

Development of GP and GEP models to estimate an environmental issue induced by blasting operation

Roohollah Shirani Faradonbeh ·
Mahdi Hasanippanah · Hassan Bakhshandeh Amnieh ·
Danial Jahed Armaghani · Masoud Monjezi

Received: 24 May 2017 / Accepted: 11 May 2018 / Published online: 21 May 2018
© Springer International Publishing AG, part of Springer Nature 2018

Abstract Air overpressure (AOp) is one of the most adverse effects induced by blasting in the surface mines and civil projects. So, proper evaluation and estimation of the AOp is important for minimizing the environmental problems resulting from blasting. The main aim of this study is to estimate AOp produced by blasting operation in Miduk copper mine, Iran, developing two artificial intelligence models, i.e., genetic programming (GP) and gene expression programming (GEP). Then, the accuracy of the GP and GEP models has been

compared to multiple linear regression (MLR) and three empirical models. For this purpose, 92 blasting events were investigated, and subsequently, the AOp values were carefully measured. Moreover, in each operation, the values of maximum charge per delay and distance from blast points, as two effective parameters on the AOp, were measured. After predicting by the predictive models, their performance prediction was checked in terms of variance account for (VAF), coefficient of determination (CoD), and root mean square error (RMSE). Finally, it was found that the GEP with VAF of 94.12%, CoD of 0.941, and RMSE of 0.06 is a more precise model than other predictive models for the AOp prediction in the Miduk copper mine, and it can be introduced as a new powerful tool for estimating the AOp resulting from blasting.

R. S. Faradonbeh
Young Researchers and Elite Club, South Tehran Branch, Islamic Azad University, Tehran, Iran
e-mail: Roohollah_Shirani@yahoo.com

M. Hasanippanah (✉)
Department of Mining Engineering, University of Kashan, Kashan, Iran
e-mail: Hasanippanah.m@grad.kashanu.ac.ir
e-mail: Hasanippanah.m@gmail.com

H. B. Amnieh
School of Mining, College of Engineering, University of Tehran, Tehran 11155-4563, Iran
e-mail: hbakhshandeh@ut.ac.ir

D. J. Armaghani
Department of Civil and Environmental Engineering, Amirkabir University of Technology, Tehran 15914, Iran
e-mail: danialarmaghani@gmail.com

M. Monjezi
Department of Mining Engineering, Tarbiat Modares University, Tehran 14115-143, Iran
e-mail: Monjezi@modares.ac.ir

Keywords Blasting · Air overpressure · Genetic programming · Gene expression programming

Abbreviations

ANN	Artificial neural network
AOp	Air overpressure
CoD	Coefficient of determination
CPs	Computer programs
D	Distance between monitoring station and blast point
EAs	Evolutionary algorithms
ETs	Expression trees
F	Function set
GA	Genetic algorithm
GEP	Gene expression programming

GP	Genetic programming
ICA	Imperialist competitive algorithm
MC	Maximum charge used per delay
MLR	Multiple linear regression
PSO	Particle swarm optimization
RMSE	Root mean square error
T	Terminal set
USBM	US Bureau of Mines
VA	Variance account for

Introduction

Blasting is a broadly used method for rock breakage in the civil and mining industries. Ground vibration, air overpressure (AOp), flyrock, and backbreak are considered as the environmental side effects induced by blasting. Hence, accurate estimation of these unwanted effects is crucial (Bhandari 1997; Monjezi et al. 2010; Rezaei et al. 2011; Mohammadnejad et al. 2013; Hasanipanah et al. 2015a, b; Jahed Armaghani et al. 2016a). Among these effects, AOp is one of the most undesirable effects and can produce damage to nearby structure (Hajihassani et al. 2015). Based on literature, AOp is produced through a large shock wave from explosive into the surface which refracted horizontally (Khandelwal and Singh 2005; Jahed Armaghani et al. 2015). The unspent energy in the explosion gases above the normal air atmospheric level is defined as AOp. Based on literature, AOp which is also considered as air waves is created based on four sources, i.e., air-pressure pulse, rock-pressure pulse, gas-pressure pulse, and stemming-release pulse. More explanations regarding these four sources can be found in the other references (Wiss and Linehan 1978; Siskind et al. 1980). AOp is measured by Pascal (Pa) or decibels (dB) and can be recognized by sound. The lowest detectable sound for humans is 20 Hz (Kuzu et al. 2009). Therefore, there is no doubt that humans may not react to sounds of less than 20 Hz. A value of 180 dB is the limit of structural damage possibility; a range of 130–180 dB can be resulted to break glass. In addition, an ultimate value of 130 dB can be damaged to window (or at least window vibrations) (Griffiths et al. 1978; Rodríguez et al. 2010). So, it is of interest if values lower than 110 dB are obtained for AOp. Many parameters affect the intensity of AOp. These parameters can be divided into three main categories,

i.e., blast design parameters, properties of explosive materials, and rock mass (Hemphill 1981; McKenzie 1990; Segarra et al. 2010). The first and second groups are considered as controllable parameters, whereas the third group is considered as natural and non-controllable parameters. Blast design parameters such as burden, stemming, spacing between blastholes, maximum charge used per delay (MC), total charge, blasthole diameter, blasthole length, delay timing, sub-drilling, blasthole inclination, and blasting type (production or controlled) are the controllable parameters. Properties of explosive materials such as type (dynamite, emulsion, and ANFO), density, and strength of explosive materials are also considered as the controllable parameters. In the third group, non-controllable parameters such as shear, compressive, and tensile strength of rock mass are related to geological and geotechnical conditions of rock mass. In the literature, many attempts have been done for AOp prediction through empirical models (EMs) (Siskind et al. 1980; Rodríguez et al. 2007). The EMs were constructed based on two influential parameters on the AOp, MC, and distance from blast points. These EMs are mainly simple and quick; nevertheless, they have low-performance capacity, as mentioned in many studies (Hasanipanah et al. 2015b; Jahed Armaghani et al. 2015). Besides, the proper estimation of the AOp is important for determination of blast safety area. The artificial intelligence (AI) methods have been widely highlighted for solving the geotechnical problems (e.g., Ahmad et al. 2017; Singh et al. 2017; Sharma et al. 2017a, b, c) and more specifically in the field of AOp estimation (Singh and Chakrapani 2015; Ravansalar and Rajaei 2015; Sankararajan et al. 2017; Deo and Şahin 2016; Singh et al. 2016; Verma et al. 2016; Naderi et al. 2017). In this regard, Khandelwal and Kankar (2011) developed support vector machine (SVM) and generalized predictor model to estimate AOp in three mines, India. They concluded that SVM can predict AOp better than generalized predictor equation. Tonnizam Mohamad et al. (2012) investigated the application of artificial neural network (ANN) to estimate blast-induced AOp. In their study, datasets were gathered from two granite quarry sites in Malaysia. As a result, it was found that the ANN can be introduced as a reasonable tool for the AOp prediction. A hybrid model of ANN and particle swarm optimization (PSO) algorithm were employed to predict AOp by Hajihassani et al. (2015). Their

results indicated that the PSO is a useful algorithm to design the ANN, and obtained results from PSO-ANN model is better than ANN and EMs. Another evolutionary algorithm, namely imperialist competitive algorithm (ICA), was used to design ANN in predicting the AOp by Jahed Armaghani et al. (2016b). They demonstrated that ICA-ANN is a more acceptable tool than ANN and other EMs.

