



An empirical-driven machine learning (EDML) approach to predict PPV caused by quarry blasting

Panagiotis G. Asteris¹ · Danial Jahed Armaghani²

Received: 25 September 2024 / Accepted: 6 March 2025 / Published online: 19 March 2025
© The Author(s) 2025

Abstract

Blasting in mining and quarrying serves multiple purposes but poses environmental challenges, notably generating shock-waves and vibrations through peak particle velocity (PPV) from explosions. Previous efforts to predict PPV values have relied on empirical equations using parameters such as maximum charge per delay (MC) and distance from the blast face (D). Numerous attempts have employed machine learning (ML) to estimate PPV with the same input parameters. This study introduces a novel approach called empirical-driven ML (EDML), which integrates empirical equations and their outcomes as inputs for PPV prediction. EDML leverages existing knowledge to enhance model performance, interpretability, and generalization. For the EDML approach, four empirical equations, namely USBM, CMRI, General Predictor, and Ambraseys-Hendron have been chosen based on prior research. These four empirical equations were selected based on their good performance as reported in the literature. Using these equations' PPV values as inputs, three advanced tree-based techniques (random forest, deep forest, and extreme gradient boosting) have been employed for model training. Comparison with the conventional ML approach (using only maximum charge per delay and distance from the blast face) reveals EDML's superior predictive capacity for PPV estimation. Note that the inputs of these databases were directly and indirectly extracted from MC and D with the same PPV values. The proposed EDML approach effectively integrates data-driven insights with domain expertise, improving accuracy and interpretability through the inclusion of PPV and blasting observations.

Keywords Blasting · Peak particle velocity · Empirical-driven machine learning · Traditional machine learning · Tree-based techniques

Abbreviations

AAN	Artificial neural network	CART	Classification and regression tree
ANFIS	Adaptive neuro-fuzzy inference system	CMRI	Central Mining Research Institute
ANFO	Ammonium nitrate and fuel oil	CSO	Cuckoo search optimization
B	Burden	D	Distance
BI	Blastability index	DF	Deep forest
B/S	Burden to spacing	Di	Diameter of hole
		DT	Delay time
		DTs	Decision trees
		E	Young's modulus
		EDML	Empirical-driven machine learning
		EPB	Explosive amount per blast
		f	The learned tree
		FCM	Fuzzy C-means clustering
		FFA	Firefly algorithm
		FL	Fuzzy logic
		FS	Fuzzy system
		GFNN	Generalized feed forward neural network
		GA	Genetic algorithm

✉ Danial Jahed Armaghani
danial.jahedarmaghani@uts.edu.au

Panagiotis G. Asteris
asteris@aspete.gr

¹ Computational Mechanics Laboratory, School of Pedagogical and Technological Education, Marousi, Athens 15122, Greece

² School of Civil and Environmental Engineering, University of Technology Sydney, Sydney, NSW 2007, Australia

GEP	Gene expression programming
GBT	Gradient boosted tree
GMDH	Group method of data handling
H_d	Depth of hole
ICA	Imperialist competitive algorithm
KNN	K-nearest neighbors
MC	Maximum charge per delay
ML	Machine learning
MR_r	Modified radial distance
N_h	Number of hole
ν	Poisson's ratio
θ	The feature needed to develop a tree
P_c	Protodyakonov coefficient
PF	Powder factor
PPV	Peak particle velocity
PSOGWO	Particle swarm optimization with grey wolf optimizer
PSO	Particle swarm optimization
R^2	Coefficient of determination
RF	Random forest
RQD	Rock quality designation
R_r	Radial distance
RVM	Relevance vector machine
S	Spacing
S_d	Sub-drilling
SVM	Support vector machine
SVR	Support vector regression
T	Stemming
T_c	Total charge
USBM	The United States Bureau of Mines
VOD	Velocity of detonation of explosive
V_p	P-wave velocity
X	A variable in XGBoost
XGBoost	eXtreme Gradient Boosting

Introduction

Blasting serves as a prevalent method for fragmenting rock in mining and quarrying, which is applied across diverse objectives including raw material creation, ore retrieval, and contouring. The material produced by blasting operations can be used for construction, mining, infrastructure development, and landscaping (Khotbehsara et al. 2014; Mehdizadeh et al. 2022; Mehdizadeh Miyandehi et al. 2023; Eftekhari Afzali et al. 2025). Nevertheless, these activities give rise to a range of environmental concerns and effects. Among these are noise pollution, vibrations, and the release of particulate matter. Consequently, meticulous assessment, analysis, and mitigation are imperative to curtail adverse effects on neighboring ecosystems and communities (Bakhtavar et al. 2021; Hajihassani et al. 2014; Khandelwal and Singh 2007).

Although all of aforementioned effects are impactful to the surrounding area, ground vibration and its common measurement, i.e., peak particle velocity (PPV), can be a more significant concern in mining, quarrying, and construction activities, especially in urban or very close to urban areas. When explosives are detonated in the ground, they release a tremendous amount of energy, causing shock waves to propagate through the earth (Hajihassani, Jahed Armaghani et al. 2015; Rai and Singh 2004). These shock waves create ground vibrations, which can be felt and measured at varying distances from the blast site. Some of the key factors that influence blast-induced PPV include i) explosive characteristics, blast design parameters, such as spacing, hole diameter and burden, iii) geology and rock properties, iv) distance from the blast, and v) blast timing and sequencing (Bhandari 1997; Davies et al. 1964; Iphar et al. 2008; New 1986; Zhou et al. 2020).

There are several techniques used to determine blast-induced ground vibrations such as empirical, statistical, and soft computing approaches. One of the most popular techniques in practice among them is the solution of empirical equations, originally suggested by some authors or research groups, which are typically based on two factors i.e., maximum charge per delay (MC) and distance (D) along with certain coefficients or constants. It should be mentioned that these equations have been developed for some specific geological conditions or rock-mass and because of that we should expect a high-performance prediction of PPV if a new data becomes available, generally. However, interesting results have been reported in literature (Hasanipanah et al. 2017a; Hosseini et al. 2019; Matidza et al. 2020; Yang et al. 2020) using these empirical equations which show the capability of equations in predicting PPV. On the other hand, a low performance level of empirical equations has been highlighted in some other studies (Khandelwal 2010; Khandelwal and Singh 2009; Tribe et al. 2021). In terms of statistical models, similar accuracy with empirical equations can be reported by majority of investigations (Khandelwal et al. 2017). Nevertheless, the last group of techniques (i.e., intelligent techniques) which are based on machine learning (ML) algorithms and concepts showed the best performance among the mentioned techniques. It is important to mention that the same ML models have been successfully applied in a wide range of geotechnical and geological engineering (Asteris et al. 2021, 2022; Fattahi et al. 2024; Ghanizadeh et al. 2023; He et al. 2024a; He, Armaghani, Tsoukalas, He et al. 2024a, b; Koopialipoor et al. 2022; Skentou et al. 2023; Yari et al. 2023). To predict PPV induced by blasting using ML, some researchers used only MC and D which are also the effective variables in empirical equations. Table 1 shows a summary of the most important studies along with their techniques, database size, and performance predictions.

Table 1 A summary of PPV studies considering only MC and D as inputs

Nr.	Size of Data	Technique	Input Factors	Performance Capacity (R^2)	Reference
1	44	ANFIS	MC, D	0.980	Iphar et al. (2008)
2	162	ANN	MC, D	0.940	Mohamed (2011)
3	150	ANN	MC, D	0.919	Khandelwal et al. (2011)
4	33	FL	MC, D	0.920	Fisne et al. (2011)
5	37	SVM	MC, D	0.890	Mohamadnejad et al. (2012)
6	80	SVM	MC, D	0.957	Hasanipanah et al. (2015)
7	109	ANFIS	MC, D	0.973	Armaghani et al. (2015)
8	75	ANN-KNN	MC, D	0.880	Amiri et al. (2016)
9	51	CART	MC, D	0.920	Khandelwal et al. (2017)
10	85	GA	MC, D	0.920	Hasanipanah et al. (2017)
11	81	ANFIS-PSO	MC, D	0.984	Shahnazar et al. (2019)
12	73	ICA	MC, D	0.940	Armaghani et al. (2018)
13	50	FS-ICA	MC, D	0.942	Hasanipanah et al. (2018)
14	68	ANN	MC, D	0.981	Nguyen et al. (2019)
15	152	PSO-KNN	MC, D	0.977	Bui et al. (2019)
16	117	GMDH	MC, D	0.857	Zeng et al. (2022)

Another group of researchers predicted PPV by considering some other factors, such as blasting pattern parameters, as inputs apart from MC and D. Table 2 summaries these studies with their inputs, size of data, and performance predictions. Note that the performance metric considered in

Tables 1 and 2 is the coefficient of determination (R^2). As can be seen from these two tables, apart from the size of the data, which is very varied, the performance predictions of these studies are quite similar. This may mean that those studies with only two input parameters (Table 1) are comparable in terms of performance capacity with those with more than two input parameters (Table 2). However, for those databases with more than two input parameters, some additional field works, or laboratory investigations are needed. For instance, in order to determine the Young’s modulus (E) of the rock and add it as one of the input parameters, there is a need to collect samples, transfer them to laboratory, prepare standard samples (Hatheway 2009), and conduct uniaxial compression tests. This process requires considerable time, cost, and energy (Armaghani et al. 2015). On the other hand, proposing a predictive model with the least number of influential factors (i.e., inputs) is of interest in modelling as well as practice. Practice engineers are always looking for a simple and practical way to accomplish different purposes, such as determining the PPV induced by blasting. It seems that using only two input parameters (i.e., MC and D) can be sufficient to get the preferred prediction capacity for estimating PPV values.

