Scientific Research Publishing

Engineering, 2025, 17(7), 370-387
https://www.scirp.org/journal/eng
ISSN Online: 1947-394X
ISSN Print: 1947-3931

Numerical Prediction of Air Overpressure during Mining Operations in an Open Pit Mine

Michael Bulangashane Ndagano[1], Alex Kalonji-Kabambi[1], Augustin Tshifung[2], Albert Agisha Ntwali[1]

[1]School of Mines, Université Officielle de Bukavu, Bukavu, Democratic Republic of Congo
[2]Minerals and Metals Group Limited-Mine Kinsevere, Lubumbashi, Democratic Republic of Congo
Email: michaelbulangashane2001@gmail.com, alexkalonjikab@gmail.com, augustin.tshifung@mmg.com, agishaalbert@gmail.com

How to cite this paper: Bulangashane Ndagano, M., Kalonji-Kabambi, A., Tshifung, A. and Agisha Ntwali, A. (2025) Numerical Prediction of Air Overpressure during Mining Operations in an Open Pit Mine. *Engineering*, **17**, 370-387.
https://doi.org/10.4236/eng.2025.177022

Received: May 25, 2025
Accepted: July 27, 2025
Published: July 30, 2025

Abstract

Blasting is considered an indispensable process in mining excavation operations. Generally, only a small percentage of the total energy of blasting is consumed in the fragmentation and displacement of the rock, and the rest of the energy is transmitted to the structures and environment surrounding the mined area. The air overpressure (AOp) induced by explosions in open-cast mines has unavoidable environmental and safety consequences, but can be minimized to an acceptable threshold to limit environmental damage and the impact on the sustainability of mining activities. The development of numerical predictive models of AOp was the main objective of this study. Thus, the methodology used to achieve this main objective was articulated around six parts: knowledge of the study area, processing and statistical analysis of the data collected, development of both numerical and empirical prediction models, and evaluation of model performance of the numerical model parameters. The results show that only numerical models are suitable for predicting AOp. Moreover, numerical models generally perform better than empirical models in predicting this phenomenon. Among these AI models, the results show that the DT model is the best suited for predicting AOp in this study, with remarkable performance results (RRSE of 0.08, RAE of 0.05, RMSE of 0.29, MAE of 0.37, MAPE of 0.07, and an R^2 of 0.994). This could therefore justify its application in practical engineering to predict blast-induced AOp in open-cast mines to reduce undesirable environmental effects.

Keywords

Numerical Prediction, Artificial Intelligence (AI), Air Overpressure (AOp), Blasting, Open Pit Mining

1. Introduction

Blasting is considered an essential process in mining excavation operations, with the aim of loosening or reducing the massive structure of high-strength bedrock to enable cost-effective mechanical excavation [1]. Drilling and blasting are the most economical and commonly used rock fragmentation techniques, not only in mining operations, but also in several civil engineering applications such as tunnelling and road construction [2]-[6]. Unfortunately, the energy used in mining operations is not only consumed in the fragmentation of the rock mass. Generally, only 20% to 30% of the total energy is consumed in the fragmentation and displacement of the rock [7], and the rest of the energy is transmitted to the rock, structures and environment surrounding the mined area in the form of ground vibrations, flying rock, noise, mound feet and back breakers [8]-[13].

The air overpressure (AOp) resulting from mining operations represents an undesirable effect of the use of explosives. It represents the shock wave that is refracted horizontally by density variations in the atmosphere and gradually dies out with time and distance [14]-[16]. This pressure wave is made up of an audible sound and a subaudible sound. The highest-frequency part (>15 Hz) of the pressure wave emerging immediately near the explosion is audible, while the subaudible part is the lowest-frequency part located in the infrasound region (<15 Hz). The inaudible part generally occurs far from the explosion site [14]. AOp is defined in terms of sound level and is measured in decibels (dB) or pascals (Pa), with 20 Hz being the lowest detectable sound level for humans [17]. Consequently, it is undeniable that there is a concussion risk for humans in the event of exposure to sounds above 20 Hz. The level of AOp likely to cause structural damage is 180 dB, that of glass breakage 130 to 180 dB, and that of window vibrations 110 to 130 dB. Consequently, many measures are taken to keep AOp below 110 dB in critical areas where the public is concerned [17] [18]. According to Khandelwal & Kankar [19], the geometry of the blast pattern, distance from the blast face, amount of explosive charge per delay, geological discontinuities, direction of blast, topography of the blast site, and vegetation are the main parameters that influence AOp. Konya and Walter [20] have indicated that AOp can be controlled by good borehole containment or tamping (in terms of material type and length). In addition, other parameters such as atmospheric conditions, overloading, brittle strata, and conditions resulting from secondary mining can have a direct influence on AOp [21] [22]. However, AOp induced by blasting is difficult to predict, as the same blast pattern can produce different results in different cases [23].

To control the undesirable effects of mining-induced AOp, predictive methods can be applied. A review of the literature has indicated that the empirical and artificial in-telligence (AI) approaches are the two main methods used for this purpose [24]. In empirical methods, the explosive charge per delay (or maximum explosive capacity) (W) and the watch distance (D) have been used to predict the AOp induced by mining operations. Numerous researchers have investigated and developed experimental equations for rapidly predicting AOp with the highest possible degree of confidence [25]-[27] (**Table 1**).

Table 1. Empirical models for predicting AOp.

Predictive approaches of AOp	Equations
National Association of Australian State (1983)	$P = \frac{140\sqrt[3]{E/200}}{d}$
Persson *et al.* (1994)	$P = 0.7 * \frac{\sqrt[3]{W}}{D}$
McKenzine (1990)	$P = 165 - 24\log\left[\frac{D}{\sqrt[3]{W}}\right]$

However, the results of previous studies have shown that empirical methods lack precision [7] [8] [15] [28]-[30]. There are several reasons for the limitations of experimental techniques. Firstly, they generally only consider the relationship between the explosive charge per delay and the monitoring distance. However, other factors influencing AOp, such as those mentioned above, are not considered [31]-[33]. Secondly, empirical techniques are often studied for specific areas. Consequently, their generalizability remains very low, as does their accuracy when applied to other areas. Thirdly, the empirical equations only consider the linear relationship between the influencing parameters. However, most of the influencing parameters have non-linear relationships with the AOp [34].

To overcome these limitations, AI has been studied and developed because of its ability to generalize and explain complex non-linear relationships [24]. Currently, AI using Big Data has become an indispensable tool in all fields, particularly research into the sustainable management of natural resources [35]-[38]. In mining, AI has been widely applied for the prediction and assessment of slope stability [39], prediction of mining problems [32] [40] [41], prediction of access roads in longwall mines [42], prediction of the health risk caused by whole-body vibrations from mining trucks [43], etc. To predict mining-induced AOp, several researchers have studied and successfully developed different AI models. Khandelwal & Singh [28] developed an artificial neural network (ANN) system to predict blast-induced AOp from 56 mining recordings collected in a magnesite mine. Initial results were significant, with a mean absolute percentage error (MAPE) of 2.74 and a corresponding coefficient of determination (R^2) of 0.96. In another study, Hajihassani *et al.* [23] trained an ANN with an evolutionary algorithm (particle swarm optimization—PSO), namely an ANR-based PSO model, using 62 AOp datasets. Their results showed that the ANR-based PSO model correctly predicted the AOp caused by mining operations, with a correlation coefficient of determination (R^2) of 0.94. Hasanipanah *et al.* [44] used adaptive neuro-fuzzy inference system (ANFIS), ANN, fuzzy systems (FS) techniques and an empirical equation to estimate the AOp induced by mining operations.