Evolutionary algorithms (EAs) are one of the most applicable and famous stochastic meta-heuristics. The literature review shows the success use of these algorithms in different real and complex problems (Gong et al. 2016; Xu et al. 2016; Qazi et al. 2016; Zou et al. 2017). Afterwards, a new computation field was named as evolutionary computation (Back et al. 1997). Genetic programming (GP) and gene expression programming (GEP) are the subset of EAs that are developed versions of the genetic algorithm (GA). The GP and GEP methods are the powerful tools for function finding in non-linear and highly complex problems (Ghotbi Ravandi et al. 2013; Shirani Faradonbeh et al. 2015, 2016a, b; Khandelwal et al. 2016, 2017). In the present research work, GP and GEP models are used to develop mathematical predictor functions for estimating the AOp in Miduk copper mine, Iran. For comparison purposes, regression analysis and three EMs are also developed.

Methodology

Genetic programming

GP paradigm is a subset of EAs which applies the Darwinian principle of evolution to hierarchical computer programs and has produced promising breakthroughs in various scientific and engineering applications. GP firstly invented by Cramer (1985) and then improved mathematically by Koza (1996). GP is an advanced version of GA, and their main difference is in the structure of individuals. The principal components of GA are binary coded strings, which are called chromosomes and demonstrate proposed solutions to a problem. In GP, solutions are in the form of computer programs (CPs) which follow a computer programming language namely LISP and can be arranged as tree structures with diverse sizes and shapes. Each CP is a combination of the terminal set (T) and function set (F). The terminals are the input parameters and the constant values that can be defined by GP designer. The function set consists of mathematical, logical, Boolean, or user-defined functions that can be determined according to the nature of the problem and the suggestions of other scholars (Koza 1996; Ferreira 2001; Hoseinian et al. 2017; Shirani Faradonbeh et al., 2017; Jahed Armaghani et al. 2018). Some of these functions are listed as follows (Ferreira 2006):

$$F = \{+, -, *, /, ^2, ^3, Q, 3Rt, \sin, \cos, \tan, \text{Asin}, \text{Acos}, \text{Atan}, \text{Exp}, \text{Ln}, \log, \text{and}, \text{or}, \text{not}, \text{nor}, \dots\} \tag{1}$$

where Q is the square root of a number, Asin is the inverse trigonometric function for \sin , and $3Rt$ is the cube root of the terminal. Figure 1 shows the process of GP modeling. GP algorithm starts with the random generation of initial population (CPs) using one of three strategies of *Grow*, *Half*, and *Full*. Then, the fitness of each CPs is evaluated based on the defined fitness function. The best individuals of the first generation will be selected by the selection operator to reproduce into the next generation. There are four different selection methods including roulette wheel selection, tournament selection, rank selection, and lexicographic parsimony pressure selection (Ferreira 2006). Roulette wheel method, which is also called fitness proportionate selection method, is more common among others. Suppose that $f(s_i(t))$ is the fitness of individual s_i in the population at

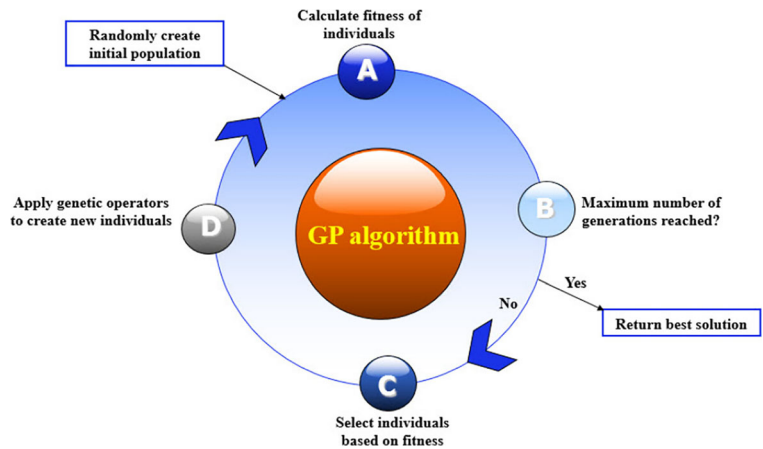
generation t ; then, under roulette wheel selection method, the individual s_i will be copied into the next generation with the probability of $P(S_j(t))$ as a result of reproduction operation (Ferreira 2006).

$$P(S_j(t)) = \frac{f(s_i(t))}{\sum_{j=1}^M f(s_j(t))} \tag{2}$$

where $\sum_{j=1}^M f(s_j(t))$ is the total amount of fitness values with M number of chromosomes. In the following, the probability values of chromosomes are calculated as a sector of the circle and placed on it, and the process of selection will be done by a selector (see Fig. 2).

Afterwards, the stopping criteria (i.e., the maximum number of generation or the favorable value for fitness function) will be checked and if are not met, the process

Fig. 1 The flow chart of the GP algorithm (Khandelwal et al. 2017)



will continue. In this stage, two important genetic operators are applied to the remained individuals to create the modified CPs. Crossover is one of the genetic operators that selects two CPs (parents) randomly and exchanges some parts of parse trees between them that lead to the creation of two new children. Another genetic operator is mutation, which has this ability to change any nodes of parse trees and create modified programs. The modified CPs caused by the applying genetic operators on the first population are copied to the next generation, and again, the fitness of the second generation will be evaluated. This process will be repeated up to meet the stopping condition. Although GP has some preferences to other algorithms, it suffers from some problems such as generation of infeasible solutions because of difficulties in applying genetic operators to the parse trees and uncontrollable code growth. More information regarding GP algorithm can be found in the study conducted by Koza (1996).

Gene expression programming

GA and GP are well-known techniques in the fields of mining and rock mechanics. Nevertheless, gene expression programming (GEP) has not been used extensively in these fields. GEP was introduced by Ferreira (2006) as a genotype-phenotype system which contains the simplicity of GA and the abilities of GP. As mentioned before, GP suffers from shortcomings that GEP can solve them. In GEP, solutions are chromosomes which follow the Karva language and then are expressed as expression trees (ETs) (Ferreira 2001, 2006). A chromosome contains one or more gene (sub-ET). In multi-genic chromosomes, these sub-ETs are connected to each other using a linking function (e.g., +, -, ×, /) to create a big and complex ET. Each gene is made from two parts of the head (*h*) which may contain any functions or terminals, and tail (*t*) that only contains terminal set. Figure 3a shows an example of a chromosome

Fig. 2 Roulette wheel selection method (Hoseinian et al. 2017)

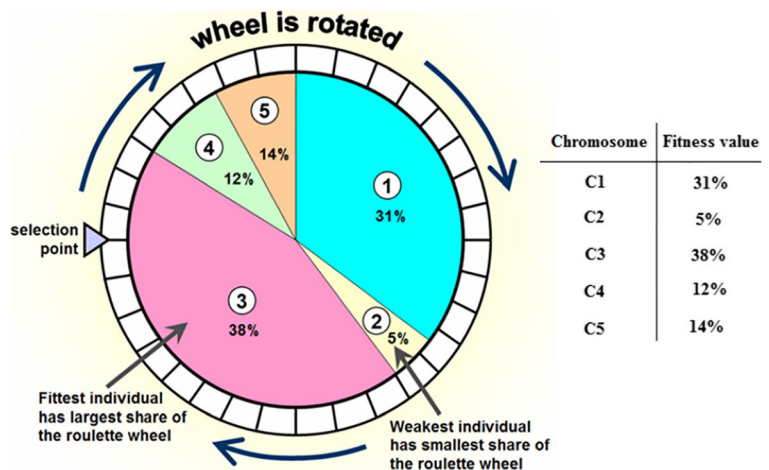
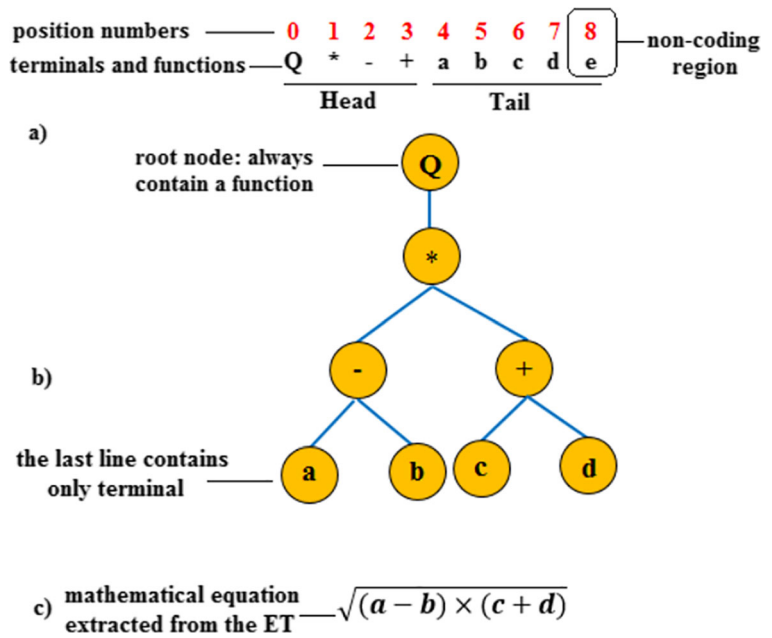