This study explores the effective utilization of empirical equations and their outcomes as input parameters to address environmental concerns arising from blasting-induced PPV with the help of ML. This new concept namely empirical-driven ML (EDML) which can be calculated using MC and D, seamlessly fuses empirical insights with the analytical power of ML. This harmonious integration examines prediction capacity level of PPV and checks if it is able to provide higher performance. The obtained results will be compared with a database with MC and D as inputs.

The applied ML models in this study are a powerful group of tree-based techniques characterized by decision

Table 2 A summary of PPV studies considering a various range of inputs apart from MC and D

Nr	Size of Data	Technique	Input Factors	Performance Capacity (R^2)	Reference
1	174	ANN	D, H_d , B, S, MC, VOD, V_p , BI, E, v	0.984	Khandelwal and Singh (2009)
2	29	ANN	T, D, MC, N_h	0.99	Bakhshandeh Amnieh et al. (2010)
3	20	ANN	MC, D, T_c	0.930	Monjezi et al. (2013)
4	95	ICA-ANN	V_p , B/S, T, PF, MC, D, E	0.980	Hajihassani et al. (2015)
5	166	ANFIS	B/S, T, PF, MC, D	0.960	Jahed Armaghani et al. (2015)
6	115	GEP	B, S, MC, T, D_i , H_d , N_h , PF, D	0.874	Faradonbeh and Monjezi (2017)
7	120	FCM-SVR	B, S, T, N_h , MC, D	0.853	Sheykhi et al. (2018)
8	37	GFNN	T_c , MC, D	0.954	Abbaszadeh Shahri and Asheghi (2018)
9	83	FFA-ANN	MC, D, S, B, PF	0.966	Shang et al. (2020)
10	70	GA-ANN	D, MR_p , MC, R_f	0.988	Azimi et al. (2019)
11	102	RF	PF, D, T, H_d , MC, B/S	0.90	Zhou et al. (2020)
12	118	CSO-ANN	T_c , EPB, D, S, B, PF	0.990	Bui et al. (2021)
13	93	GBT	PF, MC, RQD, D, S_d	0.893	Peng et al. (2021)
14	137	PSOGWO-RVM	MC, T_c , B, D, H_d , P_c , DT	0.971	Yu et al. (2022)
15	117	GMDH	S_d , PF, MC, D	0.887	Zeng et al. (2022)

trees (DTs). They excel in capturing complex relationships in data, handling various types of features, and are robust against overfitting, making them valuable tools for diverse prediction tasks. Tree-based techniques have been utilized by many researchers in different areas to solve engineering problems (Lu et al. 2020; Mohammed et al. 2021; Sarir et al. 2021). In particular, deep forest (DF), random forest (RF) and eXtreme Gradient Boosting (XGBoost) were applied on PPV data collected from a quarry site in Malaysia considering two different strategies namely EDML and ML. The subsequent sections of this paper are structured in the following manner:

Derived from the preceding content, the organizational framework of this paper, aside from the introductory segment, is structured as follows: In Sect. 2, an explication of the foundational aspects of the employed machine learning techniques and the EDML approach is presented. Section 3 delves into intricate particulars concerning the case study, blast parameters, and PPV values. The fourth Sect. 4 delineates the process of modeling, encompassing strategies, and performance criteria. Subsequently, following this, Sect. 5 provides an in-depth discussion of the results. Section 6 addresses limitations, challenges, and potential future directions. Finally, Sect. 7 contains the concluding remarks intended for the readers' contemplation.

Methodology

This section presents the methods and theories used in this research to predict PPV induced by blasting with more details. However, the relevant references available in literature can always help if in-depth understanding is needed.

Empirical-driven machine learning (EDML)

The EDML is an approach that combines ML techniques with principles and insights from theories, prior knowledge or empirical equations. In traditional ML, models are typically built based solely on data and patterns within that data such as those proposed models in Tables 1 and 2. However, EDML integrates existing knowledge, rules, constraints, or empirical equations into the learning process to improve model performance, interpretability, and generalization. Some key aspects of EDML are discussed as follows:

Incorporating domain knowledge Researchers or experts in a specific field often have deep knowledge about the underlying mechanisms governing the data. The EDML

takes advantage of this knowledge to design models that align with these principles.

Constraint-based modeling Incorporating domain theories allows the formulation of constraints and assumptions that the ML models should adhere to. This helps to ensure that the model process results are consistent with existing knowledge.

Sample efficiency By using prior knowledge or constraints, EDML can reduce the need for extensive data collection and make the learning process more efficient, especially in cases where data is limited.

Model validation and falsification The integration of theories or prior knowledge can also help in the validation and falsification of ML models. If a model produces results that contradict well-established theories or empirical equations, it could indicate possible errors in the model or in the underlying data.

As the use of empirical equations for predicting PPV values are well-established in both research and practice, this study aims to apply them during data preparation. In fact, the plan is to use the results of the most accurate ones as input parameters instead of using the direct use of MC and D. Of course, their results can be obtained from the same factors (i.e., MC and D). To do this, a performance investigation was conducted on the well-known and frequently-used empirical equations in the past where their performance is considered as R^2 . Table 3 shows the most well-known empirical equations to determine PPV values and their performance capacities based on R^2 . These empirical equations are Langefors-Kihlstrom, General Predictor, Indian Standard, Ghosh-Daemen Predictor, Central Mining Research Institute (CMRI), The United States Bureau of Mines (USBM), Ambraseys-Hendron which more information about them can be found in their original references (Ambraseys and Hendron 1968; Davies et al. 1964; Duvall and Petkof 1959; Ghosh and Daemen 1983; Langefors and Kihlström 1963; Pal Roy 1993; Standard 1973). Based on Table 3, some equations (e.g., CMRI) are consistently good if a new PPV database is applied and some equations (e.g., Langefors-Kihlstrom) have a wide variation for different PPV databases. Therefore, considering an average of their prediction capacities is a good idea to select the most influential ones. Obviously, we can use all of them as model inputs, however, the complexity level should be considered during model development. It is well-established that ML models with the lowest model inputs are of advantage and interest because of model simplicity and faster running time.

Table 3 Performance capacity for the previous PPV studies using empirical equations

Reference	Empirical Equation						
	Langefors-Kihlstrom	General Predictor	Indian Standard	Ghosh-Daemen Predictor	CMRI	USBM	Ambraseys-Hendron
	R^2						
Khandelwal and Singh (2009)	0.131	0.487	0.325	0.478	0.532	0.541	0.203
Khandelwal et al. (2010)	0.106	0.615	0.278	0.659	0.612	0.633	0.434
Mohamadnejad et al. (2012)	-	-	0.433	-	-	-	-
Armaghani et al. (2015)	0.315	0.831	0.349	0.834	0.827	0.836	-
Hasanipanah et al. (2015)	-	0.495	0.001	-	-	0.557	0.684
Monjezi et al. (2016)	0.642	-	0.465	-	-	0.614	0.524
Hasanipanah et al. (2017)	-	0.803	-	-	0.800	0.840	-
Khandelwal et al. (2017)	0.830	-	0.670	-	-	0.840	-
Faradonbeh and Monjezi (2017)	0.620	-	0.657	-	0.742	0.700	0.659
Taheri et al. (2017)	-	-	0.730	-	-	0.770	0.750
Azimi et al. (2019)	0.804	-	0.804	-	0.847	0.830	0.853
Xue (2019)	-	-	0.670	-	-	0.650	0.690
Hosseini et al. (2019)	0.700	0.780	0.510	-	0.750	0.750	0.770
Arthur et al. (2019)	0.614	-	0.571	-	-	0.580	0.557
Zhou et al. (2020)	0.275	0.654	0.293	0.635	0.642	0.672	-
Zhang et al. (2020a, b)	0.686	-	-	-	-	0.567	0.545
Matidza et al. (2020)	0.938	0.958	0.810	-	0.933	0.950	0.927
Arthur et al. (2020)	0.614	-	0.571	-	-	0.581	0.557
Armaghani et al. (2020)	-	-	0.837	-	-	0.870	0.867
Yang et al. (2020)	-	-	0.853	-	0.909	0.885	0.894
Tribe et al. (2021)	0.112	-	-	-	-	0.149	0.508
Average	0.528	0.703	0.546	0.652	0.759	0.691	0.651

Average R^2 values of 0.528, 0.703, 0.546, 0.652, 0.759, 0.691, and 0.651 are obtained for Langefors-Kihlstrom, General Predictor, Indian Standard, Ghosh-Daemen Predictor, CMRI, USBM, and Ambraseys-Hendron equations, respectively. As a result, USBM, CMRI, and General Predictor equations were directly selected to be used as inputs. Among Ambraseys-Hendron and Ghosh-Daemen predictors, which received very close average R^2 (0.651 and 0.652, respectively), the authors decided to use Ambraseys-Hendron because of its high utilisation rate based on Table 3. Therefore, four empirical equations namely USBM, CMRI, General Predictor and Ambraseys-Hendron were selected to play role of input parameters in this study to predict PPV values. Equations 1–4 show USBM, CMRI, General Predictor and Ambraseys-Hendron formulas, respectively for the conditions of this study.

$$PPV = k \left[\frac{D}{\sqrt{MC}} \right]^{-B}, k = 179.31, B = 1.09 \tag{1}$$

$$PPV = n + k \left[\frac{D}{\sqrt{MC}} \right]^{-1}, k = 168.91, n = 1.57 \tag{2}$$

$$PPV = kD^{-B}(MC)^A, k = 212.27, B = 1.09, A = 0.52 \tag{3}$$

$$PPV = k \left[\frac{D}{(MC^{0.333})} \right]^{-B}, k = 329.306, B = 0.970 \tag{4}$$

Deep forest (DF)

The DF model is an innovative and powerful ML approach that combines the strengths of random forests and deep learning techniques. It leverages the ensemble learning capability of random forests with the deep representation learning ability of deep neural networks (Z.-H. Zhou and Feng 2019). The DF model consists of multiple layers of random forests, each representing a distinct level of abstraction. The model begins with a shallow random forest layer that learns simple and low-level features from the input data. The outputs of this initial layer are then fed into subsequent random forest layers, which progressively learn more complex and high-level representations.

