



Performance evaluation of hybrid WOA-XGBoost, GWO-XGBoost and BO-XGBoost models to predict blast-induced ground vibration

Yingui Qiu¹ · Jian Zhou¹ · Manoj Khandelwal² · Haitao Yang³ · Peixi Yang¹ · Chuanqi Li¹

Received: 26 February 2021 / Accepted: 29 March 2021 / Published online: 22 April 2021
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2021

Abstract

Accurate prediction of ground vibration caused by blasting has always been a significant issue in the mining industry. Ground vibration caused by blasting is a harmful phenomenon to nearby buildings and should be prevented. In this regard, a new intelligent method for predicting peak particle velocity (PPV) induced by blasting had been developed. Accordingly, 150 sets of data composed of thirteen uncontrollable and controllable indicators are selected as input dependent variables, and the measured PPV is used as the output target for characterizing blast-induced ground vibration. Also, in order to enhance its predictive accuracy, the gray wolf optimization (GWO), whale optimization algorithm (WOA) and Bayesian optimization algorithm (BO) are applied to fine-tune the hyper-parameters of the extreme gradient boosting (XGBoost) model. According to the root mean squared error (RMSE), determination coefficient (R^2), the variance accounted for (VAF), and mean absolute error (MAE), the hybrid models GWO-XGBoost, WOA-XGBoost, and BO-XGBoost were verified. Additionally, XGBoost, CatBoost (CatB), Random Forest, and gradient boosting regression (GBR) were also considered and used to compare the multiple hybrid-XGBoost models that have been developed. The values of RMSE, R^2 , VAF, and MAE obtained from WOA-XGBoost, GWO-XGBoost, and BO-XGBoost models were equal to (3.0538, 0.9757, 97.68, 2.5032), (3.0954, 0.9751, 97.62, 2.5189), and (3.2409, 0.9727, 97.65, 2.5867), respectively. Findings reveal that compared with other machine learning models, the proposed WOA-XGBoost became the most reliable model. These three optimized hybrid models are superior to the GBR model, CatB model, Random Forest model, and the XGBoost model, confirming the ability of the metaheuristic algorithm to enhance the performance of the PPV model, which can be helpful for mine planners and engineers using advanced supervised machine learning with metaheuristic algorithms for predicting ground vibration caused by explosions.

Keywords Blasting · Ground vibration · PPV · Extreme gradient boosting · Machine learning · Metaheuristic optimization

1 Introduction

The increased advancement of surface mines and higher demand for coal and other minerals has led to the usage of a massive quantity of explosives. Till today, Drilling

and blasting combination is still an economically feasible method to fragment and displace the hard rock in geo-engineering. However, the energy used for the actual crushing and displacement of the rock mass accounts for only a small part (20–30%) of the explosive energy, and the rest of the

✉ Jian Zhou
j.zhou@csu.edu.cn; csujzhou@hotmail.com

✉ Manoj Khandelwal
m.khandelwal@federation.edu.au;
mkhandelwal1@gmail.com

Yingui Qiu
195512085@csu.edu.cn

Haitao Yang
yht03121@163.com

Peixi Yang
peixiyang@csu.edu.cn

Chuanqi Li
lichuanqi@csu.edu.cn

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 State Key Laboratory of Safety and Health for Metal Mines, Maanshan 243000, China

energy is wasted away and has many harmful adverse effects, such as ground vibration, air-blast, fly rocks, noise, back break, over break, etc. [14, 15, 21]. The adverse effects of blasting are inevitable and cannot be completely eliminated, but these adverse effects can definitely be minimized to an allowable level to avoid damage to the dwellings and other infrastructures. Research of blast vibration is very crucial compared to the other unwanted ill effects due to involvement of public residing in the close vicinity of mining sites, regulating and ground vibration standards setting agencies together with mine planners and owners [6, 24, 28, 50, 51]. In addition, the field of ground vibration has become an important parameter for the smooth development of mining projects, as the focus shifts to sustainable eco-friendly and geo-environmental activities.

To avoid socio-economic problems caused by ground vibrations and carry out cost-effective blasting operations, pre-operation planning becomes a requisite part [44]. For small-scale mining projects, economic reasons may limit continuous monitoring of ground vibration during mining operations. Further planning with the help of predictor equations by measuring vibration data prior to an actual operation may help mine management and owners. Therefore, the vibration level should be predicted before operation.

Different researchers have proposed many vibration predictors for PPV prediction. All the predictors estimate the PPV based on mainly two parameters (distance between blast face to monitoring point and maximum charge used per delay [21, 22, 25]). However, few predictors considered the attenuation/damping factor too. For the same excavation site, each delay has different values of safe PPV vis-à-vis charge which is given by different predictors. Different predictors may have different predicted results. As we all know, PPV is affected by various geotechnical, geological, explosive geometry, and explosive parameters, and these parameters have not been included in any available predictor variables. It is not proficient to predict any other significant parameters such as air overpressure, air overpressure, dust, frequency, noise and fly rocks. For geotechnical and mining construction projects, safe, smooth and environmentally friendly rock excavation are equally important and vital. Many empirical criteria are also proved that PPV plays a decisive role in safe and stable mining. It is necessary to develop a code that can contain the maximum number of influencing indicators to predict PPV with high precision.

Due to the excessive number of influencing parameters and the complex relationship between them, the empirical method may not be fully applicable to such problems. Currently, various soft computing tools, such as artificial neural network (ANN), maximum likelihood classification (MLC), genetic algorithm (GA), technique for order preference similarity to ideal solutions (TOSIS), random forest (RF), support vector machine (SVM), extreme gradient

boosting (XGBoost), gray wolf optimization (GWO), whale optimization algorithm (WOA) etc. are frequently applied in different complex engineering applications [2–4, 6, 10, 17, 24–27, 37, 49, 52, 53, 65–67].

Therefore, there is always a need for a simple technique to predict ground vibrations caused by explosions with higher accuracy through some indirect but relevant and reliable methods. Monjezi et al. [38] applied the ANN, empirical, and multiple regression models to predict blast vibrations in Siabhisheh pumped storage dam, Iran. They used 182 blasting data sets to study PPV, and in comparison with other proposed models, the ANN results provided better results in PPV prediction. Iphar et al. [19] and Jahed Armaghani et al. [2] applied the adaptive neuro-fuzzy inference system (ANFIS) to estimate the PPV induced by blasting. Fisne et al. [11] proposed a fuzzy inference system (FIS) model to evaluate and predict the PPV values obtained from the Akdaglar quarry, Turkey. Ghasemi et al. [13] used different fuzzy models to indirectly determine PPV using six different controllable input parameters. They found that using a fuzzy model can better predict PPV.

Mohamed [36] proposed both the ANN and FIS models to estimate PPV and found stronger performance capacity of FIS model in estimating the PPV. Hasanipanah et al. [18] used the SVM model to study PPV of Bakhtiari Dam, Iran. Recently, Zhang et al. [61–64] compared the predictive ability of five machine learning predictors which were selected, including chi-squared automatic interaction detection (CHAID), classification and regression trees (CART), ANN, RF and SVM for PPV analysis. They report that the performance of RF is significantly better than any other regression model. Yu et al. [57] utilized the RF model with Harris hawks optimization (HHO) algorithm and Monte Carlo simulation method to predict the blast-induced ground vibration with 137 datasets from the China open-pit mine and obtained a high prediction performance. Yu et al. [58] also proposed a new multikernel relevance vector machine model which uses the HPSOGWO algorithm to predict and control blast-induced ground vibration. Their results show that the proposed hybrid model is more suitable for estimating ground vibrations caused by explosions than empirical equations. More recently, Yu et al. [59] proposed hybrid extreme learning machines (ELMs) with the Harris hawks optimization (HHO) and grasshopper optimization algorithm (GOA) for forecasting PPV with 166 datasets collected from Malaysian quarries blasting. According to their report, the GOA-ELM model can obtain more accurate ground vibration values than the HHO-ELM model. Zhou et al. [73] developed a combination of two prediction algorithms and probability algorithms. Based on the gene expression program and Monte Carlo simulation to predict the PPV value to reduce the PPV risk caused by blasting, they found that the model can solve the ground vibration caused by

blasting with the lowest level of error and risk. In general, these applications show that soft computing technology has advantages in solving a series of complex parameters that will affect the process and results, and when the process and results are not fully understood and historical or experimental data are available.

Therefore, this paper tries to predict ground vibration induced by blasting by considering blast design, rock and explosive parameters with the help of extreme gradient boost (XGBoost), gray wolf optimization (GWO), whale optimization algorithm (WOA) and Bayesian optimization (BO) algorithm. These applications demonstrate that XGBoost, GWO, WOA and BO have advantages in solving the following problems: many complex parameters will affect the process and results, and the understanding of the process and results is not enough, and where there are historical or experimental data. The prediction of ground vibrations caused by explosions is also of this type.

2 Materials and methods

2.1 Extreme gradient boosting (XGBoost)

Extreme gradient boosting (XGBoost) is an algorithm based on a gradient boosting tree [8], which can play a powerful role in gradient enhancement. XGBoost based on the theory of classification and regression tree can be a very effective method of regression and classification problems [9, 29, 39, 45, 55, 61, 62, 66, 67, 69, 70]. And XGBoost can symbolize a soft computing library which combines the new algorithm with GBDT methods.

After optimization, the objective function of XGBoost consists of two different parts, which represent the deviation of the model and the regular term to prevent over-fitting [7]. $D = \{(x_i, y_i)\}$ represents a data set which contains n samples and m features, in which the predictive variable is an additive model which is made up of k basic models. The results of sample prediction are as follows:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in \varphi, \tag{1}$$

$$\varphi = \{f(x) = w_s(x)\} (s : R^m \rightarrow T, w_s \in R^T), \tag{2}$$

where \hat{y}_i represents the prediction label, x_i represents one of the samples, and the predicted score is $f_k(x_i)$ or the given sample, φ symbolizes the set of regression tree which is a tree structure parameters of s, $f(x)$ and w represents the weight of leaves and the number of leaves.

The objective function of XGBoost includes the traditional loss function and model complexity. It can be used to evaluate

the operational efficiency of the algorithm. In Formula (3), the first term represents the traditional loss function, while the second term represents the complexity of the model.

$$\text{Obj} = \sum_{i=1}^m l(y_i, \hat{y}_i^{(t-1)} + f_i(x_i)) + \Omega(f_k), \tag{3}$$

$$\Omega(f_k) = \gamma T + 1/2\lambda w^2. \tag{4}$$

In both of these formulas, i stands for the number of samples in the dataset, and m stands for the total amount of data imported into the k th tree. And γ and λ are used to adjust the complexity of the tree. Regularization term can smooth the final learning weight and avoid over-fitting.

