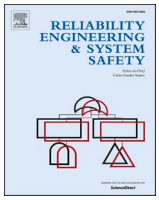
Reliability Engineering and System Safety 231 (2023) 109032

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress



Reliability and availability artificial intelligence models for predicting blast-induced ground vibration intensity in open-pit mines to ensure the safety of the surroundings

Hoang Nguyen [a,b,*], Xuan-Nam Bui [a,b], Erkan Topal [c]

[a] Department of Surface Mining, Mining Faculty, Hanoi University of Mining and Geology, 18 Vien Str., Duc Thang Ward, Bac Tu Liem Distr., Hanoi 100000, Viet Nam
[b] Innovations for Sustainable and Responsible Mining (ISRM) Group, Hanoi University of Mining and Geology, 18 Vien Str., Duc Thang Ward, Bac Tu Liem Dist., Hanoi 100000, Viet Nam
[c] Mining Engineering and Metallurgical Engineering, Western Australian School of Mines, Faculty of Science and Engineering, Curtin University, Australia

ARTICLE INFO

Keywords:
Reliability model
Mine blasting
Ground vibration
Risk assessment
Optimization
Man-made hazards

ABSTRACT

This study aims to predict ground vibration intensity in mine blasting, which is measured by peak particle velocity (PPV), using three novel intelligent models based on metaheuristic algorithms and extreme learning machine (ELM) model, including salp swarm optimization (SalSO), sparrow search optimization (SpaSO), and moth-flame optimization (MFO), named as SpaSO-ELM, SalSO-ELM, and MFO-ELM models. In this study, the SpaSO, SalSO and MFO algorithms were utilized to optimize the weights of the ELM for predicting PPV based on their different optimization mechanisms. In order to assess the performance of these models, 216 blasting records were considered and the corresponding PPV values were measured at the Coc Sau open-pit coal mine (located in the North of Vietnam). The algorithms' parameters were structured with different activation functions of the ELM model. Furthermore, in order to diagnose the improvement of the SpaSO-ELM, SalSO-ELM, and MFO-ELM models, the standalone ELM and two empirical models (linear and nonlinear models) were also investigated and evaluated. The results revealed that nonlinear models are potential candidates for predicting PPV, and the ELM-based models are robust solutions to model the nonlinear relationships of the dataset. The developed models were then also validated in practical engineering, and the findings indicated that the SpaSO-ELM model is the best intelligent model for predicting PPV in this study with an accuracy of 91.4%. The remaining hybrid models provided slightly lower performances with the accuracies in the range of 89.8%—90.5%. Although the nonlinear empirical model predicted PPV much better than the linear model; its performance is still significantly lower than the proposed hybrid intelligent models. Thus, the optimized metaheuristic-based ELM models proposed in this study are considered as the high reliability models for predicting blast-induced ground vibration intensity in open-pit mines to ensure the safety of the surroundings.

1. Introduction

Surface mining is one of the most common methods to exploit minerals, fossil fuels and metals with high mechanization and productivity. Among rock fragment methods used in open pit mines, drilling-blasting is the most common method to fragment rocks before the next unit operations are conducted, such as loading, hauling. Many reports indicated that the blasting's advantages are significant and undeniable.

Abbreviations: PPV, Peak particle velocity; ELM, Extreme learning machine; SalSO, Salp swarm optimization; SpaSO, Sparrow search optimization; MFO, Moth-flame optimization; ICA, Imperialist competitive algorithm; GA, Genetic algorithm; PSO, Particle swarm optimization; WOA, Whale optimization algorithm; BO, Bayesian optimization; JA, Jaya algorithm; SCA, Sine cosine algorithm; FA, Firefly algorithm; GWO, gray wolf optimization; HHO, Harris hawks optimization; GO, Grasshopper optimization; SaDE, Self-adaptive differential evolutionary; MVO, Multi-verse optimization; BIGV, Blast-induced ground vibration; AI, Artificial intelligence; ANN, Artificial neural network; MLP, Multi-layer perceptron neural network; SVM, Support vector machine; CART, Classification and regression trees; KNN, K-nearest neighbors; XGBoost, eXtreme gradient boosting machine; R^2, Determination coefficient; MSE, Mean squared error; RMSE, Root-mean-squared error; MAE, Mean absolute error; MAPE, Mean absolute percentage error; VAF, Variance accounted for; USBM, United States Bureau of Mines; NLE, Non-linear equation.

* Corresponding author at: Department of Surface Mining, Mining Faculty, Hanoi University of Mining and Geology, 18 Vien Str., Duc Thang Ward, Bac Tu Liem Distr., Hanoi 100000, Viet Nam.
E-mail address: nguyenhoang@humg.edu.vn (H. Nguyen).

https://doi.org/10.1016/j.ress.2022.109032
Received 24 February 2022; Received in revised form 1 December 2022; Accepted 4 December 2022
Available online 12 December 2022

Nevertheless, its adverse effects are substantial, such as blast-induced ground vibration (BIGV), flyrock, airblast, and air pollution [1–4]. Of these side effects, BIGV is considered a dangerous phenomenon that profoundly impacts the neighboring fields, especially open-pit mines located near residential areas. Although such hazardous in blasting has also been assessed and probabilistic risk-based models was proposed to deal with these problems, aiming to ensure the blast safety hazards [5]. However, blasting is still considered a complex problem with various accident risk groups can be occurred [6]. As a matter of fact, many structures have been cracked by BIGV, and a lot of slopes/benches have been subsided or became unstable due to the high intensity of BIGV in open-pit mines [7,8]. Therefore, accurate BIGV intensity prediction is imperative not only in terms of the neighboring structures but also the production aspect of open-pit mines.

For this purpose, many researchers studied and proposed a variety of empirical equations for estimating BIGV intensity since the 1950s of the 21st century, and they used the peak particle velocity (PPV) to measure the intensity of BIGV [8–11]. Such empirical equations demonstrated the convenience and ability to predict PPV from blasting parameters. However, the accuracies obtained from empirical equations, are still in very modest level.

In recent years, artificial intelligence (AI) techniques with data-driven methods have been introduced and implemented as state-of-the-art techniques not only for blasting issues but also other aspects in the mining industry as well as reliability and safety systems [12–18]. The AI-based Bayesian approach was also introduced to assess the safety of blasting operations [6]. For predicting PPV, many researchers studied and proposed various AI-based models with promising results. Monjezi, Hasanipanah [19] considered and predicted PPV using only 20 blast vibration records and an artificial neural network (ANN) model. Their results showed high acquiescence of the used ANN model in predicting PPV with an R^2 (determination coefficient) of 0.927. Similar approach was also applied at the E-Gohar iron mine (Iran) for predicting PPV with an R^2 of 0.957 and MSE (mean squared error) of 0.000722 [20]. In another study, Hasanipanah et al. [21] evaluated the feasibility of using support vector machine (SVM) for the purpose of PPV prediction, and they found that this is a potential machine learning model for this task with an outstanding testing results of R^2 and RMSE (root mean squared error) values as 0.957 and 0.34, respectively, with MAPE (mean absolute percentage error) as 6.36%. By the combination of ANN and another machine learning algorithm based on the experiences of nearest neighbors, i.e., K-nearest neighbors (KNN), Amiri et al. [22] proposed the integrated ANN-KNN model applied for predicting PPV with an accuracy improvement of ~2% as compared to traditional ANN model (92.21%). In similar way of machine learning approach, Hasanipanah et al. [23] developed the CART model (classification and regression trees) to predict PPV under the combination of decision trees in a model. Different empirical models were also compared with the CART's results, and they showed that the CART's accuracy was better ~7—15% compared to the empirical models. Among AI techniques, metaheuristic algorithms are also recommended as a robust approach to solve optimization problems in engineering, and Armaghani et al. [24], therefore, investigated the feasibility of the imperialist competitive algorithm (ICA) to predict PPV values with the consideration of the power and quadratic equations. Their results revealed that the use of metaheuristic algorithms is a new approach that can improve the accuracy of the previous machine learning/AI models in predicting accurate PPV.

Since 2019, metaheuristic algorithms have gradually been applied more commonly in PPV prediction with significantly improved results. Azimi et al. [25] considered applying the genetic algorithm (GA) in predicting PPV by optimizing an ANN model, and they found that GA could significantly support the ANN model in the improvement of PPV predictions ($R^2 = 0.988$, MAE = 1.451). Bui, Jaroonpattanapong [26] also applied another metaheuristic algorithm in predicting PPV for optimizing the KNN model, namely particle swarm optimization (PSO)-KNN model. It scored 97% of accuracy with the Tri weight kernel function. By the use of firefly algorithm (FA) for the same purposes as those used by Azimi et al. [25], Shang et al. [27] developed the FA-ANN model with an accuracy approximately 96.6%. Zhang et al. [28] also combined the PSO algorithm and XGBoost model (extreme gradient boosting machine) for predicting PPV, named as PSO-XGBoost. Similar results were also reported in the PSO-XGBoost model with R^2, RMSE and VAF (variance accounted for) were 0.968, 0.583 and 96.083, respectively. Qiu, Zhou [29] developed three XGBoost-based hybrid models for predicting PPV, including WOA (whale optimization algorithm)-XGBoost, GWO (gray wolf optimization)-XGBoost, and BO (Bayesian optimization)-XGBoost. They then evaluated the performance of these models and the results indicated that the WOA-XGBoost model yielded the best performance (RMSE = 3.054, $R^2 = 0.976$, VAF = 97.68, and MAE = 2.503). Similar approach with another metaheuristic algorithm, Jaya algorithm (JA) was also applied to optimize the XGBoost for predicting PPV by another research group [30]. Lawal et al. [31] also applied the sine cosine algorithm (SCA) to optimize an ANN model for the same purpose in five different granite quarries. An ideal result was also reported in their study with an R^2 of 0.999 for the SCA-ANN model. Besides, a variety of other hybrid models were also proposed to predict PPV with high reliability, such as MFO (Moth-flame optimization)-ANFIS (adaptive fuzzy inference neural network) [32]; HHO (harris hawks optimization)-ANN, WOA-ANN [7]; FA-SVM [33]; PSO—CRANFIS (chaos recurrent ANFIS) [34]; HHO-ELM (extreme learning machine) and GO (grasshopper optimization)-ELM [35]; MVO (multi-verse optimization)-ELM [36]; PSO-ELM [37], and SaDE (self--adaptive differential evolutionary)-ELM [38].

Despite the fact that many AI-based soft computing models/paradigms have been proposed to predict PPV induced by mine blasting, yet, they have been implemented in single site specific mine but have not been tested in other areas/mines. A review of the published works showed that AI-based soft computing models/paradigms based on the optimization characteristics of metaheuristic algorithms tend to provide better performances than those of the traditional AI models (i.e., standalone models). This study is, therefore, to propose three different novel AI-based models for predicting PPV in open-pit mines based on the ELM model and three swarm-based metaheuristic algorithms, including sparrow search optimization (SpaSO), salp swarm optimization (SalSO), and MFO, named as SpaSO-ELM, SalSO-ELM and MFO-ELM models. It must be emphasized that these AI-based hybrid models have not been considered and proposed to predict PPV before. Furthermore, to interpret the role of three metaheuristic algorithms used in the optimization of the ELM model, the standalone ELM model was also developed and compared with the SpaSO-ELM, SalSO-ELM and MFO-ELM models. Two empirical equations, including a linear equation (i.e., USBM - United States Bureau of Mines) and a nonlinear equation were also considered to estimate PPV to diagnose the characteristics of the dataset used and compared to the AI-based models. The methodology, results, and discussion will be taken up in the next sections.

2. Methodology

In this study, the enhanced ELM models are selected to predict PPV with the support of the SpaSO, SalSO, and MFO optimization algorithms, namely SpaSO-ELM, SalSO-ELM, and MFO-ELM. Thus, the principle of the ELM and three optimization algorithms, as well as the framework of the hybrid models based on ELM model will be presented in this section. Moreover, empirical method is also selected to estimate PPV and compared with the hybrid ELM-based models. The model evaluation methods are also presented in this section as a part of the methodology.