The DF model stands out due to its ability to handle both structured and unstructured data. It excels in scenarios where there is a lack of labelled training data, as it can leverage unsupervised pre-training techniques to learn useful representations from unlabeled data. Furthermore, the DF model exhibits robustness against overfitting and can effectively handle noisy or incomplete data. It achieves this by employing the randomization techniques inherent

in random forests, such as bootstrap aggregating (bagging) and feature sub-setting. These techniques help reduce the model’s sensitivity to outliers and enhance its generalization capabilities.

Conclusively, the DF model represents a novel fusion of random forests and deep learning, offering a unique blend of interpretability, robustness, and deep representation learning. Its ability to handle diverse data types and effectively deal with limited labelled data makes it a promising approach for various machine learning tasks. A simple flow-chart of DF is presented in Fig. 1.

eXtreme gradient boosting (XGBoost)

The XGBoost model is a state-of-the-art gradient-boosting ML algorithm known for its exceptional performance in various domains. It belongs to the family of boosting algorithms, which sequentially train an ensemble of weak learners to create a powerful predictive model. XGBoost leverages the concept of gradient boosting to iteratively improve the predictive accuracy of the model (Chen and Guestrin 2016). It creates a collection of DTs, with each successive tree constructed to rectify the errors of the preceding ones. This iterative approach facilitates the model in grasping intricate connections and effectively encapsulating nonlinear associations within the dataset.

One of the key features of XGBoost is its ability to optimize a specific objective function during the boosting process. It accomplishes this by utilizing a combination of gradient descent and approximate greedy algorithm techniques. The objective function incorporates both a loss

function, which measures the model’s performance on the training data, and a regularization term, which helps prevent overfitting. Additionally, XGBoost supports various advanced features, such as handling missing values within the data, which it accomplishes by learning a direction for missing values during training.

Conclusively, the XGBoost model represents a powerful gradient boosting algorithm that combines iterative ensemble learning, customizable objective functions, and advanced regularization techniques. Its exceptional performance, versatility, and scalability have established it as a prominent machine learning technique. A simple flowchart of XGBoost is presented in Fig. 2.

Random forest (RF)

The RF is a versatile and powerful ML algorithm widely used for both classification and regression tasks (Han et al. 2020; He et al. 2023; Ye et al. 2021; H. Zhang et al. 2020a, b). It falls under the ensemble learning category, which combines the predictions of multiple individual models to create a stronger and more robust final model. The algorithm was introduced by Breiman (Breiman 2001) aiming to address the limitations of single DTs and enhance their predictive accuracy. The RF builds upon the concept of DTs, which recursively partition the input data based on feature values to make predictions. However, DTs are prone to overfitting, where they capture noise in the data and perform poorly on unseen examples. The RF overcomes this by generating an ensemble of DTs and introducing randomness in their construction.

