2022

Predicting the pyrite oxidation process within coal waste piles using multiple linear regression (MLR) and teaching-learning-based optimization (TLBO) algorithm

Shima Entezam

Behshad Jodeiri Shokri

Shaghayegh Doulati Ardejani

Ali Mirzaghorbanali

Kevin McDougall

See next page for additional authors

Authors

Shima Entezam, Behshad Jodeiri Shokri, Shaghayegh Doulati Ardejani, Ali Mirzaghorbanali, Kevin McDougall, and Naj Aziz

PREDICTING THE PYRITE OXIDATION PROCESS WITHIN COAL WASTE PILES USING MULTIPLE LINEAR REGRESSION (MLR) AND TEACHING-LEARNING- BASED OPTIMIZATION (TLBO) ALGORITHM

Shima Entezam[1], Behshad Jodeiri Shokri[1], Shaghayegh Doulati Ardejani[2], Ali Mirzaghorbanali[1], Kevin McDougall[1] and Naj Aziz[1,3]

Abstract: Coal mining often leads to significant environmental hazards and health concerns when sulfide minerals, particularly pyrite, are associated with coal waste. The oxidation of pyrite typically generates acid mine drainage, a significant problem. This paper presents two mathematical relationships using a teaching-learning-based optimization (TLBO) algorithm for predicting pyrite oxidation and pH changes within a coal waste pile from Alborz-Markazi in northern Iran. A dataset was built based on historical data to achieve this goal. Some influential parameters comprising the depths of the various samples, oxygen fraction, and bicarbonate concentrations were considered as input data, while the outputs were pyrite content and pH. Then, the best statistical relationships were suggested between input and output parameters employing curve and surface fitting methods. Afterward, two multiple linear regression (MLR) models were presented for predicting pyrite content and pH. Also, two relationships have been suggested for predicting the same outputs by applying the TLBO algorithm. Comparison of the results of the latter method with the results obtained using the statistical technique, including correlation coefficient and root mean squared error (RMSE), revealed that the TLBO could predict the outcomes better than the MLR.

BACKGROUND

Tailings and wastes resulting from mining activities are the most abundant volume of solid wastes worldwide. Consequently, mining wastes and tailings cause severe environmental issues, including acid mine drainage (AMD), surface and groundwater bodies, and soil contamination where sulfidic minerals, mainly pyrite and chalcopyrite, are exposed to water and atmospheric oxygen. AMD is characterized by low pH values, high sulfate concentrations, and dissolved heavy metals (Blowes et al., 2003). In addition, the pyrite oxidation can represent the most critical chemical reaction of the AMD generation in the presence of oxygen which results in ferrous iron, sulfate, and acid (Eq. 1):

$$\mathrm{FeS_2} + \frac{7}{2}\mathrm{O_2} + \mathrm{H_2O} \rightarrow \mathrm{Fe^{2+}} + 2\mathrm{So_4^{2-}} + 2\mathrm{H^+} \quad (1)$$

Several critical factors that affect the pyrite oxidation processes are as follows: the presence of oxygen, ferric iron, temperature, presence/absence of bacteria, Eh, and pH. Indeed, these factors/ parameters play crucial roles either in determining strategies for further prevention of pyrite oxidation, creation of control, and management strategies the selection of mineral processing methods, motivating researchers to perform a vast array of studies on relevant topics that include experimental methods (e.g., static and kinetic tests) along with numerical, stochastic, and artificial intelligent (AI) techniques to estimate the parameters affecting the AMD generation over the past decades (Balci and Demirel (2018); Buckley and Woods 1987; Brown and Jurinak 1989; Chen et al. (2020); Dold (2017); Luis et al. (2020); Sasaki et al. 1995; Sebogodi et al. (2019); Wiersma and Rimstidt 1994). Doulati Ardejani et al. (2012) used the GRNN to predict the REEs resulting from neutral alkaline mine drainage (NAMD) at the Razi coal mine in northern Iran. Sadeghimirshahidi et al. (2013) applied the artificial neural networks (ANNs) technique to predict pyrite oxidation within a coal pile. Different input parameters, including the depth of the pile, initial pyrite fraction, diffused oxygen fraction throughout the wastes, and annual precipitation

