https://doi.org/10.1007/s00521-021-06776-z

Neural Computing and Applications

A hybrid metaheuristic approach using random forest and particle swarm optimization to study and evaluate backbreak in open-pit blasting

--Manuscript Draft--

Manuscript Number:	NCAA-D-21-03636R1
Full Title:	A hybrid metaheuristic approach using random forest and particle swarm optimization to study and evaluate backbreak in open-pit blasting
Article Type:	Original Article
Keywords:	Backbreak; Blasting; Random Forest; PSO algorithm.
Corresponding Author:	Manoj Khandelwal, Ph. D. Federation University Australia Ballarat, Victoria AUSTRALIA
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Federation University Australia
Corresponding Author's Secondary Institution:	
First Author:	Yong Dai
First Author Secondary Information:	
Order of Authors:	Yong Dai
	Manoj Khandelwal, Ph. D.
	Yingui Qiu
	Jian Zhou
	M. Monjezi
	Peixi Yang
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	Backbreak is a rock fracture problem that exceeds the limits of the last row of holes in an explosion operation. Excessive backbreak increases operational cost and also poses a threat to mine safety. In this regard, a new hybrid intelligence approach based on random forest (RF) and particle swarm optimization (PSO) are proposed for predicting backbreak with high accuracy to reduce the unsolicited phenomenon induced by backbreak in open-pit blasting. A data set of 234 samples with 6 input parameters including special drilling (SD), spacing (S), burden (B), hole length (L), stemming (T) and powder factor (PF)) and one output parameter backbreak (BB) is set up in this study. Seven input combinations (one with six parameters, six with five parameters) are built to generate the optimal prediction model. The PSO algorithm was integrated with the RF algorithm to find the optimal hyper-parameters ([[EQUATION]] and [[EQUATION]]) of each model and the fitness function, which is the MAE of 10 cross-validations. The performance capacities of the optimal models are assessed using mean absolute error, root mean square error, Pearson correlation coefficient and mean absolute percentage error. Findings demonstrated that the PSO-RF model combining L-S-B-T-PF with MAE of (0.0132, and 0.0568), RMSE of (0.0811, 0.1686), R 2 of (0.9990, 0.9961), and MAPE of (0.0027, and 0.0116) in training and testing phases, respectively, has optimal prediction performance. The optimal PSO-RF models are compared with the classical Artificial Neural Network (ANN), RF, Genetic Programming (GP), and Support Vector Machine (SVM) models that show that the PSO-RF model has superiority in predicting backbreak. The Gini index of each input variable has also been calculated in the RF model, which were 31.2(L), 23.1(S), 27.4(B), 36.6(T), 23.4(PF), and 16.9(SD), respectively.

Response to Reviewers:	Response Letter Dear Editor: RE: NCAA-D-21-03636 Thank you for your letter and for the reviewer's comments concerning our manuscript entitled "A hybrid metaheuristic approach using random forest and particle swarm optimization to study and evaluate backbreak in open-pit blasting" (ID: NCAA-D-21-03636). We would like to thank the reviewers for thoroughly reviewing our manuscript, and for making many thoughtful comments. The changes made to the paper can be identified by the text in bright color. Thank you for your time and kind consideration. -Response to Reviewer #1 -Response to Reviewer #3 -Response to Reviewer #5 Best regards, Jian Zhou and Manoj Khandelwal Corresponding author Response to Reviewer #1 Dear Prof. / Dr. We appreciate your precise comments. Please consider our explanations and clarifications. Comment 1: Some important results in this manuscript should be added in the Abstract. Reply: I appreciate your suggestion. Some important results are added in the Abstract. Comment 2: Please introduce the equation (7) in detail. Reply: I am glad that you brought up the question. The equation (7) is introduced in more detail. Comment 3: There are some mistakes in the references. Such as Line 514, Page 32: Li, D., Line 581, Page 35: Yu, Z., Line 593 Page 593: Zhao, C., Make sure that references are listed according the Journal standards. Reply: Thank you, all references have been properly corrected and available full DOI links are listed in the reference part. Response to Reviewer #3 Dear Prof. / Dr. We appreciate your precise comments. Please consider our explanations and clarifications. Comment 1: The author summarized the related literatures on "backbreak" prediction. It is best to divide them according to the data sets, and then sort them by year. Reply: Thanks for your question. Related literature has been divided according to the data sets, and then sort them by year. Comment 2: The introduction of the data set is not clear enough. Reply: I am glad for you brought up the question. The data set has been introduced clearer. Comment 3: The data preprocessing part does not specify whether data cleaning is required, such as deleting redundant data or missing value data, and whether it is necessary to normalize the values. Reply: Thank you. The data preprocessing part has been supplemented. Comment 4: The author didn't explain the advantages of PSO in hyperparameter selection compared with other methods, such as NSGAIII and other swarm intelligence algorithms.

	Reply: Thank you for this suggestion. The advantages of PSO in hyperparameter selection has been explained. Comment 5: The author didn't explained what advantages the random forest algorithm has compared with other classic regression algorithms. Reply: Thank you. The advantages of the random forest algorithm have been properly added. Response to Reviewer #5 Dear Prof. / Dr. We appreciate your precise comments. Please consider our explanations and clarifications. Authers proposed a hybrid intelligence approach based on random forest (RF) and particle swarm optimization (PSO) are proposed for predicting backbreak with high accuracy to reduce the unsolicited phenomenon. The author uses traditional machine learning to predict. At present, there are many traditional machine learning methods in this field. Such as https://www.researchgate.net/publication/257445416_Prediction_of_Backbreak_in_Open-Pit_Blasting_Operations_Using_the_Machine_Learning_Method. Reply: Thank you very much for reviewing our paper. Comment 1: Can the author provide more comparison with deep learning methods? Reply: I am glad for you brought up the advice. The author has proposed a convolutional neural network to predict the backbreak of the same data set. Comment 2: Can the author explain why the design method of deep learning is not adopted? Reply: Thank you. The result shows that the CNN model is not as outstanding as the traditional machine learning models proposed in the manuscript. The author thinks that the advantages of deep learning will be revealed when the sample size is large enough, so deep learning is not adopted in this paper. When the sample size is large enough, the author will give priority to the deep learning model.

A hybrid metaheuristic approach using random forest and particle swarm optimization to study and evaluate backbreak in open-pit blasting

Yong Dai[1], Manoj Khandelwal[2]*, Yingui Qiu[1], Jian Zhou[1]*, M. Monjezi[3], Peixi Yang[1]

[1] School of Resources and Safety Engineering, Central South University, Changsha 410083, China.

[2] School of Engineering, Information Technology and Physical Sciences, Federation University Australia, Ballarat, VIC 3350, Australia.

[3] Faculty of Engineering, Tarbiat Modares University, Tehran, Iran.

* To whom correspondence should be addressed; Email: m.khandelwal@federation.edu.au, j.zhou@csu.edu.cn

Abstract

Backbreak is a rock fracture problem that exceeds the limits of the last row of holes in an explosion operation. Excessive backbreak increases operational costs and also poses a threat to mine safety. In this regard, a new hybrid intelligence approach based on random forest (RF) and particle swarm optimization (PSO) is proposed for predicting backbreak with high accuracy to reduce the unsolicited phenomenon induced by backbreak in open-pit blasting. A data set of 234 samples with 6 input parameters including special drilling (SD), spacing (S), burden (B), hole length (L), stemming (T) and powder factor (PF)) and one output parameter backbreak (BB) is set up in this study. Seven input combinations (one with six parameters, six with five parameters) are built to generate the optimal prediction model. The PSO algorithm is integrated with the RF algorithm to find the optimal hyperparameters of each model and the fitness function, which is the MAE of 10 cross-validations. The performance capacities of the optimal models are assessed using mean absolute error, root mean square error, Pearson correlation coefficient and mean absolute percentage error. Findings demonstrated that the PSO-RF model combining L-S-B-T-PF with MAE of (0.0132, and 0.0568), RMSE of (0.0811, 0.1686), R^2 of (0.9990, 0.9961), and MAPE of (0.0027, and 0.0116) in training and testing phases, respectively, has optimal prediction performance. The optimal PSO-RF models were compared with the classical Artificial Neural Network (ANN), RF, Genetic Programming (GP), Support Vector Machine (SVM) and Convolutional Neural Network (CNN) models and show that the PSO-RF model has superiority in predicting backbreak. The Gini index of

each input variable has also been calculated in the RF model, which was 31.2(L), 23.1(S), 27.4(B), 36.6(T), 23.4(PF), and 16.9(SD), respectively.

Keywords: Backbreak; Blasting; Random forest; PSO algorithm; Predictive model

Introduction

Explosives are widely used to break hard rock mass in open-pit mining due to their low costs. However, the explosive energy is poorly utilized, with nearly 70 to 80 percent of the explosive energy dissipating in the ground, which may cause several detrimental influences (Berta 1990, Zhou et al. 2021c), i.e., blasting fume, ground vibration, noise, backbreak, flyrock. Particularly, backbreak, as part of side effects of the explosion, is a rock fracture phenomenon that exceeds the limits of the last row of holes in an explosion operation (Jimeno et al. 1995), which has various undesirable impacts, such as an increase in the stripping ratio, falling down the mining machinery, instability in mine walls, reduction in efficiency of drilling and lower in the overall slope angle (Gates et al. 2005; Khandelwal and Singh 2013; Sari et al. 2014; Zhou et al. 2021a). Therefore, accurate estimation of backbreak before a blasting operation is of great significance to minimize the harmful impact of backbreak.

Fig. 1 represents an open-pit bench terminology. It can be seen from Fig. 1 that explosive properties, blast design parameters and rock mass properties have certain effects on backbreak. Controllable factors, namely explosive properties and blast design parameters and uncontrollable factors, namely rock mass properties have been selected by numerous researchers to predict backbreak (Monjezi et al. 2012, 2013;

Esmaeili et al. 2014; Zhou et al. 2021a). Lundborg (1974) and Roth (1979) have attempted to predict backbreak with some empirical models. Nevertheless, these empirical models are capable of predicting backbreak under certain geo-mining conditions and are based on only a few influencing factors. Models featuring wider adaptability between backbreak, and the influencing parameters are needed to minimize production cost vis-à-vis to enhance the safety and stability of an open-pit mine.

Currently, various artificial intelligence (AI) techniques including fuzzy set theory (Wang et al. 2019; Huang and Xiao 2021), ANN (Wang et al. 2015; Ferentinou and Fakir 2018; Biourge et al. 2020), SVM (Goh and Goh 2007; Zhao et al. 2017; Li et al. 2020; Zhou et al, 2021b), GP (Beiki et al. 2010; Liu et al. 2021), CNN (He et al. 2021) and Neuro-genetic approach (Alemdag et al. 2016) utilize capturing non-linear relationships between multi-dimensional variables which have made a great success in plenty of geotechnical engineering applications, and have been showing good performance in the field of predicting rockbursts (Zhou et al. 2012, 2016; Yin et al. 2021), blast vibrations (Iphar et al. 2008; Li et al. 2012; Armaghani et al. 2014; Hosseini et al. 2019; Yu et al. 2020), ground settlement (Gong et al. 2014; Zhou et al, 2017; Moeinossadat et al. 2018; Zhang et al. 2020a). Regarding backbreak prediction, Table 1 summarizes some published literature on backbreak prediction. It has been found that ANN is mostly used to predict backbreak, however other models, such as SVM, GP, ANFIS, etc. have also been used. As a branch of ensemble learning, a random forest algorithm shows good prediction performance in a large number of databases, less over-fitting phenomenon, fast training speed, and the importance of each feature can also be evaluated internally (Ray et al. 2020; Zhang et al. 2020a). In addition, as a widely used swarm intelligence algorithm, PSO has been proved to have the advantages of fewer parameters to be adjusted, easy implementation, the use of individual local information and global information of the group to guide the search, and better hyper-parameter selection ability compared with other algorithms (Zhang et al, 2021; Zhou et al, 2021b; Nabiollahi et al. 2021). Based on this, a hybrid artificial model combining PSO and RF, namely PSO-RF is presented in this study to predict backbreak. To compare the proposed PSO-RF model prediction capability, various other AI algorithms, such as ANN, RF, GP, SVM and CNN, which are popular or potential in predicting backbreak, are also adopted in this study.