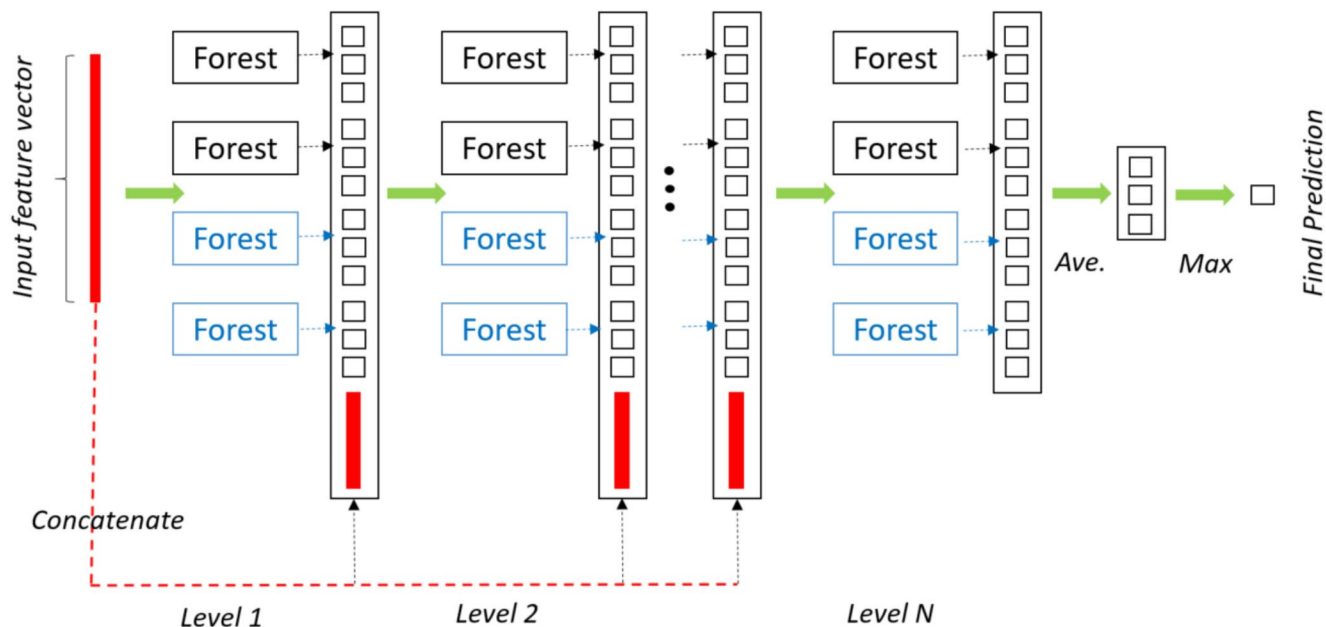


Fig. 1 A simple presentation of DF flowchart

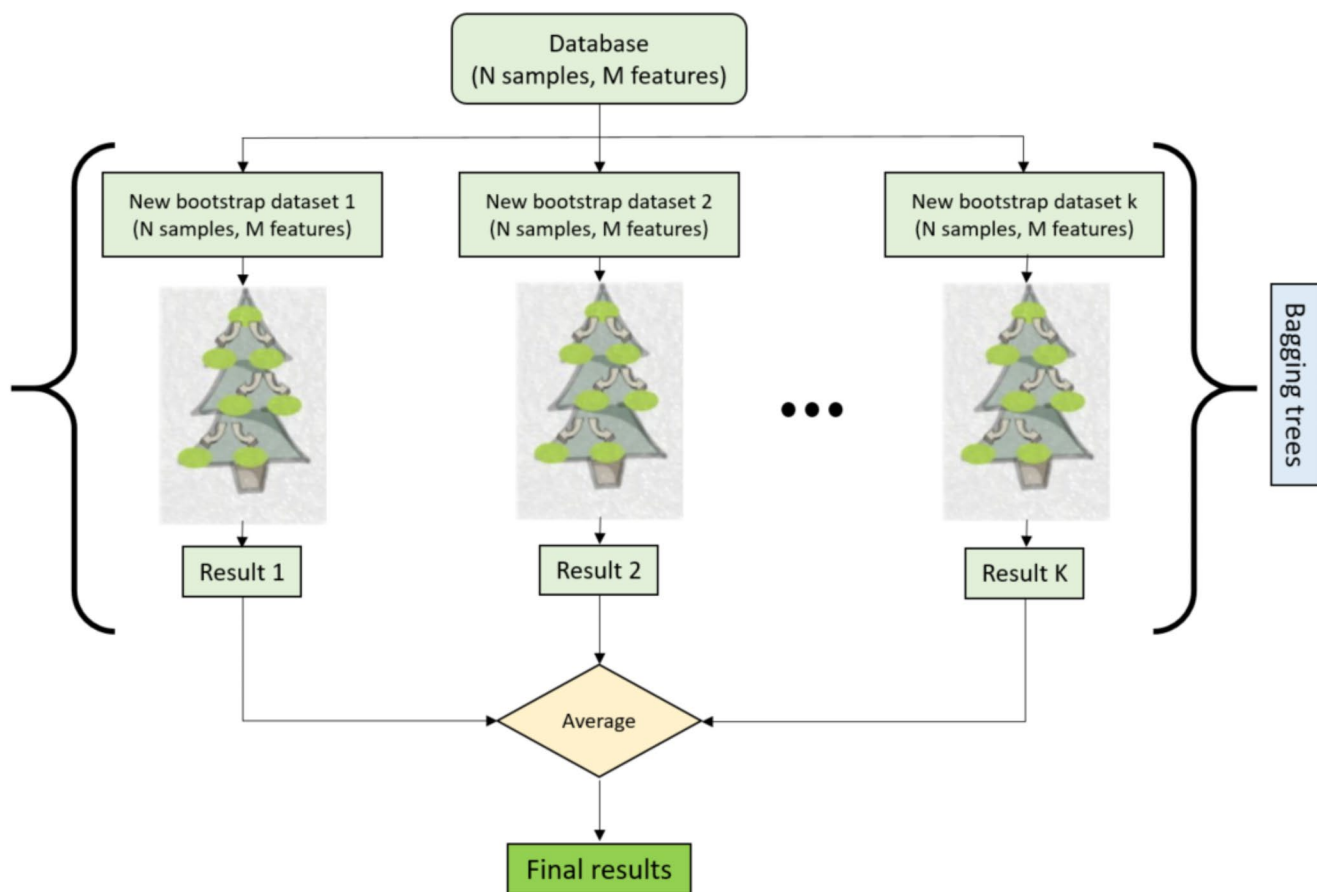


Fig. 3 A simple presentation of RF flowchart

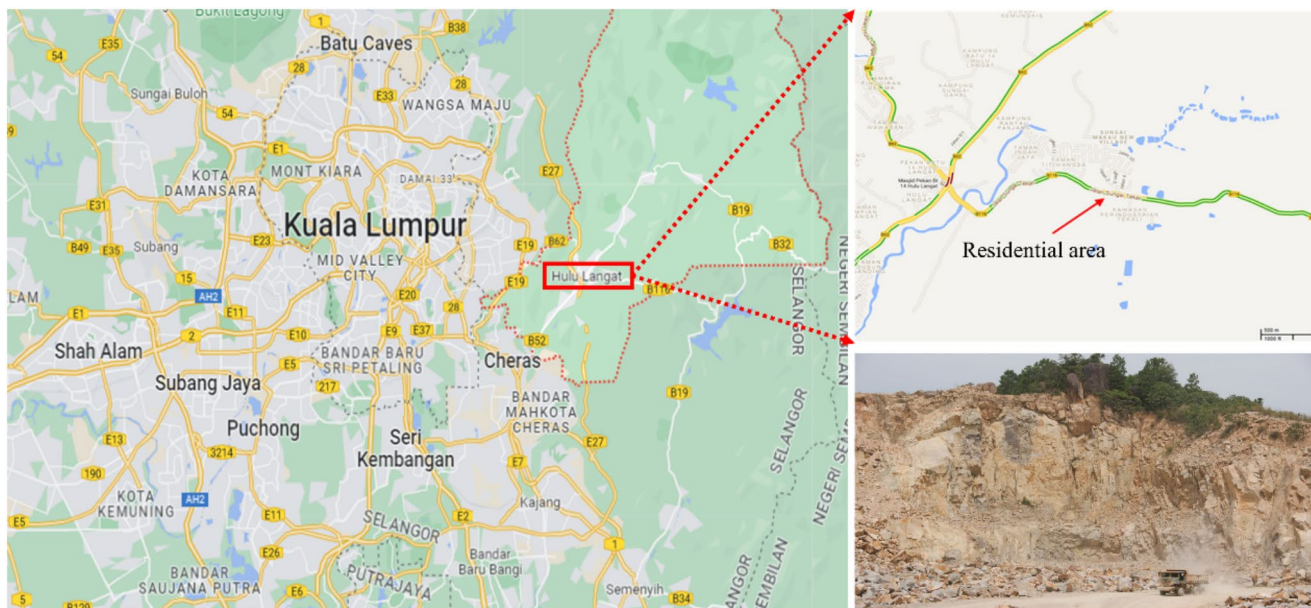


Fig. 4 Location and view of Langat granite quarry site

Table 4 Statistical information related to ML strategy

Statistical information	MC (kg)	D (m)	PPV (mm/s)
Max.	101.55	700.00	11.20
Min.	53.40	205.00	1.10
Mean	78.14	488.99	4.13
Std.	15.13	149.84	2.68

Table 5 Statistical information related to EDML strategy

Statistical information	USBM (mm/s)	General Predictor (mm/s)	CMRI (mm/s)	Ambraseys–Hendron (mm/s)	PPV (mm/s)
Max.	4.98	5.30	7.87	6.86	11.20
Min.	1.32	1.41	3.43	2.14	1.10
Mean	2.57	2.73	4.99	3.69	4.13
Std.	1.08	1.14	1.31	1.35	2.68

from the active free face varied from 205 to 700 m, ensuring effective data collection in controlled circumstances.

Model development

Data design and splitting strategy

This study considers two scenarios, namely ML and EDML, for implementing the PPV prediction task. The ML scenario refers to the condition that there are two common parameters, i.e., MC and D, as inputs to predict PPV values, while in the EDML scenario, the results of four well-known empirical equations, i.e., USBM, CMRI, General Predictor, and Ambraseys–Hendron, are used as inputs to predict PPV values. Actually, the latter also requires the same MC and D as predictors to proceed. The primary distinction between these two scenarios is just to use empirical equations to convert inputs to a form of EDML approach, and basically the requirement for both scenarios is the same. The database

(nominated as the ML) in Table 4 encompasses three key variables: MC, D, and PPV. The former two variables serve as inputs for the predictive models, while the latter represents the output to be estimated by the models.

Subsequently, the EDML database employed a process that involved the application of four empirical formulas (Eqs. 1–4). These formulas utilise MC and D as inputs, enabling the calculation of four unique PPV values. Each PPV value corresponded to the outcome obtained by employing a specific empirical formula. By leveraging this strategy, the EDML dataset was constructed, featuring multiple inputs that encapsulated the calculated PPV results derived from each of the four formulas, as shown in Table 5. This strategy facilitates an exploration of the influence of different formula-derived PPV values on the model’s predictive performance.

After the preparation of these two databases, the essential step is to select the appropriate data for model training—data splitting. Data splitting is a fundamental practice in ML and statistical analysis aimed at ensuring robust model evaluation and generalisation. This study utilises two distinct databases (i.e., ML and EDML) to train the PPV prediction models. Both databases were divided into training, testing, and validation subsets, comprising 79, 20, and 10 samples, respectively, following a consistent splitting approach. Table 6 presents the statistical information for these three sets, including the maximum, minimum, mean, and standard deviation of each variable. Notably, a common criterion was enforced during the division process: ensuring that the output variables remained identical across both databases. Despite variations in the input variables, this harmonisation of outputs ensured a consistent target variable for model training and evaluation. By aligning the output values, we aimed to assess and compare the performance of models on different input datasets without introducing confounding factors associated with varying outputs. As it

Table 6 Statistical information of training, testing, and validation sets

Set	Statistical information	EDML database				ML database			
		Input				Output	Input		Output
		USBM	General Predictor	CMRI	Ambraseys–Hendron	PPV	MC	D	PPV
Train	Max.	4.98	5.30	7.87	6.86	11.20	101.55	700	11.20
	Min.	1.32	1.41	3.43	2.14	1.10	53.40	210	1.10
	Mean	2.38	2.53	4.76	3.47	3.54	76.89	511.90	3.54
	Std.	0.95	1.01	1.16	1.21	2.33	14.35	146.19	2.33
Test	Max.	4.68	4.94	7.53	6.66	9.50	101	650	9.50
	Min.	1.42	1.51	3.56	2.29	1.20	53.40	210	1.20
	Mean	2.89	3.06	5.38	4.06	5.11	81.89	449	5.11
	Std.	1.17	1.24	1.43	1.44	2.79	16.01	141.72	2.79
Validation	Max.	4.98	5.30	7.87	6.86	11	98.49	600	11
	Min.	1.51	1.62	3.68	2.45	2.10	53.40	205	2.10
	Mean	3.42	3.63	6.01	4.74	6.84	80.53	388	6.84
	Std.	1.35	1.43	1.63	1.70	3.06	19.21	149.34	3.06

is clear from Table 6, each set in different databases covers and presents the main range of the database, which means the results obtained from modelling are reliable enough and represent the behaviour of the whole database. It is worth mentioning that the validation sets will not be used in the model development of ML and EDML strategies; they are just for representing the capacity of the developed models in real-world conditions.

Model training strategy

The present study employs three state-of-the-art tree-based models, namely DF, XGBoost, and RF, to tackle the prediction task at hand. DF is an ensemble learning technique that combines the power of DTs and RFs with the flexibility of deep learning (Z.-H. Zhou and Feng 2019). It operates by hierarchically partitioning the input space and fusing the outputs of multiple tree models. XGBoost is a powerful algorithm that utilises gradient boosting principles to construct an ensemble of weak learners, such as DTs (Chen and Guestrin 2016). It focuses on iteratively optimising a user-defined loss function, leading to enhanced model performance. RF (Breiman 2001) harnesses the strength of an ensemble of DTs through bootstrapping and feature randomness. It aggregates the predictions of individual trees to arrive at a final prediction.

The best design of these techniques requires the determination of their optimal hyperparameters. In this way, the maximum prediction capacity can be obtained. In the model training process, we employed the grid search approach to identify the optimal hyperparameters for these models. Grid search systematically explores a predefined hyperparameter space, evaluating each combination. This process allows for identifying hyperparameter values that maximise model performance and generalisation ability, ensuring robustness in our predictive models.

To further enhance model robustness and generalization, this study implemented a 10-fold cross-validation approach during model training. Cross-validation is a widely used resampling technique that partitions the dataset into ten subsets (folds), where nine folds are used for training and one for validation in each iteration. This process is repeated ten times, and the final model performance is evaluated based on the averaged results from all iterations. This method helps prevent overfitting and ensures that the trained models generalize well to unseen data.

Model evaluation criterion

This study used three evaluation metrics to examine the model's performance, including R^2 , root mean squared error

(RMSE), and mean absolute percentage error (MAPE). The following equations are used to compute these metrics.

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (5)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (6)$$

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (7)$$

where y_i represents the measured PPV, \hat{y}_i represents the predicted PPV by tree-based model, \bar{y} represents the average of y_i , and N represents the number of samples. The ranges of R^2 , RMSE, and MAPE are $[0, 1]$, $[0, \infty]$, and $[0\%, 100\%]$, respectively. An excellent model should possess a large value of R^2 (close to 1), small values of RMSE (close to 0), and MAPE (close to 0%).