2.1. Extreme learning machine (ELM)

ELM is one of the members in the ANN family, but it comprises only one hidden layer. It is an enhancement of the MLP model with the learning rate improved [39] and the model can obtain global

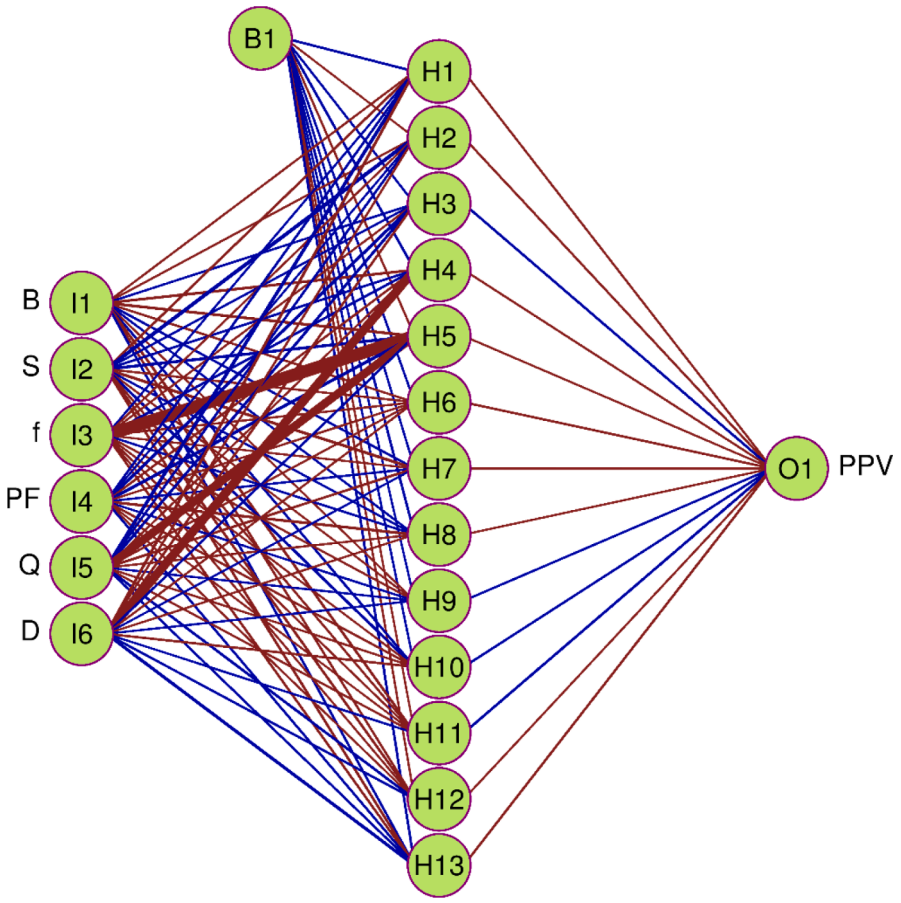


Fig. 1. Illustrating the ELM architecture for predicting PPV.

optimization based on the random weights generated [40]. The principle of the ELM model is described as follows:

Given a training dataset $\{X, Y\} = (x_i, y_i)$ with $x_i \in \mathbb{R}^V, y_i \in \mathbb{R}^C$, $i = 1, ..., N$; x_i is the i^{th}input data in the V dimensions, $x_i = [x_{i1}, x_{i2}, ..., x_{iV}]^T \in \mathbb{R}^V$; N is the number of inputs; y_i is the i^{th}output data in the V dimensions, corresponding with the x_i input, $y_i = [y_{i1}, y_{i2}, ..., y_{iC}]^T \in \mathbb{R}^C$. Subsequently, random weights (w) and biases (b) are generated with $w \in \mathbb{R}^{V \times L}$ and $b \in \mathbb{R}^L$, L denotes the hidden nodes of the ELM model. Next, the original PPV dataset is mapped to the hidden layer to get the $h(x)$ function, $h(x) \in \mathbb{R}^{1 \times L}$. The weights matrix of the output layer is represented as $\beta \in \mathbb{R}^{L \times C}$, and they can be calculated based on prediction errors and it can prevent overfitting problem concurrently. The objective function of this task can be expressed as:

$$\begin{aligned} &\min_{\beta \in \mathbb{R}^{L \times m}} \quad \frac{1}{2}\| \beta \|_F^2 + \frac{\alpha}{2}\| \varphi_i \|^2 \\ &s.t. \quad h(x_i)\beta = y_i - \varphi_i \end{aligned} \tag{1}$$

where φ_i is the regression error of the ith training sample, $\varphi_i \in \mathbb{R}^C$; αdenotes the penalty coefficient.

β can be computed as follows:

$$\beta = \begin{cases} H^T\left(HH^T + \dfrac{I_N}{C}\right)^{-1} Y & if\ N < L \\ \left(H^T H + \dfrac{I_L}{C}\right)^{-1} H^T Y & if\ N \geq L \end{cases} \tag{2}$$

with $H = [h(x_1), h(x_2), ..., h(x_N)]^T \in \mathbb{R}^{N \times L}$; $Y = [y_1, y_2, ..., y_N]^T \in \mathbb{R}^{N \times C}$, I is the identify matrix.

The overall architecture of the ELM model is demonstrated in Fig. 1.

2.2. Sparrow search optimization (SpaSO)

SpaSO is a new metaheuristic algorithm that was proposed in 2020 by Xue and Shen [41] based on the behaviors of sparrows in searching the food. In a sparrow swarm, producers and scroungers are considered as the main members of the swarm. Whereas the producers have high energy levels that can identify the food sources, the scroungers have lower levels of energy and they get the food from the producers [42]. In SpaSO, the energy levels are evaluated through the fitness values of sparrows. When the sparrow detects predators, it makes a warning signal to the individuals. Once this signal exceeds the normal threshold, the producers will lead the flock to a safer location.

In SpaSO, every sparrow can become a producer if it can find out better food sources. Yet, the ratio between the producers and scroungers in the flock are constant. To become a producer, several scroungers try to fly to other destinations to obtain higher energy levels. In addition, they can compete against the producers to get the best food sources to become the producers. When in danger, the sparrows in the outer edge will try to find better positions, whereas, the others seek random locations to converge together. The mathematical of the SpaSO can be expressed as follows:

- Step 1: Generating the matrix of sparrows' positions

$$P = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,d} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ p_{sp,1} & p_{sp,2} & \cdots & p_{sp,d} \end{bmatrix} \tag{3}$$

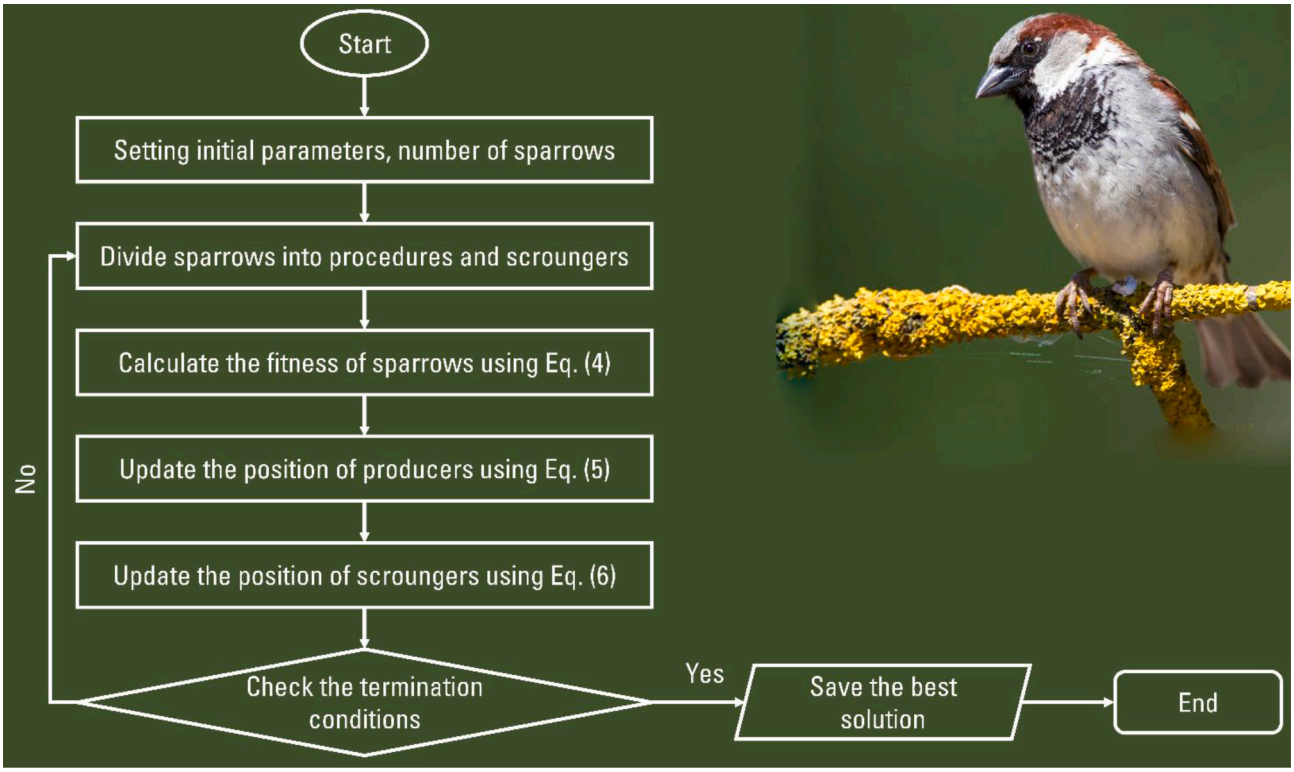


Fig. 2. The SpaSO framework for searching optimization solutions.

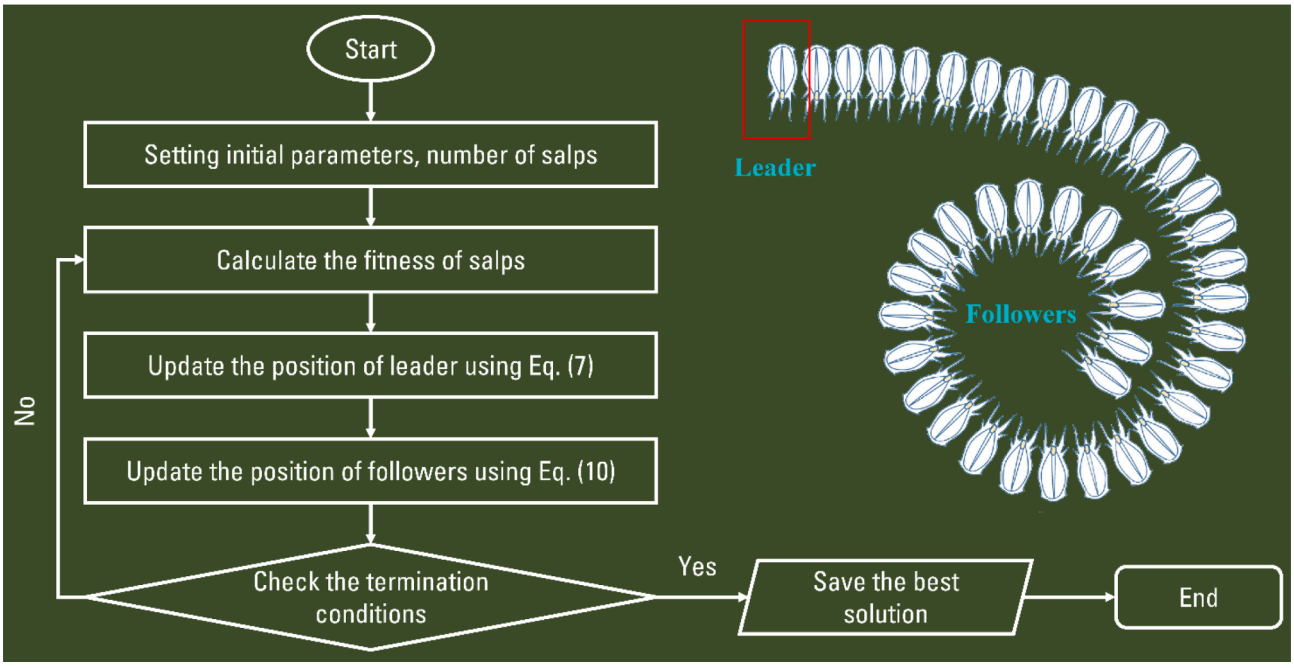


Fig. 3. The SalSO framework for searching optimization solutions.

where d denotes the dimensions of the dataset, sp is the number of sparrows.