In this article, first, the background of the proposed methodologies has been presented. Then, the framework of the proposed model is illustrated, and the establishment of datasets are introduced. After searching the optimum hyper-parameters of different combinations of PSO-RF models, the optimal combination of PSO-RF will be determined. Moreover, the optimal PSO-RF models will be compared with classical models. Finally, the Gini index will be calculated internally in the RF model to investigate the most important input variables in estimating backbreak.

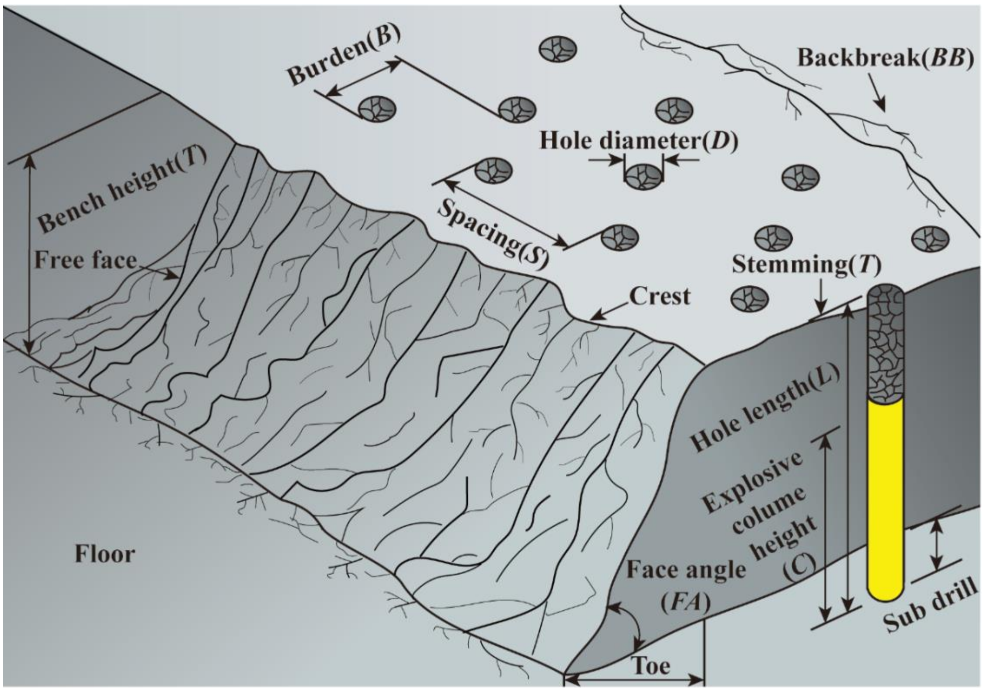

Fig. 1. The appearance of open-pit bench.

Table 1 Current literature on backbreak prediction applying AI techniques.

Method	Input parameters	No. of dataset	R^2	RMSE	References
FIS	B, L, T, S, SD, C, PF, D	--	0.95	0.44	Monjezi et al. (2010)
GA -ANN	RMR, L, HD, B, C, SD, S, PF	195	0.95	--	Monjezi et al. (2012)
SVM	HD, B, PF, S, SDT,	193	0.92	0.34	Mohammadnejad et al. (2013)
ANN	HD, UCS, S, B, T, C, PF, W, K, SD	97	0.90	--	Monjezi et al. (2013)
BP, RBF	SC, B, HL, T, SD	103	BP=0.87 RBF=0.52	BP=0.22 RBF=0.31	Sayadi et al. (2013)
ANN, ANFIS	SC, N, T, CLR	42	ANN=0.92 ANFIS=0.96	ANN=0.60 ANFIS=0.88	Esmaeili et al. (2014)
ANN	DB, B, PF, T, S, SD, R, N	--	0.87	0.49	Monjezi et al. (2014)
ABC-ANN	B, L, S, T, PF	34	0.77	0.53	Ebrahimi et al. (2016)
GP	B, PF, T, S, SR	175	0.98	0.327	Faradonbeh et al. (2016)
RT, ANFIS	B, P, S, T, K	175	RT=0.97 ANFIS=0.99	RT=0.35 ANFIS=0.08	Ghasemi et al. (2016)

PSO	B, T, PF, S, K	175	0.98	0.27	Ghasemi (2017)
PSO-ANFIS	B, S, PF, T	80	0.98	0.13	Hasanipanah et al. (2017)
PSO	S, SC, B, T, RMR	84	0.96	0.08	Eskandar et al. (2018)
GA, ICA	PF, B, MC, B/D, T/B, S/B	62	GA=0.96 ICA=0.93	GA=0.058 ICA=0.041	Hasanipanah and Bakhshandeh Amnieh (2021)
RF	H/T, ED, S/B, P-wave	40	0.98	0.87	Kumar et al. (2021)
RF	PF, B, T, S	26	0.87	0.59	Sharma et al. (2021)
SVM	SD, L, B, T, PF, S	234	0.98	--	Khandelwal and Monjezi (2013)
SCA-RF, HHO-RF	PF, L, B, T, S, SD	234	SCA-RF=0.98 HHO-RF=0.98	SCA-RF =0.09 HHO-RF=0.11	Zhou et al. (2021a)

Nomenclature: burden (B); spacing (S); hole length (L); burden (B); special drilling (SD); stemming (ST); charge per delay (C); rock density (D); powder factor (PF); uniaxial compressive strength (USC); water content (W); bench height (K); specific charge (SC); number of rows (N); charge last row (CLR); rock factor (R); delay per burden (DB); hole diameter (HD);stiffness ratio (SR); maximum charge per delay (MC); explosive density(ED); fuzzy inference system (FIS); genetic algorithm (GA); regression tree (RT); imperialist competitive algorithm (ICA); adoptive neuro-fuzzy inference system (ANFIS).

Methodology

Random forest algorithm

Random forest (Breiman 2001) is a supervised algorithm composed of an independent decision tree (DT) and bagging framework. Here, the Bootstrap sampling method is utilized to stochastically extract a certain amount of data from the training set N to form a bootstrap training set N_t. Accordingly, DTs for each bootstrap training set N_t are built. Based on bootstrap training set**s**, out-of-bag (OOB) predictors (about a third of N) is built, which contain non exist samples in the N_t. In the OOB error estimation process, OOB predictors play a test set role, so there is no need to create another test set. The essence of random forest is the integration of DTs (Zhou et al. 2019), which forms multiple DTs through the randomization of column variables and row values of the dataset, and eventually averages the results of the DTs as per Eq. (1). Random feature selection is carried out after random data selection. The double randomness reduces the correlation between DTs, decreases the phenomenon of over-fitting, and has good anti-noise ability. When constructing the DT, the procedure of pruning is not implemented to avoid inhibiting the growth of the tree. Each tree is composed of randomly selecting column variables and row observations. Single DT is difficult to predict correctly, but all DTs form a forest, making the aggregated results integrate the results of all DTs, so the overall prediction is more accurate (Zhou et al. 2020). Alongside, the mean decrease in the Gini index was also calculated showing the variable importance within the model as per Eq. (2).

$$y(x) = \frac{1}{B}\sum_{b=1}^{B} T(x, E_b) \tag{1}$$

Where, $y(x)$ is the result of a combined prediction model, B represents the overall number of DTs, $T(x, E_b)$ are the results of all DTs generated from bootstrapped training samples.

$$Gini(X_i) = \sum_{j-1}^{J} P(X_i = Y_j)(1 - P(X_i = Y_j)) = 1 - \sum_{j}^{J} P(X_i = Y_j)^2 \tag{2}$$

Where $Gini(X_i)$ is the Gini index, $P(X_i = Y_j)$ is estimated values and $X_i = Y_j$ is probabilities. The flowchart of building the random forest is shown in Fig.2.

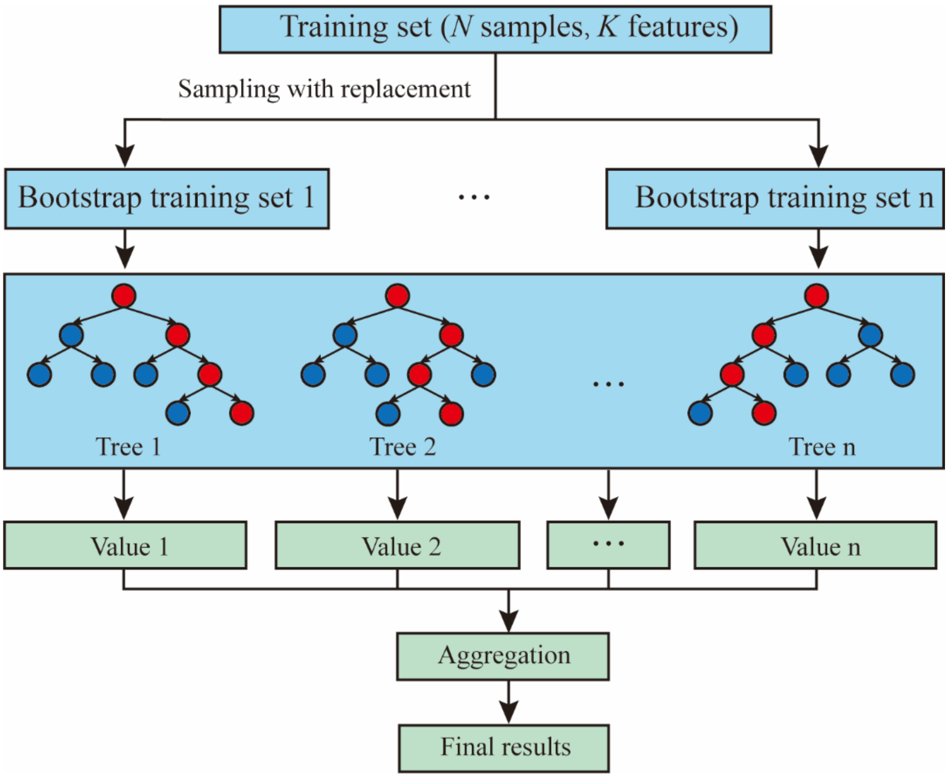

Fig. 2. The flowchart of building the random forest.

