

A comparative study of different artificial intelligence techniques in predicting blast-induced air over-pressure

Hoang Nguyen^{*1,2}, Xuan-Nam Bui^{1,2}, Panagiotis G. Asteris³, Quang Hieu-Tran¹,
Danial Jahed Armaghani⁴, Masoud Monjezi⁵, Manoj Khandelwal⁶
and Phonepaserth Sukhanouvong¹

¹Department of Surface Mining, Mining Faculty, Hanoi University of Mining and Geology,
18 Vien Street, Duc Thang Ward, Bac Tu Liem District, Hanoi, Vietnam

²Center for Mining, Electro-Mechanical research, Hanoi University of Mining and Geology,
18 Vien st., Duc Thang ward, Bac Tu Liem dist., Hanoi, Vietnam

³Computational Mechanics Laboratory, School of Pedagogical and Technological Education, 14121
Heraklion, Athens, Greece

⁴Department of Civil Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur,
Malaysia

⁵Department of Mining, Faculty of Engineering, Tarbiat Modares University, Tehran 14115-143, Iran

⁶School of Science, Engineering and Information Technology, Federation University Australia, Ballarat,
Australia

(Received April 1, 2020, Revised August 18, 2020, Accepted December 22, 2020)

Abstract. Blasting is known as the most common approach for fragmenting rock in open-pit mines. Nevertheless, its side effects are not insignificant, for example, fly rock, ground vibration, dust, toxic by-products, air over-pressure, and back-break. These effects considerably alter the circumambient environment, particularly when pressure is higher than usual. This study proposed and compared four artificial intelligence models for predicting blast-induced air over-pressure, namely multi-layer perceptron (MLP), Random Forest (RF), isotonic regression (IR), and M5-Rules. The air over-pressure was selected as the output variable based on the input variables, i.e., stemming length (T), explosive charge per delay (W), burden (B), monitoring distance (R), and spacing (S). Several statistical performance indices, including coefficient of determination (R^2), root relative squared error (RRSE), root-mean-square error (RMSE), average absolute error (MAE), and relative absolute error (RAE) were utilized to assess the models. Moreover, the ranking approaches of color intensity grouping, and the approach of general ranking were employed to further evaluate the models. The results, based on performance indices, confirm that M5-Rules is the outstanding model compared to the other techniques.

Keywords: blast-induced air over-pressure; artificial intelligence techniques; earth science; quarry mine; soft computing

1. Introduction

Mining activities (particularly blasting operations) often have dramatic effects on the

*Corresponding author, Professor, E-mail: nguyenhoang@humg.edu.vn

environment (Monjezi *et al.* 2009, Hresc *et al.* 2018, Juniah 2018, Bui *et al.* 2020, Nguyen *et al.* 2020). In previous studies, many researchers claimed that up to 80-85% of the energy explosive is wasted. Furthermore, this process is accompanied by particularly unfavorable environmental influences, such as flyrock, ground vibration, and air over-pressure, as well as back-break (Bui *et al.* 2019a, Nguyen *et al.* 2019a, Nguyen *et al.* 2020). Amongst these blasting side effects, air over-pressure is known as the most detrimental effect. It is defined as the suddenly change of the air pressure due to the effects of blasts (Nguyen and Bui 2020, Nguyen *et al.* 2020). Accordingly, explosives are activated, and chemical reactions occur that expands the air volume around the boreholes suddenly. (Do *et al.* 2020). The air volume is suddenly increased that creates the air over-pressure. It can lead to buildings'/houses' vibration, smash some doors, and harm humans (Khandelwal and Kankar 2011, Fang *et al.* 2019, Wei Gao *et al.* 2019). Therefore, predicting blast-induced air over-pressure is necessary to reduce its unfavorable influences on the environment.

To reduce the effects of blast-induced air over-pressure, some conventional approaches were applied, e.g., enhancing safety distances, as well as improving the quality of works (Army 1998). By utilizing barriers to decrease the influences of blast-induced air over-pressure, as well as ground vibration, different investigations and different methods have been adopted to calculate the damage to structures and to minimize the effects on buildings (Andersen and Nielsen 2005; Remennikov and Rose 2007, Asteris *et al.* 2017, Asteris *et al.* 2019). Nevertheless, these approaches are only qualitative and insufficient to control blast-induced air over-pressure. Researchers have tried to investigate and propose some empirical equations to estimate blast-induced air over-pressure (Kuzu *et al.* 2009). Nevertheless, only linear regression of the explosive charge per delay (W) and monitoring distance (R) was taken into account. Air over-pressure may be affected by different factors. Khandelwal and Kankar (2011) investigated blast geometry, W , and R , as well as vegetation, which affects air over-pressure. Siskind *et al.* (1980) resulted that W and R were the important influential parameters on air over-pressure. However, other parameters like atmospheric conditions, over-charging, etc. influenced air over-pressure (Segarra *et al.* 2010); thus, empirical formulas frequently lead to predicted results with low precision (Sawmliana *et al.* 2007, Khandelwal and Kankar 2011). There are disadvantages to the methods mentioned above; however, to overcome this issue, artificial intelligence (AI) has been employed in recent years by various scholars (Khaki *et al.* 2015, Amirmojahedi *et al.* 2016, Arabameri *et al.* 2017, Behmanesh and Mehdizadeh 2017, Mehdizadeh *et al.* 2017, Asteris and Mokos 2019, Cavaleri *et al.* 2019, Sarir *et al.* 2019, Duan *et al.* 2020). AI can decrease the undesirable influences produced from blasting operations (Iphar *et al.* 2008, Nguyen *et al.* 2018, Nguyen *et al.* 2019b, S. Zhang *et al.* 2019, X. Zhang *et al.* 2019). For predicting air over-pressure, Mohamed (2011) developed an artificial neural network (ANN), as well as fuzzy logic (FS). In this way, in Assay Cement Company, 162 blasting events were selected to investigate blast-induced air over-pressure. It was determined that the model of FS has a better performance compared to the ANN model. Khandelwal and Kankar (2011) also used an algorithm named support vector machine (SVM) to calculate blast-induced air over-pressure. They used three distinct limestone mines and about 75 blasting events. Hajihassani *et al.* (2014) expanded a novel ANN hybrid model, as well as an algorithm of particle swarm optimization (PSO) and obtained a proper result. By utilizing 62 blasting events, these algorithms were for estimating blast-induced air over-pressure in four quartiles of Malaysia. Hasanipanah *et al.* (2016) developed three non-linear models for estimating blast-induced air over-pressure. These models included ANN and adaptive neuro-fuzzy inference systems (ANFIS), as well as FS. In this study, 77 blasting events and a feasible technique (to compare with these models) were utilized in the Miduk Copper Mine (located in Iran).

Hasanipanah *et al.* (2017) constructed a hybrid model related to the algorithms of PSO, also supporting vector regression (SVR). Various equations such as linear (L) and quadratic (Q), as well as radial basis (RBF) kernels are used in the PSO algorithm. The PSO algorithm was selected in order to optimize the SVR model and MLR is additionally utilized to compare to the PSO-SVR models. Scholars have determined that the PSO-SVR-RBF is the best model compared to similar works.