2.2 Gray wolf optimization (GWO)

GWO [34] algorithm is based on the gray wolves’s rank and hunting mechanism in nature. It can be used to solve and optimize problems. To solve the optimization problem, there are four processes: gray wolves in nature rank, track, surround and kill their prey.

First, a random initial population was generated in the decision space. And then, according to the hierarchy from top to bottom in nature, the gray Wolf population was separated into four groups, which were $\alpha, \beta, \delta, \omega$ respectively. In Fig. 1, α, β, δ represents the three optimal solutions in each generation, which adjust the whole optimization process of GWO. ω represents the led gray wolf and accepts the leadership of the other three gray wolves.

When gray wolves are hunting, they always surround their prey, as shown in the following numerical model.

$$\mathbf{Q}(t + 1) = \mathbf{Q}_p(t) - (2\mathbf{a} \cdot \rho_1 - \mathbf{a})(2\rho_2 \cdot \mathbf{Q}_p(t) - \mathbf{Q}(t)) \tag{5}$$

$$\mathbf{a} = 2 - 2 \cdot \text{it}/\text{MAX} - \text{IT} \tag{6}$$

where t stands for the number of iterations [74], $X_p(t)$ presents position vector of the prey, $X(t)$ is position vector of

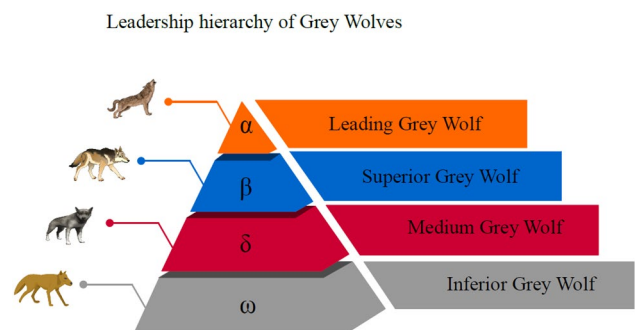


Fig. 1 Leadership hierarchy of gray wolves in nature

the current gray wolf, $X(t + 1)$ represents the position of gray wolf in the next iteration. a can be calculated by Eq. (6) and it's value decreases linearly from 2 to 0. Where it stands for the current number of iterations, and MAX_IT represents the maximum number of iterations. ρ_1 and ρ_2 are both the random numbers in $[0, 1]$.

During each iteration, the best three gray wolves are always kept in the current population by GWO, and assumes that they have strong ability to track prey [32]. And then simulated the hunting process of gray wolf. And updating the equations in GWO to update the position of other wolves.

2.3 Whale optimization algorithm

A population-based algorithm WOA was created based on the predatory behavior of humpback whales [34]. The humpback whale hunting method which is named bubble net feeding is divided into three steps: first surround the prey, then use a bubble net to attack the prey, and finally find the prey. And humpback whales use both contractive containment mechanisms and spirals to change position during bubble net attacks [16].

When surrounding the prey, humpback whales can use certain ability to judge the position of the prey and surround the prey by echo. The position update formula is as follows:

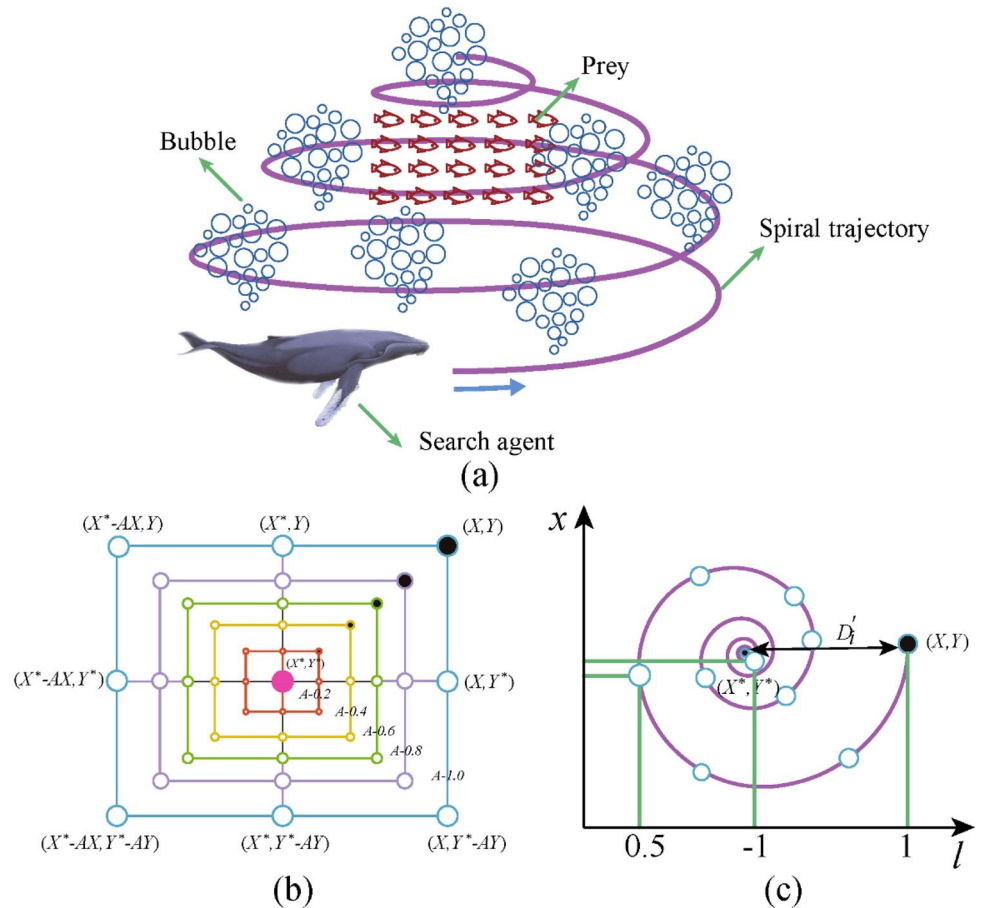
$$Q^{k+1} = Q_{best}^k - (2a \cdot \rho_1 - a) \cdot |2\rho_2 \cdot Q_{best}^k - Q^k|, \tag{7}$$

where k presents the current iteration, Q^k represents the position vector at k -th iteration, and the best solution obtained so far is indicated by $Q_{best}^k = (Q_{best1}^k, Q_{best2}^k, \dots, Q_{bestD}^k)$, D is vector dimension. ρ_1 and ρ_2 are random number in $[0, 1]$. The maximum value of a is 2 and the minimum value is 0. It can be measured by the following function:

$$a = 2 - 2k/k_{max}. \tag{8}$$

According to the previous description, humpback whales use a contractive enveloping mechanism and a spiral to update their position when using bubble nets to capture prey. With the decrease of the median value of a in Eq. (8), the prey is surrounded. The new position (Q^{k+1}) is between the current position (Q^k) and the best position (Q_{best}^k). That is, humpback whales travel around in contractions to locate their prey. The following is a digital model of the spiral updated position:

Fig. 2 Strategy and bubble-net search mechanism of the WOA [33]: **a** bubble-net feeding behavior of humpback whales; **b** shrinking encircling mechanism and **c** spiral updating position



$$Q^{k+1} = Q_{best}^k + |Q_{best}^k - Q^k| \cdot e^{bl} \cdot \cos(2\pi l). \tag{9}$$

The shape of the logarithmic spiral is regulated by b and a random number l is obtained in the interval $[-1, 1]$ [16, 75]. To achieve a better effect of bubble net attack, Mirjalili et al. [33, 35, 75] proposed that the probability of updating the existing position of humpback whales by contraction encircle mechanism or spiral position updating is 50%.

When the coefficient vector $|A| > 1$, the research on the hunting prey of humpback whales is carried out according to the following digital model:

$$Q^{k+1} = Q_{rand}^k - A \cdot |C \cdot Q_{rand}^k - Q^k|, \tag{10}$$

where Q_{rand}^k represents a position vector of whale individual which is randomly selected from the current population.

Through research, relevant scholars found that parameter A is of great significance for coordinating global exploration and local development, and it is limited to a . According to these research contents, a regulates the global detection and local development capability of WOA algorithm to a large extent (Fig. 2).

2.4 Bayesian optimization (BO) algorithm

The goal function $f(x)$ is assumed to be a known mathematical form and a convex function which can be evaluated easily by many optimizations. For hyper-parameter tuning, the goal function is an expensive non-convex function in computation which is undiscovered. So, it is hard to use optimization methods as a critical role. Moreover, the BO method has great power when the objective function is unknown and the calculation is greatly complex. The prior knowledge is used to dispose of the posterior distribution of the undiscovered

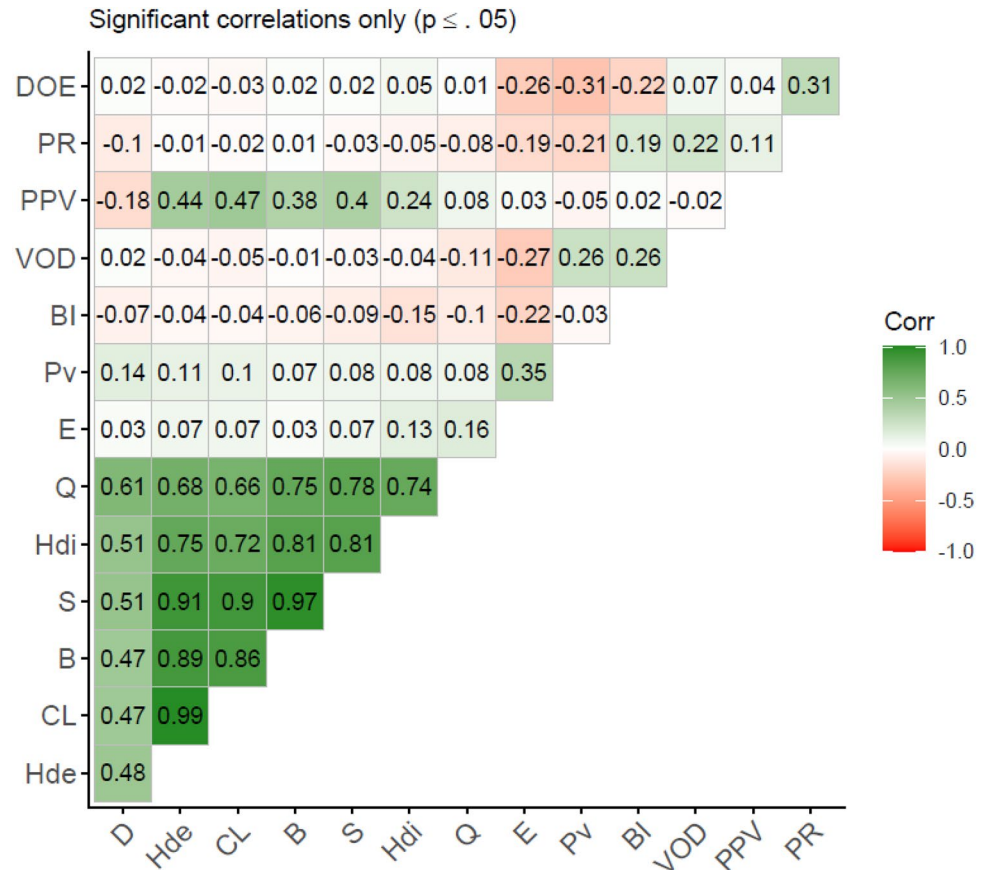
goal function by BO. And then, the next sampled hyperparameter combination will be selected with the help of the distribution.