Particle swarm optimization

The PSO algorithm was presented for solving unconstrained optimization problems (Eberhart and Kennedy 1995), which simulates the behavior of bird swarms or fish swarms. Depending on its simplicity and remarkable search efficiency, the success of PSO has been verified in many fields, such as function optimization, vehicle routing optimization, geodesy, image processing (Seo et al. 2006; Civicioglu 2012; Bhandari et al. 2015; Jamasb et al. 2017; Wu et al. 2017; Mirghasemi et al. 2019). The architecture of the PSO algorithm is presented in Fig. 3. The algorithm starts by randomly locating N particles in the search space. Each swarm particle has its unique position vector $x_{id}(t)$ and velocity vector, $v_{id}(t)$. The position and velocity of the *i*th particle in each iteration are updated as follows:

$$c_1 r_1[p_{id}(t) - x_{id}(t)] + c_2 r_2[p_{gd}(t) - x_{id}(t)] + v_{id}(t) = v_{id}(t+1) \quad (3)$$

$$v_{id}(t+1)+x_{id}(t)=x_{id}(t+1) \quad (4)$$

Where, d is the dth dimension, $x_{id}(t)$ and $v_{id}(t)$ are the position and velocity of the ith particle at the tth iteration; $p_{id}(t)$ = historical best position found by the ith particle; and $p_{gd}(t)$ = historical global optimal position found by all particles; c_1 and c_2 are constant called acceleration coefficients, both c_1 and c_2 are equal to 1.49. r_1 and r_2 are two generated random numbers in the interval [0, 1]. The size of the population is set as 20, which is sufficient to earn the optimum position vector. Herein, the optimal position is the optimal hyper-parameter that occurs at the position, where fitness is minimized and maintained constant. The maximum generation is set to 100 to obtain the optimal results. Table 2 presents the values of parameters in PSO-RF algorithms.

Table 2 Parameters in PSO algorithm.

X_{min}	X_{max}	V_{min}	V_{max}	c_1	c_2	Generation	Population size
1	300	-1	1	1.49	1.49	100	20

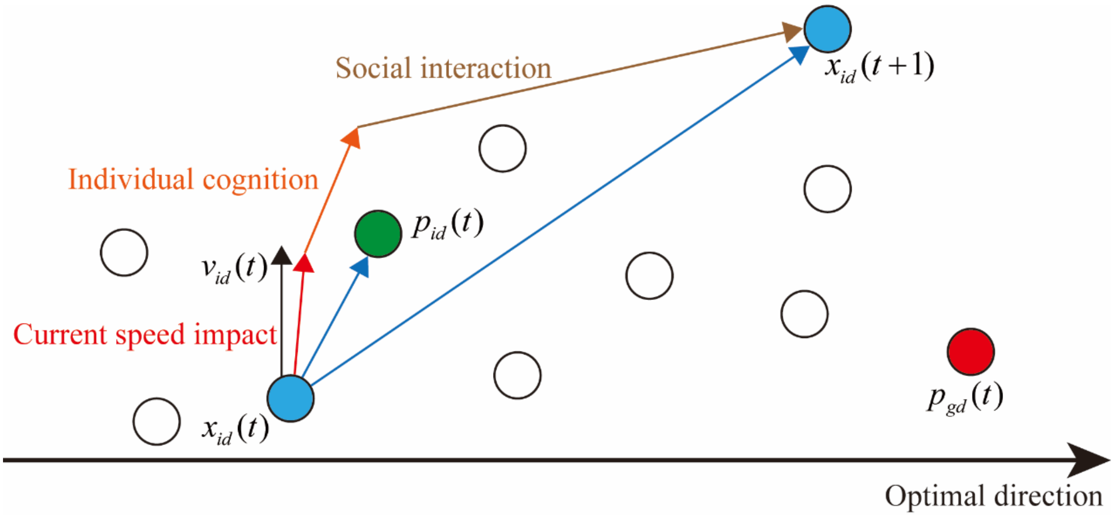

Fig. 3. The architecture of particle swarm optimization.

The hybrid PSO-RF model

Particle swarm optimization algorithm has the advantages of independent problem information, strong universality of the algorithm, few parameters to be adjusted, simple principle and easy implementation (Eberhart and Kennedy 1995). Therefore, PSO is adopted to optimize the hyper-parameters m_{try} and n_{tree} of the RF algorithm. The PSO-RF model framework is shown in Fig. 4, which has three stages: data processing, RF model training and testing.

In the data processing stage, the data set consists of several selected influential factors, and then the data is divided into eight different parameter combinations by feature selection method to obtain the best combination. After that, 80% of data is randomly assigned to train the model, while the rest of the data is used to test the model (Zhang, et al. 2020b; Zhou, et al. 2021c). The data in this paper were all measured in practice, and there are no outliers, duplicate values, and missing values, so there is no data cleaning, and all data are used in this paper. All data were normalized into [-1,1] to increase the computational efficiency and enhance the performance of the model.

In the training stage, the optimal RF models are determined when the optimal hyperparameters of the models are searched by the PSO algorithm. Firstly, the initial location and speed of the particles are randomly assigned, and the corresponding hyper-parameters of the RF model are specified. To acquire the fitness of each model, an approach called 10-fold CV is favoured. After the fitness of each round has been calculated, the local optimal position and global optimal position of the particle swarm are determined. Because the position and velocity of particles are dynamic, the best RF model can be obtained when the number of iterations reaches the maximum and the fitness value does not change.

In the testing stage, the eight optimal models were evaluated by their respective test sets. Better performance models have higher overall scores in indicators (MAE, RMSE, R^2, MAPE).

There are several ways to validate models, including the holdout method, cross-validation and bootstrap (Brenning 2012; Zhao et al. 2015; Li and Jimenez 2018). Since the data used in this paper is limited, if most of the data are used for training models, it will easily lead to model overfitting, so a cross-validation method is used (Li 2020; Wang et al. 2021). In this study, a 10-fold CV method which is widely used and has proven to have good performance is proposed to improve the credibility of the hybrid models and the MAE of 10 datasets is employed as a fitness function to quantificationally evaluate the accuracy of the hybrid model, as shown in Eq. (5) (Zhang, et al. 2020b).

$$Fitness = \frac{1}{10}\sum_{i=1}^{10} MAE_i \tag{5}$$

where, MAE_i refers mean decrease error of ith validation set.

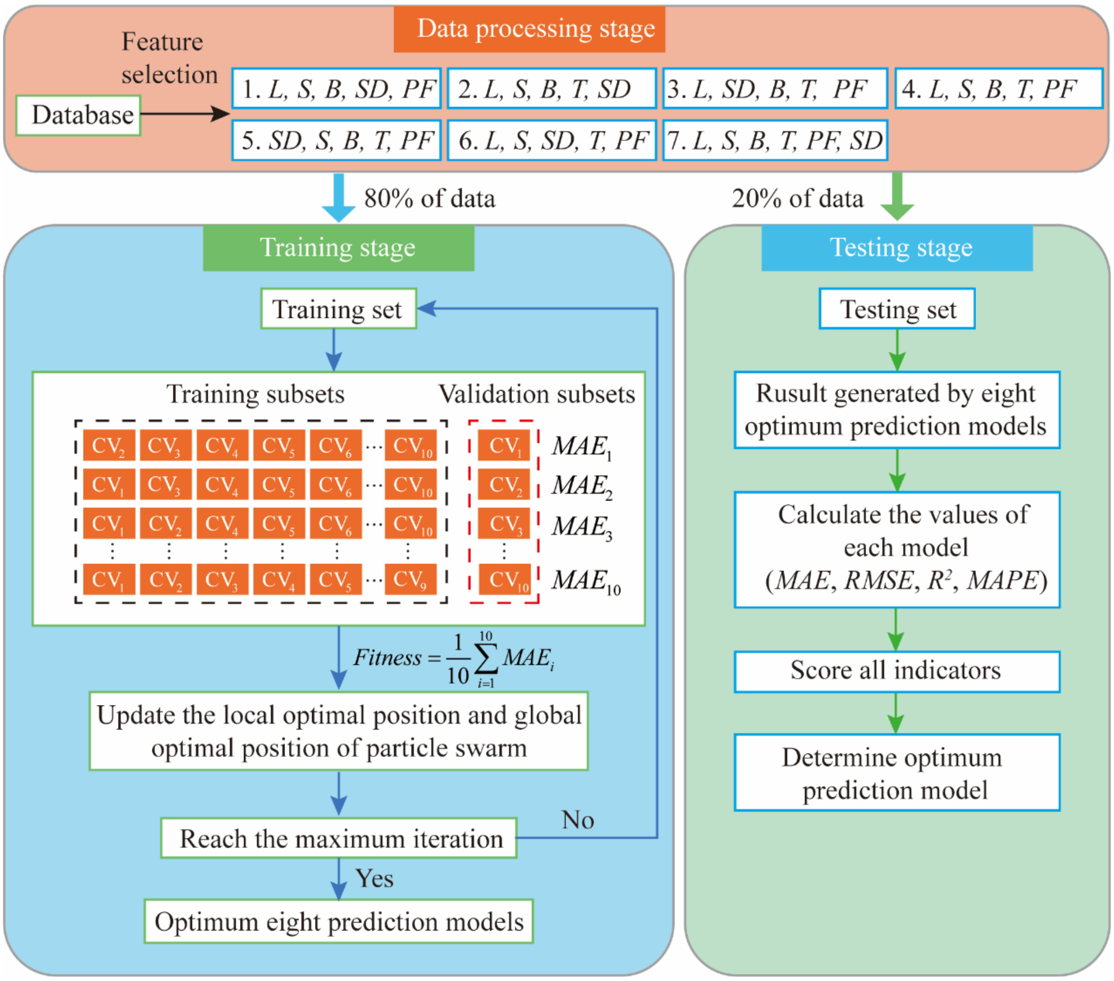

Fig. 4. Flow chart of the proposed PSO-RF model.

Grey relational analysis

As aforementioned above, BB is affected by many factors, and the correct selection of influencing factors has great effects on the prediction efficiency and accuracy of the model. The correlations among variables have been evaluated by grey correlation analysis (Zhang et al. 2013; Khan and Abdullah 2018; Li and Chen 2019), which is adopted to calculate the correlation degree between BB and the six selected factors.

The following are the calculation steps of grey relational analysis:

1) Reference variable $\delta_O = \delta_O[\delta_O(1), \delta_O(2), \ldots, \delta_O(n)]$ and compared variable $\delta_t = \delta_t[\delta_t(1), \delta_t(2), \ldots, \delta_t(n)]$ are given.

2) reduce the error of correlation analysis, interval transformation is used to make all variables dimensionless. The method is as follows:

$$\delta_t(u) = \frac{\delta_t(u) - \min\limits_u \delta_t(u)}{\max\limits_u \delta_t(u) - \min\limits_u \delta_t(u)} \tag{6}$$

where t=1, 2, …, m; u=1, 2, …, n.

3) The following is the calculation formula of grey relational degree between variables:

$$\gamma(\delta_O(u), \delta_t(u)) = \frac{\min\limits_t \min\limits_u |\Delta\delta| + \chi \max\limits_t \max\limits_u |\Delta\delta|}{|\Delta\delta| + \chi \max\limits_t \max\limits_u |\Delta\delta|} \tag{7}$$

where $\Delta\delta = \delta_O(u) - \delta_t(u)$, and the smaller it is, the bigger the correlation is.; χ = the resolving coefficient, was set to be 0.5. $\min\limits_t \min\limits_u |\Delta\delta|$ and $\max\limits_t \max\limits_u |\Delta\delta|$ denotes minimum and a maximum deviation of δ_O and δ_t, respectively, and their addition prevents grey relational degrees from being identical to 0 when one of them is 0.