- Step 2: Calculating the fitness values of sparrows

$$F(P) = \begin{bmatrix} f(p_{1,1}, p_{1,2}, \ldots, p_{1,d}) \\ f(p_{2,1}, p_{2,2}, \ldots, p_{2,d}) \\ \vdots \\ f(p_{sp,1}, p_{sp,2}, \ldots, p_{sp,d}) \end{bmatrix} \quad (4)$$

- Step 3: Updating the position of producers

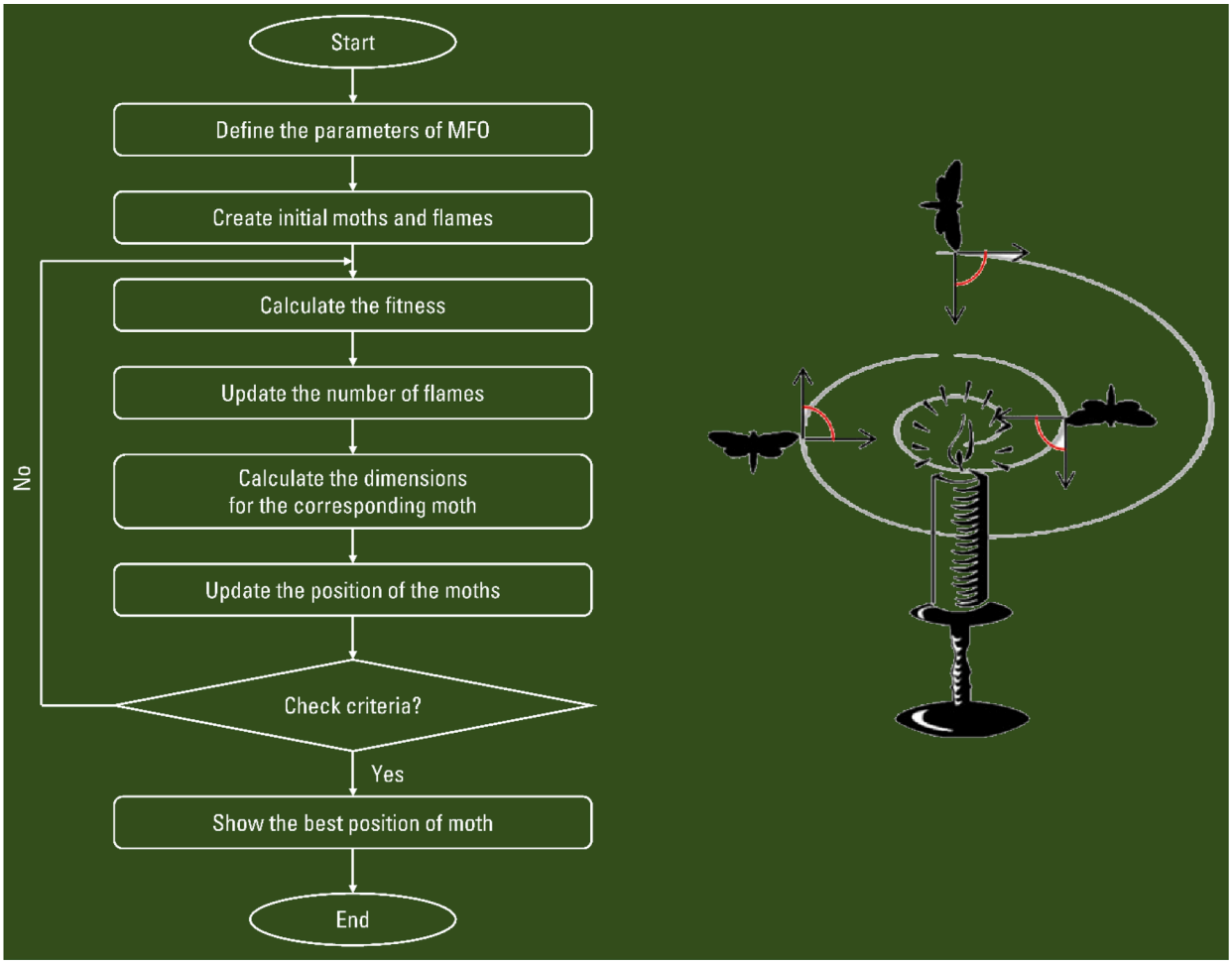


Fig. 4. The MFO framework for searching optimization solutions.

$$P_{ij}(t+1) = \begin{cases} P_{ij}(t)e^{\left(\frac{-i}{r_1 T_{\max}}\right)} & W_s < \delta \\ P_{ij}(t) + r_2 M & W_s \geq \delta \end{cases} \tag{5}$$

where $P_{ij}(t)$ stands for the position of the ith sparrow in the jth dimension at the current iteration; r_1 is a random value interval [0,1]; r_2 is a random value in normal distribution; W_s is the warning signal value interval [0,1]; δ denotes the threshold interval [0.5, 1]; M is the matrix of unity elements with *1xd* dimension.

- Step 4: Updating the position of scroungers

$$P_{ij}(t+1) = \begin{cases} P_{best}(t) + \beta\left|P_{ij}(t) - P_{best}(t)\right| & \upsilon_i > \upsilon_b \\ P_{ij}(t) + r_3\left(\frac{\left|P_{ij}(t) - P_{worst}(t)\right|}{(\upsilon_i - \upsilon_w) + \omega}\right) & \upsilon_i = \upsilon_b \end{cases} \tag{6}$$

where β denotes the step size parameter; r_3 is a random value interval [−1,1]; υ_i stands for the fitness of the ith sparrow; υ_b and υ_w are the best and worst fitness values of the ith sparrow, respectively; ω is a coefficient to prevent the zero division. The framework of the SpaSO for searching the optimization solution is shown in Fig. 2.

2.3. Salp swarm optimization (SalSO)

SalSO is a bio-inspired metaheuristic algorithm that was constructed based on the predation behaviors of salp [43]. Using salp chain search calculation, SalSO can find the optimal sources of foods. Indeed, it is depending on the position of the salp chain, as well as the individuals. They are classified to two groups: followers and leaders. Whereas leaders play a role as the trailblazers leading the salp chain, other individuals follow as followers [44]. For this reason, an initial population is necessary, like other metaheuristic algorithms. Also, an objective function is required to calculate the fitness of individuals in the chain. In the SalSO, a salp is assigned as the leader if it has the best fitness, and the remaining individuals are followers. During search food sources, followers can become the leader in case of the new food source found is better the previous one, and the fitness of the current follower is better than the current leader. Therefore, the position of leader should be updated using the equation below:

$$P_d^1 = \begin{cases} FS_d + \mu_1((UB_d - LB_d)r_4 + LB_d), & r_5 \geq 0 \\ FS_d - \mu_1((UB_d - LB_d)r_4 + LB_d), & r_5 < 0 \end{cases} \tag{7}$$

where P_d^1 is the position of leader in the dth dimension; FS_d denotes the food source position in the d^{th} dimension; μ_1 stands for the balance parameter of SalSO that can be calculated using Eq. (8); LB_d and UB_dindicate the lower and upper bounds in the d^{th} dimension; r_4 and r_5 are random numbers.

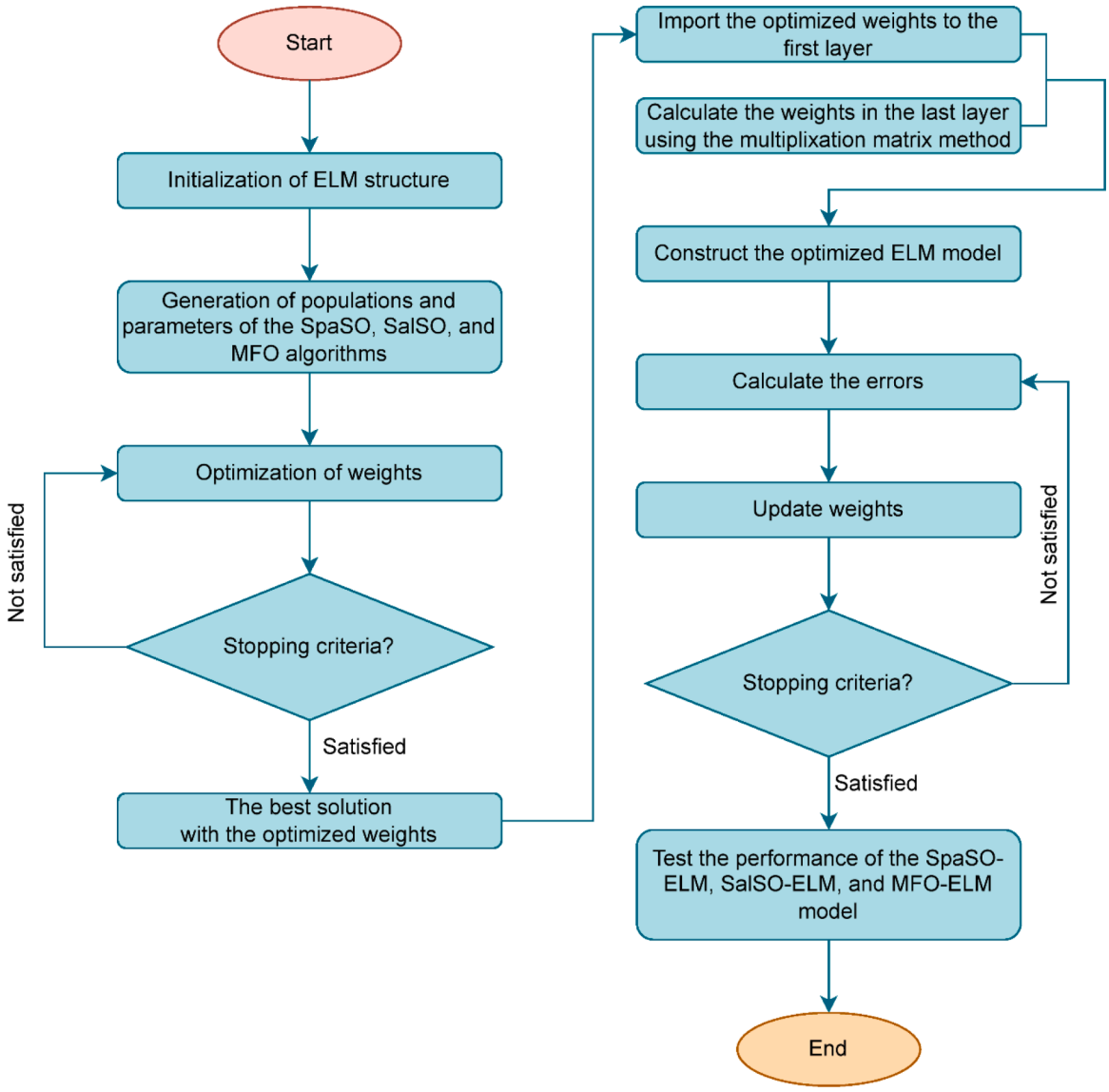


Fig. 5. Steps of the ELM-based optimization models for predicting PPV.

$$\mu_1 = 2e^{-\left(\frac{4t}{T}\right)^2} \tag{8}$$

where t is the existing iteration, and T stands for the maximum number of iterations.

Similarly, the position of followers are also updated with the equation below:

$$P_d^j = \frac{1}{2}\alpha\tau + \nu_0\tau, \quad j \geq 2 \tag{9}$$

where P_d^j refer to the position of the jth follower in the dth dimension; α is the acceleration of the algorithm; τ is the time interval between evaluations, and ν_0 is the initial velocity of the salp.

Due to the initial velocity ν_0 is set equal to 0, and τ is set equal to 1 in this algorithm; therefore, the position of followers can be rewritten as follows:

$$P_d^j = \frac{1}{2}\left(P_d^j + P_d^{j-1}\right), \quad j \geq 2 \tag{10}$$

The framework of the SalSO for searching the optimization solution is shown in Fig. 3.

2.4. Moth-flame optimization (MFO)

MFO is a population-based metaheuristic algorithm which was developed by Mirjalili [45] based on the behaviors of moths and flames. Accordingly, moths tend to fly towards the moon. However, it is often attracted to the flames in reality. The positions of moth and flame refer to a candidate solution. Meanwhile, only moths can fly around the flames, and therefore, they must update their positions continuously to get better solutions.

In MFO, two populations are used, including moth and flame. The positions of moths can be depicted through a matrix, as follows:

$$M = \begin{bmatrix} M_1 \\ M_2 \\ \vdots \\ M_n \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1d} \\ m_{21} & m_{22} & \cdots & m_{2d} \\ \vdots & \vdots & \vdots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nd} \end{bmatrix} \tag{11}$$

where n represents the number of populations (e.g., moths and flames); d represents the number of variables used (dimensions).

Next, the fitness of moths is calculated using the fitness function in Eq. (12).