Generally speaking, it is desirable to select hyperparameters to get the best performance. So, hyperparameter selection equals to a problem which is to get the optimal solution, that is, a performance function $f(x)$ with the best hyperparameter value as the independent variable. In many challenging optimization benchmark functions, it has been demonstrated that BO is better than other global optimization algorithms [20]. To use the BO technique, there is a need for us to find an efficient method to model the objective function’s distribution. For the case where x contains continuous hyper-parameters, the number of x that will be modeled for $f(x)$ will be infinite. (i.e., to construct a distribution for the objective function). In order to solve this problem, a multidimensional Gaussian distribution is generated by the Gaussian process [42, 43, 54]. And it is a high-dimensional normal distribution which can be flexible to model any objective function sufficiently. In other words, there is a function assumed by BO, which to be optimized is $f : X \rightarrow \mathbb{R}$, where $X \subset \mathbb{R}^n, n \in N$. Then, according to the acquisition function (α_t), in each iteration ($t = 1, 2, \dots, T$), $f(x_t), x_t \in X$ is obtained. And then, a noisy observation $y_t = f(x_t) + \varepsilon$ is obtained, in which ε follows the zero-mean Gaussian distribution $\varepsilon \sim \mathcal{N}(0, \sigma^2)$, and σ represents the noise variance. Then, the next iteration is performed after adding the new observation (x_t, y_t) to the observation data. To make good use of the information from the previous sampling point, BO learns the objective function. And BO finds the parameters which can enhance the result to the global optimum. The most likely point given by the posterior distribution is tested by the algorithm.

Table 1 Input parameters and their range

| Type of parameter | Name | Range | Mean | Std |
|-------------------|---|-----------|------|-------|
| Blast design | Hole Dia (Hdi) in mm | 100–311 | 191 | 58.2 |
| | Hole depth (Hde) in m | 4.8–43 | 13.2 | 11.1 |
| | Burden (B) in m | 2–10.5 | 4.93 | 2.10 |
| | Spacing (S) in m | 2.5–12.5 | 5.41 | 2.67 |
| | Maximum charge per delay (Q) in kg | 25–6000 | 1113 | 1661 |
| | Charge Length (CL) in m | 1–37.5 | 8.41 | 10.4 |
| Distance | Distance of monitoring point from blasting face (D) in m | 35–5800 | 463 | 703 |
| Geotechnical | Blastability index (BI) (compressive strength/tensile strength) | 6.2–12.9 | 8.72 | 1.46 |
| | Young’s modulus (E) in GPa | 3.88–9.67 | 6.7 | 1.4 |
| | Poisson’s ratio (PR) | 0.18–0.35 | 0.26 | 0.05 |
| | P-wave velocity (Pv) in m/s | 1853–4200 | 2898 | 630 |
| Explosive | Velocity of detonation of explosive (VOD) in km/s | 3.38–5.23 | 4.35 | 0.55 |
| | Density of Exp (DOE) in t/m^3 | 1.05–1.30 | 1.20 | 0.069 |

Fig. 3 Scatterplot matrix of PPV dataset with correlation



3 Materials

3.1 Blasting site and established database

The research was conducted at Jayant opencast mine of Northern Coalfields Limited (NCL). It is a subsidiary company of Coal India Limited which is located at Singrauli, Distt. Sidhi (M.P.), India. The area of NCL lies geographically between latitudes of 24o 0' to 24o 12' and longitudes 82o 30' to 82o 45'. It belongs to Gondwana super group. The dip of the strata is gentle and varying from 20 to 50.

The coalfield consists of two sub-basins, viz. Singrauli Main basin (1890 sq. km.) and Moher sub-basin (312 sq. km.). The oil field consists of 11 main mining areas, namely Kakri, Khdia, Amlohri, Dhudhichua, Bina, Marrack, Jayant, Nighahi, Gorbi, Moher and Jhingurdah [24].

The overlying rocks in this area mainly consist of argillaceous sandstone, coarse-grained sandstone and medium-grained to coarse-grained sandstone. The large dragline (24 m³ bucket size and 96 m boom length) is used in 40 m benches by the mine. And there are 311 mm diameter holes drilled on the bench. Normal blast consists of firing 50–60

holes; consuming 150–200 t of explosive. Each hole with a length of 35–40 m contains 300 kg of explosives. Moreover, there is about 6000 kg for the maximum charge for each delay. None and MS connectors are used for initiation. The inter-hole delay was 17–25 ms, while inter-row delay was 2–4 times the inter-hole delay.

According to ISRM standards, there are 150 blast vibration records at different vulnerable and strategic locations in and around to Jayant opencast mine which is used in the present paper. The range of values of different input variables of the model in the paper is determined according to the field investigation of Jayant opencast mine of Northern Coalfields Limited (NCL) and the literature of many researchers [23, 46–48]. The blasting design, geotechnical engineering scope, explosion parameters, and the distance from the monitoring point to the working surface are used as input parameters, while PPV and frequency are regarded as the output parameters of XGBoost prediction. Table 1 gives the range of input parameters which is based on the field and laboratory investigations.

To determine different physico-mechanical properties, the representative rock samples which were collected from

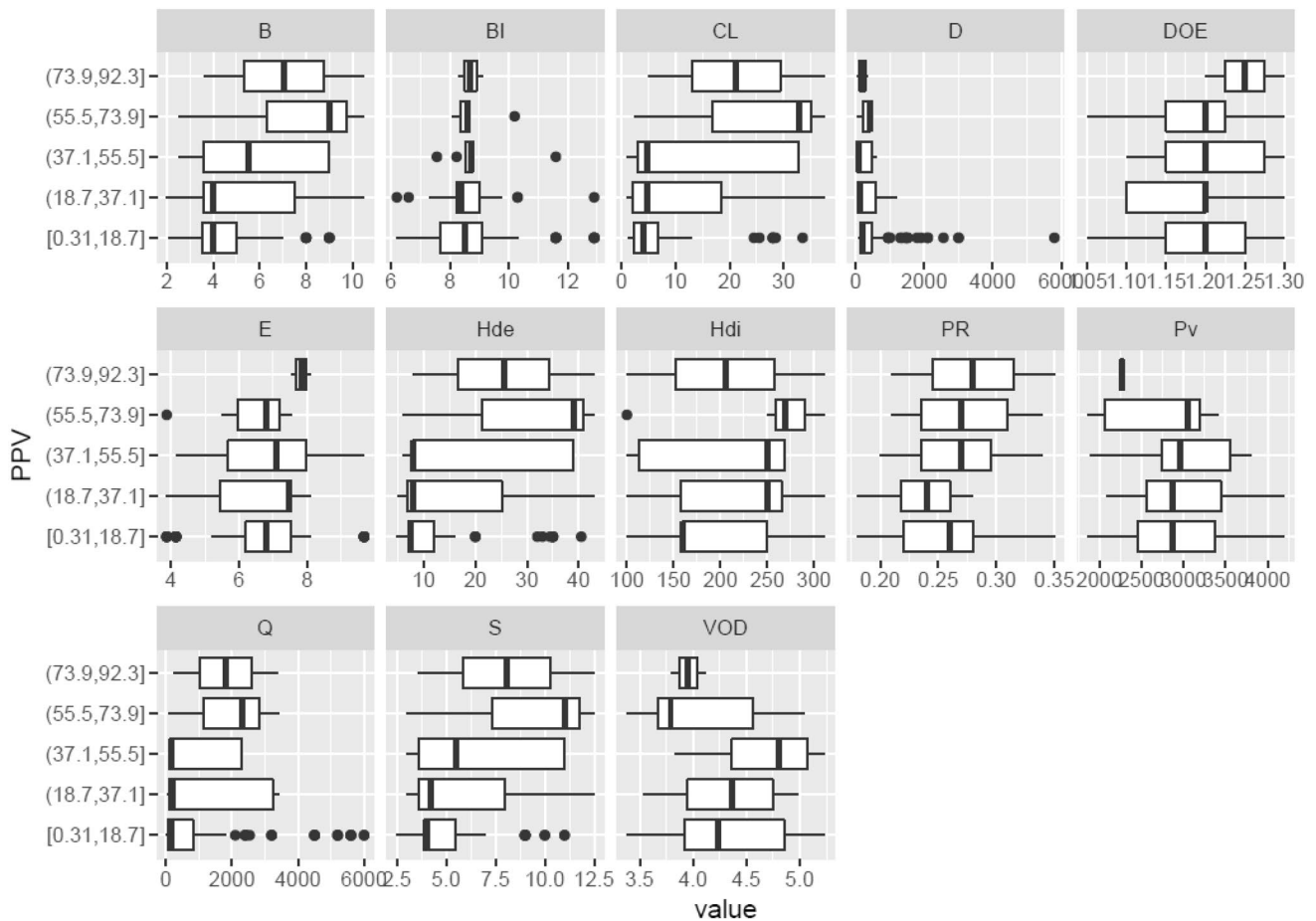


Fig. 4 View bivariate continuous distribution based on cut by PPV

the exposed coal seams of Singrauli coal fields were tested as per ISRM standards. To determine blastability index (compressive strength/tensile strength), Young’s modulus, Poisson’s ratio, and P-wave velocity, 174 rock samples were tested according to ISRM standards.

There were many parameters which were collected from the blasting site, such as burden, spacing, charge length, hole diameter and its depth, density of explosive, explosive per hole and velocity of detonation of explosive Table 1. Before blasting, the distance from the monitoring point to the blasting surface of the blasting site was also measured.

Generally speaking, one of the key parameters for the prediction of PPV is rock density [1, 24]. However, there is little difference between rock densities in the Singrauli area because of similar lithology. Therefore, during XGBoost analysis, there was no influence to be observed on PPV. Accordingly, it is not used as an input parameter in this work, but it should be used as a significant input indicator when the density changes significantly.