Fig. 3 The structure of simple chromosome with different expressions in GEP. **a** K-Expression. **b** Expression tree (ET). **c** Mathematical equation



coded in Karva language (K-Expression). In GEP algorithm, the coded chromosomes should be expressed as ETs (see Fig. 3b). Ferreira (2006) suggested four rules to the syntactically correct expression of ETs:

Rule 1: The function in the position no. 0 is the root node of ET.

Rule 2: According to the argument number of function, each node is divided into several sub-nodes. For instance, the common functions of -, +, /, and × have two arguments (sub-nodes), while Q (square root) has just one argument. Terminals have no argument.

Rule 3: According to the order of functions and terminals along the chromosome, they are embedded from top to down and left to right in each line of nodes.

Rule 4: The process of ET formation continues until a line without any function is created.

In GEP, there are some fragments along the chromosomes namely non-coding regions which cannot be expressed as ET (e.g., the position no. 8 in Fig. 3a) but have a great role in the evolution process and creating valid solutions. After generating an ET, its mathematical equation can be extracted (see Fig. 3c). The length of the head (*h*) is a parameter that there is no definite way to

determine its optimum value and there is a need to use a trial and error procedure. The length of the tail can be calculated using the following equation (Ferreira 2001, 2006; Shirani Faradonbeh et al. 2018):

$$t = h(n_{max}-1) + 1 \tag{3}$$

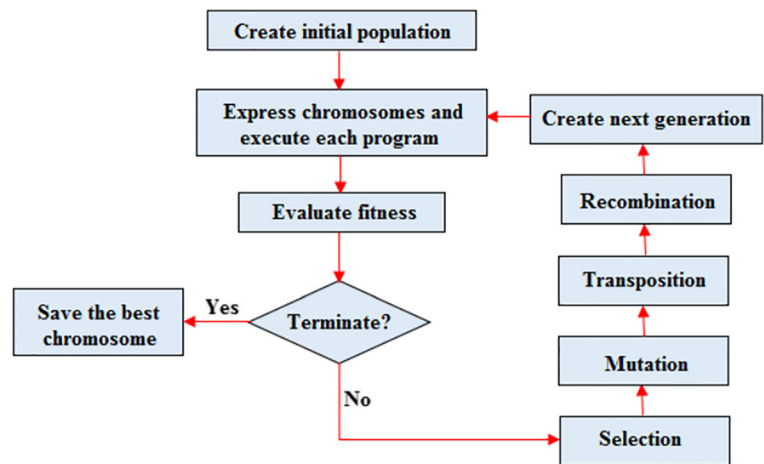
where *n_{max}* is the maximum number of arguments.

According to Fig. 3, the set of functions is $F = \{Q, *, -, +\}$ and the set of terminals is $T = \{a, b, c, d, e\}$. According to *F*, the *n_{max}* = 2 with *h* = 4, then $t = 4 \times (2 - 1) + 1 = 5$.

According to Fig. 4, the process of GEP modeling can be summarized into five below steps:

- Step 1: By combining the terminals and functions in a stochastic manner, the first population of chromosomes is generated.
- Step 2: Following the four aforementioned rules, the chromosomes are presented as ETs and their mathematical equations are achieved, respectively.
- Step 3: After evaluating the fitness of chromosomes and checking the termination condition, the fittest chromosomes are selected and copied to the next generation.
- Step 4: In this step, genetic operators which have key role in modifying the initial chromosomes are

Fig. 4 The flow chart of GEP algorithm



applied to the initial population. Ferreira introduced several genetic operators to improve the performance of GEP. By applying these operators, the next generation of population is created.

Step 5: The steps 2 to 4 are repeated until the best solution is obtained.

The next section introduces the genetic operators in GEP concisely.

Genetic operators in GEP algorithm

Mutation

Mutation is the operator with wide latitude which can occur anywhere in the chromosome length provided that the structure of the chromosome remains constant. In head, mutation can convert any function or terminal to each other. In tail, mutation can only convert terminal to another terminal. A range of (0.01 – 0.1) is suggested for this operator in the literature (e.g., Ferreira 2006; Shirani Faradonbeh and Monjezi, 2017).

Inversion

This operator which is restricted to the head of genes creates a big influence by reversing a small fragment (e.g., one to three positions). The suggested rate for this operator is 0.1 (Ferreira 2006).

Transposition

Some transposable fragments jump to other positions along a chromosome in three different methods: (1) IS transposition: short fragments with a function or terminal in the first position transpose to the head of genes except the root, (2) RIS transposition: short fragments with function in the first position transpose to the root of genes, and (3) gene transposition: an entire gene transposes itself to the beginning of the chromosome. In contrast to the other forms of transposition, in gene transposition, the gene is deleted at the place of origin. The range of 0.01 to 0.1 is suggested by scholars for transposition operator (Ferreira 2006; Khandelwal et al. 2017).

Recombination

Recombination, which is also called crossover, selects two chromosomes randomly, and using one of three different methods of one-point recombination, two-point recombination, or gene recombination replaces some material between them and generates two new individuals. The suggested value for the sum of these three recombination operators is 0.7 (Ferreira 2006).

Case study

In the present research, the studied area is Miduk copper mine. This mine is one of the most important copper mines in Iran, which is located approximately 132 km northwest of Sar-cheshmeh porphyry copper mine in Kerman province, between 30° 25' 14" latitudes and

55° 10' 2" longitudes (see Fig. 5). The mineralogical combinations in the Miduk mine are given in Table 1. The total reserve of the deposit is approximately 170 Mt with 0.85 and 0.25% for copper grade and cut of grade, respectively. The average bench height and overall slope angle are 15 m and 38°, respectively. Also, the overall stripping ratio is 2.4. Blasting operation is performed for the rock fragmentation in the Miduk copper mine. In this regard, the mainly used explosive material is ammonium nitrate fuel oil (ANFO). As mentioned earlier, AOp is considered as an undesirable effect induced by blasting operation. Therefore, prediction of this environmental effect is essential for the safety issues around the mines. For achieving the aim of this study, 92 blasting operations were monitored and the values of AOp were recorded in terms of decibel (dB) using Minimate Pulse instrument. Table 2 gives the properties of this instrument. In addition, the values of two influential parameters on the intensity of AOp, i.e., MC and D, were measured for the total of 92 blasting events. Table 3 summarizes the range of used parameters in this research.