To develop these models, a group of 77 records was used in their study. Their results revealed that ANFIS was the best-performing approach for predicting AOp. Armaghani *et al.* [12] studied 128 datasets from three quarry sites in Malaysia to predict mining-induced AOp. They successfully developed an ANFIS model with five input variables (maximum explosive load per delay, loading factor, load, tamping height, and monitoring distance) and one output parameter (AOp). Various other methods, such as empirical techniques, multiple regression (MR), and ANN, were also applied to predict blast-induced AOp and compared with the ANFIS model developed. They demonstrated that, in their study, ANFIS was the best model with an RMSE of 2.33 and an R^2 of 0.97. Bui *et al.* [45] also evaluated the performance of different AI techniques for estimating AOp in an open-cast coal mine, including Random forests (RF), boosted regression trees, KNN, SVR, GP (Gaussian process), BART (Bayesian additive regression trees), and ANN. They confirmed the feasibility of the AI techniques mentioned. The ANN model was recommended as the best model in their study for estimating AOp. Nguyen *et al.* [32] also compared and evaluated the performance of different ANN systems, including RNA, hybrid fuzzy neural inference systems (HYFIS), and regularized Bayesian neural networks (RBNN) for predicting blast-induced AOp. They used 146 explosions with 10 parameters to develop predictive models of AOp in the Deo Nai open-pit coal mine in Vietnam. Their results showed that the RNA outperformed the other models with an RMSE of 2.32 and an R^2 of 0.96. Mohamad *et al.* [46] predicted AOp induced by mining operations with an ANN-based genetic algorithm (GA), abbreviated as GA-ANN, using 76 explosions. Empirical and ANN models were also provided to predict AOp and compared with the GA-ANN model. Their results showed that the GA-ANN model outperformed both the empirical and ANN models.

Not wishing to be exhaustive in citing previous work on AOp prediction, the literature review shows that predictive models for AOp have been satisfactorily developed and proposed. Nevertheless, they cannot be applied and represented to other sites, while the effects of blast-induced AOp differ from one site to another [47]. This study aims to evaluate and predict the AOp induced by mining operations in the Kinsevere open pit mine in the Democratic Republic of Congo (DRC), using three ML algorithms: multiple linear regression (MLR), random forest (RF), and decision tree (DT). An empirical model has also been developed to predict and compare with the models presented above.

2. Methodological Approach

2.1. Overview of the Study Area

The data used in this study come from the Kinsevere MMG mine. Located between 27˚30' longitude East and 11˚15' latitude South. The Kinsevere MMG mine is a copper and cobalt mine situated in the Kipushi territory, Haut-Katanga province in the DRC, around 33 km north-north-east (NNE) of the city of Lubumbashi, around 27 km NNW of the Luano airport, and 1 km south of the Lwiswishi mine (**Figure 1**). It has copper (malachite, pseudo-malachite, azurite, chrysocole, chalchosine, chalcopyrite, bornite) and cobalt (heterogenite, caroline) mineralization in three known deposits: Central, Mashi and Kinsevere Hill, hosted in dolomitic schists (DS) and black mineral lime-stones (BML) with minor mineralization occurring within the clayey-talcous rocks (CTR) formation at the contact with the Silicified Dolomite (SD) [48].

The conventional mining method used at the Kinsevere mine is open-pit mining. Shovel-truck mining, with mining limited to 10 m and 15 m benches, is the method commonly used at Kinsevere. Drilling and blasting are the main operations used by the Kinsevere mine to excavate materials. Holes 102 mm (pre-cutting holes), 127 mm (production holes on 10 m benches), and 140 mm (production holes on 15 m benches) in diameter are drilled using an off-hole hammer drill (PANTERA DP1500I) and charged with free emulsion (P100 from Orica) and encapsulated emulsion (SPLITEX/Magnum), respectively, for the production and pre-cutting holes.

The holes are primed using a 100 g or 400 g booster coupled to a 500-millisecond bottom-hole detonator. The surface connection is made using relays whose delay varies between 17 ms and 100 ms, following a precise firing pattern (usually the "Christmas tree"). Vibrations and induced AOp are recorded at each blast using a MINI-SEIS III seismograph (White Industrial Seismology, Inc.) coupled to a sound level meter installed at a precise distance from the blast. A total of 73 mining recordings from this mine were selected to build the AOp prediction models after data processing.

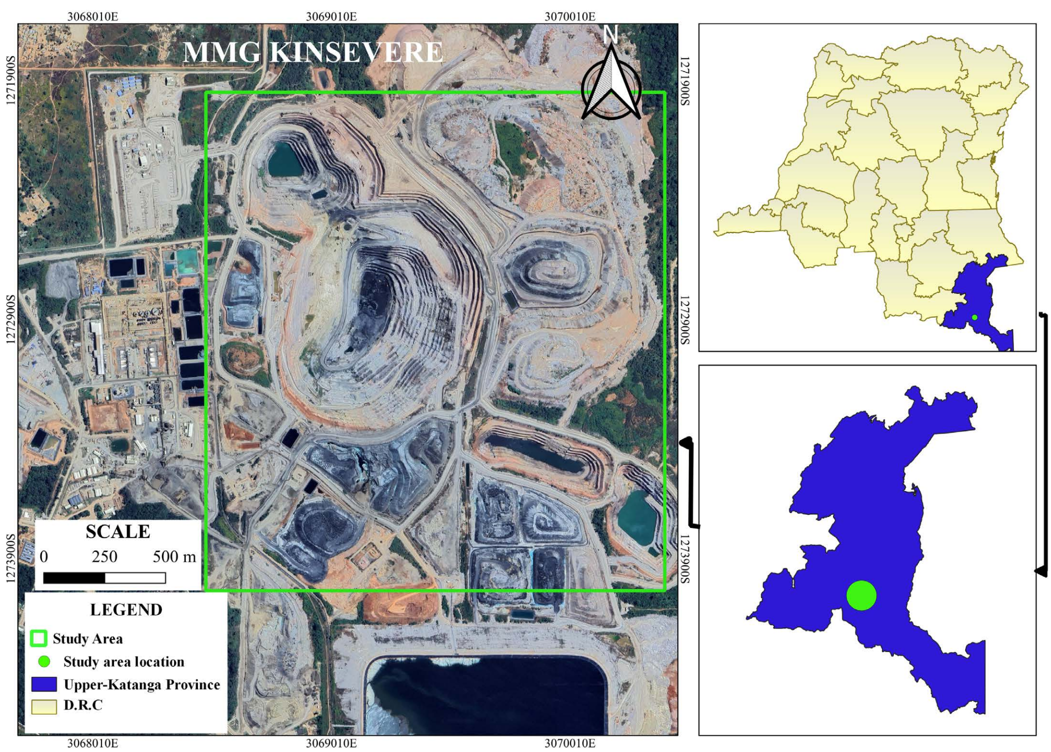

Figure 1. Study area location map.