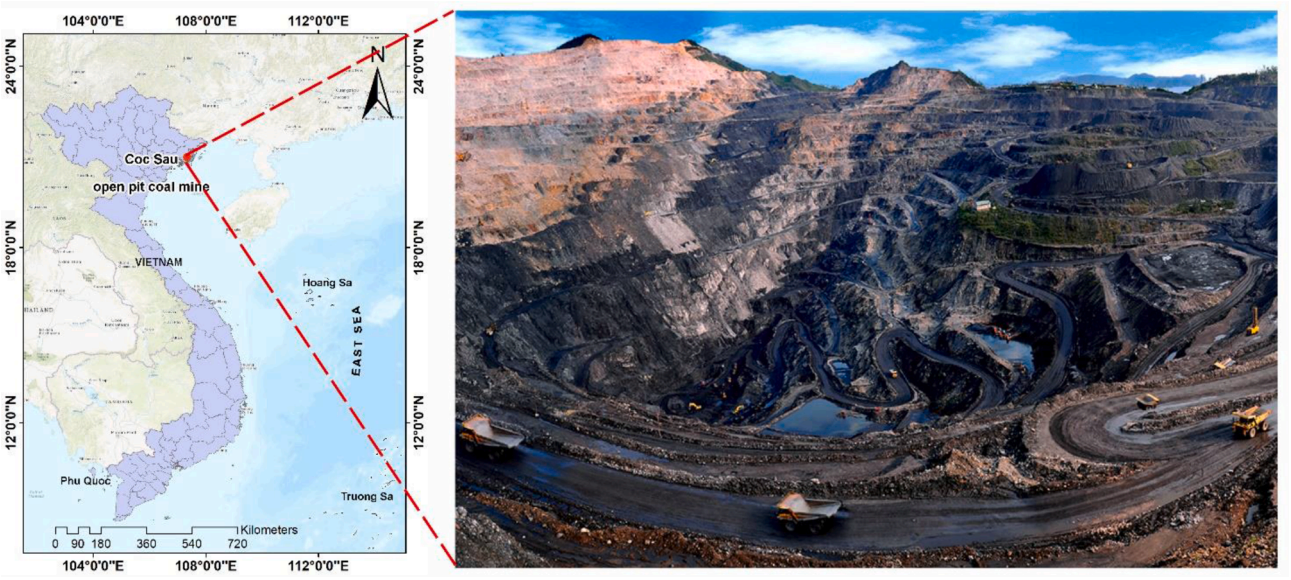


Fig. 6. Location of the selected study area with an aerial view.

$$OM = \begin{bmatrix} OM_1 \\ OM_2 \\ \vdots \\ OM_n \end{bmatrix} \tag{12}$$

where OM_i indicates the fitness value of the ith moth in the dimension, $M_i = [m_{i1}, m_{i2}, ..., m_{id}]$ with $i \in \mathbb{R}^n$.

Similarly, the positions of the flames can be described in the matrix (13) and their fitness values is computed using Eq. (14).

$$F = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_n \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1d} \\ f_{21} & f_{22} & \cdots & f_{2d} \\ \vdots & \vdots & \vdots & \vdots \\ f_{n1} & f_{n2} & \cdots & f_{nd} \end{bmatrix} \tag{13}$$

$$OF = \begin{bmatrix} OF_1 \\ OF_2 \\ \vdots \\ OF_n \end{bmatrix} \tag{14}$$

where OF_j denotes the fitness value of the jth flame in the dimension, $F_i = [f_{j1}, f_{j2}, ..., f_{jd}]$ with $j \in \mathbb{R}^n$.

Like other metaheuristic algorithms, MFO also creates an initial population. However, in the MFO, two populations (i.e., moths and flames) are created using Eq. (15).

$$m_{id} = LB_i + r_6(UB_i - LB_i) \tag{15}$$

where LB_i and UB_iindicate the lower and upper bounds of the ith variable in the d dimension; r_6 is random number in the range of 0 and 1.

The position of the ith moth at the t iteration is calculated using the following formula:

$$M_i(t+1) = \begin{cases} \chi_i \cdot e^{\varepsilon r_7} \cdot \cos 2\pi r_7 + F_i(t), & i \le f_a \\ \chi_i \cdot e^{\varepsilon r_7} \cdot \cos 2\pi r_7 + F_{f_a}(t), & i > f_a \end{cases} \tag{16}$$

where ε is a constant used to define the spiral motion form; t is the current iteration; r_7 is random number in the range of γ and 1 with $\gamma = -1 + t(\frac{-1}{T})$; The maximum number of iterations is denoted by T; f_a stands for the number of adaptive flames and it can be calculated using Eq. (18); χ_i is the distance between the ith moth and the corresponding flame calculated using Eq. (17).

$$\chi_i = \begin{cases} |F_i - M_i|, & i \le f_a \\ |F_{f_a} - M_i|, & i > f_a \end{cases} \tag{17}$$

$$f_a = round\left(n - \frac{t(n-1)}{T}\right) \tag{18}$$

It should be noted that f_a is used to identify the number of moths in each iteration, aiming to balance the exploration and exploitation stages. In the exploration phase, each moth fly around a specific flame and updates its position to avoid falling into local optima. In the exploitation phase, the fitness values of moths and flames are calculated at each iteration and they are then sorted to select the best solution at the current iteration. Based on this result, the populations of flame are determined for the next iteration.

The framework of the MFO for searching the optimization solution is presented in Fig. 4.

2.5. Proposing ELM-based optimization models for predicting PPV

As presented in the principle of ELM, this network consists of only one hidden layer, and the weights between the input and hidden layers are generated randomly. It should be underlined that the randomly generated weights in this period are constant. Meanwhile, the weights between the hidden and output layers are calculated by multiplication matrix method. Once the truth ground values are given, the multiplication matrix method can be used to find out the weights of the network. Remarkably, this method can find global optimal exactly.

However, the found global optimal point is only for the previous randomly generated weights. For another random weights, another global optimal will be defined. Therefore, the use of metaheuristic algorithms (i.e., SpaSO, SalSO, and MFO) will provide a population of solution, and each solution contains a random set of weights for the input and hidden layers of the ELM model. The weights between the hidden and output layers are still calculated using the multiplication matrix method, like the standalone ELM model. In other words, the SpaSO, SalSO, and MFO algorithms are used to optimize the weights of the ELM model to predict PPV, named as SpaSO-ELM, SalSO-ELM, and MFO-ELM models. The proposed method is shown in Fig. 5.

2.6. Empirical methods

As introduced earlier, AI models and empirical equations are

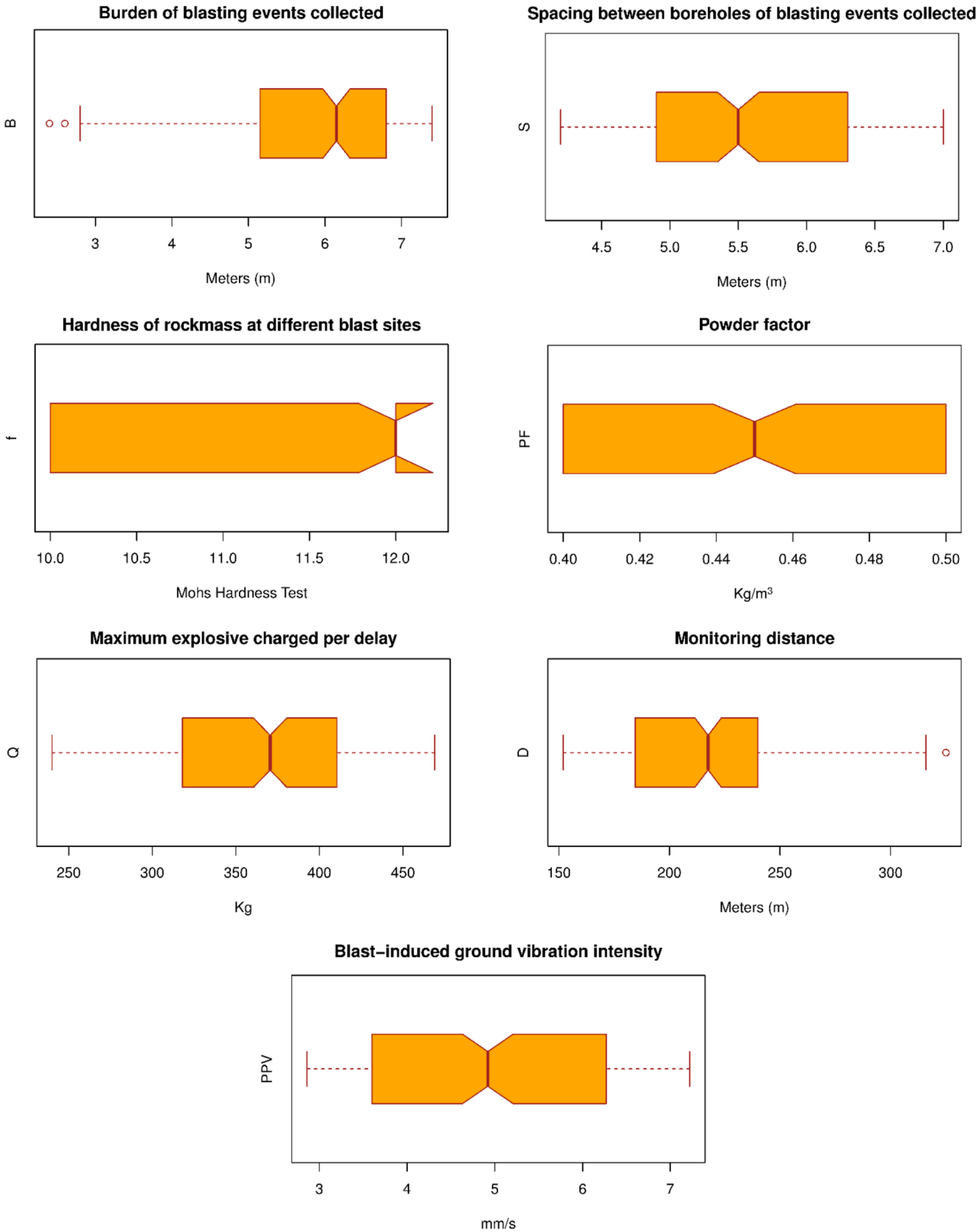


Fig. 7. Boxplots and statistics of the collected dataset.

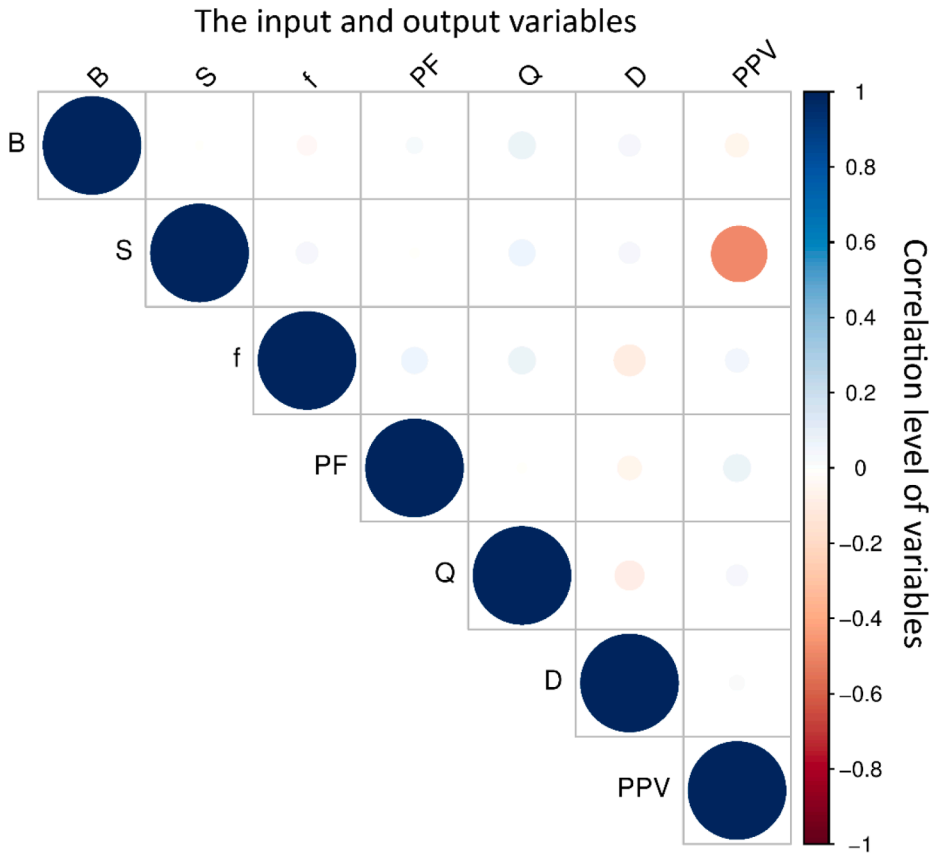


Fig. 8. Correlation matrix of the blasting parameters and PPV.

Table 1
Statistics of the prepared datasets for predicting PPV.