1 School of Civil Engineering and Surveying, University of Southern Queensland, Queensland, 4350, Australia
2 Third Faculty of Medicine, Charles University, Czech Republic
3 School of Civil, Mining and Environmental Engineering, University of Wollongong, NSW, 2500, Australia

data, were considered. Jodeiri Shokri et al. (2014a) predicted pyrite oxidation throughout coal waste particles by applying an adaptive neuro-fuzzy inference system (ANFIS). Jodeiri Shokri et al. (2014b) suggested several statistical relationships for investigating the pyrite oxidation process in the same circumstances. Bahrami and Doulati Ardejani (2016) applied the ANN-simulated annealing (SA) hybrid method to estimate pyrite contents. The results revealed that the SA could predict better estimations than ANN and statistical methods. Hadadi et al. (2020) applied a probabilistic approach to predict how acid mine drainage is generated within coal waste particles using historical data. Jodeiri Shokri et al. (2020) applied gene expression programming (GEP) for predicting the AMD generation potential throughout copper tailings. Four predictive relationships for the remaining pyrite fraction, remaining chalcopyrite fraction, sulfate concentration, and pH have been suggested by applying GEP algorithms. This article has used multiple linear regression (MLR) and the teaching-learning-based optimization (TLBO) algorithm to suggest alternative relationships for mathematical relationships to predict the AMD generation process.

SITE DESCRIPTION

The Anjir Tangeh coal washing plant is located in Zirab, 45 km south of Qaem Shahr and 3 km from the Qaem Shahr-Tehran road in the Mazandaran province northern Iran (**Figure 1**). The coal tailings dumps have resulted from coal washing at the processing plant for more than 22 years (Shahhoseiny et al., 2011).

Figure 1: A view of the coal tailings dump in Anjir Tangeh coal washing Plant

MLR

A multivariate regression model is a regression model in which more than one regression variable is used. In general, the response variable (y) may depend on n variables (x). Eq. (2) presents an MLR prediction model with n regression variables:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n + \varepsilon \quad (2)$$

where:

y: dependent variable;

ε: model error rate;

and the parameter βj, j = 0,1, ..., n is called the regression coefficients.

The predicted model is a hyperplane in the n-dimensional space of the regression variables xj. However, this method can also consider the prediction models with a more complex (non-linear) structure than Equation (3). For example, in the following model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2^3 + \beta_3 e^{x_3} + \beta_4 x_1 x_2 + \varepsilon \quad (3)$$

A non-linear relationship is sufficient to replace the used variables used with the linear variables to make it easier to analyze the Equation. In Eq. (3), assuming $z_1 = x_1$, $z_2 = x_2^3$, $z_3 = e^{x_3}$, $z_4 = x_1 x_2$, then Eq. (3) converts to the MLR (Eq. (4)):

$$y = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 + \beta_4 z_4 + \varepsilon \quad (4)$$

TLBO ALGORITHM

In (2011), the TLBO algorithm, which requires no algorithm-specific parameters, was presented by Rao et al. Compared to other algorithms, the TLBO algorithm requires only common controlling parameters like population size and the number of generations for its working. Since 2011, the TLBO algorithm has extensively been applied for optimization problems. The algorithm considers the population of solutions as a group of students in a class. It selects the best member of the population as the teacher. The teacher then attempts to train the students to add to their knowledge deals, with the students further adding to their knowledge upon training by communicating with one another. This algorithm goes through two stages: the teacher and the student stages (Rao et al., 2011).