Prediction of AOp

In this section, the development of GP, GEP, and multiple linear regression (MLR) models to predict AOp in Miduk copper mine is described. To model the predictors, the experimental datasets were categorized into two sets: train and test. Training datasets were applied to calibrate the models, while testing datasets were applied

Table 1 The mineralogical combinations in the ore of Miduk copper mine

Mineral	Description	Mineral	Description
Limonite	FeOOH ≈ 0.11%	Hematite	Fe ₂ O ₃ ≈ 0.12%
Magnetite	Fe ₃ O ₄ ≈ 0.23%	Pyrite	FeS ₂ ≈ 6%
Chalcopyrite	CuFeS ₂ ≈ 0.33%	Chalcocite	Cu ₂ S ≈ 1%

to test the validity of the models. In this research, 80 and 20% of the total experimental data (74 and 18 data) were selected randomly for training and testing aims, in order, as recommended by Nelson and Illingworth (1990) and Swingler (1996).

AOp prediction using MLR

MLR is a statistical method which is widely used to solve many problems in the fields of mining and rock engineering. In the MLR, a linear equation between two or more independent variables and one dependent variable can be fitted. As an example, Sari et al. (2014) developed multiple regression model to estimate blast-induced backbreak in the Sungun copper mine. They showed a high conformity (determination coefficient = 0.981) between the measured and predicted backbreak values by multiple regression model. In the other study, Sharma and Singh (2017) employed regression models for the predicting unconfined compressive strength of artificially structured soil, and based on their obtained results, regression models were acceptable

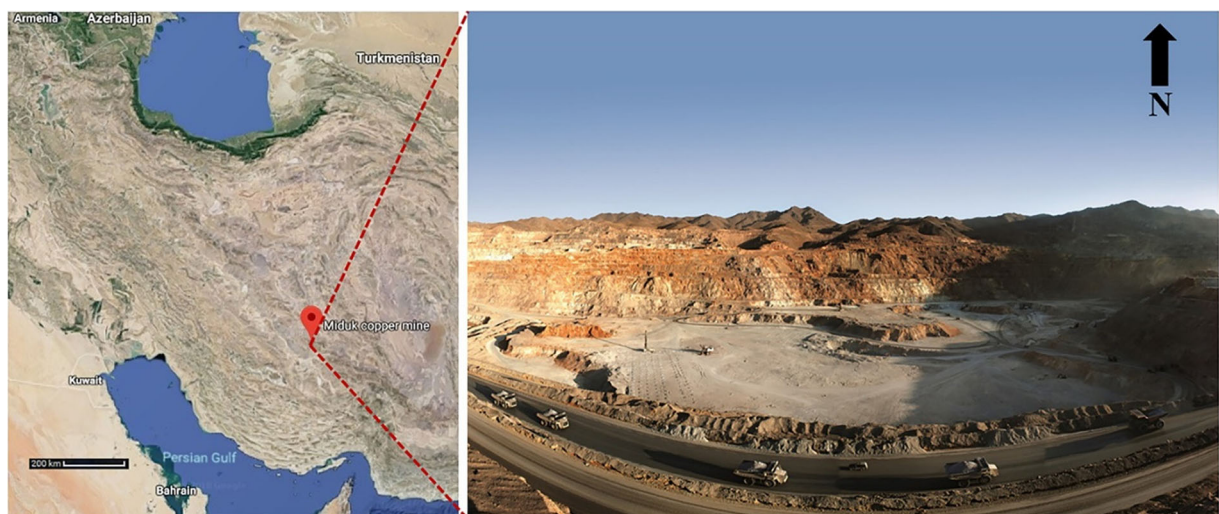


Fig. 5 The location of the Miduk copper mine

Table 2 The properties of the Minimate Pulse

Parameter	Value
Linear range	88 to 148 dB
Linear resolution	500 Pa peak
Linear accuracy	0.25 Pa
Liner frequency response	2 to 250 Hz

models in the studied field. Generally, the MLR is formulated as below:

$$Y = P_0 + P_1X_1 + \dots + P_nX_n \tag{4}$$

in which X_i ($i = 1, \dots, n$) and Y are input and output parameters, in order. Also, P_i ($i = 0, 1, \dots, n$) present regression coefficients. As stated previously, considering training datasets, Eq. 5 was constructed using SPSS v16 software:

$$AOp = 127.91 + (-0.05 \times D) + (0.014 \times MC) \tag{5}$$

In Eq. 5, MC and D are in terms of kilograms and meters, respectively. More details regarding evaluation of developed MLR equation are given later.

AOp prediction using empirical models

In the present paper, three EMs were used for predicting the AOp. National Association of Australian State (1983) developed an EM for the AOp prediction as follows:

$$P = \frac{140 \sqrt[3]{\frac{MC}{200}}}{D} \tag{6}$$

where P is overpressure in terms of kilo Pascal. Note that 1 Pa is 94 dB. In the other research, McKenzie

Table 3 The range of used parameters in this study

Parameter	Unit	Min	Max
Maximum charge weight used per delay	kg	215	715
Distance between monitoring station and blast point	m	211	593
AOp	dB	100.6	126.1

Table 4 The parameter values for GP constructed model

GP algorithm parameters		Parameter settings
Initial parameters	Population size	100
	Fitness function	RMSE
	Initialization method	Half
	Selection method	Roulette wheel
	Number of generation	200
Genetic operators	Crossover	0.9
	Mutation	0.05
	Reproduction	0.2

(1990) used the following equation for predicting the AOp:

$$dB = 165 - 24 \log \left(\frac{D}{MC^{1/3}} \right) \tag{7}$$

Also, the US Bureau of Mines (USBM) is used as the most common EM in many studies (Rosenthal and Morlock 1987; Cengiz 2008) and is formulated as follows:

$$AOp = k \left(\frac{D}{MC^{1/3}} \right)^n \tag{8}$$

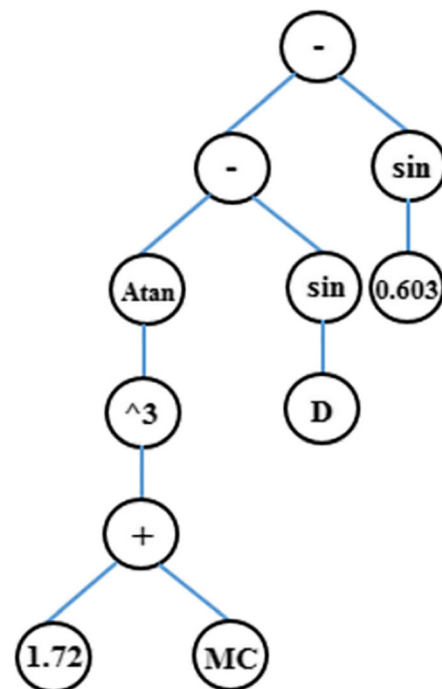


Fig. 6 Parse tree (S-Expression) of the best GP model

Table 5 The parameters of the developed GEP model

GEP algorithm parameters		Parameter settings
Initial parameters	Number of chromosomes	50
	Head size	9
	Number of genes	5
	Linking function	Addition (+)
	Fitness function	RMSE
	Number of generation	3000
Genetic operators	Mutation	0.03
	Inversion	0.2
	IS transposition	0.1
	RIS transposition	0.1
	Gene transposition	0.1
	One-point recombination	0.3
	Two-point recombination	0.3
	Gene recombination	0.1