Also, AminShokravi *et al.* (2018) employed three functions for PSO to predict air over-pressure (L, Q, and power (P) functions). An ANN model is selected and applied for comparison with the PSO algorithm. In their case study, they have shown that the used combination model (PSO-L) had the best performance to predict blast-induced air over-pressure. Alel *et al.* (2018) performed a similar study by taking into account a variety of ANNs based on PSO, as well as the algorithms of multi-swarm optimization (MSO). They determined that the MSO-Rand-Mut model had a proper performance in their case study. Nguyen *et al.* (2018) conducted a similar study. They compared three types of ANNs in the case of predicting blast-induced air over-pressure and showed that a promising result could be achieved with the ANN model. Mahdiyari *et al.* (2018) have expanded a Monte Carlo model (by using AI) to examine blast-induced air over-pressure in a case study. They used multiple linear regression (MLR) for making an equation to apply simulation of Monte Carlo and determined that MLR was proper for applying the Monte Carlo simulation in order for estimating the blast-induced air over-pressure. It was stated that the genetic algorithm (GA) is considered as one of the methods to optimize the ANN model in order to calculate blast-induced air over-pressure. The number of hidden neurons can be optimized using the GA algorithm to enhance the efficiency of the ANN model. Bui *et al.* (2019b) investigated seven AI methods to predict blast-induced air over-pressure, such as *k*-nearest neighbors (KNN), SVR, Bayesian additive regression trees (BART), gaussian process (GP) and boosted regression trees (BRT). They claimed that AI methods are the excellent methods in order to predict blast-induced air over-pressure. These techniques were also used in similar works (Nguyen and Bui 2018; Nguyen *et al.* 2019c; Nguyen *et al.* 2019d). The primary purpose of the present investigation was to make a reliable machine learning model to calculate blast-induced air over-pressure. Firstly, four different machine learning-based solutions were selected, including MLP regression (MLP), Isotonic Regression (IR) and RF as well as M5-Rules. Then, the obtained results of predictions are analyzed and discussed.

2. Research significance

Air over-pressure (AOp) is known as a dangerous effect induced by blasting operations in open-pit mines. The influential parameters have been interpreted as having a non-linear relationship with AOp. Indeed, characteristics of rock mass, as well as different geological, and geophysics conditions are the uncontrollable parameters during blasting. Also, meteorological elements (e.g., humidity, barometric pressure) were confirmed as having a significant effect on AOp. This non-linear nature of Air over-pressure, results in low reliability of deterministic models, such as empirical equations. Therefore, artificial intelligence, with meta-heuristic algorithms, based on a reliable database, seems to be the most suitable approach to reveal the non-linear behavior of AOp, in a robust and reliable manner.

It should be pointed out that a reliable database must comprise of, not only reliable data, but also of a sufficient amount of data, that covers the full range of parameter values, regarding the

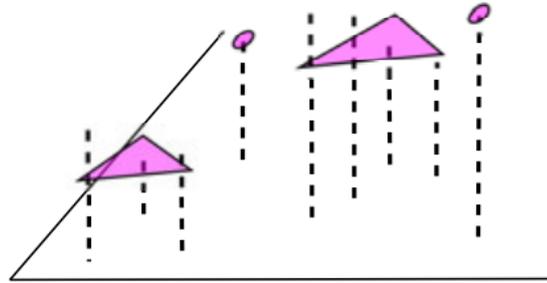


Fig. 1 Level collections to 2D IR

parameters which influence the air over-pressure phenomenon.

3. Machine learning-based solutions

3.1 Isotonic Regression

Isotonic regression (IR) is known as a considerable form of non-parametric regression. This technique can be used in different applications. It has also been used to optimize and classify (Kaufman and Tamir 1993; Stout 2015; Wei, Gao *et al.* 2019) existing problems. This method can be used in large data categories, such as learning and investigating data from as well as general data mining. In statistical algorithms, there are discrete mathematics, operations research, and computer science forums. They are utilizing methods as network flows and dynamic programming as well as parametric search.

In the present work, the case where V containing points within d -dimensional space, for $d \geq 2$, and point $p = (p_1, \dots, p_d)$ precedes points $q = (q_1, \dots, q_d)$ if $p_i \leq q_i$ for $1 \leq i \leq d$. q stands for dominant p , and the ordering is considered as domination, matrix, or dimensional ordering. It was determined that “multivariate IR” seemed proper. It has a distinct definition [35]. Any dimension requires just be a linearly ordered set. For instance, the independent variables allocated some white blood cell, age, and tumor severity systematization $I < II < III$, respectively. The dependent variable was chosen for five-year survival. In this regard, if any of the independent variables were maintained fixed, an enhanced in the third cannot reduce the survival probability, reversing the arrangement on age as well as on tumor classification. We have not considered assumptions about any possible changes if some enhance or increase. An example of IR on a category of points in an optional position, as shown in Fig. 1. As seen, some regions are undefined for it, and other is the constant regression. It indicated constant regions related to level sets. For any value of it, the L_p is average of the weighted amounts for its point. This was the weighted mean for L_2 , a weighted median for L_1 , and for L_∞ is the weighted mean of maximum violators, considered in Section five. Taken data values of f , a pair of values u , as well as v , is an offending pair when $u < v$ and $f(u) > f(v)$. IR with multi-dimension has long been investigated in various case studies [16]. However, there are difficulties in the calculation of IR as forcing them to utilize inferior replacements, as well as it can be limited to low-value data set (Dette and Scheder 2006; Luss *et al.* 2010; Saarela and Arjas 2011). There are various investigations related to the betterment of approximation. In

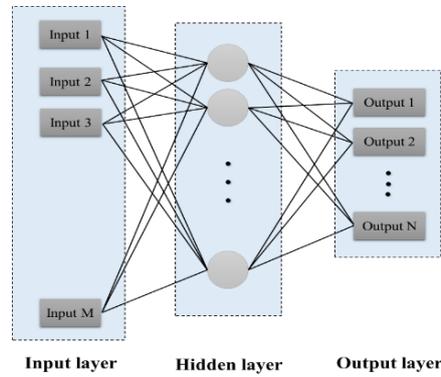


Fig. 2 MLP network architecture

this regard, scholars proposed the utilization of an additive isotonic model” [34]. The model of additive isotonic takes a sum of 1D IRs. It restricts the ability of isotonic functions representation (i.e., a sum of 1D regressions does not show the plain function, $z = x \cdot y$). In the present paper, a systematic approach is used to problems in the case of the IRs finding data at arbitrary positions in multi-dimensional places. P can be embedded in a *dag* $G = (P \cup R, E)$ that is present as $|R|, |E| = \mathcal{O}(|P|)$. Also, more discussion on the IR process has been presented in similar works (Stout 2015).