The teacher stage

The teacher improves the students' information and knowledge through teaching and training. Equation (5) describes this stage:

$$X_{new,i} = X_{old,i} + r(X_i^{best} - T_F \ M_i) \quad (5)$$

where:

r: a random number in the [0, 1] interval;

TF: teaching factor;

Xibest: the best member of the population at the ith iteration (selected as a teacher);

Mi: mean of the class at the ith iteration;

Xold, i: a member that needs to be taught;

Xnew, i: a taught member.

The Xnew, i would be accepted if it is somehow better than Xold, i (Rao et al., 2011).

The student stage

In this stage, each student exchanges information with another randomly selected student to enhance their knowledge. For the ith member of the population, a member is chosen randomly. Then, if $f(X_i) < f(X_j)$, the ith member is taught based on Equation (6), while Equation (7) is applied otherwise. This stage is performed for every member of the population.

$$X_{new.i} = X_{old.i} + r(X_i - \ X_j) \quad (6)$$

$$X_{new.i} = X_{old.i} + r(X_j - \ X_i) \quad (7)$$

where:

r: a random number in the [0, 1] interval;

Xold, i: a member that needs to be taught;

Xnew, i: a taught member.

If Xnew, i is better than Xold, i, it replaces the Xold, i (Rao et al., 2011).

DATASET

All experimental data were taken from previous studies such as Shahhosseiny et al. (2011), and Hadadi et al. (2019) was collected to build a database. The depth of the tailings within the dump, the fraction of diffused oxygen through the waste particles, concentrations of bicarbonate, and initial pyrite remaining

fractions within the dump were considered as input data. The remaining pyrite fraction and pH values were selected as output data. The range of the input and output parameters is given in **Table (1)**. Then, the MLR and the TLBO algorithms were applied for finding the best relationships between each output and all the inputs. The dataset was divided into two parts: the training and validation datasets, consisting of 70% and 30% of the database. All the datasets were chosen randomly.

Table 1: Range of the datasets

Parameter	Symbol	Change interval	Type of Parameter
Depth of the samples (m)	D	0.00-4.00	Input
Oxygen fraction (%)	O	0.00-0.21	
Bicarbonate concentration (mg/lit)	Bi	1.20-2.51	
Initial of the remaining pyrite fraction (%)	Py (i)	1.08-1.31	
the remaining pyrite fraction (%)	Py	0.9-1.31	Output
pH	pH	3.20-7.40	

MLR RESULTS

For finding the mathematical relationships for the prediction, the relationships between the parameters and the dependent variables were calculated by Table Curve v5.01 (**Tables 2** and **3**).

Table 2: Suggested relationships between input parameters and the remaining pyrite fraction

The relationship between each input parameter and the remaining pyrite fraction	Suggested relationship
$\boldsymbol{Py \propto f(D)}$	$x_1 = D$
$\boldsymbol{Py \propto f(O)}$	$x_2 = e^{-O}$
$\boldsymbol{Py \propto f(Bi)}$	$x_3 = \frac{1}{Bi^2}$
$\boldsymbol{Py \propto f(Py(i))}$	$x_4 = \frac{1}{\sqrt{Py(i)}}$
$\boldsymbol{Py \propto f(D,O)}$	$x_5 = e^{-D} + O$
$\boldsymbol{Py \propto f(D,Bi)}$	$x_6 = D + e^{-Bi}$
$\boldsymbol{Py \propto f(D,Py(i))}$	$x_7 = D + \frac{1}{\sqrt{Py(i)}}$
$\boldsymbol{Py \propto f(O,Bi)}$	$x_8 = O + \frac{1}{Bi}$
$\boldsymbol{Py \propto f(O,Py(i))}$	$x_9 = O + Py(i)$
$\boldsymbol{Py \propto f(Bi,Py(i))}$	$x_{10} = \frac{1}{Bi} + Py(i)$