The relativity between the input variables in the data set, and the relativity between the input variables and the output can be observed from the scatterplot matrix plot, which is presented in Fig. 3. Moreover, the bivariate continuous distribution of each input indicator by cut of PPV, and the analysis of outliers is displayed by boxplot which is shown in Fig. 4. Obviously, there is no outliers for all variables when the PPV is within 73.9–92.3 mm/s. In Fig. 5, there is the whole analysis process implemented in this study. According to this figure, the means of this study consists of four steps: (1) data set preparation, (2) model establishment, (3) model verification and evaluation, (4) result analysis.

In this research, three optimization methods are selected to adjust the extreme gradient boosting hyperparameters. The baseline model for forecasting the PPV is XGBoost, while the GWO, WOA and BO techniques are used to search the optimal hyperparameters of the XGBoost model. In this respect, these five parameters involved in the performance of the XGBoost model, namely, Learning_rate, Num_boosting_rounds, and

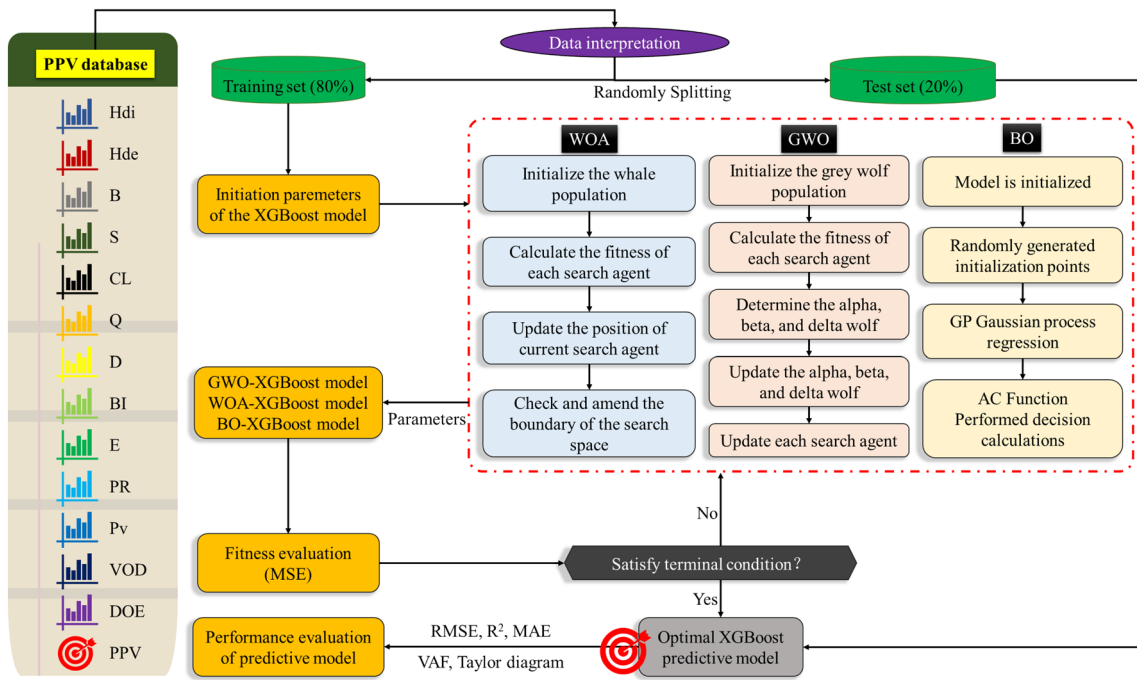


Fig. 5 The whole analysis process of three hybrid intelligence models based on XGBoost

Reg_lambda, are optimized by these three optimization algorithms. It is worth noting that the minimum value of loss mean-squared error (MSE) determines the optimal XGBoost model. The paper compares two different types of hyperparameter optimization methods for PPV prediction. The PPV which is induced by blasting from the beginning to the end is estimated by the XGBoost-based hybrid intelligence models. And its analytical process is shown in Fig. 5.

3.2 Model verification and evaluation

In order to evaluate the dependability of the hybrid models effectively in this research, several indicators such as coefficient of determination (R^2), root mean-squared error (RMSE), the variance accounted for (VAF), and mean absolute error (MAE) [29–31, 56, 60, 61, 72, 74, 75] are used to represent the relationship between the actual PPV value and the predicted PPV value, namely:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (PPV_i - \widehat{PPV}_i)^2}, \tag{11}$$

$$R^2 = 1 - \frac{\sum_i (PPV_i - \widehat{PPV}_i)^2}{\sum_i (PPV_i - \overline{PPV})^2}, \tag{12}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |PPV_i - \widehat{PPV}_i|, \tag{13}$$

$$VAF = \left[1 - \frac{\text{var}(PPV_i - \widehat{PPV}_i)}{\text{var}(PPV_i)} \right] \times 100, \tag{14}$$

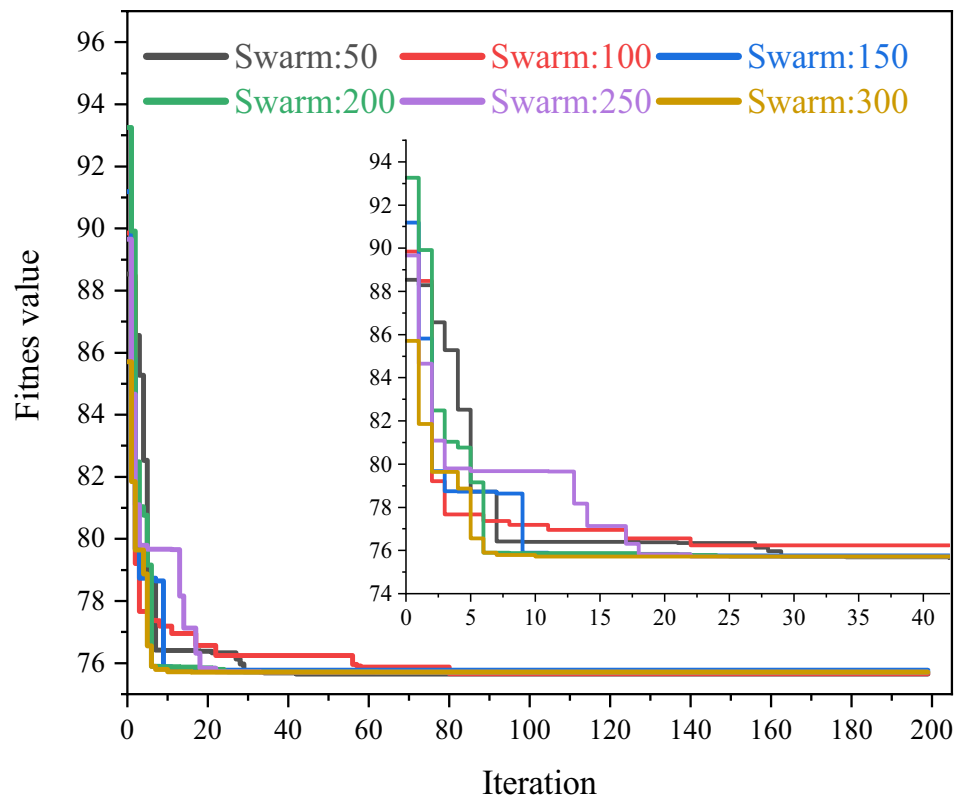
$$MSE = \frac{1}{n} \sum_{i=1}^n (PPV_i - \widehat{PPV}_i)^2. \tag{15}$$

In which n represents the number of samples in the stage of training or testing, PPV_i , \widehat{PPV}_i and \overline{PPV} represents the observed PPV value, the predicted PPV value and the average of the observed PPV values, respectively.

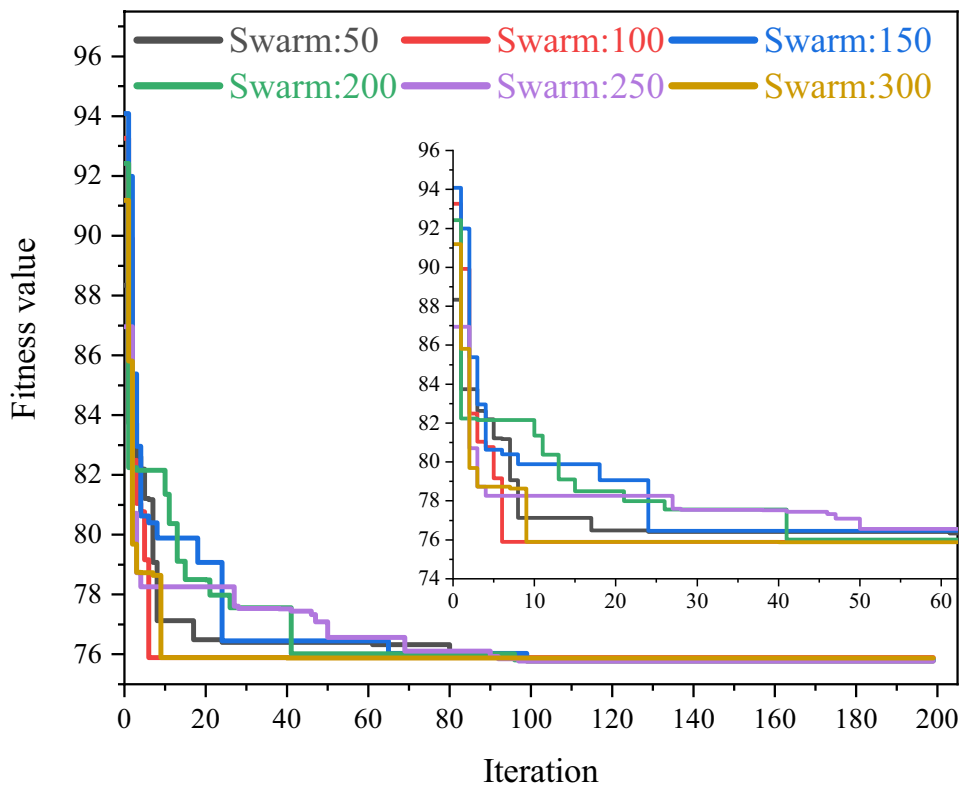
4 Results and discussion

Due to its high performance in solving problems and low requirements for feature engineering. This paper adopts an XGBoost model based on a gradient enhancement tree. XGBoost model can play a powerful role in gradient enhancement and has good adaptability to outliers and continuous variables. And it is a very effective method for

