back-propagation | traingdx | 0.64620| $y = x + 0.78$    |

**Figure 1** | Training, validation and test MSE for the LMA for the Batch gradient descent (traingd) algorithm.
where subscripts cal and exp denote calculated and experimental values of SP (set point), respectively. $N$ is the number of validation and training data.

The line of best fit for the data set used in this experiment, also known as BLE, is described by Equation (8):

$$y = bX + a$$ 

(8)

This equation allows the estimation of the value of a dependent variable ($y$) from a given independent variable ($X$), where $b$ defines the slope of the line and $a$ is the intercept. The best linear equation that was chosen after training and testing was found to be $y = x + 1.4$. It gave a clear straight line with the values for training, validation and test $R$ being 0.99908.

Data distribution
The NN model was based on the selected BP algorithm, Levenberg-Marquardt back-propagation (trainlm) for the experimental data. This was applied to train the NN. During training, the output matrix was computed by a forward pass (feed-forward

![Figure 2](image-url)  

*Figure 2* | Training, validation and test MSE for the LMA for the Batch gradient descent with momentum (traindm) algorithm.
BP NN) in which the input matrix is propagated forward through the network to compute the output value of each unit. The output matrix was then compared with the desired matrix, which results in an error signal for each output unit. In order to minimize the error, appropriate adjustments were made for each of the weights of the network. The training was stopped after iterations for the Levenberg-Marquardt algorithm (LMA) where the differences between training errors and validation errors were starting to increase. Eight iterations of training were performed with different weights being adjusted. In order to minimize the error, appropriate adjustments were made for each of the weights of the network. Figures 1–10 present the different algorithms: traingd, traindm, trainbfg, traincfg, trainoss, traincgp, trainscg, traindx, and trainlm.

The optimum algorithm was found to be trainlm for training, validation and test MSE as shown in Figure 10. The Levenberg-Marquardt back-propagation (trainlm) algorithm resulted in an $R$ value of 0.99993 during training, 0.99998 during validation, 0.9993 during testing, and a summary of all the stages resulted in an $R$ of 0.99984. This means that the output predicted by the network is nearly an exact fit to the output from the laboratory analysis, and this is shown by the MSE being 0.00041. This algorithm was found to be the one that resulted in the optimum structure because it had the smallest MSE value and the BLE showed better fit than the other algorithms. This means that the ideal algorithm to use in training the NNToolbox for the prediction of heavy metals in mine water is the Levenberg-Marquardt back-propagation, and the results of the output predicted by the network will be of great closeness to the output from the laboratory analysis by an MSE value of 0.00041. The training parameters for this algorithm are presented in Table 5.
Selection of NN structure

For the best performance of the NN structure to be determined, it was necessary for the optimal network architecture to be defined. The number of hidden layers and number of neurons in it were determined based on the minimum value of MSE of the training and prediction set. The minimum value of MSE was 0.00041 using the Levenberg-Marquardt back-propagation (trainlm). Figure 11 shows the optimal structure: 3-5-6 with three neurons in input layer, five neurons in hidden layer and six neurons in output layer.

The network was found to be fully connected. This means that there was a connection of every neuron in each layer to every neuron in the next layer. This is how the NN structure was named, based on the number of neurons in each layer.

![Figure 4](http://iwaponline.com/wst/article-pdf/84/12/3489/980144/wst084123489.pdf)

**Figure 4** | Training, validation and test MSE for the LMA for the Fletcher-Reeves conjugate gradient back-propagation (traincgf) algorithm.
Figure 12 shows the optimum NN structure in detail when the Levenberg-Marquardt back-propagation (trainlm) is applied on a three-layer NN with a tangent sigmoid transfer function (tansig) at the hidden layer and a linear transfer function (purelin) at the output layer. There is only one hidden layer, which consists of five neurons.

Selection of initial weight
According to literature, an important problem encountered when training an NN was the determination of the appropriate initial values for the connection weights (Kabuba et al. 2014). These weights are modified during utilization so as to satisfy a
criterion of performance. NN basically adds up the signal that comes from its inputs and multiplies them with the corresponding weights. If the result goes beyond the threshold, the neuron is able to fire and transmit a signal at the output using a transfer function (Kabuba et al. 2014).