Results and discussion

Determination of the model's hyperparameters

The tree-based models, i.e., DF, XGBoost, and RF, were trained both on EDML and ML databases. For each dataset, the particular procedure of model training is: the training set (79 samples) was used to train all tree-based models; the testing set (20 samples) was used to determine the optimal hyperparameters of each model in the training process; and the validation set (10 samples) was used to evaluate the model's performance on unknown data. Additionally, the R^2 value serves as an evaluation criterion to determine the optimal hyperparameters of tree-based models. The designed searching domains of hyperparameters of tree-based models are listed below (Guo et al. 2020; He et al. 2022, 2023; Yari et al. 2023b).

- For the DF model, the hyperparameter $n_estimators$ increased from 2 to 16 with 2 increments each time; the hyperparameter n_trees increased from 50 to 200 with 50 increments each time.
- For the XGBoost model, the hyperparameter $n_estimators$ increased from 25 to 200 with 25 increments each time; the hyperparameter $learning_rate$ increased from 0.01 to 0.30 with 0.05 increment each time (except the first value).
- For the RF model, the hyperparameter $n_estimators$ increased from 25 to 200 with 25 increments each time;

the hyperparameter *max_depth* increased from 4 to 10 with 1 increment each time.

Figure 5 shows the results of training tree-based models on EDML and ML databases. Taking the DF model as an example, it achieved the highest R^2 value on the EDML database when hyperparameters were: $n_estimators=12$ and $n_trees=125$; it achieved the highest R^2 value on the ML database when hyperparameters were: $n_estimators=12$ and $n_trees=50$. Similarly, we can also determine the optimal hyperparameters of the XGBoost and RF models, which are summarised in Table 7. For the DF model, increasing the number of estimators and trees initially improves performance, but excessive values may lead to overfitting. Similarly, for XGBoost, the optimal learning rate is identified as 0.05, balancing convergence speed and accuracy. The RF model demonstrates consistent performance gains with an increasing number of estimators up to a point, after which improvements plateau.

Performance evaluation

After determining the hyperparameters in each model, which was conducted on the training databases, it is time to evaluate all tree-based models using the results of PPV values in their testing and validation phases. This is a common practice since in these phases, the output (i.e., PPV) is not provided and the results are a real evaluation of the model's capacities. Figure 6 shows the comparison of measured PPV values and predicted PPV values by tree-based models, both applied to testing and validation sets. Table 8 further demonstrates the quantitative assessment of tree-based models on testing and validation sets according to the evaluation metrics R^2 , RMSE, and MAPE. Taking the DF model as an example, we found that DF models trained on the EDML database significantly outperformed DF models trained on the ML database—both for testing and validation sets. A similar conclusion holds for the XGBoost model. As for the RF model, the models trained on the EDML database showed slightly better performance than those trained on the ML database. Consequently, it is clearly confirmed that using the EDML strategy can achieve better performance than using the ML strategy for PPV prediction. In other words, prediction of real PPV values using results of PPV values obtained by empirical equations had a higher performance capacity compared to using only MC and D values.

As the aim is to propose a model with higher performance capacity and generalisation, a more detailed discussion of the obtained results can give a better understanding of the advantages of the EDML strategy. Considering the average performance condition of both strategies, which can be calculated from Table 8, values of (0.790 and 0.732) and

(1.24 and 1.40) are obtained for R^2 and RMSE of the EDML and ML strategies, respectively. This clearly shows the better prediction capacity of the EDML strategy in estimating PPV values resulting from blasting operations. This is easily achievable just by calculating and using the empirical equations of PPV instead of the direct use of MC and D, which are inputs to these equations. In this way, apart from the appreciation of utilising empirical equations and previous well-known and established theories and efforts, a higher range of prediction performance can be obtained. This is of special interest where the mine or quarry site is very close to residential areas or an area of construction development.

Moreover, this study utilises the Taylor diagram to characterise the differences in the model's performance. The Taylor diagram is a graphical tool used to assess the performance of ML models. It provides a visualisation of multiple evaluation metrics, including the correlation coefficient, centred RMSE, and standard deviation. The Taylor diagram consists of concentric circles and radial lines, and each point on the diagram represents a model. Figure 7 (featured by the Taylor diagram) allows for assessing the dissimilarity of the previously mentioned tree-based models. The coloured circle and black star denote the model's metrics and the referenced metrics, respectively. A well-performing model would possess a high correlation coefficient, a low centred RMSE, and a small difference between its standard deviation and the reference's standard deviation.

Regarding models' performance on the testing set (Fig. 7a), the DF (EDML) model performed a small difference between its standard deviation and the reference's standard deviation compared with the DF (ML); the XGBoost (EDML) model performed a higher correlation coefficient and a lower centred RMSE compared with the XGBoost (ML); and the RF (EDML) model performed a slightly smaller difference between its standard deviation and the reference's standard deviation compared with the RF (ML) model. Regarding models' performance on the validation set (Fig. 7b), DF (EDML) model performed a small difference between its standard deviation and the reference's standard deviation and a lower centred RMSE compared with DF (ML); XGBoost (EDML) model performed a higher correlation coefficient, a lower centred RMSE, and a small difference between its standard deviation and the reference's standard deviation compared with XGBoost (ML); RF (EDML) model performed slightly higher correlation and a lower root mean square error compared with RF (ML) model.

To sum up, all the above analysis reaches a consensus: the tree-based models trained on the EDML database show better predictive ability than those trained on the ML database. Thus, we conclude that integrating empirical knowledge and equations into ML models is a promising way to achieve a

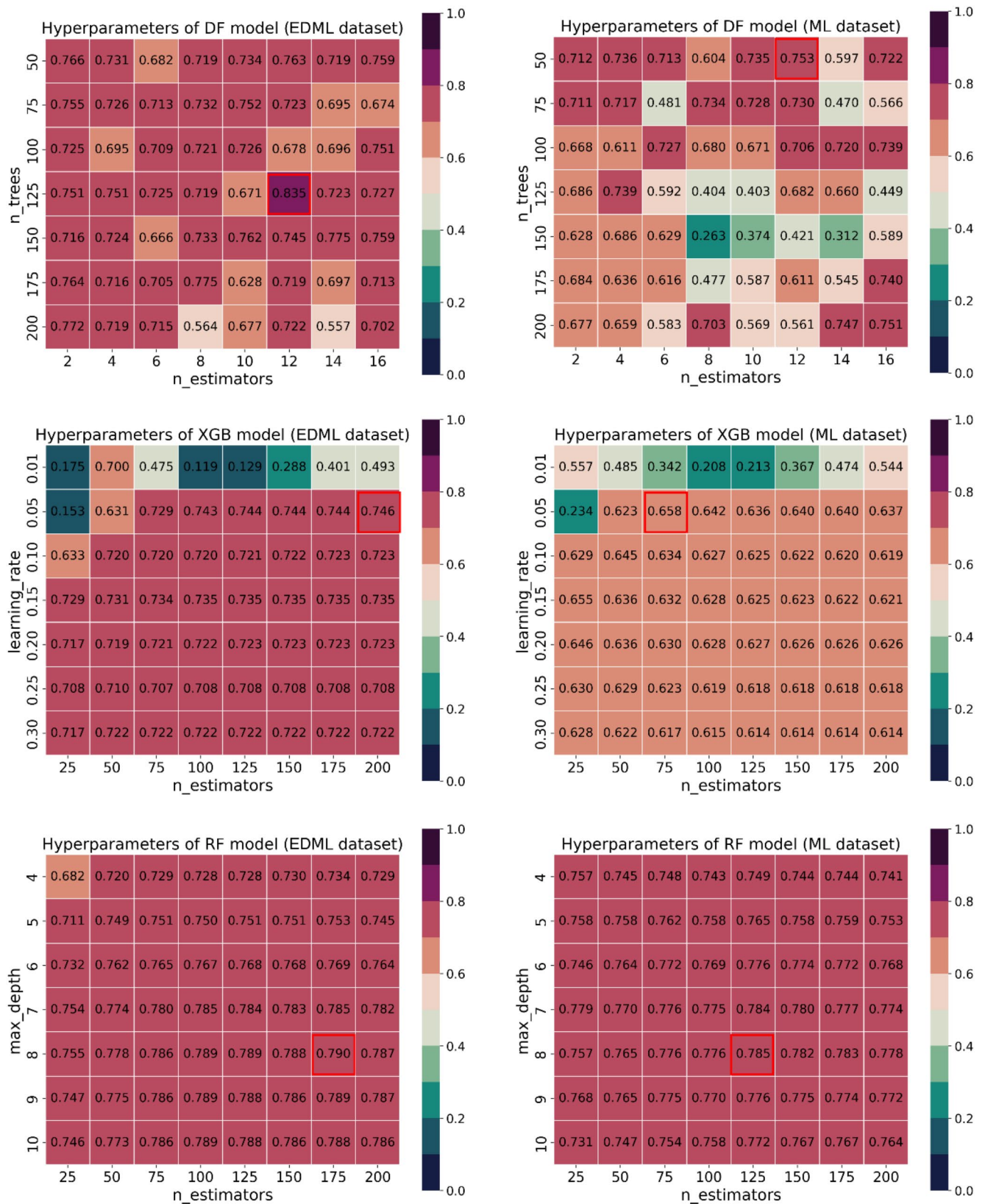
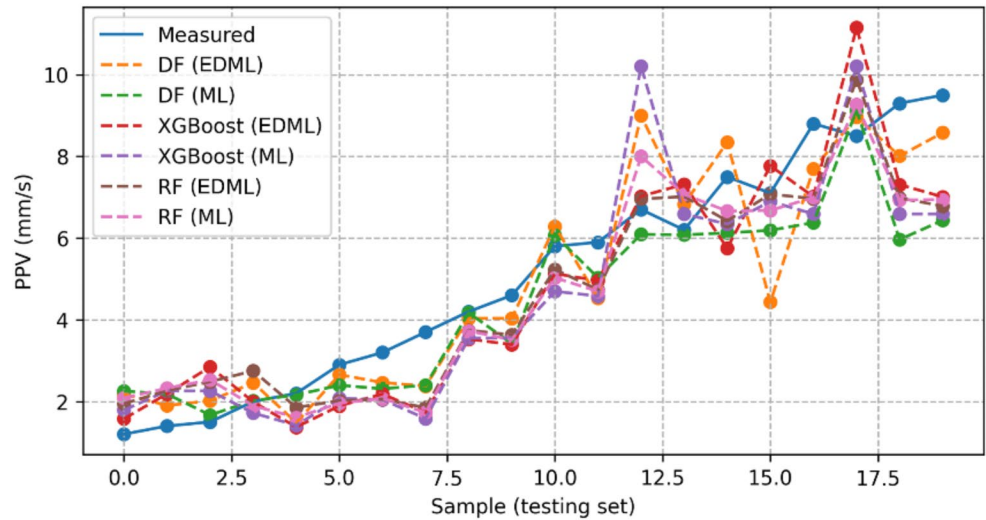