3.2 Multi-layer perceptron (MLP)

MLP is known as a widely used ANN. Multi-layer perceptron regression (MPR) is given the most appropriate possible regression when there is a collection of information samples S . It fulfilled using distributing the S for testing datasets. One introduced MLP as a composition of multiple computational units’ layers. (Nguyen *et al.* 2020). The schematic structure of a typical MLP is presented in Fig. 2:

Each node generates a local result. It also gives the nodes in the subsequent layer up to the output nodes that stayed in the layer in output, sent out the MLP answer (Çaylak and Kaftan 2014; Asteris *et al.* 2016). It assumed that the dataset included N groups of records. Equation 1 indicated the general operation performed using j^{th} neuron to calculate the output

$$O_j = F \left(\sum_{n=1}^N I_n W_{nj} + b_j \right) \quad (1)$$

I stands for the input value, b shows the bias of the node, and W stands for the weighting factor. Besides, F stands the function of activation. In this paper, the function of tangent-sigmoid activation was used that named Tansig. This function is expressed as the below relation

$$\text{Tansig}(x) = \frac{2}{1 + e^{-2x}} - 1 \quad (2)$$

3.3 Random Forest (RF)

Breiman (2001) proposed RF as an algorithm with the ability to incorporate calculations in the

case of multiple decision trees. Classification and regression algorithms were built as a planted forest as well as abundant trees. Each of the considered decision trees was utilized as a vote to predict the air over-pressure due to similar each tree in the forest (Khan *et al.* 2018). The approach of bootstrap resampling from the training dataset was utilized. In this method, the precision of the calculation trees can be synthesized. The RF algorithm robustly performs and has high precision, which can be obtained in a piercing data environment. Nevertheless, the RF was sometimes utilized in the case of blast-induced prediction. Air over-pressure was known as an open-pit mine (Longjun *et al.* 2011). In the following, we considered $T(x_i)$ as counting air over-pressure be the function demonstrating the training data sets. x_i stands the matrix of seven input parameters. Air over-pressure is considered as a predictive value. Using the model of RF indicates the decision trees number per forest to reduce the deviation among the estimated and experimental air over-pressure values. For improving the precision of the models, input parameters as well as decision tree number in each forest should be specified. To confirm predictable and objective obtained results, using I for many decision trees, the forest requires to be diversified. Thus, any number of decision trees and whole input parameters could be utilized for developing the RF estimation in the case of the air over-pressure model. We applied the RF method to predict blast-induced air over-pressure. Then we compared the obtained results to the other algorithms. In the RF algorithm, the number of trees sets to 2000 to consider the diversity of the forest (Nguyen and Bui 2018). The factor m_{try} is utilized to control the RF model quality. The tenfold cross-validation approach was utilized to three repeats to pass overfitting.

3.4 M5-Rules

In this regard, there are similar machine learning algorithms. The algorithm of tree-learning (TL) is a form of the M5-Rules model. For extraction of rules in trees model, a straight-forward mechanism can be employed that named M5-Rules. This model is used for different classification and calculation problems (Sharma *et al.* 2015; Bayzid *et al.* 2016; El-Bendary *et al.* 2016). M5-Rules as a straight-forward mechanism commonly utilize a tree learner over the training models for educating a pruned tree. In the tree model, the elite leaf can be constructed into a rule. Then, the tree is removed. This act sometimes is considered as the sole distinct among M5-Rules and regular methods, which can produce a single rule. In addition, all samples can be covered by the rule that is eliminated in the records. This method delayed when at least one rule considered covering total examples. Properly, obtaining the rules from the best leaf may cause for reducing the risk in the case over-pruning. The algorithm of partial decision trees (PART) produces determined trees, M5-Rules generates full trees. Producing partial trees causes the level of computational veracity to be increasing; however, it has no effects on the size as well as precision of the determined rules (Holmes *et al.* 1999).

4. Study area and data collection

As a general trend, it is noticed that, during the process of developing a forecast model, researchers pay particular attention to the computational model itself, while at the same time, not giving the same amount of attention to the database that is used for the development, training and validation of the model. Although research related to new computational models is of course of high importance and added value for the international scientific community, the authors believe that, since the ultimate goal is a reliable forecast, the reliability of the database should be of utmost

Table 1 The input and output parameters used in the development of AI models (All Datasets)

| Nr. | Variable | Symbol | Units | Category | Statistics | | | |
|-----|------------------------------------|--------|-------------------|----------|------------|---------|-------|--------|
| | | | | | Min | Average | Max | STD |
| 1 | Powder factor | P | Kg/m ³ | Input | 0.28 | 0.51 | 0.79 | 0.11 |
| 2 | Maximum explosive charge per delay | W | Kg | Input | 36.5 | 86.03 | 126.7 | 22.05 |
| 3 | Stemming | T | m | Input | 1.49 | 2.11 | 2.83 | 0.30 |
| 4 | Burden | B | m | Input | 1.12 | 2.30 | 3.45 | 0.44 |
| 5 | Spacing | S | m | Input | 2.32 | 3.31 | 4.17 | 0.38 |
| 6 | Monitoring distance | R | m | Input | 300 | 508.7 | 804.6 | 117.19 |
| 7 | Air over-pressure | AOp | dB | Output | 83.2 | 103.8 | 118.8 | 7.95 |

importance and should be thoroughly investigated in this regard. In fact, a reliable database must comprise of not only reliable data, but also of a sufficient amount of data, that covers the full range of parameter values, regarding the parameters which influence the problem investigated. Aiming to avoid misinterpretation of the term “sufficient”, it is highlighted that a sufficient amount of data is not necessarily a high amount of data, but rather datasets that cover a wide range of combinations of input parameter values, thus assisting in the model’s ability to simulate the problem.

Thus, a reliable database that also covers the full range of parameter values and combinations of parameter values, in addition to contributing to the development of reliable mathematical models, will also decisively contribute to the reliable comparison between different forecast computational models.

In the present paper, a quarry mine is selected as a case study (in Vietnam). The investigation area located among latitudes of 20°25’40”N - 20°26’20”N as well as longitudes of 105°53’10”E - 105°54’00”E. The primary ground of this mine is determined hilly and covered using dolomite and marl as well as limestone. In this regard, their stiffness ranges were determined from 10 to 14 conforming to Protodiakonov's hardness strength. In addition, blasting was determined as an efficient approach to fragment rocks.

In this mine, boreholes with 105 mm diameter and the ANFO was selected as the primary explosive. The capacity of explosive for any in this paper considered about 1,500 kg with the powder parameter around 0.28 to 0.79 kg/m³. There are various households with a distance of around 300 - 400 m (around the study area), so the influences of AOp is inevitable. It is essential to predict blast-induced AOp to ensure the surrounding residential and also there is a need to monitor it. In order to use the results of this study, a database with 164 blastings are considered, and a micromate facility was applied to collect AOp from 164 blasting events. To meet the monitoring distance (R), a different device, i.e., hand-held GPS receiver, was used. The remaining factors were used based on the patterns of the mine.