In USBM, k and n are the site constant and can be computed by regression analysis. For this aim, 74 datasets were applied for constructing the models, and then, 18 new datasets were applied for evaluating the performance of the models. Note that, in USBM modeling procedure of this study, the same datasets performed

in the modeling of MLR, GP, and GEP were applied. The constructed USBM model in the present paper is formulated as below:

$$AOp = 218.5 \left(\frac{D}{MC^{1/3}} \right)^{-0.17} \tag{9}$$

AOp prediction using GP

The most important characteristic of GP and GEP algorithms is the ability to developing mathematical function between the input parameters and the corresponding output, while other algorithms such as ANNs fail to do this. Hence, in the current study, GP and GEP methods were used to find functions in the form of $AOp = f(D, MC)$. The parameters of D and MC are independent parameters, while the AOp is the objective parameter. In the present study, GPLAB toolbox was utilized to develop GP model in MATLAB environment. At first, to avoid the overfitting problem, all 92 datasets compiled from Miduk copper mine were normalized between 0 and 1 using following equation:

$$X_{norm} = \frac{(X - X_{min})}{(X_{max} - X_{min})} \tag{10}$$

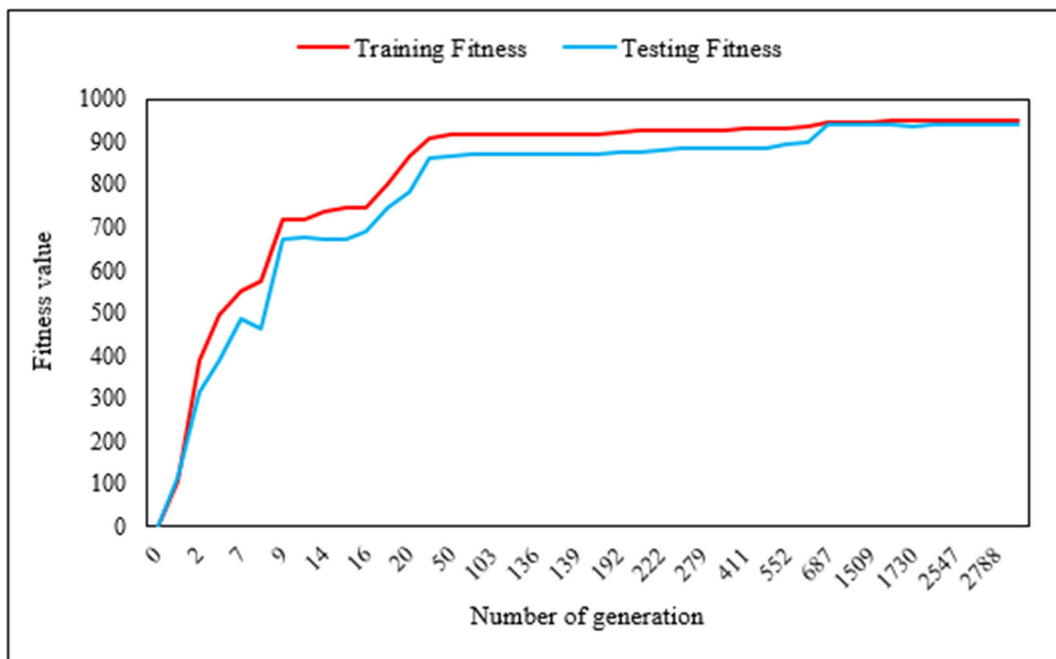


Fig. 7 Progression of fitness value for training and testing datasets

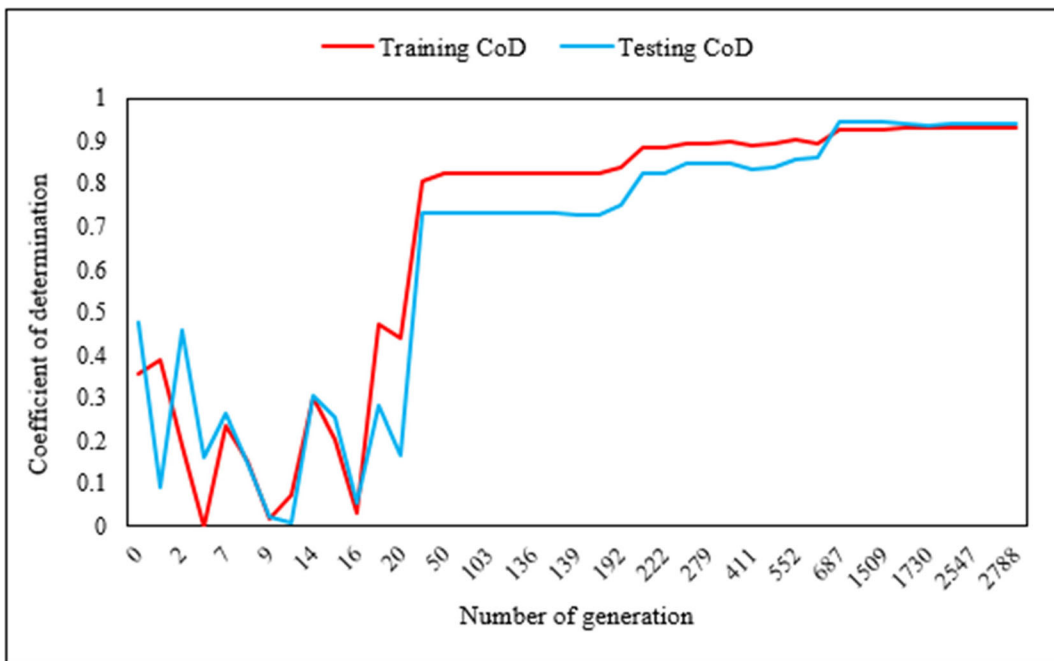


Fig. 8 Variations of CoD during 3000 generations for training and testing datasets

where X is the parameter and X_{min} and X_{max} are the minimum and maximum values for X , respectively. After normalization, the database should be separated to training and testing sections. The main parameters of GP include population size, selection method, the initialization method, the number of generation (iteration), and the values of genetic operators. There is no definite method to determine the appropriate values for these parameters, so in this study, the suggested values by other scholars and trial and error procedure were used to achieve an optimum combination of parameters. In both of GP and GEP modeling, root mean square error (RMSE) was considered as a fitness function for evaluating the fitness of individuals in each iteration.

As mentioned before, the individuals are generated by combining the terminals and functions. The

parameters of D and MC accompanied by the six random constant values were called to the software as the terminal set (T).

$$T = \{D, MC, 9.2744, 1.7261, 2.5175, 4.5290, 0.6039, 4.9850\} \quad (11)$$

Considering the relationship between input parameters in EMs and previous studies regarding AOp prediction, a function set was defined as follows:

$$F = \{+, -, *, /, ^2, ^3, 3Rt, Q, sin, cos, tan, Atan, Exp, Ln\} \quad (12)$$

The population size and the number of generations were set on 100 and 200, respectively. Several GP models with different values of parameters were developed using trial and error procedure to predict AOp. The best model with the high performance was selected. The GP parameters of the selected model are listed in