Fig. 6 Optimizing XGBoost model with WOA and GWO of different population values



(a) WAO-XGBoost



(b) GWO-XGBoost

Table 2 The performance of optimizing XGBoost models with WOA

| | Training | | | | | | | | |
|-------------|----------|------|--------|------|--------|------|-------|------|-------|
| | RMSE | Rank | R^2 | Rank | MAE | Rank | VAF | Rank | Total |
| 50 | 0.001 | 4 | 1.0 | 6 | 0.0007 | 4 | 99.99 | 6 | 20 |
| 100 | 0.103 | 3 | 1.0 | 6 | 0.0756 | 2 | 99.99 | 6 | 17 |
| 150 | 0.0006 | 6 | 1.0 | 6 | 0.0005 | 3 | 99.99 | 6 | 21 |
| 200 | 0.0008 | 5 | 1.0 | 6 | 0.0006 | 5 | 99.99 | 6 | 22 |
| 250 | 0.0008 | 5 | 1.0 | 6 | 0.0006 | 5 | 99.99 | 6 | 22 |
| 300 | 0.0006 | 6 | 1.0 | 6 | 0.0005 | 6 | 99.99 | 6 | 24 |
| Swarm (WOA) | Testing | | | | | | | | |
| | RMSE | Rank | R^2 | Rank | MAE | Rank | VAF | Rank | Total |
| 50 | 3.1541 | 3 | 0.9741 | 3 | 2.5865 | 3 | 97.52 | 3 | 12 |
| 100 | 3.2969 | 2 | 0.9717 | 2 | 2.7217 | 2 | 97.39 | 2 | 8 |
| 150 | 3.0979 | 4 | 0.975 | 4 | 2.5211 | 5 | 97.61 | 4 | 17 |
| 200 | 3.0538 | 6 | 0.9757 | 6 | 2.5032 | 6 | 97.68 | 6 | 24 |
| 250 | 3.0538 | 6 | 0.9757 | 6 | 2.5032 | 6 | 97.68 | 6 | 24 |
| 300 | 3.0925 | 5 | 0.9751 | 5 | 2.5275 | 4 | 97.62 | 5 | 19 |

dealing with regression and classification problems. In order to study the PPV predictive models' performance, there is a need to prepare the PPV database. Because of the limited sample data, there were two stages which made up the division of training/test set of this database according to the most commonly used division ratio of 80/20%. It is also known as the Pareto principle [40]. There was 80% of the data which was randomly chosen in the PPV database to train all the aforementioned models, while other 20% of the data to test the models.

The performance indicators in Eqs. (11–14) which include MAE, R^2 , VAF, and RMSE, were used to evaluate the PPV prediction models developed. For the GWO-XGBoost model, WOA-XGBoost model, BO-XGBoost model, CatB model, RandomForest model, GBR model, and XGBoost model, the test data was used for the invisible data of the models, and 30 observations of the test data set were used to predict the PPV. What's more, all prediction models used the same test data set.

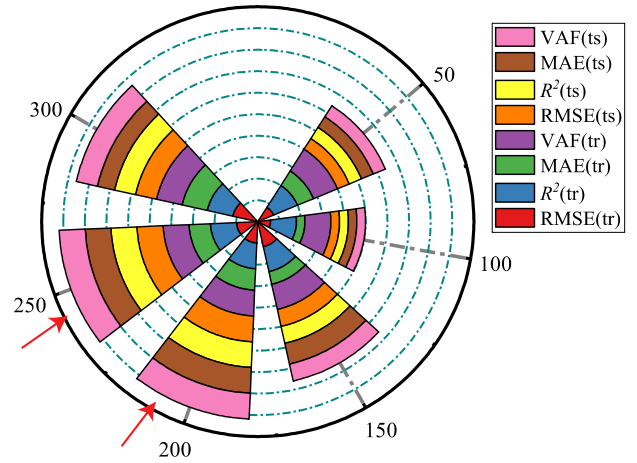
With the method proposed in Fig. 5, the WOA-XGBoost model was developed. It is difficult to determine the optimal relevant setting parameter values of the heuristic optimization algorithms in the hyper-parameter optimization process of the machine learning model. Therefore, the use of random numbers in the effective range of specific parameters of heuristic algorithms may achieve better optimization results. Firstly, the relevant hyperparameters of the XGBoost model should be initialized. And then, there was a need to set the WOA algorithm's relevant parameters: the constant b of the logarithmic spiral shape was set

to 1, L was set to a random number in $[-1, 1]$, r in $[0, 1]$ was set to a random number, and the values of swarm size were set to 50, 100, 150, 200, 250, and 300. Additionally, a technique which is called fivefold cross-validation resampling was used to enhance the reliability and performance of the optimization process.

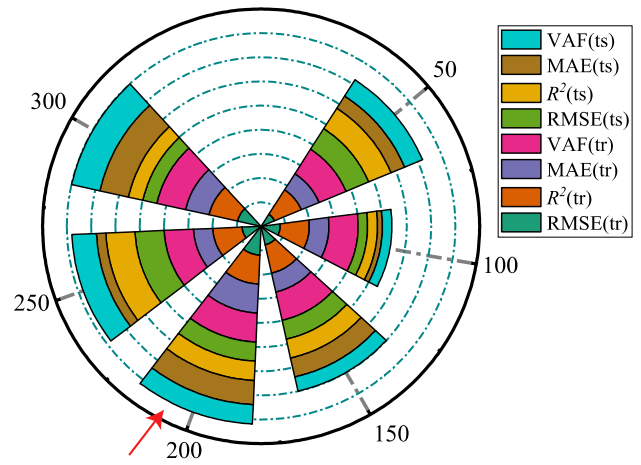
Figure 6a shows that the different population values of the WOA-XGBoost model all have reached a stable state of fitness value at 200 iterations. Notably, the different population values based on Table 2 performed very well, it is difficult to determine which population value is the best. Therefore, after WOA-XGBoost model training was completed, a comprehensive evaluation was adopted for the performance indices of the prediction model. And from Fig. 7a, the optimal populations of the WOA-XGBoost model in this condition can be seen clearly is 200(250) (RMSE = 3.0538, $R^2 = 0.9757$, VAF = 97.68, and MAE = 2.5032).

For GWO-XGBoost, after initializing the XGBoost model, the GWO algorithm's parameters were set: convergence-constant (a) was decreased linearly from 2 to 0 and the value of the population number was set to 50, 100, 150, 200, 250, and 300. Then other training and testing conditions were set to the same as WOA. Figure 6b shows change in the adaptation value during the iteration process. Finally, based on the comprehensive score in Table 3, the GWO-XGBoost model reached the best parameters combination and the best performance with populations = 200 (RMSE = 3.0954, $R^2 = 0.9751$, VAF = 97.62, and MAE = 2.5189) which is shown in Fig. 7b.

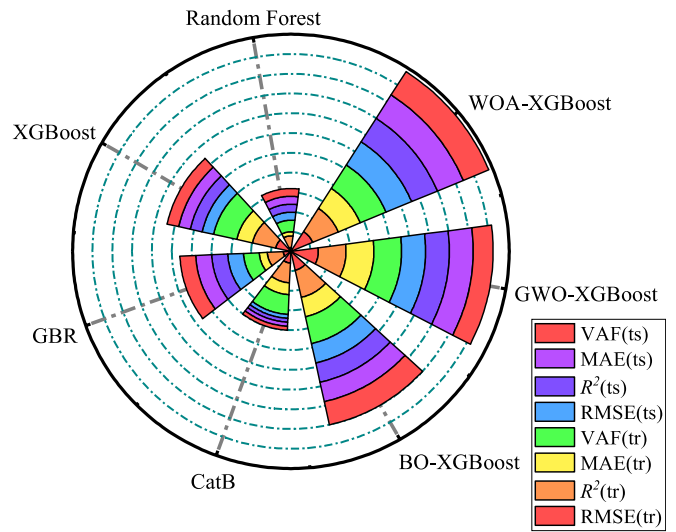
Fig. 7 Comprehensive ranking comparison of PPV prediction models



(a) WOA-XGBoost



(b) GWO-XGBoost



(c) Hybrid models

Table 3 The performance of optimizing XGBoost models with GWO

| Swarm (GWO) | Training | | | | | | | | |
|-------------|----------|------|--------|------|--------|------|-------|------|-------|
| | RMSE | Rank | R^2 | Rank | MAE | Rank | VAF | Rank | Total |
| 50 | 0.001 | 3 | 1.0 | 6 | 0.0007 | 4 | 99.99 | 6 | 19 |
| 100 | 0.0009 | 4 | 1.0 | 6 | 0.0007 | 4 | 99.99 | 6 | 20 |
| 150 | 0.0009 | 4 | 1.0 | 6 | 0.0007 | 4 | 99.99 | 6 | 20 |
| 200 | 0.0007 | 6 | 1.0 | 6 | 0.0005 | 6 | 99.99 | 6 | 24 |
| 250 | 0.0009 | 4 | 1.0 | 6 | 0.0007 | 4 | 99.99 | 6 | 20 |
| 300 | 0.0008 | 5 | 1.0 | 6 | 0.0006 | 5 | 99.99 | 6 | 22 |
| Swarm (GWO) | Testing | | | | | | | | |
| | RMSE | Rank | R^2 | Rank | MAE | Rank | VAF | Rank | Total |
| 50 | 3.0846 | 5 | 0.9752 | 5 | 2.5199 | 3 | 97.62 | 4 | 17 |
| 100 | 3.1838 | 2 | 0.9736 | 2 | 2.6249 | 1 | 97.44 | 2 | 7 |
| 150 | 3.0958 | 4 | 0.9751 | 4 | 2.5193 | 4 | 97.52 | 3 | 15 |
| 200 | 3.0954 | 4 | 0.9751 | 4 | 2.5189 | 5 | 97.62 | 4 | 17 |
| 250 | 3.0739 | 6 | 0.9754 | 6 | 2.5244 | 2 | 97.64 | 5 | 19 |
| 300 | 3.1682 | 3 | 0.9739 | 3 | 2.4467 | 6 | 97.74 | 6 | 18 |

For the establishment of BO-XGBoost, the initialization of the XGBoost model was performed first. Then the initial point was generated and Gaussian regression was performed. It is worth noting that BO is different from WOA and GWO in that no specific parameters need to be set. And other training and testing conditions were set to be the same as WOA (GWO). Finally, based on Fig. 8, it can be seen that the BO-XGBoost reached the best parameters combination and the corresponding best performance (RMSE = 3.2409, $R^2 = 0.9727$, VAF = 97.65, and MAE = 2.5867).

According to Tables 2 and 3, it is easy to find that all the WOA/GWO-XGBoost models with different population values perform very well in predicting PPV, all of which were above 0.97. However, it is difficult to determine which population value is the most appropriate. Therefore, based on the data of training, the model was established, and the model’s performance of the training and testing phases was ranked accordingly. Figure 8 shows the related ranking results.