2.2. Data Description, Processing and Statistical Analysis

The data used in this study were taken from the database collected at the Kinsevere

MMG mine on May 06, 2024. A total of 158 mining records were collected by the company. In the database, several fields were filled in, but only 4 columns caught our attention after data pre-processing. These were: maximum instantaneous load (Qmax), recording distance (Slope_dist), scaled distance (SD) and AOp.

Data processing was carried out with the aim of guaranteeing data quality, given that the performance of ML models is directly impacted by the quality of the data on which they are trained [49]. This process was carried out in several stages, from understanding the data to subdividing it, as shown in **Figure 2**.

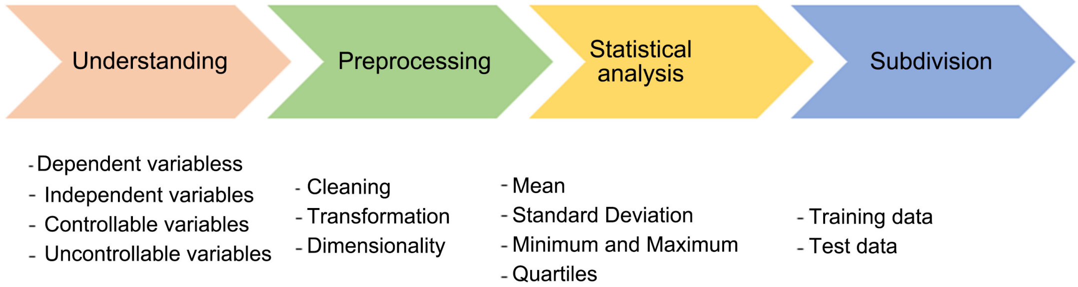

Figure 2. Illustrative diagram of data description, processing and statistical analysis.

Firstly, the understanding of the data consisted in identifying all the key factors most often involved in modelling mining operations specifically for the phenomenon of AOp. The literature consulted led to the identification of dependent and independent factors or parameters, as well as those that are controllable (linked to the design of the blast pattern and the explosive used) or uncontrollable (linked to the physical and geomechanical properties of the rock, joint properties, and rock classification). For AOp, the most popular independent variables are maximum load per delay, distance between the recording station and blast batch, scaled distance, drilling, and blast parameters.

Secondly, the pre-processing of the data to develop the various empirical and numerical models mainly involved cleaning, transforming, and reducing the dimensionality of the database. Cleaning the database involved eliminating duplicate records, managing missing fields, and identifying and rectifying or removing outliers in the data, as suggested by Chu *et al.* [50]. Data transformation involves standardizing the data. Dimensionality reduction involved the selection of the features used to identify the important ones that have an impact on the output of ML models. The non-important features identified were ignored, resulting in the reduction of input features and, consequently, in model complexity. A strategy based on the correlation coefficient was used in this work to apply this reduction in data dimensionality. Variables that correlate strongly or perfectly were discarded, as some numerical models are sensitive to strong correlations during training, which can significantly impact their performance.

Thirdly, the basic statistical analysis consisted of simple descriptive statistics of the pre-processed database. Basic statistical parameters such as mean, standard deviation, maximum, minimum, and quartiles (1st, 2nd, and 3rd) were evaluated for each variable in the database. This statistical analysis was carried out using the Python program V.3.12 (2024).

Finally, the data subdivision consisted of the creation of the training set and a test set, as frequently observed in several research studies [51]. This subdivision was carried out with the aim of assessing the generalizability of the models, *i.e.*; its ability to apply learned models to unseen data, thus mitigating over-fitting and under-fitting. The data subdivision strategy most used in several research studies [51], which is the subdivision into two sets with 80% of the data assigned to the training set and 20% to the test set, is the one employed in this study.

2.3. Data Description, Processing and Statistical Analysis

In this study, AOp was predicted using the prediction model proposed by United State Bureau of Mines model (USBM) as applied by Hajihassani *et al.* [23] and Armaghani *et al.* [30]. The characteristic equation of this model is:

$$\mathrm{AOp} = K\left[\frac{D}{\sqrt[3]{W}}\right]^{-B} = K\left(\mathrm{SD}\right)^{-B} \tag{1}$$

where D is the distance between the recording station and the blasting batch; W is the instantaneous explosive charge; K and B are site constants, and SD is the scaled distance.

K and B were obtained by linear regression analysis between AOp and SD transformed data by implementing the Python program V.3.12 (2024). The regression equation is:

$$\log\left(\mathrm{AOp}\right) = \log\left(K\right) - B\log\left(\mathrm{SD}\right) \tag{2}$$

The general form of the equation is:

$$y = ax + b \tag{3}$$

By identification $y = \log\left(\mathrm{AOp}\right)$, $x = \log\left(\mathrm{SD}\right)$, $a = -B$ et $b = \log\left(K\right)$. Thus, $K = 10^b$ et $B = -a$.

2.4. Multiple Linear Regression

Multiple linear regression (MLR) belongs to the category of supervised learning algorithms. This implies that models are trained on a set of labelled data (training data), and these models are used to predict unlabelled data (test data) [52]. The aim of MLR, as part of the family of regression algorithms, is to find relationships and dependencies between variables. It represents a model between a continuous scalar dependent variable "*y*" (label or target variable in ML terminology) and several explanatory variables "*x*" (independent variables, input variables, observed data, characteristics, observations, attributes, dimensions, etc.) [52] [53]. The characteristic equation of this type of regression is as follows [54]:

$$y = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \cdots + \beta_n \cdot x_n + e \tag{4}$$

where y represents the dependent variable, x_i denotes the independent variables, β_0 is the intercept term, β_i are weights associated with each independent variable, and e represents the error term.

2.5. Decision Tree

The decision tree (DT) algorithm is an ML algorithm that also belongs to the category of supervised learning. It is a kind of non-parametric model widely used for both classification and regression problems [53]. It follows a tree structure by making certain decisions and follows a top-down approach [55]. When building a DT model, a set of binary rules is used to predict the target value. Four key parameters are involved in the construction: the root node, the decision node, the branch, and the leaf node [55]. The root node represents the dataset or sample and then divides into different nodes. Each node of the tree represents an input entity of the data set. Each splitting node is called a decision node, and a sub-tree is called a branch. The leaf node is a node where further division ends and represents the target value [53] [55]. In the context of decision tree construction, this involves asking each decision node a series of questions based on an already chosen threshold value. Depending on the answers, it will be classified into different branches. This process occurs iteratively fashion, eventually reaching the leaf node which represents the target value [56].

2.6. Random Forests

This model uses an approach that integrates many basic learners (called decision trees) through integrated learning and uses bagging and bootstrapping techniques to train the decision trees, thus solving the problem of insufficient performance of individual decision trees when processing complex data [57]. The RF model was first introduced by Breiman [58] and became popular due to its robust non-parametric statistical approach to regression and classification problems, but on the contrary, it depends on the results obtained from the decision trees to guarantee the accuracy and efficiency of the predictions [59]. Thus, for each observation, it merges the predicted values of each decision tree in the forest to provide the optimal value [55].