Based on the above database, each input training vector p (input parameters) is of dimension 1×6 and consists of the values of the Powder factor (P), the Maximum explosive charge per delay (K), the Stemming (T), the Burden (B), the Spacing (S) and the value of Monitoring distance (R). The corresponding output training vector (output parameter) is of dimension 1×1 and consist of the value of the blast-induced Air over-pressure (AOp) of open-pit mines. Their mean values together with the minimum, maximum values as well standard deviation (STD) values are listed in Table 1. Basically, some of the Blast-Induced Air Over-Pressure variables could be dependent on

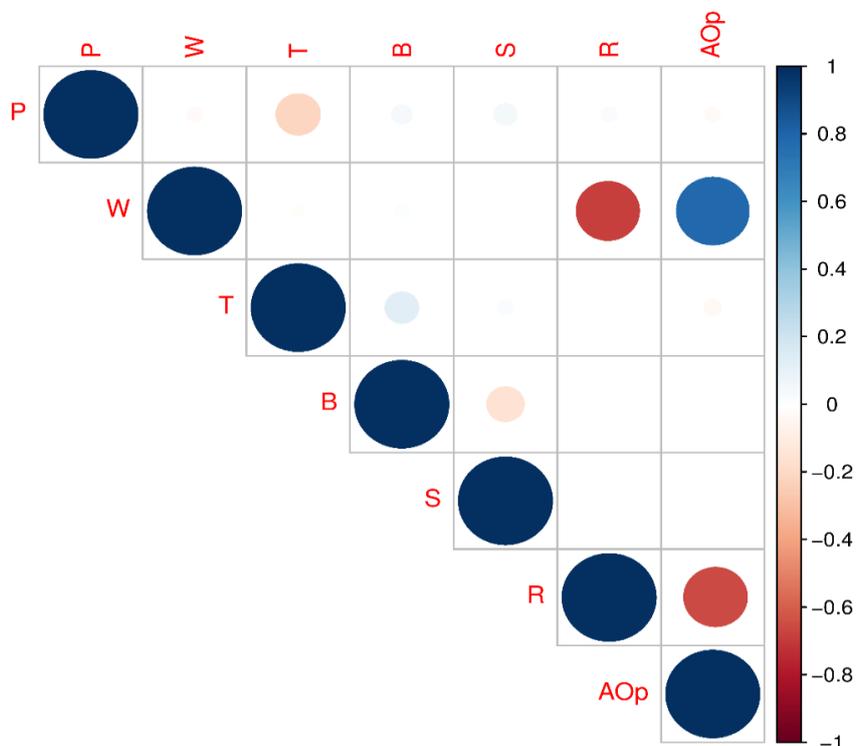


Fig. 3 Correlogram of the input and output variables

each other. High negative or positive values of the correlation coefficient between the input variables may result to poor efficiency of the methods and to the difficulty in construing the effects of the expository variables on the respond. Subsequently, the correlation coefficients between all possible variables have been specified and presented in Fig. 3. As can be seen in the table, there are not significant correlations between the independent input variables. On the other hand, in order to develop a reliable, robust and optimum artificial intelligence model the correlation coefficients between the input variables (parameters) and the output parameter of the blast-induced Air over-pressure (AOp), last line in Table 1, are needed to be as great as possible. Based on these values it is clearly shown that there is a strong relation between the output parameter of the blast-induced Air over-pressure (AOp) and the input parameters of the Maximum explosive charge per delay (W) confirmed by the correlation factor 0.784.

5. Performance indices

In this work, a procedure was performed to divide the data into two stages before expanding the models of air over-pressure predictive, and the data are split randomly. Based on similar works, the ratio of 80:20 (train/test) is the most usually utilized. This ratio was achieved from the Pareto principle suggested by Nick 2008. The upper mentioned ratio needs a proper starting point based on Swingler's 1996 suggestion, so in the first phase, 80% of the total data, 131 observations, was

approximately selected as the training dataset. For as testing database, in the second stage, the remaining data 20%, 33 observations were selected. The present work has focused on appraising the qualification of four prevalent machine learning tools named IR, MLP, RF, M5-Rules to predict the air over-pressure. The WEKA software was used for the training process that is a proper framework for data-mining as well as classification. This software has been formerly used in the previous studies for different simulating aims between 49-51. Five statistical were used (R^2 , MAE, RMSE, RAE, and RRSE) to expand a color intensity ranking for presenting a colored comparison of the obtained data. These criteria have been utilized in similar studies 37, 52-55. Eq. (3) to Eq. (7) respectively illustrates the RRSE, R^2 , MAE, RAE, and RMSE formulation.

$$R^2 = 1 - \frac{\sum_{i=1}^s (Y_{i_{\text{predicted}}} - Y_{i_{\text{observed}}})^2}{\sum_{i=1}^s (Y_{i_{\text{observed}}} - Y_{i_{\text{predicted}}})^2} \quad (3)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^s |Y_{i_{\text{observed}}} - Y_{i_{\text{predicted}}}| \quad (4)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^s \left[(Y_{i_{\text{observed}}} - Y_{i_{\text{predicted}}})^2 \right]} \quad (5)$$

$$\text{RAE} = \frac{\sum_{i=1}^s |Y_{i_{\text{predicted}}} - Y_{i_{\text{observed}}}|}{\sum_{i=1}^s |Y_{i_{\text{observed}}} - \bar{Y}_{i_{\text{observed}}}|} \quad (6)$$

$$\text{RRSE} = \sqrt{\frac{\sum_{i=1}^s (Y_{i_{\text{predicted}}} - Y_{i_{\text{observed}}})^2}{\sum_{i=1}^s (Y_{i_{\text{observed}}} - \bar{Y}_{i_{\text{observed}}})^2}} \quad (7)$$

In all the considered above equations, the actual and predicted data of response, $Y_{i_{\text{observed}}}$, and $Y_{i_{\text{predicted}}}$, were determined, respectively. The parameter of S shows the number of data, and $\bar{Y}_{i_{\text{observed}}}$ stands for the actual data mean of air over-pressure. In the following section, the accuracy of considered models to approximate the response is presented and further discussed.

6. Results and discussions

In this study, four machine learning models were taken into account to predict air over-pressure, as mentioned above. For the development of these models, Weka 3 was applied with the available packages. It is worth noting that the architecture of MLP was selected by a trial and error

Table 3 The obtained data from using proposed networks with several statistical indexes (only provided to train dataset)

| Proposed models | Network results | | | | | Ranking the predicted models | | | | | Total ranking score | Rank |
|-----------------|-----------------|-------|-------|---------|----------|------------------------------|-----|------|---------|----------|---------------------|------|
| | R ² | MAE | RMSE | RAE (%) | RRSE (%) | R ² | MAE | RMSE | RAE (%) | RRSE (%) | | |
| 1 IR | 0.991 | 0.473 | 1.083 | 7.111 | 13.657 | 3 | 3 | 3 | 3 | 3 | 15 | 2 |
| 2 MLP | 0.987 | 0.730 | 1.296 | 10.977 | 16.343 | 2 | 1 | 2 | 1 | 2 | 8 | 3 |
| 3 RF | 0.997 | 0.345 | 0.612 | 5.182 | 7.714 | 4 | 4 | 4 | 4 | 4 | 20 | 1 |
| 4 M5-Rules | 0.984 | 0.726 | 1.399 | 10.910 | 17.646 | 1 | 2 | 1 | 2 | 1 | 7 | 4 |

Table 4 The obtained data from proposed networks with several statistical indexes (only provided to test dataset)

| Proposed models | Network results | | | | | Ranking the predicted models | | | | | Total ranking score | Rank |
|-----------------|-----------------|-------|-------|---------|----------|------------------------------|-----|------|---------|----------|---------------------|------|
| | R ² | MAE | RMSE | RAE (%) | RRSE (%) | R ² | MAE | RMSE | RAE (%) | RRSE (%) | | |
| 1 IR | 0.987 | 0.717 | 1.178 | 12.640 | 16.409 | 2 | 2 | 2 | 2 | 2 | 10 | 3 |
| 2 MLP | 0.990 | 0.717 | 1.067 | 12.625 | 14.86 | 3 | 3 | 3 | 3 | 3 | 15 | 2 |
| 3 RF | 0.978 | 0.892 | 1.445 | 15.722 | 20.124 | 1 | 1 | 1 | 1 | 1 | 5 | 4 |
| 4 M5-Rules | 0.992 | 0.612 | 0.987 | 10.789 | 13.737 | 4 | 4 | 4 | 4 | 4 | 20 | 1 |