Table 3: Suggested relationships between input and pH

The relationship between each input parameter and pH	Suggested relationship
$\boldsymbol{pH \propto f(D)}$	$x_1 = e^{-D}$
$\boldsymbol{pH \propto f(O)}$	$x_2 = O$
$\boldsymbol{pH \propto f(B)}$	$x_3 = \frac{1}{Bi^2}$
$\boldsymbol{pH \propto f(D,O)}$	$x_4 = D^{2.5} + O^3$
$\boldsymbol{p}\mathbf{H} \boldsymbol{\propto f(D,B)}$	$x_5 = \sqrt{D} + e^{-D}$
$\boldsymbol{pH \propto f(O,B)}$	$x_6 = O + \frac{1}{Bi}$

These relationships entered the IBM SPSS v25 software in the next step, and the relationships were suggested in predicting the remaining pyrite fraction and pH values (**Tables 4**).

Table 4: the obtained models by using the MLR method for predicting the remaining pyrite fraction and pH

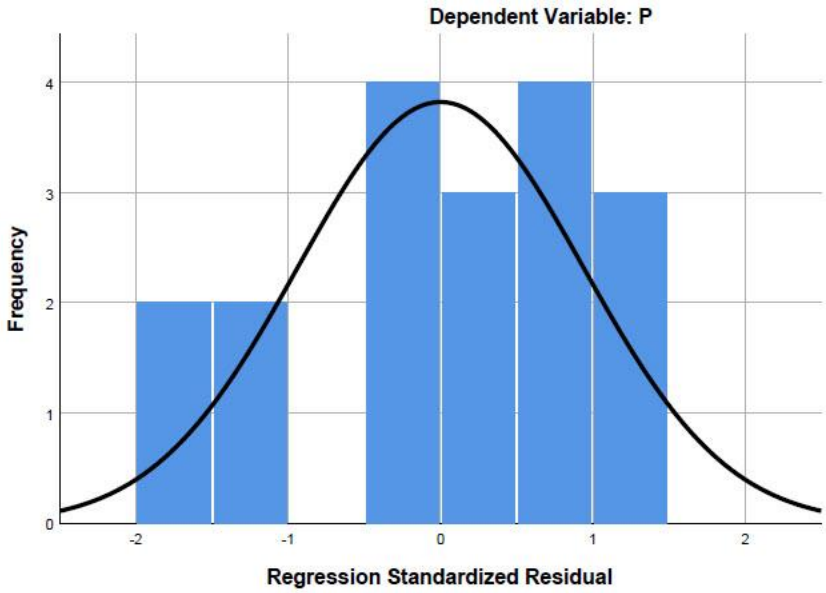
Model number	Parameter	R-Squared	Adjusted R-Squared
1	**Py**	0.938	0.919
2	**Py**	0.936	0.922
3	**Py**	0.933	0.924
1	**pH**	0.723	0.663
2	**pH**	0.722	0.685
3	**pH**	0.712	0.694

By comparing the models, model 3 were chosen as the best model for predicting both the remaining pyrite and pH according to the lower prediction error rate with R-squared and adjusted R-Squared as following relationships (8) and (9):