It can be seen that the best parameters combination of the WOA-XGBoost model was obtained with populations = 200(250) (Num_boosting_rounds = 194, Learning_rate = 0.5368, and Reg_lambda = 0.2246) according to the result. For the GWO-XGBoost, the best parameters combination was obtained with populations = 200 (Num_boosting_rounds = 199, Learning_rate = 0.5374, and Reg_lambda = 0.2261). In addition, the best parameters obtained by the BO-XGBoost model were (Num_boosting_rounds = 91, Learning_rate = 0.2725, and Reg_lambda = 1). Additionally, XGBoost [7], CatB [41], Random Forest [5, 68, 71], and gradient boosting regression (GBR) [12, 69, 70] were taken into consideration and applied for comparison of the developed multiple hybrid-XGBoost model. Table 4 and

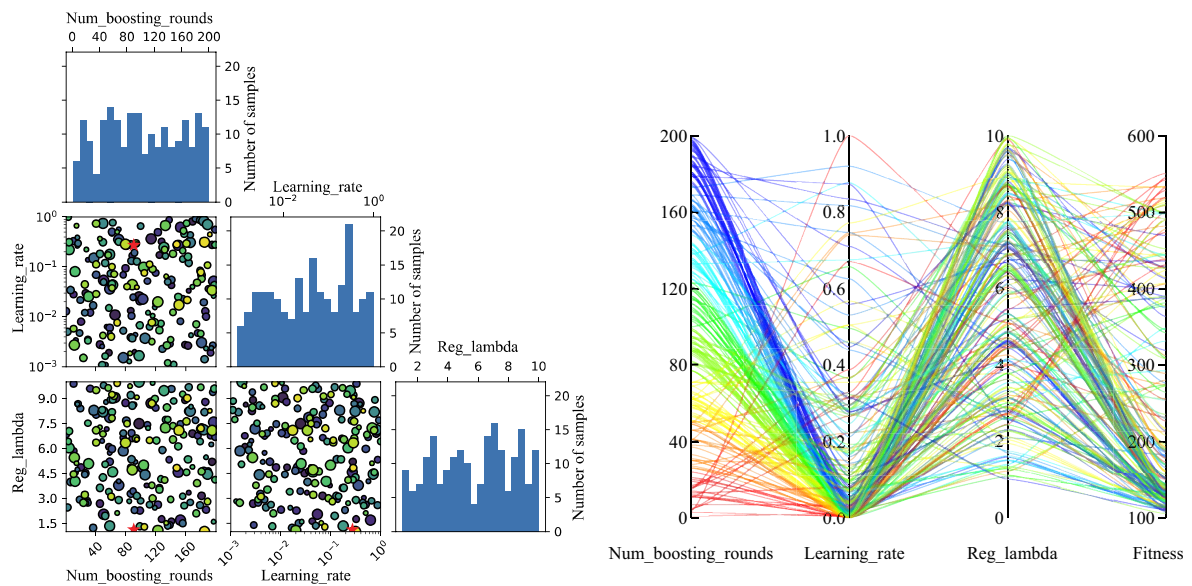
Fig. 7c list the performance results of all prediction models and their comprehensive rankings.

It is easy to find from Table 4 and Fig. 7c that the WOA-XGBoost model which was proposed before has the best performance in predicting PPV. The PPV prediction results and correlative relation were shown in Fig. 9. It is not hard to see that the distribution of measured/predicted points was close to the perfect fit line.

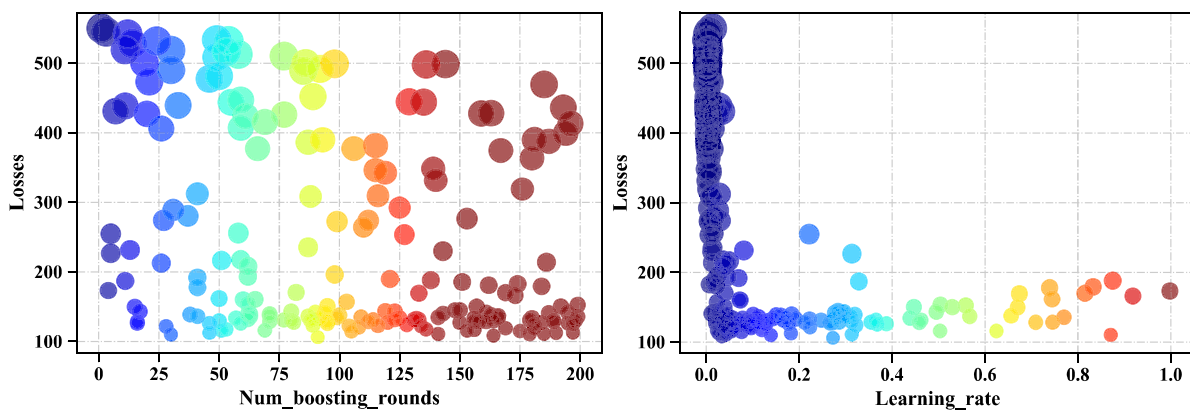
To get the further comparison and analysis of the developed models, multiple evaluation criteria were comprehensively considered through the Taylor diagram in Fig. 10. Compared with a single model evaluation index, the Taylor diagram is more intuitive for the performance between models. Taylor diagram can display the relevant information of multiple models, which is an effective method widely used in model evaluation and inspection in recent years. From the results in Fig. 10, among all the generated models, the result of WOA-XGBoost model in predicting PPV was the best. Taking the optimal hybrid XGBoost model in this paper as an example, Fig. 11 shows the prediction results of 30 sets of blast-induced PPV based on the learning training set. It can be seen that the prediction error after learning the training set is relatively small.

5 Conclusions

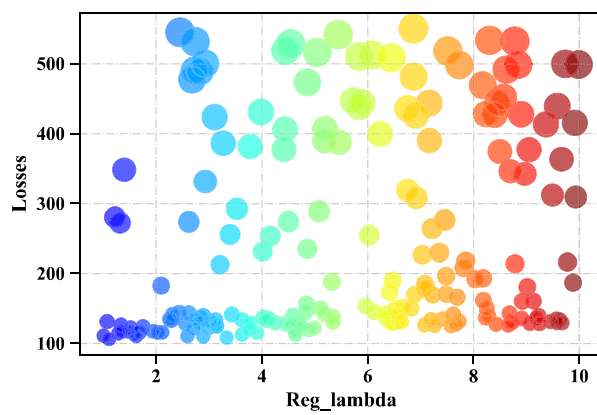
PPV caused by blasting is a commonly used parameter to evaluate ground vibration. This paper studies the XGBoost regression technology for high-precision optimization of prediction of PPV, which is of great significance to disaster prevention and mitigation in engineering.



(a) Parameters changes of optimizing model with BO in matrix diagram (b) Parameters relationship of optimizing model with BO



(c) Num_boosting_rounds changes of BO optimizing (d) Learning_rate changes of BO optimizing



(e) Reg_lambda changes of BO optimizing

Fig. 8 Optimizing XGBoost model with BO

Table 4 The performance of PPV predictive models

| Model | Training | | | | | | | | |
|---------------|----------|------|--------|------|--------|------|-------|------|-------|
| Evaluation | RMSE | Rank | R^2 | Rank | MAE | Rank | VAF | Rank | Total |
| WOA-XGBoost | 0.0008 | 6 | 1.0 | 7 | 0.0006 | 6 | 99.99 | 7 | 26 |
| GWO-XGBoost | 0.0007 | 7 | 1.0 | 7 | 0.0005 | 7 | 99.99 | 7 | 28 |
| BO-XGBoost | 0.0023 | 5 | 1.0 | 7 | 0.0017 | 5 | 99.99 | 7 | 24 |
| CatB | 0.5246 | 3 | 0.9991 | 5 | 0.4085 | 3 | 99.91 | 5 | 16 |
| GBR | 1.7914 | 2 | 0.9895 | 4 | 1.3509 | 2 | 98.95 | 4 | 12 |
| XGBoost | 0.4 | 4 | 0.9995 | 6 | 0.2954 | 4 | 99.95 | 6 | 20 |
| Random forest | 4.4244 | 1 | 0.9358 | 3 | 2.5053 | 1 | 93.58 | 3 | 8 |
| Model | Testing | | | | | | | | |
| Evaluation | RMSE | Rank | R^2 | Rank | MAE | Rank | VAF | Rank | Total |
| WOA-XGBoost | 3.0538 | 7 | 0.9757 | 7 | 2.5032 | 7 | 97.68 | 7 | 28 |
| GWO-XGBoost | 3.0954 | 6 | 0.9751 | 6 | 2.5189 | 6 | 97.62 | 5 | 23 |
| BO-XGBoost | 3.2409 | 5 | 0.9727 | 5 | 2.5867 | 5 | 97.65 | 6 | 21 |
| CatB | 5.5157 | 1 | 0.9208 | 1 | 3.9466 | 1 | 92.33 | 1 | 4 |
| GBR | 3.6375 | 4 | 0.9656 | 4 | 2.8818 | 4 | 96.86 | 4 | 16 |
| XGBoost | 3.96 | 3 | 0.9592 | 3 | 2.8917 | 3 | 96.25 | 3 | 12 |
| Random forest | 4.8576 | 2 | 0.9386 | 2 | 3.8514 | 2 | 94.26 | 2 | 8 |

The model is designed based on the blasting design, geo-technical engineering scope, explosion parameters, the distance from the monitoring point to the working surface, and other variables related to blasting. Compared with traditional prediction methods, machine learning methods can better explain multiple highly non-linear variables. Therefore, in the past few years, some scholars have conducted related research on PPV estimation and prediction through machine learning techniques. However, improving the accuracy of prediction is a major problem, and it is crucial to select the hyperparameters of the relevant model reasonably at this time.