Table 5 The obtained data from the total ranking of proposed models to estimate the response

| Proposed models | Total network result | | | | | | | | | | Total score | Rank |
|-----------------|----------------------|-----|------|---------|----------|-----------------|-----|------|-----|------|-------------|------|
| | Training dataset | | | | | Testing dataset | | | | | | |
| | R ² | MAE | RMSE | RAE (%) | RRSE (%) | R ² | MAE | RMSE | RAE | RRSE | | |
| 1 IR | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 2 | 2 | 2 | 25 | 2 |
| 2 MLP | 2 | 1 | 2 | 1 | 2 | 3 | 3 | 3 | 3 | 3 | 23 | 4 |
| 3 RF | 4 | 4 | 4 | 4 | 4 | 1 | 1 | 1 | 1 | 1 | 25 | 2 |
| 4 M5-Rules | 1 | 2 | 1 | 2 | 1 | 4 | 4 | 4 | 4 | 4 | 27 | 1 |

procedure to find out the best MLP model. Finally, the architecture of 6-12-1 was selected for the MLP model in predicting air over-pressure. For the RF modeling, the number of the trees were selected as 500 to ensure the enrich of the forest during developing RF model. Whereas, IR and M5-Rules models were developed based on the available packages of Weka.

Tables 3 and 4 show the predicted values of RRSE, RAE, R², RMSE, and MAE to estimate the reply in for training datasets, and Table 5 shows the total ranking calculated for considered models. A graphical presentation of the obtained data is too presented, based on the model of color intensity. A red collection in tables is marked for these. By using more intense red color, a higher amount of R² and less RRSE, RAE, RMSE, and MAE were addressed. As seen in tables 3-5, the ultimate ranking has defined to consider the total score taken for the models of IR, MLP, RF, M5-Rules. For each model, the proposed total score shows the aggregate of the partial scores,

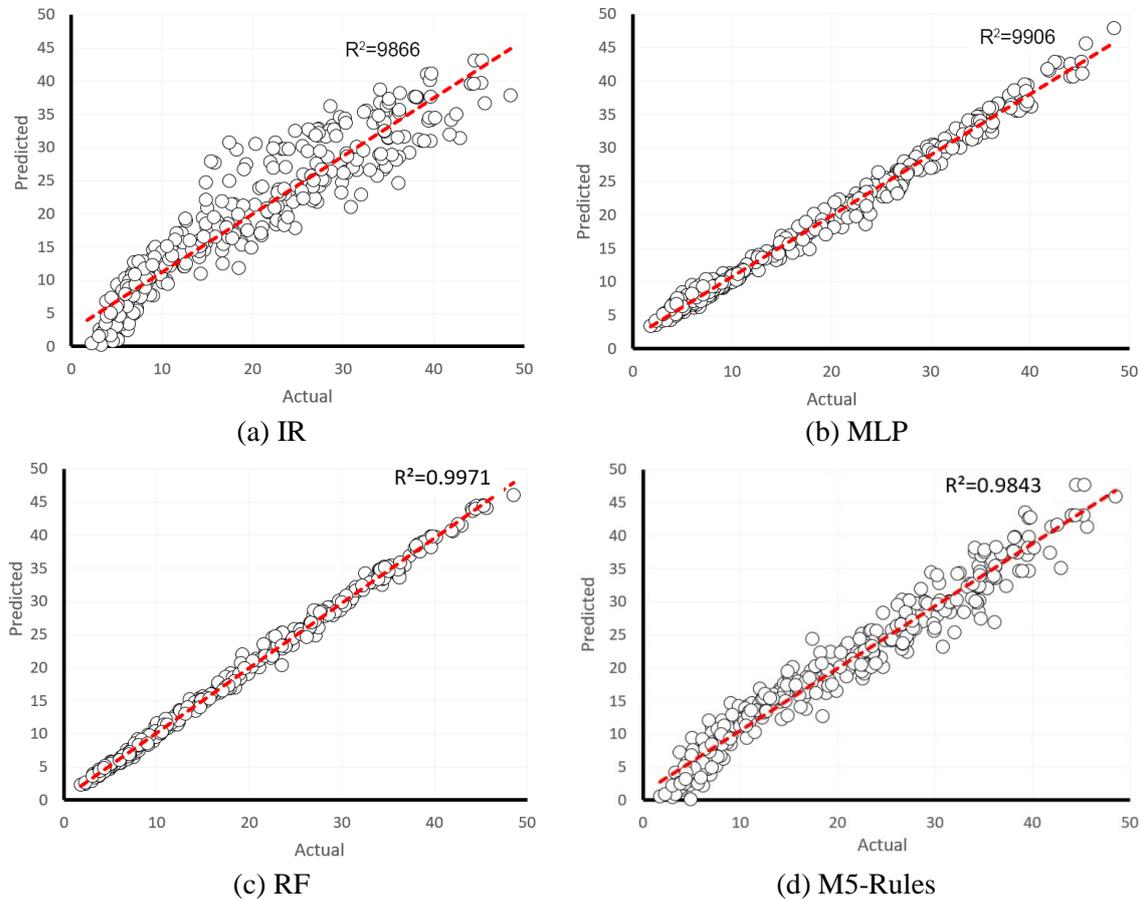


Fig. 4 The network results to train dataset; (a) IR, (b) MLP, (c) RF, and (d) M5-Rules

according to RRSE, RAE, R^2 , RMSE, and MAE. The data of statistical indicators in the case of RRSE, RAE, R^2 , RMSE, and MAE to train results in IR, MLP, RF, M5-Rules were selected as denoted in Table 3. After reviewing training and testing datasets, the RF method can be presented as a model of outstanding, because of its the highest perfection and partial scores predicted for the RF method. This predictive model is the most efficient training in comparison to similar models. In this paper, IR and MLP introduced the second and third- precise models, respectively, and we take the lowest training rate for a model of the squares regression median (M5-Rules).

Table 5 shows a comparison of considered approaches. As seen, the assumption of individual ranks obtained for each pattern is provided as a total ranking, and the RF (total score=20) obtained the excellent accuracy between four models considered in the present work. IR and MLP are illustrated as a proper performance (total score=15 and 8). The considerable point in the whole three tables is the similar partial score predicted to all statistical index using each considered model. It is found that all considered models had an equal accuracy to train and test the datasets of IR, MLP, RF, M5-Rules. Therefore, we concluded that equal accuracy is due to the existing differences among the amounts of RRSE, RAE, R^2 , RMSE, and MAE predicted for all considered phase. Figs. 4 and 5 show a comparison of the real and modeled values of responses in the case of