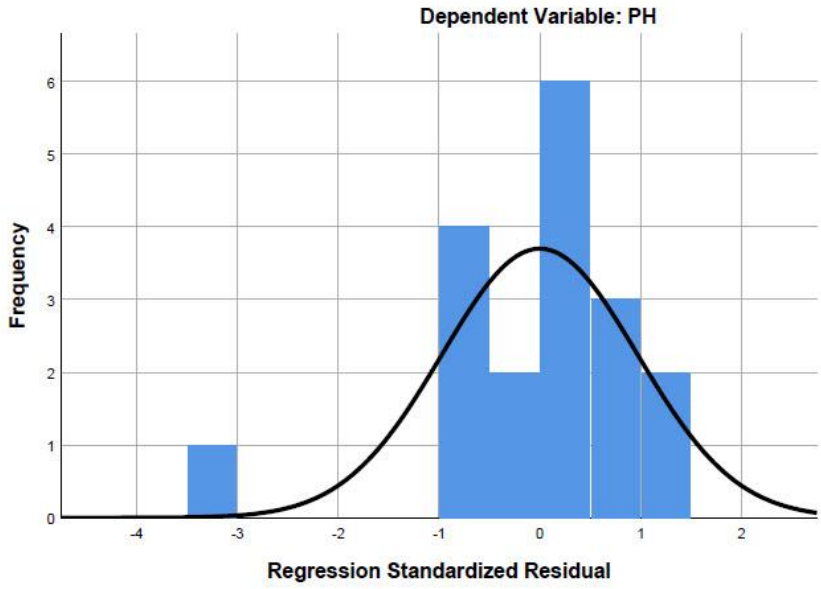
$$Py = 2.9391 + (0.0386 \times D) + (0.1084 \times e^{-O}) - \left(\frac{0.0391}{Bi^2}\right) - \left(\frac{2.1729}{\sqrt{Py\ (i)}}\right) \quad (8)$$

$$pH = 6.7462 + (0.2441 \times e^{-D}) - (14.8327 \times O) + \left(\frac{0.5555}{Bi^2}\right) \quad (9)$$

Figure (2) shows the error analysis histograms of the proposed models. The model error distribution functions are normal functions that prove the accuracy of the suggested models.



(a)



(b)

Figure 2: The error analysis histogram of the proposed model;(a) remaining pyrite fraction and (b) pH values

THE TLBO RESULTS

The TLBO algorithm was used to optimize the MLR relationships' coefficients. For this, the MLR relationships were presented in the following equations (10 and 11):

$$Py = C_1 + C_2 \times D + C_3 \times e^{-O} + \frac{C_4}{Bi^{C_6}} + \frac{C_5}{Py\ (i)^{C_7}} \quad (10)$$

$$pH = C_1 + C_2 \times e^{-D} + C_3 \times O + \frac{C_4}{Bi^{C_5}} \quad (11)$$

RMSE was selected as the objective function (Eq. 12). MATLAB 2017b was used for building the TLBO model. The trial and error method recognized the population size and the maximum number of iterations. These parameters were set to 100 and 400, respectively. The decreasing trends of RMSE were given in **Figure 3**.

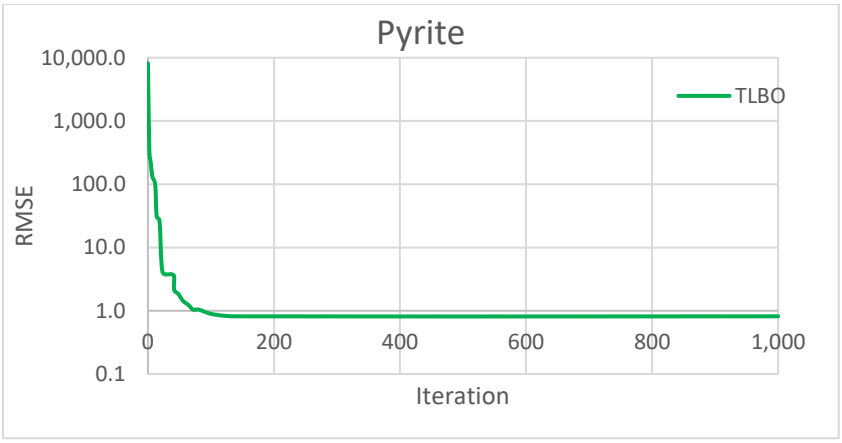
$$Z = \min(RMSE) = \min\left(\sqrt{\frac{1}{n}\ \sum_{i=1}^{n}(P_i - O_i)^2}\right) \quad (12)$$

Where:

n: number of data;

P_i: estimated data

O_i: is actual data



(a)

(b)