Category	B (m)	S (m)	f	PF (Kg/m^3)	Q (Kg)	D (m)	PPV (mm/s)
	Training dataset (80%)						
Min.	2.4	4.2	10	0.4	240	152	2.86
1st Qu.	5.1	4.9	10	0.4	319	184.8	3.572
Median	6.2	5.55	12	0.45	373.5	217	4.92
Mean	5.85	5.572	11.03	0.4526	366.6	217.1	4.96
3rd Qu.	6.8	6.2	12	0.5	410.2	238.2	6.27
Max.	7.4	7	12	0.5	464	325	7.22
	Testing dataset (20%)						
Min.	2.9	4.2	10	0.4	259	160	2.86
1st Qu.	5.475	4.9	10	0.4	316	184.5	3.953
Median	5.95	5.5	11	0.45	356	220	4.92
Mean	5.868	5.62	11	0.4534	363	221.9	4.966
3rd Qu.	6.525	6.4	12	0.5	405.2	250	6.27
Max.	7.3	7	12	0.5	469	310	7.22

considered as two common methods to estimate PPV in this study. To understand whether the proposed AI models (i.e., SpaSO-ELM, SalSO-ELM, and MFO-ELM) are good for predicting PPV, two forms of empirical equation were considered, including linear and nonlinear empirical equations. Whereas the USBM model, which is the most common empirical model, was used as the linear empirical model, the gene expression programming method was applied to generate a nonlinear model. The principle of these methods is briefly presented as below.

USBM model (linear regression model)

As a well-known empirical model for predicting PPV, the USBM empirical equation that was proposed by Duvall and Petkof [46], was used in this study, and it is described in Eq. (19).

$$PPV = k\left(\frac{D}{\sqrt{W}}\right)^{-b} \tag{19}$$

where k and b are the two coefficients that represent for the rock hardness as well as characteristics of each region. They can be calculated using the regression analysis method through the field measurements.

Nonlinear empirical model (NLE)

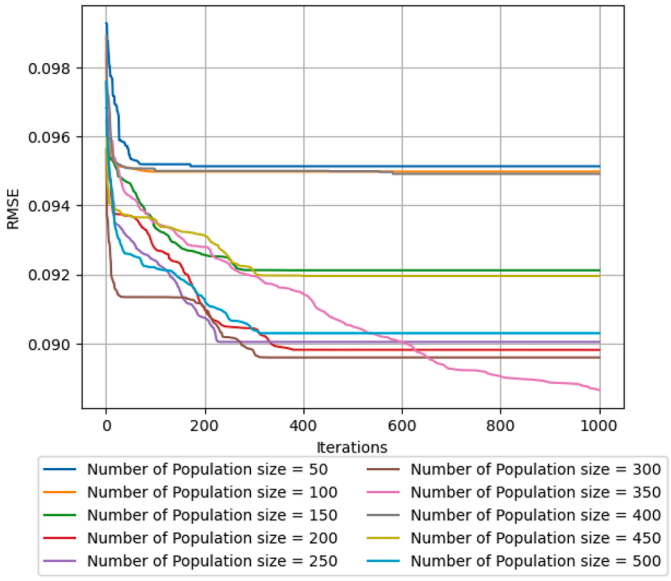
For the NLE model, genetic-based programming was used with various chromosomes, and genes were used to determine the relationship between the input variables. It used a fitness function (i.e., RMSE) to evaluate the fitness of the model with different configurations. During building the NLE model, the optimal evolution strategy was used with a variety of operators, such as function insertion, mutation, leaf mutation biased leaf mutation, fixed-root mutation, conservative mutation, permutation, inversion, gene recombination, to name a few. In addition, various functions were considered in the NLE between the input variables, such as addition, natural logarithm, subtraction, x to the power of 2, multiplication, division, cube root, exponential, inverse, to name a few. The addition linking function was used to combine the generated genes to create the NLE model. Finally, an NLE model was generated with the best RMSE (the lowest RMSE) to predict PPV as presented in Eq. (24).

2.7. Performance metrics for models' evaluation

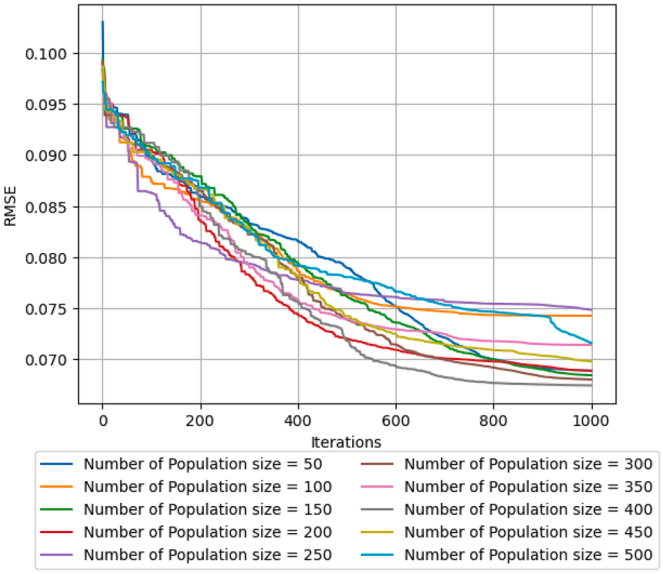
In order to evaluate the performance (e.g., accuracy, error, conformity) of the proposed AI-based and empirical models including ELM, SpaSO-ELM, SalSO-ELM, MFO-ELM, USBM, and NLE, three statistical metrics, including MAE, RMSE, and R^2 were used, and their descriptions are presented in Eqs. (20)–(22).

Table 2
Weights and biases of the traditional ELM model obtained for predicting PPV.

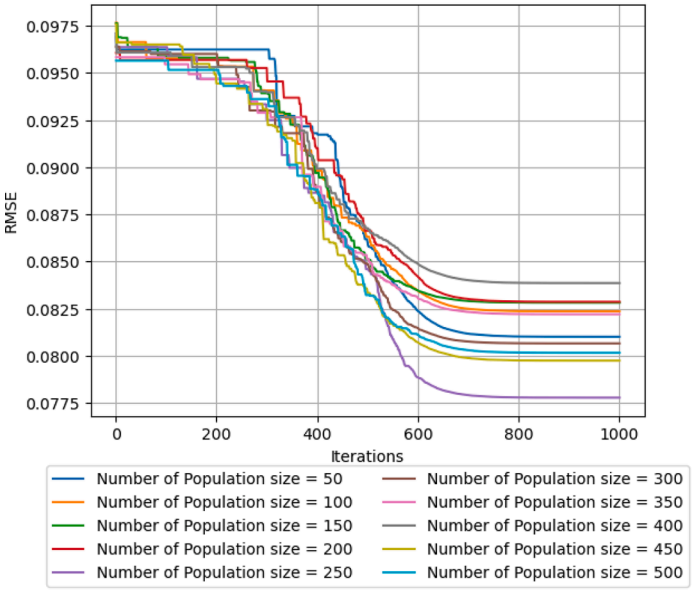
Inputs	Hidden neurons												
	1	2	3	4	5	6	7	8	9	10	11	12	13
B	0.695	0.753	0.202	0.755	0.909	0.218	0.209	0.013	0.345	0.551	0.310	0.322	0.723
S	0.477	0.720	0.797	0.629	0.626	0.820	0.162	0.469	0.049	0.917	0.532	0.265	0.133
f	0.827	0.555	0.763	0.023	0.146	0.206	0.335	0.739	0.453	0.544	0.992	0.146	0.520
PF	0.761	0.353	0.703	0.058	0.469	0.177	0.694	0.936	0.234	0.853	0.309	0.039	0.491
Q	0.300	0.204	0.939	0.547	0.651	0.738	0.825	0.920	0.458	0.705	0.301	0.971	0.461
D	0.932	0.675	0.208	0.035	0.428	0.504	0.368	0.178	0.416	0.565	0.941	0.801	0.587
Bias	0.851	0.117	0.261	0.281	0.210	0.169	0.990	0.854	0.583	0.087	0.156	0.440	0.800
Output													
PPV	−11.222	−2.063	−6.081	3.836	8.824	13.474	1.336	1.991	−13.188	4.528	4.638	−8.133	20.453



(a) SpaSO-ELM training performance



(c) MFO-ELM training performance



(b) SalSO-ELM training performance

Fig. 9. Optimization performance of the ELM-based models for predicting PPV.

Table 3
Optimized weights of the best SpaSO-ELM model for predicting PPV.

Inputs	Hidden neurons												
	1	2	3	4	5	6	7	8	9	10	11	12	13
B	−0.577	0.523	−0.940	−0.301	−0.239	−0.570	0.228	−0.871	0.759	−0.544	0.411	0.296	−0.903
S	−0.196	0.002	0.745	0.009	−0.474	−0.545	0.688	−0.508	0.302	0.669	−0.694	−1.000	−0.657
f	−0.047	0.027	−0.710	−0.078	0.309	0.423	−0.100	−0.035	0.633	−0.822	−0.470	0.294	0.726
PF	0.409	−0.274	0.465	0.554	−0.818	0.037	−0.719	0.600	−0.297	0.086	0.657	−0.194	0.006
Q	−0.512	0.117	−0.188	0.512	0.204	0.167	−0.533	−0.162	0.195	0.408	−0.233	0.359	−0.445
D	−0.211	−0.509	−0.466	−0.378	0.756	−0.553	−0.553	0.741	−0.097	0.281	0.513	−0.003	−0.340
Bias	−0.008	0.612	0.678	0.532	−0.415	0.603	−0.098	−0.812	0.340	0.398	0.365	0.290	0.012
Output													
PPV	−0.741	2.448	−0.588	−1.353	−0.514	0.907	−0.934	1.454	0.381	0.631	−0.235	−2.233	0.134

Table 4
Optimized weights of the best SalSO-ELM model for predicting PPV.

Inputs	Hidden neurons												
	1	2	3	4	5	6	7	8	9	10	11	12	13
B	−0.638	−0.715	0.224	0.336	−0.828	−0.084	0.343	0.808	−0.235	0.801	0.222	−0.653	−0.146
S	−0.535	0.576	0.055	0.936	−0.987	−0.506	0.022	−0.969	0.425	−0.604	0.562	0.764	0.198
f	0.969	−0.576	0.303	0.958	−0.818	0.786	−0.586	0.532	−0.232	−0.312	0.013	−0.137	0.009
PF	0.223	0.190	0.390	0.611	−0.894	−0.481	−0.738	0.112	0.816	0.104	−0.821	−0.025	0.200
Q	−0.467	−0.085	0.458	0.967	−0.814	−0.252	0.959	0.573	0.802	0.796	−0.423	−0.211	0.158
D	0.457	−0.938	0.027	−0.473	−0.623	0.403	0.172	0.758	−0.942	0.248	0.066	0.451	1.000
Bias	−0.170	0.233	0.411	0.504	0.514	−0.188	0.993	0.724	0.168	−0.493	0.168	0.056	−0.657
Output													
PPV	2.242	−20.427	32.542	−0.462	2.463	−9.349	1.322	−10.181	−13.031	13.550	−13.282	47.105	−45.672

Table 5
Optimized weights of the best MFO-ELM model for predicting PPV.

Inputs	Hidden neurons												
	1	2	3	4	5	6	7	8	9	10	11	12	13
B	−3.977	8.623	0.622	1.729	−13.652	−18.376	19.950	38.622	10.150	11.607	18.241	−1.592	44.872
S	702.776	−4.870	7.784	3.427	−9.813	31.432	−0.860	409.122	60.666	151.158	−10.884	80.105	−0.062
f	68.699	−5.890	7.358	−2.722	−5.215	−29.258	−2.029	17.807	19.264	−32.699	−44.945	−3.132	−286.476
PF	62.514	1.060	−12.875	−0.268	0.370	1.525	−4.749	−6.664	−6.623	−144.192	−2.029	−18.422	−13.337
Q	−15.408	−19.338	−81.086	−7.489	10.297	−147.220	−48.016	59.406	−25.525	−28.912	27.671	20.682	−25.548
D	48.473	13.173	118.830	34.974	−127.504	580.811	−50.948	44.906	21.609	63.590	−3.815	−3.335	46.509
Bias	−16.469	16.235	17.216	−10.527	28.888	−96.107	10.167	13.802	13.489	39.665	35.019	−14.588	−32.754
Output													
PPV	0.005	−0.058	0.009	0.068	−0.014	−0.005	−0.022	−0.012	0.033	−0.002	0.026	0.010	0.004

Table 6
Performance metrics of the AI-based and empirical models.