In the decision tree construction process, the RF model generates more random trees as the number of decision trees in the forest increases. Instead of selecting the optimal value step by step like a decision tree, the RF model uses its random selection capability and the decision tree's voting mechanism to quickly find the optimal value. This property allows the RF model to have better classification or regression performance and a high generalization capacity for the learning system [57]. Particularly for the regression task, when predicting the unknown output value, each decision tree gives a predicted value, and the final value is the average of all decision trees [57]. Bui *et al.* [45] show that key RF features for regression, such as the number of trees (n_tree), maximum learning depth (max_depth), subsample size (or rate), and the number of features considered at each split (mtry), significantly affect the performance of the RF model and can be tuned for optimal model fitting.

2.7. Model Construction

A total of 58 mining records (training data) were used to calculate site factors K and B, with 15 records used to test and evaluate model performance. On this basis, the Python code V.3.12 (2024). implementation of the linear regression analysis technique was carried out to define K and B. For the development of numerical models based on ML algorithms, all models used the same type of data in training and testing as those used for the development of the USBM model. However, to develop the DT model, the following training hyper-parameters were used: max_depth = 6, random_state = 4, criterion = "squared_error", max_features = 3, min_samples_leaf = 1, min_samples_split = 2. For the RF model, n_estimators = 10, random_state = 4, max_depth = 7, criterion = "squared_error", max_features = 2. These parameter values were determined through trial and error, training the models until an acceptable level of performance was achieved. The MLR model did not require any hyper-parameter tuning for training.

2.8. Model Evaluations

Six statistical criteria, including coefficient of determination (R^2), root-mean-square error (RMSE), root relative squared error (RRSE), Relative Absolute Error (RAE), mean absolute error (MAE), and mean absolute percentage error (MAPE), were used to evaluate the performance of the models and rank them according to their respective evaluation results. These criteria have been used in similar studies [7] [19] [24] [29] [30] [45] [60]. Equations 5 to 10 illustrate the mathematical formulations of R^2, RMSE, RRSE, RAE, MAE, and MAPE, respectively [19] [60].

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_{act_i} - y_{pred_i}\right)^2}{\sum_{i=1}^{n}\left(y_{act_i} - \overline{y}_{act}\right)^2} \tag{5}$$

$$\mathrm{RMSE} = \frac{1}{n}\sqrt{\sum_{i=1}^{n}\left(y_{act_i} - y_{pred_i}\right)^2} \tag{6}$$

$$\mathrm{RRSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(y_{act_i} - y_{pred_i}\right)^2}{\sum_{i=1}^{n}\left(y_{act_i} - \overline{y}_{act}\right)^2}} \tag{7}$$

$$\mathrm{RAE} = \frac{\sum_{i=1}^{n}\left|y_{act_i} - y_{pred_i}\right|}{\sum_{i=1}^{n}\left|y_{act_i} - \overline{y}_{act}\right|} \tag{8}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_{act_i} - y_{pred_i}\right| \tag{9}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_{act_i} - y_{pred_i}}{y_{act_i}}\right| * 100 \tag{10}$$

3. Results and Discussion

The basic statistical parameters for the statistical analysis of the database are presented in Table 2. Following the development of the predictive models and their re-spective evaluation, the results are shown in Table 3 and Table 4 below.

The results in Table 2 show that the quantity of explosive charge per delay (Qmax) used at MMG Kinsevere has an average of 68.84 kg, a standard deviation of 35.40, and a minimum and maximum value of 6 kg and 120 kg, respectively. It also turns out that 25% of Qmax data are ≤50 kg, 50% ≤ 70 kg, and 75% ≤ 100 kg. AOp is recorded over a distance (Slope_Dist) ranging from 30 m to 697 m. The average, maximum, and minimum values recorded for AOp are 147.02 dB, 148.2 dB, and 130.6 dB, respectively. The quartile analysis shows that 25%, 50% and 75% of the AOp data have values of 148.1 dB, 148.2 dB, and 148.2 dB, respectively.

Table 2. Statistical description of data.

	Qmax (kg)	**Slope_Dist** (m)	**SD**	**AOp** (dB)
Q_25%	50	65	16.59	148.1
Q_50%	70	91	22.28	148.2
Q_75%	100	140	40.72	148.2
Count	73	73	73	73
Max	120	697	150.16	148.2
Mean	68.84	125.84	32.79	147.02
Min	6	30	6.28	130.6
Std	35.40	121.45	27.25	3.47

After determining the site constants ($K = 158.49$ ($10^{2.2}$) and $B = -0.02$) by linear regression analysis using the empirical prediction model proposed by USBM, the characteristic equation of the prediction model becomes:

$$\mathrm{AOp} = 158.49(\mathrm{SD})^{-0.02} \tag{11}$$

Table 3 shows the performance results of the models on the training data. This table shows that empirical technique is the one with the weakest results in the evaluation, with an RMSE of 3.01, a MAE of 1.895, and an R^2 of 0.21. On the other hand, the performance of the AI models improves on the same data. For the MLR model, prediction errors are reduced with an RMSE of 0.63, a MAE of 1.32, and an R^2 of 0.60. The RF and DT models further improve their performance results compared to the two previous models, respectively with an RMSE of 0.25 and 0.29, an MAE of 0.25 and 0.17, and an R^2 of 0.94 and 0.92.

Table 4 shows the performance results of the models on the test data. The results on the test set are consistent with those obtained during training. The empirical technique remains the worst performer, with an RMSE of 2.80, a MAE of 1.75, and an R^2 of 0.41. Among the AI-based models, the MLR model achieves an RMSE of 1.51, a MAE of 1.02, and an R^2 of 0.83. The RF and DT models significantly reduce prediction errors and improve the coefficient of determination. Specifically, the RF model achieves an RMSE of 0.72, a MAE of 0.37, and an R^2 of 0.96, while the DT model obtains an RMSE of 0.29, a MAE of 0.10, and an R^2 of 0.99.

Overall, the AI models outperformed the empirical technique in both training and testing. However, it remains difficult to clearly distinguish which model is the best among the predictive models developed. Thus, a simple ranking method was applied to assess the quality of the models developed using performance indices (RRSE, RAE, RMSE, MAE, MAPE, and R^2), as shown in Table 5.

Table 5 shows that the DT model proposed in this study is the best model, with a total rank of 45, followed by the RF model with a total rank of 39. The MLR model achieved only average performance, with a total rank of 24, and finally, the empirical technique obtained the worst performance, with a total rank of 12. Figures 3-5 illustrate the performance of the models by comparing measured and predicted values in the test dataset.

Table 3. Model performance evaluation results on training data.