Figure 3: Variation of RMSE during the TLBO algorithm wasm being performed

Relationships (13) and (14) presents the modified models for predicting the remaining pyrite fraction and pH values, based on the TLBO algorithm:

$$Py = 1.042 + 0.039 \times D + 0.051 \times e^{-O} - \frac{0.195}{Bi^{21.01}} - \frac{0.49}{Py\ \ (i)^{14.875}} \quad (13)$$

$$pH = 6.895 + 0.001 \times e^{-D} - 15.388 \times O + \frac{7.096}{Bi^{17.373}} \quad (14)$$

MODELS' VALIDATION

After finding the predicted relationships, the models' validations was done. The RMSE and R-squared coefficients have been presented in **Table 5**. As seen, the TLBO relationship could predict the remaining pyrite fraction and pH values better than the MLR prediction.

Table 5: statistical results including RMSE and R-squared values in predicting the remaining pyrite fraction and pH

Predicted Parameter		**Remaining Pyrite**		**pH**	
Methods		**MLR**	**TLBO**	**MLR**	**TLBO**
RMSE	**Modelling**	0.00056	0.00058	0.341	0.303
	Validation	0.01196	0.01517	0.620	0.553
R Square	**Modelling**	0.938	0.94000	0.723	0.753
	Validation	0.801	0.80300	0.702	0.718

CONCLUSIONS

The MLR and TLBO algorithms were applied to suggest four relationships in the AMD generation process based on the pyrite process in coal tailing particles. After collecting the data, it was divided into training and validation datasets. Then, using the TLBO algorithm, two empirical relationships were proposed for predicting the remaining pyrite fraction and pH values based on samples' depth, bicarbonate concentrations, Oxygen fraction, and initial pyrite. Accordingly, the RMSE of the validation models was calculated as 0.015 and 0.553 for pyrite, and pH values, respectively, with the R-squared, being 0.803and 0.718, respectively. The results showed that the TLBO algorithm model offers a more accurate prediction than the MLR.

REFERENCES

Bahrami, S, Doulati Ardejani, F, 2016. Prediction of pyrite oxidation in a coal washing waste pile using a hybrid method, coupling artificial neural networks and simulated annealing (ANN/SA). *Journal of cleaner production*, 137: 1129-1137.

Balci, N, Demirel, C, 2018. Prediction of acid mine drainage (AMD) and metal release sources at the Küre Copper Mine Site, Kastamonu, NW Turkey. *Mine Water and the Environment,* 37(1): 56-74.

Blowes, D W, Ptacek, C J, Jambor, J L., Weisener, C G, Paktunc, D, Gould, W D, Johnson, D B, 2003. *The geochemistry of acid mine drainage*. 10.1016/B978-0-08-095975-7.00905-0

Brown, S, Jurinak, J J, 1989. Pyrite oxidation in aqueous mixtures. *Journal of Environmental Quality,* 18: 545-550.

Buckley, A N, Woods, R W, 1987. The surface oxidation of pyrite. *Applied Surface Science,* 27: 437-452.

Chen, J, Jiskani, I M, Jinliang, C, Yan, H, 2020. Evaluation and future framework of green mine construction in China based on the DPSIR model. *Sustainable Environment Research,* 30 (1): 1-10.

Dold, B, 2017. Acid rock drainage prediction: A critical review. *Journal of Geochemical Exploration,* 172: 120-132.

Doulati Ardejani, F, Rooki, R, Jodieri Shokri, B, Eslam Kish, T, Aryafar, A, Tourani, P, 2012. Prediction of rare earth elements in neutral alkaline mine drainage from Razi coal mine, Golestan Province, northeast Iran, using general regression neural network. *Journal of Environmental Engineering*,139(6): 896-907.

Ferguson K.D., Erickson PM, 1988. *Pre-Mine Prediction of Acid Mine Drainage*. In: Salomons W., Förstner U. (eds) Environmental Management of Solid Waste. Springer, Berlin, Heidelberg.

Ferreira, C, 2001. *Gene expression programming: a new adaptive algorithm for solving problems*. arXiv preprint cs/0102027.