4) The grey relational grade between the variables can thus be obtained by:

$$\gamma(\delta_O, \delta_t) = \frac{1}{n} \sum_{u=1}^{n} \gamma(\delta_O(u), \delta_t(u)) \tag{8}$$

where a large grey relational grade suggests a strong correlation between variables δ_O and δ_t.

Evaluation indicators

Four evaluation indicators namely mean absolute error (MAE), root mean square error

(RMSE), mean absolute percentage error (MAPE)and Pearson correlation coefficient (R^2) is applied in this work to evaluate the performance of the hybrid RF models. Scale-dependent indicator MAE, RMSE and scale-independent indicator R^2 and MAPE are applied to indicate the error between the measured and predicted value. MAPE can be used as a prediction index when the predicted value is not 0. If the error dispersion is high, that is, the maximum deviation is large, then the RMSE will increase because the RMSE is the square of the deviation. The range of R^2 is 0~1, the larger R^2 means the greater the correlation degree, and vice versa. The following equations depict the definition of MAE, RMSE, MAPE and R^2 (Armaghani et al. 2017; Zhang et al. 2020b; Zhou et al. 2020):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|BB_i - BB'_i\right| \tag{9}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(BB_i - BB'_i)^2} \tag{10}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{BB_i - BB'_i}{BB_i}\right| \times 100\% \tag{11}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(BB_i - BB'_i)^2}{\sum_{i=1}^{n}(BB_i - \overline{BB}_i)} \tag{12}$$

where, BB denotes the actual backbreak distance, BB' = predicted backbreak distance, n is the number of samples, $\overline{BB}$ refers to the average of actual backbreak distance.

Backbreak database and its description

The datasets used in this study were conducted by Khandelwal and Monjezi (2013) and composed of 234 blasting datasets of Sungun Copper Mine, Iran. As can be seen from Table 1, Backbreak is affected by many factors, but only the changeable parameters during the blasting operation in Sungun Copper Mine are considered in this study, and other parameters such as rock properties, hole diameter and bench height that affect BB remained constant and due to that, such constant parameters were not considered in this study. This database includes the following input parameters, namely, Burden (B), hole length (L), stemming (T), spacing (S), powder factor (PF) and special drilling (SD). Note that these input parameters have been widely used in previous models with better predictive performance (Monjezi et al. 2010, 2012; Faradonbeh et al. 2016; Khandelwal and Monjezi 2013; Ghasemi 2017; Hasanipanah et al. 2017) and thus, it is promising to build a superior model to predict the backbreak in the Sungun Copper mine. All parameters are recorded before each blasting operation. Because the bench crest is an uneven edge, several points on the bench crest are usually selected to measure the horizontal distance between them and the critical cracks, and the average value is regarded as backbreak distance. The descriptive statistics of the parameters are presented in Table 3. To visualize the distribution of input parameters, which are normalized to [0,1] and plotted in a violin diagram in Fig. 5. The datasets are randomly divided into 190 training datasets (approximately 80%) and 44 (approximately 20%) testing datasets, to develop the prediction model and evaluate the generation ability of the model.

Table 3 the descriptive statistics of the parameters.

Parameters	Mean	Standard deviation	Median	Min	Max	Range

L (m)	12.30983	1.181205	12.5	10	14	4
S (m)	4.527778	0.898882	4.5	2	6.5	4.5
B (m)	3.694444	0.81362	4	2	5	3
T (m)	3.663675	0.762366	4	1.8	4.5	2.7
PF (kg/m^3)	0.460812	0.197263	0.4	0.2	0.93	0.73
SD (m/m^3)	0.072906	0.039485	0.06	0.04	0.29	0.25
BB (m)	4.320513	2.541923	4	1	10	9

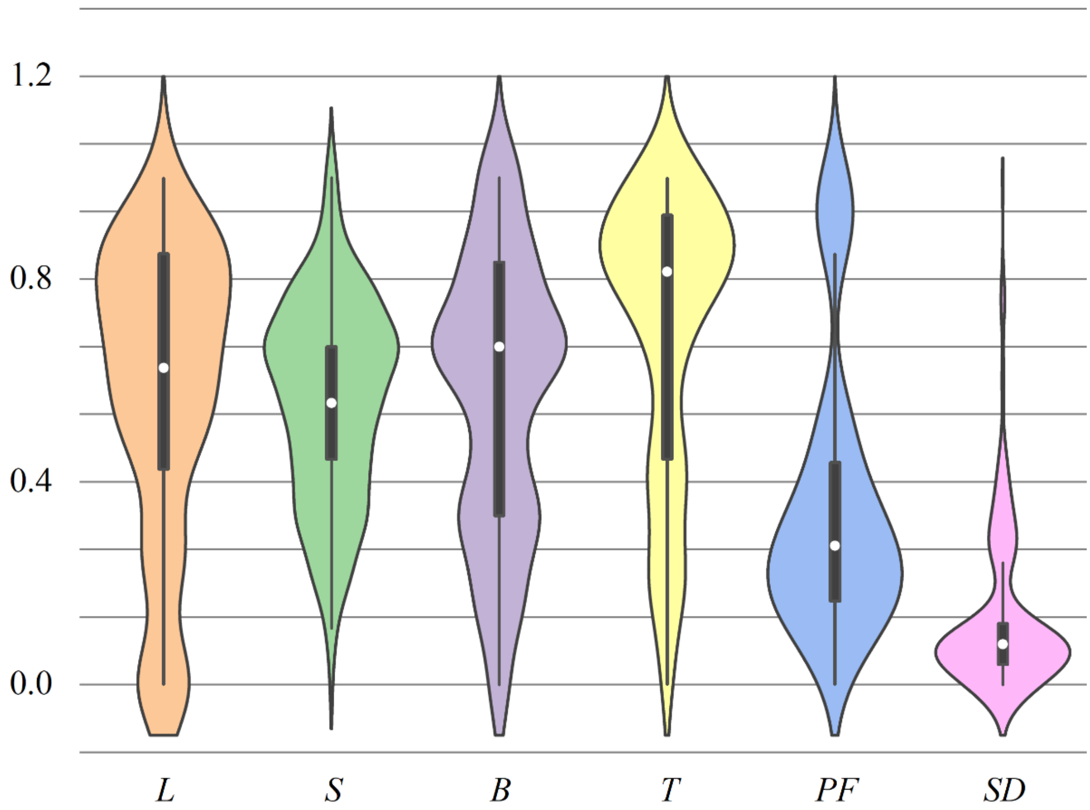

Fig. 5. Violin plots of input parameters for BB prediction

As aforementioned above, backbreak is affected by many factors, and the correct selection of influencing factors has great effects on the prediction efficiency and accuracy of the model. The correlations among variables have been evaluated by grey correlation analysis (Khan and Abdullah 2018; Li and Chen 2019), which is adopted to calculate the correlation degree between backbreak and the six selected factors.

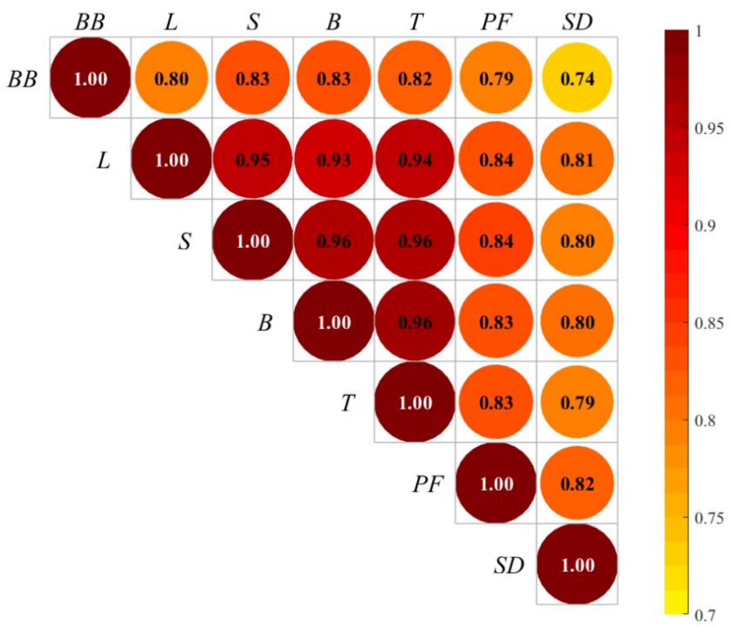

Fig. 6. The grey relational grade between variables.

It can be seen from Fig.6 that the grey relational grade between backbreak and each influential factor is in the range of 0.74-0.83, S has the highest degree of correlation between backbreak, and SD has the least degree of correlation between backbreak. Grey correlation analysis shows that the inputs and output have a high correlation, suggesting that the influential factors selected in this article are suitable for predicting backbreak.

Result

Optimal hyper-parameters

As aforementioned above, the PSO algorithm is applied to optimize the hyper-parameters (i.e., m_{try} and n_{tree}) of RF models, and MAE is applied as a fitness function to determine the optimum two hyper-parameters. Table 4 summarizes optimal hyper-parameters in each PSO-model. The fitness of 7 models is obtained by using 100 generations as the stop criteria. As shown in Fig. 7, the fitness values of the 7 models remained constant after 22 generations, suggesting that the optimal model could be obtained before 100 generations. The fitness values of RF models range from 0.6842 to 0.1026. The lowest fitness value, 0.6842, appears in the RF model combining L-S-B-T-SD, even better than the RF model combining all variables, suggesting a better performance in CV sets of the model. The biggest fitness value appears in the RF model combining L-S-B-T-PF and keeps constant from the first generation, indicating that the model was relatively worse in the CV sets.

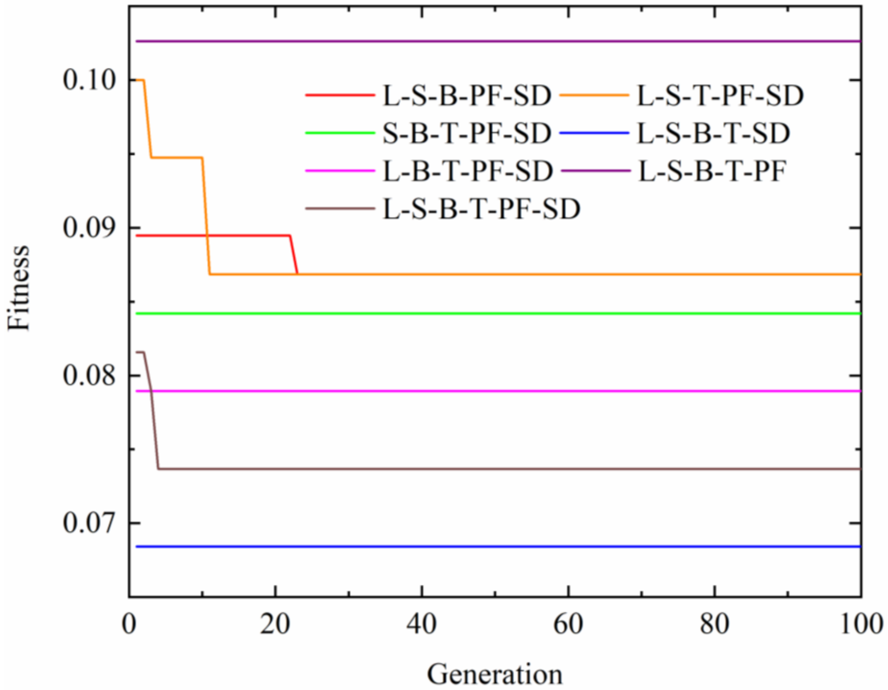

Fig. 7. Evolution of fitness value in all proposed RF models.