On Train dataset														
Models	Network results						Ranking the predicted models						Total ranking score	Rank
	RRSE	RAE	RMSE	MAE	MAPE	R_Score	RRSE	RAE	RMSE	MAE	MAPE	R_Score		
USMB	8.460	69.363	3.008	1.895	93.405	0.211	1	1	1	1	1	1	6	4
MLR	0.630	0.684	2.135	1.316	0.911	0.602	2	2	2	2	2	2	12	3
RF	0.245	0.191	0.831	0.368	0.254	0.939	4	3	4	3	3	4	21	1
DT	0.288	0.130	0.975	0.251	0.174	0.917	3	4	3	4	4	3	21	2

Table 4. Model performance evaluation results on test data.

On Test Dataset														
Models	Network results						Ranking the predicted models						Total ranking score	Rank
	RRSE	RAE	RMSE	MAE	MAPE	R_Score	RRSE	RAE	RMSE	MAE	MAPE	R_Score		
USMB	4.447	93.405	2.796	1.747	25.909	0.4138	1	1	1	1	1	1	6	4
MLR	0.414	0.4832	1.512	1.024	0.7093	0.8286	2	2	2	2	2	2	12	3
RF	0.201	0.1738	0.735	0.368	0.2531	0.9594	3	3	3	3	3	3	18	2
DT	0.078	0.0459	0.286	0.097	0.0671	0.9938	4	4	4	4	4	4	24	1

Models based on decision trees, notably those used in this study (DT and RF), are renowned for their inherent interpretability, making them very useful in fields where understanding the decision-making process is crucial [61] [62]. Unlike other machine learning (ML) algorithms, which offer black-box models that are difficult to interpret, decision trees offer transparency by representing the decision-making process as a sequence of relatively simple and intuitive rules, enabling stakeholders such as the mining industry to easily understand and interpret how the model arrives at its predictions. However, these models have difficulty handling high-dimensional data with many features [63]-[65]. Although decision trees are recognized for their interpretability compared to other machine learning algorithms, they can nevertheless be difficult to understand and explain, especially when they become complex and large. Moreover, they are highly sensitive to outliers and can easily be influenced by noisy data, leading to false predictions [66].

Table 5. Classification of model ranks by performance.

Total network result														
Models	Training dataset						Testing dataset						Total ranking score	Rank
	RRSE	RAE	RMSE	MAE	MAPE	R_Score	RRSE	RAE	RMSE	MAE	MAPE	R_Score		
USMB	1	1	1	1	1	1	1	1	1	1	1	1	12	4
MLR	2	2	2	2	2	2	2	2	2	2	2	2	24	3
RF	4	3	4	3	3	4	3	3	3	3	3	3	39	2
DT	3	4	3	4	4	3	4	4	4	4	4	4	45	1

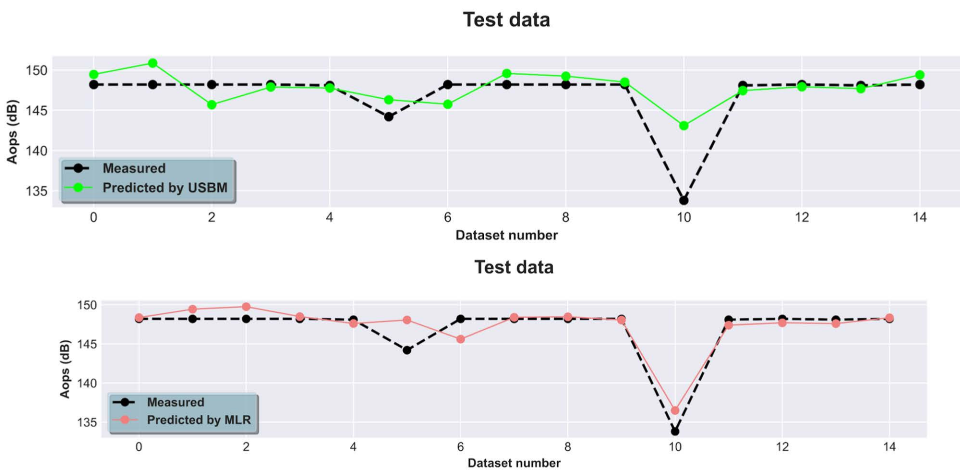

Figure 3. AOp measured and predicted by MLR and USBM models.

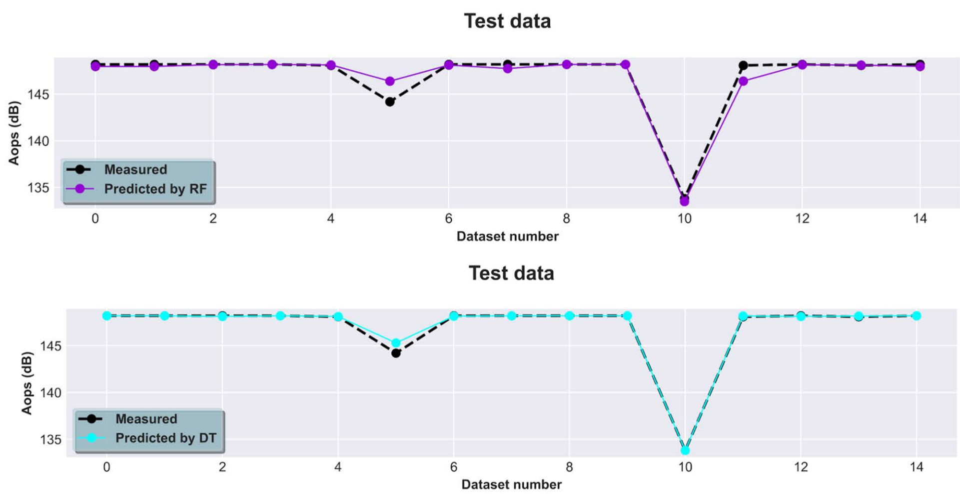

Figure 4. AOp measured and predicted by RF and DT models.

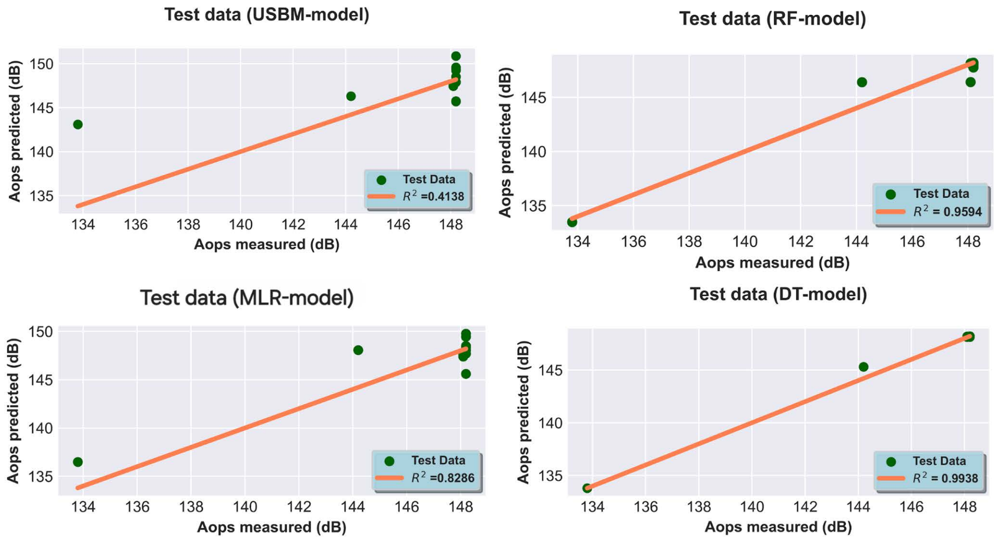

Figure 5. Coefficient of determination of AOp measured and predicted by different models.