Fig. 5 Determination of hyperparameters of tree-based models on EDML and ML strategies

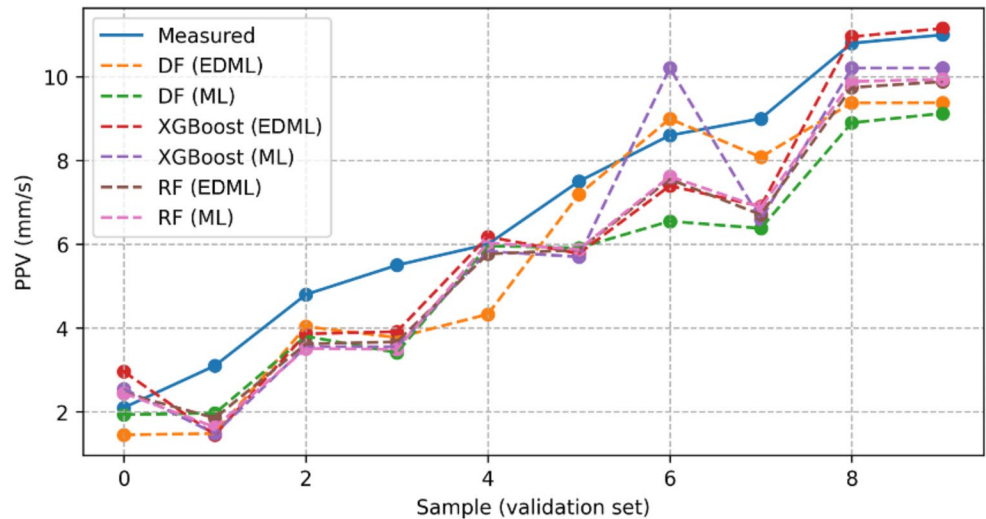
Table 7 Optimal hyperparameters of tree-based models trained on EDML and ML strategies

Strategy	DF model		XGBoost model		RF model	
	n_estimators	n_trees	n_estimators	learning_rate	n_estimators	max_depth
EDML	12	125	200	0.05	175	8
ML	12	50	75	0.05	125	8

Fig. 6 Measured PPV versus predicted PPV by tree-based models; (a) testing set; (b) validation set



(a)



(b)

Table 8 Performance evaluation of tree-based models on testing and validation sets

EDML Strategy	Testing set			Validation set		
	R ²	RMSE	MAPE	R ²	RMSE	MAPE
DF	0.835	1.106	22.881%	0.820	1.231	20.519%
XGBoost	0.746	1.372	27.985%	0.813	1.255	20.859%
RF	0.790	1.249	27.140%	0.788	1.337	19.796%
ML Strategy	R ²	RMSE	MAPE	R ²	RMSE	MAPE
DF	0.753	1.353	21.460%	0.675	1.655	21.249%
XGBoost	0.658	1.593	29.699%	0.753	1.442	21.987%
RF	0.785	1.263	28.159%	0.786	1.343	20.251%

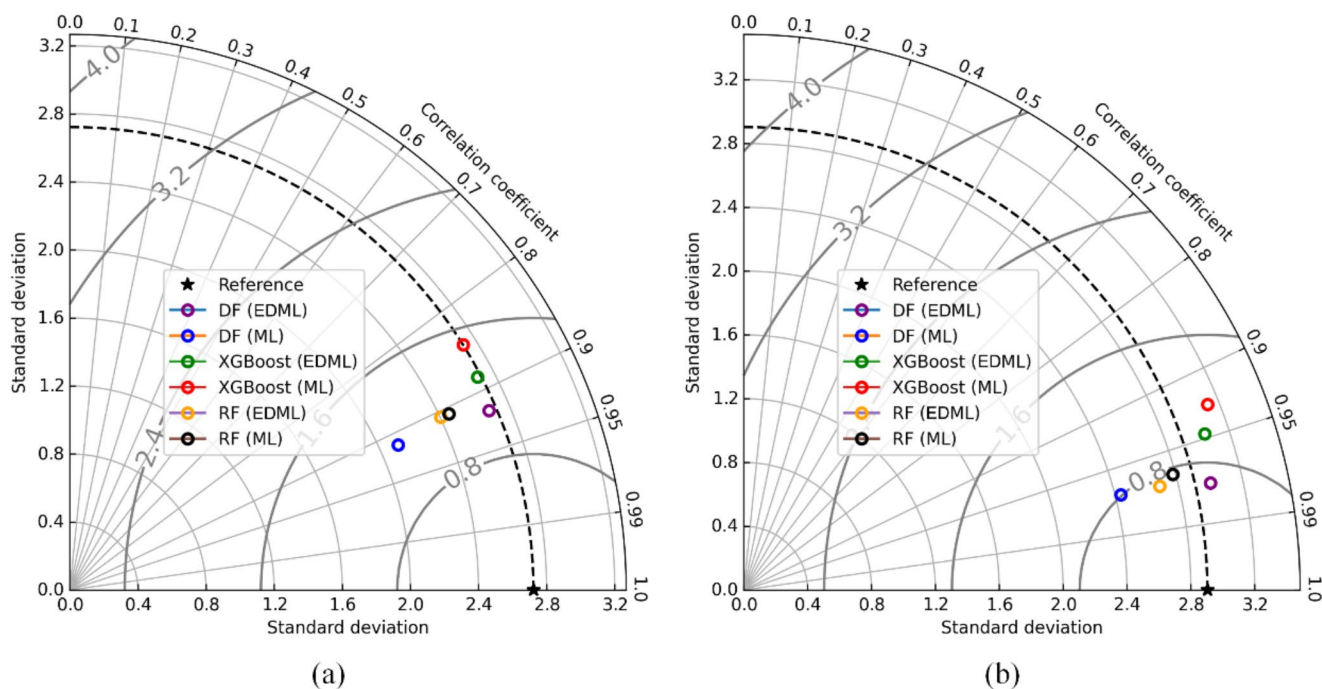


Fig. 7 Taylor diagram of tree-based model's performance on (a) testing set and (b) validation set

higher prediction level. By leveraging empirical equations within a ML framework, this model enables engineers to predict PPV more accurately, thereby reducing the risk of excessive ground vibrations that could cause structural damage to nearby buildings or infrastructure. The model's enhanced predictive capability allows for better decision-making in blast design, optimizing explosive usage while ensuring compliance with environmental safety regulations.

Limitations and future directions

While the collection of data from blasting operations offers significant insights into safety, efficiency, and environmental impacts, it comes with inherent limitations and challenges. These challenges encompass:

- I. **Data Quality and Consistency:** Ensuring the reliability and uniformity of collected data poses a challenge, affecting the accuracy of subsequent analyses.
- II. **Data Interpretation:** Interpreting data from diverse sources and formats demands careful consideration, preventing misinterpretations and inaccuracies.
- III. **Environmental Variability:** Environmental conditions vary across blasting operations, influencing data relevance and comparability.
- IV. **Resource Intensiveness:** The extensive resources required for data collection, processing, and analysis can be prohibitive.

V. **Complexity of Blasting Processes:** The intricate nature of blasting processes introduces complexity into data collection and interpretation.

As evident in Tables 1 and 2, prior studies have considered a range of 20 to 174 blasting operations. This study utilized data from 109 operations, surpassing the average number utilized for developing ML models in the literature. Nonetheless, broader application of the proposed EDML approach necessitates an expanded database encompassing increased entries and a diverse array of geological conditions and rock types. Furthermore, this approach can be extended to address other blasting-related environmental concerns, such as flyrock and air-overpressure, by combining empirical equations with ML techniques. Evaluating the significance of empirical formulas could be achieved through sensitivity analysis or parametric studies in forthcoming research. This method offers a more scientifically grounded approach for their selection, moving beyond the reliance solely on average R^2 calculation (as performed in this study) for inclusion as input parameters.

Additionally, exploring alternative ML techniques using the same dataset may yield enhanced performance, particularly hybrid intelligent models that incorporate optimization algorithms. These algorithms can autonomously identify optimal hyperparameters for base models. The primary objective of this study is to introduce the novel EDML approach to the domain of blasting environmental concerns. Naturally, other ML-driven methodologies such as risk-ML

and classification hold potential for application within the same research domain.