Table 6 K-Expression of genes related to the best GEP model

Gene no.	K-Expression
1	d1.X3.*.X3.+.Atan.Sqrt.d1.d0.d1.c1.d1.d1.d1.d0.d1.d1.d1.c0
2	Sin.d0.-./d1.X3.c0.X2.d0.d0.d0.c1.c0.c0.d1.d0.d0.c0
3	Cos.+d0.-.-.*.Cos.c0.c0.c0.d1.c0.c0.d1.d0.c0.c0
4	*.Ln.Sin.c0.+X2.-.+X3.d1.d1.d1.c1.d0.c1.d0.d0.c0.d1
5	Sin.+.*.c0.3Rt.d0.Sin.d1.-.d0.d1.d0.d1.d1.c1.d1.c1.d0.c0

d0: D, d1: MC. Italic parts show the non-coding regions

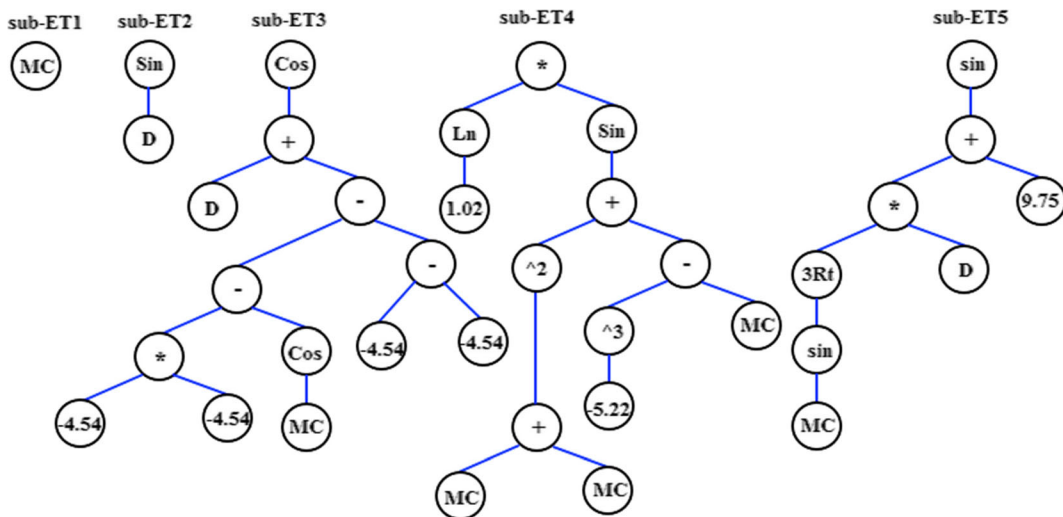


Fig. 9 Sub-ETs of the best chromosome

Table 4. Finally, the developed GP equation for AOp prediction was extracted as follows:

$$AOp = Atan[(1.7261 + MC)^3] - sin(D) - sin(0.6039) \quad (13)$$

The parse tree of the above equation is shown in Fig. 6. More discussion about the appraisal of the GP model is given later.

AOp prediction using GEP

For GEP modeling, the GeneXproTools 4.0 software was used. The same normalized training and testing datasets used in GP section were employed for GEP modeling as well. GEP algorithm in comparison to

Table 7 Performance evaluation of the developed models in this research

Predictive model	Training			Testing		
	CoD	VAF (%)	RMSE	CoD	VAF (%)	RMSE
Eq. 6	0.908	62.58	35.04	0.918	72.66	32.35
Eq. 7	0.905	70.83	11.72	0.913	80.99	9.01
USBM (Eq. 9)	0.901	90.88	1.94	0.920	91.54	2.13
MLR (Eq. 5)	0.902	90.13	1.9	0.903	86.48	2.43
GP	0.915	90.68	0.06	0.928	92.44	0.08
GEP	0.933	93.20	0.05	0.941	94.12	0.06

GP has more parameters that assigning optimum values for all of them is difficult. In this study, according to suggested values by Ferreira (2001) and other scholars (Güllü 2012; Shirani Faradonbeh et al., 2017; Khandelwal et al. 2017) for genetic operators, the values of 0.03, 0.2, 0.1, 0.3, and 0.1 were determined for mutation, inversion, three kinds

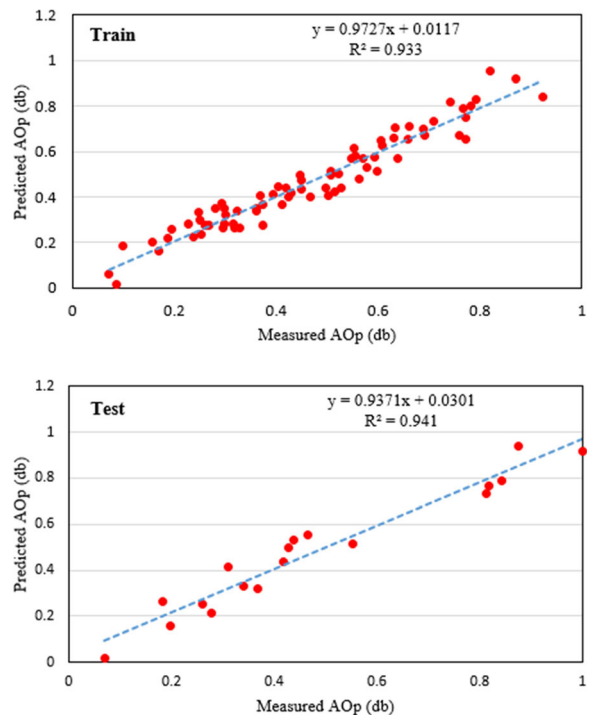


Fig. 10 Comparison of measured and predicted AOp for GEP model

of transposition, recombination (one-point and two-point), and gene recombination, respectively. Similar to GP, the RMSE was used as the fitness function to evaluate the chromosomes. GEP parameters achieved by several trial and error models are listed in Table 5. The best chromosome (solution) is a five genic chromosome which belongs to generation no. 2788 (total number of generations = 3000) with the head and tail sizes of 9 and 10, respectively. In this chromosome, the genes are connected to each other using (+) as the linking function. The variations of fitness function and coefficient of determination (CoDs) during the 3000 generations are shown in Figs. 7 and 8, respectively. According to Fig. 7, the algorithm has been stopped at generation no. 2788 after the fitness values remained constant on 949.942 and 939.362 for training and testing datasets, respectively.

As mentioned above, the best chromosome consists of five genes and each gene can be presented as K-Expression (Table 6). According to the defined rules by Ferreira (2006), these K-Expressions can be expressed in the form of sub-ETs (see Fig. 9). Eventually, the mathematical formation of these sub-ETs was extracted as Eqs. 14–18. Equation 19 shows the general AOp predictor developed by GEP algorithm.

$$\text{sub-ET 1 : } MC \tag{14}$$

$$\text{sub-ET 2 : } \text{Sin}(D) \tag{15}$$

$$\text{sub-ET 3 : } \text{Cos}\left(D + (-4.544586)^2 - \text{Cos}(MC)\right) \tag{16}$$

$$\begin{aligned} \text{sub-ET 4 : } & Ln(1.023254) \\ & \times \text{Sin}\left(4MC^2 + (-5.22644)^3 - MC\right) \end{aligned} \tag{17}$$

$$\text{sub-ET 5 : } \text{Sin}\left(\left(\sqrt[3]{\text{sin}(MC)}\right) \times D\right) + 9.757782 \tag{18}$$

$$\begin{aligned} \text{AOP} = & (\text{sub-ET1}) + (\text{sub-ET2}) + (\text{sub-ET3}) \\ & + (\text{sub-ET4}) + (\text{sub-ET5}) \end{aligned} \tag{19}$$

Results and discussion

To evaluate the effectiveness and accuracy of the proposed models in the predicting AOp, three statistical indices, i.e., variance account for (VAF), CoD, and RMSE, were computed. They are defined as follows:

$$\text{CoD} = \frac{\left[\sum_{i=1}^n (x_{im} - x_{mean})^2\right] - \left[\sum_{i=1}^n (x_{im} - x_{ip})^2\right]}{\left[\sum_{i=1}^n (x_{im} - x_{mean})^2\right]} \tag{20}$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{im} - x_{ip})^2} \tag{21}$$

$$\text{VAF} = \left[1 - \frac{\text{var}(x_{im} - x_{ip})}{\text{var}(x_{im})}\right] \times 100 \tag{22}$$

in which n is the number of datasets, var is sign of the variance, x_{mean} is mean of the measured value, and x_m and x_p are the measured and predicted AOp values, respectively. The CoD, RMSE, and VAF equal to 1, 0, and 100%, respectively, indicate the best approximation. The computed statistical indices for the training and testing datasets are shown in Table 7. From Table 7, it is found that the accuracy level of the GEP is higher than GP, USBM, and MLR, Eqs. 6 and 7, respectively. In the other words, the highest values of VAF and CoD as well as the lowest value of RMSE were obtained by implementing the GEP. As an example, Fig. 10 shows the training and testing results (CoDs) for the GEP model. According to Fig. 10 and Table 7, it can be seen that the GEP results have lower errors than those obtained from other predictive models.