4. Conclusions

Sustainable exploitation, transformation and use of natural resources are of vital importance to many countries around the world, particularly in the energy sector (transition to green energy). However, the sustainability of these activities boils down to minimizing the level of pollution in the surrounding environment, including, for example, blasting operations to break up rock in open-cast mines. Although much research has justified the effectiveness of drilling and blasting operations in fracturing rock, its secondary effects, such as AOp, remain significant; hence the need to control and accurately predict it to minimize its undesirable effects on neighbouring communities.

The primary objective of this study was to develop numerical predictive models for AOp. The methodology was structured into six key components: understanding the study area, particularly its geographical and geological characteristics; processing and statistical analysis of the collected dataset (73 records); development of both numerical and empirical prediction models, and evaluation of model performance of the numerical model parameters.

The results obtained in this study indicate that artificial intelligence (AI)-based models outperform the empirical method in predicting AOp. Among these AI models, the results show that the DT model is well-suited for predicting AOp in this study, with remarkable performance results (RRSE of 0.08, RAE of 0.05, RMSE of 0.29, MAE of 0.37, MAPE of 0.07, and an R^2 of 0.994). This could therefore justify its application in practical engineering to predict blast-induced AOp in open-cast mines to reduce undesirable environmental effects.

The dataset size of 73 records represents the maximum available data. We recommend that further research be carried out on a larger data set and that even more rigorous methods are used to fit the optimal hyper-parameters for the ML models.

Acknowledgements

We would like to express our immense gratitude to MMG Kinsevere, both for hosting the first author of the article at their mining facilities and for agreeing to carry out this scientific study using their blasting data.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Delille, F. (2012) Recherche d'une prédiction de fragmentation charge par charge pour les tirs à ciel ouvert [École nationale supérieure des mines de Panis].

[2] Ainalis, D., Kaufmann, O., Tshibangu, J.P., Verlinden, O. and Kouroussis, G. (2016) Modelling the Source of Blasting for the Numerical Simulation of Blast-Induced Ground Vibrations: A Review. *Rock Mechanics and Rock Engineering*, **50**, 171-193. https://doi.org/10.1007/s00603-016-1101-2

[3] Tan, L., Ren, T., Dou, L., Cai, X., Yang, X. and Zhou, Q. (2021) Dynamic Response and Fracture Evolution of Marble Specimens Containing Rectangular Cavities Subjected to Dynamic Loading. *Bulletin of Engineering Geology and the Environment*, **80**, 7701-7716. https://doi.org/10.1007/s10064-021-02425-w

[4] Wu, H., Ma, D., Spearing, A.J.S. and Zhao, G.Y. (2021) Fracture Phenomena and Mechanisms of Brittle Rock with Different Numbers of Openings under Uniaxial Loading. *Geomechanics and Engineering*, **25**, 481-493.

[5] Tripathy, G.R. and Gupta, I.D. (2002) Prediction of Ground Vibrations Due to Construction Blasts in Different Types of Rock. *Rock Mechanics and Rock Engineering*, **35**, 195-204. https://doi.org/10.1007/s00603-001-0022-9

[6] Valliappan, S. and Ang, K.K. (1998) Finite Element Analysis of Vibrations Induced by Propagating Waves Generated by Tunnel Blasting. *Rock Mechanics and Rock Engineering*, **21**, 53-78.

[7] Armaghani, D.J., Hajihassani, M., Sohaei, H., Mohamad, E.T., Marto, A., Motaghedi, H., *et al*. (2015) Neuro-Fuzzy Technique to Predict Air-Overpressure Induced by Blasting. *Arabian Journal of Geosciences*, **8**, 10937-10950. https://doi.org/10.1007/s12517-015-1984-3

[8] Khandelwal, M. and Singh, T.N. (2006) Prediction of Blast Induced Ground Vibrations and Frequency in Opencast Mine: A Neural Network Approach. *Journal of Sound and Vibration*, **289**, 711-725.

[9] Monjezi, M., Bahrami, A., Varjani, A.Y. and Sayadi, A.R. (2011) Prediction and Controlling of Flyrock in Blasting Operation Using Artificial Neural Network. *Arabian Journal of Geosciences*, **4**, 421-425. https://doi.org/10.1007/s12517-009-0091-8

[10] Tonnizam, E., Hajihassani, M., Armaghani, D.J. and Marto, A. (2012) Simulation of Blasting-Induced Air Overpressure by Means of Artificial Neural Networks. *International Review on Modelling and Simulations*, **5**, 2501-2506.

[11] Armaghani, D.J., Faradonbeh, R.S., Rezaei, H., Rashid, A.S.A. and Amnieh, H.B. (2018) Settlement Prediction of the Rock-Socketed Piles through a New Technique Based on Gene Expression Programming. *Neural Computing and Applications*, **29**, 1115-1125. https://doi.org/10.1007/s00521-016-2618-8

[12] Ebrahimi, E., Monjezi, M., Khalesi, M.R. and Armaghani, D.J. (2015) Prediction and Optimization of Back-Break and Rock Fragmentation Using an Artificial Neural Network and a Bee Colony Algorithm. *Bulletin of Engineering Geology and the Environment*, **75**, 27-36. https://doi.org/10.1007/s10064-015-0720-2

[13] Ghoraba, S., Monjezi, M., Talebi, N., Moghadam, M.R. and Jahed Armaghani, D. (2015) Prediction of Ground Vibration Caused by Blasting Operations through a Neural Network Approach: A Case Study of Gol-E-Gohar Iron Mine, Iran. *Journal of Zhejiang University-SCIENCE A*, **10**, 1631. https://scholar.google.com/citations?view_op=view_citation&hl=en&user=BDc_Ti-oAAAAJ&citation_for_view=BDc_TioAAAAJ:aqlVkmm33-oC

[14] Farhad, F., Farsangi, E., Ali, M. and Hamid, M. (2014) Simultaneous Investigation of Blast Induced Ground Vibration and Airblast Effects on Safety Level of Structures and Human in Surface Blasting. *International Journal of Mining Science and Technology*, **24**, 663-669. https://doi.org/10.1016/j.ijmst.2014.07.006

[15] Hasanipanah, M., Jahed, D., Hossein, A., Hassan, B.A. and Saber, G. (2015) Several Non-Linear Models in Estimating Air-Overpressure Resulting from Mine Blasting. *Engineering with Computers*, **32**, 441-455. https://doi.org/10.1007/s00366-015-0425-y

[16] Tiile, R.N. (2016) Artificial Neural Network Approach to Predict Blast-Induced Ground Vibration, Airblast and Rock Fragmentation. Missouri University of Science and Technology. https://scholarsmine.mst.edu/masters_theses/7571

[17] Kuzu, C., Fisne, A. and Ercelebi, S.G. (2009) Operational and Geological Parameters in the Assessing Blast Induced Airblast-Overpressure in Quarries. *Applied Acoustics*, **70**, 404-411. https://doi.org/10.1016/j.apacoust.2008.06.004

[18] Rodríguez, R., Lombardía, C. and Torno, S. (2010) Prediction of the Air Wave Due to Blasting Inside Tunnels: Approximation to a 'Phonometric Curve'. *Tunnelling and Underground Space Technology*, **25**, 483-489. https://doi.org/10.1016/j.tust.2010.01.011

[19] Khandelwal, M. and Kankar, P.K. (2011) Prediction of Blast-Induced Air Overpressure Using Support Vector Machine. *Arabian Journal of Geosciences*, **4**, 427-433. https://doi.org/10.1007/s12517-009-0092-7

[20] Konya, C.J. and Walter, E.J. (1990) Surface Blast Design. Prentice-Hall.