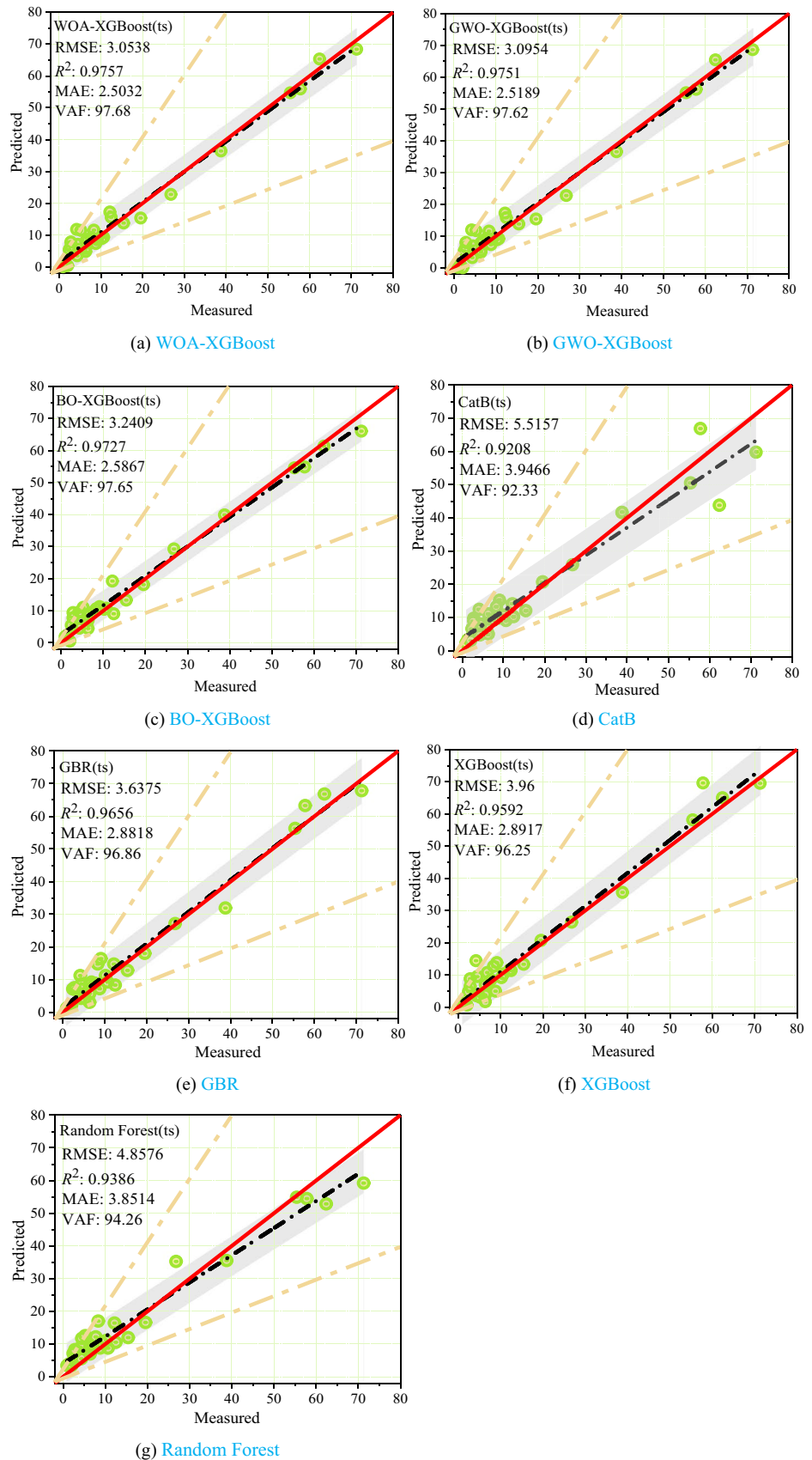
In order to select the best model hyperparameters, this paper develops a set of hybrid PPV prediction models, combining XGBoost and three optimization strategies. And obtained the best hyperparameters by fine-tuning XGBoost model hyperparameters with the three optimization algorithms of WOA, GWO, and BO algorithm. After corresponding judgments, it was compared with other single ML models to determine the technical optimization of the proposed hybrid model.

The experimental results show that the XGBoost model using this data set combined with the WOA optimization

algorithm performs best compared with other optimization models in this study, with the lowest RMSE and MAE values (RMSE = 3.0538, MAE = 2.5032), which is more suitable for predicting PPV caused by blasting. With the help of each optimization algorithm strategy, the performance of the XGBoost model has been significantly improved. The correlation coefficients/ the variance accounted for of the corresponding mixed models for PPV are 0.9757/97.68, 0.9751/97.62, and 0.9727/97.65, respectively. Considering the complicated relationship between each input variable and PPV, it is very satisfactory to obtain such a result.

According to the research results, these three optimization methods all play a role in boosting the performance of the model. Among them, WOA-XGBoost has the highest accuracy and can be applied to blast-induced PPV prediction. The limitation of using the XGBoost method to predict PPV in this article is that the training data set is relatively small, and only 150 cases are involved in the ML modeling process. And there may be other feature parameters that affect the PPV value induced by blasting. In addition, with the addition of more training data samples, improvements, feature selection, and processing data in parallel, the performance of the hybrid model may achieve a greater degree of accuracy.

Fig. 9 Correlation analyses of the measured and predicted PPV values



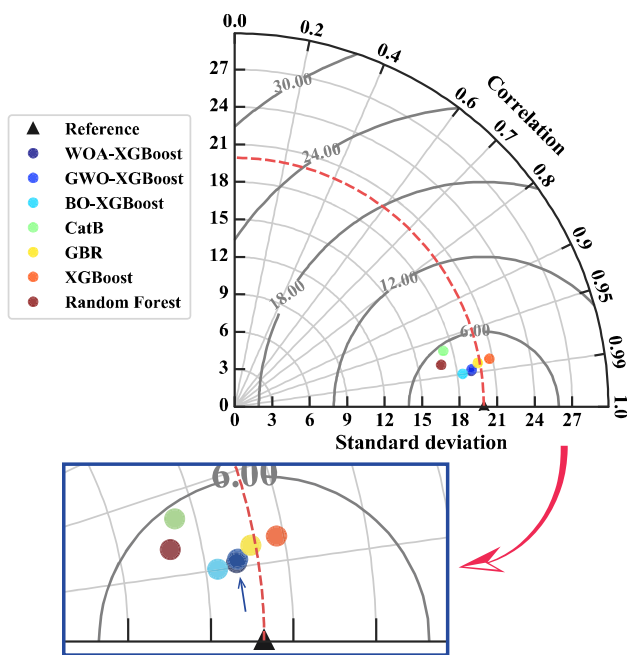


Fig. 10 Performance comparison using Taylor diagram

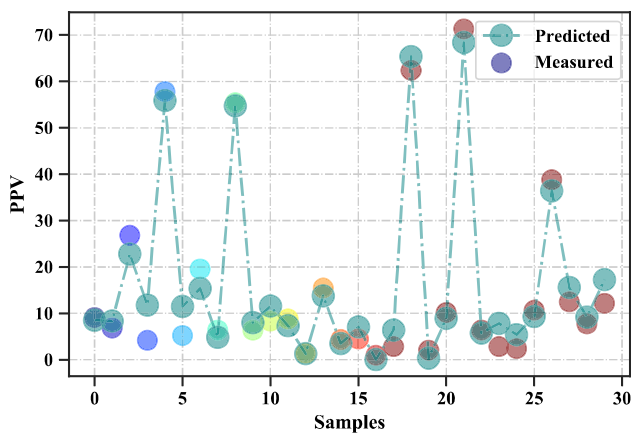


Fig. 11 The prediction results of the optimal WOA-XGBoost regressor on the test set

Acknowledgements This research was funded by the National Science Foundation of China (41807259), the National Key R&D Program of China (2017YFC0602902) and the Innovation-Driven Project of Central South University (no. 2020CX040).