Model	Training			Testing		
	MAE	RMSE	R^2	MAE	RMSE	R^2
SpaSO-ELM	0.346	0.393	0.926	0.320	0.389	0.914
SalSO-ELM	0.307	0.361	0.937	0.347	0.418	0.905
MFO-ELM	0.245	0.304	0.955	0.360	0.431	0.898
ELM	0.428	0.498	0.880	0.350	0.434	0.898
USBM	0.892	1.106	0.418	0.967	1.158	0.274
NLE	0.648	0.801	0.747	0.667	0.795	0.691

$$\mathrm{MAE} = \frac{1}{n_{blast}} \sum_{blast=1}^{n_{blast}} |PPV_i - \widehat{PPV}_i| \tag{20}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n_{blast}} \sum_{blast=1}^{n_{blast}} (PPV_i - \widehat{PPV}_i)^2} \tag{21}$$

$$\mathrm{R}^2 = 1 - \frac{\sum_{blast=1}^{n_{blast}} (PPV_i - \widehat{PPV}_i)^2}{\sum_{blast=1}^{n_{blast}} (PPV_i - \overline{PPV}_i)^2} \tag{22}$$

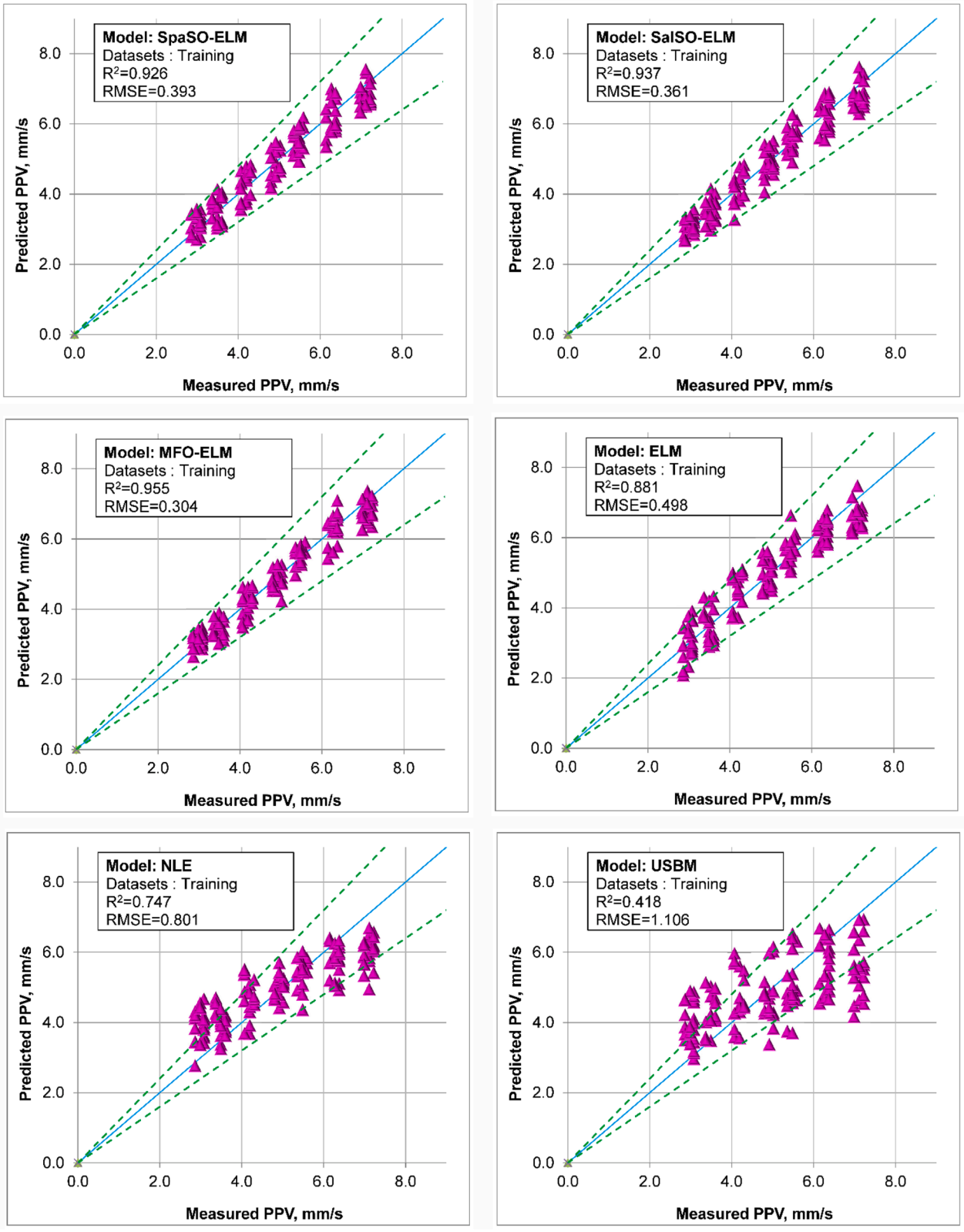
where n_{blast} is the number of blasting cases used in the dataset; PPV_i, $\widehat{PPV}_i$, and $\overline{PPV}_i$ stand for the measured PPV, predicted PPV, and mean of measured PPVs.

3. Data collection and preparation

3.1. Data collection

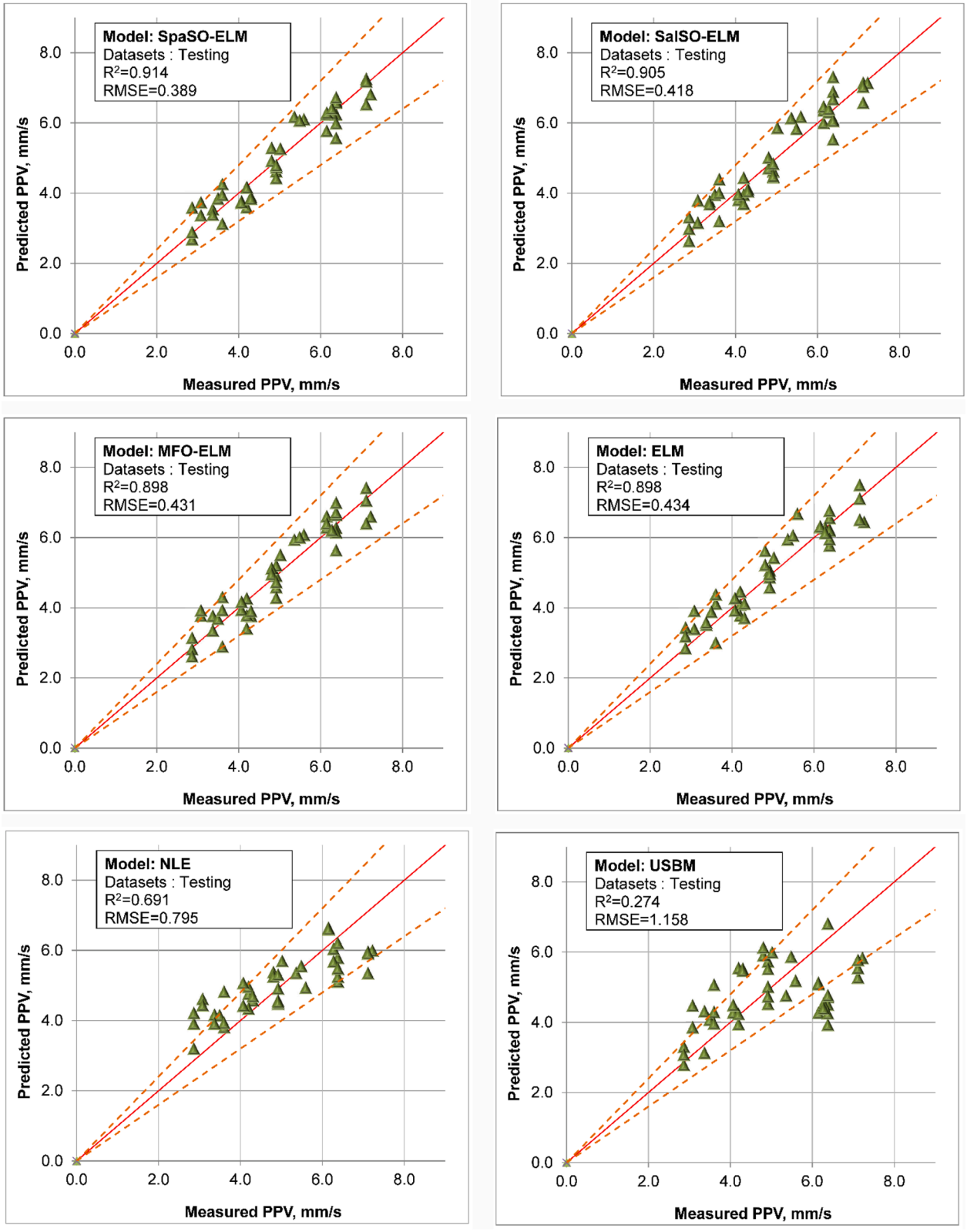
To diagnose the performance of the ELM, SpaSO-ELM, SalSO-ELM, MFO-ELM, USBM, and NLE models, 216 blasting events were collected at the Coc Sau open-pit coal mine, located in the Northern Vietnam (Fig. 6). In this coal mine, the rock hardness measured and evaluated in the range of 10 to 12 (based on Mohs hardness test), and therefore, blasting was recommended as the primary method for fragmenting rock.

According to the geological reports, the geological conditions of this area are not too complex. Furthermore, geological conditions are classified into the uncontrollable parameters group [27,47]; therefore, they



(a) Training dataset

Fig. 10. Correlation between the predicted PPVs versus measured PPVs.



(b) Testing dataset

Fig. 10. (*continued*).

have rarely been used to predict PPV. Review of related works showed that most of the previous studies used blasting parameters only for predicting PPV, and the accuracies were promising [1,22,24,38]. Hence, this study also considered five blasting parameters for predicting PPV, including burden (B), spacing (S), rock hardness (f), powder factor (PF), and maximum explosive charged per delay (Q). These parameters were

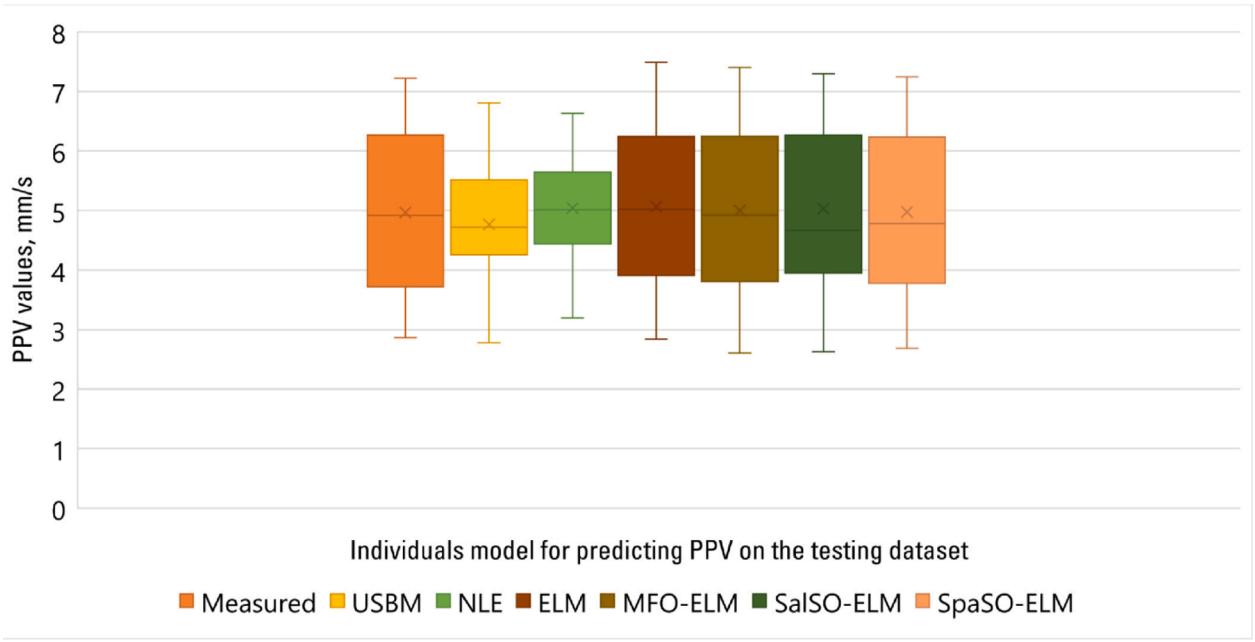


Fig. 11. Comparison of the actual models and developed models.

exported from blasting patterns. To measure PPV, the micromate seismograph (instantel – Canada) was used, and it was put at sensitive locations, such as bench face/slope, nearby households, or nearby open-pit coal mines (e.g., Cao Son and Deo Nai). Before measuring PPV, the distance between the measured locations and blast sites was determined by a GPS system. Finally, the dataset with 216 observations and 7 variables were used to predict PPV. The statistics of the dataset are illustrated through the boxplots, as shown in Fig. 7.

3.2. Data preparation

Prior to developing the prediction models, the comprehensive understanding of the relationship between the input variables can make/provide new ideas during developing the models. Thus, the correlation of the input variables, i.e., B, S, f, PF, Q, D, was analyzed, and strong correlations may be considered to eliminate aiming to avoid the insufficient data quality and their contributions to prediction. As illustrated in Fig. 8, it is evident that all the variables used in the dataset have weak correlations. In other words, they may have nonlinear relationships, especially the PPV variable. Therefore, nonlinear predictive models may be potential candidates for predicting PPV in this case. However, it is still too early to conclude which model is the best for predicting PPV at the Coc Sau open-pit coal mine. The answer will be demonstrated in the next sections.