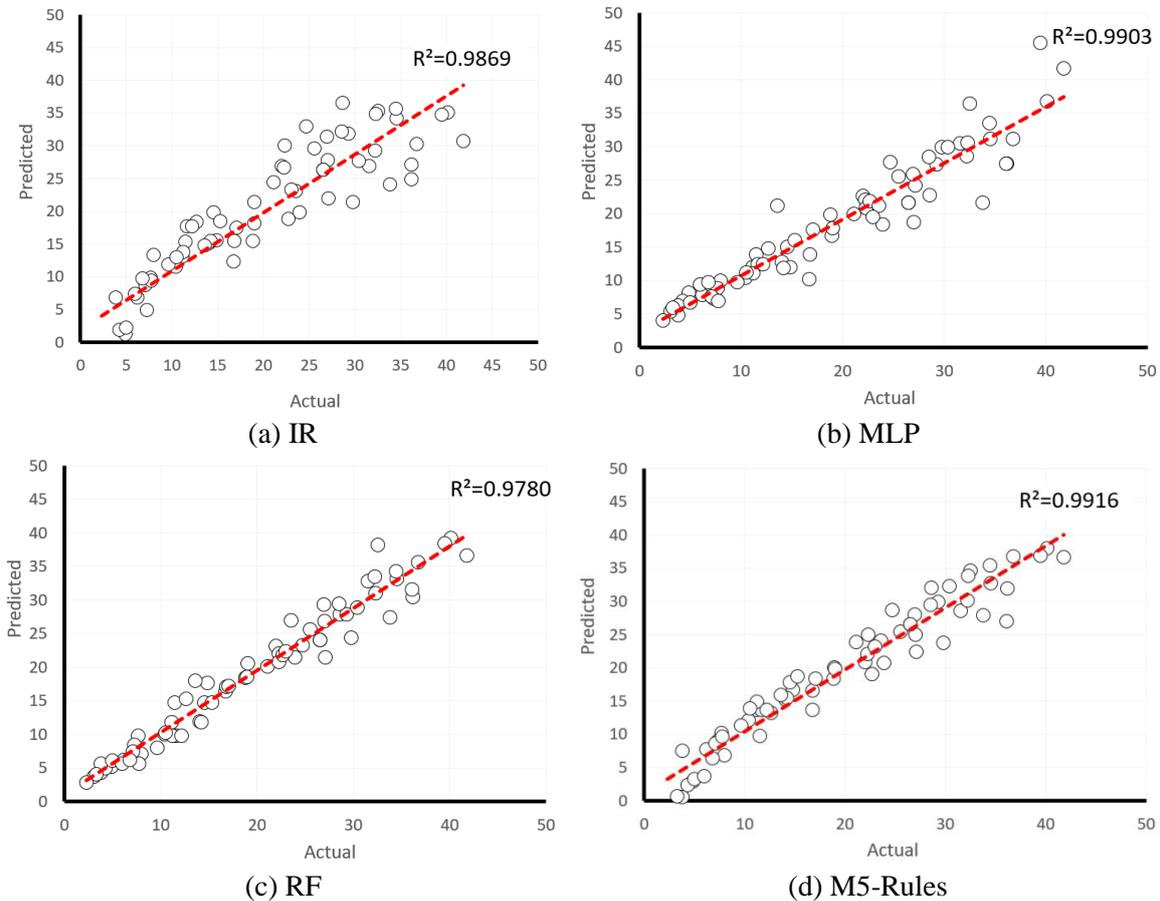


Fig. 5 The network results to test the models; (a) IR, (b) MLP, (c) RF, and (d) M5-Rules

the training and testing of IR, MLP, RF, M5-Rules datasets. As seen, the proper prediction among the results determined on the horizontal, actual data, and the vertical axis, predicted data, is shown by the line $x=y$. Figures 4 and 5 showed the RF-based solution is a proper accord compared to exact data training ($R^2=0.9971$) and ($R^2=0.978$) testing phases. The obtained results proved the advantage of proposed RF efficiency in the case of the testing “ R^2 (0.9869, 0.9903, 0.978 and 0.9916), RMSE (1.1784, 1.0672, 1.4452 and 0.9865) and RRSE (16.4086, 14.861, 20.1242 and 13.7374 %) and training R^2 (0.9906, 0.9866, 0.9971 and 0.9843), MAE (0.473, 0.7301, 0.3447, and 0.7257) and RAE (7.01109, 10.9766, 5.1821 and 10.9103 %)” , respectively, models of predicted for IR, MLP, RF, M5-Rules.

7. Conclusions

Blast-induced air-overpressure is one of the negative effects induced by blasting operations in open-pit mines. Hence, predicting and controlling it is a challenging of engineers and researchers. In this study, some following conclusions were given:

- Various artificial intelligence models, such as IR, MLP, RF, and M5-Rules have been developed in order to obtain a proper prediction performance in comparison to available predictive models.
- Based on performance indices, the M5-Rules model proved to be the most robust and reliable among the developed and trained intelligence models.
- Furthermore, the proposed optimum M5 model proved to be a useful tool for researchers, engineers, and for supporting the teaching and interpretation of the air-blast overpressure (AO_p) of blasting operations in open-pit mines.

Acknowledgments

The authors would like to thank Hanoi University of Mining and Geology (HUMG), Hanoi, Vietnam and the Center for Mining, Electro-Mechanical research of HUMG.

References

- Alel, M.N.A., Upom, M.R.A., Abdullah, R.A. and Abidin, M.H.Z. (2018), *Optimizing Blasting's Air Overpressure Prediction Model using Swarm Intelligence*. Journal of Physics: Conference Series, IOP Publishing.
- AminShokravi, A., Eskandar, H., Derakhsh, A.M., Rad, H.N. and Ghanadi, A. (2018), "The potential application of particle swarm optimization algorithm for forecasting the air-overpressure induced by mine blasting", *Eng. Comput.*, **34**(2), 277-285. <https://doi.org/10.1007/s00366-017-0539-5>.
- Amirmojahedi, M., Mohammadi, K., Shamshirband, S., Seyed Danesh, A., Mostafaiepour, A. and Kamsin, A. (2016), "A hybrid computational intelligence method for predicting dew point temperature", *Environ. Earth Sci.*, **75**(5), 415. <https://doi.org/10.1007/s12665-015-5135-7>.
- An Dinh, N., Bach Van N., Bao Dinh, T., Hoa Van P. and Tho Anh, N. (2020), "Definition of amount explosive per blast for spillway at the Núi Mot lake - Binh Dinh province (in Vietnames)", *J. Mining Earth Sci.*, **61**(5), 117-124. <https://doi.org/10.46326/JMES.KTLT2020.10>.
- Andersen, L. and Nielsen, S.R. (2005), "Reduction of ground vibration by means of barriers or soil improvement along a railway track", *Soil Dyn. Earthq. Eng.*, **25**(7-10), 701-716. <https://doi.org/10.1016/j.soildyn.2005.04.007>.
- Arabameri, A., Pourghasemi, H.R. and Yamani, M. (2017), "Applying different scenarios for landslide spatial modeling using computational intelligence methods", *Environ. Earth Sci.*, **76**(24), 832. <https://doi.org/10.1007/s12665-017-7177-5>.
- Army, U. (1998). *Technical manual design and analysis of hardened structures to conventional weapons effects*, Army TM5-855-1, Washington D.C.
- Asteris, P., Kolovos, K., Douvika, M. and Roinos, K. (2016), "Prediction of self-compacting concrete strength using artificial neural networks", *European J. Environ. Civil Eng.*, **20**(sup1), s102-s122. <https://doi.org/10.1080/19648189.2016.1246693>.
- Asteris, P., Roussis, P. and Douvika, M. (2017), "Feed-forward neural network prediction of the mechanical properties of sandcrete materials", *Sensors*, **17**(6), 1344. <https://doi.org/10.3390/s17061344>.
- Asteris, P.G. and Mokos, V.G. (2019), "Concrete compressive strength using artificial neural networks", *Neural Computing and Applications*, 1-20. <https://doi.org/10.1007/s00521-019-04663-2>.
- Asteris, P.G., Apostolopoulou, M., Skentou, A.D. and Moropoulou, A. (2019), "Application of artificial neural networks for the prediction of the compressive strength of cement-based mortars", *Comput. Concrete*, **24**(4), 329-345. <https://doi.org/10.12989/cac.2019.24.4.329>.
- Bayzid, S.M., Mohamed, Y. and Al-Hussein, M. (2016), "Prediction of maintenance cost for road