Conclusion

In the present study, the several models, i.e., GP, GEP, MLR, McKenzine, USBM, and National Association of Australian State models were employed for predicting AOp in the Miduk copper mine. For this purpose, 92 blasting events were considered and the values of AOp as well as two effective parameters on the AOp, i.e., MC and D, were measured for each event. In the analysis of this study, 74 and 18 datasets were applied to construct the models and to verify the constructed models, respectively. Then, to select the best predictive model, three statistical criteria, i.e., CoD, VAF, and RMSE, were used

and the obtained results were compared. As a result, the highest CoD (0.941) and VAF (94.12%) and the lowest RMSE (0.06) were obtained by GEP. According to the obtained results, the GEP model is more accurate than the GP, USBM, MLR, Eqs. 6 and 7, respectively. It is important to note that developed GP and GEP models in this study are applicable in estimating air overpressure values; nevertheless, more accurate results can be obtained when GEP model is utilized, and this model can be introduced as a new model in field of air overpressure prediction. It should be also noted that the presented regression and USBM models in this study are specific to Miduk copper mine and cannot be directly used in other sites.

References

- Ahmad, M., Ansari, M. K., Sharma, L. K., Singh, R., & Singh, T. N. (2017). Correlation between strength and durability indices of rocks-soft computing approach. *Process Engineering*, *191*, 458–466.
- Back, T., David, B., & Zbigniew, M. (1997). *Handbook of evolutionary computation*. Oxford University Press: IOP Publishing Ltd.
- Bhandari, S. (1997). *Engineering rock blasting operations*. Rotterdam: A.A. Balkema.
- Cengiz, K. (2008). The importance of site-specific characters in prediction models for blast-induced ground vibrations. *Soil Dynamics and Earthquake Engineering*, *28*, 405–414.
- Cramer, N. L. (1985). A representation for the adaptive generation of simple sequential programs. In *Proceedings of the international conference on genetic algorithms and their applications* (pp. 183–187). Pittsburgh: Carnegie-Mellon University.
- Ferreira, C. (2001). Gene expression programming: a new adaptive algorithm for solving problems. *Complex Syst*, *13*(2), 87–129.
- Ferreira, C. (2006). *Gene expression programming: mathematical modeling by an artificial intelligence* (2nd ed.). Germany: Springer-Verlag.
- Ghotbi Ravandi, E., Rahmamejad, R., Feili Monfared, A. E., & Ghotbi Ravandi, E. (2013). Application of numerical modeling and genetic programming to estimate rock mass modulus of deformation. *International Journal of Mining Science and Technology*, *23*, 733–737.
- Gong, M., Wang, S., Liu, W., Yan, J., & Jiao, L. (2016). Evolutionary computation in China: a literature survey. *CAAI Transactions on Intelligence Technology*, *1*(4), 334–354.
- Griffiths, M. J., Oates, J. A. H., & Lord, P. (1978). The propagation of sound from quarry blasting. *Journal of Sound and Vibration*, *291*, 358–370.
- Güllü, H. (2012). Prediction of peak ground acceleration by genetic expression programming and regression: a comparison using likelihood-based measure. *Engineering Geology*, *141*, 92–113.
- Hajihassani, M., Jahed Armaghani, D., Monjezi, M., Tonnizam, M. E., & Marto, A. (2015). Blast-induced air and ground vibration prediction: a particle swarm optimization-based artificial neural network approach. *Environment and Earth Science*, *74*, 2799–2817.
- Hasanipanah, M., Jahed Armaghani, D., Khamesi, H., Bakhshandeh Amnieh, H., & Ghoraba, S. (2015b). Several non-linear models in estimating air-overpressure resulting from mine blasting. *Engineering Computations*, *32*, 441–455. <https://doi.org/10.1007/s00366-015-0425-y>.
- Hasanipanah, M., Monjezi, M., Shahnazar, A., Jahed Armaghani, D., & Farazmand, A. (2015a). Feasibility of indirect determination of blast induced ground vibration based on support vector machine. *Measurement*, *75*, 289–297.
- Hemphill, G. B. (1981). *Blasting operations*. New York: McGraw-Hill.
- Hoseinian, F. S., Faradonbeh, R. S., Abdollahzadeh, A., Rezai, B., & Soltani-Mohammadi, S. (2017). Semi-autogenous mill power model development using gene expression programming. *Powder Technology*, *308*, 61–69.
- Jahed Armaghani, D., Hajihassani, M., Marto, A., Faradonbeh, R. S., & Tonnizam Mohamad, E. (2015). Prediction of blast-induced air overpressure: a hybrid AI-based predictive model. *Environmental Monitoring and Assessment*, *187*, 1–13.
- Jahed Armaghani, D., Mahdiyari, A., Hasanipanah, M., Faradonbeh, R. S., Khandelwal, M., & Amnieh, H. B. (2016a). Risk assessment and prediction of flyrock distance by combined multiple regression analysis and Monte Carlo simulation of quarry blasting. *Rock Mechanics and Rock Engineering*, *49*(9), 3631–3641.
- Jahed Armaghani, D., Hasanipanah, M., & Tonnizam Mohamad, E. (2016b). A combination of the ICA-ANN model to predict air overpressure resulting from blasting. *Engineering Computations*, *32*, 155–171.
- Khandelwal, M., & Kankar, P. K. (2011). Prediction of blast-induced air overpressure using support vector machine. *Arabian Journal of Geosciences*, *4*, 427–433.
- Jahed Armaghani, D., Shirani Faradonbeh, R., Momeni, E., Fahimifard, A., & Tahir, M. M. (2018). Performance prediction of tunnel boring machine through developing a gene expression programming equation. *Engineering Computers*, *34*(1), 129–141.
- Khandelwal, M., & Singh, T. N. (2005). Prediction of blast induced air overpressure in opencast mine. *Noise Vib Control Worldw*, *36*, 7–16.
- Khandelwal, M., Armaghani, D. J., Faradonbeh, R. S., Ranjith, P. G., & Ghoraba, S. (2016). A new model based on gene expression programming to estimate air flow in a single rock joint. *Environment and Earth Science*, *75*(9), 739.
- Khandelwal, M., Shirani Faradonbeh, R., Monjezi, M., Jahed Armaghani, D., Majid, M. Z. B. A., & Yagiz, S. (2017). Function development for appraising brittleness of intact rocks using genetic programming and non-linear multiple regression models. *Engineering with Computers*, *33*(1), 13–21.
- Koza, J. R. (1996). *On the programming of computers by means of natural selection*. USA: MIT Press.
- Kuzu, C., Fisne, A., & Ercelebi, S. G. (2009). Operational and geological parameters in the assessing blast induced airblast-overpressure in quarries. *Applied Acoustics*, *70*, 404–411.