For training the models, 80% of the whole dataset was randomly selected, and the remaining 20% was used for testing the accuracy of the trained models. The datasets are summarized in Table 1.

4. Designing, training, and building prediction models

4.1. ELM model

For the development of the ELM model, 13 hidden nodes (neurons) were selected for the topology network of the ELM, and four different activation functions, including relu, elu, sigmoid, and tanh were considered using the trial-and-error procedure. In addition, the MinMax scale method was employed to normalize the datasets before training the model, aiming to reduce the bias of the model. Finally, the best ELM model was identified with the sigmoid activation function (i.e., MAE = 0.428, RMSE = 0.498, R^2 = 0.880). The weights of the ELM model are presented in Table 2.

4.2. SpaSO-ELM, SalSO-ELM, and MFO-ELM models

To develop the mentioned hybrid ELM-based optimization models, the proposed framework in Fig. 5 was utilized. Accordingly, the advantages of three metaheuristic algorithms (i.e., SpaSO, SalSO, and MFO) were applied to optimize the ELM model. The same topology network of the ELM was also used during optimization process of the SpaSO, SalSO, and MFO algorithms. Also, the MinMax scale method was utilized to normalize the dataset before running optimization processes, and different activation functions were also considered.

Herein, the initial populations of the SpaSO, SalSO, and MFO algorithms were generated, and each population created a random set of weights for the ELM model. To discover the diversity of the populations and their effects on the performance of the optimization process, different population sizes were applied with 50, 100, 150, 200, 250, 300, 350, 400, 450, 500 individuals (solutions).

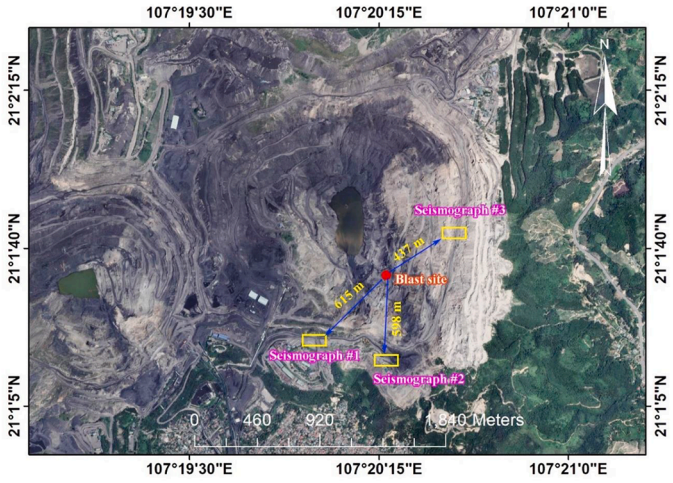
Besides, the parameters of the metaheuristic algorithms were also set up before training the ELM model, as follow:

The SpaSO's parameters:
- Number of epochs: 1000
- The safety threshold value: [0.6, 0.8]
- The number of producers: [0.1, 0.2]
- The number of sparrows who perceive the danger: 0.1

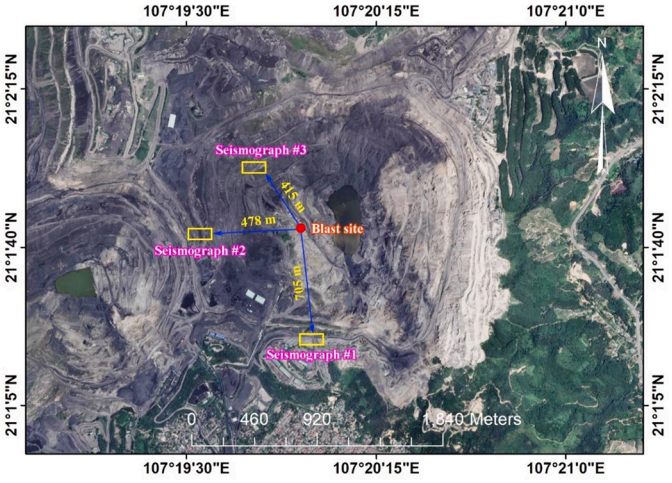
The SalSO's parameters:
- Number of epochs: 1000

The MFO's parameters:
- Number of epochs: 1000

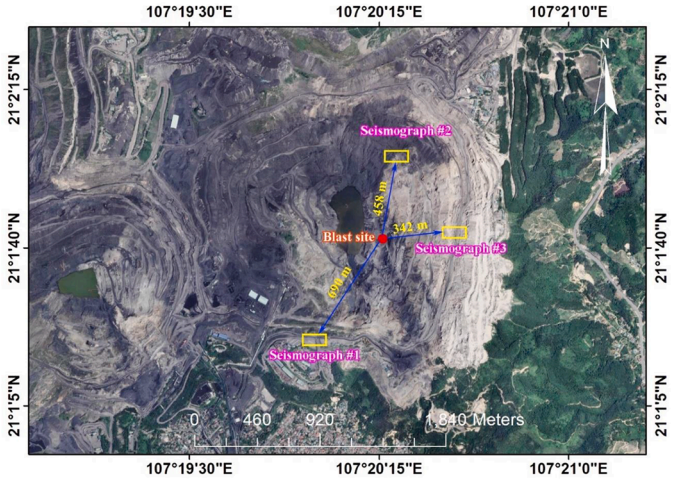
To evaluate the fitness of solutions, RMSE was used as the objective function of the models during the optimization process of the SpaSO, SalSO, and MFO algorithms. The lowest RMSE is corresponding to the best solution (the best set of weights). The optimization processes of the SpaSO-ELM, SalSO-ELM, and MFO-ELM models are presented in Fig. 9. Finally, the best SpaSO-ELM, SalSO-ELM, and MFO-ELM models were selected with the lowest RMSE values, and their weights are listed in Tables 3–5.



Blasting event #1 for validation



Blasting event #3 for validation



Blasting event #2 for validation

Fig. 12. Locations validation blasts and seismographs for measuring PPV.

Table 7
Summarize of the blasting events for validation and measured PPVs.

Seismograph	Blasting event	B	S	f	PF	Q	D	PPV
Seismograph #1	Blast #1	7	5.5	12	0.5	412	615	2.375
	Blast #2	6	5.5	10	0.45	401	690	2.105
	Blast #3	5.6	6.4	12	0.5	366	705	6.425
Seismograph #2	Blast #1	7	5.5	12	0.5	412	598	2.406
	Blast #2	6	5.5	10	0.45	401	458	2.388
	Blast #3	5.6	6.4	12	0.5	366	478	7.012
Seismograph #3	Blast #1	7	5.5	12	0.5	412	437	2.86
	Blast #2	6	5.5	10	0.45	401	342	2.761
	Blast #3	5.6	6.4	12	0.5	366	415	7.109

4.3. Empirical models

For the empirical models, the principles in Eq. (19) and gene-based programming method were applied and calculated. The empirical equations used for estimating PPV in this study are described in Eqs. (23)–(24).

(i) USBM model (linear empirical model):

$$PPV = 244.198\left(\frac{D}{\sqrt{W}}\right)^{-0.859} \tag{23}$$

(ii) Nonlinear empirical model:

Note that, the multivariate regression analysis method was applied to determine the coefficients in Eq. (23) based on the historical dataset, and

Table 8
Comparison of PPV predictions on the validation dataset in practical engineering.

Measured PPV	Predicted PPVs					
	SpaSO-ELM	SalSO-ELM	MFO-ELM	ELM	NLE	USBM
2.375	2.475	0.660	2.160	2.461	5.285	1.424
2.105	2.331	3.375	2.281	2.588	5.231	1.225
6.425	6.864	5.615	9.100	8.248	5.927	1.941
2.406	2.543	0.886	2.131	2.527	5.283	1.467
2.388	2.587	2.820	3.196	3.311	5.211	1.838
7.012	6.840	7.349	8.888	7.069	5.909	2.815
2.86	3.077	3.359	3.110	3.116	5.266	2.068
2.761	2.641	2.189	2.899	2.671	5.195	2.515
7.109	6.846	8.287	9.561	7.281	5.902	3.197

Table 9
Performance metrics of the AI-based and empirical models in practical engineering.

Model	Training		
	MAE	RMSE	R^2
SpaSO-ELM	0.208	0.229	0.990
SalSO-ELM	0.926	1.043	0.844
MFO-ELM	0.985	1.397	0.984
ELM	0.446	0.710	0.941
USBM	1.883	2.507	0.565
NLE	2.154	2.336	0.970

Considering the results obtained by the standalone ELM and optimized ELM models, to be sure, the AI-based models are much better than those of the empirical models in predicting PPV. The performance of the standalone ELM model was improved approximately by 14–20% compared to the performance obtained by the NLE model. It differs in learning theory from that of the NLE model (nonlinear equation). However, the primary objective of this study is to propose SpaSO-ELM, SalSO-ELM, and MFO-ELM models for predicting PPV. Now, we can see the results obtained by the hybrid ELM models and discover how the performance improved with three optimization algorithms. Firstly, by looking at the MFO-ELM model, evidently, the results varied somewhat in the performance and accuracy. The MFO algorithm optimized the ELM model to improve its accuracy, and the results can be observed on the training dataset in Table 6. On the testing dataset, there was some variation in accuracy of the ELM and MFO-ELM model. Secondly, considering the performance of the SalSO-ELM model in predicting PPV, it is clear that its performance is superior to those of the MFO-ELM and the other models. Compared to the standalone ELM model, the accuracy of the SalSO-ELM model was significantly improved on both training and testing datasets. Finally, beyond all shadow of doubt, the SpaSO-ELM model is the best model for predicting PPV in this study with the highest performance and accuracy. In effect, the errors of the SpaSO-ELM are lowest on both training and testing samples, and the R^2 values are highest as well. In a comparison with the ELM model, the accuracy of the SpaSO-ELM model was increased ~ 2%, and it is ~ 22% compared to the nonlinear empirical model (i.e., NLE). The accuracies of the developed models can be further interpreted in Fig. 10.

As can be seen in Fig. 10, the USBM model provided the worst convergence. In contrast, the AI-based models provided better convergences, especially the SpaSO-ELM model. Whereas several data points were detected outside of the 80% confidence level of the empirical models, the ELM-based models provided most of the data points inside of this range, especially the hybrid models. These results indicated that the metaheuristic algorithm-based ELM models yielded better predictability than those of the standalone ELM and empirical models in predicting PPV. To indicate which model is the best and the most closely the actual model, the boxplot below shows the statistics of the actual and developed models based on the median, quartiles, whiskers, fences and outliers (Fig. 11). Fig. 11 is a powerful graphical representation of the outcome predictions that give an overview and numerical summary of the outcome predictions on the testing dataset. The results revealed that the SpaSO-ELM model tends to be strongly similar to the actual model (measured model). Meanwhile, the other hybrid models provided slightly lower similarities, and the empirical models are far from the actual model. These findings are in agreement with the results and discussions in Table 6.

$$PPV = S + \left(\sqrt{3\left(\frac{1}{12977.86 - PF}\right)^2}\right)^2 + \sqrt[3]{\frac{B - \sqrt[3]{f}}{-2.054D}} + \sqrt[3]{12.807 + Q + \tanh(Q) + \tanh\left(Q^2\right)} - 1.156\left(\frac{PF}{26.686(86.118 - f)} + 7.457\right) + \tanh(f) \quad (24)$$

the gene expression programming method was applied to develop Eq. (24).

5. Results and discussion

Once the optimization process-based ELM models as well as the standalone ELM and empirical models were developed, the prediction results by these models were considered and evaluated. Indeed, three performance metrics (MAE, RMSE and R^2) were computed for not only the training samples but also testing using Eqs. (20)–(22), as shown in Table 6.

Based on the obtained performance metrics, the USBM model showed the worst performance with the highest errors (MAE = 0.892, RMSE = 1.106) on the training samples. The results also indicated that the USBM model seems not to appropriate to predict PPV in this case with an R^2 of 0.418 for the training dataset. On the testing dataset, the accuracy and determination coefficient are even worse than the performance on the training dataset. In contrast, the NLE model provided significantly superior performance compared to the USBM model. From the empirical approach point of view, the results of the NLE model are acceptable. Remarkably, R^2 of the NLE model improved up to 0.747 on the training dataset, and 0.691 on the testing dataset. This result shows that the blasting dataset seems more suitable with the NLE model. In other words, we have evidence to believe that nonlinear models are better than linear models for predicting PPV in this study. This statement is in agreement with the previous evaluation in the data preparation section.

6. Experimental risk assessment

To demonstrate the performance of the proposed models, three other blasting events at the Coc Sau open-pit coal mine were selected to validate the accuracy of the models in practical engineering after the PPV predictive models proposed. For each blasting event, three seismographs were used to measure PPV at different locations, as shown in Fig. 12. The blasting parameters used and measured PPVs by the

seismographs are summarized in Table 7. Finally, the validation dataset was import to the developed models, and the predicted PPVs by the proposed models are shown in Table 8.