Prediction of backbreak using PSO- RF

If each model's hyper-parameters can be identified positively, seven optimal RF models can be established and their performance can be evaluated. The statistical indices, MAE, RMSE, R^2 and MAPE are adopted to evaluate the accuracy level of each constructed model. It is difficult to evaluate the accuracy of the model only by these index values. Therefore, to explain the optimal model, statistical index values of the training sets and testing sets were carried out and ranked synthetically. Tables 5 and 6 respectively summarize the statistical index values of the predicted backbreak for training sets and testing sets in seven optimal models and the ranking of each model is shown in Table 7. Marginal histograms of the predicted results of seven optimal models are presented in Fig. 8, which can visually see the contribution of the data on the X and Y axes.

It can be seen from Tables 5-7 and Fig. 8 that the predicted backbreak for the training sets and testing sets are next to the P=M line. The prediction error of each optimal model is fairly small, demonstrating that seven optimal RF models developed by the PSO algorithm are promising in performance. It is worth noting that the model combining L-S-B-T-PF outperforms the remaining models, however the model combining L-B-T-PF-SD provides the lowest performance. The remaining models also perform quite well, as shown in Fig. 8. In conclusion, the combinations integrate more variables and are more robust in prediction. Therefore, the combinations L-S-B-T-PF (PSO-RF1) and L-S-B-T-PF-SD (PSO-RF2) are recommended as the optimal models for predicting backbreak in engineering practice.

Table 4 Optimal hyper-parameters in each PSO-model.

	L-S-B-T-PF	L-S-T-PF-SD	S-B-T-PF-SD	L-S-B-T-SD	L-B-T-PF-SD	L-S-B-PF-SD	L-S-B-T-PF-SD
m_{try}	2	2	2	2	2	2	2
n_{tree}	25	37	20	41	31	30	65

Table 5 Comparison of the training set performance of RF-based hybrid models.

Input variables	Training set								Score
	MAE	Score	RMSE	Score	R^2	Score	MAPE	Score	
L-S-B-T-PF	0.0132	7	0.0811	7	0.9990	7	0.0027	7	28
L-S-T-PF-SD	0.0158	6	0.0889	6	0.9988	6	0.0029	6	24
S-B-T-PF-SD	0.0290	2	0.1089	5	0.9978	2	0.0053	2	11
L-S-B-T-SD	0.0211	3	0.1026	2	0.9984	4	0.0034	4	13
L-B-T-PF-SD	0.0211	3	0.1026	2	0.9983	3	0.0037	3	11
L-S-B-PF-SD	0.0474	1	0.1147	1	0.9959	1	0.0091	1	4
L-S-B-T-PF-SD	0.0211	3	0.1026	2	0.9986	5	0.0033	5	15

Table 6 Comparison of the test set performance of RF-based hybrid models.

Input variables	Test set								Score
	MAE	Score	RMSE	Score	R^2	Score	MAPE	Score	
L-S-B-T-PF	0.0568	3	0.1686	3	0.9961	3	0.0116	3	12
L-S-T-PF-SD	0.0795	1	0.1508	4	0.9892	1	0.0147	1	7
S-B-T-PF-SD	0.0341	5	0.1306	5	0.9975	4	0.0085	4	18
L-S-B-T-SD	0.0341	5	0.1306	5	0.9975	4	0.0085	4	18
L-B-T-PF-SD	0.0795	1	0.2261	1	0.9926	2	0.0137	2	6
L-S-B-PF-SD	0.0227	7	0.1066	7	0.9983	7	0.0033	6	27
L-S-B-T-PF-SD	0.0455	4	0.1846	2	0.9982	6	0.0033	7	19

Table 7 Comparison of performance of RF-based hybrid models.

Input variables	L-S-B-T-PF	L-S-T-PF-SD	S-B-T-PF-SD	L-S-B-T-SD	L-B-T-PF-SD	L-S-B-PF-SD	L-S-B-T-PF-SD
Training score	28	24	11	13	11	4	15

Test score	12	7	18	18	6	27	19
Total score	40	31	29	31	27	31	34
Rank	1	3	6	3	7	3	2

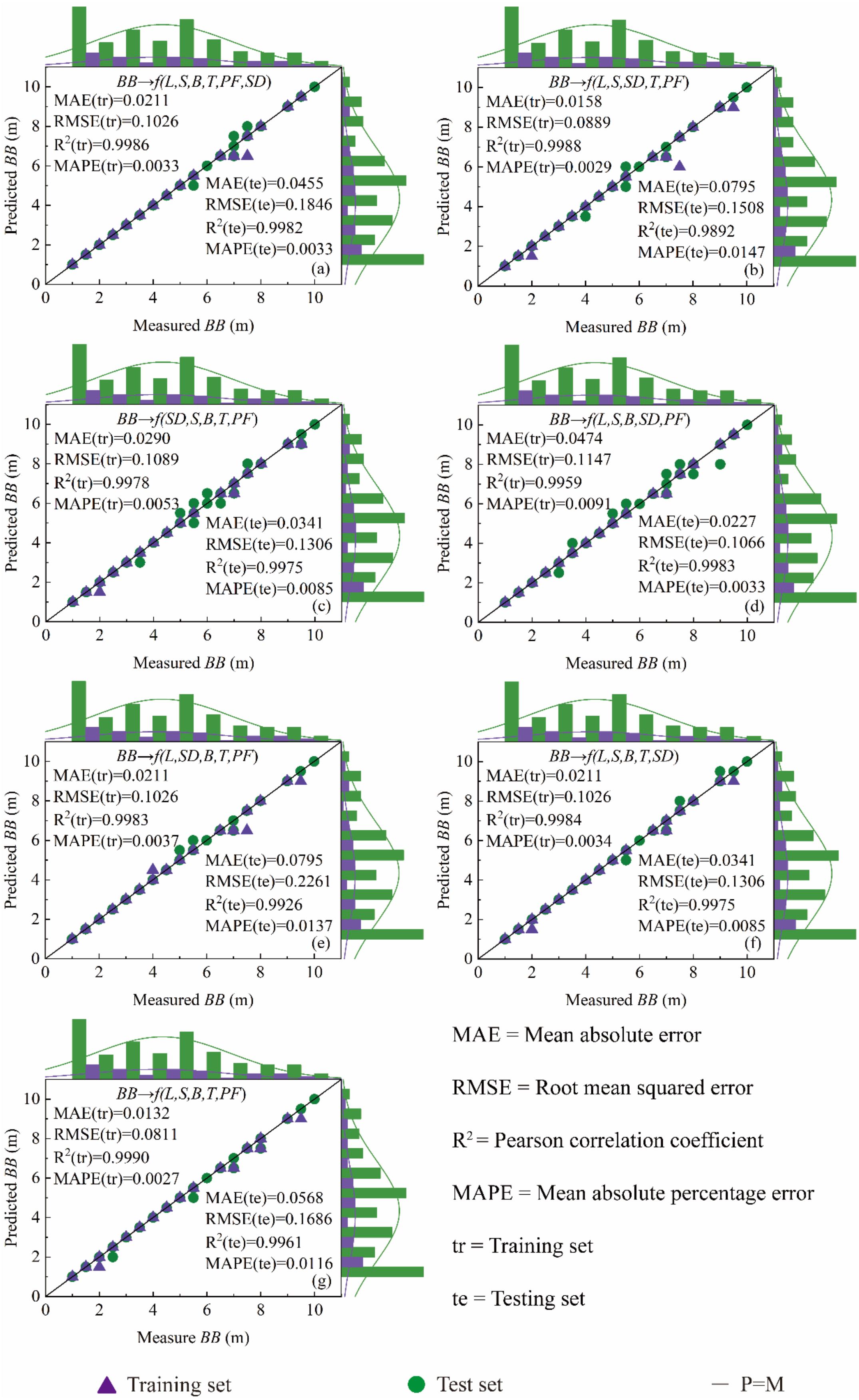

Fig. 8. Marginal histograms of the predicted results of seven optimal models.

Discussion

Comparison with different combinations of models

Fig. 9 presents the distributions of MAE values in each RF model, with lower mean MAE value (fitness value) appearing in outperformed RF models. According to the fitness values, the RF model combining L-S-B-T-SD performs best in CV sets, a little better than the RF model combining all parameters. Therefore, for the performance of CV sets, the RF model which contains all variables and the RF model combining L-S-B-T-SD are better than that of the remaining models, indicating that the backbreak prediction models combining more variables outperform those with fewer.

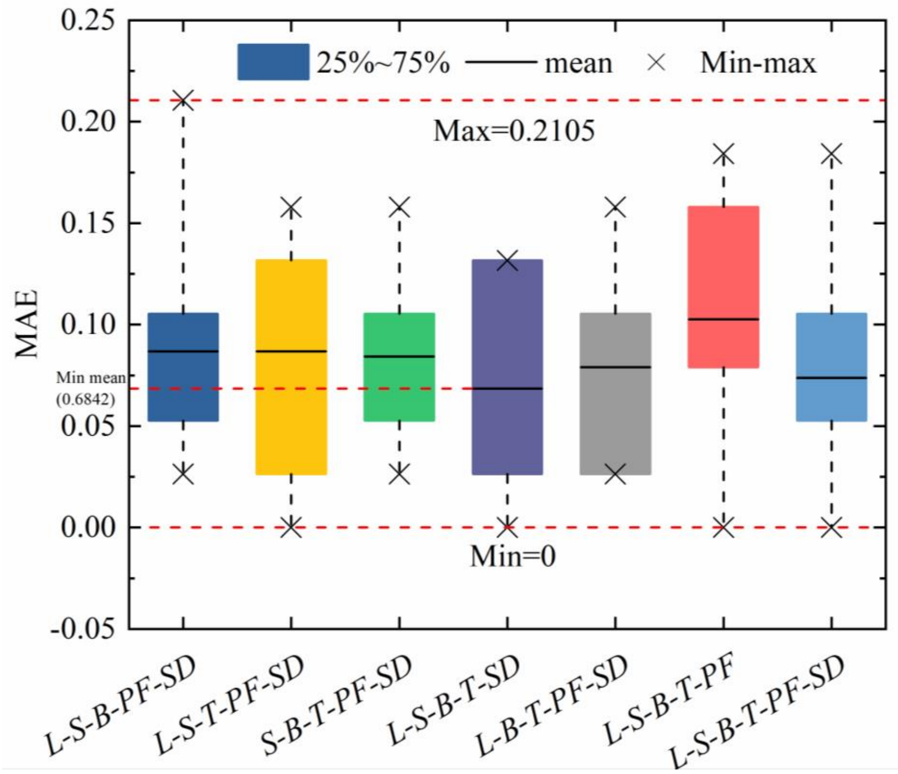

Fig. 9. Distribution of MAE values in ten CV sets.