Conclusion

The feasibility of implementing the EDML approach in practical blasting operations lies in its ability to seamlessly integrate with existing engineering workflows. By leveraging well-established empirical equations as inputs, the approach requires no additional complex data collection beyond standard blast design parameters, making it a practical tool for mining and geotechnical engineers. The research yields the following conclusions:

- By overviewing the previous studies, four empirical equations i.e., USBM, CMRI, General Predictor and Ambraseys-Hendron were found to be the most impactful and effective in predicting PPV induced by blasting.
- The EDML approach has demonstrated successful application in a quarry case study regarding prediction of PPV while the same parameters (MC and D) are needed for this approach compared to empirical equations.
- The average performance of two strategies, EDML and ML, was obtained with values for R^2 and RMSE as (0.790 and 0.732) for EDML, and (1.24 and 1.40) for ML, respectively. These values indicate that the EDML strategy has a stronger predictive ability in estimating PPV values from blasting operations compared to the ML strategy.
- The proposed EDML combines data-driven insights with domain knowledge, enhancing predictive accuracy and interpretability. It leverages PPV and blasting observations to refine ML models, making it effective for complex systems where empirical equations might fall short.
- The validation part of the analysis showed the high practical aspects of the proposed models, considering the point that some empirical equations are normally checked in mines or quarries before blasting operations. Therefore, the ML part should be implemented by the geotechnical or mining engineers before the operations to get more accurate results for identifying the blast safety zone.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions

Declarations

Conflict of interest The authors declare that there is no conflict of interest in relation to this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abbaszadeh Shahri A, Asheghi R (2018) Optimized developed artificial neural network-based models to predict the blast-induced ground vibration. *Innovative Infrastructure Solutions* 3(1):34
- Ambraseys NR, Hendron AJ (1968) Dynamic behaviour of rock masses. *Rock mechanics in engineering practice*. J. Wiley & Sons
- Amiri M, Amnieh B, Hasanipanah H, M., Mohammad Khanli L (2016) A new combination of artificial neural network and K-nearest neighbors models to predict blast-induced ground vibration and air-overpressure. *Engineering With Computers* 32:631–644
- Amnieh HB, Mozdianfard MR, Siamaki A (2010) Predicting of blasting vibrations in Sarcheshmeh copper mine by neural network. *Saf Sci* 48(3):319–325
- Armaghani DJ, Mohamad ET, Momeni E, Narayanasamy MS (2015a) An adaptive neuro-fuzzy inference system for predicting unconfined compressive strength and Young's modulus: a study on main range granite. *Bull Eng Geol Environ* 74(4):1301–1319
- Armaghani DJ, Momeni E, Abad SVANK, Khandelwal M (2015b) Feasibility of ANFIS model for prediction of ground vibrations resulting from quarry blasting. *Environ Earth Sci* 74(4):2845–2860. <https://doi.org/10.1007/s12665-015-4305-y>
- Armaghani DJ, Hasanipanah M, Amnieh HB, Mohamad ET (2018) Feasibility of ICA in approximating ground vibration resulting from mine blasting. *Neural Comput Appl* 29(9):457–465. <https://doi.org/10.1007/s00521-016-2577-0>
- Arthur CK, Temeng VA, Ziggah YY (2019) Soft computing-based technique as a predictive tool to estimate blast-induced ground vibration. *J Sustainable Min* 18(4):287–296
- Arthur CK, Temeng VA, Ziggah YY (2020) Novel approach to predicting blast-induced ground vibration using Gaussian process regression. *Engineering With Computers* 36(1):29–42
- Asteris PG, Mamou A, Hajihassani M, Hasanipanah M, Koopialipoor M, Le T-T, Kardani N, Armaghani DJ (2021) Soft computing based closed form equations correlating L and N-type Schmidt hammer rebound numbers of rocks. *Transp Geotechnics* 29:100588
- Asteris PG, Rizal FIM, Koopialipoor M, Roussis PC, Ferentinou M, Armaghani DJ, Gordan B (2022) Slope stability classification under seismic conditions using several Tree-Based intelligent techniques. *Appl Sci* 12(3):1753
- Azimi Y, Khoshrou SH, Osanloo M (2019) Prediction of blast induced ground vibration (BIGV) of quarry mining using hybrid genetic algorithm optimized artificial neural network. *Measurement* 147:106874
- Bakhtavar E, Hosseini S, Hewage K, Sadiq R (2021) Air pollution risk assessment using a hybrid fuzzy intelligent probability-based approach: mine blasting dust impacts. *Nat Resour Res* 30:2607–2627

- Bhandari S (1997) Engineering rock blasting operations. AA Balkema, 215
- Breiman L (2001) Random forests. 45:5–32
- Bui X-N, Jaronpattanon P, Nguyen H, Tran Q-H, Long NQ (2019) A novel hybrid model for predicting blast-induced ground vibration based on k-nearest neighbors and particle swarm optimization. *Sci Rep* 9(1):13971
- Bui X-N, Nguyen H, Tran Q-H, Nguyen D-A, Bui H-B (2021) Predicting ground vibrations due to mine blasting using a novel artificial neural network-based cuckoo search optimization. *Nat Resour Res* 30:2663–2685
- Chen T, Guestrin C (2016) XGBoost: A scalable tree boosting system. *Proc ACM SIGKDD Int Conf Knowl Discovery Data Min* 13–17–Aug:785–794. <https://doi.org/10.1145/2939672.2939785>
- Davies B, Farmer IW, Attewell PB (1964) Ground vibration from shallow sub-surface blasts. *Engineer* 217(5644):553–559
- Duvall WI, Petkof B (1959) Spherical propagation of explosion-generated strain pulses in rock (Issues 5481–5485). US Department of the Interior, Bureau of Mines
- Eftekhar Afzali S, Ghasemi M, Rahimiratki A, Mehdizadeh B, Yousefieh N, Asgharnia M (2025) Compaction and compression behavior of waste materials and fiber-reinforced cement-treated sand. *J Struct Des Constr Pract* 30(2):04025007
- Faradonbeh RS, Monjezi M (2017) Prediction and minimization of blast-induced ground vibration using two robust meta-heuristic algorithms. *Engineering With Computers* 33:835–851
- Fattahi H, Ghaedi H, Malekmahmoodi F, Armaghani DJ (2024) Optimizing pile bearing capacity prediction: insights from dynamic testing and smart algorithms in geotechnical engineering. *Measurement* 230:114563
- Fişne A, Kuzu C, Hüdaverdi T (2011) Prediction of environmental impacts of quarry blasting operation using fuzzy logic. *Environ Monit Assess* 174:461–470
- Ghanizadeh AR, Ghanizadeh A, Asteris PG, Fakharian P, Armaghani DJ (2023) Developing bearing capacity model for geogrid-reinforced stone columns improved soft clay utilizing MARS-EBS hybrid method. *Transp Geotechnics* 38:100906
- Ghosh A, Daemen JJK (1983) A simple new blast vibration predictor (based on wave propagation laws). *ARMA US Rock Mechanics/ Geomechanics Symposium, ARMA-83*
- Guo M, Liang S, Zhao L, Wang P (2020) Transportation mode recognition with deep forest based on GPS data. *IEEE Access* 8:150891–150901
- Hajihassani M, Jahed Armaghani D, Sohaei H, Tonnizam Mohamad E, Marto A (2014) Prediction of airblast-overpressure induced by blasting using a hybrid artificial neural network and particle swarm optimization. *Appl Acoust* 80:57–67. <https://doi.org/10.1016/j.apacoust.2014.01.005>
- Hajihassani M, Jahed Armaghani D, Marto A, Tonnizam Mohamad E (2015a) Vibrations Au Sol prédiction Dans quarry dynamitage à Travers Un Réseau neural artificiel optimisé par Une concurrence impérialiste algorithm. *Bull Eng Geol Environ* 74:873–886
- Hajihassani M, Jahed Armaghani D, Monjezi M, Mohamad ET, Marto A (2015b) Blast-induced air and ground vibration prediction: a particle swarm optimization-based artificial neural network approach. *Environ Earth Sci* 74:2799–2817
- Han H, Jahed Armaghani D, Tarinejad R, Zhou J, Tahir MM (2020) Random forest and bayesian network techniques for probabilistic prediction of flyrock induced by blasting in quarry sites. *Nat Resour Res* 29(2):655–667. <https://doi.org/10.1007/s11053-019-09611-4>
- Harandizadeh H, Jahed Armaghani D, Khari M (2019) A new development of ANFIS–GMDH optimized by PSO to predict pile bearing capacity based on experimental datasets. *Engineering With Computers* 37(1):685–700. <https://doi.org/10.1007/s00366-019-00849-3>
- Hasanipanah M, Monjezi M, Shahnazar A, Jahed Armaghani D, Farazmand A (2015) Feasibility of indirect determination of blast induced ground vibration based on support vector machine. *Measurement: J Int Meas Confederation* 75:289–297. <https://doi.org/10.1016/j.measurement.2015.07.019>
- Hasanipanah M, Faradonbeh RS, Amnieh HB, Armaghani DJ, Monjezi M (2017a) Forecasting blast-induced ground vibration developing a CART model. *Engineering With Computers* 33:307–316
- Hasanipanah M, Golzar SB, Larki IA, Maryaki MY, Ghahremanians T (2017b) Estimation of blast-induced ground vibration through a soft computing framework. *Engineering With Computers* 33:951–959
- Hasanipanah M, Bakhshandeh Amnieh H, Khamesi H, Jahed Armaghani D, Bagheri Golzar S, Shahnazar A (2018) Prediction of an environmental issue of mine blasting: an imperialistic competitive algorithm-based fuzzy system. *Int J Environ Sci Technol* 15:551–560
- Hatheway AW (2009) The complete ISRM suggested methods for rock characterization, testing and monitoring: 1974–2006. Association of Environmental & Engineering Geologists
- He B, Lai SH, Mohammed AS, Sabri MMS, Ulrikh DV (2022) Estimation of blast-induced peak particle velocity through the improved weighted random forest technique. *Appl Sci* 12(10):5019
- He B, Armaghani DJ, Lai SH (2023) Assessment of tunnel blasting-induced overbreak: A novel metaheuristic-based random forest approach. *Tunn Undergr Space Technol* 133:104979
- He B, Armaghani DJ, Lai SH, He X, Asteris PG, Sheng D (2024a) A deep dive into tunnel blasting studies between 2000 and 2023—A systematic review. *Tunn Undergr Space Technol* 147:105727
- He B, Armaghani DJ, Tsoukalas MZ, Qi C, Bhatwadekar RM, Asteris PG (2024b) A case study of resilient modulus prediction leveraging an explainable metaheuristic-based XGBoost. *Transp Geotechnics*, 101216
- 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
- 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* 56:97–107
- Jahed Armaghani D, Hajihassani M, Monjezi M, Mohamad ET, Marto A, Moghaddam MR (2015) Application of two intelligent systems in predicting environmental impacts of quarry blasting. *Arab J Geosci* 8(11):9647–9665. <https://doi.org/10.1007/s12517-015-1908-2>
- Jahed Armaghani D, Hasanipanah M, Bakhshandeh Amnieh H, Tien Bui D, Mehrahi P, Khorami M (2020) Development of a novel hybrid intelligent model for solving engineering problems using GS-GMDH algorithm. *Engineering With Computers* 36:1379–1391
- Khandelwal M (2010) Evaluation and prediction of blast-induced ground vibration using support vector machine. *Int J Rock Mech Min Sci* 47(3):509–516
- Khandelwal M, Singh TN (2007) Evaluation of blast-induced ground vibration predictors. *Soil Dyn Earthq Eng* 27(2):116–125
- Khandelwal M, Singh TN (2009) Prediction of blast-induced ground vibration using artificial neural network. *Int J Rock Mech Min Sci* 46(7):1214–1222
- Khandelwal M, Kumar L, D., Yellishetty M (2011) Application of soft computing to predict blast-induced ground vibration. *Engineering With Computers* 27:117–125
- Khandelwal M, Armaghani DJ, Faradonbeh RS, Yellishetty M, Majid MZA, Monjezi M (2017) Classification and regression tree technique in estimating peak particle velocity caused by blasting. *Engineering With Computers* 33:45–53