[21] Segarra, P., Domingo, J.F., López, L.M., Sanchidrián, J.A. and Ortega, M.F. (2010) Prediction of Near Field Overpressure from Quarry Blasting. *Applied Acoustics*, **71**, 1169-1176. https://doi.org/10.1016/j.apacoust.2010.07.008.

[22] Siskind, D.E., Stachura, V.J., Stagg, M.S. and Koop, J.W. (1980) Structure Response and Damage Produced by Airblast from Surface Mining. United States Bureau of Mines Report of Investigation 8485.

[23] Hajihassani, M., Armaghani, D.J., Sohaei, H., Mohamad, E.T. and Marto, A. (2014) Prediction of Airblast-Overpressure Induced by Blasting Using a Hybrid Artificial Neural Network and Particle Swarm Optimization. *Applied Acoustics*, **80**, 57-67. https://doi.org/10.1016/j.apacoust.2014.01.005

[24] Nguyen, H. and Bui, X.N. (2018) Predicting Blast-Induced Air Overpressure: A Robust Artificial Intelligence System Based on Artificial Neural Networks and Random Forest. *Natural Resources Research*, **28**, 893-907. https://doi.org/10.1007/s11053-018-9424-1

[25] McKenzie, C. (1990) Quarry Blast Monitoring Technical and Environmental Perspective. *Quarry Management*, **17**, 23-29.

[26] Persson, P.A., Holmberg, R. and Lee, J. (1994) Rock Blasting and Explosives Engineering. CRC Press.

[27] National Association of Australian State Road Authorities (1983) Explosives in Roadworks—A Users Guide. NAASRA.

[28] Khandelwal, M. and Singh, T.N. (2005) Prediction of Blast Induced Air Overpressure in Opencast Mine. *Noise & Vibration Worldwide*, **36**, 7-16. https://doi.org/10.1260/0957456053499095

[29] Hasanipanah, M., Shahnazar, A., Amnieh, H.B. and Armaghani, D.J. (2017) Prediction of Air-Overpressure Caused by Mine Blasting Using a New Hybrid PSO–SVR Model. *Engineering with Computers*, **33**, 23-31. https://doi.org/10.1007/s00366-016-0453-2

[30] Armaghani, D.J., Hasanipanah, M. and Mohamad, E.T. (2016) A Combination of the ICA-ANN Model to Predict Air-Overpressure Resulting from Blasting. *Engineering with Computers*, **32**, 155-171. https://doi.org/10.1007/s00366-015-0408-z

[31] Bui, X.N., Nguyen, Q.H., Tran, H., Nguyen, D.A. and Bui, H.B. (2021) Predicting Blast-induced Ground Vibration in Quarries Using Adaptive Fuzzy Inference Neural Network and Moth-Flame Optimization. *Natural Resources Research*, **30**, 4719-4734. https://link.springer.com/article/10.1007/s11053-021-09968-5

[32] Nguyen, H., Bui, X.-N., Bui, H.-B. and Mai, N.-L. (2018) A Comparative Study of Artificial Neural Networks in Predicting Blast-Induced Air-Blast Overpressure at Deo Nai Open-Pit Coal Mine, Vietnam. *Neural Computing and Applications*, **32**, 3939-3955. https://doi.org/10.1007/s00521-018-3717-5

[33] Monjezi, M., Baghestani, M., Faradonbeh, R.S., Saghand, M.P. and Armaghani, D.J. (2016) Modification and Prediction of Blast-Induced Ground Vibrations Based on Both Empirical and Computational Techniques. *Engineering with Computers*, **32**, 717-728.

[34] Chandar, K.R., Sastry, V. and Hegde, C. (2017) A Critical Comparison of Regression Models and Artificial Neural Networks to Predict Ground Vibrations. *Geotechnical and Geological Engineering*, **35**, 573-583. https://doi.org/10.1007/s10706-016-0126-3

[35] Geranian, H., Tabatabaei, S.H., Asadi, H.H. and Carranza, E.J.M. (2016) Application of Discriminant Analysis and Support Vector Machine in Mapping Gold Potential Areas for Further Drilling in the Sari-Gunay Gold Deposit, NW Iran. *Natural Resources Research*, **25**, 145-159. https://doi.org/10.1007/s11053-015-9271-2

[36] Mokhtari, M. and Behnia, M. (2018) Comparison of LLNF, ANN, and COA-ANN Techniques in Modelling the Uniaxial Compressive Strength and Static Young's Modulus of Limestone of the Dalan formation. Natural Resources Research.

[37] Tessema, A. (2017) Mineral Systems Analysis and Artificial Neural Network Modeling of Chromite Prospectivity in the Western Limb of the Bushveld Complex, South Africa. *Natural Resources Research*, **26**, 465-488. https://doi.org/10.1007/s11053-017-9344-5

[38] Zuo, R. and Xiong, Y. (2018) Big Data Analytics of Identifying Geochemical Anomalies Supported by Machine Learning Methods. *Natural Resources Research*, **27**, 5-13. https://doi.org/10.1007/s11053-017-9357-0

[39] Koopialipoor, M., Jahed Armaghani, D., Hedayat, A., Marto, A. and Gordan, B. (2018) Applying Various Hybrid Intelligent Systems to Evaluate and Predict Slope Stability under Static and Dynamic Conditions. *Soft Computing*, **23**, 5913-5929. https://doi.org/10.1007/s00500-018-3253-3

[40] Armaghani, D.J., Hasanipanah, M., Amnieh, H.B. and Mohamad, E.T. (2016) Feasibility of ICA in Approximating Ground Vibration Resulting from Mine Blasting.