Ferreira, C, 2006. *Gene expression programming: mathematical modelling by an artificial intelligence,* pp 1-478 (Springer)

Hadadi, F, Jodeiri Shokri, B, Zare Naghadehi, M, Doulati Ardejani, F, Shahhosseini, M, 2019. Probabilistic prediction of pyrite oxidation process within Anjir Tangeh coal wastes pile. *Iranian Journal of Mining Engineering,* 14(44): 55-68.

Hadadi, F., Jodeiri Shokri, B, Zare Naghadehi, M, Doulati Ardejani, F, 2020. Probabilistic prediction of acid mine drainage generation risk based on pyrite oxidation process in coal washery rejects-A Case Study. *Journal of Mining and Environment,* 12(1): 127-137.

Jodeiri Shokri, B, Zare Naghadehi, M, Doulati Ardejani, F, 2020. Prediction of Acid Mine Drainage Generation Potential of A Copper Mine Tailings Using Gene Expression Programming-A Case Study. *Journal of Mining and Environment,* 11(4): 11-27-1140.

Jodeiri Shokri, B, Ramazi, H, Doulati Ardejani, F, Sadeghiamirshahidi, M, 2014a. Prediction of pyrite oxidation in a coal washing waste pile applying artificial neural networks (ANNs) and adaptive neuro-fuzzy inference systems (ANFIS). *Mine Water and the Environment*, 33(2): 146-156.

Jodeiri Shokri, B, Ramazi, H, Doulati Ardejani, F, Moradzadeh, A, 2014b. A statistical model to relate pyrite oxidation and oxygen transport within a coal waste pile: case study, Alborz Sharghi, northeast of Iran. *Environmental Earth Sciences*, 71(11): 4693-4702.

Luis, Pino, Beltran, L, Scharwz, A, Cristina Ruiz, M, Borquez, R, 2020. Optimization of nanofiltration for treatment of acid mine drainage and copper recovery by solvent extraction. *Hydrometallurgy*, 195:105361.

Rao, R V, Savsani, V J, Vakharia, D P, 2011. Teaching-learning-based optimization: a novel method for constrained mechanical design optimization problems. *Computer-Aided Design*, 43(3): 303-315.

Rimstidt, J D, Chermak, J A, and Gagen, P M, 1994. Rates of reaction of galena, sphalerite, chalcopyrite and arsenopyrite. In: Alpers CN and Blowes DW (eds.) Environmental Geochemistry of Sulfide Oxidation, vol. 550, pp. 2–13. Washington, DC: American Chemical Society.

Sadeghiamirshahidi, M, Doulati Ardejani, F, 2013. Application of artificial neural networks to predict pyrite oxidation in a coal washing refuse pile. *Fuel*, 104, 163-169.

Sasaki, K, Tsunekawa, M, Ohtsuka, T, Konno, H, 1995. Confirmation of a sulfur-rich layer on pyrite after oxidative dissolution by Fe(III) ions around pH 2. *Geochimica et Cosmochimica Acta,* 59: 3155-3158.

Sebogodi, Keolebogile R, Johakimu, J K, Sithole, B, 2020. Beneficiation of pulp mill waste green liquor dregs: Applications in treatment of acid mine drainage as new disposal solution in South Africa. *Journal of Cleaner Production,* 246: 118979.

Shahhoseiny, M, Doulati Ardejani, F, Shafaei, S Z, Singh, R, Jodeiri Shokri, B. 2011. Geochemical characterisation of a pyrite containing coal washing refuse pile produced by the Anjir Tangeh coal washing plant in Zirab, Mazandaran province, Northern Iran. Mine Water – Managing the Challenges" IMWA 2011, Aachen, Germany. pp: 463-468.

Wiersma, C L, and Rimstidt, J D, 1984. Rates of reaction of pyrite and marcasite with ferric iron at pH 2. *Geochimica et Cosmochimica Acta,* 48: 85–92.