The effective weight initiation is associated with performance characteristics such as the time needed to successfully train the network and the generalization ability of the trained network (Adam et al. 2014). The wrong choice of initial weights can lead to an increase in the training time or can even cause the non-convergence of the training algorithm. To decide on the

![Figure 6](https://iwaponline.com/wst/article-pdf/84/12/3489/980144/wst084123489.pdf)

**Figure 6** | Training, validation and test MSE for the LMA for the Polak-Ribiere conjugate gradient back-propagation (traincgp) algorithm.
initial weight for NN training, the Garson equation was used (Aber et al. 2009).

\[
I_j = \frac{m = 1}{N_h} \sum_{m = 1}^{N_h} \left( \frac{|W_{jmn}|}{\sum_{k = 1}^{N_i} |W_{km}|} x_{j}^{W_{kmo}} \right)
\]

In Equation (9), \( W \) is the connection weight. The superscripts 'I', 'h' and 'o' refer to input, hidden and output layers, respectively and subscripts 'k', 'm' and 'n' refer to input, hidden and output neurons, respectively.

The initial weights to layer 1 from input 1 using the Levenberg-Marquardt back-propagation algorithm were the following: [1.2251; −1.5666; −1.3327; 1.4726; −0.80382; 1.7078; −1.5087; 0.18962; 1.8491; 1.6045; 1.7758; −0.056764; 0.55684; 1.9558; 1.2663]. These weights resulted in the decision to make the Levenberg-Marquardt back-propagation (\textit{trainlm})

![Figure 7](image-url) | Training, validation and test MSE for the LMA for the Powell-Boole conjugate gradient back-propagation (\textit{traincgb}) algorithm.
algorithm the optimum one because it resulted in the MSE of 0.00041. The MSE using the combination of this algorithm and weights gave the smallest MSE after testing 10 algorithms, meaning the error of this algorithm is very low. These weights also made the training time short and resulted in the BLE after training and testing being $y = x + 1.4$. It gave a clear straight line with the values for training, validation and test $R$ being 0.99908.

**CONCLUSION**

The application of NN techniques to predict the heavy metals in AMD from South African mines has been presented. The prediction of heavy metals (Zn, Fe, Mn, Si and Ni) using NN techniques with BP algorithm is presented and compared.
with the experimental data. The configuration of the BP giving the smallest mean-squared error was Levenberg Marquardt Algorithm (3-6-5) with a tangent sigmoid transfer function (tansig) at hidden layer and a linear transfer function (purelin) at output layer. NN predicted results are very close to the experimental results, with $R$ (for all training, validation and test) = 0.99984 and $MSE = 0.00041$. NN results showed that NN techniques could effectively simulate and predict the heavy metals in AMD from South African mines.

**Figure 9** | Training, validation and test MSE for the LMA for the variable learning rate back-propagation (traingdx) algorithm.
Figure 10 | Training, validation and test MSE for the LMA for the Levenberg-Marquart back-propagation (`trainlm`) algorithm.

Table 5 | Training parameters

| Parameter          | Value          |
|--------------------|----------------|
| Show window:       | true           |
| Show command line: | false          |
| Show:              | 25             |
| Epoch:             | 1,000          |
| Time:              | Inf            |
| Goal:              | 0              |
| Min_grad:          | 1e-10          |
| Max_fail:          | 6              |
| mu                 | 0.001          |
| mu_dec:            | 0.1            |
| mu_inc:            | 10             |
| mu_max:            | 10,000,000,000 |
DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY STATEMENT

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request. All relevant data are included in the paper or its Supplementary Information.

REFERENCES

Aber, S., Amani-Ghadim, A. R. & Mirzajani, V. 2009 Removal of Cr(VI) from polluted solutions by electrocoagulation: modeling of experimental results using artificial neural network. *Journal of Hazardous Materials* **171** (1–3), 484–490.

Adam, S. P., Karras, D. A., Mogoulas, F. D. & Vrahatis, M. N. 2014 Solving the linear interval tolerance problem for weight initialization of neural networks. *Neural Networks* **54**, 17–37.

Akci, A. & Koldas, S. 2005 Acid mine drainage (AMD): causes, treatment and case studies. *Journal of Cleaner Production* **14** (2006), 1139–1145.

Arendze, S. & Sibiya, M. S. 2017 Removal of aluminium and iron from surface water by conventional water treatment. *Water Supply* **17** (3), 733–734.

Ayora, C., Macias, F. & Torres, E. 2016 Recovery of rare earth elements and yttrium from passive-remediation systems of acid mine drainage. *Environmental Science Technology* **50**, 5255–5262.
Cánovas, C. R., Olías, M., Nieto, J. M., Sarmiento, A. M. & Cerón, J. C. 2007 Hydrogeochemical characteristics of the Tinto and Odiel Rivers (SW Spain). Factors controlling metal contents. Science of the Total Environment 373 (1), 363–382.

Demiral, I., Samdan, C. & Demiral, H. 2021 Enrichment of the surface functional groups of activated carbon by modification method. Surface Interfaces 22, 100873.