Comparison with different classical models

By comparing four classical regression algorithms, namely SVM, GP, RF and ANN (Goh and Goh 2007; Liu et al. 2021; Ferentinou and Fakir 2018), we can verify the advanced algorithm. Based on the conclusion that the combinations integrating more variables are more robust in predicting backbreak, all variables are adopted in these models. The scatter plots of the predicted backbreak using these models are presented in Fig. 10. It is not difficult to find that the prediction accuracy of all models is roughly the same, but the PSO-RF models outperform slightly from Figs. 8 and 10. Additionally, the accuracy of classical models is shown in Table 8 and the comprehensive prediction score is presented in Fig. 11 to visually compare the differences between models. As shown in Fig. 11, the PSO-RF model combining L-S-B-T-PF achieve the highest score, indicating the most outstanding performance (MAE=0.0132, RMSE=0.0811, R^2=0.9990, MAPE=0.0027 on the training dataset; MAE=0.0568, RMSE=0.1686, R^2=0.9961, MAPE=0.0116 on the test dataset) of it. What is worth noting is the PSO-RF model has better prediction accuracy than the remaining models. To sum up, the models proposed in this study perform commonly well in predicting and evaluating backbreak in open-pit mines, especially the proposed PSO-RF model. Therefore, the PSO-RF is recommended for predicting backbreak in engineering practice.

In previous research, Khandelwal and Monjezi (2013) developed multivariate regression analysis (MVRA) and SVM models for forecasting backbreak with the same dataset. They employed R^2 and MAE as evaluation indicators, which are (R^2=0.987, MAE=0.29) and (R^2=0.89, MAE=1.07) in the MVRA and SVM models, respectively. Compared with the results obtained by (Khandelwal and Monjezi 2013), the PSO-RF model shows more outstanding performance in predicting backbreak.

Therefore, it is recommended to use the PSO-RF model to predict backbreak in practice.

Table 8 The accuracy of classical models.

		MAE	RMSE	R^2	MAPE
SVM	Train	0.1244	0.2074	0.9686	0.0319
	Test	0.1554	0.2874	0.9507	0.0390
GP	Train	0.0849	0.1328	0.9799	0.0259
	Test	0.1155	0.1622	0.9722	0.0338
RF	Train	0.0875	0.1269	0.9808	0.0219
	Test	0.0917	0.1420	0.9757	0.0279
ANN	Train	0.1276	0.1691	0.9744	0.0343
	Test	0.1306	0.1910	0.9672	0.0352
CNN	Train	0.3166	0.4194	0.9364	0.1126
	Test	0.3254	0.4209	0.9277	0.1314

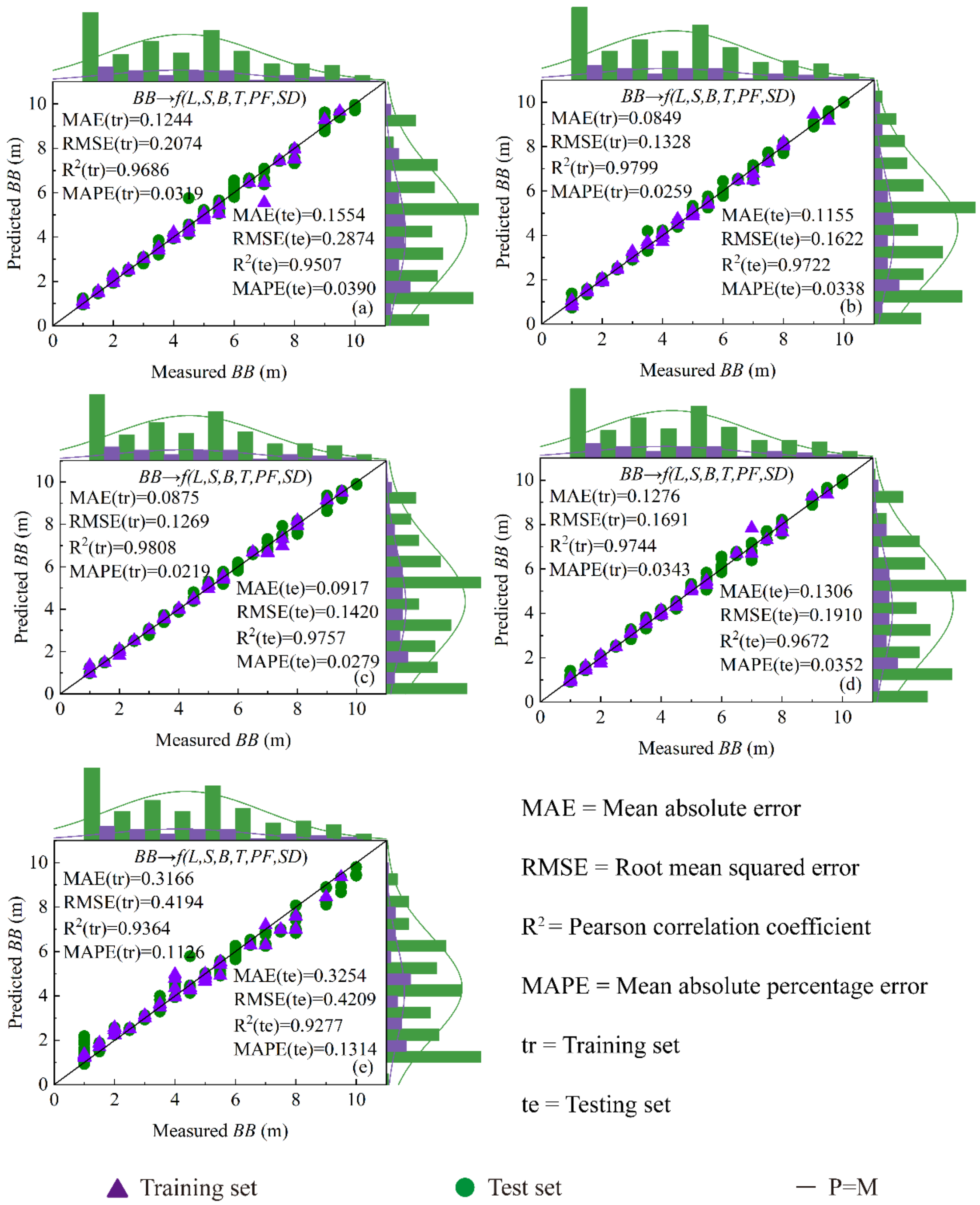

Fig. 10. Marginal histograms of the predicted results of four classical models. (a) SVM; (b) GP; (c) RF; (d) ANN; (e) CNN

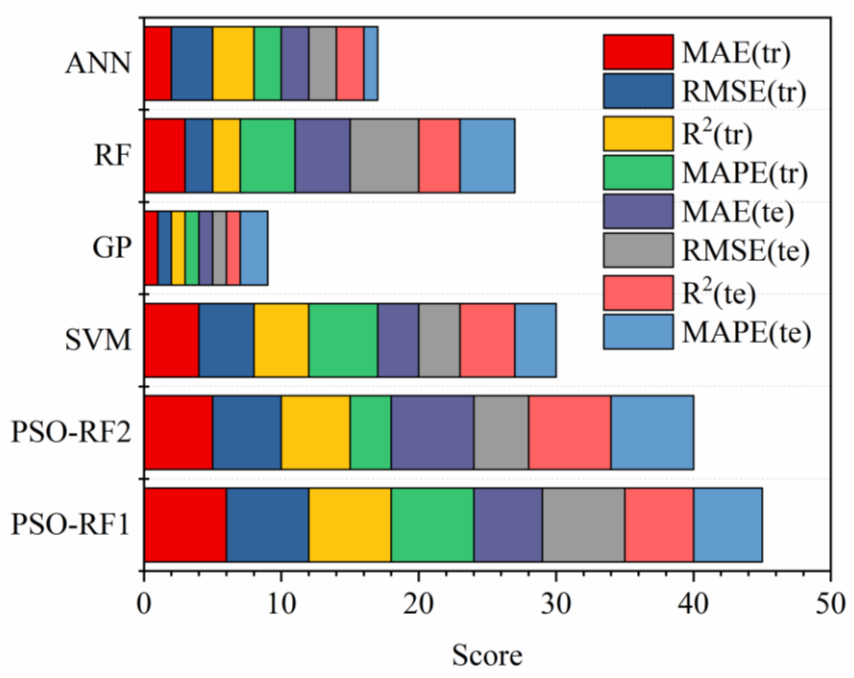

Fig. 11. Comprehensive sorted stacked graph for different models.

Sensitive analysis

The importance of the input variables largely determines the accuracy of the model. In previous studies, the importance of input variables was not analyzed. Therefore, based on the proposed PSO-RF method, the importance of input variables is studied in this paper. As the introduction explains, the Gini index of each input variable is computed internally using a random forest model to investigate the importance. Breiman et al. (1984) demonstrated variables that are more sensitive to the backbreak have a higher Gini index. The sensitivity analysis of each variable is shown in Fig. 12. It can be clearly seen that T is the most sensitive parameter to backbreak，followed by L, B, PF, S and SD. Their importance scores were 36.6, 31.2, 27.4, 23.4, 23.1 and 16.9, respectively. Some research results also show that T, L and B can greatly affect backbreak (Eskandar et al. 2018; Sari et al. 2014; Khandelwal and Monjezi 2013; Monjezi et al. 2013). In addition, Eskandar et al. (2018) reduce backbreak impact by decreasing B and T, and Sari et al. (2014) decrease backbreak distance by reducing L. In practical operation, through reasonably adjusting these sensitive parameters, and using PSO-RF model to test backbreak after adjusted blasting design, backbreak can be effectively reduced under the condition of meeting engineering requirements. Moreover, sensitivity analysis can provide a reference for selecting more important input parameters to establish a model in the future.

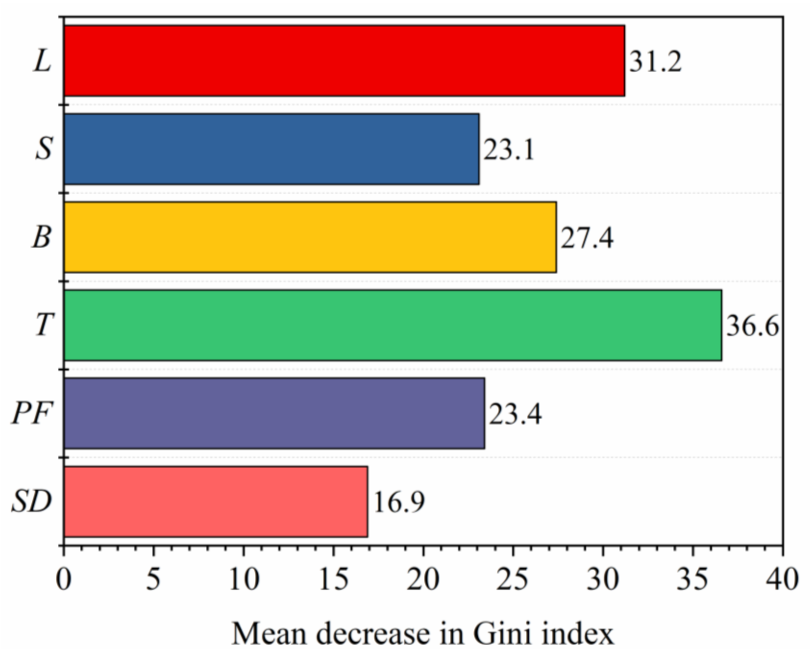

Fig. 12. The relative importance of the influenced variables.