Neural Computing and Applications, **29**, 457-465.
https://doi.org/10.1007/s00521-016-2577-0

[41] Torres, V.N., Silveira, L.G., Lopes, P.F. and de Lima, H.M. (2018) Assessing and Controlling of Bench Blasting-Induced Vibrations to Minimize Impacts to a Neighboring Community. *Journal of Cleaner Production*, **187**, 514-524.
https://doi.org/10.1016/j.jclepro.2018.03.210

[42] Mahdevari, S., Shahriar, K., Sharifzadeh, M. and Tannant, D.D. (2017) Stability Prediction of Gate Roadways in Longwall Mining Using Artificial Neural Networks. *Neural Computing and Applications*, **28**, 3537-3555.
https://doi.org/10.1007/s00521-016-2263-2

[43] Rahimdel, M.J., Mirzaei, M., Sattarvand, J., Ghodrati, B. and Mirzaei Nasirabad, H. (2017) Artificial Neural Network to Predict the Health Risk Caused by Whole Body Vibration of Mining Trucks. *Journal of Theoretical and Applied Vibration and Acoustics*, **3**, 1-14.

[44] Hasanipanah, M., Armaghani, D.J., Khamesi, H., Amnieh, H.B. and Ghoraba, S. (2016) Several Non-Linear Models in Estimating Air-Overpressure Resulting from Mine Blasting. *Engineering with Computers*, **32**, 441-455.
https://doi.org/10.1007/s00366-015-0425-y

[45] Bui, X.-N., Nguyen, H., Le, H.-A., *et al.* (2019) Prediction of Blast-Induced Air Over-Pressure in Open-Pit Mine: Assessment of Different Artificial Intelligence Techniques. *Natural Resources Research*, **29**, 571-591.
https://doi.org/10.1007/s11053-019-09461-0

[46] Mohamad, E.T., Armaghani, D.J., Hasanipanah, M., Murlidhar, B.R. and Alel, M.N.A. (2016) Estimation of Air-Overpressure Produced by Blasting Operation through a Neuro-Genetic Technique. *Environmental Earth Sciences*, **75**, Article No. 174. https://doi.org/10.1007/s12665-015-4983-5

[47] Nguyen, H., Bui, X.-N., Tran, Q.-H., Van, P., *et al.* (2019) A Comparative Study of Empirical and Ensemble Machine Learning Algorithms in Predicting Air Overpressure in Open-Pit Coal Mine. *Acta Geophysica*, **68**, 325-336.

[48] Golder Associates (2019) Updated Environmental and Social Impact Assessment and Environmental Management Plan for Kinsevere Mine. Report 1790467-320201-8.

[49] Halevy, A., Norvig, P. and Pereira, F. (2009) The Unreasonable Effectiveness of Data. *IEEE Intelligent Systems*, **24**, 8-12. https://doi.org/10.1109/mis.2009.36

[50] Chu, X., Ilyas, I.F., Krishnan, S. and Wang, J. (2016) Data Cleaning: Overview and Emerging Challenges. *Proceedings of the* 2016 *International Conference on Management of Data*, San Francisco, 26 June-1 July 2016, 2201-2206.

[51] Munagala, V., Thudumu, S., Logothetis, I., Bhandari, S., Vasa, R. and Mouzakis, K. (2024) A Comprehensive Survey on Machine Learning Applications for Drilling and Blasting in Surface Mining. *Machine Learning with Applications*, **15**, Article 100517.
https://doi.org/10.1016/j.mlwa.2023.100517

[52] Nasteski, V. (2017) An Overview of the Supervised Machine Learning Methods. *Horizons B*, **4**, 51-62.

[53] Géron, A. (2019) Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow: Concepts, Tools, and Techniques to Buil Intelligent Systems. O''Reilly Media, Inc.
https://www.oreilly.com/library/view/hands-on-machine-learning/9781492032632/

[54] Bisoyi, S.K. and Pal, B.K. (2020) Prediction of Ground Vibration Using Various Regression Analysis. *Journal of Mining Science*, **56**, 378-387.
https://doi.org/10.1134/s1062739120036665

[55] Ragam, P., Komalla, A.R. and Kanne, N. (2022) Estimation of Blast-Induced Peak Particle Velocity Using Ensemble Machine Learning Algorithms: A Case Study. *Noise & Vibration Worldwide*, **53**, 404-413. https://doi.org/10.1177/09574565221114662

[56] Amraee, T. and Ranjbar, S. (2013) Transient Instability Prediction Using Decision Tree Technique. *IEEE Transactions on Power Systems*, **28**, 3028-3037.

[57] He, M., Zhang, Z., Ren, J., Huan, J., Li, G., Chen, Y., *et al*. (2019) Deep Convolutional Neural Network for Fast Determination of the Rock Strength Parameters Using Drilling Data. *International Journal of Rock Mechanics and Mining Sciences*, **123**, Article 104084. https://doi.org/10.1016/j.ijrmms.2019.104084

[58] Breiman, L. (2001) Random Forests. *Machine Learning*, **45**, 5-32. https://doi.org/10.1023/a:1010933404324

[59] Vigneau, E., Courcoux, P., Symoneaux, R., Guérin, L. and Villière, A. (2018) Random Forests: A Machine Learning Methodology to Highlight the Volatile Organic Compounds Involved in Olfactory Perception. *Food Quality and Preference*, **68**, 135-145. https://doi.org/10.1016/j.foodqual.2018.02.008

[60] Nguyen, H., Bui, X., Asteris, P.G., Hieu-Tran, Q., Armaghani, D.J., Monjezi, M., Khandelwal, M. and Sukhanouvong, P. (2020) A Comparative Study of Different Artificial Intelligence Techniques in Predicting Blast-Induced Air Over-Pressure. *Metaheuristic Computing and Applications*, **1**, 187-202.

[61] Costa, V.G. and Pedreira, C.E. (2023) Recent Advances in Decision Trees: An Updated Survey. *Artificial Intelligence Review*, **56**, 4765-4800. https://doi.org/10.1007/s10462-022-10275-5

[62] Mahbooba, B., Timilsina, M., Sahal, R. and Serrano, M. (2021) Explainable Artificial Intelligence (XAI) to Enhance Trust Management in Intrusion Detection Systems Using Decision Tree Model. *Complexity*, **2021**, 1-11. https://doi.org/10.1155/2021/6634811

[63] Mhasawade, A., Rawal, G., Roje, P., Raut, R. and Devkar, A. (2023) Comparative Study of SVM, KNN and Decision Tree for Diabetic Retinopathy Detection. 2023 *International Conference on Computational Intelligence and Sustainable Engineering Solutions* (*CISES*), Greater Noida, 28-30 April 2023, 166-170. https://doi.org/10.1109/cises58720.2023.10183456

[64] Wang, T., Gault, R. and Greer, D. (2022) Cutting down High Dimensional Data with Fuzzy Weighted Forests (FWF). 2022 *IEEE International Conference on Fuzzy Systems* (*FUZZ-IEEE*), Padua, 18-23 July 2022, 1-8. https://doi.org/10.1109/fuzz-ieee55066.2022.9882660

[65] Azam, Z., Islam, M.M. and Huda, M.N. (2023) Comparative Analysis of Intrusion Detection Systems and Machine Learning-Based Model Analysis through Decision Tree. *IEEE Access*, **11**, 80348-80391. https://doi.org/10.1109/access.2023.3296444

[66] Xia, Y. (2019) A Novel Reject Inference Model Using Outlier Detection and Gradient Boosting Technique in Peer-to-Peer Lending. *IEEE Access*, **7**, 92893-92907. https://doi.org/10.1109/access.2019.2927602