- construction equipment: a case study”, *Canadian J. Civil Eng.*, **43**(5), 480-492. <https://doi.org/10.1139/cjce-2014-0500>.
- Behmanesh, J. and Mehdizadeh, S. (2017), “Estimation of soil temperature using gene expression programming and artificial neural networks in a semiarid region”, *Environ. Earth Sci.*, **76**(2), 76. <https://doi.org/10.1007/s12665-017-6395-1>.
- Breiman, L. (2001), “Random forests”, *Machine Learning*, **45**(1), 5-32. https://doi.org/10.1007/978-1-4419-9326-7_5.
- Bui, X.N., Choi, Y., Atrushkevich, V., Nguyen, H., Tran, Q.H., Long, N.Q. and Hoang, H.T. (2019a), “Prediction of blast-induced ground vibration intensity in open-pit mines using unmanned aerial vehicle and a novel intelligence system”, *Nat. Resources Res.*, **29**(2), 771-790. <https://doi.org/10.1007/s11053-019-09573-7>.
- Bui, X.N., Nguyen, H., Le, H.A., Bui, H.B. and Do, N.H. (2019b), “Prediction of blast-induced air overpressure in open-pit mine: Assessment of different artificial intelligence techniques”, *Nat. Resources Res.* 1-25. <https://doi.org/10.1007/s11053-019-09461-0>.
- Cavaleri, L., Asteris, P.G., Psyllaki, P.P., Douvika, M.G., Skentou, A.D. and Vaxevanidis, N.M. (2019), Prediction of surface treatment effects on the tribological performance of tool steels using artificial neural networks”, *Appl. Sci.*, **9**(14), 2788. <https://doi.org/10.3390/app9142788>.
- Çaylak, Ç. and Kaftan, İ. (2014), “Determination of near-surface structures from multi-channel surface wave data using multi-layer perceptron neural network (MLPNN) algorithm”, *Acta Geophysica*, **62**(6), 1310-1327. <https://doi.org/10.2478/s11600-014-0207-8>.
- Detle, H. and Scheder, R. (2006), “Strictly monotone and smooth nonparametric regression for two or more variables”, *Canadian J. Statistics*, **34**(4), 535-561. <https://doi.org/10.1002/cjs.5550340401>.
- Duan, J., Asteris, P.G., Nguyen, H., Bui, X.N. and Moayed, H. (2020), “A novel artificial intelligence technique to predict compressive strength of recycled aggregate concrete using ICA-XGBoost model”, *Eng. Comput.*, 1-18, <https://doi.org/10.1007/s00366-020-01003-0>.
- El-Bendary, N., Elhariri, E., Hazman, M., Saleh, S.M. and Hassanien, A.E. (2016), “Cultivation-time recommender system based on climatic conditions for newly reclaimed lands in Egypt”, *Procedia Comput. Science*, **96**, 110-119.
- Fang, Q., Nguyen, H., Bui, X.N. and Tran, Q.H. (2019), “Estimation of blast-induced air overpressure in quarry mines using cubist-based genetic algorithm”, *Nat. Resources Res.*, <https://doi.org/10.1007/s11053-019-09575-5>.
- Gao, W., Alqahtani, A.S., Mubarakali, A., Mavaluru, D. and Khalafi, S. (2019), “Developing an innovative soft computing scheme for prediction of air overpressure resulting from mine blasting using GMDH optimized by GA”, *Eng. Comput.*, **36**(2), 647-654. <https://doi.org/10.1007/s00366-019-00720-5>.
- Gao, W., Guirao, J.L.G., Abdel-Aty, M. and Xi, W. (2019), “An independent set degree condition for fractional critical deleted graphs”, *Discrete Continuous Dyn. Syst-S.*, **12**(4-5), 877-886.
- 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”, *Appl. Acoustics*, **80**, 57-67. <https://doi.org/10.1016/j.apacoust.2014.01.005>.
- 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”, *Eng. Comput.*, **32**(3), 441-455. <https://doi.org/10.1007/s00366-015-0425-y>.
- 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”, *Eng. Comput.*, **33**(1), 23-31. <https://doi.org/10.1007/s00366-016-0453-2>.
- Hoang, N., Nam Xuan, B., Hieu Quang, T. and Giang Huong Thi, L. (2020), “A novel soft computing model for predicting blast - induced ground vibration in open - pit mines using gene expression programming”, *J. Mining Earth Sci.*, **61**(5), 107-116. <https://doi.org/10.46326/JMES.KTLT2020.09>.
- Holmes, G., Hall, M. and Prank, E. (1999), “Generating rule sets from model trees”, *Australasian Joint Conference on Artificial Intelligence*, Springer.
- Hresc, J., Riley, E. and Harris, P. (2018), “Mining project’s economic impact on local communities, as a