Conclusions

A novel hybrid artificial intelligence approach using RF with PSO algorithm is presented for estimating backbreak modeling in open-pit blasting. A data set of 234 samples with six inputs and one output was built to build a prediction model. The grey relational grade between backbreak and each influential factor is in the range of 0.74-0.83, indicating the selected variables are capable of predicting backbreak. Seven input combinations were set up to obtain the optimal prediction model. The PSO algorithm was integrated with the RF algorithm for finding the optimal hyper-parameters of each model, whose fitness function is the MAE of 10-fold CV. After the prediction results of each optimal model were calculated, MAE, RMSE, R^2 and MAPE were employed as the statistical indicators. Then, the comprehensive performance of each model was evaluated by the overall score. The results indicated that the combinations of more variables are more robust in prediction. Thus, the results of this study suggested that the combinations L-S-B-T-PF and L-S-B-T-PF-SD are the optimal models for predicting backbreak in engineering practice. The optimal models recommended in the result section were compared with the proposed classical models (SVM, GP, RF, ANN, CNN), and the results showed that PSO-RF model had good performance in predicting backbreak. Finally, the RF algorithm was used to calculate the Gini index of each input variable internally, which were 31.2(L), 23.1(S), 27.4(B), 36.6(T), 23.4(PF), and 16.9(SD), respectively. This method can evaluate the sensitivity of input variables. The variables employed in this study are commonly sensitive to backbreak, suggesting that the parameter selection is reasonable. What's more, because the model is inseparable from the input parameters, the optimal models determined in this article are only recommended to predict backbreak under the same condition. Additionally, based on the advantages of the RF algorithm in predicting backbreak, it is recommended to use the RF algorithm to predict backbreak in other cases.

Acknowledgements

This study was funded by National Science Foundation of China (42177164) and the Innovation-Driven Project of Central South University (No. 2020CX040).

Conflict of Interest

All the authors declare that they have NO affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

Alemdag S, Gurocak Z, Cevik A, Cabalar AF, Gokceoglu C (2016) Modeling deformation modulus of a stratified sedimentary rock mass using neural network, fuzzy inference and genetic programming. Eng Geol 203: 70-82. https://doi.org/10.1016/j.enggeo.2015.12.002

Armaghani DJ, Hajihassani M, Mohamad ET, Marto A, Noorani SA (2014) Blasting-induced flyrock and ground vibration prediction through an expert artificial neural network based on particle swarm optimization. Arab J Geosci 7(12): 5383-5396. https://doi.org/10.1007/s12517-013-1174-0

Armaghani DJ, Mohamad ET, Narayanasamy MS, Narita N, Yagiz S (2017) Development of hybrid intelligent models for predicting TBM penetration rate in hard rock condition. Tunn Undergr Sp Tech 63: 29-43. https://doi.org/10.1016/j.tust.2016.12.009

Beiki M, Bashari A, Majdi A (2010) Genetic programming approach for estimating the deformation modulus of rock mass using sensitivity analysis by neural network. Int J Rock Mech Min 47(7): 1091-1103. https://doi.org/10.1016/j.ijrmms.2010.07.007

Berta G (1990) Explosives: an engineering tool. Italesplosivi, Millano.

Bhandari AK, Kumar A, Singh GK (2015) Modified artificial bee colony based computationally efficient multilevel thresholding for satellite image segmentation using Kapur's, Otsu and Tsallis functions. Expert Syst Appl 42(3): 1573-1601. https://doi.org/10.1016/j.eswa.2014.09.049

Biourge V, Delmotte S, Feugier A, Bradley R, McAllister M, Elliott J (2020) An artificial neural network-based model to predict chronic kidney disease in aged cats. J Vet Intern Med 34(5): 1920-1931. https://doi.org/10.1111/jvim.15892

Breiman L, Friedman J, Stone CJ, Olshen RA (1984) Classification and regression trees. CRC press.

Breiman L (2001) Random forests. Mach Learn 45(1): 5-32.

Brenning A (2012) Spatial cross-validation, bootstrap for the assessment of prediction rules in remote sensing: the R package sperrorest. 2012 IEEE International Geoscience and Remote Sensing Symposium: 5372-5375. https://doi.org/10.1109/IGARSS.2012.6352393

Civicioglu P (2012) Transforming geocentric cartesian coordinates to geodetic coordinates b y using differential search algorithm. Comput Geosci 46: 229-247. https://doi.org/10.101 6/j.cageo.2011.12.011

Eberhart R, Kennedy J (1995) A new optimizer using particle swarm theory. MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science, 4-6 Oct. 1995, New York, NY, USA, IEEE. https://doi.org/10.1109/MHS.1995.494215

Ebrahimi E, Monjezi M, Khalesi MR, Armaghani DJ (2016) Prediction and optimization of back-break and rock fragmentation using an artificial neural network and a bee colony algorithm. B Eng Geol Environ 75(1): 27-36. https://doi.org/10.1007/s10064-015-0720-2

Eskandar H, Heydari E, Hasanipanah M, Masir MJ, Derakhsh AM (2018) Feasibility of particle swarm optimization and multiple regression for the prediction of an environmental issue of mine blasting. Eng Computation 35(1): 363-376. https://doi.org/10.1108/EC-01-2017-0040

Esmaeili M, Osanloo M, Rashidinejad F, Bazzazi AA, Taji M (2014) Multiple regression, ANN and ANFIS models for prediction of backbreak in the open pit blasting. Eng Comput 30(4): 549-558. https://doi.org/10.1007/s00366-012-0298-2

Faradonbeh RS, Monjezi M, Armaghani DJ (2016) Genetic programing and non-linear multiple regression techniques to predict backbreak in blasting operation. Eng Comput 32(1): 123-133. https://doi.org/10.1007/s00366-015-0404-3

Ferentinou M, Fakir M (2018) Integrating Rock Engineering Systems device and Artificial Neural Networks to predict stability conditions in an open pit. Eng Geol 246: 293-309. https://doi.org/10.1016/j.enggeo.2018.10.010

Gates WCB, Ortiz LT, Florez RM (2005) Analysis of rockfall and blasting backbreak problems, US 550, Molas Pass, CO. 40th US Rock Mechanics Symposium: Rock Mechanics for Energy, Mineral, Infrastructure Development in the Northern Regions, ALASKA ROCKS 2005, June 25, 2005 - June 29, 2005, Anchorage, AK, United states, American Rock Mechanics

Association (ARMA).

Ghasemi E (2017) Particle swarm optimization approach for forecasting backbreak induced by bench blasting. Neural Comput Appl 28(7): 1855-1862. https://doi.org/10.1007/s00521-016-2182-2

Ghasemi E, Amnieh HB, Bagherpour R (2016) Assessment of backbreak due to blasting operation in open pit mines: a case study. Environ Earth Sci 75(7). https://doi.org/10.1007/s12665-016-5354-6

Goh ATC, Goh SH (2007) Support vector machines: Their use in geotechnical engineering as illustrated using seismic liquefaction data. Comput Geotech 34(5): 410-421. https://doi.org/10.1016/j.compgeo.2007.06.001

Gong WP, Luo Z, Juang CH, Huang HW, Zhang J, Wang L (2014) Optimization of site exploration program for improved prediction of tunneling-induced ground settlement in clays. Comput Geotech 56: 69-79. https://doi.org/10.1016/j.compgeo.2013.10.008

Hasanipanah M, Bakhshandeh AH (2021) Developing a new uncertain rule-based fuzzy approach for evaluating the blast-induced backbreak. Eng Comput 37, 1879–1893. https://doi.org/10.1007/s00366-019-00919-6

Hasanipanah M, Shahnazar A, Arab H, Golzar SB, Amiri M (2017) Developing a new hybrid-AI model to predict blast-induced backbreak. Eng Comput 33(3): 349-359. https://doi.org/10.1007/s00366-016-0477-7

He M, Zhang Z, Li N (2021) Deep Convolutional Neural Network-Based Method for Strength Parameter Prediction of Jointed Rock Mass Using Drilling Logging Data. Int J Geomech 21(7) https://doi.org/04021111. 10.1061/(ASCE)GM.1943-5622.0002074

Hosseini SA, Tavana A, Abdolahi SM, Darvishmaslak S (2019) Prediction of blast-induced ground vibrations in quarry sites: a comparison of GP, RSM and MARS. Soil Dyn Earthq Eng 119: 118-129. https://doi.org/10.1016/j.soildyn.2019.01.011

Huang G, Xiao L (2021) Failure mode and effect analysis: An interval-valued intuitionistic fuzzy cloud theory-based method. Appl Soft Comput 98: 106834. https://doi.org/

Iphar M, Yavuz M, Ak H (2008) Prediction of ground vibrations resulting from the blasting operations in an open-pit mine by adaptive neuro-fuzzy inference system. Environ Geol 56(1): 97-107. https://doi.org/10.1016/j.asoc.2020.106834

Jamasb A, Motavalli-Anbaran SH, Zeyen H (2017) Non-linear stochastic inversion of gravity data via quantum-behaved particle swarm optimisation: application to Eurasia-Arabia collision zone (Zagros, Iran). Geophys Prospect 65: 274-294. https://doi.org/10.1111/1365-2478.12558

Jimeno CJ, EL; Carcedo FJA (1995) Drilling and blasting of rocks. Balkema, Rotterdam.

Lundborg N (1974) The hazards of fly rock in rock blasting. Report DS1974, Swedish Detonic Res Found (SveDeFo), Stockholm.

Khan MSA, Abdullah S (2018) Interval-valued Pythagorean fuzzy GRA method for multiple-attribute decision making with incomplete weight information. Int J Intell Syst 33(8): 1689-1716. https://doi.org/10.1002/int.21992

Khandelwal M, Monjezi M (2013) Prediction of Backbreak in Open-pit Blasting Operations Using the Machine Learning Method. Rock Mech Rock Eng 46(2): 389-396. https://doi.org/10.1007/s00603-012-0269-3

Khandelwal M, Singh TN (2013) Application of an Expert System to Predict Maximum Explosive Charge Used Per Delay in Surface Mining. Rock Mech Rock Eng 46(6): 1551-1558. https://doi.org/10.1007/s00603-013-0368-9

Khandelwal, M., Mahdiyar, A., Armaghani, D.J. et al. (2017) An expert system based on hybrid ICA-ANN technique to estimate macerals contents of Indian coals. Environ Earth Sci 76, 399. https://doi.org/10.1007/s12665-017-6726-2

Khandelwal, M., Singh, T.N. (2011) Predicting elastic properties of schistose rocks from unconfined strength using intelligent approach. Arab J Geosci 4, 435–442. https://doi.org/10.1007/s12517-009-0093-6

Kumar S, Mishra AK, Choudhary BS (2021) Prediction of back break in blasting using random decision trees. Eng Comput 1-7 https://doi.org/10.1007/s00366-020-01280-9

Li DT, Yan JL, Zhang L (2012) Prediction of blast-induced ground vibration using support vector machine by tunnel excavation. Appl Mech Mater p.1414-1418. https://doi.org/10.4028/www.scientific.net/AMM.170-173.1414

Li N, Jimenez R (2018) A logistic regression classifier for long-term probabilistic prediction of rock burst hazard. Nat Hazards 90(1): 197-215. https://doi.org/10.1007/s11069-017-3044-7

Li N, Yi C (2020) Predicting underground rock pillar stability using Logistic Model Tree method. ISRM International Symposium - EUROCK 2020, June 14, 2020 - June 19, 2020, Trondheim, Virtual, Norway, International Society for Rock Mechanics.