References

- Adhikari GR, Singh MM, Gupta RN (1989) Influence of rock properties on blast-induced vibration. *Min Sci Technol* 8(3):297–300
- Armaghani DJ, Hajihassani M, Mohamad ET, Marto A, Noorani SA (2014) Blast-induced flyrock and ground vibration prediction through an expert artificial neural network based on particle swarm optimization. *Arab J Geosci* 7(12):5383–5396
- Armaghani DJ, Hajihassani M, Bejarbaneh BY, Marto A, Mohamad ET (2014) Indirect measure of shale shear strength parameters by means of rock index tests through an optimized artificial neural network. *Measurement* 55:487–498
- Armaghani DJ, Koopialipoor M, Bahri M, Hasanipanah M, Tahir MM (2020) A SVR-GWO technique to minimize flyrock distance resulting from blasting. *Bull Eng Geol Environ* 79:1–17
- Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32
- Cardu M, Coragliotto D, Oreste P (2019) Analysis of predictor equations for determining the blast-induced vibration in rock blasting. *Int J Min Sci Technol* 29(6):905–915
- Chen T, Guestrin C (2016) Xgboost: a scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp 785–794
- Chen T, He T (2015) Xgboost: extreme gradient boosting. R package version 0.4-2
- Ding Z, Nguyen H, Bui XN, Zhou J, Moayedi H (2020) Computational intelligence model for estimating intensity of blast-induced ground vibration in a mine based on imperialist competitive and extreme gradient boosting algorithms. *Nat Resour Res* 29:751–769
- Emary E, Zawbaa HM, Hassanien AE (2016) Binary grey wolf optimization approaches for feature selection. *Neurocomputing* 172:371–381
- Fisne A, Kuzu C, Hüdaverdi T (2011) Prediction of environmental impacts of quarry blasting operation using fuzzy logic. *Environ Monit Assess* 174:461–470
- Friedman JH (2001) Greedy function approximation: a gradient boosting machine. *Ann Stat* 29:1189–1232
- Ghasemi E, Ataei M, Hashemolhosseini H (2013) Development of a fuzzy model for predicting ground vibration caused by rock blasting in surface mining. *J Vib Control* 19(5):755–770
- Gou Y, Shi X, Huo X, Zhou J, Yu Z, Qiu X (2019) Motion parameter estimation and measured data correction derived from blast-induced vibration: new insights. *Measurement* 135:213–230
- Gou Y, Shi X, Zhou J, Qiu X, Chen X, Huo X (2020) Attenuation assessment of blast-induced vibrations derived from an underground mine. *Int J Rock Mech Min Sci* 127:104220
- Guo H, Zhou J, Koopialipoor M, Armaghani DJ, Tahir MM (2021) Deep neural network and whale optimization algorithm to assess flyrock induced by blasting. *Eng Comput* 37(1):173–186
- Hajihassani M, Armaghani DJ, Marto A, Mohamad ET (2015) Ground vibration prediction in quarry blasting through an artificial neural network optimized by imperialist competitive algorithm. *Bull Eng Geol Environ* 74(3):873–886
- Hasanipanah M, Monjezi M, Shahnazar A, Armaghani DJ, Farazmand A (2015) Feasibility of indirect determination of blast induced ground vibration based on support vector machine. *Measurement* 75:289–297
- 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
- Jones DR (2001) A taxonomy of global optimization methods based on response surfaces. *J Glob Optim* 21(4):345–383
- Khandelwal M (2012) Application of an expert system for the assessment of blast vibration. *Geotech Geol Eng* 30:205–217. <https://doi.org/10.1007/s10706-011-9463-4>
- Khandelwal M, Saadat M (2015) A dimensional analysis approach to study blast-induced ground vibration. *Rock Mech Rock Eng* 48:727–735. <https://doi.org/10.1007/s00603-014-0604-y>
- Khandelwal M, Singh TN (2006) Prediction of blast induced ground vibrations and frequency in opencast mine: a neural network approach. *J Sound Vib* 289(4–5):711–725. <https://doi.org/10.1016/j.jsv.2005.02.044>
- 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. <https://doi.org/10.1016/j.ijrmmms.2009.03.004>
25. 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* 46:1551–1558. <https://doi.org/10.1007/s00603-013-0368-9>
 26. Khandelwal M, Kankar PK, Harsha SP (2010) Evaluation and prediction of blast induced ground vibration using support vector machine. *Min Sci Technol (China)* 20(1):64–70. [https://doi.org/10.1016/S1674-5264\(09\)60162-9](https://doi.org/10.1016/S1674-5264(09)60162-9)
 27. Khandelwal M, Armaghani DJ, Faradonbeh RS, Yellishetty M, Abd Majid MZ, Monjezi M (2017) Classification and regression tree technique in estimating peak particle velocity caused by blasting. *Eng Comput* 33(1):45–53. <https://doi.org/10.1007/s00366-016-0455-0>
 28. Lawal AI, Kwon S, Hamed OS, Idris MA (2021) Blast-induced ground vibration prediction in granite quarries: an application of gene expression programming, ANFIS, and sine cosine algorithm optimized ANN. *Int J Min Sci Technol* 31(2):265–277
 29. Le LT, Nguyen H, Zhou J, Dou J, Moayed H (2019) Estimating the heating load of buildings for smart city planning using a novel artificial intelligence technique PSO-XGBoost. *Appl Sci* 9(13):2714
 30. Li C, Zhou J, Jahed Armaghani D et al (2020) Stability analysis of underground mine hard rock pillars via combination of finite difference methods, neural networks, and Monte Carlo simulation techniques. *Undergr Space*. <https://doi.org/10.1016/j.undsp.2020.05.005>
 31. 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*. <https://doi.org/10.1007/s00366-020-01014-x>
 32. Liu B, Wang R, Zhao G, Guo X, Wang Y, Li J, Wang S (2020) Prediction of rock mass parameters in the TBM tunnel based on BP neural network integrated simulated annealing algorithm. *Tunn Undergr Sp Technol* 95:103103
 33. Mirjalili S, Lewis A (2016) The whale optimization algorithm. *Adv Eng Softw* 95:51–67
 34. Mirjalili S, Mirjalili SM, Lewis A (2014) Grey wolf optimizer. *Adv Eng Softw* 69:46–61
 35. Mirjalili S, Mirjalili SM, Hatamlou A (2016) Multi-Verse Optimizer: a nature-inspired algorithm for global optimization. *Neural Comput Appl* 27:495–513
 36. Mohamed MT (2011) Performance of fuzzy logic and artificial neural network in prediction of ground and air vibrations. *J Eng Sci* 39(2):425–440
 37. Monjezi M, Mohamadi HM, Barati B, Khandelwal M (2014) Application of soft computing in predicting rock fragmentation to reduce environmental blasting side effects. *Arab J Geosci* 7(2):505–511. <https://doi.org/10.1007/s12517-012-0770-8>
 38. Monjezi M, Singh TN, Khandelwal M, Sinha S, Singh V, Hosseini I (2006) Prediction and analysis of blast parameters using artificial neural network. *Noise Vib Worldwide* 37(5):8–16. <https://doi.org/10.1260/095745606777630323>
 39. Nguyen H, Drebenstedt C, Bui XN, Bui DT (2020) Prediction of blast-induced ground vibration in an open-pit mine by a novel hybrid model based on clustering and artificial neural network. *Nat Resour Res* 29(2):691–709
 40. Nick N (2008) Joseph Juran, 103, Pioneer in quality control. *Dies New York Times* 3:3
 41. Prokhorenkova L, Gusev G, Vorobev A, Drogush AV, Gulina A (2017) CatBoost: unbiased boosting with categorical features. arXiv preprint arXiv: 1706.09516
 42. Rasmussen CE (2004) Gaussian processes in machine learning. *Advanced lectures on machine learning*. Springer, Berlin, pp 63–71
 43. Rasmussen CE, Williams CK (2005) *Gaussian processes for machine learning*, vol 2. MIT, Cambridge, p 4
 44. Saadat M, Hasanzade A, Khandelwal M (2015) Differential evolution algorithm for predicting blast induced ground vibrations. *Int J Rock Mech Min Sci* 77:97–104. <https://doi.org/10.1016/j.ijrmmms.2015.03.020>
 45. Shi XZ, Zhou J, Wu BB, Huang D, Wei W (2012) Support vector machines approach to mean block size of rock fragmentation due to bench blasting prediction. *Trans Nonferrous Metals Soc China* 22(2):432–441
 46. Singh DP, Chopra RK (1977) A comparison of static and dynamic properties of Singrauli rock. *J Mines Metals Fuels* 23(8):228–231
 47. Singh DP, Sastry VR (1986) Rock fragmentation by blasting influence of joint filling material. *J Explos Eng*:18–27
 48. Singh DP, Singh A (1975) A study of physical properties of Singrauli rocks. *J Mines Metals Fuels* 23(2):100–107
 49. Teymen A, Mengüç EC (2020) Comparative evaluation of different statistical tools for the prediction of uniaxial compressive strength of rocks. *Int J Min Sci Technol* 30(6):785–797
 50. Wang M, Shi X, Zhou J, Qiu X (2018) Multi-planar detection optimization algorithm for the interval charging structure of large-diameter longhole blasting design based on rock fragmentation aspects. *Eng Optim* 50(12):2177–2191
 51. Wang M, Shi X, Zhou J (2018) Charge design scheme optimization for ring blasting based on the developed Scaled Heelan model. *Int J Rock Mech Min Sci* 110:199–209
 52. 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
 53. Wei W, Li X, Liu J, Zhou Y, Li L, Zhou J (2021) Performance evaluation of hybrid WOA-SVR and HHO-SVR models with various kernels to predict factor of safety for circular failure slope. *Appl Sci* 11(4):1922
 54. Williams CK (1998) Prediction with Gaussian processes: From linear 579 regression to linear prediction and beyond. In: *Learning in graphical 580 models*. Dordrecht, Springer Netherlands, pp. 599–621
 55. Xu H, Zhou J, Asteris PG, Jahed Armaghani D, Tahir MM (2019) Supervised machine learning techniques to the prediction of tunnel boring machine penetration rate. *Appl Sci* 9(18):3715
 56. Yu Z, Shi X, Zhou J, Chen X, Qiu X (2020) Effective assessment of blast-induced ground vibration using an optimized random forest model based on a Harris hawks optimization algorithm. *Appl Sci* 10(4):1403
 57. 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*. <https://doi.org/10.1007/s00366-020-01136-2>
 58. Yu Z, Shi X, Zhou J, Chen X, Miao X, Teng B, Ipangelwa T (2020) Prediction of blast-induced rock movement during bench blasting: use of gray wolf optimizer and support vector regression. *Nat Resour Res* 29:843–865
 59. Yu C, Koopialipoor M, Murlidhar BR, Mohammed AS, Armaghani DJ, Mohamad ET, Wang Z (2021) Optimal ELM–Harris Hawks optimization and ELM–Grasshopper optimization models to forecast peak particle velocity resulting from mine blasting. *Nat Resour Res*. <https://doi.org/10.1007/s11053-021-09826-4>
 60. Zhang W, Zhang R, Wu C, Goh ATC, Lacasse S, Liu Z, Liu H (2019) State-of-the-art review of soft computing applications in underground excavations. *Geosci Front* 11:1095

61. Zhang W, Zhang R, Wu C, Goh AT, Wang L (2020) Assessment of basal heave stability for braced excavations in anisotropic clay using extreme gradient boosting and random forest regression. *Undergr Sp*. <https://doi.org/10.1016/j.undsp.2020.03.001>
62. Zhang W, Wu C, Zhong H, Li Y, Wang L (2020) Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization. *Geosci Front* 12:469
63. Zhang H, Zhou J, Armaghani DJ, Tahir MM, Pham BT, Huynh VV (2020) A combination of feature selection and random forest techniques to solve a problem related to blast-induced ground vibration. *Appl Sci* 10(3):869
64. Zhang X, Nguyen H, Bui XN, Tran QH, Nguyen DA, Bui DT, Moayedi H (2020) 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
65. Zhou J, Li XB, Shi XZ (2012) Long-term prediction model of rockburst in underground openings using heuristic algorithms and support vector machines. *Saf Sci* 50(4):629–644
66. Zhou J, Li X, Mitri HS (2015) Comparative performance of six supervised learning methods for the development of models of hard rock pillar stability prediction. *Nat Hazards* 79(1):291–316
67. Zhou J, Li X, Mitri HS (2016) Classification of rockburst in underground projects: comparison of ten supervised learning methods. *J Comput Civ Eng* 30(5):04016003
68. 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):04016129
69. Zhou J, Li E, Wang M, Chen X, Shi X, Jiang L (2019) Feasibility of stochastic gradient boosting approach for evaluating seismic liquefaction potential based on SPT and CPT case histories. *J Perform Constr Facil* 33(3):04019024
70. Zhou J, Li E, Yang S, Wang M, Shi X, Yao S, Mitri HS (2019) Slope stability prediction for circular mode failure using gradient boosting machine approach based on an updated database of case histories. *Saf Sci* 118:505–518
71. Zhou J, Qiu Y, Zhu S, Armaghani DJ, Khandelwal M, Mohamad ET (2020) Estimation of the TBM advance rate under hard rock conditions using XGBoost and Bayesian optimization. *Undergr Sp*. <https://doi.org/10.1016/j.undsp.2020.05.008>
72. 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* 139:106390
73. Zhou J, Li C, Koopialipoor M, Jahed Armaghani D, Thai Pham B (2021) Development of a new methodology for estimating the amount of PPV in surface mines based on prediction and probabilistic models (GEP-MC). *Int J Min Reclam Environ* 35(1):48–68
74. Zhou J, Qiu Y, Zhu S, Armaghani DJ, Li C, Nguyen H, Yagiz S (2021) Optimization of support vector machine through the use of metaheuristic algorithms in forecasting TBM advance rate. *Eng Appl Artif Intell* 97:104015
75. Zhou J, Qiu Y, Armaghani DJ, Zhang W, Li C, Zhu S, Tarinejad R (2021) Predicting TBM penetration rate in hard rock condition: a comparative study among six XGB-based metaheuristic techniques. *Geosci Front* 12(3):101091

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.