- Khotbehsara MM, Zadshir M, Miyandehi BM, Mohseni E, Rahmannia S, Fathi S (2014) Rheological, mechanical and durability properties of self-compacting mortar containing nano-TiO₂ and fly ash. *J Am Sci* 10(11):222–228
- Koopialipoor M, Asteris PG, Mohammed AS, Alexakis DE, Mamou A, Armaghani DJ (2022) Introducing stacking machine learning approaches for the prediction of rock deformation. *Transp Geotechnics* 34:100756
- Langefors U, Kihlström B (1963) The modern technique of rock blasting. (No Title)
- Lu S, Koopialipoor M, Asteris PG, Bahri M, Armaghani DJ (2020) A novel feature selection approach based on tree models for evaluating the punching shear capacity of steel fiber-reinforced concrete flat slabs. *Materials* 13(17):3902
- Matidza MI, Jianhua Z, Gang H, Mwangi AD (2020) Assessment of blast-induced ground vibration at Jinduicheng molybdenum open pit mine. *Nat Resour Res* 29:831–841
- Mehdizadeh B, Vessalas K, Ben B, Castel A, Deilami S, Asadi H (2022) Advances in characterization of carbonation behavior in slag-based concrete using nanotomography. In: International conference on variability of the sun and sun-like stars: from astero-seismology to space weather. Springer Nature, Singapore, pp 297–308
- Mehdizadeh Miyandehi B, Vessalas K, Castel A, Mortazavi M (2023) Investigation of carbonation behaviour in high-volume ggbs concrete for rigid road pavements. *ASCP (Australian Society for Concrete Pavements)*
- Mohamadnejad M, Gholami R, Ataei M (2012) Comparison of intelligence science techniques and empirical methods for prediction of blasting vibrations. *Tunn Undergr Space Technol* 28:238–244
- Mohamed MT (2011) Performance of fuzzy logic and artificial neural network in prediction of ground and air vibrations. *JES J Eng Sci* 39(2):425–440
- Mohammed AS, Asteris PG, Koopialipoor M, Alexakis DE, Lemonis ME, Armaghani DJ (2021) Stacking ensemble tree models to predict energy performance in residential buildings. *Sustainability* 13(15):8298
- Monjezi M, Hasanipanah M, Khandelwal M (2013) Evaluation and prediction of blast-induced ground vibration at Shur river dam, Iran, by artificial neural network. *Neural Comput Appl* 22:1637–1643
- Monjezi M, Baghestani M, Shirani Faradonbeh R, Pourghasemi Saghand M, Armaghani J, D (2016) Modification and prediction of blast-induced ground vibrations based on both empirical and computational techniques. *Engineering With Computers* 32:717–728
- New BM (1986) Ground vibration caused by civil engineering works
- Nguyen H, Bui X-N, Moayedi H (2019) A comparison of advanced computational models and experimental techniques in predicting blast-induced ground vibration in open-pit coal mine. *Acta Geophys* 67(4):1025–1037
- Pal Roy P (1993) Putting ground vibration predictions into practice. *Colliery Guardian Redhill* 241(2):63–67. [https://doi.org/10.1016/0148-9062\(93\)92499-g](https://doi.org/10.1016/0148-9062(93)92499-g)
- Peng K, Zeng J, Armaghani DJ, Hasanipanah M, Chen Q (2021) A novel combination of gradient boosted tree and optimized ANN models for forecasting ground vibration due to quarry blasting. *Nat Resour Res* 30:4657–4671
- Rai R, Singh TN (2004) A new predictor for ground vibration prediction and its comparison with other predictors
- Sarir P, Chen J, Asteris PG, Armaghani DJ, Tahir MM (2021) Developing GEP tree-based, neuro-swarm, and Whale optimization models for evaluation of bearing capacity of concrete-filled steel tube columns. *Engineering With Computers* 37:1–19
- Shang Y, Nguyen H, Bui X-N, Tran Q-H, Moayedi H (2020) A novel artificial intelligence approach to predict blast-induced ground vibration in open-pit mines based on the firefly algorithm and artificial neural network. *Nat Resour Res* 29(2):723–737
- Sheykhi H, Bagherpour R, Ghasemi E, Kalhori H (2018) Forecasting ground vibration due to rock blasting: a hybrid intelligent approach using support vector regression and fuzzy C-means clustering. *Engineering With Computers* 34:357–365
- Skentou AD, Bardhan A, Mamou A, Lemonis ME, Kumar G, Samui P, Armaghani DJ, Asteris PG (2023) Closed-form equation for estimating unconfined compressive strength of granite from three non-destructive tests using soft computing models. *Rock Mech Rock Eng* 56(1):487–514
- Standard I (1973) Criteria for safety and design of structures subjected to under ground blast. ISI., IS-6922, 1973
- Taheri K, Hasanipanah M, Golzar SB, Majid MZA (2017) A hybrid artificial bee colony algorithm-artificial neural network for forecasting the blast-produced ground vibration. *Engineering With Computers* 33:689–700
- Tribe J, Koroznikova L, Khandelwal M, Giri J (2021) Evaluation and assessment of blast-induced ground vibrations in an underground gold mine: a case study. *Nat Resour Res* 30:4673–4694
- Xue X (2019) Neuro-fuzzy based approach for prediction of blast-induced ground vibration. *Appl Acoust* 152:73–78
- Yang H, Rad N, Hasanipanah H, Amnieh MB, H., Nekouie A (2020) Prediction of vibration velocity generated in mine blasting using support vector regression improved by optimization algorithms. *Nat Resour Res* 29(2):807–830
- Yari M, Armaghani DJ, Maraveas C, Ejlali AN, Mohamad ET, Asteris PG (2023a) Several Tree-Based solutions for predicting flyrock distance due to mine blasting. *Appl Sci* 13(3):1345
- Yari M, He B, Armaghani DJ, Abbasi P, Mohamad ET (2023b) A novel ensemble machine learning model to predict mine blasting-induced rock fragmentation. *Bull Eng Geol Environ* 82(5):187
- Ye J, Koopialipoor M, Zhou J, Armaghani DJ, He X (2021) A novel combination of tree-based modeling and Monte Carlo simulation for assessing risk levels of flyrock induced by mine blasting. *Nat Resour Res* 30:225–243
- Yu Z, Shi X, Zhou J, Gou Y, Huo X, Zhang J, Armaghani DJ (2022) A new multikernel relevance vector machine based on the HPSOGWO algorithm for predicting and controlling blast-induced ground vibration. *Engineering With Computers*, 1–16
- Zeng J, Mohammed AS, Mirzaei F, Moosavi SMH, Armaghani DJ, Samui P (2022) A parametric study of ground vibration induced by quarry blasting: an application of group method of data handling. *Environ Earth Sci* 81(4):127
- Zhang H, Zhou J, Armaghani DJ, Tahir MM, Pham BT, Van Huynh V (2020a) A combination of feature selection and random forest techniques to solve a problem related to blast-induced ground vibration. *Appl Sci (Switzerland)* 10(3). <https://doi.org/10.3390/app10030869>
- Zhang X, Nguyen H, Bui X-N, Tran Q-H, Nguyen D-A, Bui DT, Moayedi H (2020b) Novel soft computing model for predicting blast-induced ground vibration in open-pit mines based on particle swarm optimization and XGBoost. *Nat Resour Res* 29:711–721
- Zhou Z-H, Feng J (2019) Deep forest. *Natl Sci Rev* 6(1):74–86
- 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(May). <https://doi.org/10.1016/j.soildyn.2020.106390>