Edraki, M., Golding, S. D., Baublys, K. A. & Lawrence, M. G. 2005 Hydro-Chemistry, mineralogy and sulfur isotope geochemistry of acid mine drainage at the Mt. Morgan mine environment, Queensland, Australia. Applied Geochemistry 20, 789–805.

Elmolla, E. S., Chaudhuri, M. & Eltoukhay, M. M. 2010 The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solutions by the Fenton process. Journal of Hazardous Materials 179, 127–134.

Fosso-Kankeu, E. 2018 Synthesized of-PFCl and GG-g-P(AN)/TEOS hydrogel composite used in hybridized technique applied for AMD treatment. Physics and Chemistry of the Earth 105, 170–176.

Fu, F. & Wang, Q. 2011 Removal of heavy metal ions from wastewaters: a review. Journal of Environmental Management 92, 407–418.

Ghadimi, F. 2015 Prediction of heavy metals contamination in the groundwater of Arak region using artificial neural network and multiple linear regression. Journal of Tethys 3 (3), 203–215.

Kabuba, J. & Banza, M. 2020 Ion-exchange process for the removal of Ni (II) and Co (II) from wastewater using modified clinoptilolite: modeling by response surface methodology and artificial neural network. Results in Engineering 8, 100189.

Kabuba, J., Mulaba-Bafubiani, A. & Battle, K. 2014 Neural network technique for modeling of Cu (II) removal from aqueous solution by clinoptilolite. Arabian Journal of Science and Engineering 39, 6793–6803.

Kefeni, K. K., Msangati, T. A. M. & Mamba, B. B. 2017 Acid mine drainage: prevention, treatment options, and resource recovery: a review. Journal of Cleaner Production 151 (2017), 475–493.

McCarthy, T. S. 2011 The impact of acid mine drainage in South Africa. South African Journal of Science 107 (5/6), 1–7.

Moharbi, S. S. A., Devi, M. G., Sangeetha, B. M. & Johan, S. 2020 Removal of copper ions from industrial effluent by Azadirachta indica powder. Applied Water Science 10, 23.

Moodley, Y., Russo, I. R., Dalton, D., Kotzé, A., Muya, S., Haubensak, P., Bálint, B., Munimanda, G. K., Deimel, C., Setzer, A., Dicks, K., Herzig-Straschil, B., Kalthoff, D. C., Siegismund, H. R., Robovský, J., Donoghue, P. & Bruford, M. W. 2017 Extinctions, genetic erosion and conservation options for the black rhinoceros (Diceros bicornis). Scientific Reports 7, 41417.

Nieto, J. M., Sarmiento, A. M. & Canovas, C. R. 2013 Acid mine drainage in the Iberian Pyrite Belt: 1. Hydrochemical characteristics and pollutant load of the Tinto and Odiel rivers. Environmental Science Pollution Research 20, 7509–7519.

Olesik, J. W., Kinzer, J. A. & Harkeroad, B. 1994 Inductively coupled plasma optical emission spectrometry using nebulizers with widely different sample consumption rates. Analytical Chemistry 66, 13.

Penner, G. B., Johnson, J. A., Sutherland, B. D., Clark, L. P. & Elford, C. J. 2020 Effects of drinking water sulfate concentrations on feed and water intake, growth, and serum mineral concentration in growing beef heifers. Applied Animal Science 36 (2), 201–207.

Rodríguez-Galán, A., Fernández-Messina, L. & Sánchez-Madrid, F. 2019 Control of immunoregulatory molecules by miRNAs in T cell activation. Frontiers in Immunology 9, 1–10.

Rooki, R., Ardejani, F. D., Aryafar, A. & Asadi, A. B. 2011 Prediction of heavy metals in acid mine drainage using artificial. Environmental Earth Sciences 64 (1), 1303–1316.

Ryu, S., Naidu, G. & Hasan Joir, M. A. 2019 Acid mine drainage treatment by integrated submerged membrane distillation-sorption system. Chemosphere 218, 953–965.

Toma, F. L., Guessasna, S., Klein, D., Montavon, G., Bertrand, G. & Coddet, C. 2004 Neural computation to predict TiO₂ photocatalytic efficiency for nitrogen oxides removal. Journal of Photochemistry and Photobiology A Chemistry 165, 91–96.

Vital, B., Bartzec, J., Ortega-Bravo, J. C. & Jeison, D. 2018 Treatment of acid mine drainage, West Coast Region, South Island, New Zealand. Journal of Geology and Geophysics 55, 227–237.

First received 30 August 2021; accepted in revised form 1 November 2021. Available online 12 November 2021