- McKenzie, C. (1990). Quarry blast monitoring: technical and environmental perspectives. *Quarry Management*, 17, 23–34.
- Mohammadnejad, M., Gholami, R., Sereshki, F., & Jamshidi, A. (2013). A new methodology to predict backbreak in blasting operation. *International Journal of Rock Mechanics and Mining Sciences*, 60, 75–81.
- Monjezi, M., Bahrami, A., & Yazdian Varjani, A. (2010). Simultaneous prediction of fragmentation and flyrock in blasting operation using artificial neural networks. *International Journal of Rock Mechanics and Mining Sciences*, 47, 476–480.
- Naderi, A., Delavar, M. A., Kaboudin, B., & Askari, M. S. (2017). Assessment of spatial distribution of soil heavy metals using ANN-GA, MSLR and satellite imagery. *Environmental Monitoring and Assessment*, 189, 214.
- National Association of Australian State Road Authorities. (1983). *Explosives in roadworks—a users guide*. Sydney: NAASRA.
- Nelson, M., & Illingworth, W. T. (1990). *A practical guide to neural nets*. Reading: Addison-Wesley.
- Ravansalar, M., & Rajaei, T. (2015). Evaluation of wavelet performance via an ANN-based electrical conductivity prediction model. *Environmental Monitoring and Assessment*, 187, 366.
- Deo, R. C., & Şahin, M. (2016). An extreme learning machine model for the simulation of monthly mean streamflow water level in eastern Queensland. *Environmental Monitoring and Assessment*, 188, 90.
- Rezaei, M., Monjezi, M., & Yazdian Varjani, A. (2011). Development of a fuzzy model to predict flyrock in surface mining. *Safety Science*, 49, 298–305.
- Rodríguez, R., Lombardía, C., & 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.
- Rodríguez, R., Toraño, J., & Menéndez, M. (2007). Prediction of the airblast wave effects near a tunnel advanced by drilling and blasting. *Tunnelling and Underground Space Technology*, 22, 241–251.
- Rosenthal, M.F. & Morlock, G.L. (1987). *Blasting guidance manual, office of surface mining reclamation and enforcement*. US Department of the Interior.
- Sankararajan, V., Neelakandhan, N., & Chandrasekaran, S. (2017). Modeling of constructed wetland performance in BOD5 removal for domestic wastewater under changes in relative humidity using genetic programming. *Environmental Monitoring and Assessment*, 189, 164.
- Sari, M., Ghasemi, E., & Ataei, M. (2014). Stochastic modeling approach for the evaluation of backbreak due to blasting operations in open pit mines. *Rock Mechanics and Rock Engineering*, 47(2), 771–783.
- Segarra, P., Domingo, J. F., López, L. M., Sanchidrián, J. A., & Ortega, M. F. (2010). Prediction of near field overpressure from quarry blasting. *Applied Acoustics*, 71, 1169–1176.
- Sharma, L. K., Singh, R., Umrao, R. K., Sharma, K. M., & Singh, T. N. (2017a). Evaluating the modulus of elasticity of soil using soft computing system. *Engineering Computations*, 33(3), 497–507.
- Sharma, L. K., & Singh, T. N. (2017). Regression based models for the prediction of unconfined compressive strength of artificially structured soil. *Engineering Computations*, 34, 175–186. <https://doi.org/10.1007/s00366-017-0528-8>.
- Sharma, L. K., Vishal, V., & Singh, T. N. (2017b). Developing novel models using neural networks and fuzzy systems for the prediction of strength of rocks from key geomechanical properties. *Measurement*, 102, 158–169.
- Sharma, L. K., Vishal, V., & Singh, T. N. (2017c). Predicting CO2 permeability of bituminous coal using statistical and adaptive neurofuzzy analysis. *Journal of Natural Gas Science and Engineering*, 42, 216–225. <https://doi.org/10.1016/j.jngse.2017.02.037>.
- Singh, R., Umrao, R. K., Ahmad, M., Ansari, M. K., Sharma, L. K., & Singh, T. N. (2017). Prediction of geomechanical parameters using soft computing and multiple regression approach. *Measurement*, 99, 108–119.
- Shirani Faradonbeh, R., & Monjezi, M. (2017). Prediction and minimization of blast-induced ground vibration using two robust meta-heuristic algorithms. *Engineering Computations*, 33, 835–851. <https://doi.org/10.1007/s00366-017-0501-6>.
- Shirani Faradonbeh, R., Salimi, A., Monjezi, M., Ebrahimabadi, A., & Moormann, C. (2017). Roadheader performance prediction using genetic programming (GP) and gene expression programming (GEP) techniques. *Environment and Earth Science*, 76(16), 584.
- Shirani Faradonbeh, R., Jahed Armaghani, D., & Monjezi, M. (2016a). Development of a new model for predicting flyrock distance in quarry blasting: a genetic programming technique. *Bulletin of Engineering Geology and the Environment*, 75, 993–1006.
- Shirani Faradonbeh, R., Jahed Armaghani, D., Monjezi, M., & Mohamad, E. T. (2016b). Genetic programming and gene expression programming for flyrock assessment due to mine blasting. *International Journal of Rock Mechanics and Mining Sciences*, 88, 254–264.
- Shirani Faradonbeh, R., Monjezi, M., & Jahed Armaghani, D. (2015). Genetic programming and non-linear multiple regression techniques to predict backbreak in blasting operation. *Engineering Computations*, 32, 123–133. <https://doi.org/10.1007/s00366-015-0404-3>.
- Shirani Faradonbeh, R., Jahed Armaghani, D., Bakhshandeh Amnieh, H., & Mohamad, E. T. (2018). Prediction and minimization of blast-induced flyrock using gene expression programming and firefly algorithm. *Neural Computing and Applications*, 29(6), 269–281.
- Singh, J., Verma, A. K., Banka, H., Singh, T. N., & Maheshwar, S. (2016). A study of soft computing models for prediction of longitudinal wave velocity. *Arabian Journal of Geosciences*, 9, 224.
- Singh, N., & Chakrapani, G. J. (2015). ANN modelling of sediment concentration in the dynamic glacial environment of Gangotri in Himalaya. *Environmental Monitoring and Assessment*, 187, 494.
- Siskind, D.E., Stachura, V.J., Stagg, M.S. & Koop, J.W. (1980). In: Siskind DE, editor. *Structure response and damage produced*

- by *airblast from surface mining*. United States Bureau of Mines.
- Swingler, K. (1996). *Applying neural networks: a practical guide*. New York: Academic Press.
- Tonnizam Mohamad, E., Hajihassani, M., Jahed Armaghani, D., & Marto, A. (2012). Simulation of blasting-induced air overpressure by means of artificial neural networks. *International Review on Modelling and Simulations*, 5, 2501–2506.
- Verma, A. K., Singh, T. N., Kumar, N., & Sarkar, C. (2016). A hybrid FEM–ANN approach for slope instability prediction. *Journal of The Institution of Engineers (India): Series A*, 97, 171–180.
- Wiss, J. F., & Linehan, P. W. (1978). *Control of vibration and blast noise from surface coal mining*. Northbrook: Wiss, Janney, Elstner and Associates Inc..
- Qazi, K. I., Lam, H. K., Xiao, B., Ouyang, G., & Yin, X. (2016). Classification of epilepsy using computational intelligence techniques. *CAAI Transactions on Intelligence Technology*, 1(2), 137–149.
- Xu, X., Law, R., Chen, W., & Tang, L. (2016). Forecasting tourism demand by extracting fuzzy Takagi–Sugeno rules from trained SVMs. *CAAI Transactions on Intelligence Technology*, 1(1), 30–42.
- Zou, J., Bui, K. T. T., Xiao, Y., & Doan, C. V. (2017). Dam deformation analysis based on BPNN merging models. *Geo-spatial Information Science*, 1–9.