As indicated in the validation results, we can see that the predicted PPVs by the hybrid ELM models are very favorable, especially those obtained by the SpaSO-ELM model. The results showed that the SpaSO-ELM model also yielded the most dominant accuracy in experimental validation. In contrast, the empirical models are in poor stale for predicting PPV in practical engineering. Whereas the USBM model predicted too low PPVs compared to the actual PPVs, the NLE model predicted PPVs with the mean of PPV is around 5.2 to 5.9 mm/s in practical engineering. These evaluations are interpreted more clear in Table 9 through the performance metrics of the models in practical engineering.

Indeed, the performance indicators of the models were tested in practical engineering through nine blasts showed which model is the best in practice, i.e., the SpaSO-ELM model, with an RMSE of 0.229, R^2 of 0.990 and MAE of 0.208 only. Meanwhile, the other models provided the accuracies in the range of 0.565 to 0.984 for R^2 and RMSE in the range of 0.710 to 2.507. Remarkably, the performance of the ELM seems to be better than the optimized models SalSO-ELM and MFO-ELM models. This showed that not all AI models optimized by metaheuristic algorithms provide higher accuracy than traditional model (without optimization). And in this case, the SpaSO seems to be better than the other metaheuristic algorithms in optimization of the ELM model for predicting PPV in open-pit mines.

7. Conclusions

Blasting and its adverse effects are considerable concerns in open-pit mining, especially PPV. Controlling PPV in open-pit mines, especially near residential areas, is challenging. This study, therefore, proposes three AI-based models (i.e., SpaSO-ELM, SalSO-ELM, and MFO-ELM) for predicting PPV induced by bench blasting, and they were applied at the Coc Sau coal mine with high accuracy and reliability. Of those, the SpaSO-ELM model provided the best performance and its accuracy was superior 20 — 22% compared to the empirical models. The optimization role of the used metaheuristic algorithms was also interpreted in this study with the improvements of the standalone ELM model, especially the SpaSO algorithm. Conspicuously enough, the three proposed ELM-based hybrid models were thus far novel AI models in predicting PPV. The applications in practical engineering for predicting PPV were also conducted and validated with high accuracy obtained. They should be used towards controlling PPV and its adverse effects in open-pit mines. It is to be noted that the SpaSO-ELM model is proposed as the best choice for this task in practical engineering. In addition, the consideration of the optimization of blasting parameters based on the proposed SpaSO-ELM model is a further research recommendation to minimize PPV in mine blasting.

CRediT authorship contribution statement

Hoang Nguyen: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Xuan-Nam Bui:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Investigation, Funding acquisition, Conceptualization. **Erkan Topal:** Writing – review & editing, Writing – original draft, Validation, Formal analysis.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

References

[1] Abbaszadeh SA, Asheghi R. Optimized developed artificial neural network-based models to predict the blast-induced ground vibration. Innov Infrastruct Solut 2018; 3:1–10.
[2] Ainalis D, Kaufmann O, Tshibangu JP, Verlinden O, Kouroussis G. Modelling the source of blasting for the numerical simulation of blast-induced ground vibrations: a review. Rock Mech Rock Eng 2017;50:171–93.
[3] Monjezi M, Ahmadi M, Sheikhan M, Bahrami A, Salimi A. Predicting blast-induced ground vibration using various types of neural networks. Soil Dyn Earthq Eng 2010;30:1233–6.
[4] Marks NA, Stewart MG, Netherton MD, Stirling CG. Airblast variability and fatality risks from a VBIED in a complex urban environment. Reliab Eng Syst Saf 2021;209: 107459.
[5] Stewart MG, Netherton MD. A probabilistic risk-acceptance model for assessing blast and fragmentation safety hazards. Reliab Eng Syst Saf 2019;191:106492.
[6] Gerassis S, Albuquerque MTD, García JF, Boente C, Giráldez E, Taboada J, et al. Understanding complex blasting operations: a structural equation model combining Bayesian networks and latent class clustering. Reliab Eng Syst Saf 2019; 188:195–204.
[7] Nguyen H, Bui XN, Tran QH, Nguyen DA, Hoa LTT, Le QT, et al. Predicting blast-induced ground vibration in open-pit mines using different nature-inspired optimization algorithms and deep neural network. Nat Resour Res 2021;30: 4695–717.
[8] Nicholls HR, Johnson CF, Duvall WI. Blasting vibrations and their effects on structures. US Government Printers; 1971.
[9] Murmu S, Maheshwari P, Verma HK. Empirical and probabilistic analysis of blast-induced ground vibrations. Int J Rock Mech Min Sci 2018;103:267–74.
[10] Mesec J, Kovač I, Soldo B. Estimation of particle velocity based on blast event measurements at different rock units. Soil Dyn Earthquake Eng 2010;30:1004–9.
[11] Ak H, Iphar M, Yavuz M, Konuk A. Evaluation of ground vibration effect of blasting operations in a magnesite mine. Soil Dyn Earthquake Eng 2009;29:669–76.
[12] Bui XN, Bui HB, Nguyen H. A review of artificial intelligence applications in mining and geological engineering. In: Proceedings of the international conference on innovations for sustainable and responsible mining. Springer; 2021. p. 109–42.
[13] Jung D, Choi Y. Systematic review of machine learning applications in mining: exploration, exploitation, and reclamation. Minerals 2021;11:148.
[14] Kim M, Ismail L, Kwon S. Review of the application of artificial intelligence in blasting area. Explos Blast 2021;39:44–64.
[15] Yong W, Zhang W, Nguyen H, Bui X-N, Choi Y, Nguyen-Thoi T, et al. Analysis and prediction of diaphragm wall deflection induced by deep braced excavations using finite element method and artificial neural network optimized by metaheuristic algorithms. Reliab Eng Syst Saf 2022;221:108335.
[16] Liu D, Wang S, Cui X. An artificial neural network supported Wiener process based reliability estimation method considering individual difference and measurement error. Reliab Eng Syst Saf 2022;218:108162.
[17] Yang Z, Baraldi P, Zio E. A method for fault detection in multi-component systems based on sparse autoencoder-based deep neural networks. Reliab Eng Syst Saf 2022;220:108278.
[18] Seo SK, Yoon YG, Lee JS, Na J, Lee CJ. Deep neural network-based optimization framework for safety evacuation route during toxic gas leak incidents. Reliab Eng Syst Saf 2022;218:108102.
[19] Monjezi M, Hasanipanah M, Khandelwal M. Evaluation and prediction of blast-induced ground vibration at Shur River Dam, Iran, by artificial neural network. Neural Comput Appl 2013;22:1637–43.
[20] Saadat M, Khandelwal M, Monjezi M. An ANN-based approach to predict blast-induced ground vibration of Gol-E-Gohar iron ore mine, Iran. J Rock Mech Geotech Eng 2014;6:67–76.
[21] Hasanipanah M, Monjezi M, Shahnazar A, Armaghani DJ, Farazmand A. Feasibility of indirect determination of blast induced ground vibration based on support vector machine. Measurement 2015;75:289–97.
[22] Amiri M, Bakhshandeh Amnieh H, Hasanipanah M, Mohammad Khanli L. A new combination of artificial neural network and K-nearest neighbors models to predict blast-induced ground vibration and air-overpressure. Eng Comput 2016;32: 631–44.
[23] Hasanipanah M, Faradonbeh RS, Amnieh HB, Armaghani DJ, Monjezi M. Forecasting blast-induced ground vibration developing a CART model. Eng Comput 2017;33:307–16.
[24] Armaghani DJ, Hasanipanah M, Amnieh HB, Mohamad ET. Feasibility of ICA in approximating ground vibration resulting from mine blasting. Neural Comput Appl 2018;29:457–65.
[25] Azimi Y, Khoshrou SH, Osanloo M. Prediction of blast induced ground vibration (BIGV) of quarry mining using hybrid genetic algorithm optimized artificial neural network. Measurement 2019;147:106874.
[26] Bui XN, Jaroonpattanapong P, Nguyen H, Tran QH, Long NQ. A novel hybrid model for predicting blast-induced ground vibration based on k-nearest neighbors and particle swarm optimization. Sci Rep 2019;9:13971.

[27] Shang Y, Nguyen H, Bui XN, Tran QH, Moayedi H. 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 2020;29:723–37.
[28] Zhang X, Nguyen H, Bui XN, Tran QH, Nguyen DA, Bui DT, et al. Novel soft computing model for predicting blast-induced ground vibration in open-pit mines based on particle swarm optimization and XGBoost. Nat Resour Res 2020;29:711–21.
[29] Qiu Y, Zhou J, Khandelwal M, Yang H, Yang P, Li C. Performance evaluation of hybrid WOA-XGBoost, GWO-XGBoost and BO-XGBoost models to predict blast-induced ground vibration. Eng Comput 2021.
[30] Zhou J, Qiu Y, Khandelwal M, Zhu S, Zhang X. Developing a hybrid model of Jaya algorithm-based extreme gradient boosting machine to estimate blast-induced ground vibrations. Int J Rock Mech Min Sci 2021;145:104856.
[31] Lawal AI, Kwon S, Hammed OS, Idris MA. Blast-induced ground vibration prediction in granite quarries: an application of gene expression programming, ANFIS, and sine cosine algorithm optimized ANN. Int J Mining Sci Technol 2021;31:265–77.
[32] Bui XN, Nguyen H, Tran QH, Nguyen DA, Bui HB. Predicting blast-induced ground vibration in quarries using adaptive fuzzy inference neural network and moth–flame optimization. Nat Resour Res 2021;30:4719–34.
[33] Ding X, Hasanipanah M, Rad HN, Zhou W. Predicting the blast-induced vibration velocity using a bagged support vector regression optimized with firefly algorithm. Eng Comput 2020:1–12.
[34] Zhu W, Rad HN, Hasanipanah M. A chaos recurrent ANFIS optimized by PSO to predict ground vibration generated in rock blasting. Appl Soft Comput 2021;108:107434.
[35] Yu C, Koopialipoor M, Murlidhar BR, Mohammed AS, Armaghani DJ, Mohamad ET, et al. Optimal ELM–harris hawks optimization and ELM–grasshopper optimization models to forecast peak particle velocity resulting from mine blasting. Nat Resour Res 2021;30:2647–62.
[36] Zhang X, Nguyen H, Choi Y, Bui XN, Zhou J. Novel extreme learning machine-multi-verse optimization model for predicting peak particle velocity induced by mine blasting. Nat Resour Res 2021;30:4735–51.
[37] Armaghani DJ, Kumar D, Samui P, Hasanipanah M, Roy B. A novel approach for forecasting of ground vibrations resulting from blasting: modified particle swarm optimization coupled extreme learning machine. Eng Comput 2021;37:3221–35.
[38] Arthur CK, Temeng VA, Ziggah YY. A Self-adaptive differential evolutionary extreme learning machine (SaDE-ELM): a novel approach to blast-induced ground vibration prediction. SN Appl Sci 2020;2:1845.
[39] Ali M, Prasad R. Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition. Renew Sustain Energy Rev 2019;104:281–95.
[40] Han S, Zhu K, Wang R. Improvement of evolution process of dandelion algorithm with extreme learning machine for global optimization problems. Expert Syst Appl 2021;163:113803.
[41] Xue J, Shen B. A novel swarm intelligence optimization approach: sparrow search algorithm. Syst Sci Control Eng 2020;8:22–34.
[42] Fathy A, Alanazi TM, Rezk H, Yousri D. Optimal energy management of micro-grid using sparrow search algorithm. Energy Rep 2022;8:758–73.
[43] Mirjalili S, Gandomi AH, Mirjalili SZ, Saremi S, Faris H, Mirjalili SM. Salp swarm algorithm: a bio-inspired optimizer for engineering design problems. Adv Eng Software 2017;114:163–91.
[44] Zheng X, Nguyen H, Bui XN. Exploring the relation between production factors, ore grades, and life of mine for forecasting mining capital cost through a novel cascade forward neural network-based salp swarm optimization model. Resour Policy 2021;74:102300.
[45] Mirjalili S. Moth-flame optimization algorithm: a novel nature-inspired heuristic paradigm. Knowl Based Syst 2015;89:228–49.
[46] Duvall WI, Petkof B. Spherical propagation of explosion-generated strain pulses in rock: us department of the interior. Bureau of Mines; 1958.
[47] Hasanipanah M, Bakhshandeh Amnieh H, Khamesi H, Jahed Armaghani D, Bagheri Golzar S, Shahnazar A. Prediction of an environmental issue of mine blasting: an imperialistic competitive algorithm-based fuzzy system. Int J Environ Sci Technol 2018;15:551–60.