- social determinant of health: A documentary analysis of environmental impact statements”, *Environ. Impact Assessment Rev.*, **72**, 64-70. <https://doi.org/10.1016/j.eiar.2018.05.009>.
- Iphar, M., Yavuz, M. and Ak, H. (2008), “Prediction of ground vibrations resulting from the blasting operations in an open-pit mine by adaptive neuro-fuzzy inference system”, *Environ. Geology*, **56**(1), 97-107.
- Juniah, R. (2018), “Study of carbon value of the allotment of former coal mining land of PT Samantaka Batubara for sustainable mining environment”, *J. Sustain. Dev*, **11**(4), 213.
- Kaufman, Y. and Tamir, A. (1993), “Locating service centers with precedence constraints”, *Discrete Appl. Mathem.*, **47**(3), 251-261. [https://doi.org/10.1016/0166-218X\(93\)90130-G](https://doi.org/10.1016/0166-218X(93)90130-G).
- Khaki, M., Yusoff, I. and Islami, N. (2015), “Simulation of groundwater level through artificial intelligence system”, *Environ. Earth Sci.*, **73**(12), 8357-8367. <https://doi.org/10.1007/s12665-014-3997-8>.
- Khan, A.R., Mir, A.A., Saeed, S., Rafique, M., Asim, K.M., Iqbal, T., Jabbar, A. and Rahman, S.U. (2018). “Classification of rocks radionuclide data using machine learning techniques”, *Acta Geophysica*, **66**(5), 1073-1079. <https://doi.org/10.1007/s11600-018-0190-6>.
- Khandelwal, M. and Kankar, P. (2011), “Prediction of blast-induced air overpressure using support vector machine”, *Arabian J. Geosci.*, **4**(3-4), 427-433. <https://doi.org/10.1007/s12517-009-0092-7>.
- Kuzu, C., Fisne, A. and Ercelebi, S. (2009), “Operational and geological parameters in the assessing blast induced airblast-overpressure in quarries”, *Applied Acoustics*, **70**(3), 404-411. <https://doi.org/10.1016/j.apacoust.2008.06.004>.
- Long Quoc, N., Dung Ngoc, V. and My Chi, V. (2020), “Advanced mining geomatic technologies serving open-pit mining operation in Vietnam (in Vietnamese)”, *J. Mining Earth Sci.*, **61**(5), 125-133. <https://doi.org/10.46326/JMES.KTLT2020.11>.
- Longjun, D., Xibing, L., Ming, X. and Qiyue, L. (2011), “Comparisons of random forest and support vector machine for predicting blasting vibration characteristic parameters”, *Procedia Eng.*, **26**, 1772-1781. <https://doi.org/10.1016/j.proeng.2011.11.2366>.
- Luss, R., Rosset, S. and Shahar, M. (2010), “Decomposing isotonic regression for efficiently solving large problems”, *In Advances in Neural Information Processing Systems*.
- Mahdiyari, A., Marto, A. and Mirhosseini, S.A. (2018), “Probabilistic air-overpressure simulation resulting from blasting operations”, *Environm. Earth Sci.*, **77**(4),123. <https://doi.org/10.1007/s12665-018-7293-x>.
- Mehdizadeh, S., Behmanesh, J. and Khalili, K. (2017), “Evaluating the performance of artificial intelligence methods for estimation of monthly mean soil temperature without using meteorological data”, *Environ. Earth Sci.*, **76**(8), 325. <https://doi.org/10.1007/s12665-017-6607-8>.
- Mohamed, M.T. (2011), “Performance of fuzzy logic and artificial neural network in prediction of ground and air vibrations”, *Int. J. Rock Mech. Mining Sci.*, **48**(5), 845-851.
- Monjezi, M., Shahriar, K., Dehghani, H. and Namin, F.S. (2009), “Environmental impact assessment of open pit mining in Iran”, *Environ. Geology*, **58**(1), 205-216. <https://doi.org/10.1007/s00254-008-1509-4>.
- 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”, *Nat. Resources Res.*, 1-15. <https://doi.org/10.1007/s11053-018-9424-1>.
- Nguyen, H. and Bui, X.N. (2020), “Soft computing models for predicting blast-induced air over-pressure: A novel artificial intelligence approach”, *Appl. Soft Comput.*, **92**, 106292. <https://doi.org/10.1016/j.asoc.2020.106292>.
- Nguyen, H., Bui, X.N. and Moayedi, H. (2019b), “A comparison of advanced computational models and experimental techniques in predicting blast-induced ground vibration in open-pit coal mine”, *Acta Geophysica*, **67**(4), 1025-1037. <https://doi.org/10.1007/s11600-019-00304-3>.
- Nguyen, H., Bui, X.N., Bui, H.B. and Cuong, D.T. (2019a), “Developing an XGBoost model to predict blast-induced peak particle velocity in an open-pit mine: A case study”, *Acta Geophysica*, **67**(2), 477-490. <https://doi.org/10.1007/s11600-019-00268-4>.
- 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 Comput. Appl.* 1-17. <https://doi.org/10.1007/s00521-018-3717-5>.

- Nguyen, H., Bui, X.N., Tran, Q.H. and Mai, N.L. (2019c), "A new soft computing model for estimating and controlling blast-produced ground vibration based on hierarchical K-means clustering and cubist algorithms", *Appl. Soft Comput.*, **77**, 376-386. <https://doi.org/10.1016/j.asoc.2019.01.042>.
- Nguyen, H., Bui, X.N., Tran, Q.H., Van Hoa, P., Nguyen, D.A., Hoa, L.T.T., Le, Q.T., Do, N.H., Bao, T.D., Bui, H.B. and Moayedi, H. (2020). "A comparative study of empirical and ensemble machine learning algorithms in predicting air over-pressure in open-pit coal mine", *Acta Geophysica*, <https://doi.org/10.1007/s11600-019-00396-x>.
- Nguyen, H., Drebenstedt, C., Bui, X.N. and Bui, D.T. (2019d), "Prediction of blast-induced ground vibration in an open-pit mine by a novel hybrid model based on clustering and artificial neural network", *Nat. Resources Res.*, <https://doi.org/10.1007/s11053-019-09470-z>.
- Remennikov, A.M. and Rose, T.A. (2007), "Predicting the effectiveness of blast wall barriers using neural networks", *Int. J. Impact Eng.*, **34**(12), 1907-1923. <https://doi.org/10.1016/j.ijimpeng.2006.11.003>.
- Saarela, O. and Arjas, E. (2011), "A method for Bayesian monotonic multiple regression", *Scandinavian J. Statistics*, **38**(3), 499-513. <https://doi.org/10.1111/j.1467-9469.2010.00716.x>.
- Sarir, P., Chen, J., Asteris, P.G., Armaghani, D.J. and Tahir, M.M. (2019), "Developing GEP tree-based, neuro-swarm, and whale optimization models for evaluation of bearing capacity of concrete-filled steel tube columns", *Eng. Comput.*, 1-19. <https://doi.org/10.1007/s00366-019-00808-y>.
- Sawmliana, C., Roy, P.P., Singh, R. and Singh, T. (2007), "Blast induced air overpressure and its prediction using artificial neural network", *Mining Technol.*, **116**(2), 41-48. <https://doi.org/10.1179/174328607X191065>.
- Segarra, P., Domingo, J., López, L., Sanchidrián, J. and Ortega, M. (2010), "Prediction of near field overpressure from quarry blasting", *Appl. Acoustics*, **71**(12), 1169-1176. <https://doi.org/10.1016/j.apacoust.2010.07.008>.
- Sharma, R., Kumar, S. and Maheshwari, R. (2015), "Comparative analysis of classification techniques in data mining using different datasets", *Int. J. Comput. Sci. Mobile Comput.*, **4**(12), 125-134.
- Siskind, D.E., Stachura, V.J., Stagg, M.S. and Kopp, J.W. (1980), "Structure response and damage produced by airblast from surface mining", Citeseer.
- Stout, Q.F. (2015). "Isotonic regression for multiple independent variables", *Algorithmica*, **71**(2), 450-470. <https://doi.org/10.1007/s00453-013-9814-z>.
- Tuoc Ngoc, D., Giao Si, H., Xuan Manh, T., Thanh Van, D. and Nam Duy, B. (2020), "Mining technology for exploiting deep open-pit mines in Vietnam (in Vietnamese)", *J. Mining Earth Sci.*, **61**(5), 47-57. <https://doi.org/10.46326/JMES.KTLT2020.04>.
- Zhang, S., Bui, X.N., Trung, N.T., Nguyen, H. and Bui, H.B. (2019), "Prediction of rock size distribution in mine bench blasting using a novel ant colony optimization-based boosted regression tree technique", *Nat. Resources Res.*, <https://doi.org/10.1007/s11053-019-09603-4>.
- Zhang, X., Nguyen, H., Bui, X.N., Tran, Q.H., Nguyen, D.A., Bui, D.T. and Moayedi, H. (2019), "Novel soft computing model for predicting blast-induced ground vibration in open-pit mines based on particle swarm optimization and XGBoost", *Nat. Resources Research*, 1-11. <https://doi.org/10.1007/s11053-019-09492-7>.