Li Z, Chen L (2019) A novel evidential FMEA method by integrating fuzzy belief structure and grey relational projection method. Eng Appl Artif Intel 77: 136-147. https://doi.org/10.1016/j.engappai.2018.10.005

Li E, Zhou J, Shi X, Armaghani DJ, Yu Z, Chen X, Huang P (2020) Developing a hybrid model of salp swarm algorithm based support vector machine to predict the strength of fiber reinforced cemented paste backfill. Eng Comput 1-22. https://doi.org/10.1007/s00366-020-01014-x

Liang WZ, Zhao GY, Wang X, Zhao J, Ma CD (2019) Assessing the rockburst risk for deep shafts via distance-based multi-criteria decision making approaches with hesitant fuzzy information. Eng Geol 260: 12. https://doi.org/10.1016/j.enggeo.2019.105211

Liao, X., Khandelwal, M., Yang, H. et al. (2020) Effects of a proper feature selection on prediction and optimization of drilling rate using intelligent techniques. Engineering with Computers 36, 499–510. https://doi.org/10.1007/s00366-019-00711-6

Liu Y, Gu Z, Hughes DJ, Ye J, Hou X (2021) Understanding mixed mode ratio of adhesively bonded joints using genetic programming (GP). Compos Struct 258: 113389. https://doi.org/10.1016/j.compstruct.2020.113389

Mirghasemi S, Andreae P, Zhang MJ (2019) Domain-independent severely noisy image segmentation via adaptive wavelet shrinkage using particle swarm optimization and fuzzy C-means. Expert Syst Appl 133: 126-150. https://doi.org/10.1016/j.eswa.2019.04.050

Moeinossadat SR, Ahangari K, Shahriar K (2018) Modeling maximum surface settlement due to EPBM tunneling by various soft computing techniques. Innov Infrastruct So 3(1): 13. https://doi.org/10.1007/s41062-017-0114-3

Mohammadnejad M, Gholami R, Sereshki F, Jamshidi A (2013) A new methodology to predict backbreak in blasting operation. Int J Rock Mech Min 60: 75-81. https://doi.org/10.1016/j.ijrmms.2012.12.019

Monjezi, M., Mohamadi, H.A., Barati, B., Khandelwal M. (2014) Application of soft computing in predicting rock fragmentation to reduce environmental blasting side effects. Arab J Geosci 7, 505–511. https://doi.org/10.1007/s12517-012-0770-8

Monjezi M, Rizi SH, Majd VJ, Khandelwal M (2014) Artificial neural network as a tool for backbreak prediction. Geotech Geol Eng 32(1), 21-30. https://doi.org/10.1007/s10706-

013-9686-7

Monjezi M, Ahmadi Z, Varjani AY, Khandelwal M (2013) Backbreak prediction in the Chadormalu iron mine using artificial neural network. Neural Comput Appl 23(3-4): 1101-1107. https://doi.org/10.1007/s00521-012-1038-7

Monjezi M, Singh TN, Khandelwal M, Sinha S, Singh V, Hosseini I. (2006) Prediction and Analysis of Blast Parameters Using Artificial Neural Network. Noise & Vibration Worldwide 37(5), 8-16. doi:10.1260/095745606777630323

Monjezi M, Khoshalan HA, Varjani AY (2012) Prediction of flyrock and backbreak in open pit blasting operation: a neuro-genetic approach. Arab J Geosci 5(3): 441-448. https://doi.org/10.1007/s12517-010-0185-3

Monjezi M, Rezaei M, Yazdian A (2010) Prediction of backbreak in open-pit blasting using fuzzy set theory. Expert Syst Appl 37(3): 2637-2643. https://doi.org/10.1016/j.eswa.2009.08.014

Nabiollahi K, Taghizadeh-Mehrjardi R, Shahabi A, Heung B, Amirian-Chakan A, Davari M, Scholten T (2021) Assessing agricultural salt-affected land using digital soil mapping and hybridized random forests. Geoderma, 385, 114858. https://doi.org/10.1016/j.geoderma.2020.114858

Ray U, Chouhan U, Verma N (2020) Comparative study of machine learning approaches for classification and prediction of selective caspase-3 antagonist for Zika virus drugs. Neural Comput Appl 1-18. https://doi.org/10.1007/s00521-019-04626-7

Roth J (1979) A model for the determination of flyrock range as a function of shot condition. US Bureau of Mines Contract J0387242. Management Science Associates: p61.

Sari M, Ghasemi E, Ataei M (2014) Stochastic Modeling Approach for the Evaluation of Backbreak due to Blasting Operations in Open Pit Mines. Rock Mech Rock Eng 47(2): 771-783. https://doi.org/10.1007/s00603-013-0438-z

Sayadi A, Monjezi M, Talebi N, Khandelwal M (2013) A comparative study on the application of various artificial neural networks to simultaneous prediction of rock fragmentation and backbreak. J Rock Mech Geotech 5(4), 318-324. https://doi.org/10.1016/j.jrmge.2013.05.007

Seo JH, Im CH, Heo CG, Kim JK, Jung HK, Lee CG (2006) Multimodal function optimization based on particle swarm optimization. IEEE T Magn 42(4): 1095-1098. https://doi.org/10.1109/TMAG.2006.871568

Sharma M, Choudhary BS, Agrawal H (2021) Prediction and Assessment of Back Break by Multivariate Regression Analysis, and Random Forest algorithm in hot strata/fiery seam of open-pit coal mine. https://doi.org/10.21203/rs.3.rs-267513/v1

Wang H, Zhang YM, Yang Z (2019) A risk evaluation method to prioritize failure modes based on failure data and a combination of fuzzy sets theory and grey theory. Eng Appl Artif Intel 82: 216-225. https://doi.org/10.1016/j.engappai.2019.03.023

Wang SM, Zhou J, Li CQ, Armaghani DJ, Li XB, Mitri HS (2021) Rockburst prediction in hard rock mines developing bagging and boosting tree-based ensemble techniques. J Cent South Univ 28(2): 527-542. https://doi.org/10.1007/s11771-021-4619-8

Wang Y, Lu C, Zuo C (2015) Coal mine safety production forewarning based on improved BP neural network. Int J Min Sci Techno 25(2): 319-324. https://doi.org/10.1016/j.ijmst.2015.02.023

Wu QH, Song T, Liu HM, Yan XS (2017) Particle swarm optimization algorithm based on parameter improvements. J Comput Methods Sci 17(3): 557-568. https://doi.org/10.3233/JCM-170742

Yin X, Liu QS, Pan YC, Huang X, Wu J, Wang XY (2021) Strength of Stacking Technique of Ensemble Learning in Rockburst Prediction with Imbalanced Data: Comparison of Eight Single and Ensemble Models. Nat Resour Res 30(2): 1795-1815. https://doi.org/10.1007/s11053-020-09787-0

Yu Z, Shi X, Zhou J, Gou Y, Huo X, Zhang J, Armaghani DJ (2020) A new multikernel relevance vector machine based on the HPSOGWO algorithm for predicting and controlling blast-induced ground vibration. Eng Comput 1-16. https://doi.org/10.1007/s00366-020-01136-2

Zhang P, Wu HN, Chen RP, Chan TH (2020a) Hybrid meta-heuristic and machine learning algorithms for tunneling-induced settlement prediction: A comparative study. Tunn Undergr Sp Tech 99, 103383. https://doi.org/10.1016/j.tust.2020.103383

Zhang P, Yin ZY, Jin YF, Chan THT (2020b) A novel hybrid surrogate intelligent model for creep index prediction based on particle swarm optimization and random forest. Eng Geol 265: 105328. https://doi.org/10.1016/j.enggeo.2019.105328

Zhang X, Jin F, Liu P (2013) A grey relational projection method for multi-attribute decision making based on intuitionistic trapezoidal fuzzy number. Appl Math Model 37(5): 3467-3477. https://doi.org/10.1016/j.apm.2012.08.012

Zhang H, Nguyen H, Bui XN, Pradhan B, Mai NL, Vu DA (2021) Proposing two novel hybrid intelligence models for forecasting copper price based on extreme learning machine and meta-heuristic algorithms. Resour Policy 73, 102195. https://doi.org/ 10.1016/j.resourpol.2021.102195

Zhao C, He J, Zhang X, Qi X, Chen A (2015) Recognition of driving postures by nonsubsampled contourlet transform and k-nearest neighbor classifier. Comput Syst Sci Eng 30(3), 233-241. https://doi.org/10.1049/iet-its.2011.0116

Zhao H, Li S, Ru Z (2017) Adaptive reliability analysis based on a support vector machine and its application to rock engineering. Appl Math Model 44: 508-522. https://doi.org/10.1016/j.apm.2017.02.020

Zhou J, Li X, Mitri HS (2016) Classification of rockburst in underground projects: Comparison of ten supervised learning methods. J Comput Civil Eng 30(5), 04016003. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000553

Zhou J, Li XB, Shi XZ (2012) Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines. Safety Sci 50(4): 629-644. https://doi.org/10.1016/j.ssci.2011.08.065

Zhou J, Li E, Wei H, Li C, Qiao Q, Armaghani DJ (2019) Random forests and cubist algorithms for predicting shear strengths of rockfill materials. Appl Sci 9(8):1621. https://doi.org/10.3390/app9081621

Zhou J, Asteris PG, Armaghani DJ, Pham BT (2020) Prediction of ground vibration induced by blasting operations through the use of the Bayesian Network and random forest models. Soil Dyn Earthq Eng 139, p.106390. https://doi.org/10.1016/j.soildyn.2020.106390

Zhou J, Shi X, Du K, Qiu X, Li X, Mitri HS (2017) Feasibility of random-forest approach for prediction of ground settlements induced by the construction of a shield-driven tunnel. Int J Geomech 17(6), p.04016129. https://doi.org/10.1061/(ASCE)GM.1943-5622.0000817

Zhou J, Dai Y, Khandelwal M, Monjezi M, Yu Z, Qiu Y (2021a). Performance of Hybrid SCA-RF and HHO-RF Models for Predicting Backbreak in Open-Pit Mine Blasting Operations. Nat Resour Res 1-19. https://doi.org/10.1007/s11053-021-09929-y

Zhou J, Qiu Y, Armaghani DJ, Zhang W, Li C, Zhu S, Tarinejad R (2021b) Predicting TBM penetration rate in hard rock condition: A comparative study among six XGB-based metaheuristic techniques. Geosci Front 12(3), 101091. https://doi.org/10.1016/j.gsf.2020.09.020

Zhou J, Qiu Y, Zhu S, Armaghani DJ, Li C, Nguyen H, Yagiz S (2021c) Optimization of support vector machine through the use of metaheuristic algorithms in forecasting TBM advance rate. Eng Appl Artif Intel 97, p.104015. https://doi.org/10.1016/j.engappai.2020.104015